Handbook of Sustainable Finance

Roncalli, Thierry

University of Paris-Saclay

1 January 2024

Handbook of Sustainable Finance

Thierry Roncalli

Université Paris-Saclay

Contents

Preface

Teaching sustainable finance

I started teaching finance at the University of Bordeaux in 1995. My first course was called *Stochastic Finance* and was dedicated to option pricing and stochastic calculus. At that time I was also working as a consultant for a bank to develop numerical methods for some exotic payoffs. I realized that this academic course could benefit from this professional position. Indeed, I learnt that the main problem of traders is not the calculation of the option price, but the definition of the hedging strategy. Combining academic theory with professional practice has been a constant theme of my teaching career. In the 2000s, my risk management courses made extensive use of the risk management knowledge acquired at the Groupe de Recherche Opérationnelle (GRO) at Crédit Lyonnais and Crédit Agricole between 1999 and 2005. In the 2010s, my courses in asset management were largely based on the professional experience developed at SGAM AI and Lyxor Asset Management between 2005 and 2016. This sustainable finance course follows the same path. It is largely based on the ESG and climate investing research I have conducted at Amundi Asset Management since 2018.

Although this course has many features in common with the previous ones, sustainable finance is not as mature as option pricing, risk management and asset management. In particular, regulation is in its infancy and not yet stabilised, academic models are still in their infancy, data are noisy, biased and of poor quality, and even concepts are not well defined. In this context, all actors (investors, issuers, financial institutions, regulators, etc.) are adopting a *learning-by-doing* approach. This has important implications for the development of a training course. Each year, the course needs to be updated by incorporating new advances in modelling, adjusting definitions, changing the structure of the lecture notes, removing obsolete sections and adding new paragraphs. Since its inception, the course has been a continuous work in progress.

The issue of sustainable economic growth is a major change in the way economics and finance are taught. However, most academic models do not take climate change into account because they were developed in the 19th and 20th centuries. Taking climate change into account requires a complete overhaul of economic theory and many related concepts: economic growth, negative externality, moral hazard, labour productivity, economic rationality, consumption maximization, social welfare, Pareto optimality, invisible hand, market efficiency, utility function, homo economicus, capital allocation, risk theory, golden rule, overlapping generations model, steady state, non-cooperative games, etc. As teachers, it is difficult to adapt to this new world because we lack distance and cannot benefit from well-established textbooks. However, we cannot wait as time is running out. Therefore, we have a heavy responsibility to educate students about these important issues without adequate tools and a normative framework. This explains why there are thousands of ways to teach sustainable finance. It is not a science. Then be careful about the content, because it necessarily reflects our personal beliefs and experiences. And so I claim the right to be biased.

About these lecture notes

In January 2019, I started to introduce some elements of ESG and climate investing in my course *Advanced Asset Management*. At that time, it was only about portfolio optimisation with ESG scores and carbon footprints. Over the years, the part dedicated to sustainable finance has increased and now makes up 25% of the course. In 2021, we had some discussions at the University of Paris-Saclay about the future development of the Masters in Finance. We decided to create a full and comprehensive course in sustainability finance, going beyond portfolio allocation. For the 2021/2022 academic year, I then put together a set of 770 slides for the first course at Paris-Saclay University. During the year, Peter Tankov also offered to share his course *Green Finance* at ENSAE Paris. At the same time, Emmanuel Gobet and Gilles Pagès proposed that I create a mathematical course on sustainable finance in the Master of Probability & Finance at Sorbonne University. Today, sustainable finance is taught in four master's programmes: *Master in Risk and Asset Management* (Paris-Saclay University), *Master in Banking & Finance* (Paris-Saclay University), *Master in Statistics, Finance and Actuarial Science* (ENSAE Paris and Paris Cité University) and *Master in Probability & Finance* (Sorbonne University). These four courses differ in the number of hours and the scientific approach (qualitative/quantitative and economics/mathematics). Unfortunately, creating four different lecture notes was not effective because there was a lot of overlap between them. Therefore, I decided to write a single set of lecture notes and use them for each course. This explains why some parts of this handbook are highly descriptive, but other parts are technical and require a mathematical background in probability, statistics, machine learning, linear algebra, optimization and stochastic analysis.

The publication of my previous lecture notes on risk management (Roncalli, 2020a, Handbook of Financial Risk Management) was a great disappointment. In fact, the publisher has set a high price that is prohibitive for my students. Therefore, I have decided to publish this manual for free. Whatever happens, the electronic version of this manual is and will remain free for my students and everyone else. It is available at the following web site:

<div align="center">

http://www.thierry-roncalli.com/SustainableFinanceBook.html

</div>

Acknowledgments

I would like to thank the various people who helped me organise these academic courses, in alphabetical order: Michel Guillard, Emmanuel Gobet, Eleni Iliopulos, Claire Loupias, Gilles Pagès, Fabrice Pansard, Peter Tankov and Fabien Tripier. I would also like to thank all the Masters students who participated in my Sustainable Finance courses.

I would also like to thank all my co-authors at Amundi Asset Management, in particular Inès Barahhou, Mohamed Ben Slimane, Leila Bennani, Amina Cherief, Angelo Drei, Théo Le Guenedal, Frédéric Lepetit, Edmond Lezmi, François Lombard, Noureddine Oulid Azouz, Théo Roncalli, Raphaël Semet, Lauren Stagnol, Takaya Sekine and Jiali Xu. I have been learning about the practice of sustainable finance since I joined Amundi's quantitative research team. The writing of these lecture notes is therefore heavily influenced by the materials and models developed at Amundi.

I would also like to thank Mohamed Ben Slimane, Théo Roncalli and Jiali Xu for reading some parts of this book and testing the exercises.

Paris, October 2023   Thierry Roncalli

List of Symbols and Notations

Symbol Description

$\times$	Arithmetic multiplication
$\cdot$	Scalar, vector and matrix multiplication
$*$	Convolution
$\circ$	Hadamard product $(x \circ y)_i = x_i y_i$
$\otimes$	Kronecker product $A \otimes B$
$|\mathcal{E}|$	Cardinality of the set $\mathcal{E}$
$\prec$	Preference ordering
$\langle x, x' \rangle$	Inner product of x and x'
$\mathbf{0}$ (or $\mathbf{0}_n$)	Vector of zeros
$\mathbf{1}$ (or $\mathbf{1}_n$)	Vector of ones
$\mathbb{1}\{\mathcal{A}\}$	The indicator function is equal to 1 if $\mathcal{A}$ is true, 0 otherwise
$\mathbb{1}_{\mathcal{A}}\{x\}$	The characteristic function is equal to 1 if $x \in \mathcal{A}$, 0 otherwise
$(A_{i,j})$	Matrix A with entry $A_{i,j}$ in row i and column j
A^{-1}	Inverse of the matrix A
$A^{\top}$	Transpose of the matrix A
A^{+}	Moore-Penrose pseudo-inverse of the matrix A
$\mathcal{AE}$	Avoided emissions
$\mathrm{AS}\left(w \mid b\right)$	Asset-based active share
$\mathrm{AS}_{\boldsymbol{\mathcal{S}}ector}\left(w \mid b\right)$	Sector-based active share
α_p	Planetary albedo (default value $= 0.29$)
b	Vector of weights $(b_1, \ldots, b_n)$ for the benchmark b
$\textit{b}$	$\textit{b}$-score (or Beta score)
$B_{\lambda}\left(\lambda, \mathcal{T}\right)$	Planck radiation law
$\mathcal{B}\left(p\right)$	Bernoulli distribution with parameter p
$\mathcal{B}\left(n, p\right)$	Binomial distribution with parameters n and p
$\mathcal{B}ates\left(n\right)$	Bates distribution with parameter n
$\mathcal{B}\left(\alpha, \beta\right)$	Beta distribution with parameters α and β
β_i	Beta of asset i with respect to portfolio w
$\beta_i\left(w\right)$	Another notation for the symbol β_i
$\beta\left(w \mid b\right)$	Beta of portfolio w when the benchmark is b
c	Specific heat capacity
$\mathrm{cov}\left(X\right)$	Covariance of the random vector X
$\mathcal{C}_i$	Carry of bond i
$\mathbb{C}$ (or ρ)	Correlation matrix
$\mathbb{C}_n\left(\rho\right)$	Constant correlation matrix of size n with uniform correlation ρ
$\boldsymbol{\mathcal{CB}}\left(t_1, t_2\right)$	Carbon budget between t_1 and t_2

$\mathcal{CC}$	Concentration rate
$\mathcal{CE}$	Carbon emissions
$\mathcal{CI}$	Carbon intensity
$\mathcal{CI}^{\star}$	Maximum carbon intensity value or threshold
$\mathcal{CM}$	Carbon momentum
$\mathcal{CM}^{Long}$	Long-term carbon momentum
$\mathcal{CM}^{Short}$	Short-term carbon momentum
CPI	Consumer price index

d	Distance
$\det(A)$	Determinant of the matrix A
$\operatorname{diag}(a)$	Diagonal matrix with elements $(a_1, \ldots, a_n)$
D	Covariance matrix of idiosyncratic risks
$\mathcal{D}^{\eta}(p)$	Hill number (or diversity index of order η)
$\mathcal{D}(x)$	Discriminant curve
$\mathcal{D}(a)$	Diagonal matrix $\mathcal{D}(a) = \operatorname{diag}(a)$ with elements $(a_1, \ldots, a_n)$
DTS_i	Duration-times-spread factor of bond i
$\mathrm{DTS}(w)$	Duration-times-spread factor of portfolio w
$\delta\mathrm{D}$	Deuterium isotope ratio ($^2\mathrm{H}/^1\mathrm{H}$)
$\delta^{13}\mathrm{C}$	Carbon isotope ratio ($^{13}\mathrm{C}/^{12}\mathrm{C}$)
$\delta^{18}\mathrm{O}$	Oxygen isotope ratio ($^{18}\mathrm{O}/^{16}\mathrm{O}$)

$\mathbf{e}_i$	The value of the vector is 1 for the row i and 0 elsewhere
$\exp(A)$	Exponential of the matrix A
$\mathbb{E}[X]$	Mathematical expectation of the random variable X
$\mathcal{E}$	Set of edges in a graph
$\mathcal{E}(\lambda)$	Exponential probability distribution with parameter λ
$\mathcal{E}$	Energy flux (or radiation flux)
$\mathcal{EF}$	Emission factor
ε	Emissivity

$f(x)$	Probability density function
F	Climate forcing
F_{solar}	Incoming solar radiation (default value $= 242.82\,\mathrm{W/m^2}$)
$\mathbf{F}(x)$	Cumulative distribution function
$\mathbf{F}^{-1}(\alpha)$	Quantile function
$\mathcal{F}$	Vector of risk factors $(\mathcal{F}_1, \ldots, \mathcal{F}_m)$
$\mathcal{F}_j$	Risk factor j
$\mathcal{F}_t$	Filtration

$\boldsymbol{g}$	Greenium
$\mathcal{G}$	Greenness measure
$\mathcal{G}(p)$	Geometric distribution with parameter p
$\mathcal{G} = (\mathcal{V}, \mathcal{E})$	Graph with vertices $\mathcal{V}$ and edges $\mathcal{E}$
$\mathcal{GI}$	Green intensity
$\mathcal{G}ini$	Gini coefficient
$\mathcal{GRS}$	Green revenue share
γ	Risk-tolerance coefficient
$\bar{\gamma}$	Risk-aversion coefficient
γ_1	Skewness
γ_2	Excess kurtosis

$\mathcal{H}(p)$	Herfindahl index

i	Asset (or component) i
I_n	Identity matrix of dimension n
$\mathcal{I}(p)$	Shannon entropy of the distribution p
$\mathcal{I}(X)$	Shannon entropy of the random variable X
$\mathcal{I}(X,Y)$	Shannon entropy of the random vector (X,Y)
$\mathcal{I}(X \cap Y)$	Mutual information of the random vector (X,Y)
$\mathcal{I}^\star(p)$	Shannon diversity index
$\mathcal{IG}(\mu,\lambda)$	Inverse Gaussian distribution with parameters μ and λ
$\mathrm{IR}(w \mid b)$	Information ratio of portfolio w with respect to the benchmark b
$\mathbb{K}$	State space $(1,\dots,K)$
$\boldsymbol{\kappa}_j^{\mathrm{up}}$	Upstreamness index of sector j
$\boldsymbol{\kappa}_j^{\mathrm{down}}$	Downstreamness index of sector j
$\ln(A)$	Logarithm of the matrix A
$\boldsymbol{\ell}(\theta)$	Log-likelihood function with θ the vector of parameters to estimate
ℓ_t	Log-likelihood function for the observation t
$\mathcal{L}(x;\lambda)$	Lagrange function, whose Lagrange multiplier is λ
$\mathcal{L}$	Leontief inverse matrix
$\tilde{\mathcal{L}}$	Dual inverse matrix (or upstream multiplier matrix)
$\check{\mathcal{L}}$	Downstream multiplier matrix
$\lambda(p)$	Simpson index
m	m-score (or min-max score)
$\mathcal{M}ap$	Map function
MD_i	Modified duration of bond i
$\mathrm{MD}(w)$	Modified duration of portfolio w
μ	Vector of expected returns $(\mu_1,\dots,\mu_n)$
μ_i	Expected return of asset i
μ_m	Expected return of the market portfolio
$\hat{\mu}$	Empirical mean
$\mu(w)$	Expected return of portfolio w
$\mu(X)$	Mean of the random vector X
$\mu_m(X)$	m-th centered moment of the random vector X
$\mu'_m(X)$	m-th moment of the random vector X
n_S	Number of scenarios or simulations
$\mathcal{N}(\mu,\sigma^2)$	Normal distribution with mean μ and standard deviation σ
$\mathcal{N}(\mu,\Sigma)$	Multivariate normal distribution with mean μ and covariance matrix Σ
$\mathcal{NRD}(\mu_x,\sigma_x^2,\mu_y,\sigma_y^2)$	Ratio distribution of $Z=X/Y$ where $X \sim \mathcal{N}(\mu_x,\sigma_x^2)$ and $Y \sim \mathcal{N}(\mu_y,\sigma_y^2)$
Ω	Covariance matrix of risk factors
$p(k)$	Probability mass function of an integer-valued random variable
P	Markov transition matrix
$P(t)$	Transition matrix at time t
$P(\Sigma)$	Cholesky decomposition of Σ
$\mathcal{P}(\lambda)$	Poisson distribution with parameter λ
$\mathcal{P}(x)$	Performance curve
PPI	Producer price index
π	Vector of risk premia
$\pi^\star$	Stationary distribution
π_{cpi}	Inflation rate (consumer price index)
π_m	Market risk premium
π_n^-	1-diversity distribution

π_n^+	n-diversity distribution
π_{ppi}	Inflation rate (producer price index)
$\phi(x)$	Probability density function of the standardized normal distribution
$\phi_2(x_1, x_2; \rho)$	Probability density function of the bivariate normal distribution with correlation ρ
$\phi_n(x; \Sigma)$	Probability density function of the multivariate normal distribution with covariance matrix Σ
$\Phi(x)$	Cumulative density function of the standardized normal distribution
$\Phi^{-1}(\alpha)$	Inverse of the cdf of the standardized normal distribution
$\Phi_2(x_1, x_2; \rho)$	Cumulative density function of the bivariate normal distribution with correlation ρ
$\Phi_n(x; \Sigma)$	Cumulative density function of the multivariate normal distribution with covariance matrix Σ
q	q-score (or quantile score)
$\mathcal{QF}(x; Q, R, c)$	Quadratic form
r	Radius
r	Return of the risk-free asset
R	Vector of asset returns $(R_1, \ldots, R_n)$
R_i	Return of asset i
$R_{i,t}$	Return of asset i at time t
$R_{m,t}$	Return of the market portfolio at time t
$R(w)$	Return of portfolio w
$\mathcal{R}(t)$	Rating of the entity at time t
$\mathcal{R}(w)$	Risk measure of portfolio w
$\boldsymbol{\mathcal{R}}$	Reduction rate of carbon emissions
$\boldsymbol{\mathcal{R}}(w \mid b)$	Carbon footprint reduction rate of portfolio w wrt benchmark b
$\mathfrak{R}^2$	Coefficient of determination
$\mathcal{RN}(\mu, \sigma^2)$	Reciprocal normal distribution of $Y = 1/X$ where $X \sim \mathcal{N}(\mu, \sigma^2)$
ρ (or $\mathbb{C}$)	Correlation matrix of asset returns
$\rho_{i,j}$	Correlation between asset returns i and j
$\rho(x, y)$	Correlation between portfolios x and y
s	Credit spread
$\boldsymbol{s}_j$	Mapping vector of sector j
$\boldsymbol{\mathcal{S}ector}_j$	Sector j
$\boldsymbol{\mathcal{S}}$	ESG score
$\boldsymbol{\mathcal{S}}^\star$	Minimum ESG score or threshold
S_0	Total solar irradiance or solar constant (default value $= 1\,368\,\mathrm{W/m}^2$)
$\mathcal{S}(x)$	Selection curve
$\boldsymbol{\mathcal{SC}}_1$	Scope 1
$\boldsymbol{\mathcal{SC}}_2$	Scope 2
$\boldsymbol{\mathcal{SC}}_3^{\mathrm{up}}$	Upstream scope 3
$\boldsymbol{\mathcal{SC}}_3^{\mathrm{down}}$	Downstream scope 3
$\boldsymbol{\mathcal{SC}}_3$	Scope 3 ($= \boldsymbol{\mathcal{SC}}_3^{\mathrm{up}} + \boldsymbol{\mathcal{SC}}_3^{\mathrm{down}}$)
$\boldsymbol{\mathcal{SC}}_{1-2}$	Scope 1 and 2
$\boldsymbol{\mathcal{SC}}_{1-3}^{\mathrm{up}}$	Scope 1, 2 and upstream 3 ($= \boldsymbol{\mathcal{SC}}_1 + \boldsymbol{\mathcal{SC}}_2 + \boldsymbol{\mathcal{SC}}_3^{\mathrm{up}}$)
$\boldsymbol{\mathcal{SC}}_{1-3}$	Scope 1, 2 and 3
$\mathrm{SR}(w \mid r)$	Sharpe ratio of portfolio w when the risk-free rate is equal to r
σ	Stefan-Boltzmann constant (default value $= 5.67 \times 10^{-8}\,\mathrm{W/m}^2\,\mathrm{K}^{-4}$)
σ_i	Volatility of asset i
σ_m	Volatility of the market portfolio
$\tilde{\sigma}_i$	Idiosyncratic volatility of asset i
$\hat{\sigma}$	Empirical volatility
$\sigma(w)$	Volatility of portfolio w

$\sigma\left(w \mid b\right)$	Tracking error volatility of portfolio w wrt benchmark b
$\sigma_{\mathrm{AS}}\left(w \mid b\right)$	Active share active risk of portfolio w wrt benchmark b
$\sigma_{\mathrm{MD}}\left(w \mid b\right)$	Duration active risk of portfolio w wrt benchmark b
$\sigma_{\mathrm{DTS}}\left(w \mid b\right)$	DTS active risk of portfolio w wrt benchmark b
$\sigma\left(X\right)$	Standard deviation of the random variable X
Σ	Covariance matrix
$\hat{\Sigma}$	Empirical covariance matrix
t_{ν}	Student's t distribution with ν degrees of freedom
$t_n\left(\Sigma, \nu\right)$	Multivariate t distribution with ν degrees of freedom and covariance matrix Σ
$\mathrm{trace}\left(A\right)$	Trace of the matrix A
$\mathbf{T}\left(x; \nu\right)$	Cumulative density function of the Student's t distribution with ν degrees of freedom
$\mathbf{T}^{-1}\left(\alpha; \nu\right)$	Inverse of the cdf of the Student's t distribution with ν degrees of freedom
$\mathbf{T}_n\left(x; \Sigma, \nu\right)$	Cumulative density function of the t distribution with parameters Σ and ν
$\mathbf{T}_2\left(x_1, x_2; \rho, \nu\right)$	Cumulative density function of the bivariate t distribution with parameters ρ and ν
$\mathcal{T}\left(u\right)$	Matrix $\mathcal{T}\left(u\right) = uu^{\top}$ of dimension $n \times n$ where a is a $n \times 1$ vector
$\mathcal{T} = \left(\mathcal{V}, \mathcal{E}\right)$	Tree with vertices $\mathcal{V}$ and edges $\mathcal{E}$
$\mathcal{T}$	Temperature
$\mathcal{T}_a$	Air/atmospheric temperature
$\mathcal{T}_e$	Effective temperature (default value $= degK 255.81$ or $-17.34°\mathrm{C}$)
$\mathcal{T}_s$	Earth's surface temperature
$\mathcal{TS}$	Temperature score
$\boldsymbol{\tau}$	Hitting time
θ	Vector of parameters
$\hat{\theta}$	Estimator of θ
$\mathcal{U}_{[a,b]}$	Uniform distribution between a and b
$\boldsymbol{\mathcal{U}}\left(W\right)$	Utility function of the wealth W
$\mathrm{var}\left(X\right)$	Variance of the random variable X
$\mathcal{V}$	Set of vertices in a graph
w	Vector of weights $\left(w_1, \ldots, w_n\right)$ for portfolio w
$w^{\star}$	Mean-variance optimized portfolio
w^{*}	Tangency portfolio
w_i	Weight of asset i in portfolio w
w_{gmv}	Global minimum variance portfolio
w_m	Market portfolio
W	Wealth
x^{+}	Maximum value between x and 0
X	Random variable
$X_{i:n}$	i^{th} order statistic of a sample of size n
y	Yield to maturity
z	Altitude
z	z-score (or Gaussian score)

Other scientific conventions

YYYY-MM-DD	We use the international standard date notation where YYYY is the year in the usual Gregorian calendar, MM is the month of the year between 01 (January) and 12 (December), and DD is the day of the month between 01 and 31.
BP	Before present (1950)
Kyr/kyr/ka	1 000 years
Myr/myr/Ma	1 000 000 years
Gyr/byr/Ga	1 000 000 000 years
USD (or \$)	US dollar
EUR (or €)	Euro
KUSD	One thousand dollars
\$1 mn/bn/tn	One million/billion/trillion dollars
%	Percent or 0.01
‰	Per mil or 0.1%
bp	Basis point or 0.01%
ppm	Part per million
ppmv	Part per million by volume
ppb	Part per billion
ppbv	Part per billion by volume
H_2O	Water vapor
CO_2	Carbon dioxide
CH_4	Methane
N_2O	Nitrous oxide
CO_2e	Carbon dioxide equivalent
Hz	Frequency (Hertz or s^{-1})
J	Energy (Joule or $m^2\,kg\,s^{-2}$ or $N\,m$)
K	Temperature (Kelvin)
kg	Mass (kilogram)
m	Length (meter)
m^2	Area (square meter)
$m\,s^{-2}$	Acceleration (meter per square second)
N	Force (Newton or $m\,kg\,s^{-2}$)
Pa	Pressure (Pascal or $m^{-1}\,kg\,s^{-2}$ or $N\,m^{-2}$)
s	Time (second)
W	Power (Watt or $m^2\,kg\,s^{-3}$ or $J\,s^{-1}$)
$W\,m^{-2}$	Irradiance (Watt per square meter or $kg\,s^{-3}$)
Ton	Imperial unit of weight equivalent to 1 016.047 kilograms
Tonne	Metric unit of weight equivalent to 1 000 kilograms (also known as a metric ton)
gCO_2e	One gram of CO_2e
$kgCO_2e$	One kilogram of CO_2e (= 1 000 gCO_2e)
tCO_2e	One tonne of CO_2e (= 1 000 $kgCO_2e$)
$ktCO_2e$	One kilotonne of CO_2e (= 1 000 tCO_2e)
$MtCO_2e$	One megatonne of CO_2e (= 10^6 tCO_2e)
$GtCO_2e$	One gigatonne of CO_2e (= 10^9 tCO_2e)
$gCO_2e/\$$	One gram of CO_2e per one dollar
$tCO_2e/\$ \text{ mn}$	One tonne of CO_2e per one million of dollar

Abbreviations

ACPR Autorité de contrôle prudentiel et de résolution
AGM Annual general meeting
AI Artificial intelligence
AMOC Atlantic meridional overturning circulation
APS Announced pledges scenario
AUM Assets under management

BAU Business as usual
BCBS Basel committee on banking supervision
BECCS Bio-energy carbon capture and storage
BIS Bank for international settlements
BMR EU benchmark regulation
BoE Bank of England

CAPM Capital asset pricing model
CAT Cap-and-trade
CBD Convention on biological diversity
CBI Climate bonds initiative
CBIRC China banking and insurance regulatory commission
CCF Corporate carbon footprint
CCUS Carbon capture, use, and storage
CDP Carbon disclosure project
CDR Carbon dioxide removal
CDSB Climate disclosure standards board
CE Carbon emissions
CEO Chief executive officer
Ceres Coalition for environmentally responsible economies
CFP Corporate financial performance
CI Carbon intensity
COP Conference of the parties
CPI Consumer price index
CRA Credit rating agency
CSDDD Corporate sustainability due diligence directive

CSP Corporate social performance
CSR Corporate social responsibility
CSRD Corporate sustainability reporting directive
CTB Climate transition benchmark

DAC Direct air capture
DACCS Direct air carbon capture and storage
DDQ Due diligence questionnaire
DICE Dynamic integrated climate-economy model
DNSH Do no significant harm
DTS Duration-times-spread factor

EBA European banking authority
EC European Commission
ECB European central bank
ECS Equilibrium climate sensitivity
EDGAR Emission Database for Global Atmospheric Research
EEIO Environmentally-extended input-output model
EET European ESG template
EFDB Emission factor database
EFRAG European financial reporting advisory group
EIB European investment bank
EIOPA European insurance and occupational pensions authority
ENCORE Exploring natural capital opportunities, risks and exposure
ENSO El Niño Southern Oscillation
EPICA European project for ice coring in Antarctica
ERA Extra-financial rating agency
ERF Effective radiative forcing
ESAs European supervisory authorities
ESFS European system of financial supervision

ESG Environmental, social and governance
ESM European stability mechanism
ESMA European securities and markets authority
ETS Emissions trading scheme
EUGBR EU green bonds regulation
Eurosif European sustainable investment forum
EUTR European Union taxonomy regulation for sustainable activities

FAO Food and Agriculture Organization
FAs Financial advisers
FDIC Federal deposit insurance corporation
FIO Federal insurance office
FMP Financial market participant
FRB Board of governors of the federal reserve system
FSB Financial stability board
FSOC Financial stability oversight council
FUND Climate framework for uncertainty, negotiation, and distribution

GB Green bond
GBP Green bond principles
GC Global Compact
GCM General circulation model
GEVA Greenhouse gas emissions per unit of value added
GFANZ Glasgow financial alliance for net zero
GHG Greenhouse gaz
GICS Global industry classification standard
GIIN Global impact investing network
GISP Greenland ice sheet project
GLP Green loans principles
GMO Genetically modified organism
GMV Global minimum variance portfolio
GQE Green quantitative easing
GRI Global reporting initiative
GRIP Greenland ice core project
GRS Green revenue share
GSAT Global mean surface air temperature
GSIA Global sustainable investment alliance
GSIR Global sustainable investment review
GSS Green, social and sustainability bonds
GTAP Global trade analysis project
GTB Green-to-brown ratio
GTP global temperature potential
GTS Geologic time scale

GWP Global warming potential

HCIS High climate impact sector
HKMA Hong Kong monetary authority
HLEG High-level expert group on sustainable finance

IAIS International association of insurance supervisors
IAM Integrated assessment model
IASB International accounting standards board
IBAT Integrated Biodiversity Assessment Tool
ICMA International capital market association
IDD Insurance distribution directive
IEA International energy agency
IFC International finance corporation
IFRS International financial reporting standards
IIASA International institute for applied systems analysis
IIRC International integrated reporting council
ILO International labour organization
IMF International monetary fund
IOPS International organisation of pensions supervisors
IOSCO International organization of securities commissions
IOT Input-output table
IPCC Intergovernmental panel on climate change
ISSB International sustainabilityÂăstandards board
ITR Implied temperature rating
IUCN International union for conservation of nature

JSIF Japan sustainable investment forum

KF Kalman filter
KPI Key performance indicator

LCA Life cycle assessment
LP Linear programming
LPI Living planet index
LSEG London Stock Exchange Group

MAC Marginal abatement cost
MACC Marginal abatement cost curve
MAS Monetary authority of Singapore
MD Modified duration

MiFID Markets in financial instruments directive

MRIO Multi-regional input-output model

MSA Mean species abundance

MVO Mean-variance optimized portfolio

NACE Nomenclature statistique des activités économiques dans la Communauté Européenne

NCA National competent authority

NDC Nationally determined contribution

NET Negative emissions technology

NFRD Non-financial reporting directive

NGFS Network of central banks and supervisors for greening the financial system

NGO Non-governmental organization

NGRIP North Greenland ice core project

NICE Nested inequalities climate-economy model

NIR National inventory report

NLP Natural langage processing

NRSRO Nationally recognized statistical rating organization

NZAM Net zero asset managers initiative

NZAOA Net zero asset owner alliance

NZBA Net zero banking alliance

NZE Net zero emissions scenario

NZFSPA Net zero financial service providers alliance

NZIA Net zero insurance alliance

NZICI Net zero investment consultants initiative

OCC Office of the comptroller of the currency

OCR Office of credit ratings

OECD Organisation for economic cooperation and development

OLR Outgoing longwave radiation

OLS Ordinary least squares

OPS One planet summit

OPSWF One planet sovereign wealth fund

ORSE Observatoire de la responsabilité sociétale des entreprises

PAB Paris aligned benchmark

PAGE Policy analysis of the greenhouse gas effect

PAI Principal adverse impact

PAII Paris aligned investment initiative

PBOC People's Bank of China

PCAF Partnership for carbon accounting financials

PCF Product carbon footprint

PDB Pee Dee Belemnite

PPI Producer price index

PRI Principles for responsible investment

QP Quadratic programming

RCP Representative concentration pathway

RIAA Responsible investment association Australasia

RICE Regional integrated climate-economy model

RLI Red list index

RLS Recursive least squares

RTS Regulatory technical standard

SASB Sustainability accounting standards board

SB Social bond

SBP Social bond principles

SBTi Science-based targets initiative

SCC Social cost of carbon

SDA Sectoral decarbonisation approach

SDGs Sustainable development goals

SDTF Sudan divestment task force

SEC Securities and exchange commission

SF Sustainable finance

SFDR Sustainable finance disclosure reporting

SIB Social impact bond

SIF Sustainable investment forum

SLB sustainability-linked bond

SMOW Standard mean ocean water

SPO Second party opinion

SREP Supervisory review and evaluation process

SRI Socially responsible investing

SRP Supervisory review process

SSP Shared socioeconomic pathway

STAR Species Threat Abatement and Restoration metric

STEPS Stated policies scenario

SWF Sovereign wealth fund

TCFD Task force on climate-related financial disclosures

TEG Technical expert group on sustainable finance

TFP Total factor productivity

TSI Total solar irradiance

UK SIF UK sustainable investment and finance association

UN United Nations

UN PRI UN principles for responsible investment

UNECE United Nations Economic Commission for Europe

UNEP United Nations environment program

UNEP FI United Nations environment program finance initiative

UNFCCC UN framework convention on climate change

UNICEF United Nations children's fund

US SIF Forum for sustainable & responsible investment

VBDO Vereniging van beleggers voor duurzame ontwikkeling

VCM Voluntary carbon market

WACI Weighted average carbon intensity

WAIS West Antarctic ice sheet project

WBCSD World business council for sustainable development

WCW Who cares wins

WDPA World Database on Protected Areas

WHO World health organization

WIOD World input-output database

WMO World meteorological organization

WRI World resources institute

Chapter 1

Introduction

In this chapter, we first define the concept of sustainable finance (SF) and discuss its historical origins, particularly the motivations of responsible investors. We also present the ecosystem of responsible investing and the corresponding regulatory framework. Finally, we provide some figures on the sustainable finance market.

1.1 Definition

The European Commission defines the concept of sustainable finance as follows[1]:

> "*Sustainable finance refers to the process of taking* **environmental, social and governance (ESG) considerations** *into account when making investment decisions in the financial sector, leading to more long-term investments in sustainable economic activities and projects.* **Environmental considerations** *might include climate change mitigation and adaptation, as well as the environment more broadly, for instance the preservation of biodiversity, pollution prevention and the circular economy.* **Social considerations** *could refer to issues of inequality, inclusiveness, labour relations, investment in human capital and communities, as well as human rights issues. The* **governance** *of public and private institutions — including management structures, employee relations and executive remuneration — plays a fundamental role in ensuring the inclusion of social and environmental considerations in the decision-making process.*"

In this definition, the EC also introduces the concept of ESG (Environmental, Social and Governance), which is very popular among asset owners and managers. For instance, in contrast to business-as-usual (BAU) or traditional investing, the goal of ESG investing is to take into account extra-financial analysis when performing asset selection. Nevertheless, the frontier between SF and ESG is not very clear. This is also the case with other terms that are frequently used such as responsible investment (RI), sustainable investing (SI) and socially responsible investing (SRI). We report here some definitions we can find in financial textbooks:

- **Responsible investment** is an approach to investment that explicitly acknowledges the relevance to the investor of environmental, social and governance factors, and of the long-term health of the market as a whole.

- **Sustainable investing** is an investment approach that considers environmental, social and governance factors in portfolio selection.

---

[1]See the webpage https://finance.ec.europa.eu/sustainable-finance/overview-sustainable-finance_en.

- **Socially responsible investing** is an investment strategy that is considered socially responsible, because it invests in companies that have ethical practices.

- **ESG** refers to the factors that measure the sustainability of an investment.

In fact, it is really difficult to make the difference between all these concepts, because they both encompass the same underlying idea. Therefore, we can consider them as the same subject (Figure 1.1). We can complete this list by other expressions such as green finance, climate finance, blue finance, etc. Generally, the term green finance is reserved for the environmental pillar, whereas blue finance is an emerging area in climate finance and concerns the ocean economy (IFC, 2022).

Figure 1.1: Many words, one concept

1.2   Short history of responsible and ethical investing

From an historical point of view, we observe three stages. In the 1990s and 2000s, the word "*sustainable finance*" is not really used. The term "*responsible investment*" is preferred because of the ethical considerations of some final investors and asset owners. In the 2010s, "*ESG investing*" takes the lead because it gains momentum in the asset management industry. Moreover, ESG rating agencies adopted the break down of the extra-financial information into environmental, social and governance pillars. Finally, the concept of ESG spreads across all financial actors and sectors (e.g., corporates, banks, regulators, policy makers and central banks). In this context, the investment side is not only concerned, but it also affects financing, regulation, society and public policies. Therefore, it is better to use the term "*sustainable finance*", which is more generic than responsible or ESG investing.

The previous evolution (responsible investment → ESG → sustainable finance) can be explained by the history of ethical investment. Religious motivations explained the first examples of responsible (or faith-based) investing. In 1758, the Quaker Philadelphia yearly meeting prohibited members from participating in the slave trade (buying or selling humans). They are followed by religious groups (e.g., Muslims, or Methodists), which invited people and members to avoid investing in companies linked to weapons, tobacco, alcohol, or gambling. According to Beabout and Schmiesing (2003), the first SRI mutual fund (Pioneer Fund) was created in 1928 by Philip Carret for Evangelical

Protestants. The years 1930-1960 saw the emergence of several doctrines about responsible investing. In particular, a number of corporate scandals lead to more focus on governance issues. During the Vietnam War, shareholders organized resolutions against the production of napalm[2] and Agent Orange[3]. Therefore, we observe the development of engagement policies besides exclusion policies. In 1971, two members of the United Methodist Church (Luther Tyson and Jack Corbett) and the portfolio manager Tony Brown launched the Pax World fund, which may also claim to be the first sustainable mutual fund in the United States. Indeed, the strategy of the fund mixed both financial and social criteria. This is a step forward since the fund considers selection screening and not only exclusion screening. Moskowitz (1972) published a first list of socially responsible stocks, including Chase Manhattan, Johnson Products, Levi Strauss, New York Times, Whirlpool and Xerox. These stocks are challenged by Vance (1975), who found that they have largely underperformed the Dow Jones from 1972 to 1975. The concept of "*sin stocks*" was born, and the relationship between responsible investment and profitability led to many academic publications on these topics. This first period of sustainable finance may be summarized as follows:

> "*Do no harm. That is the central concept of traditional faith-based investing and, to some degree, the central concept of socially responsible investing: Avoiding products or industries that conflict with a set of moral values.*" (Townsend, 2020, page 2).

The question of moral values is also the main factor explaining the development of corporate social responsibility (CSR). This theory begins with the publication of "*Social Responsibilities of the Businessman*" by Bowen (1953). In this book, the author analyzed the responsibilities to society that companies are expected to assume. Considered as the "*Father of Corporate Social Responsibility*" (Carroll, 1999), Howard Bowen assumed that "*CSR can help business reach the goals of social justice and economic prosperity by creating welfare for a broad range of social groups, beyond the corporations and their shareholders.*" Regarded as an alternative to socialism and pure capitalism, CSR is rejected by neoclassical economists. One of the most famous opponents is Milton Friedman:

> "*There is one and only one social responsibility of business — to use its resources and engage in activities designed to increase its profits so long as it stays within the rules of the game, which is to say, engages in open and free competition without deception or fraud.*" (Friedman, 1962).

In particular, his article published in New York Times (Friedman, 1970) has a big impact on the shareholder vs. stakeholder debate. The stakeholder theory suggests that the real success of a company lies in satisfying all its stakeholders, not just the shareholders (Freeman, 2004). The stakeholder ecosystem involves customers, suppliers, employees, local communities, governmental agencies, financiers, and others. In this theory, each business entity creates, and sometimes destroys, value for each stakeholder group. Again, many academic research have been published on this topic, in particular how to define corporate social performance (CSP), and its relationship with corporate financial performance (CFP). Nevertheless, even if the debate is still raging, the stakeholder theory has profoundly changed the vision for the business. Indeed, there's today a wide consensus that business objective should not just be about profit maximization. An example is the Global Compact (GC) initiative created by the UN Secretary-General Kofi Annan on July 2000. It is a voluntary

---

[2]In 1968, the Medical Committee for Human Rights acquired shares in Dow Chemical in order to prohibit sales of napalm.

[3]Agent Orange is a mixture of two herbicides. It was used by the US military to defoliate forests and terrorize populations in South Vietnam (Townsend, 2020, page 3).

initiative based on CEO commitments to implement a set of human rights, labour, environmental, and anti-corruption principles[4]. The 10 principles are:

- Human rights

    1. Businesses should support and respect the protection of internationally proclaimed human rights; and

    2. Make sure that they are not complicit in human rights abuses.

- Labour

    3. Businesses should uphold the freedom of association and the effective recognition of the right to collective bargaining;

    4. The elimination of all forms of forced and compulsory labour;

    5. The effective abolition of child labour; and

    6. The elimination of discrimination in respect of employment and occupation.

- Environment

    7. Businesses should support a precautionary approach to environmental challenges;

    8. Undertake initiatives to promote greater environmental responsibility; and

    9. Encourage the development and diffusion of environmentally friendly technologies.

- Anti-corruption

    10. Businesses should work against corruption in all its forms, including extortion and bribery.

From 2004 to 2008, the UN Global Compact, the International Finance Corporation (IFC) and the Swiss government sponsored a series of annual conferences for investment professionals, asset managers, and financial institutions to develop guidelines and recommendations on how to better integrate environmental, social and corporate governance issues. The term ESG was first coined in the 2004 conference report "*Who Cares Wins — Connecting Financial Markets to a Changing World*" (WCW, 2004) and was popularized by the next four reports[5].

Socially responsible investing does not only concern corporations, but also sovereigns and countries. For instance, the US Congress passed in 1986 the "*Comprehensive Anti-Apartheid Act*", banning new investment in South Africa. Similarly, the Sudan Divestment Task Force (SDTF) was formed in 2005 to coordinate and provide resources for the Sudan divestment campaign in response to the genocide occurring in the Darfur region. The US "*Sudan Accountability and Divestment Act*"came into force in December 2007. It authorized a state or local governments to divest assets in companies that are conducting business operations in Sudan that include power production activities, mineral extraction activities, oil-related activities, or the production of military equipment.

---

[4]The Global compact initiative takes its root in the code of conduct for companies developed in 1977 by Leon Sullivan, a clergyman and civil rights leader. The original Sullivan Principles consisted of seven requirements a corporation operating in South Africa must satisfy. They were a response to apartheid and an alternative to complete divestment, which was perceived as a costly strategy (Grossman and Sharpe, 1986; Rudd, 1979).

[5]The titles of the four conference reports are "*Investing for Long-Term Value — Integrating Environmental, Social and Governance Value Drivers in Asset Management and Financial Research*" (2005), "*Communicating ESG Value Drivers at the Company-Investor Interface*" (2006), "*New Frontiers in Emerging Markets Investment*" (2007) and "*Future Proof? Embedding Environmental, Social and Governance Issues in Investment Markets*" (2008).

According to Townsend (2020), the current concept of sustainable finance mixes "*the traditional North American model for socially responsible investing, and ESG, which first took hold in Europe*". It is true that the Who Cares Wins (WCW) conferences had rather a European orientation, with participants mainly coming from European asset owners and managers, especially the 2005 conference (WCW, 2005). While SRI is more an exclusion and qualitative process at its inception in North America, ESG is a best-in-class and quantitative process when it is implemented at the beginning of the 2000s. The growth of ESG data and ESG rating agencies[6] largely explains this shift. One reason is "*the strong intellectual and legal debate on the relationship between fiduciary duty and issues of sustainability*" (Townsend, 2020, page 6). In 2005, UNEP invited the law firm Freshfields Bruckhaus Deringer to produce a report about the legal use to integrate ESG issues by pension funds, insurance companies and asset managers. The objective of the report was to answer the following question:

> "*Is the integration of environmental, social and governance issues into investment policy (including asset allocation, portfolio construction and stock-picking or bond-picking) voluntarily permitted, legally required or hampered by law and regulation; primarily as regards public and private pension funds, secondarily as regards insurance company reserves and mutual funds?*" (Freshfields Bruckhaus Deringer, 2005, page 6).

The 154-pages report analyzed the legal framework for institutions in Australia, Canada, France, Germany, Italy, Japan, Spain, the UK and the US. While the analysis is very technical and the results depends on the jurisdiction, the report concluded that integrating ESG issues is consistent with fiduciary duty if ESG factors impact the investment value and long-term risks. In this context, we observe an increasing change of European institutional investors, who consider that their fiduciary duties require them to incorporate ESG into investment analysis. As institutional investors are sophisticated investors and they base their decisions on an in-depth quantitative analysis, this has implied to transform the original qualitative approach based on discretionary exclusions to a more systematic model based on extra-financial quantitative data.

A second reason that explains the shift from a qualitative-oriented SRI to a quantitative-oriented ESG is the climate change factor. In response to global warming, the Intergovernmental Panel on Climate Change (IPCC) was established in 1988 by the World Meteorological Organization (WMO) and United Nations Environment Programme (UNEP). The IPCC prepares assessment reports about knowledge on climate change. The first assessment report (AR1) was published in March 1990, whereas the synthesis of the last assessment report (AR6) is expected in December 2022. These reports are extensively used in the United Nations Climate Change conferences. In June 1992, the Earth Summit held in Rio de Janeiro produced two important legal agreements: the Convention on Biological Diversity (CBD) and the UN Framework Convention on Climate Change (UNFCCC). The objective of this international treaty is to reduce environmental impacts across the globe. The implementation of the UNFCCC to address global warming is an on-going process. For instance, the Kyoto Protocol negotiated in 1997 and the Paris Agreement adopted in 2015 are certainly the two famous international treaties on climate change. On the investor side, the Coalition for environmentally responsible economies (Ceres) was founded in 1989 with the aim of changing corporate environmental practices. Following the Exxon Valdez oil spill, Ceres created the Valdez Principles[7]. In 2000, it also launched the Global Reporting Initiative (GRI) to standardize corporate disclosure on ESG issues.

---

[6]This topic will be elaborated in the next chapter.

[7]The 10 principles are (1) protection of the biosphere, (2) sustainable use of natural resources, (3) reduction and disposal of wastes, (4) energy conservation, (5) risk reduction, (6) safe products and services, (7) environmental restoration, (8) informing the public, (9) management commitment, and (10) audits and reports.

The last twenty years have strengthened the place of ESG in finance, not only on the investment side, but also on the financing side. Regulators are now involved, accounting standards have been developed, climate change is recognized as a key risk factor, controversies may harm corporate reputation, social pressure impacts corporate governance, etc. In the next chapters, we will extensively document the evolution of sustainable finance during the last period. The motivations to implement a socially responsible investment are now multiple. In Figure 1.2, we give some reasons. They can be classified into two groups. The first one (economic sustainability, moral values and social pressure) is related to the "*do not harm principle*", and can be applied to many situations or decisions. The second group, which includes financial performance, fiduciary duty and risk management, is related to investment principles. The underlying idea is that ESG risks have to be managed and can not be ignored within portfolio construction.

Figure 1.2: The raison d'être of sustainable finance

1.3   The ESG ecosystem

As we have just seen, the ESG landscape involves many financial actors. First, investors are in the center of the ecosystem. Generally, we distinguish asset owners and managers. Asset owners correspond to end-investors and include pension funds, institutional investors, sovereign wealth funds (SWF), insurance companies, endowments and foundations, family offices, retail investors, etc. On the contrary, the asset management industry manages funds for end-investors. In this category, we have mutual funds, hedge funds, private equity funds, infrastructure funds, third-party distributors, etc. We could also mention ESG index providers since they are essential for passive management. Asset managers act then as financial intermediary between the financial markets (e.g., equity and fixed-income markets) and the saving of households, companies and organizations. While asset owners and managers constitutes the investing side, banks and issuers form the financing side. Therefore, sustainable finance also concerns the emission of debt and the structuring of ESG products such as green bonds.

1.3.1 Sustainable investment forum

The sustainable investment forums (SIF) are membership-based sustainable and responsible investment organisations. They work to promote a broader adoption of sustainable and responsible investment practices and more generally for a broader adoption of sustainability matters into financial markets and the investment chain. They are organized by countries or regions. For instance, the European Sustainable Investment Forum (Eurosif) was launched in 2001 and groups together Forum per la Finanza Sostenibile (Italy), Forum Nachhaltige Geldanlagen (Germany), Forum pour l'Investissement Responsable (France), Foro de Inversión Sostenible (Spain), Sustainable Finance Ireland, Swiss Sustainable Finance (Switzerland) and UK Sustainable Investment and Finance Association. Other SIFs are the Responsible Investment Association Australasia[8] (RIAA), the Responsible Investment Association Canada (RIA Canada), The Forum for Sustainable & Responsible Investment (US SIF), the Dutch Association of Investors for Sustainable Development[9] (VBDO) and the the Japan Sustainable Investment Forum (JSIF). All these organizations are members of the Global Sustainable Investment Alliance (GSIA).

These forums have been created at different dates, reflecting the evolution of sustainable finance. For example, US SIF was founded in 1984 and is the oldest SIF. It is followed by RIA Canada in 1990, UK SIF in 1991 and VBDO in 1995. Most of European forums were established in 2001 (e.g., Germany, France, Italy). The main activities of these sustainable investment forums are public policy, education, training, research and promoting sustainable investing best practices. Founded in 2010, GSIA is in charge of aggregating responsible investment market data from its members in order to analyze the global sustainable investment market and the evolution of ESG trends. In particular, it publishes a biennial Global Sustainable Investment Review or GSIR (GSIA, 2013, 2015, 2017, 2019, 2021). The 2022 GSIR edition is expected mi-year 2023. We will extensively used these reports in Section 1.5 on page 38 when we will analyse the market of ESG investing.

Figure 1.3: 2018 & 2020 GSIA reports

1.3.2 Initiatives

In this section, we present the most relevant initiatives (PRI, Climate action 100+ and net zero alliances). We also list other initiatives that participate in the ESG ecosystem. Some of them will be detailed further in the next chapters.

---

[8] It groups together Australia and New Zealand.
[9] The Dutch name is Vereniging van Beleggers voor Duurzame Ontwikkeling.

Principles for responsible investment

In early 2005, the UN Secretary-General Kofi Annan invited a group of the world's largest institutional investors to join a process to develop the Principles for Responsible Investment[10] (PRI). A 20-person investor group drawn from institutions in 12 countries was supported by a 70-person group of experts from the investment industry, intergovernmental organisations and civil society. The PRI were launched in April 2006 at the New York Stock Exchange.

<div align="center">Box 1.1: PRI signatories' commitment</div>

"As institutional investors, we have a duty to act in the best long-term interests of our beneficiaries. In this fiduciary role, we believe that environmental, social, and corporate governance (ESG) issues can affect the performance of investment portfolios (to varying degrees across companies, sectors, regions, asset classes and through time).
We also recognise that applying these Principles may better align investors with broader objectives of society. Therefore, where consistent with our fiduciary responsibilities, we commit to the following:

- Principle 1: We will incorporate ESG issues into investment analysis and decision-making processes.

- Principle 2: We will be active owners and incorporate ESG issues into our ownership policies and practices.

- Principle 3: We will seek appropriate disclosure on ESG issues by the entities in which we invest.

- Principle 4: We will promote acceptance and implementation of the Principles within the investment industry.

- Principle 5: We will work together to enhance our effectiveness in implementing the Principles.

- Principle 6: We will each report on our activities and progress towards implementing the Principles.

The Principles for Responsible Investment were developed by an international group of institutional investors reflecting the increasing relevance of environmental, social and corporate governance issues to investment practices. The process was convened by the United Nations Secretary-General.
In signing the Principles, we as investors publicly commit to adopt and implement them, where consistent with our fiduciary responsibilities. We also commit to evaluate the effectiveness and improve the content of the Principles over time. We believe this will improve our ability to meet commitments to beneficiaries as well as better align our investment activities with the broader interests of society.
We encourage other investors to adopt the Principles."

Source: https://www.unpri.org.

---

[10]UN PRI and PRI are two interchangeable terms. For example, the website url is https://www.unpri.org. Nevertheless, PRI is the official term.

Signatories' commitment is reported in Box 1.1. The principles are voluntary and aspirational, and offer a set of possible actions for incorporating ESG issues into investment practice. For instance, here are some possible actions for Principle 1:

- Address ESG issues in investment policy statements;

- Support development of ESG-related tools, metrics, and analyses;

- Advocate ESG training for investment professionals;

- Etc.

Becoming a signatory requires to pay an annual fee[11], but there are no other formal requirements when the membership agreement is signed. Nevertheless, signatories are required to report on their responsible investment activities annually. The answers of the members, which form the transparency report, are public and available to anyone[12]. Since 2019, members must also fill in a climate transparency report, which contains specific indicators regarding the management of risks and opportunities related to climate change. These indicators are modelled on the disclosure framework of the Task Force on Climate-related Disclosures (TCFD). Based on the transparency report, PRI produces an assessment report for each member, which consists of a series of scores on several dimensions (from 0 to 100) and a rating system (from one to five stars). The assessment report is confidential, except if the member choose to make it public. It is also used by PRI to verify that signatories meet minimum requirements. If it is not the case, PRI engage with the member (one-on-one sessions, action plans, etc.). Delisting is a last resort if a signatory has not met the requirements after the two-year period. Since 2018, 165 signatories have been identified as not meeting the minimum requirements. PRI has delisted 5 signatories, and 23 other members of the 165 identified have been delisted on a voluntarily basis.

In Figure 1.4, we show the PRI growth. At the inception date, most of the 63 founding signatories were asset owners[13] with a few asset managers[14] and data providers[15]. They were mainly located in the US, Canada, UK, France[16], the Netherlands and the Nordics. As of September 2022, the PRI has 5 020 signatories, representing approximately \$121 trillion of assets under management (AUM). Investment managers is the most represented category (76%) followed by asset owners (14%) and service providers (10%). We observe a rapid evolution since 2015, and even an acceleration since 2021, especially in Asia and emerging markets.

Climate action 100+

Climate Action 100+ is an investor initiative to ensure the world's largest corporate greenhouse gas emitters take necessary action on climate change. It was formed in the wake of the 2015

---

[11]The 2022/23 fee goes from £478 to £14,222 depending of the size and the category (asset owners, investment managers and service providers) of the signatory.

[12]The reports from 2014 to 2020 are available at the webpage https://www.unpri.org/signatories/reporting-and-assessment/public-signatory-reports, whereas the more recent reports can be downloaded in the data portal (or PRI's central depository for signatories' reporting data): https://ctp.unpri.org/dataportalv2.

[13]The most important were AP2, BT Pension Scheme, CDC, CDPQ, CalPERS, CPPIB, ERAFP, FRR, NYCERS, NZSF, NGPF, PGGM, TIAA-CREF, UNJSPF and USS.

[14]The best known asset managers were ABN AMRO Asset Management, Aviva Investors, BNP PAM, Candriam, CAAM (now Amundi), Daiwa AM, Henderson Global Investors and Threadneedle AM.

[15]e.g., Ethix, Trucost and Vigeo.

[16]There are 8 founding signatories: BNP PAM, CDC, CAAM (now Amundi), ERAFP, FRR, Groupama AM, Macif Gestion and Vigeo.

Figure 1.4: PRI signatory growth 2006–2021
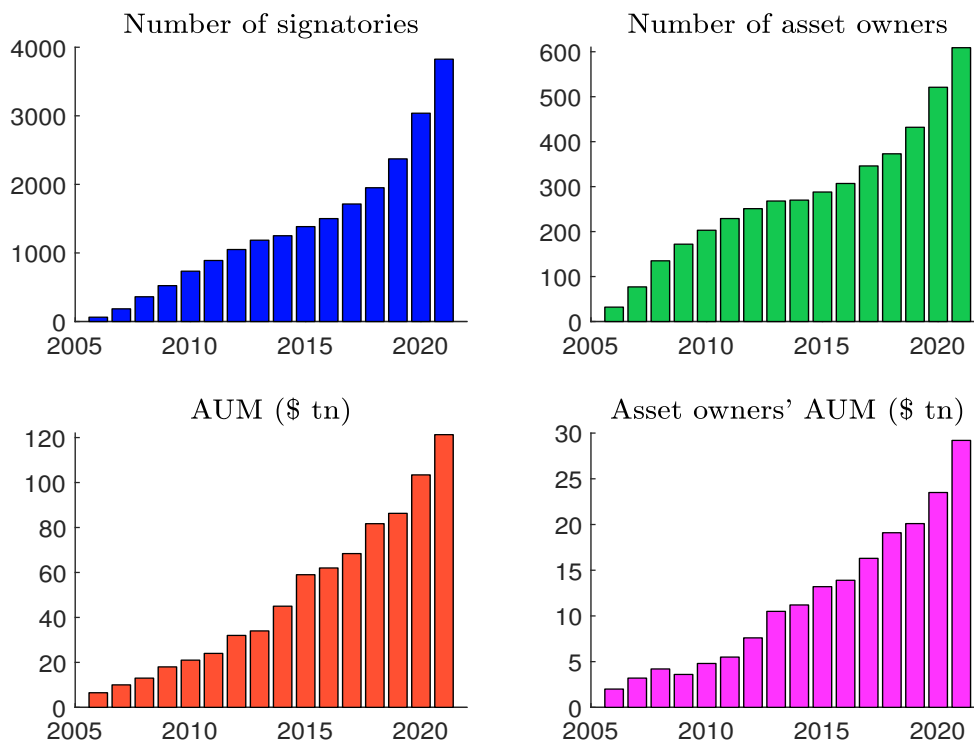
Paris Agreement, and is launched in December 2017. It is supported by 700 investors, responsible for over \$68 trillion in assets under management. Climate Action 100+ focuses on engagement, and coordinates the efforts of the investor signatories. In a nutshell, engagement corresponds to the active dialogue between the investor and the company by discussing sustainability risks and providing the investor's expectations of corporate behavior. The main objectives are improving the climate performance of the company, reducing GHG emissions across the value chain and ensuring transparent disclosure. The engagement process van de described as follows:

- Engagement with focus company executives and board members is spearheaded by a lead investor or investors, who work cooperatively with a number of collaborating investors and are supported by technical experts.

- When signing on to the initiative, investors are asked to nominate which focus companies they wish to engage with and whether this is as a lead investor or collaborating investor.

- Engagement takes several forms, e.g., holding meetings with companies, making a statement at a company AGM, supporting shareholder resolutions on climate change, voting for the removal of directors who have failed in their accountability of climate change risks.

Climate Action 100+ engagement focuses on 166 companies, accounting for up to 80% of global corporate industrial greenhouse gas emissions. The geographic breakdown is the following: 1.8% in Africa, 20.4% in Asia, 9.0% in Australasia, 33.5% in Europe, 32.3% in North America, and 3.0% in South America. The sector distribution is reported in Table 1.1, where we indicate the number of companies and the market capitalization[17]. For example, the 5 focus companies for the airlines

---

[17]The market capitalization is computed as of 15 December 2020.

sector are Air France, American Airlines, Delta Air Lines, Qantas Airways, United Airlines, the 12 focus companies for the automobiles sector are BMW, Ford, General Motors, Honda, Mercedez-Benz, Nissan, Renault, SAIC, Stellantis, Suzuki, Toyota, Volkswagen, etc.

Table 1.1: Sector breakdown of Climate Action 100+ engagement

Sector	Frequency		Market capitalisation	
	Number	in %	in $ bn	in %
Airlines	5	3.0%	57.7	0.6%
Automobiles	12	7.2%	737	7.1%
Cement	11	6.6%	181	1.8%
Chemicals	7	4.2%	287	2.8%
Coal mining	4	2.4%	68.5	0.7%
Consumer goods & services	12	7.2%	1 900	18.4%
Diversified mining	10	6.0%	484	4.7%
Electric utilities	30	18.1%	1 000	9.7%
Oil & gas	39	23.5%	3 700	35.8%
Oil & gas distribution	5	3.0%	160	1.5%
Other industrials	13	7.8%	1 100	10.6%
Other transportation	7	4.2%	501	4.8%
Paper	2	1.2%	33.6	0.3%
Shipping	1	0.6%	39	0.4%
Steel	8	4.8%	85	0.8%
Total	166	100.0%	10 334	100.0%

Source: https://www.climateaction100.org/whos-involved/companies.

Net-zero alliances

Net-zero emissions refers to a state in which the greenhouse gases going into the atmosphere are balanced by removal out of the atmosphere. This is a condition to stop global warming. According to IPCC (2018), global temperature increase needs to be limited to 1.5°C pre-industrial levels in order to mitigate the worst impacts of climate change and preserve a livable planet. Generally, we assume that net zero emissions must be achieved by 2050 IEA (2021), otherwise multiple tipping points could be triggered with irreversible impacts.

The concept of "*Net Zero Alliance*" starts with the launch of the Net Zero Asset Owner Alliance (NZAOA) in September 2019 under the umbrella of UNEP FI. In September 2022, the Alliance counts 74 members, accounting for $10.6 tn in AUM (UNEP, 2022). These members must satisfy a common protocol to target setting and reporting based on four components:

1. Engagement targets

   - Engage with 20 companies focusing on those with highest owned emissions or those responsible for combined 65% owned emissions in portfolio.

2. Sub-portfolio emission targets

   - 22 to 32% CO_2e reduction by 2025 (per IPCC 1.5°C SR scenarios);
   - 49 to 65% CO_2e reduction by 2030 (per IPCC 1.5°C SR scenarios);

- Cover portfolio scope $1+2$ emissions, tracking of scope 3, and use absolute or intensity-based reduction KPIs.

3. Sector targets

  - Use absolute or intensity-based reductions on all material sectors;
  - Scope 3 to be included wherever possible;
  - Sector specific intensity KPIs recommended.

4. Financing transition targets

  - Reporting progress on a climate-positive trend for all Alliance members internally to the Alliance;
  - Build solutions or enhance climate solution reporting.

For example, the targets[18] defined by Munich Re are the following: (1) concentrate on and engage with large contributors of financed emissions within the listed equities and corporate bond portfolio; (2) reduce the absolute emissions of listed equities, corporate bond and real estate portfolio by $25 - 29\%$ (scope $1+2$ emissions of investee companies) by 2025; (3) reduce emissions for listed equities and corporate bonds for thermal coal (-35%) and oil & gas (-25%); (4) double the renewable portfolio (equity and debt) from €1.6 bn to €3 bn.

In June 2020, UNFCCC launches the "*Race to Zero*" campaign, which have definitively accelerated the net zero commitments. For example, the Glasgow Financial Alliance for Net Zero (GFANZ) is created in April 2021 by Mark Carney[19] and the COP26 presidency to coordinate efforts across all sectors of the financial system to accelerate the transition to a net zero global economy[20]. GFANZ is an umbrella organisation covering seven net zero initiatives: NZAOA, the Net Zero Asset Managers initiative (NZAM), the Paris Aligned Investment Initiative (PAII), the Net Zero Banking Alliance (NZBA), the Net Zero Insurance Alliance (NZIA), Net Zero Financial Service Providers Alliance (NZFSPA) and the Net Zero Investment Consultants Initiative (NZICI).

Other initiatives

There are a growing list of initiatives that are related to ESG issues. Here are a few examples with respect to the three pillars:

- Environmental
  Asia Investor Group On Climate Change (AIGCC), Finance for Biodiversity Pledge, Finance for Tomorrow, Institutional Investors Group on Climate Change (IIGCC), Montreal Carbon Pledge, One Planet Sovereign Wealth Fund (OPSWF), Portfolio Decarbonization Coalition, etc.

---

[18] The reader can consult the web page https://www.unepfi.org/net-zero-alliance/resources/member-targets to retrieve the 2025 member targets.

[19] Mark Carney was the governor of the Bank of Canada from 2008 to 2013, the governor of the Bank of England from 2013 to 2020 and the chairman of the Financial Stability Board from 2011 to 2018. Since 2020, he is a United Nations special envoy for climate action and finance.

[20] We report here the press release of November 3, 2021 during the COP 26 summit:

> "*Today, through the Glasgow Financial Alliance for Net Zero (GFANZ), over \$130 trillion of private capital is committed to transforming the economy for net zero. These commitments, from over 450 firms across 45 countries, can deliver the estimated \$100 trillion of finance needed for net zero over the next three decades.*"

- Social
  Platform Living Wage Financials (PLWF), PRI Human Rights Engagement, Tobacco-Free Finance Pledge, Workforce Disclosure Initiative (WDI), etc.

- Governance
  Australian Shareholders' Association (ASA), European Corporate Governance Institute (ECGI), International Corporate Governance Network (ICGN), Say on Climate, etc.

1.3.3 Regulators

While regulators and supervisors were absent from the ESG ecosystem for a long time, they are now at the forefront of the ESG debate. The main reason is the phenomenal growth of ESG and climate investing, and the change of motivations. As long as responsible investing was driven by moral values, ESG investing concerned a small market of investors. Today, ESG has become a marketing argument and the risk of ESG-washing and greenwashing has become very high. We must distinguish two types of risk:

- Explicit & deliberate greenwashing;

- Unintentional greenwashing.

Deliberate greenwashing is a mis-selling risk, which is a subject of close scrutiny from supervisors. An example is the DWS scandal[21]. Unintentional greenwashing is a misinterpretation risk, which must be clarified by regulators[22]. An example is the definition of a net zero investment policy. In this context, clients must be protected from both types of greenwashing risk. Another reason that explains the recent interest of regulators is the political will to mitigate global warming. Indeed, financial regulation is certainly one of the most important instruments to achieve this goal. Therefore, it is no coincidence if the financial sector is expected to play a key role in helping to decarbonize the corporate sector.

Table 1.2: The supervision institutions in finance

	Banks	Insurers	Markets	All sectors
Global	BCBS	IAIS	IOSCO	FSB
EU	EBA/ECB	EIOPA	ESMA	ESFS
US	FDIC/FRB	FIO	SEC	FSOC

Regulators in charge of sustainable risk are the same than those in charge of traditional risks (e.g., market risk, credit risk or liquidity risk). In Table 1.2, we have reported a list of supervision institutions in finance. At the global level, four international authorities have primary responsibility of the financial regulation: the Basel Committee on Banking Supervision (BCBS), the International Association of Insurance Supervisors (IAIS), the International Organization of Securities Commissions (IOSCO) and the Financial Stability Board (FSB). The FSB, which is in charge of the systemic risk regulation, has identified climate risk at a very early stage. The speech "*Breaking the tragedy*

---

[21]See the Financial Times' article on litigation issues of ESG investing: `https://www.ft.com/content/1094d5da-70bf-40b5-98f4-725d50620a5a`.

[22]Here, we make the difference between regulation and supervision from a risk management viewpoint (Roncalli, 2020a, page 12). The regulator is responsible of setting rules and policy guidelines. The supervisor evaluates the safety and soundness of financial institutions and verifies that the regulation rules are applied. For example, in Europe, the regulator of the banking sector is EBA while the supervisor is ECB.

of the horizon" by Mark Carney, Chairman of the FSB, at London, 29 September 2015, marked a turning-point in the recognition of climate change as a big risk for the financial stability (Carney, 2015). According to Mark Carney, the financial stability can be affected through three channels: physical risk (the impact on insurance liabilities and financial assets that arise today from climate extreme events), liability risk (the impacts that could arise tomorrow if parties who have suffered from climate change seek compensation from those they hold responsible) and transition risk (the financial risk that could result from the process of adjustment towards a lower-carbon economy). Following the G20 Antalya Summit, the FSB proposed then to "*establish an industry-led disclosure task force, to design and deliver voluntary standards for effective disclosures that meet the needs of investors and creditors*" (FSB, 2015). The Task Force on Climate-related Financial Disclosures (TCFD) is created in December 2015 under the chairmanship of Michael Bloomberg. In June 2017, TCFD released its final climate-related financial disclosure recommendations. The first status report on disclosure practices is published in September 2018. Disclosures is the first pillar of the FSB roadmap, which covers three other areas: (1) data, (2) vulnerability analysis and (3) regulatory and supervisory practices and tools (Figure 1.5).

Figure 1.5: Stylised overview of the FSB roadmap for addressing climate-related financial risks
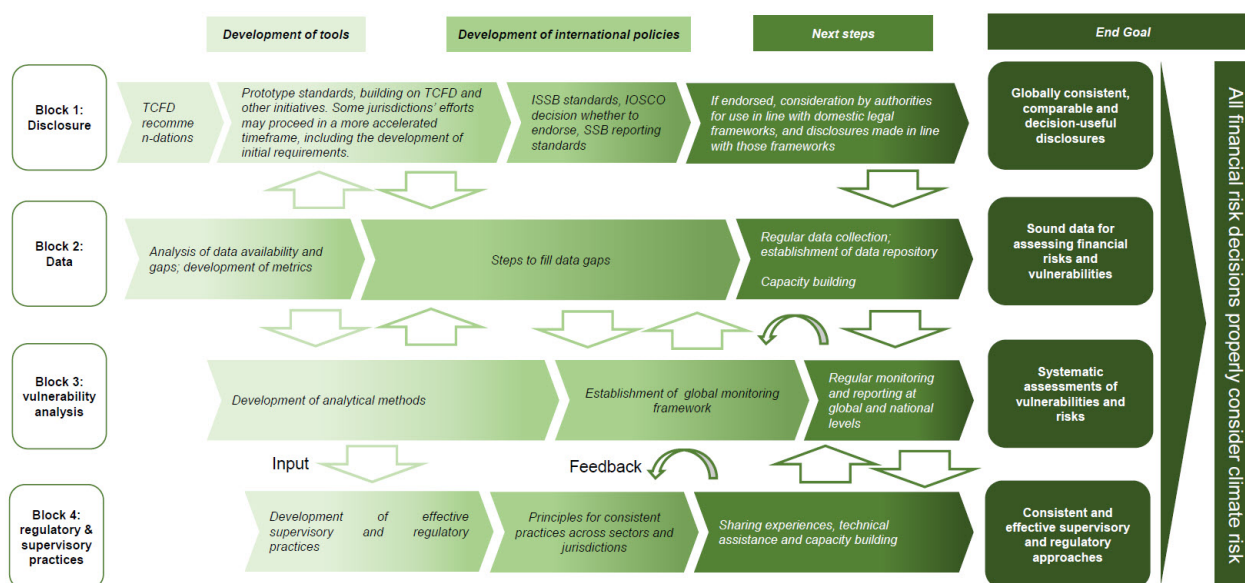
The Basel Committee on Banking Supervision (BCBS) provides a forum for regular cooperation on banking supervisory matters. Its main objective is to improve the quality of banking supervision worldwide. Its first publication on climate-related financial risks dates back to April 2020. In June 2022, BCBS released its first guidelines on this topic (BCBS, 2022). These guidelines includes 12 principles for the effective risk management of climate risk and 6 principles for the supervisory review process (SRP). We report here the first principle, which states that climate risk must be managed such as financial risks (e.g., market risk or credit risk):

> "*Banks should develop and implement a sound process for understanding and assessing the potential impacts of climate-related risk drivers on their businesses and on the envi-*

ronments in which they operate. Banks should consider material climate-related financial risks that could materialise over various time horizons and incorporate these risks into their overall business strategies and risk management frameworks."(BCBS, 2022, page 2).

Since we know that BCBS is able to rapidly develop new regulatory frameworks, we can expect the publication of new standards including climate risk in the coming years. Concerning the two other global supervision institutions, IOSCO has produced a report on ESG rating agencies and data providers (IOSCO, 2021) whereas IAIS is more focusing on climate risk and its supervision in the insurance sector (IAIS, 2021). It is interesting to notice that supervisors of the asset management industry are more focused on ESG data while supervisors of the insurance industry are more concerned by the physical risk.

We also observe a rapid transformation of the regulatory framework at the regional or national level. For instance, a new SEC rule requires all registrants to disclose information on climate risks[23]. The sustainable finance roadmap 2022-2024 identifies three priorities for ESMA[24]: (1) tackling greenwashing and promoting transparency, (2) building NCAs' and ESMA's capacities and (3) (3) monitoring, assessing and analysing ESG markets and risks. The banking supervision has already conducted several climate stress testing programs (ACPR, 2021; Bank of England, 2022; ECB, 2022). Central banks are also very active. Thus, the Network of Central Banks and Supervisors for Greening the Financial System (NGFS) is launched at the Paris One Planet Summit (OPS) on December 2017. Its 8 founding members are Banco de Mexico, BoE, Banque de France, Dutch Central Bank, Deutsche Bundesbank, Swedish FSA, Hong Kong Monetary Authority (HKMA), Monetary Authority of Singapore (MAS) and The People's Bank of China (PBOC). As of October 3rd 2022, the NGFS consists of 121 members and 19 observers[25]. In addition to its mythological publications, the NGFS is well known for its database on climate scenarios[26].

Sustainable finance is not only regulated by financial regulators. In Figures 1.6 and 1.7, we have reported charts from the MSCI website, that list the ESG regulations by type of regulatory agency or by type of regulated party. We observe that the number of regulations is greater for issuers than investors. Moreover, other bodies than financial regulators are involved in the ESG regulation landscape, especially governments. For instance, the French law "*Climat et Résilience*" (climate and resilience) of 22 August 2021 translates part of the 146 proposals of the Citizen's Climate Convention adopted by the French government, to reduce greenhouse gas emissions by 40% by 2030 in a spirit of social justice. In Europe, most of ESG regulations are defined by the European Commission (EC) and the European Parliament. This is for example the case of the EU Non-Financial Reporting Directive (NFRD, 2014), the EU Taxonomy Regulation for sustainable activities (EUTR, 2020) and the EU Sustainable Finance Disclosure Regulation (SFDR, 2021). All these policy initiatives are part of the "*European Green Deal*", whose aim is making the European Union climate neutral in 2050. New legislations on the circular economy, building renovation, biodiversity, farming and innovation are under way. To define these directives, the EC is supported by technical working groups such as the High-Level Expert Group on sustainable finance (HLEG) or the Technical Expert Group on sustainable finance (TEG). For example, the EC has mandated in 2020 the European Financial Reporting Advisory Group (EFRAG) to undertake preparatory work for the elaboration of the new EU Corporate Sustainability Reporting Directive (CSRD) that will amend the current NFRD.

---

[23]Visit https://www.sec.gov/news/press-release/2022-46.
[24]Visit https://www.esma.europa.eu/policy-activities/sustainable-finance.
[25]Including BIS, BCBS, ESM, FSB, IAIS, IMF, IOPS, IOSCO and OECD.
[26]Visit https://www.ngfs.net/ngfs-scenarios-portal.

Figure 1.6: Who will regulate ESG? — The regulators viewpoint (MSCI, 2022)

Source: https://www.msci.com/who-will-regulate-esg.
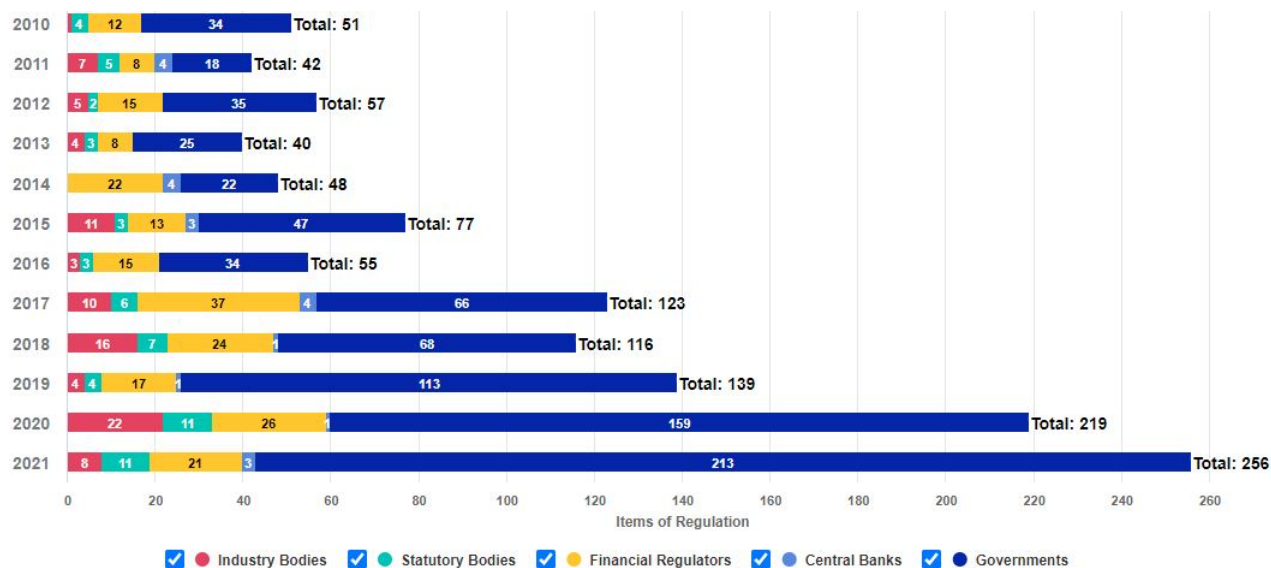
Figure 1.7: Who will regulate ESG? — The regulated viewpoint (MSCI, 2022)

Source: https://www.msci.com/who-will-regulate-esg.
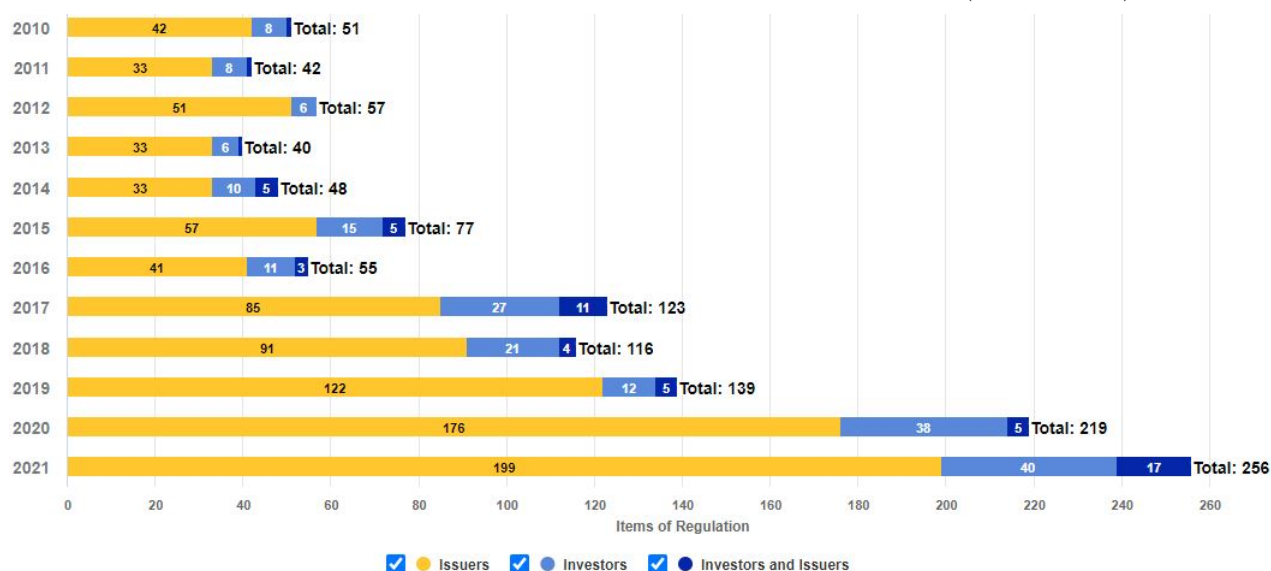
1.3.4 Reporting frameworks

As we have already seen, reporting is a key element to understand ESG and climate policies of issuers and investors. In the past 20 years, we are seeing more and more new reporting frameworks. In Table 1.3, we list the best-known ones. For each reporting, we give the creation date of the initiative and the implementation date of the standards. The two first reporting frameworks were those of the Global Reporting Initiative (GRI) and the GHG Protocol. The most recent is the International SustainabilityÂăStandards Board (ISSB). In what follows, we distinguish sustainability general reporting and climate specific reporting.

Table 1.3: List of the main reporting frameworks

Perimeter	Acronym	Name	Dates
General	GC	UN Global Compact Initiative	2000/2000
	GRI	Global Reporting Initiative	1997/2000
	IIRC	International Integrated Reporting Council	2010/2013
	ISSB	International SustainabilityÂăStandards Board	2021/2023
	SASB	Sustainability Accounting Standards Board	2011/2016
	SDGs	UN Sustainable Development Goals	2015/2016
Climate	CDP	Carbon Disclosure Project	2000/2000
	CDSB	Climate Disclosure Standards Board	2007/2015
	GHG Protocol	Greenhouse Gas Protocol	1998/2001
	PCAF	Partnership for Carbon Accounting Financials	2019/2020
	SBTi	Science Based Targets initiative	2015/2015
	TCFD	Task Force on Climate-Related Financial Disclosures	2015/2017

Sustainability reporting

International Sustainability Standards Board After a decade of framework proliferation, the landscape of sustainability reporting has changed significantly over the past two years. In June 2021, SASB and IIRC definitively merged into one organization to form the Value Reporting Foundation (VRF). On 3 November 2021, the IFRS Foundation Trustees announced the creation of the International Sustainability Standards Board (ISSB) chaired by Emmanuel Faber[27], with the objective to deliver a comprehensive global baseline of sustainability-related disclosure standards. On 31 January 2022, the Climate Disclosure Standards Board (CDSB) was consolidated into the IFRS Foundation to support the work of ISSB. On 1 August 2022, the IFRS Foundation completes a new consolidation with VRF. Even if the previous frameworks continue to exist and can still be used by companies, it will exist only one sustainability reporting standard in the future.

On 31 March 2022, ISSB published the drafts of its first proposed standards:

- IFRS S1 general requirements for disclosure of sustainability-related financial information (ISSB, 2022a);

- IFRS S2 climate-related disclosures (ISSB, 2022b).

The IFRS S1 draft requires companies to identify sustainability-related risks and opportunities until the SASB standards are replaced by IFRS Sustainability Disclosure Standards. The IFRS S2 draft

---

[27]Emmanual Faber was CEO and Chair of the Board at multi-national food products company Danone.

builds on the TCFD recommendations. On October 2022, the ISSB decided to include the scope 3 GHG emissions in the climate reporting according to the fifteen scope 3 categories described in the GHG Protocol.

Sustainable Development Goals   The SDGs are a collection of 17 interlinked global goals designed to be a "*blueprint to achieve a better and more sustainable future for all*". They were set up in 2015 by the United Nations and are intended to be achieved by 2030. The 17 SDGs are given in Table 1.4.

Figure 1.8: The SDGs icons

Source: https://sdgs.un.org/goals#icons.

Each goal is defined by specific targets, and the progress toward each target is measured by indicators. A total of 69 targets and 231 unique indicators are then considered. The numbering system Goal.Target.Indicator is used to structure the tree map of the SDGs. For instance, the first target of the first goal is: 1.1 — By 2030, eradicate extreme poverty for all people everywhere, currently measured as people living on less than \$1.25 a day. This target is measured by only one indicator: 1.1.1 — Proportion of the population living below the international poverty line by sex, age, employment status and geographic location. The fifth target of the first goal is: 1.5 — By 2030, build the resilience of the poor and those in vulnerable situations and reduce their exposure and vulnerability to climate-related extreme events and other economic, social and environmental shocks and disasters. This target is measured by three indicators: 1.5.1 — Number of deaths, missing persons and directly affected persons attributed to disasters per 100 000 population, 1.5.2 — Direct economic loss attributed to disasters in relation to global gross domestic product and 1.5.3 — Number of countries that adopt and implement national disaster risk reduction strategies in line with the Sendai Framework for Disaster Risk Reduction 2015-2030. Initially, the SDGs are built for assessing the progress of each country on the different pillars. We can then analyze the evolution of each indicator per country and year. Synthetic scores are also available at the country or goal level. A compilation of these scores[28] can be found in Sachs *et al.* (2022).

---

[28]They are also available at the web page https://dashboards.sdgindex.org.

Table 1.4: The 17 SDGs

#	Name	Description	E	S	G
1	No poverty	End poverty in all its forms everywhere		✓	
2	Zero hunger	End hunger, achieve food security and improved nutrition and promote sustainable agriculture		✓	
3	Good health and well-being	Ensure healthy lives and promote well-being for all at all ages		✓	
4	Quality education	Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all		✓	
5	Gender equality	Achieve gender equality and empower all women and girls		✓	✓
6	Clean water and sanitation	Ensure availability and sustainable management of water and sanitation for all	✓	✓	
7	Affordable and clean energy	Ensure access to affordable, reliable, sustainable and modern energy for all	✓		
8	Decent work and economic growth	Promote sustained, inclusive and sustainable economic growth, full and productive employment and decent work for all		✓	✓
9	Industry, innovation and infrastructure	Build resilient infrastructure, promote inclusive and sustainable industrialization and foster innovation	✓	✓	✓
10	Reduced inequality	Reduce inequality within and among countries		✓	
11	Sustainable cities and communities	Make cities and human settlements inclusive, safe, resilient and sustainable	✓		✓
12	Responsible consumption and production	Ensure sustainable consumption and production patterns	✓	✓	✓
13	Climate action	Take urgent action to combat climate change and its impacts	✓		✓
14	Life below water	Conserve and sustainably use the oceans, seas and marine resources for sustainable development	✓		
15	Life on land	Protect, restore and promote sustainable use of terrestrial ecosystems, sustainably manage forests, combat desertification, and halt and reverse land degradation and halt biodiversity loss	✓		
16	Peace, justice, and strong institutions	Promote peaceful and inclusive societies for sustainable development, provide access to justice for all and build effective, accountable and inclusive institutions at all levels		✓	✓
17	Partnerships for the goals	Strengthen the means of implementation and revitalize the Global Partnership for Sustainable Development			✓

Source: https://sdgs.un.org/goals.

The SDGs has been quickly used by financial institutions as a framework for impact investing. In Table 1.4, we map the 17 SDGs and the 3 ESG pillars. Therefore, we can assign the SDGs targets to each ESG dimension. An example applied to artificial intelligence companies is provided by Sætra (2022). The SDGs have also been used to evaluate the ESG objectives of sustainable financial products. For example, ICMA has published a mapping[29] to SDGs for green and social bonds (ICMA, 2022), where targets are associated with GBP and SBP categories.

Climate reporting

GHG Protocol   The GHG Protocol has been created by WRI and WBCSD in 1998 with the aim of "*establishing comprehensive global standardized frameworks to measure and manage greenhouse gas emissions from private and public sector operations, value chains and mitigation actions*". First published in 2001, the standard defines the accounting and reporting of six greenhouse gases covered by the Kyoto Protocol, including carbon dioxide (CO_2) and methane (CH_4).

The GHG Protocol corporate standard classifies a company's greenhouse gas emissions in three scopes (GHG Protocol, 2004):

- Scope 1 denotes direct GHG emissions occurring from sources that are owned and controlled by the issuer.

- Scope 2 corresponds to the indirect GHG emissions from the consumption of purchased electricity, heat or steam.

- Scope 3 are other indirect emissions (not included in scope 2) of the entire value chain.

Scope 2 emissions can be computed using two methods[30] (GHG Protocol, 2015):

1. the energy mix of the countries (location-based);

2. the energy mix of the utility companies supplying the electricity (market-based).

Scope 3 is based on 15 sub-categories (GHG Protocol, 2011, 2013), which are divided into two main categories[31]:

- Upstream scope 3 emissions are defined as indirect carbon emissions related to the upstream value chain. More precisely, the upstream scope 3 is based on 8 sub-categories: (1) purchased goods and services, (2) capital goods, (3) fuel and energy related activities, (4) upstream transportation and distribution, (5) waste generated in operations, (6) business travel, (7) employee commuting and (8) upstream leased assets.

- Downstream scope 3 emissions are defined as indirect carbon emissions related to the downstream value chain. They correspond to these next 7 sub-categories: (9) downstream transportation and distribution, (10) processing of sold products, (11) use of sold products, (12) end-of-life treatment of sold products, (13) downstream leased assets, (14) franchises and (15) investments.

Scope 1 emissions are also called direct emissions, whereas indirect emissions encompass both scope 2 and 3 GHG emissions. Unlike scope 1 and 2, scope 3 is an optional reporting category.

---

[29]The link to the Excel mapping file is https://www.icmagroup.org/assets/documents/Sustainable-finance/2022-updates/Mapping-to-SDGs_June-2022-280622.xlsx.

[30]The exact definitions are the following: "*a location-based method reflects the average emissions intensity of grids on which energy consumption occurs (using mostly grid-average emission factor data*", while "*a market-based method reflects emissions from electricity that companies have purposefully chosen (or their lack of choice)*".

[31]The upstream value chain includes all activities related to the suppliers whereas the downstream value chain refers to post-manufacturing activities.

Carbon Disclosure Project   The CDP (formerly the Carbon Disclosure Project) is a UK-based not-for-profit charity[32] co-founded by Paul Dickinson and Tessa Tennant in 2000. CDP runs a global disclosure system for investors, companies, cities, states and regions to manage their environmental impacts. Each year, CDP sends a questionnaire to organizations and collects information on three environmental dimensions:

1. Climate change (based on the GHG Protocol[33]).

2. Forest management;

3. Water security.

In particular, the CDP database[34] is extensively used to measure the carbon footprint of companies, cities and governments. In 2022, more than 18 700 companies and 1 100 cities, states and regions have filled in the questionnaire. This represents half of global market capitalization. Nevertheless, more than 29 500 companies (20% of market capitalization) didn't respond to the disclosure request.

Remark 1 *As CDP is the most comprehensive reporting database for carbon emissions[35], the CDP data are extensively used by commercial data providers (e.g., Trucost and MSCI) when providing carbon footprint estimates.*

TCFD   The Task Force on Climate Related Financial Disclosures (TCFD) is established by the FSB in 2015 to develop a set of voluntary and consistent disclosure recommendations for use by companies in providing information to investors, lenders and insurance underwriters about their climate-related financial risks. The TCFD consists of 31 members from the G20, representing both preparers and users of financial disclosures and is chaired by Michael Bloomberg. The TCFD framework is published in June 2017 and the 11 recommendations are structured around 4 core elements: (1) governance, (2) strategy, (3) risk management, and (4) metrics and targets (Table 1.5). The first core element describes the organization's governance around climate-related risks and opportunities, whereas the second one lists the actual and potential impacts of climate-related risks and opportunities on the organization's businesses, strategy, and financial planning. The processes used by the organization to identify, assess, and manage climate-related risks are specified in the risk management tag. Finally, the last core element defines the metrics and targets used to measure and manage relevant climate-related risks and opportunities. The implementation of the reporting framework is extensively described in TCFD (2021a,b), and many examples can be found in CDSB (2021c) and TCFD (2022).

Contrary to the other climate frameworks (e.g., GHG Protocol and CDP), the TCFD framework is a risk reporting, and not only a carbon emission reporting. For instance, we report below some examples of recommended metrics (TCFD, 2022, pages 16-17):

- GHG emissions (absolute scope 1, scope 2, and scope 3 GHG emissions; financed emissions by asset class; weighted average carbon intensity);

---

[32]The global budget of CDP is about $30 mn. CDP's funding comes mainly from philanthropic grants (32%), service based memberships (30%) and government grants (12%).

[33]The differences between the GHG Protocol and CDP reporting templates are the following. The GHG Protocol reporting is more focused on figures, while the CDP reporting contains more open questions and comments. Moreover, the CDP reporting is a little more comprehensive, because it also concern forest management and Water security.

[34]It is available at https://www.cdp.net/en/data.

[35]The fact that the CDP reporting is an Excel file may explain that it had more success than the GHG Protocol reporting, which is a Word file. However, they are very similar regarding carbon emissions disclosure.

Table 1.5: The 11 recommended disclosures (TCFD, 2017)

Recommendation	#	Recommended Disclosure
Governance	1	Board oversight
	2	Management's role
Strategy	3	Risks and opportunities
	4	Impact on organization
	5	Resilience of strategy
Risk management	6	Risk ID and assessment processes
	7	Risk management processes
	8	Integration into overall risk management
Metrics and targets	9	Climate-related metrics
	10	Scope 1, 2, 3 GHG emissions
	11	Climate-related targets

Source: https://www.fsb-tcfd.org.

- Transition risks (volume of real estate collaterals highly exposed to transition risk; concentration of credit exposure to carbon-related assets; percent of revenue from coal mining);

- Physical risks (number and value of mortgage loans in 100-year flood zones; revenue associated with water withdrawn and consumed in regions of high or extremely high baseline water stress; proportion of property, infrastructure, or other alternative asset portfolios in an area subject to flooding, heat stress, or water stress; proportion of real assets exposed to 1:100 or 1:200 climate-related hazards);

- Climate-related opportunities (net premiums written related to energy efficiency and low-carbon technology; revenues from products or services that support the transition to a low-carbon economy; proportion of green buildings);

- Capital deployment (percentage of annual revenue invested in R&D of low-carbon products/services; investment in climate adaptation measures);

- Internal carbon prices (internal carbon price, shadow carbon price);

- Remuneration (portion of employee's annual discretionary bonus linked to investments in climate-related products; weighting of climate goals on long-term incentive; scorecards for executive directors).

Similarly, targets are also more general, and are not limited to carbon emission reduction. For instance, they can concern the amount of executive management remuneration impacted by climate considerations by 2025, the internal carbon price by 2030, the amount invested in green buildings by 2035, etc.

Remark 2 *Examples of TCFD reporting are given in Figure 1.9. We can generally find TCFD and climate reports by using the* Google *search bar with the keywords* year + "TCFD report" + corporate name *or* year + "climate report" + corporate name. *As we can observe, the formats of TCFD reports are diverse. They can correspond to a powerpoint file or a written document, the number of pages ranges from 3 to 100, etc.*

Figure 1.9: Examples of TCFD reports

The TCFD framework is supported by many international bodies and supervisors: European Commission, IFRS, IOSCO, Singapore Exchange Regulation, Central Bank of Brazil, Australian Prudential Regulatory Authority, Canadian Securities Administrators, etc. In this context, it has become the most popular reporting framework from the viewpoint of regulation. Nevertheless, much progress remains to be done, since this reporting is voluntary and not mandatory. For fiscal year 2021 reporting, "80% *of companies disclosed in line with at least one of the 11 recommended disclosures; however, only* 4% *disclosed in line with all 11 recommended disclosures and only around* 40% *disclosed in line with at least five* " (TCFD, 2022, page 5). The average level of disclosure is 60% for European companies, 36% for Asia Pacific companies and 29% for North American companies. Nearly 50% of asset managers and 75% of asset owners reported information aligned with at least five of the 11 recommended disclosures. The most popular recommended disclosures were (#3) risks and opportunities (61%), (#4) impact on organization (47%), and (#9) climate-related metrics (47%), while the less popular items were (#5) resilience of strategy (16%), (#2) management's role (22%) and (#1) board oversight (29%).

1.3.5   Rating agencies and data providers

To implement ESG strategies, we need extra-financial data. In the 1980s and 1990s, several research companies are then established to provide research on responsible investing. These small-sized firms are generally specialized on a specific dimension and a region. Some of them are focused on the environmental pillar, but the majority of them are specialized in the social pillar. In addition to research and advisory activities, they begin to collect a lot of extra-financial data and build sustainable scores. After an initial period of expansion and innovation, they structure themselves as global rating agencies using the model of credit rating agencies (CRA). Then, we observe a concentration in this industry and a period of consolidation in the 2010s.

The early stage of extra-financial rating agencies

In 2001, the French observatory centre for the corporate social responsibility[36] ORSE published a guide of entities specialized in ESG analysis (ORSE, 2001). This guide has been updated several times until 2012. For instance, in the 2005 edition, ORSE listed 34 sustainable research organizations of which 25 are located in Europe, 5 in North America and 4 in the rest of the world (Australia, Japon, South Korea). This number does not change much during the 2000s. Indeed, most of extra-financial rating agencies were created in the 1990s. Here is a list[37] of some well-know entities[38]: Ethical Investment Research Services Ltd. (Eiris[39], 1983, UK), Institutional Shareholder Services (ISS[40], 1985, UK), Kinder, Lydenberg, Domini & Co. (KLD[41], 1988, US), Jantzi Research (Jantzi Research[42], 1992, Canada), Global Engagement Services (GES[43], 1992, Sweden), Innovest Strategic Value Advisors (Innovest[44], 1995, US), Sustainable Asset Management Ltd. (SAM[45], 1995, Switzerland), RepRisk (RepRisk[46], 1998, Switzerland), Oekom Research AG (Oekom[47], 1999, Germany), Ethix SRI Advisors (Ethix[48], 1999, Sweden), Trucost Plc (Trucost[49], 2000, UK), Inrate (Inrate[50],

---

[36]The French name is Observatoire de la Responsabilité Sociétale des Entreprises.

[37]This list is based on the works of Eccles and Stroehle (2018) and the company profiles provided by ORSE (2007).

[38]The date of creation and the country are provided in parentheses.

[39]EIRIS was founded in 1983 by charities and churches as the UK's first independent research service for ethical investors.

[40]Institutional Shareholder Services was originally founded in 1985 by Robert Monks, an ESG advocate. It began to provide voting services in 1992.

[41]KLD was founded in 1989 by Amy Domini, Peter Kinder and Steve Lydenberg to offer institutional investors social research on US companies. In May 1990, it launched the Domini 400 Social Index (DSI).

[42]Founded by Michael Jantzi in 1992, Jantzi Research became a pioneer in the field of ESG research in Canada. In January 2000, it launched the Jantzi Social Index (JSI) consisting of 50 Canadian stocks.

[43]Caring Capitalism was founded in 1992 and renamed Global Engagement Services in 2003.

[44]Innovest was created by Matthew Kierman and Hewson Batlzell as "*a green analogy to Moody's*" (Eccles *et al.*, 2020). In the first years, it focused on environmental screening. Later, it created the IVA ratings using the credit-like rating scale (AAA, AA, A, BBB, BB, B, and CCC).

[45]SAM was founded in 1995. In 1999, SAM and Dow Jones launched the Dow Jones Sustainability Indexes. In 2001, SAM created the first global sustainable water fund (SAM Sustainable Water Fund). In 2006, SAM is acquired by Robeco and is renamed RobecoSAM in 2013.

[46]RepRisk was founded in 1998. The RepRisk's database was launched later and began in January 2007.

[47]The environmental publishing house ökom was founded in 1989. In 1993, ökom GmbH was created for providing environmental research. Oekom research AG became independent in 1999 and focused on corporate responsibility ratings.

[48]Established in 1999, Ethix developed norm-based screening in 2000 and extended it to emerging markets in 2005.

[49]Trucost was established in 2000 to help organisations, investors and governments understand and quantify the environmental impacts of business activities. The Trucost's database was launched later and began with the 2005 financial year.

[50]Inrate is officialy created in 2001, but its roots dated back to 1990 with the foundation of Centre Info.

2001, Switzerland), Vigeo (Vigeo[51], 2002, France), Dutch Sustainability Research (DSR[52], 2002, Netherlands), EthiFinance (EthiFinance[53], 2004, France).

The consolidation of the industry

As shown by Eccles and Stroehle (2018) and Demartini (2020), we are seeing a consolidation period in the 2010s. Here are some examples:

- Vigeo and Eiris merged in October 2015 to form Vigeo-Eiris (V.E), which is acquired by Moody's in April 2019.

- In September 2015 and March 2018, ISS acquired Ethix SRI Advisors and Oekom to form ISS ESG solutions (ISS-ethix, ISS-climate and ISS-oekom). In November 2020, ISS is majority owned by Deutsche Börse Group.

- In February and November 2009, RiskMetrics acquired Innovest and KLD. RiskMetrics is bought by MSCI in 2010, which creates MSCI ESG Research LLC.

- In September 2009, DSR and Jantzi Research merged to form Sustainalytics. In the 2010s, Sustainalytics acquired Responsible Research (Singapore), ESG Analytics (Switzerland), Solaron (India) and GES (Sweden). In April 2020, Sustainalytics becomes a wholly-owned subsidiary of Morningstar.

- S&P Global acquired Trucost in October 2016 and the ESG ratings business of RobecoSAM in November 2019.

Today, the industry of extra-financial analysis and ESG ratings is dominated by ISS-Oekom, MSCI, Refinitiv[54], Reprisk, S&P Global, Sustainalytics and Moody's.

Remark 3 *In Chapter 2, we will see that these ESG rating agencies are specialized and do not provide the same solutions. For instance, on controversy risk, the major players are MSCI, Reprisk and Sustainalytics. On climate risk, CDP, MSCI and Trucost are the leader agencies, while Verisk Maplecroft is the specialized agency in sovereign ESG risk.*

Remark 4 *Even if we observe a consolidation, this does not mean that we observe a convergence of ESG methodologies (Chatterji et al., 2016; Berg et al., 2022). This point will be discussed in the next chapter.*

The current business of extra-financial data

As noticed by Demartini (2020), the industry is mainly made up of large Anglo-Saxon companies (US and UK) and small European start-up firms. More precisely, we can classify them into three main categories:

---

[51]Founded in 1997, Arese was the first SRI rating agency in France. In June 2002, it became Vigeo and was lead-managed by Nicole Notat, the former secretary general of the labor union CFDT. In June 2005, Vigeo merged with the Belgian agency Ethibel.

[52]DSR was the research team of Triodos Bank, a Dutch niche player in sustainable finance founded in 1980. In 2008, it changed its name and became Sustainalytics.

[53]EthiFinance was founded in 2004. In 2017, it merged with Spread Research.

[54]Refinitiv is the former financial and risk unit of Thomson Reuters (including Eikon and Datastream). It is now part of the London Stock Exchange Group (LSEG), which has also acquired FTSE Russell (and Beyond Ratings).

1. Market data providers
   This category comprises financial information providers (Bloomberg, Morningstar), index sponsors (Bloomberg, FTSE Russell, MSCI, Solactive) and stock exchanges (LSEG, Deutsche Börse Group).

2. Financial rating agencies
   Moody's, S&P Global and Fitch are now involved in the ESG landscape.

3. Specialized ESG companies
   In this category, we generally find some pioneer companies such as Inrate, ISS ESG, RepRisk and Sustainalytics.

4. Technology start-up firms
   Most of new entrants use artificial intelligence (AI), big data, natural language processing (NLP), sentiment analysis and quantitative approaches. Some examples are Arabesque, Covalence, OWL ESG, and Truvalue Labs.

This explains the discrepancy between the companies in terms of ESG analysts. About 20% of extra-financial rating agencies have more than 200 ESG analysts[55], while 30% have less than 20 analysts (Demartini, 2020, page 11). Another difference concerns the wide scope of activities: provision of raw data, provision of processed data (indicators, scores and ratings), production of ESG indexes, ESG screening, portfolio analysis, normative analysis, ESG controversy tracking, engagement monitoring, proxy advisory services, consultancy, etc. (Demartini, 2020). These activities explain that the business model of extra-financial rating is based on the investor-pays principle, contrary to credit rating agencies whose historical model was mainly driven by the issuer-pays principle. This means that investors pay a fee to access data, but issuers don't pay to be rated.

The question of certification and supervision is on everyone's lips. For instance, credit rating agencies registered in the EU are supervised by ESMA[56] (Regulation 462/2013/EU and Directive 2013/14/EU). In the US, the Office of Credit Ratings (OCR) assists the Security and Exchange Commission (SEC) to oversight those registered as nationally recognized statistical rating organizations[57] (NRSRO). As the market of ESG ratings is expected to grow, the supervision of this industry and the protection of investors are becoming an unavoidable topic. For instance, on April 2022, the EC launched a targeted consultation on the functioning of the ESG rating market in the European Union and on the consideration of ESG factors in credit ratings (EC, 2022a). The summary report based on 204 responses[58] is published in August 2022. Its main conclusions are the following:

> "The large majority of respondents (over 84%) consider that the market is not functioning well today. On the quality of ESG ratings, two thirds of respondents consider the quality to be fine to very good, with about one third considering it poor. A large majority of respondents (83%) consider that the lack of transparency on the methodologies used by the providers is a problem in the ESG ratings market. The vast majority of respondents (91%) also consider that there are significant biases with the methodology used by providers [...] Almost all respondents (94%) consider that intervention is necessary,

---

[55]For instance, Sustainalytics has more than 800 ESG analysts (source: https://www.sustainalytics.com/about-us).

[56]The list of certified CRAs is available at https://www.esma.europa.eu/supervision/credit-rating-agencies/risk.

[57]The list of certified NRSROs is available at https://www.sec.gov/ocr/ocr-current-nrsros.html.

[58]Including 21 ESG rating providers, 48 rating users (investors), 49 rated companies and 18 rating users (company).

> *of which the large majority (80%+) support a legislative intervention with the remain-der supporting the development of non-regulatory intervention in the form of guidelines, code of conduct. Respondents largely indicated (90%+) that the main element to be addressed by the intervention should be improving transparency on the methodology used by ESG rating provider [...] The vast majority of respondents (82%) consider that ESG rating providers should be subject to some form of authorisation/registration regime in order to offer their services in the EU."* (EC, 2022b, pages 3-4).

These results confirm previous analyzes that found that the most common shortcomings are (1) a lack of coverage, (2) data quality, and (3) a lack of transparency around methodologies used by ESG rating providers (Boffo and Patalano, 2020). We can then anticipate that supervisors will certainly introduce regulatory safeguards for using ESG ratings in the near future.

1.4 Regulatory framework

The number of ESG regulations has dramatically increased over the last years. In Figure 1.10, we report its global evolution and the breakdown by region. According to PRI (2022b), there are 868 policy tools and guidance around the world, which encourage or require investors to consider ESG factors. Most of them have been developed since 2000, and we observe an acceleration which coincides with the Paris Agreement for climate change. The breakdown by region is reported in Figure 1.11. We observe that ESG regulations have gained the greatest momentum in Europe, but they are increasing in the other regions too. By analyzing the PRI's regulation database, we obtain the following results[59]:

- Policies are mainly issued by governments and regulators (78.8%). 19% are released by industry associations, including stock exchanges. Finally, less than 3% are due to international organisations (OECD, UN, ILO, etc.).

- Most of these policies are mandatory. Nevertheless, the number of voluntary-based approaches is significant[60] since they represent 33.2% of the sample.

- Four types of policy dominate: corporate ESG disclosure (61.5%), investor ESG disclosure (24.2%), investor ESG integration (20.3%) and national sustainable finance strategy (10.8%). The other types are sector specific policy (5.6%), financial products[61] (4.5%), stewardship code (2.6%) and taxonomy (1.6%).

- The application of these policies mainly concern corporations (72.4%), asset owners (60.9%), investment managers (40%) and insurance companies (28.2%). The other categories are financial service providers (16.2%) and credit rating agencies (6.5%).

- If we focus on countries, we obtain the following ranking: China (49 policies), Germany (31), European Union (29), Italy (28), Spain (26), France (24), US (23), Netherlands (22), Japan (22) and UK (20).

Most of policies are national. They can also concern a specific sector. For instance, if we focus on regulations aiming to promote the improvement of the energy performance of buildings, and their

---

[59]For each item, we indicate the frequency in %. Since each policy can concern several items, the frequencies may not add up to 100%. For example, a policy can be applicable to both asset owners, issuers and asset managers.

[60]Especially when they are issued by the industry.bodies.

[61]This category include green bonds, social bonds, green labels, etc.

Figure 1.10: Total number of ESG regulations
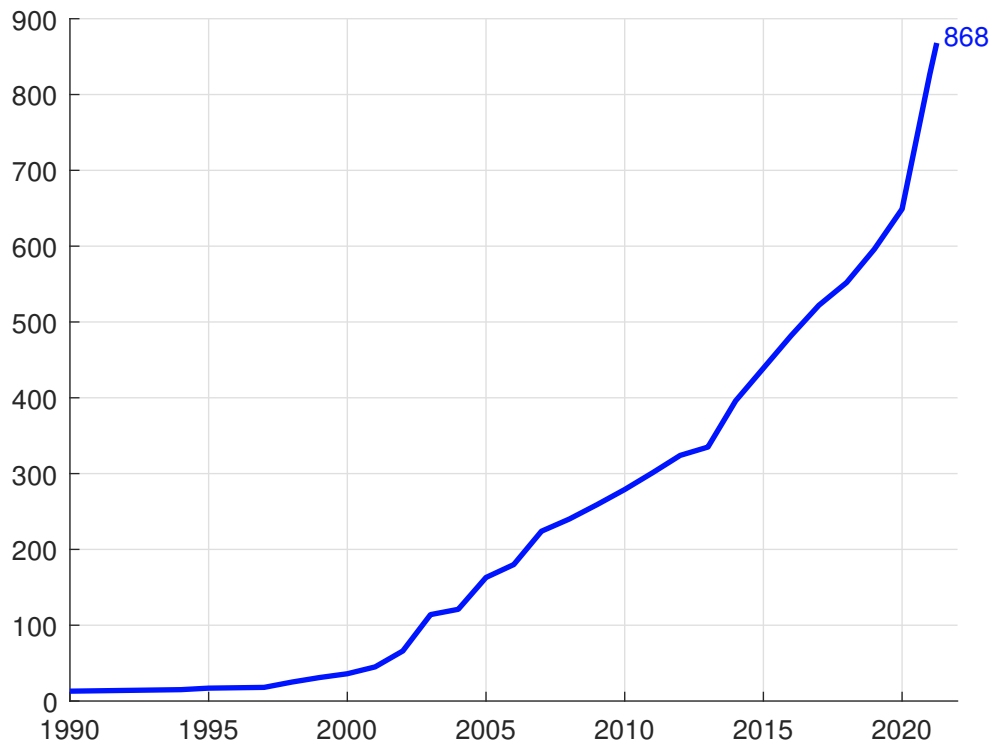
Figure 1.11: Number of ESG regulations per region
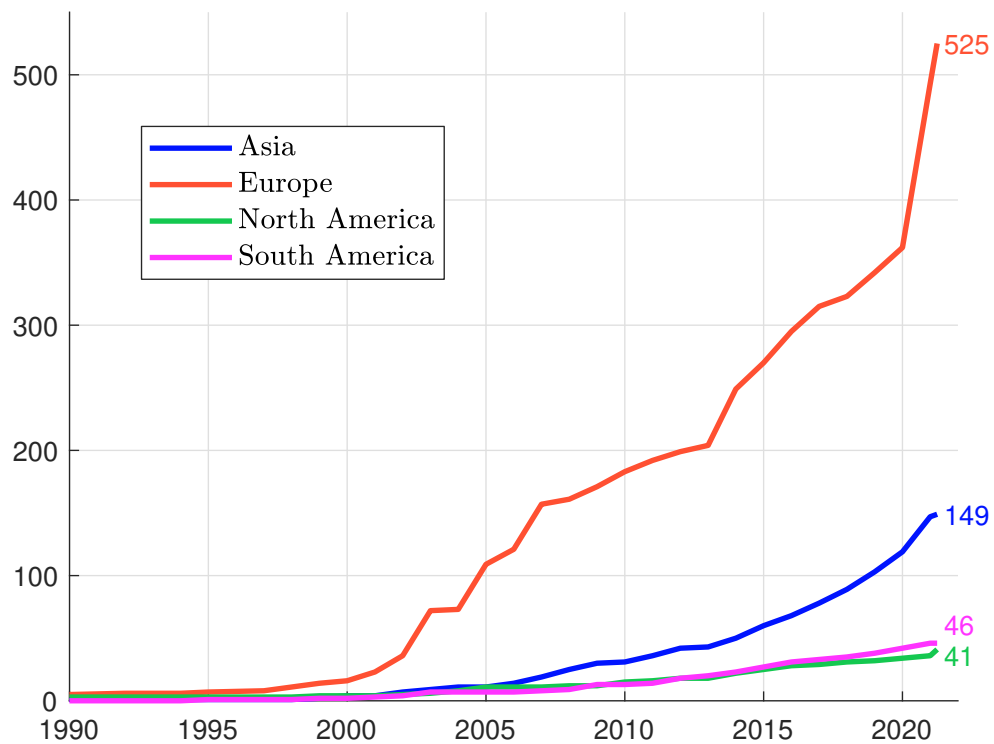
reduction of GHG emissions and energy consumption, we find many legislative policies, for example the Energy Performance of Buildings Directive (EPBD) in Europe, the French high environmental quality (HQE) certification, the German buildings energy act (Gebäudeenergiegesetz or GEG), the German sustainable building certification (Deutsche Gesellschaft für Nachhaltiges Bauen or DGNB), the Italian energy-efficient construction and renovation certification (CasaClima), the Spanish climate change and energy transition law, etc. In this context, it is not realistic to have an overview of the regulations in the World. This is why we will focus on the European Union[62].

The European regulatory framework is articulated around a set of policy initiatives by the European Commission:

- The action plan on sustainable finance (May 2018);

- The European Green Deal (December 2019);

- The Fit-for-55 package (July 2021);

- The REPowerEU plan or energy security package (May 2022).

In December 2016, the EC established a high-level expert group on sustainable finance (HLEG), consisted of 20 senior experts. HLEG (2018) published its final report on 31 January 2018 with several recommendations and proposals: (1) a classification system (or taxonomy) to provide market clarity on what is sustainable, (2) clarifying the duties of investors when it comes to achieving a more sustainable financial system, (3) improving disclosure by financial institutions and companies on how sustainability is factored into their decision-making, (4) an EU-wide label for green investment funds, (5) making sustainability part of the mandates of the European Supervisory Authorities (ESAs) and (6) a European standard for green bonds. All these recommendations are endorsed by the EC and form the basis of the action plan on sustainable finance adopted by the EC in March 2018 and the European Parliament in May 2018. In July 2018, a technical expert group on sustainable finance (TEG) is established to assist the EC in developing an EU green taxonomy (1), guidance to improve corporate disclosure of climate-related information (3), an EU green bond standard (6) and methodologies for EU climate benchmarks. In December 2019, the EC proposed a set of climate change policies, including biodiversity, circular economy, construction, energy, food, forests, transport, etc. The overarching aim of this European Green Deal is for the European Union to become the world's first climate-neutral continent by 2050. To finance this climate change strategy, the EC adopted in July 2021 the Renewed Sustainable Finance Strategy, which is a package of measures to help improve the flow of money towards financing the transition to a sustainable economy. The goal is to mobilize at least €1 tn financing over the decade. At the same time, a new cycle of legislative package are proposed under the European Green Deal framework. In particular, the EC adopted the Fit-for-55 package, a set of policies to reach the objective of cutting GHG emissions by 55% by 2030 versus 1990[63]. The plan relies on four pillars:

1. a more pronounced industrial transformation, with a wider application of the EU Emissions Trading System[64] (ETS) to new sectors, along with a tightening of the ETS itself;

---

[62]We will not speak about the situation in the US, because it is not stabilized, in particular with the recent emergence of the anti-ESG movement. The subject apparently seems to be highly controversial. On May 25, 2022, the US Securities and Exchange Commission (SEC) proposed new rules and reporting forms to enhance the regulatory framework for disclosures concerning funds' and advisers' incorporation of ESG factors. Nevertheless, these rules are still under discussion. At the same time, we observe US political moves against ESG investing (e.g., Texas, Florida). Therefore, these backlashes place the US in an uncertain ESG environment.

[63]With existing measures, the EU's carbon emissions are expected to fall 36% only below 1990 levels.

[64]Emissions trading systems are presented in Chapter 8 on page 349.

2. a faster transition to clean mobility and air transport;

3. a significant growth of renewable energies[65] and energy efficiency;

4. the restoration of natural ecosystems and forestry to absorb carbon from natural sinks.

On 18 May 2022, the EC published the REPowerEU plan that contains a suite of measures to phase out Russian fossil fuels by 2027 and boost the EU's renewable energy production and energy efficiency measures. The REPowerEU plan is presented as "*the response of the EC to the hardships and global energy market disruption caused by Russia's invasion of Ukraine*". The objective is clearly to cut the gas dependency of the European Union. It is also an extension project of the Fit-for-55 package with four objectives: energy savings, diversifying supplies, accelerating the rollout of renewable energies and reducing fossil fuel consumption in industry and transport.

Remark 5 *The European Green Deal, the Fit-for-55 package and the REPowerEU plan forms the global climate strategy of the European Union. At first sight, we may think that they concern almost exclusively climate investing, and not ESG investing. Nevertheless, the EU climate strategy supports a just transition mechanism. According to ILO, "a just transition means greening the economy in a way that is as fair and inclusive as possible to everyone concerned, creating decent work opportunities and leaving no one behind". This implies that the* **S** *and* **G** *pillars of ESG factors cannot be disregarded. For instance, to ensure a socially fair transition, the EC proposed to create a social climate fund of €144.4 bn. The objective of this fund is to protect the poorest citizens that are most impacted by energy and mobility costs.*

The implementation timeline, which is reported in Figures 1.12 and 1.12, demonstrates that the European ESG regulation is a continuous work in progress, implying that most of frameworks discussed below are not finished or can change.

1.4.1   EU taxonomy regulation

The purpose of a green financial taxonomy is to define what is green, and its objective is to inform investors about the greenness of their investments. Therefore, they can evaluate whether these levels satisfy or not their expectations. According to the European Commission[66], the EU taxonomy for sustainable activities is "*a classification system, establishing a list of environmentally sustainable economic activities. [...] The EU taxonomy would provide companies, investors and policymakers with appropriate definitions for which economic activities can be considered environmentally sustainable. In this way, it should create security for investors, protect private investors from greenwashing, help companies to become more climate-friendly, mitigate market fragmentation and help shift investments where they are most needed*". In this context, the EU taxonomy is a common base for other ESG regulations (BMR, SFDR, MiFID II, IDD, CSRD), acting as a "*common language*" around sustainable economic activities.

Developed by the Technical Expert Group on sustainable finance (TEG, 2020), the EU green taxonomy defines economic activities which make a substantive contribution to at least one of the following six environmental objectives:

1. Climate change mitigation

---

[65]The target of renewables share is set to 40% in 2030.
[66]See the EU website: https://ec.europa.eu/info/business-economy-euro/banking-and-finance/sustainable-finance_en.

Figure 1.12: Sustainable finance — implementation timeline (Reference ESMA34-45-1580)

Figure 1.13: Sustainable finance — implementation timeline (Reference ESMA34-45-1580)

SUSTAINABLE FINANCE
Implementation timeline for SFDR | TR | CSRD | MiFID | IDD | UCITS | AIFMD

Legend

- Taxonomy Regulation (TR) L1
- Taxonomy Regulation Article 8 Delegated Act (DA)
- Sustainable Finance Disclosures Regulation (SFDR) L1
- SFDR RTS - Joint ESAs draft Regulatory Technical Standards (RTS)
- MiFID and IDD DAs
- UCITS and AIFMD DAs
- Corporate Sustainability Reporting Directive (CSRD) – final text

- European Commission evaluation Reports
- ESAs Report on voluntary disclosures under SFDR

- **'First FMP PAI statement':** First reference period for the Financial Market Participant (FMP) first Principal Adverse Impact (PAI) statement on 30 June 2023 must be 1 Jan – 31 Dec 2022
- **'First two environmental objectives':** Point (a) (climate change mitigation) and point (b) (climate change adaptation) of environmental objectives under Art 9 TR
- **'All environmental objectives':** In addition to point (a) and (b) above, point (c) (the sustainable use and protection of water and marine resources), point (d) (the transition to a circular economy), point (e) (pollution prevention and control) and point (f) (the protection and restoration of biodiversity and ecosystems) of environmental objectives under Article 9 TR
- **'Art 5 and Art 6 TR':** Transparency of environmentally sustainable investments (Article 5) and of financial products that promote environmental characteristics (Article 6) in pre-contractual disclosures and in periodic reports
- **'Art 8 TR DA':** Transparency of undertakings in non-financial statements
- **'COM adopted DA bundling SFDR and TR RTSs':** COM bundled all 13 RTS of the SFDR, including the new empowerments for RTS introduced by the TR in one single DA (Commission Delegated Regulation EU 2022/1288)
- **Companies currently subject to NFRD** (Non-Financial Reporting Directive)**:** requirements apply to financial years (FYs) starting on / after 1 January 2024, first reporting in 2025
- **Large companies not currently subject to NFRD:** requirements apply to FYs starting on / after 1 January 2025, first reporting in 2026
- **Listed SMEs:** requirements apply to FYs starting on / after 1 January 2026, first reporting in 2027 (opt-out possible until 2029)
- **3rd country companies:** requirements apply to FYs starting on / after 1 January 2028, first reporting in 2029

ESMA34-45-1580

2. Climate change adaptation

3. Sustainable use and protection of water and marine resources

4. Transition to a circular economy

5. Pollution prevention and control

6. Protection and restoration of biodiversity and ecosystem

To qualify as sustainable, a business activity must also meet two other criteria. Indeed, the activity must do no significant harm to the other environmental objectives (DNSH constraint) and comply with minimum social safeguards[67] (MS constraint). Figure 1.14 summarizes the different steps.

Figure 1.14: EU taxonomy for sustainable activities

1a. SC	**Substantially contribute to at least one of the six objectives**
1b. TSC	Comply with **Technical Screening Criteria**
2. DNSH	**Do No Significant Harm** to any other five objectives
3. MS	Comply with **Minimum (Social) Safeguards**
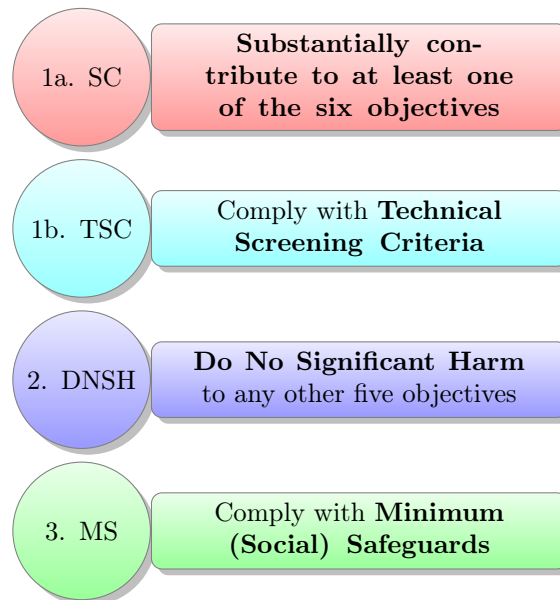
In Table 1.6, we have reported the activities eligible for the first two environmental objectives[68] (climate change mitigation and climate change adaptation). For instance, the activity "*Human health and social work activities*" is eligible for the adaptation objective, but not for the mitigation objective. For each activity, we have also indicated the number of sub-activities that are concerned[69]. For instance, the activity "*Financial and insurance activities*" has only two eligible sub-activities: #10.1 Non-life insurance: underwriting of climate-related perils and #10.2 Reinsurance. For each sub-activity, the taxonomy also indicates the corresponding NACE sectors[70], and the different criteria (technical screening and DNSH) for the eligibility certification[71].

---

[67]For example, the UN guiding principles on business and human rights.

[68]The finalization of the four other environmental objectives is expected on 1 January 2023 (Figure 1.12).

[69]When there are two numbers, the first one is for the mitigation objective whereas the second concerns the adaptation objective.

[70]NACE is the industry standard classification system used in the European Union. It is the abbreviation for the French term "*Nomenclature statistique des Activités économiques dans la Communauté Européenne*" (in English, statistical classification of economic activities in the European Community).

[71]All these informations can be found in the EU Taxonomy Compass Excel file, which corresponds to Annex 1 and Annex 2 of the Delegated Act on the climate objectives (Delegated Regulation2021/2139 of 4 June 2021).

Table 1.6: Activities eligible for the first two objectives (mitigation and adaptation)

Activity name	#	Objective (1)	(2)	Activity number
Arts, entertainment and recreation	13		✓	3
Construction and real estate	7	✓	✓	7
Education	11		✓	1
Energy	4	✓	✓	31
Environmental protection and restoration activities	2	✓	✓	1
Financial and insurance activities	10		✓	2
Forestry	1	✓	✓	4
Human health and social work activities	12		✓	1
Information and communication	8	✓	✓	2/3
Manufacturing	3	✓	✓	17
Professional, scientific and technical activities	9	✓	✓	3/2
Transport	6	✓	✓	17
Water supply, sewerage, waste management and remediation	5	✓	✓	12

Source: EU Taxonomy Compass, https://ec.europa.eu/sustainable-finance-taxonomy.

Remark 6 *The EU Taxonomy Regulation cannot be reduced to a green taxonomy. Indeed, the environmental taxonomy is the most advanced area, but the objective is to cover other topics. In particular, the development of brown and social taxonomies are currently discussed by the EC and European regulators.*

1.4.2   Climate benchmarks

In September 2019, the EU Technical Expert Group on sustainable finance (TEG) proposed to create two climate benchmark labels[72]: climate transition benchmark (CTB) and Paris aligned benchmark (PAB). These labels are structured along the following common principles:

1. A year-on-year self-decarbonization of 7% on average per annum, based on scope 1, 2 and 3 emissions;

2. A minimum carbon intensity reduction $\mathcal{R}^-$ compared to the investable universe;

3. A minimum exposure to sectors highly exposed to climate change.

For the CTB label, the minimum reduction $\mathcal{R}^-$ is set to 30% whereas it is equal to 50% for the PAB label. Other constraints are also imposed such as issuer exclusions (controversial weapons and societal norms violators), a minimum green revenue share ratio (or green-to-brown ratio[73]) or some activity exclusions. These climate labels are now part of the EU Benchmarks Regulation (BMR), which also specifies ESG disclosure requirements for all benchmarks[74]. In particular, an index sponsor must disclose whether its benchmarks pursue ESG objectives and provide an explanation of the methodology incorporating ESG factors used by these benchmarks.

---

[72]According to the TEG (2019a), "*a climate benchmark is defined as an investment benchmark that incorporates — next to financial investment objectives — specific objectives related to greenhouse gas (GHG) emission reductions and the transition to a low-carbon economy through the selection and weighting of underlying benchmark constituents*".

[73]The implementation of GRS or GTB ratios is delayed, because it requires a comprehensive definition of green/brown taxonomies.

[74]With the exception of interest rate and currency benchmarks.

1.4.3   Sustainable finance disclosure regulation

The SFDR is a European disclosure regulation[75] that applies at entity level and product level. It concerns websites of financial market participants (Article 4), remuneration policies in relation to the integration of sustainability risks (Article 5), the disclosure of principal adverse impacts (Article 7), the promotion of ESG on websites (Article 10), and periodic and annual reports (Article 11). The disclosure level depends on the ESG degree of the product and the following product/fund classification:

- Article 6 (or non-ESG products)
  It covers standard financial products that cannot be Article 8 or Article 9.

- Article 8 (or ESG products)
  It corresponds to financial products which "*promote, among other characteristics, environmental or social characteristics, or a combination of those characteristics, provided that the companies in which the investments are made follow good governance practices*".

- Article 9 (or sustainable products)
  In addition to the points covered by Article 8, these financial products have a sustainable investment objective.

For Article 8 and Article 9 products, the SFDR implies the disclosure of qualitative and quantitative ESG information (ex-ante requirements for KIID, prospectus and websites, and ex-post requirements for annual reports and MiFID client reports). In particular, pre-contractual documents need to indicate ex-ante minimum and planned percentage of sustainable investment (SI) according to following breakdown:

- SI with environmental objective

    - in economic activities that are taxonomy-aligned
    - in economic activities that are not taxonomy-aligned

- SI with social objective

They also need to indicate how the portfolio manager takes into account principal adverse impacts (PAI). Among the 64 PAI indicators, some of them are mandatory while other are voluntary. In Table 1.7, we report the 18 mandatory PAI indicators, which depend on the investment type (exposure on corporations, investment on sovereign and supranational securities, real estate assets). Beside these mandatory indicators, the SFDR RTS[76] also defines 22 and 24 optional PAI indicators for **E** and **S** pillars[77].

The first level (SFDR Level 1) has come into effect on 10 March 2021. It required FMPs to disclose general SFDR information at entity level and SFDR classification at product level. On 1st

---

[75]Regulation (EU) 2019/2088 of the European Parliament and of the Council of 27 November 2019 on sustainability-related disclosures in the financial services sector.

[76]In European Union, a Regulatory Technical Standard RTS is a delegated act, technical, prepared by a European Supervisory Authority. It should further develop, specify and determine the conditions for consistent harmonisation of the rules included in the basic legislative act. For instance, the SFDR RTS has been developed by ESMA, EBA, EIOPA and the ESAs' Joint Committee.

[77]The comprehensive list of PAI indicators and their associated metrics is given in Chapter 4 on page 205.

[78]Final Report on draft Regulatory Technical Standards with regard to the content, methodologies and presentation of disclosures pursuant to Article 2a(3), Article 4(6) and (7), Article 8(3), Article 9(5), Article 10(2) and Article 11(4) of Regulation (EU) 2019/2088

Table 1.7: The 18 mandatory PAI indicators

Corporates		
Climate and other environment-related indicators		
1	GHG emissions	
2	Carbon footprint	
3	GHG intensity of investee companies	
4	Exposure to companies active in the fossil fuel sector	
5	Share of non renewable energy consumption and production	
6	Energy consumption intensity per high impact climate sector	
7	Activities negatively affecting biodiversity sensitive areas	
8	Emissions to water	
9	Hazardous waste ratio	
Social and employee, respect for human rights, anti-corruption and anti-bribery matters		
10	Violations of UN Global Compact principles and OECD Guidelines for Multinational Enterprises	
11	Lack of processes and compliance mechanisms to monitor compliance with UN Global Compact principles and OECD Guidelines for MNEs	
12	Unadjusted gender pay gap	
13	Board gender diversity	
14	Exposure to controversial weapons (anti personnel mines, cluster munitions, chemical weapons and biological weapons)	
Sovereigns and supranationals		
Climate and other environment-related indicators		
15	GHG intensity	
Social and employee, respect for human rights, anti-corruption and anti-bribery matters		
16	Investee countries subject to social violations	
Real estate assets		
Climate and other environment-related indicators		
17	Exposure to fossil fuels through real estate assets	
18	Exposure to energy-inefficient real estate assets	

Source: SFDR RTS[78] (2 February 2021).

January 2023, SFDR Level 2 comes into effect, implying the publication of PAI indicators for Article 8 and Article 9 products. PAI reporting at the entity level for year 2022 must be published in June 2023.

Since August 2022, financial advisors (FAs) have to assess the sustainability preferences of their clients (MiFID II & IDD). For that, FinDatEx has developed the European ESG Template (EET) in order to facilitate the exchange of ESG-related data between market participants. The EET is an Excel file that contains qualitative information (e.g., fund's name, isin, currency, SFDR classification) and quantitative data, especially PAI indicators and taxonomy figures. The EET could be viewed as a SFDR/Taxonomy template.

1.4.4   MiFID II and sustainable preferences

MiFID is the Markets in Financial Instruments Directive (2004/39/EC). It has been applicable across the European Union to investment advice and portfolio management activity since November 2007.

Its aim is to standardize practices across the EU for investment services and activities and to ensure a high degree of harmonised protection for investors in financial instruments. MiFID II is a revised version of the original MiFID and came into force in 2018. It covers organisational requirements for investment firms, regulatory reporting to avoid market abuse, OTC trading, transparency of costs, etc.

Concerning investor protection, financial advisors must make a suitability and appropriateness assessment for individual portfolio management or advice regarding financial instruments. This implies FAs must obtain information from the client before it provides investment advice or individual portfolio management. The MiFID II Suitability Test includes questions about investors' knowledge and experience, their financial position, and their investment objectives. In September 2022, ESMA has published its guidelines on integrating ESG risks and factors in MiFID II (ESMA, 2022). There are two main consequences:

1. Integration of sustainability preferences to define the suitable product;

2. Integration of ESG criteria in the product governance.

The first point ensures that the product is in line with investors' values when providing financial advice and portfolio management services. This implies a new version of the suitability and appropriateness assessment (profiling questionnaire, suitability test, adequacy report). The second point covers the product offering of FMPs. Indeed, manufacturers and distributors must specify their target markets and the sustainability-related objectives with which the product is compatible.

"*Sustainability preferences*" is the key concept when selling an ESG product[79]. If the client has any sustainability preferences (yes/no), it has to choose one or a combination of the criteria below:

1. Minimum percentage in environmentally sustainable investments aligned to the EU Taxonomy.

2. Minimum percentage invested in sustainable investments as defined in the SFDR (Articles 8 and 9).

3. Quantitative/qualitative elements of principal adverse impacts defined by the client.

Once the choice is done, the financial adviser can sell a product to the client only after ensuring that the product matches the sustainability preferences of the client.

Remark 7 *The integration of sustainability preferences is not limited to financial investment products and MiFID II. It also applies to insurance-based investment products and the Insurance Distribution Directive (IDD).*

1.4.5 Corporate sustainability reporting directive

The Corporate Sustainability Reporting Directive (CSRD) makes mandatory for corporates to disclose sustainability information in their financial reports. It applies to all listed companies in the EU and all large European companies meeting at least two of the following criteria: (1) 250 employees, (2) €40 mn turnover and (3) €20 mn total assets. This represents about 50,000 corporates and 75% of total corporates' turnover in the EU. The CSRD will replace the NFRD and is planned to come into effect on 1 January 2025. According to EFRAG (2022), the sustainable reporting standards shall taking into account the following topics:

---

[79]Client's sustainability preferences are required since August 2022.

- (**E**)nvironmental factors: (1) climate change mitigation; (2) climate change adaptation; (3) water and marine resources; (4) resource use and circular economy; (5) pollution; (6) biodiversity and ecosystems.

- (**S**)ocial factors: (1) equal opportunities for all; (2) working conditions; (3) respect for human rights.

- (**G**)overnance factors: (1) role and composition of administrative, management and supervisory bodies; (2) business ethics and corporate culture, including anti-corruption and anti-bribery; (3) political engagements of the undertaking, including its lobbying activities; (4) management and quality of relationships with business partners.

For the environmental factors, we recognize the 6 objectives of the green taxonomy. It is no coincidence, given that CSRD must be in line with SFDR and EUTR. In fact, it must help financial institutions to compute the ESG and climate metrics in a more robust and effective manner. Beside current KPIs, the CSRD also requires the company to measure and assess its targets. For instance, this type of forward-looking information is helpful for investors to define their net zero investment policies. Nevertheless, the CSRD has another ambition by considering double materiality. It is a first step in developing a comprehensive extra-financial/climate accounting statement[80]. Materiality is an accounting principle which states that an information on a company is material if it is reasonably likely to impact investors' decision-making. This is why it must be recorded or reported in financial statements. It is now widely accepted that climate-related impacts on a company can be material and therefore require disclosure. This approach is called financial or single materiality. The concept of double materiality is an extension of the single materiality by also considering the negative externalities of the company. In this case, we must consider two materiality perspectives:

- How sustainable factors impact the financial value of the company?

- How the company affects the environment, the society and people?

The first one corresponds to the financial (or outside-in) materiality, while the second defines the impact (or inside-out) materiality. For example, the SASB framework is based on the financial materiality. On the contrary, the GRI framework has adopted an inside-out materiality by reporting companies' impact on people and the planet. In the case of the CSRD, EFRAG has made the choice to consider the double materiality assessment.

1.5   The market of ESG investing

In this section, we present a global overview of the ESG market from the investment viewpoint. First, we define the different ESG strategies and provide some examples. Then, we give some figures about the ESG market and its growth.

1.5.1   ESG strategies

In Figure 1.15, we define the different types of ESG strategies. This list is based on several reports (Eurosif, 2018; GSIA, 2021; PRI, 2020). Depending on the region and the organization, these

---

[80]The CSRD standards are mainly defined by EFRAG (European Financial Reporting Advisory Group), which represents the European accounting profession and the European voice in financial reporting. In particular, it participates in IASB's standard setting process and develops the European views concerning international accounting standards (IFRS).

Figure 1.15: The 7 categories of ESG strategies

1. Exclusion	**Exclusion policy & negative (or worst-in-class) screening**
2. Values	**Norms-based screening**
3. Selection	**Positive (or best-in-class) screening**
4. Thematic	**Sustainability themed investing (e.g. green bonds)**
5. Integration	**ESG scoring is fully integrated in portfolio management**
6. Engagement	**Voting policy & shareholder activism**
7. Impact	**Impact investing**

Source: Eurosif (2021), https://www.eurosif.org/responsible-investment-strategies.

Table 1.8: Comparison of Eurosif, GSIA and PRI classifications

#	Eurosif	GSIA	PRI
1	Exclusions	Negative/exclusionary screening	Negative screening
2	Norms-based screening	Norms-based screening	Norms-based screening
3	Best-in-class	Best-in-class/positive screening	Positive screening
4	Sustainability themed	Sustainability themed/thematic investing	Thematic
5	ESG integration	ESG integration	Integration of ESG issues
6	Engagement & voting	Corporate engagement & shareholder action	Engagement/proxy voting
7	Impact investing	Impact/community investing	Sustainability impact

categories can take different names. For instance, in the case of the seventh category, the term "*community investing*" is extensively used in North America (RIA Canada, US SIF), while we prefer the term "*microfinance*" in Europe.

Most of these categories are based on the screening concept, which refers to investment filters. Negative screening is an approach that excludes specific investments or classes of investment from the investible universe such as companies, sectors, or countries that do not comply with specific ESG criteria. When applied to companies, it is also called the worst-in-class exclusion strategy. In this case, the investor can systematically exclude issuers that have the worst rating grade (e.g., companies rated CCC). This category also concerns sector or sub-industry exclusion (e.g., coal & consumable fuels, fossil fuels production, conventional weapons, civilian firearms). This approach can also include certain activities, because they do not comply with the values of the investor (e.g., pornography, tobacco, alcohol, gambling, genetically modified). Most of the time, the investor define an exclusion list of individual issuers.

Norm-based screening consists in excluding companies that have been called into question because they have violated international standards and norms on social or environmental issues, such as those issued by the OECD, ILO, UN Global Compact and UNICEF. The first category is closed to this second category, but this later is based on international values while the former is based on individual values. For instance, a company, which complies with all the minimum standards of business practice based on international norms, can be excluded in the negative screening approach because it has a very bad ESG rating or it belongs to a sector that the investor does not want to finance. By considering the first two categories, the top exclusion criteria in Europe are (1) controversial weapons (Ottawa and Oslo treaties), (2), tobacco, (3) all weapons, (4) gambling, (5) pornography, (6) nuclear energy, (7) alcohol, (8) GMO[81] and (9) animal testing (Eurosif, 2018).

The third category invests in issuers, sectors, or projects selected for positive ESG performance relative to industry peers. This is why the selection category is also called the positive screening approach. For example, the best-in-class ESG strategy selects issuers with the best ESG ratings (e.g. AAA, AA and A), while the ESG momentum strategy selects issuers that have improved their ESG rating.

The aim of sustainability themed investing is to invest in companies whose activity is related to sustainability, for example clean energy, green technology, sustainable agriculture, sustainable infrastructure, natural ressources, biodiversity. ESG thematic investing considers all the ESG issues, not only the environmental pillar. For example, it concerns investment in social topics (e.g., health, food security, diversity) and governance topics (e.g., gender equality, inclusive boards). Generally, these thematic investments are implemented in mutual funds.

ESG integration means the systematic and explicit inclusion by investment managers of ESG factors into financial analysis and asset allocation. This strategy can be viewed as an extension of the exclusion and best-in-class strategies. For example, the stock or bond picking score may be a mix of the fundamental score and the ESG score. Some asset managers also impose funds to have an ESG score greater than the ESG score of their benchmarks.

The sixth category uses shareholder power and active ownership to influence corporate behavior, including through direct corporate engagement (i.e., communicating with senior management and/or boards of companies), filing or co-filing shareholder proposals, and proxy voting that is guided by ESG guidelines. Examples of engagement activities are voting policy, public divestment, engagement with target companies on a specific subject (e.g., pay ratio, living wage), proposing shareholder resolutions, public litigation, etc.

---

[81]Genetically modified organism.

Impact investing are investments made with the intention to generate positive, measurable social and environmental impact alongside a financial return. Contrary to thematic investing which mainly considers stocks and bonds of companies, impact investing considers assets and securities financing specific projects. For example, impact investing includes microfinance, community investing, social entrepreneurship funds, funds with a social impact objective, green and social bonds. Since the goal is to achieve positive, social and environmental impacts, this requires measuring and reporting against these extra-financial impacts. Extra-financial reporting is then the key element of impact investing, because it must clearly define and measure the ESG objectives (e.g., GHG avoided emissions per €1 mn invested per year, percentage of water consumption saving).

Remark 8 *We notice that these seven categories can be split into two strategy groups. The first one considers ESG scores when building an investment portfolio (exclusion, selection, thematic and integration). The main objective of these strategies remains the financial performance of the portfolio. The second group places a high priority on ethical conduct (norm-based screening, engagement and impact investing) and can be related to the signaling theory. In this approach, investors send a negative signal to the market and corporations when they apply norm-based screening or they engage with a company[82]. On the contrary, investors send a positive signal when they implement impact investing.*

1.5.2 The market share of ESG investing

In this section, we present a global view of the market growth of ESG investing[83] based on the Global Sustainable Investment Reviews (GSIA, 2017, 2019, 2021). For each report, Figures 1.16, 1.17 and 1.18 gives the AUM of responsible investing, the corresponding market share, the global growth and the breakdown by countries. for each report. According to GSIA (2021), sustainable investments represented \$35.3 tn of assets under management at the start of 2020, representing a market share of 36%. They are continuing to grow in most regions, with Canada experiencing the largest increase (48% growth), followed by the United States (42% growth). If we consider a regional analysis, the regional market share of sustainable investments is equal to 62% in Canada, 42% in Europe, 38% in Australasia, 33% in the United States and 24% in Japan.

Remark 9 *The case of Europe (13% decline in the growth of sustainable investment assets in 2018 to 2020) is due to a changed measurement methodology from which European data are collected, and European regulations, especially the SFDR. Therefore, we must be cautious before drawing conclusions. Indeed, these data do not take into account how ESG factors are really implemented.*

Figure 1.19 shows changes in asset values by ESG categories. For many years, negative screening and exclusion dominates the other strategies. Since 2020, ESG integration has become the most implemented approach. We also notice that some categories are less represented: thematic investing, best-in-class/selection and impact investing. However, we observe that thematic investing is the category with the highest growth of asset values[84].

If we focus on asset ownership, the ESG market is mainly driven by institutional investors. They currently represent 75% of the market, while the remaining part (25%) corresponds to retail assets. Nevertheless, we observe a basic shift, because the split was 89%/11% in 2012 (GSIA, 2021, page 13).

---

[82]Because the investor is not satisfied by the environmental, social or governance policies of the company.

[83]The figures concerning specific segments (mutual funds, ETFs, green financing, etc.) are given in Chapters 2 and 10.

[84]In the last study period, but also since 2014.

Figure 1.16: Sustainable investment assets at the start of 2016
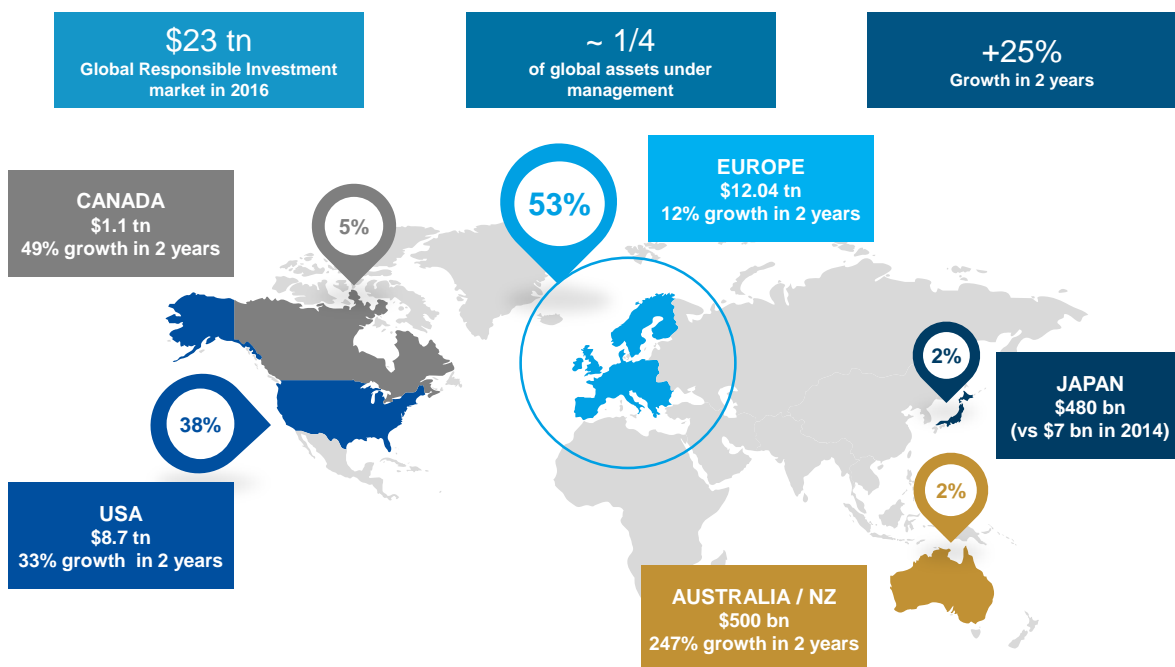
Source: GSIA (2017).

Figure 1.17: Sustainable investment assets at the start of 2018
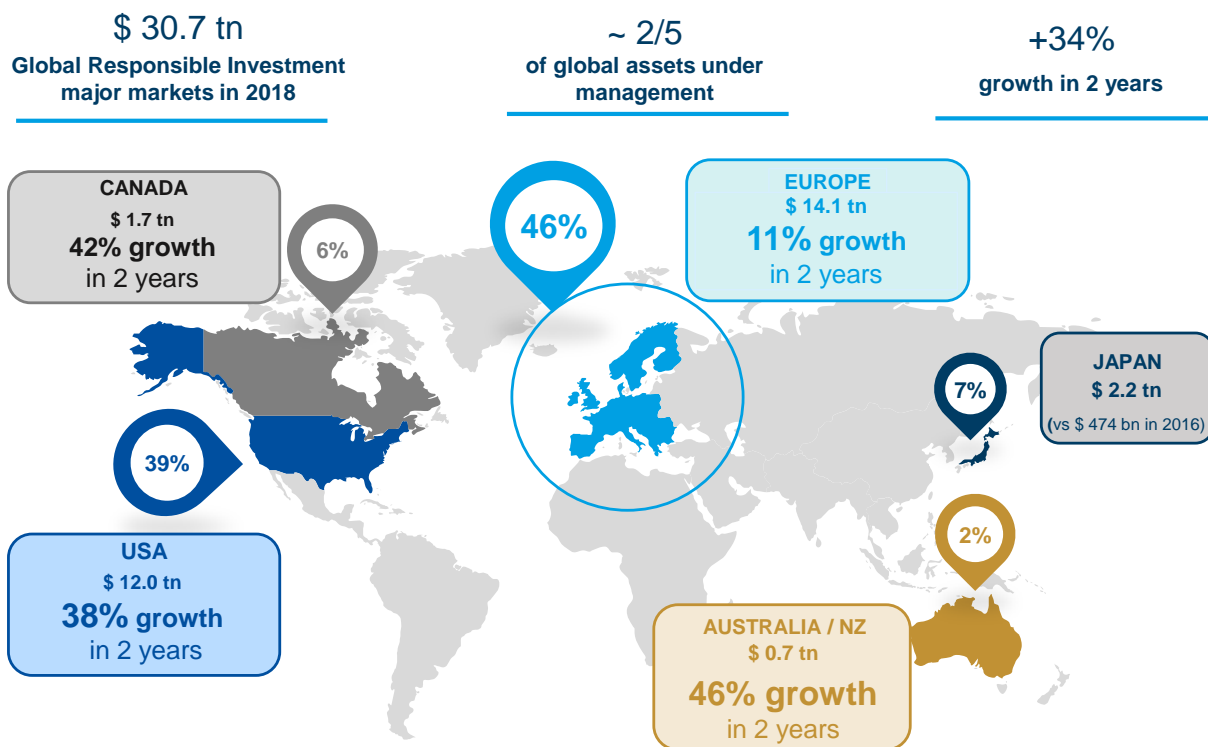
Source: GSIA (2019).

Figure 1.18: Sustainable investment assets at the start of 2020

| **$35.3tn** | **35.9%** | **+15%** |
| Global RI market in 2020 | of total global AUM | Growth in 2 years |

7%*

34%*

EUROPE
$12tn
-13% growth
in 2 years

8%*

JAPAN
$2.8tn
34% growth
in 2 years

48%*

CANADA
$2.4tn
+48% growth in
2 years

AUS / NZ
$906bn
25% growth in
2 years

3%*

USA
$17tn
+42% growth in
2 years
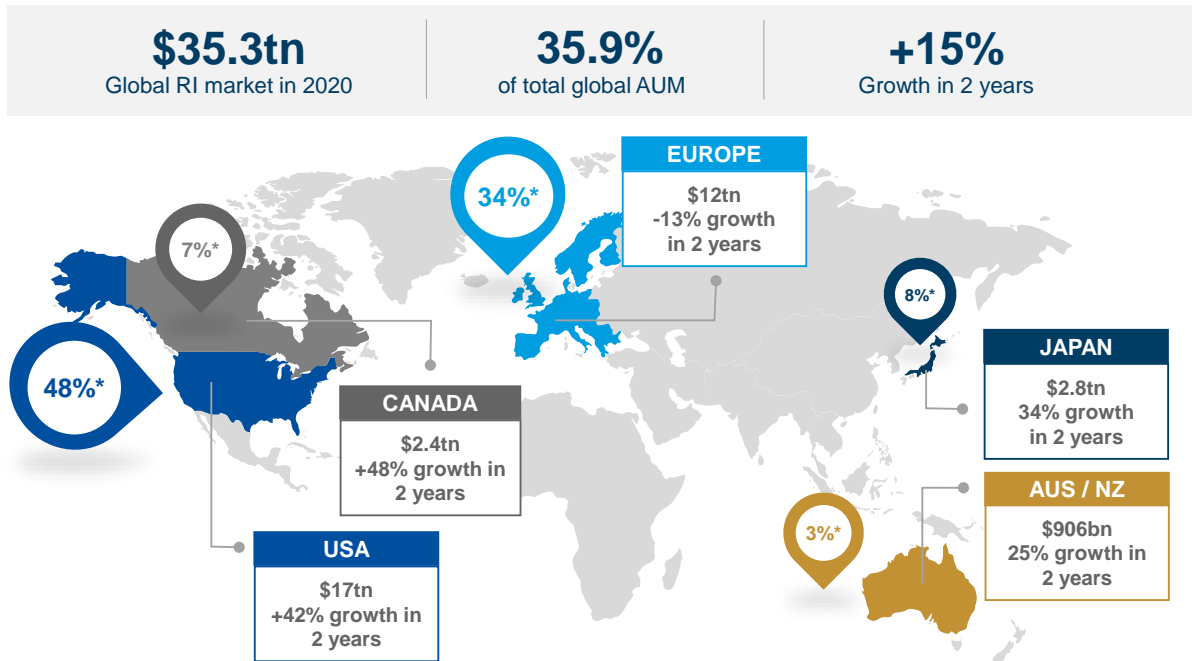
Source: GSIA (2021).

Figure 1.19: Asset values of ESG strategies between 2014 and 2018

Asset values (in $ tn)

Source: GSIA (2015, 2017, 2019, 2021) & Author's calculations.

Table 1.9: ESG asset growth

#	ESG strategy	Asset growth			2020 AUM (in $ bn)
		2014-2016	2016-2018	2018-2020	
1	Exclusion	11.7%	14.6%	−24.0%	15 030
2	Values/Norms-based	19.0%	−13.1%	−11.5%	4 140
3	Selection	7.6%	50.1%	−24.9%	1 384
4	Thematic Investing	55.1%	92.0%	91.4%	1 948
5	Integration	17.4%	30.2%	43.6%	25 195
6	Engagement	18.9%	8.3%	6.8%	10 504
7	Impact Investing	56.8%	33.7%	−20.8%	352

Source: GSIA (2015, 2017, 2019, 2021) & Author's calculations.

1.6   Conclusion

This little-boring introduction provides a global overview of the sustainability landscape. As a student, you need to understand who does what, the outlines of the various regulations, and some figures about sustainable finance. However, a course in sustainable finance cannot only be summarized by a set of acronyms (Figure 1.20). We also need to understand the data we are manipulating. Since the data are very noisy and not exhaustive, probability and quantitative modeling is important for two main reasons. The first is to estimate the limits of what professionals do. The second reason is to measure the relationships between the various factors from an ex-ante point of view, and to make some convictions, because data and numbers can lie to us.

Figure 1.20: The ESG world of acronyms

Part I

ESG Risk

Chapter 2

ESG Scoring

To develop ESG analysis, we need extra-financial data provided by companies, reporting frameworks, academic institutions, research centers, international bodies, etc. In general, this heterogeneous data is collected by ESG data providers. These data are available for two levels of use. First, they are widely used by ESG analysts to assess the sustainability risks of companies and countries. Second, they provide the raw material for ESG scoring systems. Like credit scoring models, such systems are of paramount importance for risk assessment and decision making. In the case of credit scoring, the decision is whether or not to grant credit. Therefore, credit scoring models are at the center of the lending process. In the case of ESG scoring, the issue is a bit different. Of course, we can use ESG scores to decide whether or not to invest in a company, but exclusion is not the only strategy, as we saw in the previous chapter. For example, ESG scores are fully embedded in the strategy of ESG integration. In this approach, they play the role of screening rules for portfolio selection. ESG scoring is therefore more than a traditional scoring model. Nevertheless, the analogy between ESG scoring and credit scoring remains essential and poses several challenges in terms of performance evaluation, score consistency and backtesting. This is particularly true as ESG risk ratings are generated from these scoring systems. From this perspective, the concept of ESG model validation takes on a new dimension. We recall that any internal risk model must comply with an independent model validation process that is highly binding and formal from a regulatory standpoint (FRB, 2011; EBA, 2022). Moreover, the validation process is not limited to credit, market, operational and liquidity risks. For example, it also applies to compliance risks: statistical models developed for anti-money laundering detection, transaction monitoring, anomaly detection scenarios, list filtering, etc. The systematic validation approach of any model (risk-based, behavioral, rules-based and AI-based) has recently been reinforced in the US with the interagency guidance on model risk management (FRB, 2021). As ESG scoring models become more widely used by financial institutions, we can easily predict that they will be regulated in the near future, just like the other risk models. Therefore, in this chapter, we adopt the basic idea that a scoring system needs to be validated from an ex-post viewpoint. To do so, we make extensive use of the mathematical and statistical tools available to us in the fields of scoring theory and Markov-based rating methods (Roncalli, 2020a, Section 3.3.3 and Chapter 15). This chapter is then organized as follows. Section 1 presents the ESG data and the sources of extra-financial information. The construction of ESG scores and the performance evaluation of ESG scoring models are discussed in Section 2. Finally, we examine ESG rating systems and assess the consistency of ESG migration matrices.

2.1   Data and variables

In this section, we list the most important variables or indicators that are used in an ESG scoring system. For that, we distinguish between sovereign and corporate data since the sources are not the same and their access is more or less easy. As the adage says that "*we can only measure what we can define*", we must first specify the meaning of the three ESG factors, because the objective of an ESG score is to measure the risk and opportunities of an entity with respect to environmental, social and governance dimensions. If we consider the definition on page 1, we notice that each factor is defined by encompassing several issues:

(E) climate change mitigation, climate change adaptation, preservation of biodiversity, pollution prevention, circular economy;

(S) inequality, inclusiveness, labor relations, investment in human capital and communities, human rights;

(G) management structure, employee relations, executive remuneration.

Of course, this list is non-exhaustive and must be adapted to show the difference between sovereign and corporate entities. Let us consider an example. According to the Universal Declaration of Human Rights[1], States have obligations and duties under international law to respect, protect and fulfill human rights:

> "*The obligation to respect means that States must refrain from interfering with or curtailing the enjoyment of human rights. The obligation to protect requires States to protect individuals and groups against human rights abuses. The obligation to fulfill means that States must take positive action to facilitate the enjoyment of basic human rights.* "

For a sovereign, the issue of human rights concerns both social (e.g., access to health and education, labor rights) and governance (e.g., safeguarding of civil and political rights) pillars, while it is more related to the social pillar (e.g., ethical supply chain, employment conditions) of a company. Therefore, it is certainly easier to define the three dimensions with examples.

2.1.1   Sovereign ESG data

The World Bank framework

The World Bank database dedicated to sovereign ESG indicators is certainly the easiest way to understand the most common topics relevant to ESG analysis. The database is available at the following webpage: https://datatopics.worldbank.org/esg. It contains 67 ESG indicators grouped into 17 themes. In Table 2.1, we have reported these categories and the corresponding number of indicators. For example, the **(E)** pillar is made up of 5 categories, such as emissions and pollution that contains 5 indicators. In total, we have 5 **(E)**, 6 **(S)** and 6 **(G)** themes. We observe that global warming and its consequences are the main drivers of the environmental pillar. If we analyse the 27 indicators (Table 2.2), two categories are related to the measurement of climate change (emissions & pollution, environment/climate risk & resilience), one category is related to the mitigation risks of climate change (natural capital endowment & management, energy use & security), and the last category concerns the impact of climate change on food security. If we focus on the **(S)** pillar, the sources of social risks are related to inclusiveness and inequality (education & skills, poverty & inequality, health & nutrition, access to services), in particular the literacy rate, the school enrollment,

---

[1]See https://www.ohchr.org/en/ohchr_homepage.

Table 2.1: The World Bank database of sovereign ESG indicators

Environmental (27)

- Emissions & pollution (5)
- Natural capital endowment & management (6)
- Energy use & security (7)
- Environment/climate risk & resilience (6)
- Food security (3)

Social (22)

- Education & skills (3)
- Employment (3)
- Demography (3)
- Poverty & inequality (4)
- Health & nutrition (5)
- Access to services (4)

Governance (18)

- Human rights (2)
- Government effectiveness (2)
- Stability & rule of law (4)
- Economic environment (3)
- Gender (4)
- Innovation (3)

Table 2.2: Indicators of the environmental pillar (World Bank database)

- **Emissions & pollution** (1) CO_2 emissions (metric tons per capita); (2) GHG net emissions/removals by LUCF (Mt of CO_2 equivalent); (3) Methane emissions (metric tons of CO_2 equivalent per capita); (4) Nitrous oxide emissions (metric tons of CO_2 equivalent per capita); (5) PM2.5 air pollution, mean annual exposure (micrograms per cubic meter);

- **Natural capital endowment & management**: (1) Adjusted savings: natural resources depletion (% of GNI); (2) Adjusted savings: net forest depletion (% of GNI); (3) Annual freshwater withdrawals, total (% of internal resources); (4) Forest area (% of land area); (5) Mammal species, threatened; (6) Terrestrial and marine protected areas (% of total territorial area);

- **Energy use & security**: (1) Electricity production from coal sources (% of total); (2) Energy imports, net (% of energy use); (3) Energy intensity level of primary energy (MJ/$2011 PPP GDP); (4) Energy use (kg of oil equivalent per capita); (5) Fossil fuel energy consumption (% of total); (6) Renewable electricity output (% of total electricity output); (7) Renewable energy consumption (% of total final energy consumption);

- **Environment/climate risk & resilience**: (1) Cooling degree days (projected change in number of degree Celsius); (2) Droughts, floods, extreme temperatures (% of population, average 1990-2009); (3) Heat Index 35 (projected change in days); (4) Maximum 5-day rainfall, 25-year return level (projected change in mm); (5) Mean drought index (projected change, unitless); (6) Population density (people per sq. km of land area)

- **Food security**: (1) Agricultural land (% of land area); (2) Agriculture, forestry, and fishing, value added (% of GDP); (3) Food production index (2004-2006 = 100);

Source: https://datatopics.worldbank.org/esg/framework.html.

Table 2.3: Indicators of the social pillar (World Bank database)

- **Education & skills**: (1) Government expenditure on education, total (% of government expenditure); (2) Literacy rate, adult total (% of people ages 15 and above); (3) School enrollment, primary (% gross);

- **Employment**: (1) Children in employment, total (% of children ages 7-14); (2) Labor force participation rate, total (% of total population ages 15-64) (modeled ILO estimate); (3) Unemployment, total (% of total labor force) (modeled ILO estimate);

- **Demography**: (1) Fertility rate, total (births per woman); (2) Life expectancy at birth, total (years); (3) Population ages 65 and above (% of total population);

- **Poverty & inequality**: (1) Annualized average growth rate in per capita real survey mean consumption or income, total population (%); (2) Gini index (World Bank estimate); (3) Income share held by lowest 20%; (4) Poverty headcount ratio at national poverty lines (% of population);

- **Health & nutrition**: (1) Cause of death, by communicable diseases and maternal, prenatal and nutrition conditions (% of total); (2) Hospital beds (per 1,000 people); (3) Mortality rate, under-5 (per 1,000 live births); (4) Prevalence of overweight (% of adults); (5) Prevalence of undernourishment (% of population);

- **Access to services**: (1) Access to clean fuels and technologies for cooking (% of population); (2) Access to electricity (% of population); (3) People using safely managed drinking water services (% of population); (4) People using safely managed sanitation services (% of population);

Table 2.4: Indicators of the governance pillar (World Bank database)

- **Human rights**: (1) Strength of legal rights index (0 = weak to 12 = strong); (2) Voice and accountability (estimate);

- **Government effectiveness**: (1) Government effectiveness (estimate); (2) Regulatory quality (estimate);

- **Stability & rule of law**: (1) Control of corruption (estimate); (2) Net migration; (3) Political stability and absence of violence/terrorism (estimate); (4) Rule of law (estimate)

- **Economic environment**: (1) Ease of doing business index (1 = most business-friendly regulations); (2) GDP growth (annual %); (3) Individuals using the internet (% of population);

- **Gender**: (1) Proportion of seats held by women in national parliaments (%); (2) Ratio of female to male labor force participation rate (%) (modeled ILO estimate); (3) School enrollment, primary and secondary (gross), gender parity index (GPI); (4) Unmet need for contraception (% of married women ages 15-49);

- **Innovation**: (1) Patent applications, residents; (2) Research and development expenditure (% of GDP); (3) Scientific and technical journal articles;

Source: https://datatopics.worldbank.org/esg/framework.html.

the Gini index[2], the income share held by the lowest 20%, etc. We also notice that the integration of some indicators from the categories employment and demography is disturbing. Indeed, we may wonder how the fertility rate is related to the social pillar. For instance, does a high fertility rate increase or decrease social risk? The **G** pillar includes two classical governance categories (government effectiveness, stability & rule of law), two economic categories (economic development, innovation) and two social-based categories (human rights, gender). In this last case, the frontier between social and governance is blurred. For instance, we can classify the four indicators of the gender category in the social pillar as a non-discrimination category.

Remark 10 *The definition of each indicator can be found on the website* https://datatopics.worldbank.org/esg/framework.html. *Most of these variables are intuitive and easy to understand. Some of them are more technical and less comprehensible, especially some technical variables of the governance pillar. Therefore, we report in footnotes the definition provided by the World Bank for the following indicators: strength of legal rights index*[3]; *voice and accountability*[4], *government effectiveness*[5]; *regulatory quality*[6]; *rule of law*[7].

Certainly, one of the difficulties when building an ESG score is the data gathering, which requires the use of many internal and external sources. In the case of the World Bank framework, the data comes from[8]:

- National accounts statistics collected by Eurostat, OECD (https://www.oecd-ilibrary.org/statistics), United Nations Statistics Division (UNSD, https://unstats.un.org) and the World Bank;

- Internal departments and specialized databases of the World Bank: World Bank Open Data (https://data.worldbank.org), Business Enabling Environment (BEE), Climate Change Knowledge Portal (CCKP, https://climateknowledgeportal.worldbank.org), Global Database of Shared Prosperity (GDSP), Global Electrification Database (GEP), and Poverty and Inequality Platform (PIP, https://pip.worldbank.org);

- International organizations: Emission Database for Global Atmospheric Research (EDGAR, https://edgar.jrc.ec.europa.eu), Food and Agriculture Organization (FAO, https://www.fao.org/faostat), International Energy Agency (IEA, https://www.iea.org/data-and-statistics), International Labour Organization (ILO, https://ilostat.ilo.org), International Renewable Energy Agency (IRENA, https://www.irena.org/Data), UNESCO Institute for Statistics (UIS, http://uis.unesco.org), United Nations Population

---

[2]The Gini index is a measure of income inequality among individuals. It is based on the comparison of cumulative proportions of the population against cumulative proportions of income they receive, and it ranges between 0 in the case of perfect equality and 1 in the case of perfect inequality. Its computation is derived from the Lorenz curve.

[3]"*Strength of legal rights index measures the degree to which collateral and bankruptcy laws protect the rights of borrowers and lenders and thus facilitate lending*".

[4]"*Voice and accountability captures perceptions of the extent to which a country's citizens are able to participate in selecting their government, as well as freedom of expression, freedom of association, and a free media*".

[5]"*Government effectiveness captures perceptions of the quality of public services, the quality of the civil service and the degree of its independence from political pressures, the quality of policy formulation and implementation, and the credibility of the government's commitment to such policies*".

[6]"*Regulatory quality captures perceptions of the ability of the government to formulate and implement sound policies and regulations that permit and promote private sector development*".

[7]"*Rule of law captures perceptions of the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence*".

[8]The list is not exhaustive.

Division (DESA, `https://population.un.org`), World Health Organization (WHO, `https://www.who.int/data`), World Intellectual Property Organization (WIPO, `https://www.wipo.int/ipstats`);

- National agencies and non-governmental organizations: Climate Watch (`https://www.climatewatchdata.org`), Netherlands Environmental Assessment Agency (PBL, `https://www.pbl.nl`), World Database on Protected Areas (WDPA, `https://www.protectedplanet.net/en/thematic-areas/wdpa`);

- Academic resources: Kaufmann *et al.* (2010), Cohen *et al.* (2017), and the international disasters database (EM-DAT) of the Centre for Research on the Epidemiology of Disasters (CRED, Université Catholique de Louvain).

Some of these databases are more relevant than others. If we would like to focus on a small number, our preferences are CCKP, EDGAR and Climate Watch for the **(E)** pillar, FAO, ILO and WHO for the **(S)** pillar, and the Worldwide Governance Indicators (WGI, `https://info.worldbank.org/governance/wgi` produced by Daniel Kaufmann and Aart Kraay for the **(G)** pillar.

Table 2.5: Sovereign ESG taxonomy

Environmental	Social	Governance
• Biodiversity & land use • CO_2e emissions • Compliance with environnemental standards • Energy security & renewables • Emissions reduction targeting • Food security • Fossil fuel dependency • Green economy • Physical risk exposure • Pollution & waste management • Temperature • Transition risk • Water management	• Civil unrest • Demography • Education • Gender • Health • Income inequality & poverty • Labour rights & working conditions • Living standards • Migration • Human rights & local communities • Non-discrimination • Social cohesion • Water and electricity access	• Business & economic environment • Corruption & money laundering • Governance effectiveness • Infrastructure and mobility • International relations • Justice • National security • Political stability & institutional strength • Personal freedom & civil liberties • Rights of shareholders • Rule of law

Source: Author's research based on the works of Bouyé and Menville (2021),
Gratcheva *et al.* (2020) and Semet *et al.* (2021).

Other frameworks

Most ESG rating agencies provide sovereign ESG data. The most known are FTSE (Beyond Ratings), Moody's (Vigeo-Eiris), MSCI, Sustainalytics and RepRisk. One of the most comprehensive databases is certainly Verisk Mapplecroft (https://www.maplecroft.com), which is a global company covering country risk. If we make the union of the different categories, we obtain a taxonomy that looks like the one in Table 2.5. There are many categories, much greater than for the World Bank framework or the PRI taxonomy[9]. If we consider the indicators, the number of variables is large, much greater than 400. Nevertheless, as explained by Bouyé and Menville (2021) and Semet *et al.* (2021), they are highly correlated. If we perform a principal component analysis, there are few independent dimensions (less than 10). In fact, many of these indicators are correlated to the GDP. For instance, Gratcheva *et al.* (2020) found that the average correlation between sovereign ESG scores and national income is equal to 81% for aggregate ESG, 51% for (E) pillar, 85% for (S) pillar, and 70% for (G) pillar. If we consider correlations for ESG providers, the lowest correlations are obtained for the E pillar of ISS (7%), MSCI (10%) and V.E (23%), and the G pillar of RepRisk (37%) and V.E (39%), but they are generally high. Therefore, although all these providers use very different indicators, we notice a relative convergence between them[10].

Table 2.6: Correlation of ESG scores with country's national income (GNI per capita)

Factor	ESG	E	S	G
ISS	68%	7%	86%	77%
FTSE (Beyond Ratings)	91%	74%	88%	84%
MSCI	84%	10%	90%	77%
RepRisk	78%	79%	75%	37%
RobecoSAM	89%	82%	85%	85%
Sustainalytics	95%	83%	94%	93%
V.E	60%	23%	79%	39%
Total	81%	51%	85%	70%

Source: Gratcheva *et al.* (2020, Table 3.1, page 32).

The mushrooming growth of data

We observe among ESG data providers a mushrooming of indicators and data sources. This concerns the first well-established variables. For a very long time, income inequality was mainly measured by the Gini coefficient or the Lorenz curve, even if there were many other academic measures. It seems that data providers have recently rediscovered and embraced the academic literature. Thus, income inequality may also be measured by the Palma ratio, the S80/S20 (or 20:20) ratio, the Atkinson index, the percentile ratios (P90/P10, P90/P50, P50/P10), the Pietra index, the coefficient of variation or the Theil index. Nevertheless, the mushrooming growth of data mainly concerns non-economic variables. We provide some examples in Figures 2.1–2.4 with palm oil production and consumption,

---

[9]PRI (2019a) identifies 4 environmental factors (natural resources, physical risks, energy transition risk, energy security), 4 social factors (demographic change, education and human capital, living standards and income inequality, social cohesion) and 4 governance factors (institutional strength, political stability, government effectiveness, regulatory effectiveness).

[10]Gratcheva *et al.* (2020, Table 2.3, page 27) found that the average cross-correlation between these providers is equal to 85% for the ESG score, 42% for the environmental score, 85% for the social score and 71% for the governance score. These results are confirmed by the study of Bouyé and Menville (2021, Table 4, page 14).

Figure 2.1: Palm oil production (2019)

Source: Our World in Data, https://ourworldindata.org/palm-oil.
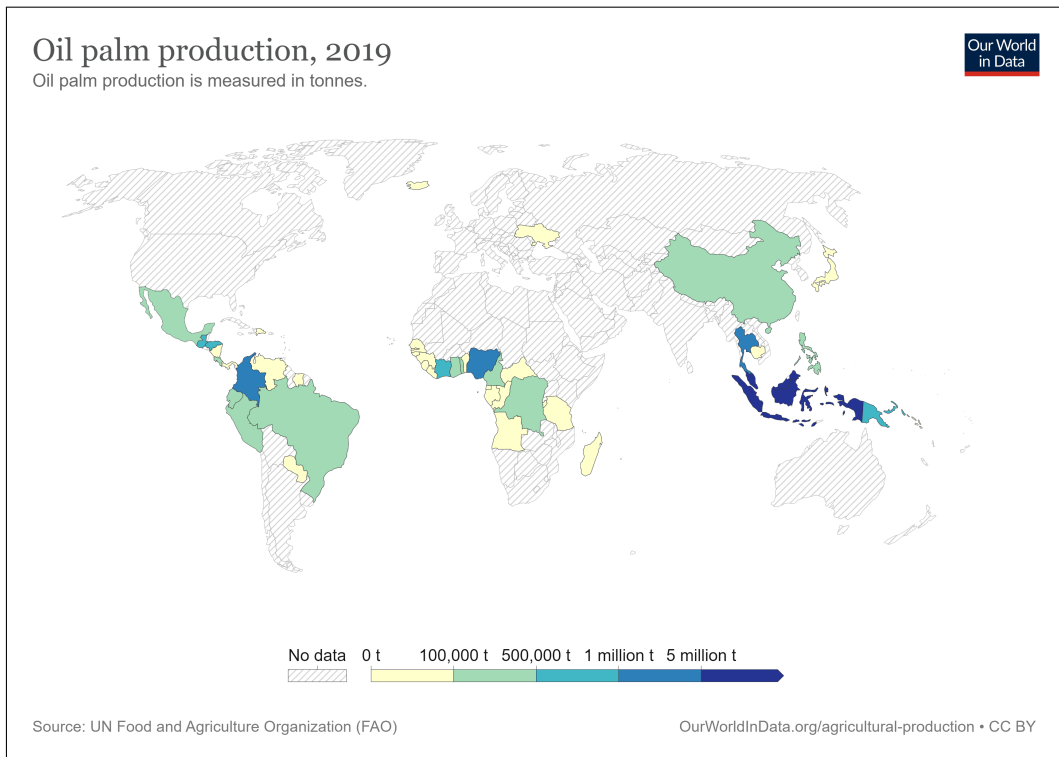
Figure 2.2: Palm oil imports (2019)

Source: Our World in Data, https://ourworldindata.org/palm-oil.
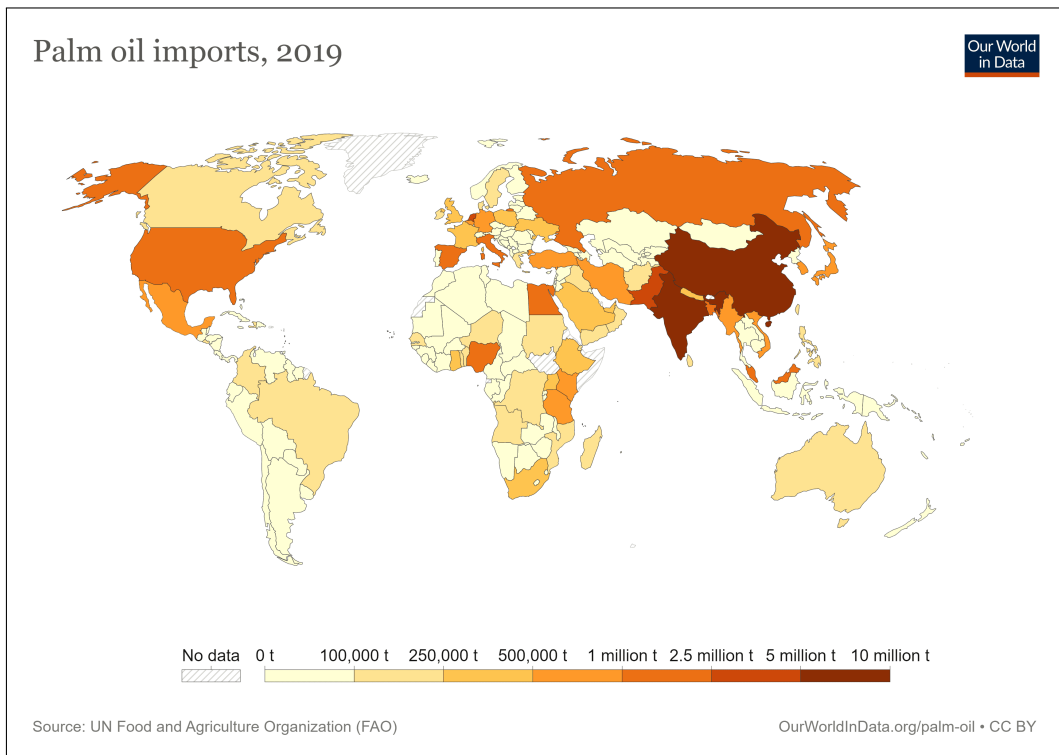
Figure 2.3: Share of global annual deforestation (2015)
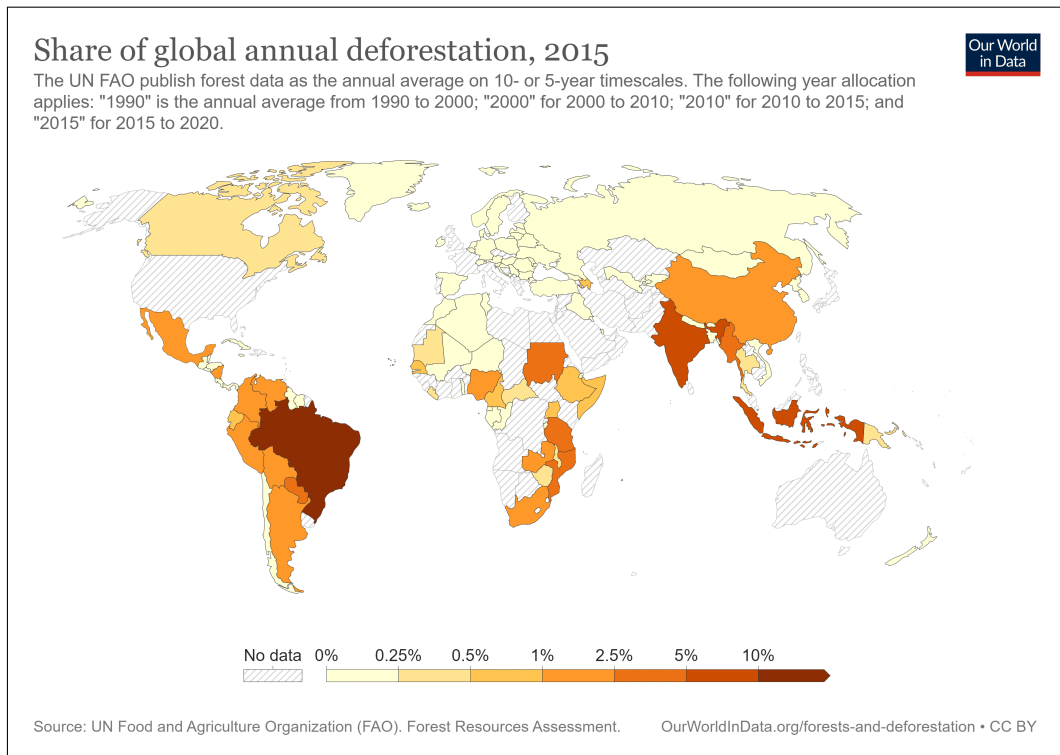
Figure 2.4: Threatened mammal species (2018)
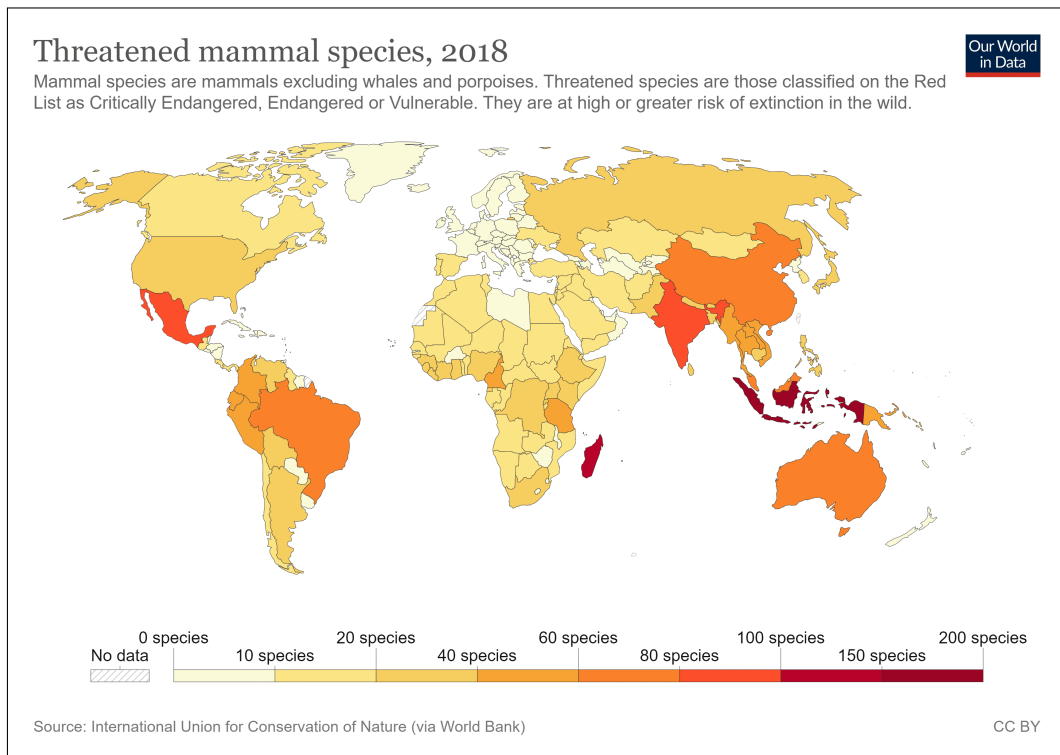
deforestation and threatened mammal species. In particular, we notice the increasing use of geo-location data, real-time data, or satellite data, for example, the data provided by the World Resources Institute (WRI) and its different data platforms (https://www.wri.org/data/data-platforms). The most interesting are Ocean Watch (data on ocean economies and management), Aqueduct (cutting-edge data to identify and evaluate water risks), Global Forest Watch (data on forest economies and management) and LandMark (global data of indigenous and community lands). For instance, we can collect data on coastal eutrophication risk, mangrove extent change, coral reef locations, seagrass, salt marshes, soil erosion, chlorophyll-a concentration, etc.

Figure 2.5: Global living planet index
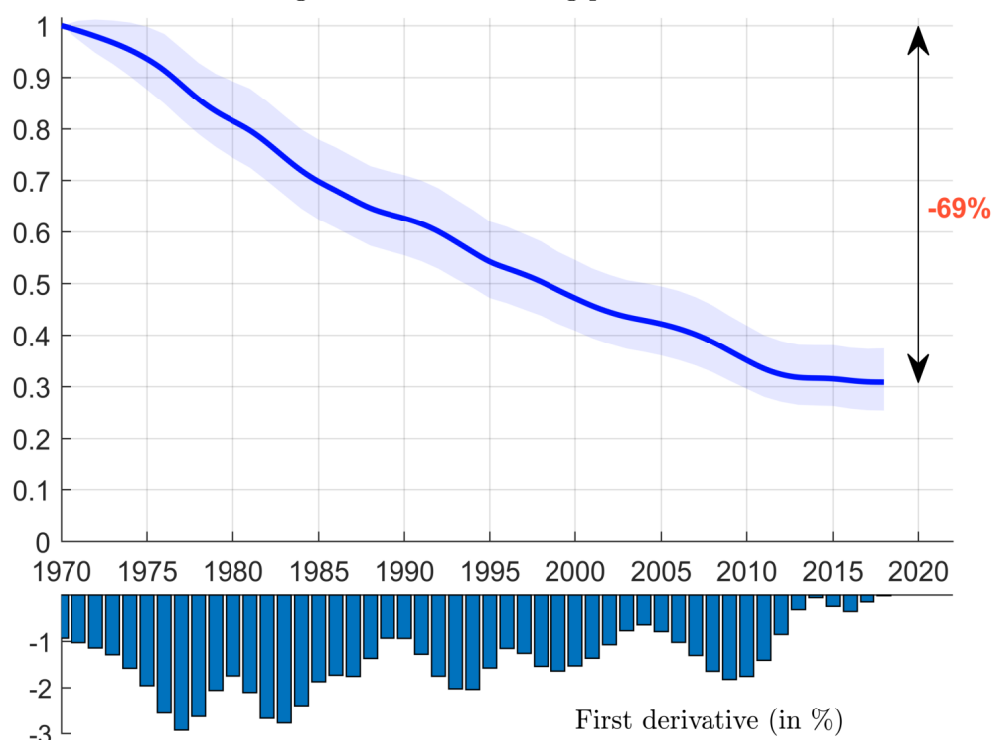
Source: https://livingplanetindex.org/latest_results & Author's calculations.

One of the hot topics is currently the biodiversity. A quick search on the web produces dozen of internet pages[11]. Financial institutions have also launched another initiative: Finance For Biodiversity Pledge (https://www.financeforbiodiversity.org). The UN Biodiversity Conference (COP 15), which is organized by the CBD in Montreal, Canada from 7 to 19 December 2022, has certainly given a special impulse, and may explain this new interest. However, biodiversity loss[12] is a very old topic and has been scientifically documented since the 1990s (Cardinale et al., 2012). According to Almond et al. (2022), biodiversity, as measured by the Living Planet Index[13], has decreased by 69%

---

[11]For example, "Why biodiversity is moving to top of mind for investors" (Lazard Asset Management, February 2022), "Investors grapple with complexities of biodiversity" (Financial Times, September 2022), "Asset Management, a lever for preserving biodiversity" (BNP Paribas, September 2022), "Biodiversity quickly rises up the ESG investing agenda" (Financial Times, September 2022), "Biodiversity: why investors should care" (Pictet AM, October 2022), "More asset owners and managers sign biodiversity pledge" (Pensions&Investments, October 2022), etc.

[12]Biodiversity loss describes the decline in the number, genetic variability, and variety of species, and the biological communities in a given area.

[13]The LPI is computed using a subset of 31 821 populations of 5 230 species and a statistical model (Westveer et al., 2022, page 28-31).

on average since 1970, but with a lot of heterogeneity across regions[14]. Even if the biodiversity loss has decreased these last years (Figure 2.5), this will inevitably result in negative consequences on global wealth in the long run. The seminal work of Costanza *et al.* (1997) estimated that the annual economic value of natural capital is on average two times the annual economic value of global GNP, explaining that "*ecosystem services provide a significant portion of the total contribution to human welfare on this planet*".

<div style="text-align:center">Box 2.1: Ecological diversity indexes</div>

Let $p = (p_1, \ldots, p_{n_S})$ be the proportion vector of species where p_i is the relative abundance[a] of the ith specie. The Hill diversity coefficient of order $\eta \geq 0$ is defined as:

$$\mathcal{D}^\eta (p) = \left(\sum_{i=1}^{n_S} p_i^\eta \right)^{1/(1-\eta)}$$

We can show that $1 \leq \mathcal{D}^\eta (p) \leq n_s$ and the bounds are reached for the 1- and n-diversity distributions[b] π_n^- and π_n^+. The Hill number measures the "*effective number of species*", meaning that the system holds a diversity equivalent to $\mathcal{D}^\eta (p)$ equally distributed species. The parameter η defines the sensitivity of the true diversity to rare versus. abundant species by modifying the weight given to specie abundances. When $\eta = 0$, $\mathcal{D}^\eta (p)$ is equal the current number n_S of species or the richness of species. When $\eta \to 1$, we obtain the Shannon diversity index $\mathcal{I}^\star (p)$, which is equal to the exponential of the Shannon entropy $\mathcal{I} (p)$:

$$\mathcal{D}^1 (p) = \mathcal{I}^\star (p) = \exp (\mathcal{I} (p)) = \exp \left(- \sum_{i=1}^{n_S} p_i \ln p_i \right)$$

When $\eta = 2$, we obtain:

$$\mathcal{D}^2 (p) = \left(\sum_{i=1}^{n_S} p_i^2 \right)^{-1}$$

We recognize the inverse of the Herfindahl index $\mathcal{H} (p) = \sum_{i=1}^{n_S} p_i^2$ (also called the Simpson index $\lambda (p)$ in ecology). Finally, when $\eta \to \infty$, the Hill index converges to the proportional abundance of the most abundant specie:

$$\mathcal{D}^\infty (p) = \max_i p_i$$

$\mathcal{D}^\infty (p)$ is then equal to the infinite norm of p.

---

[a]It is equal to number of individuals in the ith specie relative to the total number of individuals in the population.

[b]See their definition on page 658.

The sudden interest of financial institutions in biodiversity may be explained by climate change, but also by a greater awareness of its critical functions (food security, health, etc.). Moreover, the seventh mandatory PAI indicator requires reporting the share of investments that negatively affect biodiversity sensitive areas[15], and the sixth objective of the EUTR is dedicated to the "*protec-*

---

[14]This figure is respectively equal to -18% in Europe and Central Asia, -20% in North America, -55% in Asia and the Pacific, -66% in Africa, and -94% in Latin America.

[15]See Table 1.7 on page 36.

Figure 2.6: Aggregate national RLI
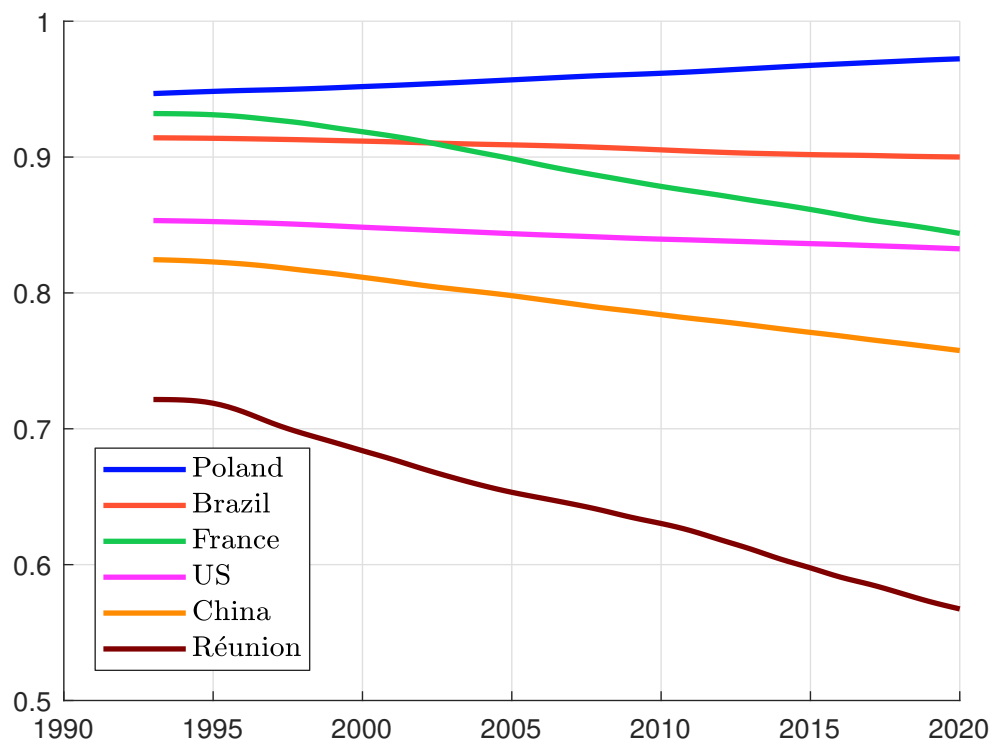
tion and restoration of biodiversity and ecosystem". All this obviously creates a high demand for biodiversity data and new opportunities for data providers, but as mentioned by Bowker (2000), biodiversity implies data diversity. Again, we are dealing with a huge amount of data[16]. For example, Icerberg Data Lab[17] (corporate biodiversity footprint or CBF), Carbon 4 (global biodiversity score[18] or GBS), CDC Biodiversité (global biodiversity score for financial institutions[19] or GBSFI), ISS ESG[20] (mean species abundance or MSA, potentially disappeared fraction of species or PDF) and Verisk Maplecroft (biodiversity and protected areas index score[21] or BPAI) have already developed biodiversity scores. For countries, most biodiversity data are open source[22] (Stephenson and Stengel, 2020):

- the Red List Index (RLI, https://www.iucnredlist.org)
  The RLI is an index of extinction risk for species of plants and animals. It is computed[23] by the IUCN and is available for five taxonomic groups: birds, mammals, amphibians, cycads and warm-water reef-forming corals. It can be disaggregated in various ways: subset of species (pollinator species, forest-specialist species, invasive alien species, etc.), country, region, etc.

---

[16]And also with a lot of diversity measures (Bandeira *et al.*, 2013; Ohlmann *et al.*, 2019).

[17]https://icebergdatalab.com.

[18]https://www.carbon4finance.com/product/biodiversity-impacts.

[19]https://www.cdc-biodiversite.fr/le-global-biodiversity-score.

[20]https://www.issgovernance.com/esg/biodiversity-impact-assessment-tool.

[21]https://www.maplecroft.com/insights/analysis/mining-operations-face-growing-biodiversity-risks.

[22]See the *Guide on Biodiversity Measurement Approaches* produced by the Finance for Biodiversity Pledge for a comparison of commercial and open source databases (https://www.financeforbiodiversity.org/publications/guide-on-biodiversity-measurement-approaches).

[23]The methodology is described in Butchart *et al.* (2007).

For instance, we report in Figure 2.6 the aggregate RLI for Brazil, China, France, Poland, La Réunion and US.

- World Database on Protected Areas (WDPA, https://www.protectedplanet.net)

- Integrated Biodiversity Assessment Tool (IBAT, https://www.ibat-alliance.org), including the Species Threat Abatement and Restoration metric (STAR)

- Exploring Natural Capital Opportunities, Risks and Exposure (ENCORE, https://encore.naturalcapital.finance)

- Etc.

Remark 11 *Biodiversity risk is a key element for impact investing. We refer to Chapter 5 for an extensive study of this risk (Section 5.4 on page 225).*

2.1.2 Corporate ESG data

Compared to sovereign ESG data, the collection, understanding and use of corporate ESG data is much more complicated and requires a lot of time and resources. In the first case, we have about 200 countries in the world, many international organizations that produce country data for decades, and vast academic research on this topic. For instance, the economic literature on income inequality starts in the early twentieth century with the seminal publications of Lorenz (1905) and Gini (1921). Since that time, the number of research studies on income inequality has grown exponentially[24]. In the case of corporate ESG data, data production has just become very recently, and the data dimension is not comparable. According to the World Federation of Exchanges (WFE), there are nearly 58 200 listed companies in the world at the end of Q1 2022. Moreover, data collection is not easier because it has concerned private data for a very long time. It is only recently that extra-financial reporting frameworks for corporate issuers exist, and most of them are voluntary. Finally, the last issue when using corporate ESG data is that most of indicators does not have a universal definition. In a nutshell, there are 3 main challenges and barriers to corporate ESG data:

1. Data coverage (how to collect data for all the listed companies?);

2. Data sourcing (where to find the data?);

3. Data quality (what is the accuracy of the collected data?)

Main indicators

As we have previously seen, we must distinguish several levels of data. Indeed, raw data are generally transformed into ESG metrics, and these metrics are used to define an ESG score. The main difficulty is collecting the raw data. In the case of corporates, this process is time-consuming and manual. The main sources of raw data are:

1. Corporate publications (self-reporting)

   (a) Annual reports

   (b) Corporate sustainability reports

---

[24]According to Google scholar, there are more than 5 700 published papers on this topic from 1950 to 1980, and already about 270 before 1950.

Figure 2.7: From raw data to ESG pillars
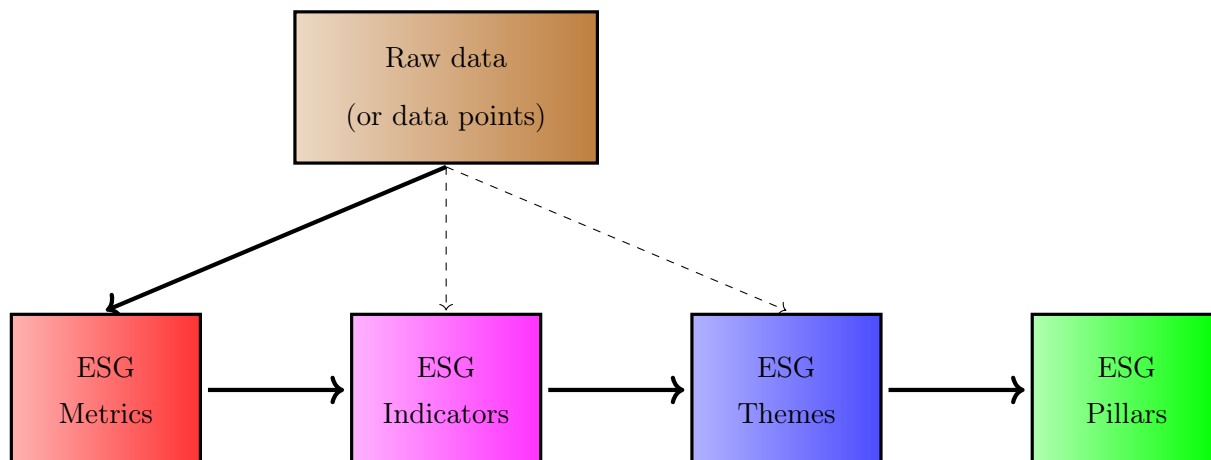
2. Financial and regulatory filings (standardized reporting)

    (a) Mandatory reports (SFDR, CSRD, EUTR, etc.)
    (b) Non-mandatory frameworks (PRI, TCFD, CDP, etc.)

For instance, the CDP database is the basic raw material used by all ESG rating agencies for measuring the carbon footprint of issuers. It can be viewed as the entry point and gives a first picture. Then, rating agencies will complete these data by gathering information from annual reports and other sources such as:

3 News and other media

4 NGO reports and websites

5 Company assessment and due diligence questionnaire (DDQ)

For example, S&P Global uses a 230-pages company questionnaire[25] called *Corporate Sustainability Assessment*. At the end of the collection data process, we may have missing data, noisy data or heterogeneous data. Therefore, the data are completed or adjusted by considering internal statistical models, e.g., industry-based clustering methods. This can be considered as a sixth source of data:

6 Internal models

Once these raw data are collected and cleaned, they can be used to calculate ESG metrics (second level of ESG variables). They are then grouped to define ESG indicators (third level of ESG variables), which are combined to form the basic ESG themes (fourth level of ESG variables). In the sequel, an ESG criterion is a generic term to name ESG variables: it may be an ESG metric, an ESG indicator or an ESG theme. Finally, ESG Pillars are generally based on a few number of ESG themes. This slicing method of ESG variables is illustrated in Figure 2.7. ESG rating agencies do not publicly disclose the raw data they use. Generally, they stop on the theme stage, sometimes on the indicator stage. To better understand the slotting method, we report an example of ESG criteria in Table 2.7. In this example, the (E) (S) and (G) pillars are made up of 9 environmental themes, 7 social themes and 8 governance themes. Here, we do not have access to the ESG indicators. We

---

[25]It can be downloaded at https://portal.csa.spglobal.com/survey/documents/CSA_Companion.pdf.

notice that some criteria are global and concerns all the issuers (e.g., carbon emissions), but other are specific to a given industry (e.g., green cars for the automobile sector, green financing for the banking sector). The choice of the themes/indicators will be done by the ESG rating agency, and the distinction between the two levels is not always obvious. For instance, if the board diversity is measured by the male-female ratio[26], the ESG theme is measured by a single ESG indicator. In this case, it is difficult to make a distinction between the two levels. Another issue concerns the classification of ESG themes. Let us consider the supply chain for example. It is a social issue if we would like to measure whether or not the suppliers of the company respect human rights and labor standards, but it can be an environmental issue if we would like to measure the impact of the suppliers on climate change and pollution. In Table 2.7, we also observe that the choice of ESG themes may be subjective. For example, we can merge the two categories *Pollution* and *Waste disposal* into one category *Pollution & waste disposal*, we can name *Corporate ethics* instead of *Corporate behaviour*, we can split *Biodiversity* into two categories (fauna & wildlife conservation; flora & land management), etc.

Table 2.7: An example of ESG criteria

Environmental	**S**ocial	**G**overnance
• Biodiversity	• Access to medicine	• Audit and control
• Carbon emissions	• Community involvement & human rights	• Board diversity
• Green cars*	• Customer concern & responsibility	• Board independence
• Green financing*	• Diversity	• Corporate behaviour
• Energy use	• Employment conditions & labor standards	• CSR strategy
• Pollution	• Gender equality	• Executive compensation
• Renewable energy	• Supply chain	• Management compensation
• Waste disposal		• Shareholder' rights
• Water use		

In what follows, we give some insight into the themes and indicators used by rating agencies. This public information about the ESG criteria has been collected from their website, and varies considerably between providers.

- Bloomberg rates 11 800 public companies. They use more than 120 ESG indicators and 2 000+ data points.

- ISS ESG rates about 10 000 issuers. They use more than 800 indicators and apply approximately 100 indicators per company[27].

---

[26]It is the ratio of men to women or the proportion of women in the company board.

[27]E.g, climate change strategy, eco-efficiency, energy management, environmental impact of product portfolio, environmental management, water risk and impact for the **E** pillar; equal opportunities, freedom of association, health and safety, human rights, product responsibility, social impact of product portfolio, supply chain management, taxes for the **S** pillar; business ethics, compliance, independence of the board, voting rights, shareholder participation, remuneration for the **G** pillar.

Table 2.8: MSCI ESG key issue hierarchy

Pillar	#	Theme	#	Indicator
Environment	1	Climate Change	1	Carbon Emissions
			2	Product Carbon Footprint
			3	Financing Environmental Impact
			4	Climate Change Vulnerability
	2	Natural Capital	5	Water Stress
			6	Biodiversity & Land Use
			7	Raw Material Sourcing
	3	Pollution & Waste	8	Toxic Emissions & Waste
			9	Packaging Material & Waste
			10	Electronic Waste
	4	Environmental Opportunities	11	Opportunities in Clean Tech
			12	Opportunities in Green Building
			13	Opportunities in Renewable Energy
Social	5	Human Capital	14	Labor Management
			15	Health & Safety
			16	Human Capital Development
			17	Supply Chain Labor Standards
	6	Product Liability	18	Product Safety & Quality
			19	Chemical Safety
			20	Consumer Financial Protection
			21	Privacy & Data Security
			22	Responsible Investment
			23	Health & Demographic Risk
	7	Stakeholder Opposition	24	Controversial Sourcing
			25	Community Relations
	8	Social Opportunities	26	Access to Communications
			27	Access to Finance
			28	Access to Health Care
			29	Opportunities in Nutrition & Health
Governance	9	Corporate Governance	30	Ownership & Control
			31	Board
			32	Pay
			33	Accounting
	10	Corporate Behavior	34	Business Ethics
			35	Tax Transparency

Source: MSCI (2022, Exhibit 2, page 4).

Table 2.9: Refinitiv materiality matrix

Pillar	#	Theme	Metrics	#	Indicator
Environment	1	Emissions	28	1	Emissions
				2	Waste
				3	Biodiversity
				4	Environmental Management Systems
	2	Innovation	20	5	Product Innovation
				6	Green Revenues, Green R&D and Green CapEx
	3	Resource Use	20	7	Water
				8	Energy
				9	Sustainable Packaging
				10	Environmental Supply Chain
Social	4	Community	14	11	Community
	5	Human Rights	8	12	Human Rights
	6	Product Responsibility	10	13	Responsible Marketing
				14	Product Quality
				15	Data privacy
	7	Workforce	30	16	Diversity & Inclusion
				17	Career Development & Training
				18	Working Conditions
				19	Health & Safety
Governance	8	CSR Strategy	9	20	CSR Strategy
				21	ESG Reporting & Transparency
	9	Management	35	22	Structure (independence, diversity, committees)
				23	Compensation
	10	Shareholders	12	24	Shareholder Rights
				25	Takeover Defenses

Source: Refinitiv (2022, page 10).

- FTSE Russell rates about 7 200 securities. They use more than 300 indicators and 14 themes: biodiversity, climate change, pollution and resources, supply chain and water security for the (E) pillar; customer responsibility, health and safety, human rights and community, labor standards and supply chain for the (S) pillar; anti-corruption, corporate governance, risk management and tax transparency for the (G) pillar. Each theme contains 10 to 35 indicators, and an average of 125 indicators are applied per company.

- Moody's V.E rates more than 5 000 companies. They consider six pillars (corporate governance, business behavior, environment, human rights, human resources, community involvement) and 38 ESG indicators[28].

---

[28]Moody's has also developed a methodology for assessing ESG risks in credit analysis based on 15 themes: carbon transition, physical climate risks, water management, waste & pollution, natural capital for the (E) pillar; customer relations, human capital, demographic & societal trends, health & safety, responsible production for the (S) pillar; financial strategy & risk management, management credibility & track record, organizational structure, compliance & reporting, board structure & procedures for the (G) pillars.

- MSCI (2022) rates 10 000 companies (14 000 issuers including subsidiaries) and 680 000 securities globally. Using 1000+ data points, they consider two families of metrics: 80 exposure metrics (how exposed is the company to each material issue?) and 250+ management metrics (how is the company managing each material issue?). These metrics are then combined into 35 key issues selected annually for each industry. These key metrics are reported in Table 2.8 and are combined to build 10 main themes.

- Refinitiv (2022) rates 12 000 public and private companies. They consider 10 themes: resource use, emissions and innovation for the **E** pillar; workforce, human rights, community and product responsibility for the **S** pillar; management, shareholders and responsibility (CSR) strategy for the **G** pillar. These themes are built using 186 metrics and 630+ data points. Table 2.9 shows the materiality matrix of themes, indicators and the number of metrics per theme.

- S&P Dow Jones Indices uses between 16 to 27 criteria scores, a questionnaire-based analysis process with 80-120 industry-specific questions and 1 000 data points.

- Sustainalytics rates more than 16 300 companies. They consider 20 material ESG issues, based on 350+ indicators.

Remark 12 *Contrary to sovereign issuers, raw data for corporate issuers are more difficult to find, because they are not in open source data or they can only be manually collected (e.g., annual reporting). The ESG Data Cartography[29], which has been developed by the Louis Bachelier Institute, proposes a comprehensive list of ESG data with 140+ data sources. The user can filter the databases by accessibility (free, open source, partially free and proprietary).*

Exercise 1 *Berg et al. (2022) consider a common taxonomy based on 64 indicators to compare the different ESG rating providers: access to basic services; access to healthcare; animal welfare; anti-competitive practices; audit; biodiversity; board; board diversity; business ethics; chairperson-CEO separation; child labor; climate risk management; clinical trials; collective bargaining; community & society; corporate governance; corruption; customer relationship; diversity; ESG incentives; electromagnetic fields; employee development; employee turnover; energy; environmental fines; environmental management system; environmental policy; environmental reporting; financial inclusion; forests; GHG emissions; GHG policies; GMOs; Global Compact membership; green buildings; green products; HIV programs; hazardous waste; health & safety; human rights; indigenous rights; labor practices; lobbying; non-GHG air emissions; ozone-depleting gases; packaging; philanthropy; privacy & IT; product safety; public health; remuneration; reporting quality; resource efficiency; responsible marketing; shareholders; site closure; supply chain; sustainable finance; systemic risk; taxes; toxic spills; unions; waste; water. For each pillar, give the list of indicators that fall in the category. We consider a very basic ESG classification matrix with 12 themes:*

E	**S**	**G**
Global warming	*Health*	*Board*
Green opportunities	*Human rights*	*Corporate ethics*
Natural resource	*Workforce*	*CSR strategy*
Transition risk	*Social responsibility*	*Shareholder*

For each indicator, associate the right ESG theme[30].

---

[29]The website is https://www.institutlouisbachelier.org/en/esg-data-cartography.
[30]There may be no or several valid answers.

The race for alternative data

Alternative data corresponds to data that is not available through traditional channels (corporate publications, sustainable reporting, etc.). It includes non-structured data such as images or textual contents. The case of ESG ratings mainly concerns three types of data:

- Internet traffic, browsing activity, web scraping, product reviews, social media and sentiment data;

- Satellite imagery, geotracking data, sensor data[31];

- Supply-chain data;

Brière *et al.* (2022) discuss several uses of alternative data sets. The most famous application is the tracking and measurement of ESG controversies. A controversy risk occurs when allegations concerning a company could lead to reputational risk[32] and financial losses. Everybody knows the famous quotes of Warren Buffet about building and destroying a reputation:

> "*It takes 20 years to build a reputation and five minutes to ruin it. If you think about that, you'll do things differently. [...] We can afford to lose money — even a lot of money. But we can't afford to lose reputation — even a shred of reputation. [...] Should you find yourself in a chronically leaking boat, energy devoted to changing vessels is likely to be more productive than energy devoted to patching leaks. [...] Lose money for the firm, and I will be understanding. Lose a shred of reputation for the firm, and I will be ruthless.*"

The allegations can be reported by media, NGOs, social networks and stakeholders. Data providers generally use text mining and natural language processing (NLP) to analyze an enormous amount of information, detect controversial events and measure the severity of the reputational risk. For example, Refinitiv (2022) completes the traditional ESG score with a controversy score for the 10 ESG themes presented in Table 2.9 on page 63. This score is updated on a weekly basis. The ESG controversies score is calculated based on 23 ESG controversy topics:

- Community: (1) anti-competition controversy, (2) business ethics controversies, (3) intellectual property controversies, (4) critical countries controversies, (5) public health controversies, (6) tax fraud controversies;

- Human rights: (7) child labour controversies, (8) human rights controversies;

- Management: (9) management compensation controversies count;

- Product responsibility: (10) consumer controversies, (11) customer health and safety controversies; (12) privacy controversies; (13) product access controversies; (14) responsible marketing controversies; (15) responsible R&D controversies;

- Resource use: (16) environmental controversies;

- Shareholders: (17) accounting controversies count, (18) insider dealings controversies, (19) shareholder rights controversies;

---

[31]E.g., temperature, humidity, pressure, chemical levels.

[32]Some famous examples are the *Mexico oil spill* (BP, 2010), *dieselgate* affair (Volkswagen, 2015), the gender pay gap (BBC, 2017), the Cambridge Analytica scandal (Facebook, 2018), the opioid epidemic (Purdue Pharma, 2019), the Ehpad scandal (Orpea, 2021), the Pegasus software (NSO, 2021), the greenwashing (DWS, 2022), etc.

- Workforce: (20) diversity and opportunity controversies; (21) employee health and safety controversies; (22) wages or working conditions controversies; (23) strikes.

One of the most famous controversy data providers is the Swiss company RepRisk (https://www.reprisk.com), which was created in Zurich in 1998. They are specialized in ESG data science and machine learning. In November 2021, they published their comprehensive methodology (RepRisk, 2022) and Jupyter Notebooks. To identify and classify ESG risks consistent with how key international standards and norms define ESG, they consider a 3-step process:

1. Daily, they collect 500 000+ documents from 100 000+ sources[33] in 23 languages;

2. These documents are scraped from online sources and fed to machine learning (ML) applications, which predict relevant and unique ESG risk incidents. Results are sent to the ML reducer, in particular, irrelevant results are discarded and predictions fed to the multilingual queue;

3. Then, documents are sorted in priority order. A team of 150+ human analysts confirm and correct ML predictions, assess severity, reach, and novelty, and write risk incident summaries; Final results are incorporated into RepRisk databases[34].

Exercise 2 *RepRisk (2022) uses 73 controversial topics: abusive/illegal fishing; access to products and services; agricultural commodity speculation; airborne pollutants; alcohol; animal transportation; arctic drilling; asbestos; automatic and semi-automatic weapons; biological weapons; chemical weapons; cluster munitions; coal-fired power plants; conflict minerals; coral reefs; cyberattack; deep sea drilling; depleted uranium munitions; diamonds; drones; economic impact; endangered species; energy management; epidemics/pandemics; forest burning; frocking; fur and exotic animal skins; gambling; gender inequality; genetically modified organisms (GMOs); genocide/ethnic cleansing; greenhouse gas (GHG) emissions; health impact; high conservation value forests; human trafficking; hydropower (dams); illegal logging; indigenous people; involuntary resettlement; land ecosystems; land grabbing; land mines; lobbying; marijuana/cannabis; marine/coastal ecosystems; migrant labor; monocultures; mountaintop removal mining; negligence; nuclear power; nuclear weapons; offshore drilling; oil sands; opioids; palm oil; plastics; pornography; predatory lending; privacy violations; protected areas; racism/racial inequality; rare earths; salaries and benefits; sand mining and dredging; seabed mining; security services; ship breaking and scrapping; soy; tax havens; tobacco; wastewater management; water management; water scarcity. For each topic, associate the right ESG pillar.*

Besides controversy risk, text mining and NLP techniques became recently an essential ML tool for different ESG applications. For example, they are more and more used for assessing company disclosures and verifying their credibility. Friederich *et al.* (2021) use the language model BERT[35] to automatically identify disclosures of climate-related risks from corporates' annual reports. In a similar way, Bingler *et al.* (2022) analyse climate risk disclosures along the TCFD categories and conclude that "*the firms' TCFD support is mostly cheap talk and that firms cherry-pick to report primarily non-material climate risk information*". Always using the same machine learning model

---

[33]These include government agencies, news sites, newsletters, NGOs, print media, regulators, research firms, social media blogs, think tanks and twitter messages.

[34]As of July 2022, the RepRisk dataset includes more than 205 000 companies that are associated with risk incidents. Of these 205 000 companies, approximately 7% are listed companies and 93% are non-listed companies (RepRisk, 2022, page 5).

[35]Bidirectional Encoder Representations from Transformers (BERT) is a transformer-based machine learning technique for NLP pre-training developed by Google.

BERT, Kölbel *et al.* (2022) consider the impact of climate risk disclosures on the CDS market and find that disclosing transition risks increases CDS spreads, which is not the case for physical risks.

Figure 2.8: Geolocation of world power plants by energy source
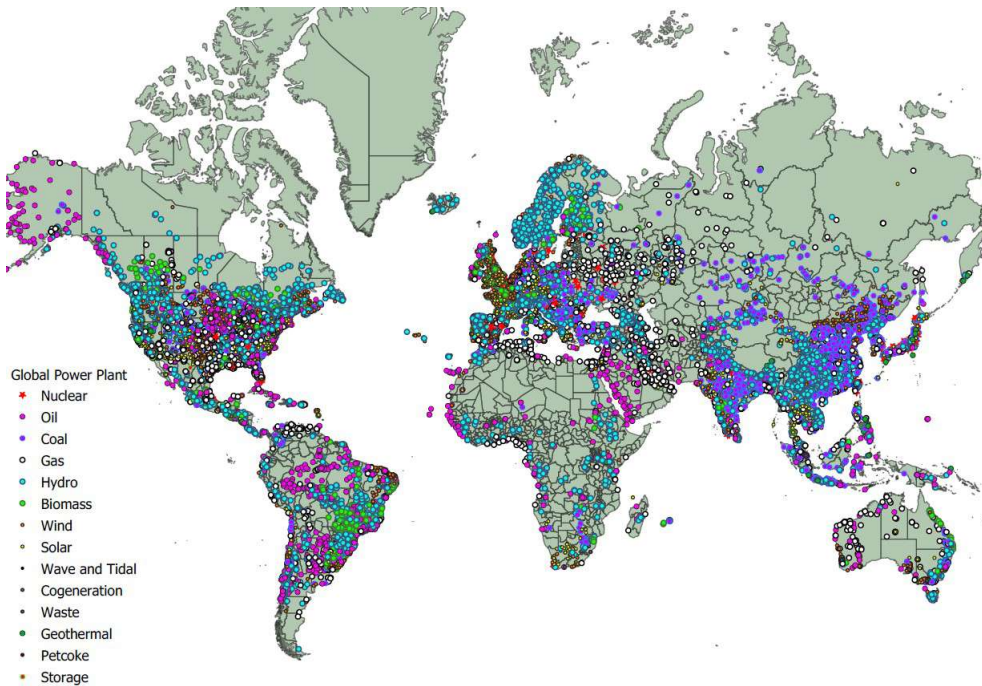
Another application of alternative data is the estimation of physical risk exposures. They correspond to the potential financial losses that companies can suffer, and includes droughts, floods, storms, etc. This risk is more difficult to quantify, and its evaluation requires multidisciplinary methodologies: climate modeling, physical asset geolocation, financial loss estimation, etc. In this case, asset tracking is really the crux of physical risk modeling. An example of spatial data is provided in Figure 2.8. This type of geolocalized data is extensively used by Le Guenedal *et al.* (2021, 2022) when developing a fully integrated methodology to measure cyclone-related physical risk. Until now, most of the models have been developed for countries and regions (Burke *et al.*, 2021)(Burke *et al.*, 2021). When dealing with corporate ESG, data providers generally use input-output matrices in order to compute the risk exposure or contribution of each firm. It is for example the case for biodiversity risk. Nevertheless, we have recently observed some initiatives to provide geospatial data and asset tracking directly at the company level. Even if these solutions are not yet mature, they are very promising[36].

Remark 13 *Apart from controversies and physical risk, we also notice a third application of alternative data, which consists in building more reactive or real-time ESG scores. Ben Dor* et al. *(2022) propose to monitor planned sustainability-related corporate activities based on firms' actions, rather than relying solely on their announcements. For that, they use job postings and NLP to identify ESG-related openings and ESG-related activities of firms. This technique can be used to understand the dynamics of sustainability within a firm.*

---

[36]You can visit the website of the French technology company Kayrros (https://www.kayrros.com), which received the Financial TimesâĂŹ Tech Champions award for its innovation in the IT & Software sector. Kayrros uses satellite observation and AI to analyse trends in emissions and deforestation.

Figure 2.9: ESG rating disagreement
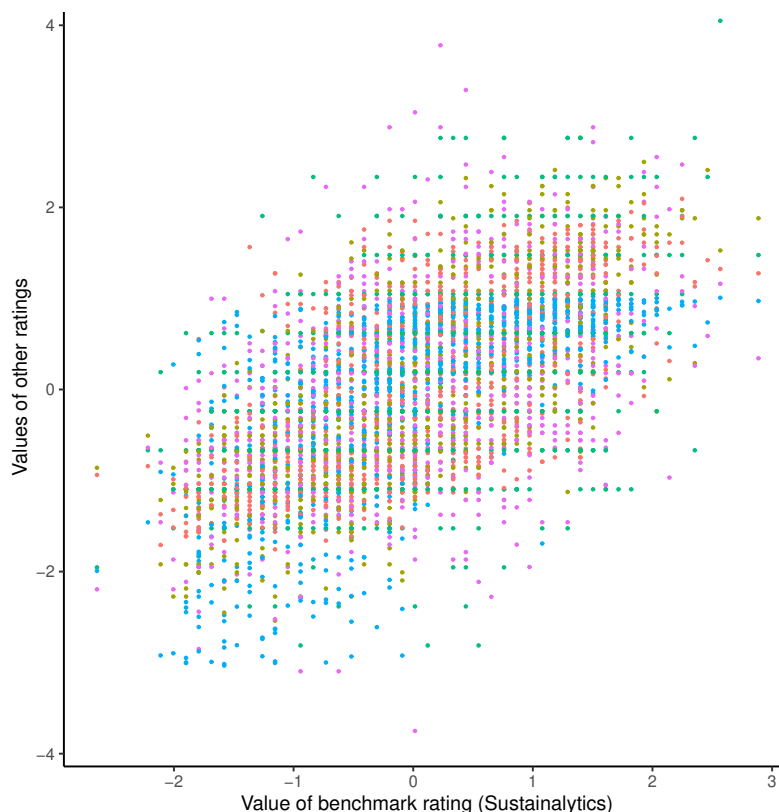
"*This graph illustrates the ESG rating divergence. The horizontal axis indicates the value of the Sustainalytics rating as a benchmark for each firm ($n = 924$). Rating values by the other five raters are plotted on the vertical axis in different colors*. For each rater, the distribution of values has been normalized to zero mean and unit variance. The Sustainalytics rating has discrete values that show up visually as vertical lines where several companies have the same rating value.*"

The divergence of corporate ESG ratings

Corporate ESG data are very different than sovereign ESG data in terms of standardization. Therefore, we must expect more discrepancies in ESG ratings. In Figure 2.9, we have reported one of the most famous illustrations[37] about this rating disagreement extracted from the pioneering research of Berg *et al.* (2022). These authors investigate the divergence of ESG ratings from six prominent rating agencies: KLD, Moody's ESG, MSCI, Refinitiv, S&P Global and Sustainalytics. Using the ESG metrics of these data providers, they reconstruct synthetic ratings based on a common taxonomy of 64 indicators[38]. They identify three sources of divergence:

1. "**Measurement** *divergence refers to situation where rating agencies measure the same indicator using different ESG metrics;*

2. **Scope** *divergence refers to situation where ratings are based on different set of ESG indicators*;

---

[37]We have the following mapping: *S&P Global, Moody's ESG, KLD, Refinitiv, MSCI.
[38]See Exercise 1 on page 64 for the list of ESG indicators

3. **Weight** *divergence emerges when rating agencies take different views on the relative importance of ESG indicators*".

They find that measurement contributes to 56% of the divergence, scope 38% and weight 6%.

Since this publication, the standardization issue of data and methodologies has been an ongoing discussion among practitioners and academics. For instance, Billio *et al.* (2021) analyze ESG ratings and indexes agreement and find that "*it is extremely difficult to measure the ability of a fund manager if financial performances are strongly conditioned by the chosen ESG benchmark*" and "*disagreement in the scores provided by the rating agencies disperses the effect of preferences of ESG investors on asset prices*". In Table 2.10, we report their rank correlation matrix. On average, they obtain a mean correlation of 58% for corporate ESG ratings vs. 85% for sovereign ESG scores (Gratcheva *et al.*, 2020).

Table 2.10: Rank correlation among ESG ratings

	MSCI	Refinitiv	S&P Global	
MSCI	100%			
Refinitiv	43%	100%		
S&P Global	45%	69%	100%	
Sustainalytics	53%	64%	69%	100%

Source: Billio *et al.* (2021, Table 3, page 1432) .

2.2 Scoring system

A scoring model is a mathematical model that forms the basis for risk stratification. For example, credit scoring refers to statistical models used to measure the creditworthiness of a company or individual (Roncalli, 2020a). In particular, the Altman Z score is probably the most famous score for predicting the bankruptcy of commercial companies (Altman, 1968). However, scoring models can be found in many areas. For example, anti-money laundering (AML) scoring is a rating model used to assess the risk profile of customers (Chen *et al.*, 2018). The goal of trauma and field triage scoring systems is to predict the severity of injury or estimate the prognosis of trauma patients (Senkowski and McKenney, 1999). The Apgar score assesses the physical condition of newborns shortly after birth (Finster *et al.*, 2005). In the case of medicine, we find many scoring systems: ACR score (rheumatoid arthritis symptoms), Alvarado score (appendicitis), Framingham and QRISK scores (cardiovascular risk), Geneva score (pulmonary embolism), etc.

At first glance, we might think that ESG scoring is an extension of credit scoring using extra-financial data instead of financial data. And there are many similarities between the two concepts: ESG ratings vs. credit ratings, ESG materiality vs. credit materiality, ESG risk vs. credit risk, etc. However, from a mathematical point of view, they are two different concepts. In fact, ESG scoring is an unsupervised approach to risk materiality, while credit scoring is a supervised approach to risk materiality[39]. In the case of credit, we want to measure the one-year probability of default.

---

[39]Unsupervised learning is a branch of statistical learning in which the test data does not contain a response variable. It is in contrast to supervised learning, where the goal is to predict the value of the response variable Y given a set of explanatory variables X. In unsupervised learning, we only know the X values, because the Y values do not exist or are not observed. Supervised and unsupervised learning methods are also called "*learning with/without a teacher*" (Hastie *et al.*, 2009). This metaphor means that in supervised learning we have access to the correct answer provided by the supervisor (or teacher). In unsupervised learning, we have no feedback on the correct answer. For example,

Therefore, credit scoring models are calibrated with a historical database of borrower default events. The response variable is then a binary variable that is 1 if the borrower has defaulted and 0 otherwise. In the case of ESG, we want to measure the sustainability of issuers, but we face an endogenous puzzle because the ESG score is already the sustainability measure. The big problem with ESG scoring systems is how to define the response variable. In most cases, the scoring model is not calibrated and is a simple rule-based method. For this reason, we generally consider ESG scoring to be an unsupervised statistical approach. This has several drawbacks in terms of performance evaluation, score consistency and backtesting. Nevertheless, we will see that we can define some proxies for the response variable and use traditional statistical tools to assess the quality of ESG scores.

2.2.1   ESG and scoring theory

The goal of scoring models is to produce a numeric score $\mathcal{S}$. This score can take values between a lower bound $\mathcal{S}^-$ and an upper bound $\mathcal{S}^+$. We generally assume that a high value of $\mathcal{S}$ is a good risk, while a low value of $\mathcal{S}$ is a bad risk. Scoring systems are then used for two main types of decisions: selection and exclusion. A selection process consists of selecting the good risks so that $\mathcal{S} \geq s_1$, while an exclusion process consists of excluding the bad risks so that $\mathcal{S} \leq s_2$. The choice of the thresholds s_1 and s_2 is important and depends on the process. For example, suppose the score is between 0 and 1 000. We can formulate the following rule:

$$\begin{cases} \mathcal{S} \geq 500 \Rightarrow \text{good risk} \Rightarrow \text{selection} \\ \mathcal{S} \leq 500 \Rightarrow \text{bad risk} \Rightarrow \text{exclusion} \end{cases}$$

Obviously, separating the population into two groups (bad and good risks) is a difficult task. Another approach is to use stricter decision rules:

$$\begin{cases} \mathcal{S} \geq 600 \Rightarrow \text{good risk} \Rightarrow \text{selection} \\ \mathcal{S} \leq 300 \Rightarrow \text{bad risk} \Rightarrow \text{exclusion} \end{cases}$$

Here we see that there is an asymmetry in the definition of good and bad risks. The choice of the threshold is as important as the construction of the score. However, we have an endogenous problem because the threshold depends on the model itself. In the case of ESG, scores can be designed to define an exclusion process or a selection process, but in most cases they are used to perform both or they are used to define an integration process. This implies that the previous binary choice is replaced by the preference ordering:

$$\mathcal{S}_1 \succ \mathcal{S}_2 \Leftrightarrow \text{issuer } \#1 \text{ is a better ESG risk than issuer } \#2$$

This approach to scoring is different from the traditional approach. We face a problem here because ESG scores are generally the result of an unsupervised statistical method. Therefore, the development of ESG scoring has been done outside the framework of scoring theory. In fact, unsupervised statistical methods are generally used for clustering and classification. They are then adapted to a multimodal statistical problem, such as a binary statistical model (bad risk versus good risk). ESG scoring is more of an expert system than a true scoring model. In what follows, we try to provide a theoretical framework for ESG scoring, but we must be careful. This theoretical framework is limited for all the reasons mentioned above.

Remark 14 *The fact that ESG scoring is designed as an unsupervised statistical process makes it difficult to assess the quality of ESG rating systems, particularly with respect to materiality.*

---

linear regression is a typical supervised learning model, while principal component analysis is an unsupervised learning approach.

2.2.2 Tree-based scoring methods

Tree structure

To understand a tree-based scoring model, we first consider the one-level tree structure. Let $X_1, \ldots, X_m$ be m features. These metrics are linearly combined to obtain a score:

$$\boldsymbol{S} = \sum_{j=1}^{m} \omega_j X_j$$

where ω_j is the weight of the jth metric. Generally, the weights are normalized such that $\sum_{j=1}^{m} \omega_j = 1$. This is the most simple scoring model. For instance, the original bankruptcy score of Altman (1968) was equal to:

$$Z = 1.2 \cdot X_1 + 1.4 \cdot X_2 + 3.3 \cdot X_3 + 0.6 \cdot X_4 + 1.0 \cdot X_5$$

where the variables X_j represent the following financial ratios:

X_j	Ratio
X_1	Working capital / Total assets
X_2	Retained earnings / Total assets
X_3	Earnings before interest and tax / Total assets
X_4	Market value of equity / Total liabilities
X_5	Sales / Total assets

If we note Z_i the score of the firm i, we can calculate the normalized score $Z_i^{\star} = (Z_i - m_z)/\sigma_z$ where m_z and σ_z are the mean and standard deviation of the observed scores. $Z_i^{\star}$ can then be compared to the quantile of the Gaussian distribution or the empirical distribution. A low value of $Z_i^{\star}$ (for instance $Z_i^{\star} < -2.5$) indicates that the firm has a high probability of default, while companies with high scores above (for instance $Z_i^{\star} > 3$) are not likely to go bankrupt.

We can extend the previous approach to a two-level tree structure. We first begin to compute intermediary scores:

$$\boldsymbol{S}_k^{(1)} = \sum_{j=1}^{m} \omega_{j,k}^{(1)} X_j$$

Then we obtain a set of $m_{(1)}$ intermediary scores ($k = 1, \ldots, m_{(1)}$), which are combined to obtain the final score:

$$\boldsymbol{S} := \boldsymbol{S}_1^{(0)} = \sum_{k=1}^{m_{(1)}} \omega_k^{(0)} \boldsymbol{S}_k^{(1)}$$

The exponents (0) and (1) indicate the level of the tree. An example of two-level tree structure is given in Figure 2.10. For the first level, we have:

$$\begin{cases} \boldsymbol{S}_1^{(1)} = 0.5 \cdot X_1 + 0.25 \cdot X_2 + 0.25 \cdot X_3 \\ \boldsymbol{S}_2^{(1)} = 0.5 \cdot X_4 + 0.5 \cdot X_5 \\ \boldsymbol{S}_3^{(1)} = X_6 \end{cases}$$

The final score is the average of the three intermediary scores:

$$\boldsymbol{S} = \frac{\boldsymbol{S}_1^{(1)} + \boldsymbol{S}_2^{(1)} + \boldsymbol{S}_3^{(1)}}{3}$$

Figure 2.10: A two-level tree
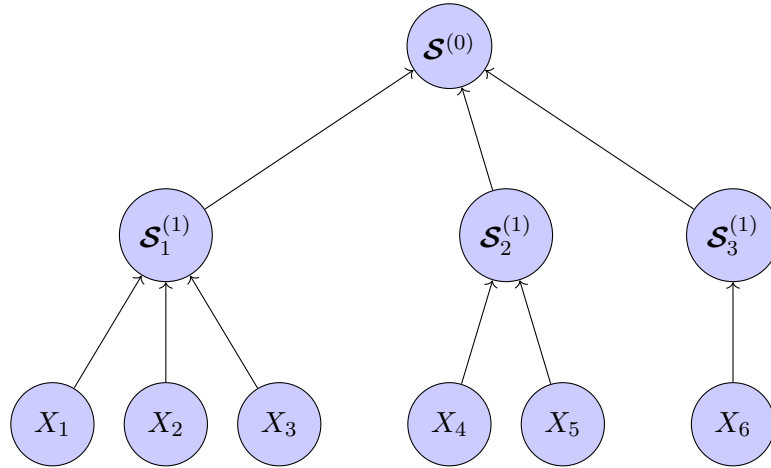
- Level 1: $\omega_{1,1}^{(1)} = 50\%$; $\omega_{2,1}^{(1)} = 25\%$; $\omega_{3,1}^{(1)} = 25\%$; $\omega_{4,2}^{(1)} = 50\%$; $\omega_{5,2}^{(1)} = 50\%$; $\omega_{6,3}^{(1)} = 100\%$;
- Level 0: $\omega_1^{(0)} = \omega_2^{(0)} = \omega_3^{(0)} = 33.33\%$;

Figure 2.11: A two-level overlapping graph
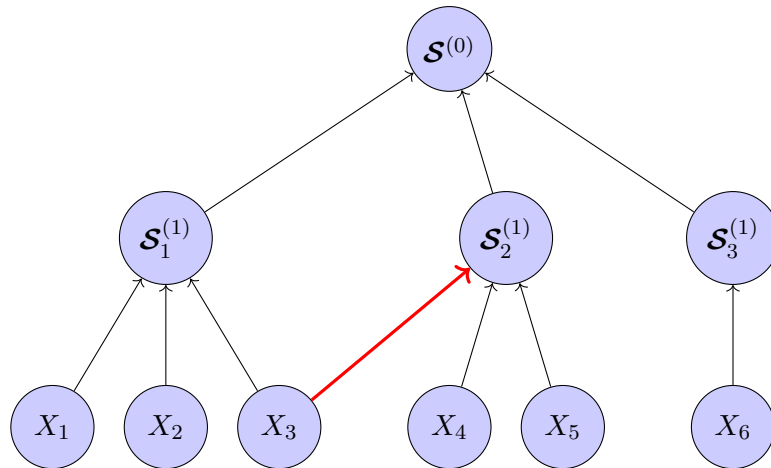
- Level 1: $\omega_{1,1}^{(1)} = 50\%$; $\omega_{2,1}^{(1)} = 25\%$; $\omega_{3,1}^{(1)} = 25\%$; $\omega_{3,2}^{(1)} = 25\%$; $\omega_{4,2}^{(1)} = 25\%$; $\omega_{5,2}^{(1)} = 50\%$; $\omega_{6,3}^{(1)} = 100\%$;
- Level 0: $\omega_1^{(0)} = \omega_2^{(0)} = \omega_3^{(0)} = 33.33\%$;

This tree is a non-overlapping graph because each child node is related to a single parent node, otherwise it is an overlapping graph (but it is not a tree). For example, if we assume that the score $\mathcal{S}_2^{(1)}$ also depends on the metric X_3, we obtain the overlapping graph structure given in Figure 2.11. In this case, the first level becomes:

$$\begin{cases} \mathcal{S}_1^{(1)} = 0.5 \cdot X_1 + 0.25 \cdot X_2 + 0.25 \cdot X_3 \\ \mathcal{S}_2^{(1)} = 0.25 \cdot X_3 + 0.25 \cdot X_4 + 0.5 \cdot X_5 \\ \mathcal{S}_3^{(1)} = X_6 \end{cases}$$

Box 2.2: Tree and graph theory

A scoring tree is a special case of a tree data structure, which is defined as a collection of nodes that are organized in a hierarchical structure (see Figure 2.A). A tree is also a connected graph without any circuits[a]. Therefore, the terminology of trees derives from the graph theory. A node (or a vertex) is the basic unit that may contain data and links to other nodes. A connection between two nodes is called an edge. In a tree, edges are directed and are also called arcs or arrows. For instance, our example has 13 nodes $\mathcal{V} = (A, \ldots, M)$ and 12 edges $\mathcal{E} = (\{K, H\}, \{L, J\}, \ldots, \{C, A\}, \{D, A\})$. Mathematically, the tree $\mathcal{T}$ is defined by the set $\mathcal{V}$ of nodes (or vertices) and the set $\mathcal{E}$ of edges: $\mathcal{T} = (\mathcal{V}, \mathcal{E})$. In a tree, the first node is called the root node (A). Any node within a tree can be viewed as a root of its own subtree. By definition, The subtree $\mathcal{T}_{(v)}$ with v as its root is also a tree consisting of v and its descendants. $\mathcal{T}_{(v)}$ is defined by $(\mathcal{V}_{(v)}, \mathcal{E}_{(v)})$, where $\mathcal{V}_{(v)}$ and $\mathcal{E}_{(v)}$ are the sets of vertices and edges of the subgraph. Each edge (or directed path) has a child and a parent (or an internal node). For example, C is the parent node of (H, I) and H is a child node of C. Our example tree has then 6 parent nodes $\mathcal{P} = (A, B, C, D, H, J)$ and 12 children[b]. B has three children (E, F, G) and B is a child of A. Child nodes with the same parent are sibling nodes, while a leaf node (or external node) is a node without child nodes. In the example tree, (B, C, D), (E, F, G), (H, I) are siblings. The leaf nodes are (E, F, G, I, K, L, M). The degree of a node is its number of children. It is equal to 3 for (A, B), 2 for (C, J) and 1 for (H, D) and 0 for the leaves. The degree of a tree is equal to the maximum degree of nodes. Our example tree has a degree 3. The level of a node refers to the distance between the node and the root. We deduce that the root node is at level 0. Children of the root are at level 1 (B, C, D). Level 2 corresponds to the nodes (E, F, G, H, I, J) and we have 3 nodes at level 3 (K, L, M). The depth of a tree is the level of the deepest node. It is equal to 3 in our example tree.

Figure 2.A: Tree data structure

---
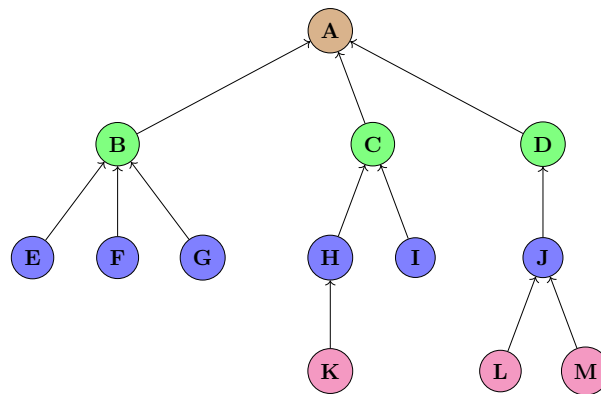
[a]We have the following properties:

- There is one and only one path between every pair of nodes in a tree;
- A tree with n nodes has exactly $n - 1$ edges;
- A graph is a tree if and only if it is minimally connected;
- Any connected graph with n nodes and $n - 1$ edges is a tree.

[b]In a tree, the number of children is exactly equal to the number of edges.

The two-level tree structure can be extended to multi-level tree structures. Let L be the number of levels. At level ℓ, the value of the kth node is given by:

$$\boldsymbol{\mathcal{S}}_k^{(\ell)} = \sum_{j=1}^{m_{(\ell+1)}} \omega_{j,k}^{(\ell)} \boldsymbol{\mathcal{S}}_j^{(\ell+1)} \tag{2.1}$$

where $m_{(\ell+1)}$ is the number of scores at level $\ell + 1$, $\boldsymbol{\mathcal{S}}_j^{(\ell+1)}$ is the jth score at level $\ell + 1$ and $\omega_{j,k}^{(\ell)}$ is the weight of the jth score at level $\ell + 1$ for the kth score at level ℓ. By construction, the scores at level L are exactly equal to the features: $\boldsymbol{\mathcal{S}}_j^{(L)} = X_j$. We verify that the final score $\boldsymbol{\mathcal{S}}$ corresponds to the root score $\boldsymbol{\mathcal{S}}_1^{(0)}$. It can be computed by Algorithm 1. If we would like to target a specific level $\ell^\star$, we replace the **for** statement $t = 1 : L$ by $t = 1 : L + 1 - \ell^\star$.

Algorithm 1 Recursive tree-based algorithm for computing the final score

Compute the final score $\boldsymbol{\mathcal{S}}_1^{(1)}$
Input: L the number of levels, $(X_1, \ldots, X_m)$ the vector of the metrics and $\left\{\omega_{j,k}^{(\ell)}\right\}$ the weight tensor
Initialize $m_{(L)} = m$
for $j = 1 : m_{(L)}$ **do**
$\quad \boldsymbol{\mathcal{S}}_j^{(L)} \leftarrow X_j$
end for
for $t = 1 : L$ **do**
$\quad$ {Change the value of L by $L - \ell^\star$ if you target the level $\ell^\star$}
$\quad \ell \leftarrow L - t$
$\quad$ **for** $k = 1 : m_{(\ell)}$ **do**
$\quad\quad \boldsymbol{\mathcal{S}}_k^{(\ell)} \leftarrow 0$
$\quad\quad$ **for** $j = 1 : m_{(\ell+1)}$ **do**
$\quad\quad\quad \boldsymbol{\mathcal{S}}_k^{(\ell)} \leftarrow \boldsymbol{\mathcal{S}}_k^{(\ell)} + \omega_{j,k}^{(\ell)} \boldsymbol{\mathcal{S}}_j^{(\ell+1)}$
$\quad\quad$ **end for**
$\quad$ **end for**
end for
$\boldsymbol{\mathcal{S}} \leftarrow \boldsymbol{\mathcal{S}}_1^{(0)}$
return $\boldsymbol{\mathcal{S}}$

The multi-level tree structure is very popular for computing ESG scores. For instance, the final ESG score corresponds to level 0. It is the weighted average of the 🟢**E**, 🔵**S** and 🔵**G** scores, which form the first level. Each pillar depends on a number of ESG themes, which constitutes the second level. As we have seen previously, an ESG theme is made up of several indicators. These last ones are located at the third level of the scoring tree. The computation of indicators requires some ESG metrics. Therefore, an ESG scoring model has at least four levels. For example, we have reported in Figure 2.12 an example of a tree from the MSCI scoring model. *Carbon emissions management and exposure* are two metrics (level 4). They are combined to form the indicator *carbon emissions* (level 3). MSCI uses this indicator and four others to define the *climate change* theme (level 2). It is one of the four themes of the environmental pillar (level 1).

Figure 2.12: An example of ESG scoring tree (MSCI methodology)
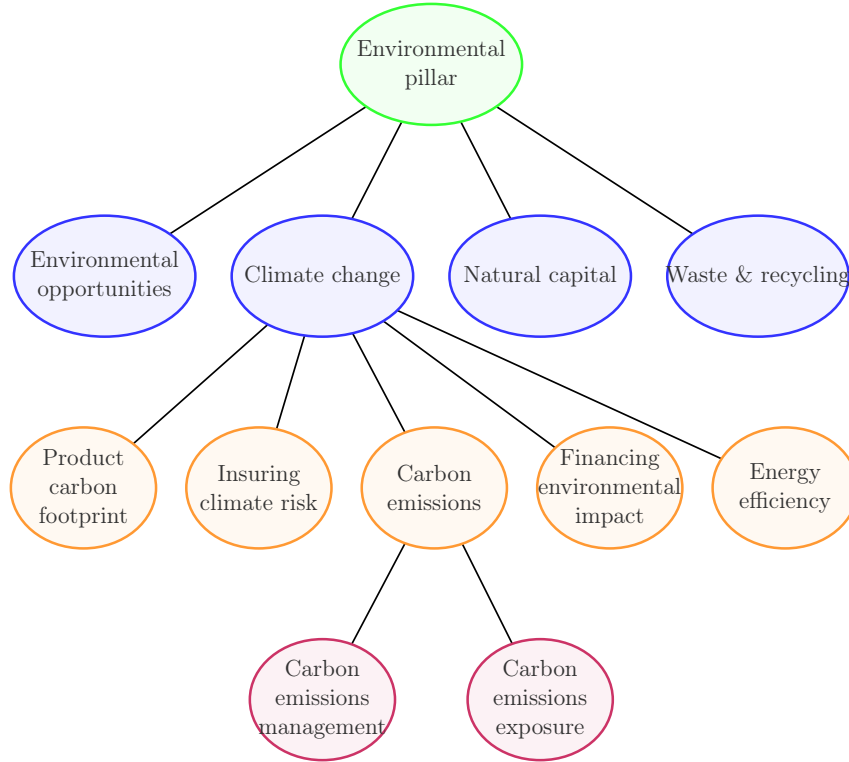
Score normalization

Why we need to normalize   Let $\omega_{(\ell)}$ be the $m_{(\ell+1)} \times m_{(\ell)}$ matrix, whose elements are $\omega_{j,k}^{(\ell)}$ for $j = 1, \ldots, m_{(\ell+1)}$ and $k = 1, \ldots, m_{(\ell)}$. We note $\boldsymbol{S}^{(\ell)} = \left(\boldsymbol{S}_1^{(\ell)}, \ldots, \boldsymbol{S}_{m_{(\ell)}}^{(\ell)}\right)$ the vector of scores at the tree level ℓ. We have:

$$\boldsymbol{S}^{(\ell)} = \omega_{(\ell)}^{\top} \boldsymbol{S}^{(\ell+1)}$$

At level 1, we obtain $\boldsymbol{S}_1^{(1)} = \omega_{(1)}^{\top} \boldsymbol{S}^{(2)}$. Since we have $\boldsymbol{S}^{(2)} = \omega_{(2)}^{\top} \boldsymbol{S}^{(3)}$, we deduce that $\boldsymbol{S}_1^{(1)} = \omega_{(1)}^{\top} \omega_{(2)}^{\top} \boldsymbol{S}^{(3)}$. By iterating the previous equation and noting that $\boldsymbol{S}^{(L)} = X$, the final score is equal to:

$$\boldsymbol{S} = \omega^{\top} X \tag{2.2}$$

where:

$$\omega = \omega_{(L-1)} \cdots \omega_{(1)} \omega_{(0)}$$

If we are interested in an intermediary score, we proceed in a similar way and we have:

$$\begin{aligned}
\boldsymbol{S}_k^{(\ell)} &= \mathbf{e}_k^{\top} \boldsymbol{S}^{(\ell)} \\
&= \omega^{\top} X
\end{aligned}$$

where:

$$\omega = \omega_{(L-1)} \cdots \omega_{(\ell-1)} \omega_{(\ell)}$$

We conclude that all the scores are a weighted average of initial metrics.

Let us consider the scoring tree given in Figure 2.11. We have:

$$\omega_{(1)} = \begin{bmatrix} 0.5 & 0 & 0 \\ 0.25 & 0 & 0 \\ 0.25 & 0.25 & 0 \\ 0 & 0.25 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and:

$$\omega_{(0)} = \frac{1}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

We deduce that:

$$\omega = \omega_{(1)}\omega_{(0)} = \frac{1}{12} \begin{bmatrix} 2 \\ 1 \\ 2 \\ 1 \\ 2 \\ 4 \end{bmatrix}$$

The expression of the final score is:

$$\boldsymbol{\mathcal{S}} = \frac{2X_1 + X_2 + 2X_3 + X_4 + 2X_5 + 4X_3}{12}$$

Let us assume that X follows a multivariate probability distribution $\mathbf{F}$. We deduce that $\boldsymbol{\mathcal{S}}$ follows the univariate probability distribution $\mathbf{G}$ defined by:

$$
\begin{aligned}
\mathbf{G}(s) &= \Pr\{\boldsymbol{\mathcal{S}} \le s\} \\
&= \Pr\left\{\omega^\top X \le s\right\} \\
&= \int \cdots \int \mathbb{1}\left\{\omega^\top x \le s\right\} \, d\mathbf{F}(x) \\
&= \int \cdots \int \mathbb{1}\left\{\sum_{j=1}^{m} \omega_j x_j \le s\right\} \, d\mathbf{F}(x_1, \ldots, x_m) \\
&= \int \cdots \int \mathbb{1}\left\{\sum_{j=1}^{m} \omega_j x_j \le s\right\} \, d\mathbf{C}(\mathbf{F}_1(x_1), \ldots, \mathbf{F}_m(x_m))
\end{aligned}
$$

Therefore, the distribution $\mathbf{G}$ depends on the copula function $\mathbf{C}$ and the marginals $(\mathbf{F}_1, \ldots, \mathbf{F}_m)$ of $\mathbf{F}$.

We first investigate the independent case. It follows that:

$$\mathbf{G}(s) = \int \cdots \int \mathbb{1}\left\{\sum_{j=1}^{m} \omega_j x_j \le s\right\} \prod_{j=1}^{m} d\mathbf{F}_j(x_j)$$

We deduce that $\mathbf{G}$ is a convolution probability distribution. In some cases, it corresponds to a well-known probability distribution. For example, if $X_j \sim \mathcal{N}\left(\mu_j, \sigma_j^2\right)$, we have $\omega_j X_j \sim \mathcal{N}\left(\omega_j \mu_j, \omega_j^2 \sigma_j^2\right)$. We deduce that:

$$\boldsymbol{\mathcal{S}} \sim \mathcal{N}\left(\sum_{j=1}^m \omega_j \mu_j, \sum_{j=1}^m \omega_j^2 \sigma_j^2\right) \equiv \mathcal{N}\left(\omega^\top \mu, \omega^\top \Sigma \omega\right)$$

where $\mu = (\mu_1, \ldots, \mu_m)$ and $\Sigma = \text{diag}\left(\sigma_1^2, \ldots, \sigma_m^2\right)$. In Figure 2.13, we have reported the density function of the intermediary and final scores for the tree 2.10 when $X_j \sim \mathcal{N}(0, 1)$. The four scores $\boldsymbol{\mathcal{S}}_1^{(1)}$, $\boldsymbol{\mathcal{S}}_2^{(1)}$, $\boldsymbol{\mathcal{S}}_3^{(1)}$ and $\boldsymbol{\mathcal{S}}$ are Gaussian, centered at 0 with different standard deviations[40]. We face here an issue because we cannot compare the different scores and it is impossible to have a homogeneous rule to assess whether a score is good or not.

Figure 2.13: Probability distribution of the scores (Tree 2.10)
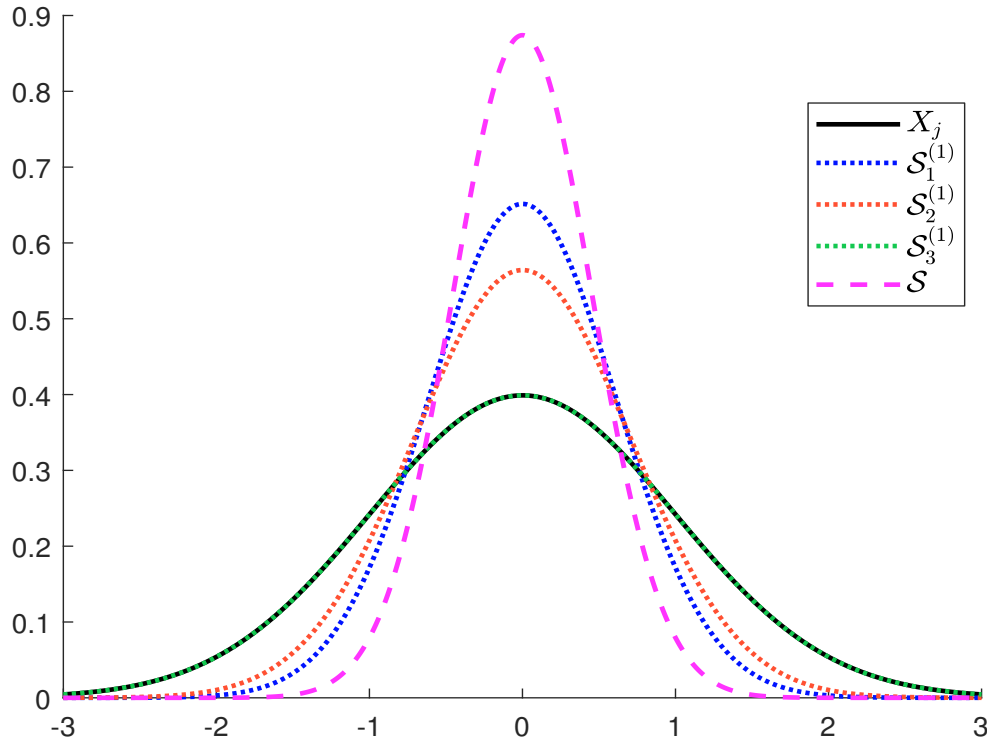
Example 3 *We assume that $X_1 \sim \mathcal{U}_{[0,1]}$ and $X_2 \sim \mathcal{U}_{[0,1]}$ are two independent random variables. We consider the score $\boldsymbol{\mathcal{S}}$ defined as:*

$$\boldsymbol{\mathcal{S}} = \frac{X_1 + X_2}{2}$$

We have $\boldsymbol{\mathcal{S}} \in [0, 1]$. In Figure 2.14, we consider a geometric interpretation of the probability mass function $\Pr\{\boldsymbol{\mathcal{S}} \leq s\} = \Pr\{X_1 + X_2 \leq 2s\}$. We distinguish two cases. The first case (a) corresponds to $s \leq 0.5$. The probability mass corresponds then to the right triangle with vertices $(0, 0)$, $(0, 2s)$ and $(2s, 0)$. It is equal to one-half the square, whose length is $2s$. The second case (b) corresponds to $0.5 \leq s \leq 1$. Here the probability mass is equal to the area of the polygon shape, whose vertices

---

[40]They are equal to 0.6124 for $\boldsymbol{\mathcal{S}}_1^{(1)}$, 0.7071 for $\boldsymbol{\mathcal{S}}_2^{(1)}$, 1 for $\boldsymbol{\mathcal{S}}_3^{(1)}$ and 0.4564 for $\boldsymbol{\mathcal{S}}$.

Figure 2.14: Geometric interpretation of the probability mass function (Example 3)
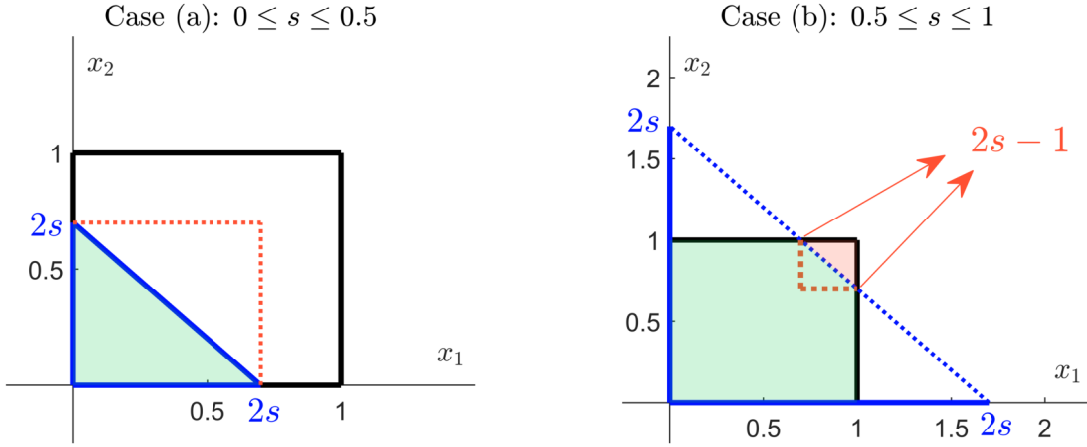
are $(0,0)$, $(1,0)$, $(1, 2s-1)$, $(2s-1, 1)$ and $(0,1)$. This is equivalent to computing the area of the unit square minus the right triangle with vertices $(1, 2s-1)$, $(1,1)$ and $(2s-1, 1)$. We notice that the area of the right triangle is equal to one-half the square, whose length is $2-2s$. We deduce that:

$$\Pr\{\boldsymbol{S} \le s\} = \begin{cases} \dfrac{1}{2}(2s)^2 & \text{if } 0 \le s \le \dfrac{1}{2} \\ 1 - \dfrac{1}{2}(2-2s)^2 & \text{if } \dfrac{1}{2} \le s \le 1 \end{cases}$$

Finally, we obtain:

$$\mathbf{G}(s) = \begin{cases} 2s^2 & \text{if } 0 \le s \le \dfrac{1}{2} \\ -1 + 4s - 2s^2 & \text{if } \dfrac{1}{2} \le s \le 1 \end{cases}$$

The density function is then:

$$g(s) = \begin{cases} 4s & \text{if } 0 \le s \le \dfrac{1}{2} \\ 4 - 4s & \text{if } \dfrac{1}{2} \le s \le 1 \end{cases}$$

The previous results can be extended to the case $m > 2$. We can show that the score follows the Bates distribution[41]:

$$\boldsymbol{S} = \frac{X_1 + X_2 + \cdots + X_m}{m} \sim \mathfrak{Bates}(m)$$

In Figures 2.15 and 2.16, we report the density and distribution functions of the score for several values of m. We verify that the shape of the score depends on the number m of features. In particular, we observe that the score tends to the Dirac distribution $\delta(x - 1/2)$.

Let us now investigate the dependent case. If we assume that $X \sim \mathcal{N}(\mu, \Sigma)$, the score $\boldsymbol{S} = \omega^\top X$ is normally distributed: $\boldsymbol{S} \sim \mathcal{N}(\omega^\top \mu, \omega^\top \Sigma \omega)$. We consider that $\mu_j = 0$, $\sigma_j = 1$ and $\rho_{j,k} = \rho$ for $j \ne k$. Since the covariance matrix is the constant correlation matrix $\mathbb{C}_m(\rho)$, we deduce that

---

[41]See Exercise 2.4.1 on page 126.

Figure 2.15: Probability density function of $\boldsymbol{\mathcal{S}}$ (uniform distribution)
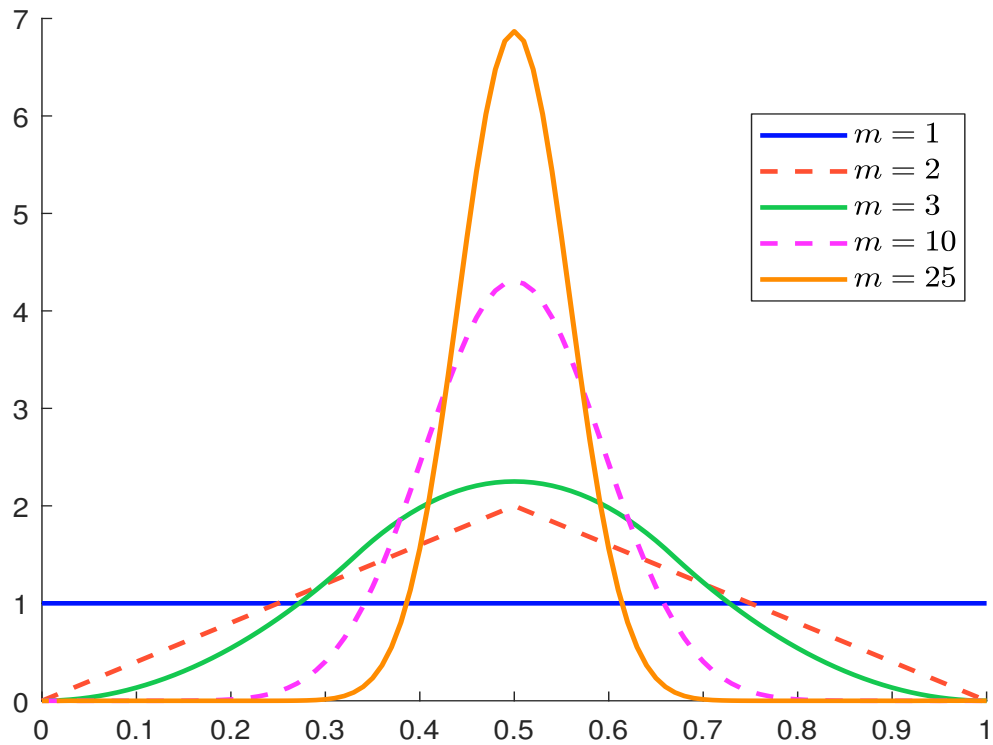
Figure 2.16: Cumulative distribution function of $\boldsymbol{\mathcal{S}}$ (uniform distribution)
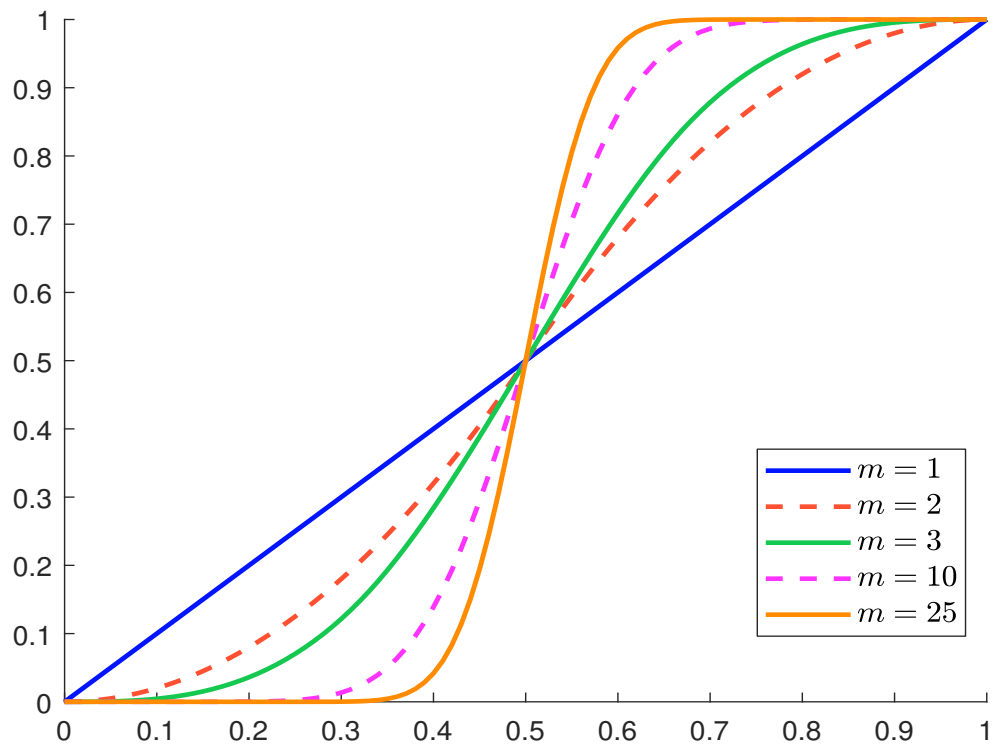
$\mathbb{E}\left[\boldsymbol{S}\right] = 0$ and[42]:

$$
\begin{aligned}
\mathrm{var}\left(\boldsymbol{S}\right) &= \omega^{\top}\mathbb{C}_m\left(\rho\right)\omega \\
&= \sum_{j=1}^{m}\sum_{k=1}^{m}\omega_j\omega_k\rho_{j,k} + \left(\sum_{j=1}^{m}\omega_j^2\rho - \sum_{j=1}^{m}\omega_j^2\rho\right) \\
&= \rho\sum_{j=1}^{m}\sum_{k=1}^{m}\omega_j\omega_k + \left(1-\rho\right)\sum_{j=1}^{m}\omega_j^2 \\
&= \rho\mathcal{S}^2\left(w\right) + \left(1-\rho\right)\mathcal{H}\left(\omega\right)
\end{aligned}
$$

where $\mathcal{S}\left(w\right) = \sum_{j=1}^{m}\omega_j$ is the sum index and $\mathcal{H}\left(\omega\right) = \sum_{j=1}^{m}\omega_j^2$ is the Herfindahl index. Generally, the weights are normalized and we obtain $\boldsymbol{S}\sim\mathcal{N}\left(0,\sigma_{\boldsymbol{S}}^2\right)$ where:

$$
\sigma_{\boldsymbol{S}} = \sqrt{\rho + \left(1-\rho\right)\mathcal{H}\left(\omega\right)}
$$

Since $0\leq\mathcal{H}\left(\omega\right)\leq 1$, $\sigma_{\boldsymbol{S}}\in\left[\sqrt{\rho},1\right]$. When the weights are equal ($\omega_j = 1/m$), the previous formula reduces to[43]:

$$
\sigma_{\boldsymbol{S}} = \sqrt{\rho + \frac{\left(1-\rho\right)}{m}}
$$

Let us build the confidence interval of the score at the confidence level α.   We have $\Pr\left\{\boldsymbol{S}\in\left[\boldsymbol{S}^-\left(m,\alpha\right),\boldsymbol{S}^+\left(m,\alpha\right)\right]\right\} = \alpha$. Since the expectation of the score is equal to zero, we consider an interval centered at 0. We deduce that:

$$
\boldsymbol{S}^+\left(m,\alpha\right) = \Phi^{-1}\left(\frac{1+\alpha}{2}\right)\sqrt{\rho + \frac{\left(1-\rho\right)}{m}}
$$

and $\boldsymbol{S}^-\left(m,\alpha\right) = -\boldsymbol{S}^+\left(m,\alpha\right)$.

In Figure 2.17, we illustrate the shrinkage issue of the score when α is set to 99.75%. At this confidence level, a standard normal random variable lies between -3 and $+3$. This is the range that we have in mind when we build a z-score. We observe that the shrinkage begins when the score is made up of three features and a correlation lower than 80%. The shrinkage issue increases with the number of features. Let us consider for example an ESG score that depends on 20 ESG metrics, that have an average correlation of 20%. While the range of the features is between -3 and $+3$, the aggregate ESG score lies between -1.5 and $+1.5$, meaning that the support of the score has been divided by a factor of two!

How to normalize?   The previous analysis implies that we must normalize the raw data and the scores at the different levels such that they follow the same probability distribution $\mathbf{F}_{\boldsymbol{S}}$. Equation (2.1) is no longer valid and we deduce that the node values of the multi-level tree structure are equal to:

$$
\boldsymbol{S}_k^{(\ell)} = \varphi\left(\sum_{j=1}^{m_{(\ell+1)}}\omega_{j,k}^{(\ell)}\boldsymbol{S}_j^{(\ell+1)}\right) \tag{2.3}
$$

where $\varphi\left(s\right)$ is the normalization function. We also have $\boldsymbol{S}_j^{(L)} = \varphi\left(X_j\right)$.

Let X be a variable to normalize and $\{x_1,\ldots,x_n\}$ a sample of n observations. In practice, there are three main approaches:

---

[42]We note $\mathcal{S}\left(w\right) = \sum_{j=1}^{m}\omega_j$. We deduce that $\left(\sum_{j=1}^{m}\omega_j\right)^2 = \sum_{j=1}^{m}\sum_{k=1}^{m}\omega_j\omega_k = \mathcal{S}^2\left(w\right)$.

[43]We have $\mathcal{H}\left(\omega\right) = \sum_{j=1}^{m}\left(1/m\right)^2 = 1/m$.

Figure 2.17: Upper and lower bounds of the aggregate score when $\alpha = 99.75\%$
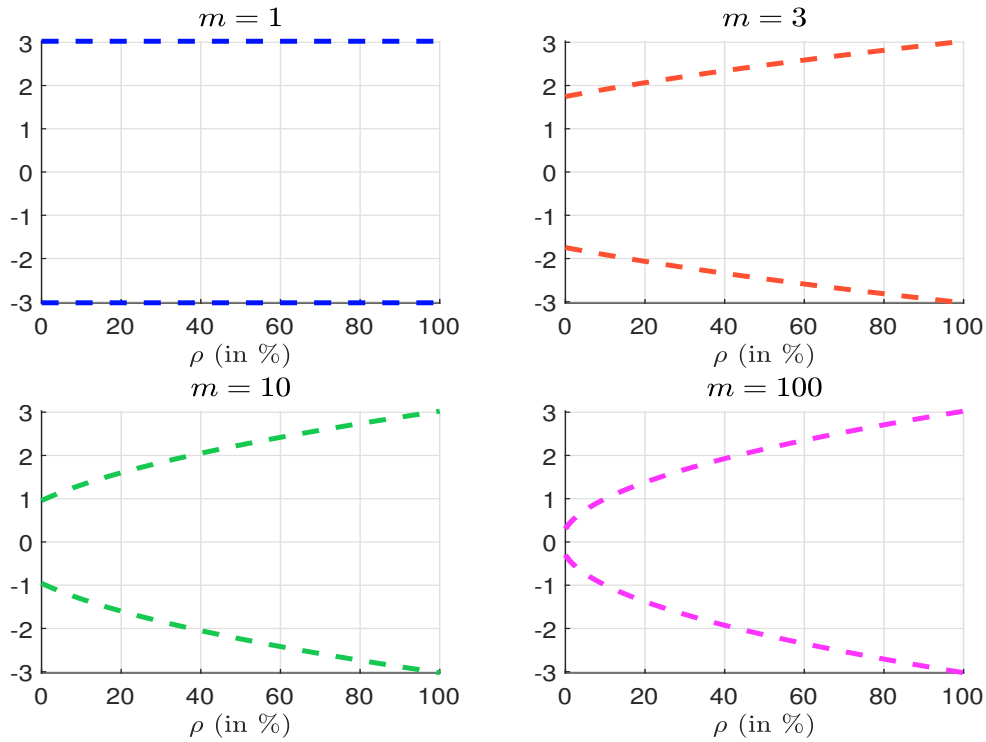
1. The first one is the m-score (or min-max) normalization:

$$m_i = \frac{x_i - x^-}{x^+ - x^-}$$

where $x^- = \min x_i$ and $x^+ = \max x_i$. This is the most naive approach to obtain a $0/1$ normalization.

2. The second approach is the q-score (or quantile) normalization:

$$q_i = \mathbf{H}(x_i)$$

where $\mathbf{H}$ is the distribution function of X. When the distribution of X is unknown, we replace $\mathbf{H}$ by the empirical distribution $\hat{\mathbf{H}}$: $q_i = \hat{\mathbf{H}}(x_i)$. In both cases, we obtain a $0/1$ normalization.

3. The third method is the famous z-score normalization:

$$z_i = \frac{x_i - \mu}{\sigma}$$

where μ and σ are the mathematical expectation and standard deviation of X. Again, when the distribution of X is unknown, we use the empirical mean and standard deviation of the sample:

$$z_i = \frac{x_i - \hat{\mu}}{\hat{\sigma}}$$

By construction, we have $z_i \in (-\infty, \infty)$. Nevertheless, we have seen that $\Pr\{-3 \leq N(0,1) \leq 3\} \approx 99.75\%$. Generally, we consider that the z-score method produces a $-3/+3$ normalization.

Among the three approaches, only the second approach satisfies the consistency property. Indeed, if $X \sim \mathbf{H}$ and is continuous, we know that $Y = \mathbf{H}(X)$ is a uniform random variable[44]. $Y = \mathbf{H}(X)$ defines the *probability integral transform*, which plays an important role in statistics and probability. The min-max approach is consistent only if $X \sim \mathcal{U}_{[x^-, x^+]}$, whereas the z-score normalization is consistent if the original data are normally distributed.

<div align="center">

Box 2.3: Computing the empirical distribution $\hat{\mathbf{H}}$

</div>

Let $\{x_1, x_2, \ldots, x_n\}$ be the sample. We have:

$$q_i = \hat{\mathbf{H}}(x_i) = \Pr\{X \leq x_i\} = \frac{\#\{x_j \leq x_i\}}{n_q}$$

We can use two normalization factors: $n_q = n$ or $n_q = n + 1$. For example, if $n = 4$, we have $q_i \in \{0.25, 0.5, 0.75, 1\}$ if $n_q = n$, and $q_i \in \{0.2, 0.4, 0.6, 0.8\}$ if $n_q = n + 1$. The second solution is better because $q_i \in \,]0, 1[$. Therefore, we can transform q_i into a random variable Y with probability distribution $\mathbf{G}$ by considering the inverse probability integral transform:

$$Y = \mathbf{G}^{-1}(q_i) \sim \mathbf{G}$$

For example, we can transform a q-score into a z-score by considering the Gaussian quantile function:

$$z = \Phi^{-1}(q)$$

With the second solution, we are sure that $z \notin (-\infty, \infty)$.

To obtain an a/b normalization, we consider the following property:

$$\mathcal{U}_{[a,b]} = a + (b - a)\,\mathcal{U}_{[0,1]}$$

Therefore, we apply the following transform to obtain the new score:

$$q' = a + (b - a)\,q$$

The q-score is distributed according to the uniform distribution. It has the advantage to be normalized between 0 and 1. However, it has the disadvantage to be a flat score, meaning that the extreme scores have the same probability to occur as the mean score. The z-score is distributed according to the Gaussian distribution. It has a bell-curve shape, meaning that the extreme scores have a lower probability to occur than the mean score. Therefore, it is interesting to combine the two properties to obtain a discriminant score between 0 and 1. This is done by considering the $\textit{b}$-score using a Beta distribution $\mathcal{B}(\alpha, \beta)$. In this case, we have:

$$\textit{b}_i = \mathfrak{B}^{-1}(\mathbf{H}(x_i); \alpha, \beta)$$

---

[44]We have $Y \in [0, 1]$ and:

$$
\begin{aligned}
\Pr\{Y \leq y\} &= \Pr\{\mathbf{H}(X) \leq y\} \\
&= \Pr\{X \leq \mathbf{H}^{-1}(y)\} \\
&= \mathbf{H}(\mathbf{H}^{-1}(y)) \\
&= y
\end{aligned}
$$

We deduce that $Y \sim \mathcal{U}_{[0,1]}$.

When $\alpha = \beta$, This creates a symmetric score around 0.5. To obtain a left-skewed score (when having more best-in-class issuers than worst-in-class issuers), we use $\alpha > \beta$. The case $\alpha < \beta$ produces a right-skewed score (when having more worst-in-class issuers than best-in-class issuers). The standard values for parameters are $\alpha = \beta = 2$.

Example 4 *The data are normally distributed with mean $\mu = 5$ and standard deviation $\sigma = 2$. To map these data into a $0/1$ score, we consider the following transform:*

$$s := \varphi(x) = \mathfrak{B}^{-1}\left(\Phi\left(\frac{x-5}{2}\right); \alpha, \beta\right)$$

In Figure 2.18, we report the transform function $\varphi(x)$ and the final distribution for three sets of parameters (α, β).

Figure 2.18: Transforming data into b-score using the Beta distribution
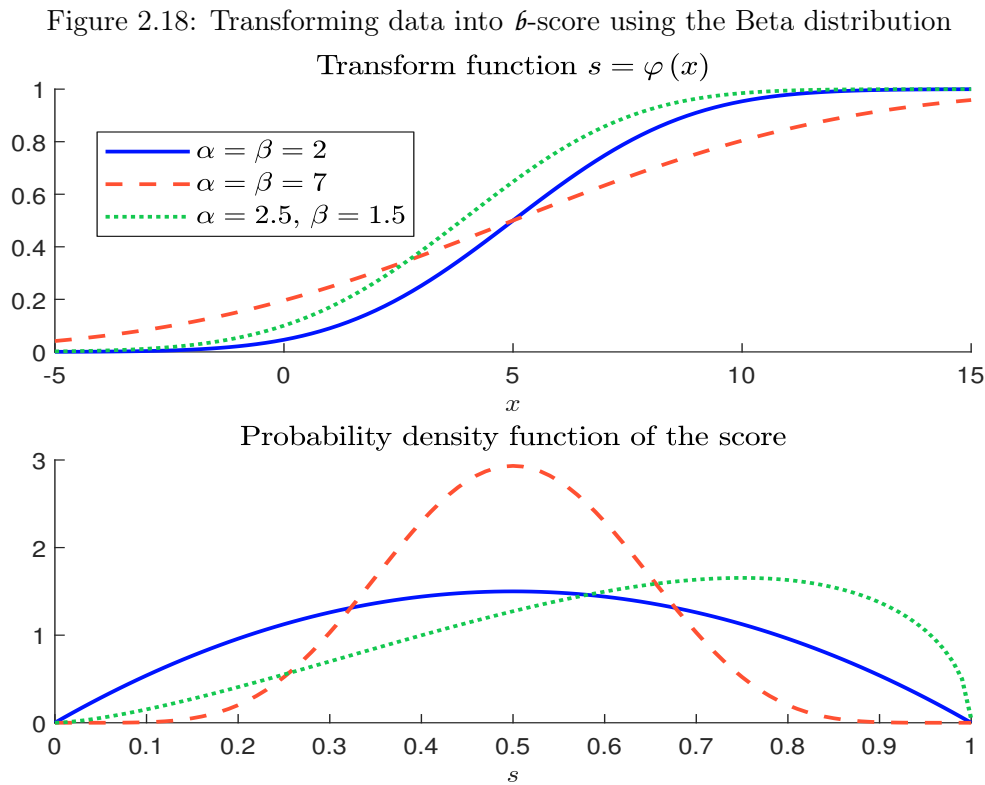
Remark 15 *ESG data providers do not publish their statistical methodology for computing ESG scores[45]. It is then difficult to know which normalization approach is used. Nevertheless, they give the scale of the scores. Bloomberg, S&P Global and Sustainalytics use a range from 0 to 100, Refinitiv from 1 to 100, MSCI from 0 to 10, ISS ESG from 1 to 10, etc. Some asset managers use a scale between -3 and $+3$.*

---

[45]Only S&P Dow Jones Indices (2022, page 8) indicates that "*the normalization is performed by a sigmoid-function on a standard z-score*":

$$\boldsymbol{S} = \frac{2}{1 + e^{-z}} - 1 \in [-1, 1]$$

Example 5 *We consider the raw data of* 9 *companies in Table 2.11. These companies belong to the same industry. The first variable measures the carbon intensity of scope* $1 + 2$ *in 2020, while the second variable is the variation of carbon emissions between 2015 and* 2020. *We would like to create the score* $\boldsymbol{\mathcal{S}} \equiv 70\% \cdot X_1 + 30\% \cdot X_2$.

Table 2.11: Raw data of 9 companies (carbon emissions and carbon momentum)

Firm	Carbon intensity (in tCO$_2$e/\$ mn)	Carbon momentum (in %)
1	94.0	−3.0
2	38.6	−5.5
3	30.6	5.6
4	74.4	−1.3
5	97.1	−16.8
6	57.1	−3.5
7	132.4	8.5
8	92.5	−9.1
9	64.9	−4.6

To create the synthetic score, we must analyze the data. An ESG investor prefers to be exposed to low-carbon companies than to high-carbon companies. Similarly, he favors firms that have reduced their carbon emissions in the past. If we consider the raw variables as the two features, a high value of the score will indicate a worst-in-class company while a low value of the score will indicate a best-in-class company. If we prefer that high scores correspond to best-in-class scores, we need to take the opposite of these data. We consider the first choice. In Table 2.12, we report the computation of the score by using a q-score 0/100 normalization. For instance, company #7 has the highest carbon emission and then the highest score $q_1 = 90$. It is followed by company #5 which has a score of 80. We verify that the mean of q_1 is equal to $(0 + 100)/2$ and its standard deviation is approximately[46] equal to $\sqrt{(100 − 0)^2/12}$. For the carbon momentum, the best issuer is company #5 with a trend of $−16.8\%$, and its q-score is equal to 10. We compute then the score $s = 0.7 \cdot q_1 + 0.3 \cdot q_2$. This indicates that the ESG investor is more sensitive to carbon intensity than carbon intensity. Said differently, he would like to build a score primarily based on the carbon intensity, but he would like to penalize companies that increase their carbon emissions. For the first firm, we obtain $s = 0.7 \times 70 + 0.3 \times 60 = 67$. We notice that the standard deviation of the variable s is equal to 20.60, which is lower than 27.39. Again, we have to normalize the variable s and we obtain the final score. We also report the rank $\mathfrak{R}$ and we obtain the following ordering:

$$\#2 \succ \#3 \succ \#6 \succ \#9 \succ \#8 \succ \#4 \succ \#5 \succ \#1 \succ \#7$$

We deduce that the best-in-class and worst-in-class issuers are respectively companies #2 and #7. If we consider the z-score normalization, the results are given in Table 2.13. For example, z_1 and z_2 are equal to $(92.5 − 75.73)/31.95 = 0.525$ and $(−9.10 + 3.30)/7.46 = −0.778$ for company #8. We deduce that $s = 0.7 \times 0.525 + 0.3 \times (−0.778) = 0.134$. Again we observe that the variable s is not normalized since its standard deviation is not equal to 1. We deduce that $\boldsymbol{\mathcal{S}} = (0.134 − 0.000)/0.759 = 0.177$. Finally, we obtain the following ordering:

$$\#2 \succ \#3 \succ \#6 \succ \#9 \succ \#5 \succ \#4 \succ \#8 \succ \#1 \succ \#7$$

We conclude that the two approaches produce two different rankings.

Table 2.12: Computation of the score $\boldsymbol{S} \equiv 70\% \cdot X_1 + 30\% \cdot X_2$ (q-score 0/100 normalization)

#	X_1	q_1	X_2	q_2	s	$\boldsymbol{S}$	$\mathfrak{R}$
1	94.00	70.00	−3.00	60.00	67.00	80.00	8
2	38.60	20.00	−5.50	30.00	23.00	10.00	1
3	30.60	10.00	5.60	80.00	31.00	20.00	2
4	74.40	50.00	−1.30	70.00	56.00	60.00	6
5	97.10	80.00	−16.80	10.00	59.00	70.00	7
6	57.10	30.00	−3.50	50.00	36.00	30.00	3
7	132.40	90.00	8.50	90.00	90.00	90.00	9
8	92.50	60.00	−9.10	20.00	48.00	50.00	5
9	64.90	40.00	−4.60	40.00	40.00	40.00	4
Mean	75.73	50.00	−3.30	50.00	50.00	50.00	
Std-dev.	31.95	27.39	7.46	27.39	20.60	27.39	

Table 2.13: Computation of the score $\boldsymbol{S} \equiv 70\% \cdot X_1 + 30\% \cdot X_2$ (z-score normalization)

#	X_1	z_1	X_2	z_2	s	$\boldsymbol{S}$	$\mathfrak{R}$
1	94.00	0.572	−3.00	0.040	0.412	0.543	8
2	38.60	−1.162	−5.50	−0.295	−0.902	−1.188	1
3	30.60	−1.413	5.60	1.193	−0.631	−0.831	2
4	74.40	−0.042	−1.30	0.268	0.051	0.067	6
5	97.10	0.669	−16.80	−1.810	−0.075	−0.099	5
6	57.10	−0.583	−3.50	−0.027	−0.416	−0.548	3
7	132.40	1.774	8.50	1.582	1.716	2.261	9
8	92.50	0.525	−9.10	−0.778	0.134	0.177	7
9	64.90	−0.339	−4.60	−0.174	−0.290	−0.382	4
Mean	75.73	0.000	−3.30	0.000	0.000	0.000	
Std-dev.	31.95	1.000	7.46	1.000	0.759	1.000	

Table 2.14: Comparison of the different scoring methods

#	q		z		qz		zq		bz		$bz^\star$	
	$\boldsymbol{S}$	$\mathfrak{R}$	$\boldsymbol{S}$	$\mathfrak{R}$	$\boldsymbol{S}$	$\mathfrak{R}$	$\boldsymbol{S}$	$\mathfrak{R}$	$\boldsymbol{S}$	$\mathfrak{R}$	$\boldsymbol{S}$	$\mathfrak{R}$
1	80.00	8	0.54	8	76.27	8	0.84	8	0.66	8	0.81	8
2	10.00	1	−1.19	1	9.19	1	−1.28	1	0.20	1	0.30	1
3	20.00	2	−0.83	2	21.37	2	−0.84	2	0.29	2	0.38	2
4	60.00	6	0.07	6	54.13	5	0.25	6	0.52	6	0.70	6
5	70.00	7	−0.10	5	56.65	6	0.52	7	0.51	5	0.64	5
6	30.00	3	−0.55	3	24.42	3	−0.52	3	0.34	3	0.50	3
7	90.00	9	2.26	9	98.04	9	1.28	9	0.93	9	0.96	9
8	50.00	5	0.18	7	60.39	7	0.00	5	0.56	7	0.72	7
9	40.00	4	−0.38	4	30.96	4	−0.25	4	0.39	4	0.56	4
Mean	50.00		0.00		47.94		0.00		0.49		0.62	
Std-dev.	27.39		1.00		28.79		0.82		0.22		0.21	

In Table 2.14, we compare the scores and the ranks obtained for different scoring schemes. Besides the q- and z-scores, we consider the following transforms:

- The qz-score is defined as:

$$qz = c \cdot \Phi\left(z\right) \in [0, c]$$

  where $c = 100$ is the scaling factor.

- The zq-score is defined as:

$$zq = \Phi^{-1}\left(\frac{q}{c}\right) \in [-3, 3]$$

- The $\textit{bz}$-score is defined as:

$$\textit{bz} = \mathfrak{B}^{-1}\left(\Phi\left(z\right); \alpha, \beta\right) \in [0, 1]$$

  where $\alpha = \beta = 2$.

- The $\textit{bz}^\star$-score is a modification of the $\textit{bz}$-score by using $\alpha = 2.5$ and $\beta = 1.5$.

We verify that the $\textit{bz}^\star$-score is left-skewed and the mean is above $1/2$.

Remark 16 *Most ESG scoring systems are sector neutral, meaning that the normalization is done at the sector (or industry) level, not at the issuer universe level. ESG scores are then **relative** scores (with respect to the sector/industry), not **absolute** scores. This is the concept of best-in-class/worst-in-class issuers. A best-in-class company is then not a best-in-universe issuer. Let us consider the example where the score of corporate A is $\boldsymbol{\mathcal{S}}_A = +2$ and the score of corporate B is $\boldsymbol{\mathcal{S}}_B = +1$. We have:*

- *If A and B belong to the same sector, we have $A \succ B$;*

- *If A and B belong to two different sectors, we may have $A \succ B$ or $B \succ A$.*

The preference ordering $\succ$ is then partial and not total.

An example with the CEO pay ratio   The CEO pay ratio is calculated by dividing the CEO's compensation by the pay of the median employee. It is one of the key metrics for the (G) pillar. It has been imposed by the Dodd-Frank Act, which requires that publicly traded companies disclose:

1. the median total annual compensation of all employees other than the CEO;

2. the ratio of the CEO's annual total compensation to that of the median employee;

3. the wage ratio of the CEO to the median employee.

According to the American Federation of Labor and Congress of Industrial Organizations (AFL-CIO, https://aflcio.org), the average S&P 500 company's CEO-to-worker pay ratio was 324-to-1 in 2021. In Table 2.15, we have reported some data collected by this organization (P is the median worker pay (in \$) and R is the CEO pay ratio).

Computing the scores for the pay ratio is a real challenge, because the probability distribution of the pay ratio has both a high skew and a big kurtosis. To illustrate these issues, Figure 2.19 shows the histogram of the CEO pay ratios for all US public companies[47] for the fiscal year 2021. If we

---

[46]Since we have only 9 observations, we observe a small difference between the true and the sample values.

[47]We use the database *Fiscal 2021 CEO Pay Ratios* by Mark Siciliano, which can be downloaded at the University of Alabama: https://ir.ua.edu/handle/123456789/8639.

Table 2.15: Examples of CEO pay ratio (June 2021)

Company name	P	R	Company name	P	R
Abercrombie & Fitch	1 954	4,293	Netflix	202 931	190
McDonald's	9 291	1,939	BlackRock	133 644	182
Coca-Cola	11 285	1,657	Pfizer	98 972	181
Gap	6 177	1,558	Goldman Sachs	138 854	178
Alphabet	258 708	1,085	MSCI	55 857	165
Walmart	22 484	983	Verisk Analytics	77 055	117
Estee Lauder	30 733	697	Facebook	247 883	94
Ralph Lauren	21 358	570	Invesco	125 282	92
NIKE	25 386	550	Boeing	158 869	90
Citigroup	52 988	482	Citrix Systems	181 769	80
PepsiCo	45 896	368	Harley-Davidson	187 157	59
Microsoft	172 512	249	Amazon.com	28 848	58
Apple	57 596	201	Berkshire Hathaway	65 740	6

Source: <https://aflcio.org> (June 2021)

compute the z-score directly from the pay ratio, we obtain the blue histogram in Figure 2.20. This z-score has a mean around 0 and a standard deviation of 1, but we have $z \in (-0.338, 38.669)$. If we do the same exercise with the logarithm of the CEO pay ratio, we obtain the red histogram. In this case, we have $z \in (-3.561, 4.545)$. This is better, but it is not perfect. This example demonstrates that we must conduct a deep analysis of each data before applying a blind scoring approach. Most ESG data are skewed with fat tails, some of them are binary, others take discrete values, etc. In this context, data analysis is essential to choose the right normalization and scoring transform.

Exercise 6 *The database of the CEO pay ratios for all US public companies contains several sector/industry variables. Compute the z-score of the CEO pay ratio and its logarithm at the sector level. Identify the most problematic sectors. Same question if we consider the industry level.*

2.2.3 Other statistical methods

Since ESG scoring is an unsupervised statistical method, the use of alternative statistical methods is limited. Indeed, we can use K-means clustering, hierarchical clustering, and principal component analysis to perform dimension reduction of the variables or to create synthetic features. However, it is difficult to go further with these methods. The use of classical parametric supervised methods, such as discriminant analysis and binary choice models, is not well suited except when the goal is to identify bad ESG risks. They are not used in practice. In fact, the only alternative statistical method that adds some value is the lasso method, which is described in Appendix A.2.3 on page 679. The underlying idea is to perform the penalized linear regression $\boldsymbol{\mathcal{S}} = \beta^\top X + U$ subject to the $\mathcal{L}_1$-norm constraint $\|\beta\|_1 \le \tau$ and identify the most important variables of the ESG scoring system. The lasso regression is then used as an ex-post statistical method to estimate the contribution of each variable. In practice, we find that a complex ESG scoring system based on more than 100 variables can be replicated with fewer variables.

2.2.4 Performance evaluation criteria

This section is dedicated to the performance assessment of a score. Backtesting an ESG score is relatively close to backtesting a credit score. Even if the tree-based scoring model is an unsuper-
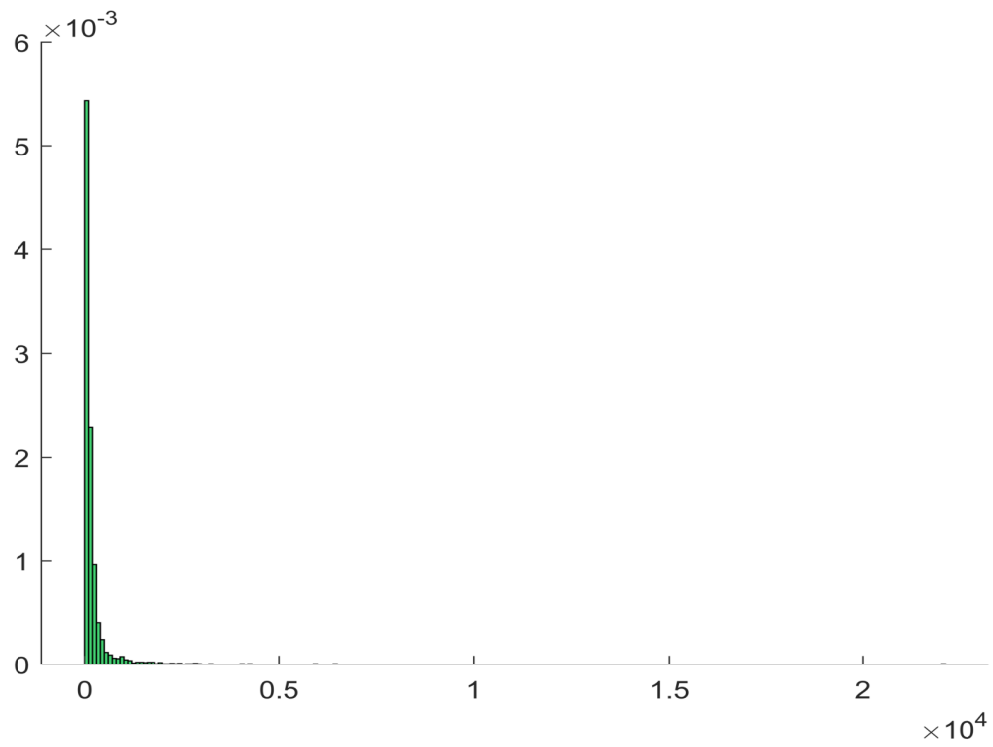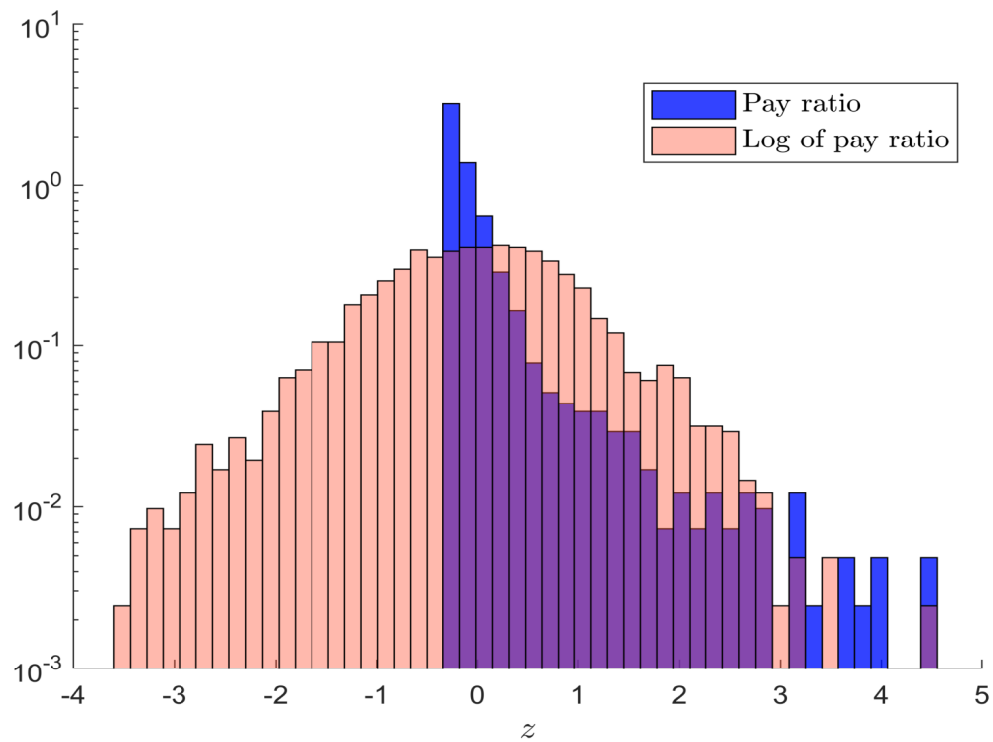
Figure 2.19: Histogram of the CEO pay ratio

Figure 2.20: Histogram of z-score applied to the CEO pay ratio

vised learning approach, we have seen that it is possible to build supervised models by using a control variable. The response function can be exogenous such as the controversy index $\mathcal{C}(t)$ or the controversy indicator $I(t) = \mathbb{1}\{\mathcal{C}(t) > 0\}$. It can also be endogenous by considering control variables based on the score $\mathcal{S}(t + \delta)$ in the future. Examples are the best-in-class risk indicator $\mathcal{G}(t, \delta) = \mathbb{1}\{\mathcal{S}(t + \delta) \geq s^\star\}$ or the worst-in-class risk indicator $\mathcal{B}(t, \delta) = \mathbb{1}\{\mathcal{S}(t + \delta) \leq s^\star\}$. They are noted $\mathcal{G}$ and $\mathcal{B}$ by analogy with the risk theory that distinguishes good and bad risk. For instance, credit scoring models are mainly based on bad risk detection. This is also the same thing in the case of responsible investing. Nevertheless, we might also want to have a statistical model, whose main objective is to select the good issuers. In the case of the ESG momentum strategy, the response variable depends on the improvement of the ESG score. It can be defined as[48] $\mathcal{M}(t, \delta) = \mathbb{1}\{\mathcal{S}(t + \delta) - \mathcal{S}(t) \geq \Delta_s^\star\}$.

Remark 17 *In this section, the score $\mathcal{S}$ is not necessarily the final ESG score. It corresponds to any score or screening rule derived from the ESG scoring model. For instance, it may corresponds to a selection score, an exclusion score, etc.*

We notice that most control variables are binary. Therefore, we can use the classical tool of credit scoring and follow Roncalli (2020a, Chapter 15) to assess the performance of scoring models. In the first paragraph, we use information theory to know if the scoring system is informative or not. The second paragraph presents the graphical tools to measure the classification accuracy of the score. We then define the different statistical measures to estimate the score performance. Finally, we consider the assessment tools when there is no response function.

Shannon entropy

Definition and properties The entropy is a measure of unpredictability or uncertainty of a random variable. Let (X, Y) be a random vector where $p_{i,j} = \Pr\{X = x_i, Y = y_j\}$, $p_i = \Pr\{X = x_i\}$ and $p_j = \Pr\{Y = y_j\}$. The Shannon entropy of the discrete random variable X is given by[49]:

$$\mathcal{I}(X) = -\sum\nolimits_{i=1}^{n} p_i \ln p_i$$

We have the property $0 \leq \mathcal{I}(X) \leq \ln n$. The entropy is equal to zero if there is a state i such that $p_i = 1$ and is equal to $\ln n$ in the case of the uniform distribution ($p_i = 1/n$). The Shannon entropy is a measure of the average information of the system. The lower the Shannon entropy, the more informative the system. For a random vector (X, Y), we have:

$$\mathcal{I}(X, Y) = -\sum\nolimits_{i=1}^{n} \sum\nolimits_{j=1}^{n} p_{i,j} \ln p_{i,j}$$

We deduce that the conditional information of Y given X is equal to:

$$\begin{aligned} \mathcal{I}(Y \mid X) &= \mathbb{E}[\mathcal{I}(Y \mid X = x)] \\ &= -\sum\nolimits_{i=1}^{n} \sum\nolimits_{j=1}^{n} p_{i,j} \ln \frac{p_{i,j}}{p_i} \\ &= \mathcal{I}(X, Y) - \mathcal{I}(X) \end{aligned}$$

We have the following properties:

---

[48]Alternative measures are $\mathcal{M}(t, \delta) = \mathbb{1}\{\mathcal{S}(t + \delta) - \mathcal{S}(t) > \Delta_s^\star | \mathcal{S}(t) \geq s^\star\}$ when we prefer to select companies among the best-in-class issuers and $\mathcal{M}(t, \delta) = \mathbb{1}\{\mathcal{S}(t + \delta) - \mathcal{S}(t) > \Delta_s^\star | \mathcal{S}(t) \leq s^\star\}$ when we want to detect worst-in-class issuers that will improve their ESG score.

[49]We use the convention $p_i \ln p_i = 0$ when p_i is equal to zero.

- if X and Y are independent, we have $\mathcal{I}(Y \mid X) = \mathcal{I}(Y)$ and $\mathcal{I}(X,Y) = \mathcal{I}(Y) + \mathcal{I}(X)$;

- if X and Y are perfectly dependent, we have $\mathcal{I}(Y \mid X) = 0$ and $\mathcal{I}(X,Y) = \mathcal{I}(X)$.

The amount of information obtained about one random variable, through the other random variable is measured by the mutual information:

$$
\begin{aligned}
\mathcal{I}(X \cap Y) &= \mathcal{I}(Y) + \mathcal{I}(X) - \mathcal{I}(X,Y) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} p_{i,j} \ln \frac{p_{i,j}}{p_i p_j}
\end{aligned}
$$

Figure 2.21: Examples of Shannon entropy calculation

1/36	1/36	1/36	1/36	1/36	1/36
1/36	1/36	1/36	1/36	1/36	1/36
1/36	1/36	1/36	1/36	1/36	1/36
1/36	1/36	1/36	1/36	1/36	1/36
1/36	1/36	1/36	1/36	1/36	1/36
1/36	1/36	1/36	1/36	1/36	1/36

$$\mathcal{I}(X) = \mathcal{I}(Y) = 1.792$$
$$\mathcal{I}(X,Y) = 3.584$$
$$\mathcal{I}(X \cap Y) = 0$$

1/6					
	1/6				
		1/6			
			1/6		
				1/6	
					1/6

$$\mathcal{I}(X) = \mathcal{I}(Y) = 1.792$$
$$\mathcal{I}(X,Y) = 1.792$$
$$\mathcal{I}(X \cap Y) = 1.792$$

1/24	1/24				
1/24	1/24	1/24	1/48		
	1/24	1/6	1/24	1/48	
	1/48	1/24	1/6	1/24	
		1/48	1/24	1/24	1/24
				1/24	1/24

$$\mathcal{I}(X) = \mathcal{I}(Y) = 1.683$$
$$\mathcal{I}(X,Y) = 2.774$$
$$\mathcal{I}(X \cap Y) = 0.593$$

					1/12
1/8			1/8		
	1/24				
5/24		1/24			
3/24				1/24	
3/24	1/24	1/24			

$$\mathcal{I}(X) = 1.658, \; \mathcal{I}(Y) = 1.328$$
$$\mathcal{I}(X,Y) = 2.236$$
$$\mathcal{I}(X \cap Y) = 0.750$$

Figure 2.21 shows some examples of Shannon entropy calculation. For each example, we indicate the probabilities $p_{i,j}$ and the values taken by $\mathcal{I}(X)$, $\mathcal{I}(Y)$, $\mathcal{I}(X,Y)$ and $\mathcal{I}(X \cap Y)$. The top/left panel corresponds to a diffuse system. The value of $\mathcal{I}(X,Y)$ is maximum, meaning that the system is extremely disordered. The top/right panel represents a highly ordered system in the bivariate case and a diffuse system in the univariate case. We have $\mathcal{I}(X \mid Y) = \mathcal{I}(Y \mid X) = 0$, implying that the knowledge of X is sufficient to find the state of Y. Generally, the system is not perfectly ordered or perfectly disordered. For instance, in the case of the system described in the bottom/left panel, the knowledge of X informs us about the state of Y. Indeed, if X is in the third state, then we

know that Y cannot be in the first or sixth state. Another example is provided in the bottom/right panel.

Remark 18 *If we apply the Shannon entropy to the transition matrix of a Markov chain, we set $X = \mathcal{R}(s)$ and $Y = \mathcal{R}(t)$ where $\mathcal{R}(t)$ is the state variable at the date t. We obtain:*

$$\mathcal{I}\left(\mathcal{R}\left(t\right) \mid \mathcal{R}\left(s\right)\right) = -\sum_{i=1}^{K} \pi_i^{\star} \sum_{j=1}^{K} p_{i,j}^{(t-s)} \ln p_{i,j}^{(t-s)}$$

where $p_{i,j} = \Pr\left\{\mathcal{R}\left(t+1\right) = j \mid \mathcal{R}\left(t\right) = i\right\}$, $\mathbb{K} = \{1, 2, \ldots, K\}$ is the state space of the Markov chain and $\pi^{\star}$ is the associated stationary distribution.

Application to scoring   Let $\mathcal{S}$ and Y be the score and the control variable. For instance, Y is a binary random variable that may indicate a bad ESG risk ($Y = 0$) or a good ESG risk ($Y = 1$). Y may also correspond to classes defined by some quantiles. With Shannon entropy, we can measure the information of the system $(\mathcal{S}, Y)$. We can also compare two scores $\mathcal{S}_1$ and $\mathcal{S}_2$ by using the statistical measures $\mathcal{I}(\mathcal{S}_1 \cap Y)$ and $\mathcal{I}(\mathcal{S}_2 \cap Y)$. Let $\mathcal{S}_3$ be the aggregated score obtained from the two individual scores $\mathcal{S}_1$ and $\mathcal{S}_2$. We can calculate the information contribution of each score with respect to the global score. Therefore, we can verify that a score really adds an information.

We consider the following decision rule:

$$\begin{cases} \mathcal{S} \leq 0 \Rightarrow \mathcal{S}^{\star} = 0 \\ \mathcal{S} > 0 \Rightarrow \mathcal{S}^{\star} = 1 \end{cases}$$

We note $n_{i,j}$ the number of observations such that $\mathcal{S}^{\star} = i$ and $Y = j$. We obtain the following system $(\mathcal{S}^{\star}, Y)$:

	$Y = 0$	$Y = 1$
$\mathcal{S}^{\star} = 0$	$n_{0,0}$	$n_{0,1}$
$\mathcal{S}^{\star} = 1$	$n_{1,0}$	$n_{1,1}$

where $n = n_{0,0} + n_{0,1} + n_{1,0} + n_{1,1}$ is the total number of observations. The hit rate is the ratio of good bets:

$$H = \frac{n_{0,0} + n_{1,1}}{n}$$

This statistic can be viewed as an information measure of the system $(\mathcal{S}, Y)$. When there are more states, we can consider the Shannon entropy. In Figure 2.22, we report the contingency table of two scores $\mathcal{S}_1$ and $\mathcal{S}_2$ for 100 observations[50]. We have $\mathcal{I}(\mathcal{S}_1 \cap Y) = 0.763$ and $\mathcal{I}(\mathcal{S}_2 \cap Y) = 0.636$. We deduce that $\mathcal{S}_1$ is more informative than $\mathcal{S}_2$.

Graphical methods

We assume that the control variable Y can takes two values: $Y = 0$ corresponds to a bad risk (or bad signal) while $Y = 1$ corresponds to a good risk (or good signal). Gouriéroux (1992) introduced three graphical tools for assessing the quality of a score: the performance curve, the selection curve and the discrimination curve[51]. In the following, we assume that the probability $\Pr\{Y = 1 \mid \mathcal{S} \geq s\}$ is increasing with respect to the level[52] $s \in [0, 1]$, which corresponds to the rate of acceptance. We deduce that the decision rule is the following:

---

[50]Each score is divided into 6 intervals $(s_1, \ldots, s_6)$ while the dependent variable is divided into 5 intervals $(y_1, \ldots, y_5)$.

[51]See also Gouriéroux and Jasiak (2007).

[52]We assume that the score is based on the 0/1 normalization, but this assumption is not important since we can always map a general score into a 0/1 score.

Figure 2.22: Scorecards $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$ (100 observations)

	y_1	y_2	y_3	y_4	y_5
s_1	10	9			
s_2	7	9			
s_3	3		7	2	
s_4		2	10	4	5
s_5				10	2
s_6			3	4	13

	y_1	y_2	y_3	y_4	y_5
s_1	7	10			
s_2	10	8			
s_3			5	4	3
s_4	3		10	6	4
s_5	2			5	8
s_6			5	5	5

$$\mathcal{I}\left(\boldsymbol{S}_1\right) = 1.767 \qquad\qquad \mathcal{I}\left(\boldsymbol{S}_2\right) = 1.771$$
$$\mathcal{I}\left(Y\right) = 1.609 \qquad\qquad \mathcal{I}\left(Y\right) = 1.609$$
$$\mathcal{I}\left(\boldsymbol{S}_1, Y\right) = 2.614 \qquad\qquad \mathcal{I}\left(\boldsymbol{S}_2, Y\right) = 2.745$$
$$\mathcal{I}\left(\boldsymbol{S}_1 \cap Y\right) = 0.763 \qquad\qquad \mathcal{I}\left(\boldsymbol{S}_2 \cap Y\right) = 0.636$$

- if the score of the observation is above the threshold s, the observation is selected;

- if the score of the observation is below the threshold s, the observation is not selected.

If s is equal to one, we select no observation. If s is equal to zero, we select all the observations. In a scoring system, the threshold s is given. Below, we assume that s is varying and we analyze the relevance of the score with respect to this parameter.

Performance curve, selection curve and discriminant curve   The performance curve is the parametric function $y = \mathcal{P}\left(x\right)$ defined by:

$$\left\{ \begin{array}{l} x\left(s\right) = \Pr\left\{\boldsymbol{S} \geq s\right\} \\ y\left(s\right) = \dfrac{\Pr\left\{Y = 0 \mid \boldsymbol{S} \geq s\right\}}{\Pr\left\{Y = 0\right\}} \end{array} \right.$$

where $x\left(s\right)$ corresponds to the proportion of selected observations and $y\left(s\right)$ corresponds to the ratio between the proportion of selected bad risks and the proportion of bad risks in the population. The score is efficient if the ratio is below one. If $y\left(s\right) > 1$, the score selects more bad risks than those we can find in the population[53]. If $y\left(s\right) = 1$, the score is random and the performance is equal to zero. In this case, the selected population is representative of the total population.

The selection curve is the parametric curve $y = \mathcal{S}\left(x\right)$ defined by:

$$\left\{ \begin{array}{l} x\left(s\right) = \Pr\left\{\boldsymbol{S} \geq s\right\} \\ y\left(s\right) = \Pr\left\{\boldsymbol{S} \geq s \mid Y = 0\right\} \end{array} \right.$$

where $y\left(s\right)$ corresponds to the ratio of observations that are wrongly selected. By construction, we would like that the curve $y = \mathcal{S}\left(x\right)$ is located below the bisecting line $y = x$ in order to verify that $\Pr\left\{\boldsymbol{S} \geq s \mid Y = 0\right\} < \Pr\left\{\boldsymbol{S} \geq s\right\}$.

---

[53]In this case, we have $\Pr\left\{Y = 0 \mid \boldsymbol{S} \geq s\right\} > \Pr\left\{Y = 0\right\}$.

Remark 19 *The performance and selection curves are related as follows[54]:*

$$\mathcal{S}(x) = x\mathcal{P}(x)$$

The discriminant curve is the parametric curve $y = \mathcal{D}(x)$ defined by:

$$\mathcal{D}(x) = g_1\left(g_0^{-1}(x)\right)$$

where:

$$g_y(s) = \Pr\{\boldsymbol{\mathcal{S}} \geq s \mid Y = y\}$$

It represents the proportion of good risks in the selected population with respect to the proportion of bad risks in the selected population. The score is said to be discriminant if the curve $y = \mathcal{D}(x)$ is located above the bisecting line $y = x$.

Some properties We first notice that the previous parametric curves do not depend on the probability distribution of the score $\boldsymbol{\mathcal{S}}$, but only on the ranking of the observations. They are then invariant if we apply an increasing function to the score. Gouriéroux (1992) also established the following properties:

1. the performance curve (respectively, the selection curve) is located below the line $y = 1$ (respectively, the bisecting line $y = x$) if and only if $\mathrm{cov}\,(f(Y), g(\boldsymbol{\mathcal{S}})) \geq 0$ for any increasing functions f and g;

2. the performance curve is increasing if and only if:

$$\mathrm{cov}\,(f(Y), g(\boldsymbol{\mathcal{S}}) \mid \boldsymbol{\mathcal{S}} \geq s) \geq 0$$

for any increasing functions f and g, and any threshold level s;

3. the selection curve is convex if and only if $\mathbb{E}\,[f(Y) \mid \boldsymbol{\mathcal{S}} = s]$ is increasing with respect to the threshold level s for any increasing function f.

Remark 20 *The first property is the least restrictive. It allows us to verify that the score $\boldsymbol{\mathcal{S}}$ is better than a random score. We can show that $(3) \Rightarrow (2) \Rightarrow (1)$. The last property is then the most restrictive.*

A score is perfect or optimal if there is a threshold level $s^\star$ such that $\Pr\{Y = 1 \mid \boldsymbol{\mathcal{S}} \geq s^\star\} = 1$ and $\Pr\{Y = 0 \mid \boldsymbol{\mathcal{S}} < s^\star\} = 1$. It separates the population between good and bad risks. Graphically, the selection curve of a perfect score is equal to:

$$y = \mathbb{1}\{x > \Pr\{Y = 1\}\} \cdot \left(1 + \frac{x - 1}{\Pr\{Y = 0\}}\right)$$

---

[54]We have:

$$
\begin{aligned}
\Pr\{\boldsymbol{\mathcal{S}} \geq s \mid Y = 0\} &= \frac{\Pr\{\boldsymbol{\mathcal{S}} \geq s, Y = 0\}}{\Pr\{Y = 0\}} \\
&= \Pr\{\boldsymbol{\mathcal{S}} \geq s\} \cdot \frac{\Pr\{\boldsymbol{\mathcal{S}} \geq s, Y = 0\}}{\Pr\{\boldsymbol{\mathcal{S}} \geq s\}\Pr\{Y = 0\}} \\
&= \Pr\{\boldsymbol{\mathcal{S}} \geq s\} \cdot \frac{\Pr\{Y = 0 \mid \boldsymbol{\mathcal{S}} \geq s\}}{\Pr\{Y = 0\}}
\end{aligned}
$$

Using the relationship $\mathcal{S}(x) = x\mathcal{P}(x)$, we deduce that the performance curve of a perfect score is given by:

$$y = \mathbb{1}\left\{x > \Pr\{Y = 1\}\right\} \cdot \left(\frac{x - \Pr\{Y = 1\}}{x \cdot \Pr\{Y = 0\}}\right)$$

For the discriminant curve, a perfect score satisfies $\mathcal{D}(x) = 1$. When the score is random, we have $\mathcal{S}(x) = \mathcal{D}(x) = x$ and $\mathcal{P}(x) = 1$. In Figure 2.23, we have reported the performance, selection and discriminant curves of a given score S. We also show the curves obtained with an optimal (or perfect) score and a random score. A score must be located in the area between the curve computed with a random score and the curve computed with a perfect score, except if the score ranks the observations in a worst way than a random score.

   Gouriéroux (1992) also established two properties for comparing two scores $\boldsymbol{\mathcal{S}}_1$ and $\boldsymbol{\mathcal{S}}_2$:

- the score $\boldsymbol{\mathcal{S}}_1$ is more performing on the population P_1 than the score S_2 on the population P_2 if and only if the performance (or selection) curve of $(\boldsymbol{\mathcal{S}}_1, P_1)$ is below the performance (or selection) curve of $(\boldsymbol{\mathcal{S}}_2, P_2)$;

- the score $\boldsymbol{\mathcal{S}}_1$ is more discriminatory on the population P_1 than the score $\boldsymbol{\mathcal{S}}_2$ on the population P_2 if and only if the discriminant curve of $(\boldsymbol{\mathcal{S}}_1, P_1)$ is above the discriminant curve of $(\boldsymbol{\mathcal{S}}_2, P_2)$.

Figure 2.24 illustrates the case where the score $\boldsymbol{\mathcal{S}}_1$ is better than the score $\boldsymbol{\mathcal{S}}_2$. However, the order is only partial. Most of the time, the two scores cannot be globally compared. An example is provided in Figure 2.25. The second score is not very good to distinguish good and bad risks when it takes small values, but it is close to a perfect score when it takes high values.
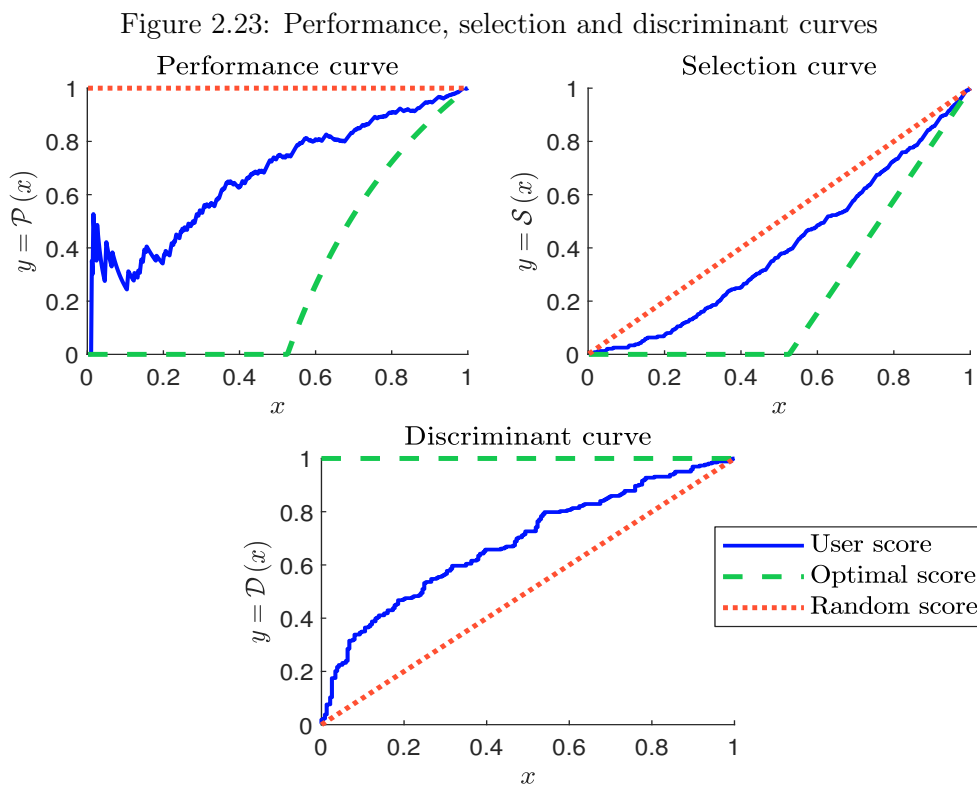
Figure 2.23: Performance, selection and discriminant curves

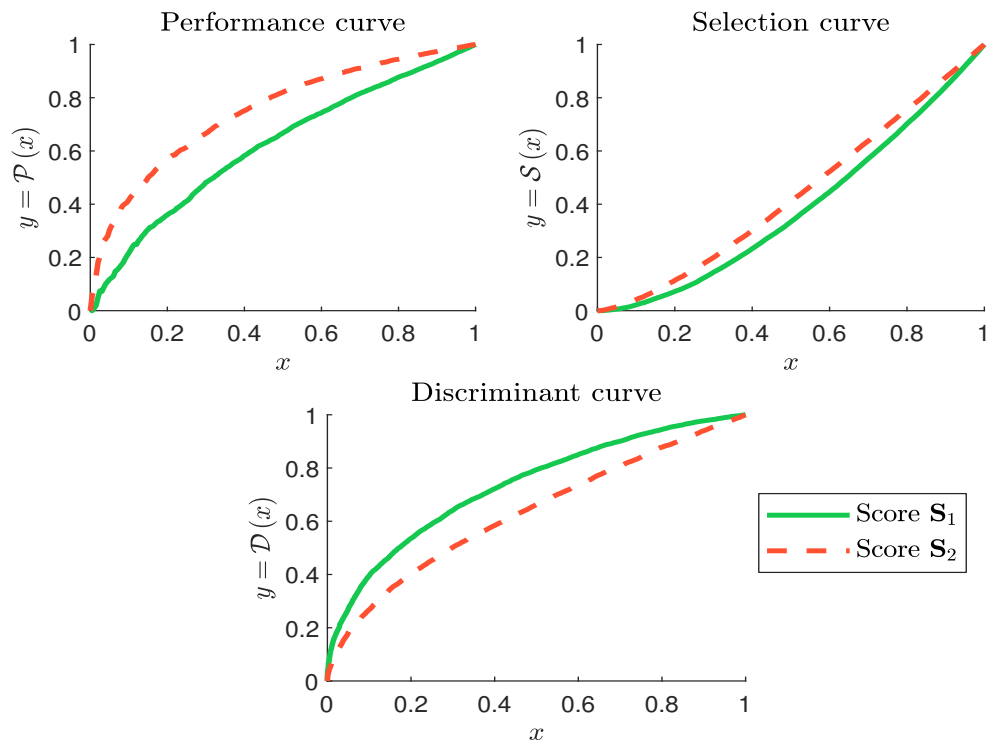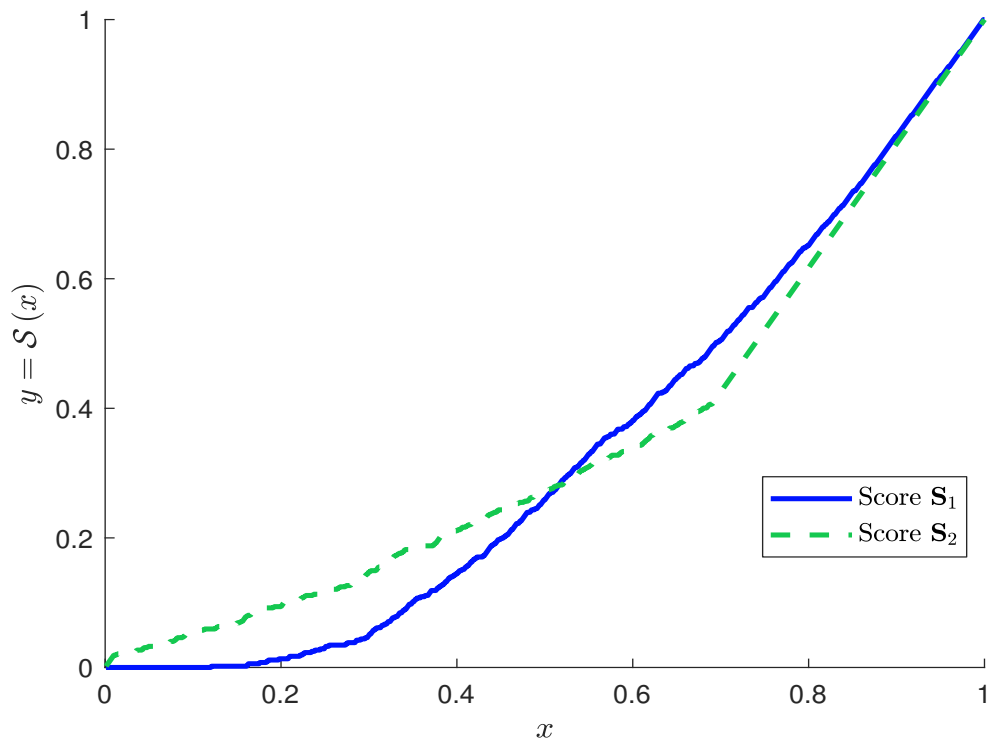Figure 2.24: The score S_1 is better than the score S_2

Figure 2.25: Illustration of the partial ordering between two scores

Statistical methods

Since the quantitative tools for comparing two scores are numerous, we focus on two non-parametric measures: the Kolmogorov-Smirnov test and the Gini coefficient.

Kolmogorov-Smirnov test   We consider the cumulative distribution functions:

$$\mathbf{F}_0\left(s\right) = \Pr\left\{\boldsymbol{\mathcal{S}} \leq s \mid Y = 0\right\}$$

and:

$$\mathbf{F}_1\left(s\right) = \Pr\left\{\boldsymbol{\mathcal{S}} \leq s \mid Y = 1\right\}$$

The score $\boldsymbol{\mathcal{S}}$ is relevant if we have the stochastic dominance order $\mathbf{F}_0 \succ \mathbf{F}_1$. In this case, the score quality is measured by the Kolmogorov-Smirnov statistic:

$$\mathrm{KS} = \max_s \left|\mathbf{F}_0\left(s\right) - \mathbf{F}_1\left(s\right)\right|$$

It takes the value 1 if the score is perfect. The KS statistic may be used to verify that the score is not random. We then test the assumption $\mathcal{H}_0 : \mathrm{KS} = 0$ by using the tabulated critical values[55]. In Figure 2.26, we give an example with 5 000 observations. The KS statistic is equal to 36%, which implies that $\mathcal{H}_0$ is rejected at the confidence level 1%.

Gini coefficient

   The Lorenz curve   The Gini coefficient is the statistic, which is the most used for measuring the performance of a score. It is related to the concept of Lorenz curve, which is a graphical representation of the concentration. Let X and Y be two random variables. The Lorenz curve $y = \mathcal{L}\left(x\right)$ is the parametric curve defined by:

$$\left\{ \begin{array}{l} x = \Pr\left\{X \leq x\right\} \\ y = \Pr\left\{Y \leq y \mid X \leq x\right\} \end{array} \right.$$

In economics, x represents the proportion of individuals that are ranked by income while y represents the proportion of income. In this case, the Lorenz curve is a graphical representation of the distribution of income and is used for illustrating inequality of the wealth distribution between individuals. For example, we observe that 70% of individuals have only 34% of total income in Figure 2.27.

   Definition of the Gini coefficient   The Lorenz curve has two limit cases. If the wealth is perfectly concentrated, one individual holds 100% of the total wealth. If the wealth is perfectly allocated between all the individuals, the corresponding Lorenz curve is the bisecting line. We define the Gini coefficient by:

$$\mathcal{G}ini\left(\mathcal{L}\right) = \frac{A}{A+B}$$

---

[55]The critical values at the 5% confidence level are equal to:

n	10	50	100	500	5000
CV	40.9%	18.8%	13.4%	6.0%	1.9%

More generally, the null hypothesis is rejected at the confidence level α if we have:

$$\mathrm{KS} > \sqrt{\frac{1}{2n} \ln\left(\frac{2}{\alpha}\right)}$$

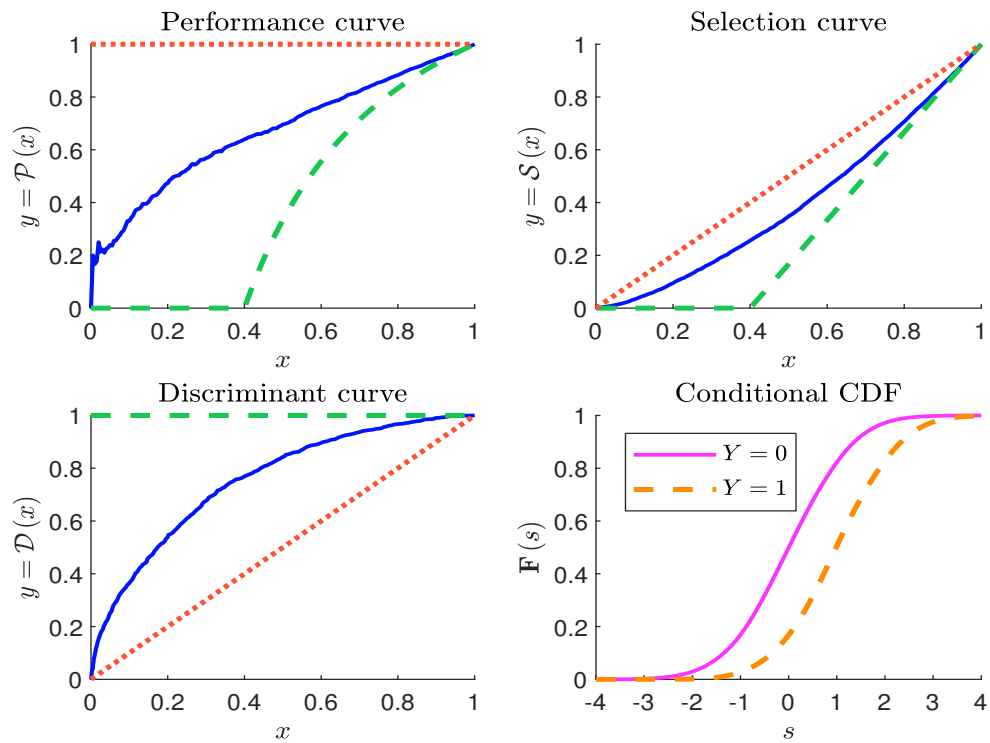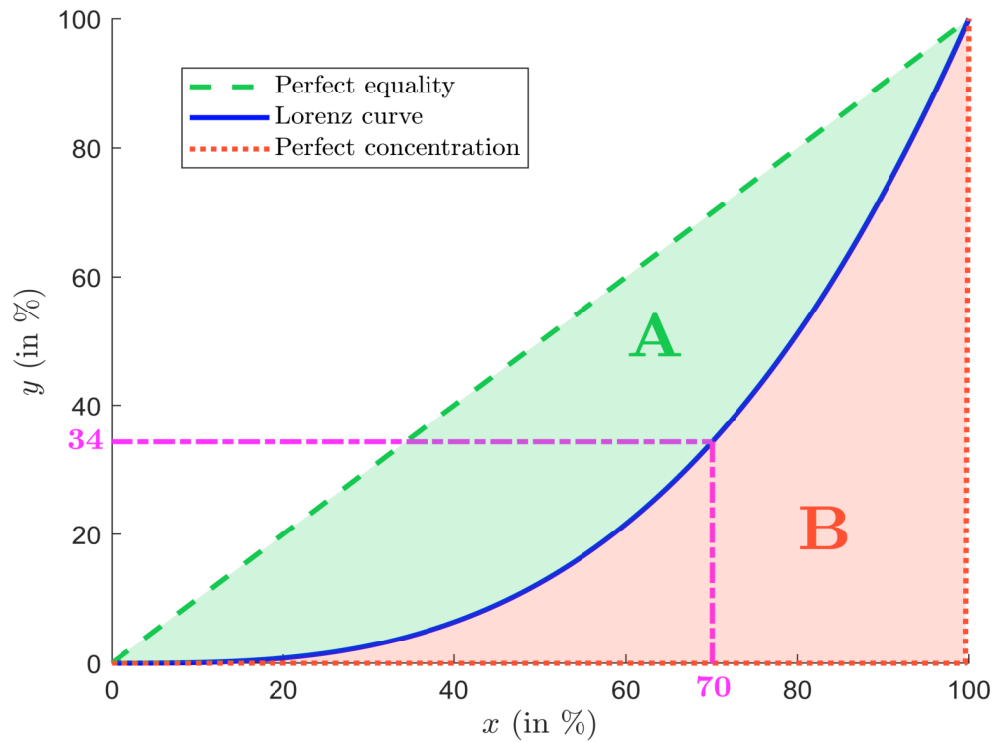Figure 2.26: Comparison of the distributions $\mathbf{F}_0(s)$ and $\mathbf{F}_1(s)$

Figure 2.27: An example of Lorenz curve

where A is the area between the Lorenz curve and the curve of perfect equality, and B is the area between the curve of perfect concentration and the Lorenz curve. By construction, we have $0 \leq \mathcal{G}ini\left(\mathcal{L}\right) \leq 1$. The Gini coefficient is equal to zero in the case of perfect equality and one in the case of perfect concentration. We have:

$$\mathcal{G}ini\left(\mathcal{L}\right) = 1 - 2\int_0^1 \mathcal{L}\left(x\right)\,\mathrm{d}x$$

Application to scoring   We can interpret the selection curve as a Lorenz curve. We recall that $\mathbf{F}\left(s\right) = \Pr\left\{\boldsymbol{\mathcal{S}} \leq s\right\}$, $\mathbf{F}_0\left(s\right) = \Pr\left\{\boldsymbol{\mathcal{S}} \leq s \mid Y = 0\right\}$ and $\mathbf{F}_1\left(s\right) = \Pr\left\{\boldsymbol{\mathcal{S}} \leq s \mid Y = 1\right\}$. The selection curve is defined by the following parametric coordinates:

$$\begin{cases} x\left(s\right) = 1 - \mathbf{F}\left(s\right) \\ y\left(s\right) = 1 - \mathbf{F}_0\left(s\right) \end{cases}$$

The selection curve measures the capacity of the score for not selecting bad risks. We could also build the Lorenz curve that measures the capacity of the score for selecting good risks:

$$\begin{cases} x\left(s\right) = \Pr\left\{\boldsymbol{\mathcal{S}} \geq s\right\} = 1 - \mathbf{F}\left(s\right) \\ y\left(s\right) = \Pr\left\{\boldsymbol{\mathcal{S}} \geq s \mid Y = 1\right\} = 1 - \mathbf{F}_1\left(s\right) \end{cases}$$

It is called the precision curve. Another popular graphical tool is the receiver operating characteristic (or ROC) curve (Powers, 2011), which is defined by:

$$\begin{cases} x\left(s\right) = \Pr\left\{\boldsymbol{\mathcal{S}} \geq s \mid Y = 0\right\} = 1 - \mathbf{F}_0\left(s\right) \\ y\left(s\right) = \Pr\left\{\boldsymbol{\mathcal{S}} \geq s \mid Y = 1\right\} = 1 - \mathbf{F}_1\left(s\right) \end{cases}$$

An example for a given score $\boldsymbol{\mathcal{S}}$ is provided in Figure 2.28. For all the three curves, we can calculate the Gini coefficient. Since the precision and ROC curves are located above the bisecting line, the Gini coefficient associated to the Lorenz curve $\mathcal{L}$ becomes[56]:

$$\mathcal{G}ini\left(\mathcal{L}\right) = 2\int_0^1 \mathcal{L}\left(x\right)\,\mathrm{d}x - 1$$

The Gini coefficient of the score $\boldsymbol{\mathcal{S}}$ is then computed as follows:

$$\mathcal{G}ini^\star\left(\boldsymbol{\mathcal{S}}\right) = \frac{\mathcal{G}ini\left(\mathcal{L}\right)}{\mathcal{G}ini\left(\mathcal{L}^\star\right)}$$

where $\mathcal{L}^\star$ is the Lorenz curve associated to the perfect score.

Remark 21 *The Gini coefficient is not necessarily the same for the three curves. However, if the population is homogeneous, we generally obtain very similar figures*[57].

[56]An alternative to the Gini coefficient is the AUC measure, which corresponds to the area under the ROC curve. However, they give the same information since they are related by the equation:

$$\mathcal{G}ini\left(\mathrm{ROC}\right) = 2 \times \mathrm{AUC}\left(\mathrm{ROC}\right) - 1$$

[57]For instance, we obtain the following results with the score $\boldsymbol{\mathcal{S}}$ that has been used in Figure 2.28:

Curve	$\mathcal{G}ini\left(\mathcal{L}\right)$	$\mathcal{G}ini\left(\mathcal{L}^\star\right)$	$\mathcal{G}ini^\star\left(S\right)$
Selection	20.41%	40.02%	51.01%
Precision	30.62%	59.98%	51.05%
ROC	51.03%	100.00%	51.03%

Figure 2.28: Selection, precision and ROC curves
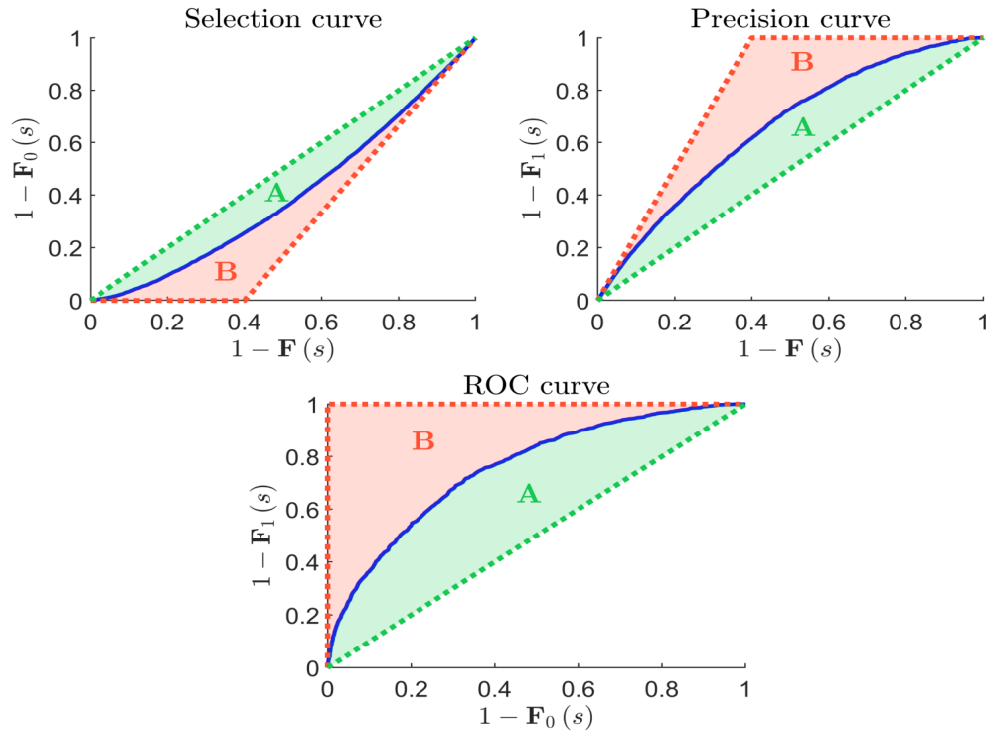
Choice of the optimal cut-off   The choice of the optimal cut-off $s^\star$ depends on the objective function. For instance, we can calibrate $s^\star$ in order to achieve a minimum universe size of ESG assets. We can also fix a given selection rate. From a statistical point of view, we must distinguish the construction of the scoring model and the decision rule. In statistical learning, we generally consider three datasets: the training set, the validation set and the test set. The training set is used for calibrating the model and its parameters whereas the validation set helps to avoid overfitting. But the decision rule is based on the test set.

Confusion matrix   A confusion matrix is a special case of contingency matrix. Each row of the matrix represents the frequency in a predicted class while each column represents the frequency in an actual class. Using the test set, it takes the following form:

	$Y = 0$	$Y = 1$
$\boldsymbol{\mathcal{S}} < s$	$n_{0,0}$	$n_{0,1}$
$\boldsymbol{\mathcal{S}} \geq s$	$n_{1,0}$	$n_{1,1}$
	$n_0 = n_{0,0} + n_{1,0}$	$n_1 = n_{0,1} + n_{1,1}$

where $n_{i,j}$ represents the number of observations of the cell (i, j). The interpretation of the confusion matrix is given in Table 2.16. The cells $(\boldsymbol{\mathcal{S}} < s, Y = 0)$ and $(\boldsymbol{\mathcal{S}} \geq s, Y = 1)$ correspond to observations that are well-classified: true negative (TN) and true positive (TP). The cells $(\boldsymbol{\mathcal{S}} \geq s, Y = 0)$ and $(\boldsymbol{\mathcal{S}} < s, Y = 1)$ correspond to two types of errors:

1. A false positive (FP) can induce a future loss, because the risk can materialize: this is a type I error;

2. A false negative (FN) potentially corresponds to a an opportunity cost: this is a type II error.

Table 2.16: Interpretation of the confusion matrix

	$Y = 0$	$Y = 1$
$\mathcal{S} < s$	It is rejected and it is a bad risk (true negative)	It is rejected, but it is a good risk (false negative)
$\mathcal{S} \geq s$	It is accepted, but it is a bad risk (false positive)	It is accepted and it is a good risk (true positive)
	(negative)	(positive)

Classification ratios   Binary classification defines many metrics for measuring the performance of the classifier[58] (Fawcett, 2006):

$$\text{True Positive Rate} \quad \text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{False Negative Rate} \quad \text{FNR} = \frac{\text{FN}}{\text{FN} + \text{TP}} = 1 - \text{TPR}$$

$$\text{True Negative Rate} \quad \text{TNR} = \frac{\text{TN}}{\text{TN} + \text{FP}}$$

$$\text{False Positive Rate} \quad \text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}} = 1 - \text{TNR}$$

The true positive rate (TPR) is also known as the sensitivity or the recall. It measures the proportion of real good risks that are correctly predicted good risk. Fawcett (2006) also defines the precision or the positive predictive value (PPV):

$$\text{PPV} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

It measures the proportion of predicted good risks that are correctly real good risk. Besides these metrics, statisticians also use three generic metrics:

1. The accuracy considers the classification of both negatives and positives:

$$\text{ACC} = \frac{\text{TP} + \text{TN}}{\text{P} + \text{N}} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FN} + \text{TN} + \text{FP}}$$

2. The F_1 score is the harmonic mean of precision and sensitivity:

$$\begin{aligned} F_1 &= \frac{2}{1/\text{precision} + 1/\text{sensitivity}} \\ &= \frac{2 \cdot \text{PPV} \cdot \text{TPR}}{\text{PPV} + \text{TPR}} \end{aligned}$$

---

[58]We rewrite the confusion matrix as follows:

	$Y = 0$	$Y = 1$
$\mathcal{S} < s$	TN	FN
$\mathcal{S} \geq s$	FP	TP
	N = TN + FP	P = FN + TP

3. The ϕ coefficient or the Matthews correlation coefficient (MCC) is a measure of association between $\mathcal{S}$ and Y:

$$\phi = \text{MCC} = \frac{\text{TP} \cdot \text{TN} - \text{FP} \cdot \text{FN}}{\sqrt{(\text{TP} + \text{FP}) \cdot (\text{TP} + \text{FN}) \cdot (\text{TN} + \text{FP}) \cdot (\text{TN} + \text{FN})}}$$

$\mathcal{S}$ and Y are positively associated if most of the observations fall along the diagonal cells.

Table 2.17: Confusion matrix of three scoring systems and three cut-off values s

Score	$s = 100$		$s = 200$		$s = 500$	
$\mathcal{S}_1$	386	616	698	1 304	1 330	3 672
	1 614	7 384	1 302	6 696	670	4 328
$\mathcal{S}_2$	372	632	700	1 304	1 386	3 616
	1 628	7 368	1 300	6 696	614	4 384
$\mathcal{S}_3$	382	616	656	1 344	1 378	3 624
	1 618	7 384	1 344	6 656	622	4 376
Perfect	1 000	0	2 000	0	2 000	3 000
	1 000	8 000	0	8 000	0	5 000

Example 7 *We consider three scoring systems that have been calibrated on a training set. These systems produce a score between 0 and 1 000. A low value predicts a bad risk while a high value predicts a good risk. In order to calibrate the cut-off, we consider a test set, which is composed of 10 000 new observations. In Table 2.17, we report the confusion matrix of each scoring system for different cut-off values (100, 200 and 500).*

Table 2.18: Binary classification ratios (in %) of the three scoring systems

Score	s	TPR	FNR	TNR	FPR	PPV	ACC	F_1
	100	92.3	7.7	19.3	80.7	82.1	77.7	86.9
$\mathcal{S}_1$	200	83.7	16.3	34.9	65.1	83.7	73.9	83.7
	500	54.1	45.9	66.5	33.5	86.6	56.6	66.6
	100	92.1	7.9	18.6	81.4	81.9	77.4	86.7
$\mathcal{S}_2$	200	83.7	16.3	35.0	65.0	83.7	74.0	83.7
	500	54.8	45.2	69.3	30.7	87.7	57.7	67.5
	100	92.3	7.7	19.1	80.9	82.0	77.7	86.9
$\mathcal{S}_3$	200	83.2	16.8	32.8	67.2	83.2	73.1	83.2
	500	54.7	45.3	68.9	31.1	87.6	57.5	67.3
	100	100.0	0.0	50.0	50.0	88.9	90.0	94.1
Perfect	200	100.0	0.0	100.0	0.0	100.0	100.0	100.0
	500	62.5	37.5	100.0	0.0	100.0	70.0	76.9
Best	100	$\mathcal{S}_1/\mathcal{S}_3$	$\mathcal{S}_1/\mathcal{S}_3$	$\mathcal{S}_1$	$\mathcal{S}_1$	$\mathcal{S}_1$	$\mathcal{S}_1$	$\mathcal{S}_1$
scoring	200	$\mathcal{S}_1/\mathcal{S}_2$	$\mathcal{S}_1/\mathcal{S}_2$	$\mathcal{S}_2$	$\mathcal{S}_2$	$\mathcal{S}_2$	$\mathcal{S}_2$	$\mathcal{S}_2$
system	500	$\mathcal{S}_2$	$\mathcal{S}_2$	$\mathcal{S}_2$	$\mathcal{S}_2$	$\mathcal{S}_2$	$\mathcal{S}_2$	$\mathcal{S}_2$

Using confusion matrices given in Table 2.17, we calculate the different classification ratios and report them in Table 2.18. In addition to the three scoring systems, we have also considered a

perfect score in order to show what the best value is for each classification ratio, and we indicate the best scoring system. We notice that it depends on the ratio and on the value of the cut-off. For instance, if we want to maximize the true positive ratio or minimize the false negative ratio, $\mathcal{S}_1$ is the best scoring system for low value of s while $\mathcal{S}_2$ is better when s is equal to 500. For the other ratios, $\mathcal{S}_1$ seems to be the best scoring system when $s = 100$, otherwise $\mathcal{S}_2$ dominates $\mathcal{S}_1$ and $\mathcal{S}_3$ when $s = 200$ or $s = 500$.

Remark 22 *We recall that* $\mathbf{F}_0(s) = \Pr\{\mathcal{S} \leq s \mid Y = 0\}$ *and* $\mathbf{F}_1(s) = \Pr\{\mathcal{S} \leq s \mid Y = 1\}$. *We deduce that* $\mathrm{TNR} = \mathbf{F}_0(s)$, $\mathrm{FNR} = \mathbf{F}_1(s)$, $\mathrm{FPR} = 1 - \mathbf{F}_0(s)$ *and* $\mathrm{TPR} = 1 - \mathbf{F}_1(s)$. *Therefore, the ROC curve is the parametric curve, where the x-coordinates are the false positive rates and the y-coordinates are the true positive rates. Generally, we note* α *and* β *the type I and II errors. We may also interpret the ROC curve as the relationship of* $1 - \beta(s)$ *with respect to* $\alpha(s)$.

Backtesting of unsupervised scoring systems

To understand how to implement a backtesting procedure, we take the example of credit scoring models. Let $\mathcal{S}_i(t)$ be the credit score of individual/company i at time t. The response variable is the default indicator variable $Y_i(t)$:

$$Y_i(t) = \mathbb{1}\{\tau_i \geq t + \delta\} = \mathbb{1}\{D_i(t + \delta) = 0\}$$

where τ_i and D_i are the default time and the default indicator function, and δ is the time horizon (*e.g.*, one year). In this problem, $\mathcal{S}_i(t)$ is used to predict the default time τ_i and the decision process is:

$$\begin{cases} \mathcal{S}_i(t) \leq s^\star \Rightarrow i \text{ is a bad risk} \Rightarrow Y_i(t) = 0 \\ \mathcal{S}_i(t) > s^\star \Rightarrow i \text{ is not a bad risk} \Rightarrow Y_i(t) = 1 \end{cases}$$

The calibration problem for the credit scoring model is therefore:

$$\Pr\{Y_i(t) = 0\} = f(\mathcal{S}_i(t))$$

where f is an increasing function. We obtain a binary choice model that can be calibrated using discriminant analysis, clustering methods or logit/probit regression. The backtesting procedure consists of verifying that $\hat{Y}_i(t) = \mathbb{1}\{\mathcal{S}_i(t) > s^\star\}$ is a good estimator of $Y_i(t)$. Building a backtesting procedure for ESG scores then requires defining an ad hoc response variable.

Static analysis   Let $\mathcal{S}_i(t)$ be the ESG score of company i at time t. The endogenous response variable can be defined as follows:

Scoring system	Risk class	$Y_i(t)$
Best-in-class oriented	Good risk	$\mathbb{1}\{\mathcal{S}_i(t + \delta) \geq s^\star\}$
Worst-in-class oriented	Bad risk	$\mathbb{1}\{\mathcal{S}_i(t + \delta) \leq s^\star\}$

where $s^\star$ is the best-in-class/worst-in-class threshold to determine. $Y_i(t)$ is endogenous because it depends on the future value of the score. Here, the backtesting procedure can be seen as a stability test of the ESG scoring system. An alternative is to use an exogenous response variable based on controversies. For example, to predict bad risks, we can use the binary response $Y_i(t) = \mathbb{1}\{\mathcal{C}_i(t + \delta) \geq 0\}$ where $\mathcal{C}_i(t)$ is the controversy index.

Dynamic analysis   The static analysis can be extended to the dynamic case. Instead of using the ESG score, we consider the past momentum $\mathcal{M}_i(t, h) = \mathcal{S}_i(t) - \mathcal{S}_i(t - h)$ where h is typically the year, while the response variable is based on the future momentum $\mathcal{S}_i(t + \delta) - \mathcal{S}_i(t)$.

Illustration using an ESG scoring system   We consider the scoring system of an ESG rating agency. We normalize the published scores in order to obtain z-scores and apply the backtesting procedure to these latter numbers. We consider four risk classes:

Risk class	Definition	$Y_i(t)$
Worst-in-class	$\boldsymbol{\mathcal{S}}_i(t) \leq \hat{\mathbf{F}}^{-1}(20\%)$	$\mathbb{1}\{\boldsymbol{\mathcal{S}}_i(t+\delta) \leq s^\star\}$
Bad risk	$\boldsymbol{\mathcal{S}}_i(t) \leq \bar{\boldsymbol{\mathcal{S}}}$	$\mathbb{1}\{\boldsymbol{\mathcal{S}}_i(t+\delta) \leq s^\star\}$
Good risk	$\boldsymbol{\mathcal{S}}_i(t) \geq \bar{\boldsymbol{\mathcal{S}}}$	$\mathbb{1}\{\boldsymbol{\mathcal{S}}_i(t+\delta) \geq s^\star\}$
Best-in-class	$\boldsymbol{\mathcal{S}}_i(t) \geq \hat{\mathbf{F}}^{-1}(80\%)$	$\mathbb{1}\{\boldsymbol{\mathcal{S}}_i(t+\delta) \geq s^\star\}$

where $\hat{\mathbf{F}}$ is the empirical distribution of the score and $\bar{\boldsymbol{\mathcal{S}}}$ is the average of scores. For each risk class, we compute the classification ratios ACC, F_1, and ϕ and the Shannon entropy $\mathcal{I}(\boldsymbol{\mathcal{S}} \cap Y)$ with respect to the cut-off value s. The results, expressed in %, are shown in Figures 2.29 to 2.32 when we consider the MSCI World and MSCI EM universes.

Table 2.19: Optimal cut-off $s^\star$ (MSCI World)

Risk class	$\delta = 3$ months				$\delta = 12$ months			
	ACC	F_1	ϕ	$\mathcal{I}(\boldsymbol{\mathcal{S}} \cap Y)$	ACC	F_1	ϕ	$\mathcal{I}(\boldsymbol{\mathcal{S}} \cap Y)$
Worst-in-class	−0.91	−0.61	−0.68	−0.58	−0.96	−0.58	−0.67	−0.54
Bad risk	−0.01	0.18	0.02	0.02	0.01	0.24	0.04	0.05
Good risk	−0.02	−0.18	0.01	0.02	−0.01	−0.20	0.03	0.04
Best-in-class	1.05	0.79	0.85	0.77	1.08	0.76	0.83	0.72

Table 2.20: Optimal cut-off $s^\star$ (MSCI EM)

Risk class	$\delta = 3$ months				$\delta = 12$ months			
	ACC	F_1	ϕ	$\mathcal{I}(\boldsymbol{\mathcal{S}} \cap Y)$	ACC	F_1	ϕ	$\mathcal{I}(\boldsymbol{\mathcal{S}} \cap Y)$
Worst-in-class	−1.87	−1.19	−1.29	−1.15	−2.00	−1.17	−1.30	−1.12
Bad risk	0.13	0.23	−0.03	−0.10	0.16	0.28	−0.05	−0.14
Good risk	0.13	−0.15	−0.03	−0.14	0.16	−0.22	−0.05	−0.24
Best-in-class	0.48	0.14	0.22	0.11	0.56	0.13	0.24	0.11

We can then estimate the optimal cut-off $s^\star$, which is the value s that maximizes the backtesting metric. Results are given in Tables 2.19 and 2.20. Theoretically, the optimal cut-off is $s^\star = \Phi^{-1}(20\%) = -0.8416$ for the worst-in-class category, $s^\star = \mathbb{E}[\mathcal{N}(0,1)] = 0$ for the bad-risk and good-risk categories and $s^\star = \Phi^{-1}(80\%) = 0.8416$ for the best-in-class category, because the backtesting procedure concerns z-scores. For the MSCI World universe, the estimated cut-offs are not that far from the theoretical cut-offs. Moreover, we do not observe a large difference between the three-month and the twelve-month horizons. For the MSCI World universe, we face two problems. First, the backtesting of the worst-in-class category implies an optimal cut-off that is well below the theoretical cut-off, meaning that the worst-in-class universe is not adequately defined. Second, the scoring system is not really able to discriminate between bad risk and worst-in-class categories because the optimal cut-off for the latter category is close to zero.

Remark 23 *The previous example illustrates that the robustness of an ESG scoring system can vary depending on the investment universe.*

Figure 2.29: Backtesting of ESG scores (worst-in-class & bad risk, MSCI World)
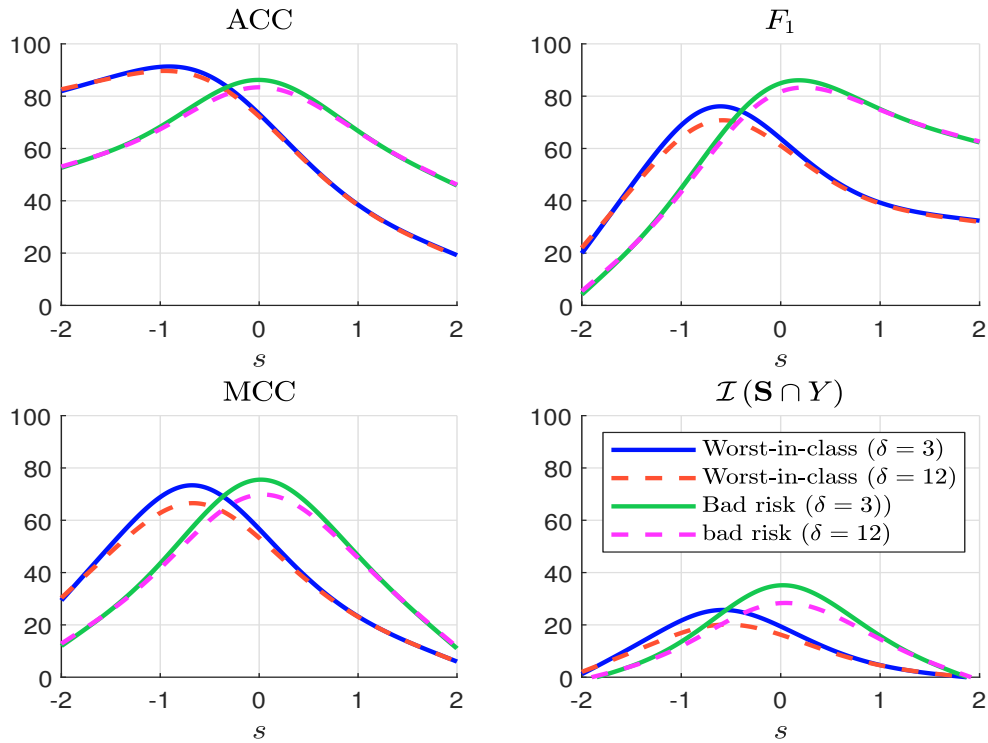

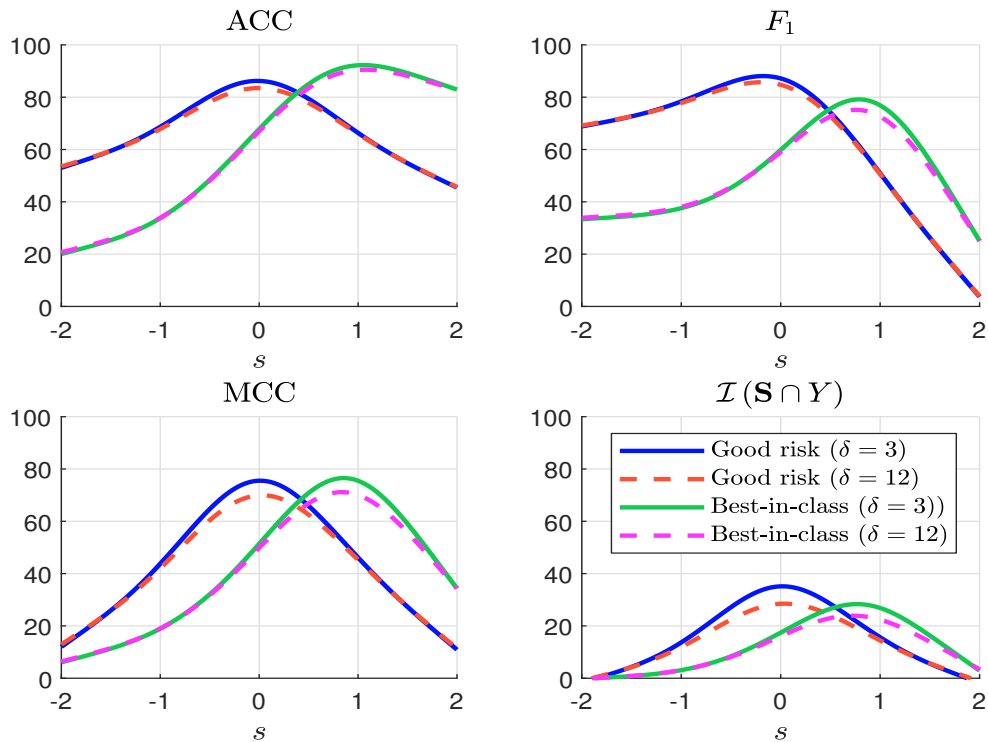Figure 2.30: Backtesting of ESG scores (best-in-class & good risk, MSCI World)

Figure 2.31: Backtesting of ESG scores (worst-in-class & bad risk, MSCI EM)
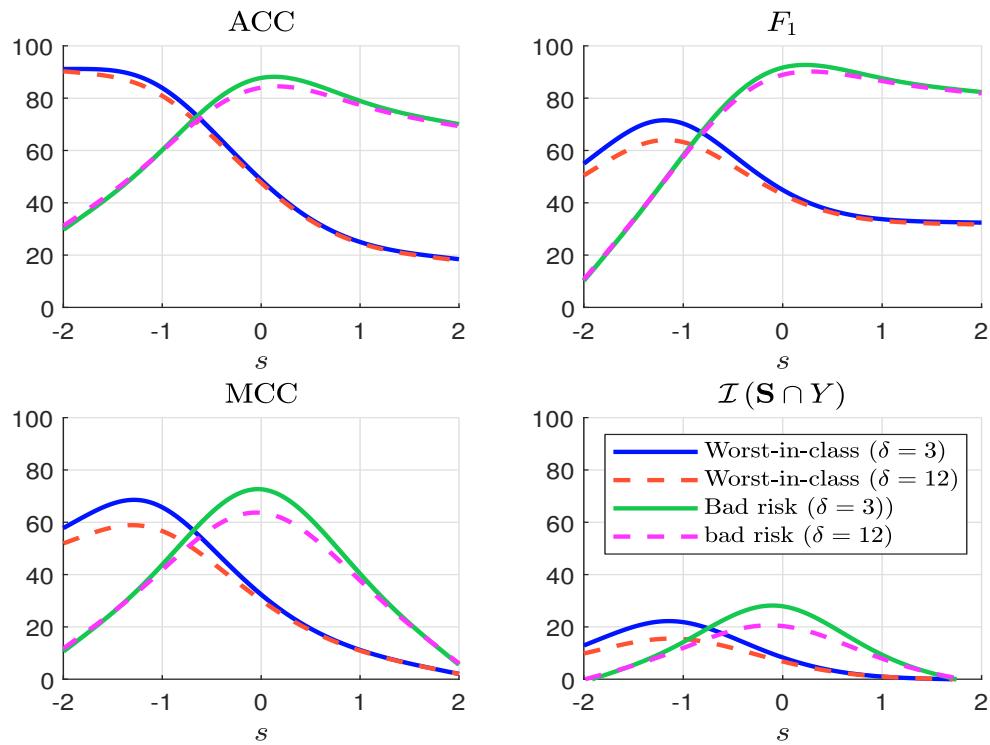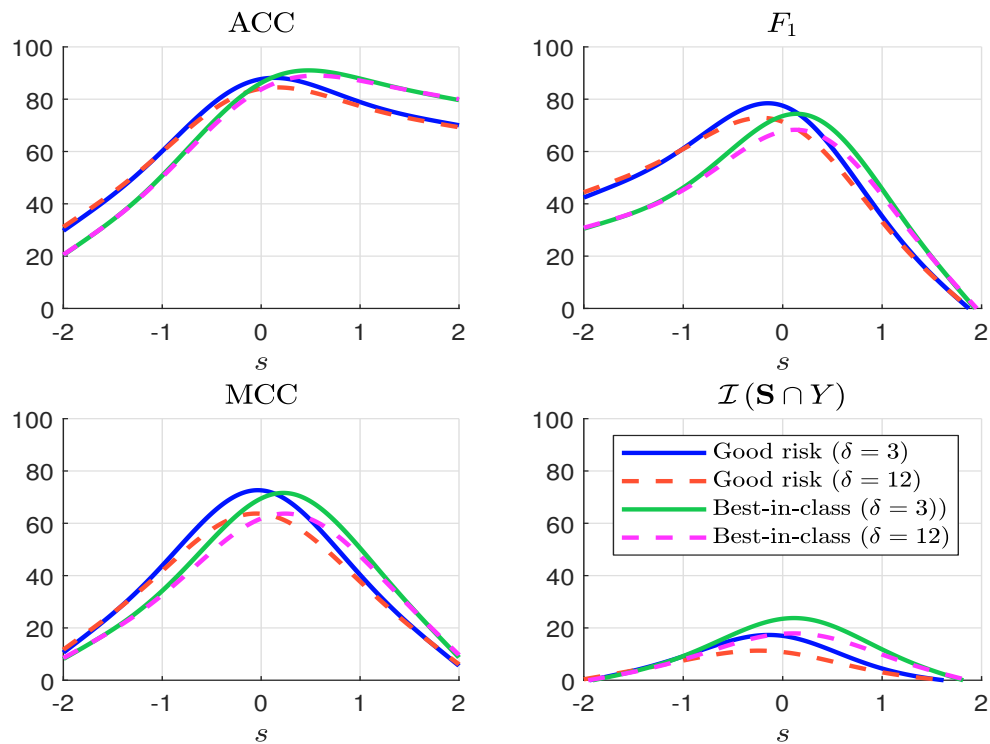

Figure 2.32: Backtesting of ESG scores (best-in-class & good risk, MSCI EM)

2.3   Rating system

As we have seen, a scoring model provides an automatic and statistical score. It is a pure quantitative approach. There may be the intervention of an analyst, but it is limited to data quality checks or forcing of input data[59]. A rating (or a notation) is different from a score, because it implies a quality scale. Since it implies a value judgement, a rating is generally produced by an analyst. For example, this is the case of credit ratings, which are made by an analyst who takes into account several quantitative scores, qualitative data, private and meeting information. Nevertheless, the balance between quantitative and qualitative judgements depends on the type of issuers. For retail borrowers, the rating is mainly explained by the scoring model. For blue chip and mega-cap companies, the rating highly depends on the qualitative assessment of the credit risk. If we consider ESG risk, the rating process shares similar patterns. The ESG score is generally the key component of the ESG rating. It is validated by an extra-financial analyst and it may be *forced* based on his qualitative information and experience. This explains why extra-financial analysis is organized with respect to sectors. An ESG analyst, who is specialized in a given sector, can then have a better view of all the ratings produced for this sector. This is particularly true for the strategic sectors: automobiles, coal, ciment, oil & gas, fertilizers & agricultural chemicals, utilities, etc. Nevertheless, there is a strong difference between credit and ESG rating processes from an investor viewpoint. Indeed, the number of rated companies for ESG analysis bears no comparison with the number of rated companies for credit analysis, because only few corporates have access to bond markets. On the contrary, a comprehensive ESG rating system must encompass all the securities and assets, notably all listed corporates and some private companies. In this context, the impact and the implication of the extra-financial analyst decrease with the firm size. This is why there is a strong small size bias in ESG rating systems.

2.3.1   Definition

ESG rating systems are based on the terminology of credit ratings (Box 2.4). For example, MSCI (2022) uses a 7-grade rating scale based on the grades AAA, AA, A, BBB, BB, B and CCC (Table 2.21). The number of grades, the rating symbols (letter, numeric) and the ordering of the system (low/worst rating to high/best rating vs. high/best to low/worst rating) differs from one provider to another provider[60]:

- Amundi: A (high), B,... to G (low) — 7-grade scale

- FTSE Russell: 0 (low), 1,... to 5 (high) — 6-grade scale

- ISS ESG: 1 (high), 2,... to 10 (low) — 10-grade scale

- MSCI: AAA (high), AA,... to CCC (low) — 7-grade scale

- Refinitiv: A+ (high), A, A-, B+,... to D- (low) — 12-grade scale

- RepRisk: AAA (high), AA,... to D (low) — 8-grade scale

- Sustainanalytics: 1 (low), 2,... to 5 (high) — 5-grade scale

We notice the high heterogeneity of rating scales. Nevertheless, we observe that they are less granular than those used by credit rating agencies. On average, an ESG rating system is made up of 7 grades vs. 20 grades for a credit rating system.

---

[59]In some cases, the analyst may also validate the score.

[60]In this list, we have included the asset manager Amundi, because ESG ratings are not only built by ESG rating agencies. Some investors (asset owners and managers) have defined their own internal ESG ratings.

Box 2.4: Terminology of credit ratings

A rating system is a symbolic or numeric classification according to grade, which indicates a degree or step in a scale. For example, the credit rating systems of S&P, Moody's and Fitch is reported in Table 2.B. They are all based on a rating scale of 20 grades[a]. The symbolic rank AA+ (or BBB) is then a grade or a rating in the S&P classification. A notch means the difference between a particular rating and the next lower. For example, in the case of Moody's, the difference between Baa1 and Baa2 constitutes one Notch, whereas the difference between Aaa and Aa2 corresponds to two notches. When a credit rating agency revises the credit risk of a company, it may upgrade its rating by one notch (+1 notch), two notches (+2 notches), etc. or it may downgrade its rating by one notch (−1 notch), two notches (−2 notches), etc.

Table 2.B: Credit rating system of S&P, Moody's and Fitch

	Prime Maximum Safety	High Grade High Quality			Upper Medium Grade			Lower Medium Grade		
S&P/Fitch	AAA	AA+	AA	AA−	A+	A	A−	BBB+	BBB	BBB−
Moody's	Aaa	Aa1	Aa2	Aa3	A1	A2	A3	Baa1	Baa2	Baa3

	Non Investment Grade Speculative			Highly Speculative			Substantial Risk	In Poor Standing		Extremely Speculative
S&P/Fitch	BB+	BB	BB−	B+	B	B−	CCC+	CCC	CCC−	CC
Moody's	Ba1	Ba2	Ba3	B1	B2	B3	Caa1	Caa2	Caa3	Ca

[a]Or 21 grades if we include the issuer default. Nevertheless, D is not considered as a rating.

Table 2.21: ESG rating system of Moody's

Leader		Average			Laggard	
AAA	AA	A	BBB	BB	B	CCC

Source: MSCI (2022, Exhibit 8, page 12).

Remark 24 *In the sequel, we use the 7-grade scale based on the ratings* AAA, AA, A, BBB, BB, B *and* CCC. *We think that it is easier to manipulate and understand from a pedagogical point of view. A company rated* AAA *is a good company (or a good ESG risk) and a company rated* CCC *is a bad company (or a bad ESG risk).*

2.3.2 ESG rating process

The construction of ESG ratings follows the same process than credit ratings (Figure 2.33). We need to define the map function that converts an ESG score into an ESG rating. In the case of credit risk, the estimate of the one-year probability of default is converted into credit ratings[61]. In the case of ESG risk, the ESG score is converted into an ESG rating such that the best scores correspond to the best ratings and the worst scores correspond to the worst ratings.

---

[61]For instance, a CCC-rated company has a one-year probability of default of 25%; a B-rated company has a 5% probability to default in the next year; for a BB-rated company, the one-year probability of default is equal to 1%; etc.

Figure 2.33: From ESG score to ESG rating
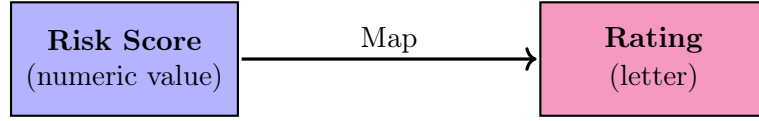
The first step consists in specifying the map function:

$$\mathcal{M}ap: \begin{array}{ccc} \Omega_{\boldsymbol{\mathcal{S}}} & \longrightarrow & \Omega_{\mathcal{R}} \\ \boldsymbol{\mathcal{S}} & \longmapsto & \mathcal{R} = \mathcal{M}ap\left(\boldsymbol{\mathcal{S}}\right) \end{array}$$

where $\Omega_{\boldsymbol{\mathcal{S}}}$ is the support of ESG scores, $\Omega_{\mathcal{R}}$ is the ordered state space of ESG ratings and $\mathcal{R}$ is the ESG rating. By construction, $\mathcal{M}ap$ is a monotone function in order to preserve the preference ordering. In the case where $\mathcal{M}ap$ is increasing, we verify that:

$$\boldsymbol{\mathcal{S}}_2 > \boldsymbol{\mathcal{S}}_1 \Leftrightarrow \mathcal{M}ap\left(\boldsymbol{\mathcal{S}}_2\right) \succ \mathcal{M}ap\left(\boldsymbol{\mathcal{S}}_1\right)$$

The second step is the validation (and the possible *forcing*) of the rating by the analyst.

Let us see some examples. MSCI (2022, page 12) explains that they use they use a uniform map function where $\Omega_{\boldsymbol{\mathcal{S}}} = [0, 10]$ and $\Omega_{\mathcal{R}} = \{\mathsf{CCC}, \mathsf{B}, \mathsf{BB}, \mathsf{BBB}, \mathsf{A}, \mathsf{AA}, \mathsf{AAA}\}$. The score is then divided into 7 equally-sized intervals and we have:

$$\mathcal{M}ap\left(s\right) = \begin{cases} \mathsf{CCC} & \text{if } \boldsymbol{\mathcal{S}} \in [0, {}^{10}/_7] & (0 - 1.429) \\ \mathsf{B} & \text{if } \boldsymbol{\mathcal{S}} \in [{}^{10}/_7, {}^{20}/_7] & (1.429 - 2.857) \\ \mathsf{BB} & \text{if } \boldsymbol{\mathcal{S}} \in [{}^{20}/_7, {}^{30}/_7] & (2.857 - 4.286) \\ \mathsf{BBB} & \text{if } \boldsymbol{\mathcal{S}} \in [{}^{30}/_7, {}^{40}/_7] & (4.286 - 5.714) \\ \mathsf{A} & \text{if } \boldsymbol{\mathcal{S}} \in [{}^{40}/_7, {}^{50}/_7] & (5.714 - 7.143) \\ \mathsf{AA} & \text{if } \boldsymbol{\mathcal{S}} \in [{}^{50}/_7, {}^{60}/_7] & (7.143 - 8.571) \\ \mathsf{AAA} & \text{if } \boldsymbol{\mathcal{S}} \in [{}^{60}/_7, 10] & (8.571 - 10) \end{cases}$$

For instance, if the ESG score of the company is equal to 5, we assign the grade BBB, a score of 8 corresponds to the grade AA, etc. Refinitiv (2022, page 7) also considers a uniform map function, implying that $\Omega_{\boldsymbol{\mathcal{S}}}$ is divided by 12 equally-sized intervals:

> "[...] 'D' score (D-, D and D+) *indicates poor relative ESG performance and insufficient degree of transparency in reporting material ESG data publicly. 'C' score (C-, C and C+) indicates satisfactory relative ESG performance and moderate degree of transparency in reporting material ESG data publicly. 'B' score (B-, B and B+) indicates good relative ESG performance and above average degree of transparency in reporting material ESG data publicly. 'A' score (A-, A and A+) indicates excellent relative ESG performance and high degree of transparency in reporting material ESG data publicly."*

We assume that the map function is an increasing piecewise function, $\boldsymbol{\mathcal{S}} \sim \mathbf{F}$ and $\boldsymbol{\mathcal{S}} \in (s^-, s^+)$. We note $s_1^\star, \ldots, s_{K-1}^\star$ the knots of the piecewise function, K the number of ratings and $\Omega_{\mathcal{R}} = \{R_1, \ldots, R_K\}$ the set of grades. We set $s_0^\star = s^-$ and $s_K^\star = s^+$. We deduce that:

$$\begin{aligned} p_k &= \Pr\left\{\mathcal{R} = R_k\right\} \\ &= \Pr\left\{s_{k-1}^\star \leq \boldsymbol{\mathcal{S}} < s_k^\star\right\} \\ &= \mathbf{F}\left(s_k^\star\right) - \mathbf{F}\left(s_{k-1}^\star\right) \end{aligned}$$

Using this equation, we can then compute the frequency distribution of the ratings. The set of frequencies $\{p_1, \ldots, p_K\}$ is denoted by $\mathbb{P}$. If we don't know the distribution $\mathbf{F}$, we consider the empirical distribution $\hat{\mathbf{F}}$ and the estimated frequency is equal to $\hat{p}_k = \hat{\mathbf{F}}\left(s_k^\star\right) - \hat{\mathbf{F}}\left(s_{k-1}^\star\right)$. If we would like to build a rating system with pre-defined frequencies $(p_1, \ldots, p_K)$, we have to solve the following equation:

$$\mathbf{F}\left(s_k^\star\right) - \mathbf{F}\left(s_{k-1}^\star\right) = p_k$$

We deduce that:

$$
\begin{aligned}
\mathbf{F}\left(s_k^\star\right) &= p_k + \mathbf{F}\left(s_{k-1}^\star\right) \\
&= p_k + p_{k-1} + \mathbf{F}\left(s_{k-2}^\star\right) \\
&= \left(\sum_{j=1}^{k} p_j\right) + \mathbf{F}\left(s_0^\star\right)
\end{aligned}
$$

Since $\mathbf{F}\left(s_0^\star\right) = 0$, we conclude that:

$$s_k^\star = \mathbf{F}^{-1}\left(\sum_{j=1}^{k} p_j\right)$$

Remark 25 *The discrete probability space of the rating system is denoted by* $(\Omega_\mathcal{R}, \Omega_\mathcal{R}, \mathbb{P})$ *and we have:*

$$\mathcal{E} := \Omega_\mathcal{R} \times \mathbb{P} = \{(R_1, p_1), \ldots, (R_K, p_K)\}$$

Let us consider a uniform score $\mathcal{S} \sim \mathcal{U}_{[a,b]}$. We have $\mathbf{F}(s) = (s-a)/(b-a)$. The rating system consists in K equally-sized intervals. The knots of the map function are then equal to:

$$s_k^\star = a + \frac{(b-a)}{K}k$$

It follows that:

$$
\begin{aligned}
p_k &= \mathbf{F}\left(a + \frac{(b-a)}{K}k\right) - \mathbf{F}\left(a + \frac{(b-a)}{K}(k-1)\right) \\
&= \frac{a + \dfrac{(b-a)}{K}k - a}{b-a} - \frac{a + \dfrac{(b-a)}{K}(k-1) - a}{b-a} \\
&= \frac{1}{K}
\end{aligned}
$$

We obtain a trivial result: the rating frequencies are all equal. In the case where we impose pre-defined frequencies $(p_1, \ldots, p_K)$, the knots of the map function are equal to[62]:

$$s_k^\star = a + (b-a)\left(\sum_{j=1}^{k} p_j\right)$$

If we consider a $0/100$ uniform score, we deduce that $s_k^\star = 100 \cdot \sum_{j=1}^{k} p_j$. For example, if $\Omega_\mathcal{R} \times \mathbb{P} = \{(\mathsf{CCC}, 5\%), (\mathsf{B}, 10\%), (\mathsf{BB}, 20\%), (\mathsf{BBB}, 30\%), (\mathsf{A}, 20\%), (\mathsf{AA}, 10\%), (\mathsf{AAA}, 5\%)\}$, we obtain the trivial piecewise function where[63] $s_\mathsf{CCC}^\star = 5$, $s_\mathsf{B}^\star = 15$, $s_\mathsf{BB}^\star = 35$, $s_\mathsf{BBB}^\star = 65$, $s_\mathsf{A}^\star = 85$ and $s_\mathsf{AA}^\star = 95$.

Figure 2.34: Map function of a z-score (equal-space ratings)
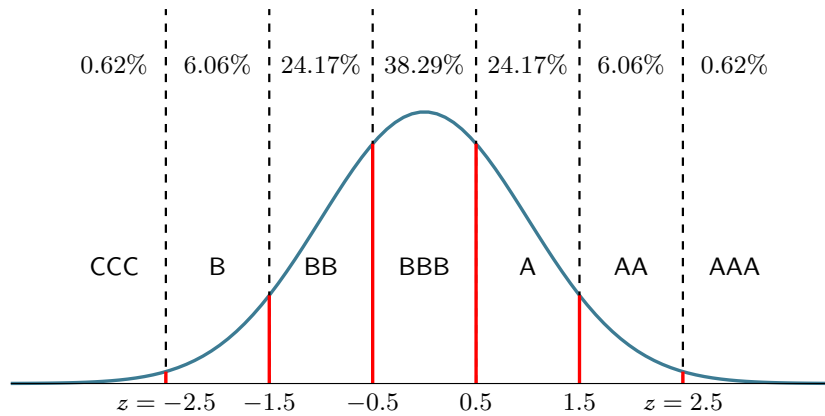
For a z-score system, we assume that $\boldsymbol{\mathcal{S}} \sim \mathcal{N}(0,1)$ and we obtain:

$$p_k = \Phi\left(s_k^\star\right) - \Phi\left(s_{k-1}^\star\right)$$

If we consider the 7-grade rating system with the classical knots $(-2.5, -1.5, -0.5, 0.5, 1.5, 2.5)$, we obtain the map function given in Figure 2.34 where[64]:

$$p_k = \Phi\left(-3.5 + k\right) - \Phi\left(-4.5 + k\right)$$

The rating system with equal frequencies is obtained by using the following knots:

$$s_k^\star = \Phi^{-1}\left(\frac{k}{K}\right) \qquad \text{for } k = 1, \dots, K$$

In the case $K = 7$, the map function is given in Figure 2.35.

Figure 2.35: Map function of a z-score (equal-frequency ratings)
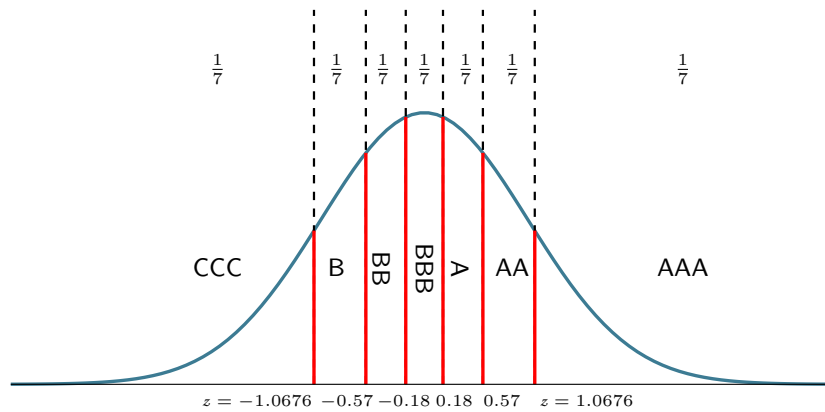
Remark 26 *We recall that the role of the ESG analyst is to verify the consistency of the rated companies. This is why we generally observe forced ratings (or scores).*

---

[62]We have $\mathbf{F}(s) = u \Leftrightarrow (s - a)/(b - a) = u$. We deduce that $s = \mathbf{F}^{-1}(u) = a + (b - a)u$.

[63]By construction, the knot $s_{\mathsf{AAA}}^\star = 100$ is not necessarily to be defined because we always have $s_{\mathsf{AAA}}^\star = s^+$.

[64]We obtain the following results: $p_{\mathsf{CCC}} = \Phi(-2.5) = 0.62\%$, $p_{\mathsf{B}} = \Phi(-1.5) - \Phi(-2.5) = 6.06\%$, $p_{\mathsf{BB}} = \Phi(-0.5) - \Phi(-1.5) = 24.17\%$, $p_{\mathsf{BBB}} = \Phi(0.5) - \Phi(-0.5) = 38.29\%$, $p_{\mathsf{A}} = \Phi(1.5) - \Phi(0.5) = 24.17\%$, $p_{\mathsf{AA}} = \Phi(2.5) - \Phi(1.5) = 6.06\%$ and $p_{\mathsf{AAA}} = 1 - \Phi(2.5) = 0.62\%$.

Box 2.5: Asymmetric rating system

In a symmetric rating system, the probability of the kth rating class is equal to the probability of the $(K - k + 1)^{\text{th}}$ rating class[a], e.g., $p_{\text{CCC}} = p_{\text{AAA}}$, $p_{\text{B}} = p_{\text{A}}$ and $p_{\text{BB}} = p_{\text{A}}$. To obtain an asymmetric rating system, the first approach is to define the frequencies p_k such that $\exists k : p_k \neq p_{K-k+1}$. We note $P_{worst} = \sum_{k=1}^{\lfloor K/2 \rfloor} p_k$ and $P_{best} = \sum_{\lceil K/2+1 \rceil}^{K} p_k$ the probability to be below and average the median rating. The rating process is said to be a *losing-oriented* system if $P_{worst} \geq P_{best}$, otherwise it is a *winning-oriented* system. This means that companies with bad ESG risk are more prevalent than companies with good ESG risk. For instance, $\Omega_{\mathcal{R}} \times \mathbb{P} = \{(\text{CCC}, 5\%), (\text{B}, 10\%), (\text{BB}, 25\%), (\text{BBB}, 40\%), (\text{A}, 15\%), (\text{AA}, 4\%), (\text{AAA}, 1\%)\}$ is a *losing-oriented* system. The choice of an asymmetric rating system may be motivated by the underlying ESG strategy. For instance, implementing an exclusion ESG policy is not equivalent to considering a selection ESG policy. The investor may then want to adapt his rating system to take into account the objective of the strategy. The second approach to obtain an asymmetric rating system is to consider a *b*-score normalization with $\alpha \neq \beta$. Some examples are provided in Table 2.C.

Table 2.C: Frequency distribution of ESG ratings (in %)

Parameters		Rating						
α	β	CCC	B	BB	BBB	A	AA	AAA
1	1	14.3	14.3	14.3	14.3	14.3	14.3	14.3
2	2	5.5	14.3	19.5	21.3	19.5	14.3	5.5
3	3	2.3	12.1	22.3	26.4	22.3	12.1	2.3
0.25	0.25	33.9	7.5	5.9	5.5	5.9	7.5	33.9
2.5	1.5	1.5	6.4	12.4	18.1	22.3	23.2	16.0
1.5	2.5	16.0	23.2	22.3	18.1	12.4	6.4	1.5
0.75	1	23.2	15.8	13.9	12.8	12.0	11.4	10.9

[a]We reiterate that we only consider rating systems that satisfy a comprehensive preference ordering: $\forall k : \mathcal{R}_k \succ \mathcal{R}_{k-1}$.

2.3.3 Rating migration matrix

One important issue concerns the consistency of the rating system. In particular, we may wonder whether it is relevant to use an equal-frequency, an equal-space or an asymmetric rating scheme. In the case of credit rating systems, we generally observe that medium risk classes have a higher frequency than extreme (low/high) risk classes. For instance, there are more BBB-rated companies than CCC-rated companies, the less frequent class is by far the AAA rating (less than 1%), etc. In the case of ESG rating systems, there is no consensus. Therefore, to assess the consistency and robustness of the rating system, we need to use probabilistic tools based on transition probability matrices (Norris, 1997).

Discrete-time modeling

Markov chain model We consider a time-homogeneous Markov chain $\mathcal{R}$, whose transition matrix is $P = (p_{i,j})$. We note $\Omega_{\mathcal{R}} = \{R_1, \ldots, R_K\}$ the state space of the chain and $\mathbb{K} = \{1, \ldots, K\}$ the corresponding index set. $p_{i,j}$ is the probability that the entity migrates from rating R_i to rating R_j.

The matrix P satisfies the following properties:

- $\forall i, j \in \mathbb{K}$, $p_{i,j} \geq 0$;

- $\forall i \in \mathbb{K}$, $\sum_{j=1}^{K} p_{i,j} = 1$.

Let $\mathcal{R}(t)$ be the value of the state at time t. We define $p(s, i; t, j)$ as the probability that the entity reaches the state R_j at time t given that it has reached the state R_i at time s. We have:

$$p(s, i; t, j) = \Pr\{\mathcal{R}(t) = R_j \mid \mathcal{R}(s) = R_i\} = p_{i,j}^{(t-s)}$$

This probability only depends on the duration between s and t because of the Markov property. Therefore, we can restrict the analysis by calculating the n-step transition probability:

$$p_{i,j}^{(n)} = \Pr\{\mathcal{R}(t + n) = R_j \mid \mathcal{R}(t) = R_i\}$$

and the associated n-step transition matrix $P^{(n)} = \left(p_{i,j}^{(n)}\right)$. For $n = 2$, we obtain:

$$
\begin{aligned}
p_{i,j}^{(2)} &= \Pr\{\mathcal{R}(t + 2) = R_j \mid \mathcal{R}(t) = R_i\} \\
&= \sum_{k=1}^{K} \Pr\{\mathcal{R}(t + 2) = R_j, \mathcal{R}(t + 1) = R_k \mid \mathcal{R}(t) = R_i\} \\
&= \sum_{k=1}^{K} \Pr\{\mathcal{R}(t + 2) = R_j \mid \mathcal{R}(t + 1) = R_k\} \cdot \Pr\{\mathcal{R}(t + 1) = R_k \mid \mathcal{R}(t) = R_i\} \\
&= \sum_{k=1}^{K} p_{i,k} \cdot p_{k,j}
\end{aligned}
$$

In a similar way, we obtain:

$$p_{i,j}^{(n+m)} = \sum_{k=1}^{K} p_{i,k}^{(n)} \cdot p_{k,j}^{(m)} \qquad \forall n, m > 0 \tag{2.4}$$

This equation is called the forward Chapman-Kolmogorov equation. In matrix form, we have:

$$P^{(n+m)} = P^{(n)} \cdot P^{(m)}$$

with the convention $P^{(0)} = I$. In particular, we have:

$$
\begin{aligned}
P^{(n)} &= P^{(n-1)} \cdot P^{(1)} \\
&= P^{(n-2)} \cdot P^{(1)} \cdot P^{(1)} \\
&= \prod_{t=1}^{n} P^{(1)} \\
&= P^n
\end{aligned}
$$

We deduce that:

$$p(t, i; t + n, j) = p_{i,j}^{(n)} = \mathbf{e}_i^\top P^n \mathbf{e}_j \tag{2.5}$$

When we apply this framework to ESG risk, $\mathcal{R}(t)$ denotes the rating (or the risk class) of the company at time t and $p_{i,j}$ is the one-period transition probability from rating R_i to rating R_j.

Table 2.22: ESG migration matrix #1 (one-year transition probability in %)

	AAA	AA	A	BBB	BB	B	CCC
AAA	92.76	5.66	0.90	0.45	0.23	0.00	0.00
AA	4.15	82.73	11.86	0.89	0.30	0.07	0.00
A	0.18	15.47	72.98	10.46	0.82	0.09	0.00
BBB	0.07	1.32	19.60	69.49	9.03	0.42	0.07
BB	0.04	0.19	1.55	19.36	70.88	7.75	0.23
B	0.00	0.05	0.24	1.43	21.54	74.36	2.38
CCC	0.00	0.00	0.22	0.44	2.21	13.24	83.89

In Table 2.22, we report an example of transition probability matrix. We read the figures as follows[65]: a company rated AAA has a one-year probability of 92.76% to remain AAA; its probability to become AA is 5.66%; a company rated CCC has a probability of 16.11% to improve its rating, etc. In Tables 2.23 and 2.24, we have reported the two-year and five-year transition matrices. We detail below the calculation of $p_{\mathsf{AAA,AAA}}^{(2)}$:

$$
\begin{aligned}
p_{\mathsf{AAA,AAA}}^{(2)} &= p_{\mathsf{AAA,AAA}} \times p_{\mathsf{AAA,AAA}} + p_{\mathsf{AAA,AA}} \times p_{\mathsf{AA,AAA}} + p_{\mathsf{AAA,A}} \times p_{\mathsf{A,AAA}} + p_{\mathsf{AAA,BBB}} \times p_{\mathsf{BBB,AAA}} + \\
&\quad p_{\mathsf{AAA,BB}} \times p_{\mathsf{BB,AAA}} + p_{\mathsf{AAA,B}} \times p_{\mathsf{B,AAA}} + p_{\mathsf{AAA,CCC}} \times p_{\mathsf{CCC,AAA}} \\
&= 0.9276^2 + 0.0566 \times 0.0415 + 0.0090 \times 0.0018 + 0.0045 \times 0.0007 + 0.0023 \times 0.0004 \\
&= 86.28\%
\end{aligned}
$$

Table 2.23: Two-year transition probability in % (migration matrix #1)

	AAA	AA	A	BBB	BB	B	CCC
AAA	86.28	10.08	2.25	0.92	0.44	0.02	0.00
AA	7.30	70.52	18.68	2.67	0.66	0.15	0.00
A	0.95	24.24	57.16	15.20	2.19	0.25	0.01
BBB	0.21	5.06	28.22	52.11	12.93	1.33	0.14
BB	0.09	0.79	6.07	27.45	53.68	11.37	0.55
B	0.01	0.18	0.98	6.26	31.47	57.28	3.82
CCC	0.00	0.05	0.50	1.32	6.31	21.13	70.70

Table 2.24: Five-year transition probability in % (migration matrix #1)

	AAA	AA	A	BBB	BB	B	CCC
AAA	70.45	18.69	6.97	2.61	1.08	0.18	0.01
AA	13.13	50.21	26.03	7.90	2.22	0.48	0.03
A	4.35	33.20	37.78	17.99	5.52	1.08	0.09
BBB	1.50	16.49	32.49	30.90	14.61	3.63	0.38
BB	0.50	5.98	17.83	30.10	31.35	12.85	1.39
B	0.15	1.90	7.40	18.95	35.11	31.26	5.23
CCC	0.05	0.64	2.55	6.93	17.96	38.54	43.33

---

[65]The rows represent the initial rating whereas the columns indicate the final rating.

<div style="text-align: center;">Box 2.6: Stationary distribution of a Markov chain</div>

We note $\pi_k^{(n)}$ the probability of the state R_k at time n:

$$\pi_k^{(n)} = \Pr\left\{\mathcal{R}(n) = R_k\right\}$$

and $\pi^{(n)} = \left(\pi_1^{(n)}, \ldots, \pi_K^{(n)}\right)$ the probability distribution. By construction, we have:

$$\pi^{(n+1)} = P^\top \pi^{(n)}$$

The Markov chain $\mathcal{R}$ admits a stationary distribution $\pi^\star$ if[a]:

$$\pi^\star = P^\top \pi^\star$$

In this case, $\pi_k^\star$ is the limiting probability of state R_k:

$$\lim_{n\to\infty} p_{i,k}^{(n)} = \pi_k^\star \qquad \forall i$$

We can interpret $\pi_k^\star$ as the average duration spent by the chain $\mathcal{R}$ in the state R_k. Let $\mathcal{T}_k$ be the return period[b] of state R_k:

$$\mathcal{T}_k = \inf\left\{n : \mathcal{R}(n) = R_k \mid \mathcal{R}(0) = R_k\right\}$$

The average return period is then equal to:

$$\boldsymbol{\tau}_k := \mathbb{E}\left[\mathcal{T}_k\right] = \frac{1}{\pi_k^\star}$$

---

[a]Not all Markov chains behave in this way, meaning that $\pi^\star$ does not necessarily exist.
[b]$\mathcal{T}_k$ is a stopping time. It is also called the *first-passage time*.

We compute the stationary distribution[66] and we obtain:

$$\pi^\star = (17.78\%, 29.59\%, 25.12\%, 15.20\%, 8.35\%, 3.29\%, 0.67\%)$$

The average return periods are then equal to 5.6, 3.4, 4.0, 6.6, 12.0, 30.4 and 149.0 years. The interpretation of these results is the following. In the long term, the probability to observe a AAA-rated company is equal to 17.78%, while the probability to observe a CCC-rated company is equal to 0.67%. The probability $\pi_k^\star$ is then the long-term equivalent of the current (or sample) frequency p_k. Similarly, the expected time to reach the worst-in-class state is equal to 149 years. These results show that the rating system #1 is clearly a winning-oriented system, where more than 70% of corporates are expected to have a rating above BBB.

---

[66]There are three numerical approaches to compute $\pi^\star$. The first one is to approximate $P^{(\infty)}$ by $P^{(n)}$ with n sufficient large ($n > 100$) and take any rows of the matrix $P^{(\infty)}$. The second method is to compute the eigendecomposition $V\Lambda V^{-1}$ of $P^\top$ and return the left eigenvector whose eigenvalue is exactly equal to 1. This second approach uses the fact that $\pi^\star = P^\top \pi^\star$ defines an eigenvalue problem $\left(P^\top - \lambda I_K\right)\pi^\star = \mathbf{0}$ with $\lambda = 1$. Finally, the third method directly solves the equation $\left(P^\top - I_K\right)\pi^\star = \mathbf{0}$ by computing an orthonormal basis for the null space of $P^\top - I_K$. For the last two methods, we normalize the solution such that $\mathbf{1}^\top \pi^\star = 1$.

Table 2.25: ESG migration matrix #2 (one-month transition probability in %)

	AAA	AA	A	BBB	BB	B	CCC
AAA	93.50	5.00	0.50	0.50	0.50	0.00	0.00
AA	2.00	93.00	4.00	0.50	0.50	0.00	0.00
A	0.00	3.00	93.00	3.90	0.10	0.00	0.00
BBB	0.00	0.10	2.80	94.00	3.00	0.10	0.00
BB	0.00	0.00	0.10	3.50	94.50	1.80	0.10
B	0.00	0.00	0.00	0.10	3.70	96.00	0.20
CCC	0.00	0.00	0.00	0.40	0.50	0.60	98.50

In Table 2.25, we now consider the ESG migration matrix #2, which has been computed on a monthly basis. If we would like to compare the two rating systems, we can compute the one-year probability transition matrix (Table 2.26). We observe that the two transition matrices are very different. Indeed, the second rating system is more reactive than the first rating system. If we compute the stationary distribution of the second rating system, we obtain:

$$\pi^\star = (3.11\%, 10.10\%, 17.46\%, 27.76\%, 25.50\%, 12.68\%, 3.39\%)$$

implying that the average return periods are equal to 32.2, 9.9, 5.7, 3.6, 3.9, 7.9 and 29.5 years. These results show that the rating system #2 is a balanced system.

Table 2.26: One-year probability transition in % (migration matrix #2)

	AAA	AA	A	BBB	BB	B	CCC
AAA	48.06	29.71	10.34	6.42	4.95	0.49	0.03
AA	11.65	49.25	24.10	9.60	4.87	0.49	0.03
A	2.02	17.51	49.67	24.72	5.52	0.54	0.03
BBB	0.27	3.53	17.46	55.50	20.21	2.88	0.16
BB	0.03	0.60	4.21	23.43	57.45	13.27	1.01
B	0.00	0.08	0.74	5.94	27.10	64.18	1.96
CCC	0.00	0.07	0.57	4.22	5.77	5.85	83.51

Table 2.27: One-month probability transition in % (migration matrix #1)

	AAA	AA	A	BBB	BB	B	CCC
AAA	99.36	0.53	0.05	0.04	0.02	0.00	0.00
AA	0.39	98.31	1.26	0.01	0.03	0.01	0.00
A	−0.02	1.65	97.14	1.21	0.02	0.01	0.00
BBB	0.01	−0.07	2.28	96.72	1.06	−0.01	0.01
BB	0.00	0.02	−0.12	2.29	96.92	0.88	0.01
B	0.00	0.00	0.04	−0.15	2.45	97.42	0.25
CCC	0.00	0.00	0.02	0.04	0.05	1.37	98.53

Remark 27 *Another approach to analyze the two rating systems is to compute the monthly transition matrix associated to the migration matrix #1. In this case, we have to find the matrix M such*

that $M^{(12)} = P$. The solution[67] *is given by $M = P^{1/12}$ and reported in Table 2.27. We can compare it with the matrix in Table 2.25. Because M has some negative probabilities, it is not a transition matrix, which indicates that the rating system #1 does not satisfy the Markov property*[68].

<div align="center">Box 2.7: Mean hitting time</div>

Let $\mathcal{A} \subset \mathbb{K}$ be a given subset. The first hitting time of $\mathcal{A}$ is given by:

$$\mathcal{T}(\mathcal{A}) = \inf \{n : \mathcal{R}(n) \in \mathcal{A}\}$$

$\mathcal{T}(\mathcal{A})$ measures how long it takes to reach the target states $j \in \mathcal{A}$. We can show that it is a stopping time. The mean first hitting (or passage) time to target $\mathcal{A}$ from state k is defined as:

$$\boldsymbol{\tau}_k(\mathcal{A}) = \mathbb{E}[\mathcal{T}(\mathcal{A}) \mid \mathcal{R}(0) = R_k]$$

Let $\boldsymbol{\tau}(\mathcal{A}) = (\boldsymbol{\tau}_1(\mathcal{A}), \dots, \boldsymbol{\tau}_K(\mathcal{A}))$ be the vector of mean first hitting times. Norris (1997) showed that:

$$\boldsymbol{\tau}_k(\mathcal{A}) = 1 + \sum_{j=1}^{K} p_{k,j} \boldsymbol{\tau}_j(\mathcal{A})$$

By construction, we have $\boldsymbol{\tau}_k(\mathcal{A}) = 0$ if $k \in \mathcal{A}$. In fact, $\boldsymbol{\tau}_k(\mathcal{A})$ is the minimal non negative solution to the previous system. It follows that $\|\boldsymbol{\tau}(\mathcal{A})\| = \sum_{k=1}^{K} |\boldsymbol{\tau}_k(\mathcal{A})| = \sum_{k=1}^{K} \boldsymbol{\tau}_k(\mathcal{A})$ because $\boldsymbol{\tau}_k(\mathcal{A}) \geq 0$. We deduce that:

$$\boldsymbol{\tau}(\mathcal{A}) \quad = \quad \arg\min \sum_{k=1}^{K} x_k$$
$$\text{s.t.} \quad \begin{cases} x_k = 0 & \text{if } k \in \mathcal{A} \\ x_k = 1 + \sum_{j=1}^{K} p_{k,j} x_j & \text{if } k \notin \mathcal{A} \\ x_k \geq 0 \end{cases}$$

We obtain a linear programming problem with $K + 1$ equally constraints and K lower bounds:

$$\boldsymbol{\tau}(\mathcal{A}) \quad = \quad \arg\min \sum_{k=1}^{K} x_k$$
$$\text{s.t.} \quad \begin{cases} x_k = 0 & \text{if } k \in \mathcal{A} \\ \sum_{j \notin \mathcal{A}} p_{k,j} x_j = -1 \\ \sum_{j \notin \mathcal{A} \cup \{k\}} p_{k,j} x_j + (p_{k,k} - 1) x_k = -1 & \text{if } k \notin \mathcal{A} \\ x_k \geq 0 \end{cases}$$

---

[67]Since $f(x) = x^{\alpha}$ with $\alpha > 1$ is a transcendental function, we use the Schur decomposition $P = QTQ^*$ to compute numerically the matrix M. Using Appendix A.1.1, we deduce that $M = QT^{1/12}Q^*$.

[68]The Markov property of ESG ratings is discussed later on page 121.

Let $\mathcal{B} = \{\mathsf{AAA}, \mathsf{AA}, \mathsf{A}\}$ and $\mathcal{W} = \{\mathsf{BB}, \mathsf{B}, \mathsf{CCC}\}$ be the best-in-class and worst-in-class sets. We obtain the following mean hitting times (in years) for the two rating systems:

Rating	$\mathcal{W}$-target				$\mathcal{B}$-target			
system	AAA	AA	A	BBB	BBB	BB	B	CCC
#1	79.21	70.04	62.34	46.54	7.50	13.28	17.58	22.68
#2	10.24	9.92	9.13	6.68	8.68	11.99	14.26	17.54

Estimation of the transition matrix Using Bayes theorem, we have:

$$
\begin{aligned}
p_{i,j} &= \Pr\{\mathcal{R}(t+1) = R_j \mid \mathcal{R}(t) = R_i\} \\
&= \frac{\Pr\{\mathcal{R}(t+1) = R_j, \mathcal{R}(t) = R_i\}}{\Pr\{\mathcal{R}(t) = R_i\}}
\end{aligned}
$$

We reiterate that[69] $\mathcal{R}(t) = R_k \Leftrightarrow \boldsymbol{S}(t) \in [s_{k-1}^{\star}, s_k^{\star}]$. We have seen that:

$$
\Pr\{\mathcal{R}(t) = R_k\} = \mathbf{F}(s_k^{\star}) - \mathbf{F}(s_{k-1}^{\star}) = p_k
$$

where $\mathbf{F}(s)$ is the probability distribution of the score $\boldsymbol{S}(t)$. We assume that[70]:

$$
\Pr\{\boldsymbol{S}(t) \leq s, \boldsymbol{S}(t+1) \leq s'\} = \mathbf{C}(\mathbf{F}(s), \mathbf{F}(s'))
$$

where $\mathbf{C}$ is the copula function of the random vector $(\boldsymbol{S}(t), \boldsymbol{S}(t+1))$. We deduce that:

$$
\begin{aligned}
\Pr\{\mathcal{R}(t+1) = R_j, \mathcal{R}(t) = R_i\} &= \Pr\{s_{i-1}^{\star} \leq \boldsymbol{S}(t) \leq s_i^{\star}, s_{j-1}^{\star} \leq \boldsymbol{S}(t+1) \leq s_j^{\star}\} \\
&= \mathbf{C}(\mathbf{F}(s_i^{\star}), \mathbf{F}(s_j^{\star})) - \mathbf{C}(\mathbf{F}(s_{i-1}^{\star}), \mathbf{F}(s_j^{\star})) - \\
&\quad \mathbf{C}(\mathbf{F}(s_i^{\star}), \mathbf{F}(s_{j-1}^{\star})) + \mathbf{C}(\mathbf{F}(s_{i-1}^{\star}), \mathbf{F}(s_{j-1}^{\star}))
\end{aligned}
$$

Finally, we obtain:

$$
p_{i,j} = \frac{\mathbf{C}\left(\mathbf{F}(s_i^{\star}), \mathbf{F}(s_j^{\star})\right) - \mathbf{C}\left(\mathbf{F}(s_{i-1}^{\star}), \mathbf{F}(s_j^{\star})\right) - \mathbf{C}\left(\mathbf{F}(s_i^{\star}), \mathbf{F}(s_{j-1}^{\star})\right) + \mathbf{C}\left(\mathbf{F}(s_{i-1}^{\star}), \mathbf{F}(s_{j-1}^{\star})\right)}{\mathbf{F}(s_i^{\star}) - \mathbf{F}(s_{i-1}^{\star})}
$$

This is the theoretical expression of the probability transition $p_{i,j}$. In practice, we do not know the probability functions $\mathbf{F}$ and $\mathbf{C}$. Therefore, we can estimate them and the estimated value of $p_{i,j}$ is equal to:

$$
\hat{p}_{i,j} = \frac{\hat{\mathbf{C}}\left(\hat{\mathbf{F}}(s_i^{\star}), \hat{\mathbf{F}}(s_j^{\star})\right) - \hat{\mathbf{C}}\left(\hat{\mathbf{F}}(s_{i-1}^{\star}), \hat{\mathbf{F}}(s_j^{\star})\right) - \hat{\mathbf{C}}\left(\hat{\mathbf{F}}(s_i^{\star}), \hat{\mathbf{F}}(s_{j-1}^{\star})\right) + \hat{\mathbf{C}}\left(\hat{\mathbf{F}}(s_{i-1}^{\star}), \hat{\mathbf{F}}(s_{j-1}^{\star})\right)}{\hat{\mathbf{F}}(s_i^{\star}) - \hat{\mathbf{F}}(s_{i-1}^{\star})}
$$

This parametric estimation approach is interesting when we specify the parametric functions $\mathbf{F}(s; \theta_1)$ and $\mathbf{C}(s, s'; \theta_2)$, and we estimate the parameters θ_1 and θ_2.

Generally, we have no idea about the probability functions $\mathbf{F}$ and $\mathbf{C}$. We can then adopt a non-parametric estimation approach. The first idea is to replace $\mathbf{F}$ and $\mathbf{C}$ by the empirical distribution of $\boldsymbol{S}(t)$ and the empirical copula of $(\boldsymbol{S}(t), \boldsymbol{S}(t+1))$. In practice, we can simplify this approach by estimating directly the empirical probability. Thanks to the Bayes theorem, we have:

$$
\hat{p}_{i,j}(t) = \frac{\#\{\mathcal{R}(t+1) = R_j, \mathcal{R}(t) = R_i\}}{\#\{\mathcal{R}(t) = R_i\}}
$$

---

[69]In this analysis, we have the following correspondance: $R_1 = \mathsf{CCC}, R_2 = \mathsf{B}, \ldots, R_K = \mathsf{AAA}$.
[70]There is no reason that the probability distribution of $\boldsymbol{S}(t+1)$ is different than this of $\boldsymbol{S}(t)$.

We consider a cohort of issuers for a given period $[t, t+1]$. Let $n_{i,\cdot}(t)$ be the number of issuers rated R_i at the beginning of the period t. Let $n_{i,j}(t)$ be the number of issuers rated R_i at the beginning of the period t and R_j at the end of the period t. We deduce that $\hat{p}_{i,j}$ is the ration between the two quantities:

$$\hat{p}_{i,j}(t) = \frac{n_{i,j}(t)}{n_{i,\cdot}(t)}$$

When the period is the year YYYY, the cohort starts on 1 January YYYY and ends on 31 December YYYY. If we have several annual cohorts, we can average the empirical probabilities:

$$\hat{p}_{i,j} = \frac{1}{T}\sum_{t=1}^{T}\hat{p}_{i,j}(t) = \frac{1}{T}\sum_{t=1}^{T}\frac{n_{i,j}(t)}{n_{i,\cdot}(t)}$$

Another approach is to use the pooling method:

$$\hat{p}_{i,j} = \frac{\sum_{t=1}^{T}n_{i,j}(t)}{\sum_{t=1}^{T}n_{i,\cdot}(t)}$$

The two approaches give different results. In the first case, each annual cohort has the same weight. In the second case, the approach puts more weight on the year which is more representative[71].

Table 2.28: Number of observations $n_{i,j}$ (migration matrix #1)

$n_{i,j}$	AAA	AA	A	BBB	BB	B	CCC	$n_{i,\cdot}(t)$	$\hat{p}_{i,\cdot}(t)$
AAA	2 050	125	20	10	5	0	0	2 210	3.683%
AA	280	5 580	800	60	20	5	0	6 745	11.242%
A	20	1 700	8 020	1 150	90	10	0	10 990	18.317%
BBB	10	190	2 820	10 000	1 300	60	10	14 390	23.983%
BB	5	25	200	2 500	9 150	1 000	30	12 910	21.517%
B	0	5	25	150	2 260	7 800	250	10 490	17.483%
CCC	0	0	5	10	50	300	1 900	2 265	3.775%
$n_{\cdot,j}(t)$	2 365	7 625	11 890	13 850	12 875	9 175	2 190	60 000	
$\hat{p}_{\cdot,j}(t)$	3.942%	12.708%	19.817%	23.133%	21.458%	15.292%	3.650%		100.00%

In Table 2.28, we report all the information[72] for estimating the migration matrix #1. We have used the pooling method with 60 000 observations. For 2 050 observations, the initial rating on 1 January YYYY is AAA and the final rating on 31 December YYYY is AAA. We observe 125 cases where a AAA-rated issuer has been downgraded by one notch. If we compute the sum, we obtain 2 210 AAA-rated observations at the beginning of the year and 2 365 AAA-rated observations at the end of the year. We can then compute the transition probabilities: $\hat{p}_{\mathsf{AAA},\mathsf{AAA}} = \dfrac{2\,050}{2\,210} = 92.76\%$, $\hat{p}_{\mathsf{AAA},\mathsf{AA}} = \dfrac{125}{2\,210} = 5.66\%$, ..., $\hat{p}_{\mathsf{CCC},\mathsf{CCC}} = \dfrac{1\,900}{2\,265} = 83.89\%$.

Previously, we have seen that the stationary distribution of the migration matrix #1 is equal to:

$$\pi^\star = (17.78\%, 29.59\%, 25.12\%, 15.20\%, 8.35\%, 3.29\%, 0.67\%)$$

---

[71]From a theoretical viewpoint, this second method is biased. However, it is extensively used in particular when the number of observations is low for each period.

[72]We have $n(t) = \sum_{i=1}^{K}n_{i,\cdot}(t)$, $\hat{p}_{i,\cdot}(t) = n_{i,\cdot}(t)/n(t)$, $n'(t) = \sum_{j=1}^{K}n_{j,\cdot}(t)$, $\hat{p}_{j,\cdot}(t) = n_{j,\cdot}(t)/n'(t)$.

In Table 2.28, we observe that the initial empirical distribution of ratings is:

$$\hat{\pi}^{(0)} = (3.683\%, 11.242\%, 18.317\%, 23.983\%, 21.517\%, 17.483\%, 3.775\%)$$

We conclude that the long-term dynamics of the Markov chain has dramatically change the initial probability distribution. In Table 2.28, we also observe that the final distribution of ratings after one year is:

$$\hat{\pi}^{(1)} = (3.942\%, 12.708\%, 19.817\%, 23.133\%, 21.458\%, 15.290\%, 3.650\%)$$

We reiterate that the Kolmogorov equation applied to the distribution[73] $\pi^{(n)}$ is given by $\pi^{(n+1)} = P^\top \pi^{(n)}$. In particular, we verify that $\hat{\pi}^{(1)} = \hat{P}^\top \hat{\pi}^{(0)}$ where $\hat{P} = (\hat{p}_{i,j})$. In Figure 2.36, we have reported the dynamics of $\pi^{(n)}$ with $\pi^{(0)} = \hat{\pi}^{(0)}$. We conclude that the distribution of the score $\mathcal{S}(t)$ is not stationary.

Figure 2.36: Dynamics of the probability distribution $\pi^{(n)}$ (migration matrix #1)
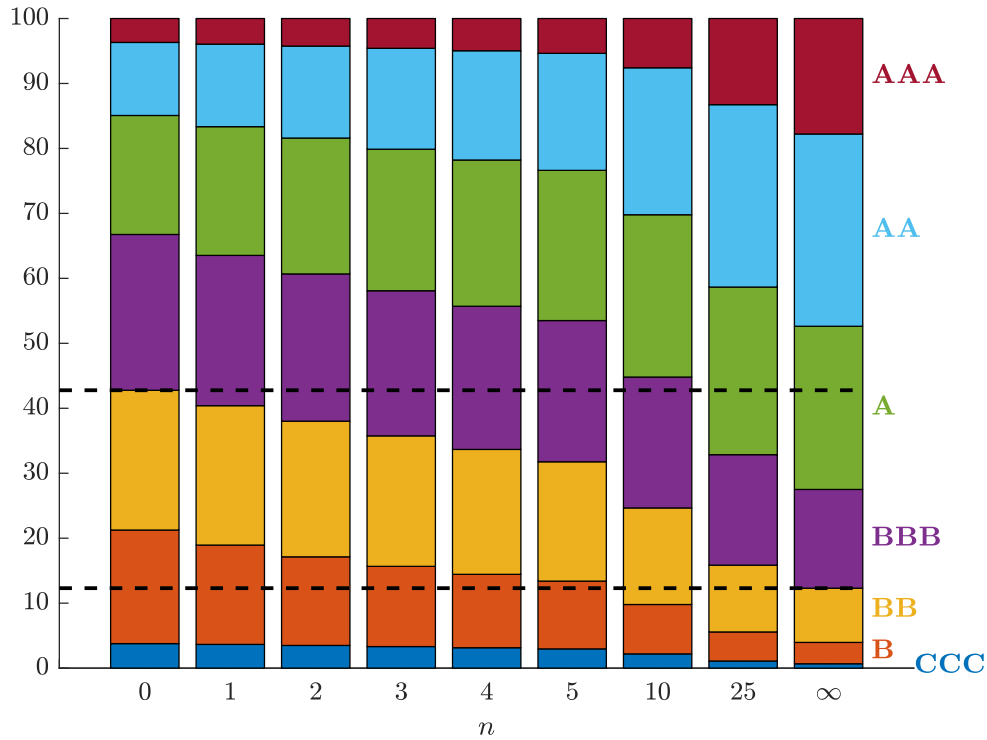
Continuous-time modeling

Markov generator We now consider the case $t \in \mathbb{R}_+$. We note $P(s;t)$ the transition matrix defined as follows:

$$P_{i,j}(s;t) = p(s, i; t, j) = \Pr\{\mathcal{R}(t) = R_j \mid \mathcal{R}(s) = R_i\}$$

Assuming that the Markov chain is time-homogenous, we have $P(t) = P(0;t)$. Jarrow *et al.* (1997) introduce the generator matrix $\Lambda = (\lambda_{i,j})$ where $\lambda_{i,j} \geq 0$ for all $i \neq j$ and $\lambda_{i,i} = -\sum_{j \neq i}^{K} \lambda_{i,j}$. In this case, the transition matrix satisfies the following relationship:

$$P(t) = \exp(t\Lambda) \tag{2.6}$$

---

[73]We have $\pi_k^{(n)} = \Pr\{\mathcal{R}(n) = R_k\}$.

where $\exp(A)$ is the matrix exponential of A. Let us give a probabilistic interpretation of Λ. If we assume that the probability of jumping from rating R_i to rating R_j in a short time period Δt is proportional to Δt, we have $p(t, i; t + \Delta t, j) = \lambda_{i,j} \Delta t$. The matrix form of this equation is $P(t; t + \Delta t) = \Lambda \Delta t$. We deduce that:

$$P(t + \Delta t) = P(t) P(t; t + \Delta t) = P(t) \Lambda \Delta t$$

and:

$$dP(t) = P(t) \Lambda \, dt$$

Because we have $\exp(\mathbf{0}) = I$, we obtain the solution $P(t) = \exp(t\Lambda)$. We then interpret $\lambda_{i,j}$ as the instantaneous transition rate of jumping from rating R_i to rating R_j.

Remark 28 *In Appendix A.1.1, we present the matrix exponential function and its mathematical properties. In particular, we have $e^{A+B} = e^A e^B$ and $e^{A(s+t)} = e^{As} e^{At}$ where A and B are two square matrices such that $AB = BA$ and s and t are two real numbers.*

Example 8 *We consider a rating system with three states:* A *(good rating),* B *(average rating) and* C *(bad rating). The Markov generator is equal to:*

$$\Lambda = \begin{pmatrix} -0.30 & 0.20 & 0.10 \\ 0.15 & -0.40 & 0.25 \\ 0.10 & 0.15 & -0.25 \end{pmatrix}$$

The one-year transition probability matrix is equal to:

$$P(1) = e^{\Lambda} = \begin{pmatrix} 75.63\% & 14.84\% & 9.53\% \\ 11.63\% & 69.50\% & 18.87\% \\ 8.52\% & 11.73\% & 79.75\% \end{pmatrix}$$

For the two-year maturity, we get:

$$P(2) = e^{2\Lambda} = \begin{pmatrix} 59.74\% & 22.65\% & 17.61\% \\ 18.49\% & 52.24\% & 29.27\% \\ 14.60\% & 18.76\% & 66.63\% \end{pmatrix}$$

We verify that $P(2) = P(1) \cdot P(1)$. This derives from the property of the matrix exponential:

$$P(t) = e^{t\Lambda} = \left(e^{\Lambda}\right)^t = P(1)^t$$

The continuous-time framework allows to calculate transition matrices for non-integer maturities, which do not correspond to full years. For instance, the one-month transition probability matrix of the previous example is equal to:

$$P\left(\frac{1}{12}\right) = e^{\frac{1}{12}\Lambda} = \begin{pmatrix} 97.54\% & 1.62\% & 0.83 \\ 1.22\% & 96.74\% & 2.03 \\ 0.82\% & 1.22\% & 97.95 \end{pmatrix}$$

Box 2.8: Estimation of the Markov generator

One of the issues with the continuous-time framework is to estimate the Markov generator Λ. One solution consists in using the empirical transition matrix $\hat{P}(t)$, which have been calculated for a given time horizon t. In this case, the estimate $\hat{\Lambda}$ must satisfy the relationship $\hat{P}(t) = \exp\left(t\hat{\Lambda}\right)$. We deduce that:

$$\hat{\Lambda} = \frac{1}{t}\ln\left(\hat{P}(t)\right)$$

where $\ln A$ is the matrix logarithm of A. However, the matrix $\hat{\Lambda}$ cannot verify the Markov conditions $\hat{\lambda}_{i,j} \geq 0$ for all $i \neq j$ and $\sum_{j=1}^{K} \lambda_{i,j} = 0$. Therefore, Israel *et al.* (2001) propose two estimators to obtain a valid generator:

1. the first approach consists in adding the negative values back into the diagonal values:
$$\left\{\begin{array}{ll} \bar{\lambda}_{i,j} = \max\left(\hat{\lambda}_{i,j}, 0\right) & i \neq j \\ \bar{\lambda}_{i,i} = \hat{\lambda}_{i,i} + \sum_{j \neq i}\min\left(\hat{\lambda}_{i,j}, 0\right) \end{array}\right.$$

2. in the second method, we carry forward the negative values on the matrix entries which have the correct sign:
$$\left\{\begin{array}{l} G_i = \left|\hat{\lambda}_{i,i}\right| + \sum_{j \neq i}\max\left(\hat{\lambda}_{i,j}, 0\right) \\ B_i = \sum_{j \neq i}\max\left(-\hat{\lambda}_{i,j}, 0\right) \\ \tilde{\lambda}_{i,j} = \left\{\begin{array}{ll} 0 & \text{if } i \neq j \text{ and } \hat{\lambda}_{i,j} < 0 \\ \hat{\lambda}_{i,j} - B_i\left|\hat{\lambda}_{i,j}\right|/G_i & \text{if } G_i > 0 \\ \hat{\lambda}_{i,j} & \text{if } G_i = 0 \end{array}\right. \end{array}\right.$$

Markov property   The Markov property refers to the lack of memory of stochastic processes. This implies that the probability distribution of future states of the process conditional on both past and present values depends only upon the present state. Therefore, given the present, the future does not depend on the past. In order to better understand the implications of this property, we consider the following example with three companies:

$t-2$		$t-1$		t		$t+1$
AAA	$\longrightarrow$	BBB	$\longrightarrow$	BBB	$\longrightarrow$	?
BBB	$\longrightarrow$	BBB	$\longrightarrow$	BBB	$\longrightarrow$	?
BB	$\longrightarrow$	BB	$\longrightarrow$	BBB	$\longrightarrow$	?

Today, the three companies are rated BBB. We would like to predict the ESG rating of those companies at time $t+1$. If the ESG ratings are Markovian, these entities are equivalent and have the same conditional probabilities to become AAA, AA, etc. Otherwise, this means that the conditional probabilities depend on the past trajectory. In this case, we have:

$$\Pr\left\{\mathcal{R}_{c_1}(t+1) = R_j \mid \mathcal{R}_{c_1}(t) = R_i\right\} \neq \Pr\left\{\mathcal{R}_{c_2}(t+1) = R_j \mid \mathcal{R}_{c_2}(t) = R_i\right\}$$

for two different companies c_1 and c_2. In our example, the firms have different past trajectories. They don't have the same transition matrix if the rating process has not the Markov property.

To verify the Markov property, we compute the matrix $\Lambda' = \ln(P)$ and measure whether Λ' is a Markov generator or not. Using the rating migration matrix #1, we obtain the results given in Table 2.29. We notice that $\ln P$ is not a Makov generator since 11 off-diagonal elements are not positive. Using the first method of Israel *et al.* (2001) described in Box 2.8, we transform this matrix into a Markov generator[74] $\bar{\Lambda}$ (Table 2.30). We recompute the one-year transition matrix $\bar{P}(1) = \exp(\bar{\Lambda})$ and observe some small differences with the original transition matrix (see Table 2.31 vs. Table 2.22).

Table 2.29: Non-Markov generator $\Lambda' = \ln(P)$ of the migration matrix #1 (in %)

	AAA	AA	A	BBB	BB	B	CCC
AAA	-7.663	6.427	0.542	0.466	0.245	-0.016	-0.000
AA	4.770	-20.604	15.451	-0.001	0.318	0.066	-0.001
A	-0.267	20.259	-35.172	14.953	0.152	0.083	-0.008
BBB	0.102	-1.051	28.263	-40.366	13.100	-0.128	0.080
BB	0.032	0.307	-1.762	28.351	-37.889	10.832	0.129
B	-0.005	-0.008	0.503	-2.240	30.227	-31.482	3.006
CCC	0.000	-0.024	0.194	0.469	0.365	16.806	-17.810

Table 2.30: Markov generator of the migration matrix #1 (in %)

	AAA	AA	A	BBB	BB	B	CCC
AAA	-7.679	6.427	0.542	0.466	0.245	0.000	0.000
AA	4.770	-20.606	15.451	0.000	0.318	0.066	0.000
A	0.000	20.259	-35.447	14.953	0.152	0.083	0.000
BBB	0.102	0.000	28.263	-41.545	13.100	0.000	0.080
BB	0.032	0.307	0.000	38.351	-39.651	10.832	0.129
B	0.000	0.000	0.503	0.000	30.227	-33.735	3.006
CCC	0.000	0.000	0.194	0.469	0.365	16.806	-17.834

Table 2.31: ESG migration Markov matrix #1 (one-year transition probability in %)

	AAA	AA	A	BBB	BB	B	CCC
AAA	92.75	5.66	0.90	0.45	0.23	0.01	0.00
AA	4.17	82.73	11.85	0.89	0.30	0.07	0.00
A	0.40	15.51	72.79	10.39	0.81	0.10	0.01
BBB	0.12	2.11	19.60	68.69	8.91	0.50	0.07
BB	0.04	0.43	2.79	19.25	69.65	7.61	0.23
B	0.01	0.09	0.65	2.98	21.21	72.71	2.35
CCC	0.00	0.02	0.25	0.58	2.19	13.09	83.87

---

[74]The matrix $\bar{\Lambda}$ is the best Markov generator that minimize the $\mathcal{L}_1$-norm distance to P.

Dynamic analysis We have now all the tools to conduct a dynamic analysis of the ESG rating system. There is tremendous potential. For instance, we compute the probability to reach the states $\mathcal{A}$ with the following formula:

$$\pi_k\left(t, \mathcal{A}\right) = \Pr\left\{\mathcal{R}\left(t\right) \in \mathcal{A} \mid \mathcal{R}\left(0\right) = k\right\} = \sum_{j \in \mathcal{A}} \mathbf{e}_k^\top e^{t\Lambda} \mathbf{e}_j^\top$$

Some examples are given in Figure 2.37. We can also use the continuous-time framework to investigate the probability density function of conditional events, the probability over a given interval, the m-order derivative of time functions, etc. We use the properties $\partial_t \exp\left(\Lambda t\right) = \Lambda \exp\left(\Lambda t\right)$, $\partial_t^m \exp\left(\Lambda t\right) = \Lambda^m \exp\left(\Lambda t\right)$ and[75] $\int_0^t e^{\Lambda s}\,\mathrm{d}s = \left(e^{\Lambda t} - I_K\right)\Lambda^{-1}$. For example, we have:

$$\pi_k^{(m)}\left(t, \mathcal{A}\right) := \frac{\partial\,\pi_k\left(t, \mathcal{A}\right)}{\partial\,t^m} = \sum_{j \in \mathcal{A}} \mathbf{e}_k^\top \Lambda^m e^{t\Lambda} \mathbf{e}_j^\top$$

$\pi_k^{(1)}\left(t, \mathcal{A}\right)$ may be interpreted as a "*time density function*". In Figure 2.37, we report $\pi_k\left(t, \mathsf{AAA}\right)$, $\pi_k^{(1)}\left(t, \mathsf{AAA}\right)$, $\pi_k\left(t, \mathsf{CCC}\right)$ and $\pi_k^{(1)}\left(t, \mathsf{CCC}\right)$. We observe the strange behavior of the CCC rating towards the AAA rating.

Figure 2.37: Probability $\pi_k\left(t, \mathcal{A}\right)$ to reach $\mathcal{A}$ at time t (migration matrix #1)
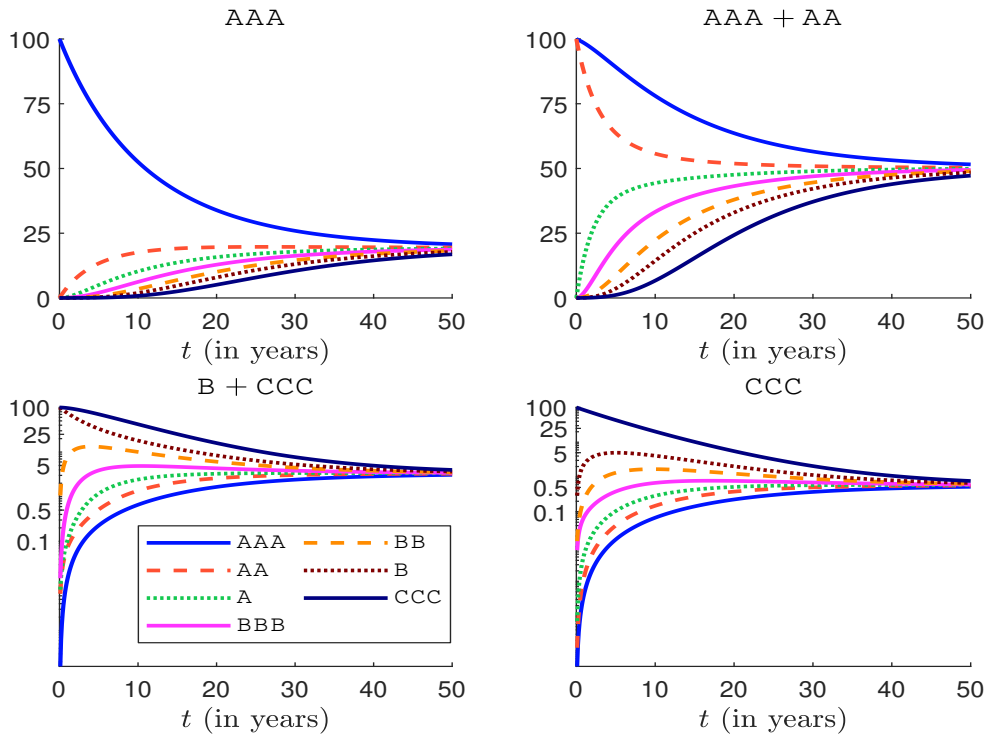
Remark 29 *The previous analysis can be used to check consistency of ratings. In particular, the fact that ratings satisfy ordering preferences implies that we must generally observe a monotone behavior of quantities that are a non-decreasing and concave function of ratings.*

---

[75] For more general integrals of type $\int_0^t e^{\Lambda s} Q f\left(s\right)\,\mathrm{d}s$, we use the numerical algorithms developed by Van Loan (1978).

Figure 2.38: Time functions $\pi_k(t, \mathsf{AAA})$, $\pi_k^{(1)}(t, \mathsf{AAA})$, $\pi_k(t, \mathsf{CCC})$ and $\pi_k^{(1)}(t, \mathsf{CCC})$ (migration matrix #1)
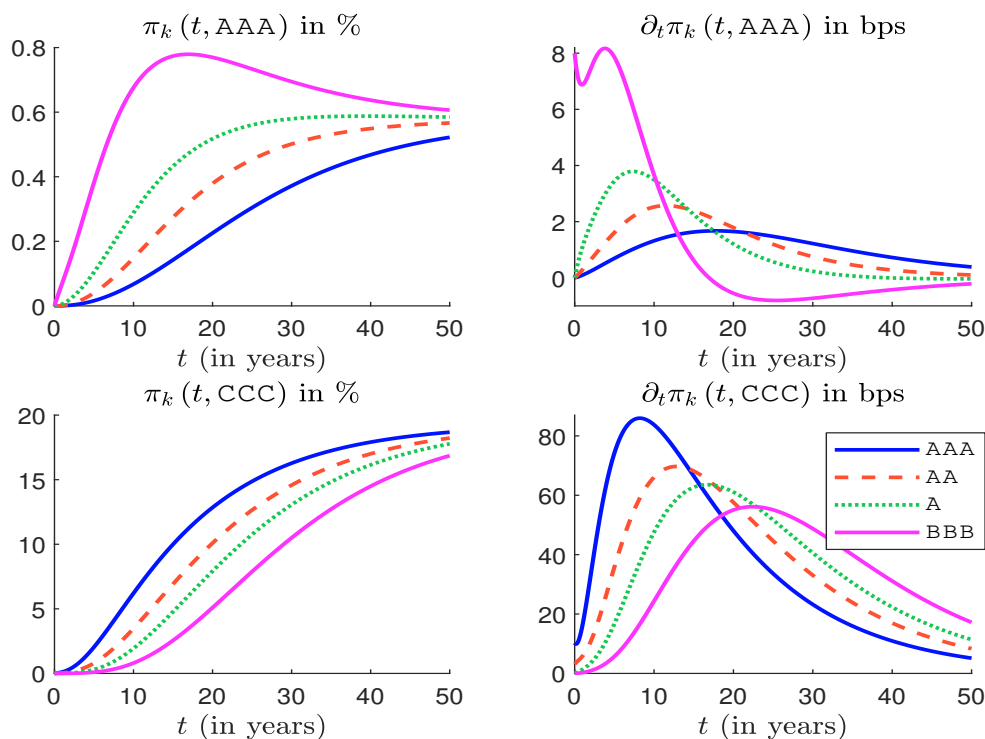
Box 2.9: Computing statistical moments with continuous-time Markov chains

The distribution $\pi(t)$ follows the Kolmogorov equation:

$$\frac{\mathrm{d}\pi(t)}{\mathrm{d}t} = \Lambda \pi(t)$$

It follows that $\pi(t) = e^{\Lambda t} \pi(0)$. Let $Y(t) = \phi(\mathcal{R}(t))$ be a random variable that depends on the ratings. We have:

$$\mu(t) = \sum_{k=1}^{K} \phi(R_k) \pi_k(t)$$

and:

$$\sigma^2(t) = \sum_{k=1}^{K} (\phi(R_k) - \mu(t))^2 \pi_k(t)$$

2.3.4 Comparison with credit ratings

The modeling of credit ratings is similar than this of ESG ratings, but there is one important difference. The states include the default state. This means that R_K is the absorbing state, implying that any entity which has reached this state remains in this state. In this case, $p_{i,K}$ is the one-period default probability of rating R_i and we have $p_{K,K} = 1$. An example of credit migration matrix is given in Table 2.32. It is the S&P one-year transition probability matrix for corporate bonds estimated by Kavvathas (2001) for the period 1960-1998. More recent credit migration matrices[76] are given in Table 2.33.

Table 2.32: Example of credit migration matrix (one-year probability transition in %)

	AAA	AA	A	BBB	BB	B	CCC	D
AAA	92.82	6.50	0.56	0.06	0.06	0.00	0.00	0.00
AA	0.63	91.87	6.64	0.65	0.06	0.11	0.04	0.00
A	0.08	2.26	91.66	5.11	0.61	0.23	0.01	0.04
BBB	0.05	0.27	5.84	87.74	4.74	0.98	0.16	0.22
BB	0.04	0.11	0.64	7.85	81.14	8.27	0.89	1.06
B	0.00	0.11	0.30	0.42	6.75	83.07	3.86	5.49
CCC	0.19	0.00	0.38	0.75	2.44	12.03	60.71	23.50
D	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100.00

Source: Kavvathas (2001).

Table 2.33: Credit migration matrix in % (Moody's, 1983-2021)

	Aaa	Aa	A	Baa	Ba	B	Caa	W	D
				Sovereign issuers					
Aaa	96.99	2.87	0.03	0.08	0.00	0.00	0.00	0.03	0.00
Aa	2.73	93.52	2.56	0.62	0.09	0.00	0.00	0.47	0.00
A	0.00	3.60	92.17	3.19	0.98	0.06	0.00	0.00	0.00
Baa	0.00	0.00	5.43	89.17	4.98	0.39	0.03	0.00	0.00
Ba	0.00	0.00	0.00	6.91	85.72	6.53	0.29	0.10	0.44
B	0.00	0.00	0.00	0.00	4.31	88.49	4.50	0.26	2.43
Caa	0.00	0.00	0.00	0.00	0.06	13.60	73.24	0.75	12.35
				Corporates issuers					
Aaa	87.16	8.05	0.45	0.08	0.03	0.00	0.00	4.23	0.00
Aa	0.70	85.02	8.57	0.42	0.06	0.04	0.02	5.17	0.02
A	0.05	2.44	86.84	5.15	0.45	0.10	0.04	4.88	0.05
Baa	0.02	0.12	3.73	86.43	3.42	0.65	0.16	5.31	0.15
Ba	0.00	0.03	0.38	6.02	75.95	7.19	0.86	8.78	0.77
B	0.01	0.03	0.12	0.42	4.73	73.61	7.34	10.79	2.95
Caa	0.00	0.01	0.02	0.07	0.26	5.58	70.41	14.82	8.83

Source: Moody's (2020).

---

[76]W means that the issuer has required to stop the rating (withdrawn).

Since there are few research on ESG ratings, credit migration matrices can be used as a benchmark to compare the two rating systems. For that, we consider the trace statistics:

$$\lambda\left(t\right) = \frac{\text{trace}\left(e^{t\Lambda}\right)}{K}$$

It is the average of the diagonal transition probabilities. It measures the average probability to remain in its state[77]. Results are reported in Figure 2.39. Even if the two ESG rating systems used here are fictitious examples, we generally conclude that ESG rating systems are less stable than credit rating systems, and the time horizon of ESG ratings for prediction is shorter.

Figure 2.39: Trace statistics of credit and ESG migration matrices
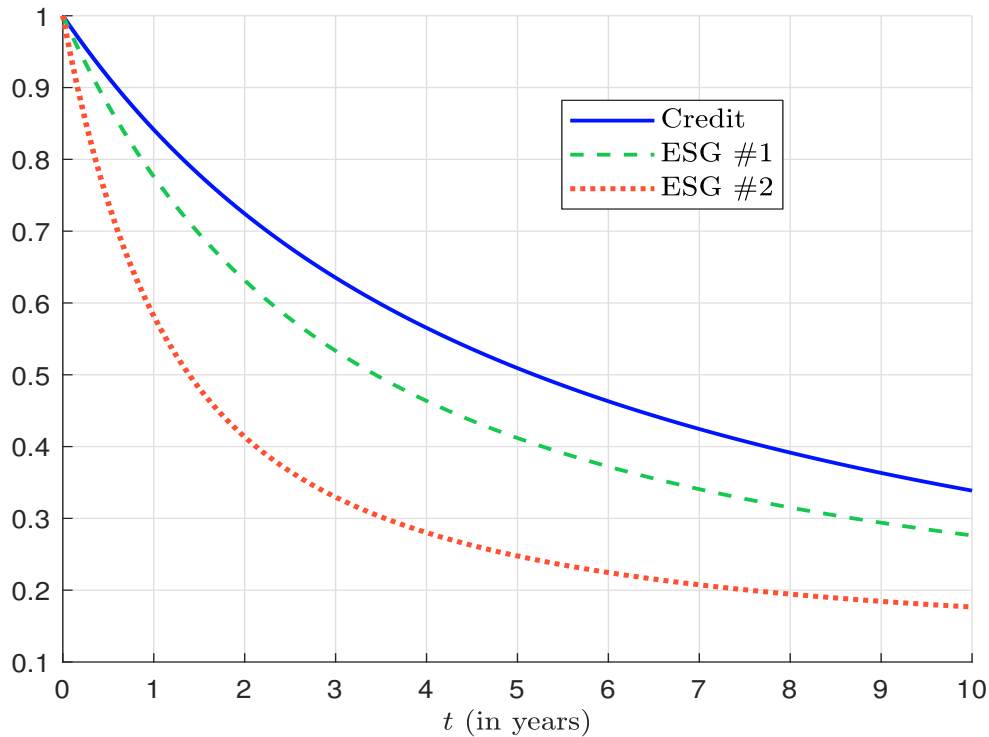
2.4   Exercises

2.4.1   Score normalization when the features are independent

We study the behavior of the score defined as:

$$\boldsymbol{\mathcal{S}} = \frac{X_1 + X_2 + \cdots + X_m}{m}$$

where $X_1, \ldots, X_m$ is a sequence of *iid* uniform random variables, whose distribution is $\mathbf{F}$.

1. We consider the case $m = 2$ and $X_j \sim \mathcal{U}_{[0,1]}$. The score is a weighted average of X_1 and X_2:

$$\boldsymbol{\mathcal{S}} = \omega X_1 + \left(1 - \omega\right) X_2$$

where $\omega \in [0, 1]$.

---

[77]For a credit migration matrix, we consider all the states except the absorbing state. In this case, we have $\lim_{t \to \infty} \lambda\left(t\right) = 0$.

(a) Let $s \in [0, 1]$. Find the points of intersection between the curve $x_2 = (s - \omega X_1) / (1 - \omega)$ and the unit square. Discuss the different cases.

(b) For each case, compute the area $\mathcal{A}(s)$ defined as:

$$\mathcal{A}(s) = \iint_{\Omega(s)} \mathrm{d}x_1 \, \mathrm{d}x_2$$

where $\Omega(s) = \left\{ (x_1, x_2) \in [0, 1]^2 : \omega x_1 + (1 - \omega) x_2 \leq s \right\}$. Deduce the cumulative distribution function $\mathbf{G}$ of the score.

(c) Compute the density function g.

(d) Find $\mathbf{G}$ and g when $X_j \sim \mathcal{U}_{[a,b]}$ where $b > a$.

2. We consider the case $m = 3$ and $X_j \sim \mathcal{U}_{[0,1]}$. The volume $\mathcal{V}(s)$ is equal to:

$$\mathcal{V}(s) = \iiint_{\Omega(s)} \mathrm{d}x_1 \, \mathrm{d}x_2 \, \mathrm{d}x_3$$

where $\Omega(s) = \left\{ (x_1, x_2, x_3) \in [0, 1]^3 : x_1 + x_2 + x_3 \leq s \right\}$.

(a) Compute the volume $\mathcal{V}(s)$ when $0 \leq s \leq 1$.

(b) Compute the volume difference $\mathcal{V}(s) - \mathcal{V}(1)$ when $1 \leq s \leq 2$.

(c) Compute the volume difference $\mathcal{V}(s) - \mathcal{V}(2)$ when $1 \leq s \leq 2$.

(d) Deduce the cumulative distribution function $\mathbf{G}$ of the score.

(e) Compute the density function g.

3. We consider that $X_j \sim \mathcal{U}_{[0,1]}$ and $m \geq 1$. We note $\mathbf{G}_m(s)$ the probability $\Pr\{\mathcal{S} \leq s\}$.

(a) Give the expression of $\mathbf{G}_m(s)$ and the associate density function $g_m(s)$.

(b) We assume that $X_j \sim \mathcal{U}_{[a,b]}$ where $b > a$. Deduce the expressions of the density and distribution functions of the score $\mathcal{S}$.

4. We assume that $X_j \sim \mathcal{G}(\alpha_j, \beta)$ where $\alpha_j > 0$ and $\beta > 0$.

(a) Compute the cumulative distribution function $\mathbf{G}$ of the score.

(b) Deduce the density function g.

(c) Compute the mean and the variance of $\mathcal{S}$.

(d) We assume that $\alpha_j = 2$ and $\beta = 2$.

    i. Draw the functions $\mathbb{E}[\mathcal{S}]$ and $\mathrm{var}(\mathcal{S})$ with respect to the number m of features.

    ii. Find the value $m^+(p, \varepsilon)$ such that:

$$m^+(p, \varepsilon) = \{\inf m : \Pr\{2 - \varepsilon \leq \mathcal{S} \leq 2 - \varepsilon\} \leq p\}$$

    for the pairs $(p, \varepsilon) = (99\%, 5\%)$.

    iii. Draw the function $m^+(p, \varepsilon)$ with respect to p when $\varepsilon = 1\%$.

    iv. Draw the function $m^+(p, \varepsilon)$ with respect to ε when $p = 99.99\%$.

2.4.2   Score normalization when the features are correlated

2.4.3   Construction of a sovereign ESG score

2.4.4   Probability distribution of an ESG score

1. We consider an investment universe of 8 issuers with the following ESG scores:

Issuer	#1	#2	#3	#4	#5	#6	#7	#8
E	−2.80	−1.80	−1.75	0.60	0.75	1.30	1.90	2.70
S	−1.70	−1.90	0.75	−1.60	1.85	1.05	0.90	0.70
G	0.30	−0.70	−2.75	2.60	0.45	2.35	2.20	1.70

   (a) Calculate the ESG score of the issuers if we assume the following weighting scheme: 40% for **E**, 40% for **S** and 20% for **G**.

   (b) Calculate the ESG score of the equally-weighted portfolio x_{ew}.

2. We assume that the ESG scores are *iid* and follow a standard Gaussian distribution:

$$\boldsymbol{\mathcal{S}}_i \sim \mathcal{N}(0,1)$$

   (a) We note $x_{\mathrm{ew}}^{(n)}$ the equally-weighted portfolio composed of n issuers. Calculate the distribution of the ESG score $\boldsymbol{\mathcal{S}}\left(x_{\mathrm{ew}}^{(n)}\right)$ of the portfolio $x_{\mathrm{ew}}^{(n)}$.

   (b) What is the ESG score of a well-diversified portfolio?

   (c) We note $T \sim \mathbf{F}_\alpha$ where $\mathbf{F}_\alpha(t) = t^\alpha$, $t \in [0,1]$ and $\alpha \geq 0$. Draw the graph of the probability density function $f_\alpha(t)$ when α is respectively equal to 0.5, 1.5, 2.5 and 70. What do you notice?

   (d) We assume that the weights of the portfolio[78] $x = (x_1, \ldots, x_n)$ follow a power-law distribution $\mathbf{F}_\alpha$:

$$x_i \sim cT_i$$

   where $T_i \sim \mathbf{F}_\alpha$ are *iid* random variables and c is a normalization constant. Explain how to simulate the portfolio weights $x = (x_1, \ldots, x_n)$. Represent one simulation of the portfolio x for the previous values of α. Comment on these results. Deduce the relationship between the Herfindahl index $\mathcal{H}_\alpha(x)$ of the portfolio weights x and the parameter α.

   (e) We assume that the weight x_i and the ESG score $\boldsymbol{\mathcal{S}}_i$ of the issuer i are independent. How to simulate the portfolio's score $\boldsymbol{\mathcal{S}}(x)$? Using 50 000 replications, estimate the probability distribution function of $\boldsymbol{\mathcal{S}}(x)$ by the Monte Carlo method. Comment on these results.

   (f) We now assume that the weight x_i and the ESG score $\boldsymbol{\mathcal{S}}_i$ of the issuer i are positively correlated. More precisely, the dependence function between x_i and $\boldsymbol{\mathcal{S}}_i$ is the Normal copula function with parameter ρ. Show that this is also the copula function between T_i and $\boldsymbol{\mathcal{S}}_i$. Deduce an algorithm to simulate $\boldsymbol{\mathcal{S}}(x)$.

   (g) Using 50 000 replications, estimate the probability distribution function of $\boldsymbol{\mathcal{S}}(x)$ by the Monte Carlo method when the correlation parameter ρ is set to 50%. Comment on these results.

   (h) Estimate the relationship between the correlation parameter ρ and the expected ESG score $\mathbb{E}[\boldsymbol{\mathcal{S}}(x)]$ of the portfolio x. Comment on these results.

---

[78]We use $n = 50$ in the rest of the exercise.

(i) How are the previous results related to the size bias of ESG scoring?

3. Let $\boldsymbol{S}$ be the ESG score of the issuer. We assume that the ESG score follows a standard Gaussian distribution:

$$\boldsymbol{S} \sim \mathcal{N}(0, 1)$$

The ESG score $\boldsymbol{S}$ is also converted into an ESG rating $\mathcal{R}$, which can take the values[79] **A**, **B**, **C** and **D**.

(a) We assume that the breakpoints of the rating system are -1.5, 0 and $+1.5$. Compute the frequencies of the ratings.

(b) We would like to build a rating system such that each category has the same frequency. Find the mapping function.

(c) We would like to build a rating system such that the frequency of the median ratings **B** and **C** is 40% and the frequency of the extreme ratings **A** and **D** is 10%. Find the mapping function.

4. Let $\boldsymbol{S}(t)$ be the ESG score of the issuer at time t. The ESG scoring system is evaluated every month. The index time t corresponds to the current month, whereas the previous month is $t - 1$. We assume that:

(a)   i. The ESG score at time $t - 1$ follows a standard Gaussian distribution:

$$\boldsymbol{S}(t - 1) \sim \mathcal{N}(0, 1)$$

  ii. The variation of the ESG score is Gaussian between two months:

$$\Delta \boldsymbol{S}(t) = \boldsymbol{S}(t) - \boldsymbol{S}(t - 1) \sim \mathcal{N}\left(0, \sigma^2\right)$$

  iii. The ESG score $\boldsymbol{S}(t - 1)$ and the variation $\Delta \boldsymbol{S}(t)$ are independent.

The ESG score $\boldsymbol{S}(t)$ is converted into an ESG rating $\mathcal{R}(t)$, which can take following grades:

$$\mathcal{R}_1 \prec \mathcal{R}_2 \prec \cdots \prec \mathcal{R}_k \prec \cdots \prec \mathcal{R}_{K-1} \prec \mathcal{R}_K$$

We assume that the breakpoints of the rating system are $(s_1, s_2, \ldots, s_{K-1})$. We also note $s_0 = -\infty$ and $s_K = +\infty$.

(a) Compute the bivariate probability distribution of the random vector $(\boldsymbol{S}(t - 1), \Delta \boldsymbol{S}(t))$.

(b) Compute the bivariate distribution of the random vector $(\boldsymbol{S}(t - 1), \boldsymbol{S}(t))$.

(c) Compute the probability $p_k = \Pr\{\mathcal{R}(t - 1) = \mathcal{R}_k\}$.

(d) Compute the joint probability $\Pr\{\mathcal{R}(t) = \mathcal{R}_k, \mathcal{R}(t - 1) = \mathcal{R}_j\}$.

(e) Compute the transition probability $p_{j,k} = \Pr\{\mathcal{R}(t) = \mathcal{R}_k \mid \mathcal{R}(t - 1) = \mathcal{R}_j\}$.

(f) Compute the monthly turnover $\mathcal{T}(\mathcal{R}_k)$ of the ESG rating $\mathcal{R}_k$.

(g) Compute the monthly turnover $\mathcal{T}(\mathcal{R}_1, \ldots, \mathcal{R}_K)$ of the ESG rating system.

(h) For each rating system given in Questions 3.a, 3.b and 3.c, compute the corresponding migration matrix and the monthly turnover of the rating system if we assume that σ is equal to 10%. What is the best ESG rating system if we would like to control the turnover of ESG ratings?

---

[79]**A** is the best rating and **D** is the worst rating.

(i) Draw the relationship between the parameter σ and the turnover $\mathcal{T}(\mathcal{R}_1, \ldots, \mathcal{R}_K)$ for the three ESG rating systems.

(j) We consider a uniform ESG rating system where:

$$\Pr\{\mathcal{R}(t-1) = \mathcal{R}_k\} = \frac{1}{K}$$

Draw the relationship between the number of notches K and the turnover $\mathcal{T}(\mathcal{R}_1, \ldots, \mathcal{R}_K)$ when the parameter σ takes the values 5%, 10% and 25%.

(k) Why is an ESG rating system different than a credit rating system? What do you conclude from the previous analysis? What is the issue of ESG exclusion policy and negative screening?

2.4.5   Markov generator of ESG migration matrix

2.4.6   Properties of Markov chains

Chapter 3

Impact of ESG Investing on Asset Prices and Portfolio Returns

The question of ESG performance is on everyone's lips. This question is linked to several other questions, which can be summarized as follows: What is the impact of ESG on corporate financial performance? What is the impact of ESG investing on risk premia? How does ESG screening affect portfolio returns? Is there a difference between ESG investing and climate investing? In fact, we can multiply the questions because the term *ESG performance* covers different topics, and we need to be more precise when we talk about it. First, we need to distinguish between operational performance, social performance, accounting performance, market performance, etc. For example, it is not the same thing to measure performance based on financial statements (balance sheet and income statement) or share price evolution. Second, we can measure performance from an investor's or issuer's perspective. The third ambiguity concerns the type of financial asset. Are they stocks or bonds? Because we know that fixed-income and equity markets react differently. Another important source of discrepancy is the choice of financial instruments. We can compare the performance of securities, mutual funds, asset owners or backtests. For example, simulated performance must be different from live performance. The fifth issue is the investment universe and sample. We can imagine that the impact of ESG is different from one region to another, from one sector to another, one period to another, etc. Finally, if we focus on the financial performance of ESG strategies, the final question is the implementation of the portfolio strategy. Are we talking about an exclusion, selection, integration or momentum strategy? Are we talking about active or passive management? Moreover, since ESG scores vary widely from one rating agency to another, we are not sure that we are capturing the performance of the ESG market, but perhaps some idiosyncratic patterns. There are therefore many factors to consider, and it is no coincidence that there are many academic studies with different conclusions. It is impossible to cite all of them, even the most famous ones. They are described in meta-analyses, *e.g.*, Orlitzky *et al.* (2003), Margolis *et al.* (2009), Friede *et al.* (2015), Atz *et al.* (2022) and Coqueret (2022). Rather than delve deeply into all these empirical studies, we take a different approach. Indeed, ESG investing did not exist or was so marginal fifteen or twenty years ago. Moreover, ESG data are certainly not robust or relevant before 2010. Therefore, it is better to focus on theoretical research when analyzing the performance of ESG investing. This first section is mainly based on the work of Pástor *et al.* (2021) and Pedersen *et al.* (2021). It will help us to understand when, where and why ESG investing may underperform or overperform business-as-usual investing. The second section is dedicated to empirical studies, but we make a selection to illustrate the theoretical results and concepts defined in the first section. Finally, the third section examines the impact of ESG on the cost of capital.

3.1   Theoretical models

Before discussing the impact of ESG on the theory of risk premium and security selection, we summarize the main results of the modern portfolio theory as presented in Roncalli (2013).

3.1.1   A primer on modern portfolio theory

The concept of the market portfolio has a long history and dates back to the seminal work of Markowitz (1952). He showed that an efficient portfolio is the portfolio that maximizes the expected return for a given level of risk. Markowitz concluded that there is not only one optimal portfolio, but a set of optimal portfolios which is called the efficient frontier. By studying the liquidity preference, Tobin (1958) showed that the efficient frontier becomes a straight line in the presence of a risk-free asset. In this case, optimal portfolios correspond to a combination of the risk-free asset and one particular efficient portfolio named the tangency portfolio. Sharpe (1964) summarized the results of Markowitz and Tobin as follows: "*the process of investment choice can be broken down into two phases: first, the choice of a unique optimum combination of risky assets*[1]; *and second, a separate choice concerning the allocation of funds between such a combination and a single riskless asset*". This two-step procedure is today known as the *mutual fund separation theorem*. In this seminal research paper, Sharpe developed the CAPM theory and highlighted the relationship between the risk premium of the asset (the difference between the expected return and the risk-free rate) and its beta (the systematic risk with respect to the tangency portfolio). By assuming that the market is at equilibrium, he showed that the prices of assets are such that the tangency portfolio is the market portfolio, which is composed of all risky assets in proportion to their market capitalization.

The efficient frontier

The optimization problem   Seventy years ago, Markowitz introduced the concept of the efficient frontier. We consider a universe of n assets. Let $w = (w_1, \ldots, w_n)$ be the vector of weights in the portfolio. We assume that the portfolio is fully invested meaning that $\sum_{i=1}^{n} w_i = \mathbf{1}^\top w = 1$. We denote $R = (R_1, \ldots, R_n)$ the vector of asset returns where R_i is the return of asset i. The return of the portfolio is then equal to $R(w) = \sum_{i=1}^{n} w_i R_i = w^\top R$. Let $\mu = \mathbb{E}[R]$ and $\Sigma = \mathbb{E}\left[(R - \mu)(R - \mu)^\top\right]$ be the vector of expected returns and the covariance matrix of asset returns. The expected return $\mu(w) := \mathbb{E}[R(w)]$ of the portfolio is equal to:

$$\mu(w) = \mathbb{E}\left[w^\top R\right] = w^\top \mathbb{E}[R] = w^\top \mu$$

whereas its variance $\sigma^2(w) := \mathrm{var}(R(w))$ is given by:

$$
\begin{aligned}
\sigma^2(w) &= \mathbb{E}\left[(R(w) - \mu(w))(R(w) - \mu(w))^\top\right] \\
&= \mathbb{E}\left[w^\top (R - \mu)(R - \mu)^\top w\right] \\
&= w^\top \Sigma w
\end{aligned}
$$

We can then formulate the investor's financial problem as follows:

1. Maximizing the expected return of the portfolio under a volatility constraint (σ-problem):

$$\max \mu(w) \quad \text{s.t.} \quad \sigma(w) \leq \sigma^\star \tag{3.1}$$

---

[1]It is precisely the tangency portfolio.

2. Or minimizing the volatility of the portfolio under a return constraint (μ-problem):

$$\min \sigma(w) \quad \text{s.t.} \quad \mu(w) \geq \mu^{\star} \tag{3.2}$$

By considering all the portfolios belonging to the simplex set defined by $\{w \in [0,1]^n : \mathbf{1}^{\top} w = 1\}$, we can compute the expected return and volatility bounds of the portfolios: $\mu^{-} \leq \mu(w) \leq \mu^{+}$ and $\sigma^{-} \leq \sigma(w) \leq \sigma^{+}$. There is also a solution to the first problem if $\sigma^{\star} \geq \sigma^{-}$. The second problem has a solution if $\mu^{\star} \leq \mu^{+}$. If these two conditions are verified, the inequality constraints becomes $\sigma(w) = \min(\sigma^{\star}, \sigma^{+})$ and $\mu(w) = \max(\mu^{-}, \mu^{\star})$.

The key idea of Markowitz (1956) was to transform the original non-linear optimization problem (3.1) into a quadratic optimization problem which is easier to solve numerically. For that, he introduced the mean-variance (or quadratic) utility function:

$$\boldsymbol{\mathcal{U}}(w) := \mathbb{E}[R(w)] - \frac{\bar{\gamma}}{2} \operatorname{var}(R(w)) = w^{\top} \mu - \frac{\bar{\gamma}}{2} w^{\top} \Sigma w$$

where $\bar{\gamma}$ is the absolute risk-aversion parameter. We obtain the following problem:

$$\begin{aligned} w^{\star}(\bar{\gamma}) &= \arg\max \left\{ \boldsymbol{\mathcal{U}}(w) = w^{\top} \mu - \frac{\bar{\gamma}}{2} w^{\top} \Sigma w \right\} \\ \text{s.t.} \quad & \mathbf{1}^{\top} w = 1 \end{aligned} \tag{3.3}$$

If $\bar{\gamma} = 0$, the optimized portfolio is the one that maximizes the expected return and we have $\mu(w^{\star}(0)) = \mu^{+}$. If $\bar{\gamma} = \infty$, the risk-aversion parameter is maximum, we obtain the global minimum variance (GMV) portfolio:

$$\begin{aligned} w^{\star}(\infty) &= \arg\min \frac{1}{2} w^{\top} \Sigma w \\ \text{s.t.} \quad & \mathbf{1}^{\top} w = 1 \end{aligned}$$

and we have $\sigma(w^{\star}(\infty)) = \sigma^{-}$. In practice, we formulate the optimization problem (3.3) as follows:

$$\begin{aligned} w^{\star}(\gamma) &= \arg\min \frac{1}{2} w^{\top} \Sigma w - \gamma w^{\top} \mu \\ \text{s.t.} \quad & \mathbf{1}^{\top} w = 1 \end{aligned} \tag{3.4}$$

where $\gamma = \bar{\gamma}^{-1}$ is the inverse of the risk-aversion parameter and is called the risk-tolerance coefficient. The reason is that this new formulation is a standard quadratic programming (QP) problem[2]. From a numerical point of view, it is therefore better to use Problem (3.4). In this case, the minimum variance portfolio corresponds to $\gamma = 0$. The set of solutions $\{w^{\star}(\gamma), \gamma \geq 0\}$ corresponds to mean-variance optimized (MVO) portfolios.

Example 9 *We consider an investment universe of five assets. Their expected returns are equal to 5%, 7%, 6%, 10% and 8% while their volatilities are equal to 18%, 20%, 22%, 25% and 30%. The correlation matrix of asset returns is given by the following matrix:*

$$\mathbb{C} = \begin{pmatrix} 100\% & & & & \\ 70\% & 100\% & & & \\ 20\% & 30\% & 100\% & & \\ -30\% & 20\% & 10\% & 100\% & \\ 0\% & 0\% & 0\% & 0\% & 100\% \end{pmatrix}$$

Figure 3.1: Efficient frontier (Example 9)
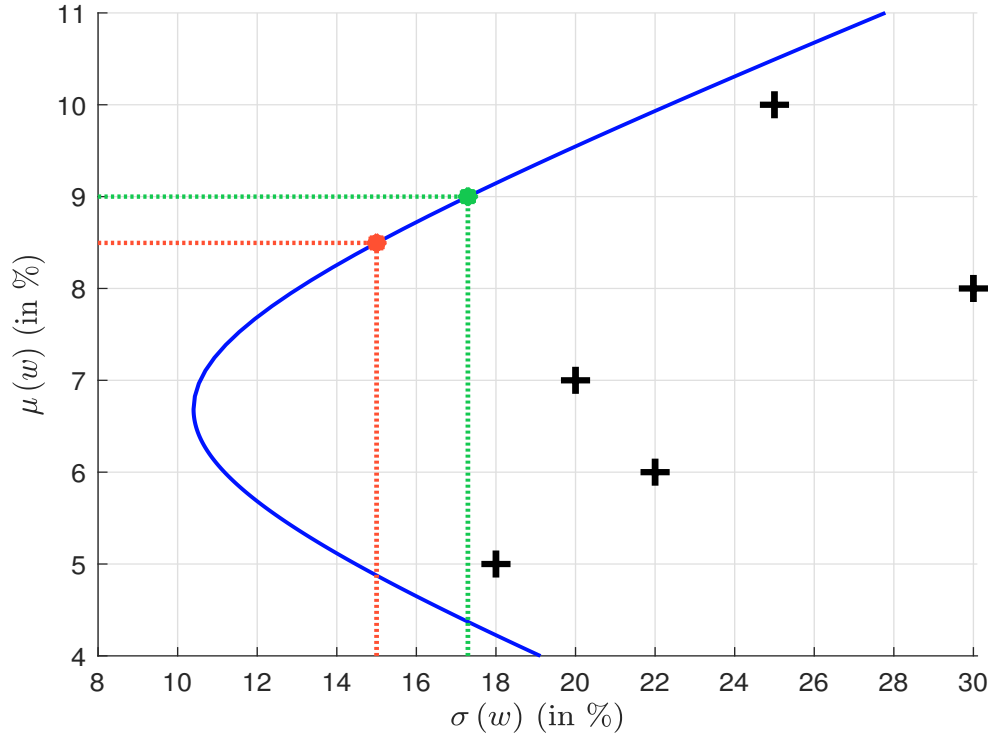
In Figure 3.1, we report the efficient frontier $\{\sigma(w^\star(\gamma)), \mu(w^\star(\gamma))\}$ by considering several values[3] of $\gamma \in [-0.5, 1]$. We note that optimized portfolios substantially improve the risk/return profile with respect to the five assets, which are represented by a cross symbol. Some special solutions are given in Table 3.1. The portfolio weights, its return and its volatility are expressed in %. For instance, the GMV portfolio is obtained with $\gamma = 0$. The solution is $(66.35\%, -28.52\%, 15.31\%, 34.85\%, 12.02\%)$ and it is not possible to find a portfolio whose volatility is lower than 10.40%.

Table 3.1: Solution of the Markowitz optimization problem (in %)

γ	0.00	0.10	0.20	0.50	1.00	5.00
$w_1^\star(\gamma)$	66.35	58.25	50.14	25.84	-14.67	-338.72
$w_2^\star(\gamma)$	-28.52	-22.67	-16.82	0.74	30.00	264.12
$w_3^\star(\gamma)$	15.31	13.30	11.30	5.28	-4.74	-84.93
$w_4^\star(\gamma)$	34.85	37.65	40.44	48.82	62.78	174.50
$w_5^\star(\gamma)$	12.02	13.48	14.94	19.32	26.62	85.03
$\mu(w^\star(\gamma))$	6.69	6.97	7.25	8.09	9.49	20.71
$\sigma(w^\star(\gamma))$	10.40	10.53	10.93	13.35	19.71	84.38

Solving the μ-problem or the σ-problem is equivalent to finding the optimal value of γ such that $\mu(w^\star(\gamma)) = \mu^\star$ or $\sigma(w^\star(\gamma)) = \sigma^\star$. We know that the functions $\mu(w^\star(\gamma))$ and $\sigma(w^\star(\gamma))$

---

[2]See Appendix A.1.2 on page 652.

[3]When $\gamma < 0$, $w^\star(\gamma)$ is not a MVO portfolio since it has a lower expected return than the GMV portfolio with a higher volatility. In fact, Problem (3.4) defines the convex hull $\{\mu(w), \sigma(w)\}$ of all possible portfolios $\{w : \mathbf{1}^\top w = 1\}$.

are increasing with respect to γ and are bounded. The optimal value of γ can then be easily computed using the bisection algorithm described on page A.1.2. This is the approach used in practice, because it benefits from the numerical efficiency of quadratic programming solvers[4]. For instance, if we target a portfolio with $\sigma^\star = 15\%$, we know that $\gamma \in [0.5, 1]$. The optimal solution $w^\star$ is $(14.06\%, 9.25\%, 2.37\%, 52.88\%, 21.44\%)$. The bisection algorithm returns $\gamma = 0.6455$. In this case, we obtain $\mu(w^\star(\gamma)) = 8.50\%$. Let us now consider a μ-problem with $\mu^\star = 9\%$. We find $\gamma = 0.8252$, $w^\star = (-0.50\%, 19.77\%, -1.23\%, 57.90\%, 24.07)$ and $\sigma(w^\star(\gamma)) = 17.30\%$.

Adding some constraints The Lagrange function of the optimization problem (3.4) is equal to:

$$\mathcal{L}(w; \lambda_0) = \frac{1}{2} w^\top \Sigma w - \gamma w^\top \mu + \lambda_0 \left(\mathbf{1}^\top w - 1 \right)$$

where λ_0 is the Lagrange coefficients associated with the constraint $\mathbf{1}^\top w = 1$. The solution $w^\star$ verifies the following first-order conditions:

$$\begin{cases} \partial_w \mathcal{L}(w; \lambda_0) = \Sigma w - \gamma \mu + \lambda_0 \mathbf{1} = \mathbf{0} \\ \partial_{\lambda_0} \mathcal{L}(w; \lambda_0) = \mathbf{1}^\top w - 1 = 0 \end{cases}$$

We obtain $w = \Sigma^{-1}(\gamma \mu - \lambda_0 \mathbf{1})$. Because $\mathbf{1}^\top w - 1 = 0$, we have $\gamma \mathbf{1}^\top \Sigma^{-1} \mu - \lambda_0 \mathbf{1}^\top \Sigma^{-1} \mathbf{1} = 1$. It follows that:

$$\lambda_0 = \frac{\gamma \mathbf{1}^\top \Sigma^{-1} \mu - 1}{\mathbf{1}^\top \Sigma^{-1} \mathbf{1}}$$

The solution is then:

$$\begin{aligned} w^\star(\gamma) &= \frac{\Sigma^{-1} \mathbf{1}}{\mathbf{1}^\top \Sigma^{-1} \mathbf{1}} + \gamma \frac{\left(\mathbf{1}^\top \Sigma^{-1} \mathbf{1}\right) \Sigma^{-1} \mu - \left(\mathbf{1}^\top \Sigma^{-1} \mu\right) \Sigma^{-1} \mathbf{1}}{\mathbf{1}^\top \Sigma^{-1} \mathbf{1}} \\ &= w_{\text{gmv}} + \gamma w_{\text{lsp}} \end{aligned}$$

where $w_{\text{gmv}} = \left(\Sigma^{-1} \mathbf{1}\right) / \left(\mathbf{1}^\top \Sigma^{-1} \mathbf{1}\right)$ is the global minimum variance portfolio and w_{lsp} is a long/short cash-neutral portfolio[5] such that $\mathbf{1}^\top w_{\text{lsp}} = 0$.

We deduce that a QP solver is not required to find the solution of the optimization problem (3.4). For instance, the analytical calculus gives $w_{\text{gmv}} = (66.35\%, -28.52\%, 15.31\%, 34.85\%, 12.02\%)$ and $w_{\text{lsp}} = (-81.01\%, 58.53\%, -20.05\%, 27.93\%, 14.60\%)$. Using numerical results in Table 3.1, we verify that the equation $w^\star(\gamma) = w_{\text{gmv}} + \gamma w_{\text{lsp}}$ is satisfied. Nevertheless, these solutions are not realistic, because they correspond to leveraged long/short portfolios, but most of investors can not have short positions. Moreover, short selling can only be implemented a few number of assets, which are very liquid and highly tradable. Otherwise, the cost of short selling is huge. This is why portfolio optimization in practice considers other constraints:

$$\begin{aligned} w^\star(\gamma) &= \arg \min \frac{1}{2} w^\top \Sigma w - \gamma w^\top \mu \\ \text{s.t.} \quad &\begin{cases} \mathbf{1}^\top w = 1 \\ w \in \Omega \end{cases} \end{aligned} \tag{3.5}$$

where $w \in \Omega$ corresponds to the set of restrictions. The most frequent constraints are certainly the no short-selling restriction and asset bounds. In the first case, $w_i \geq 0$ and $\Omega = [0, 1]^n$. The second

---

[4]From a numerical point of view, it is generally better to solve several QP problems than one non-linear optimization problem.

[5]We have $\mathbf{1}^\top w^\star(\gamma) = 1 \Leftrightarrow \mathbf{1}^\top w_{\text{gmv}} + \gamma \mathbf{1}^\top w_{\text{lsp}} = 1 \Leftrightarrow \mathbf{1}^\top w_{\text{lsp}} = 0$ because $\mathbf{1}^\top w_{\text{gmv}} = 1$.

constraint imposes $w_i \leq w^+$, in order to be sure that the portfolio is not concentrated in a few number of assets.

Let us introduce some constraints in Example 9. In Figure 3.2, we have reported two constrained efficient frontiers, the first one by imposing no short-selling and the second one by imposing that the weights are between 0% and 33%. We verify that investment constraints may substantially reduce opportunity arbitrages.

Figure 3.2: Impact of constraints on the efficient frontier (Example 9)
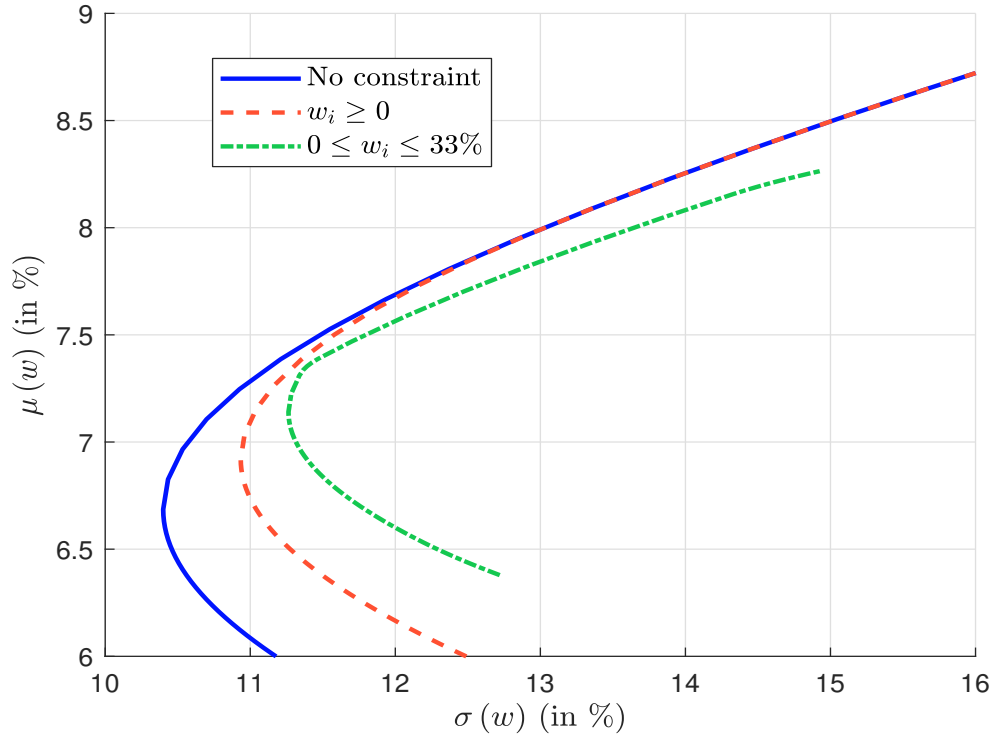
The tangency portfolio

Two-fund separation theorem   We recall that in the view of Markowitz, there is a set of optimized portfolios. However, Tobin (1958) showed that one optimized portfolio dominates all the others if there is a risk-free asset. Let us consider a combination of the risk-free asset and a portfolio w. We denote r the return of the risk-free asset. We have[6]:

$$R(\tilde{w}) = (1 - \alpha) r + \alpha R(w)$$

where $\tilde{w} = (\alpha w, 1 - \alpha)$ is a vector of dimension $(n + 1)$ and $\alpha \geq 0$ is the proportion of the wealth invested in the risky portfolio. It follows that:

$$\mu(\tilde{w}) = (1 - \alpha) r + \alpha \mu(w) = r + \alpha(\mu(w) - r)$$

and:

$$\sigma^2(\tilde{w}) = \alpha^2 \sigma^2(w)$$

---

[6]We have $n + 1$ assets in the universe where the first n assets correspond to the previous risky assets and the last asset is the risk-free asset.

We deduce that:

$$\mu\left(\tilde{w}\right) = r + \frac{\left(\mu\left(w\right) - r\right)}{\sigma\left(w\right)}\sigma\left(\tilde{w}\right) \tag{3.6}$$

It is the equation of a linear function between the volatility and the expected return of the combined portfolio $\tilde{w}$. In Figure 3.3, we reported the previous (unconstrained) efficient frontier. The dashed line corresponds to the combination between the risk-free asset (r is equal to 3%) and the optimized portfolio $w^{\star}$. Nevertheless this combination is suboptimal, because it is dominated by other combinations. We note that a straight line dominates all the other straight lines and the efficient frontier. This line is tangent to the efficient frontier and is called the capital market line. It implies that one optimized risky portfolio dominates all the other risky portfolios, namely the tangency portfolio. We denote it by w^{*}.

Figure 3.3: Capital market line (Example 9)
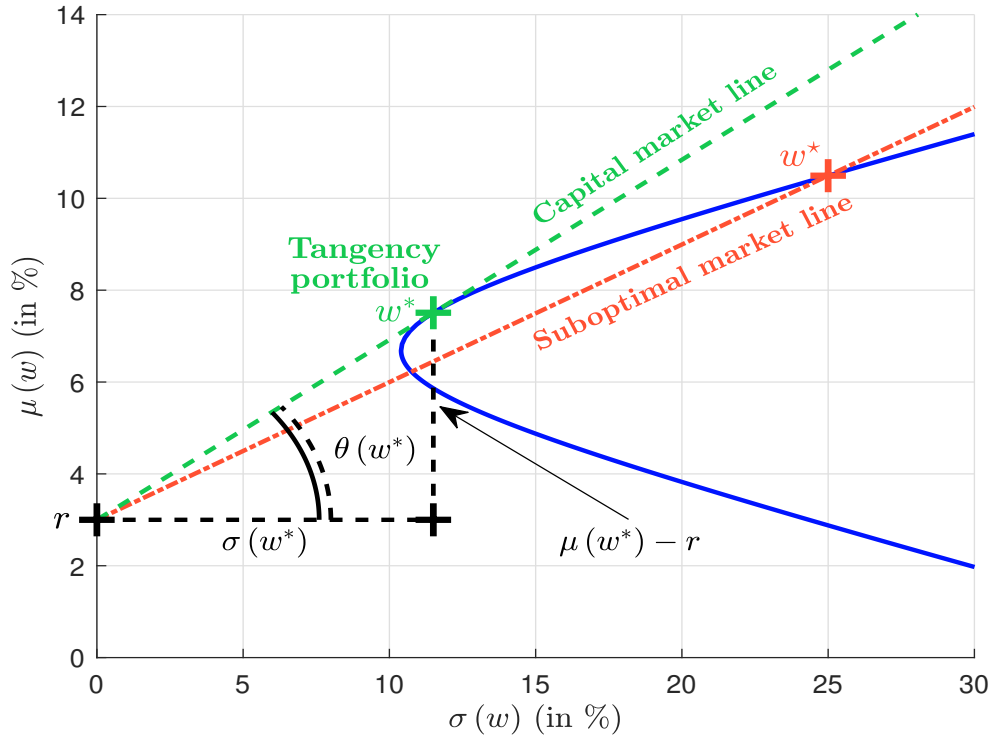
Let $\mathrm{SR}\left(w \mid r\right)$ be the Sharpe ratio of portfolio w:

$$\mathrm{SR}\left(w \mid r\right) = \frac{\mu\left(w\right) - r}{\sigma\left(w\right)}$$

We note that we can write Equation (3.6) as follows:

$$\frac{\mu\left(\tilde{w}\right) - r}{\sigma\left(\tilde{w}\right)} = \frac{\mu\left(w\right) - r}{\sigma\left(w\right)} \Leftrightarrow \mathrm{SR}\left(\tilde{w} \mid r\right) = \mathrm{SR}\left(w \mid r\right)$$

We deduce that the tangency portfolio is the one that maximizes the angle $\theta\left(w\right)$ or equivalently $\tan\theta\left(w\right)$ which is equal to the Sharpe ratio. The tangency portfolio is also the risky portfolio corresponding to the maximum Sharpe ratio. We also note that any portfolio which belongs to the capital market line has the same Sharpe ratio. If we consider our example with $r = 3\%$, the composition of the tangency portfolio w^{*} is $(42.57\%, -11.35\%, 9.43\%, 43.05\%, 16.30\%)$ and we have $\mu\left(w^{*}\right) = 7.51\%$, $\sigma\left(w^{*}\right) = 11.50\%$, $\mathrm{SR}\left(w^{*} \mid r\right) = 0.39$ and $\theta\left(w^{*}\right) = 21.40$ degrees.

Augmented optimization problem   When the risk-free asset belongs to the investment universe, the optimization problem becomes:

$$\tilde{w}^\star(\gamma) \quad = \quad \arg\min \frac{1}{2}\tilde{w}^\top \tilde{\Sigma}\tilde{w} - \gamma \tilde{w}^\top \tilde{\mu} \tag{3.7}$$

$$\text{s.t.} \quad \begin{cases} \mathbf{1}^\top \tilde{w} = 1 \\ \tilde{w} \in \Omega \end{cases}$$

where $\tilde{w} = (w, w_r)$ is the augmented allocation vector of dimension $n + 1$. It follows that:

$$\tilde{\Sigma} = \begin{pmatrix} \Sigma & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix} \quad \text{and} \quad \tilde{\mu} = \begin{pmatrix} \mu \\ r \end{pmatrix}$$

In the case where $\Omega = \mathbb{R}^{n+1}$, Roncalli (2013, pages 13-14) showed that the optimal solution is equal to:

$$\tilde{w}^\star(\gamma) = \underbrace{\alpha \cdot \begin{pmatrix} w^* \\ 0 \end{pmatrix}}_{\text{risky assets}} + \underbrace{(1-\alpha) \cdot \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix}}_{\text{risk-free asset}}$$

where w^* is the tangency portfolio:

$$w^* = \frac{\Sigma^{-1}(\mu - r\mathbf{1})}{\mathbf{1}^\top \Sigma^{-1}(\mu - r\mathbf{1})}$$

and the proportion of risky assets is equal to $\alpha = \gamma \mathbf{1}^\top \Sigma^{-1}(\mu - r\mathbf{1})$. It follows that the risk-tolerance coefficient associated to the tangency portfolio is given by:

$$\gamma(w^*) = \frac{1}{\mathbf{1}^\top \Sigma^{-1}(\mu - r\mathbf{1})}$$

When $\alpha \neq 1$, the weights $\tilde{w}^\star(\gamma)$ of the optimal portfolio are proportional to the weights w^* of the tangency portfolio whereas the wealth invested in the risk-free asset is the complement $(1 - \alpha)$ to obtain a total exposure equal to 100%. We retrieve then the *two-fund separation theorem*.

Market equilibrium and CAPM

Risk premium and beta   Based on the results of Markowitz and Tobin, Sharpe (1964) developed the capital asset pricing model (CAPM). Let w^* be the tangency portfolio. On the efficient frontier, we have seen that any portfolio w satisfies the capital market line:

$$\mu(w) = r + \frac{\sigma(w)}{\sigma(w^*)}(\mu(w^*) - r)$$

We consider a portfolio x with a proportion ω invested in the asset i and a proportion $(1 - \omega)$ invested in the tangency portfolio w^*. We have[7] $\mu(x) = \omega\mu_i + (1-\omega)\mu(w^*)$ and $\sigma^2(x) = \omega^2\sigma_i^2 + (1-\omega)^2 \sigma^2(w^*) + 2\omega(1-\omega)\rho(\mathbf{e}_i, w^*)\sigma_i\sigma(w^*)$. It follows that:

$$\frac{\partial \mu(x)}{\partial \sigma(x)} = \frac{\mu_i - \mu(w^*)}{\left(\omega\sigma_i^2 + (\omega - 1)\sigma^2(w^*) + (1 - 2\omega)\rho(\mathbf{e}_i, w^*)\sigma_i\sigma(w^*)\right)\sigma^{-1}(x)}$$

---

[7]$\mathbf{e}_i$ is the unit vector with 1 in the ith position and 0 elsewhere. It corresponds then to the portfolio fully invested in asset i.

When $\omega = 0$, the portfolio x is the tangency portfolio w^* and the previous derivative is equal to the Sharpe ratio $\text{SR}(w^* \mid r)$:

$$\lim_{\omega \to 0} \frac{\partial \mu(x)}{\partial \sigma(x)} = \tan \theta(w^*) = \frac{\mu(w^*) - r}{\sigma(w^*)}$$

We deduce that:

$$\frac{(\mu_i - \mu(w^*)) \sigma(w^*)}{\rho(\mathbf{e}_i, w^*) \sigma_i \sigma(w^*) - \sigma^2(w^*)} = \frac{\mu(w^*) - r}{\sigma(w^*)}$$

which is equivalent to:

$$\pi_i := \mu_i - r = \beta_i(\mu(w^*) - r) \tag{3.8}$$

where π_i is the risk premium of the asset i and:

$$\beta_i = \frac{\rho(\mathbf{e}_i, w^*) \sigma_i}{\sigma(w^*)} = \frac{\text{cov}(R_i, R(w^*))}{\text{var}(R(w^*))} \tag{3.9}$$

The coefficient β_i is the ratio of the covariance between the return of asset i and the return of the tangency portfolio upon the variance of the tangency portfolio return. Equation (3.8) tells us that the risk premium of the asset i is equal to its beta times the excess return of the tangency portfolio. It is easy to show that this relationship remains valid for any portfolio w and not only for the assets that compose the tangency portfolio.

<div style="text-align:center">Box 3.1: Computation of the beta coefficient</div>

Let $R_{i,t}$ and $R_t(w)$ be the returns of asset i and portfolio w at time t. We consider the linear regression:

$$R_{i,t} = \alpha_i + \beta_i R_t(w) + \varepsilon_{i,t}$$

where $\varepsilon_{i,t}$ is a white noise process. The OLS coefficient $\hat{\beta}_i$ is an estimate of the beta β_i of the asset i. We can generalize this approach to estimate the beta of one portfolio x with respect to another portfolio w. We have:

$$R_t(x) = \alpha + \beta R_t(w) + \varepsilon_t$$

Another way to compute the beta is to use the following relationship:

$$\beta(x \mid w) = \frac{\sigma(x, w)}{\sigma^2(w)} = \frac{x^\top \Sigma w}{w^\top \Sigma w}$$

We deduce that the expression of the beta of asset i is also:

$$\beta_i = \beta(\mathbf{e}_i \mid w) = \frac{\mathbf{e}_i^\top \Sigma w}{w^\top \Sigma w} = \frac{(\Sigma w)_i}{w^\top \Sigma w}$$

It follows that the beta of a portfolio is the weighted average of the beta of the assets that compose the portfolio:

$$\beta(x \mid w) = \frac{x^\top \Sigma w}{w^\top \Sigma w} = x^\top \frac{\Sigma w}{w^\top \Sigma w} = \sum_{i=1}^{n} x_i \beta_i$$

The relationship (3.8) is very important and highlights the role of the beta coefficient. However, this result is not the only main finding of Sharpe (1964). In his article, Sharpe showed also that

if the market is at the equilibrium, the prices of assets are such that the tangency portfolio w^* is the market portfolio w_m (or the market-cap portfolio). With this result, the characterization of the tangency portfolio does not depend upon the assumptions about expected returns, volatilities and correlations.

In the case of Example 9, we have seen that the composition of the tangency portfolio is $w^* = (42.57\%, -11.35\%, 9.43\%, 43.05\%, 16.30\%)$. Since its expected return is $\mu(w^*) = 7.51\%$ and $r = 3\%$, we deduce that the market risk premium is equal to 4.51%. In Table 3.2, we report the beta of each asset and two portfolios: the equally weighted (EW) portfolio w_{ew} and the GMV portfolio w_{gmv}. We compute the associated expected return $\mu(w) = w^\top \mu$ and the risk premium explained by the tangency portfolio $\pi(w \mid w^*) = \beta(w \mid w^*)(\mu(w^*) - r)$. We verify the relationship $\pi(w \mid w^*) = \mu(w) - r$. For instance, the beta of the first asset is equal to 0.444 and we have $0.444 \times 4.51\% = 2\%$, which is also equal to the difference between 5% and 3%. For the EW portfolio, the risk premium is equal to $0.932 \times 4.51\% = 4.20\%$. We also verify that it is equal to the difference between the expected return 7.20% and the risk-free rate 3%.

Table 3.2: Computation of the beta and risk premia (Example 9)

Portfolio	$\mu(w)$	$\mu(w) - r$	$\beta(w \mid w^*)$	$\pi(w \mid w^*)$
$\mathbf{e}_1$	5.00%	2.00%	0.444	2.00%
$\mathbf{e}_2$	7.00%	4.00%	0.887	4.00%
$\mathbf{e}_3$	6.00%	3.00%	0.665	3.00%
$\mathbf{e}_4$	10.00%	7.00%	1.553	7.00%
$\mathbf{e}_5$	8.00%	5.00%	1.109	5.00%
w_{ew}	7.20%	4.20%	0.932	4.20%
w_{gmv}	6.69%	3.69%	0.817	3.69%

Risk premium and alpha return   Jensen (1968) analyzed the performance of active management by using the following regression model:

$$R_{j,t} - r = \alpha_j + \beta_j (R_t(w_m) - r) + \varepsilon_{j,t}$$

where $R_{j,t}$ is the return of the mutual fund j at time t, $R_t(w_m)$ is the return of the market portfolio and $\varepsilon_{j,t}$ is an idiosyncratic risk. If the mutual fund outperforms the market portfolio, the assumption $\alpha_j > 0$ is not rejected. However, Jensen rejected this assumption for most mutual funds and concluded that active management did not create alpha. More generally, the alpha is defined by the difference between the risk premium $\pi(w)$ of portfolio w and the beta[8] $\beta(w)$ of the portfolio times the market risk premium π_m:

$$
\begin{aligned}
\alpha &= (\mu(w) - r) - \beta(w \mid w_m)(\mu(w_m) - r) \\
&= \pi(w) - \beta(w)\pi_m
\end{aligned}
$$

If we now impose a no short-selling constraint by using a lower bound $x_i \geq 0$, the tangency portfolio becomes $w^\star = (33.62\%, 0\%, 8.79\%, 40.65\%, 16.95\%)$ in our previous example. We verify that the portfolio has not a short exposure on the second asset. Since we have $\mu(w^\star) = 7.63\%$ and $r = 3\%$, we deduce that the market risk premium is equal to 4.63%. It is higher than in the unconstrained case. We report the beta and the risk premium in Table 3.3. We notice that the equality $\mu(w) - r =$

---

[8]The notation $\beta(w)$ means that the beta is computed with respect to the market portfolio.

Box 3.2: Computing the implied risk premia of investors

Let us consider the optimization problem[a]:

$$w^\star = \arg\min \frac{1}{2} w^\top \Sigma w - \gamma w^\top (\mu - r\mathbf{1})$$
$$\text{s.t.} \quad \begin{cases} \mathbf{1}^\top w = 1 \\ w \in \Omega \end{cases}$$

If we omit the constraints, the solution is $w^\star = \gamma \Sigma^{-1} (\mu - r\mathbf{1})$. In the Markowitz model, the unknown variable is the vector w of weights. We now suppose that the investor has a current asset allocation w_0. By construction, w_0 is the optimal portfolio for this investor, otherwise he will change its investment policy. We deduce that[b]:

$$w_0 = \gamma \Sigma^{-1} (\mu - r\mathbf{1}) \Leftrightarrow \tilde{\pi} = \mu - r\mathbf{1} = \frac{1}{\gamma} \Sigma w_0 \qquad (3.10)$$

We may interpret $\tilde{\pi}$ as the vector of risk premia which is coherent with the portfolio w_0 (Black and Litterman, 1991, 1992). $\tilde{\pi}$ is then the risk premium priced by the portfolio manager. The computation of $\tilde{\pi}$ requires to specify the risk tolerance of the investor. Let us assume that the investor targets a Sharpe ratio $\text{SR} (w_0 \mid r)$ for his portfolio. We deduce that:

$$\text{SR} (w_0 \mid r) = \frac{\mu(w_0) - r}{\sigma(w_0)} = \frac{w_0^\top (\mu - r\mathbf{1})}{\sqrt{w_0^\top \Sigma w_0}} = \frac{1}{\gamma} \sqrt{w_0^\top \Sigma w_0}$$

Finally, we obtain:

$$\tilde{\pi} = \text{SR} (w_0 \mid r) \cdot \frac{\Sigma w_0}{\sqrt{w_0^\top \Sigma w_0}}$$

Let us consider Example 9. We suppose that the current allocation w_0 is equal to $(35\%, 25\%, 15\%, 15\%, 10\%)$. The volatility of the portfolio is then equal to $\sigma(x_0) = 12.52\%$. The objective of the portfolio manager is to target a Sharpe ratio equal to 0.25. The implied risk tolerance is $\gamma = 0.50$ and the implied risk premia are $\tilde{\pi} = (3.36\%, 4.45\%, 2.83\%, 1.59\%, 1.80\%)$.

---

[a]We notice that the excess expected return is equal to $w^\top (\mu - r\mathbf{1}) = w^\top \mu - r$. Adding the risk-free rate has then no impact on the mean-variance utility function.

[b]From this equation, we also deduce the following relationship: $\tilde{\pi}(w_0) = \gamma^{-1} \sigma^2(w_0)$.

$\beta(w)(\mu_m - r)$ is not always satisfied. This is particularly true for the second asset, which has a negative alpha of 49 bps. We know that the true risk premium of this asset is 4%. Nevertheless, investors are constrained and they can not short this asset. From a theoretical point of view, the optimal demand for this asset must be negative. Because of the lower bound $x_i \geq 0$, the market demand is higher than the expected demand deduced from the CAPM. Therefore, there is a price pressure on this asset due to a lack of arbitrage. The risk premium perceived by the market is then higher, creating a negative alpha because the price is overvalued. As the equally-weighted portfolio is long on this asset, it has also a negative alpha. This is not the case of the GMV portfolio, which is short on this asset (its weight is equal to -28.52% — see Table 3.1 on page 134).

Table 3.3: Computation of the alpha return (Example 9)

Portfolio	$\mu(w)$	$\mu(w) - r$	$\beta(w \mid w^*)$	$\pi(w \mid w^*)$	$\alpha(w \mid w^*)$
$\mathbf{e}_1$	5.00%	2.00%	0.432	2.00%	0.00%
$\mathbf{e}_2$	7.00%	4.00%	0.970	4.49%	-0.49%
$\mathbf{e}_3$	6.00%	3.00%	0.648	3.00%	0.00%
$\mathbf{e}_4$	10.00%	7.00%	1.512	7.00%	0.00%
$\mathbf{e}_5$	8.00%	5.00%	1.080	5.00%	0.00%
w_{ew}	7.20%	4.20%	0.929	4.30%	-0.10%
w_{gmv}	6.69%	3.69%	0.766	3.55%	0.14%

The previous analysis can be applied to a more general framework. There are two main explanations of alpha generation. The first one concerns the assumptions of the CAPM. In particular, this model assumes that investors face no constraints in terms of leverage, short selling, transaction costs, etc. In practice, investors are highly constrained, especially large institutional investors. Since they can not leverage their portfolios, they do not use the tangency portfolio. They will prefer a portfolio with a lower Sharpe ratio, but with a higher expected return. This explains that the demand for high beta assets is greater than the demand predicted by the CAPM. Therefore, we observe a positive alpha return for low-beta assets and a negative alpha return for high-beta assets (Black, 1972; Frazzini and Pedersen, 2014). The second explanation is the existence of other risk factors, which are not priced by the CAPM (Ross, 1976). The development of factor investing and alternative risk premia in the aftermath of the 2008 global financial crisis is related to this issue. If investors use systematic strategies with the same approach and these strategies are very popular, they may impact asset prices (Roncalli, 2017). In both cases, alpha generation takes its root in the imbalance between supply and demand and the dynamics of investment flows.

Portfolio optimization in the presence of a benchmark

Utility function revisited   The Markowitz approach for portfolio optimization assumes that the investor has a mean-variance utility function without any reference to a given investment policy. We now extend the optimization problem when a strategic asset allocation imposes a benchmark, which is represented by a portfolio b. The tracking error between the active portfolio w and its benchmark b is the difference between the return of the portfolio and the return of the benchmark:

$$\epsilon = R(w) - R(b) = \sum_{i=1}^{n} w_i R_i - \sum_{i=1}^{n} b_i R_i = w^\top R - b^\top R = (w - b)^\top R$$

The tracking error ϵ is a stochastic random variable. The expected excess return is equal to:

$$\mu(w \mid b) := \mathbb{E}[\epsilon] = (w - b)^\top \mu$$

whereas the volatility of the tracking error is defined as:

$$\sigma(w \mid b) := \sigma(e) = \sqrt{(w - b)^\top \Sigma (w - b)}$$

The objective of the investor is then to maximize the expected tracking error with a constraint on the tracking error volatility:

$$
\begin{aligned}
w^\star \quad &= \quad \arg\max \mu(w \mid b) \\
\text{s.t.} \quad &\begin{cases} \mathbf{1}^\top x = 1 \\ \sigma(w \mid b) \le \sigma^\star \end{cases}
\end{aligned}
\tag{3.11}
$$

Like the Markowitz problem, we transform this σ-problem into a γ-problem:

$$w^\star (\gamma) = \arg\min f (w \mid b)$$

where:

$$
\begin{aligned}
f (w \mid b) &= \frac{1}{2}\sigma^2 (w \mid b) - \gamma\mu (w \mid b) \\
&= \frac{1}{2} (w - b)^\top \Sigma (w - b) - \gamma (w - b)^\top \mu \\
&= \frac{1}{2}w^\top \Sigma w - w^\top (\gamma\mu + \Sigma b) + \underbrace{\frac{1}{2}b^\top \Sigma b + \gamma b^\top \mu}_{\text{constant}}
\end{aligned}
$$

Again, we recognize a quadratic programming problem. The efficient frontier is then the parametric curve $(\sigma (w^\star (\gamma) \mid b), \mu (w^\star (\gamma) \mid b))$ with $\gamma \geq 0$.

Remark 30 *Using Equation (3.10), we notice that $w^\top (\gamma\mu + \Sigma b) = 2\gamma w^\top \left(\dfrac{\pi + \tilde{\pi}}{2} \right) - \gamma$ where $\tilde{\pi}$ is the implied risk premia associated to the benchmark b. We obtain a Markowitz problem where the vector of expected returns is replaced by an average between the true and implied risk premia.*

Example 10 *We consider an investment universe of four assets. Their expected returns are equal to 5%, 6.5%, 8% and 6.5% while their volatilities are equal to 15%, 20%, 25% and 30%. The correlation matrix of asset returns is given by the following matrix:*

$$
\mathbb{C} = \begin{pmatrix}
100\% & & & \\
10\% & 100\% & & \\
40\% & 70\% & 100\% & \\
50\% & 40\% & 80\% & 100\%
\end{pmatrix}
$$

We consider Example 10 with the benchmark $b = (60\%, 40\%, 20\%, -20\%)$. In Figure 3.4, we have represented the corresponding efficient frontier. We verify that it is a straight line when there is no restriction (Roll, 1992). If we impose that $w_i \geq -10\%$, the efficient frontier is moved to the right. For the third case, we assume that the weights are between a lower bound and an upper bound: $w_i^- \leq w_i \leq w_i^+$ with $w_i^+ = 50\%$. For the first three assets, the lower bound w_i^- is set to 0, whereas it is equal to -20% for the fourth asset.

Information ratio To compare the performance of different portfolios, a better measure than the Sharpe ratio is the information ratio which is defined as follows:

$$\text{IR} (w \mid b) = \frac{\mu (w \mid b)}{\sigma (w \mid b)} = \frac{(w - b)^\top \mu}{\sqrt{(w - b)^\top \Sigma (w - b)}}$$

If we consider a combination of the benchmark b and the active portfolio w, the composition of the portfolio is:

$$x = (1 - \alpha) b + \alpha w$$

where $\alpha \geq 0$ is the proportion of wealth invested in the portfolio w. It follows that:

$$\mu (x \mid b) = (x - b)^\top \mu = \alpha\mu (w \mid b)$$

Figure 3.4: Efficient frontier with a benchmark (Example 10)
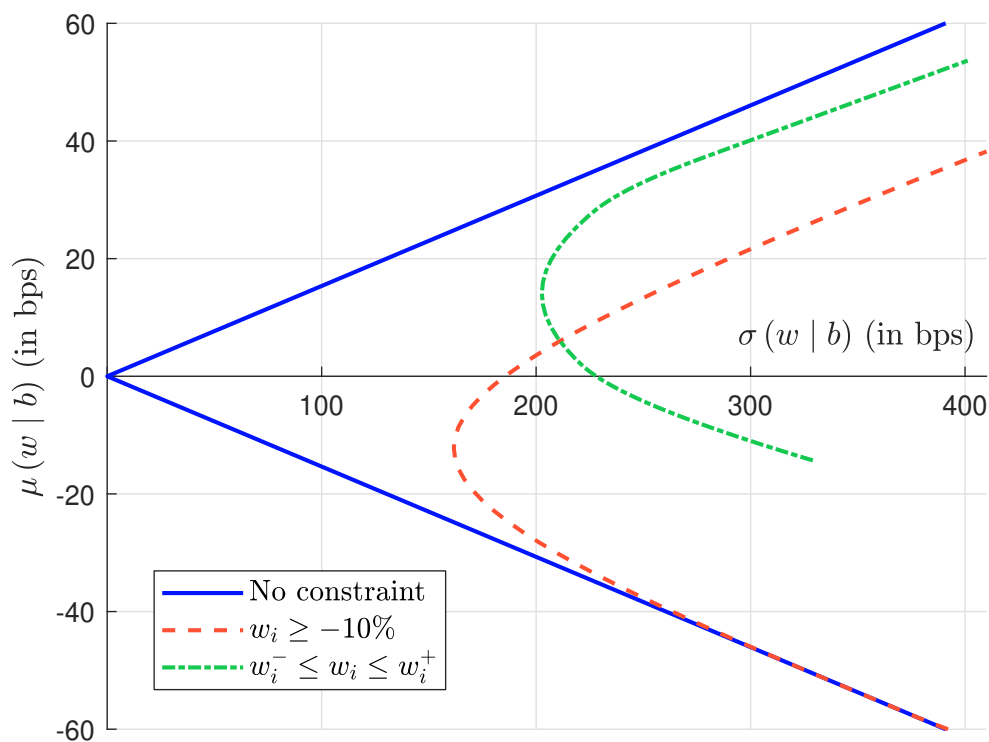
Figure 3.5: Tangency portfolio with respect to a benchmark (Example 10)
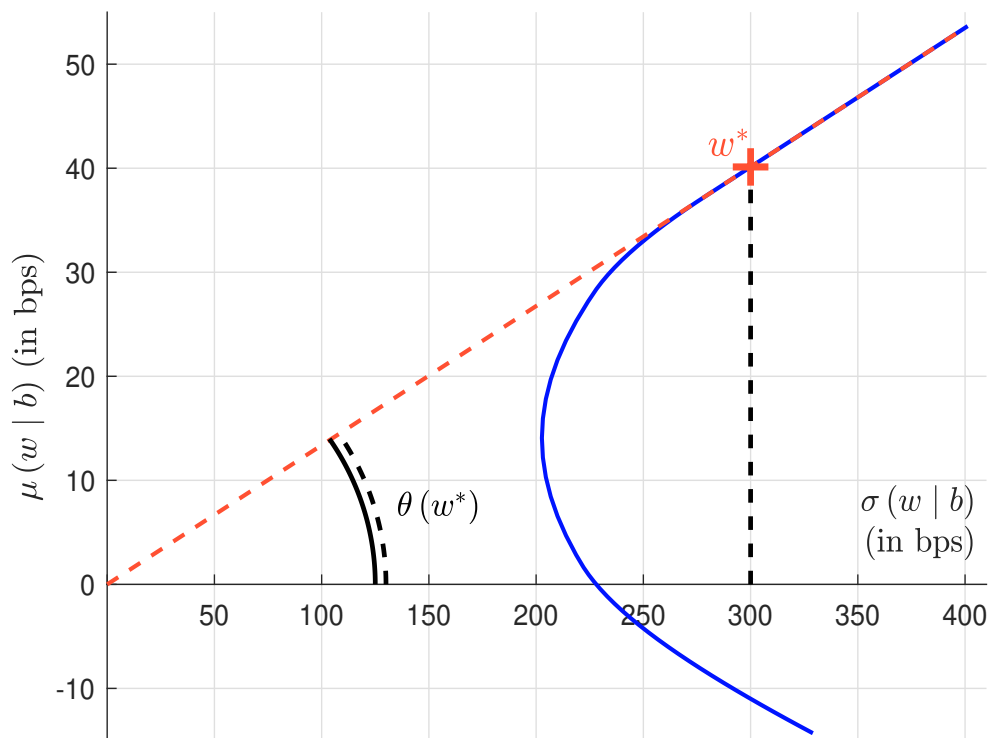
and:

$$\sigma^2 \left(x \mid b\right) = \left(x - b\right)^\top \Sigma \left(x - b\right) = \alpha^2 \sigma^2 \left(w \mid b\right)$$

We deduce that:

$$\mu \left(x \mid b\right) = \mathrm{IR} \left(w \mid b\right) \cdot \sigma \left(x \mid b\right)$$

It is the equation of a linear function between the tracking error volatility and the expected tracking error of the portfolio x. It implies that the efficient frontier is a straight line:

> "*If the manager is measured solely in terms of excess return performance, he or she should pick a point on the upper part of this efficient frontier. For instance, the manager may have a utility function that balances expected value added against tracking error volatility. Note that because the efficient set consists of a straight line, the maximal Sharpe ratio is not a usable criterion for portfolio allocation*" (Jorion, 2003, page 172).

If we add some other constraints to the portfolio optimization problem (3.11), the efficient frontier is no longer a straight line. In this case, one optimized portfolio dominates all the other portfolios. It is the portfolio which belongs to the efficient frontier and the straight line which is tangent to the efficient frontier. It is also the portfolio which maximizes the information ratio.

Let us look at the previous efficient frontier when we impose lower and upper bounds (third case). When we combine it with the benchmark, we obtain the straight line produced in Figure 3.5 and the tangency portfolio is equal to $(46.56\%, 33.49\%, 39.95\%, -20.00\%)$.

3.1.2 ESG risk premium

We now analyze the impact of ESG investing in the CAPM. However, it is important to reiterate that the risk premium is the expected excess return earned by investors because they are exposed to a systematic risk. Therefore, we must differentiate between expected (or required) returns and historical (or realised) returns. Moreover, it is not very clear whether the risk premium is a specific requirement from investors or the long-term performance. This is the difference between the unconstrained risk premium π_i and the implied risk premium $\tilde{\pi}_i$.

The Pastor-Stambaugh-Taylor model

In this section, we present the model developed by Pástor *et al.* (2021) (hereafter, PST model). It is a direct extension of the CAPM and has the advantage to highlights many intuitive stylized facts.

Model settings Pástor *et al.* (2021) consider an investment universe of n assets corresponding to the shares of n firms. They assume that the asset excess returns $\tilde{R} = R - r = \left(\tilde{R}_1, \ldots, \tilde{R}_n\right)$ are normally distributed — $\tilde{R} \sim \mathcal{N}\left(\pi, \Sigma\right)$, and the firms produce social impact. Each firm has an ESG characteristic $\mathcal{G}_i$, which is positive for *esg-friendly* (or *green*) firms and negative for *esg-unfriendly* (or *brown*) firms. This means that $\mathcal{G}_i > 0$ induces positive social impact, while $\mathcal{G}_i < 0$ induces negative externalities on the society. They consider an economy with a continuum of agents ($j = 1, 2, \ldots, \infty$). We note $w_{i,j}$ the fraction of the wealth invested by agent j in stock i, and $w_j = (w_{1,j}, \ldots, w_{n,j})$ the allocation vector of agent j. The relationship between the initial and terminal wealth W_j and $\tilde{W}_j$ is given by:

$$\tilde{W}_j = \left(1 + r + w_j^\top \tilde{R}\right) W_j$$

Pástor *et al.* (2021) assume that the economic agent j has an exponential CARA utility function:

$$\boldsymbol{\mathcal{U}}\left(\tilde{W}_j, w_j\right) = -\exp\left(-\bar{\gamma}_j \tilde{W}_j - w_j^\top b_j W_j\right)$$

where $\bar{\gamma}_j$ is the absolute risk-aversion and $b_j = \varphi_j \boldsymbol{\mathcal{G}}$ is the vector of nonpecuniary benefits that depends on the green intensity $\boldsymbol{\mathcal{G}}$ and the ESG preference coefficient $\varphi_j \geq 0$ of the economic agent.

Optimal portfolio   The expected utility is equal to:

$$
\begin{aligned}
\mathbb{E}\left[\boldsymbol{\mathcal{U}}\left(\tilde{W}_j, w_j\right)\right] &= \mathbb{E}\left[-\exp\left(-\bar{\gamma}_j \tilde{W}_j - w_j^\top b_j W_j\right)\right] \\
&= \mathbb{E}\left[-\exp\left(-\bar{\gamma}_j\left(1 + r + w_j^\top \tilde{R}\right)W_j - w_j^\top b_j W_j\right)\right] \\
&= -e^{-\bar{\gamma}_j(1+r)W_j}\mathbb{E}\left[\exp\left(-\bar{\gamma}_j w_j^\top W_j\left(\tilde{R} + \bar{\gamma}_j^{-1}b_j\right)\right)\right] \\
&= e^{-\bar{\Gamma}_j(1+r)}\mathbb{E}\left[\exp\left(-\bar{\Gamma}_j w_j^\top\left(\tilde{R} + \bar{\gamma}_j^{-1}b_j\right)\right)\right]
\end{aligned}
\tag{3.12}
$$

where $\bar{\Gamma}_j = \bar{\gamma}_j W_j$ is the nominal risk aversion. We notice that $\tilde{R} + \bar{\gamma}_j^{-1}b_j \sim \mathcal{N}\left(\pi + \bar{\gamma}_j^{-1}b_j, \Sigma\right)$ and:

$$
-\bar{\Gamma}_j w_j^\top\left(\tilde{R} + \bar{\Gamma}_j^{-1}b_j\right) \sim \mathcal{N}\left(-\bar{\Gamma}_j w_j^\top\left(\pi + \bar{\gamma}_j^{-1}b_j\right), \bar{\Gamma}_j^2 w_j^\top \Sigma w_j\right)
$$

Using the mathematical expectation formula of the log-normal distribution[9], we deduce that:

$$
\mathbb{E}\left[\boldsymbol{\mathcal{U}}\left(\tilde{W}_j, w_j\right)\right] = e^{-\bar{\Gamma}_j(1+r)}\exp\left(-\bar{\Gamma}_j w_j^\top\left(\pi + \bar{\gamma}_j^{-1}b_j\right) + \frac{1}{2}\bar{\Gamma}_j^2 w_j^\top \Sigma w_j\right)
$$

The first-order condition is equal to:

$$
-\bar{\Gamma}_j\left(\pi + \bar{\gamma}_j^{-1}b_j\right) + \bar{\Gamma}_j^2 \Sigma w_j = 0
$$

Finally, Pástor *et al.* (2021) conclude that the optimal portfolio is:

$$
w_j^\star = \Gamma_j \Sigma^{-1}\left(\pi + \gamma_j b_j\right)
\tag{3.13}
$$

where $\Gamma_j = \bar{\Gamma}_j^{-1}$ and $\gamma_j = \bar{\gamma}_j^{-1}$ are the relative nominal and unitary risk-tolerance coefficients. This is the unconstrained optimal portfolio where asset returns include the green sentiment $\gamma_j b_j = \gamma_j \varphi_j \boldsymbol{\mathcal{G}}$.

Remark 31 *We assume that $W_j = 1$. Since we have $\mathbf{1}^\top w_j = 1$, $w_j^\top r = r$ and $\bar{\Gamma}_j = \bar{\gamma}_j$, we deduce that:*

$$
\begin{aligned}
-\ln\mathbb{E}\left[\boldsymbol{\mathcal{U}}\left(\tilde{W}_j, w_j\right)\right] &= \bar{\gamma}_j\left(1 + r\right) + \bar{\gamma}_j w_j^\top\left(\pi + \bar{\gamma}_j^{-1}b_j\right) - \frac{1}{2}\bar{\gamma}_j^2 w_j^\top \Sigma w_j \\
&\propto w_j^\top\left(\pi + r\mathbf{1} + \bar{\gamma}_j^{-1}b_j\right) - \frac{1}{2}\bar{\gamma}_j w_j^\top \Sigma w_j \\
&= w_j^\top\left(\mu + \bar{\gamma}_j^{-1}b_j\right) - \frac{1}{2}\bar{\gamma}_j w_j^\top \Sigma w_j
\end{aligned}
$$

Maximizing the expected utility is then equivalent to solve the classical Markowitz QP problem:

$$
\begin{aligned}
w_j^\star\left(\gamma_j\right) &= \arg\min \frac{1}{2}w_j^\top \Sigma w_j - \gamma_j w_j^\top \mu' \\
&s.t. \quad \mathbf{1}^\top w_j = 1
\end{aligned}
$$

where $\gamma_j = \bar{\gamma}_j^{-1}$ is the relative risk tolerance and $\mu' = \mu + \gamma_j b_j$ is the vector of modified expected returns that takes into account the ESG sentiment of the economic agent concerning the social impact of firms.

---

[9]See Appendix A.2.1 on page 660.

Example 11 *We consider a universe of n risky assets, where n is an even number. The risk-free rate r is set to 3%. We assume that the Sharpe ratio of these assets is the same and is equal to 20%. The volatility of asset i is equal to $\sigma_i = 0.10 + 0.20 \cdot e^{-n^{-1}\lfloor 0.5i \rfloor}$. The correlation between asset returns is constant: $\mathbb{C} = \mathbb{C}_n(\rho)$. The social impact of the firms is given by the vector $\mathcal{G}$. When $\mathcal{G}$ is not specified, it is equal to the cyclic vector $(+1\%, -1\%, +1\%, \ldots, +1\%, -1\%)$. This implies that half of the firms (green firms) have a positive social impact while the others (brown firms) have a negative impact.*

We consider the case $n = 6$ and $\rho = 25\%$, and we assume that we can not be short on the assets. We calibrate the risk-tolerance parameter γ such that the long-only optimized portfolio of the non-ESG investor has a volatility of 20%. We find $\gamma = 1.5456$ and obtain the results reported in Table 3.4. We verify that the optimized portfolio depends on the ESG preference coefficient φ. We consider a second set of ESG characteristics: $\mathcal{G} = (10\%, 5\%, 2\%, 3\%, 25\%, 30\%)$. Since $\mathcal{G}_i > 0$, we can consider that this investment universe has been filtered in order to keep only the best-in-class issuers and implement an ESG selection strategy. Again, we measure the impact of φ on the optimized portfolios. In Figure 3.6, we report the efficient frontier when the investment universe is made up of 20 assets. We verify that the expected returns of the efficient frontier are reduced when considering ESG preferences, and this reduction depends on the ESG preference coefficient φ. We also notice that all these efficient frontiers start at the same point since the global minimum variance portfolio is not affected by the ESG taste of the investor.

Table 3.4: Mean-variance optimized portfolios with ESG preferences (Example 11, $n = 6$, $\rho = 25\%$)

	$\mathcal{G} = (1\%, -1\%, 1\%, -1\%, 1\%, -1\%)$				$\mathcal{G} = (10\%, 5\%, 2\%, 3\%, 25\%, 30\%)$			
φ	0.00%	1.00%	5.00%	50.00%	0.00%	0.50%	1.00%	2.00%
$w_1^\star$	44.97%	48.87%	58.65%	67.48%	44.97%	46.83%	28.69%	0.00%
$w_2^\star$	44.97%	41.06%	19.60%	0.00%	44.97%	37.06%	9.17%	0.00%
$w_3^\star$	5.03%	9.82%	21.75%	32.52%	5.03%	0.00%	0.00%	0.00%
$w_4^\star$	5.03%	0.25%	0.00%	0.00%	5.03%	0.00%	0.00%	0.00%
$w_5^\star$	0.00%	0.00%	0.00%	0.00%	0.00%	0.83%	16.62%	21.09%
$w_6^\star$	0.00%	0.00%	0.00%	0.00%	0.00%	15.28%	45.53%	78.91%
$\mu(w^\star)$	8.33%	8.33%	8.27%	8.22%	8.33%	8.23%	7.79%	7.43%
$\sigma(w^\star)$	20.00%	20.09%	20.07%	21.56%	20.00%	19.33%	16.70%	19.17%
$\text{SR}(w^\star \mid r)$	0.27	0.27	0.26	0.24	0.27	0.27	0.29	0.23

Remark 32 *In this numerical example, the impact of ESG preferences is low because the assets have similar financial characteristics: same Sharpe ratio and same cross-correlation values. This explains that the optimized portfolios are different, but their Sharpe ratios are very close. Nevertheless, the expected return is always lower when implementing an ESG strategy[10].*

Risk premium The market total wealth W is equal to $\int W_j \, dj$. Let $\omega_j = W_j/W$ be the market share of the economic agent j. His amount $W_{i,j}$ invested in stock i is equal to $W_{i,j} = w_{i,j}^\star W_j = w_{i,j}^\star \omega_j W$. The total dollar amount invested in stock i is then equal to

$$W_i = \int_j W_{i,j} \, dj = \int_j w_{i,j}^\star \omega_j W \, dj$$

---

[10]In what follows, we are seeing that it is not always the case since it depends on the sign of the aggregate ESG preference $w_m^\top \mathcal{G}$ where w_m is the market portfolio.

Figure 3.6: Efficient frontier with ESG preferences (Example 11, $n = 20$, $\rho = 25\%$)
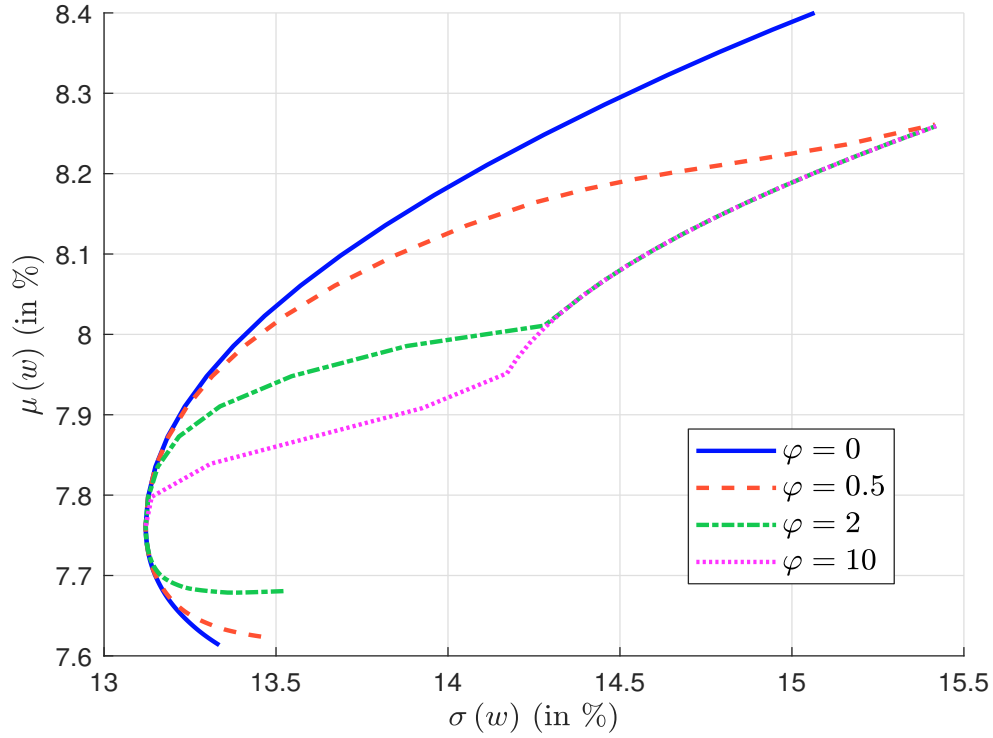
Let $w_m = (w_{1,m}, \ldots, w_{n,m})$ be the market portfolio. We have:

$$w_{i,m} = \frac{W_i}{W} = \int_j w_{i,j}^{\star} \omega_j \, \mathrm{d}j$$

and $\int_j \omega_j \, \mathrm{d}j = 1$. Pástor *et al.* (2021) deduce that the market clearing condition satisfies:

$$
\begin{aligned}
w_m &= \int_j \omega_j w_j^{\star} \, \mathrm{d}j \\
&= \int_j \omega_j \Gamma_j \Sigma^{-1} \left(\pi + \gamma_j b_j \right) \, \mathrm{d}j \\
&= \int_j \omega_j \Gamma_j \Sigma^{-1} \left(\pi + \gamma_j \varphi_j \boldsymbol{\mathcal{G}} \right) \, \mathrm{d}j \\
&= \left(\int_j \Gamma_j \omega_j \, \mathrm{d}j \right) \Sigma^{-1} \pi + \left(\int_j \omega_j \Gamma_j \psi_j \, \mathrm{d}j \right) \Sigma^{-1} \boldsymbol{\mathcal{G}}
\end{aligned}
$$

where $\psi_j = \gamma_j \varphi_j$. Let $\Gamma_m = \int_j \Gamma_j \omega_j \, \mathrm{d}j$ and $\psi_m = \Gamma_m^{-1} \left(\int_j \omega_j \Gamma_j \psi_j \, \mathrm{d}j \right)$ be the average risk tolerance and the weighted average of ESG preferences. The expression of the market portfolio is then equal to:

$$w_m = \Gamma_m \Sigma^{-1} \pi + \Gamma_m \psi_m \Sigma^{-1} \boldsymbol{\mathcal{G}}$$

We deduce that the asset risk premia are equal to:

$$\pi = \frac{1}{\Gamma_m} \Sigma w_m - \psi_m \boldsymbol{\mathcal{G}}$$

while the market risk premium is defined as:

$$
\begin{aligned}
\pi_m &= w_m^\top \pi \\
&= \frac{1}{\Gamma_m} w_m^\top \Sigma w_m - \psi_m w_m^\top \boldsymbol{\mathcal{G}} \\
&= \frac{1}{\Gamma_m} \sigma_m^2 - \psi_m \boldsymbol{\mathcal{G}}_m
\end{aligned}
$$

where $\sigma_m = \sqrt{w_m^\top \Sigma w_m}$ and $\boldsymbol{\mathcal{G}}_m = w_m^\top \boldsymbol{\mathcal{G}}$ are the volatility and the green intensity (or greenness) of the market portfolio. On page 141 (see footnote b.), we have seen that $\pi_m^{\mathrm{capm}} = \Gamma_m^{-1} \sigma_m^2$ is the risk premium deduced from the CAPM. Since $\Gamma_m \geq 0$ and $\psi_m \geq 0$, Pástor *et al.* (2021) notice that:

- The risk premium including the ESG sentiment is lower than the CAPM risk premium if the market ESG intensity is positive:

$$
\boldsymbol{\mathcal{G}}_m > 0 \Longrightarrow \pi_m \leq \pi_m^{\mathrm{capm}}
$$

- It is greater than the CAPM risk premium if the market ESG intensity is negative:

$$
\boldsymbol{\mathcal{G}}_m < 0 \Longrightarrow \pi_m \geq \pi_m^{\mathrm{capm}}
$$

- The gap $\Delta \pi_m^{\mathrm{esg}} := |\pi_m - \pi_m^{\mathrm{capm}}|$ is an increasing function of the market ESG sentiment ψ_m:

$$
\psi_m \nearrow \Longrightarrow \Delta \pi_m^{\mathrm{esg}} \nearrow
$$

If we assume that $\boldsymbol{\mathcal{G}}_m \approx 0$, we have $\Gamma_m = \sigma_m^2 / \pi_m$ and:

$$
\pi = \beta \pi_m - \psi_m \boldsymbol{\mathcal{G}} \tag{3.14}
$$

because $\beta(w_m) = \left(w_m^\top \Sigma w_m\right)^{-1} \Sigma w_m$ is the vector of asset betas with respect to the market portfolio. This is the most important result of Pástor *et al.* (2021). It follows that the alpha of asset i is equal to:

$$
\alpha_i = \pi_i - \beta_i \pi_m = -\psi_m \boldsymbol{\mathcal{G}}_i
$$

Pástor *et al.* (2021) conclude that if $\psi_m > 0$, "*green stocks have negative alphas, and brown stocks have positive alphas. Moreover, greener stocks have lower alphas*".

Example 12 *We consider Example 11. The market is made up of two long-only investors ($j = 1, 2$): a non-ESG investor ($\varphi_1 = 0$) and an ESG investor ($\varphi_2 > 0$). We assume that they have the same risk tolerance γ. We note W_1 and W_2 their financial wealth, which is entirely invested in the risky assets. We assume that $W_1 = W_2 = 1$.*

If the market is at the equilibrium, we have to compute the market portfolio. If there is no short-selling constraint, we have seen that the weights of the tangency portfolio are equal to:

$$
w^* = \frac{\Sigma^{-1} (\mu - r\mathbf{1})}{\mathbf{1}^\top \Sigma^{-1} (\mu - r\mathbf{1})}
$$

We obtain $w^* = (15.04\%, 15.04\%, 16.65\%, 16.65\%, 18.31\%, 18.31\%)$. Since there is no short position, this is the market portfolio[11] without ESG preferences. It follows that the optimal portfolio $w_1^\star$ of

---

[11]Otherwise, we have to solve the Markowitz QP problem subject to the constraints $\mathbf{1}^\top w = 1$ and $w_i \geq 0$.

the first economic agent is equal to w^*. Then, we deduce the risk-tolerance coefficient of this agent and find:

$$\gamma_1 = \frac{1}{\mathbf{1}^\top \Sigma^{-1} (\mu - r\mathbf{1})} = 0.4558$$

We can now compute the optimal portfolio of the second economic agent by assuming that $\gamma_2 = \gamma_1$ and considering the following optimization problem:

$$
\begin{aligned}
w_2^\star \quad &= \quad \arg\min \frac{1}{2} w^\top \Sigma w - \gamma_2 w^\top (\mu + \gamma_2 \varphi_2 \boldsymbol{\mathcal{G}}) \\
\text{s.t.} \quad &\left\{ \begin{array}{l} \mathbf{1}^\top w = 1 \\ w \geq \mathbf{0} \end{array} \right.
\end{aligned}
$$

It is important to use the QP program and not the analytical formula, because the ESG-tilted returns $\mu' = \mu + \gamma_2 \varphi_2 \boldsymbol{\mathcal{G}}$ may be very different from the asset returns μ. In this example, the long-only market portfolio is equal to the long/short tangency portfolio because we consider a uniform correlation of 25% and a constant Sharpe ratio of 20%. The ESG preference φ_2 combined with the greenness vector $\boldsymbol{\mathcal{G}}$ may dramatically change the Sharpe ratio of the assets when it is computed with the ESG-tilted returns. We obtain $w_2^\star = (18.86\%, 11.22\%, 21.33\%, 11.97\%, 23.96\%, 12.65\%)$. The market portfolio is then equal to:

$$
\begin{aligned}
w_m \quad &= \quad \frac{W_1}{W} w_1^\star + \frac{W_2}{W} w_2^\star \\
&= \quad (1 - \omega^{\mathrm{esg}}) \cdot w_1^\star + \omega^{\mathrm{esg}} \cdot w_2^\star
\end{aligned}
$$

where $W = W_1 + W_2$ and ω^{esg} is the wealth share of ESG investors. When $W_1 = W_2 = 1$, we obtain $w_m = (16.95\%, 13.13\%, 18.99\%, 14.31\%, 21.13\%, 15.48\%)$, $\mu_m = 7.86\%$ and $\sigma_m = 14.93\%$. It follows that the beta values are equal to $\beta = (1.15, 1.05, 1.04, 0.95, 0.95, 0.86)$. We deduce that the risk premia are $\pi = (5.58\%, 5.12\%, 5.06\%, 4.61\%, 4.62\%, 4.17\%)$. Finally, we conclude that the alpha vector expressed in bps is $\alpha = (-19.09, 26.19, -19.43, 25.84, -19.72, 25.55)$. A summary of these results is given in Table 3.5.

Table 3.5: Computation of alpha returns (Example 12, $n = 6$, $\rho = 25\%$)

i	Portfolio $w_1^\star$			Portfolio $w_2^\star$				Portfolio w_m			
	w_i (in %)	β_i	π_i (in %)	w_i (in %)	β_i	π_i (in %)	α_i (in bps)	w_i (in %)	β_i	π_i (in %)	α_i (in bps)
1	15.04	1.11	5.39	18.86	1.17	5.69	-30	16.95	1.15	5.58	-19
2	15.04	1.11	5.39	11.22	0.99	4.80	58	13.13	1.05	5.12	26
3	16.65	1.00	4.87	21.33	1.07	5.18	-32	18.99	1.04	5.06	-19
4	16.65	1.00	4.87	11.97	0.88	4.30	57	14.31	0.95	4.61	26
5	18.31	0.91	4.43	23.96	0.98	4.76	-33	21.13	0.95	4.62	-20
6	18.31	0.91	4.43	12.65	0.80	3.87	56	15.48	0.86	4.17	26

Remark 33 *In Figure 3.7, we show the evolution of the alpha return α_i with respect to the market share ω^{esg} of ESG investors. It increases in absolute value because the deviation of the market portfolio including ESG preferences increases with ω^{esg}. We notice that $\alpha_1 \approx \alpha_3 \approx \alpha_5$ and $\alpha_2 \approx \alpha_4 \approx \alpha_6$ because of the specification of the exercise problem. If the Sharpe ratio of assets is different and the correlation is not uniform, the alpha returns are more diffuse.*

Figure 3.7: Evolution of the alpha return with respect to the market share of ESG investors (Example 12, $n = 6$, $\rho = 25\%$)
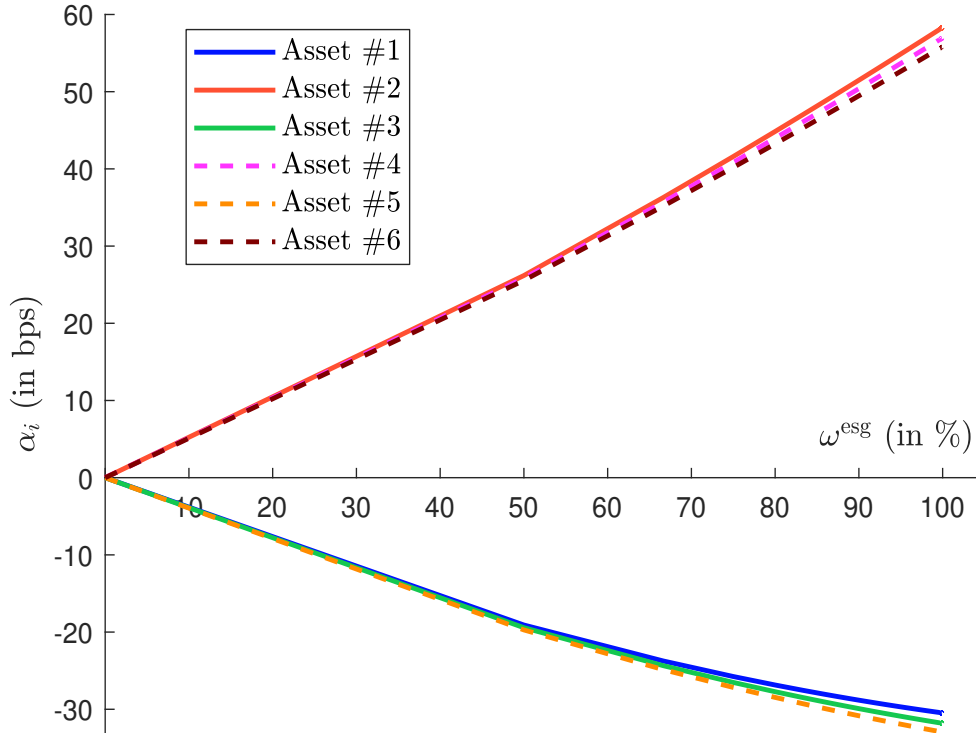
Interpretation of the results   As we have already mentioned, we must differentiate expected returns and realized returns. From a theoretical point of view, there is a scientific consensus that the risk premium of brown assets is positive, implying that the risk premium of green assets is negative (Zerbib, 2019; Ben Slimane *et al.*, 2020; Bolton and Kacperczyk, 2021). This is because there is a systematic market risk when investing in brown assets due to several factors, including carbon pricing, regulation, reputational, asset stranding and climate hedging risks. Moreover, it is obvious that high demand for green assets from ESG investors lowers their expected returns. However, we must be careful because the positive expected excess return of brown assets does not necessarily imply that the performance of green assets is lower than the performance of brown assets:

> "*In equilibrium, green assets have low expected returns because investors enjoy holding them and because green assets hedge climate risk. Green assets nevertheless outperform when positive shocks hit the ESG factor, which captures shifts in customers' tastes for green products and investors' tastes for green holdings.*" (Pástor *et al.*, 2021).

The important word in this quote is *equilibrium*, meaning that green assets have low expected returns in the long run. In this case, investors will need to earn an additional return to compensate for the risk they take when investing in brown assets. In the short term however, when the market is not at equilibrium, green assets can outperform brown assets, in particular when we observe a supply/demand imbalance.

We may wonder what does equilibrium mean? In fact, it refers to a certain long-term period. In order to quantify the long run more precisely, we consider the one-factor risk model:

$$R_i - r = \alpha_i + \beta_i \left(R_m - r \right) + \varepsilon_i$$

where $R_m \sim \mathcal{N}\left(\mu_m, \sigma_m^2\right)$ is the stochastic market return, $\varepsilon_i \sim \mathcal{N}\left(0, \tilde{\sigma}_i^2\right)$ is the idiosyncratic risk and $\varepsilon_i \perp \varepsilon_j$. It follows that (R_i, R_j) follows a bivariate Gaussian distribution:

$$
\begin{pmatrix} R_i \\ R_j \end{pmatrix} = \mathcal{N}\left(\begin{pmatrix} \mu_i \\ \mu_j \end{pmatrix}, \begin{pmatrix} \sigma_i^2 & \sigma_{i,j} \\ \sigma_{i,j} & \sigma_j^2 \end{pmatrix} \right)
$$

where $\mu_i = r + \alpha_i + \beta_i\left(\mu_m - r\right)$, $\sigma_i^2 = \beta_i^2 \sigma_m^2 + \tilde{\sigma}_i^2$ and $\sigma_{i,j} = \beta_i \beta_j \sigma_m^2$. We deduce that $R_i - R_j = \mathcal{N}\left(\mu_{i-j}, \sigma_{i-j}^2\right)$ where:

$$
\mu_{i-j} = (\alpha_i - \alpha_j) + (\beta_i - \beta_j)\left(\mu_m - r\right)
$$

and:

$$
\sigma_{i-j}^2 = (\beta_i - \beta_j)^2\, \sigma_m^2 + \tilde{\sigma}_i^2 + \tilde{\sigma}_j^2
$$

Let us assume that the two assets have the same systematic risk: $\beta_i = \beta_j$. We obtain:

$$
R_i - R_j = \mathcal{N}\left(\alpha_i - \alpha_j, \tilde{\sigma}_i^2 + \tilde{\sigma}_j^2\right)
$$

In the standard CAPM, the alpha returns are equal to zero and we deduce that:

$$
\Pr\left\{R_i < R_j\right\} = \frac{1}{2}
$$

If two assets have the same systematic risk, the probability that one asset underperforms the other is equal to 50%. Let us now take into account the ESG preferences by considering that asset i is the green asset and asset j is the brown asset. The one-year underperformance probability becomes:

$$
p_u\left(\Delta\alpha\right) = \Pr\left\{R_i < R_j\right\} = \Phi\left(\frac{\alpha_j - \alpha_i}{\sqrt{\tilde{\sigma}_i^2 + \tilde{\sigma}_j^2}}\right) > \frac{1}{2}
$$

because we have $\Delta\alpha = \alpha_j - \alpha_i > 0$. We can extend this formula to a greater holding period than one year. If we assume that the dynamics of asset returns are Brownian motions, we obtain:

$$
p_u\left(\Delta\alpha, t\right) = \Phi\left(\frac{(\alpha_j - \alpha_i)\sqrt{t}}{\sqrt{\tilde{\sigma}_i^2 + \tilde{\sigma}_j^2}}\right)
$$

where t is the holding period. Using this formula, we can find the holding period to achieve a given underperformance probability p_u:

$$
t\left(\Delta\alpha, p_u\right) = \frac{\left(\tilde{\sigma}_i^2 + \tilde{\sigma}_j^2\right)}{(\alpha_j - \alpha_i)^2} \Phi^{-1}\left(p_u\right)^2
$$

In Figure 3.8, we report the relationship between $\Delta\alpha$ and the underperformance probability $p_u\left(\Delta\alpha, t\right)$ for several values of the holding period[12]. For plausible values of $\Delta\alpha$ (less than 200 bps), we notice that the probability is lower than 55% for a one-year holding period. It increases until 70% for a ten-year time period, which is not very high. Therefore, it follows that the values of $t\left(\Delta\alpha, p_u\right)$ are very high at the asset level. Let us now consider the same exercise at the portfolio level. We consider an equally-weighted portfolio of 500 green assets and 500 brown assets. Results are given in the bottom/right panel. If the alpha difference is equal to 40 bps, an underperformance probability of 90% is achieved in two years.

---

[12]We assume a typical value of 10% for the idiosyncratic volatility: $\tilde{\sigma}_i = \tilde{\sigma}_j = 10\%$.

Figure 3.8: Impact of alpha returns on the underperformance probability
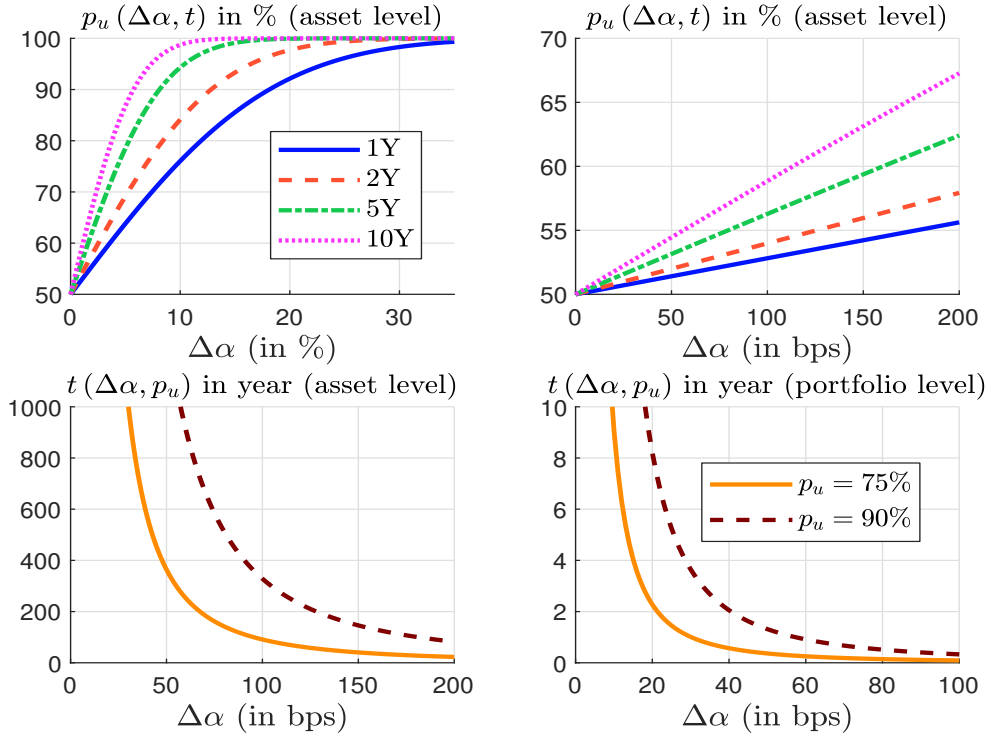
Remark 34 *All these results show that the term equilibrium refers to long holding periods. At the asset level, alpha returns require at least ten years to observe a significant difference. At the portfolio level, a three-year holding period is necessary for the materialization of alpha returns.*

We may wonder whether the PST model considers the ESG risk premium or it is more adapted to assess the green risk premium. Indeed, the sustainable characteristic of the firm i is measured by a non-random metric $\mathcal{G}_i$. For instance, $\mathcal{G}_i$ may correspond to the carbon footprint or the green intensity[13] of the firm. If we apply this model with ESG characteristics, $\mathcal{G}_i$ is the ESG score $\mathcal{S}_i$ of the firm. In this case, assuming that all investors have the same view on the ESG score is a strong hypothesis. In particular, we have already seen that there are a high divergence between ESG scoring models (Berg *et al.*, 2022). In this context, the original formulation of the PST model is certainly more adapted to assess the climate risk premium than the ESG risk premium.

Extension of the model

ESG uncertainty The previous issue has been solved by Avramov *et al.* (2022) (hereafter, ACLT model), who analyze the impact of ESG score uncertainty on the ESG risk premium. For that, they assume that ESG scores are stochastic and may be correlated to asset excess returns:

$$\begin{pmatrix} \tilde{R} \\ \mathcal{S} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \pi \\ \mu_s \end{pmatrix}, \begin{pmatrix} \Sigma & \Sigma_{\pi,s} \\ \Sigma_{s,\pi} & \Sigma_s \end{pmatrix} \right)$$

where $\tilde{R}$ and $\mathcal{S}$ are the random vectors of excess returns and ESG scores. It follows that $\tilde{b}_j = \varphi_j \mathcal{S}$ is stochastic and not constant. Using Equation (3.12), we deduce that the expected utility of the

---

[13]Measured by the green revenue share for instance.

economic agent j is equal to:

$$\mathbb{E}\left[\boldsymbol{\mathcal{U}}\left(\tilde{W}_j, w_j\right)\right] = e^{-\bar{\Gamma}_j(1+r)}\mathbb{E}\left[\exp\left(-\bar{\Gamma}_j w_j^\top \left(\tilde{R} + \psi_j \boldsymbol{\mathcal{S}}\right)\right)\right]$$

where $\psi_j = \gamma_j \varphi_j$. Since we have $\tilde{R} + \psi_j \boldsymbol{\mathcal{S}} \sim \mathcal{N}\left(\breve{\mu}_j, \breve{\Sigma}_j\right)$ where $\breve{\mu}_j = \pi + \psi_j \mu_s$ and $\breve{\Sigma}_j = \Sigma + \psi_j^2 \Sigma_s + 2\psi_j \Sigma_{\pi,s}$, a new expression of the expected utility is:

$$\mathbb{E}\left[\boldsymbol{\mathcal{U}}\left(\tilde{W}_j, w_j\right)\right] = e^{-\bar{\Gamma}_j(1+r)}\exp\left(-\bar{\Gamma}_j w_j^\top \breve{\mu}_j + \frac{1}{2}\bar{\Gamma}_j^2 w_j^\top \breve{\Sigma}_j w_j\right)$$

The first-order condition is equal to $-\bar{\Gamma}_j \breve{\mu}_j + \bar{\Gamma}_j^2 \breve{\Sigma}_j w_j = 0$, implying that the optimal portfolio is:

$$w_j^\star = \Gamma_j \breve{\Sigma}_j^{-1} \tilde{\mu}_j = \Gamma_j \breve{\Sigma}_j^{-1}\left(\pi + \psi_j \mu_s\right)$$

This is exactly the same expression than Equation (3.13) where the asset covariance matrix Σ is replaced by the augmented covariance matrix $\breve{\Sigma}_j$ and the greenness vector $\boldsymbol{\mathcal{G}}$ is equal to the vector μ_s of expected ESG scores. Avramov et al. (2022) introduce the matrix $\Omega_j = \breve{\Sigma}_j^{-1} - \Sigma^{-1}$ and rewrite the optimal solution as follows:

$$w_j^\star = \underbrace{\Gamma_j \Sigma^{-1}\left(\pi + \psi_j \mu_s\right)}_{\text{PST solution}} + \underbrace{\Gamma_j^{-1}\Omega_j\left(\pi + \psi_j \mu_s\right)}_{\text{ESG uncertainty}}$$

Therefore, the optimal portfolio is made up of two components. The first one is the optimal portfolio of the PST model. The second component is another portfolio due to the uncertainty on ESG scores. The two portfolios have the same expression, except that the second portfolio depends on the matrix Ω_j and not the covariance matrix Σ. Using market clearing conditions, Avramov et al. (2022) derive the CAPM-like relationship:

- If there is no ESG uncertainty ($\boldsymbol{\mathcal{S}} = \mu_s$ and $\Sigma_s = \boldsymbol{0}$), the vector of risk premia is given by:

$$\begin{aligned}\pi^{\text{esg}} &= \beta\pi_m - \psi_m\left(\mu_s - \beta\bar{\boldsymbol{\mathcal{S}}}_m\right) \\ &= \pi^{\text{capm}} - \psi_m\left(\mu_s - \beta\bar{\boldsymbol{\mathcal{S}}}_m\right)\end{aligned}$$

where:

$$\begin{cases}\pi_m = \dfrac{1}{\Gamma_m}\sigma_m^2 - \psi_m\bar{\boldsymbol{\mathcal{S}}}_m \\ \sigma_m^2 = w_m^\top \Sigma w_m \\ \beta = \dfrac{\Sigma w_m}{w_m^\top \Sigma w_m} \\ \bar{\boldsymbol{\mathcal{S}}}_m = w_m^\top \mu_s\end{cases}$$

Here, π_m, σ_m and $\bar{\boldsymbol{\mathcal{S}}}_m$ are the risk premium, the volatility and the ESG score of the market portfolio w_m, and β is the vector of beta coefficients. Γ_m and ψ_m are the aggregate risk tolerance and ESG preference across the investors:

$$\begin{cases}\Gamma_m = \dfrac{\int_j \omega_j \Gamma_j \, \mathrm{d}j}{\int_j \omega_j \, \mathrm{d}j} = \int_j \omega_j \Gamma_j \, \mathrm{d}j \\ \psi_m = \dfrac{\int_j \omega_j \Gamma_j \psi_j \, \mathrm{d}j}{\int_j \omega_j \Gamma_j \, \mathrm{d}j} = \dfrac{\int_j \omega_j \Gamma_j \psi_j \, \mathrm{d}j}{\Gamma_m}\end{cases}$$

The authors retrieve the formula (3.14) when the market is ESG neutral[14].

---

[14]Pástor et al. (2021) assume that $\boldsymbol{\mathcal{G}}_m = 0$ and find that $\pi = \beta\pi_m - \psi_m\boldsymbol{\mathcal{G}}$.

- If there is an uncertainty on ESG scores ($\boldsymbol{\mathcal{S}} \neq \mu_s$ and $\Sigma_s \neq \mathbf{0}$), the vector of risk premia becomes:

$$
\begin{aligned}
\breve{\pi}^{\text{esg}} &= \breve{\beta}\breve{\pi}_m - \psi_m \left(\breve{\mu}_s - \breve{\beta}\breve{\boldsymbol{\mathcal{S}}}_m \right) \\
&= \beta\pi_m + \left(\breve{\beta} - \beta \right) \pi_m - \psi_m \left(\breve{\mu}_s - \breve{\beta}\breve{\boldsymbol{\mathcal{S}}}_m \right)
\end{aligned}
$$

where β and ψ_m are the values defined previously and:

$$
\begin{cases}
\breve{\pi}_m = \dfrac{1}{\Gamma_m} \breve{\sigma}_m^2 - \psi_m \breve{\boldsymbol{\mathcal{S}}}_m \\
\breve{\sigma}_m^2 = w_m^\top \breve{\Sigma}_m w_m \\
\breve{\beta} = \dfrac{\breve{\Sigma}_m w_m}{w_m^\top \breve{\Sigma}_m w_m} \\
\breve{\mu}_s = \dfrac{\Psi_m \mu_s}{\psi_m} \\
\breve{\boldsymbol{\mathcal{S}}}_m = w_m^\top \tilde{\mu}_s
\end{cases}
$$

Here, $\breve{\pi}_m$, $\breve{\sigma}_m$ and $\breve{\boldsymbol{\mathcal{S}}}_m$ are the risk premium, the volatility and the ESG score of the market portfolio w_m, $\breve{\beta}$ is the vector of effective beta coefficients, and $\breve{\mu}_s$ is the vector of modified average ESG scores. These quantities depend on $\breve{\Sigma}_m$ and Ψ_m:

$$
\begin{cases}
\breve{\Sigma}_m = \left(\dfrac{\int_j \omega_j \Gamma_j \breve{\Sigma}_j^{-1} \, \mathrm{d}j}{\Gamma_m} \right)^{-1} \\
\Psi_m = \left(\int_j \omega_j \Gamma_j \breve{\Sigma}_j^{-1} \, \mathrm{d}j \right)^{-1} \int_j \omega_j \Gamma_j \breve{\Sigma}_j^{-1} \psi_j \, \mathrm{d}j
\end{cases}
$$

The relationship $\breve{\pi}^{\text{esg}} = \breve{\beta}\breve{\pi}_m - \psi_m \left(\breve{\mu}_s - \breve{\beta}\breve{\boldsymbol{\mathcal{S}}}_m \right)$ obtained with ESG uncertainty is very close to the equilibrium formula $\pi^{\text{esg}} = \beta\pi_m - \psi_m \left(\mu_s - \beta\bar{\boldsymbol{\mathcal{S}}}_m \right)$ obtained without ESG uncertainty. In fact, the ESG uncertainty changes the risk perception of the investors. Therefore, the ESG-tilted covariance matrix $\breve{\Sigma}_j = \Sigma + \psi_j^2 \Sigma_s + 2\psi_j \Sigma_{\pi,s}$ is no longer equal to the asset covariance matrix Σ. It impacts the quantities related to the market portfolio.

In order to better understand the impact of ESG uncertainty on the alpha returns, Avramov et al. (2022) consider the special case in which agents have homogeneous preferences ($\bar{\gamma}_j = \bar{\gamma}$, $\varphi_j = \varphi$) and the same wealth ($W_j = 1$), the covariance matrix of the ESG score is diagonal ($\Sigma_s = \mathrm{diag}\left(\sigma_{1,s}^2 \ldots, \sigma_{n,s}^2 \right)$) and the returns are independent from the ESG scores ($\Sigma_{\pi,s} = \mathbf{0}$). They deduce that $\Gamma_m = \gamma$, $\psi_m = \psi = \gamma\varphi$, $\breve{\Sigma}_m = \Sigma + \psi^2 \Sigma_s$, $\Psi_m = \psi I_n$, $\breve{\sigma}_m^2 = \sigma_m^2 + \psi^2 \sigma_s^2$, $\sigma_m^2 = w_m^\top \Sigma w_m$, $\sigma_s^2 = w_m^\top \Sigma_s w_m$, $\breve{\mu}_s = \mu_s$ and $\breve{\boldsymbol{\mathcal{S}}}_m = w_m^\top \mu_s$. The vector of the effective beta is equal to:

$$
\begin{aligned}
\breve{\beta} &= \frac{\sigma_m^2}{\breve{\sigma}_m^2}\beta + \psi^2 \frac{\sigma_s^2}{\breve{\sigma}_m^2}\beta_s \\
&= \beta + \psi^2 \frac{\sigma_s^2}{\breve{\sigma}_m^2}\left(\beta_s - \beta \right)
\end{aligned}
$$

where:

$$
\beta_s = \frac{\Sigma_s w_m}{\sigma_s^2}
$$

[Avramov et al.](#) ([2022](#)) find that:

$$
\begin{aligned}
\breve{\alpha}^{\mathrm{esg}} &= \breve{\pi}^{\mathrm{esg}} - \beta \pi_m \\
&= \psi^2 \frac{\sigma_s^2}{\breve{\sigma}_m^2} (\beta_s - \beta) \pi_m - \psi_m \left(\mu_s - \left(\beta + \psi^2 \frac{\sigma_s^2}{\breve{\sigma}_m^2} (\beta_s - \beta) \right) \breve{\boldsymbol{S}}_m \right) \\
&= \psi^2 \frac{\sigma_s^2}{\breve{\sigma}_m^2} (\beta_s - \beta) \left(\pi_m + \psi_m \breve{\boldsymbol{S}}_m \right) - \psi_m \left(\mu_s - \beta \breve{\boldsymbol{S}}_m \right)
\end{aligned}
$$

If we consider the asset i, we obtain:

$$
\beta_{i,s} = w_{i,m} \frac{\sigma_{i,s}^2}{\sigma_s^2}
$$

$\beta_{i,s}$ increases with the volatility $\sigma_{i,s}$ of the score $\boldsymbol{S}_i$. We deduce that:

$$
\frac{\partial \breve{\alpha}_i^{\mathrm{esg}}}{\partial \sigma_{i,s}} > 0
$$

This implies that alpha increases with ESG uncertainty. The authors also study the impact of ESG uncertainty on the demand and test the model using the standard deviations from ESG rating agencies as a proxy for ESG uncertainty. Their conclusion is the following:

> "*In equilibrium, the market premium increases and demand for stocks declines under ESG uncertainty. In addition, the CAPM alpha and effective beta both rise with ESG uncertainty and the negative ESG-alpha relation weakens.*" [Avramov et al.](#) ([2022](#), page 642).

Example 13 *We consider an investment universe of four assets. Their expected returns are equal to 5%, 6%, 7% and 8% while their volatilities are equal to 15%, 20%, 30% and 30%. The correlation matrix of asset returns is given by the following matrix:*

$$
\mathbb{C} = \begin{pmatrix}
100\% & & & \\
10\% & 100\% & & \\
40\% & 60\% & 100\% & \\
50\% & 40\% & 80\% & 100\%
\end{pmatrix}
$$

The risk-free return is set to 2%. The average ESG scores are respectively equal to +3%, −2%, +1% and −1%, whereas the standard deviation of ESG score is the same for all the assets and is equal to 20%. We assume that the ESG preference φ of the long-only investor is equal to 0.50 while his risk tolerance corresponds to the market risk tolerance.

We first begin by computing the market risk tolerance $\gamma_m = \left(\mathbf{1}^\top \Sigma^{-1} (\mu - r\mathbf{1}) \right)^{-1} = 0.4654$ and the weights of the market portfolio:

$$
w_m = \gamma_m \Sigma^{-1} (\mu - r\mathbf{1}) = \begin{pmatrix}
50.08\% \\
49.98\% \\
-22.52\% \\
23.47\%
\end{pmatrix}
$$

Then, we consider the following optimization problem:

$$
\begin{aligned}
w^\star (\psi, \mu_s, \Sigma_s) &= \arg\min \frac{1}{2} w^\top \left(\Sigma + \psi^2 \Sigma_s \right) w - \gamma_m (\mu + \psi \mu_s) \\
&\quad s.t. \begin{cases} \mathbf{1}^\top w = 1 \\ w \geq \mathbf{0} \end{cases}
\end{aligned}
$$

where $\Sigma_s = 4\% \times I$, $\mu_s = (3\%, -2\%, 1\%, -1\%)$ and $\psi = \gamma_m \varphi = 0.4954 \times 0.50 = 0.2327$. In order to decompose the risk premium of the long-only ESG investor, we compute the long-only portfolio $w^\star(0, \mathbf{0}_n, \mathbf{0}_{n,n})$, the long-only portfolio without ESG uncertainty $w^\star(\psi, \mu_s, \mathbf{0}_{n,n})$ and the long-only portfolio with ESG uncertainty $w^\star(\psi, \mu_s, \Sigma_s)$. For each portfolio, we compute the beta coefficient $\beta(w^\star(\psi, \mu_s, \Sigma_s) \mid w_m)$ with respect to the market portfolio. We deduce the risk premium $\pi(w^\star(\psi, \mu_s, \Sigma_s)) = \beta(w^\star(\psi, \mu_s, \Sigma_s) \mid w_m) \cdot \pi_m$ where $\pi_m = \mu(x_m) - r$. We obtain the following decomposition:

$$\pi\left(w^\star\left(\psi, \mu_s, \Sigma_s\right)\right) = \underbrace{\pi\left(w_m\right)}_{\text{market risk premium}} +$$
$$\underbrace{\pi\left(w^\star\left(0, \mathbf{0}_n, \mathbf{0}_{n,n}\right)\right) - \pi\left(w_m\right)}_{\text{long-only alpha}} +$$
$$\underbrace{\pi\left(w^\star\left(\psi, \mu_s, \mathbf{0}_{n,n}\right)\right) - \pi\left(w^\star\left(0, \mathbf{0}_n, \mathbf{0}_{n,n}\right)\right)}_{\text{ESG score alpha}} +$$
$$\underbrace{\pi\left(w^\star\left(\psi, \mu_s, \Sigma_s\right)\right) - \pi\left(w^\star\left(\psi, \mu_s, \mathbf{0}_{n,n}\right)\right)}_{\text{ESG uncertainty alpha}}$$

The optimal weights and risk statistics are given in Tables 3.6 and 3.7. We have:

$$\pi\left(w^\star\left(\psi, \mu_s, \Sigma_s\right)\right) = 3.74\% - 0.96 \text{ bps} - 19.80 \text{ bps} + 5.41 \text{ bps} = 3.50\%$$

In this example, the cost of the long-only constraint is -0.96 bps, the ESG score alpha is equal to -19.80 bps whereas the ESG uncertainty alpha is equal to 5.41 bps.

Table 3.6: Weights of optimized portfolios (Example 13)

Asset	w_m	$w^\star(0, \mathbf{0}_n, \mathbf{0}_{n,n})$	$w^\star(\psi, \mu_s, \mathbf{0}_{n,n})$	$w^\star(\psi, \mu_s, \Sigma_s)$
#1	50.08%	52.00%	64.03%	61.87%
#2	48.98%	39.65%	31.51%	32.05%
#3	-22.52%	0.00%	0.00%	0.00%
#4	23.47%	8.35%	4.47%	6.09%

Table 3.7: Risk statistics of optimized portfolios (Example 13)

Portfolio	$\mu(w)$	$\sigma(w)$	$\beta(w \mid w_m)$	$\pi(w)$	$\boldsymbol{\mathcal{S}}(w)$
w_m	5.74%	13.20%	1.0000	3.74%	0.06%
$w^\star(0, \mathbf{0}_n, \mathbf{0}_{n,n})$	5.65%	13.33%	0.9743	3.65%	0.68%
$w^\star(\psi, \mu_s, \mathbf{0}_{n,n})$	5.45%	12.86%	0.9214	3.45%	1.25%
$w^\star(\psi, \mu_s, \Sigma_s)$	5.50%	12.99%	0.9358	3.50%	1.15%

Risk factor model In the capital asset pricing model, the asset return R_i satisfies the one-factor risk model:

$$R_i - r = \alpha_i + \beta_i (R_m - r) + \varepsilon_i$$

where $\alpha_i = 0$, β_i is the CAPM beta of asset i, $R_m \sim \mathcal{N}\left(\mu_m, \sigma_m^2\right)$ is the market return, $\varepsilon_i \sim \mathcal{N}\left(0, \tilde{\sigma}_i^2\right)$ is the residual. Moreover, we have $\varepsilon_i \perp R_m$ and $\varepsilon_i \perp \varepsilon_j$. In matrix form, we obtain:

$$R = r + \beta (R_m - r) + \varepsilon$$

where $R \sim \mathcal{N}(\mu, \Sigma)$, $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n) \sim \mathcal{N}(\mathbf{0}, D)$ and $D = \mathrm{diag}\left(\tilde{\sigma}_1^2, \ldots, \tilde{\sigma}_n^2\right)$. We deduce that:

$$\mu = \mathbb{E}[R] = r + \beta(\mu_m - r)$$

or:

$$\pi = \mu - r = \beta \pi_m$$

It follows that:

$$R - \mu = \beta(R_m - \mu_m) + \varepsilon$$

Therefore, the expression of the covariance matrix is:

$$
\begin{aligned}
\Sigma &= \mathbb{E}\left[(R - \mu)(R - \mu)^{\top}\right] \\
&= \mathbb{E}\left[(\beta(R_m - \mu_m) + \varepsilon)(\beta(R_m - \mu_m) + \varepsilon)^{\top}\right] \\
&= \mathbb{E}\left[\beta(R_m - \mu_m)(R_m - \mu_m)\beta^{\top} + \varepsilon\varepsilon^{\top} + 2\beta(R_m - \mu_m)\varepsilon^{\top}\right] \\
&= \sigma_m^2 \beta\beta^{\top} + D
\end{aligned}
$$

When we introduce ESG preferences, we obtain a two-factor model:

$$R = r + \beta(R_m - r) + \beta_{\mathrm{esg}}R_{\mathrm{esg}} + \varepsilon$$

where $R \sim \mathcal{N}(\mu, \Sigma)$, $R_m \sim \mathcal{N}\left(\mu_m, \sigma_m^2\right)$ is the market return, $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n) \sim \mathcal{N}(\mathbf{0}, D)$ and $D = \mathrm{diag}\left(\tilde{\sigma}_1^2, \ldots, \tilde{\sigma}_n^2\right)$. Here, $R_{\mathrm{esg}} \sim \mathcal{N}\left(\mu_{\mathrm{esg}}, \sigma_{\mathrm{esg}}^2\right)$ is the return of the ESG portfolio $w_{\mathrm{esg}} \propto \Sigma^{-1}\mu_s$, which is a zero-beta strategy(Pástor *et al.*, 2021):

$$\beta^{\top} w_{\mathrm{esg}} = 0$$

This is why we assume that $R_{\mathrm{esg}} \perp R_m$ and $R_{\mathrm{esg}} \perp \varepsilon$. We deduce that:

$$\mu = r + \beta(\mu_m - r) + \beta_{\mathrm{esg}}\mu_{\mathrm{esg}}$$

or:

$$\pi = \beta\pi_m + \beta_{\mathrm{esg}}\mu_{\mathrm{esg}}$$

For the covariance matrix, we obtain:

$$\Sigma = \sigma_m^2 \beta\beta^{\top} + \sigma_{\mathrm{esg}}^2 \beta_{\mathrm{esg}}\beta_{\mathrm{esg}}^{\top} + D$$

Even if the ESG portfolio has a zero-return ($\mu_{\mathrm{esg}} = 0$), we notice that it may have a big impact on the structure of the covariance matrix.

Remark 35 *The CAPM and ESG coefficients β_i and $\beta_{i,\mathrm{esg}}$ do not have the same status. Indeed, we generally assume that $\beta_i \geq 0$. Otherwise, the asset risk premium is negative: $\pi_i = \beta_i \pi_m$. Moreover, the CAPM average beta $\bar{\beta}$ is close to 1 because the beta of the market portfolio is equal to 1:*

$$
\begin{aligned}
\beta(w_m) = 1 \quad &\Leftrightarrow \quad \frac{w_m^{\top}\Sigma w_m}{w_m^{\top}\Sigma w_m} = 1 \\
&\Leftrightarrow \quad w_m^{\top}\frac{\Sigma w_m}{w_m^{\top}\Sigma w_m} = 1 \\
&\Leftrightarrow \quad \sum_{i=1}^{n} w_{i,m}\beta_i = 1
\end{aligned}
$$

In practice, we assume that $\beta_i \in [0,3]$ in the equity market. This is not the case of the ESG beta because the ESG factor is a zero-beta long/short portfolio. This means that $\beta_{i,\mathrm{esg}}$ can be positive and negative. In practice, the ESG factor[15] is build such that $\beta_{i,\mathrm{esg}} \in [-1,1]$.

In order to measure the impact of ESG on the covariance matrix, we compare the one- and two-factor models. The expression of the asset variance are respectively $\sigma_i^2 \left(\mathrm{capm}\right) = \sigma_m^2 \beta_i^2 + \tilde{\sigma}_i^2$ for the one-factor model and $\sigma_i^2 \left(\mathrm{esg}\right) = \sigma_m^2 \beta_i^2 + \sigma_{\mathrm{esg}}^2 \beta_{i,\mathrm{esg}}^2 + \tilde{\sigma}_i^2$ for the two factor model. We deduce that:

$$\sigma_i^2 \left(\mathrm{esg}\right) - \sigma_i^2 \left(\mathrm{capm}\right) = \sigma_{\mathrm{esg}}^2 \beta_{i,\mathrm{esg}}^2 \geq 0$$

From a theoretical point of view, the introduction of the ESG factor increases asset volatilities. The reason lies in the fact that a new risk is priced in by the market. From a practical point of view, the impact may be lower:

$$\sigma_i^2 \left(\mathrm{esg}\right) - \sigma_i^2 \left(\mathrm{capm}\right) \leq \sigma_{\mathrm{esg}}^2 \beta_{i,\mathrm{esg}}^2$$

because the idiosyncratic volatility in the two-factor model may be reduced. Indeed, we can assume that the ESG factor may capture a part of the CAPM residual risk. If we focus on the correlation, we obtain:

$$\rho_{i,j} \left(\mathrm{esg}\right) = \frac{\sigma_m^2 \beta_i \beta_j + \sigma_{\mathrm{esg}}^2 \beta_{i,\mathrm{esg}} \beta_{j,\mathrm{esg}}}{\sigma_i \left(\mathrm{esg}\right) \sigma_j \left(\mathrm{esg}\right)}$$

and:

$$\rho_{i,j} \left(\mathrm{esg}\right) - \rho_{i,j} \left(\mathrm{capm}\right) = \underbrace{\left(\frac{1}{\sigma_i \left(\mathrm{esg}\right) \sigma_j \left(\mathrm{esg}\right)} - \frac{1}{\sigma_i \left(\mathrm{capm}\right) \sigma_j \left(\mathrm{capm}\right)} \right) \sigma_m^2 \beta_i \beta_j}_{\text{negative}} + \underbrace{\frac{\sigma_{\mathrm{esg}}^2 \beta_{i,\mathrm{esg}} \beta_{j,\mathrm{esg}}}{\sigma_i \left(\mathrm{esg}\right) \sigma_j \left(\mathrm{esg}\right)}}_{\text{not signed}}$$

We have two effects:

1. Since asset volatilities increase, the contribution of the CAPM covariance factor $\beta_i \beta_j$ in the two factor model decreases.

2. The second component depends on the sign of the two ESG beta coefficients. If $\beta_{i,\mathrm{esg}}$ and $\beta_{j,\mathrm{esg}}$ are both positive or negative, the contribution is positive, otherwise it is negative. This implies that the ESG factor increases the correlation between ESG-friendly (or green) assets. The correlation also increases between ESG-unfriendly (or brown) assets. However, the correlation is decreases between a green asset and a brown asset.

Example 14 *We consider an investment universe, which is made up of five assets. Their market beta is respectively equal to 0.9, 0.8, 1.2, 0.7 and 1.3 whereas their specific volatility is 4%, 12%, 5%, 8% and 5%. The market portfolio volatility is equal to 25%. Concerning the ESG factor, we have $\sigma_{\mathrm{esg}} = 10\%$ whereas the ESG sensitivity values are set to -0.5, 0.7, 0.2, 0.9 and -0.3.*

The covariance matrices are reported in Tables 3.8 and 3.9. We verify that asset volatilities have increased with the two-factor model: $\sigma_i \left(\mathrm{esg}\right) > \sigma_i \left(\mathrm{capm}\right)$. We notice that most of correlations have decreased except the cross-correlation $\rho_{2,4}$ between the second and fourth assets. These assets

---

[15]Since the ESG factor is long/short, we can always scale its volatility by leveraging or deleveraging.

have the largest ESG sensitivities: $\beta_{2,\text{esg}} = 0.7$ and $\beta_{4,\text{esg}} = 0.9$. The cross-correlation $\rho_{1,5}$ is not reduced although we have $\beta_{1,\text{esg}} = -0.5$ and $\beta_{5,\text{esg}} = -0.3$. In fact, the product $\beta_{1,\text{esg}}\beta_{5,\text{esg}}$ is not enough large to compensate the increase of the volatilities: $\sigma_1 (\text{esg}) - \sigma_1 (\text{capm}) = 54$ bps and $\sigma_5 (\text{esg}) - \sigma_5 (\text{capm}) = 14$ bps.

Table 3.8: CAPM covariance matrix (Example 14)

Asset	σ_i (in %)	$\rho_{i,j}$ (in %)				
#1	22.85	100.00	84.43	97.12	89.54	97.31
#2	23.32	84.43	100.00	84.58	77.99	84.75
#3	30.41	97.12	84.58	100.00	89.71	97.49
#4	19.24	89.54	77.99	89.71	100.00	89.89
#5	32.88	97.31	84.75	97.49	89.89	100.00

Table 3.9: Two-factor covariance matrix (Example 14)

Asset	σ_i (in %)	$\rho_{i,j}$ (in %)				
#1	23.39	100.00	72.85	93.27	70.18	96.61
#2	24.35	72.85	100.00	82.72	79.84	78.23
#3	30.48	93.27	82.72	100.00	83.87	96.28
#4	21.24	70.18	79.84	83.87	100.00	77.24
#5	33.02	96.61	78.23	96.28	77.24	100.00

We consider the portfolio $w = \vartheta \Sigma^{-1} \eta$ where η is a $n \times 1$ vector and $\vartheta = 1/\left(\mathbf{1}^\top \Sigma^{-1} \eta\right)$ is the scalar such that $\mathbf{1}^\top w = 1$. Using results in Box 3.3, we deduce that:

$$
\begin{aligned}
w &= \vartheta D^{-1} \eta - \vartheta M^{-1} \eta \\
&= \vartheta D^{-1} \eta - \vartheta \omega_1 \tilde{\beta}\tilde{\beta}^\top \eta - \vartheta \omega_2 \tilde{\beta}_{\text{esg}} \tilde{\beta}_{\text{esg}}^\top \eta + \vartheta \omega_3 \left(\tilde{\beta}_{\text{esg}} \tilde{\beta}^\top + \tilde{\beta} \tilde{\beta}_{\text{esg}}^\top \right) \eta \\
&= \underbrace{\vartheta \left(D^{-1} - \omega_1 \tilde{\beta}\tilde{\beta}^\top \right) \eta}_{\text{capm}} + \underbrace{\vartheta \left(\omega_3 \tilde{\beta}_{\text{esg}} \tilde{\beta}^\top + \omega_3 \tilde{\beta} \tilde{\beta}_{\text{esg}}^\top - \omega_2 \tilde{\beta}_{\text{esg}} \tilde{\beta}_{\text{esg}}^\top \right) \eta}_{\text{esg}}
\end{aligned}
\tag{3.15}
$$

Therefore, we can derive analytical formulas for GMV ($\eta = \mathbf{1}$), MVO ($\eta = \gamma\mu$) and tangency ($\eta = \mu - r\mathbf{1}$) portfolios. For instance, if we consider the minimum variance portfolio, Roncalli *et al.* (2020) showed that:

$$
\omega_{i,\text{gmv}} = \frac{\sigma^2 (\omega_{\text{gmv}})}{\tilde{\sigma}_i^2} \max \left(1 - \frac{\beta_i}{\beta^\star} - \frac{\beta_{i,\text{esg}}}{\beta_{\text{esg}}^\star}, c \right)
$$

where $c = -\infty$ if there is no constraint and $c = 0$ in the no short-selling case. Since the mean of beta coefficients is close to one, $\beta^\star$ is positive, implying that the asset weight is a decreasing function of the asset beta. Therefore, the minimum variance portfolio is a low-beta strategy. The impact of the ESG factor is more complex because the mean of ESG beta coefficients is close to zero, implying that the threshold $\beta_{\text{esg}}^\star$ can be positive or negative. We conclude that the asset weight can be a decreasing or increasing function of the asset ESG beta.

Example 15 *We consider an investment universe, which is made up of five assets. Their market beta is respectively equal to* 0.7, 0.8, 0.9, 1.2 *and* 1.5 *whereas their specific volatility is* 10%, 8%, 3%, 5% *and* 4%. *The market portfolio volatility is equal to* 25%. *Concerning the ESG factor, we have* $\sigma_{\text{esg}} = 10\%$ *whereas the ESG sensitivity values are set to* -0.5, -0.7, -0.5, 0.9 *and* 1.3.

Box 3.3: One- and two-factor precision matrices

In portfolio optimization, several variables (weights, risk premium, etc.) are expressed with respect to the inverse of the covariance matrix, which is called the precision matrix. In the case of the one-factor model, we apply the Sherman-Morrison-Woodbury formula[a] with $A = D$ and $u = v = \sigma_m\beta$, and we obtain:

$$\Sigma^{-1} = D^{-1} - \frac{\sigma_m^2}{1 + \sigma_m^2\tilde{\beta}^\top\beta}\tilde{\beta}\tilde{\beta}^\top$$

where $\tilde{\beta}_i = \beta_i/\tilde{\sigma}_i^2$.

For the two-factor model, we use the generalized Sherman-Morrison-Woodbury formula with $A = D$, $u_1 = v_1 = \sigma_m\beta$ and $u_2 = v_2 = \sigma_{\text{esg}}\beta_{\text{esg}}$. It follows that the inverse of the covariance matrix is equal to:

$$\Sigma^{-1} = D^{-1} - D^{-1}US^{-1}V^\top D^{-1}$$

where $U = V = \begin{pmatrix} \sigma_m\beta & \sigma_{\text{esg}}\beta_{\text{esg}} \end{pmatrix}$ and:

$$S = \begin{pmatrix} 1 + \sigma_m^2\beta^\top D^{-1}\beta & \sigma_m\sigma_{\text{esg}}\beta^\top D^{-1}\beta_{\text{esg}} \\ \sigma_m\sigma_{\text{esg}}\beta^\top D^{-1}\beta_{\text{esg}} & 1 + \sigma_{\text{esg}}^2\beta_{\text{esg}}^\top D^{-1}\beta_{\text{esg}} \end{pmatrix}$$

Roncalli *et al.* (2020) showed that:

$$\Sigma^{-1} = D^{-1} - M^{-1}$$

where:

$$M^{-1} = \omega_1\tilde{\beta}\tilde{\beta}^\top + \omega_2\tilde{\beta}_{\text{esg}}\tilde{\beta}_{\text{esg}}^\top - \omega_3\left(\tilde{\beta}_{\text{esg}}\tilde{\beta}^\top + \tilde{\beta}\tilde{\beta}_{\text{esg}}^\top\right)$$

and:

$$\begin{cases} \tilde{\beta}_i = \beta_i/\tilde{\sigma}_i^2 \\ \tilde{\beta}_{i,\text{esg}} = \beta_{i,\text{esg}}/\tilde{\sigma}_i^2 \\ \omega_0 = 1 + \sigma_m^2\tilde{\beta}^\top\beta + \sigma_{\text{esg}}^2\tilde{\beta}_{\text{esg}}^\top\beta_{\text{esg}} + \sigma_m^2\sigma_{\text{esg}}^2\left(\left(\tilde{\beta}^\top\beta\right)\left(\tilde{\beta}_{\text{esg}}^\top\beta_{\text{esg}}\right) - \left(\tilde{\beta}^\top\beta_{\text{esg}}\right)^2\right) \\ \omega_1 = \omega_0^{-1}\sigma_m^2\left(1 + \sigma_{\text{esg}}^2\tilde{\beta}_{\text{esg}}^\top\beta_{\text{esg}}\right) \\ \omega_2 = \omega_0^{-1}\sigma_{\text{esg}}^2\left(1 + \sigma_m^2\tilde{\beta}^\top\beta\right) \\ \omega_3 = \omega_0^{-1}\sigma_m^2\sigma_{\text{esg}}^2\left(\tilde{\beta}^\top\beta_{\text{esg}}\right) \end{cases}$$

---

[a]See Appendix A.1.1 on page 646.

In Tables 3.10 and 3.11, we report the weights of the GMV and long-only MV portfolios and compare the allocation between the one- and two-factor models. If we consider Example 14, adding the ESG factor increases (resp. decreases) the weights of assets with negative (resp. positive) values of $\beta_{i,\text{esg}}$. The reason lies in the fact that the threshold $\beta_{\text{esg}}^\star$ is positive. In the case of Example 15, $\beta_{\text{esg}}^\star$ is equal to -3.5677 for the GMV portfolio and -7.5752 for the long-only MV portfolio. The relationship between $\beta_{i,\text{esg}}$ and $\omega_{i,\text{gmv}}$ becomes more complex. Indeed, the long exposure condition is $\beta_i/\beta^\star + \beta_{i,\text{esg}}/\beta_{\text{esg}}^\star \leq 1$. If $\beta_i \leq \beta^\star$, $\omega_{i,\text{gmv}}$ may be positive if $\beta_{i,\text{esg}}$ is greater than the bound

$\beta_{\mathrm{esg}}^{\star}(1 - \beta_i/\beta^{\star})$, which is negative. Therefore, both positive and negative values of $\beta_{i,\mathrm{esg}}$ can lead to a long exposure. If $\beta_i \geq \beta^{\star}$, the bound is positive and only an asset with a positive ESG sensitivity has a positive weight.

Table 3.10: Minimum variance portfolios (Example 14)

Asset	β_i	$\beta_{i,\mathrm{esg}}$	One-factor		Two-factor	
			GMV	MV	GMV	MV
#1	0.90	-0.50	147.33%	0.00%	166.55%	33.54%
#2	0.80	0.70	24.67%	9.45%	21.37%	1.46%
#3	1.20	0.20	-49.19%	0.00%	-58.80%	0.00%
#4	0.70	0.90	74.20%	90.55%	65.06%	64.99%
#5	1.30	-0.30	-97.01%	0.00%	-94.18%	0.00%
$\sigma(w)$			11.45%	19.19%	11.54%	20.40%
$\beta(w)$			0.1913	0.7095	0.1954	0.7686
$\beta_{\mathrm{esg}}(w)$			0.2965	0.8811	0.0674	0.4274
$\beta^{\star}$			1.0972	0.8307	1.0906	0.8667
$\beta_{\mathrm{esg}}^{\star}$					19.7724	9.7394

Table 3.11: Minimum variance portfolios (Example 15)

Asset	β_i	$\beta_{i,\mathrm{esg}}$	One-factor		Two-factor	
			GMV	MV	GMV	MV
#1	0.70	-0.50	26.21%	66.96%	57.34%	73.46%
#2	0.80	-0.70	32.17%	33.04%	19.57%	26.54%
#3	0.90	-0.50	166.32%	0.00%	10.31%	0.00%
#4	1.20	0.90	-7.55%	0.00%	130.35%	0.00%
#5	1.50	1.30	-117.15%	0.00%	-117.58%	0.00%
$\sigma(w)$			8.10%	19.69%	17.21%	20.47%
$\beta(w)$			0.0899	0.7330	0.4513	0.7265
$\beta_{\mathrm{esg}}(w)$			-2.7786	-0.5661	-0.8306	-0.5531
$\beta^{\star}$			1.1664	0.8462	1.0505	0.9227
$\beta_{\mathrm{esg}}^{\star}$					-3.5677	-7.5752

Remark 36 *The previous examples illustrate that the global minimum variance portfolio can have a positive or negative ESG beta. In fact, it depends on the correlation between CAPM betas and ESG betas. Generally, the GMV portfolio has a positive ESG beta if there is a negative correlation between β_i and $\beta_{i,\mathrm{esg}}$.*

If the market risk and ESG factors are uncorrelated, we can assume that[16] $\beta^{\top}\beta_{\mathrm{esg}} \approx 0$. If we consider mean-variance optimized portfolios, Equation (3.15) becomes:

$$w = \vartheta \left(D^{-1} - \omega_1 \tilde{\beta}\tilde{\beta}^{\top} - \omega_2 \tilde{\beta}_{\mathrm{esg}}\tilde{\beta}_{\mathrm{esg}}^{\top} \right) \mu$$

---

[16]Otherwise, it means that green assets are generally associated to high beta assets if $\beta^{\top}\beta_{\mathrm{esg}} \gg 0$ or low beta assets if $\beta^{\top}\beta_{\mathrm{esg}} \ll 0$.

We deduce that:

$$w_i \propto \omega_\eta \frac{\mu_i}{\tilde{\sigma}_i^2} - \omega_\beta \frac{\beta_i}{\tilde{\sigma}_i^2} - \omega_{\beta_{\mathrm{esg}}} \frac{\beta_{i,\mathrm{esg}}}{\tilde{\sigma}_i^2}$$

where:

$$\begin{cases} \omega_\eta = 1 + \sigma_m^2 \tilde{\beta}^\top \beta + \sigma_{\mathrm{esg}}^2 \tilde{\beta}_{\mathrm{esg}}^\top \beta_{\mathrm{esg}} \geq 0 \\ \omega_\beta = \sigma_m^2 \left(1 + \sigma_{\mathrm{esg}}^2 \tilde{\beta}_{\mathrm{esg}}^\top \beta_{\mathrm{esg}}\right) \sum_{j=1}^n \frac{\beta_j \mu_j}{\tilde{\sigma}_j^2} \geq 0 \\ \omega_{\beta_{\mathrm{esg}}} = \sigma_{\mathrm{esg}}^2 \left(1 + \sigma_m^2 \tilde{\beta}^\top \beta\right) \sum_{j=1}^n \frac{\beta_{j,\mathrm{esg}} \mu_j}{\tilde{\sigma}_j^2} \lessgtr 0 \end{cases}$$

We deduce that w_i is an increasing function of μ_i and a decreasing function of β_i and $\tilde{\sigma}_i$. Like the minimum variance portfolio, w_i can be a decreasing or increasing function of $\beta_{i,\mathrm{esg}}$ because $\omega_{\beta_{\mathrm{esg}}}$ can be positive or negative.

3.1.3 ESG efficient frontier

Pedersen *et al.* (2021) propose an extension of the Markowitz optimization model by considering ESG preferences (hereafter, PFP model). Even if the model settings are similar, the PFP model slightly differs from the PST model, because it is more focused on the efficient frontier.

Model settings

The investment universe is made up of n assets. We have $\tilde{R} = R - r \sim \mathcal{N}(\pi, \Sigma)$. The assets have an ESG score given by $\boldsymbol{\mathcal{S}} = (\boldsymbol{\mathcal{S}}_1, \ldots, \boldsymbol{\mathcal{S}}_n)$. Let $w = (w_1, \ldots, w_n)$ be the portfolio of the investor. His initial wealth is W whereas his terminal wealth is given by $\tilde{W} = \left(1 + r + w^\top \tilde{R}\right) W$. The model uses the mean-variance utility function, which is tilted by the ESG score of the portfolio:

$$\begin{aligned} \boldsymbol{\mathcal{U}}\left(\tilde{W}, w\right) &= \mathbb{E}\left[\tilde{W}\right] - \frac{\bar{\gamma}}{2} \mathrm{var}\left(\tilde{W}\right) + \zeta\left(\boldsymbol{\mathcal{S}}(w)\right) W \\ &= \left(1 + r + w^\top \pi - \frac{\bar{\gamma}}{2} w^\top \Sigma w + \zeta\left(w^\top \boldsymbol{\mathcal{S}}\right)\right) W \end{aligned}$$

where ζ is a function that depends on the investor. Optimizing the utility function is equivalent to find the mean-variance-esg optimized portfolio:

$$\begin{aligned} w^\star &= \arg\max w^\top \pi - \frac{\bar{\gamma}}{2} w^\top \Sigma w + \zeta\left(w^\top \boldsymbol{\mathcal{S}}\right) \\ &\text{s.t.} \quad \mathbf{1}^\top w = 1 \end{aligned}$$

Let $\sigma(w) = \sqrt{w^\top \Sigma w}$ and $\boldsymbol{\mathcal{S}}(w) = w^\top \boldsymbol{\mathcal{S}}$. The optimization problem can be decomposed as follows:

$$w^\star = \arg\left\{\max_{\bar{\boldsymbol{\mathcal{S}}}} \left\{\max_{\bar{\sigma}} \left\{\max_w \left\{f(w; \pi, \Sigma, \boldsymbol{\mathcal{S}}) \text{ s.t. } w \in \Omega\left(\bar{\sigma}, \bar{\boldsymbol{\mathcal{S}}}\right)\right\}\right\}\right\}\right\} \tag{3.16}$$

where:

$$f(w; \pi, \Sigma, \boldsymbol{\mathcal{S}}) = w^\top \pi - \frac{\bar{\gamma}}{2} \sigma^2(w) + \zeta\left(\boldsymbol{\mathcal{S}}(w)\right)$$

and:

$$\Omega = \left\{w \in \mathbb{R}^n : \mathbf{1}^\top w = 1, \sigma(w) = \bar{\sigma}, \boldsymbol{\mathcal{S}}(w) = \bar{\boldsymbol{\mathcal{S}}}\right\}$$

The optimal portfolio

We consider the first optimization sub-problem, which is a $\sigma - \boldsymbol{S}$ problem:

$$w^{\star}\left(\bar{\sigma},\bar{\boldsymbol{S}}\right) \quad = \quad \arg\max w^{\top}\pi$$

$$\text{s.t.} \quad \left\{ \begin{array}{l} \mathbf{1}^{\top}w = 1 \\ \sigma\left(w\right) = \sqrt{w^{\top}\Sigma w} = \bar{\sigma} \\ \boldsymbol{S}\left(w\right) = w^{\top}\boldsymbol{S} = \bar{\boldsymbol{S}} \end{array} \right.$$

Pedersen *et al.* (2021) rewrite the last two equations as $w^{\top}\Sigma w - \bar{\sigma}^2 = 0$ and $w^{\top}\left(\boldsymbol{S} - \bar{\boldsymbol{S}}\mathbf{1}\right) = 0$ because[17] $\mathbf{1}^{\top}w = 1$. Therefore, the Lagrange function is:

$$\mathcal{L}\left(w;\lambda_1,\lambda_2\right) = w^{\top}\pi + \lambda_1\left(w^{\top}\Sigma w - \bar{\sigma}^2\right) + \lambda_2\left(w^{\top}\left(\boldsymbol{S} - \bar{\boldsymbol{S}}\mathbf{1}\right)\right)$$

The first-order condition is:

$$\frac{\partial\,\mathcal{L}\left(w;\lambda_1,\lambda_2\right)}{\partial\,w} = \pi + 2\lambda_1\Sigma w + \lambda_2\left(\boldsymbol{S} - \bar{\boldsymbol{S}}\mathbf{1}\right) = \mathbf{0}$$

We deduce that the optimal portfolio is given by:

$$w = -\frac{1}{2\lambda_1}\Sigma^{-1}\left(\pi + \lambda_2\left(\boldsymbol{S} - \bar{\boldsymbol{S}}\mathbf{1}\right)\right)$$

The second constraint $w^{\top}\left(\boldsymbol{S} - \bar{\boldsymbol{S}}\mathbf{1}\right) = 0$ implies that:

$$(*) \quad \Leftrightarrow \quad \left(\boldsymbol{S} - \bar{\boldsymbol{S}}\mathbf{1}\right)^{\top}\frac{1}{2\lambda_1}\Sigma^{-1}\left(\pi + \lambda_2\left(\boldsymbol{S} - \bar{\boldsymbol{S}}\mathbf{1}\right)\right) = 0$$

$$\Leftrightarrow \quad \lambda_2 = -\frac{\left(\boldsymbol{S} - \bar{\boldsymbol{S}}\mathbf{1}\right)^{\top}\Sigma^{-1}\pi}{\left(\boldsymbol{S} - \bar{\boldsymbol{S}}\mathbf{1}\right)^{\top}\Sigma^{-1}\left(\boldsymbol{S} - \bar{\boldsymbol{S}}\mathbf{1}\right)}$$

$$\Leftrightarrow \quad \lambda_2 = \frac{\bar{\boldsymbol{S}}\left(\mathbf{1}^{\top}\Sigma^{-1}\pi\right) - \boldsymbol{S}^{\top}\Sigma^{-1}\pi}{\boldsymbol{S}^{\top}\Sigma^{-1}\boldsymbol{S} - 2\bar{\boldsymbol{S}}\left(\mathbf{1}^{\top}\Sigma^{-1}\boldsymbol{S}\right) + \bar{\boldsymbol{S}}^2\left(\mathbf{1}^{\top}\Sigma^{-1}\mathbf{1}\right)}$$

$$\Leftrightarrow \quad \lambda_2 = \frac{C_{1,\pi}\bar{\boldsymbol{S}} - C_{s,\pi}}{C_{s,s} - 2C_{1,s}\bar{\boldsymbol{S}} + C_{1,1}\bar{\boldsymbol{S}}^2}$$

where $C_{x,y}$ is the compact notation for $x^{\top}\Sigma^{-1}y$ — $C_{1,\pi} = \mathbf{1}^{\top}\Sigma^{-1}\pi$, $C_{s,\pi} = \boldsymbol{S}^{\top}\Sigma^{-1}\pi$, $C_{s,s} = \boldsymbol{S}^{\top}\Sigma^{-1}\boldsymbol{S}$, $C_{1,s} = \mathbf{1}^{\top}\Sigma^{-1}\boldsymbol{S}$ and $C_{1,1} = \mathbf{1}^{\top}\Sigma^{-1}\mathbf{1}$. Using the first constraint $w^{\top}\Sigma w - \bar{\sigma}^2 = 0$, we deduce that:

$$\begin{aligned} \bar{\sigma}^2 \quad &= \quad -\frac{1}{2\lambda_1}w^{\top}\Sigma\Sigma^{-1}\left(\pi + \lambda_2\left(\boldsymbol{S} - \bar{\boldsymbol{S}}\mathbf{1}\right)\right) \\ &= \quad -\frac{1}{2\lambda_1}\left(w^{\top}\pi + \lambda_2 w^{\top}\left(\boldsymbol{S} - \bar{\boldsymbol{S}}\mathbf{1}\right)\right) \\ &= \quad -\frac{1}{2\lambda_1}w^{\top}\pi \\ &= \quad \frac{1}{4\lambda_1^2}\pi^{\top}\Sigma^{-1}\left(\pi + \lambda_2\left(\boldsymbol{S} - \bar{\boldsymbol{S}}\mathbf{1}\right)\right) \end{aligned}$$

---

[17] This last constraint $\mathbf{1}^{\top}w = 1$ is not used in the sequel, implying that the proportion of the wealth invested in the risk-free asset is equal to $w_r = 1 - \mathbf{1}^{\top}w$.

The first Lagrange coefficient is then equal to:

$$
\begin{aligned}
\lambda_1 &= -\frac{1}{2\bar{\sigma}}\sqrt{\pi^\top \Sigma^{-1}\pi + \lambda_2 \left(\pi^\top \Sigma^{-1}\boldsymbol{S} - \bar{\boldsymbol{S}}\left(\pi^\top \Sigma^{-1}\mathbf{1}\right)\right)} \\
&= -\frac{1}{2\bar{\sigma}}\sqrt{C_{\pi,\pi} - \frac{\left(C_{1,\pi}\bar{\boldsymbol{S}} - C_{s,\pi}\right)^2}{C_{s,s} - 2C_{1,s}\bar{\boldsymbol{S}} + C_{1,1}\bar{\boldsymbol{S}}^2}}
\end{aligned}
$$

where $C_{\pi,\pi} = \pi^\top \Sigma^{-1}\pi$. Pedersen *et al.* (2021) notice that the optimal portfolio is the product of the volatility $\bar{\sigma}$ and the vector $\varrho\left(\bar{\boldsymbol{S}}\right)$:

$$
\begin{aligned}
w^\star\left(\bar{\sigma}, \bar{\boldsymbol{S}}\right) &= -\frac{1}{2\lambda_1}\Sigma^{-1}\left(\pi + \lambda_2\left(\boldsymbol{S} - \bar{\boldsymbol{S}}\mathbf{1}\right)\right) \\
&= \bar{\sigma} \cdot \varrho\left(\bar{\boldsymbol{S}}\right)
\end{aligned}
$$

where:

$$
\varrho\left(\bar{\boldsymbol{S}}\right) = \frac{1}{\lambda_1'}\Sigma^{-1}\left(\pi + \lambda_2\left(\boldsymbol{S} - \bar{\boldsymbol{S}}\mathbf{1}\right)\right)
$$

and:

$$
\lambda_1' = \sqrt{C_{\pi,\pi} - \frac{\left(C_{1,\pi}\bar{\boldsymbol{S}} - C_{s,\pi}\right)^2}{C_{s,s} - 2C_{1,s}\bar{\boldsymbol{S}} + C_{1,1}\bar{\boldsymbol{S}}^2}}
$$

Example 16 *We consider an investment universe of four assets. Their expected returns are equal to 6%, 7%, 8% and 10% while their volatilities are equal to 15%, 20%, 25% and 30%. The correlation matrix of asset returns is given by the following matrix:*

$$
\mathbb{C} = \begin{pmatrix} 100\% & & & \\ 20\% & 100\% & & \\ 30\% & 50\% & 100\% & \\ 40\% & 60\% & 70\% & 100\% \end{pmatrix}
$$

The risk-free rate is set to 2%. The ESG score vector is $\boldsymbol{S} = (3\%, 2\%, -2\%, -3\%)$.

We obtain $C_{1,\pi} = 2.4864$, $C_{s,\pi} = 0.0425$, $C_{s,s} = 0.1274$, $C_{1,s} = 1.9801$, $C_{1,1} = 64.1106$ and $C_{\pi,\pi} = 0.1193$. If we target $\bar{\sigma} = 20\%$ and $\bar{\boldsymbol{S}} = 1\%$, we deduce that $\lambda_1 = -0.8514$ and $\lambda_2 = -0.1870$. The optimal portfolio is then:

$$
w^\star\left(\bar{\sigma}, \bar{\boldsymbol{S}}\right) = \begin{pmatrix} 59.31\% \\ 29.52\% \\ 21.76\% \\ 20.72\% \end{pmatrix}
$$

It follows that the portfolio is leveraged since we have $w_r = 1 - \mathbf{1}^\top w = -31.31\%$. We verify that $\sqrt{w^\star\left(\bar{\sigma}, \bar{\boldsymbol{S}}\right)^\top \Sigma w^\star\left(\bar{\sigma}, \bar{\boldsymbol{S}}\right)} = 20\%$ and $\left(w^\star\left(\bar{\sigma}, \bar{\boldsymbol{S}}\right)^\top \boldsymbol{S}\right) \big/ \left(\mathbf{1}^\top w^\star\left(\bar{\sigma}, \bar{\boldsymbol{S}}\right)\right) = 1\%$. We also notice that:

$$
\varrho\left(\bar{\boldsymbol{S}}\right) = \begin{pmatrix} 2.9657 \\ 1.4759 \\ 1.0881 \\ 1.0358 \end{pmatrix}
$$

and verify that $w^\star\left(\bar{\sigma}, \bar{\boldsymbol{S}}\right) = \bar{\sigma} \cdot \varrho\left(\bar{\boldsymbol{S}}\right)$. The portfolio is then leveraged when $\bar{\sigma} \geq 1/\left(\mathbf{1}^\top \varrho\left(\bar{\boldsymbol{S}}\right)\right) = 17.75\%$.

The Sharpe ratio of the optimal portfolio

Let us rewrite the first-order condition as:

$$
\begin{aligned}
(*) \quad &\Leftrightarrow \quad \pi + 2\lambda_1 \Sigma w + \lambda_2 \left(\boldsymbol{S} - \bar{\boldsymbol{S}} \mathbf{1} \right) = \mathbf{0} \\
&\Leftrightarrow \quad w^\top \pi + 2\lambda_1 w^\top \Sigma w + \lambda_2 w^\top \left(\boldsymbol{S} - \bar{\boldsymbol{S}} \mathbf{1} \right) = \mathbf{0} \\
&\Leftrightarrow \quad w^\top \pi + 2\lambda_1 \bar{\sigma}^2 = \mathbf{0} \\
&\Leftrightarrow \quad \lambda_1 = -\frac{1}{2} \frac{w^\top \pi}{\bar{\sigma}^2} \\
&\Leftrightarrow \quad \lambda_1 = -\frac{1}{2} \frac{\mathrm{SR}\left(w \mid r \right)}{\bar{\sigma}}
\end{aligned}
$$

We deduce that the Sharpe ratio of the optimal portfolio $w^\star \left(\bar{\sigma}, \bar{\boldsymbol{S}} \right)$ is equal to:

$$
\mathrm{SR} \left(w^\star \left(\bar{\sigma}, \bar{\boldsymbol{S}} \right) \mid r \right) = \sqrt{ C_{\pi,\pi} - \frac{ \left(C_{1,\pi} \bar{\boldsymbol{S}} - C_{s,\pi} \right)^2 }{ C_{s,s} - 2C_{1,s} \bar{\boldsymbol{S}} + C_{1,1} \bar{\boldsymbol{S}}^2 } } = \mathrm{SR} \left(\bar{\boldsymbol{S}} \mid \pi, \Sigma, \boldsymbol{S} \right)
$$

Therefore, it depends on the asset parameters π, Σ, $\boldsymbol{S}$, the ESG objective $\bar{\boldsymbol{S}}$ of the investor, but not the volatility target $\bar{\sigma}$.

Using Example 16, we deduce that the Sharpe ratio of the optimal portfolio $w^\star \left(20\%, 1\% \right)$ is equal to 0.3406. More generally, we verify that $\mathrm{SR} \left(w^\star \left(\bar{\sigma}, \bar{\boldsymbol{S}} \right) \mid r \right)$ does not depends on the value $\bar{\sigma}$. For instance, we have $\mathrm{SR} \left(w^\star \left(\bar{\sigma}, -3\% \right) \mid r \right) = 0.2724$, $\mathrm{SR} \left(w^\star \left(\bar{\sigma}, -2\% \right) \mid r \right) = 0.2875$, $\mathrm{SR} \left(w^\star \left(\bar{\sigma}, -1\% \right) \mid r \right) = 0.3052$, $\mathrm{SR} \left(w^\star \left(\bar{\sigma}, 0\% \right) \mid r \right) = 0.3242$, $\mathrm{SR} \left(w^\star \left(\bar{\sigma}, 1\% \right) \mid r \right) = 0.3406$, $\mathrm{SR} \left(w^\star \left(\bar{\sigma}, 2\% \right) \mid r \right) = 0.3443$, and $\mathrm{SR} \left(w^\star \left(\bar{\sigma}, 3\% \right) \mid r \right) = 0.3221$. In Figure 3.9, we report the relationship between the target value $\bar{\boldsymbol{S}}$ and the Sharpe ratio $\mathrm{SR} \left(w^\star \left(\bar{\sigma}, \bar{\boldsymbol{S}} \right) \mid r \right)$.

The ESG-SR frontier

Since the objective function is equal to:

$$
\begin{aligned}
f \left(w^\star \left(\bar{\sigma}, \bar{\boldsymbol{S}} \right) ; \pi, \Sigma, \boldsymbol{S} \right) &= \left(\frac{ w^\star \left(\bar{\sigma}, \bar{\boldsymbol{S}} \right)^\top \pi }{ \bar{\sigma} } \right) \bar{\sigma} - \frac{\bar{\gamma}}{2} \bar{\sigma}^2 + \zeta \left(\bar{\boldsymbol{S}} \right) \\
&= \mathrm{SR} \left(\bar{\boldsymbol{S}} \mid \pi, \Sigma, \boldsymbol{S} \right) \bar{\sigma} - \frac{\bar{\gamma}}{2} \bar{\sigma}^2 + \zeta \left(\bar{\boldsymbol{S}} \right)
\end{aligned}
$$

the σ-problem becomes:

$$
\max_{\bar{\sigma}} \left\{ \max_w \left\{ f \left(w; \pi, \Sigma, \boldsymbol{S} \right) \text{ s.t. } w \in \Omega \left(\bar{\sigma}, \bar{\boldsymbol{S}} \right) \right\} \right\} = \max_{\bar{\sigma}} \left\{ \mathrm{SR} \left(\bar{\boldsymbol{S}} \mid \pi, \Sigma, \boldsymbol{S} \right) \bar{\sigma} - \frac{\bar{\gamma}}{2} \bar{\sigma}^2 + \zeta \left(\bar{\boldsymbol{S}} \right) \right\}
$$

The first-order condition is $\mathrm{SR} \left(\bar{\boldsymbol{S}} \mid \pi, \Sigma, \boldsymbol{S} \right) - \bar{\gamma} \bar{\sigma} = 0$ or $\bar{\sigma} = \bar{\gamma}^{-1} \mathrm{SR} \left(\bar{\boldsymbol{S}} \mid \pi, \Sigma, \boldsymbol{S} \right)$, and we have:

$$
\begin{aligned}
f \left(w^\star \left(\bar{\sigma}, \bar{\boldsymbol{S}} \right) ; \pi, \Sigma, \boldsymbol{S} \right) &= \bar{\gamma}^{-1} \mathrm{SR}^2 \left(\bar{\boldsymbol{S}} \mid \pi, \Sigma, \boldsymbol{S} \right) - \frac{1}{2} \bar{\gamma}^{-1} \mathrm{SR}^2 \left(\bar{\boldsymbol{S}} \mid \pi, \Sigma, \boldsymbol{S} \right) + \zeta \left(\bar{\boldsymbol{S}} \right) \\
&= \frac{1}{2} \bar{\gamma}^{-1} \left(\mathrm{SR}^2 \left(\bar{\boldsymbol{S}} \mid \pi, \Sigma, \boldsymbol{S} \right) + 2 \bar{\gamma} \zeta \left(\bar{\boldsymbol{S}} \right) \right)
\end{aligned}
$$

We conclude that the $\boldsymbol{S}$-problem becomes:

$$
\boldsymbol{S}^\star = \arg \max_{\bar{\boldsymbol{S}}} \left\{ \mathrm{SR}^2 \left(\bar{\boldsymbol{S}} \mid \pi, \Sigma, \boldsymbol{S} \right) + 2 \bar{\gamma} \zeta \left(\bar{\boldsymbol{S}} \right) \right\} \tag{3.17}
$$

Figure 3.9: Relationship between $\bar{\boldsymbol{S}}$ and $\mathrm{SR}\left(\bar{\boldsymbol{S}} \mid \pi, \Sigma\right)$
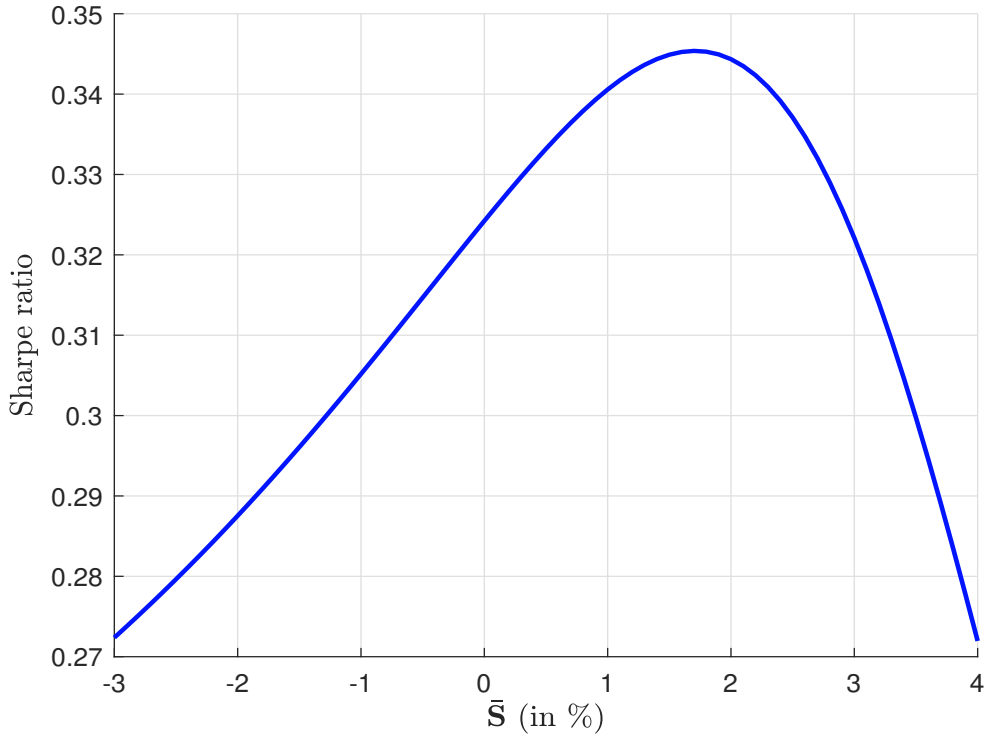
and the optimal portfolio is:

$$w^{\star} = w^{\star}\left(\sigma^{\star}, \boldsymbol{S}^{\star}\right)$$

where $\boldsymbol{S}^{\star}$ is the solution of the $\boldsymbol{S}$-problem and $\sigma^{\star} = \bar{\gamma}^{-1}\,\mathrm{SR}\left(\boldsymbol{S}^{\star} \mid \pi, \Sigma, \boldsymbol{S}\right)$. Pedersen *et al.* (2021) distinguish three groups of investors:

- Type-U or ESG-unware investors have no ESG preference and do not use the information of ESG scores;

- Type-A or ESG-aware investors have no ESG preference, but they use the ESG scores to update their views on the risk premia;

- Type-M or ESG-motivated investors have ESG preferences, implying that they would like to have a high ESG score.

Type-U investors hold the same portfolio, which is the standard tangency portfolio computed without the information of ESG scores:

$$w_U^{\star} = \frac{\Sigma^{-1}\pi}{\mathbf{1}^{\top}\Sigma^{-1}\pi}$$

Type-A investors choose the optimal portfolio with the highest Sharpe ratio. This is equivalent to set $\zeta\left(s\right) = 0$ in Equation (3.17) and we note $\boldsymbol{S}_A^{\star}$ the optimal ESG score. Finally, type-M investors choose an optimal portfolio on the ESG-SR efficient frontier, which has an ESG score greater than the optimal ESG score: $\boldsymbol{S}_M^{\star} \geq \boldsymbol{S}_A^{\star}$. In this case, we have $\mathrm{SR}\left(\boldsymbol{S}_M^{\star} \mid \pi, \Sigma, \boldsymbol{S}\right) \leq \mathrm{SR}\left(\boldsymbol{S}_A^{\star} \mid \pi, \Sigma, \boldsymbol{S}\right)$. Therefore, type-M investors reduce their Sharpe ratio in order to reach a better ESG score. While the optimal portfolio is the same for all type-A investors, it is different for two type-M investors who do not have the same risk-aversion coefficient $\bar{\gamma}$ and the same ESG utility function $\zeta\left(s\right)$.

Figure 3.10: Optimal portfolio for type-U investors (Example 16)
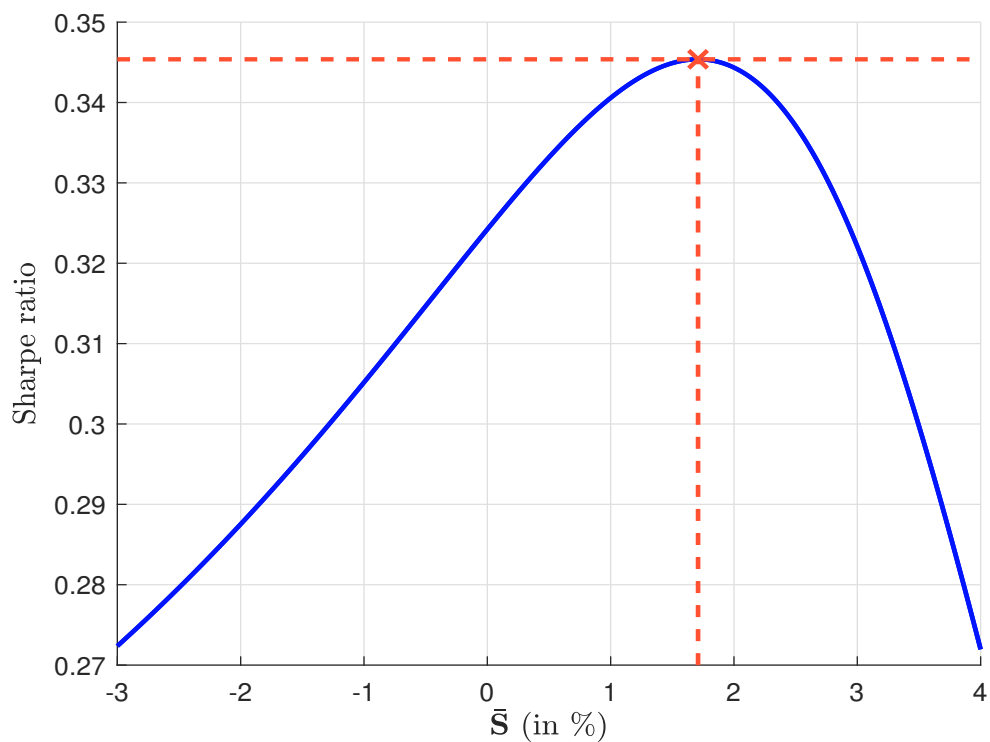
Figure 3.11: Optimal portfolio for type-A investors (Example 16)
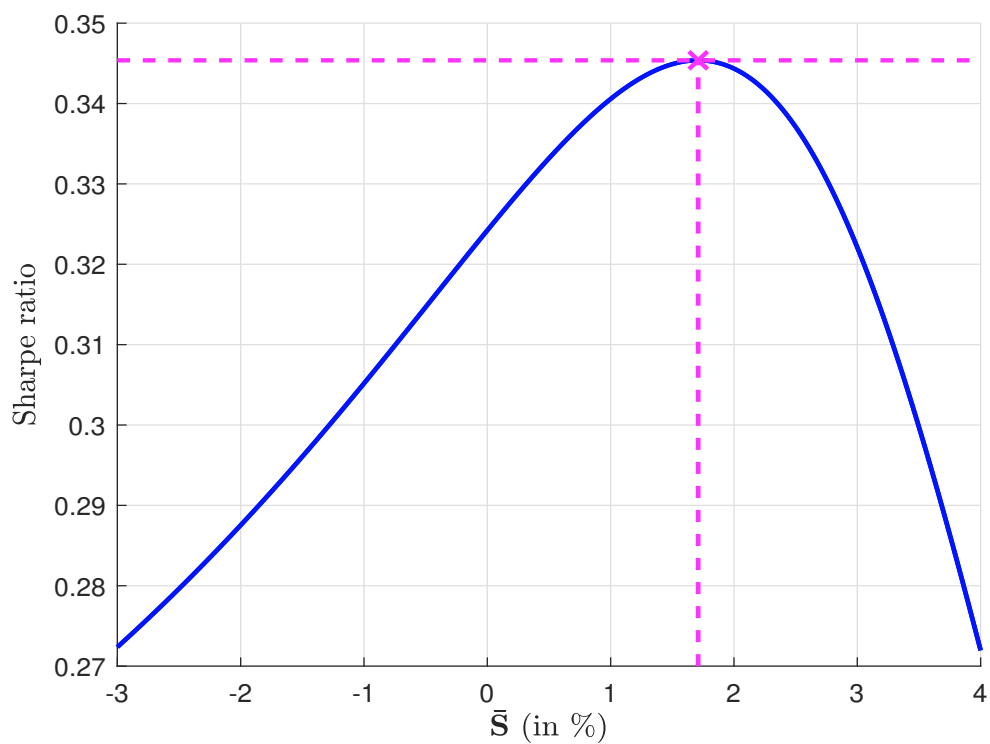
Figure 3.12: Optimal portfolio for type-M investors when $\zeta(s) = s$ (Example 16)
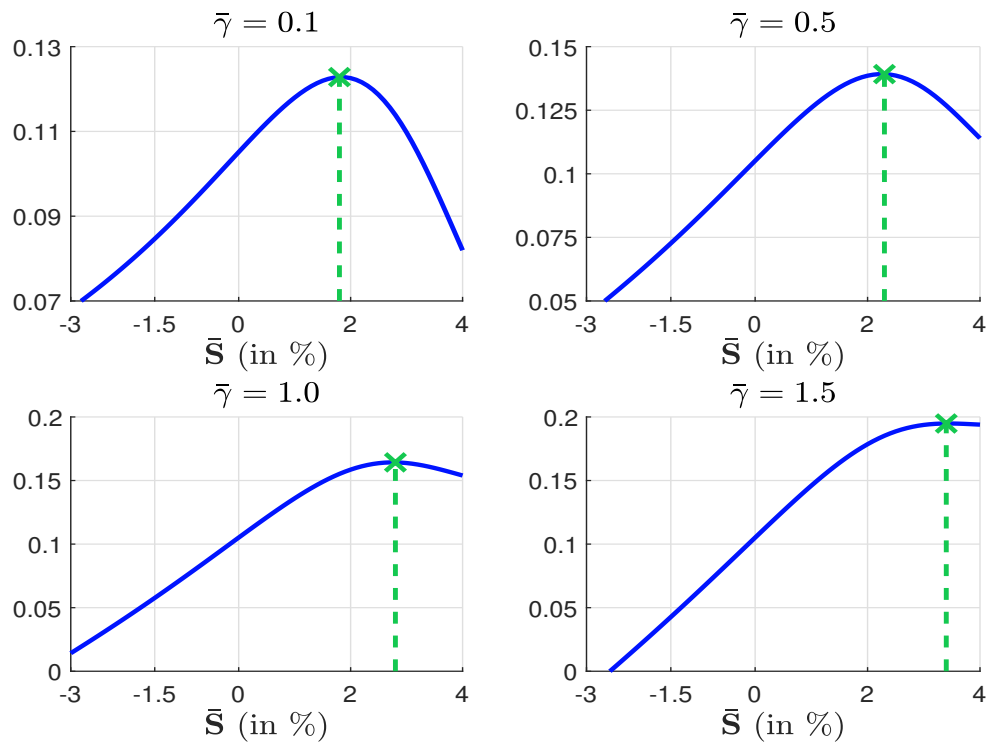
Figure 3.13: Optimal portfolio for type-M investors when $\zeta(s) = 0.2\sqrt{\max(s,0)}$ (Example 16)
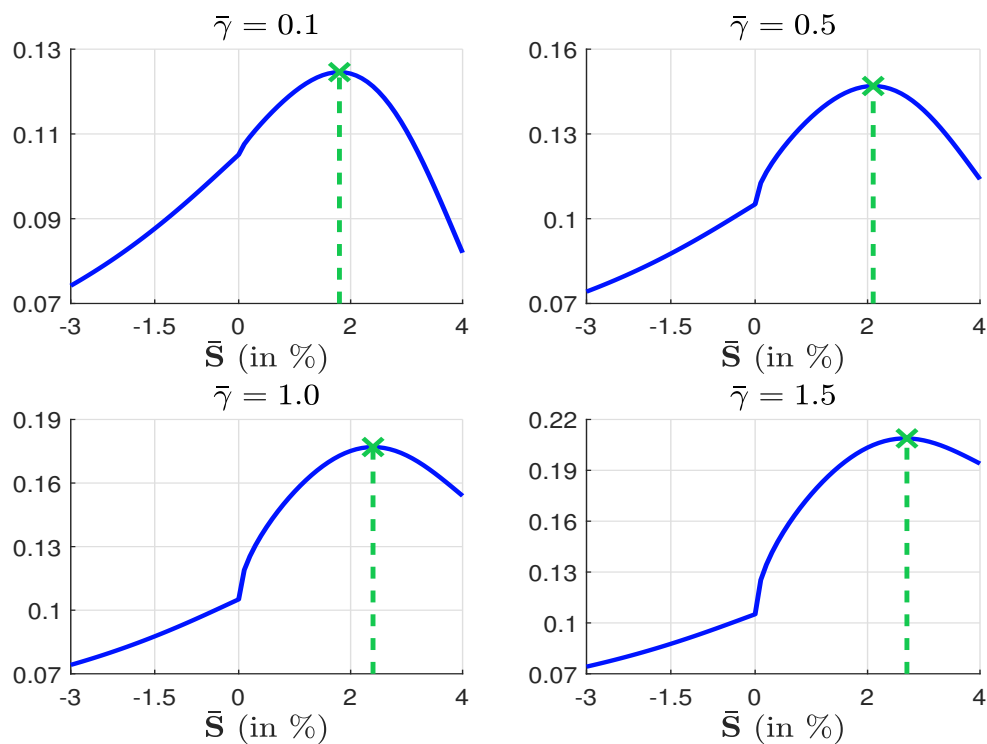
We consider Example 16. We compute the optimal portfolio for type-U investors. In this case, the previous analysis is not necessarily since the optimal portfolio is the traditional tangency portfolio. In the case of type-A investors, we must find the portfolio corresponding to the maximal Sharpe ratio of the ESG-SR efficient frontier (Figure 3.11). For type-M investors, we first compute the function $\xi\left(\bar{\boldsymbol{S}}\right)$:

$$\xi\left(\bar{\boldsymbol{S}}\right) = \mathrm{SR}^2\left(\bar{\boldsymbol{S}} \mid \pi, \Sigma, \boldsymbol{S}\right) + 2\bar{\gamma}\zeta\left(\bar{\boldsymbol{S}}\right)$$

Then, the optimal portfolio corresponds to the optimal ESG score that maximizes $\xi\left(\bar{\boldsymbol{S}}\right)$. Two examples are provided in Figures 3.12 and 3.13. Results are summarized in Table 3.12. For instance, if $\bar{\gamma} = 1.5$ and $\zeta\left(s\right) = 0.2\sqrt{\max\left(s,0\right)}$, the optimal portfolio is equal to $w_M^\star = (107.2\%, 66.0\%, 6.5\%, -7.9\%)$ and its Sharpe ratio is 0.332. This portfolio is obtained for an ESG score of 2.7%.

Table 3.12: Optimal portfolios (Example 16)

Statistics	Type-U	Type-A	Type-M					
			$\zeta\left(s\right) = s$			$\zeta\left(s\right) = 0.2\sqrt{\max\left(s,0\right)}$		
$\bar{\gamma}$			0.500	1.000	1.500	0.500	1.000	1.500
$\boldsymbol{S}\left(w^\star\right)$	0.017	0.017	0.023	0.028	0.034	0.021	0.024	0.027
$\sigma\left(w^\star\right)$	0.139	0.100	0.682	0.329	0.203	0.687	0.339	0.221
$\mathrm{SR}\left(w^\star \mid r\right)$	0.345	0.345	0.341	0.329	0.305	0.343	0.339	0.332
$w_1^\star$	0.524	0.378	3.028	1.623	1.090	2.900	1.542	1.072
$w_2^\star$	0.289	0.208	1.786	1.009	0.718	1.673	0.919	0.660
$w_3^\star$	0.120	0.086	0.383	0.073	-0.056	0.464	0.169	0.065
$w_4^\star$	0.067	0.048	-0.012	-0.144	-0.178	0.106	-0.035	-0.079
$w_r^\star$	0.000	0.280	-4.184	-1.562	-0.574	-4.143	-1.596	-0.718

Impact on asset returns

Pedersen *et al.* (2021) use the previous framework to analyze the dynamics of asset prices. They show that the impact of ESG highly depends on the relative proportion of the three types of investors. Let ω^U, ω^A and ω^M be the wealth share of type-U, type-A and type-M investors. The authors assume that the security dividend payoff is given by the vector $v = (v_1, \ldots, v_n)$ and depends on the ESG scores:

$$\mathbb{E}\left[v \mid \boldsymbol{S}\right] = \hat{\mu} + \theta\left(\boldsymbol{S} - \boldsymbol{S}_m\right)$$

where $\boldsymbol{S}_m$ is the ESG score of the market portfolio and the parameter θ determines how informative ESG scores are for future profits. In particular, $\theta = 0$ if ESG scores are non-informative. Otherwise, we can assume that $\theta > 0$, implying that firms with better ESG scores are more profitable on average. Pedersen *et al.* (2021) derive the following propositions:

- If $\omega^U = 1$ and $\omega^A = \omega^M = 0$, then unconditional expected returns are given by the CAPM:

$$\mathbb{E}\left[R_i\right] - r = \beta_i\left(\mathbb{E}\left[R_m\right] - r\right)$$

but conditional expected returns depend on the ESG scores:

$$\mathbb{E}\left[R_i \mid \boldsymbol{S}\right] - r = \beta_i\left(\mathbb{E}\left[R_m\right] - r\right) + \theta\frac{\boldsymbol{S}_i - \boldsymbol{S}_m}{P_i}$$

where P_i is the asset price of asset i. Two assets with the same beta do not have necessarily the same conditional risk premium.

- If $\omega^A = 1$ and $\omega^U = \omega^M = 0$, then the informational value of ESG scores is fully incorporated into asset prices, and we have:

$$\mathbb{E}\left[R_i \mid \boldsymbol{\mathcal{S}}\right] - r = \tilde{\beta}_i \left(\mathbb{E}\left[R_m \mid \boldsymbol{\mathcal{S}}\right] - r\right)$$

  where $\tilde{\beta}_i$ is the ESG-adjusted beta coefficient.

- If $\omega^M = 1$ and $\omega^U = \omega^A = 0$, then the conditional expected return is given by:

$$\mathbb{E}\left[R_i \mid \boldsymbol{\mathcal{S}}\right] - r = \tilde{\beta}_i \left(\mathbb{E}\left[R_m \mid \boldsymbol{\mathcal{S}}\right] - r\right) + \lambda_2 \left(\boldsymbol{\mathcal{S}}_i - \boldsymbol{\mathcal{S}}_m\right)$$

The best case for an ESG investor is $\omega^U = 1$ and $\omega^A = \omega^M = 0$ when all the others investors are ESG-unaware. The adjustment of market prices depends then on the growth of type-A and type-M investors. More generally, negative and/or positive alpha returns are explained by asymmetric information, supply/demand imbalance and trading motivations. Therefore, there is no obvious conclusion:

> "*If all types of investors exist, then several things can happen. If a security has a higher ESG score, then, everything else equal, its expected return can be higher or lower. A higher ESG score increases the demand for the stock from type-M investors, leading to a higher price and, therefore, a lower required return [...] Companies with poor ESG scores that are down-weighted by type-M investors will have lower prices and higher cost of capital. [...] Furthermore, the force that can increase the expected return is that the higher ESG could be a favorable signal of firm fundamentals, and if many type-U investors ignore this, the fundamental signal perhaps would not be fully reflected in the price [...] A future increase in ESG investing would lead to higher prices for high-ESG stocks [...]. If these flows are unexpected (or not fully captured in the price for other reasons), then high-ESG stocks would experience a return boost during the period of this repricing of ESG. If these flows are expected, then expected returns should not be affected.*" (Pedersen *et al.*, 2021).

In this context, it is difficult to predict whether ESG investing will outperform or underperform in the short run, since it depends on many factors. In particular, the PST and PFP models use the efficient market hypothesis (EMH), implying that asset prices must reflect all available information. For instance, as seen above, one consequence of EMH is that expected returns are not affected if the investment flows of ESG investing are expected. In the real life, this type of assumption is difficult to verify because we know that asset prices do not instantaneously react. Assuming that the dynamics of asset prices only depend on unexpected events in the short term also limits the validity of the theoretical analysis. At the end, asset prices are driven by trading orders whatever the real motivations of investors. These motivations can be rational or not rational, related to fundamental or extra-financial information, etc. Moreover, the PST and PFP models consider a specific trading strategy that mimics the ESG integration strategy as defined on page 38. The previous results do not necessarily hold if we consider[18] a worst-in-class exclusion strategy, a best-in-class selection strategy or an ESG momentum strategy. For all these reasons, the performance of ESG investing remains an intensive debate from a professional point of view. Nevertheless, these models are very useful because they give a normative framework and help to understand the mess of empirical results.

---

[18]See for instance Zerbib (2022).

3.2   Empirical results

As already said, the number of empirical research on the performance of ESG investing is impressive. Nevertheless, there is no obvious consensus, because there are so many factors that must be considered. First, ESG investing has evolved since the last thirteen years. The data are not the same — most of them didn't exist ten years ago — the practice of ESG scoring has definitively changed over time, the use of ESG considerations is new for many investors, etc. Backtesting an ESG strategy on a long history does not make sense. Second, we can not consider that the relationship between ESG and performance is static (positive or negative). Rather, we must accept that the relationship between ESG and performance is dynamic. Sometimes, ESG may create performance, but sometimes not. It was the case in the past, it will be the case in the future. Because the relationship mainly depends on the investment and trading flows of investors. Third, the performance of ESG investing depends on the portfolio implementation. This is not the same thing to consider an exclusion filter, add an ESG score to an existing asset picking model, implement a selection screening, etc. Finally, the relationship differs because it depends on the country, the asset class, the security universe, the ESG definition, etc. Let us illustrate with some examples. When we speak about the ESG performance, do we speak about the ESG global score or one of the pillars (**E**, **S** and **G**)? Do we speak about specific securities such as green bonds? Do we speak about American, European, Japanese or EM assets? Since we can multiply the questions endless, we focus more on the why than the whether. Why ESG investing has created or destroyed value for a specific investment universe during a given period?

3.2.1   Equity markets

The relationship between ESG and performance has been extensively investigated in stock markets. According to Coqueret (2022, Sections 4.2-4.5, pages 51-66), we can classify them into four categories: (1) ESG improves performance, (2) ESG does not impact performance, (3) ESG is financially detrimental and (4) it depends on many factors. According to Friede *et al.* (2015), the first category dominates the other categories:

> "*[...]   The results show that the business case for ESG investing is empirically very well founded. Roughly 90% of studies find a nonnegative ESG–CFP relation. More importantly, the large majority of studies reports positive findings. We highlight that the positive ESG impact on CFP appears stable over time. Promising results are obtained when differentiating for portfolio and non-portfolio studies, regions, and young asset classes for ESG investing such as emerging markets, corporate bonds, and green real estate.*" (Friede *et al.*, 2015, page 2010).

In fact, their findings are not obvious to accept since the concept of corporate financial performance covers many dimensions and is not limited to the financial performance in the equity market. For instance, CFP can also concern the cost of capital (El Ghoul *et al.*, 2011). Moreover, a large part of these studies focus on the **G** pillar (Gompers *et al.*, 2003) or use some proxy variables other than ESG scores or ratings (Edmans, 2011). We can also find many studies, whose conclusion is more neutral or negative (Barnett and Salomon, 2006; Fabozzi *et al.*, 2008; Hong and Kacperczyk, 2009; Johnson *et al.*, 2009; Capelle-Blancard and Monjon, 2014; Matos, 2020).

Since these different publications, a consensus has emerged among professionals. Like other investment styles, ESG investing has its good and bad times, and the relationship between ESG and performance is not straightforward and depends on many factors. Understanding these factors is the key challenge for investors rather than having a set of strong predetermined beliefs.

Simulated results

In what follows, we summarize the results obtained by Bennani *et al.* (2018) and Drei *et al.* (2019), who analyzed the impact of ESG on three equity portfolio management approaches: active management, passive management and factor investing.

Sorted portfolios Bennani *et al.* (2018) use the Amundi scoring system. For each company and each date, they access the ESG global score and its three components (**E**, **S** and **G**). The scores are normalized sector by sector in order to obtain a z-score shape, implying that they have a range roughly between -3 and $+3$. This also means that the scores are sector-neutral are distributed as a standard Gaussian probability distribution.

Box 3.4: The method of characteristic-sorted portfolios

Portfolio sorting has been popularized by Fama and French (1993) to test the impact of characteristics in asset pricing and to identify profitable investment strategies. The underlying idea is to sort individual assets into portfolios with respect to a given variable. If each portfolio has roughly the same number of constituents and only differs in the level of the sorting variable, the differences in the performance can then be attributed to the impact of the sorting variable. Generally, each portfolio is equally- or value-weighted in order to maximize the diversification. In the univariate case, the most popular approach is the quintile method, where the breakpoints for the sorting variable correspond to the 20th, 40th, 60th and 80th percentiles.

Table 3.A: An illustrative example

Asset	$\mathcal{S}_i$	Rank	Q_i	Weight
#1	-0.3	6	Q_3	$+50\%$
#2	0.2	5	Q_3	$+50\%$
#3	-1.0	7	Q_4	$+50\%$
#4	1.5	3	Q_2	$+50\%$
#5	-2.9	10	Q_5	$+50\%$
#6	0.8	4	Q_2	$+50\%$
#7	-1.4	8	Q_4	$+50\%$
#8	2.3	2	Q_1	$+50\%$
#9	2.8	1	Q_1	$+50\%$
#10	-2.2	9	Q_5	$+50\%$

We consider the example below, where the sorting variable is an ESG score. Since the investment universe is made up of 10 assets, each sorted portfolio has two assets. Portfolio Q_1 corresponds to the highest scores, while Portfolio Q_5 corresponds to the lowest scores. Finally, we obtain $Q_1 = (\#8, \#9)$, $Q_2 = (\#4, \#6)$, $Q_3 = (\#1, \#2)$, $Q_4 = (\#3, \#7)$ and $Q_5 = (\#5, \#10)$.

For building the active management strategy, the authors use the sorting portfolio method. Every quarter, they rank the stocks with respect to their score, and form five quintile portfolios[19]. Portfolio Q_1 corresponds to the 20% best-ranked stocks, whereas Portfolio Q_5 corresponds to the 20% worst-rated stocks. The selected stocks are then equally-weighted and each portfolio is invested the first

---

[19]Given a universe of stocks, each portfolio is then composed of 20% of assets.

trading day of the quarter and is held for three months. Quarterly rebalancing is implemented in order to limit the turnover.

Figure 3.14: Annualized return of ESG-sorted portfolios (MSCI North America, global score)
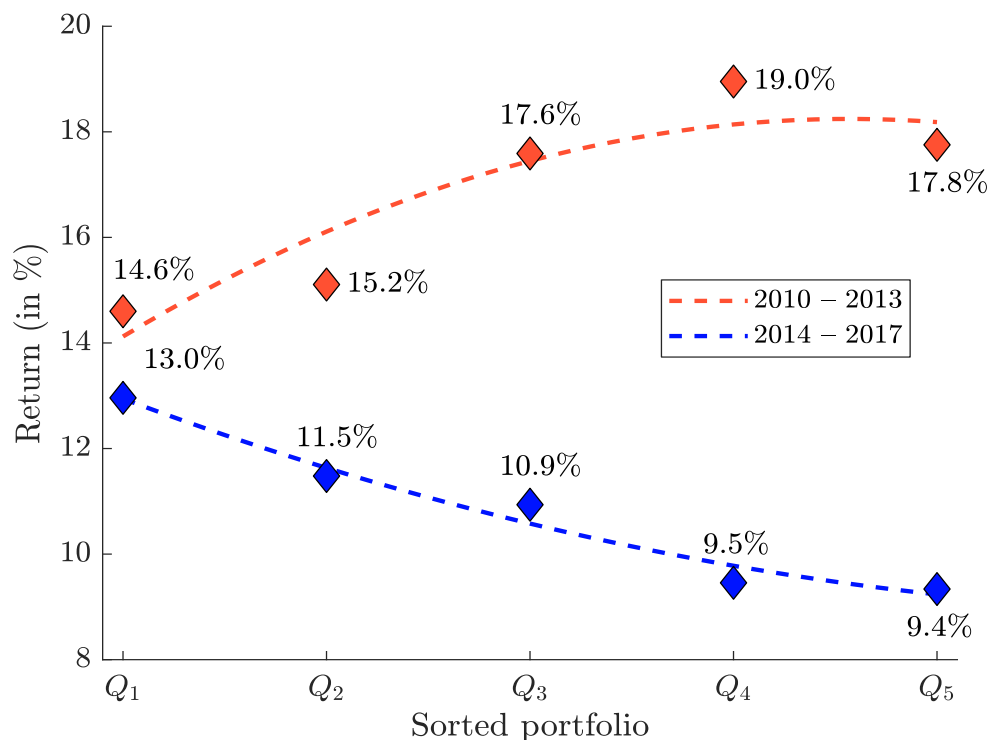
Source: Bennani *et al.* (2018).

They consider five investment universes using the following MSCI indexes: North America, EMU, Europe-ex-EMU, Japan and World. For each universe and each quintile portfolio, they calculate the gross performance without taking into account transaction costs. By analyzing the results, the authors observe a break during the 2010–2017 study period. Typically, the first half of the period is less favorable to ESG screening than the second period. In Figure 3.14, we report their results obtained for North American stocks. During the period 2010–2013, Portfolio Q_1 displays a gross return of 14.6% whereas Portfolio Q_5 shows a gross return of 17.8%. We observe an increasing function between the return and the quintile. During this period, best-in-class stocks underperformed worst-in-class stocks. The story is different when we focus on the 2014–2017 period. Portfolio Q_1 displays a performance of 13.0% whereas Portfolio Q_5 shows a performance of 9.4%. Clearly best-in-class stocks outperformed worst-in-class stocks during this second period. If we consider individual pillars, Bennani *et al.* (2018) obtained very similar results in Figure 3.15. (**E**), (**S**) and (**G**) stock picking negatively impacted performance between 2010 and 2013, whereas the impact of (**E**), (**S**) and (**G**) stock picking on performance is positive between 2014 and 2017. During the 2014–2017 period, the environmental screening produces the best result, followed by the governance scoring. However, for the governance component, the performance difference between Portfolios Q_1, Q_2, Q_3 and Q_4 is not significant. Only Portfolio Q_5 underperforms substantially, meaning that worst-rated stocks are penalized, but best-rated stocks are not necessarily rewarded.

These results clearly show that ESG active management was penalized during the 2010–2013 period, whereas it created an excess performance between 2014 and 2017. In the case of the Eurozone,

Figure 3.15: Annualized return of ESG-sorted portfolios (MSCI North America, individual pillars)
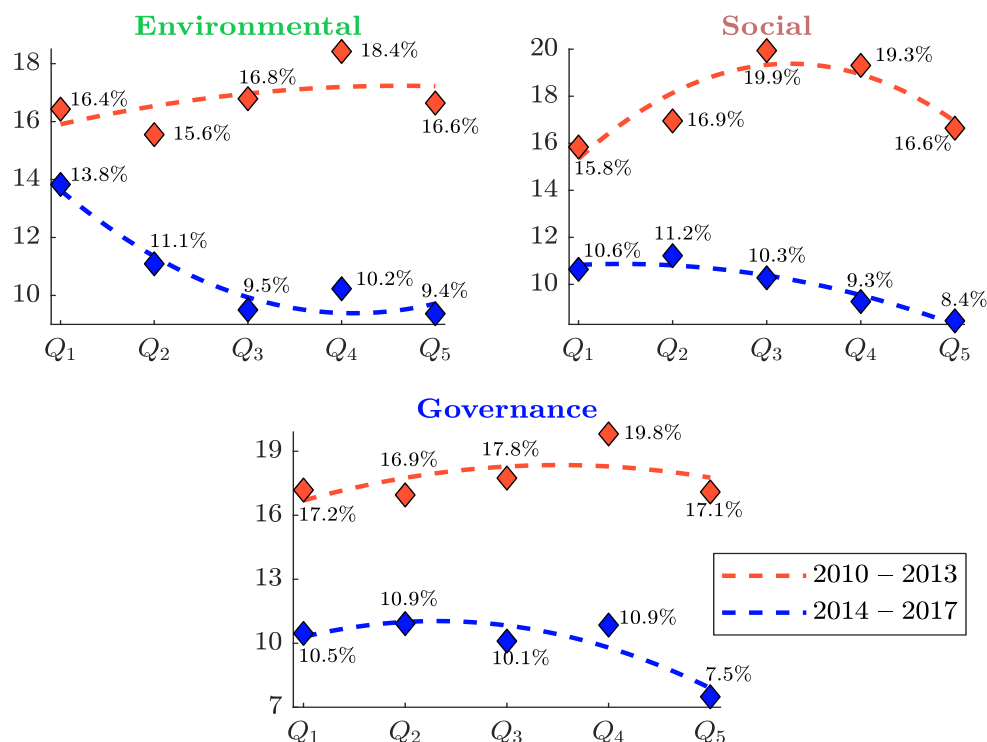
Source: Bennani *et al.* (2018).

the conclusion is the same for the ESG global score, and its three components. For instance, Portfolio Q_1 generated a return of 8.6% whereas Portfolio Q_5 generated a return of 10.0% between 2010 and 2013 (Figure 3.16). On the contrary, the performance was respectively 14.7% and 7.5% for Portfolios Q_1 and Q_5 during the 2014–2017 period. Therefore, the first period is characterized by a U-shape, whereas best-in-class stocks far outperformed worst-in-class stocks over the second period. We notice that the performance difference mainly concerns Portfolios Q_1 and Q_5, but not Portfolios Q_2, Q_3 and Q_4, implying that worst-in-class stocks are penalized and best-in-class stocks are rewarded. If we consider the individual pillars, governance is the most discriminant component (Figure 3.17). The difference between Q_1 and Q_5 Portfolios exceeds 7% during the last period. For the **(E)** score, we observe a U-shape behavior between 2010 and 2013. Since 2014, the relationship between the quintile portfolios and their returns is clearly decreasing. It is less impressive than for the **(G)** score, but it affects all the portfolios[20]. The integration of the social pillar is the least convincing.

For the other investment universes, the results are more heterogeneous. In the case of the Europe-ex-EMU universe, ESG integration is country specific, meaning that the performance is highly dependent on the overweight or underweight of each country. For example, the **(G)** screening largely overweights UK stocks if we consider a $Q_1 - Q_5$ long/short portfolio. On the contrary, **(E)** or **(S)** screenings promote Swedish stocks. The case of Japan is puzzling. Indeed, ESG screening was less favorable during the 2014–2017 period. When we consider the universe of the MSCI World index, the results are similar to those obtained for North America and the Eurozone. These different results are summarized in Table 3.13, where we have reported the impact of ESG screening (**(E)**,

---

[20]For the **(G)** score, the difference mainly concerns Portfolios Q_1 and Q_5, and less so the median portfolios.

Figure 3.16: Annualized return of ESG-sorted portfolios (MSCI EMU, global score)
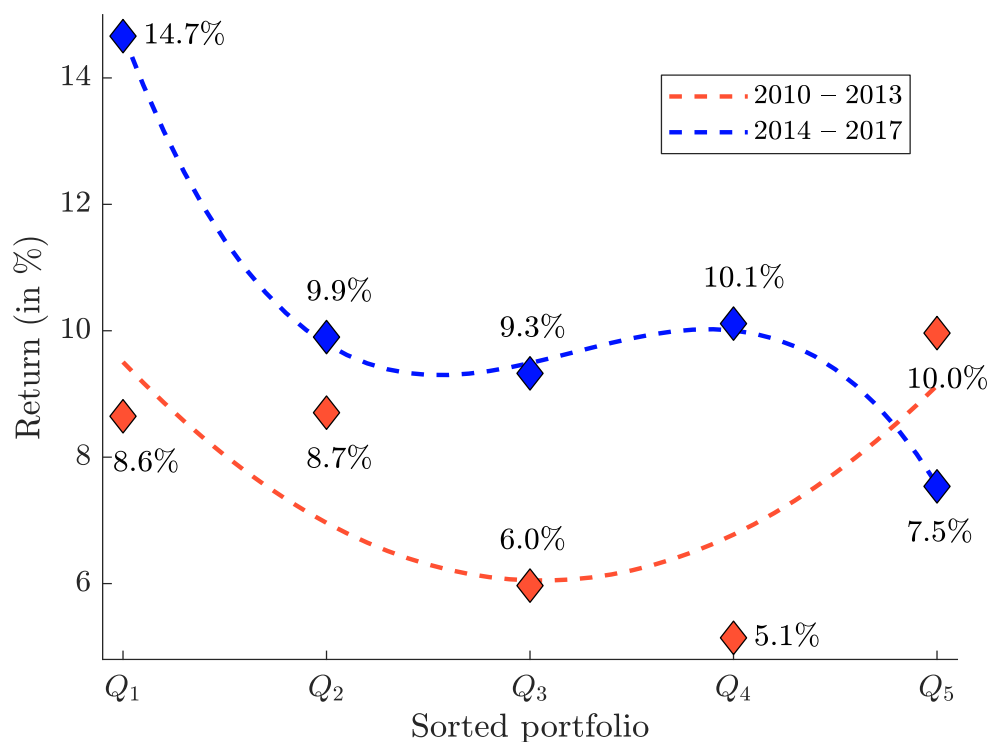
Figure 3.17: Annualized return of ESG-sorted portfolios (MSCI EMU, individual pillars)
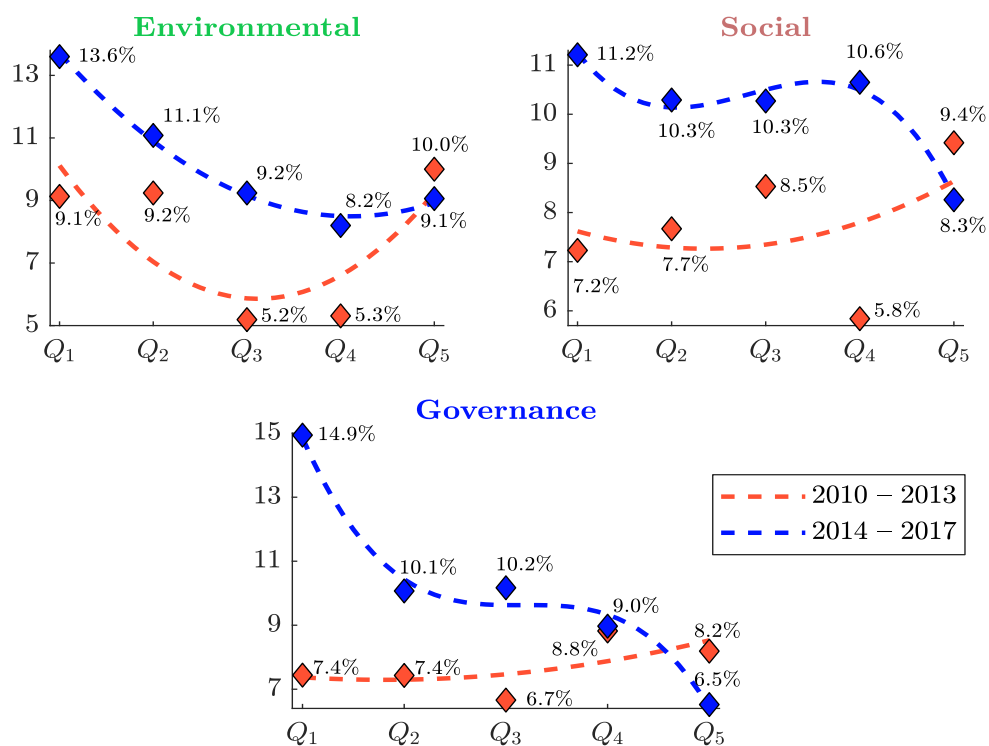
(S), (G) and ESG) on the returns of sorted portfolios. Again, the results illustrates the contrast between the two periods. To summarize, Bennani *et al.* (2018) concluded that the relationship between performance and ESG is time-varying and depend on several factors, especially the region and the ESG pillar. They also noticed that some investment universes present ESG-country biases, implying that the relationship between performance and ESG cannot be analyzed. This is the case of the MSCI Europe-ex-EMU index, but such bias is not also excluded for the MSCI World index.

Table 3.13: Impact of ESG screening on sorted portfolio returns (2010–2017)

Period	Pillar	North America	EMU	Europe-ex-EMU	Japan	World
2010–2013	ESG	−−	−	0	+	0
	E	−	0	+	−	0
	S	−	−	0	−	−
	G	−	0	+	0	+
2014–2017	ESG	++	++	0	−	+
	E	++	++	−	+	++
	S	+	+	0	0	+
	G	+	++	0	+	++

Source: Bennani *et al.* (2018).

The study of Drei *et al.* (2019) is an update of the analysis of Bennani *et al.* (2018) when considering the recent period 2018–2019. They use exactly the same data, the same investment universes and the same methodology. Their main results are reported in Figures 3.18 and 3.19.

Box 3.5: Computing the performance of long/short portfolios

Let $w_{\mathcal{L}ong}$ and $w_{\mathcal{S}hort}$ be the long and short portfolio. We note $R_t(w)$ the annualized return of the portfolio w between $t-1$ and t. The performance of the long/short portfolio $w_{\mathcal{L}ong} - w_{\mathcal{S}hort}$ satisfies the following definition:

$$(1 + R_t(w_{\mathcal{L}ong})) = (1 + \alpha_t(w_{\mathcal{L}ong} - w_{\mathcal{S}hort})) \cdot (1 + R_t(w_{\mathcal{S}hort}))$$

where $\alpha_t(w_{\mathcal{L}ong} - w_{\mathcal{S}hort})$ is the alpha return of $w_{\mathcal{L}ong} - w_{\mathcal{S}hort}$. We deduce that:

$$\alpha_t(w_{\mathcal{L}ong} - w_{\mathcal{S}hort}) = \frac{R_t(w_{\mathcal{L}ong}) - R_t(w_{\mathcal{S}hort})}{1 + R_t(w_{\mathcal{S}hort})} \qquad (3.18)$$

Looking at the $Q_1 - Q_5$ long-short portfolios in North America (Figure 3.18) and the Eurozone (Figure 3.19), we can see the evolution of the integration of ESG and its pillars in both markets. In the 2010–2013 period, sustainable investors were penalized, as seen by the negative return of the $Q_1 - Q_5$ long-short portfolios. In the 2014–2017 period, after the radical break in ESG integration, ESG investing gained momentum and yielded positive returns on all pillars on both sides of the Atlantic. However, after eight years of parallel development, Drei *et al.* (2019) observed a contradictory trend in ESG investing between North America and the Eurozone between 2018 and 2019. Indeed, the last period is marked by a squeeze in alpha returns on all dimensions in North America, and even a loss on the (E) pillar. This loss is important because it is the first long-short portfolio with a negative return since the 2014 ESG turning point in these two investment universes. Moreover, they observe

Figure 3.18: Annualized return of long/short $Q_1 - Q_5$ sorted portfolios (MSCI North America)
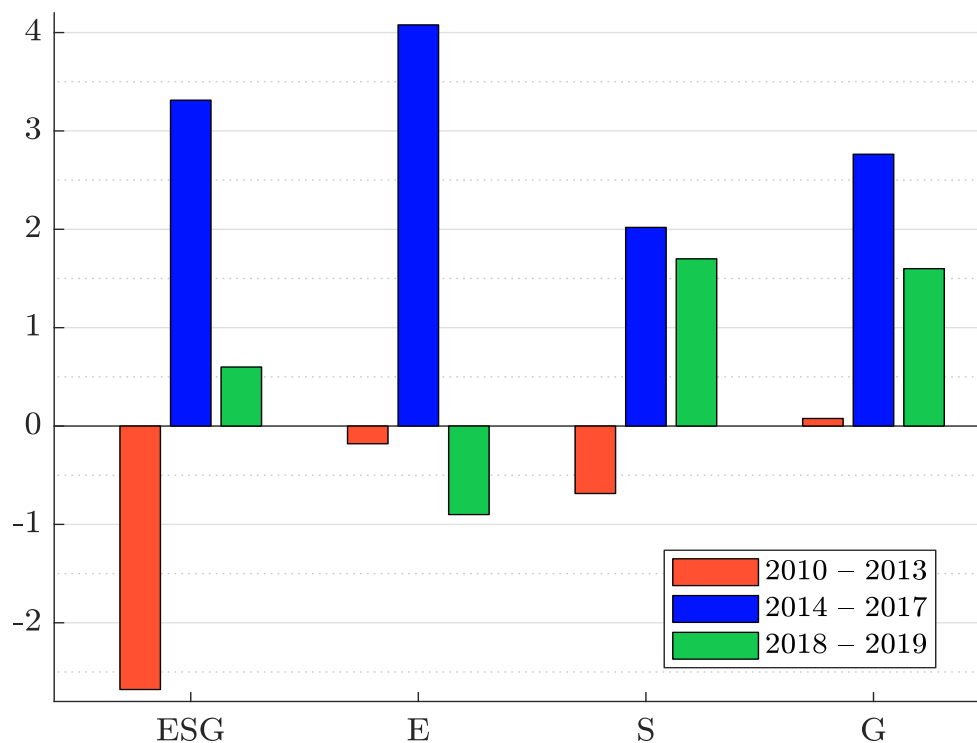
Source: Drei *et al.* (2019).

Figure 3.19: Annualized return of long/short $Q_1 - Q_5$ sorted portfolios (MSCI EMU)
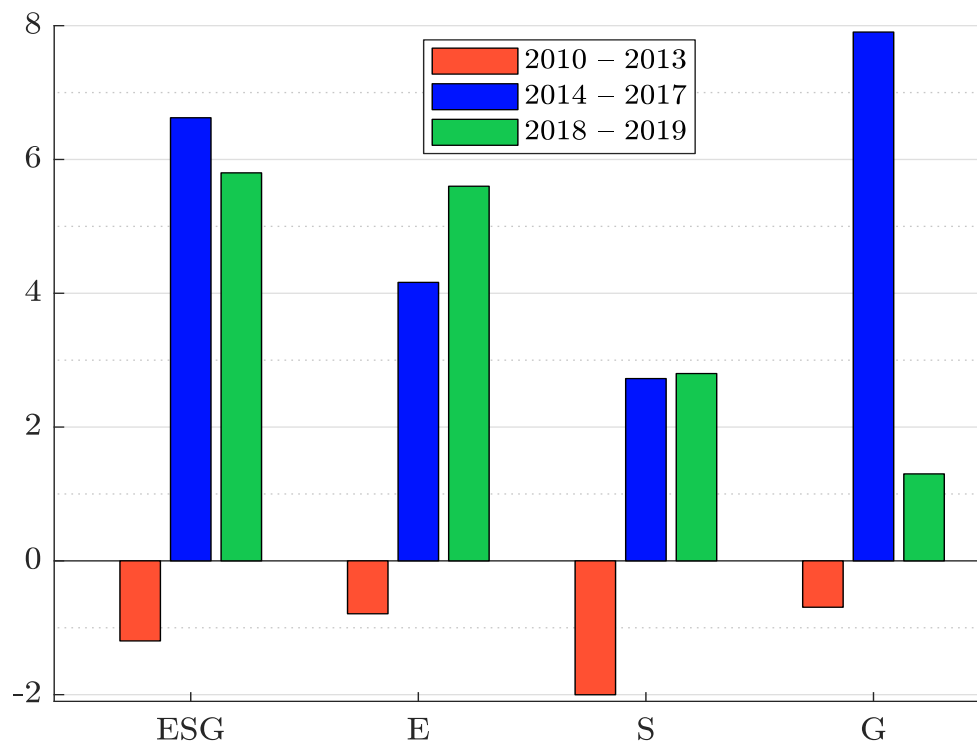
Source: Drei *et al.* (2019).

a performance reduction of (S) and (G) pillars during the 2018–2019 period. If we consider the global ESG score, its performance remains positive but it is divided by a factor of six compared to the 2014–2017 period. On the Eurozone side, the verdict is more positive. All long-short portfolio returns are positive. During the 2018–2019 period, (E) and (S) pillars yield even stronger returns comparatively to the previous period. The decline of the (G) long-short portfolio return can be partly attributed to a mean-reversion effect after an extraordinary period of impressive performance[21]. Drei *et al.* (2019) concluded that the 2018–2019 period is in line with the previous period for the Eurozone investment universe since the two periods post an annualized return around 6% in the case of the global ESG score.

How to explain these different results? Bennani *et al.* (2018) assumed that two main effects contributed to the ESG performance from 2014 to 2017: the selection effect of ESG screening and the demand effect of ESG assets. By selection effect, we think about the direct impact of extra-financial information on stock prices. By considering other risk dimensions, the ESG investor may select corporations that are better managed from social, environmental and governance viewpoints, or may avoid corporations that present extra-financial weaknesses. The underlying idea is that sooner or later these extra-financial risks have a financial impact on the performance of the corporation. The second effect is related to the supply/demand balance. Indeed, a price is the equilibrium between the supply and the demand for this stock. Bennani *et al.* (2018) found that ESG investment flows that have been observed since 2014 have largely contributed to the good performance of ESG investing over the 2014–2017 period, while the contribution of the selection effect is marginal. How to explain the 2014 break? In November 2013, the Norwegian Sovereign Wealth Fund adopted a new responsible investment policy (Dimson *et al.*, 2013). At approximately the same time, we observe a strong mobilisation of the largest European institutional investors (APG, PGGM, ERAFP, FRR, etc.), which are massively invested in European and America stocks[22]. The good performance of ESG investing during the 2014–2017 period is mainly explained by the portfolio rebalancing of these European tier-one institutional investors. The 2018–2019 period is different. Indeed, this first mobilization is followed by another mobilization of medium (or tier-two) European institutional investors, while the implication of US investors continues to be weak. Nevertheless, this second wave of investors has a low exposure on the North-American stocks. The transatlantic divided, which was observed between 2018 and 2019, is then mainly explained by the strategic asset allocation of these tier-two institutional investors. They have rebalanced their portfolios, but the trading operations mainly concerned European stocks and not American stocks. A first explanation of the American setback can then be found in these engagement differences between European and American investors. Beside the two effects (selection effect and supply/demand balance), Drei *et al.* (2019) suggested that a third factor may contribute to the ESG performance: the political and regulatory environment. The bad performance of the (E) pillar in the US may be explained by the announced withdrawal of the United States from the Paris Climate Agreement and some of the changes at the US Environmental Protection Agency (EPA). More generally, another justification of the transatlantic divided could be the public policy of the Trump administration in terms of its ESG roadmap.

Remark 37 *The original idea of the Amundi Institute studies (Bennani* et al.*, 2018; Drei* et al.*, 2019) was to frequently update the empirical relationship between ESG and performance. Nevertheless, the first study showed that there are country biases that are difficult to control. This is why the second study has focused on the investment universe of MSCI North American and EMU indexes,*

---

[21]Indeed, the annualized return was 7.9% between 2014 and 2017, compared to just 1.3% for the 2018–2019 period.

[22]For instance, the Norwegian Sovereign Wealth Fund had an exposure on US stocks greater than the exposures of the three largest US pension funds (CalPERS, CalSTRS and NYSCRF).

Figure 3.20: The monotonous assumption of the ESG-performance relationship
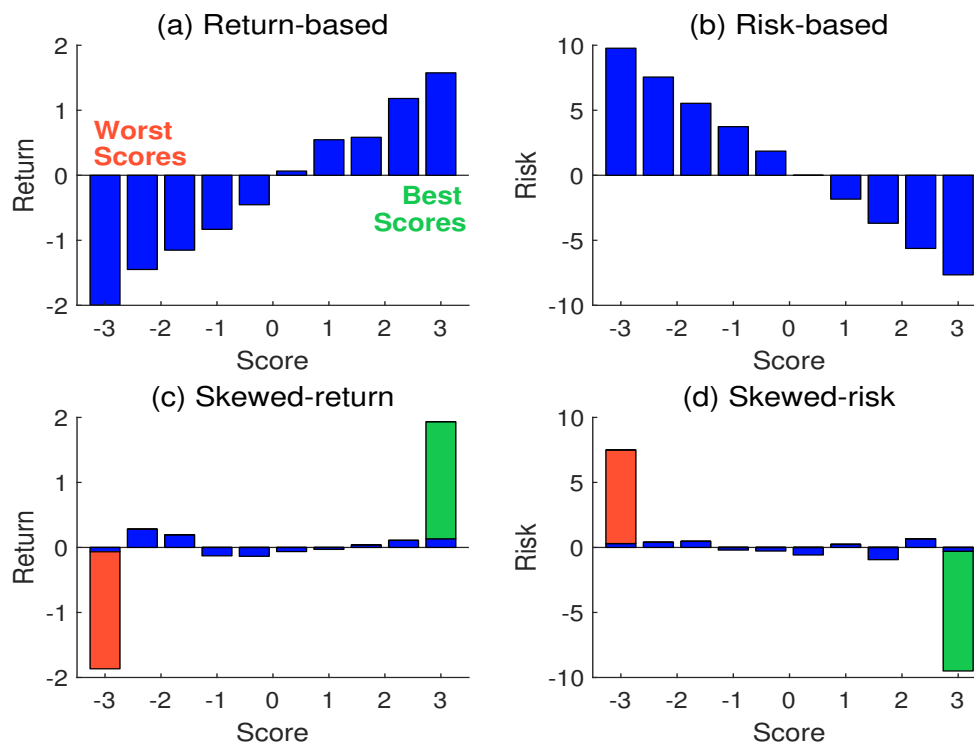
and considered the empirical relationship ESG-performance as monotonous. For instance, if the relationship is positive, we must observe $R_t(Q_1) \geq R_t(Q_2) \geq R_t(Q_3) \geq R_t(Q_4) \geq R_t(Q_5)$, while a negative relationship implies $R_t(Q_1) \leq R_t(Q_2) \leq R_t(Q_3) \leq R_t(Q_4) \leq R_t(Q_5)$. More generally, we must observe some patterns in order to interpret the results. Some of them are given in Figure 3.20. For instance, the implementation of an ESG exclusion strategy has a return-based rationale if the relationship is positive or if portfolio Q_5 underperforms the other quintile portfolios. In a similar way, the implementation of an ESG selection strategy has a return-based rationale if the relationship is positive or if portfolio Q_1 overperforms the other quintile portfolios. However, Drei et al. (2019) noticed that most of monotonous relationships that were observed during the 2010–2017 periods were no longer valid between 2018 and 2019. For instance, they found that $R_t(Q_1) \geq R_t(Q_5)$ for the ESG global score in the Eurozone, but they also found that $R_t(Q_4) \geq R_t(Q_1)$, which is a puzzling result. This ranking disorder goes beyond the binary outcome in which $Q_1 \succ Q_5$ holds or does not. Drei et al. (2019) considered that this puzzle marked the emergence of new ESG investment strategies. The $Q_1 - Q_5$ approach is representative of a static view of ESG scores, when best-in-class stocks remain best-in-class stocks and worst-in-class stocks remain worst-in-class stocks, while playing intermediary quintiles, especially the fourth quintile, seems to be related to the strategy of ESG momentum (Figure 3.21) and a dynamic view of ESG scores. During the 2018–2019 period, ESG strategies have become more complex, and this may explain the ranking disorder. This finding is in line with the results reported by GSIA (2019). In its 2018 investment review, the organization documents that the most common way to participate in sustainable investing (as measured by assets under management allocated to each strategy) is to implement negative screening, but this approach is closely tailed by ESG integration and corporate engagement strategies. Similarly, Eurosif (2018) found similar results a year before, and stated that "the main strategy is exclusion, but in the last two years the growth rate of this strategy slowed down. In contrast, best-in-class and ESG integration have had a high

growth rate". Investment strategies based on the dynamics of ESG ratings do not clearly correspond to negative or positive screening, but they are more related to ESG integration. In this approach, an improvement of an extra-financial criterion may lead to portfolio rebalancing, exactly as we observe for financial ratios. The convergence between the extra-financial approach and the traditional security analysis certainly increases the focus on the dynamics and momentum of ESG ratings, and not just their static level. In this context, analyzing the relationship between the performance of ESG investing and the static level of ESG scores is certainly outdated.

Figure 3.21: How to play ESG momentum?
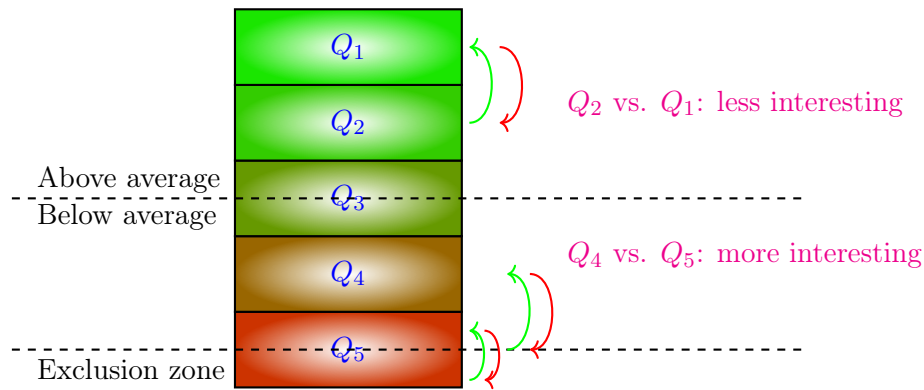
Optimized portfolios Many institutional investors implement ESG policy through passive management. In this case, they use two techniques: exclusion and optimization. The first approach consists in reducing the universe of the stock index by excluding the worst rated stocks, and then applying a capitalization-weighted scheme to form the investment portfolio. The second approach consists in improving the score of the investment portfolio with respect to the score of the benchmark portfolio, while controlling the tracking error risk. The first solution can be approximated by using the second method, implying that optimized portfolios can be used to simulate the performance of ESG passive management. This approach has been extensively used by Bennani *et al.* (2018) and Drei *et al.* (2019).

Remark 38 *Let us compare the ESG-optimized approach with the ESG-sorting method. We note $\mathbf{F}$ the probability distribution of the score $\boldsymbol{S}$. Portfolio Q_1 corresponds to best-in-class stocks $\left\{ i \in Q_1 \Leftrightarrow \boldsymbol{S}_i \geq \mathbf{F}^{-1}(80\%) \right\}$, whereas Portfolio Q_5 corresponds to worst-in-class stocks $\left\{ i \in Q_5 \Leftrightarrow \boldsymbol{S}_i \leq \mathbf{F}^{-1}(20\%) \right\}$. Moreover, the weights are uniform: $w_i(Q_j) = (5n)^{-1}$ where n is the total number of assets in the investment universe. In the case of the ESG-optimized approach, we have $w^\star(\gamma) = \Sigma^{-1}(\gamma \boldsymbol{S} + \Sigma b)$, implying that $w_i^\star(\gamma) = b_i + \gamma \left(\Sigma^{-1} \boldsymbol{S}\right)_i$. Therefore, the benchmark weights are tilted by the inverse of the covariance matrix times the vector of ESG scores. For example, if we assume that the covariance matrix is diagonal and $\boldsymbol{S}_i \sim \mathcal{N}(0, 1)$, we obtain:*

$$w_i^\star(\gamma) = b_i + \gamma \frac{\boldsymbol{S}_i}{\sigma_i^2}$$

A positive score increases the benchmark weight whereas γ controls the discrepancy. If $\gamma = 0$, the optimal portfolio is the benchmark. When γ tends to $+\infty$, the optimal weight is proportional to the score divided by the variance. If we add the long-only constraint, the optimization problem selects the stocks such that the ratio $\boldsymbol{S}_i/\sigma_i^2$ is greater than a threshold that depends on the parameter γ.

Box 3.6: ESG-optimized portfolios

We note b the benchmark, $\boldsymbol{S}$ the vector of ESG scores and Σ the covariance matrix. We consider the following optimization problem:

$$w^\star(\gamma) = \arg\min \frac{1}{2}\sigma^2(w \mid b) - \gamma\boldsymbol{S}(w \mid b)$$

where $\sigma^2(w \mid b) = (w - b)^\top \Sigma (w - b)$ and $\boldsymbol{S}(w \mid b)$ are the ex-ante tracking error variance and the ESG excess score of portfolio w with respect to the benchmark b. Since we have:

$$\boldsymbol{S}(w \mid b) = (w - b)^\top \boldsymbol{S} = \boldsymbol{S}(w) - \boldsymbol{S}(b)$$

we obtain the following optimization function:

$$w^\star(\gamma) = \arg\min \frac{1}{2}w^\top \Sigma w - w^\top(\gamma\boldsymbol{S} + \Sigma b)$$

The ESG-variance efficient frontier is defined by the parametric curve $\left(\sigma^2(w^\star(\gamma) \mid b), \boldsymbol{S}(w^\star(\gamma) \mid b)\right)$ with $\gamma \geq 0$. The QP form is given by $Q = \Sigma$ and $R = \gamma\boldsymbol{S} + \Sigma b$. If we target an ESG excess score, for instance $\boldsymbol{S}(w \mid b) \geq \Delta\boldsymbol{S}^\star$, we set $\gamma = 0$ and add the inequality constraint $\boldsymbol{S}^\top w \geq \boldsymbol{S}(b) + \Delta\boldsymbol{S}^\star$. The QP form is given by $Q = \Sigma$, $R = \Sigma b$, $C = -\boldsymbol{S}^\top$ and $D = -(\boldsymbol{S}(b) + \Delta\boldsymbol{S}^\star)$. If we use the traditional long-only constraint ($\mathbf{1}^\top w = 1$ and $w_i \geq 0$), we have $A = \mathbf{1}^\top$, $B = 1$ and $w^- = \mathbf{0}$.

We can then compute the value of γ in order to retrieve the stock selection[23] given by portfolios Q_1, $Q_1 + Q_2$, etc. This is why ESG-optimized and ESG-sorting approaches are related and generally give similar results.

In Figure 3.22, we report the ESG-variance efficient frontier estimated by Bennani *et al.* (2018). It represents the relationship between the excess score and the tracking error volatility for the MSCI World universe. For example, improving the score[24] of the index portfolio by 0.5 implies accepting a tracking error of 32 bps on average, and an excess score of 1.0 leads to a tracking error of 85 bps. Using a risk attribution analysis, the authors also show that the governance pillar generates more tracking error than the environmental and social pillars[25]. These results mean that ESG passive management requires taking on a significant tracking error risk with respect to capitalization-weighted benchmarks.

Figure 3.24 presents the performance of ESG optimized portfolios with respect to the excess score. We notice that the integration of ESG in passive management reduced its performance between 2010 and 2013, whereas it improved its annualized return between 2014 and 2017. For instance, an excess score of 1.0 led to an excess return of -34 bps during the first period and $+45$ bps during the second period. We also notice that the relationship between excess score and excess return is not necessarily monotonous. For instance, targeting an excess score of 1.5 instead of 1.0 results in reducing the excess return from 45 bps to 19 bps in the second period. This is most likely due to the diversification effect. Indeed, by increasing the excess score, we reduce the

---

[23]See Exercise 3.4.1 on page 203.

[24]We recall that these studies use z-scores, meaning that the range is between -3 and $+3$.

[25]On average, optimized portfolios with the Ⓖ score have a 50% larger tracking error than with Ⓔ and Ⓢ scores (see Figure 3.23).

Figure 3.22: Efficient frontier of ESG-optimized portfolios (MSCI World, 2010-2017, global score)

Figure 3.23: Efficient frontier of ESG-optimized portfolios (MSCI World, 2010–2017, individual pillars)

Figure 3.24: Annualized excess return of ESG-optimized portfolios (MSCI World, 2010–2017, global score)

Figure 3.25: Annualized excess return of ESG-optimized portfolios (MSCI World, 2010–2013, individual pillars)

Figure 3.26: Annualized excess return of ESG-optimized portfolios (MSCI World, 2014–2017, individual pillars)
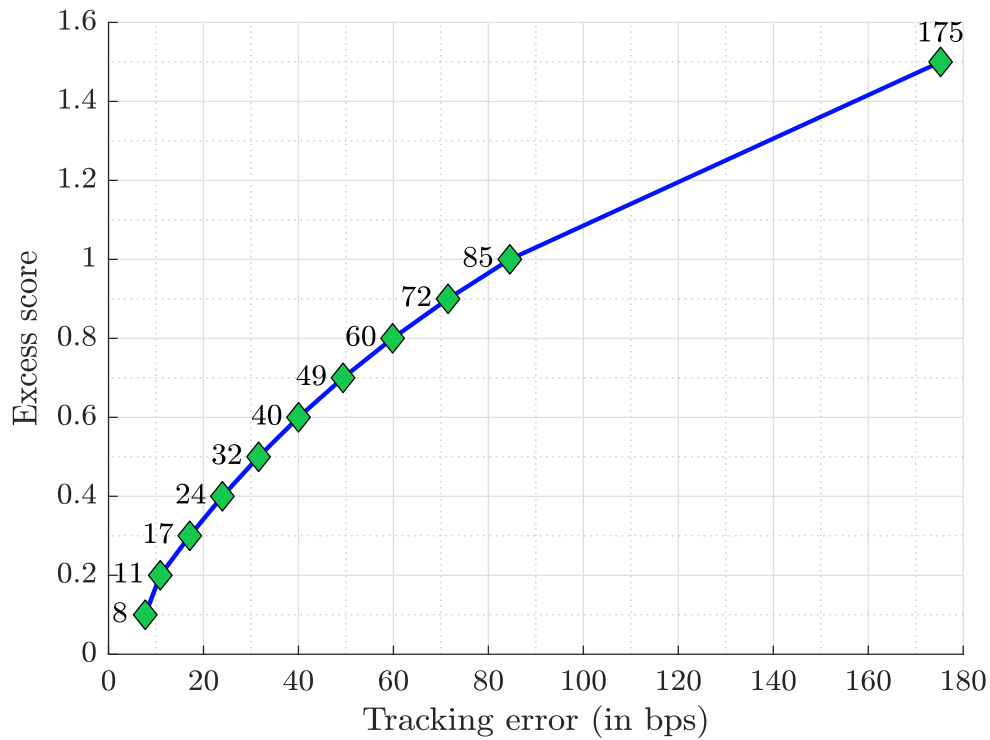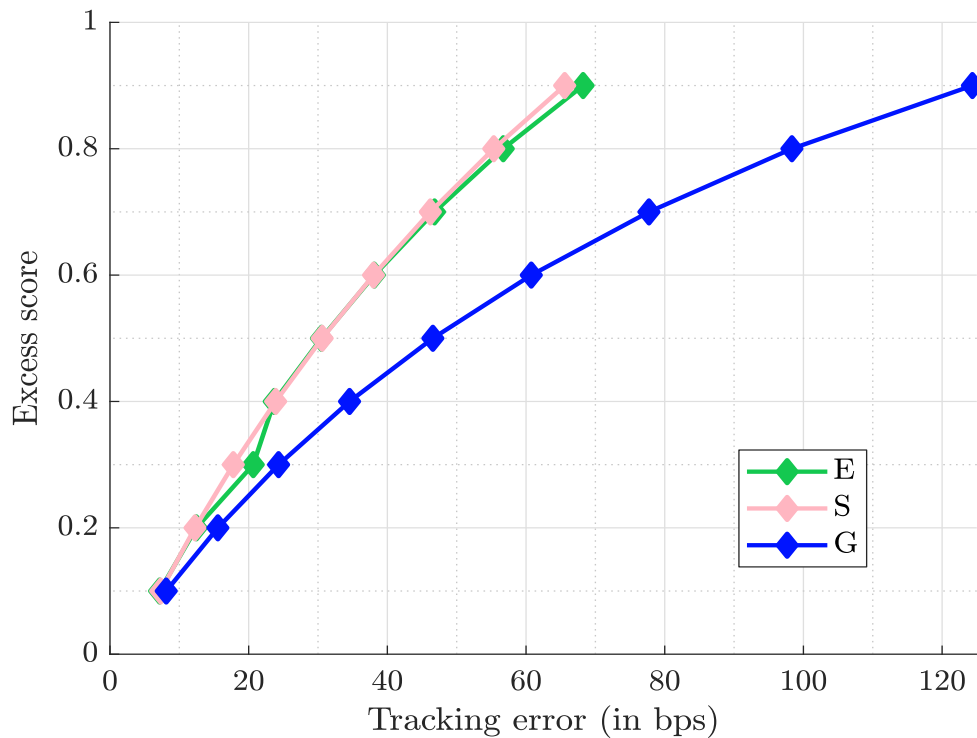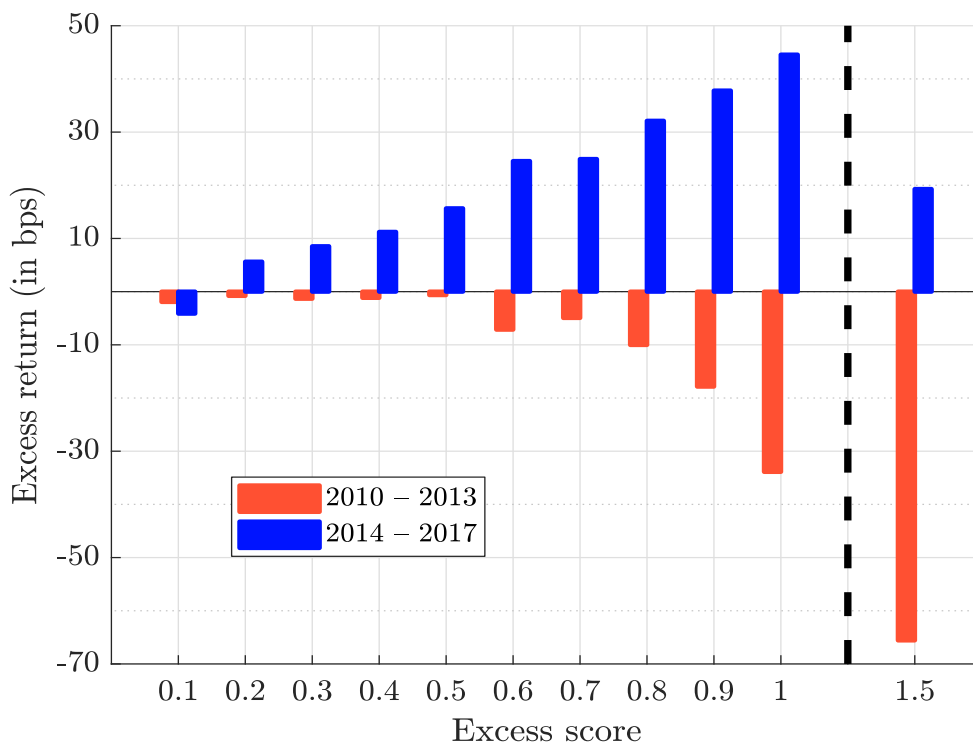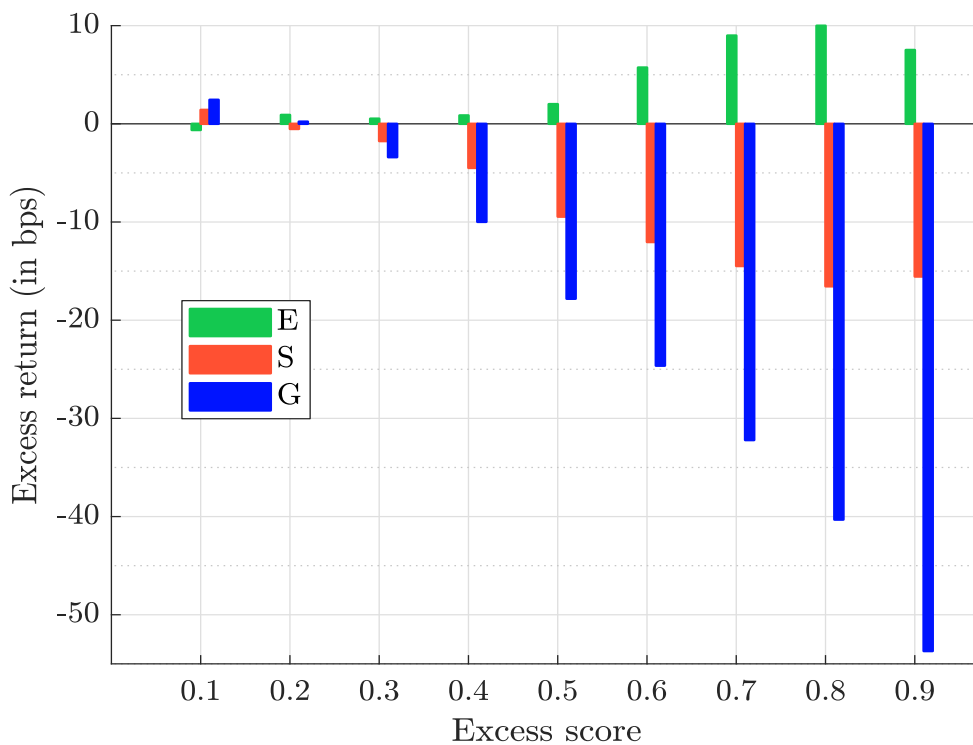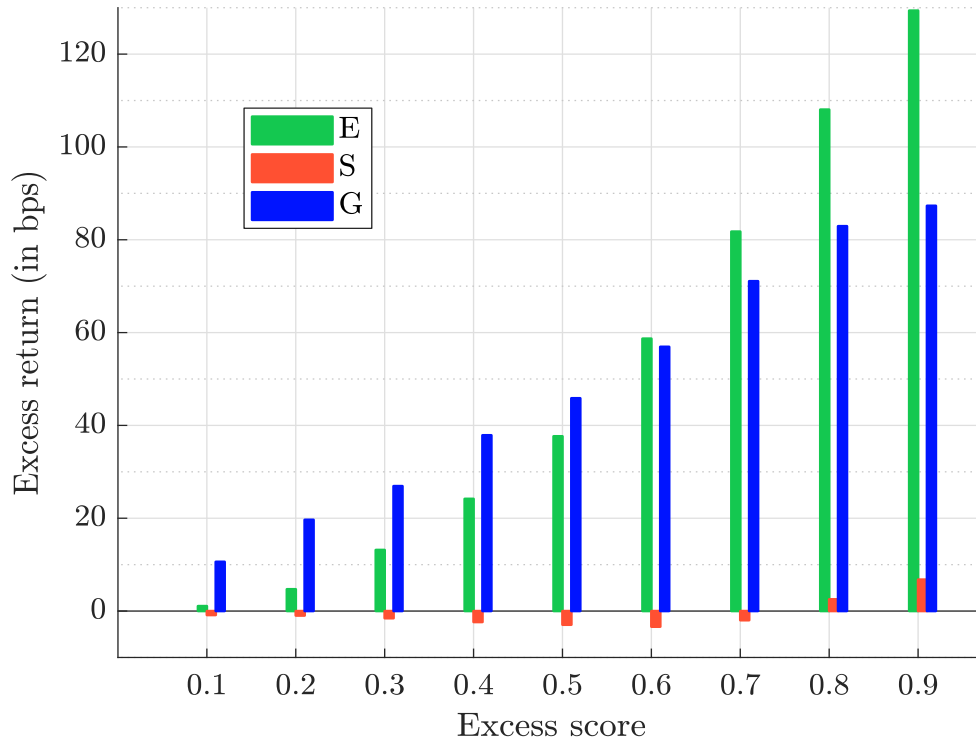
Source: Bennani *et al.* (2018).

Figure 3.27: Annualized excess return in bps of ESG-optimized portfolios (MSCI North America and EMU, 2010–2017)
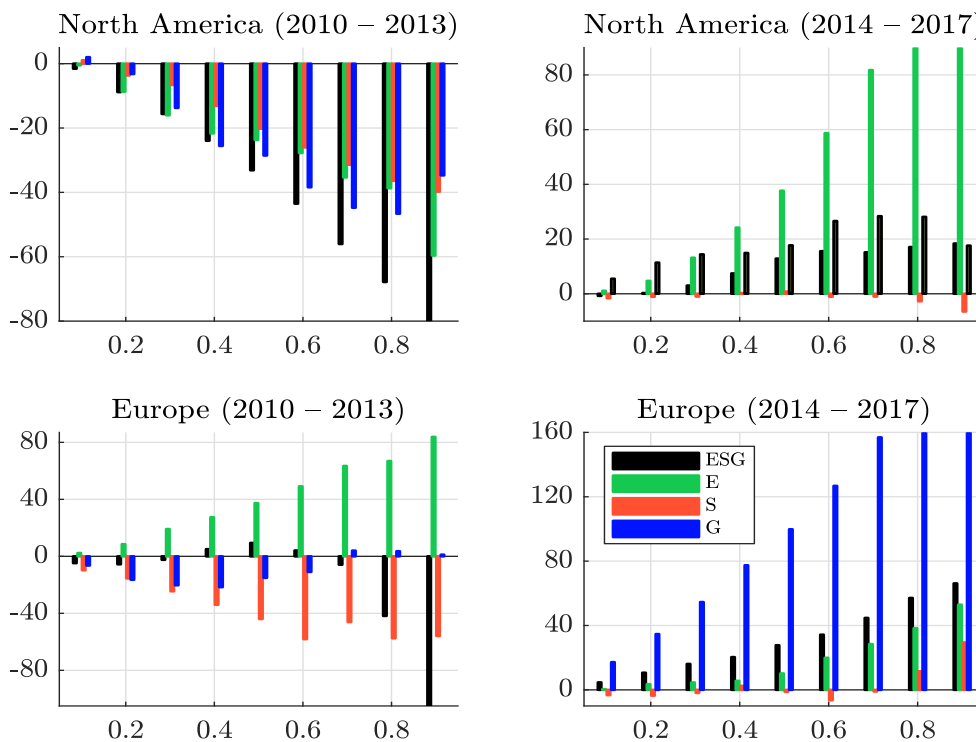
Source: Bennani *et al.* (2018).

number of positions in the invested portfolios. There comes a threshold where the gains from the ESG screening are offset by the losses resulting from the diversification reduction. If we consider the individual pillars, Bennani *et al.* (2018) retrieve the main conclusions that they have found for active management. For the MSCI World universe, all the pillars destroyed value between 2010–2013, except the environmental pillar for which results are neutral or slightly positive. This is particularly true for the governance pillar, whose underperformance is about two/three times greater the underperformance of the overall ESG score (Figure 3.25). For the 2014–2017 period, the story changes. Every score creates an outperformance, except the social pillar (Figure 3.26). If we consider the North America and Europe investment universes[26], the performance of optimized portfolios is in line with the performance of stock picking portfolios (Figure 3.27). During the 2010–2013 period, only the **(E)** score would have generated outperformance in Europe. In this region, the authors found that the performance of the ESG score was also neutral when targeting low tracking error risk (less than 60 bps) or low excess score (less than 0.8). In all other cases, we observe a negative excess return, especially in North America. Between 2014 and 2017, we obtain opposite results. All the scores generate an outperformance, except the social pillar. The results are more significant in Europe than in North America. To summarize, the two big winners were the environmental pillar in North America and the governance pillar in Europe between 2014 and 2017.

The updated study of Drei *et al.* (2019) has confirmed most of the results found by Bennani *et al.* (2018), especially the trade-off between excess score and tracking error risk, and the reversal phenomenon of the ESG-performance relationship, which is negative when targeting a high excess score. This reversal phenomenon is most likely due to the diversification effect. Indeed, by increasing the excess score, we reduce the number of positions held in the managed portfolio. Therefore, there comes a threshold where the gains from the ESG screening are offset by the losses resulting from the diversification reduction. Since the relationship between quintile portfolios and performance is not monotonous, Drei *et al.* (2019) noticed that the performance of ESG-optimized portfolios is less impressive between 2018 and 2019 than during the 2014–2017 period. This is why they observe a reduction in the maximum excess return. Focusing on the Eurozone investment universe, where the loss of diversification is reached faster than in North America, they conclude that "*optimized portfolios generate poorer results (except for the social pillar)*", and more generally that "*risk-return profiles are less interesting than before*". Therefore, the dynamic view of ESG investing implies that the performance of ESG-optimized portfolios is not necessarily in line with the performance of $Q_1 - Q_5$ sorted portfolios, because of the impact of the other sorted portfolios, in particular the fourth quintile portfolio.

A new risk factor?   Previously, we have seen that the long/short $Q_1 - Q_5$ strategy has generated a positive alpha between 2014 and 2019, whereas ESG investing has penalized ESG investors between 2010 and 2013. When we speak about alpha generation, we generally refer to factor investing. Indeed, factor investing makes the difference between the financial performance coming from systematic factors and the financial performance coming from specific factors. Said differently, factor investing makes the difference between alpha and beta returns. In factor investing, beta (or systematic) factors correspond to the common risk factors that explain a significant part of the cross-section of stock returns. Since ESG changes the landscape of asset management, we may wonder whether ESG has become a new risk factor and must be integrated into a factor investing framework, or whether it remains an alpha strategy. To answer this question, Roncalli (2020b) use the single-factor model:

$$R_{i,t} = \alpha_{i,j} + \beta_{i,j}\mathcal{F}_{j,t} + \varepsilon_{i,t}$$

---

[26]Bennani *et al.* (2018) have merged Eurozone and Europe-ex-EMU stocks into the same investment universe because the tracking error optimization forces the portfolio to be more or less country-neutral.

where $R_{i,t}$ is the return of stock i at time t, $\mathcal{F}_{j,t}$ is the value of the jth common risk factor at time t and $\varepsilon_{i,t}$ is the idiosyncratic risk. The coefficients $\alpha_{i,j}$ and $\beta_{i,j}$ are estimated by the method of ordinary least squares. For each stock, we compute the coefficient of determination:

$$\mathfrak{R}_{i,j}^2 = 1 - \frac{\text{var}\left(\varepsilon_{i,t}\right)}{\text{var}\left(R_{i,t}\right)}$$

We can then calculate the average proportion of the return variance explained by the common factor:

$$\bar{\mathfrak{R}}_j^2 = \frac{1}{n}\sum_{i=1}^{n}\mathfrak{R}_{i,j}^2$$

We consider the standard factors derived from a factor investing framework: size, value, momentum, low-volatility and quality. These factors $\mathcal{F}_{j,t}$ are built using the Fama-French methodology of sorted portfolios. Contrary to the academic literature, a long-only framework is used, which is the usual approach of institutional investors. This means that the factors correspond to Q_1 portfolios or best-in-class stocks. Moreover, we consider the traditional market factor, which corresponds to the capitalization-weighted portfolio. All the analyses use weekly returns. Results are given in Table 3.14. We read these figures as follows: between 2010 and 2013, the market risk factor explains 40.8% of the dispersion of North American stock returns, this figure is 39.3% if we consider the size factor, etc. We observe that ESG has been a strong contender as a standalone factor and competes with the market risk factor. On average, since 2014, the market risk factor explains 28.6% of the cross-section variance, whereas the ESG factor has an explanatory power of 27.4% in North America. In the Eurozone, these figures are respectively 36.3% and 35.3%. Moreover, it has more explanatory power than the other risk factors both in North America and the Eurozone during the two periods: 2010–2013 and 2014–2019.

Table 3.14: Results of cross-section regression with long-only risk factors (single-factor linear regression model, average $\mathfrak{R}^2$)

Factor	North America		Eurozone	
	2010–2013	2014–2019	2010–2013	2014–2019
Market	40.8%	28.6%	42.8%	36.3%
Size	39.3%	26.1%	37.1%	23.3%
Value	38.9%	26.7%	41.6%	33.6%
Momentum	39.6%	26.3%	40.8%	34.1%
Low-volatility	35.8%	25.1%	38.7%	33.4%
Quality	39.1%	26.6%	42.4%	34.6%
ESG	40.1%	27.4%	42.6%	35.3%

Source: Roncalli (2020b).

We now consider a multi-factor model:

$$R_{i,t} = \alpha_i + \sum_{j=1}^{m}\beta_{i,j}\mathcal{F}_{j,t} + \varepsilon_{i,t}$$

where m is the number of risk factors. In this approach, we compare the CAPM or one-factor model, the standard five-factor model based on size, value, momentum, low-volatility and quality risk factors, and the six-factor model, which consists in adding the ESG factor to the universe of

the five alternative risk factors. In Table 3.15, we verify that the five-factor model increases the proportion of systematic risk with respect to the CAPM. For example, the CAPM and the 5F model explain respectively 28.6% and 38.4% of the cross-section variance in North America during the second period. Adding the ESG factor has a minor impact between 2014 and 2019: 39.7% versus 38.4% in North America and 45.8% versus 45.0% in the Eurozone. This means that the ESG factor does not significantly improve the five-factor model. However, if we apply statistical tests of significance to the six-factor model, we find that ESG is statistically significant in the Eurozone, but not in North America. We may conclude that ESG could be a risk factor in the Eurozone, but not in North America.

Table 3.15: Results of cross-section regression with long-only risk factors (multi-factor linear regression model, average $\mathfrak{R}^2$)

Model	North America		Eurozone	
	2010–2013	2014–2019	2010–2013	2014–2019
CAPM	40.8%	28.6%	42.8%	36.3%
5F model	46.1%	38.4%	49.5%	45.0%
6F model (5F + ESG)	46.7%	39.7%	50.1%	45.8%

Source: Roncalli (2020b).

The previous results may be disturbing. Indeed, cross-section regressions show that ESG is a very good single factor, but the added value of ESG in a multi-factor framework is limited. The difference between the two approaches is the cross-correlation between risk factors that are taken into account into the cross-section multi-factor regression. In order to better understand these results, Roncalli (2020b) consider a factor picking (or a factor selection) approach. This approach is similar to the multi-factor approach, but a lasso penalized regression is used in place of the traditional least squares regression:

$$\left\{ \hat{\alpha}_i, \hat{\beta}_{i,1}, \ldots, \hat{\beta}_{i,m} \right\} = \arg\min \left\{ \frac{1}{2} \operatorname{var}\left(\varepsilon_{i,t}\right) + \lambda \left\| \beta_i \right\|_1 \right\}$$

The advantage is that we can control the factor intensity of the multi-factor portfolio. Therefore, we obtain a factor selection procedure. Beginning with a low-factor intensity ($\lambda \approx \infty$), we can determine which risk factor is the most important. Then, we increase the factor intensity in order to establish an ordering between risk factors. When the factor intensity reaches 100% ($\lambda = 0$), we obtain the same results calculated previously with the linear regression. The results are reported in Figures 3.28 and 3.29 for the period 2014–2019. In North America, we notice that quality is the first selected factor, followed by ESG, momentum, value, and finally low-volatility. Therefore, ESG is the second selected factor in North America. Thus, ESG should be a significant factor when building a multi-factor portfolio. However, we observe that the ESG beta first increases and then decreases when we increase the factor intensity. When the factor intensity reaches 100%, ESG represents a low exposure. Therefore, a part of the ESG exposure has been replaced by an exposure to other risk factors. This means that ESG has a high contribution in a low-diversified portfolio, but it is somewhat redundant in an already well-diversified portfolio. In the case of the Eurozone, we face a different situation. ESG is the first selected factor and remains an important factor even if we increase the factor intensity. For instance, it is more significant than momentum and low-volatility.

These different results (single-factor, multi-factor and factor picking) show that ESG investing remains an alpha strategy in North America. It may have generated outperformance, but the ESG risk factor cannot explain the dispersion of stock returns better than the standard five-factor

Figure 3.28: Factor picking (MSCI North America, 2014–2019, global score)
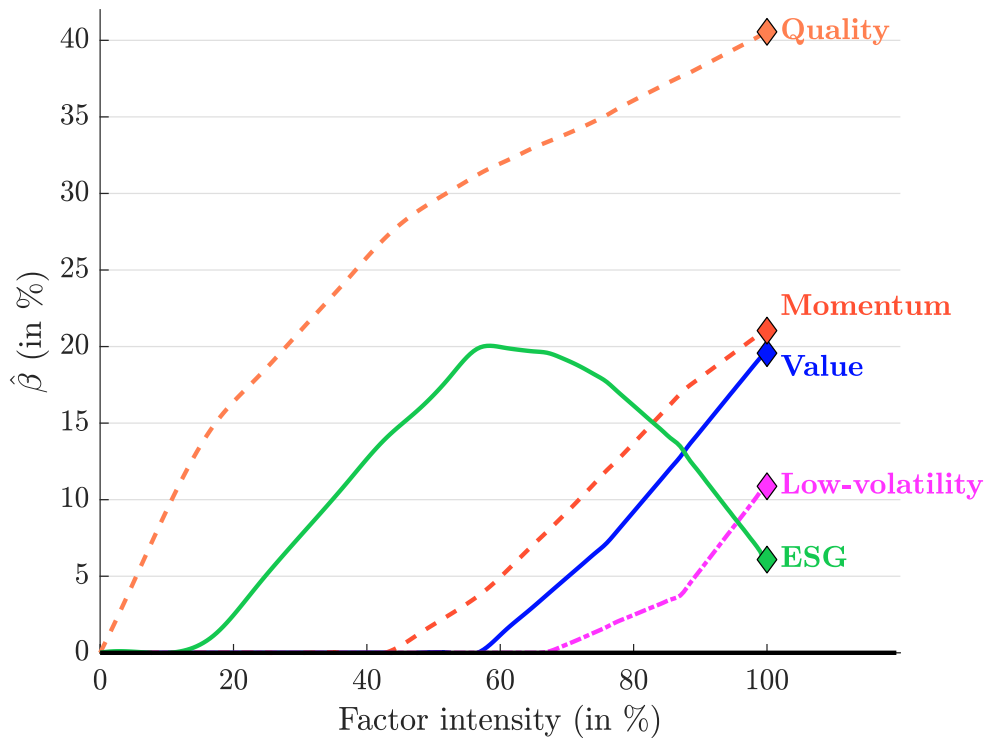
Figure 3.29: Factor picking (MSCI EMU, 2014–2019, global score)
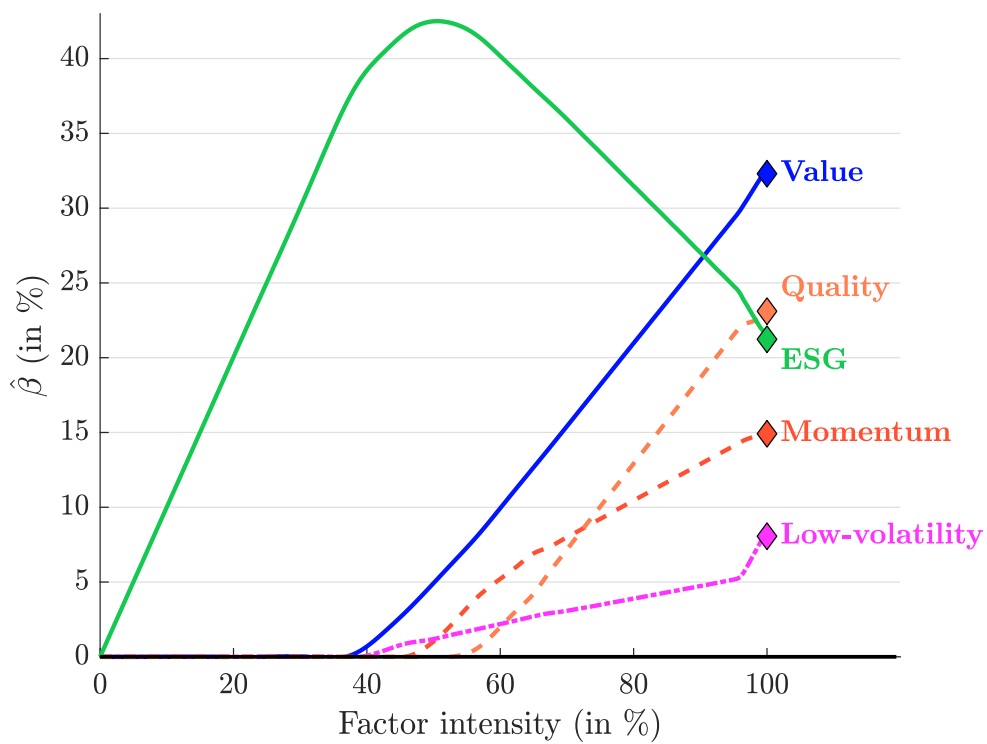
risk model. This implies that introducing ESG in a multi-factor portfolio, which is already well-diversified, adds very little value. This is clearly the definition of an alpha strategy. On the contrary, we notice that ESG is a significant factor in a Eurozone multi-factor portfolio. We may then improve the diversification of multi-factor portfolios by integrating an ESG risk factor. As such, in the Eurozone, it seems that an ESG strategy is more a beta strategy than an alpha strategy.

Remark 39 *These last observations can be related to the development of factor investing, for example low-volatility and quality risk factors (Roncalli, 2017). Low-volatility strategies have been known for many years, but they primarily emerged in the asset management industry between 2003 and 2004 after the* dot.com *bubble. Initially, low-volatility strategies were considered as alpha strategies. After the 2008 Global Financial Crisis, they were massively implemented, thereby becoming beta strategies. The case of the quality anomaly is similar. This shows that there is not a clear boundary between alpha and beta. When an alpha strategy is massively invested, it has an enough impact on the structure of asset prices to become a risk factor. The alpha/beta status of ESG strategies is related to investment flows. Indeed, an alpha strategy becomes a common market risk factor once it represents a significant part of investment portfolios and explains the cross-section dispersion of asset returns. This may explain that ESG is more a risk factor in the Eurozone than in North America.*

Table 3.16: Performance of ESG equity indexes (MSCI World, 2010–2022)

Year	Return (in %)			Alpha (in bps)	
	BM	ESG	SRI	ESG	SRI
2010	11.8	10.7	10.6	−109	−114
2011	−5.5	−5.4	−5.5	12	2
2012	15.8	14.5	13.2	−135	−258
2013	26.7	27.6	27.4	89	71
2014	4.9	4.9	3.9	−6	−102
2015	−0.9	−1.1	−1.6	−23	−71
2016	7.5	7.3	7.7	−26	18
2017	22.4	21.0	23.6	−142	124
2018	−8.7	−7.8	−6.7	94	199
2019	27.7	28.2	29.8	48	209
2020	15.9	15.3	19.9	−61	396
2021	21.8	24.7	27.0	288	523
2022	−18.1	−19.6	−22.5	−143	−436
3Y	4.9	5.0	5.7	2	73
5Y	6.1	6.4	7.4	31	125
7Y	8.5	8.5	9.6	1	110
10Y	8.9	8.9	9.5	5	64

Equity indexes

Another way to illustrate that the time-varying property of the performance of ESG investing is to analyze the annualized return of equity indexes. In Table 3.16, we compare the MSCI World capitalization-weighted index (BM) with the MSCI World ESG Leaders and SRI indexes[27]. For

---

[27]*"MSCI ESG Leaders indexes target companies that have the highest ESG rated performance in each sector of the parent index. MSCI SRI indexes are designed to represent the performance of companies with high ESG ratings.*

each index, we report the annualized return in % and we also compute the alpha in bps with respect to the CW parent index. For instance, the 2022 return was -18.1% for the MSCI World index (BM), -19.6% for the MSCI World ESG Leaders index and -22.5% for the MSCI World SRI index. Therefore, the alpha of ESG and SRI indexes is negative and is respectively equal to -143 and -436 bps. If we consider the last thirteen years, the benchmark index has overperformed the ESG index eight times, and the SRI index only five times. The 3Y, 5Y, 7Y and 10Y annualized returns are greater for the ESG and SRI indexes than for the BM index. These results clearly highlights that ESG investing may create alpha in some periods. Moreover, the relative performance depends on the construction of the ESG index. Indeed, we do not observe the same patterns between ESG Leaders and SRI indexes.

In Table 3.17, we confirm that the performance of ESG investing depends on several factors. For instance, the region has a significant impact. The MSCI EMU SRI index has created a positive alpha every year between 2015 and 2021 with an average of 393 bps per year. For the MSCI EM ESG Leaders index, the period of euphoria is between 2010 and 2017 with an average alpha of 423 bps per year. The choice of the ESG scoring model is another factor. Indeed, if we consider the S&P 500 ESG index, it has generated a positive alpha in 2014, 2020 and 2022 whereas the MSCI USA ESG Leaders index posted a negative alpha in these years (19 vs. -49 bps in 2014, 138 vs. -251 bps in 2020 and 43 vs. -73 bps in 2022).

Figure 3.30: Alpha return of several ESG equity indexes (in bps)

Remark 40 *In Figure 3.30, we have reported the distribution of alpha returns of ESG equity indexes over time. This perfectly illustrates that "ESG investing has its good and bad times".*

They employ a best-in-class selection approach to target the top 25% companies in each sector according to their MSCI ESG Ratings" (www.msci.com/our-solutions/indexes/esg-indexes).

Table 3.17: Performance of ESG equity indexes

Year	MSCI USA					MSCI EMU					MSCI EM			S&P 500		
	Return			Alpha		Return			Alpha		Return		Alpha	Return		Alpha
	BM	ESG	SRI	ESG	SRI	BM	ESG	SRI	ESG	SRI	BM	ESG	ESG	BM	ESG	ESG
2010	14.8	11.0	11.0	−373	−373	2.4	3.6	1.6	120	−83	18.9	25.5	665	14.4		
2011	1.4	0.6	0.1	−79	−129	−14.9	−13.3	−14.9	162	−3	−18.4	−13.1	536	1.5		
2012	15.3	12.6	10.2	−270	−510	19.3	19.5	23.0	20	371	18.2	21.2	299	15.2		
2013	31.8	32.8	31.7	98	−6	23.4	20.4	23.1	−301	−27	−2.6	1.3	386	31.5	29.7	−182
2014	12.7	12.2	11.8	−49	−90	4.3	3.7	2.1	−57	−218	−2.2	4.8	703	13.0	13.2	19
2015	0.7	−2.0	−3.7	−264	−438	9.8	14.0	14.7	415	493	−14.9	−12.3	262	0.7	−0.2	−91
2016	10.9	11.0	12.6	10	174	4.4	7.8	9.4	345	498	11.2	13.4	221	11.2	11.8	53
2017	21.2	19.8	23.4	−137	221	12.5	12.4	14.0	−9	154	37.3	40.4	309	21.1	20.5	−62
2018	−5.0	−3.7	−1.9	137	314	−12.7	−10.2	−8.0	248	472	−14.6	−15.0	−42	−4.9	−4.5	40
2019	30.9	30.9	32.5	2	159	25.5	26.3	29.9	85	444	18.4	19.8	134	30.7	32.5	184
2020	20.7	18.2	25.8	−251	508	−1.0	0.3	0.0	127	100	18.3	20.1	178	17.8	19.1	138
2021	26.5	31.2	35.3	473	882	22.2	22.4	28.0	28	588	−2.5	−2.1	48	28.2	31.2	305
2022	−19.8	−20.6	−23.6	−73	−373	−12.5	−12.5	−16.0	−6	−349	−20.1	−22.8	−270	−18.5	−18.1	43
3Y	7.0	7.2	9.2	23	220	1.9	2.4	2.5	49	56	−2.7	−3.2	−47	7.1	8.6	146
5Y	8.7	9.2	11.1	46	232	3.0	4.0	5.2	102	216	−1.4	−1.6	−16	8.8	10.1	128
7Y	10.7	10.9	13.0	17	223	4.5	5.7	7.0	121	248	5.2	5.7	51	10.9	11.8	93
10Y	11.8	11.7	12.8	−15	97	6.8	7.7	8.7	94	199	1.4	3.2	176	11.9	12.3	42

Source: MSCI, Factset & Author's calculations.

3.2.2 Fixed-income markets

Compared to stock markets, there are few studies that analyze the impact of ESG screening on fixed-income markets. For instance, Menz (2010) investigated the relationship between the valuation of Euro corporate bonds and corporate social responsibility and concluded that "*CSR has apparently not yet been incorporated into the pricing of corporate bonds*". In a similar way, a neutral or slightly positive effect of socially responsible investment was demonstrated by Derwall and Koedijk (2009) when they compared the performance of SRI and conventional bond funds. This overall neutrality, sometimes associated with a lack of maturity in incorporating ESG information into the bond market, was also highlighted by Goldreyer *et al.* (1999), Bauer *et al.* (2005) and Cortez *et al.* (2009). In the CAPM approach, active management performance is captured by measuring the alpha. However, Lin *et al.* (2019) constructed industry- and credit rating-controlled quintile portfolios but found no significant evidence of ESG factor contribution to a positive alpha in the bond market. On the contrary, Oikonomou *et al.* (2014) found that corporate social performance is rewarded on the corporate debt market. Results of Leite and Cortez (2016) are slightly positive, but highly dependent on the country. For instance, they concluded that "*French SRI bond funds match the performance of their conventional peers, German funds slightly outperform and UK funds significantly underperform conventional funds*". Polbennikov *et al.* (2016) also noted a slight outperformance of high ESG rated over low ESG rated bonds after controlling for varying risk exposures. More recently, Gerard (2019) and Pereira *et al.* (2019) found that there is no link between ESG and performance in the corporate bond markets.

Simulated results

There are three main reasons that the relationship between ESG and bond returns could be different than the relationship between ESG and stock returns:

1. The first reason concerns the sensitivity to ESG. Indeed, ESG criteria are very important for a stockholder, because ESG risks can affect the long-term business risk and can strongly impact the value of the equity. In the case of a debtholder, his main objective is to manage the default risk, which is more a short-run risk. Therefore, the concept of active ownership does not really apply to fixed-income instruments. For instance, the absence of voting rights reduces dramatically the impact of engagement policies when ESG investors focus on the bond market. From a theoretical point of view, it is then generally accepted that stockholders are more sensitive to ESG factors than bondholders. This may explain the lower integration of ESG in fixed-income markets. Other factors can explain the difference to ESG sensitivity. For instance, stock prices react faster and more sharply than bond prices to negative events, news flows and market sentiment, while bond prices are mainly driven by long-term fundamentals. Moreover, the liquidity difference between the two markets and the buy-and-hold strategy imply that equities are most commonly traded than bonds. In this context, investors may have the view that fixed-income markets do not price in ESG issues.

2. The second reason concerns the impact of ESG investment flows. In the case of stocks, they create a pressure on stock prices. In the case of bonds, we observe two effects. Investment flows can of course impact the dynamics of credit spreads and then create a pressure on bond prices, but they have also an impact on the primary market by reducing or increasing the coupon. ESG investment flows have then an effect on the carry, and a high ESG score generally implies a carry reduction because of the supply/demand balance.

3. The third reason is related to the correlation between ESG ratings and credit ratings. A

part of extra-financial information is already incorporated into credit ratings.   In a bond investment universe, the conditional probability distribution of ESG scores is then different to the unconditional probability distribution of ESG scores.   Most of the times, we distinguish IG and HY bonds.   Since we observe a positive correlation between ESG and credit ratings, it follows that there are more worst-in-class issuers in the HY universe than in the IG universe. So, we observe a distortion of the ESG scores, which are no longer sector-neutral in fixed-income.   Moreover, the average ESG-score will certainly not be equal to zero.   It is positive for IG bonds and negative for HY bonds[28].

These several reasons explain that ESG scoring is more incorporated in equity portfolio management, while ESG integration is generally limited to exclusions in bond portfolio management.   Moreover, the development of pure-play ESG securities has generally induced a segmentation in the construction of bond portfolios.   On one hand, we find a core portfolio, which corresponds to a global aggregate fixed-income strategy and uses minimum ESG criteria without really promoting ESG analysis.   On the other hand, a satellite portfolio is invested on green and social bonds with the goal to implement an impact investing strategy.

Figure 3.31: Probability density function of ESG scores
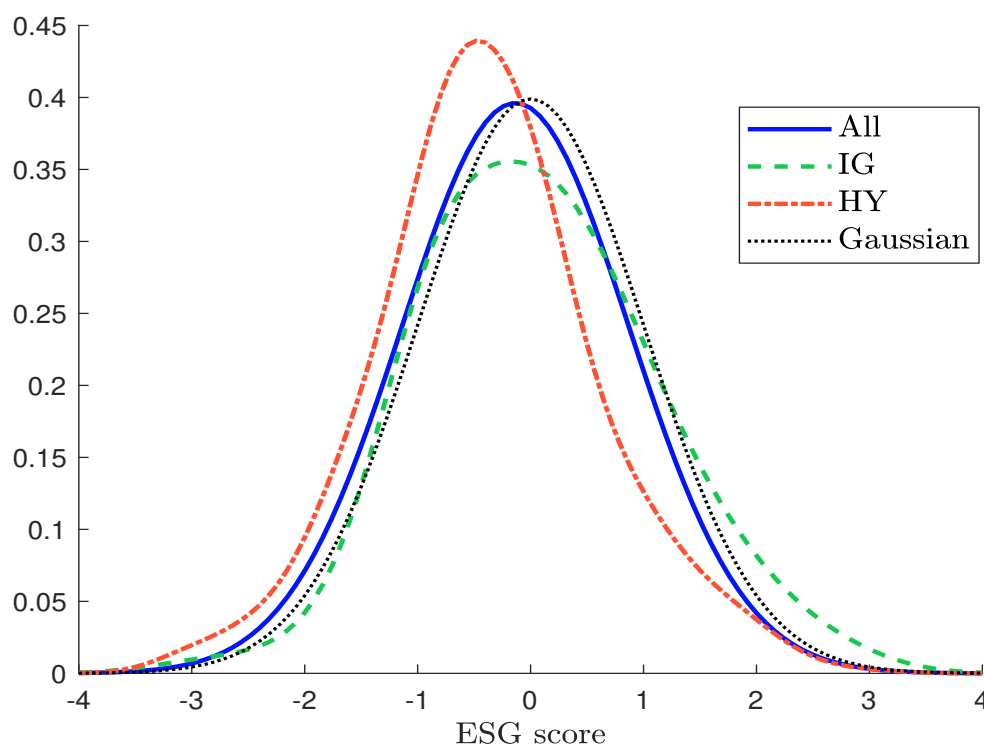
Source: Ben Slimane *et al.* (2019b).

Sorted portfolios   Ben Slimane *et al.* (2019b) has applied the sorted and optimized approach of Bennani *et al.* (2018) to the universe of bonds from the Intercontinental Exchange Bank of America Merrill Lynch large cap investment grade corporate bond index.   The study period is from January

---

[28]In Figure 3.31, we report the estimated density function of Amundi ESG z-scores within IG and HY universes. The empirical standard deviation of z-scores is equal to 1.01 whatever the bond universe. On the contrary, we find that the empirical mean is respectively equal to 0.02 and -0.38 for IG and HY bonds.

2010 to August 2019. Every month, they rank the bonds with respect to their ESG score, and form five quintile portfolios. Portfolio Q_1 corresponds to the 20% best-ranked bonds, whereas portfolio Q_5 corresponds to the 20% worst-rated bonds. By construction, sorted portfolios are sector-neutral. However, the authors perform some clustering because some sectors are small. Sorted portfolios are then built using the same structure of weights as the benchmark, while the selected bonds are equally-weighted within a sector. Moreover, each portfolio is rebalanced on a monthly basis, implying that the portfolio is invested the first trading day of the month and is held for the entire month. In Figure 3.32, we report the difference of returns between the best-in-class portfolio and the worst-in-class portfolio. We use two performance measures. The total return (TR) corresponds to the mark-to-market return of the portfolio, including bond price variations and coupon effects. The credit return (CR) indicates the return in excess of the total return of a risk-matched basket of governments or interest rate swaps, thus neutralizing the interest rate and yield curve risk and isolating the portion of performance attributed solely to credit and optionality risks. If we focus on the total return measure, all $Q_1 - Q_5$ EUR-denominated portfolios had a positive performance, both in 2010–2013 and 2014–2019. If we consider the credit return measure, ESG and Ⓢ long/short portfolios exhibited a negative performance before 2014, whereas all the portfolios have posted a positive performance during the second period. If we consider the universe of USD-denominated investment grade corporate bonds, the worst-in-class portfolio has generally outperformed the best-in-class portfolio except for the Governance pillar between 2010 and 2013.

Figure 3.32: Annualized return in bps of the long short $Q_1 - Q_5$ strategy (IG, 2010–2019)
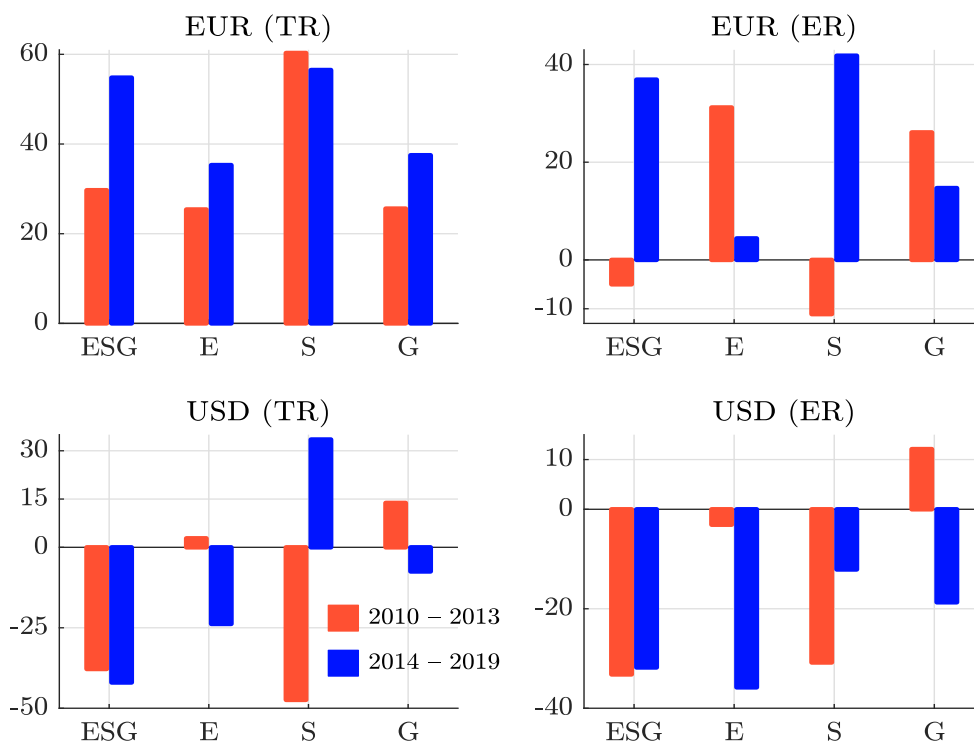
Source: Ben Slimane *et al.* (2019b).

We can draw the following conclusions from the previous results. First, the quintile portfolios present some strong bias in terms of duration risk. This is why the results based on credit return measures are not coherent with those based on total return measures. Second, these results confirm

that the period 2014–2019 has generated a better performance in terms of ESG investing, but only for EUR-denominated corporate bonds. For USD-denominated corporate bonds, the performance remains negative. Finally, we need to be careful about the feasibility to capture the ESG alpha. For instance, if we consider the long/short $Q_1 - Q_5$ strategy, the credit return is equal to 36.8 bps during the second period. We can break down this ESG alpha into the long leg and the short leg. In Figure 3.33, we report the raw performance of sorted portfolios. The carry statistics in bps are:

Period	Q_1	Q_5	$Q_1 - Q_5$
2010–2013	175	192	−17
2014–2019	113	128	−15

Therefore, the positive credit return of the long/short $Q_1 - Q_5$ cannot be explained by the carry exposure. It is due to the mark-to-market component and the dynamics of credit spreads and bond prices. This implies that an ESG strategy has a short carry position. On the one hand, buy-and-hold ESG investors may then suffer from this structural exposure, but they can increase the credit risk of their buy-and-hold portfolio to compensate the lower carry. On the other hand, active ESG investors may overperform only if they are able to rebalance their portfolio. Liquidity issues in the corporate bond market is then a barrier to create ESG alpha in the fixed-income universe. In particular, if we impose turnover constraints when building the previous simulated portfolios, we notice a decrease of the alpha return.

Figure 3.33: Annualized credit return in bps of ESG sorted portfolios (EUR IG, 2010–2019)
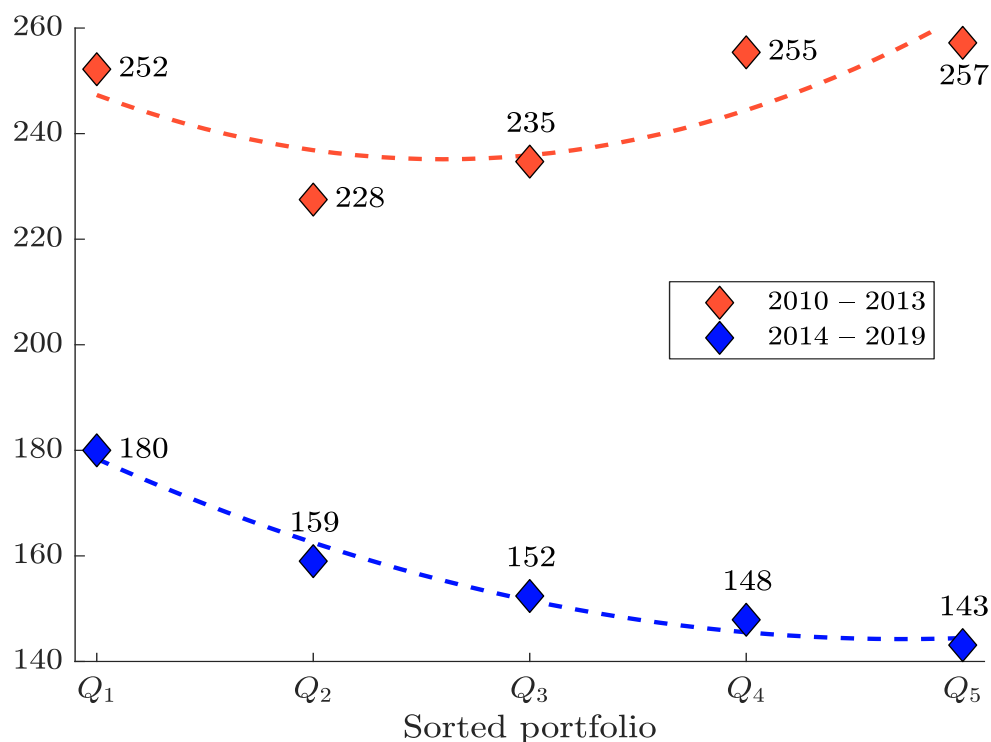
Source: Ben Slimane et al. (2019b).

Optimized portfolios In the sorted portfolios, there is no control of the duration or the credit spread. Therefore, it is difficult to know whether the alpha return is explained by the ESG scoring or the duration/spread biases. In fact, the method of sorted portfolios is not really relevant when it is implemented in the fixed-income universe. A better method is to consider the optimization approach, when the portfolio manager imposes some constraints on the active risk that he could take with respect to the benchmark index. Ben Slimane *et al.* (2019b) define the active risk measure as the weighted average of the duration and credit risks (see Box 3.7). Then, they implement an optimization program, which consists in minimizing the active risk while controlling the ESG excess score of the tilted portfolios. Starting from an ESG excess score equal to zero, they progressively increase the ESG score of the optimized portfolio until they reach one. They found that the relationship between the ESG excess score and the ex-post tracking error volatility is approximately linear. On average, targeting an excess score of one requires accepting a tracking error of 25 bps.

Using the ICE (BofAML) Large Cap IG EUR Corporate Bond index, Figures 3.34 and 3.35 show the impact of the ESG integration on the excess credit return of optimized portfolios for the periods 2010–2013 and 2014–2019. During the first period, the excess return of ESG optimized portfolios is negative, meaning that ESG investors were penalized. This is particularly true when optimized portfolios targeted high excess scores. For instance, an ESG excess score of $+1$ has produced an underperformance of -35 bps per year. During the second period, we observe slight positive outperformance that peaks at $+4$ bps when the ESG tilt is set to $+1$. We also notice that the relationship between the ESG excess score and the excess credit return is increasing. If we now consider the individual pillars, **(E)**, **(S)** and **(G)** optimized portfolios underperform during the 2010–2013 period. Among the three pillars, environmental is the best pillar and its excess return slides down until -22 bps when the targeted excess score is set to $+1$. Governance is the worst pillar, and its excess return reaches -49 bps for the same tilt. After 2014, excess credit returns are between -3 and $+9$ bps. Social is the winning pillar and exhibits significant outperformance that peaks at $+9$ bps. The recent period is then more favorable to ESG investors than before 2014.

If we consider the universe of USD investment grade corporate bonds, the results are different than those obtained with EUR-denominated corporate bonds. ESG investing has not created positive alpha for the entire 2010–2019 period (see Figures 3.36 and 3.37). Nevertheless, the substantial underperformance during the 2010–2013 period has been dramatically reduced since 2014. For instance, the excess return is close to zero for social-optimized portfolios between 2014 and 2019. Another interesting remark is the behavior of the governance pillar. In many academic studies, linkage between ESG and corporate financial performance is generally justified by the governance transmission channel. The results of Ben Slimane *et al.* (2019a) show that the governance pillar is not necessarily the most important factor, and investing in bonds with a good governance score is not fundamentally better than using the other pillars.

Remark 41 *The authors wonder if the transatlantic divide really concerns the currency of issued bonds or if it is more a regional issue. For instance, a EUR-denominated bond can be issued by an European corporate, but also by a firm which is located outside Europe. In a similar way, a USD-denominated bond can be issued by an American corporate, but also by a firm which is located outside America. Ben Slimane et al. (2019b) calculated the contribution to credit return of the different regions (Europe, Noth America and others). They noticed that Europe had a systematic positive contribution whereas North America had a systematic negative contribution whatever the currency (EUR and USD). Therefore, this transatlantic divide shows that ESG investing was a source of outperformance when it concerned IG bonds of European issuers, but a source of underperformance when it concerned IG bonds of American issuers.*

Figure 3.34: Annualized excess return in bps of ESG optimized portfolios (EUR IG, 2010–2013)
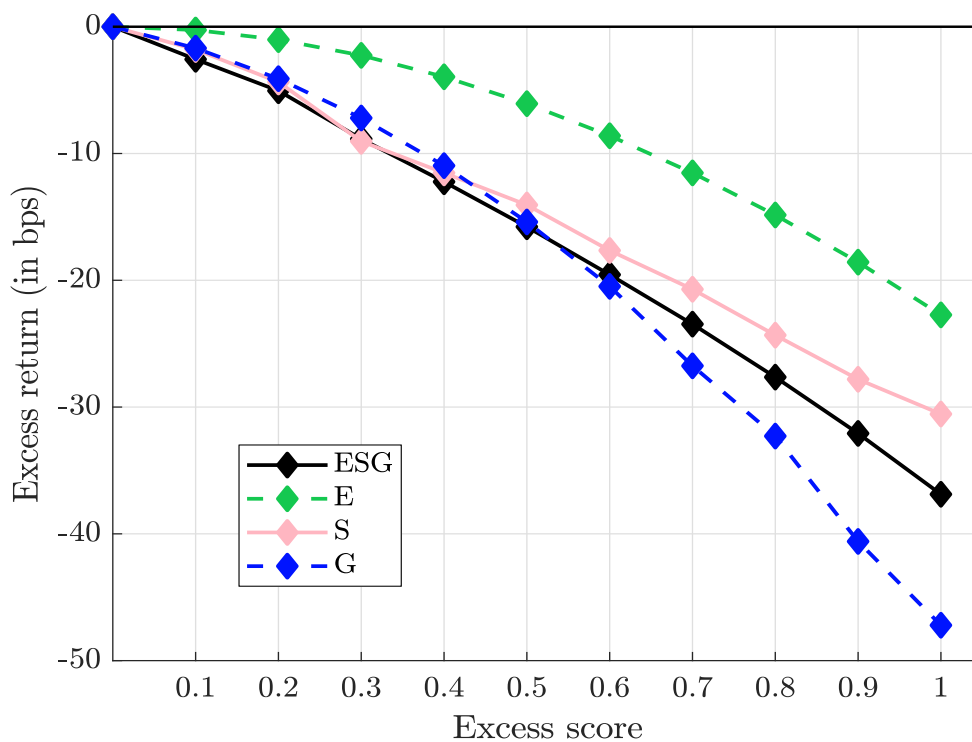
Figure 3.35: Annualized excess return in bps of ESG optimized portfolios (EUR IG, 2014–2016)
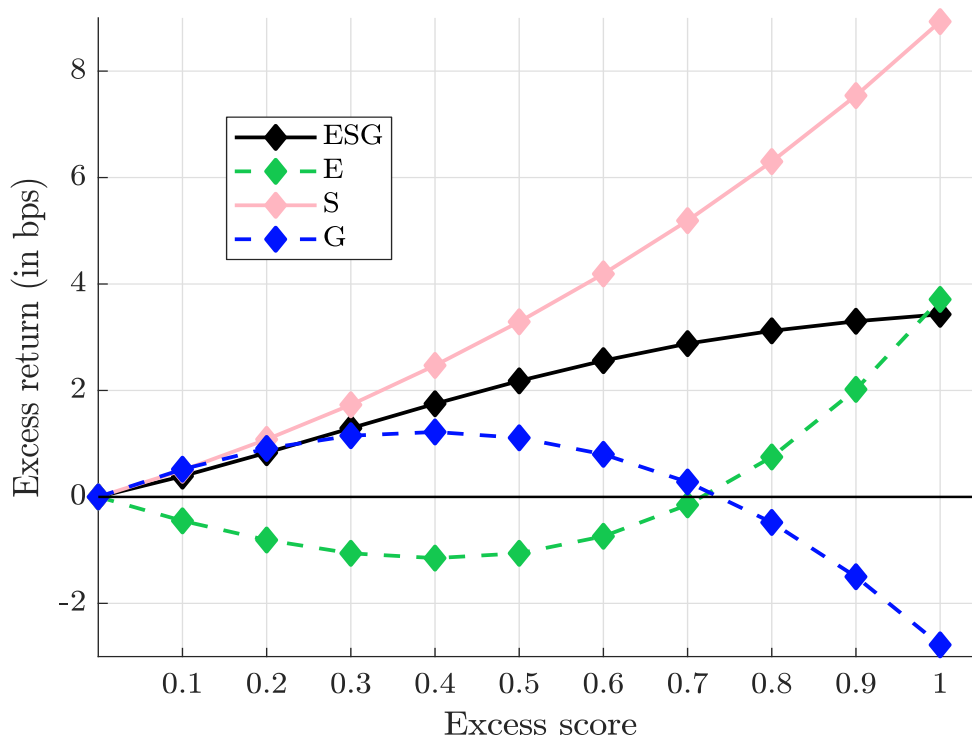
Source: Ben Slimane *et al.* (2019b).

Figure 3.36: Annualized excess return in bps of ESG optimized portfolios (USD IG, 2010–2013)
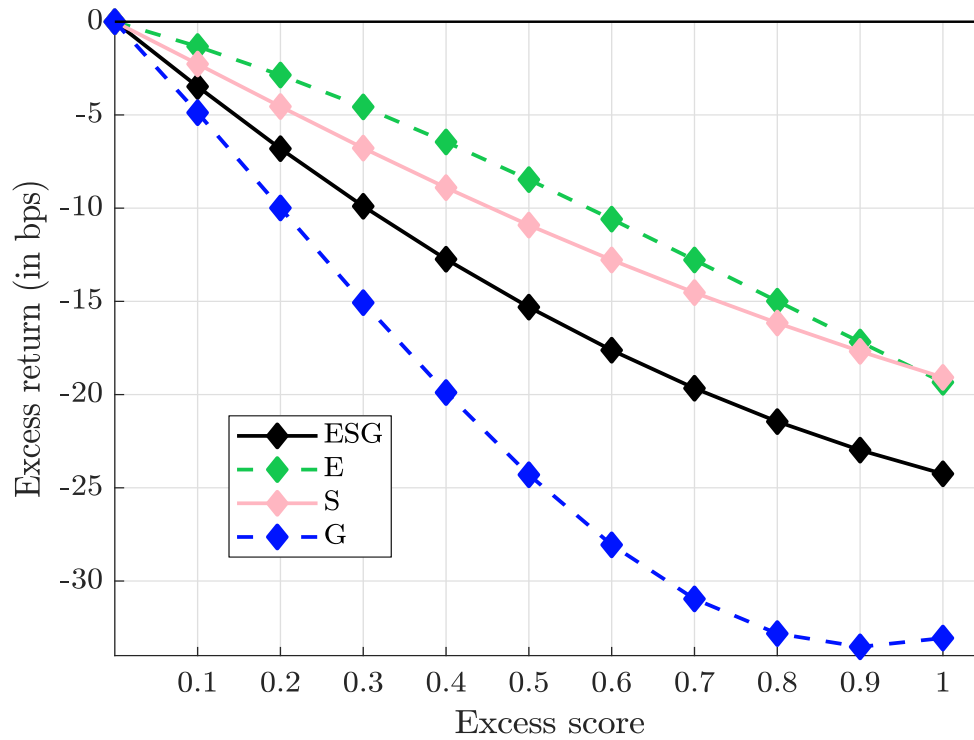
Figure 3.37: Annualized excess return in bps of ESG optimized portfolios (USD IG, 2014–2016)
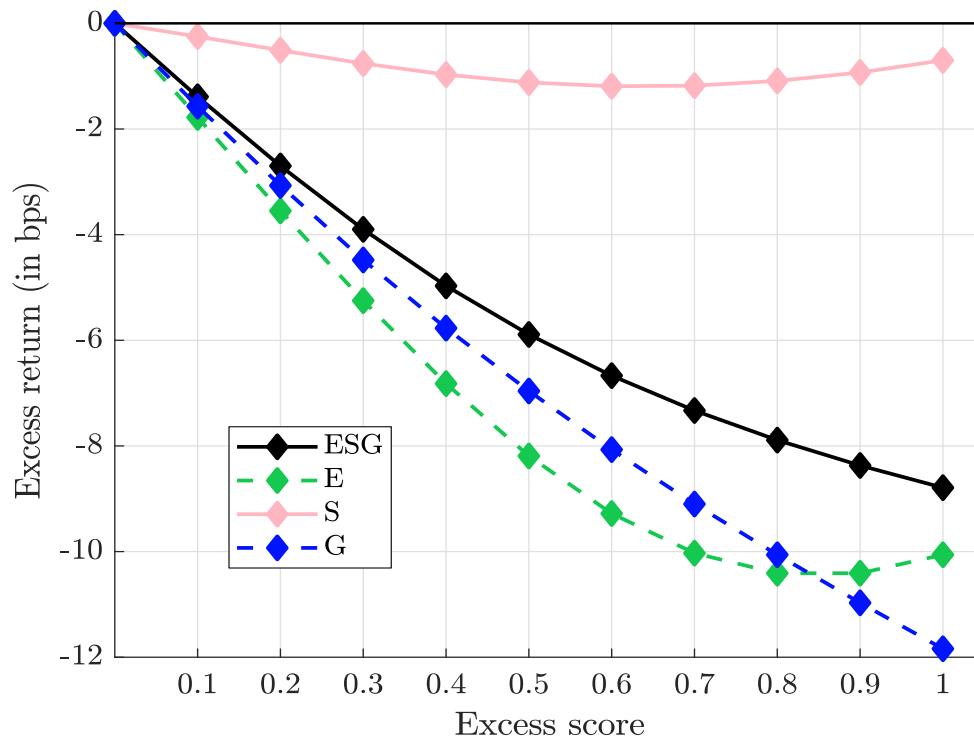
Source: Ben Slimane *et al.* (2019b).

Box 3.7: Building an optimized ESG portfolio in the fixed-income universe

The ESG score of the portfolio $w = (w_1, \ldots, w_n)$ is the weighted average of the individual scores: $\boldsymbol{\mathcal{S}}(w) = \sum_{i=1}^{n} w_i \boldsymbol{\mathcal{S}}_i$. If we consider a benchmark $b = (b_1, \ldots, b_n)$, we deduce that the ESG excess score of portfolio w with respect to benchmark b is equal to:

$$\boldsymbol{\mathcal{S}}(w \mid b) = \sum_{i=1}^{n} (w_i - b_i) \boldsymbol{\mathcal{S}}_i = \boldsymbol{\mathcal{S}}(w) - \boldsymbol{\mathcal{S}}(b)$$

When we use z-scores, we observe that $\boldsymbol{\mathcal{S}}(b) \approx 0$ because there is no reason that a capitalization-weighted index has a positive or a negative ESG score. It is generally a neutral ESG portfolio. On the contrary, an ESG portfolio w aims to have a better ESG score than the benchmark: $\boldsymbol{\mathcal{S}}(w \mid b) > 0$. When building an optimized ESG portfolio, there is of course a trade-off between the ESG excess score $\boldsymbol{\mathcal{S}}(w \mid b)$ and the active or tracking risk $\mathcal{R}(w \mid b)$ with respect to the benchmark. For instance, if the active risk is equal to zero, the ESG excess score will be equal to zero. If we consider a high ESG score (e.g., larger than 1.5), we also have to incur a high active risk. Therefore, the optimization problem becomes:

$$w^{\star}(\gamma) = \arg\min \frac{1}{2} \mathcal{R}(w \mid b) - \gamma \boldsymbol{\mathcal{S}}(w \mid b)$$

If γ is set to zero, the optimized portfolio $w^{\star}(0)$ is the benchmark portfolio b. If γ is set to infinity, the optimized portfolio $w^{\star}(\infty)$ corresponds to the bond with the largest z-score. The parameter γ can then be calibrated in order to target a given excess score $\mathcal{S}^{\star}$. The issue is the choice of the tracking risk metric. In the case of bonds, we generally use two measures. First, we can match the modified duration of the sectors. It follows that the modified duration risk of portfolio w with respect to benchmark b is:

$$\mathcal{R}_{\mathrm{MD}}(w \mid b) = \sum_{j=1}^{n_S} \left(\left(\sum_{i \in \boldsymbol{\mathcal{S}}ector(j)} w_i \, \mathrm{MD}_i \right) - \left(\sum_{i \in \boldsymbol{\mathcal{S}}ector(j)} b_i \, \mathrm{MD}_i \right) \right)^2$$

where n_S is the number of sectors and MD_i is the modified duration of bond i. An alternative is to use the DTS risk measure:

$$\mathcal{R}_{\mathrm{DTS}}(w \mid b) = \sum_{j=1}^{n_S} \left(\left(\sum_{i \in \boldsymbol{\mathcal{S}}ector(j)} w_i \, \mathrm{DTS}_i \right) - \left(\sum_{i \in \boldsymbol{\mathcal{S}}ector(j)} b_i \, \mathrm{DTS}_i \right) \right)^2$$

where DTS_i is the duration-times-spread (DTS) factor of bond i. We can also define an hybrid approach, where the risk measure is an average of the MD and DTS active risks:

$$\mathcal{R}(w \mid b) = \mathcal{R}_{\mathrm{MD}}(w \mid b) + \mathcal{R}_{\mathrm{DTS}}(w \mid b)$$

In fact, we can interpret $\mathcal{R}_{\mathrm{MD}}(w \mid b)$ as an interest rate risk measure and $\mathcal{R}_{\mathrm{DTS}}(w \mid b)$ as a credit risk measure, while $\mathcal{R}(w \mid b)$ is an integrated interest rate/credit risk measure.

Table 3.18: Performance of ESG bond indexes

Year	FTSE WGBI Return BM	FTSE WGBI Return ESG	FTSE WGBI Alpha ESG	FTSE EGBI Return BM	FTSE EGBI Return ESG	FTSE EGBI Alpha ESG	BEAC Return BM	BEAC Return SRI	BEAC Return S-SRI	BEAC Return ESG-S	BEAC Alpha SRI	BEAC Alpha S-SRI	BEAC Alpha ESG-S
2010	4.61	4.31	−30	0.61	4.14	353	3.07	2.93	2.96		−13	−10	
2011	6.35	7.05	69	3.41	7.31	391	1.49	1.17	1.43		−32	−5	
2012	1.65	3.06	141	10.65	7.39	−326	13.59	13.99	12.96		40	−63	
2013	−4.00	−2.95	105	2.21	−1.40	−362	2.37	2.49	2.36		12	−1	
2014	−0.48	−0.22	26	13.19	11.44	−175	8.40	8.31	8.49		−8	10	
2015	−3.57	−4.85	−128	1.65	0.39	−126	−0.56	−0.59	−0.50	−0.59	−3	6	−3
2016	1.60	1.02	−59	3.20	4.00	81	4.73	4.60	4.44	4.60	−13	−29	−13
2017	7.49	8.16	67	0.15	−0.47	−62	2.41	2.47	2.48	2.47	6	6	6
2018	−0.84	−1.41	−57	0.88	1.65	78	−1.25	−1.12	−1.11	−1.12	13	14	13
2019	5.90	5.56	−34	6.72	4.45	−227	6.24	6.01	5.92	6.01	−24	−32	−24
2020	10.11	10.90	79	5.03	4.11	−92	2.77	2.69	2.70	2.52	−8	−7	−25
2021	−6.97	−7.15	−17	−3.54	−3.76	−21	−0.97	−0.96	−0.99	−0.99	1	−2	−2
2022	−18.26	−20.00	−173	−18.52	−19.06	−54	−13.65	−13.62	−13.48	−13.48	3	16	17
3Y	−5.75	−6.26	−51	−6.19	−6.74	−55	−4.21	−4.22	−4.18	−4.23	−1	3	−2
5Y	−2.54	−3.03	−49	−2.33	−2.95	−61	−1.61	−1.63	−1.62	−1.64	−2	−1	−3
7Y	−0.58	−0.93	−35	−1.21	−1.63	−42	−0.16	−0.19	−0.20	−0.19	−3	−4	−3
10Y	−1.22	−1.46	−24	0.77	−0.17	−94	0.88	0.86	0.86		−2	−1	

Bloomberg US Corporate

Year	Return BM	Return SRI	Return S-SRI	Return ESG-S	Alpha SRI	Alpha S-SRI	Alpha ESG-S
2019							
2020							
2021	−1.04	−1.55	9.56	2.34	−51	1060	338
2022	−15.76	−15.12	−1.10	−13.86	64	1467	190

Bloomberg Global High Yield

Year	Return BM	Return SRI	Return SUS	Alpha SRI	Alpha SUS
2019	1.00	0.96		−4	
2020	2.47	2.80	2.87	32	40
2021	1.10	0.40	0.21	−70	−89
2022	−5.00	−5.95	−5.73	−95	−72

Source: FTSE, Bloomberg & Author's calculations.

Bond indexes

In Table 3.18, we compare the annualized returns of some famous bond indexes (FTSE World Government Bond Index or WGBI, FTSE EMU Government Bond Index or EGBI, Bloomberg Euro Aggregate Corporate Total Return Index, Bloomberg US Corporate Total Return Index and Bloomberg Global High Yield Total Return Index) with the annualized returns of their ESG equivalent indexes[29]. Results with bond indexes are close to those obtained with equity indexes. Indeed, ESG investing has its good and bad times. Nevertheless, we notice that the results are less balanced than previously. It seems that ESG investing creates positive alpha less frequently in the bond market. In fact, the main reason is that ESG bond indexes has lower carry because they have a better credit rating[30].

3.3   Cost of capital

3.3.1   ESG premium

Equity

Corporate debt

Sovereign debt

3.3.2   Carbon premium

---

[29]For the FTSE WGBI and EGBI, we consider the FTSE ESG World Government Bond and FTSE ESG Select EMU Government Bond indexes. For the Bloomberg Euro Aggregate Corporate TR Index, we use the MSCI Euro Corporate SRI TR Index (SRI), Bloomberg MSCI Euro Corporate Sustainable SRI TR Index (S-SRI), and Bloomberg MSCI Euro Corporate ESG Sustainability SRI (ESG-S). For the Bloomberg Euro Aggregate Corporate TR Index, we use the MSCI Euro Corporate SRI TR Index (SRI), Bloomberg MSCI Euro Corporate Sustainable SRI TR Index (S-SRI), and Bloomberg MSCI Euro Corporate ESG Sustainability SRI (ESG-S). For the Bloomberg US Corporate TR Index, the comparison is done with the Bloomberg MSCI US Corporate SRI Select Index (SRI), Bloomberg MSCI US Corporate Sustainability SRI Index (S-SRI) and Bloomberg MSCI US Corporate ESG Sustainability SRI (ESG-S). Finally, the Bloomberg Global High Yield Total Return Index is compared with the Bloomberg MSCI Global High Yield SRI Index (SRI) and Bloomberg MSCI Global High Yield Sustainability Index (SUS).

[30]The relationship between ESG and credit ratings is investigated on page 202.

3.4 Exercises

3.4.1 Equity portfolio optimization with ESG scores

We consider the CAPM model:

$$R_i - r = \beta_i (R_m - r) + \varepsilon_i$$

where R_i is the return of asset i, R_m is the return of the market portfolio w_m, r is the risk free asset, β_i is the beta of asset i with respect to the market portfolio and ε_i is the idiosyncratic risk of asset i. We have $R_m \perp \varepsilon_i$ and $\varepsilon_i \perp \varepsilon_j$. We note σ_m the volatility of the market portfolio. Let $\tilde{\sigma}_i$, μ_i and $\boldsymbol{\mathcal{S}}_i$ be the idiosyncratic volatility, the expected return and the ESG score of asset i. We use a universe of 6 assets with the following parameter values:

Asset i	1	2	3	4	5	6
β_i	0.10	0.30	0.50	0.90	1.30	2.00
$\tilde{\sigma}_i$ (in %)	17.00	17.00	16.00	10.00	11.00	12.00
μ_i (in %)	1.50	2.50	3.50	5.50	7.50	11.00
$\boldsymbol{\mathcal{S}}_i$	1.10	1.50	2.50	-1.82	-2.35	-2.91

and $\sigma_m = 20\%$. The risk-free return r is set to 1% and the expected return of the market portfolio w_m is equal to $\mu_m = 6\%$.

1. We assume that the CAPM is valid.

   (a) Calculate the vector μ of expected returns.

   (b) Compute the covariance matrix Σ. Deduce the volatility σ_i of the asset i and find the correlation matrix $\mathbb{C} = (\rho_{i,j})$ between asset returns.

   (c) Compute the tangency portfolio w^*. Calculate $\mu(w^*)$ and $\sigma(w^*)$. Deduce the Sharpe ratio and the ESG score of the tangency portfolio.

   (d) Compute the beta coefficient $\beta_i(w^*)$ of the six assets with respect to the tangency portfolio w^*, and the implied expected return $\tilde{\mu}_i$:

   $$\tilde{\mu}_i = r + \beta_i(w^*)(\mu(w^*) - r)$$

   (e) Deduce the market portfolio w_m. Comment on these results.

2. We consider long-only portfolios and we also impose a minimum threshold $\boldsymbol{\mathcal{S}}^\star$ for the portfolio ESG score:

   $$\boldsymbol{\mathcal{S}}(w) = w^\top \boldsymbol{\mathcal{S}} \geq \boldsymbol{\mathcal{S}}^\star$$

   (a) Let γ be the risk tolerance. Write the mean-variance optimization problem.

   (b) Find the QP form of the MVO problem.

   (c) Compare the efficient frontier when (1) there is no ESG constraint ($\boldsymbol{\mathcal{S}}^\star = -\infty$), (2) we impose a positive ESG score ($\boldsymbol{\mathcal{S}}^\star = 0$) and (3) the minimum threshold is set to 0.5 ($\boldsymbol{\mathcal{S}}^\star = 0.5$). Comment on these results.

   (d) For each previous cases, find the tangency portfolio w^* and the corresponding risk tolerance γ^*. Compute then $\mu(w^*)$, $\sigma(w^*)$, $\mathrm{SR}(w^* \mid r)$ and $\boldsymbol{\mathcal{S}}(w^*)$. Comment on these results.

(e) Draw the relationship between the minimum ESG score $\mathcal{S}^{\star}$ and the Sharpe ratio $\mathrm{SR}\,(w^{*} \mid r)$ of the tangency portfolio.

(f) We assume that the market portfolio w_m corresponds to the tangency portfolio when $\mathcal{S}^{\star} = 0.5$.

    i. Compute the beta coefficient $\beta_i\,(w_m)$ and the implied expected return $\tilde{\mu}_i\,(w_m)$ for each asset. Deduce then the alpha return α_i of asset i. Comment on these results.

    ii. We consider the equally-weighted portfolio w_{ew}. Compute its beta coefficient $\beta\,(w_{\mathrm{ew}} \mid w_m)$, its implied expected return $\tilde{\mu}\,(w_{\mathrm{ew}})$ and its alpha return $\alpha\,(w_{\mathrm{ew}})$. Comment on these results.

3. The objective of the investor is twice. He would like to manage the tracking error risk of his portfolio with respect to the benchmark $b = (15\%, 20\%, 19\%, 14\%, 15\%, 17\%)$ and have a better ESG score than the benchmark. Nevertheless, this investor faces a long-only constraint because he cannot leverage his portfolio and he cannot also be short on the assets.

(a) What is the ESG score of the benchmark?

(b) We assume that the investor's portfolio is $w = (10\%, 10\%, 30\%, 20\%, 20\%, 10\%)$. Compute the excess score $\mathcal{S}\,(w \mid b)$, the expected excess return $\mu\,(w \mid b)$, the tracking error volatility $\sigma\,(w \mid b)$ and the information ratio $\mathrm{IR}\,(w \mid b)$. Comment on these results.

(c) Same question with the portfolio $w = (10\%, 15\%, 30\%, 10\%, 15\%, 20\%)$.

(d) In the sequel, we assume that the investor has no return target. In fact, the objective of the investor is to improve the ESG score of the benchmark and control the tracking error volatility. We note γ the risk tolerance. Give the corresponding esg-variance optimization problem.

(e) Find the matrix form of the corresponding QP problem.

(f) Draw the esg-variance efficient frontier $(\sigma\,(w^{\star} \mid b), \mathcal{S}\,(w^{\star} \mid b))$ where $w^{\star}$ is an optimal portfolio.

(g) Find the optimal portfolio $w^{\star}$ when we target a given tracking error volatility $\sigma^{\star}$. The values of $\sigma^{\star}$ are 0%, 1%, 2%, 3% and 4%.

(h) Find the optimal portfolio $w^{\star}$ when we target a given excess score $\mathcal{S}^{\star}$. The values of $\mathcal{S}^{\star}$ are 0, 0.1, 0.2, 0.3 and 0.4.

(i) We would like to compare the efficient frontier obtained in Question 3(f) with the efficient frontier when we implement a best-in-class selection or a worst-in-class exclusion. The selection strategy consists in investing only in the best three ESG assets, while the exclusion strategy implies no exposure on the worst ESG asset. Draw the three efficient frontiers. Comment on these results.

(j) Which minimum tracking error volatility must the investor accept to implement the best-in-class selection strategy? Give the corresponding optimal portfolio.

3.4.2   Bond portfolio optimization with ESG scores

3.4.3   Minimum variance portfolio with climate risk

3.4.4   Cost of capital and green sentiment

3.4.5   Strategic asset allocation with ESG preferences

Chapter 4

Sustainable Financial Products

4.1 The market of ESG mutual funds

4.1.1 Greenwashing issues

4.1.2 Classification of ESG investment funds

4.1.3 The case of index funds

4.1.4 Market growth statistics

4.2 Green and social bonds

Beside traditional investment vehicles that incorporate ESG criteria, we can find a category of securities that is entirely dedicated to sustainable finance. It corresponds to specific sustainable debt instruments. The two most famous assets are green bonds and social bonds, but the list of sustainable fixed-income assets is much longer: sustainable bonds, transition bonds, sustainable-linked bonds, green loans, green notes, green ABCP notes, etc. In Table 4.1, we report the segmentation of the sustainable fixed-income market. According to CBI (2022b), the cumulative total GSS+ issuance stands at \$3.3 tn at the end of June 2022. About \$3 tn comes from GSS assets (Figure 4.1). We have the following breakdown: 62.3% for green bonds, 17% for social bonds and 20.7% for sustainability bonds. The market is dominated by DM issuers (approximatively 63%), while the remaining part is half of EM issuers and half of supranational issuers.

Table 4.1: Sustainable fixed-income market

	Theme	Label	Format
GSS+	GSS	Green	Use of proceeds
		Social	Use of proceeds
		Sustainability	Use of proceeds
	Transition	Sustainability-Linked	Entity KPI-linked
		Transition	Use of proceeds

Source: CBI (2022b).

Figure 4.1: Issuance of GSS securities (in $ bn)

Source: https://www.climatebonds.net/market/data.
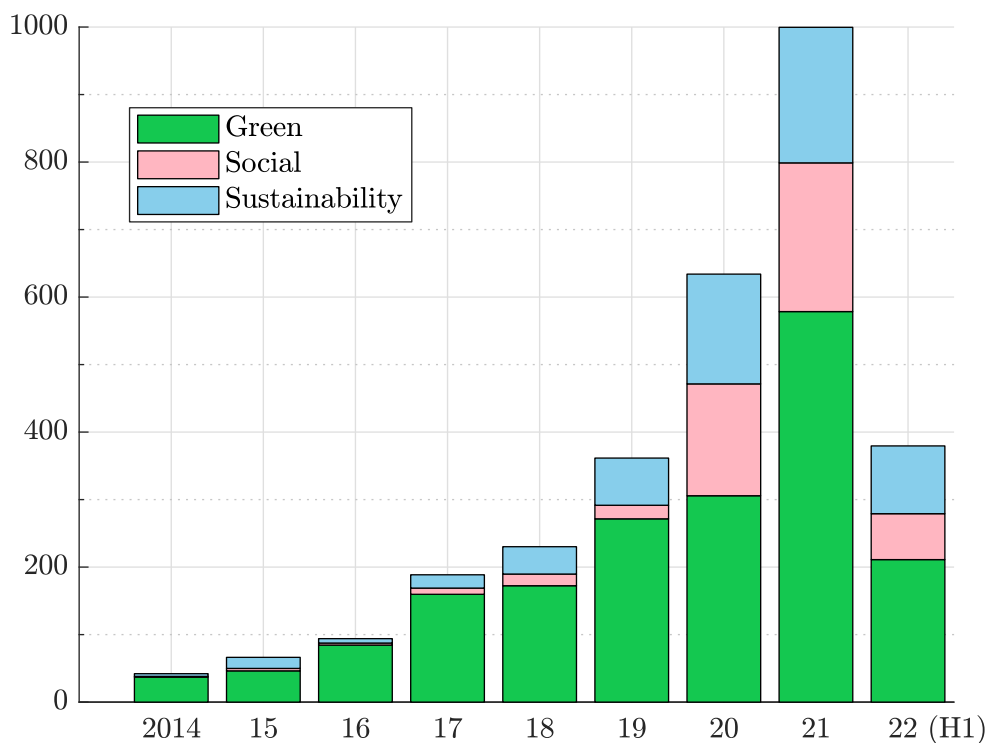
4.2.1 Green bonds

Definition

Green bonds are fixed-income securities, which finance investments with an environmental objective. They differentiate from regular bonds, because they are labelled green by issuers or external third-party entities and there is a commitment to use the funds for financing green projects. However, there is no legally-binding definition of a green bond. Most of market participants[1] have then adopted the definition of the GBP framework[2]:

> "*Green bonds are any type of bond instrument where the proceeds or an equivalent amount will be exclusively applied to finance or re-finance, in part or in full, new and/or existing eligible green projects and which are aligned with the four core components of the Green Bond Principles (GBP).*" (ICMA, 2021a, page 3).

The four core components of the GBP are:

1. Use of proceeds

2. Process for project evaluation and selection

3. Management of proceeds

4. Reporting

The utilisation of the proceeds (or the funds) should be affected to eligible green projects, e.g., renewable energy, energy efficiency, pollution prevention (GHG control, soil remediation, waste recycling), sustainable management of living natural resources (sustainable agriculture, sustainable forestry, restoration of natural landscapes), terrestrial and aquatic biodiversity conservation (protection of coastal, marine and watershed environments), clean transportation, sustainable water management, climate change adaptation, circular economy and eco-efficient products, green buildings. With respect to the process for project evaluation and selection, the issuer of a green bond should clearly communicate the environmental sustainability objectives, the eligible projects and the related eligibility criteria. The third component deals with the management of proceeds and includes the tracking of the "*balance sheet*" and the allocation of funds[3]. It also recommends an external review by a third-party entity. Finally, the reporting must be based on the following pillars: transparency, description of the projects, allocated amounts and expected impacts, qualitative performance indicators and quantitative performance measures[4].

Remark 42 *GBP is not the only green bond framework. The other popular guidelines are:*

- *China Green Bond Principles[5] (PBOC, CBIRC, July 2022)*

- *Climate Bonds Standard[6] (CBI, 2019)*

---

[1]According to IFC (2020, page 5), the Green Bond Principles are endorsed by 95% of issuers.

[2]The first version of Green Bond Principles was issued on January 2014.

[3]The proceeds should be credited to a sub-account.

[4]Here are some examples: energy capacity, electricity generation, GHG emissions reduced/avoided, number of people provided with access to clean power, decrease in water use, reduction in the number of cars required, etc.

[5]They replace China's Green Bond Standards published by PBOC in 2015.

[6]The first version is released in November 2011. A new version has been drafted for public consultation and will certainly be available in 2023 (CBI, 2022c).

- *ASEAN Green Bond Standards*[7] *(ACMF, 2018)*

- *EU Green Bond Standard*[8]

All these guidelines are based on a common framework and are closed to the GBP. The CBI approach, which also uses the GBP, is perhaps more comprehensive and gives more details for issuing a green bond[9].

Green debt instruments can be issued in different formats. They differ in the collateral assets, the recourse process in case of default, etc. For instance, the most common instrument is the green regular or "*Use of Proceeds*" bond (UoP bond), which carries the same credit rating than a conventional bond, because the bondholders have recourse to all the assets of the bond issuer. In the case of a green revenue bond, the collateral for the debt comes from cash flows of the revenue streams collected by the issuer. A green project bond is a bond dedicated to a given green project, implying that the recourse process only concerns the assets related to the project. Green loans are loans that finance green projects and may be secured or unsecured. These four instruments (regular bond, revenue bond, project bond and green loan) are called asset-linked bond structures, because they are related to a specific asset/project. Asset-backed bond structures are made up of securitization and covered bonds. They both involves a group of projects. Securitized bonds can use ABS/MBS/CLO/CDO securitization structures, while covered bonds are German debt instruments (Pfandbriefe) that use a dual recourse process based on the issuer and the cover pool[10].

The certification process (or external review) is an important step when issuing green bonds since it is related to the greenwashing issue. We recall that GBP are voluntary process guidelines. They recommend that issuers appoint an external review provider to obtain pre-issuance assessment of the green project and an external auditor to have a post-issuance validation of funding management. Certainly, the most popular form of external review is the second party opinion or SPO (IFC, 2020, page 19). In this case, the objective of the issuer is to obtain a "*green bond label*" or the approval of its green project by a competent and independent entity, which is recognized by financial markets and investors. For instance, Ehlers and Packer (2017) and IFC (2020) listed the following forms of green bond certification[11]:

- Second party opinion provided by ESG rating agencies (ISS, Sustainalytics, Vigeo-Eiris);

- Certification by specialized green bond entities (CBI, CICERO, DNV);

- Green bond assessment by statistical rating organizations (Moody's, S&P).

Some examples

The market of climate-related bonds begins in 2001 with the issuance of a revenue bond (known as the solar bond) by the City of San Francisco. The objective was to finance 140 acres with

---

[7]The first version is published in 2017.

[8]The proposal for a regulation on European green bonds has been released in 2021. It is based on four key requirements: (1) taxonomy-alignment; (2) transparency on how the bond proceeds are allocated through detailed reporting requirements; (3) all European green bonds must be checked by an external reviewer; (4) external reviewers must be supervised by the ESMA.

[9]See also the Green Bond Handbook published by IFC (2020).

[10]Investors may have recourse to the issuer, but if the issuer is unable to pay its debt, then bondholders gain recourse to the cover pool.

[11]Ehlers and Packer (2017) added green bond indexes as a fourth form of certification, since including a bond in a green bond index is a market recognition that the bond is green.

solar panels, which could power renewable energy on homes and business buildings. A second step was reached in 2007, when the European Investment Bank (EIB) issued the world's first Climate Awareness Bond (CAB) in order to finance renewable energy and energy efficiency projects. Finally, the first green bond has been created by the World Bank and the Swedish bank SEB in 2008 for a group of Scandinavian institutional investors. Since this date, the market of green bonds has been growing at a fast pace. The issued amount has soared from less then $3 bn in 2013 to more than $500 bn in 2021. According to Baker *et al.* (2022), the years 2013 and 2014 marked the development of the green bond market. For instance, the first corporate green bonds were issued by the French utility company EDF ($1.8 bn) and the Swedish real estate company Vasakronan ($120 bn). Toyota introduced the auto industry's first-ever asset-backed green bond in 2014 ($1.75 bn), while the Commonwealth of Massachusetts successfully completed the first municipal green bond in 2013 ($100 mn). The development of sovereign green bonds begins with the issuance of Poland in December 2016 ($1 bn) and France[12] in January 2017 ($10 bn).

According to CBI (2022a), the largest corporate issuers in 2021 were China Three Gorges ($7.2 bn), Iberdrola ($3.3 bn), CTP ($3 bn), Ardagh ($2.8 bn), Engie ($2.6 bn), Ford Motor ($2.5 bn), EDP ($2.4 bn), State Grid Corporation of China ($2.4 bn), Mondelez International ($2.4 bn) and Liberty Global ($2.3 bn). Since 2014, the top three corporate issuers are Engie (France), Iberdrola (Spain) and TenneT (Netherlands) with about $17 bn of cumulative issuance for each company. If we focus on sovereign green bonds, the five largest issuers are France ($43.6 bn), Germany ($25.1 bn), UK ($21.9 bn), Italy ($10.0 bn) and Netherlands ($10 bn). The NextGenerationEU program of the European Commision plans to issue $250 bn of green bonds from 2020 to 2030. This will make the European Commission the largest green bonds issuer in the world[13].

Remark 43 *The post-issuance management of a green bond may be an issue. For example, Mexico City Airport Trust issued $6 bn of green bonds in 2016 and 2017 in order to finance the construction of a new airport. It met ICMA GBP and obtained a second party opinion from Sustainalytics as well as green bond assessments from rating agencies Moody's and S&P. However, in October 2018, the new Mexican government announced to halt the construction of the airport and launched a buyback package, capped at $1.8 bn.*

The green bond market

Statistics From 2007 to the first half of 2022, CBI estimate that a total of 10 800 green bonds have been issued in the world[14]. The geographic repartition is the following: 52% in North America, 23% in Europe, 17% in Asia-Pacific and 8% in the rest of the world (including supranational entities). The distribution of deal size is highly skewed. Indeed, 70% of green bonds have a notional less than $100 mn, whereas 3.2% of them have a deal size greater than $1 bn. If we focus on the number of issuers, we obtain the following top five ranking[15]: 500 in US, 404 in China, 156 in Japan, 104 in Sweden and 63 in Norway. If we analyze the amount issued, the size of the green bond market is roughly equal to $1.9 tn. In Figures 4.2 and 4.3, we have reported the issuance and notional outstanding (or cumulative issuance) by market type and region from 2014 to 2022. The market is lead by Europe (46%), followed by Asia-Pacific[16] (25%) and North America (20%). The issuance of green bonds

---

[12]Green OAT 1.75% 25 June 2039.

[13]In October 2021, the EC issued a 15-year green bond and raised $12.8 bn (for $135 bn of orders), making it the world's largest green bond transaction to date.

[14]This number is skewed because Fannie Mae has been a very frequent issuer of relatively small green MBS deals (less than $100 mn).

[15]The number of green bond issuers in France and Germany is respectively equal to 58 and 50.

[16]China represents more than 50% of the Asia-Pacific green bond market.

Figure 4.2: Issuance and notional outstanding of green debt by market type
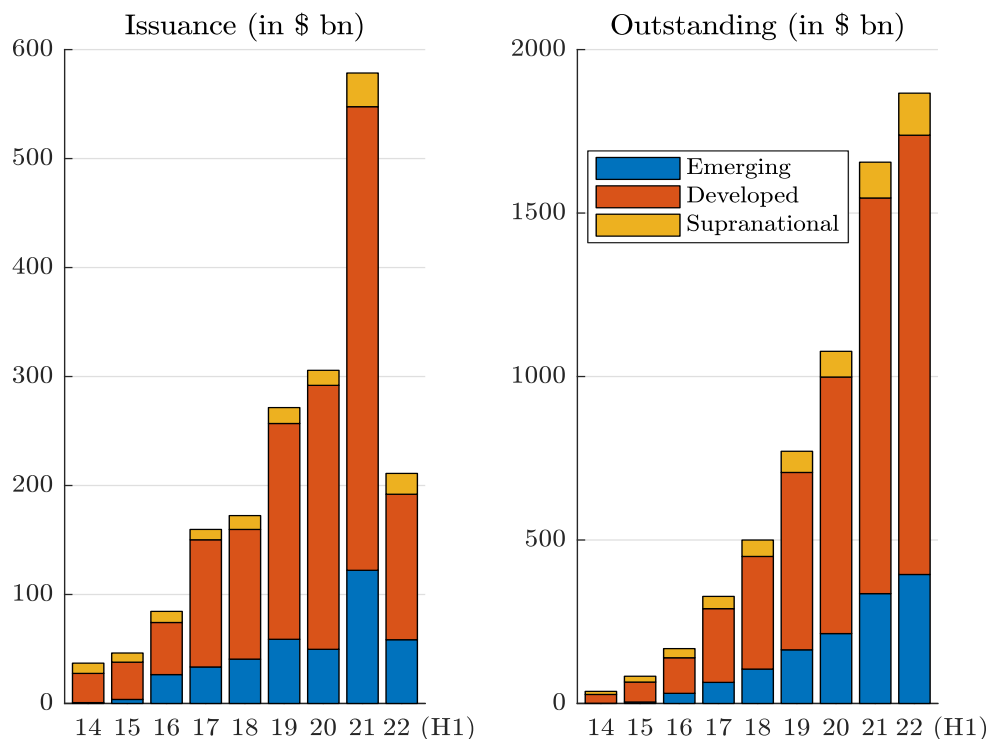
Figure 4.3: Issuance and notional outstanding of green debt by region

Source: https://www.climatebonds.net/market/data.
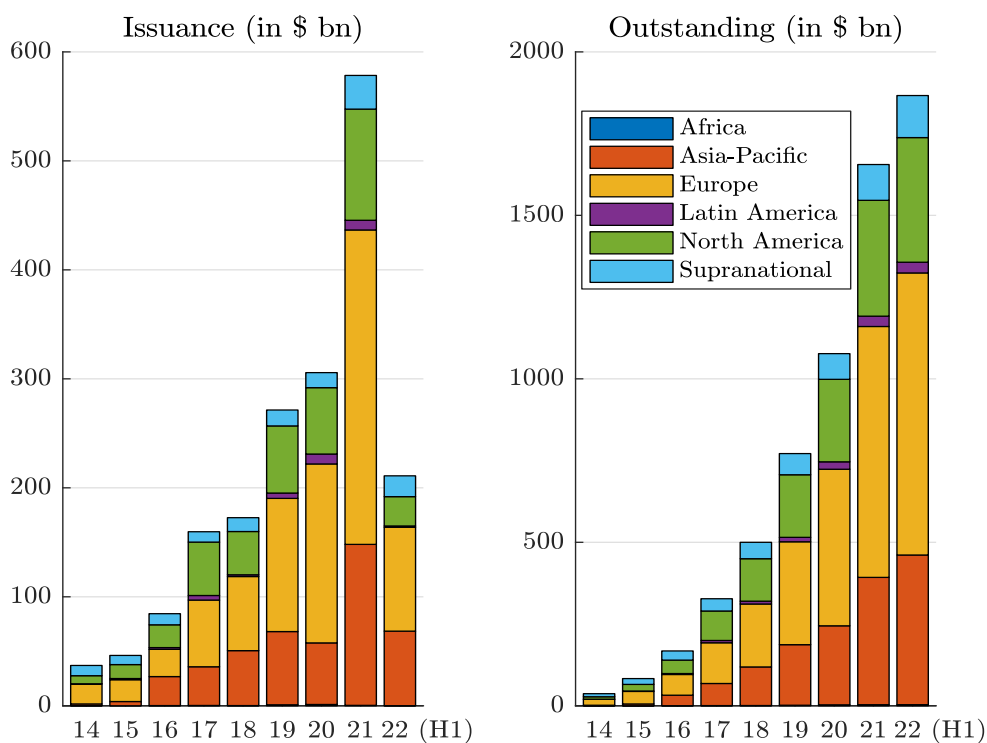
Figure 4.4: Issuance and notional outstanding of green debt by use of proceeds
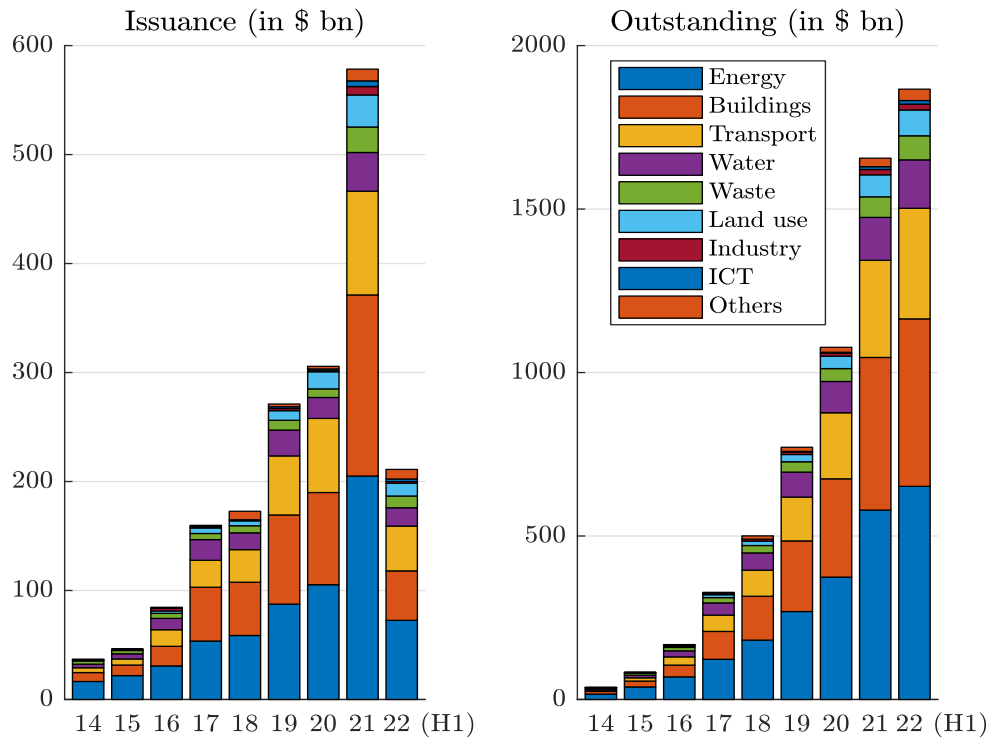
Figure 4.5: Issuance and notional outstanding of green debt by issuer type

Source: https://www.climatebonds.net/market/data.
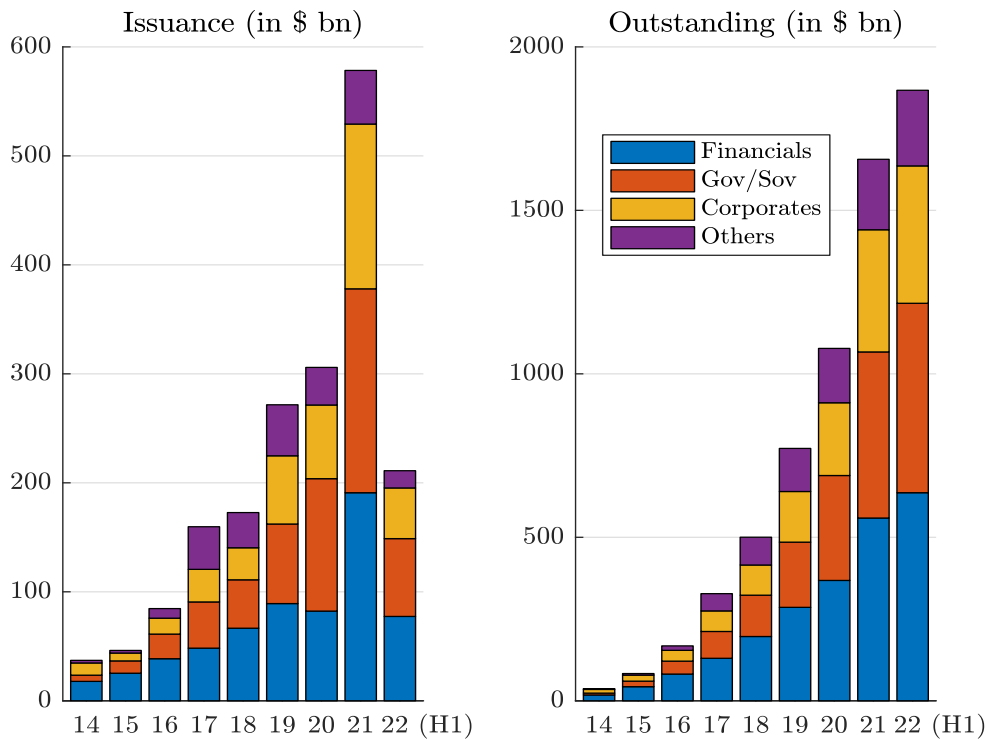
mainly concerns four sectors: energy, buildings, transport and water. They represent 88% of the market (Figure 4.4). An analysis by issuer type shows the market is approximately balanced between financials (development banks and financial corporates), governement/sovereign issuers (including government-backed entities, local governments and states) and non-financial corporations (Figure 4.5).

How to invest in green bonds   There are several ways to invest in green bonds. We can consider a mutual fund (active management), an ETF (passive management) or a direct investment[17]. Here are some examples of active mutual funds: Allianz IG green bond fund, Amundi RI impact green bonds, AXA WF ACT green bonds, BNP Paribas green bond, Calvert green bond fund, Mirova global green bond fund, TIAA-CREF green bond fund. Beside these investment vehicles, institutionals can also invest in some closed funds: Amundi planet emerging green one, Conservation fund green bonds, Foresight Italian green bond fund, etc. Nevertheless, the largest investments in green bonds are done using passive management, especially ETFs: Franklin Liberty green bond, iShares global green bond, Lyxor green bond, VanEck Vectors green bond, Xtrackers corporate green bond, etc. In this case, an ETF use a green bond index such as:

- Bloomberg Barclays MSCI Global Green Bond Index (global green bonds)

- S&P Green Bond Index (global green bonds)

- Solactive Green Bond Index (global green bonds)

- ChinaBond China Climate-Aligned Bond Index (chinese green bonds)

- SSE Green Corporate Bond Index and SSE Green Bond Index (green bonds listed on the Shanghai Stock Exchange)

- ICE BofA Green Index (global green bonds)

The economics of green bonds

Rationale for issuing green bonds   Green bonds are very different from ESG portfolios and funds, since the objective is to finance a specific green project. Therefore, the choice to invest in a green bond goes beyond CSR or SRI values (Maltais and Nykvist, 2020). From the issuer viewpoint, it is a signal and a visible endorsement that the entity participates to the green economy (Flammer, 2021; Daubanes *et al.*, 2021). From the investor viewpoint, it is a way to implement relatively easily an impact investing program. Furthermore, green bonds are more climate-related assets than ESG-related assets. They strongly participate to financing the climate transition. For instance, sovereign green bond issuance is generally presented as a response to climate change. If we consider the NextGenerationEU program of green bonds, the objective of the European Commission is to "*achieve the goal of climate neutrality by 2050*". Denmark issued its first green bond on January 2022 and the funds will be allocated to "*support the production of renewable energy sources and the green transition of the transport sector*". The success of the Republic of the Philippines is explained as the strong recognition and confidence from investors to "*achieving sustainable development and mitigating climate change, notably the pledge to reduce our greenhouse gas emissions by 75% by 2030*". Therefore, a green bond is a signaling tool to show that governments and corporations

---

[17]Only largest institutional investors have access to the primary green bond market. Nevertheless, they can trade green bonds in the secondary market.

respond to climate change[18]. This is the main conclusion of the research conducted by Caroline Flammer:

> [...] "*I show that investors respond positively to the issuance announcement, a response that is stronger for first-time issuers and bonds certified by third parties. The issuers improve their environmental performance post-issuance (i.e., higher environmental ratings and lower CO_2 emissions) and experience an increase in ownership by long-term and green investors. Overall, the findings are consistent with a signaling argument — by issuing green bonds, companies credibly signal their commitment toward the environment.*" (Flammer, 2021, page 499)

From an economic viewpoint, green bonds can be seen as a second-best instrument in the absence of a global carbon pricing scheme (carbon tax), which is the Pigovian solution to the carbon externality (Ehlers and Packer, 2017; Daubanes *et al.*, 2021; Baker *et al.*, 2022). In this perspective, green bonds are the response to the net-zero financing issue:

> "*Capital spending on physical assets for energy and land-use systems in the net-zero transition between 2021 and 2050 would amount to about $275 trillion, or $9.2 trillion per year on average, an annual increase of as much as $3.5 trillion from today*" (McKinsey, 2022, page viii).

This figure of $3.5 trillion is approximately equal to $1/2$ of global corporate profits, $1/4$ of total tax revenue, or 4.1% of world GDP. Therefore, the gap between current and expected green investments is huge. Of course, green bonds help to finance the climate transition, but they are a partial solution since they represented less than $600 mn of investment in 2021. Therefore, the second-best instrument is not currently the solution to climate change.

The last remark questions us whether the green bond market is driven by the supply or the demand. Indeed, if green bonds are a second-best solution, we should observe a greater supply. The issue is that there is apparently no economic incentive with the exception of green signaling. In this case, the temptation is to conclude that the green bond market is driven by the demand of green assets. It is true that we observe a systematic oversubscription when issuing a green bond. We have already seen that the issuance of the EC in October 2021 has been oversubscribed 11 times. Such events are not rare. For example, the Italian green BTP was 9 times oversubscribed in December 2020, the German green bond was more than 5 times oversubscribed in September 2020, etc. Because of this supply/demand imbalance, we could think that green and conventional bond prices are different for the same issuer even if the green and conventional bonds share the same characteristics (same coupon, same maturity, same seniority, same payment schedule). In particular, we expect a large negative premium of the green bond with respect to the conventional bond. Below, we are going to see that the difference is relatively low, which is a market anomaly. In Section 8.3 on page 350, we will learn that a global and fair carbon tax implies a strong distortion of the economic profitability across companies and sectors. On the contrary, green bond policies have little impact, meaning that green bonds are not really a second-best instrument. They help to capture investment flows and finance the climate transition because of the huge demand from investors, but the cost of greening the economy remains relatively high, because they have low impact on negative carbon externalities, adverse selection and moral hazard. In this context, the development of green bonds is disappointing if the goal is to fight the climate change and reduce dramatically carbon emissions. Nevertheless, the development of green bonds is also positive because it participates to the emergence and the diffusion of the green sentiment (Brière and Ramelli, 2021).

---

[18]For instance, we observe a high issuance activity just before and during the organization of a Conference of Parties (COP) to the UNFCCC.

Estimation of the greenium   The green bond premium (or greenium) is the difference in pricing between green and regular bonds. Financial theory tells us that the yield of a bond depends on its characteristics (maturity, cash flow schedule, coupon rate, seniority, liquidity), the term structure of the interest rates and the default risk of the bond issuer (Roncalli, 2020a, pages 131-136). Therefore, if we compare the yield of a green bond with the yield of a regular bond issued by the same issuer, the difference must be equal to zero if the two bonds have the same characteristics or if they are twin bonds (Box 4.1). In practice, this is generally not the case. From a mathematical viewpoint, the greenium is defined as:

$$\boldsymbol{g} = y\,(\text{GB}) - y\,(\text{CB}) \tag{4.1}$$

where $y\,(\text{GB})$ is the yield (or return) of the green bond and $y\,(\text{CB})$ is the yield (or return) of the conventional twin bond[19]. Let $s = y - y^\star$ be the difference between the yield with default risk and the yield without default risk. Another expression of the greenium is:

$$
\begin{aligned}
\boldsymbol{g} &= s\,(\text{GB}) - s\,(\text{CB}) + \underbrace{y^\star\,(\text{CB}) - y^\star\,(\text{GB})}_{\approx 0} \\
&\approx s\,(\text{GB}) - s\,(\text{CB})
\end{aligned}
\tag{4.2}
$$

Therefore, we can also define the greenium as the spread difference between the green bond and the conventional bond.

<div align="center">Box 4.1: Green twin bonds</div>

> The twin bond concept has been introduced in 2020 by Germany. The underlying idea is that investors in German green bonds may swap their holdings with a conventional German government bond with the same maturity and coupon at any time, but not vice-versa. The objective is to increase the marketability of green bonds and improve the liquidity of the green bond market. The ability to compare two bonds from the same issuer with an equivalent maturity and coupon provides a direct measure of the greenium.
>
> On 3 September 2020, the 10-year German green bond with a coupon of 0.00% was priced 1 basis point below the 10-year conventional German bond. On 19 January 2022, Denmark issued a 10-year green bond with the same maturity, interest payment dates and coupon rate as the conventional 2031 Danish bond. The effective yield of the green bond was 5 basis points below the twin regular bond.

Remark 44 *The concept of bond yield (or bond return) is relatively complex, because there is not a unique approach to compute the financial return of a bond. Generally, we use the yield to maturity, but we can also use the credit spread if we prefer to measure the excess return. Another popular measure is the current yield, which is equal to the next coupon value divided by the current market price of the bond.*

Example 17 *We consider a 10-year green bond* GB_1 *whose current price is equal to* 91.35. *The corresponding conventional twin bond is a 20-year regular bond, whose remaining maturity is exactly equal to ten years and its price is equal to* 90.07. *We assume that the two bonds have the same coupon level[20], which is equal to* 4%.

---

[19]This means that the conventional bond has the same characteristics than the green bond.

[20]This assumption is not realistic since the regular bond has been issued 10 years before the green bond. In this case, we expect that the coupon of the regular bond was higher than the coupon of the green bond.

Box 4.2: Bond pricing

We consider that the bond pays coupons $C\left(t_m\right)$ with fixing dates t_m and the notional N (or the par value) at the maturity date T. The cash flow scheme is reported in Figure 4.A. Knowing the yield curve, the price of the bond without default risk at the date t satisfies the following relationship (Roncalli, 2020a, Equation 3.2):

$$P_t + AC_t = \sum_{t_m \geq t} C\left(t_m\right) B_t\left(t_m\right) + N\, B_t\left(T\right)$$

where $B_t\left(t_m\right)$ is the discount factor at time t for the maturity date t_m and AC_t is the accrued coupon. $P_t + AC_t$ is called the *dirty price* whereas P_t refers to the *clean price*. The yield to maturity of the bond is the discount rate which returns its market price:

$$\sum_{t_m \geq t} C\left(t_m\right) e^{-\left(t_m - t\right)y} + N e^{-\left(T - t\right)y} = P_t + AC_t$$

Figure 4.A: Cash flows of a bond without default risk
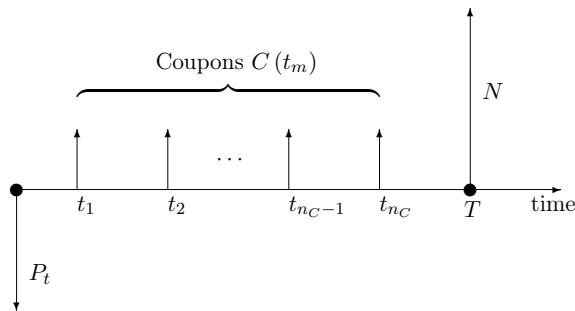
By introducing the credit risk of the issuer, the cash flows may be different because the issuer may default at time $\tau < T$ (Figure 4.B). Roncalli (2020a, Equation 3.3) shows that:

$$P_t + AC_t = \sum_{t_m \geq t} C\left(t_m\right) B_t\left(t_m\right) \mathbf{S}_t\left(t_m\right) + N\, B_t\left(T\right) \mathbf{S}_t\left(T\right) + \mathcal{R}\, N \int_t^T B_t\left(u\right) f_t\left(u\right)\,\mathrm{d}u$$

where $\mathcal{R}$ is the recovery rate, $\mathbf{S}_t\left(u\right)$ is the survival function at time u and $f_t\left(u\right)$ the associated density function. The yield to maturity of the defaultable bond is computed exactly in the same way as without default risk. The credit spread $s = y - y^\star$ is then defined as the difference of the yield to maturity y with default risk and the yield to maturity $y^\star$ without default risk.

Figure 4.B: Cash flows of a bond with default risk
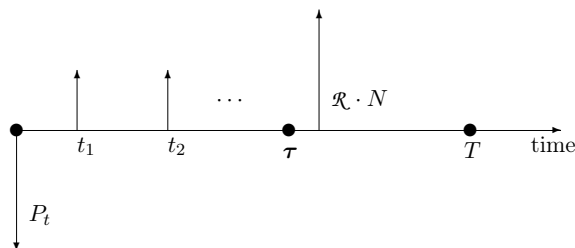
Let us consider Example 17. The computation of the yield to maturity[21] gives $y\,(\text{GB}) = 5\%$ and $y\,(\text{CB}) = 5.169\%$. We deduce that the greenium is equal to -16.9 bps. If we assess the bond return with the current yield, we have $y\,(\text{GB}) = 4/91.35 = 4.379\%$ and $y\,(\text{CB}) = 4/90.07 = 4.441\%$. In this case, we obtain $g = -6.2$ bps. We notice that the two measures are different even if the greenium is negative in both cases.

Figure 4.6: Greenium in bps of the German green bond (DBR 0% 15/08/2030)
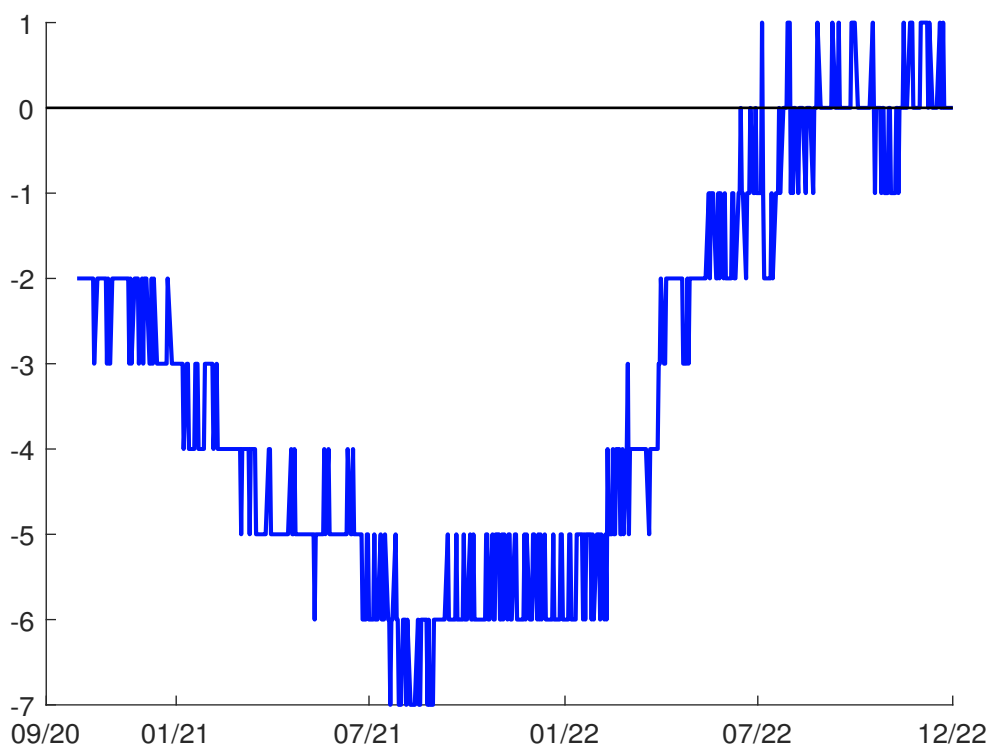
In the case of twin bonds, we can easily compute the greenium since the green and regular bonds have exactly the same characteristics and are issued at the same date. In Figure 4.6, we report the dynamics of the greenium for the German Bund 0% 15/08/2030. This analysis comes from the research study of Pástor *et al.* (2022). We observe that the greenium is always negative since the inception date (08/09/2020). On average, the greenium is equal to -3.27 bps. Its range is between -7 and $+1$ bps. Another illustration of the greenium is provided by Zerbib (2019), who analyzed the perpetual 5.5 year callable green hybrid bond that was issued by Iberdrola on 14 November 2017:

> "*At the beginning of the day, the coupon price was estimated at 2.2%–2.375%. The issue was quickly oversubscribed to 3.3 billion euros [compared to the initial offering of 1 billion euros], and the final coupon was eventually priced at 1.875%, i.e., 5 bps below the conventional benchmark*". (Zerbib, 2019).

In both cases (German and Iberdrola bonds), we must distinguish the greenium observed in the primary market (when bonds are originated) and the greenium priced in the secondary market (when bonds are traded). In the primary market, a negative greenium implies that the investor has bought

---

[21]We solve the equation $\sum_{t=1}^{10} 4e^{-ty} + 100e^{-10y} = P$ where $P = 91.35$ for the green bond and $P = 90.07$ for the conventional bond.

a green bond with a lower coupon rate compared to the coupon rate offered by the conventional bond. In the secondary market, a negative greenium implies that the investor has bought a green bond with a higher price compared to the market price of the conventional bond. Let c be the coupon rate. Mathematically, we have $c\,(\mathrm{GB}) \leq c\,(\mathrm{CB})$ in the first case and $P_t\,(\mathrm{GB}) \geq P_t\,(\mathrm{CB})$ in the second case.

The estimation of the greenium is a difficult task. First, we have to distinguish primary and secondary markets. Most of academic studies concern the secondary market, because there are few observations to compute the greenium in the primary market. Indeed, we can only have one observation per green bond in this last case (the day when the green bond is issued). Nevertheless, the different academic studies generally estimate a negative greenium between 5 and 15 bps on the primary market (Ehlers and Packer, 2017; Gianfrate and Peri, 2019; Fatica *et al.*, 2021; Kapraun *et al.*, 2021; Löffler *et al.*, 2021; Baker *et al.*, 2022). These results confirm the professional consensus that the greenium is negative and significant at the issuance date. However, the persistence of the negative greenium in the secondary market is an issue and an open debate. We observe two opposing sides: those who consider that the negative greenium persists and remains statistically significant, and those who think that the negative greenium vanishes. For instance, Zerbib (2019) found a greenium of -2 bps for EUR and USD global bonds from July 2013 to December 2017. While they measured a greenium of -18 bps in the primary market, the estimates of Gianfrate and Peri (2019) are respectively -11, -13 and -5 bps for three trading dates[22]. Generally, when academics estimate the greenium both in the primary and secondary markets, they observe that the discount premium is higher at the issuance date. However, some academic studies consider that the greenium in the secondary market is zero and statistically insignificant for municipal bonds (Larcker *et al.*, 2020) and corporate bonds (Tang and Zhang, 2020; Flammer, 2021).

The second issue when estimating the greenium is the choice of the bond yield return. In Figure 4.7, we report the value of the greenium computed with different spread measures in the case of the German green bond. All these measures give different results. If we consider the spread to worst, the average greenium is -2.73 bps versus -3.27 bps for the effective yield.

The last issue is the estimation method. According to Ben Slimane *et al.* (2020), there are two main approaches:

1. The bottom-up matching approach consists in computing the yield difference at the bond level. This means that we compare the green bond of an issuer with a synthetic conventional bond of the same issuer that has the same characteristics in terms of currency, seniority and duration (Zerbib, 2019). This matching methodology may be relaxed by considering a conventional bond of another issuer in the same country and industry and with the same rating (Flammer, 2021).

2. The top-down replication approach consists in computing the yield difference at the portfolio level. The underlying idea is to consider a diversified portfolio of green bonds and replicate it with a portfolio of conventional bonds. The objective of the replication process is to avoid biases in terms of currency, sector, credit rating, maturity, etc. Therefore, the greenium is the difference between the yield of the green bond portfolio and the yield of the replication portfolio.

In the bottom-up approach, we first filter all the conventional bonds, which has the same issuer, the same currency, and the same seniority of the green bond GB. Then, we select the two conventional

---

[22]They are 10 January, 7 July and 14 December 2017

Figure 4.7: Impact of the yield measure on the greenium (DBR 0% 15/08/2030)
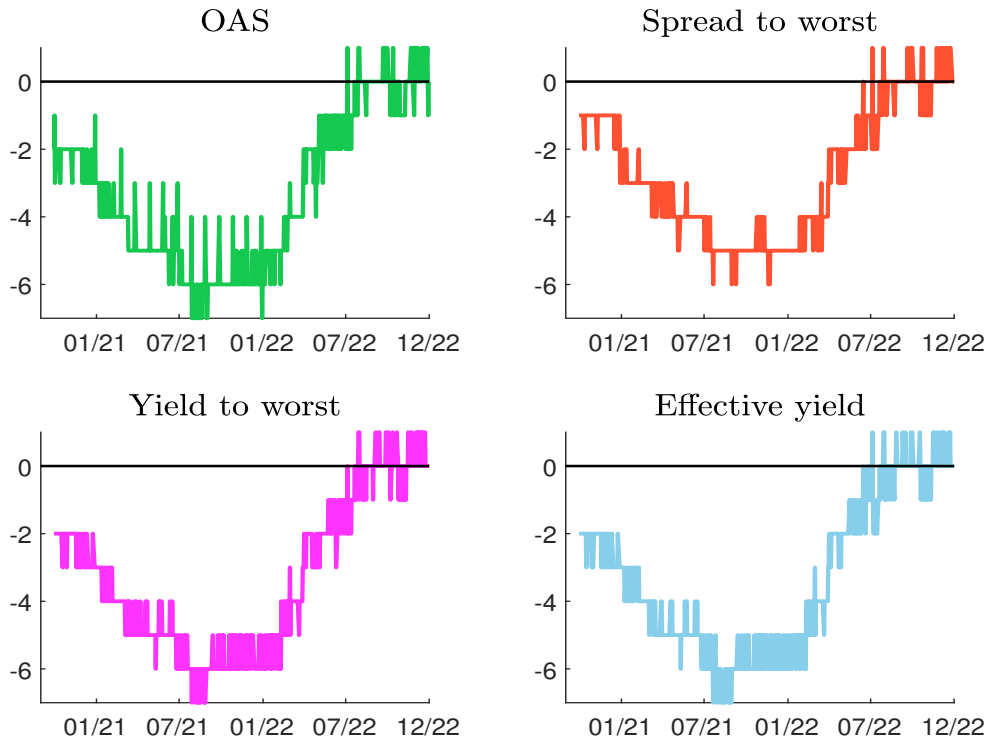
Source: ICE.

bonds CB_1 and CB_2 which are the nearest in terms of modified duration:

$$\left| \text{MD}\left(\text{GB}\right) - \text{MD}\left(\text{CB}_j\right) \right|_{j \neq 1,2} \geq \sup_{j=1,2} \left| \text{MD}\left(\text{GB}\right) - \text{MD}\left(\text{CB}_j\right) \right|$$

Finally, we perform the linear interpolation/extrapolation of the two yields $y\left(\text{CB}_1\right)$ and $y\left(\text{CB}_2\right)$ such that the modified duration of the synthetic conventional bond is exactly equal to the modified duration of the green bond. For instance, by assuming that $\text{MD}\left(\text{CB}_1\right) \leq \text{MD}\left(\text{CB}_2\right)$, the synthetic yield is:

$$y\left(\text{CB}\right) = y\left(\text{CB}_1\right) + \frac{\text{MD}\left(\text{GB}\right) - \text{MD}\left(\text{CB}_1\right)}{\text{MD}\left(\text{CB}_2\right) - \text{MD}\left(\text{CB}_1\right)} \left(y\left(\text{CB}_2\right) - y\left(\text{CB}_1\right) \right)$$

Example 18 *We consider a green bond, whose modified duration is 8 years. Its yield return is equal to* 132 *bps. We can surround the green bond by two conventional bonds with modified duration 7 and 9.5 years. The yield is respectively equal to* 125 *and* 148 *bps. The interpolated yield is equal to:*

$$
\begin{aligned}
y\left(\text{CB}\right) &= 125 + \frac{8-7}{9.5-7}\left(148 - 125\right) \\
&= 134.2 \text{ bps}
\end{aligned}
$$

It follows that the greenium is equal to -2.2 *bps:*

$$\boldsymbol{g} = 132 - 134.2 = -2.2 \text{ bps}$$

In the second approach proposed by Fender *et al.* (2019), we consider a portfolio w of green bonds. We have:

$$w = (w_1, \ldots, w_n)$$

where n is the number of green bonds. Then, we perform a clustering analysis by considering the 4-uplets (Currency $\times$ Sector $\times$ Credit quality $\times$ Maturity). Let (C_h, S_j, R_k, M_l) be an observation for the 4-uplet (e.g. EUR, Financials, AAA, 1Y-3Y). We compute its weight:

$$\omega_{h,j,k,l} = \sum_{i \in (C_h, S_j, R_k, M_l)} w_i$$

The greenium is then defined as the weighted excess yield:

$$\boldsymbol{g} = \sum_{h,j,k,l} \omega_{h,j,k,l} \left(y_{h,j,k,l} \left(\text{GB} \right) - y_{h,j,k,l} \left(\text{CB} \right) \right)$$

where $\omega_{h,j,k,l}$ is the weight of the cluster (C_h, S_j, R_k, M_l) in the green portfolio, $y_{h,j,k,l} \left(\text{GB} \right)$ is the yield of the cluster in the green portfolio and $y_{h,j,k,l} \left(\text{CB} \right)$ is the yield of the cluster in the benchmark portfolio.

Figure 4.8 shows the evolution of the greenium (expressed in bps), which has been computed by Ben Slimane *et al.* (2020) with the top-down approach. On average, it is negative and equal to -4 bps. The 95% confidence interval corresponds to the range -7.9 to -1.3 bps. Since the covid-19 crisis, we observe that the greenium tends to decrease in absolute value. Nevertheless, we notice that the greenium highly depends on the currency. The greenium of EUR-denominated bond is lower than the greenium of USD-denominated bond on average (-5.6 vs. -3.3 bps), but this is not the case in 2022 (-1.9 vs. -9.3 bps). The correlation between EUR and USD premia changes over time and is not very high in absolute value. For instance, it is equal to 29% since 2020. These differences do not only concern currencies, but also sectors, maturities, regions and ratings. For example, the greenium is not statistically significant for many sectors. Ben Slimane *et al.* (2020) also found that the volatility of green bond portfolios are lower than the volatility of conventional bond portfolios, implying that green and conventional bonds have identical Sharpe ratio since the last five years. To summarize, we can assume that the greenium is slightly negative, but the order of magnitude is relatively low.

4.2.2 Social bonds

Definition

In the mid of 2010s, the concept of green bonds has been extended to social objectives. The first social bond is issued in January 2015 by Spanish Instituto de Credito in order to help finance SMEs in economically depressed regions of Spain. In September 2015, Kutxabank issued the first social covered bond and the proceeds of the bond were used for financing and subsiding social housing projects in the Basque region. Since 2015, the framework of social bonds has evolved, but it is now a copy/paste version of the green bond framework. For example, the definition of a social bond is exactly the same as for green bonds:

> "*Social Bonds are any type of bond instrument where the proceeds, or an equivalent amount, will be exclusively applied to finance or re-finance in part or in full new and/or existing eligible social projects and which are aligned with the four core components of the Social Bond Principles (SBP).*" (ICMA, 2021b, page 3).

Again, the four core components are principles are based the use of proceeds, the process for project evaluation and selection, the management of proceeds and the reporting. The social project cate-

Figure 4.8: Evolution of the greenium (in bps)
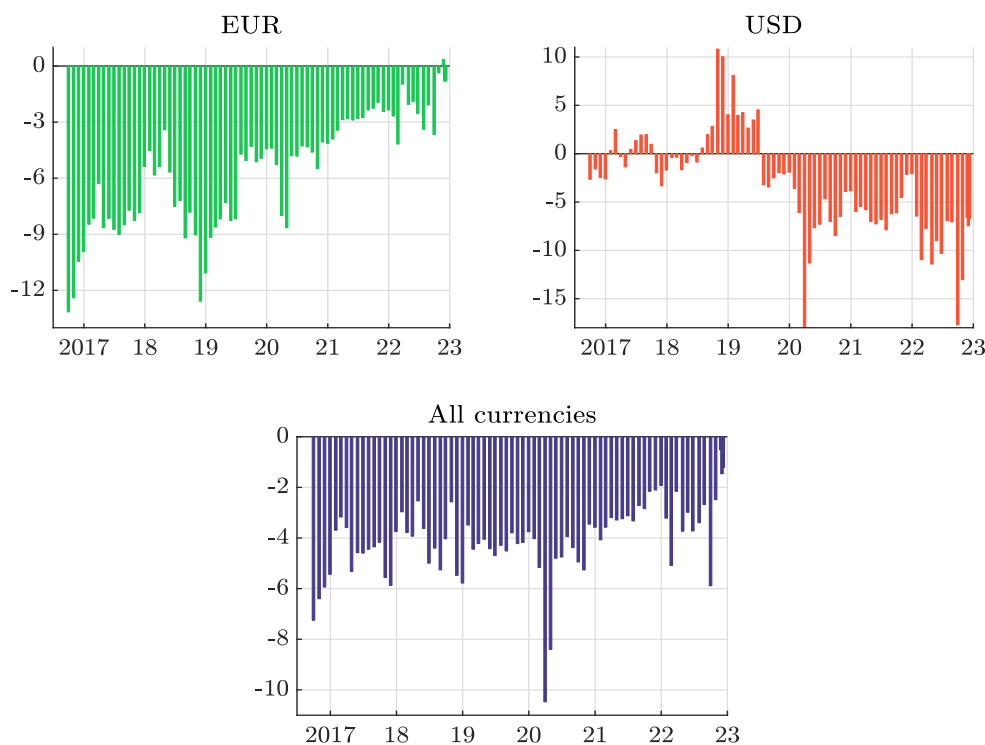
gories are affordable basic infrastructure[23], access to essential services[24], affordable housing, employ-
ment generation[25], food security and sustainable food systems[26] and socioeconomic advancement
and empowerment[27]. The use of proceeds also introduces the concept of target population, meaning
that the objective of a social bond is defined by both a social project category and a target popu-
lation. Examples of target populations are: (1) living below the poverty line, (2) excluded and/or
marginalised populations and/or communities, (3) people with disabilities, (4) migrants and/or dis-
placed persons, (5) undereducated, (6) underserved, owing to a lack of quality access to essential
goods and services, (7) unemployed, (8) women and/or sexual and gender minorities, (9) aging pop-
ulations and vulnerable youth and (10) other vulnerable groups, including as a result of natural
disasters. The three other components corresponds to the ones described in the Green Bond Princi-
ples. The only significant difference is that the SBP require that quantitative performance measures
include the number of beneficiaries, especially from target populations.

---

[23]E.g., clean drinking water, sewers, sanitation, transport, energy.

[24]E.g., health, education and vocational training, healthcare, financing and financial services.

[25]Including programs designed to prevent and/or alleviate unemployment stemming from socioeconomic crises,
SME financing and microfinance.

[26]E.g., physical, social, and economic access to safe, nutritious, and sufficient food that meets dietary needs and
requirements; resilient agricultural practices; reduction of food loss and waste; and improved productivity of small-
scale producers.

[27]E.g., equitable access to and control over assets, services, resources, and opportunities; equitable participation
and integration into the market and society, including reduction of income inequality.

The social bond market

According to CBI, the cumulative issuance of social debt amounts to \$515 bn at the end of June 2022. In Figure, we report the dynamics of the debt issuance. We notice a high growth in 2020, which is due to the issuance of social bonds to finance the covid debt. According to CBI, the market is dominated by European issuers (46%) and supranational issuers (29%). Most of social bonds are issued by the public sector (72%) followed by financials (15.7%), development banks (7.7%) and corporations (4.6%).

Figure 4.9: Issuance of social bonds

Source: https://www.climatebonds.net/market/data.
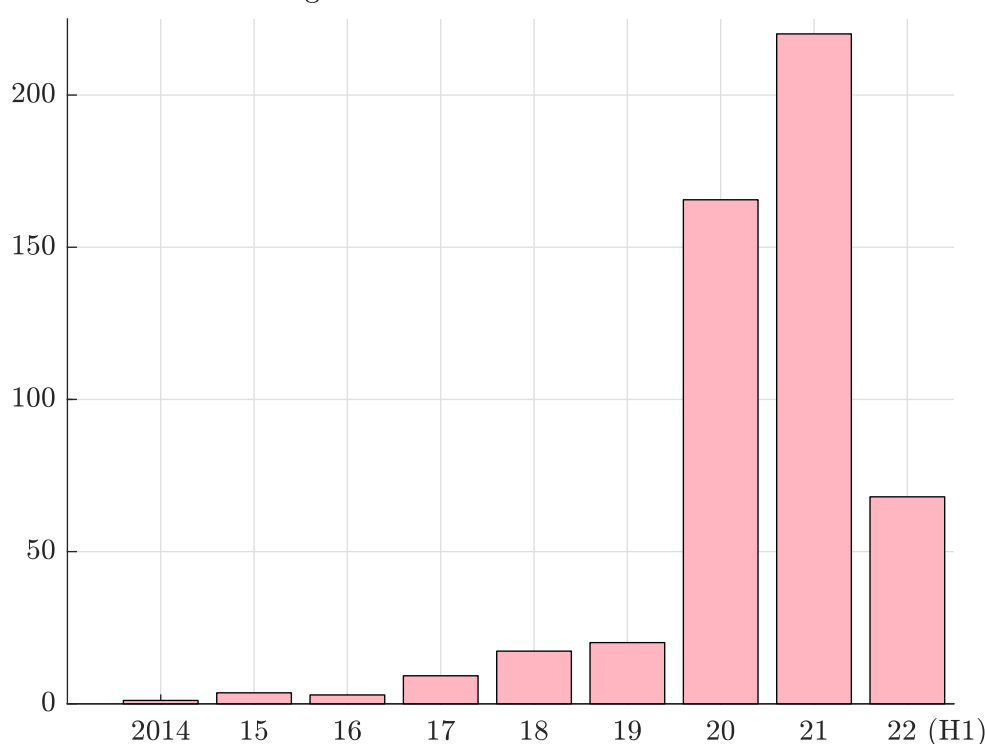
4.2.3 Other sustainability-related instruments

Sustainability bonds

Sustainability bonds are debt instruments that are issued to finance projects or activities that have both positive environmental and social impacts. The underlying idea is that some social projects may also have environmental co-benefits, and vice-versa. Sustainability bonds are aligned with the four core components of both the GBP and SBP. An example is the Series 2021 sustainability bonds issued by the American Museum of Natural History. The environmental objective is to partially finance the Gilder Center (green buildings), while the social objectives are to expand access to critical science education, and promote biocultural diversity and research. The social benefits accrue to target populations that include K-12 STEM education shortage areas and the general public.

Remark 45 *According to CBI, the cumulative issuance of sustainability bonds reaches \$620 bn at the end of June 2022.*

Box 4.3: Examples of social project categories and target populations

- Instituto de Crédito Oficial (Spanish state-owned bank, March 2020)
  "*The Social Bond proceeds under ICO's Second — Floor facilities will be allocated to loans to finance small, medium and micro enterprises with an emphasis on employment creation or employment retention in: (1) specific economically underperforming regions of Spain; (2) specific municipalities of Spain facing depopulation; (3) regions affected by a natural disaster. [...] The target populations are SMEs in line with European Union's standards.*"

- Pepper Money (non-bank lender in Australia and New Zealand, April 2022)
  "*The positive social impact of a Pepper Money eligible social project derives from its direct contribution to improving access to financial services and socio-economic empowerment, by using proprietary systems to make flexible loan solutions available to applicants who are not served by traditional banks. [...] Pepper Money is seeking to achieve positive social outcomes for a target population of Australians that lack access to essential financial services and experience inequitable access to and lack of control over assets. Pepper Money directly aims to address the positive social outcome of home ownership for borrowers who may have complexity in their income streams, gaps in their loan documentation or have adverse credit history. Traditionally, this cohort has been underserved by banks that rely on inflexible algorithmic loan application processing.*"

- Danone (French multinational food-products corporation, March 2018)
  "*The eligible project categories are: (1) research & innovation for advanced medical nutrition (target populations: infants, pregnant women, patients and elderly people with specific nutritional needs), (2) social inclusiveness (target poputions: farmers, excluded and/or marginalised populations and/or communities, people living under the poverty line, rural communities in developing countries), (3) responsible farming and agriculture (target populations: milk producers, farmers), etc.*"

- Korian (European care group, October 2021)
  "*The proceeds of any instrument issued under the framework will be used [...] to provide services, solutions, and technologies that will enable Korian to meet at least one of its social objectives: (1) to increase and improve long-term care nursing home capacity for dependent older adults; (2) to increase and improve medical capacity for people in need of medical support; (3) to increase and improve access to alternative, nonmedical services, technologies, and housing solutions that facilitate the retention of older adults' autonomy; and (4) to improve the daily provision of care to and foster a safer living environment for its patients. [...] Furthermore, Korian's target populations are older adults, which Korian defines as being over 65 years of age, and those who are dependent on others for some degree of care, which is defined by the health authorities or insurance system of the respective country.*"

- JASSO (Japan Student Services Organization, July 2022)
  "*The social project categories concern the financing of the 'Category 2 Scholarship Loans' (interest-bearing scholarship loans that have to be repaid) while the target population is made up of students with financial difficulties.*"

Source: Collected from the websites of the organizations.

Sustainability-linked bonds

A sustainability-linked bond (SLB) is a sustainability bond (green/social) plus a step-up coupon if the sustainability KPI is not satisfied. Therefore, a SLB belongs to the family of forward-looking performance-based instruments. The financial characteristics of the bond depends on whether the issuer achieves predefined ESG objectives. Those objectives are:

1. measured through predefined Key Performance Indicators (KPI);

2. assessed against predefined Sustainability Performance Targets (SPT).

Let us give some examples. In September 2019, ENEL issued a general purpose SDG linked bond based on the following SDGs: 7 (affordable and clean energy), 13 (climate action), 9 (industry, innovation and infrastructure) and 11 (sustainable cities and communties). The KPI is renewables installed capacity RIC as of December 31, 2021 while the SPT is equal to 55%. If the SDG 7 objective is not achieved[28] — $RIC < 55\%$, ENEL must pay a one-time step-up coupon of 25 bps. On April 2022, the independent report produced by KPMG certifies that "*the renewables installed capacity percentage as of December 31, 2021 is equal to* 57.5%". Since 2019, Enel has issued other sustainability-linked bonds[29].

On 18 February 2021, H&M issued a 8.5-year sustainability-linkedbond[30] for a notional of €500 mn. The annual coupon rate is 25 bps and the objectives to achieve by 2025 are:

KPI_1 Increase the share of recycled materials used to 30% (SPT_1);

KPI_2 Reduce emissions from the Group's own operations (scope 1+2) by 20% with 2017 as a baseline (SPT_2);

KPI_3 Reduce scope 3 emissions from fabric production, garment manufacturing, raw materials and upstream transport by 10% with 2017 as a baseline (SPT_3).

The global step-up rate is equal to:

$$\begin{aligned} r \quad = \quad & 40\% \times \mathbb{1}\left\{KPI_1 < SPT_1\right\} + \\ & 20\% \times \mathbb{1}\left\{KPI_2 < SPT_2\right\} + \\ & 40\% \times \mathbb{1}\left\{KPI_3 < SPT_3\right\} \end{aligned}$$

If the three objectives will be achieved, the step-up rate is equal to zero and the step-up coupon is not paid. Otherwise, H&M will pay a step-up coupon proportional to the step-up rate, which can takes the value 20% (KPI_2 is not achieved), 40% (KPI_1 or KPI_3 is not achieved), 60% (KPI_2 is not achieved, and KPI_1 or KPI_3 is not achieved) 80% (KPI_1 and KPI_3 are not achieved) or 100% (KPI_1, KPI_2 and KPI_3 are not achieved). The H&M sustainability-linked bond generated great interest, since it was 7.6 times oversubscribed.

According to (Berrada *et al.*, 2022), "*the SLB market has grown strongly since its inception. [...] Bloomberg identifies a total of 434 outstanding bonds flagged as 'sustainability-linked' as of February 2022. In contrast, in 2018, there was only a single SLB. The amount raised through the single 2018 SLB issue was $0.22 bn, whereas the total amount raised through all SLBs issued in 2021 was approximately $160 bn*". These authors also found that the large majority of SLB issues

---

[28]As of 30 June 2019, the KPI was equal to 45.9%
[29]See www.enel.com/investors/investing/sustainable-finance/sustainability-linked-finance/sustainability-linked-bonds.
[30]See https://hmgroup.com/investors/sustainability-linked-finance.

address exclusively **(E)** issues (65%) or a combination of **(E)**, **(S)** and **(G)** issues (17%) or **(E)** and **(G)** issues (3%). The other combinations are marginal (less than 1% for each). They also noticed that the most frequent KPI concerns GHG emissions (40 %), followed by the issuer's global ESG score (14 %).

Transition bonds

They are financial instruments to support the transition of an issuer, which has significant current carbon emissions. These bonds are typically used to fund projects such as renewable energy developments, energy efficiency upgrades, and other projects that are aimed at transitioning to a more low-carbon economy. The final objective of the bond issuer is always to reduce their carbon emissions. For example, transition bonds can be used to switch diesel powered ships to natural gas or to implement carbon capture and storage.

4.3   Sustainable real assets

4.3.1   Green infrastructure

4.3.2   Green real estate

4.3.3   ESG private equity and debt funds

Chapter 5

Impact Investing

5.1 Definition

- Barber *et al.* (2021)
- Caseau and Grolleau (2020)
- Geczy *et al.* (2021)

5.2 Thematic funds

5.3 Measurement tools

(CDSB, 2021a,b)

5.4 An example with the biodiversity risk

Chapter 6

Engagement & Voting Policy

Following GSIA (2021), corporate engagement & shareholder action is one of the seven categories of ESG strategies. It is defined as "*employing shareholder power to influence corporate behaviour, including through direct corporate engagement (i.e., communicating with senior management and/or boards of companies), filing or co-filing shareholder proposals, and proxy voting that is guided by comprehensive ESG guidelines*". Even if this category of ESG strategies can be found under different names — engagement and voting on sustainability matters for Eurosif, active ownership for the PRI — the scope of ESG engagement is well-marked. It refers to the process of interacting with companies to encourage them to adopt sustainable and socially responsible practices. This may involve discussing issues related to the corporate social responsibility and the sustainability impact of the business with company management and board members, and working with them to develop and implement practices that are aligned with the ESG principles of the shareholder. ESG engagement can be conducted by asset owners, asset managers, or organizations that seek to influence corporate behavior (e.g., Climate Action 100+). However, the ultimate goal of ESG engagement is always to align the ESG practices of corporations with the ESG expectations of investors. ESG engagement is often confused with the term ESG stewardship. In fact, ESG engagement is part of ESG stewardship, but this last one is a broader concept and refers to all the actions that asset owners and managers take to encourage companies to adopt sustainable and socially responsible practices. Of course, engagement is the cornerstone of stewardship because shareholder engagement and voting are the most direct ways to influence companies. Nevertheless, companies are also impacted when an investor implements an ESG scoring or publishes its exclusion list. Increasingly, we notice that the term stewardship replaces the term engagement. For instance, the publication of ESG stewardship reports by asset owners and managers has replaced the publication of ESG engagement in the last three years. In February 2021, the PRI published the guide "Stewardship" on active ownership:

> "*It guides investors on how to implement the PRI's Principle 2, which sets out signatories' commitment to stewardship, stating: we will be active owners and incorporate ESG issues into our ownership policies and practices. [...] The PRI defines stewardship as the use of influence by institutional investors to maximise overall long-term value including the value of common economic, social and environmental assets, on which returns and clients' and beneficiaries' interests depend.*" (PRI, 2021a).

Even if ESG engagement and ESG stewardship are closely related concepts, ESG stewardship is generally interpreted by investors as their ESG policy, also including rating models, strategies, organizations, etc. In this chapter, we focus on active ownership in a first section and defines the various forms of ESG shareholder activism. In a second section, we study the voting policy of ESG investors.

6.1   Active ownership

6.1.1   Definition

From an ESG viewpoint, the terms active ownership, engagement, and shareholder activism are interchangeable. Shareholder activism is certainly most frequently used by academics, while professionals prefer to speak about engagement and active ownership. A shareholder activist is a shareholder who uses his equity stake in a (listed) company in order to influence the board members, and change the way the firm is managed. For instance, Gillan and Starks (2000) define active shareholders as "*investors who, dissatisfied with some aspect of a company's management or operations, try to bring about change within the company without a change in control*". Changes may concern the business model, the strategy or the ESG policy. On the opposite, passive shareholders are investor who hold shares of the company, but they have no intention to have an influence on the company. When we refer to ESG engagement, the issues focus on sustainable and socially responsible practices.

Conflicting interests between shareholders and management are well-documented and are central in the theory of the firm (Williamson, 1970; Jensen and Meckling, 1976). Indeed, shareholders and management may pursue different objectives. In particular, the concept of "*managerial entrenchment*" refers to the tendency of managers to act in their own self-interest with a short-term time horizon rather than in the interests of shareholders, whose objective is to improve the long-term performance of the company. In this context, managers can seek to maximise their own utility and not the utility of the shareholders, who are the owners of the firm. For example, managers have incentives that firms grow beyond the optimal size, because it increases their power, the resources under their control and also their compensation (Murphy, 1985). In a similar way, the separation of ownership and control, the social responsibility of business, and the definition of a corporate objective function may result in misalignment of preferences between shareholders and the company board (Jensen and Meckling, 1976). In this context, active shareholders can serve as effective monitors of management behavior, especially when agency costs arise (Jensen, 1993) or negative externalities may generate large potential costs.

The debate about the separation between ownership and control is amplified when we introduce the stakeholder theory (Freeman, 2004). In fact, there may be some conflicts between shareholders and stakeholders, because shareholders do not have necessarily ESG preferences. On page 3, we have already mentioned the debate between the classical shareholder theory and the stakeholder theory. For Friedman (1970), "*the social responsibility of business is to increase its profits*". Nevertheless, the shareholder vs. stakeholder debate has changed over time, because more and more investors, especially institutional investors, have now sustainability preferences. Moreover, the corporate social responsibility has evolved quite spectacularly since the seminal publication of Bowen (1953) — see for instance Drucker (1954), Jones (1980), Mintzberg (1983), Drucker (1984), Wood (1991), Mitchell *et al.* (1997), Carroll (1999), Crowther and Aras (2008), Carroll and Shabana (2010), Jha and Cox (2015) and Yuan *et al.* (2020). It is now well-accepted that corporations have also social responsibilities. It is summarized by Peter Drucker as follows: "*leaders in every single institution and in every single sector ... have two responsibilities. They are responsible and accountable for the performance of their institutions, and that requires them and their institutions to be concentrated, focused, limited. They are responsible also, however, for the community as a whole*". For corporations, this implies an unlimited liability clause with respect to their employees, their consumers, and the society as a whole. One of the main objectives is then to minimize the legal issues. This is specially true for the negative externalities on the environment caused by the company. This echoes the liability risks defined Carney (2015):

> "*[...] the impacts that could arise tomorrow if parties who have suffered loss or damage*

from the effects of climate change seek compensation from those they hold responsible.
Such claims could come decades in the future [...]."

The current context is then complectly different than the 1970's, when Milton Friedman wrote his famous article. Today, corporate social responsibility is no longer an option. Nevertheless, this does not mean that there is a full alignment between management, shareholders and stakeholders.

6.1.2   The various forms of active ownership

According to Bekjarovski and Brière (2018), active ownership can take various forms:

1. Engage behind the scenes with management and the board;

2. Propose resolutions (shareholder proposals)

3. Vote (form coalition, express dissent, call back lent shares);

4. Voice displeasure publicly (in the media);

5. Initiate a takeover (acquire a sizable equity share)

6. Exit (sell shares, take an offsetting bet, divestment);

The first two approaches take the form of dialogue between investors and companies, and involving a direct communication with company management and board members to discuss ESG issues. This is particularly true with the first approach. It is also the case of the second approach, because resolutions are discussed before annual general meetings (AGM). The third approach can be viewed as a one of the financial duties of investors. This is certainly the most common way to participate in the life of the company one a year. The last three approaches take the form of an action or a response when investors do not approve the sustainable policy of the company or the outcomes of the annual general meeting. We generally find all these six approaches (except initiate a takeover) in the engagement/sustainable/stewardship reports of asset owners and managers (Figure 6.1).

Remark 46 *The first five approaches are complementary, while exit corresponds to a non-return situation. In the last case, engagement process between the investor and the company is stopped since the investor is no longer a shareholder of the company.*

Engage behind the scenes

Meeting the management and the board of a company was very rare twenty years ago. In addition to public meetings following financial performance announcements, communication between investors and investees were generally organized as a morning breakfast event, where a selection of equity or credit analysts is invited to discuss with a representative of the firm, who was generally the chief financial officer. This situation has changed over time, especially with the development of extra-financial analysis. Today, private communications between ESG investors and companies have become very common. Engage behind the scenes corresponds to all the meetings or actions made by the active shareholder to better understand the ESG strategy of firms and challenge companies on some sustainable issues:

"*Behind the curtain engagement involves private communication between activist share-*
holders and the firm's board or management, that tends to precede public measures
such as vote, shareholder proposals and voice. In a sense, the existence of other forms

Figure 6.1: Example of engagement/sustainable/stewardship reports
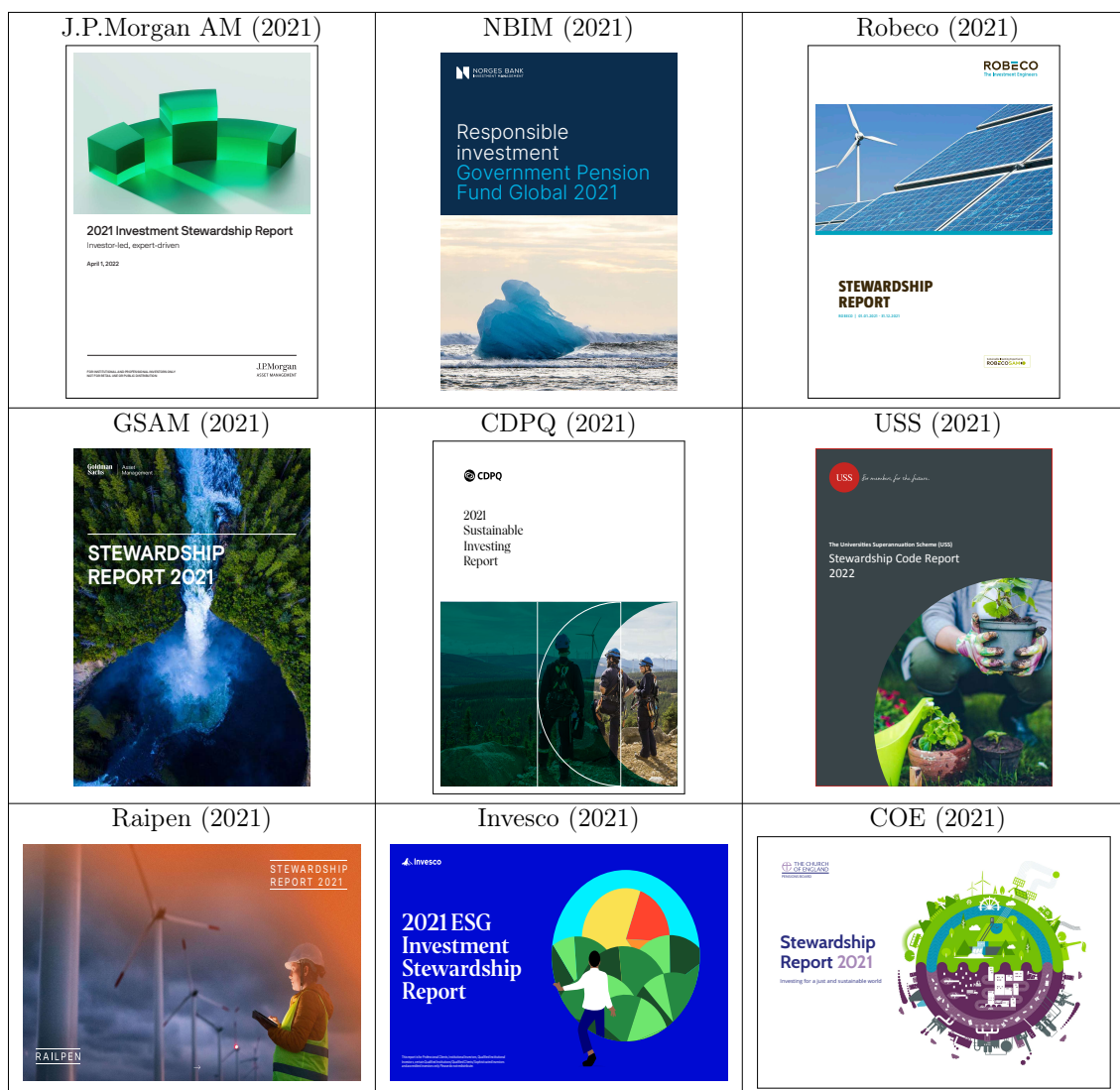
of public activism can be taken as a signal that behind the scene engagements were unsuccessful. When it comes to environmental and social issues, writing to the board or management is a common method though which shareholders can express concern and attempt to influence corporate policy behind the curtain; alternatively, face to face meetings with management or non-executive directors are a more common behind the scene engagement method when it comes to governance." (Bekjarovski and Brière, 2018, page 10)

In fact, we can classify these face-to-face meetings and more formal exchanges (letters and position statements) into three families:

1. on-going engagement, where the goal for investors is to explain their ESG policy and collect information from the company. For instance, they can encourage companies to adopt best ESG practices, alert companies on ESG risks or better understand sectorial ESG challenges;

2. engagement for influence (or protest), where the goal is to express dissatisfaction with respect to some ESG issues, make recommendations to the firm and measure/control ESG progress of companies;

3. pre-AGM engagement, where the goal is to discuss with companies any resolution items that the investor may vote against.

For many years, engage behind the scenes was very informal and mainly depends on ESG analysts. This is no longer true. Most of investors have now established an engagement process, which is generally based on three steps. First, a list of engagement issues is produced with a few items. The underlying idea is to focus on very important topics, and not to cover anything and everything. Second, a screening is performed for each engagement item in order to identify the most serious cases. Finally, the engagement can begin with the targeted companies. As noticed by the PRI (2019b), policies and processes are important to be formalized, but the ultimate goal of active ownership is obtaining outcomes. Therefore, investors must track their engagements. The different stages are:

- Issues are raised to the company;

- Issues are acknowledged by the company;

- The company develops a strategy to address the issues;

- The company implements changes and the issues are resolved;

- The company did not solve the issues and the engagement failed.

Remark 47 *Even if investors claim that transparency is the pillar of stewardship and engagement practices, they never publish the list of targeted companies and the corresponding issues. Sometimes, they give anonymous examples in their engagement reports.*

It is very difficult to obtain public statistics about the engagement behind the scenes and its trend. For instance, McCahery *et al.* (2016) noticed that "63% *of the respondents state that, in the past five years, they have engaged in direct discussions with management, and* 45% *have had private discussions with a company's board outside of management's presence*", but these results are based on a survey of 143 respondents relative to 3 300 invitations that were sent between December 2012 and July 2013. More recently, the study of Barko *et al.* (2022) is based on a proprietary database, which has been provided by a large European asset manager. We can always find figures from stewardship and engagement reports, but it gives a partial view of this topic. For example, Amundi (French asset manager, €2 064 tn in assets under management) reports 2 334 engagements in 2021 with the following breakdown by ESG themes[1]: dialogue to foster a stronger voting exercise and a sounder corporate governance (1 033), transition towards a low carbon economy (547), strong governance for sustainable development (287), social cohesion through the protection of direct and indirect employees (222), natural capital preservation (165) and product, client, societal responsibility (80). The first category corresponds to the 2021 pre-AGM dialogue statistics, meaning that the Amundi corporate governance team conducted dialogue with 1 033 issuers in 2021. The 2 334 engagements can be split into 397 soft engagements, 904 active engagements, 654 voting alerts and 379 pre-AGM dialogues. Amundi also indicate that they engaged with 1 364 unique companies in 2021. A second example of asset managers is Robeco, which is a Dutch asset manager with €200.7 bn in assets under

---

[1]Source: Amundi (2022), 2021 Engagement Report, https://about.amundi.com/esg-documentation.

Figure 6.2: Difference between stewardship and engagement reports

Amundi Stewardship
Report (2021)

Amundi Engagement
Report (2021)

Source: Amundi corporate website, https://about.amundi.com/esg-documentation.

management at the end of 2021. They reported 270 engagement cases[2] (79 environmental issues, 76 social issues, 52 governance issues, 35 SDG issues and 28 global controversy issues), while the number of engagement activities was 942 including 393 conference calls, 402 written correspondence, and 4 physical meetings. Concerning asset owners, PGGM (Dutch pension fund with €293.5 bn in assets under management at the end of 2021), reported 154 company engagements in their 2021 integrated annuel report[3]. A second example of asset owners is the Norway's sovereign wealth fund, which managed €1.24 tn at the end of 2021. They held a total of 2 628 meetings with 1 163 companies[4], and they had written communication with 486 companies in 2021. They also gave the breakdown by topics[5]: Ⓔ climate change (797), circular economy (190), biodiversity (48), ocean sustainability (18), etc.; Ⓢ children's rights (40), data privacy (34), customer interests (129), etc.; Ⓖ effective boards (267), remuneration (183), protection of shareholders (68), etc. These four examples concern large investors, but small investors also communicate on their engagement policy. Platypus Asset Management, an Australian boutique firm with $5 bn in assets under management, conducted 66 one-on-one meetings on ESG issues in 2021. Their 2021 engagement report[6] is very transparent since they listed the date of each meeting, the name of the company, the ESG issue and a summary of the meeting[7].

---

[2]Source: Robeco (2022), Active Ownership Report Q4 2021, www.robeco.com/en-int/sustainable-investing/influence.

[3]Source: PGGM (2022), 2021 Integrated Annual Report, www.pggm.nl/en/integrated-report.

[4]Source: NBIM (2022), Government Pension Fund Global 2021 — Responsible Investment, www.nbim.no/en/publications/reports/2021/responsible-investment-2021.

[5]The sum is greater than 2 628, because several topics can be discussed during one meeting.

[6]Source: Platypus Asset Management (2022), Engagement Report 2021, www.platypusassetmanagement.com.au/~/media/platypus/documents/media/engagement-report.ashx.

[7]For instance, on 9 June 2021, they met FPH on general governance. Here is the summary: "*Discussed Fisher and Paykel's approach to ESG including carbon — embodied emissions, very impressive science-based reduction targets including Scope 3, dovetails nicely with focus on product quality/need to scrutinise supply chain. The CEO is focused*

Propose resolutions

According to the SEC (Securities Exchange Act Rule 14a-8, §240), "*a shareholder proposal or resolution is a recommendation or requirement that the company and/or its board of directors take action, which the shareholder intend to present at a meeting of the company's shareholders. The proposal should state as clearly as possible the course of action that the shareholder believes the company should follow. If the proposal is placed on the company's proxy card, the company must also provide in the form of proxy means for shareholders to specify by boxes a choice between approval or disapproval, or abstention.*" Generally, shareholder resolutions are opposed by the management. Nevertheless, all submitted proposals to the management are not necessarily accepted. In the US, if a company want to exclude a shareholder proposal, they can write a letter to the SEC explaining that the proposal violates one or more conditions of SEC Rule 14a-8 (Matsusaka *et al.*, 2016). Shareholder resolutions can be excluded because of procedural requirements[8] or substantive bases[9]. Then, the SEC may or not accept the exclusion. If the SEC accepts the exclusion, the company receives a "*no-action*" letter, indicating that the SEC will take no action if the company leaves the proposal out of the proxy statement[10]. In other countries, the filing of shareholder resolutions and the acceptance rules are different. In France, Germany and UK, shareholders can submit a draft resolution to the company if they own at least 5% of the capital. This threshold is equal to 2.5% in Italy, 0.33% in the Netherlands and 3% in Spain. If shareholders organise together, the rule is the same in France, Germany, Italy and the Netherlands. In Spain, there is no provision for shareholders to organise together. In the UK, they must represent 100 shareholders and at least £10,000 in investment.

A shareholder resolution can be viewed as an escalation in the context of failed engagement or a lack of responsiveness by the company. It is a way for investors to publicly display that they are not happy with the management. Nevertheless, even if the resolution is approved by a majority of shareholders, its implementation may be an issue. For instance, shareholder resolutions are not binding in the US. This is not the case in Europe, but the management of the company can always delay the implementation or consider a partial implementation. In this context, shareholder resolutions are more viewed as a negative signal or a dissent sent to the company and the market, especially when they get media coverage and attention. Nevertheless, voted shareholder resolutions

---

on emissions and potential cost impact of the transition on the business. Also discussed gender diversity and modern slavery".

[8]The procedural requirements are described in Rules 14a-8(b)-(h): proponent must have held stock worth $2 000 or 1% of firm value continuously for at least one year before submitting proposal and must continue to hold them through meeting date; proponent may only submit one proposal per meeting; proposal and supporting statement may not exceed 500 words; proposal must be submitted no less than 120 days before proxy statement is mailed; proponent or representative must be present at meeting (Matsusaka *et al.*, 2016, Table 1).

[9]The thirteen substantive bases are described in Rule 14a-8(i): improper subject for action under state law; will cause the company to violate state, federal, or foreign law to which it is subject; proposal and supporting statement are materially false or misleading; relates to redress of a personal claim or grievance, or be designed to provide a benefit to proponent that is not shared by the other shareholders at large; relates to operations that account for less than 5 percent of company assets or sales; company lacks the power to implement; deals with ordinary business operations; would disqualify a director candidate, remove a director from office, question competence of director or nominee, seek to include specific nominee, or otherwise affect the outcome of director election; conflicts with company's own proposal; company has already substantially implemented proposal; substantially duplicates another proposal; deals with substantially the same subject as another proposal from previous years that received (specified) low support from shareholders; relates to specific amounts of dividends (Matsusaka *et al.*, 2016, Table 1).

[10]On page 3, we have seen that shareholders organized resolutions against the production of napalm during the Vietnam War. For instance, the Medical Committee for Human Rights filed a shareholder proposal in 1969 to force the Dow Chemical Company to stop its production of napalm. The SEC allowed Dow to exclude this shareholder proposal from its 1970 proxy voting, and therefore the proposal was not presented to the shareholders at the 1970 annual general meeting.

are exceptional. For example, we report below some figures[11] for companies in the Russell 3000 index as of proxy season 2022 (from 1st June 2021 to 30 June 2022):

- 98% of proposals are filed by the management, while less than 2% corresponds to shareholder resolutions;

- Only 60% of shareholder resolutions are voted; The other 40% are omitted, not presented, withdrawn or pending;

- The average number of proposals per company is around two;

- The proponents of shareholder resolutions are concentrated on a small number of investors or organisations (15 proponents were responsible of 75% of shareholder proposals);

- The repartition of shareholder proposals voted in 2022 was the following: 11% related to **E** issues, 41% related to **S** issues and 48% related to **G** issues.

These figures show that there are few shareholder proposals and they are genially filed by the same group of investors. In fact, many shareholder proposals that are submitted to proxy voting are withdrawn following negotiations between shareholders and the management before the annual general meeting[12]. Therefore, few investors have a true experience of an shareholder resolution that is voted during the AGM.

Vote

In just five years, voting at annual general meetings has become the norm for ESG investors. It is now considered as a fiduciary duty of asset owners and managers and the evidence that an investor is socially responsible. Today, voting choices are under scrutiny and are analyzed by many associations and NGOs. For instance, new editions of the Voting Matters series published by ShareAction[13] are very much awaited with fear and apprehension by asset managers.

   The voting landscape has evolved considerably these last years, especially in Europe. In the US, there is a long history that mutual funds vote at general meetings, because the SEC has always seen voting as one of the main fiduciary duties of collective funds[14]. This explains that US mutual funds have a higher voting participation than European mutual funds. In fact, the ESG fad has changed in many different ways the activity of voting, which is now one of the priorities from investors. First, the gap between asset owners and asset managers has been considerably reduced whereas the voting participation of asset managers was poor ten years ago (Brière *et al.*, 2020). Second, voting infrastructures have been strongly developed in Europe, while they exist in the US for many

---

[11]We use the reports of Rosati *et al.* (2022) and Tonello (2022).

[12]An example can be found in the 2021 Stewardship report of Robeco: "*At the end of 2020, we filed a shareholder resolution at ADM's 2021 shareholder meeting, asking the company to step up its efforts to eliminate deforestation in its soy supply chain. After several weeks of intense negotiations, spanning across multiple meetings with ADM's head of sustainability and corporate secretary, we managed to get the company to agree to most of our key requests and so we withdrew the proposal. Our achievement was to ensure that ADM published a revised no-deforestation policy, and committed to eliminate deforestation from all its supply chains by 2030.*"(Robeco, 2021, page 32). ADM is a company specialized in food, pet and animal nutrition. Other examples can be found in the stewardship report of Amundi, Candriam, Groupama Asset Management, etc.

[13]ShareAction is a UK-based registered charity that promotes responsible investment.

[14]You may consult the speech "*Every Vote Counts: The Importance of Fund Voting and Disclosure*" by the Commissioner Allison Herren Lee at the 2021 ICI Mutual Funds and Investment Management Conference (March 17, 2021), www.sec.gov/news/speech/lee-every-vote-counts.

Figure 6.3: Voting Matters series of ShareAction
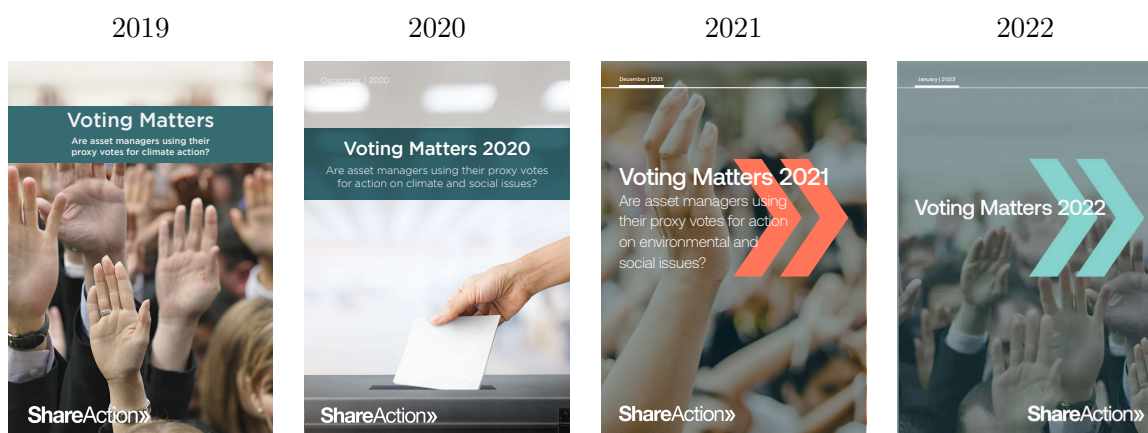
years. Nevertheless, the proxy advisory market is dominated by two US proxy advisory agencies[15] (Institutional Shareholder Services or ISS, and Glass Lewis), that represent 97% of the market. Third, asset managers have a high pressure to vote if they want to be credible. Voting has become a central pillar of any responsible investment policies, and it is now part of the ESG credentials when asset owners select asset managers.

In 2002, the United Kingdom adopt a legislation requiring companies to allow the shareholders to have a mandatory but non-binding vote on the executive compensation at each annual general meeting. This concept is called *say on pay*. According to (Rosati *et al.*, 2022), results for 2022 season has showed an increasing of shareholder opposition on this topic. For instance, support for executive remuneration was equal to 87% for Russell 3000 companies in 2022 compared to 89% in 2021. The lowest value was reached by Norwegian Cruise Line Holdings with only 15.4% of votes for. In Europe, we observe the same phenomenon. In Germany, 25% of say on pay votes were contested. In France and Spain, the most contested resolution was remuneration policy proposals. For example, about 50% of say on pay resolutions received at least 10% shareholder resolution of France. To have a point of comparison, the average support rate for management proposals is generally very high and greater than 95%.

Say on climate is inspired by say on pay proposals. This initiative[16] was launched by the hedge fund activist investor Chris Hohn through the Children's Investment Fund Foundation (CIFF) in 2020. Since then, it has 54 investor members ($14 tn AUM) such as Boussard & Gavaudan, CDPQ, Oxford University Endowment Management, Sarasin & Partners, Soros Fund Management or TCI Fund Managament, and it is supported by CIFF, CDP, ShareAction and the Australasian Centre for Corporate Responsibility (ACRR). More generally, say on climate is any shareholder or management resolution on the climate strategy of the company. When it is filed by the company's management, they expect that shareholders will vote for (shareholder approval of the climate strategy). When it is filed by shareholders, the resolution is against the climate strategy of the company. In 2021, 26 companies[17] have submitted a climate strategy at their annual general meetings[18] and there were 88

---

[15]Their websites are www.issgovernance.com/solutions/proxy-voting-services and www.glasslewis.com.

[16]The website is www.sayonclimate.org.

[17]19 were at European companies (3 in France, 5 in Spain, 1 in Switzerland and 10 in the UK), 3 in North America (1 in Canada and 2 in the US), 3 in South Africa and 1 in Australia.

[18]In 2022, the number of companies has increased and reached 36. Among them, we find seven oil & gas companies

climate-related shareholder proposals that were voted (ISS Governance, 2022). The average support rate was 93% for resolutions filed by the company and 32.7% when resolutions have been proposed by shareholders.

As we have previously seen, there are few shareholder proposals per company. The number of voted shareholder proposals is even more smaller. For instance, there were only 555 shareholder resolutions that have been voted in 2022 among the Russell 3000 companies, meaning less than 1 resolution for 5 companies. In Figure 6.4, we have reported the average support rate. In 2022, it is less than 40% for governance topics and it falls to 29% for environmental and social issues. Finally, the number of voted proposals receiving majority support is equal to 82. This means that there was one shareholder resolution that was adopted for 37 companies. Therefore, we may question the impact and the effectiveness of vote. This explains that vote is another form of voice for many academics.

Figure 6.4: Average support rate of shareholder proposals (Russell 3000 companies)
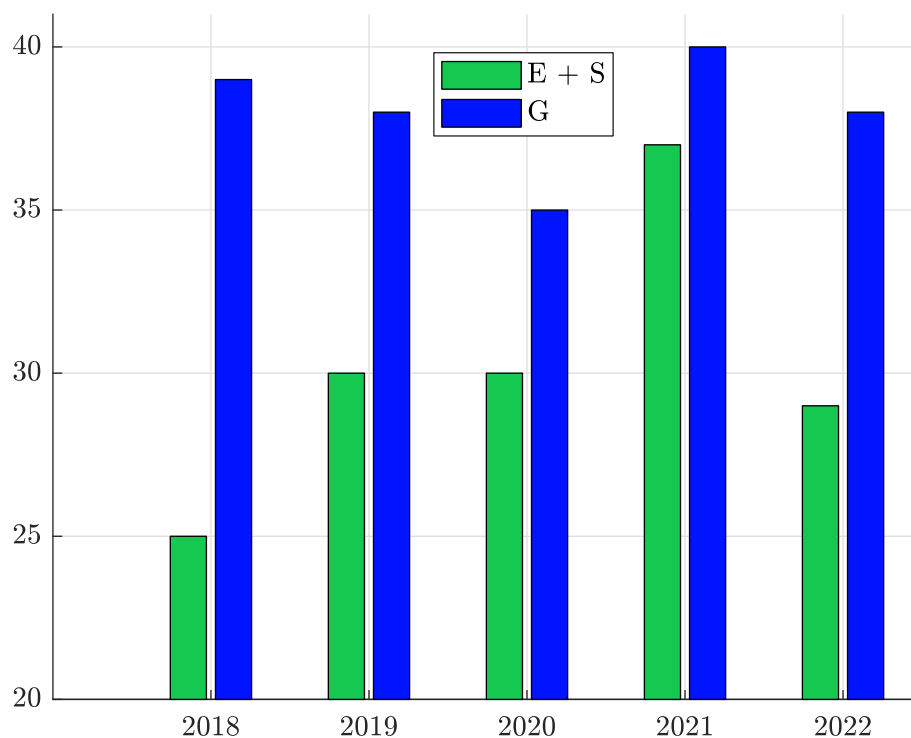
Source: PwC's Governance Insights Center (2022).

If we focus on the year 2022 and the US, the main topics of shareholder proposals were emissions reduction targets and scope 3 inclusion for the (**E**) pillar, diversity, civil rights and racial equity audits, human rights and workforce harassment for the (**S**) pillar, and executive compensation and political lobbying for the (**G**) pillar (Rosati *et al.*, 2022). In Europe, most of shareholder resolutions concern the remuneration of the management.

Remark 48 *The voting behavior of asset managers is analyzed in Section 6.2.4 on page 243.*

(BP, Equinor, Repsol, Santos, Shell, TotalEnergies, Woodside Petroleum).

Figure 6.5: Pass rate of shareholder proposals (Russell 3000 companies)
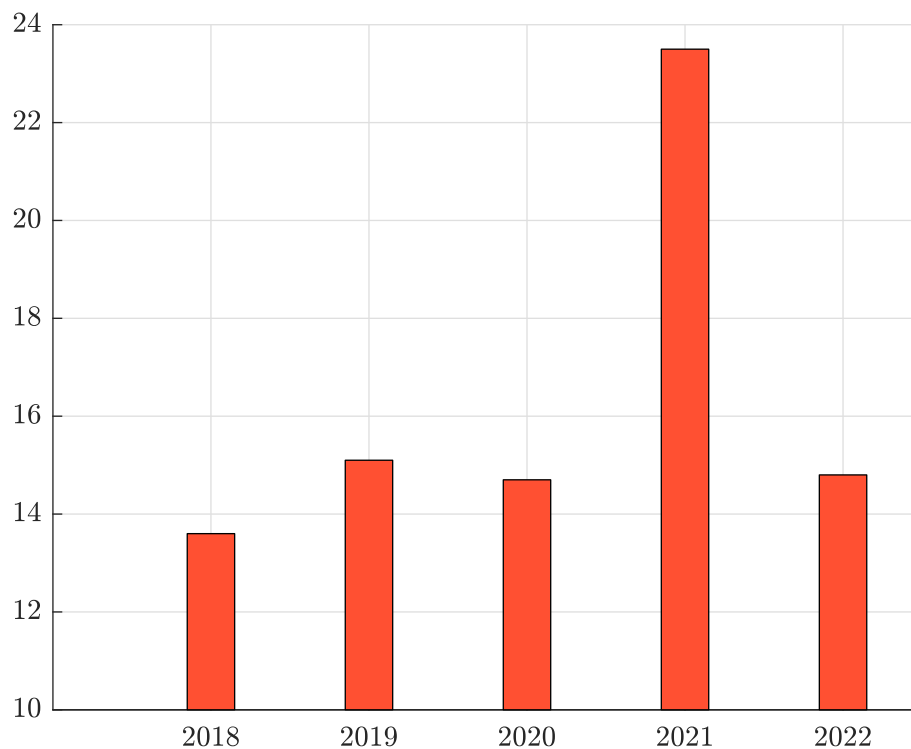
Voice

The term voice has been popularized in 1970 by the economist Albert Hirschman, when he published its book "*Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States*". The book states that members of an organization have two main possible responses when they consider that the organization fails. They can exit (withdraw from the relationship) or they can voice (change the relationship through communication of the complaint). From a socio-economic viewpoint, exit is associated with Adam Smith's invisible hand while voice is by nature political and may be confrontational. The exit-voice model of Hirschman (1970) can be applied to many situations: protest movements, migration, class action, and even corporate governance (Kostant, 1999).

In the theory of shareholder activism, voice may refer to several forms of engagement, for instance engage behind the scenes or propose resolutions (Edmans, 2014). In this chapter, voice refers to a situation where the communication between activist shareholders and the company becomes public. Like resolution proposals, voice can be interpreted as a form of escalation. Indeed, the debate and the disagreement become publicly known, implying that other stakeholders are informed and can participate, e.g., other shareholders, media, politics and the society. The company faces then the risk to be placed in the political spotlight and respond to public criticism and scrutiny. The situation can become uncomfortable for the company, especially when non-technical stakeholders (media and the society) look at the issue.

According to McCahery *et al.* (2016), voice is rarely used by individual institutional investors. One of the reason is that public communications may be ineffective, justifying the prevalence of behind the scenes communications (Levit, 2019). However, voice has increased during the recent

years due to two main factors. The first one is that collaborative engagement between investors are now quite common thanks to initiatives such as PRI or Climate Action 100+. In this context, they require more transparency by tracking progress of each member. For instance, when signing on to the initiative Climate Action 100+, investors are asked to nominate which focus companies they wish to engage with and whether this is as a lead investor or collaborating investor. The second factor is the increasing involvement of NGOs in the debate on engagement and greenwashing[19]. In this context, investors must sometimes publicly answer to greenwashing suspicions or allegations for supporting and financing some companies that are considered by these NGOs as detrimental to the environment or harmful to human health. All these elements explain why voice against companies, but also voice against investors, has recently gained in importance.

Initiate a takeover

Acquiring a large proportion of the company' shares is an agressive form of shareholder activism. The underlying idea is to form a coalition with other large shareholders in order to obtain a board set and to ultimately control the company's board. This strategy is typically implemented by activist hedge funds (Gantchev, 2013). Generally, managers perceive activist events as a hostile act, because their compensation and job security are threatened Gantchev *et al.* (2019). Most of the time, this form of shareholder activism has a big impact on the firm (change in business strategy, implementation of the proposed solutions, etc.). Nevertheless, it has never been used by ESG investors. Therefore, it is solely motivated by financial underperformance, and not extra-financial issues.

Exit

Exit refers to the process of selling off investments in a particular company or industry. Divestment is a more general term that implies a significant exposure reduction. In this case, we speak about partial divestment while exit corresponds to a full divestment[20]. For example, investors may decide to divest from a company if they are not satisfied with its ESG performance, in particular if the company policy implies social issues or it does not take sufficient action to address environmental concerns. Investors may also decide to divest from a sector such as fossil fuels or tobacco. Therefore, the exit policy is related to two ESG investing strategies[21]: exclusion and norms-based screening. Nevertheless, divestment could not be equated to these ESG strategies. Indeed, the divestment/exit concept implies that investors are currently invested in the company and they decide to sell their participation, because they are not agreed with the ESG policy of the firm. In this context, divestment is the ultimate engagement action, meaning that investors believe that they can not influence the company's behavior by staying invested. In this case, divestment is "*the final step in an escalation strategy*" (PRI, 2022a) and may be viewed as a failure of the engagement action on the part of investors. When it concerns an industry, divestment can be motivated because there are high risks or poor opportunities that the industry transition to a more sustainable business model. For example, divesting from coal reserves or mining sector may reduce the exposure of investors to the risk of stranded assets[22]. As we have seen in the first chapter, exit can also be motivated by moral values. In this case, it coincide with the norms-based screening strategy.

---

[19]For instance, we can cite the following examples among others: As You Sow, Citizens' Climate Lobby, Climate Alliance, Friends of the Earth, Fund Our Future, Oxfam, Reclaim Finance and Sunrise Movement.

[20]Nevertheless, we notice that the two terms are often used interchangeably.

[21]These strategies are described on page 40.

[22]Stranded assets are studied in Section 10.2 on page 544.

Box 6.1: Case study: the Cambridge University endowment fund

In their research paper "To Divest or to Engage? A Case Study of Investor Responses to Climate Activism", Chambers *et al.* (2020) studied the interesting case study of the Cambridge University endowment fund:

> "*A dilemma faced by an increasing number of investors is whether to divest from environmentally damaging businesses or whether to enter into a dialogue with them. This predicament now has its epicentre in Cambridge, England, where the ancient University of Cambridge faces great pressure from students and staff to respond to the threat of climate breakdown. Having already received two reports on its approach to responsible investment, the university has appointed a new chief investment officer (CIO) who, alongside University Council and the wider university community, needs to consider the question of whether to divest from or to engage with fossil-fuel firms.*"

This paper tells the story of the Cambridge's fossil-fuel divestment movement from 2012 to 2020. This includes the creation of the Positive Investment Cambridge group, the establishment of the Ethical Investment Working Group, the publication of several reports, many Zero Carbon Society petitions in favour of fossil fuel divestment, the formation of the Divestment Working Group, the activism of students and faculty staff, the involvement of many respected scholars, including prestigious professors (e.g., astronomer Royal officer, Nobel laureate, chief scientific adviser, distinguished statisticians, Fields medallists, fellows of the Royal Academy and the Royal Society), etc. The research paper reveals how pressures upon institutional investors to respond to climate activism can accelerate. It concludes that "*investment professionals need to understand the forces for change*", "*a head-in-the-sand response is counter-productive*" and "*changes in investment policy should be evidence-based*".

The case of fossil fuel sector is certainly the most symbolic divestment theme. According to the Global Divestment Commitments Database, about 1 500 institutions[23] have committed to divest from fossil fuels as of October 2021. In Figure 6.6, we report the break-down by types of organization. The three most important categories are faith-based organizations, educational institutions and philanthropic foundations, representing 63.10% of institutions divesting. They are followed by pension funds. Commitments can be classified into four categories: (1) thermal coal only, (2) coal and tar sands, (3) partial divestment from some but not all types of fossil fuel companies, (4) full divestment from any fossil fuel company (thermal coal, oil, gas). Some investors are already "*fossil free*", meaning that they currently have no investment in fossil fuel companies and are committed to avoid any fossil fuel investments in the future. The debate on fossil fuel divestment is tough, because it is not only an investor debate, but also a matter of society. In Box 6.1, we report the tense relationship between the two sides. We notice that the choice to divest or not can be motivated by other considerations than rational evidence. The pressure of anti-fossil fuel movements may be one factor, the fear of greenwashing is another one. Moreover, "*exit relaxes the tension between the activist and the board*" (Levit, 2019). In some ways, exit is a confortable position. The investor can continue to communicate with the company, but the situation is completely different. Before exit,

---

[23]Including Axa Insurance, La Banque Postale, Harvard University, the State of Maine, the Norwegian Sovereign Wealth Fund, and the University of Oxford.

the investor has to convince the company to move. After exit, the company must prove that they will change and they must convince the investor to come back. Therefore, the roles are reversed.

Figure 6.6: What kinds of institutions are divesting from fossil fuel?
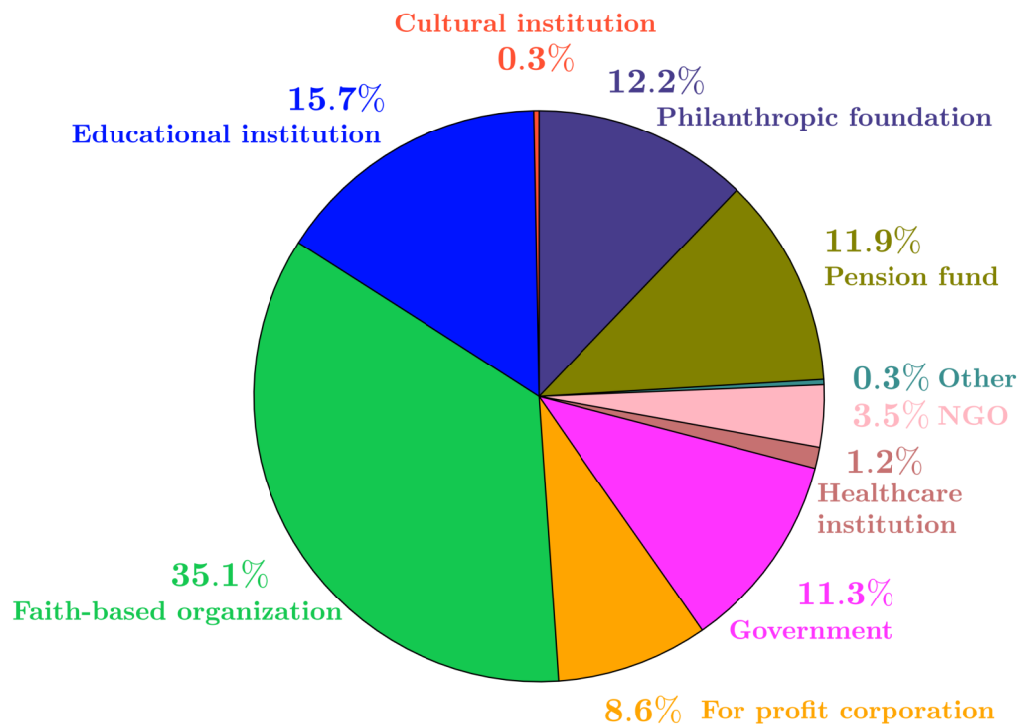
We have already seen that vote is a form of voice. For Admati and Pfleiderer (2009), exit is a form of voice too. Most of academic studies show that the impact of exit is mixed, because the cost of divestment may be high and its effectiveness is limited (Jahnke, 2019; Levit, 2019; Broccardo *et al.*, 2022; Edmans *et al.*, 2022). For some academics, the impact of divestment may also be negative:

> "*Large divestment campaigns are undertaken in part to depress share prices of firms that investors see as engaged in harmful activities. We show that, if successful, investors who divest earn lower and riskier returns than those that do not, leading them to control a decreasing share of wealth over time. Divestment therefore has only a temporary price impact. Further, we show that, for standard managerial compensation schemes, divestment campaigns actually provide an incentive for executives to increase, not reduce, the harm that they create. Therefore, divestment is both counter-productive in the short run, and self-defeating in the long run.*"(Davies and Van Wesep, 2018, page 558).

In fact, we must distinguish the impact on primary and secondary markets. From a theoretical viewpoint, divestment decreases the price of the company share price on the secondary market, implying that the returns increase. On the primary market, this induces a higher cost of capital because of the lower demand. Therefore, the rationale for exit is the effect on the cost of capital. Nevertheless, it is difficult to verify empirically this effect in the last 10 years. Academics are more consensual when they study the impact before the exit. Indeed, according to Edmans (2009), the threat of exit has more impact than divestment itself, and "*the power of loyalty relies on the threat of exit*". Again, we see that voice and exit are complementary instead of being mutually exclusive.

Box 6.2: Case studies of fossil fuel divestment

Church of England Pensions Board

By the end of 2021, the CoE Pensions Board was responsible for almost £3.7 bn of assets across three pension schemes. In July 2018, the General Synod of the Church of England voted on a motion to ensure that by 2023 the CoE pension funds have disinvested from fossil fuel companies that are not prepared to align with the goals of the Paris Agreement. In 2020, they engaged with 21 companies. At the end of the process, 12 companies were supposed to make sufficient progress, while 9 companies were added to the list of restricted investments. These divestments totalled £32.23 mn.

Source: Church of England Pensions Board, Stewardship Report 2020.

The Universities Superannuation Scheme (USS)

The Universities Superannuation Scheme (USS) is the pension scheme for university staff in the UK. They manage about £90 bn. In 2020, USS undertook a review of sectors in which the scheme invests. They concluded that, in several cases, the outcomes predicted by the market did not appropriately consider the potential financial impact of certain specific risks, including ESG. As a result, they excluded certain sectors: tobacco manufacturing; thermal coal mining (coal to be burned for electricity generation), specifically where they made up more than 25% of revenues, and certain controversial weapons. The first exclusion was announced in May 2020. Two years after, divestment from these sectors is completed. According to Ethics for USS (a group of USS members committed to reforming USS and ensuring an investment strategy that protects the planet, respects human rights, invests responsibly and ensures good pensions), "*USS still has large investments in the industries responsible for the climate emergency [...], while they recognise that USS has made plans to decarbonise its investment portfolio*". Ethics for USS estimated that "*USS continue to invest £570 mn in 48 leading fossil fuels companies*", and USS should extend its divestment policy "*to include other companies that have not committed to a credible path towards zero emissions*".

Source: USS, Stewardship Code Report 2022 & https://divestuss.org.

6.1.3 Individual versus collaborative engagement

Academic references are Dimson *et al.* (2021).

6.1.4 The role of institutional investors

Academic references are Appel *et al.* (2016), Krueger *et al.* (2020), Chen *et al.* (2020) and Gillan and Starks (2000).

6.1.5 Impact of active ownership

Academic references are Dimson *et al.* (2015), Grewal *et al.* (2016), and Broccardo *et al.* (2022).

6.2   ESG voting

6.2.1   Voting process

Each country has its own voting process. Nevertheless, they share some common ground. According to NBIM (2020), they typically include the following steps:

- *"The company sets the agenda for the annual shareholder meeting;*

- *The custodian confirms the identity of the shareholders and the number of shares eligible for voting — often for a specific date ahead of the meeting (record date);*

- *Shareholders receive the meeting materials from the company (may be before or after the record date);*

- *Shareholders procuring proxy advisory services receive voting recommendations;*

- *Shareholders instruct the custodian on how to vote, often through a proxy voting service provider, within a deadline ahead of the shareholder meeting (cut-off date);*

- *Voting takes place at the shareholder meeting;*

- *Shareholders receive confirmation from the service provider that their voting instructions have been carried out."*

6.2.2   Proxy voting

Etiam vel ipsum. Morbi facilisis vestibulum nisl. Praesent cursus laoreet felis. Integer adipiscing pretium orci. Nulla facilisi. Quisque posuere bibendum purus. Nulla quam mauris, cursus eget, convallis ac, molestie non, enim. Aliquam congue. Quisque sagittis nonummy sapien. Proin molestie sem vitae urna. Maecenas lorem. Vivamus viverra consequat enim.

6.2.3   Defining a voting policy

Integer placerat. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Sed in massa. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Phasellus tempus aliquam risus. Aliquam rutrum purus at metus. Donec posuere odio at erat. Nam non nibh. Phasellus ligula. Quisque venenatis lectus in augue. Sed vestibulum dapibus neque.

6.2.4 Statistics about ESG voting

In this section, we analyze the behavior of investors when they vote ESG resolutions. In the case of asset managers, we use the annual reports published by ShareAction. For asset owners, we use various surveys and complete the assessment by studying engagement reports of some pension funds.

Asset managers

ShareAction started the Voting Matters series in 2019. At that time, the report reviewed how largest asset managers have voted on shareholder resolutions linked to climate change (ShareAction, 2019). In 2020 and 2021, they extended the analysis by including the (**S**) pillar in addition to the (**E**) pillar (ShareAction, 2020, 2021). Finally, the 2022 edition analyzed shareholder-filed governance resolutions that directly relate executive compensation and political spending policies to environmental and social issues (ShareAction, 2023). After having merged the four datasets and clean the data[24], we obtain a database with 84 unique asset managers with the following frequencies: 35 in 2019, 54 in 2020, 65 in 2021 and 68 in 2022. There is only 26 asset managers[25], which are present every year. The scores calculated by ShareAction were based on 65 selected shareholder proposals in 2019, 102 in 2020, 146 in 2021 and 252 in 2022. In Table 6.1, we report some statistics of the selected shareholder resolutions. Let s_j be the support rate of resolution j. The resolution has the majority support if $s_j \geq 50\%$. For instance, among the 64 shareholder proposals in 2019[26], only three have obtained the majority, implying a success rate of 4.7%. During the period 2019–2022, this success rate is around 15%. We have also reported the 10%, 25%, 75% and 90% percentiles[27]. The interquartile range is between 12% and 43%. Finally, we notice that the average support rate is greater for environmental resolutions than for social resolutions.

Table 6.1: Statistics of success rate shareholder resolutions

Year		2019	2020	2021	2022
Number of resolutions		64	102	144	249
Resolutions with majority support		3	15	29	37
Success rate (in %)		4.7	14.7	20.1	14.9
Average support rate (in %)		28.2	29.9	32.9	29.9
	10%	6.5	9.2	7.2	9.4
Percentile of	25%	17.0	13.1	12.0	13.5
support rate (in %)	75%	37.7	42.6	42.8	40.3
	90%	41.8	55.2	81.2	57.6
	(**E**)	28.2	35.8	41.8	31.6
Average support rate (in %)	(**S**)		24.5	28.8	27.4

Source: ShareAction (2019, 2020, 2021, 2023) & Author's calculations.

For each asset manager, the support rate is calculated as:

$$\text{support rate} = \frac{\#\{\text{for}\}}{\#\{\text{for} + \text{against} + \text{abstention} + \text{dit-not-vote} + \text{split-vote}\}}$$

[24]In particular, some asset managers have merged or changed their name.
[25]They are APG AM, AXA IM, Abrdn, Allianz GI, Amundi AM, BNP PAM, BlackRock, Capital Group, DWS, Fidelity Investments, Generali Insurance AM, Goldman Sachs AM, HSBC Global AM, Invesco, J.P. Morgan AM, Legal & General, M&G IM, Ninety One, Northern Trust AM, Nuveen AM, SSGA, Schroders, T. Rowe Price, UBS AM, Vanguard, and Wellington Management.
[26]Some resolutions are excluded from the analysis because we don't have the figures.
[27]See Figure 6.7 for the empirical histogram.

Figure 6.7: Histogram (in %) of support rates
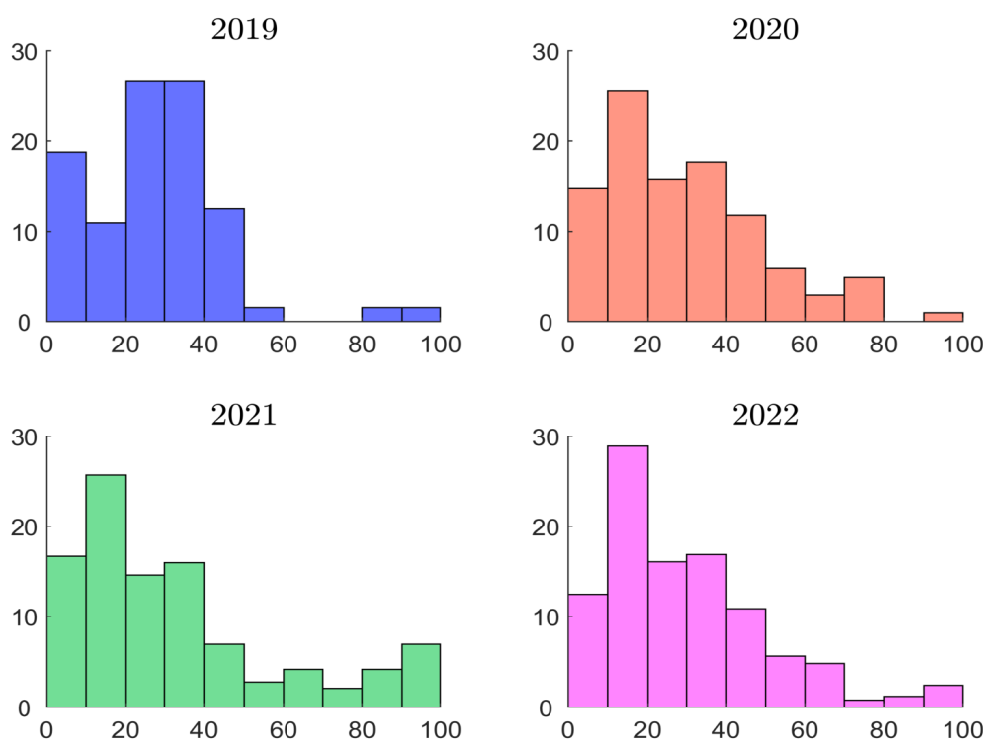
Source: ShareAction (2019, 2020, 2021, 2023) & Author's calculations.

In Table 6.2, we compute the average support rate of shareholder resolutions with two methods. The arithmetic mean is the simple average with respect to all asset managers, while the contribution of each asset manager is proportional to its assets under management for the weighted mean. We observe that the arithmetic mean of the support rates increases continuously since 2019 if we consider the overall score and the (S) pillar. This is not true for the (E) pillar for which we observe a lower value in 2022 compared to 2021. Moreover, the figure of 65% that we found in 2022 for the overall score is due to the introduction of the pay & politics topic. If we focus on the weighted mean, the figures are lower. For instance, the average support rate in 2022 is equal to 46.5% instead of 65%. This means that largest asset managers are voting for shareholder resolutions less than the others.

Table 6.2: Average support rate in % for ESG resolutions

Topic	Method	2019	2020	2021	2022
Overall	Arithmetic	45.8	57.4	58.9	65.0
	Weighted	32.7	42.1	47.6	46.5
Environment	Arithmetic	45.8	61.0	66.0	64.8
	Weighted	32.7	44.7	55.8	48.8
Social	Arithmetic		53.3	55.2	62.7
	Weighted		39.0	43.7	44.3
Pay & politics	Arithmetic				71.5
	Weighted				47.8

Source: ShareAction (2019, 2020, 2021, 2023) & Author's calculations.

Figure 6.8: Arithmetic average support rate in % per country and year
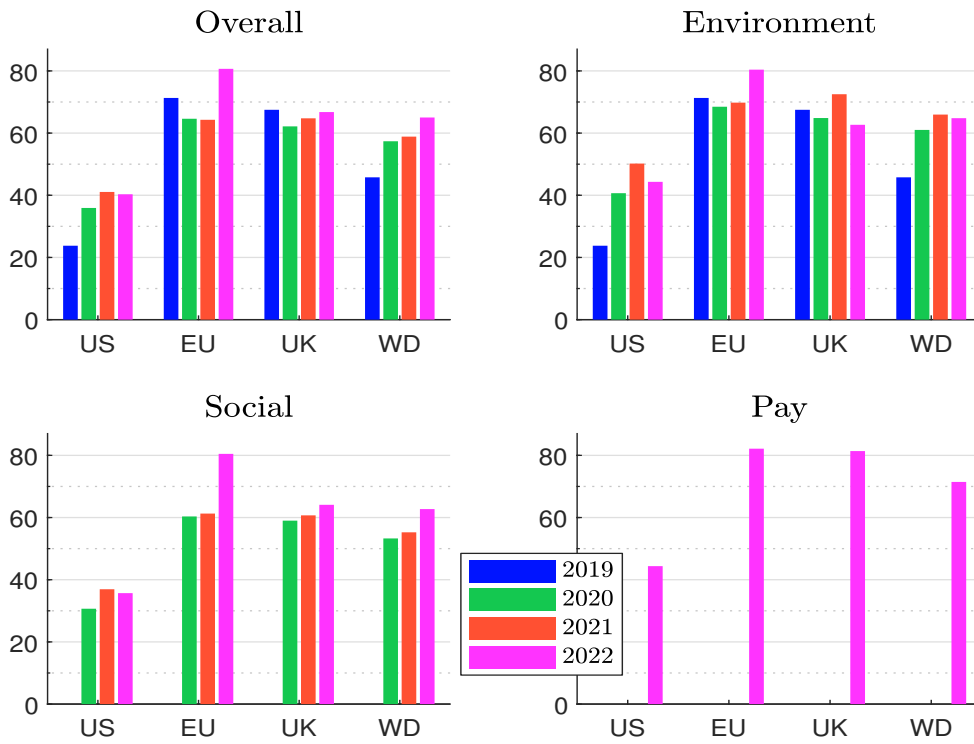
Source: ShareAction (2019, 2020, 2021, 2023) & Author's calculations.

Figure 6.9: Weighted average support rate in % per country and year
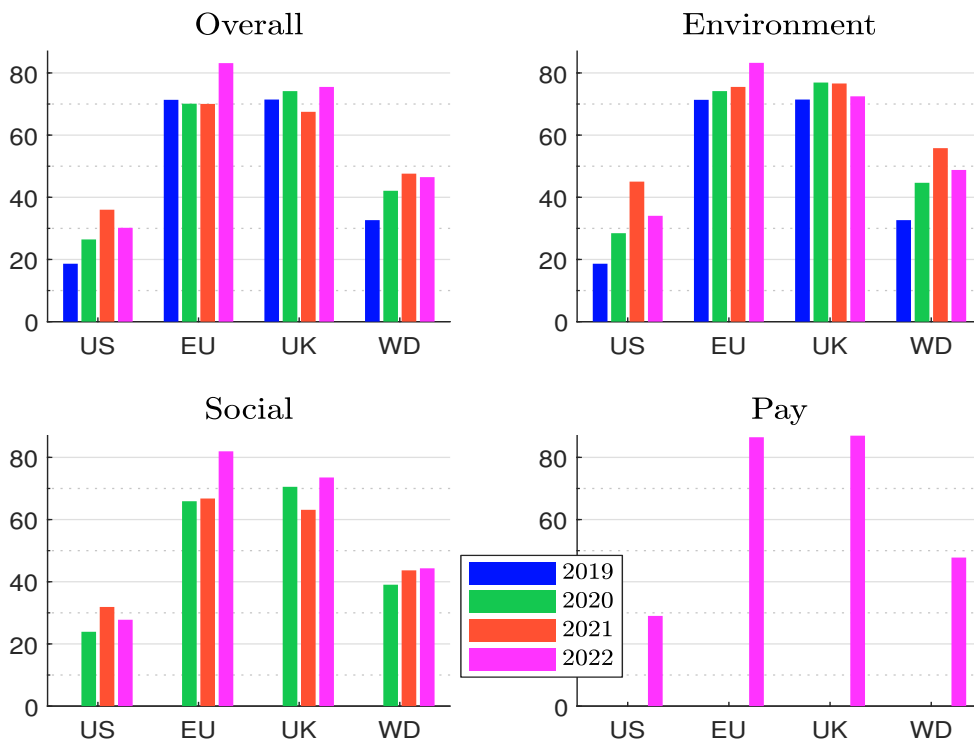
Source: ShareAction (2019, 2020, 2021, 2023) & Author's calculations.

If we perform a region analysis (Figures 6.8 and 6.9), we observe the following facts. British and European[28] asset managers have a similar voting behavior. Their support rate is significantly greater than this found in the US. Nevertheless, American asset managers have improved their ESG voting policy in 2020 and 2021, which explains the increase of the overall score for the world. In 2022, the increase of the European support was offset by the American setback.

Table 6.3: Best performers (2022, overall)

Rank	Name	Country	AUM	Overall	Ⓔ	Ⓢ	Pay
1	Achmea IM	Netherlands	251	100	100	100	100
1	Impax AM	UK	56	100	100	100	100
3	BNP PAM	France	761	99	97	100	100
3	MN	Netherlands	193	99	97	100	100
5	Candriam	Luxembourg	180	98	97	99	100
6	PGGM	Netherlands	331	97	93	100	97
7	Man	UK	149	96	98	94	98
8	Robeco	Netherlands	228	95	94	94	100
9	Aviva Investors	UK	363	93	88	96	100
10	Amundi AM	France	2 348	93	93	92	98
11	Nordea AM	Finland	333	91	93	89	90
12	Aegon AM	Netherlands	466	90	85	94	90
13	Federated Hermes	UK	672	89	88	87	90
14	Pictet AM	Switzerland	284	88	85	90	91
15	Legal & General	Switzerland	1 923	86	84	84	98

Source: ShareAction (2023) & Author's calculations.

In Tables 6.3 and 6.4, we report the ranking of the best fifteen and worst ten asset mangers when we consider the 2022 overall score. For each row, we indicate the rank, the name, the country, the assets under management (in $ bn) and the 2022 scores expressed in percentage. Two asset managers obtain a score of 100%: Achmea Investment Management and Impax Asset Management Group. The Top 15 ranking is dominated by asset managers located in the Netherlands, the UK and France, whereas there are seven American asset managers in the Bottom 10.

Table 6.4: Worst performers (2022, overall)

Rank	Name	Country	AUM	Overall	Ⓔ	Ⓢ	Pay
59	Goldman Sachs AM	US	2 218	35	56	24	24
60	Baillie Gifford	UK	455	31	29	29	45
61	SSGA	US	4 140	29	30	31	22
62	BlackRock	US	10 014	24	28	24	15
63	T. Rowe Price	US	1 642	17	26	11	18
64	Fidelity Investments	US	4 520	17	23	19	2
65	Vanguard	US	8 274	10	12	9	9
66	Dimensional Fund Advisors	US	679	4	6	5	0
67	Santander AM	Spain	220	4	0	5	6
68	Walter Scott & Partners	UK	95	3	0	6	0

Source: ShareAction (2023) & Author's calculations.

---

[28]This includes the following countries: Finland, France, Germany, Italy, Luxembourg, Netherlands, Spain, Sweden and Switzerland.

Table 6.5: Ranking of the 25 largest asset managers (2022, overall)

| Rank | Name | Country | AUM | Overall | | | | Environment | | | Social | | | Pay |
				2019	2020	2021	2022	2020	2021	2022	2020	2021	2022	2022
22	BlackRock	US	10014	7	12	40	24	11	53	28	12	34	24	15
25	Vanguard	US	8274	8	14	26	10	15	38	12	12	20	9	9
23	Fidelity Investments	US	4520	9	31	29	17	20	23	23	44	33	19	2
21	SSGA	US	4140	26	35	32	29	40	42	30	29	27	31	22
18	J.P. Morgan AM	US	2742	7	43	37	37	51	50	43	34	31	25	53
16	Capital Group	US	2716	5	8	28	45	12	26	37	4	31	47	58
2	Amundi AM	France	2348	66	89	93	93	91	97	93	88	90	92	98
20	Goldman Sachs AM	US	2218	37	45	47	35	48	57	56	43	40	24	24
3	Legal & General	UK	1923	82	96	77	86	96	87	84	95	73	84	98
24	T. Rowe Price	US	1642	5	22	31	17	27	44	26	17	25	11	18
15	Invesco	US	1611	34	37	37	47	52	51	54	19	28	37	61
12	Morgan Stanley IM	US	1566			55	64		59	64		53	65	63
14	Wellington Management	US	1426	10	51	44	48	62	60	63	39	37	41	36
7	Northern Trust AM	US	1348	21	70	60	83	79	68	83	59	57	78	96
13	Nuveen AM	US	1271	62	63	56	59	71	76	57	56	48	52	79
8	UBS AM	Switzerland	1216	90	79	75	83	91	72	84	67	75	80	87
4	DWS	Germany	1055	74	66	85	86	66	92	86	65	80	80	100
10	AXA IM	France	1009	79	71	55	73	85	72	69	55	45	68	95
6	Schroders	UK	991	56	62	73	85	63	78	81	60	70	87	87
17	AllianceBernstein	US	779				43			40			39	60
5	Allianz GI	Germany	766	89	81	77	86	89	81	80	73	76	91	85
1	BNP PAM	France	761	48	72	98	99	65	96	97	80	99	100	100
19	Columbia Threadneedle	US	754				37			43			25	53
9	Manulife IM	Canada	723				75			77			68	89
11	APG AM	Netherlands	721	72	70	59	72	80	65	89	59	56	57	83

Source: ShareAction (2019, 2020, 2021, 2023) & Author's calculations.

If we focus on the largest asset managers, we obtain the 2022 ranking reported in Table 6.5. For each row, we indicate the rank according to the 2022 overall score, the name, the country, the assets under management (in $ bn) as of December 2022, the overall scores expressed in percentage for the four reporting years (2019 to 2022). We also report the thematic scores (environmental, social, and pay & politics). In this case, the six best performers are BNP Paribas Asset Management, Amundi, Legal & General, DWS, Allianz GI and Schroders (respectively two British, French and German asset managers). We also observe that some asset managers have made large improvement. For instance, the score of BNP PAM improves from 48% in 2019 to 99% in 2022. For Northern Trust AM, we have 21% in 2019 vs. 83% in 2022. If we consider the four largest US asset managers, they have improved their score between 2019 and 2021, but they backed fewer resolutions in 2022 than they did in 2021. This is certainly explained by the political pressure and the anti-ESG movement in the US[29]. If we extend the analysis to 84 asset managers of the database, we obtain the dynamics given in Figure 6.10. Overall, trends are upward even if we notice some setbacks, especially in the US.

Figure 6.10: Evolution of the support rate in % per asset manager
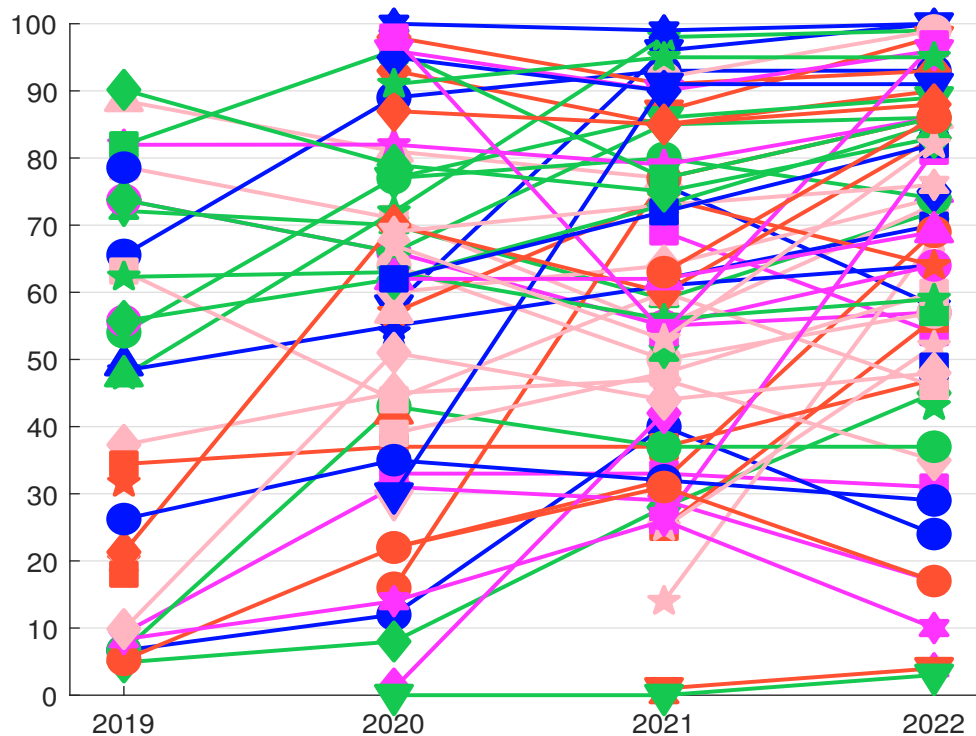
Source: ShareAction (2019, 2020, 2021, 2023) & Author's calculations.

According to ShareAction (2023), the main findings are the following:

1. "*49 additional resolutions would have received majority support if the largest asset managers had voted in favour of them.*

2. *Voting performance has been stagnant in the US and the UK compared to 2021, while European asset managers have shown a large improvement.*

---

[29]See Footnote 62 on page 29

3. *Asset managers across the board are hesitant to back action-oriented resolutions, which would have the most transformative impact on environmental and social issues.*"

Other interesting results can be found in the four reports. For instance, we learn that only six asset managers filed or co-filed a shareholder resolution at any companies assessed in 2021 (ShareAction, 2021). In fact, it seems that most shareholder resolutions are filed by civil society organizations, small impact-focused asset managers, local governmental pension funds, and charitable or faith-based investors. For a comprehensive list of resolutions, the reader may consult the Insightia database (`www.insightia.com`), the Climate Action 100+ flagged shareholder votes (`www.climateaction100.org`), the Say on Climate resolutions (`www.sayonclimate.org`) or the proxy resolutions & voting guide of Interfaith Center on Corporate Responsibility (`www.iccr.org`).

Remark 49 *In addition to the previous analysis, the 2022 ShareAction report includes statistics on say on climate resolutions. In this case, shareholders express approval or disapproval of the company's global climate strategy. Among the 36 selected resolutions, only six were filed by share-holders, requesting the company to adopt an annual advisory vote on the company's climate plan. The remaining thirty resolutions were management-sponsored standing votes requesting shareholders to approve the company's climate plan. The shareholder- and management-sponsored votes received respectively* 21.98% *and* 89.17% *support on average[30]. In Figure 6.11, we report the distribution of the support rate. We clearly see a big difference between shareholder- and management-sponsored votes.*

Figure 6.11: Ranking of the 36 say on climate resolutions with respect to the support rate in %

---
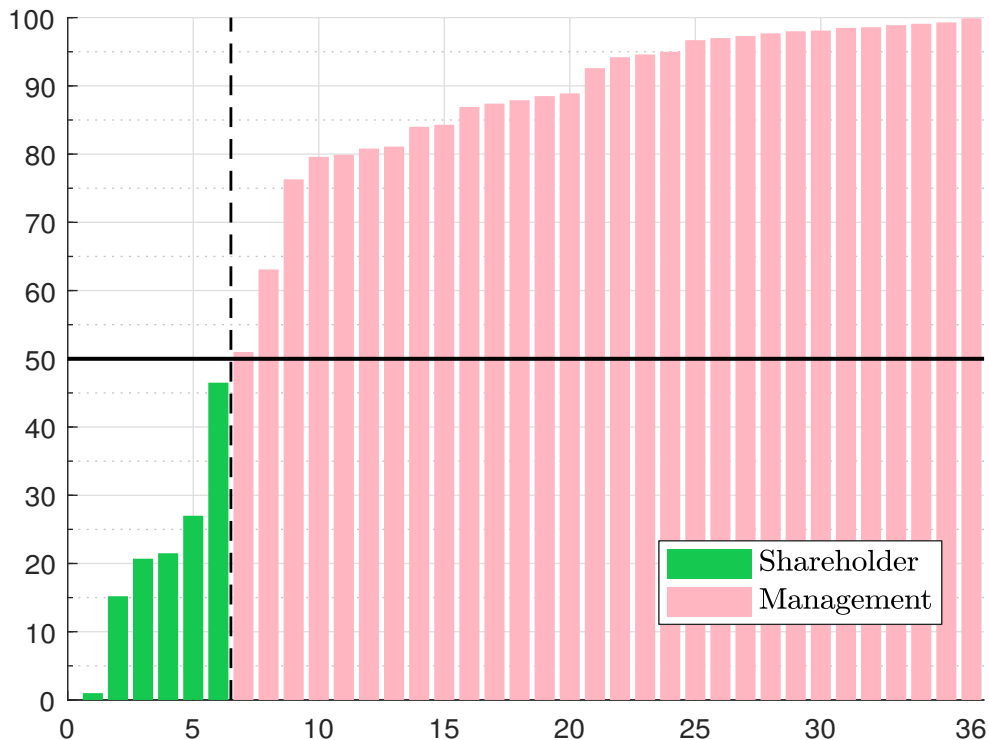
[30]The range is between 1% and 46.5% for shareholder-sponsored resolutions, whereas it is between 51% and 99.9% for shareholder-sponsored resolutions.

Box 6.3: Case studies: Barclays, EDF and Woodside Energy say on climate resolutions

Electricité de France or EDF (French energy company) filed a management-sponsored say on climate resolution at its AGM on 12 May. The group's climate transition plan to achieve carbon neutrality by 2050 was approved by 99.1%.

The management of Barclays (British bank) filed a say on climate resolution at the 2022 AGM on 4 May. They asked to approve the climate strategy, targets and progress of Barclays. Despite the plan's insufficiencies regarding coal, the result was 80.8% for and 19.2% against. In certain special situations, some asset managers prefer to vote for with the hope that the company will improve its plan and return with an improved say on climate resolution next year. It seems that it was the case for Barclays according to some NGOs.

Woodside Energy Group Ltd. (Australian energy company) filed a say on climate resolution at its 2022 AGM on 19 May. The management asked to approve their climate report. The support rate was 51.03%, meaning that 48.97% of the company's shareholders voted against its climate transition plan. These disappointed results may force the management to propose a new resolution in 2023 with a high risk of failure. Therefore, these results may also stop any new say on climate resolution filed by the management for the coming yeas.

Asset owners

Chapter 7

Extra-financial Accounting

7.1 Historical perspectives

7.2 Single vs. double materiality

7.3 Environmental accounting

7.3.1 National environmental accounts

7.3.2 Corporate environmental accounts

7.4 Sustainability accounting

7.4.1 Social issues

7.4.2 Governance factors

Part II

Climate Risk

Chapter 8

The Physics and Economics of Climate Change

8.1 Awareness of climate change impacts

In this section, we review some important dates in climate science, from the discovery of the greenhouse effect in the early nineteenth century to the evidence of global warming in the late eighties. We will then examine natural and anthropogenic climate change. In particular, we will ask: Are human activities responsible for global warming? Finally, we will present some physical climate models that help to understand climate dynamics and the balance of greenhouse gas emissions.

8.1.1 Scientific evidence of global warming

The greenhouse effect

The greenhouse effect is the natural warming of the Earth that occurs when gases in the atmosphere trap heat from the sun that would otherwise escape into space (Figure 8.1). Without the greenhouse effect, the average temperature on Earth would be -18 degrees Celsius instead of the $+15$ degrees Celsius it is today.

Figure 8.1: Diagram showing how the greenhouse effect works
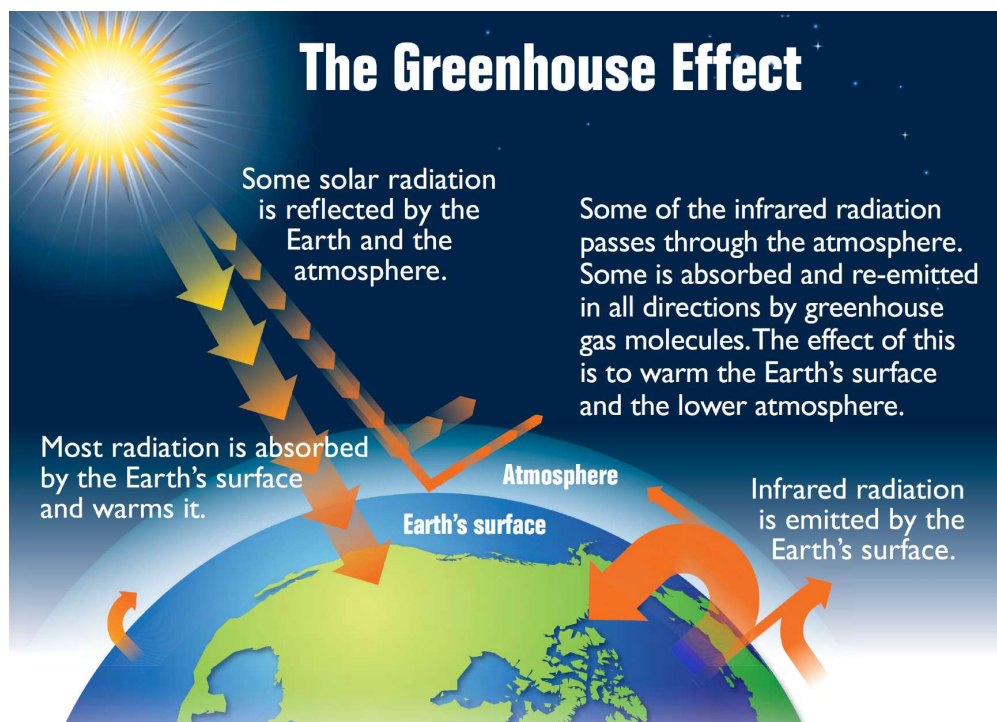
Source: US EPA (2012), https://energyeducation.ca/encyclopedia/Greenhouse_effect.

The founding text of the greenhouse effect is a French scientific publication written by Joseph Fourier in 1824: *Remarques générales sur les températures du globe terrestre et des espaces planétaires*[1]. Using some theoretical mathematical models he had developed earlier, Fourier (1824) calculated that the Earth should not be as warm as it is, given its distance from the sun and its size. He then defined the various sources of heat on the Earth's surface:

> "*The heat of the Earth's surface derives from three sources, which must first be distinguished:*

---

[1] *General Remarks on Global and Planetary Temperatures.*

1. *The Earth is heated by the sun's rays, whose uneven distribution produces the diversity of climates.*

2. *The Earth's temperature depends on the common temperature of planetary spaces, as it is exposed to the irradiation of the innumerable stars that surround the solar system on all sides.*

3. *The earth has retained within its mass some of the primitive heat it contained when the planets were formed.*"

He showed that the third effect is not significant and that the heat on the Earth's surface comes mainly from the sun (first factor). He distinguished luminous heat[2] ("*chaleur lumineuse*") and dark radiant heat[3] ("*chaleur rayonnante obscure*"), and explained that the difference between the two types of heat is due to the presence of the atmosphere. Joseph Fourier also referred to the work of the Swiss physicist and geologist Horace Bénédict de Saussure. In 1774, Saussure found that exposing a dark box with a pane of glass to the sun increased the temperature inside the box by several degrees compared to the temperature outside in the shade. He named this phenomenon "*l'effet de serre*" (the greenhouse effect in English). For Joseph Fourier, the interposition of the atmosphere is responsible for this greenhouse effect and is the factor that keeps the planet warmer because "*heat in the state of light meets with less resistance in penetrating the air than in returning to the air when it is converted into non-luminous heat*".

Pouillet (1838) developed a pyrheliometer to measure the solar irradiance at the surface as a function of the atmospheric air mass. He found that the solar constant S_0 is equal to $1\,230\ Wm^{-2}$, while the current estimate is about $1\,368\ Wm^{-2}$. He also estimated the temperature in space to be $-142°C$. In fact, the temperature in space approaches absolute zero (or $-273°C$). According to NASA[4], it can reach $120°C$ during the lunar day at the moon's equator and drop to $-130°C$ at night. At the moon's poles, temperatures can drop even lower, reaching $-250°C$. Like Joseph Fourier, Claude Pouillet explained that the greater absorption of the sun's rays and the relatively small variations in temperature were mainly due to the atmospheric stratum.

John Tyndall's work marks a new milestone. As an experienced mountaineer with a passion for glacier formation and melting, he studied climate-induced changes in the ice caps. In 1859, Tyndall showed that water vapor has a high heat absorption capacity. He went on to show that carbon dioxide and other gases could also absorb and radiate heat. Tyndall was not only interested in theory and equations, he also carried out many experiments to calculate the heat capacity of many gases using a measuring device based on thermopile technology. He realized the implications of his findings for climate change[5]:

> "*De Saussure, Fourier, Pouillet, and Hopkins regard this interception of the terrestrial rays as exercising the most important influence on climate. Now if, as the above experiments indicate, the chief influence be exercised by the aqueous vapour, every variation of this constituent must produce a change of climate. Similar remarks would apply to the carbonic acid diffused through the air; while an almost inappreciable admixture of any of the hydrocarbon vapours would produce great effects on the terrestrial rays and produce corresponding changes of climate. It is not therefore necessary to assume alterations in the density and height of the atmosphere, to account for different amounts of heat being preserved to the earth at different times; a slight change in its variable*

---

[2]This corresponds to the ultra-violet light.

[3]This corresponds to the thermal infrared radiation.

[4]See https://science.nasa.gov/mission/lro.

[5]Carbonic acid was the historical name given to CO_2 in the nineteenth century.

constituents would suffice for this. Such changes in fact may have produced all the mu-
tations of climate which the researches of geologists reveal. However this may be, the
facts above cited remain; they constitute true causes, the extent alone of the operation
remaining doubtful." (Tyndall, 1861, pages 28-29).

Tyndall is the first scientist to prove that greenhouse gases exist and are responsible for the green-
house effect.

In fact, Tyndall was not the first to discover the existence of greenhouse gases. Three years earlier,
in 1856 and 1857, the American scientist Eunice Newton Foote had published two research papers
with experiments showing that water vapor and carbon dioxide absorb heat from solar radiation.
However, her work was forgotten until it was rediscovered by Raymond Sorenson in 2011. Foote's
work led to debates about the paternity of greenhouse gases (Ortiz and Jackson, 2022).

Figure 8.2: The pioneers of the greenhouse effect

| Joseph Fourier | Eunice Newton Foote | John Tyndall | Svante Arrhenius |
| (1766-1830) | (1819-1888) | (1820-1893) | (1859-1927) |
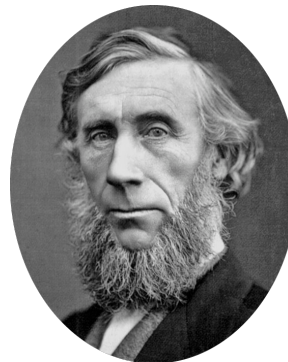
Svante Arrhenius was the first scientist to calculate the effect of a change in atmospheric CO_2
on ground temperature. Svante Arrhenius (1859-1927) was a renowned Swedish chemist who won
the Nobel Prize[6] in Chemistry in 1903 for his ionic dissociation theory. In 1896, he argued that
variations in atmospheric carbon dioxide could significantly affect the Earth's heat balance. Making
extensive use of the empirical works of Knut Ångström, Alexander Buchan, F. Paschen, Charles
Pouillet, and especially the data collected by Samuel Pierpont Langley, he came to the following
conclusion:

"*We may now inquire how great the variation of the carbonic acid in the atmosphere be*
to cause a given change of the temperature. [...] Thus if the quantity of carbonic acid
increases in geometric progression, the augmentation of the temperature will increase
nearly in arithmetic progression." (Arrhenius, 1896, Section IV, pages 265–267).

This rule has often been interpreted to mean that a doubling of the CO_2 concentration leads to an
increase in the average global temperature of about 5°C. This is the first time that an estimate of

---

[6]Arrhenius was involved in the establishment of the Nobel Institute from 1901, when he was elected a member of
the Royal Swedish Academy of Sciences. From that year he was a member of the Nobel Committee for Physics and
also for Chemistry. When the Nobel Institute for Physical Research was established in Stockholm in 1905, he was
appointed rector of the Institute, a position he held until 1927. Svante Arrhenius is credited with many stories and
intrigues about the awarding or not of the Nobel Prize, including the conflicts he had with Paul Ehrlich (theory of
immunology) and Dmitri Mendeleev (periodic table of elements).

climate sensitivity has been calculated[7]. In the final section, devoted to the geological consequences, his calculations showed that:

> "*One may now ask, How much the carbonic acid vary according to our figures, in order that the temperature should attain the same values as in the Tertiary and Ice ages respectively? A simple calculation shows that the temperature of the Arctic regions would rise about 8°C to 9°C, if the carbonic acid increased 2.5 or 3 times its present value. In order to get the temperature of the ice age between the 40th and 50th parallels, the carbonic acid in the air should sink to 0.62–0.55 of present value (lowering the temperature 4°C–5°C).*" (Arrhenius, 1896, Section V, page 268).

The last part of the article is remarkable. He discussed the work of his friend and colleague Arvid Hogborm (1857-1940) on the influence of natural carbon dioxide cycles on the Earth's ice ages. In particular, he argued that these cycles were a much better explanation of climate change on a geological scale than the theory of James Croll (1821-1890), who assumed that it was due to the variations in the Earth's orbit[8]. He also discussed the impact of human activity on the concentration of CO_2 in the atmosphere, and in particular the impact of coal mining and production.

While Arrhenius studied only the heat absorption capacity of CO_2, Tyndall calculated the heat power of several gases: water vapor, carbon dioxide, oxygen, nitrogen, hydrogen, nitrous oxide, etc. However, Tyndall assumed that water vapor was the main contributor to the greenhouse effect and that the effect of carbon dioxide was negligible. We now know that this is not true. Since the 1970s, many scientific studies have been carried out to investigate the other greenhouse gases and their effects on the environment and global warming. Ozone has received much attention (World Meteorological Organization, 1982), followed by chlorofluorocarbons or CFCs (Ramanathan, 1975), nitrates and sulfates. The survey of Ramanathan *et al.* (1985) summarizes the state of scientific knowledge in the 1980s. It was at this point that a number of scientists began to reintroduce methane as a growing threat to climate change. For example, Kvenvolden (1988) analyzed the process of permafrost warming and methane release from gas hydrates and possible destabilization in the Arctic.

Table 8.1: List of greenhouse gases

Greenhouse gas	Formula	Kyoto Protocol
Water vapor	H_2O	
Carbon dioxide	CO_2	✓
Methane	CH_4	✓
Nitrous oxide	N_2O	✓
Ozone	O_3	
Fluorinated or F-gases		
Sulfur hexafluoride	SF_6	✓
Nitrogen trifluoride	NF_3	✓
Chlorofluorocarbons	CFCs (CFC-11, CFC-12, etc.)	
Hydrofluorocarbons	HFCs (HFC-23, HFC-32, etc.)	✓
Hydrochlorofluorocarbons	HCFCs (HCFC-12, etc.)	
Perfluorocarbons	PFCs (CF_4, C_2F_6, etc.)	✓

---

[7]Climate sensitivity measures the global temperature increase that will occur in response to a doubling of atmospheric CO_2 concentrations compared to pre-industrial levels.

[8]Today it is accepted that the two theories co-exist and explain the phenomenon of global glaciation (Hays *et al.*, 1976; Delmas *et al.*, 1980).

Global warming

The discovery of the greenhouse effect implies that the temperature of the Earth depends on the concentration of greenhouse gases in the atmosphere. It does not mean that we are observing a global warming of the Earth, and even less that this global warming is due to human activities. For that, we need to show that the concentration of greenhouse gases has recently increased due to human influence. We also need to calculate the net effect of each gas. For example, the loss of ozone that we have observed in the lower stratosphere has a cooling effect on the Earth's surface, while the increase in ozone that has occurred in the troposphere has a warming effect on the Earth's surface.

Guy Stewart Callendar was a British steam engineer. In 1938, he linked the increased burning of fossil fuels to rising global temperatures. He compiled weather data and estimated a global temperature increase of about $0.25°C$ over the last fifty years. He also showed that the concentration of CO_2 in the atmosphere had increased by 10% during the same period, and he argued that this rising carbon dioxide content of the atmosphere was due to the burning of fossil fuels. Thus was born the "*Callendar Effect*", which links global warming to man's artificial production of carbon dioxide:

> "*By fuel combustion man has added about 150 000 million tons of carbon dioxide to the air during the past half century. The author estimates from the best available data that approximately three quarters of this has remained in the atmosphere. The radiation absorption coefficients of carbon dioxide and water vapour are used to show the effect of carbon dioxide on sky radiation. From this the increase in mean temperature, due to the artificial production of carbon dioxide, is estimated to be at the rate of $0.003°C$ per year at the present time. The temperature observations at zoo meteorological stations are used to show that world temperatures have actually increased at an average rate of $0.005°C$ per year during the past half century.*" (Callendar, 1938, page 223).

Guy Stewart Callendar, however, thought that this was not a bad thing. Like Svante Arrhenius[9], he believed that higher temperatures would be beneficial to civilization:

> "*In conclusion it may be said that the combustion of fossil fuel [...] is likely to prove beneficial to mankind in several ways, besides the provision of heat and power. For instance the above mentioned small increases of mean temperature would be important at the northern margin of cultivation, and the growth of favourably situated plants is directly proportional to the carbon dioxide pressure [...] In any case the return of the deadly glaciers should be delayed indefinitely.*" (Callendar, 1938, page 236).

According to Archer and Pierrehumbert (2011) (Table 8.1), three major research papers were written in the fifties. First, from 1956, the Canadian physicist Gilbert Norman Plass published a series of papers on the absorption of infrared radiation. His work confirmed that more carbon dioxide

---

[9]Thirty years earlier, in his book *Worlds in the Making: The Evolution of the Universe*, Arrhenius argued that a rise in global temperature could be beneficial to agriculture, providing more food for the world's population:

> "*We often hear lamentations that the coal stored up in the earth is wasted by the present generation without any thought of the future, and we are terrified by the awful destruction of life and property which has followed the volcanic eruptions of our days. We may find a kind of consolation in the consideration that here, as in every other case, there is good mixed with the evil. By the influence of the increasing percentage of carbonic acid in the atmosphere, we may hope to enjoy ages with more equable and better climates, especially as regards the colder regions of the earth, ages when the earth will bring forth much more abundant crops than at present, for the benefit of rapidly propagating mankind.*" (Arrhenius, 1908, Chapter II, page 63).

Box 8.1: The collected papers on global warming by David Archer and Raymond Pierrehumbert (1800-1980)

- 1824
  On the Temperatures of the Terrestrial Sphere and Interplanetary Space (Fourier)

- 1861
  On the Absorption and Radiation of Heat by Gases and Vapours, and on the Physical Connexion of Radiation, Absorption, and Conduction (Tyndall)

- 1896
  On the Influence of Carbonic Acid in the Air upon the Temperature of the Ground (Arrhenius)

- 1938
  The Artificial Production of Carbon Dioxide and its Influence on Temperature (Callendar)

- 1956
  The Influence of the 15μ Carbon-dioxide Band on the Atmospheric Infra-red Cooling Rate (Plass)

- 1957
  Carbon Dioxide Exchange Between Atmosphere and Ocean and the Question of an Increase of Atmospheric CO_2 during the Past Decades (Revelle and Suess)

- 1958
  Distribution of Matter in the Sea and Atmosphere: Changes in the Carbon Dioxide Content of the Atmosphere and Sea due to Fossil Fuel Combustion (Bolin and Eriksson)

- 1960
  The Concentration and Isotopic Abundances of Carbon Dioxide in the Atmosphere (Keeling)

- 1967
  Thermal Equilibrium of the Atmosphere with a Given Distribution of Relative Humidity (Manabe and Wetherald)

- 1969
  The Effect of Solar Radiation Variations on the Climate of the Earth (Budyko)
  A Global Climatic Model Based on the Energy Balance of the Earth-Atmosphere System (Sellers)

- 1970
  Is Carbon Dioxide from Fossil Fuel Changing Man's Environment? (Keeling)

- 1972
  Man-Made Carbon Dioxide and the Greenhouse Effect (Sawyer)

- 1975
  The Effects of Doubling the CO_2 Concentration on the Climate of a General Circulation Model (Manabe and Wetherald)

- 1977
  Changes of Land Biota and Their Importance for the Carbon Cycle (Bolin)
  Neutralization of Fossil Fuel CO_2 by Marine Calcium Carbonate (Broecker and Takahashi)

- 1979
  Carbon Dioxide and Climate: A Scientific Assessment (Charney, Arakawa, Baker *et al.*)

Source: Archer and Pierrehumbert (2011).

Box 8.2: The collected papers on global warming by David Archer and Raymond Pierrehumbert (1980-2005)

- 1984
  Climate Sensitivity: Analysis of Feedback Mechanisms (Hansen, Lacis, Rind *et al.*)

- 1985
  Evidence From Polar Ice Cores for the Increase in Atmospheric CO_2 in the Past Two Centuries (Neftel, Moor, Oeschger and Stauffer)

- 1986
  Global Temperature Variations Between 1861 and 1984 (Jones, Wigley and Wright)

- 1987
  Vostok Ice Core Provides 160,000-Year Record of Atmospheric CO_2 (Barnola, Raynaud, Korotkevich and Lorius)

- 1990
  Observational Constraints on the Global Atmospheric CO_2 Budget (Tans, Fung and Takahashi)

- 1991
  Abrupt Deep-Sea Warming, Palaeoceanographic Changes and Benthic Extinctions at the End of the Palaeocene (Kennett and Stott)

- 1992
  Effects of Fuel and Forest Conservation on Future Levels of Atmospheric Carbon Dioxide (Walker and Kasting)

- 1995
  Climate Response to Increasing Levels of Greenhouse Gases and Sulphate Aerosols (Mitchell, Johns, Gregory and Tett)

- 1999
  Northern Hemisphere Temperatures During the Past Millennium: Inferences, Uncertainties, and Limitations (Mann, Bradley and Hughes)

- 2000
  Acceleration of Global Warming Due to Carbon-Cycle Feedbacks in a Coupled Climate Model (Cox, Betts, Jones, Spall and Totterdell)
  Reduced Calcification of Marine Plankton in Response to Increased Atmospheric CO_2 (Riebesell, Zondervan, Rost *et al.*)

- 2002
  Surface Melt-Induced Acceleration of Greenland Ice-sheet Flow (Zwally, Abdalati, Herring *et al.*)

- 2003
  Anthropogenic Carbon and Ocean pH (Caldeira and Wickett)

- 2004
  Contribution of Stratospheric Cooling to Satellite-Inferred Tropospheric Temperature Trends (Fu, Johanson, Warren and Seidel)

- 2005
  Earth's Energy Imbalance: Confirmation and Implications (Hansen, Nazarenko, Ruedy *et al.*)

Source: Archer and Pierrehumbert (2011).

would have a warming effect and showed that: (1) a doubling of CO_2 would warm the planet by 3.6°C; (2) halving the amount of CO_2 would reduce the surface temperature by 3.8°C; (3) human activities have increased the concentration of CO_2 in the atmosphere at a rate of 30% per century; (4) this increase implies a warming of 1.1°C per century (Plass, 1956, pages 152-153). Secondly, Revelle and Suess (1957) showed that the average lifetime of a CO_2 molecule in the atmosphere before it is dissolved in the ocean is of the order of 10 years. The resistance to atmospheric carbon dioxide being absorbed by the surface layer of the ocean is known as the "*buffer factor*" or "*Revelle factor*". They concluded that most of the CO_2 released by artificial fuel combustion since the beginning of the industrial revolution must have been absorbed by the oceans. However, they felt that the exponential increase in industrial fuel combustion could pose a problem in the future and dramatically increase atmospheric CO_2. The mechanism of the ocean's buffering capacity was explored in a third research paper by Bolin and Eriksson (1958), who developed a dynamic model of atmosphere-ocean interactions with respect to CO_2. They estimated that the upper ocean must have absorbed less than 10% of fossil fuel emissions and acted as a barrier to the transport of CO_2 to the deep ocean. They concluded that if industrial production continues to rise, atmospheric CO_2 is likely to increase by 25% by the end of the century. In addition to the impact of CO_2 on ocean acidification, they highlighted the impact of land use and biosphere:

> "*Because of the small buffering effect of the sea it seems likely that the biosphere on land may play a more important role for the changes actually occurring in the atmosphere due to the release of CO_2 by combustion than previously believed.*" (Bolin and Eriksson, 1958, page 130).

In Revelle and Suess (1957), the authors concluded that "*present data on the total amount of CO_2, in the atmosphere, on the rates and mechanisms of CO_2, exchange between the sea and the air and between the air and the soils, and on possible fluctuations in marine organic carbon, are insufficient to give an accurate base line for measurement of future changes in atmospheric CO_2.*" Several initiatives were launched in the 1950s, particularly in Scandinavia. However, the data collected were extremely noisy and the CO_2 measurements were not stable. Roger Revelle invited Charles David Keeling to join the Scripps Institution of Oceanography (SIO) in 1956. Keeling received his Ph.D. in chemistry from Northwestern University in 1953 and was a postdoctoral fellow in geochemistry at CalTech until 1956, where he developed a stable instrument capable of measuring carbon dioxide in atmospheric samples. He decided to install his equipment in Antarctica (Little America), California (La Jolla) and on the Mauna Loa volcano in Hawaii and started collecting carbon dioxide samples in 1958. In 1960, he published a first report and tabulated the monthly average concentrations of atmospheric CO_2 at the three stations. He concluded that there were strong seasonal variations in CO_2 levels in the northern hemisphere (La Jolla, California and Mauna Loa, Hawaii), and a small but persistent increase in concentration was found in Antarctica (Keeling, 1960, page 200). In 1961, he presented new data showing a steady increase in CO_2 levels and introduced the famous "*Keeling curve*", a measurement of atmospheric carbon dioxide concentration made at Mauna Loa since 1958 (Keeling, 1978). The data was collected by Keeling until his death in 2005, and then by his son. The current version of the Keeling curve is shown in Figure 8.3. In March 1958, the CO_2 concentration was equal to 315.71 ppm[10] or 611.8 mg/m³. The limit of 400 ppm was reached in February 2015. In June 2023, the measure was 423.39 ppm or 820.5 mg/m³.

In the 1960s, a small number of scientists developed simple mathematical models of the planet's climate system. After completing his PhD at the University of Tokyo in 1959, Syukuro Manabe

---

[10]One part per million (ppm) denotes one part per 10^6 parts. It is equal to 10^{-6}. So one milliliter of gas in 1 000 liters of air would be 1 ppm.

Figure 8.3: Keeling curve: monthly mean CO_2 concentration in Mauna Loa (1958-2023)
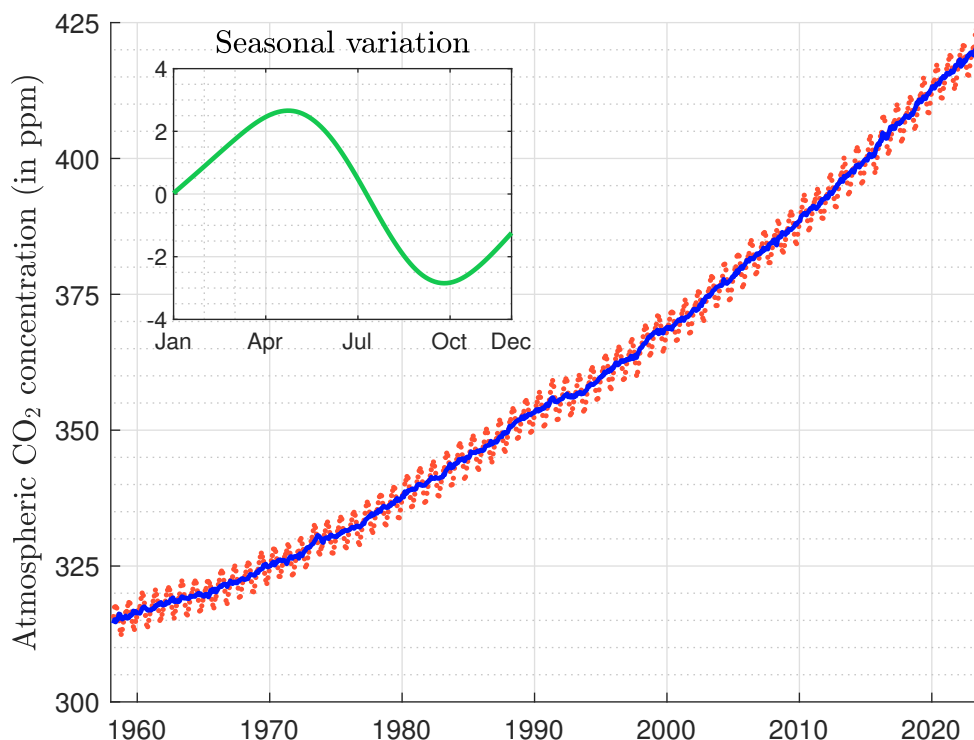
Source: Keeling *et al.* (2001) and https://scrippsco2.ucsd.edu.

went to the United States to work at the National Oceanographic and Atmospheric Administration (NOAA). Together with co-authors, Manabe published several research papers in 1961, 1964 and 1965 to develop a climate model. With the help of computer simulations, Manabe and Wetherald (1967) were able to obtain a first, but basic, model for predicting temperature by incorporating the hydrologic cycle (convective adjustment, feedback effect of water vapor, distribution of humidity, cloudiness). They found that as the concentration of carbon dioxide in the atmosphere increases, the temperature at the Earth's surface and in the troposphere rises, while it falls in the stratosphere. They also estimated the climate sensitivity to be between 1.3°C and 2.3°C. Manabe continued his research in the seventies and eighties. He and Kirk Bryan developed a general circulation model (GCM) and published the results in 1969. Manabe and Wetherald later used this original model to simulate the first three-dimensional experiment to test the idea of global warming. Their results were published in 1975.

In 2021, Syukuro Manabe and Klaus Hasselmann were awarded the Nobel Prize for their work on climate models. Klaus Hasselmann is a German oceanographer. He is known for his stochastic prediction model of climate change (Hasselmann, 1976; Frankignoul and Hasselmann, 1977). In fact, weather and climate forecasting has a long history, beginning with the book of Richardson (1922). However, it was the development of digital computers that accelerated the science of weather forecasting. Using the pioneering ENIAC computer and with the help of John von Neumann, Jules Charney and his team created a two-dimensional weather simulation model (Charney *et al.*, 1950), but the first general circulation model (GCM) is generally attributed to Phillips (1956). A GCM is a mathematical representation of the interactions between the components of the climate system (atmosphere, land surface, ocean and sea ice). In addition to the works of Manabe and Hasselmann,

we can also cite the research of Mikhail Budyko (1920-2001) and William Sellers (1928-2014), who introduced energy balance models[11] and warned of how positive and negative feedbacks (i.e., loops of mutually reinforcing effects) could amplify human impacts on the global climate. In particular, Budyko (1969) and Sellers (1969) developed two models of catastrophic ice-albedo feedbacks. Today, these feedback loops are fully integrated into GCMs, and dozens of teams around the world use sophisticated GCM models and supercomputers to predict weather and climate change[12]. Another milestone was the Charney Report, *Carbon Dioxide and Climate: A Scientific Assessment*, prepared for the American Academy of Sciences, whose conclusions were presented to President Carter in 1979. On July 23-27, 1979, Jule Charney formed a study group to "*assess the scientific basis for projection of possible future climatic changes resulting from man-made releases of carbon dioxide into the atmosphere*" (Charney *et al.*, 1979, page iv). The study group had 13 members, including eminent scientists Akio Arakawa, Bert Bolin, Henry Stommel, etc. The report examined the results of five global climate models that simulate the climatic response to an increase in atmospheric CO2: three by Manabe and his colleagues at NOAA's Geophysical Fluid Dynamics Laboratory and two by James Hansen and his colleagues at NASA's Goddart Institute for Space Studies. The report estimated a climate sensitivity of 3°C, with an error of ±1.5°C.

Figure 8.4: The fathers of the concept of global warming

| Guy Stewart Callendar (1898-1964) | Roger Revelle (1909-1991) | Charles David Keeling (1928-2005) | Wallace Broecker (1931-2019) |

The term "*global warming*" was popularized by Broecker (1975). Wallace Broecker was a famous American geochemist and the author of more than 500 scientific articles and many books. His works on the role of the ocean in climate change, the global ocean circulation map, radiocarbon dating, etc. are considered the foundation of carbon cycle science. While the term global warming appears less than 10 times before 1975, many research papers will use the term extensively after Broecker's publication. Between 1975 and 1980, more than 2 500 scientific articles made reference to global warming. In the 1980s, Broecker and other scientists warned politicians about the dangers of climate change. In 1984, for example, Broecker testified at the Congressional hearing on climate change, declaring that carbon dioxide was the number one long-term environmental problem. On

---

[11]Scientists distinguish between Energy Balance Models (EBM), Earth System Models of Intermediate Complexity (EMIC), and General Circulation Models (GCM), from the simplest to the most complex. However, the distinction between these climate models is generally not obvious.

[12]Nowadays, GCMs are generally classified into three types of families: (1) Atmospheric General Circulation Model (AGCM), (2) Oceanic General Circulation Model (OGCM), and (3) Atmosphere-Ocean General Circulation Model (AOGCM).

June 23, 1988, James Hansen was invited to another Congressional hearing on climate change[13], organized by Senator Al Gore, and stated[14]:

> "*Mr. Chairman and committee members, thank you for the opportunity to present the results of my research on the greenhouse effect which has been carried out with my colleagues at the NASA Goddard Institute for Space Studies. I would like to draw three main conclusions. Number one, the earth is warmer in 1988 than at any time in the history of instrumental measurements. Number two, the global warming is now large enough that we can ascribe with a high degree of confidence a cause and effect relationship to the greenhouse effect. And number three, our computer climate simulations indicate that the greenhouse effect is already large enough to begin to affect the probability of extreme events such as summer heat waves.*"

So in the late eighties, climate change began to become a political issue. In 1988, the United Nations Environment Programme (UNEP) and the World Meteorological Organization (WMO) established the Intergovernmental Panel on Climate Change (IPCC) to provide policy makers with regular scientific assessments of climate change, its impacts and potential future risks, and to recommend options for adaptation and mitigation. Twenty years later, the IPCC and Al Gore were awarded the 2007 Nobel Peace Prize.

Remark 50 *The reader who wants to go further and delve into the history of climate change research will find a vast amount of material and references on the website developed by Weart (2023): `https: // history. aip. org/ climate/ index. htm` . The PDF version of the contents has more than 350 pages. Among the thirty sections, the most interesting are: Introduction and Summary, The Carbon Dioxide Greenhouse Effect, Roger Revelle's Discovery, General Circulation Models of Climate, and Past Climate Cycles: Ice Age Speculations.*

8.1.2 From the Holocene to the Anthropocene?

Definition

The Anthropocene is a proposed geological epoch that dates from the beginning of significant human impacts on Earth's geology and ecosystems, including but not limited to human-induced climate change. The Earth was formed 4.6 billion years ago. Its history can be traced using the geologic time scale (GTS) shown in Figure 8.5. Earth's history is divided into five subdivisions: eon, era, period, epoch, and age. The first three eons (Hadean, Archean and Proterozoic) can be grouped into a supereon called the Precambrian (covering the first 4 billion years). The Phanerozoic runs from the Cambrian to the present. We are in the Cenozoic era, Quaternary period, Holocene epoch and Meghalayan age. It is on the scale of ages that the question of the transition from the Holocene to the Anthropocene is currently being asked.

According to Lewis (2015), the term Anthropocene was popularized by Paul Crutzen[15] and Eugene Stoermer[16] in 2000 (Crutzen and Stoermer, 2000). In 2009, the Anthropocene Working

---

[13]Other scientists attended this congressional hearing, notably future lead authors for chapters of IPCC reports: Syukuro Manabe, Michael Oppenheimer, professor at Princeton University, and William Moomaw, professor at Tufts University.

[14]Source: `https://babel.hathitrust.org/cgi/pt?id=uc1.b5127807`, pages 39-41.

[15]Paul Crutzen (1933-2021) was awarded the 1995 Nobel Prize in Chemistry for his work in atmospheric chemistry, and in particular for his efforts to study the formation and decomposition of atmospheric ozone.

[16]Eugene Stoermer (1934-2012) was a professor of biology who specialized in the study of freshwater species and diatoms.

Figure 8.5: Geologic time scale
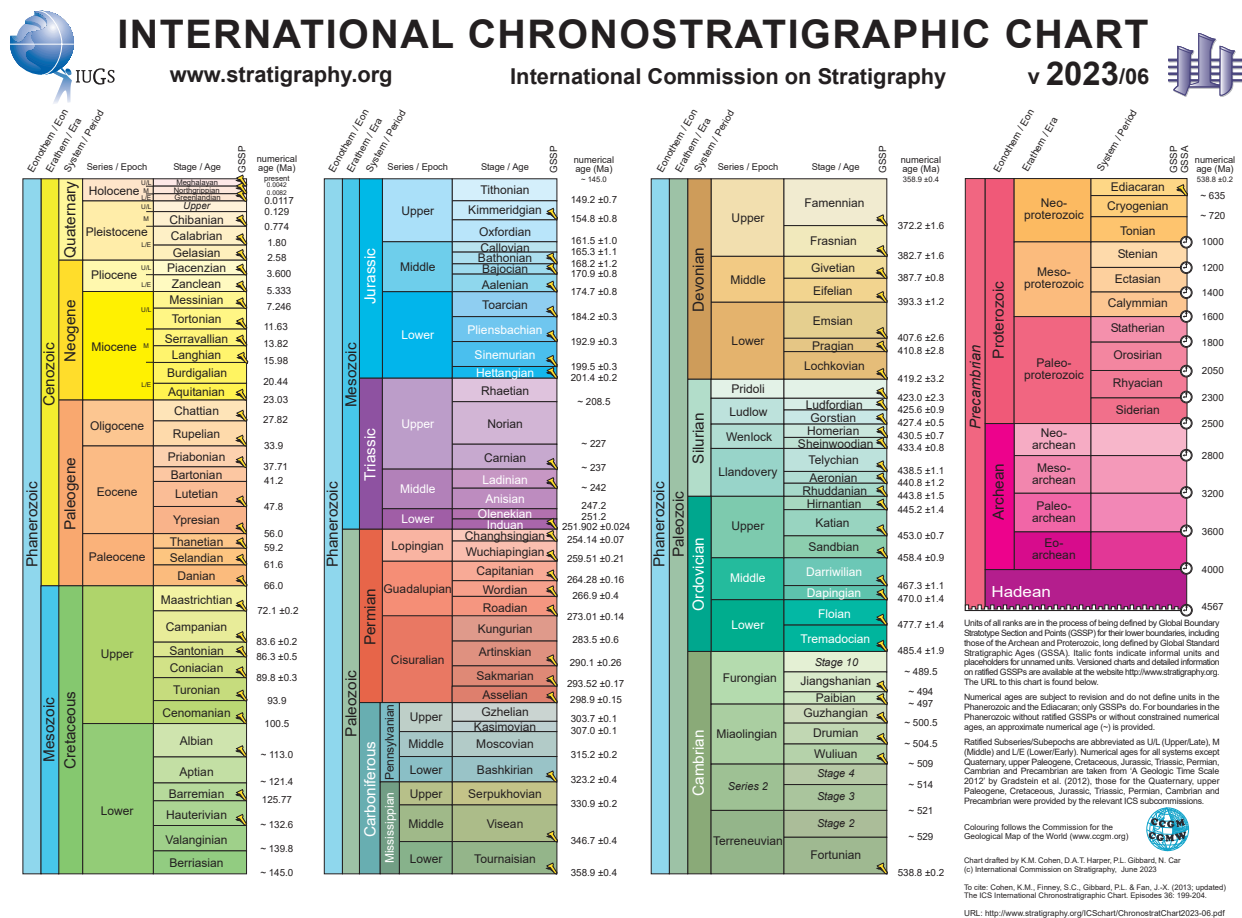
Source: International Commission on Stratigraphy (2023), https://stratigraphy.org/chart.

Group (AWG) was established as part of the quaternary stratigraphy subcommission[17], a constituent body of the International Commission on Stratigraphy (ICS). In 2019, the AWG agreed to submit a formal proposal to the ICS to establish the Anthropocene epoch, but the Anthropocene has not yet been ratified by the ICS. Indeed, the definition of the Anthropocene as a geological epoch rather than a geological event remains controversial and difficult. For example, scientists do not agree on the starting point of this epoch, known as the Global Stratotype Section and Point or GSSP. Lewis (2015) listed nine potential start dates[18]: (1) Megafauna extinction, 50 000–10 000 yr BP; (2) Origin of farming, ∼11 000 yr BP, (3) Extensive farming, ∼8 000 yr BP to present (4) Rice production, 6 500 yr BP to present; (5) Anthropogenic soils, ∼3 000–500 yr BP; (6) New-old world collision, 1 492–1 800; (7) Industrial Revolution, 1 760 to present; (8) Nuclear weapon detonation, 1 945 to present; (9) Persistent industrial chemicals, ∼1 950 to present. In 2016, the AWG proposed that the Anthropocene began in the 1950s, while in 2023 it chose Crawford Lake in Ontario, Canada, to best capture the geological impact of the Anthropocene. According to Waters (2016), the effects of human impact are diverse and of different types. They listed several lines of evidence: the

---

[17]The AWG website is http://quaternary.stratigraphy.org/working-groups/anthropocene.

[18]In the calendar marking system, BP is "*Before Present*", while BC is "*Before Christ*". According to standard practice, the present is set to 1950 Anno Domini (AD).

alteration of sedimentary processes, the altered geochemical signatures in recent sediments and ice, the radiogenic signatures and radionuclides in sediments and ice, the carbon cycle evidence from ice cores, the climate change and rates of sea level change since the end of the last ice age, and the loss of biodiversity. However, the debate about the status of the Anthropocene is reignited by the publication of Gibbard *et al.* (2022):

> "*Over the course of the last decade the concept of the Anthropocene has become widely established within and beyond the geoscientific literature but its boundaries remain undefined. Formal definition of the Anthropocene as a chronostratigraphical series and geochronological epoch following the Holocene, at a fixed horizon and with a precise global start date, has been proposed, but fails to account for the diachronic nature of human impacts on global environmental systems during the late Quaternary. By contrast, defining the Anthropocene as an ongoing geological* event *more closely reflects the reality of both historical and ongoing human-environment interactions, encapsulating spatial and temporal heterogeneity, as well as diverse social and environmental processes that characterize anthropogenic global changes. Thus, an Anthropocene Event incorporates a substantially wider range of anthropogenic environmental and cultural effects, while at the same time applying more readily in different academic contexts than would be the case with a rigidly defined Anthropocene Series/Epoch.*" (Gibbard *et al.*, 2022, page 395).*

It is certainly too early to consider a new era explained by human activity, but it remains relevant that human activity has an impact on climate.

Geological history of the climate

Precambrian and the age of early life   The Earth's climate patterns change naturally on time scales ranging from decades to millions of years[19]. During the Hadean eon (4.6–4.0 Gyr BP), the planet was characterized by volcanism and asteroid impacts. It was very hot, with temperatures certainly averaging around 80°C. The Moon was formed during this period. The Archean eon (4.0–2.5 Gyr BP) is the period when life on Earth began and oceans probably formed. Carbon dioxide emissions were abundant, and this high concentration probably gave rise to the greenhouse effect. The surface temperature decreased and may have been between 0°C and 40°C (Catling and Zahnle, 2020). The Proterozoic eon lasted from 2.5 billion to 540 million years ago. This was a time of oxygen accumulation in the Earth's atmosphere. The temperature history of the Proterozoic is controversial. There is evidence that the first glaciations occurred during the Proterozoic. However, the average temperature during the Proterozoic is estimated to have been between 10°C and 30°C.

Table 8.2: Units of time

Symbol	Definition	(in year)	Symbol	Name
Kyr/kyr	Thousand/Kilo years	10^3	ka	Kiloannus
Myr/myr	Mega/Million years	10^6	Ma	Megaannus
Gyr/byr	Giga/Billion years	10^9	Ga	Gigaannus

To estimate temperatures during the Precambrian, scientists use indirect methods, including geochemical proxies (chemical properties of rocks and minerals), paleontological studies (type and distribution of fossils and sedimentary rocks), and general circulation models. One of the most

---

[19]in Table 8.2, report the two time units commonly used in geology (yr vs. a, *e.g.*, Gyr vs. Ga).

Box 8.3: Faint young Sun paradox & Snowball Earth hypothesis

The faint young Sun paradox describes the apparent contradiction between observations of liquid water early in Earth's history and the astrophysical expectation that the Sun's output during the Archean eon would be only 75% as intense as it is today. If the Sun was fainter when the Earth was young, we might expect the Earth to be completely frozen, but there is evidence for the presence of liquid water on the Earth's surface at that time (Feulner, 2012). Many explanations have been proposed: ammonia, methane, reduced albedo, tidal heating, etc. However, the most credible assumption is the high concentration of carbon dioxide during this period of weaker solar radiation:

> "[...] we argue that the faint young Sun problem for Earth has essentially been solved. Unfrozen Archean oceans were likely maintained by higher concentrations of CO_2, consistent with the latest geological proxies, potentially helped by additional warming processes. This reinforces the expected key role of the carbon cycle for maintaining the habitability of terrestrial planets." (Charnay et al., 2020, page 1).

Other curious events during the Precambrian are the Proterozoic glaciations. The Snowball Earth hypothesis proposes that the planet's surface was completely or nearly completely frozen during these icehouse climates, especially between 750 and 550 Ma (Hoffman et al., 1998; Hoffman and Schrag, 2002). The Snowball Earth hypothesis can be explained by a runaway ice-albedo feedback. Albedo is the fraction of sunlight reflected back into space by a surface. For example, ice and snow have a high albedo, which means they reflect a lot of sunlight. As more ice and snow cover the Earth's surface, more sunlight is reflected back into space, causing the Earth to cool. This cooling leads to more ice and snow formation, which further cools the Earth, and so on. This is the runaway ice-albedo feedback. This hypothesis is currently a topic of debate among scientists, with alternative explanations also presented in the academic literature, including the high-obliquity hypothesis and the volcanic ash hypothesis. The high-obliquity hypothesis suggests that during a glacial period, the Earth had a much larger axial tilt than it does today[a]. The result would have been a decrease in the amount of solar radiation reaching the poles and an increase in the amount of solar radiation absorbed at the equator. As a result, glaciers would have melted at the equator and ice caps would have formed at the poles. The volcanic ash hypothesis suggests that volcanic ash filled the Earth's atmosphere, resulting in a cooling effect caused by the blocking of sunlight.

---

[a]This explanation refers to the Milankovitch cycles, which are variations in the Earth's orbit around the Sun that affect the amount of sunlight the planet receives. The main cycles concern the shape of the Earth's ellipsoid orbit (eccentricity), the axial tilt or angle of the Earth's axis (obliquity), and the direction of the Earth's axis of rotation (precession).

common methods is the clumped isotope thermometer, which is based on the analysis of the isotopic composition of certain molecules, such as carbonate minerals. For example, we can use the ratio of the two carbon isotopes ^{12}C and ^{13}C in carbonate rocks to infer the temperature at which the rocks formed. Its definition is:

$$\delta^{13}C = 1000 \times \left(\frac{^{13}C}{^{12}C}_{\text{sample}} \bigg/ \frac{^{13}C}{^{12}C}_{\text{standard}} - 1 \right) \tag{8.1}$$

The unit of $\delta^{13}C$ is parts per thousand[20] (per mil or ‰). Databases of $\delta^{13}C$ can be found in the permanent repository ClumpDB[21] or scientific journals that publish research papers on geological temperatures. Below we consider the Precambrian Marine Carbonate Isotope Database (PMCID) provided by Shields and Veizer (2002), which can be downloaded from `https://earthref.org/ERDA/48`. The PMCID is a compilation of strontium, carbon, and oxygen isotope compositions of about 10 000 marine carbonate rocks of Archean to Ordovician age (between 3 800 Ma and 450 Ma), and includes data from 150 published articles and books. In Figure 8.6, we have reproduced the graph of the evolution of the carbon isotopes of marine carbonate obtained by Shields and Veizer (2002). We observe a high variation of $\delta^{13}C$, which indicates that the temperatures on Earth have varied a lot during the Precambrian. This also implies a large uncertainty in the estimates, and we must be careful with the figures calculated for this period.

Figure 8.6: Carbon isotopic evolution of marine carbonate
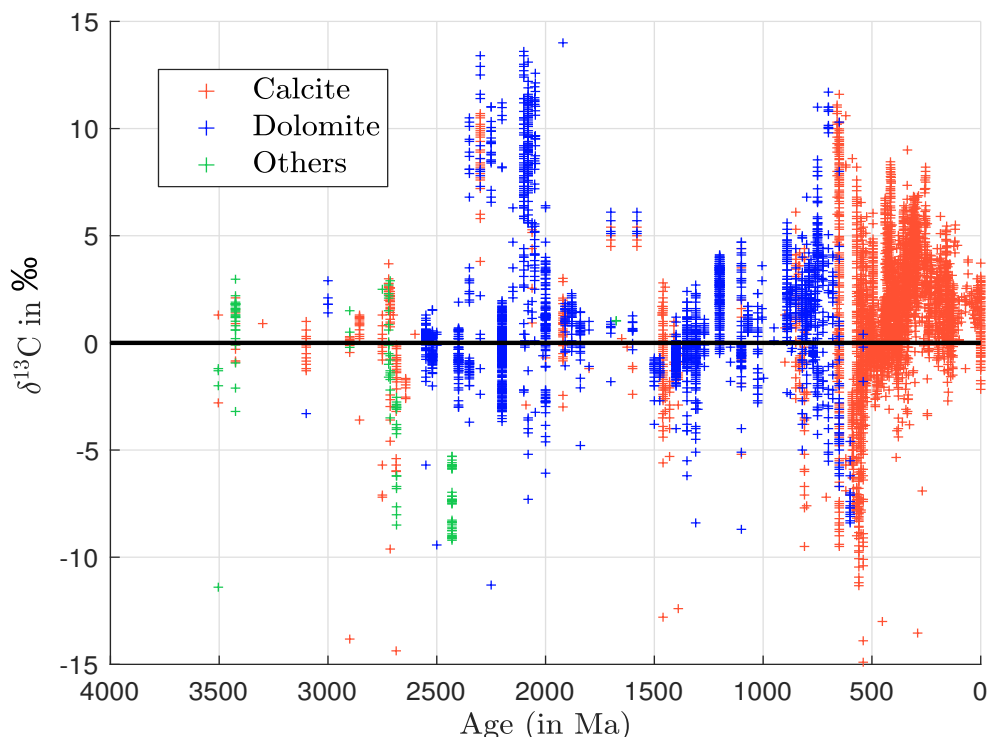
Source: Shields and Veizer (2002, Figure 2, page 5) & `https://earthref.org/ERDA/48`.

Another controversy is the Snowball Earth hypothesis presented in Box 8.3. How do we explain the massive perturbations in the carbon cycle during the Proterozoic (Figure 8.7), and how do we explain the snow and glaciers on all continents between 750 Ma and 550 Ma, when they seem to have been near the equator? Again, we must be careful because there may be multiple answers and a combination of factors. Moreover, it is difficult to imagine the time when we refer to the palaeoclimate, because the magnitude of these periods is so different from what we have experienced or learned in history, even if we consider the last 3500 years, which corresponds to the first written records. To better understand that these developments are very long, we generally use a tool called the cosmic calendar (Figure 8.8). This is a scale that maps the 13.8 billion year age of the universe

---

[20]The standard for $\delta^{13}C$ is the Pee Dee Belemnite (PDB), which has a ratio $\frac{^{13}C}{^{12}C}$ equal to 0.011238.

[21]The web address is `www.earthchem.org/communities/clumpdb`.

Figure 8.7: Neoproterozoic carbon isotope data compilation
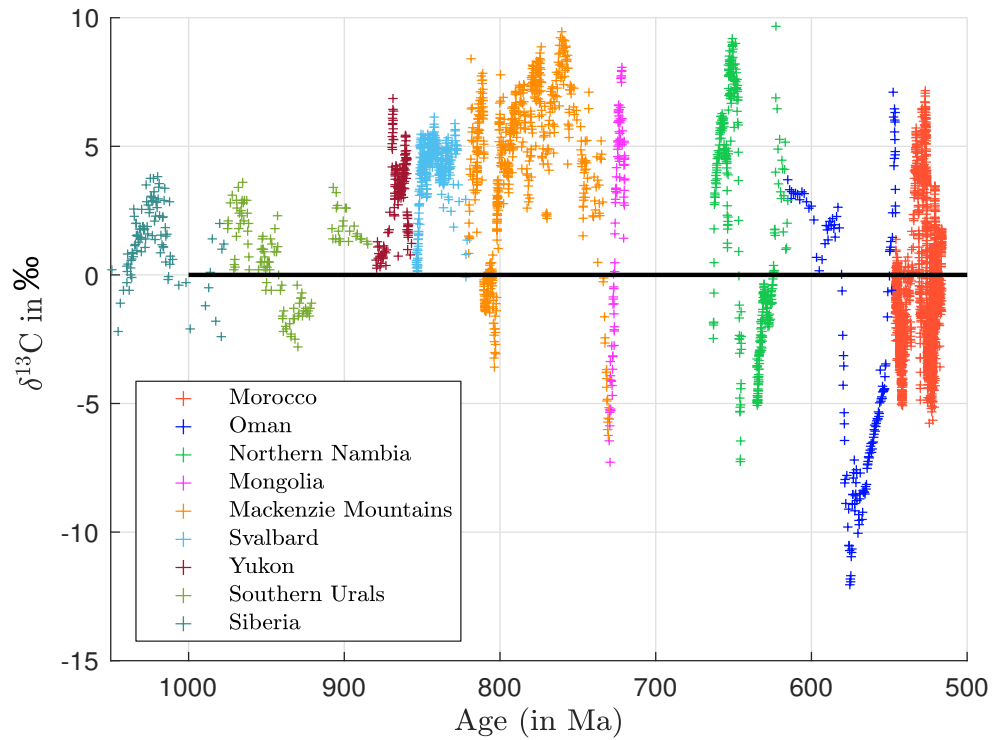
Figure 8.8: Cosmic calendar

to a single year to make it more intuitive for educational purposes. On this scale, the Big Bang took place at midnight on January 1, and the present time is mapped to midnight on December 31. The time in days is then equal to:

$$t_{\text{days}} = \left(1 - \frac{t_{\text{Ga}}}{13.797}\right) \times 365 \text{ days}$$

According to Sagan (1986), the Precambrian period begins on September 14 and ends on December 17. The first primates appear on December 29 and humans on December 31 at 22:30. They invent agriculture at 23:59:20. The invention of writing is at 23:59:45, and the voyage of Christopher Columbus takes place in the last second at 23:59:59. Modern human civilization then occurs in the last twenty minutes.

<div align="center">Box 8.4: Temperature scales</div>

Three different scales are commonly used to measure temperature: Celsius, Kelvin, and Fahrenheit. Their symbols are °C, K, and °F, respectively. The Fahrenheit scale was developed by the German physicist Gabriel Fahrenheit in 1724. The Fahrenheit scale is defined by two fixed points: 32°F for the freezing point of water at sea level and 212°F for the boiling point of water at sea level. The interval between these two points is divided into 180 steps. The Celsius scale was developed by the Swedish astronomer Anders Celsius in 1742. This scale divides the interval between the two fixed points into 100 steps. The Celsius scale is also known as the centigrade scale. Finally, the Kelvin scale, developed by Lord Kelvin in 1848, is a modification of the Celsius scale in which $0\,\text{K}$ is absolute zero. The relationships between the Celsius and Kelvin scales are $\mathcal{T}_{\circ\text{C}} = \mathcal{T}_{\text{K}} - 273.15$ and $\mathcal{T}_{\text{K}} = \mathcal{T}_{\circ\text{C}} + 273.15$. For Celsius and Fahrenheit, we have $\mathcal{T}_{\circ\text{C}} = \frac{5}{9}\left(\mathcal{T}_{\circ\text{F}} - 32\right)$ and $\mathcal{T}_{\circ\text{F}} = \frac{9}{5}\mathcal{T}_{\circ\text{F}} + 32$. Absolute zero is $-273.15°\text{C}$, $0\,\text{K}$ and $-459.67°\text{F}$, implying that $\mathcal{T} \geq -273.15°\text{C}$, $\mathcal{T} \geq 0\,\text{K}$, and $\mathcal{T} \geq -459.67°\text{F}$. The melting point (at standard pressure) is obtained at temperatures of 0°C, $273.15\,\text{K}$ and 32°F, while the boiling point of water corresponds to temperatures of 100°C, $373.15\,\text{K}$ and 212°F.

Palaeoclimate during the Phanerozoic   Studying the climate of the Phanerozoic eon is generally easier than studying the Precambrian, because this period, which begins 541 million years ago, is characterized by a more abundant and diverse fossil record. In addition, we can use other materials, such as ice cores and ocean sediments. Figure 8.9 shows the global picture of temperature evolution during the Phanerozoic[22]. It was produced in 2014 by Glen Fergus, who has done a remarkable job of collecting data from a dozen published research papers[23]. It shows several episodes of hothouse scenarios and glacial cycles. The average temperature range compared to today is between −6°C and +15°C. The graph is divided into five panels. The first panel corresponds to the Palaeozoic (Cambrian, Ordovician, Silurian, Devonian, Mississippian, Pennsylvanian and Permian) and Mesozoic (Triassic, Jurassic and Cretaceous) eras. It covers the period from 541 Ma to 66 Ma. The second panel corresponds to a part of the Tertiary period and includes the Paleocene, Eocene, Oligocene and Miocene epochs (from 66 Ma to 5.3 Ma). The third and fourth panels show the average temperature of the Pliocene and Pleistocene. The Quaternary, which begins at 2.6 Ma BP, is characterised by many glaciations, while the Holocene is marked by an increase in temperature.

---

[22]Scientists use the Celsius, Kelvin and Fahrenheit scales indifferently. That is why we have given their definition in Box 8.4.

[23]We have removed some data from the original figure.

Figure 8.9: Earth temperature since 500 Myr BP (°C vs. 1960-1990 average)
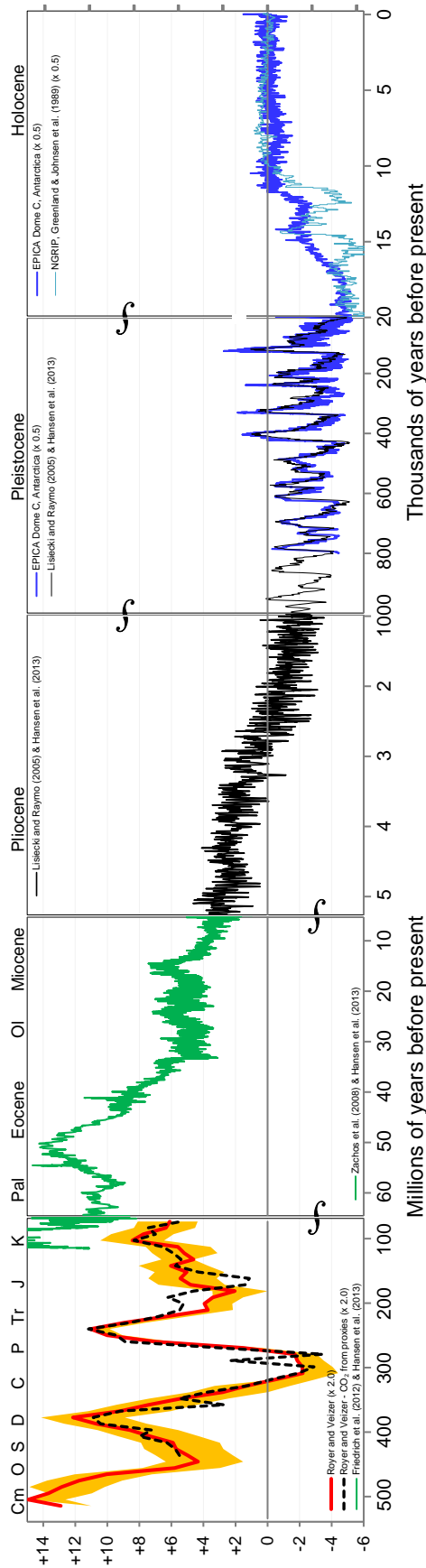
This shows estimates of global average surface air temperature over the 540 Myr of the Phanerozoic eon, since the first major proliferation of complex life forms on our planet. A major achievement of the last 30 years of climate science has been the production of a large set of actual measurements of temperature history (from physical proxies), replacing much of the earlier geological induction (*i.e.*, informed guesses). The graph shows selected proxy temperature estimates. Because many proxy temperature reconstructions indicate local rather than air temperature – or ocean rather than air temperature – considerable approximation may be involved in deriving these global temperature estimates. As a result, the relativities of some of the estimates plotted are approximate, especially the early ones. Surface air temperature is plotted as anomalies (differences) from the average over the reference interval 1960-1990 (which is about 14°C).

Time is plotted forward to the present, taken as 2015 CE. It connects five separate linearly scaled segments, expanding by about an order of magnitude at each vertical break. The first segment groups the Cambrian (C), Ordovician (O), Silurian (S), Devonian (D), Carboniferous (C), Permian (P), Triassic (Tr), Jurassic (J), and Cretaceous (K). The first break is at the Mesozoic-Cenozoic boundary (~65 Myr ago). This is the K-T boundary (called Cretaceous-Paleogene) where the dinosaurs became extinct. The second segment groups the Paleocene, Eocene, OLigocene, and Miocene. The second break is at the Miocene-Pliocene boundary (~5.3 Myr ago). The last three segments are Pliocene, Pleistocene, and Holocene. The last two breaks are at 1 Myr BP and 20 kyr BP.

Source: Glen Fergus (2014), en.wikipedia.org/wiki/File:All_palaeotemps.png & https://gergs.net/all_palaeotemps-2.

Let us know how these temperatures are estimated. To do this, we need to go back to the scientific history of climate change presented in section 8.1.1. Paleoclimate research began in the 1950s with several advances: radiocarbon dating with carbon 14, analysis of deep-sea cores and marine sediments, drilling of ice cores in Greenland and Antarctica, etc. Whichever method we use, we need two modelling tools: an age dating model and a formula that converts the calculated metric into a temperature, because the latter is not observed.

The history of ice core science is studied in Langway (2008) (2008) and Jouzel (2013). One of the first attempts to explore the interior of an ice sheet was made by Ernst Sorge during the Alfred Wegener Expedition to Eismitte in central Greenland from July 1930 to August 1931. Using a 15-meter deep pit, he investigated the individual limits of annual snow accumulation. Between 1949 and 1952, the first ice cores were extracted separately by three international research teams at depths of between 100 and 150 meters, but the quality of the ice cores was poor. With several initiatives (Site 2, Byrd Station and Little America), the period 1957-1958 is generally considered to be the starting point of ice core research (Jouzel, 2013). A new step was achieved in 1961, when the first ice core longer than 1 000 meters was taken at Camp Century. Since then, many projects have been carried out[24], but the most famous remains the ice-core drilling in Vostock, Antarctica (Petit *et al.*, 1999). Indeed, the depth record is held by the Vostock station, where scientists have drilled to a depth of 3 770 meters. In fact, it is very difficult to go beyond this figure because we reach the surface of oceans or lakes. In addition, the bottom of the ice core at very deep depths is modified by the water at the ocean surface and does not contain any palaeoclimate information.

Table 8.3: Recovered deep and very deep ice cores

Greenland			Antarctica		
Site 2	1956	305 m	Byrd Station	1957-1958	307 m
Site 2	1957	411 m	Little America	1958-1959	264 m
Camp Century	1961-1966	1387 m	Byrd Station	1966-1968	2164 m
Dye 3	1971	372 m	Vostock	1990-1998	3623 m
Milcent	1973	398 m	Dome Fuji	1994-1997	2503 m
Crete	1974	405 m	Vostock	2005-2007	3658 m
Dye 3	1979-1981	2037 m	Dome Fuji	2003-2007	3035 m
GRIP	1989-1992	3029 m	Dome C	1999-2005	3270 m
GISP 2	1989-1993	3057 m	Kohnen Station	2001-2006	2774 m
NGRIP	1996-2004	3090 m	WAIS	2006-2011	3405 m

Source: Langway (2008, Appendix A).

Analysis of trapped air bubbles in ice cores provides a direct record of the composition of the atmosphere at the time the ice formed. For example, we can measure the concentrations of greenhouse gases such as carbon dioxide and methane in the air bubbles to reconstruct past climate conditions. However, these measurements are not very robust and scientists prefer to use indirect measures. Dansgaard's seminal paper in 1964 discussed methods based on stable isotopes (Dansgaard, 1964). He defined the relative deviation δ of the heavy isotope content as follows:

$$\delta = 1000 \times \left(\frac{R_{\text{sample}} - R_{\text{standard}}}{R_{\text{standard}}} \right) \tag{8.2}$$

---

[24]In Table 8.3, we have listed some of these projects and the corresponding depth of the drilling. The locations of the projects are shown in Figures 8.10 and 8.11.

Figure 8.10: Greenland deep drilling sites
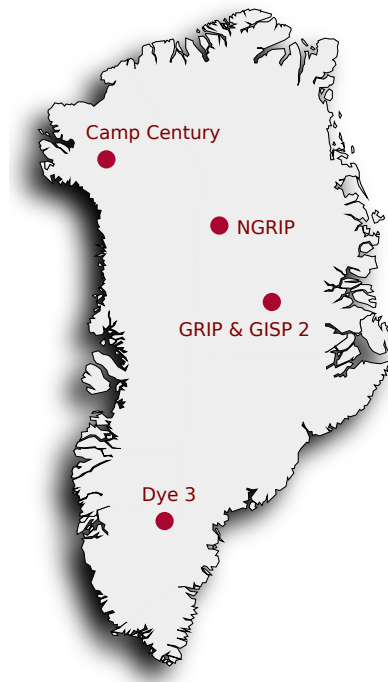
Source: Langway (2008, Appendix A), created with paintmaps.com.

Figure 8.11: Antarctica deep drilling sites

Source: Langway (2008, Appendix A), created with paintmaps.com.

Box 8.5: Isotopes of chemical elements

Isotopes are atoms of the same element. They have the same number of protons but differ in the number of neutrons. To write the symbol for an isotope, we use the standard AZE notation[a] ^A_ZE, where E is the chemical symbol for the element, A is the atomic mass number (the total number of protons and neutrons in the nucleus of the atom) and Z is the atomic number (the number of protons in the nucleus of the atom). For example, hydrogen has seven isotopes: $^1_1$H (protium), $^2_1$H or D (deuterium), $^3_1$H or T (tritium), $^4_1$H, $^5_1$H, $^6_1$H, and $^7_1$H. We also distinguish between stable and unstable isotopes. Stable isotopes have a stable nucleus and the forces holding the protons and neutrons together are strong enough to prevent the nucleus from decaying. The most stable isotopes have a neutron to proton ratio close to one. Unstable (or radioactive) isotopes have an unstable nucleus and they can decay into other elements. The most frequent isotope of an element is called the major (or principal) isotope. In the case of hydrogen, protium and deuterium are the two stable isotopes and protium is the main isotope, accounting for about 99.98% of all occurring hydrogen atoms. In addition to these two isotopes, five other isotopes of hydrogen are unstable. The table below lists the stable and unstable isotopes of hydrogen, carbon, oxygen and uranium. The main isotope and its abundance are also given.

Element	Stable isotopes	Unstable isotopes	Major isotope
Hydrogen	$^1_1$H and $^2_1$H	$^3_1$H – $^7_1$H	Protium (99.98%)
Carbon	$^{12}_6$C and $^{13}_6$C	$^8_6$C – $^{11}_6$C and $^{14}_6$C – $^{22}_6$C	Carbon-12 (98.90%)
Oxygen	$^{16}_8$O and $^{18}_8$O	$^{17}_8$O and $^{19}_8$O – $^{27}_8$O	Oxygen-16 (99.76%)
Uranium		$^{232}_{92}$U – $^{242}_{92}$U	Uranium-238 (99.27%)

A heavy isotope is an isotope of an element that has more neutrons than the most abundant isotope. In general, heavy isotopes are more stable, and can have different chemical properties. The heavy isotopes of the previous elements are $^2_1$H, $^{13}_6$C, $^{18}_8$O and $^{238}_{92}$U.

[a]We generally simplify the notation with ^{A}E because all the isotopes of the same element have the same atomic number.

where R is the absolute content. δ is measured in ‰. This is the general form of Equation (8.1) given on page 269, if we replace the ratio of the carbon isotope by R. As the chemical formula of water is H_2O, climate reconstruction from ice cores is based on the analysis of hydrogen and oxygen. In the case of hydrogen, the common isotope is ^{1}H, while the heavy isotope is ^{2}H (also called the deuterium or D). The ratio R is then $\dfrac{^2H}{^1H}$ and the relative variation is written as δD. In the case of oxygen, the common isotope is ^{16}O, while the heavy isotope is ^{18}O. The ratio R is then $\dfrac{^{18}O}{^{16}O}$ and the relative variation is written as δ^{18}O. To compute δD and δ^{18}O, we need the standard values of ^{2}H/^{1}H and ^{18}O/^{16}O. Using the standard mean ocean water (SMOW), they are equal to 0.00015576 and 0.0020052. Dansgaard (1964) assumed that the isotopic fractionation of water implies that δD $= (8.1 \pm 0.1)\,\delta^{18}$O $+ (11 \pm 1)\,‰$ at equilibrium[25]. More generally, Merlivat and Jouzel (1979) showed that the relationship is linear: δD $= s\,\delta^{18}$O $+ d$ where the slope s and the deuterium excess d may depend on several factors: temperature, regional characteristics, latitude, humidity, etc.

---

[25]This relationship has already been established by Craig (1961), who found that δD $= 8.0\,\delta^{18}$O $+ 10\,‰$. It is known as the meteoric water line.

Figure 8.12: Part of an ice core at WAIS Divide Field Camp

The aim of ice core analysis is to estimate the temperature function $t \longmapsto \mathcal{T}(t)$ with respect to the time age t. The raw analysis provides two measurements: the depth d of the ice core drilling and the isotope ratio measure δ, which is calculated using an instrument called an isotope ratio mass spectrometer (IRMS). We therefore observe the isotope function $d \longmapsto \delta(d)$. To obtain the temperature function, we proceed in two steps:

1. First, we transform the depth d of the ice core drilling into the time age t:

$$t = \varphi_t(d)$$

2. We then estimate the temperature $\mathcal{T}$ associated with the isotope ratio $\delta(d)$:

$$\mathcal{T} = \varphi_{\mathcal{T}}(\delta(d))$$

Combining the two previous equations gives the desired parametric function $t \longmapsto \mathcal{T}(t)$. To illustrate the temperature reconstruction, we consider the research of Petit *et al.* (1999), who estimated the climate and atmospheric history of the past $420\,000$ years using the Vostok ice cores. To do this, we download the file `deutnat-noaa.txt` from https://www.ncei.noaa.gov/access/paleo-search/study/2453. Below we report a sample of this dataset:

```
# Vostok - Isotope and Gas Data and Temperature Reconstruction
#-----------------------------------------------------------------------
#                  World Data Service for Paleoclimatology, Boulder
#                                    and
#                       NOAA Paleoclimatology Program
#              National Centers for Environmental Information (NCEI)
#-----------------------------------------------------------------------
```

```
[...]

#--------------------
# Variables format: Short_name <tab>
what, material, error, units, seasonality, data_type, detail,
method, data_format, additional_information
#  Note: the Short_name does not require a PaST term.
#
## depth_m     depth, , , meter, , ice cores;
             climate reconstructions, , , N,
## ice_ageBP  ice age, , , calendar year before present, , ice cores;
             climate reconstructions, , , N, GT4 chronology
## deltaD     delta 2H, bulk ice, , per mil SMOW, , ice cores, interpolated,
             isotope ratio mass spectrometry, N,
## deltaTS    surface temperature, delta 2H, , degree Celsius, annual, ice cores;
             climate reconstructions, anomalized, , N, linear regression; anomaly from recent
#
#--------------------
# Data:
# Data lines follow (have no #)
# Data line format - tab-delimited text, variable short name as header
# Missing_Values:
depth_m ice_ageBP deltaD deltaTS
   0         0     -438.0   0.00
   1        17     -438.0   0.00
   2        35     -438.0   0.00
   3        53     -438.0   0.00
   4        72     -438.0   0.00
   5        91     -438.0   0.00
   6       110     -438.0   0.00
   7       129     -438.0   0.00
   8       149     -442.9  -0.81
   9       170     -437.9   0.02
  10       190     -435.8   0.36
  11       211     -443.7  -0.95
  12       234     -449.1  -1.84
  13       258     -444.6  -1.09
  14       281     -442.5  -0.75
  15       304     -439.3  -0.22
  16       327     -440.9  -0.48
  17       351     -442.5  -0.75
  18       375     -436.6   0.23
  19       397     -430.0   1.33
  20       420     -435.9   0.35

[...]

3300    416872   -430.8   1.36
3301    417419   -430.3   1.43
3302    417969   -430.4   1.40
3303    418526   -431.1   1.27
3304    419095   -433.0   0.94
3305    419682   -435.5   0.51
3306    420281   -435.2   0.54
3307    420888   -436.4   0.32
3308    421507   -437.3   0.15
3309    422135   -437.6   0.08
3310    422766   -436.6   0.23
```

This dataset has four variables: depth d (meter), age t (yr BP), deuterium isotope ratio δD (‰) and surface temperature ΔT (°C). We have included the first 20 observations and the last 10 observations. Note that the measurement is made at every meter of the ice core. For example, the 11th observation corresponds to a depth of 10 meters and the authors have calculated a deuterium isotope ratio δD of -435.8‰. Using an age dating model[26], this depth corresponds to an age of 190 years BP. Using a climate model[27], the authors obtained a surface temperature variation $\Delta \mathcal{T}$ of 0.36 degrees Celsius. In the first panel of Figure 8.13, we plot the function $d \longmapsto \delta D(d)$. The maximum depth is 3 310 meters while $\delta D(d)$ varies between -500‰ and -400‰. Several peaks and cycles are observed. The results of the age dating model are shown in the second panel. When the depth is less than 2 000 meters, we have a linear relationship between d and the age t. Beyond this threshold, the function is convex. The mapping between δD and $\Delta \mathcal{T}$ depends on the model chosen. In our case, we use a linear function $\Delta \mathcal{T} = 78.2361 + 0.1794 \times \delta D$, which is shown in the third panel. This gives the relationship between depth d and temperature $\Delta \mathcal{T}$ in the fourth panel. Finally, the temperature function $t \longmapsto T(t)$ is shown in Figure 8.14. In Petit *et al.* (1999), this corresponds to panel b (isotopic temperature of the atmosphere) of Figure 3 on page 431.

Figure 8.13: Isotopic reconstruction of Vostok ice cores
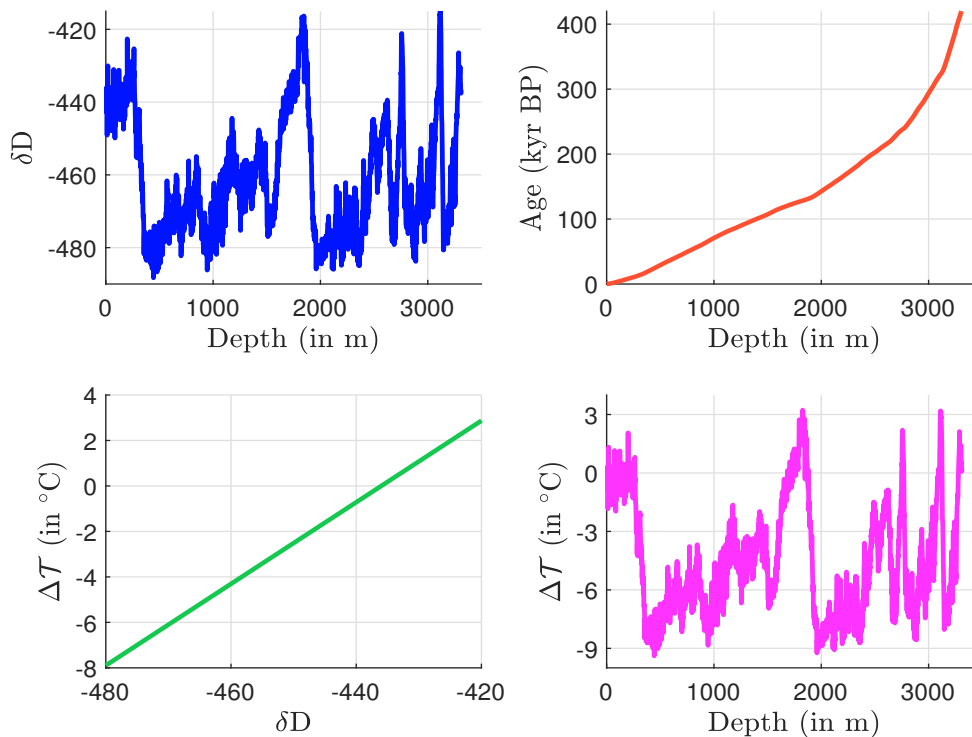
Source: Petit *et al.* (1999), https://www.ncei.noaa.gov/access/paleo-search/study/2453 & Author's calculations.

Remark 51 *Other variables can be found in the dataset. For example, we report the gas concentration of CH_4 (in ppbv or parts per billion by volume) and CO_2 (in ppmv or parts per million by volume) and also the isotope ratio $\delta^{18}O$ in Figure 8.15.*

---

[26]See Parrenin *et al.* (2007) for a review of the construction of the time scale with a focus on the EDC3 chronology, the importance of age markers and the mapping process.

[27]Described in the second section on page 431 of the article.

Figure 8.14: Temperature reconstruction of Vostok ice cores

Source: Petit *et al.* (1999) & https://www.ncei.noaa.gov/access/paleo-search/study/2453.
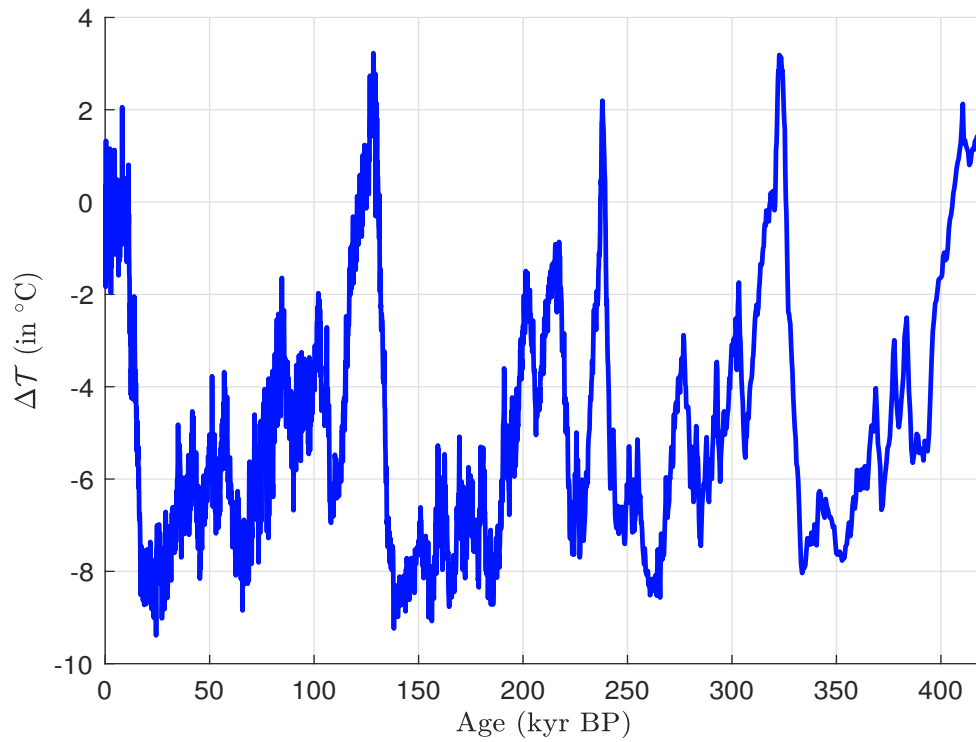
Figure 8.15: Gas concentration of Vostok ice cores

Source: Petit *et al.* (1999) & https://www.ncei.noaa.gov/access/paleo-search/study/2453.
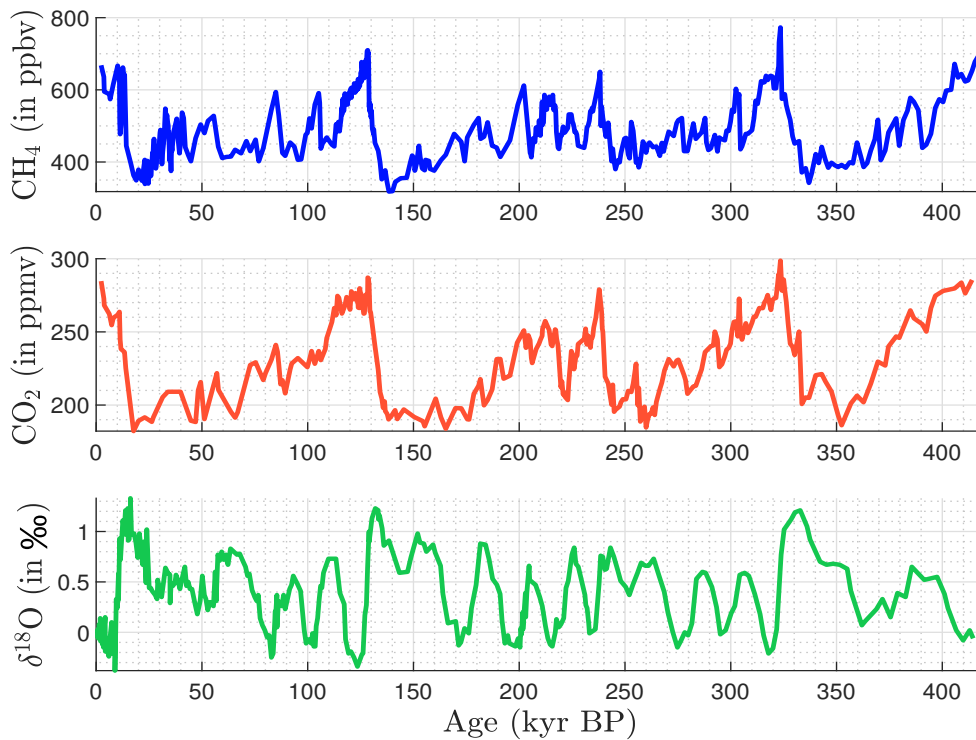
In addition to ice cores, the previous analysis can be applied to rocks, lake and ocean sediments, and tree rings. Let's go back to Figure 8.9 and look at the different scientific references used to calculate the Earth's temperature since 500 Myr BP:

- The works of Royer *et al.* (2004) are related to the works of Veizer *et al.* (1999) and Veizer *et al.* (2000). Veizer *et al.* (1999) collected a total of 2 128 calcitic and phosphatic shells, mainly brachiopods with some conodonts and belemnites, and measured their $^{87}Sr/^{86}Sr$, $\delta^{13}C$ and $\delta^{18}O$ values. It is a statistical study where the authors calculated correlograms, trends and performed a factor analysis to show that the $^{87}Sr/^{86}Sr$, $\delta^{13}C$ and $\delta^{18}O$ system is driven by three factors: the first two are tectonic factors while the third is a biologically mediated redox balance of carbon and sulfur cycles. In 2000, Jàn Veizer and two co-authors used the previous database to write a controversial paper entitled "*Evidence for decoupling of atmospheric CO_2 and global climate during the Phanerozoic eon*". In particular, they concluded:

  "*[...] our data conflict with a temperature reconstruction using an energy balance model that is forced by reconstructed atmospheric carbon dioxide concentrations. The results can be reconciled if atmospheric carbon dioxide concentrations were not the principal driver of climate variability on geological timescales for at least one-third of the Phanerozoic eon, or if the reconstructed carbon dioxide concentrations are not reliable.*" (Veizer *et al.*, 2000, page 698).

  The research by Dana Royer and his co-authors consisted of examining the dataset and the reconstructed temperature curve, and proposing a new reconstructed curve by applying a correction based on seawater pH:

  "*Recent studies have purported to show a closer correspondence between reconstructed Phanerozoic records of cosmic ray flux and temperature than between CO_2 and temperature. The role of the greenhouse gas CO_2 in controlling global temperatures has therefore been questioned.[...] We explore the possible influence of seawater pH on the $\delta^{18}O$ record and find that a pH-corrected record matches the glacial record much better. Periodic fluctuations in the cosmic ray flux may be of some climatic significance, but are likely of second-order importance on a multimillion-year timescale.*" (Royer *et al.*, 2004, page 4).

- In 2004, the North Greenland Ice Core Project (NGRIP) members published an article in *Nature*, with two main contributions. First, they extensively studied the late Eemian period, which is the last interglacial period which began about 130 000 years ago and ended about 115 000 years ago. They revealed a previously unrecognised warm period initiated by an abrupt climate warming about 115 000 years ago, before glacial conditions had fully developed North Greenland Ice Core Project members (2004, pages 2-3). Second, they observed significant regional climate differences when analyzing ice cores from northern (NGRIP) and central (GRIP/GISP 2) Greenland regions, and suggested that "*the extent of ice in the Northern Hemisphere modulated the latitudinal temperature gradients in Greenland*" North Greenland Ice Core Project members (2004, page 1).

- The research carried out by Lisiecki and Raymo (2005) is remarkable. They studied the LR04 database, which contains over 38 000 individual $\delta^{18}O$ measurements from 57 ocean drilling sites[28]. Benthic $\delta^{18}O$ records are a climate proxy used to reconstruct past ocean temperatures (and also ice volume). Benthic foraminiferal calcite is the calcium carbonate material that

---

[28]Most of them were in the Atlantic Ocean, a few in the Pacific Ocean and even fewer in the Asian Ocean.

makes up the shells of benthic foraminifera, tiny organisms that live on the ocean floor. The calcium carbonate incorporates oxygen from the seawater, and the ratio δ^{18}O in foraminiferal shells depends on the temperature of the seawater when the shell was formed. Cooler water has higher δ^{18}O values, while warmer water has lower δ^{18}O values. The authors used various alignment and stacking techniques to synchronise the 57 time series. They were then able to derive a graphical correlation and create a unique series showing the common pattern of the 38 000 individual δ^{18}O measurements.

• The data labelled EPICA Dome C, Antarctica correspond to those calculated by Jouzel *et al.* (2007). A preliminary version of these data was published three years earlier (EPICA community members, 2004). This article focused on the comparison of the deuterium isotope ratio δD between the Vostok and Dome C ice cores. The article by Jouzel *et al.* (2007) went further by converting the δD data into a temperature record[29]. They also found that the combined effects of the precession and obliquity Milankovitch cycles drove the dynamics of the ice ages, implying that climate variability over the past 800 000 years cannot be explained by the radiative forcing of CO_2 and CH_4 alone.

• James Zachos and his co-authors have collected several sources of data (Zachos *et al.*, 2001, 2008). They have used similar data to those studied by Lisiecki and Raymo (2005), but they have included both the isotope ratios δ^{13}C and δ^{18}O. By crossing these two pieces of information, they deduced the partial pressure of carbon dioxide[30] pCO_2 in the ocean over the past 65 million years. Another important contribution of these papers is the consideration of feedbacks and new estimates of climate sensitivity.

• Friedrich *et al.* (2012) compiled a new dataset of benthic foraminifera δ^{13}C and δ^{18}O ratios for the middle to late Cretaceous, spanning 55 million years. They highlighted the role of ocean circulation in greenhouse Earth episodes. For example, they showed that:

 1. "*There was widespread formation of bottom waters with temperatures above* $20°$C *during the Cretaceous hothouse world*;

 2. *These bottom waters filled the North Atlantic and probably originated as thermocline or intermediate waters in the tropical oceans*;

 3. *The interbasin δ^{13}C gradient was unusually large during the Cretaceous hot greenhouse, probably because the North Atlantic sills prevented the free exchange of waters in the deep basin*;

 4. *The hot greenhouse ended when the Equatorial Atlantic Gateway opened sufficiently to flood the deep North Atlantic with relatively cool polar waters formed in the Southern Ocean.*"

According to this study, plate tectonic movements are then an important factor for understanding the temperature dynamics during the Mesozoic, because they have influenced the circulation of warm and cold waters.

---

[29]The original data from Jouzel *et al.* (2007) can be accessed at www.ncei.noaa.gov/access/paleo-search/study/6080 (file edc3deuttemp2007.txt). To obtain all the EPICA Dome C published data (80 files), the reader can explore the directory ftp.ncdc.noaa.gov/pub/data/paleo/icecore/antarctica/epica_domec.

[30]pCO_2 is a measure of the amount of carbon dioxide dissolved in a gas or liquid. It is measured in ppm or ppmv (parts per million by volume).

- [Hansen](#) *et al.* (2013) proposed a simplified model to investigate the state dependence of climate sensitivity, sea level and atmospheric CO_2 through the Cenozoic. Their basic idea is to estimate Cenozoic sea level and temperature from empirical data, filtering the data with minimal assumptions and modelling. In fact, the paper contains only five equations. For example, the equation for the lower ocean temperature is:

$$
\mathcal{T}_{\mathrm{LO}} =
\begin{cases}
5 - 8 \left(\dfrac{\delta^{18}\mathrm{O} - 1.75}{3} \right) & \text{if } \delta^{18}\mathrm{O} < 3.25 \\[2mm]
1 - 4.4 \left(\dfrac{\delta^{18}\mathrm{O} - 3.25}{3} \right) & \text{otherwise}
\end{cases}
$$

while the surface air temperature is equal to $\mathcal{T}_{\mathrm{AT}} = 2\mathcal{T}_{\mathrm{LO}} + 12.25$ during the Pleistocene. The relationship between the oxygen isotope ratio $\delta^{18}\mathrm{O}$, the deep ocean temperature $\mathcal{T}_{\mathrm{LO}}$ and the atmospheric temperature $\mathcal{T}_{\mathrm{AT}}$ is shown in Figure 8.16. Their results are very impressive, especially Figures 2c and 6, which compare the raw estimates of sea level and global mean surface temperature anomalies with the fitted estimates calculated with the equations. In Figure 8.16 we can see that the atmospheric temperature varies between $9.35°C$ and $14.37°C$. This illustrates how an impact of $\pm 3°C$ can have a large effect on the Earth's climate.

Figure 8.16: Relationship between $\delta^{18}\mathrm{O}$, $\mathcal{T}_{\mathrm{LO}}$ and $\mathcal{T}_{\mathrm{AT}}$ during the Pleistocene
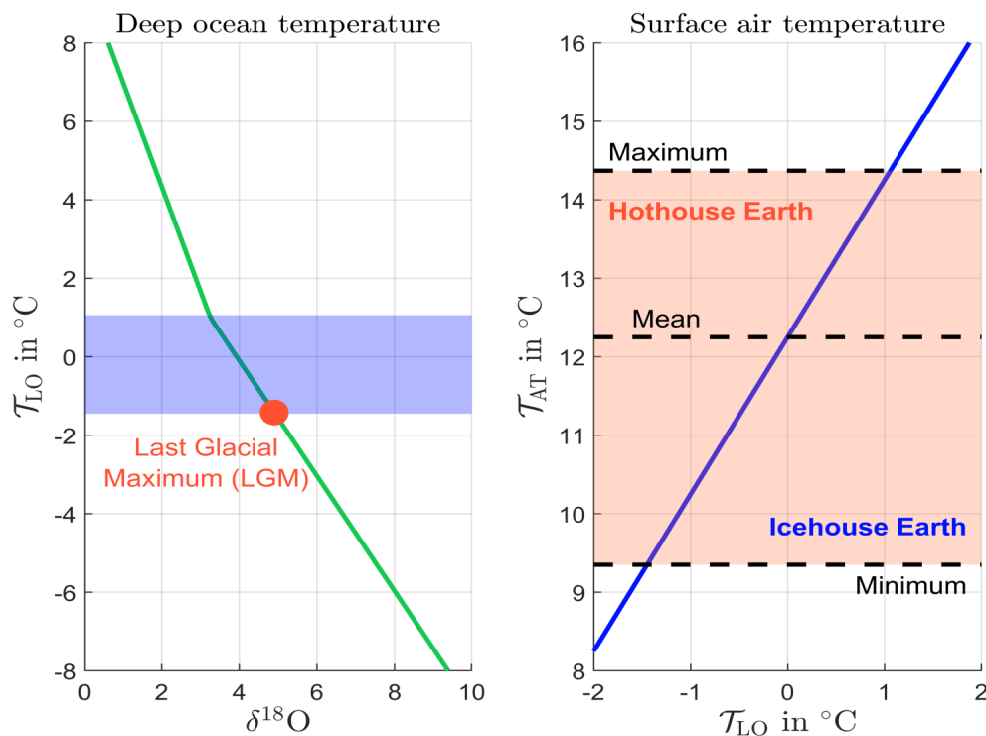
A comparison of Figure 8.14 with Figure 8.15 shows a high correlation between CO_2 concentration and temperature. But we also see that there is a high correlation when we look at the CH_4 concentration. If we consider a longer period, these correlations are lower. For example, we report the temporal evolution of atmospheric CO_2 reconstructed by [Foster](#) *et al.* (2017). Using a sample of $1\,500$ estimated atmospheric CO_2, they have performed a Loess regression. Their results for the last 420 million years are shown in Figure 8.17. Comparison with Figure 8.9 suggests that the evolution of CO_2 and $\Delta\mathcal{T}$ may be asynchrone in some periods, because CO_2 is not the only explanatory factor.

Figure 8.17: Evolution of the atmospheric CO_2 during the last 420 million years
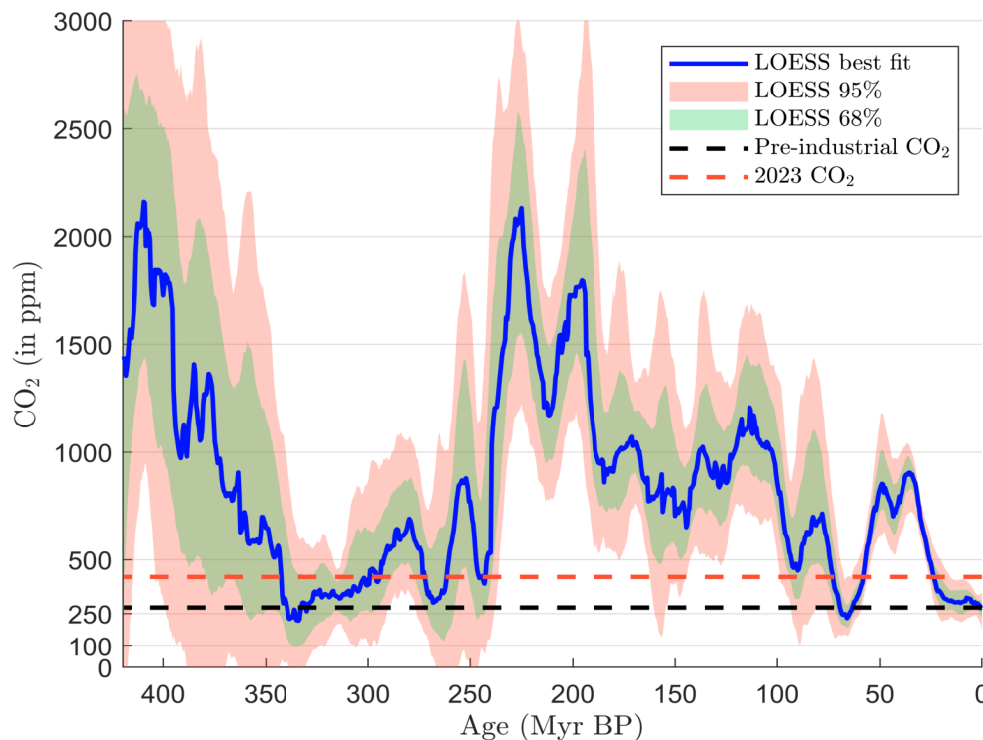
Source: Foster *et al.* (2017, Figure 1, page 3) & Supplementary Data 2 (www.nature.com/articles/ncomms14845).

Anthropogenic factors of climate change

The term "*anthropogenic*" is derived from the Greek words "*anthropos*" (human) and "*genesis*" (origin) and refers to anything caused or influenced by human activity. This includes both direct and indirect impacts, such as climate change, pollution, deforestation, and biodiversity loss. In this section, we examine the human impact on climate change since the beginning of the Industrial Revolution, which is assumed to have occurred in 1760. First, we examine the temperature anomaly, then we consider anthropogenic greenhouse gas emissions, and finally we discuss the planetary boundaries framework. Chapter 12 on page 633, on physical risks, complements this section and extends the analysis to other climate impacts such as flooding, sea-level rise, hurricanes, wildfires, heat waves, and agricultural productivity.

Temperature anomaly   We define the temperature anomaly at time t as the difference between the temperature at time t and the temperature for a reference period:

$$\Delta \mathcal{T}(t) = \mathcal{T}(t) - \mathcal{T}_{\mathfrak{B}ase}$$

where $\mathcal{T}_{\mathfrak{B}ase}$ is the reference temperature. It is generally the average of the temperature of the reference period:

$$\mathcal{T}_{\mathfrak{B}ase} = \frac{\sum_{j \in \mathfrak{B}ase} \mathcal{T}(j)}{n_{\mathfrak{B}ase}}$$

For example, the reference temperature can be the average temperature of the 20th century (from 1901 to 2000) or the pre-industrial period. In Figure 8.18, we report the global temperature anomaly

obtained with the HadCRUT5 database[31]. The reference period for the calculation of $\mathcal{T}_{\mathfrak{B}ase}$ is 1961-1990. The blue circle corresponds to the annual average, while the red line is the 10-year moving average. For each year, we have also provided the 95% confidence interval with a black vertical line. In 2023, the temperature anomaly was +1.04°C with a 95% confidence interval between 0.97°C and 1.10°C, while the 10-year moving average was 0.86°C. Since 1975 there has been a clear upward trend. If we estimate the linear trend between 1975 and 2023, we get the dashed black line. A simple projection gives the following forecasts: +1.43°C in 2050, +1.93°C in 2075 and +2.42°C in 2100.

Figure 8.18: Global average land-ocean temperature anomaly relative to 1961-1990 average

Source: Morice *et al.* (2021) & https://www.metoffice.gov.uk/hadobs/hadcrut5.
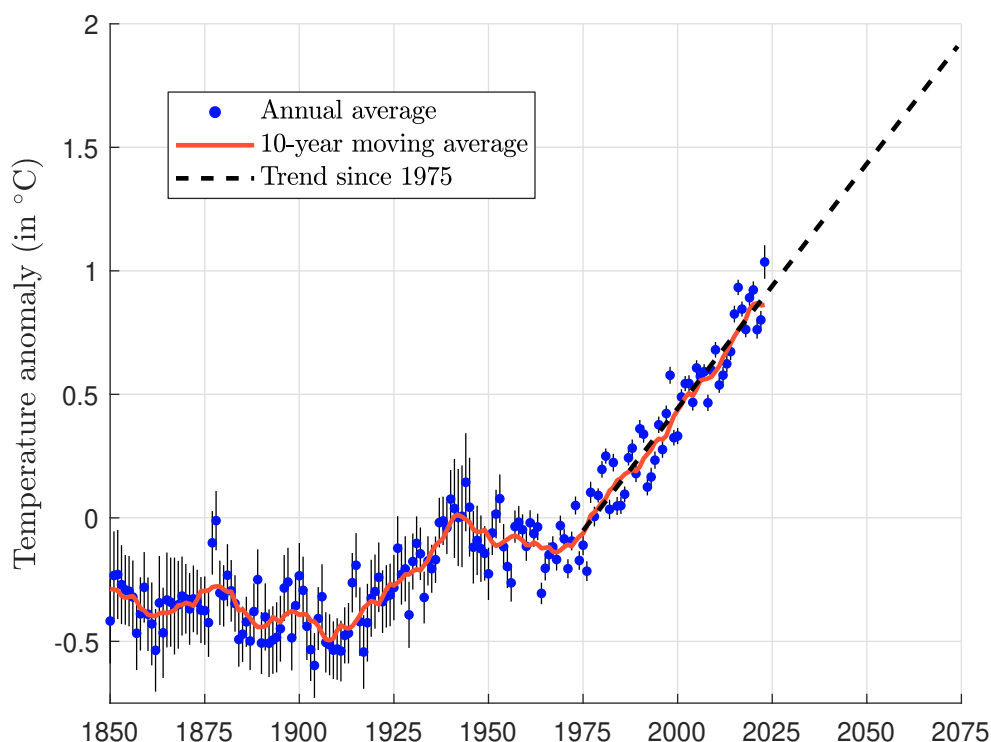
The previous analysis is a global picture. In fact, the temperature anomaly is calculated for a given location defined by a latitude, a longitude and an altitude. Let $\Delta \mathcal{T}_i(t)$ be the temperature anomaly for location i. The calculation of the temperature anomaly for a region is the weighted average of all locations belonging to the region grid:

$$\Delta \mathcal{T}_{\mathcal{R}egion}(t) = \sum_{i \in \mathcal{G}rid(\mathcal{R}egion)} w_i \cdot \Delta \mathcal{T}_i(t)$$

where w_i is the weight of location i. So we can calculate the temperature for a country, a continent and even a hemisphere. In Figure 8.19, we show the temperature anomaly for the northern and southern hemispheres, calculated by Morice *et al.* (2021), using a regular 5° latitude by 5° longitude grid of the Earth. In fact, the global temperature anomaly shown in Figure 8.18 is just the simple average of the northern and southern hemisphere temperature anomalies shown in Figure 8.19. We can see that the temperature in the northern hemisphere has risen faster than in the southern

---

[31]This database is produced and maintained by the Met Office Hadley Centre, which is also responsible for many other datasets that can be found at https://www.metoffice.gov.uk/hadobs.

Figure 8.19: Average land-ocean temperature anomaly in the northern and southern hemispheres relative to the 1961-1990 average

Source: Morice *et al.* (2021) & https://www.metoffice.gov.uk/hadobs/hadcrut5.
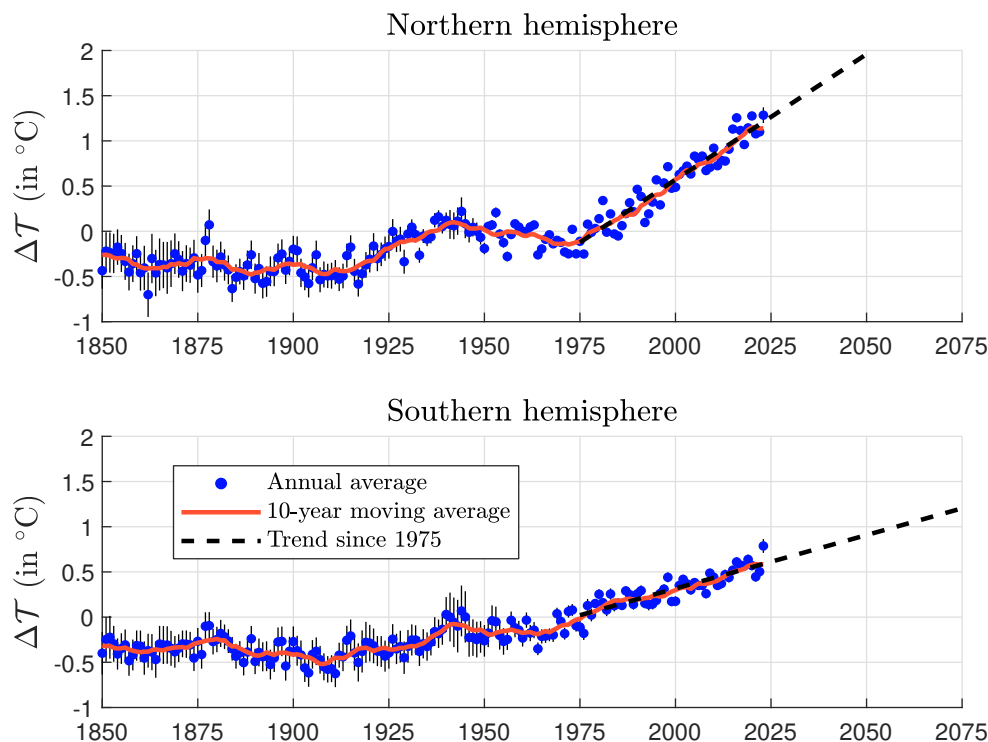
hemisphere. The slope of the trend since 1975 is $+2.78°C$ per century for the northern hemisphere, which is higher than the slope of $+1.18°C$ per century for the southern hemisphere. We therefore project a temperature anomaly of $+3.35°C$ degrees and $+1.50°C$ degrees respectively by 2100. All these results are summarized in Table 8.4.

Table 8.4: Linear projection of land-ocean temperature anomaly (in °C)

Year	HadCRUT5			NOAAGlobalTemp v5.1		
	Global	Northern	Southern	Global	Northern	Southern
2050	1.4336	1.9595	0.9078	1.4576	1.9894	0.9247
2075	1.9288	2.6540	1.2035	1.9185	2.6715	1.1642
2100	2.4239	3.3486	1.4992	2.3795	3.3536	1.4038
Slope	0.0198	0.0278	0.0118	0.0184	0.0273	0.0096

Source: https://www.ncei.noaa.gov/access/monitoring/global-temperature-anomalies/anomalies, https://www.metoffice.gov.uk/hadobs/hadcrut5 & Author's calculations.

The previous figures were obtained using the HadCRUT5 dataset, which is based on several historical datasets, models, and reanalyses. A reanalysis is a process that combines historical weather observations with a computer model to produce a consistent and complete record of past weather conditions. This is necessary because historical records are not available in a grid. Reanalysis then allows the creation of a homogeneous database of temperature records for a given grid and time period. However, the temperature records depend on the database. Below we list five land-ocean

gridded temperature datasets, their web addresses, and the scientific reference that explains how these temperature records were calculated:

1. Berkeley Earth, `berkeleyearth.org/data` (Rohde and Hausfather, 2020);

2. ERA5, `cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels` (Hersbach *et al.*, 2020);

3. HadCRUT5, `www.metoffice.gov.uk/hadobs/hadcrut5` (Morice *et al.*, 2021);

4. GISTEMP v4, `data.giss.nasa.gov/gistemp` (Lenssen *et al*, 2019);

5. NOAAGlobalTemp v5.1, `www.ncei.noaa.gov/products/land-based-station/noaa-global-temp` (Vose *et al.*, 2021).

A comparison of these datasets shows a high degree of consistency between the estimates, but we must be careful because it also shows that the results can be different, especially if we look at specific locations (e.g., a point on the grid or a country) rather than global locations.

Table 8.5: Linear projection of land and ocean temperature anomalies (in °C)

Year	Global		Northern		Southern	
	Land	Ocean	Land	Ocean	Land	Ocean
2050	2.4212	1.0238	2.8388	1.3482	1.4719	0.7972
2075	3.2386	1.3243	3.8176	1.8061	1.9222	0.9875
2100	4.0560	1.6247	4.7964	2.2641	2.3725	1.1779
Slope	0.0327	0.0120	0.0392	0.0183	0.0180	0.0076

Source: NOAAGlobalTemp v5.1 & Author's calculations.

We look at NOAAGlobalTemp v5.1 and display the land-ocean temperature anomaly in Figure 8.20. Note that the time series are very close to those plotted in Figures 8.18 and 8.19. The correlation between HadCRUT and NOAAGlobalTemp is 99.04% for the global surface, 99.27% for the northern hemisphere, and 97.74% for the southern hemisphere. Figure 8.20 also shows the decomposition of the land-ocean temperature into land and ocean temperatures. It can be seen that the temperature anomalies are more important for the land surface than for the ocean surface, implying that global warming is more important for the land than for the oceans. This is especially true in the northern hemisphere. We have calculated the trend[32] as before and reported the estimates in Tables 8.4 and 8.5. By the end of this century, the projected land temperature anomaly is +4.06°C for the world, 4.80°C for the northern hemisphere and 2.37°C for the southern hemisphere.

The discrepancy between the northern and southern hemispheres is even more pronounced when we look at countries. To illustrate this, we look at the temperature anomalies[33] on land as calculated by the FAO. As before, we calculate the trend with the temperature anomalies between 1975 and 2022 by country and project the trend to 2100. The results are shown in Figure 8.21. The range is from 0.64°C to 7.63°C. According to these results, Greenland and Russia[34] will be most affected, followed by Europe, Canada and the Middle East.

---

[32]The base period used by the NOAA model is different and is 1901-2000. To reconcile the two datasets, we need to adjust the HadCRUT5 data to have the same reference period as the NOAA data. This is equivalent to adding 0.10°C, 0.06°C, and 0.14°C degrees to the global surface, northern hemisphere, and southern hemisphere temperature anomalies of the HadCRUT5 data, respectively.

[33]We filter the database at `https://www.fao.org/faostat/en/#data/ET` using the item meteorological year and the element temperature change.

[34]For Russia and the countries of the former Soviet Union, historical data start in 1992 and may explain an overestimation of the trend.

Figure 8.20: Average temperature anomaly (land-ocean, land and ocean)

Source: Vose *et al.* (2021) &
https://www.ncei.noaa.gov/access/monitoring/global-temperature-anomalies/anomalies.
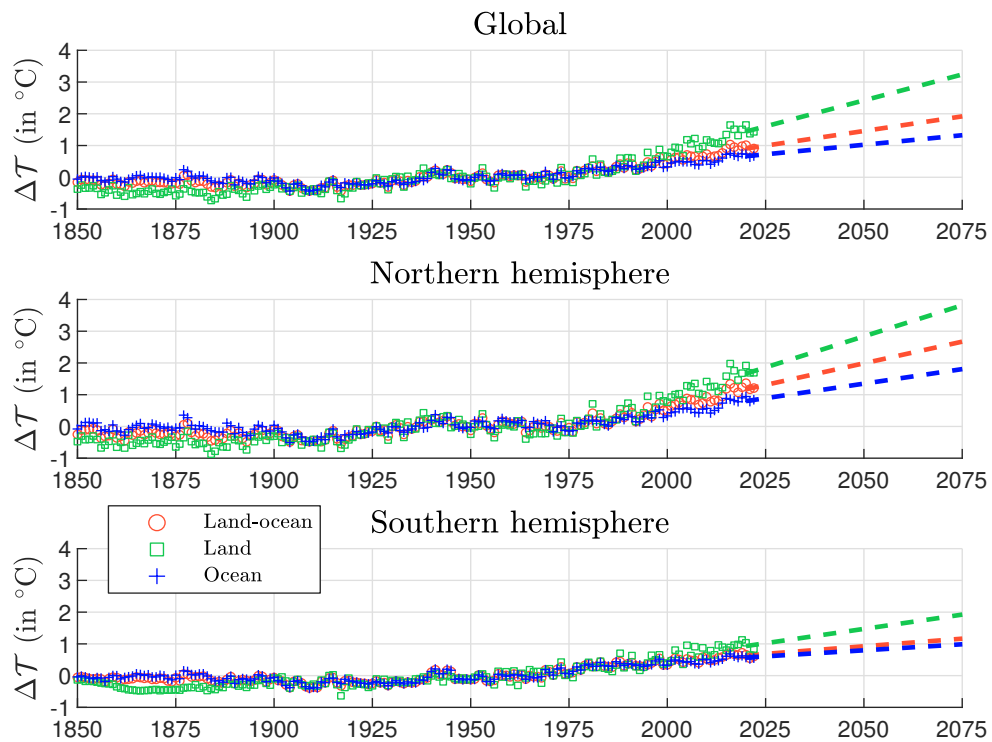
Figure 8.21: Projection of temperature anomaly by 2100 (in °C)
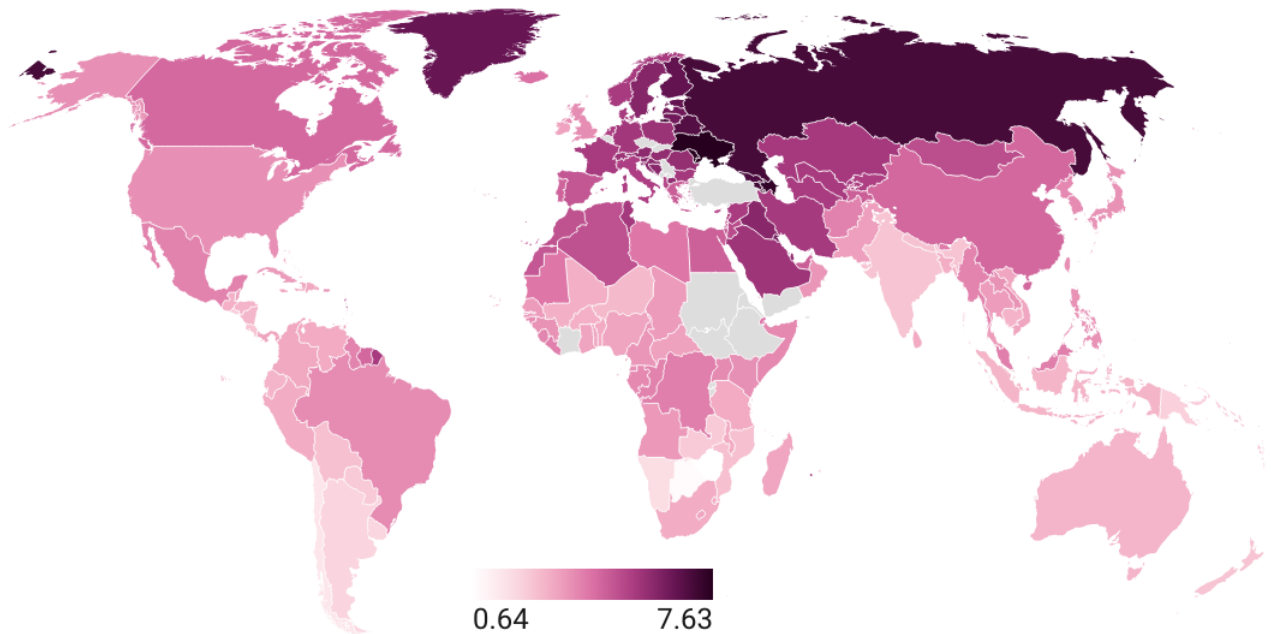
Source: https://www.fao.org/faostat/en/#data/ET & Author's calculations (created by Datawrapper).

Anthropogenic GHG emissions In order to study anthropogenic greenhouse gas emissions, we need to introduce some definitions and clarify some concepts of the global greenhouse gas budget. Let us first consider anthropogenic carbon dioxide emissions. According to Friedlingstein *et al.* (2022), the global carbon budget has five main components:

1. Fossil fuel combustion and oxidation from all energy and industrial processes, including cement production and carbonation;

2. Emissions from land-use change;

3. The growth rate of atmospheric CO_2 concentration;

4. The uptake of CO_2 by the oceans (ocean sink);

5. The uptake of CO_2 by the land (land sink).

These components are denoted $\mathcal{CE}_{\text{Industry}}$, $\mathcal{CE}_{\text{Land}}$, $\mathcal{CE}_{\text{AT}}$, $\mathcal{CS}_{\text{Ocean}}$, $\mathcal{CS}_{\text{Land}}$ respectively. From a theoretical point of view, we have the following identity:

$$\mathcal{CE}_{\text{AT}} = \underbrace{(\mathcal{CE}_{\text{Industry}} + \mathcal{CE}_{\text{Land}})}_{\text{Positive emissions}} - \underbrace{(\mathcal{CS}_{\text{Ocean}} + \mathcal{CS}_{\text{Land}})}_{\text{Negative emissions}}$$

As Friedlingstein *et al.* (2022) explained, this equality is not verified because of imperfect spatial and temporal data coverage, estimation errors and omission of minor terms. The authors therefore defined the estimated budget imbalance, which measures the mismatch between estimated emissions and estimated changes in the atmosphere, land and oceans:

$$\mathcal{CB}_{\text{Imbalance}} = \mathcal{CE}_{\text{Industry}} + \mathcal{CE}_{\text{Land}} - (\mathcal{CE}_{\text{AT}} + \mathcal{CS}_{\text{Ocean}} + \mathcal{CS}_{\text{Land}})$$

In the database provided by the authors[35], they split the carbon emissions $\mathcal{CE}_{\text{Industry}}$ into gross industrial emissions excluding carbonation minus cement carbonation sink[36]: $\mathcal{CE}_{\text{Industry}} = \mathcal{CE}_{\text{Industry}}^{\star} - \mathcal{CS}_{\text{Cement}}$.

In 2021, the authors estimate the following figures expressed in gigatonnes of carbon: $\mathcal{CE}_{\text{Industry}}^{\star} = 10.13$, $\mathcal{CE}_{\text{Land}} = 1.08$, $\mathcal{CE}_{\text{AT}} = 5.23$, $\mathcal{CS}_{\text{Ocean}} = 2.88$, $\mathcal{CS}_{\text{Land}} = 3.45$, $\mathcal{CS}_{\text{Cement}} = 0.23$, and $\mathcal{CB}_{\text{Imbalance}} = -0.58$. Expressed in gigatonnes of CO_2 we obtain: $\mathcal{CE}_{\text{Industry}}^{\star} = 37.12$, $\mathcal{CE}_{\text{Land}} = 3.94$, $\mathcal{CE}_{\text{AT}} = 19.14$, $\mathcal{CS}_{\text{Ocean}} = 10.55$, $\mathcal{CS}_{\text{Land}} = 12.64$, $\mathcal{CS}_{\text{Cement}} = 0.84$, and $\mathcal{CB}_{\text{Imbalance}} = -2.12$. Anthropogenic CO_2 emissions are therefore 36.28 $GtCO_2$ for industrial processes ($\mathcal{CE}_{\text{Industry}}$) and 3.94 $GtCO_2$ for land-use change. Since the total is 40.22 $GtCO_2$, 26.23% and 31.43% of the total anthropogenic CO_2 emissions have been absorbed by oceans and land, respectively, while 47.60% remain in the atmosphere. In Figure 8.22, we show the evolution of $\mathcal{CE}_{\text{Industry}}$ and $\mathcal{CE}_{\text{Land}}$. While carbon dioxide emissions from land-use change are relatively stable, carbon dioxide emissions from

---

[35]The full data set can be found at https://www.icos-cp.eu and https://globalcarbonbudgetdata.org. The main variables are available at https://globalcarbonbudget.org/wp-content/uploads/Global_Carbon_Budget_2022v1.0.xlsx. Carbon emissions are expressed in billion tonnes of carbon per year (GtC/yr). To convert these figures to billion tonnes of carbon dioxide ($GtCO_2$/yr), we multiply them by $\frac{44.011}{12.011}$ or 3.664. In fact, the atomic mass of carbon is 12.011 g/mol, and the atomic mass of oxygen is 15.999 g/mol. The molecular weight of CO_2 is therefore $12.011 + 2 \times 15.999 = 44.011$ g/mol.

[36]Cement carbonation sink is a natural process in which concrete absorbs carbon dioxide from the atmosphere. This process occurs when the calcium hydroxide in cement reacts with carbon dioxide in the air to form calcium carbonate:

$$\text{Ca(OH)}_2 + CO_2 \longrightarrow \text{CaCO}_3 + \text{H}_2\text{O}$$
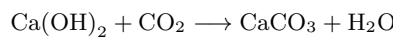
Figure 8.22: Annual CO_2 emissions (in $GtCO_2$)
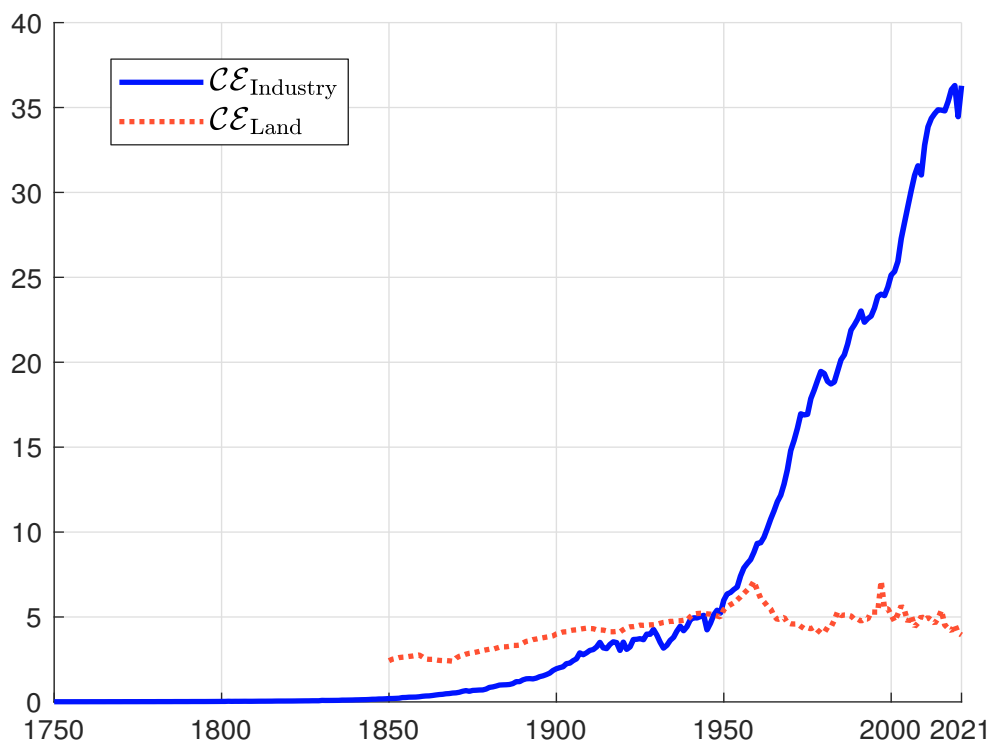
Figure 8.23: Cumulative CO_2 emissions and carbon sinks (in $GtCO_2$)
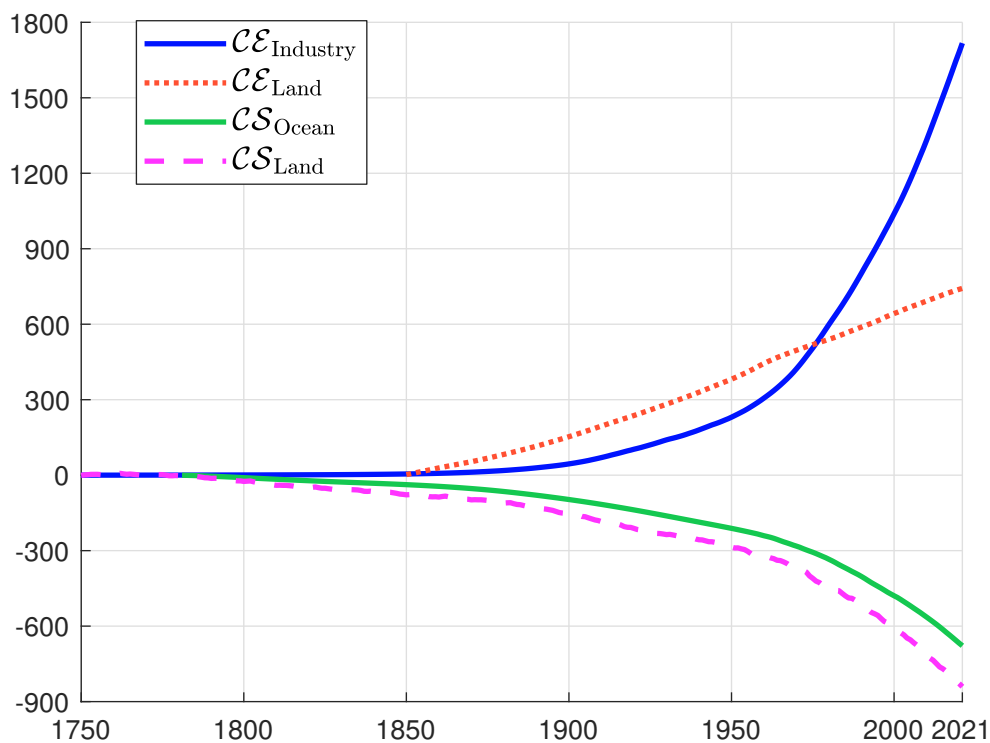
Figure 8.24: Cumulative CO_2 budget imbalance in atmosphere (in $GtCO_2$)
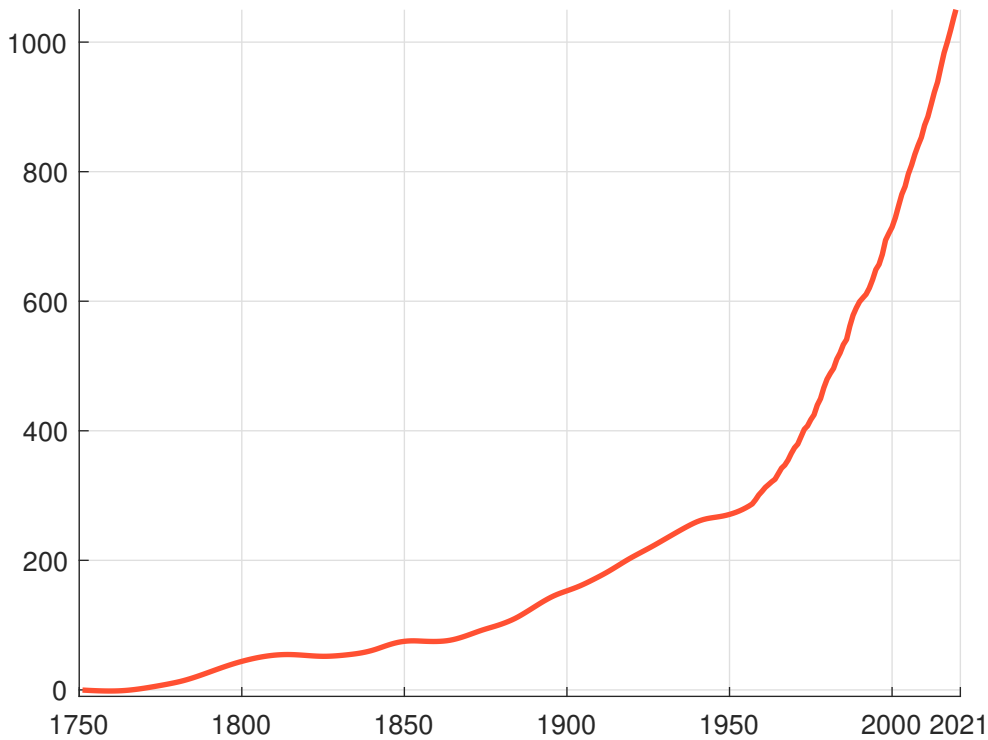
Source: Friedlingstein *et al.* (2022), https://globalcarbonbudget.org & Author's calculations.

industrial processes are growing faster since 1950. Looking at cumulative CO_2 emissions, we get Figure 8.23. We also show the negative emissions due to the two carbon sinks. From 1750, industrial processes and land-use change have emitted a total of 1 717 and 742 $GtCO_2$, respectively, and ocean and land sinks have absorbed a total of 678 and 841 $GtCO_2$, respectively. We can then calculate the cumulative CO_2 emissions remaining in the atmosphere in Figure 8.24. From 1750 to 2021, the net balance is 1 076 $GtCO_2$ of anthropogenic emissions. Friedlingstein *et al.* (2022) defined the airbone fraction as the ratio of the atmospheric CO_2 growth rate to total anthropogenic emissions. For the period 1750-2021, this ratio is equal to:

$$\text{AF} = \frac{\mathcal{CE}_{\text{AT}}}{\mathcal{CE}_{\text{Industry}} + \mathcal{CE}_{\text{Land}}} = \frac{1\,076}{1\,717 + 742} = 43.8\%$$

This means that 43.8% of anthropogenic emissions have not been absorbed by natural carbon sinks. The authors concluded that:

> "*The observed stability of the airborne fraction over the 1960-2020 period indicates that the ocean and land CO_2 sinks have on average been removing about 55% of the anthropogenic emissions.*" (Friedlingstein *et al.*, 2022, page 4834).

Let us now look at the contribution of different energy sources. Table 8.6 shows the breakdown of anthropogenic CO_2 emissions between coal, oil, gas, cement, flaring[37], and others. In 1850,

---

[37]Gas flaring is the burning of the natural gas associated with oil extraction. This gas is often burned because it is not economically feasible to collect and transport it. The top five flaring countries are Russia, Iraq, Iran, the United States, and Nigeria.

Table 8.6: Breakdown of anthropogenic CO_2 emission by energy source (in %)

	Energy	1850	1900	1950	1975	2000	2020
Annual CO_2 emissions	Coal	100.00	95.85	65.10	34.74	36.18	40.11
	Oil	0.00	3.55	26.79	47.64	40.42	33.23
	Gas	0.00	0.60	6.09	13.17	18.61	20.38
	Cement	0.00	0.00	1.13	1.99	2.87	4.33
	Flaring	0.00	0.00	0.81	2.18	1.05	1.12
	Other	0.00	0.00	0.08	0.29	0.86	0.83
Cumulative CO_2 emissions	Coal	100.00	97.58	85.24	63.49	50.00	46.29
	Oil	0.00	2.05	12.13	27.43	35.10	35.02
	Gas	0.00	0.37	2.15	6.88	11.72	14.52
	Cement	0.00	0.00	0.40	1.17	1.77	2.56
	Flaring	0.00	0.00	0.03	0.87	1.03	1.06
	Other	0.00	0.00	0.06	0.16	0.39	0.55

Source: Friedlingstein *et al.* (2022), `https://globalcarbonbudget.org` & Author's calculations.

Figure 8.25: Energy source breakdown of anthropogenic cumulative CO_2 (in %)
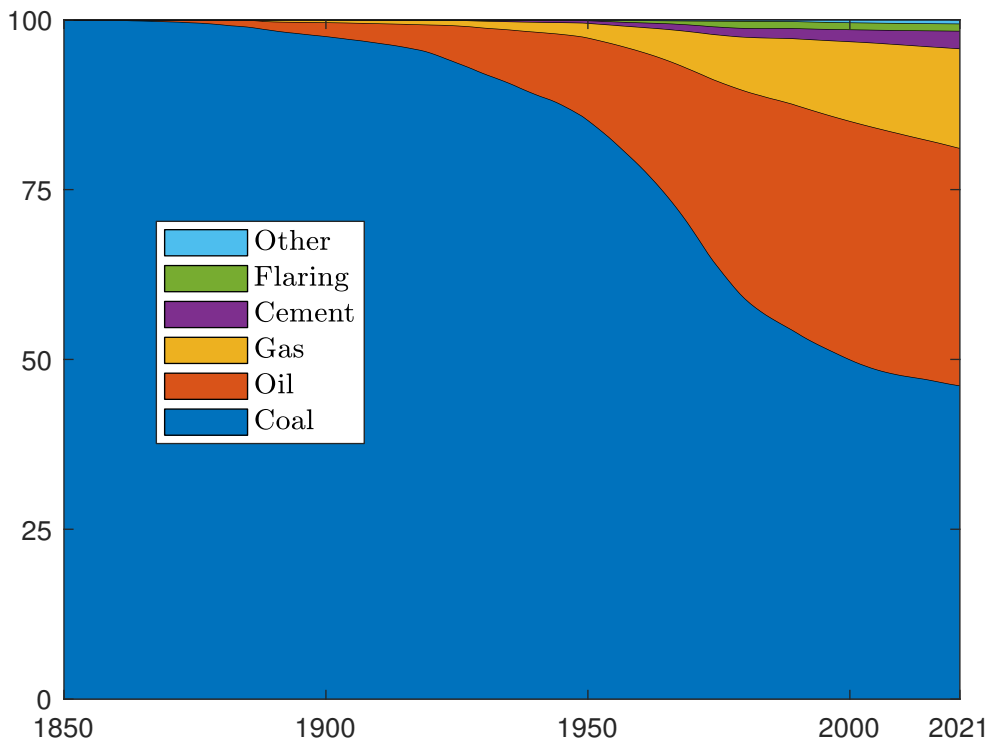
Source: Friedlingstein *et al.* (2022), `https://globalcarbonbudget.org` & Author's calculations.

100% of carbon emissions came from coal. Then we started burning oil, and later gas. Today, coal still accounts for 40% of current carbon emissions, while oil and gas account for 33% and 20%, respectively. The evolution of the distribution of cumulative CO_2 by energy source shows that the share of coal has decreased while the share of oil and gas has increased over the last century (Figure 8.25). In total, 46% of cumulative CO_2 emissions are due to coal, followed by oil and gas at 35% and 15%, respectively.

We also look at the breakdown of CO_2 emissions by region and country. The results are shown in Figures 8.26 to 8.29. Europe and North America have long been the main contributors, but Asia is now the world's largest CO_2 emitter. The country analysis shows that the top five emitters are China, India, Japan, Russia and the United States.

<div align="center">Box 8.6: Kaya identity</div>

In the early seventies, the IPAT equation was proposed to describe the factors of the human impact on the environment (Chertow, 2000):

$$I = P \times A \times T$$

where I is the level of human impact on climate, P is the population, A is the affluence[a] and T is the technology. In the 1990s, this equation was supplanted by the Kaya identity, which states:

$$\text{Anthropogenic } CO_2 \text{ emissions} = \text{Population} \times \frac{\text{GDP}}{\text{Population}} \times \frac{\text{Energy}}{\text{GDP}} \times \frac{CO_2 \text{ emissions}}{\text{Energy}}$$

Using the notations of Kaya and Yokobori (1997), this identity is generally expressed as[b]:

$$F = P \times \frac{G}{P} \times \frac{E}{G} \times \frac{F}{E}$$

Therefore, the key drivers of anthropogenic CO_2 emissions include four main factors: the population (demographics), the GDP per capita (economics), the energy intensity of the GDP (engineering) and the carbon intensity (physics). Over the last century, we have observed that the first two factors have increased while the third factor has decreased. Looking into the future, we can expect these trends to continue. Decarbonizing the economy therefore means drastically reducing the carbon intensity of the energy supply. For example, if $P \nearrow 30\%$, $G/P \nearrow 20\%$, $E/G \searrow 15\%$, the carbon intensity must decrease by 32.6% to not increase the level of carbon emissions and decrease by 62.3% to reduce the carbon emissions by two.

---

[a]The affluence represents the average consumption of each person in the population. A common proxy for measuring affluence is the GDP per capita or the wealth per capita.

[b]See for instance Davis and Caldeira (2010).

The Kaya identity was popularised by Raupach *et al.* (2007). It is a mathematical equation for understanding the factors that explain CO_2 emissions (Box 8.6). According to this identity, the four drivers of emissions are population growth, GDP per capita, energy intensity (quantity of energy consumed per unit of GDP) and carbon intensity (amount of CO_2 emitted per unit of energy consumed). In Figure 8.30, we plot the CO_2 per unit energy (kgCO_2 per kWh), the energy per GDP (kWh per dollar), the CO_2 per GDP (kgCO_2 per dollar of GDP) and the GDP per capita (1000

Figure 8.26: Share of CO_2 emissions by region (in % of total)
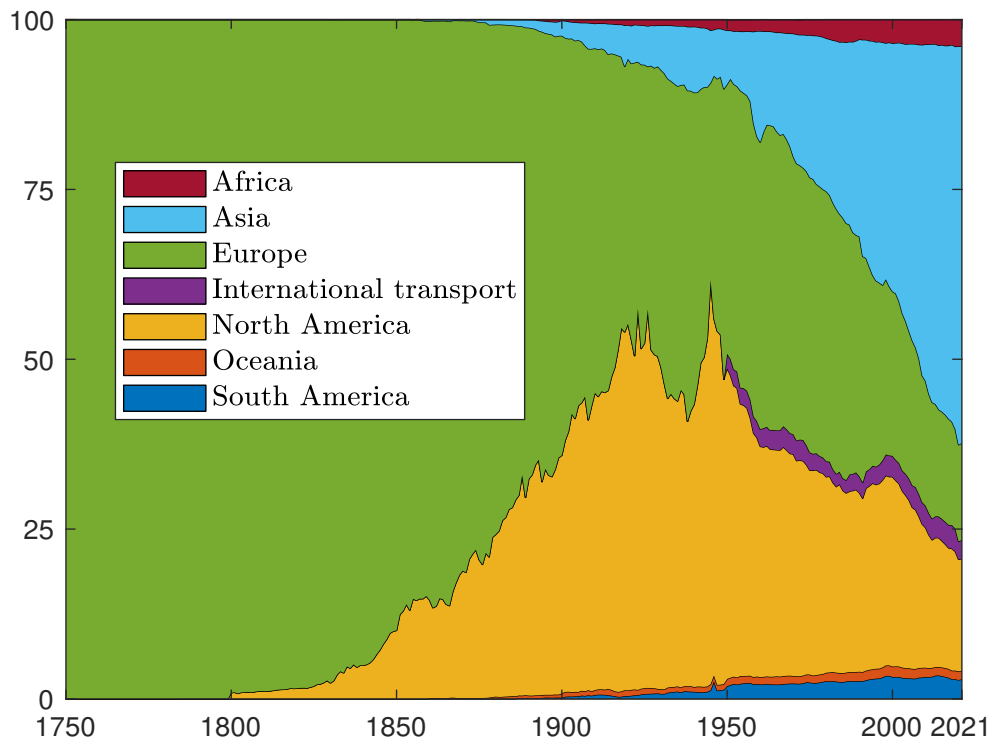
Figure 8.27: Share of cumulative CO_2 emissions by region (in % of total)
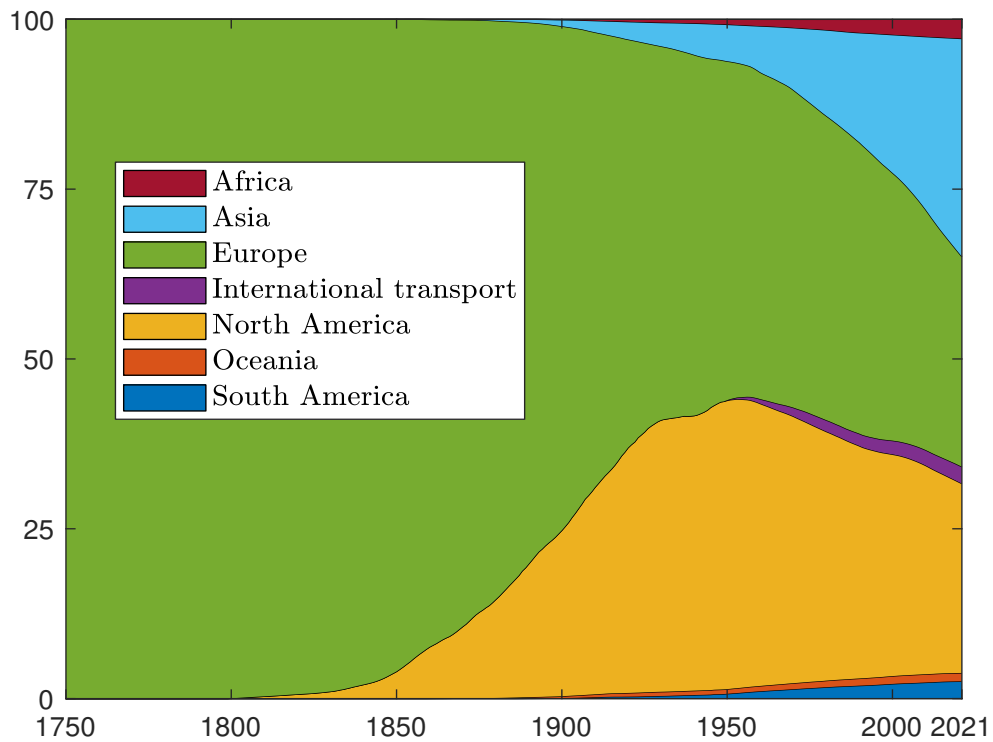
Figure 8.28: Share of CO$_2$ emissions by region (in % of total)

Source: Friedlingstein *et al.* (2022), https://globalcarbonbudget.org & https://github.com/owid/co2-data.
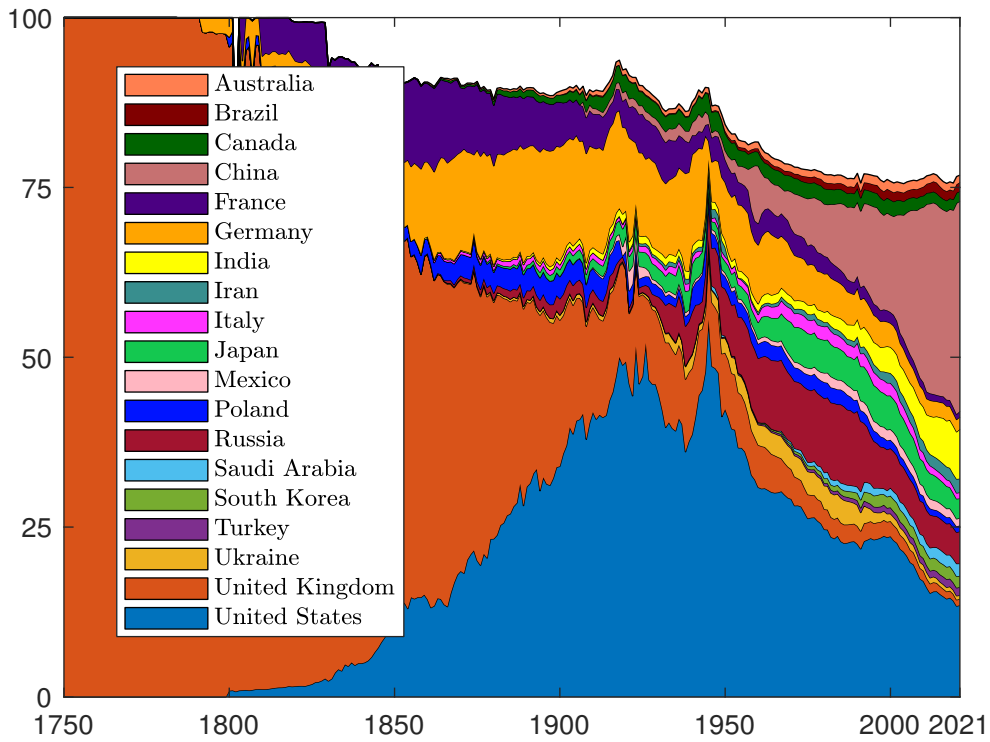
Figure 8.29: Share of cumulative CO$_2$ emissions by country (in % of total)

Source: Friedlingstein *et al.* (2022), https://globalcarbonbudget.org & https://github.com/owid/co2-data.
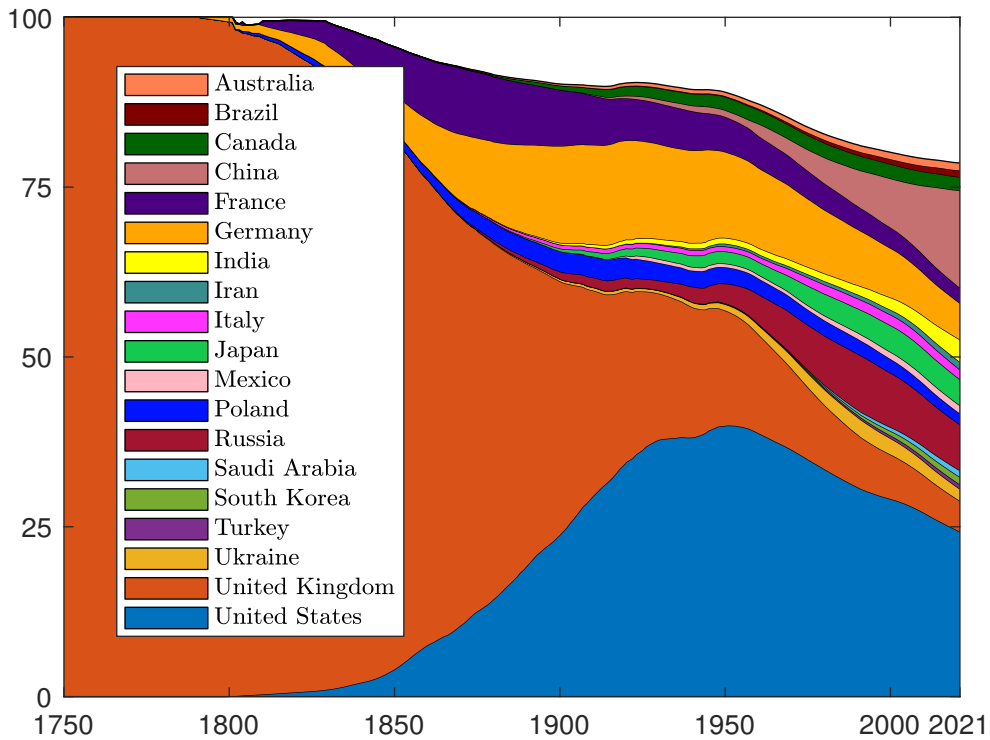
Figure 8.30: Key drivers of the Kaya identity
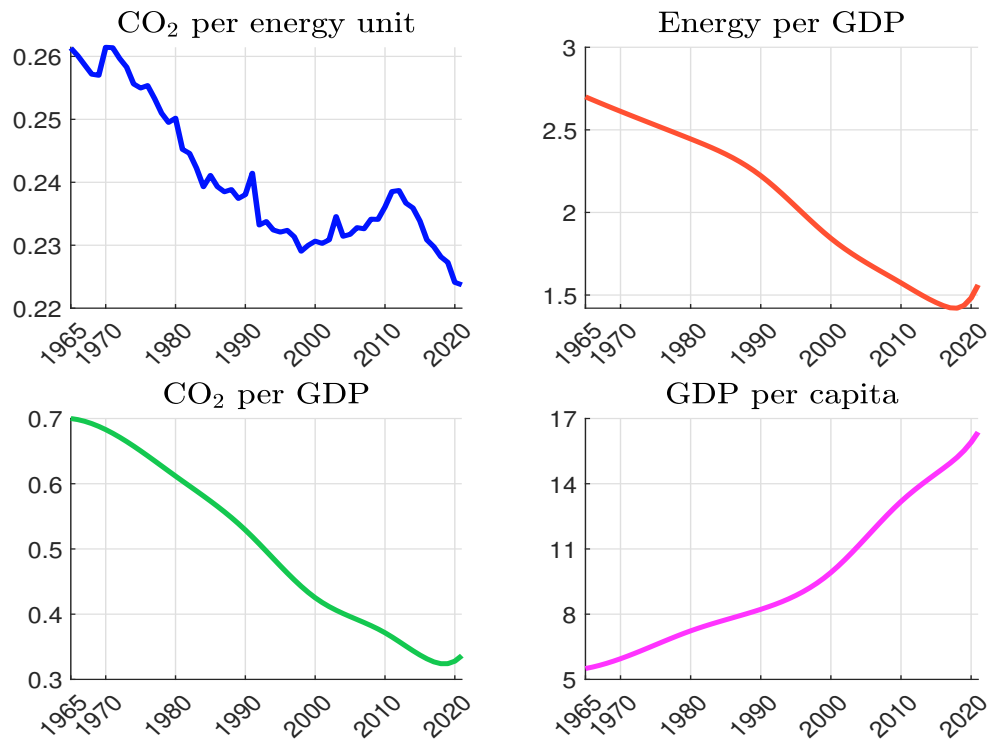
Figure 8.31: CO_2 emissions per capita (in tCO_2 per person)
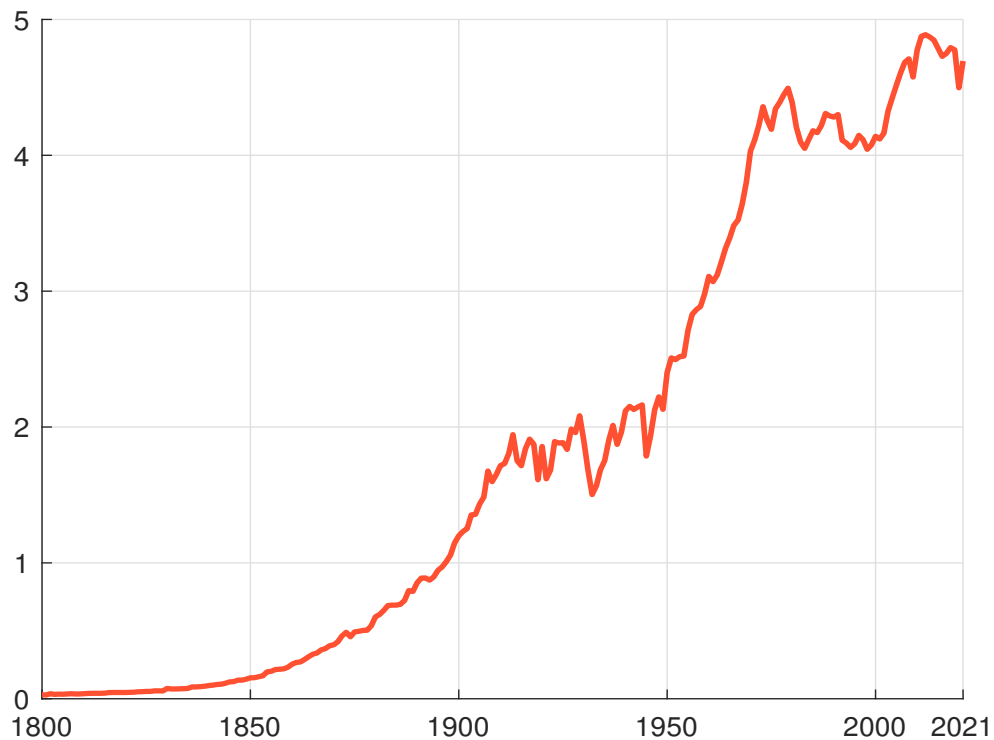
dollars per person). The first three measures fall, but GDP per capita does not. As people get richer, they consume more goods and services and emit more CO_2. This is the main driver of CO_2 emissions. This can be summarised by the trend in CO_2 per capita in Figure 8.31. We observe five periods: a continuous increase between 1800 and 1913, a flattening out between 1914 and 1945, a new increase between 1946 and 1979, a slight decrease between 1980 and 1999, and an upward trend again since 2000. Today, we produce about 5 tonnes of CO_2 per person.

Figure 8.32: GHG emissions (in GtCO$_2$e)
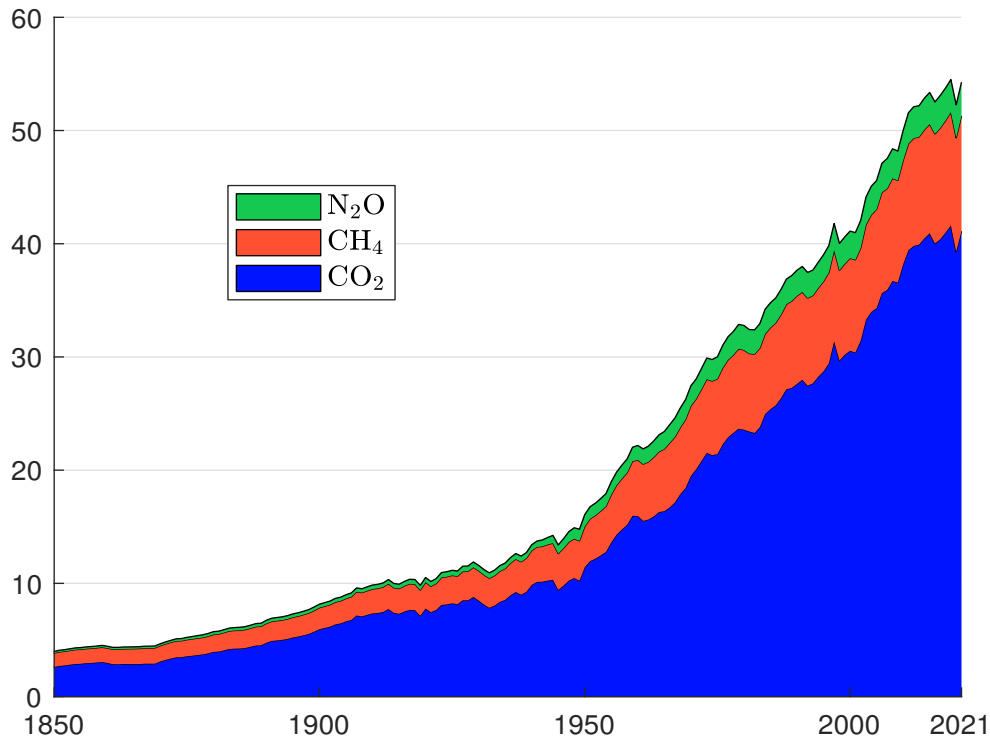
Source: Jones *et al.* (2023), https://zenodo.org/records/7636699#.ZFCy4exBweZ & Author's calculations.

The above analysis can be extended to other GHG emissions such as methane and nitrous oxide emissions. For each gas, it is important to specify the scope (e.g., including or excluding land-use change and forestry). Using the data set calculated by Jones *et al.* (2023), we plot the evolution of greenhouse gas emissions[38] since 1850 and their decomposition into CO_2, CH_4 and N_2O in Figure 8.32. GHG emissions include both industrial and land-use change sources. For the year 2021, Jones *et al.* (2023) estimated the following figures: 41.12 GtCO$_2$ for carbon dioxide, 10.18 GtCO$_2$e for methane and 2.97 GtCO$_2$e for nitrous oxide. The repartition is then 75.8%, 18.8% and 5.5%.

Table 8.7: 2021 greenhouse gas emissions (in GtCO$_2$e)

	CH_4			CO_2			N_2O	
$\mathcal{CE}_{\text{Industry}}$	$\mathcal{CE}_{\text{Land}}$	$\mathcal{CE}_{\text{Total}}$	$\mathcal{CE}_{\text{Industry}}$	$\mathcal{CE}_{\text{Land}}$	$\mathcal{CE}_{\text{Total}}$	$\mathcal{CE}_{\text{Industry}}$	$\mathcal{CE}_{\text{Land}}$	$\mathcal{CE}_{\text{Total}}$
6.23	3.95	10.18	37.11	4.00	41.12	0.79	2.18	2.97
(61.2%)	(38.8%)	(18.8%)	(90.3%)	(9.7%)	(75.8%)	(26.7%)	(73.3%)	(5.5%)

Source: Jones *et al.* (2023), https://zenodo.org/records/7636699#.ZFCy4exBweZ & Author's calculations.

---

[38]To calculate the CO_2 equivalent of CH_4 and N_2O, we use a global warming potential of 27.9 and 273 (see Section 9.1.1 on page 470.).

Planetary boundaries   In addition to climate change, the impact of human activity on the Earth system affects many other issues. They can be summarised in terms of planetary boundaries, a concept developed by Rockström *et al.* (2009). We can define planetary boundaries as a normative framework for setting limits to the impact of human activities on the Earth system. Beyond these boundaries, the environment may no longer be able to regulate itself and the Earth system may leave the Holocene period of stability during which human society developed. The planetary boundaries framework has been updated twice, in 2015 (Steffen *et al.*, 2015) and 2023 (Richardson *et al.*, 2023). We report the details of the current framework in Table 8.8 and Box 8.7. For each critical Earth process system, Richardson *et al.* (2023) defined a set of control variables and associated metrics. The baseline metric m_{1750} corresponds to the pre-industrial Holocene base value. The planetary boundary is defined by a threshold m_{Boundary}, while m_{Upper} indicates the upper end of the increasing risk. Finally, m_{2023} is the current value of the control variable. If we assume that the risk increases with the metric, we have $m_{1750} \leq m_{\text{Boundary}} \leq m_{\text{Upper}}$. If $m_{2023} \geq m_{\text{Boundary}}$, the boundary is exceeded. If $m_{2023} \geq m_{\text{Upper}}$, this indicates a high risk. According to Richardson *et al.* (2023), six of the nine planetary boundaries have been crossed (Figure 8.33). They are climate change, biodiversity loss, biogeochemical flows, land-use change, freshwater change, and novel entities. The ocean acidification planetary boundary can be reached in a few years, while atmospheric aerosol loading and stratospheric ozone depletion are not at risk.

Figure 8.33: GHG emissions (in GtCO$_2$e)
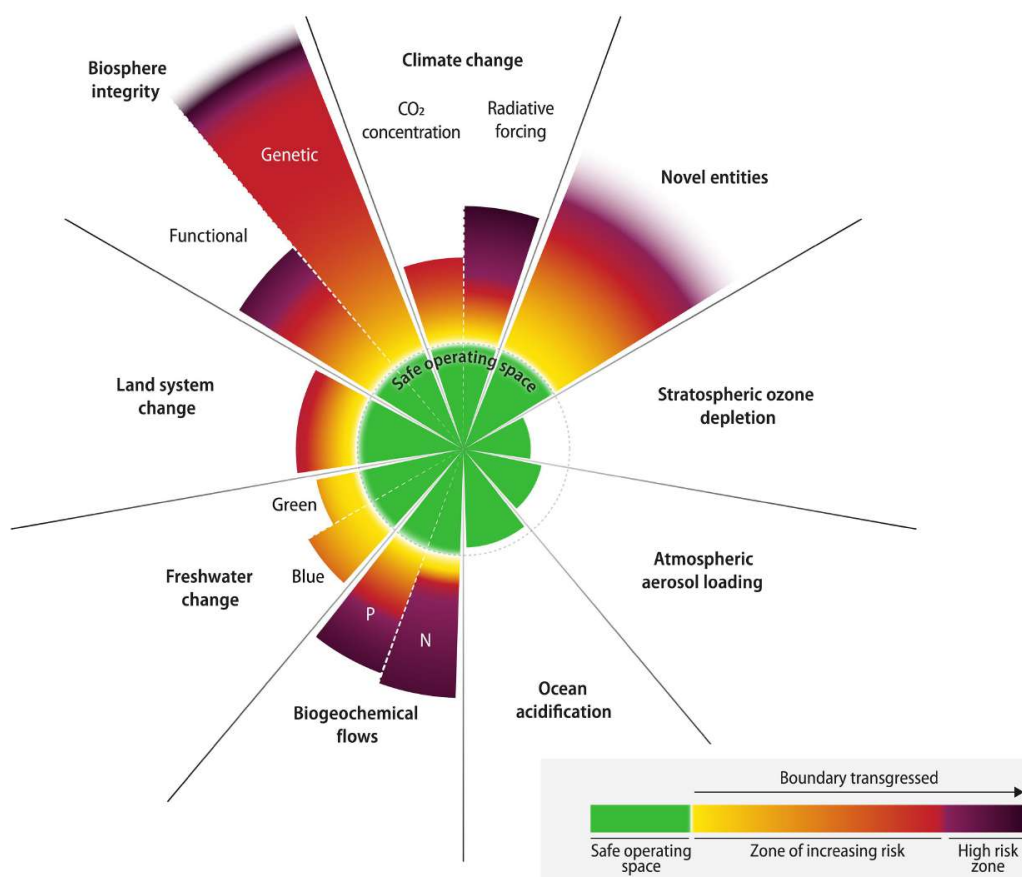
Source: Richardson *et al.* (2023, Figure 1, page 4).

Table 8.8: Current status of planetary boundaries

No	Earth process system	Control variable	m_{1750}	$m_{Boundary}$	m_{upper}	m_{2023}	Crossed
1	Climate change	Atmospheric CO_2 concentration (ppm)	280	350	450	417	✓
		Atmospheric radiative forcing (W/m^2)	0	1.0	1.5	2.91	✓
2	Biodiversity loss	Genetic diversity (E/MSY)	1	10	100	>100	✓
		Functional integrity (% HANPP)	1.9	10	20	30	✓
3	Stratospheric ozone depletion	Stratospheric O_3 concentration (DU)	290	276	261	284.6	
4	Ocean acidification	Carbonate ion concentration (Ω_{arg})	3.44	2.752	2.75	2.8	
5	Biogeochemical flows	Phosphate flow (TgP/yr) – global	0	11	100	22.6	✓
		Phosphate flow (TgP/yr) – regional	0	6.2	11.2	17.5	✓
		Nitrogen flow (TgN/yr)	0	62	82	190	✓
6	Land-use change	Area of forested land (%) – global	100	75	54	60	✓
		% area remaining – tropical	100	85	60	58.6	✓
		% area remaining – temperate	100	50	30	41.1	✓
		% area remaining – boreal	100	85	60	63.5	✓
7	Freshwater change	Blue water (%)	9.4	10.2	50	18.2	✓
		Green water (%)	9.8	11.1	50	15.8	✓
8	Atmospheric aerosol loading	Inter-hemispheric difference (AOD)	0.03	0.10	0.25	0.076	
9	Novel entities	Synthetic chemicals (%)	0	0	NA	>0	✓

Source: Richardson et al. (2023, Table 1, pages 4-5).

Box 8.7: Planetary boundaries defined by Richardson *et al.* (2023, Table 1, pages 4-5)

The planetary boundaries for climate change are defined by two control variables. First, the atmospheric CO_2 concentration $\mathcal{CC}_{AT}$ must be less than 280 ppm. Second, the total anthropogenic radiative forcing at the top of the atmosphere must be less than $+1.0\,\mathrm{W/m^2}$. For biodiversity loss, there are again two targets. The genetic diversity target implies less than 10 extinctions per million species-years (E/MSY). Functional integrity measures the amount of energy produced by photosynthesis (net primary production or NNP). The control variable is herbivore-adjusted net primary production (HANPP), which is the amount of energy available to herbivores after plants have been consumed. The variation in HANPP must be less than 10% of the pre-industrial Holocene NPP. Stratospheric ozone depletion is measured by the stratospheric O_3 concentration $\mathcal{CC}_{O_3}$ or the total amount of ozone in a vertical column through the atmosphere (Dobson units or DU). The limit is $\mathcal{CC}_{O_3} \leq 276$ DU, which is 5% less than the pre-industrial level. For ocean acidification, the authors use the carbonate ion concentration $\left[CO_3^{2-}\right]$, which is a measure of the amount of dissolved carbonate ions in seawater. The control variable is the average global surface ocean saturation state with respect to aragonite $\Omega_{\mathrm{arag}} = \dfrac{\left[CO_3^{2-}\right]}{\kappa_{sp}}$ where κ_{sp} is the solubility product constant of aragonite. The limit Ω_{arag} is set to 80% of the pre-industrial Holocene baseline ($\Omega_{\mathrm{arag}} = 3.44$). Biogeochemical flows reflect anthropogenic disturbances to the phosphorus (P) and nitrogen (N) cycles. The P cycle is assessed both globally (P flow from freshwater systems to the ocean) and regionally (P flow from fertilizers to erodible soils). The global limit is set at 11 Tg phosphorus per year, while the regional limit is set at 6.2 Tg phosphorus per year. Perturbation of the N cycle is measured by industrial and intentional nitrogen fixation. Its limit is set at 62 Tg nitrogen per year. The sixth Earth system process is assessed with both global (*e.g.*, deforestation, agriculture and urbanization) and biome-specific (*e.g.*, transformation of forests, grasslands, savannas, and deserts) land system changes. The global figure is calculated as the area of forest as a percentage of the original forest cover, while the biome figure is the area of forest as a percentage of the potential forest (% remaining area). The limits are 75% for the global figure, 85% for the tropical biome, 50% for the temperate biome and 85% for the boreal biome. Freshwater change is divided into blue water and green water. Blue water refers to surface water available for human use, while green water refers to water stored in the soil and vegetation (*i.e.* water available to plants). The blue water control variable is the 95th percentile of global land area with deviations greater than the pre-industrial period, and its limit is set at 10.2%. The green water metric is the percentage of land area with deviations from pre-industrial variability. The pre-industrial Holocene baseline is 9.8%, and the limit is 11.1%. The eighth Earth system process is the atmospheric aerosol loading, the control variable is the inter-hemispheric difference in aerosol optical depth[a] (AOD), and the limit is 0.1 mean annual inter-hemispheric difference. The final planetary boundary concerns novel entities such as industrial chemicals, pesticides, microplastics, nanoparticles, nuclear waste and genetically modified organisms (GMO). The associated control variable is the percentage of synthetic chemicals released into the environment without adequate safety testing. This ninth planetary boundary is not well defined, because it is difficult to collect data on.

---

[a]AOD measures the amount of light absorbed by aerosols in the atmosphere. The inter-hemispheric difference in AOD is the difference between the average AOD in the Northern Hemisphere and the average AOD in the Southern Hemisphere.

8.1.3 The physics of climate change

In this section, energy and forcing are two interchangeable terms, meaning that $F_{\text{solar}} := \boldsymbol{\mathcal{E}}_{\text{solar}}$, $F_{\text{thermal}} := \boldsymbol{\mathcal{E}}_{\text{infrared}}$, etc. Moreover, we have made extensive use of Hartmann's Handbook on Climate Physics, particularly chapters 2, 3 and 10 (Hartmann, 2016).

Simple energy balance models

The Earth's temperature is closely related to its energy balance, which is the balance between the energy it receives from the Sun and the energy it emits:

> "*Temperature [...] is a measure of the energy contained in the movement of molecules. Therefore, to understand how the temperature is maintained, one must consider the energy balance that is formally stated in the first law of thermodynamics. The basic global energy balance of Earth is between energy coming from the Sun and energy returned to space by Earth's radiative emission. The generation of energy in the interior of Earth has a negligible influence on its energy budget.*" (Hartmann, 2016, page 25).

The first law of thermodynamics, also known as the law of conservation of energy, is a principle of physics which states that energy cannot be created or destroyed, it can only be transferred or converted from one form to another. This means that the total amount of energy in a closed system is conserved.

Basics To calculate the Earth's temperature, we need three basic tools of physics. They are the total solar irradiance, the Stefan-Boltzmann law and the emission temperature of a planet.

Total solar irradiance According to Hartmann (2016, page 29), the total amount of electromagnetic energy emitted by the Sun, also called the solar luminosity is $L_{\odot} = 3.828 \times 10^{26}$ watts. Using the first law of thermodynamics, the amount of energy emitted by any sphere with the Sun at its centre should be equal to the total energy flux from the Sun. Total solar irradiance (TSI) is defined as:

$$S_d = \frac{L_{\odot}}{4\pi d^2} \tag{8.3}$$

where d is the distance of the sphere from the Sun in meters. For the Earth, the distance is between 147.1 and 152.1 million kilometers because the Earth's orbit around the Sun is not a perfect circle. Using a mean value of 149.6 million kilometers, we get:

$$S_0 = \frac{3.828 \times 10^{26}}{4\pi \left(149.6 \times 10^9\right)^2} = 1372.11 \text{ W/m}^2$$

A direct measurement by astrophysicists gives 1368 W/m^2 and we use this number in the sequel.

Stefan-Boltzmann law The Stefan-Boltzmann law describes the relationship between the total amount of radiation $\boldsymbol{\mathcal{E}}$ emitted by a body and its temperature $\mathcal{T}$:

$$\boldsymbol{\mathcal{E}} = \varepsilon \sigma \mathcal{T}^4 \tag{8.4}$$

where $\varepsilon \in [0, 1]$ is the emissivity of the body and $\sigma = 5.67 \times 10^{-8} \text{ W/m}^2 \text{K}^{-4}$ is the Stefan-Boltzmann constant. For an ideal black body, we have $\varepsilon = 1$ and Equation (8.4) is related to the

Planck distribution[39] B_ν (or the intensity of the black body radiation):

$$
\begin{aligned}
\mathcal{E} &= \pi \int_0^\infty B_\nu\left(\nu, \mathcal{T}\right) \, \mathrm{d}\nu \\
&= \frac{2\pi^5 k^4}{15 c^2 h^3} \mathcal{T}^4 \\
&= \sigma \mathcal{T}^4
\end{aligned}
$$

where h is Planck's constant, c is the speed of light in a vacuum and k is Boltzmann's constant.

Box 8.8: Planck radiation law and spectral density of electromagnetic radiation

In physics, Planck's law describes the spectral distribution of electromagnetic radiation emitted by a black body in thermal equilibrium at a given temperature. The expression for the spectral density function is:

$$
B_\nu\left(\nu, \mathcal{T}\right) = \frac{2h\nu^3}{c^2} \frac{1}{\exp\left(\dfrac{h\nu}{k\mathcal{T}}\right) - 1}
$$

where $h = 6.62607015 \times 10^{-34} \, \mathrm{J\,Hz^{-1}}$ is Planck's constant, $c = 299\,792\,458 \, \mathrm{m\,s^{-1}}$ is the speed of light in a vacuum, $k = 1.380649 \times 10^{-23} \, \mathrm{J\,K^{-1}}$ is Boltzmann's constant, $\mathcal{T}$ is the temperature measured in Kelvin, and ν is the frequency in hertz. Alternatively, the law can be written in terms of the wavelength λ expressed in meters:

$$
\lambda = \frac{c}{\nu} \qquad \left(\sim \frac{\mathrm{m\,s^{-1}}}{\mathrm{Hz}} = \mathrm{m}\right)
$$

We deduce that:

$$
B_\lambda\left(\lambda, \mathcal{T}\right) = \frac{2hc^2}{\lambda^5} \frac{1}{\exp\left(\dfrac{hc}{\lambda k \mathcal{T}}\right) - 1}
$$

Since we have $\dfrac{\mathrm{J\,Hz^{-1}m\,s^{-1}}}{\mathrm{m\,J\,K^{-1}K}} = 1$ and $\dfrac{\mathrm{J\,Hz^{-1}\left(m\,s^{-1}\right)^2}}{\mathrm{m^5}} = \mathrm{J\,s^{-1}m^{-2}m^{-1}} = \mathrm{W\,m^{-2}m^{-1}}$, the unit of $B_\lambda\left(\lambda, \mathcal{T}\right)$ is $\mathrm{W/m^2\,m^{-1}}$. In Figure 8.34, we show the spectral density for two values of temperature. We also show the visible spectrum, which is the band of the electromagnetic spectrum that is visible to the human eye. The electromagnetic spectrum also includes ultraviolet and infrared wavelengths. We use the following classification: UV (less than 380 nm), visible (380 nm to 780 nm), and infrared (greater than 780 nm). The temperature has a high impact on emitted radiations, since the spectral density is completly different for $\mathcal{T} = 5\,000\,\mathrm{K}$ and $\mathcal{T} = 3\,000\,\mathrm{K}$. Now let us illustrate the difference between solar and terrestrial radiation. In Figure 8.35, we scale the spectral density of the Earth's surface by a factor of 2.6×10^6 so that the two peaks are at the same level. Note that the spectrum of sunlight is shortwave radiation, while the spectrum of the Earth's surface is longwave radiation. We also see that the atmosphere is mainly transparent to a large part of the visible spectrum.

---

[39]The Planck distribution describes the probability of a photon being emitted with a particular energy.

Figure 8.34: Spectral density function $B_\lambda\left(\lambda, \mathcal{T}\right)$ (in $10^{12}\,\mathrm{W/m^2\,m^{-1}}$)
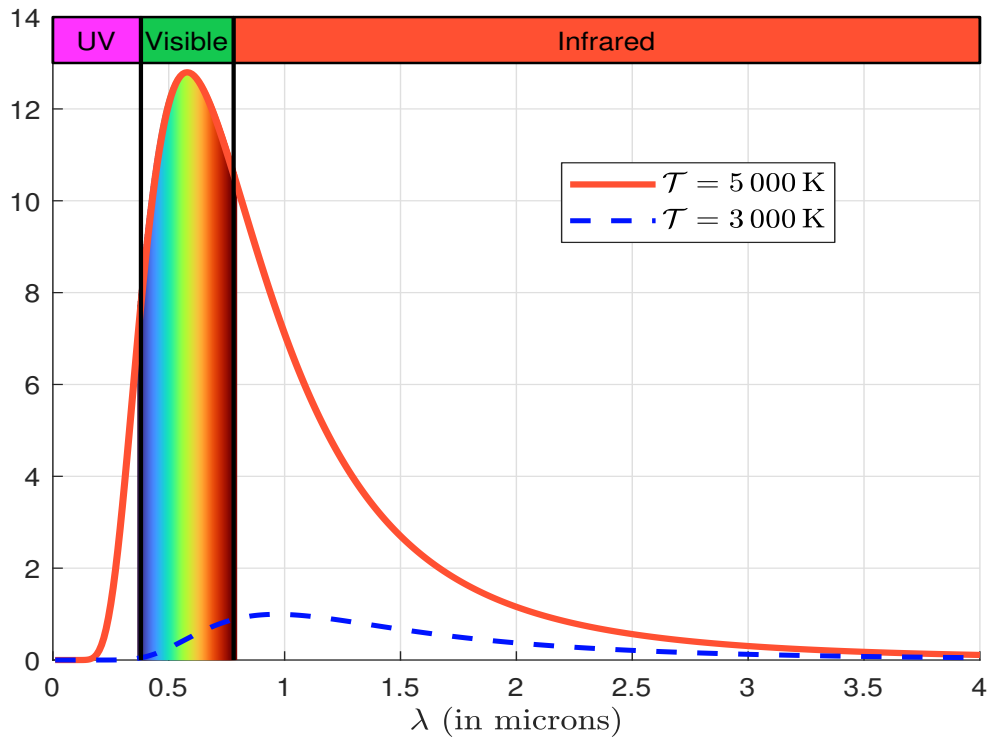
Figure 8.35: Comparison of the radiation spectra of sunlight and the Earth's surface (in $10^{12}\,\mathrm{W/m^2\,m^{-1}}$)
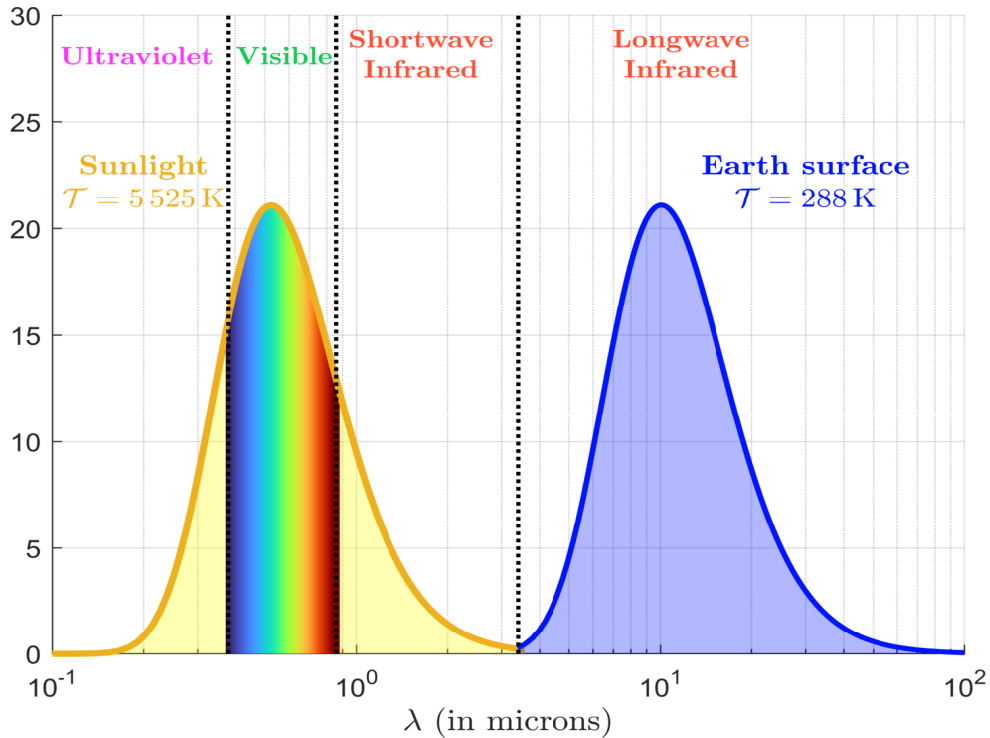
Effective temperature of stars   Since the radius of the Sun $R_{\odot}$ is about $696\,342$ kilometers, the solar irradiance at the photosphere[40] is equal to:

$$S_{\odot} = \frac{L_{\odot}}{4\pi R_{\odot}^2} = \frac{3.828 \times 10^{26}}{4\pi \left(696\,342 \times 10^3\right)^2} = 62\,822\,741 \text{ W/m}^2$$

If we assume that the sun is a perfect black body[41], then we have:

$$\sigma \mathcal{T}_{\odot}^4 = S_{\odot} \Leftrightarrow \mathcal{T}_{\odot} = \sqrt[4]{\frac{S_{\odot}}{\sigma}}$$

The numerical calculation gives:

$$\mathcal{T}_{\odot} = \sqrt[4]{\frac{62\,822\,741}{5.67 \times 10^{-8}}} = 5\,769 \text{ K}$$

This effective temperature is slightly overestimated compared to a direct physical measurement. The previous analysis can be extended to other stars. Let R_{star} be the stellar radius of the star. Since we have $L_{\text{star}} = 4\pi R_{\text{star}}^2 S_{\text{star}}$ and $S_{\text{star}} = \mathcal{E} = \sigma \mathcal{T}^4$, we get:

$$\mathcal{T}_{\text{star}} = \sqrt[4]{\frac{L_{\text{star}}}{4\pi R_{\text{star}}^2 \sigma}} \tag{8.5}$$

$\mathcal{T}_{\text{star}}$ is defined as the temperature of a black body radiating the same amount of energy per unit area as the star. It may differ from the actual temperature of a star, which depends on its kinetic energy.

Zero-order model of the atmosphere   Let us see how the effective temperature of the Earth is calculated.

Incoming solar radiation   The incoming solar radiation is equal to:

$$F_{\text{solar}} = \frac{1}{4} \left(1 - \alpha_p\right) S_0$$

where α_p is the planetary albedo, which measures the amount of reflected sunlight. For example, α_p is equal to zero for a perfect black body, and one for a perfect white body. The ratio $\frac{1}{4}$ comes from the fact that no point on the planet receives the sun's energy continuously during a full day. This is particularly true at night, but also during the day, as the point is not necessarily at the sun's zenith. On average, we can show that a point on the planet receives $\frac{1}{4}$ of the solar energy, which is the ratio of the projected area of the sphere ($\mathcal{A}rea = \pi r^2$) divided by the surface area of the sphere ($\mathcal{A}rea = 4\pi r^2$). This incoming solar radiation is illustrated in Figure 8.36.

In the case of the Earth, we have $\alpha_p \approx 0.29$ (Stephens *et al.*, 2015, page 141). We deduce that:

$$F_{\text{solar}} = \frac{1}{4} \left(1 - 0.29\right) \times 1\,368 = 242.82 \text{ W/m}^2$$

To get an idea of how much energy is involved, consider a room with a surface area of x square meters and receiving an energy $\mathcal{E}$ expressed in watts. The radiation per square meter received by

---

[40]This is the visible surface of the Sun.
[41]In fact, the emissivity of the sun is $\varepsilon = 0.96$.

Figure 8.36: Incoming solar radiation
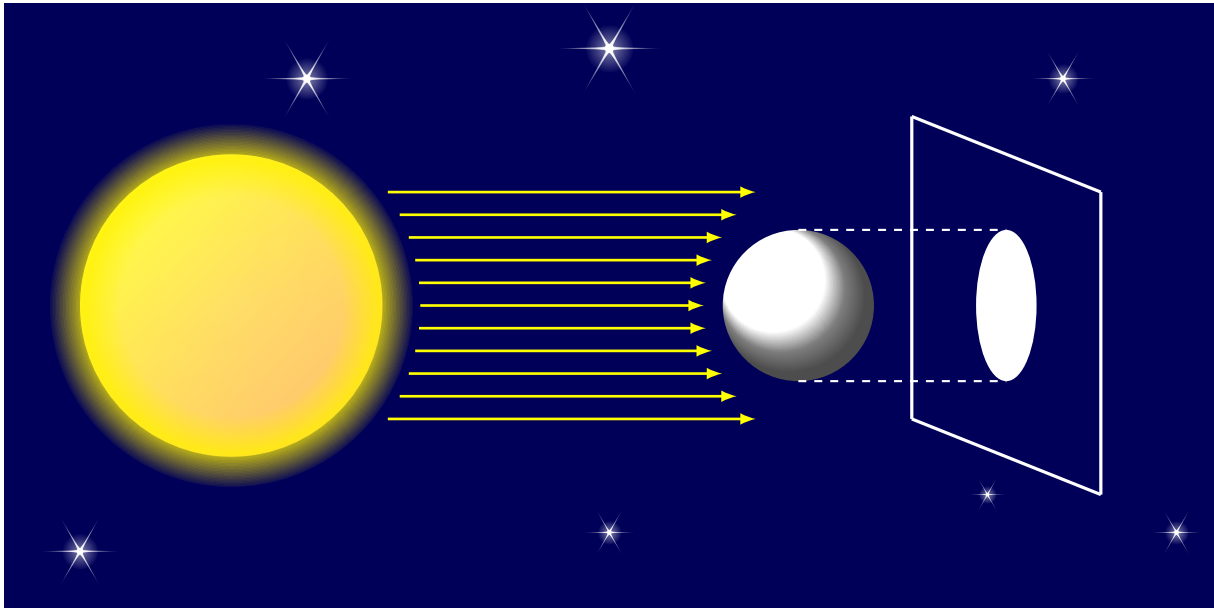
this room is equal to $\mathcal{E}/x$. If the room receives the same equivalent solar radiation F_{solar}, the energy must be equal to:

$$\mathcal{E} = x \cdot F_{\text{solar}}$$

Using a standard 200 watt lamp, we can calculate the number of lamps required to achieve the same equivalent solar radiation. The results are shown below:

x (in m^2)	1	5	10	20	50	100
$\mathcal{E}$ (in watts)	242.8	1 214	2 428	4 856	12 141	24 282
# lights	1.2	9	12	24	61	121

For a room of 20 m^2, we need 24 lights.

Effective temperature of the Earth   The Earth receives the incoming solar radiation F_{solar}, while the black body radiation is given by the Stefan-Boltzmann law. The schematic energy flow diagram is given in Figure 8.37. We deduce that:

$$\sigma \mathcal{T}_e^4 = \frac{1}{4}\left(1 - \alpha_p\right) S_0 \Leftrightarrow \mathcal{T}_e = \sqrt[4]{\frac{\left(1 - \alpha_p\right) S_0}{4\sigma}}$$

The numerical calculation gives:

$$
\begin{aligned}
\mathcal{T}_e &= \sqrt[4]{\frac{(1 - 0.29) \times 1\,368}{4 \times 5.67 \times 10^{-8}}} \\
&= 255.81\,\text{K} \\
&= 255.81°\text{C} - 273.15°\text{C} \\
&= -17.34°\text{C}
\end{aligned}
$$

The effective temperature of the Earth is then close to $-17.34°$C. Without greenhouse gases, the surface temperature of the Earth should be equal to the effective temperature. However, we observe

$$\mathcal{T}_s \approx +15°\text{C} \gg \mathcal{T}_e \approx -17°\text{C}$$

Figure 8.37: Zero-order model
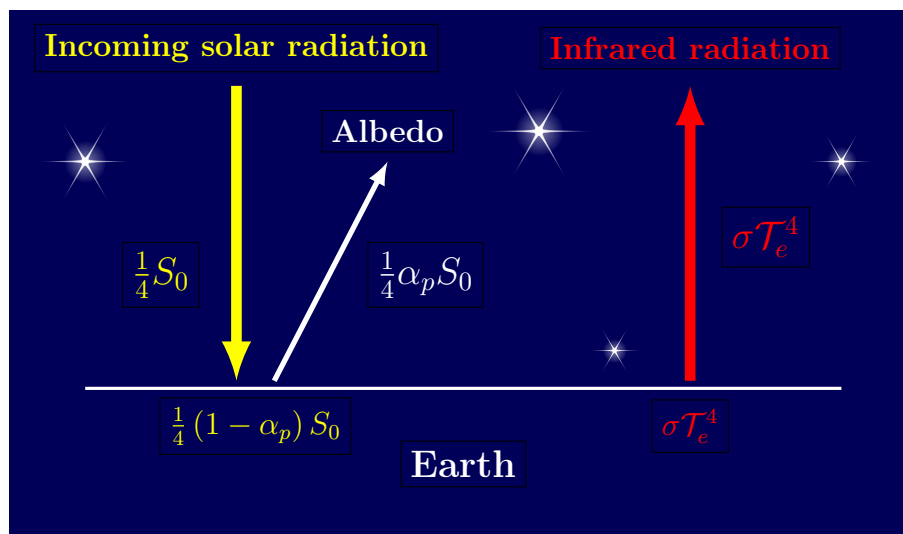
To explain this difference of 32°C, we need to introduce the greenhouse effect.

Impact of the greenhouse effect   To illustrate the greenhouse effect, Hartmann (2016) proposed to include an atmosphere in the global energy balance. He assumed that the atmosphere is a black body for terrestrial radiation but transparent to solar radiation:

> "*Since solar radiation is mostly visible and near infrared, and Earth emits primarily thermal infrared radiation, the atmosphere may affect solar and terrestrial radiation very differently. [...] Since the atmospheric layer absorbs all of the energy emitted by the surface below it and emits like a blackbody, the only radiation emitted to space is from the atmosphere in this model.*" (Hartmann, 2016, pages 32-32).

From Figure 8.38, we can deduce that the energy balance for the Earth's surface is:

$$F_{\text{solar}} + \sigma \mathcal{T}_{\text{a}}^4 = \sigma \mathcal{T}_s^4$$

while the radiation balance for the atmosphere verifies:

$$\sigma \mathcal{T}_s^4 = 2\sigma \mathcal{T}_{\text{a}}^4$$

It follows that:

$$F_{\text{solar}} + \sigma \mathcal{T}_{\text{a}}^4 = 2\sigma \mathcal{T}_{\text{a}}^4$$

or[42]:

$$F_{\text{solar}} = \sigma \mathcal{T}_{\text{a}}^4 = \sigma \mathcal{T}_{\text{e}}^4$$

We conclude that:

$$\begin{cases} \mathcal{T}_{\text{a}} = \mathcal{T}_{\text{e}} \\ \mathcal{T}_s = \sqrt[4]{2}\,\mathcal{T}_{\text{e}} \end{cases}$$

---

[42]We use the fact that the effective temperature $\mathcal{T}_{\text{e}}$ is defined by the relation $\sigma \mathcal{T}_e^4 = \dfrac{1}{4}\left(1 - \alpha_p\right)S_0 := F_{\text{solar}}$.

Using the previous numerical values, we obtain:

$$\begin{cases} \mathcal{T}_a = 255.81\,\text{K} = -17.34°\text{C} \\ \mathcal{T}_s = 304.22\,\text{K} = 31.07°\text{C} \end{cases}$$

We find that the surface temperature is warmer than the observed global mean surface temperature. This is because the assumption that the atmosphere absorbs all the heat radiated from the surface is not true.

Figure 8.38: Zero-order model with greenhouse effect
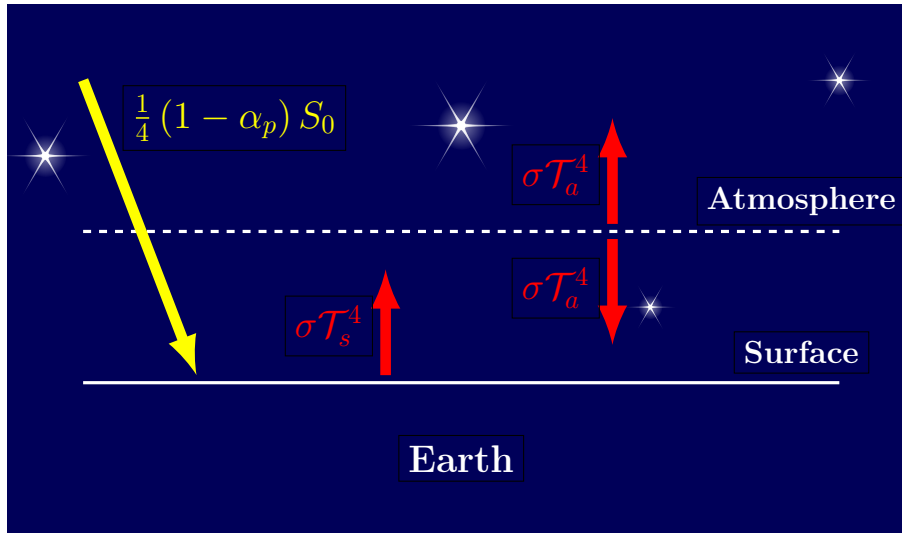
Another way to illustrate the greenhouse effect is to estimate the reflection parameter γ_p, which measures the net thermal radiation of the atmosphere with respect to the black body energy (Dommenget, 2022). We deduce that the balance $\mathcal{E}_{\text{net}}$ is:

$$\mathcal{E}_{\text{net}} = F_{\text{solar}} - \sigma \mathcal{T}_s^4 - \gamma_p \sigma \mathcal{T}_s^4$$

Solving the equation $\mathcal{E}_{\text{net}} = 0$ gives:

$$\gamma_p = 1 - \frac{F_{\text{solar}}}{\sigma \mathcal{T}_s^4} = 1 - \frac{(1 - \alpha_p) S_0}{4 \sigma \mathcal{T}_s^4}$$

Using a surface temperature of 15°C, we get $\gamma_p = 0.3788$. This means that only 62% of the infrared radiation goes into space and 38% stays on the surface.

Multi-layer model of the atmosphere Let us now consider the energy balance when we have two layers. This extension is motivated by the simplifying assumption that an atmosphere can be represented as a single black body:

> "*A layer of atmosphere that is almost opaque for longwave radiation can be crudely approximated as a blackbody that absorbs all terrestrial radiation that is incident upon it and emits like a blackbody at its temperature. For an atmosphere with a large infrared optical depth, the radiative transfer process can be represented with a series of blackbodies arranged in vertical layers. Two layers centered at 0.5 km and 2.0 km altitudes provide a simple approximation for Earth's atmosphere.*" (Hartmann, 2016, page 71).

Figure 8.39: Two-layer model
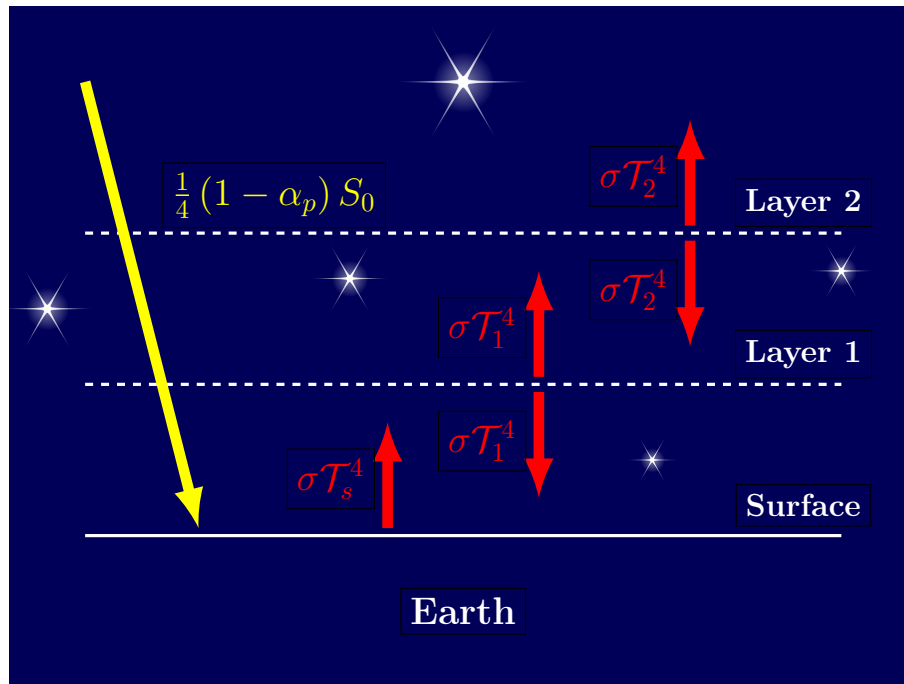
If we assume that the atmospheric layers are transparent to solar radiation, we have the schematic energy flow diagram shown in Figure 8.39. We have:

$$\begin{cases} F_{\text{solar}} + \sigma\mathcal{T}_1^4 = \sigma\mathcal{T}_s^4 \\ \sigma\mathcal{T}_2^4 + \sigma\mathcal{T}_s^4 = 2\sigma\mathcal{T}_1^4 \\ \sigma\mathcal{T}_1^4 = 2\sigma\mathcal{T}_2^4 \end{cases}$$

By replacing F_{solar} by $\sigma\mathcal{T}_e^4$ and dividing the equations by σ, we get:

$$\begin{cases} \mathcal{T}_s^4 = 3\,\mathcal{T}_e^4 \\ \mathcal{T}_1^4 = 2\,\mathcal{T}_e^4 \\ \mathcal{T}_2^4 = 1\,\mathcal{T}_e^4 \end{cases}$$

The solution is then equal to:

$$\begin{cases} \mathcal{T}_s = \sqrt[4]{3}\,\mathcal{T}_e = 336.67\,\text{K} = 63.52°\text{C} \\ \mathcal{T}_1 = \sqrt[4]{2}\,\mathcal{T}_e = 304.22\,\text{K} = 31.07°\text{C} \\ \mathcal{T}_2 = \sqrt[4]{1}\,\mathcal{T}_e = 255.81\,\text{K} = -17.34°\text{C} \end{cases}$$

It is easy to generalize these results to the multi-layer atmosphere layer. Let n be the total number of layers and $\mathcal{T}_k$ be the temperature at layer k. We have:

$$\begin{cases} \mathcal{T}_s = \mathcal{T}_0 = \sqrt[4]{n+1}\,\mathcal{T}_e \\ \mathcal{T}_k = \sqrt[4]{n+1-k}\,\mathcal{T}_e \qquad \text{for } k = 0, 1, \ldots, n \end{cases} \tag{8.6}$$

The multi-layer model is important because the atmosphere cannot be considered as a homogeneous black body. If this were true, the temperature of the atmosphere would be the same regardless of altitude. In reality, we observe that the temperature is highest at the Earth's surface and decreases

Table 8.9: Layers of the Earth's atmosphere

Index	Layer	Altitude (in Km)
1	Troposphere	12
2	Stratosphere	50
3	Mesosphere	80
4	Thermosphere	500
5	Exosphere	6 200

as we move towards the vacuum of space. This explains why the Earth's atmosphere is generally divided into five layers based on temperature and composition, as shown in Table 8.9. The solution of the multi-layer model given in Equation (8.6) satisfies the property that the temperature decreases with the layer index:

$$\frac{\partial \mathcal{T}_k}{\partial k} = -\frac{1}{4} \left(n + 1 - k \right)^{-3/4} \mathcal{T}_e \leq 0$$

Let z be the altitude in km. Since z is an increasing function of the layer index k, we deduce from the previous property that:

$$\frac{\partial \mathcal{T}(z)}{\partial z} \leq 0$$

With this simple multi-layer model, we then showed that temperature is a decreasing function of altitude, that the Earth is warmer at the surface, and that the global temperature surface depends on the multi-layer structure of the atmosphere. So it also depends on the composition of the atmosphere, because that is the main determinant of the structure of the atmosphere. However, the previous multi-layer model is not realistic because the estimated temperatures are far from the observed mean temperatures.

Emissivity model of the atmosphere The previous model is too simple for several reasons. The first reason is that the real atmosphere is not opaque. This means that it is more of a gray body than a black body. In addition to radiation, energy is also transported by convection[43]. Therefore, we cannot ignore the effect of emissivity.

We consider the one-layer model described in Figure 8.38, and introduce the emissivity ε of the atmosphere (Figure 8.40). The energy balance equilibrium begins:

- The balance at the top of the atmosphere is:

$$F_{\text{solar}} - \left(1 - \varepsilon \right) \sigma \mathcal{T}_s^4 - \varepsilon \sigma \mathcal{T}_a^4 = 0$$

- The balance of the atmosphere is:

$$\varepsilon \sigma \mathcal{T}_s^4 - 2\varepsilon \sigma \mathcal{T}_a^4 = 0$$

- The balance at the surface is:

$$F_{\text{solar}} + \varepsilon \sigma \mathcal{T}_a^4 - \sigma \mathcal{T}_s^4 = 0$$

The first equation is equivalent to:

$$\left(1 - \varepsilon \right) \sigma \mathcal{T}_s^4 + \varepsilon \sigma \mathcal{T}_a^4 = F_{\text{solar}} = \sigma \mathcal{T}_e^4$$

---

[43]Another important missing factor is ocean transportation.

Figure 8.40: One-layer model with atmospheric emissivity
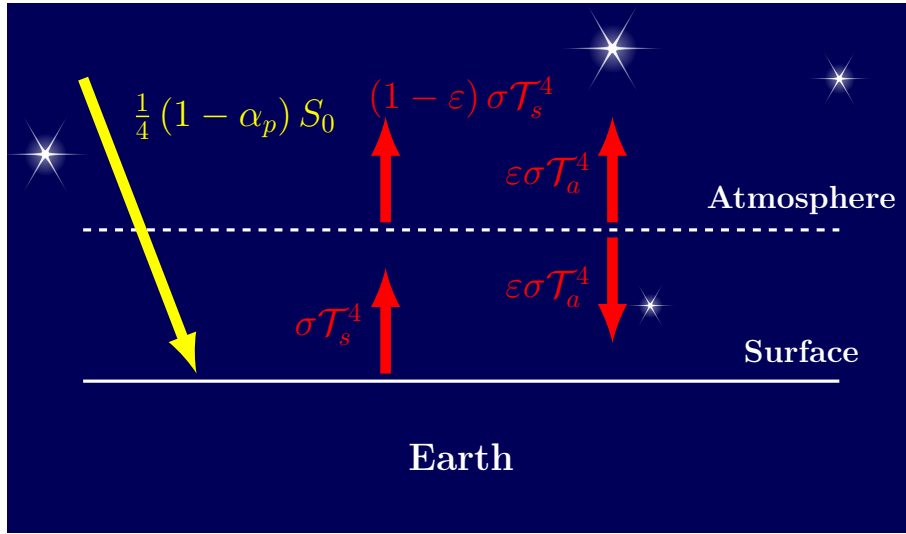
Using the second equation, it follows that:

$$\sigma \mathcal{T}_e^4 = (1 - \varepsilon) \sigma \mathcal{T}_s^4 + \frac{1}{2} \varepsilon \sigma \mathcal{T}_s^4 = \left(1 - \frac{1}{2}\varepsilon\right) \mathcal{T}_s^4$$

Finally, we conclude that[44]:

$$\mathcal{T}_s = \sqrt[4]{\frac{2}{2 - \varepsilon}} \mathcal{T}_e$$

and:

$$\mathcal{T}_a = \sqrt[4]{\frac{1}{2 - \varepsilon}} \mathcal{T}_e$$

The relationship between atmospheric emissivity ε and temperature ($\mathcal{T}_s$ and $\mathcal{T}_a$) is shown in Figure 8.41. It is an increasing function. Therefore, for a given temperature $\mathcal{T}_s^\star$, we can find the unique value of the emissivity:

$$\varepsilon^\star = 2 - 2 \left(\frac{\mathcal{T}_e}{\mathcal{T}_s^\star}\right)^4$$

Since $\mathcal{T}_s^\star \approx 15°\text{C}$, the emissivity of the atmosphere is 78%. This model then predicts an atmospheric temperature of $-30.8°\text{C}$.

Remark 52 *More generally, the Earth's energy balance is the sum of net shortwave radiation and net longwave radiation (Liang et al., 2019):*

$$\boldsymbol{\mathcal{E}}_{\text{net}} = \underbrace{\boldsymbol{\mathcal{E}}_{\text{down}}^{\text{short}} - \boldsymbol{\mathcal{E}}_{\text{up}}^{\text{short}}}_{\text{Net shortwave}} + \underbrace{\boldsymbol{\mathcal{E}}_{\text{down}}^{\text{long}} - \boldsymbol{\mathcal{E}}_{\text{up}}^{\text{long}}}_{\text{Net longwave}}$$

where $\boldsymbol{\mathcal{E}}_{\text{down}}^{\text{short/long}}$ is shortwave/longwave downward radiation and $\boldsymbol{\mathcal{E}}_{\text{up}}^{\text{short/long}}$ is shortwave/longwave upward radiation. In the previous model, we have $\boldsymbol{\mathcal{E}}_{\text{net}}^{\text{short}} = \frac{1}{4}(1 - \alpha_p) S_0$ and $\boldsymbol{\mathcal{E}}_{\text{net}}^{\text{long}} = \varepsilon \sigma \mathcal{T}_a^4 - \sigma \mathcal{T}_s^4$. Using more realistic assumptions, we get the Earth's energy balance shown in Figure 8.42. This type of representation was popularized by Kiehl and Trenberth (1997) and Trenberth et al. (2009).

---

[44]We also check that the third equation is satisfied.

Figure 8.41: Relationship between atmospheric emissivity and temperature
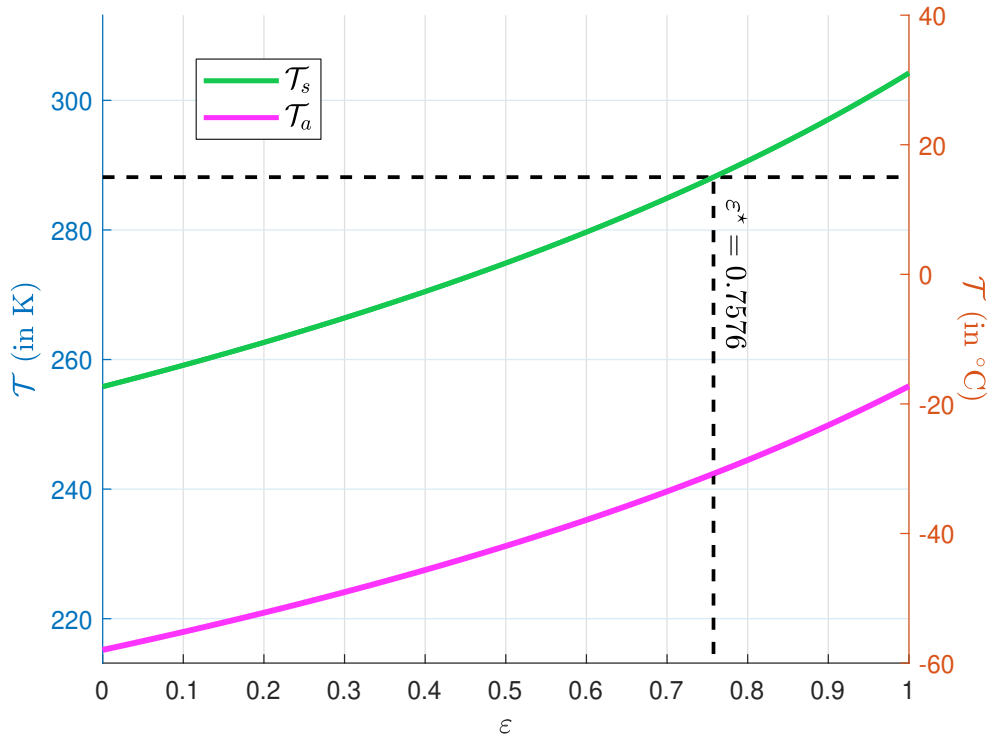
Figure 8.42: Earth's Energy Budget
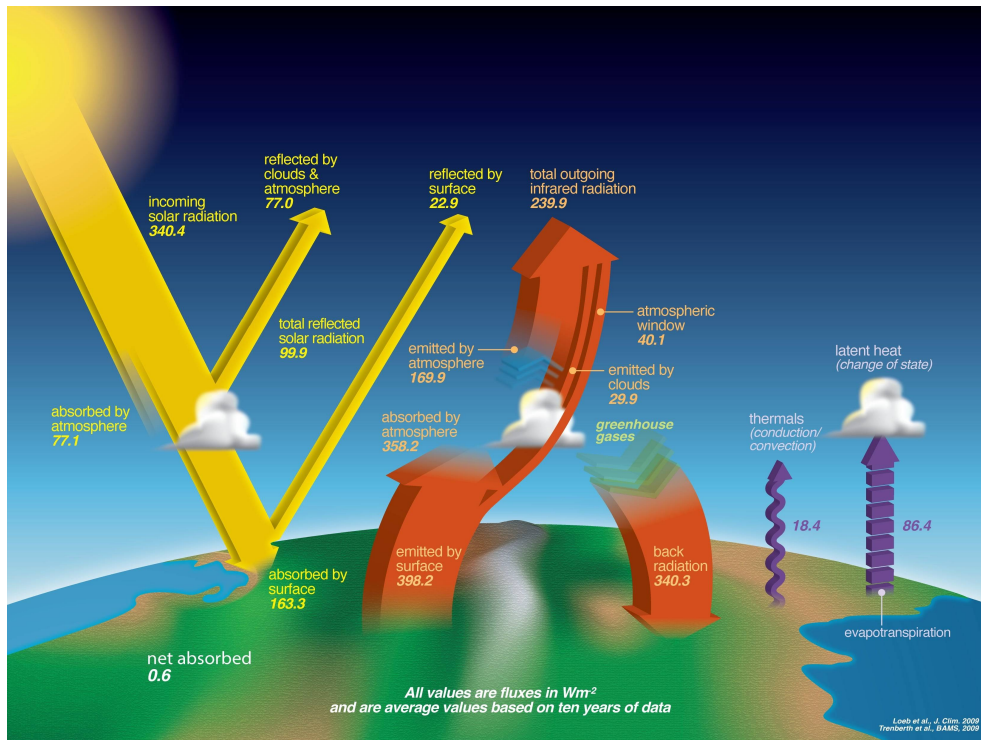
Climate sensitivity and feedback

Specific heat capacity The specific heat capacity c of a substance is the heat capacity C of the substance divided by the mass of the substance:

$$c = \frac{C}{M} = \frac{1}{M}\frac{\Delta\mathcal{E}}{\Delta\mathcal{T}} \tag{8.7}$$

where $\Delta\mathcal{E}$ is the amount of heat required to raise the temperature of the substance by $\Delta\mathcal{T}$. Here, M is the mass of the substance in kilograms (kg), $\Delta\mathcal{E}$ is the change in energy in joules (J), $\Delta\mathcal{T}$ is the change in temperature in Kelvin (K) and c is the specific heat capacity in joules per kilogram per Kelvin ($\mathrm{J\,kg^{-1}\,K^{-1}}$). For example, the specific heat capacity of water is $4\,186\,\mathrm{J\,kg^{-1}\,K^{-1}}$. From Equation (8.7), we deduce that:

$$\Delta\mathcal{E} = Mc\,\Delta\mathcal{T} \tag{8.8}$$

In this case, the amount of energy required to raise the temperature of $1\,\mathrm{m}^3$ of water by $10°\mathrm{C}$ is equal to:

$$\Delta\mathcal{E} = 10^3 \times 4\,186 \times 10 = 4\,186\,000\,\mathrm{J}$$

The specific heat capacity of air is about $1\,000\,\mathrm{J\,kg^{-1}\,K^{-1}}$, while the mass of the atmosphere is $5.148 \times 10^{18}\,\mathrm{kg}$. Therefore, the amount of energy required to raise the temperature of the atmosphere by $1°\mathrm{C}$ is:

$$
\begin{aligned}
\Delta\mathcal{E} &= \left(5.148 \times 10^{18}\,\mathrm{kg}\right) \times \left(1\,000\,\mathrm{J\,kg^{-1}\,K^{-1}}\right) \times 1\,\mathrm{K} \\
&= 5.148 \times 10^{21}\,\mathrm{J}
\end{aligned}
$$

We can also write Equation (8.7) as follows:

$$\Delta\mathcal{T} = \frac{\Delta\mathcal{E}}{Mc} \tag{8.9}$$

This equation gives the change in temperature for a change in energy. For example, adding $1\,000$ joules of energy to one liter of water will increase its temperature by about $0.239°\mathrm{C}$:

$$\Delta\mathcal{T} = \frac{1000}{1 \times 4\,186} = 0.239$$

Equations (8.7)–(8.9) can be modified by scaling the mass M of the substance to standardize the required energy $\Delta\mathcal{E}$. A possible scaling factor can be the surface area:

$$m = \frac{M}{\mathcal{A}rea}$$

and we get:

$$\Delta\mathcal{E} = mc\,\Delta\mathcal{T} \tag{8.10}$$

For the atmosphere, we have:

$$m = \frac{M}{\mathcal{A}rea} = \frac{5.148 \times 10^{18}\,\mathrm{kg}}{510.0645 \times 10^6 \times 10^6\,\mathrm{m}^2} = 1.0093 \times 10^4\,\mathrm{kg\,m^{-2}}$$

because the radius r of the Earth is $6\,371\,\mathrm{km}$ and the surface of the Earth is approximately $\mathcal{A}rea = 4\pi r^2 = 510.0645$ million km^2. In some climate modeling textbooks, Equation (8.10) is expressed using other formulas for the mass of the atmosphere per unit area. For example, m can be replaced

by the product of height h and density ρ, or by the ratio of pressure p to gravitational acceleration g. However, all these quantities are equivalent because we have:

$$m = h\rho = \left(8.2 \times 10^3 \,\text{m}\right) \times \left(1.225 \,\text{kg m}^{-3}\right) = 1.0045 \times 10^3 \,\text{kg m}^{-2}$$

and:

$$m = \frac{p}{g} = \frac{101\,325 \,\text{Pa}}{9.81 \,\text{m s}^{-2}} = \frac{101\,325 \,\text{m}^{-1} \,\text{kg s}^{-2}}{9.81 \,\text{m s}^{-2}} = 1.0329 \times 10^4 \,\text{kg m}^{-2}$$

where $h = 8.2 \,\text{km}$ is the height of the atmosphere[45], $\rho = 1.225 \,\text{kg m}^{-3}$ is the density of the atmosphere, $p = 101\,325 \,\text{Pa}$ is the standard atmospheric pressure at sea level on Earth, and $g = 9.81 \,\text{m s}^{-2}$ is the acceleration due to gravity at the Earth's surface. Therefore, we obtain the following equivalent formulas:

$$mc\,\Delta\mathcal{T} = h\rho c\,\Delta\mathcal{T} = \frac{p}{g}c\,\Delta\mathcal{T} = \Delta\mathcal{E}$$

Radiative relaxation timescale We transform Equation (8.10) into a differential equation:

$$mc\frac{\mathrm{d}\mathcal{T}}{\mathrm{d}t} = \frac{\mathrm{d}\mathcal{E}}{\mathrm{d}t}$$

We consider a black body. At the equilibrium, we have:

$$F_{\text{solar}} - \sigma\mathcal{T}^4 = 0$$

We deduce that:

$$\begin{aligned}
\mathcal{E} &= F_{\text{solar}} - \sigma\mathcal{T}^4 \\
&= \sigma\mathcal{T}_e^4 - \sigma\mathcal{T}^4
\end{aligned}$$

Let us assume that $\mathcal{T} = \mathcal{T}_e + \Delta\mathcal{T}$. It follows that:

$$mc\frac{\mathrm{d}\Delta\mathcal{T}}{\mathrm{d}t} = -4\sigma\mathcal{T}_e^3\Delta\mathcal{T}$$

because:

$$\frac{\partial}{\partial\Delta\mathcal{T}}\left(\sigma\mathcal{T}_e^4 - \sigma\mathcal{T}^4\right) = -4\sigma\mathcal{T}_e^3$$

Let τ_e be the radiative relaxation timescale defined as:

$$\tau_e = \frac{mc}{4\sigma\mathcal{T}_e^3}$$

We have:

$$\begin{cases}
\dfrac{\mathrm{d}\Delta\mathcal{T}}{\mathrm{d}t} = -\dfrac{1}{\tau_e}\Delta\mathcal{T} \\
\Delta\mathcal{T}\left(0\right) = \Delta\mathcal{T}_0
\end{cases}$$

The solution of this ordinary differential equation is well-known and we get:

$$\Delta\mathcal{T}\left(t\right) = \exp\left(-\frac{t}{\tau_e}\right)\Delta\mathcal{T}_0$$

---

[45]The height of the atmosphere depends on the choice of layer. Here we consider the troposphere, which is the lowest layer of the atmosphere at the polar regions.

$\Delta\mathcal{T}(t)$ gives the impulse response of an initial temperature shock of $\Delta\mathcal{T}_0$. Because $\tau_e > 0$, we conclude that the system is stable:

$$\lim_{t\to\infty} \Delta\mathcal{T}(t) = 0$$

We note that the equation for $\Delta\mathcal{T}(t)$ describes an exponential survival function with parameter τ_e^{-1}. We deduce that the radiative relaxation timescale τ_e is the mean lifetime.

Using the previously obtained values for the atmosphere ($m = 1.0093 \times 10^4\,\mathrm{kg\,m^{-2}}$, $c = 1\,000\,\mathrm{J\,kg^{-1}\,K^{-1}}$ and $\mathcal{T}_e = 255.81\,\mathrm{K}$), the radiative relaxation timescale is equal to 31 days[46]:

$$
\begin{aligned}
\tau_e &= \frac{\left(1.0093 \times 10^4\,\mathrm{kg\,m^{-2}}\right) \times \left(1\,000\,\mathrm{J\,kg^{-1}\,K^{-1}}\right)}{4 \times \left(5.67 \times 10^{-8}\,\mathrm{W\,m^{-2}\,K^{-4}}\right) \times \left(255.81\,\mathrm{K}\right)^3} \\
&= 2\,658\,427\,\mathrm{J\,W^{-1}} \\
&= \frac{2\,658\,427\,\mathrm{s}}{24 \times 3\,600\,\mathrm{s}} \\
&= 30.77\,\mathrm{days}
\end{aligned}
$$

Figure 8.43 shows the impulse response function for $\Delta\mathcal{T}_0 = +1°\mathrm{C}$. We check that the effect of the initial shock disappears after a period of a few weeks. However, the memory of the temperature perturbation depends strongly on the effective temperature $\mathcal{T}_e$, as illustrated in the second panel of Figure 8.43. This explains why a temperature shock does not have the same effect all over the terrestrial globe, *e.g.*, in the polar regions or in the equatorial regions.

Figure 8.43: Impulse response function for $\Delta\mathcal{T}_0 = +1°\mathrm{C}$ and a black body

---
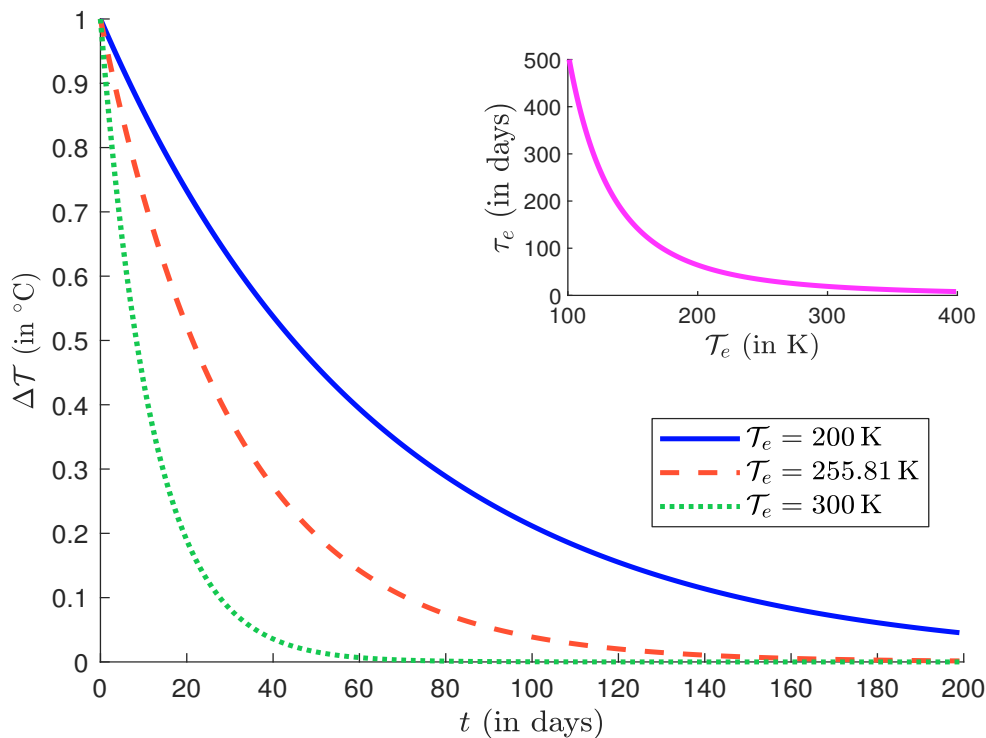
[46]We use the fact that $\mathrm{J\,W^{-1}} = \mathrm{s}$.

Figure 8.44: Impulse response function for $\Delta \mathcal{T}_0 = +1°C$ and a gray body
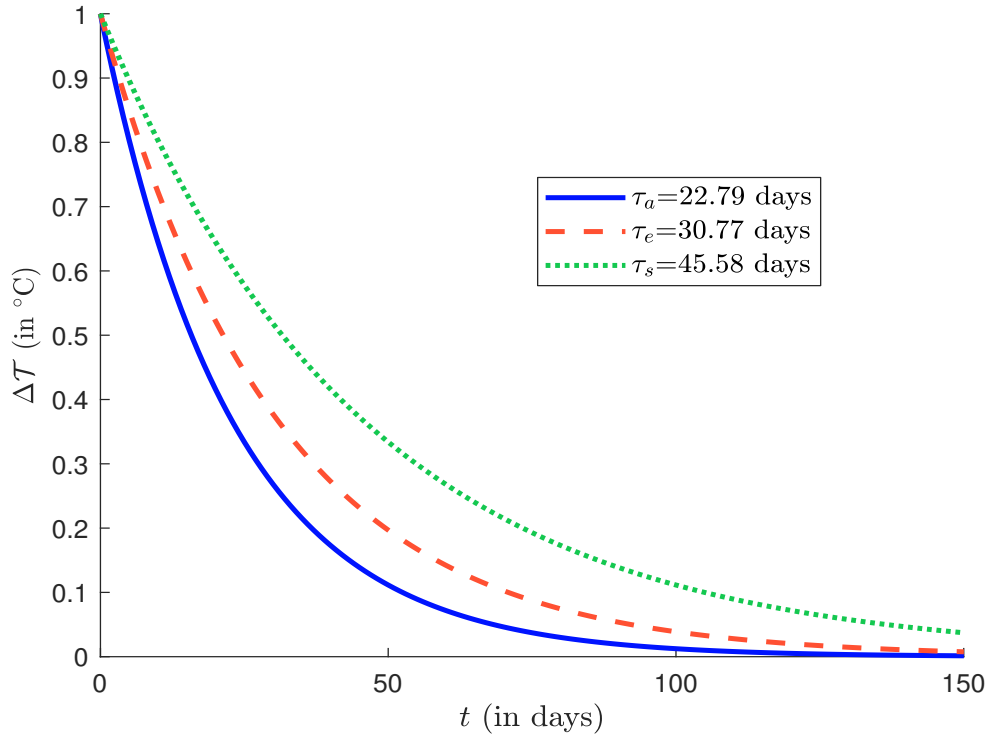
Now consider the case of a gray body and the one-layer model with emissivity. On page 310, we found that:

$$\sigma \mathcal{T}_e^4 = (2 - \varepsilon) \sigma \mathcal{T}_a^4 = \left(\frac{2 - \varepsilon}{2} \right) \sigma \mathcal{T}_s^4$$

We deduce that:

$$mc \frac{d\mathcal{T}_a}{dt} = \frac{d}{dt} \left(\sigma \mathcal{T}_e^4 - (2 - \varepsilon) \sigma \mathcal{T}_a^4 \right)$$

and:

$$mc \frac{d\mathcal{T}_s}{dt} = \frac{d}{dt} \left(\sigma \mathcal{T}_e^4 - \left(\frac{2 - \varepsilon}{2} \right) \sigma \mathcal{T}_s^4 \right)$$

Using the same reasoning as above, we obtain:

$$\begin{cases} \Delta \mathcal{T}_a (t) = \exp \left(-\frac{t}{\tau_a} \right) \Delta \mathcal{T}_0 \\ \Delta \mathcal{T}_s (t) = \exp \left(-\frac{t}{\tau_s} \right) \Delta \mathcal{T}_0 \end{cases}$$

where:

$$\tau_a = \frac{mc}{4 (2 - \varepsilon) \sigma \mathcal{T}_e^3} \qquad \text{and} \qquad \tau_s = \frac{mc}{2 (2 - \varepsilon) \sigma \mathcal{T}_e^3}$$

We consider an emissivity value of 65%, and find that $\tau_a = 22.79$ and $\tau_s = 45.58$ days. The corresponding impulse response functions are given in Figure 8.44. Although these calculations are rough and approximate[47], they clearly show that the response at the Earth's surface is not the same as in the different layers of the atmosphere.

---

[47]In fact, it is not realistic to use the same value for m, c and $\mathcal{T}_e$.

Mathematical definition of climate sensitivity and feedback   Before giving the mathematical expression of these two concepts, here is how Hartman introduces them in his book:

> "[...] Climate forcing is a change to the climate system that can be expected to change the climate. Examples would be doubling the CO_2, increasing the total solar irradiance (TSI) by 2%, introducing volcanic aerosols into the stratosphere, etc. Climate forcings are usually quantified in terms of how many W/m^2 they change the energy balance when imposed. For example, instantaneously doubling the CO_2 changes the energy balance at the top of the atmosphere by about $4\,W/m^2$. A feedback process is a response of the climate system to surface warming that then alters the energy balance in such a way as to change the temperature response to the forcing. A positive feedback makes the forced response bigger, and a negative feedback makes it smaller. Classic examples of positive feedbacks are ice-albedo feedback and water-vapor feedback. When it warms, ice melts, and this reduces Earth's albedo and causes further warming. When it cools, ice grows, and this increases Earth's albedo, causing further cooling [...]." (Hartmann, 2016, page 294).

Mathematically, we assume that some extra energy ΔF is added to the system. ΔF is called the radiative forcing and is measured in W/m^2. The climate response is generally measured as the change in the surface temperature $\Delta \mathcal{T}_s$. The climate sensitivity is defined as:

$$\phi := \frac{\Delta \mathcal{T}_s}{\Delta F}$$

Using differential notation, we have:

$$\phi := \frac{\mathrm{d}\mathcal{T}_s}{\mathrm{d}F}$$

implying that:

$$\mathrm{d}\mathcal{T}_s = \phi \, \mathrm{d}F$$

Following Hartmann (2016, page 295), we assume that the perturbation $\mathrm{d}F$ depends on the temperature $\mathcal{T}_s$ and some exogenous factors x_i:

$$\mathrm{d}F = \frac{\partial F}{\partial \mathcal{T}_s}\mathrm{d}\mathcal{T}_s + \sum_{i=1}^{n} \frac{\partial F}{\partial x_i}\mathrm{d}x_i$$

We conclude that:

$$\left(1 - \phi \frac{\partial F}{\partial \mathcal{T}_s}\right)\mathrm{d}\mathcal{T}_s = \phi \sum_{i=1}^{n} \frac{\partial F}{\partial x_i}\mathrm{d}x_i$$

We then get a feedback mechanism, because the temperature dynamics depend on the factors x_i, but also on the temperature response.

Let us apply this framework to the one-layer model with emissivity. At the equilibrium, we have $\mathcal{E} = 0$ where:

$$\begin{aligned} \mathcal{E} &= F_{\text{solar}} - \left(\frac{2-\varepsilon}{2}\right)\sigma\mathcal{T}_s^4 \\ &= \frac{1}{4}\left(1 - \alpha_p\right)S_0 - \left(\frac{2-\varepsilon}{2}\right)\sigma\mathcal{T}_s^4 \end{aligned}$$

The first-order Taylor Series expansion of $\mathcal{E} = 0$ gives:

$$\Delta\mathcal{E} = \frac{1}{4}\left(1 - \alpha_p\right)\Delta S_0 - \frac{1}{4}S_0\Delta\alpha_p + \frac{1}{2}\sigma\mathcal{T}_s^4\Delta\varepsilon - 4\left(\frac{2-\varepsilon}{2}\right)\sigma\mathcal{T}_s^3\Delta\mathcal{T}_s$$

Remember that each perturbation Δy depends on the temperature $\mathcal{T}_s$ and some exogenous factors x_i:

$$\Delta y = \frac{\partial y}{\partial \mathcal{T}_s} \Delta \mathcal{T}_s + \sum_{i=1}^n \frac{\partial y}{\partial x_i} \Delta x_i$$

We have:

$$
\begin{aligned}
\Delta \mathcal{E} &= \frac{1}{4}\left(1 - \alpha_p\right)\Delta S_0 - \frac{1}{4}S_0\left(\frac{\partial \alpha_p}{\partial \mathcal{T}_s}\Delta \mathcal{T}_s + \sum_{i=1}^n \frac{\partial \alpha_p}{\partial x_i}\Delta x_i\right) + \\
&\quad \frac{1}{2}\sigma \mathcal{T}_s^4\left(\frac{\partial \varepsilon}{\partial \mathcal{T}_s}\Delta \mathcal{T}_s + \sum_{i=1}^n \frac{\partial \varepsilon}{\partial x_i}\Delta x_i\right) - 4\left(\frac{2 - \varepsilon}{2}\right)\sigma \mathcal{T}_s^3 \Delta \mathcal{T}_s
\end{aligned}
$$

We deduce that[48]:

$$\Delta \mathcal{E} = \lambda \Delta \mathcal{T}_s + \sum_{i=0}^n \Delta F_i \tag{8.11}$$

where $\Delta F_0 := \Delta F_{\text{solar}} = \frac{1}{4}\left(1 - \alpha_p\right)\Delta S_0$,

$$\Delta F_i = \left(-\frac{1}{4}S_0\frac{\partial \alpha_p}{\partial x_i} + \frac{1}{2}\sigma \mathcal{T}_s^4\frac{\partial \varepsilon}{\partial x_i}\right)\Delta x_i \qquad \text{for } i \geq 1$$

and:

$$\lambda = -\frac{1}{4}S_0\frac{\partial \alpha_p}{\partial \mathcal{T}_s} + \frac{1}{2}\sigma \mathcal{T}_s^4\frac{\partial \varepsilon}{\partial \mathcal{T}_s} - 4\left(\frac{2 - \varepsilon}{2}\right)\sigma \mathcal{T}_s^3$$

Since $\Delta \mathcal{E}$ and ΔF_i are measured in W/m^2 and $\Delta \mathcal{T}_s$ is measured in Kelvin, we deduce that λ is measured in $\text{W m}^{-2}\,\text{K}^{-1}$. As before, we transform $\Delta \mathcal{E}$ into $\Delta \mathcal{T}_s$ by considering the heat capacity c expressed in $\text{W m}^{-2}\,\text{K}^{-1}\,\text{s}$. Then Equation (8.11) becomes:

$$c\frac{\mathrm{d}\Delta \mathcal{T}_s}{\mathrm{d}t} = \lambda \Delta \mathcal{T}_s + \Delta F \tag{8.12}$$

where $\Delta F = \sum_{i=0}^n F_i$. The climate feedback parameter λ can be positive or negative, and we have the following mathematical properties:

- If $\lambda > 0$, the system is unstable;

- If $\lambda < 0$, the system is stable and the equilibrium is reached when:

$$\Delta \mathcal{T}_s = \Delta \mathcal{T}_s^\star = -\frac{\Delta F}{\lambda} = -\phi \Delta F \tag{8.13}$$

We see that the climate feedback parameter can be decomposed into three components:

$$\lambda = \lambda_0 + \lambda_{\alpha_p} + \lambda_\varepsilon$$

where:

1. λ_0 is the Planck feedback or the black body response:

$$\lambda_0 = -4\left(\frac{2 - \varepsilon}{2}\right)\sigma \mathcal{T}_s^3$$

---

[48]$\lambda = \phi^{-1}$ is the inverse of the climate sensitivity parameter ϕ defined by Hartmann (2016).

2. λ_{α_p} is the surface albedo feedback:

$$\lambda_{\alpha_p} = -\frac{1}{4} S_0 \frac{\partial \alpha_p}{\partial \mathcal{T}_s}$$

3. λ_{ε} is the emissivity feedback:

$$\lambda_{\varepsilon} = \frac{1}{2} \sigma \mathcal{T}_s^4 \frac{\partial \varepsilon}{\partial \mathcal{T}_s}$$

Since $\varepsilon < 1$, the Planck feedback is negative, meaning that it stabilizes the climate and counteracts global warming. In fact, as the Earth warms, it emits more thermal radiation into space, and this increased longwave radiation acts as a natural cooling mechanism. Conversely, as the Earth cools, it emits less thermal radiation into space, and this decreased longwave radiation acts as a natural warming mechanism. Earlier we found that $\varepsilon = 0.78$. So the estimate of λ_0 is[49]:

$$
\begin{aligned}
\lambda_0 &= -4 \left(\frac{2 - 0.78}{2} \right) \times \left(5.67 \times 10^{-8} \right) \times \left(273.15 + 15 \right)^3 \\
&= -3.310 \, \mathrm{W \, m^{-2} \, K^{-1}}
\end{aligned}
$$

The sign of the surface albedo feedback depends on the sign of $\dfrac{\partial \alpha_p}{\partial \mathcal{T}_s}$. Two main factors play a role in determining α_p:

> "*The albedo of the planet for solar radiation is primarily determined by the clouds and surface, with the main variable component of the latter being the ice/snow cover.*" (Hansen et al., 1984, page 166).

The feedback is positive for the ice-albedo mechanism because $\dfrac{\partial \alpha_p}{\partial \mathcal{T}_s} < 0$. In fact, decreased temperature leads to increased sea ice and snow cover. This increases the albedo and decreases the absorption of shortwave radiation. This mechanism is satisfied globally, but we observe a lot of regional and seasonal variation (Stephens et al., 2015). If the Earth were completely covered in ice, its albedo α_p would be about 84%, meaning it would reflect most of the sunlight that hits it. On the other hand, if the Earth were covered by a dark green forest canopy, the albedo would be about 14% and most of the sunlight would be absorbed[50]. This explains why the ice-albedo feedback is greatest in the polar regions (Goosse et al., 2018). Cloud albedo feedback occurs because changes in cloud cover, cloud altitude or cloud properties affect the amount of reflected shortwave radiation. This feedback can be either positive or negative, depending on the specific change. The overall effect is a positive feedback (IPCC, 2021, Chapter 7, page 926). According to IPCC AR6, the value of the global surface albedo feedback is estimated to be $0.35 \, \mathrm{W \, m^{-2} \, K^{-1}}$, with a 90% confidence interval from 0.10 to $0.60 \, \mathrm{W \, m^{-2} \, K^{-1}}$ (IPCC, 2021, Chapter 7, page 971). This means that:

$$\frac{\partial \alpha_p}{\partial \mathcal{T}_s} = -\frac{4 \lambda_{\alpha_p}}{S_0} = -\frac{4 \times 0.35}{1368} = -1.023 \times 10^{-3}$$

---

[49]For comparison, the most recent estimate is $-3.22 \, \mathrm{W \, m^{-2} \, K^{-1}}$, while the 90% confidence interval is from -3.4 to $-3.0 \, \mathrm{W \, m^{-2} \, K^{-1}}$ (IPCC, 2021, Chapter 7, page 968).
[50]Source: https://earthobservatory.nasa.gov/images/84499/measuring-earths-albedo.

Box 8.9: Ice-albedo feedback modeling

We have seen that the surface albedo feedback is equal to $\lambda_{\alpha_p} = -\frac{1}{4}S_0\frac{\partial \alpha_p}{\partial \mathcal{T}_s}$. So we need to parameterize the temperature dependence of the albedo. If we assume that $\alpha_p = f_{\text{albedo}}(\mathcal{T}_s)$, then we get $\lambda_{\alpha_p} = -\frac{1}{4}S_0 f'_{\text{albedo}}(\mathcal{T}_s)$. For example, Sellers (1969) suggested:

$$\alpha_p = \begin{cases} b(\phi) - 0.009\mathcal{T}_s & \text{if } \mathcal{T}_s \leq 283.16\,\text{K} \\ b(\phi) - 2.548 & \text{if } \mathcal{T}_s \geq 283.16\,\text{K} \end{cases}$$

where $b(\phi)$ is an estimated coefficient that depends on the latitude ϕ. Sellers (1969) found $b(\phi) \in [2.798, 2.992]$. Below, we show the function $b(\phi)$ with respect to the spherical coordinate ϕ. On average we have $\overline{b(\phi)} = 2.8811$. We deduce that:

$$\lambda_{\alpha_p} = \frac{1}{4} \times 1\,368 \times 0.009 = 3.078\,\text{W m}^{-2}\,\text{K}^{-1}$$

It is obvious that this positive feedback has been overestimated. The reason is that snow and sea ice cover about 10% of the Earth's surface. Therefore, we get $\lambda_{\alpha_p} \approx 0.3\,\text{W m}^{-2}\,\text{K}^{-1}$.
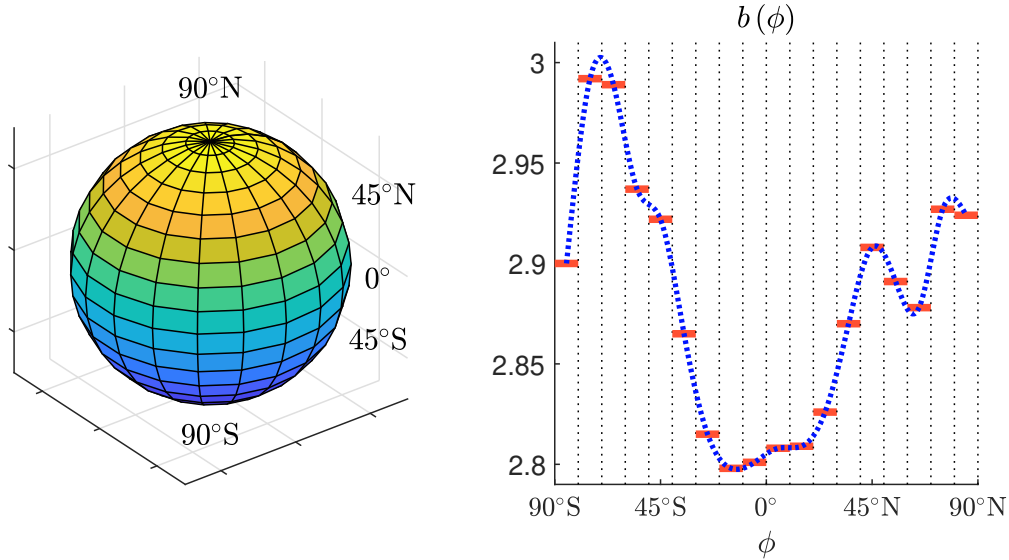
Budyko (1969) proposed another famous ice-albedo model, assuming that the land and oceans are completely covered by snow and sea ice when the temperature is below a threshold $\mathcal{T}_{\text{cold}}$, and that the planet is free of ice and snow when the temperature is above another threshold $\mathcal{T}_{\text{warm}}$:

$$\alpha_p = \begin{cases} \alpha_{\text{cold}} & \text{if } \mathcal{T}_s \leq \mathcal{T}_{\text{cold}} \\ \alpha_{\text{warm}} + (\alpha_{\text{cold}} - \alpha_{\text{warm}})\left(\dfrac{\mathcal{T}_{\text{warm}} - \mathcal{T}_s}{\mathcal{T}_{\text{warm}} - \mathcal{T}_{\text{cold}}}\right)^\eta & \text{if } \mathcal{T}_{\text{cold}} \leq \mathcal{T}_s \leq \mathcal{T}_{\text{warm}} \\ \alpha_{\text{warm}} & \text{if } \mathcal{T}_s \geq \mathcal{T}_{\text{warm}} \end{cases}$$

where $\eta \geq 1$. It follows that:

$$\lambda_{\alpha_p}(\mathcal{T}_s) = \frac{1}{4}\eta S_0 (\alpha_{\text{cold}} - \alpha_{\text{warm}})\frac{(\mathcal{T}_{\text{warm}} - \mathcal{T}_s)^{\eta-1}}{(\mathcal{T}_{\text{warm}} - \mathcal{T}_{\text{cold}})^\eta} \cdot \mathbb{1}\{\mathcal{T}_{\text{cold}} \leq \mathcal{T}_s \leq \mathcal{T}_{\text{warm}}\}$$

Using the values $\alpha_{\text{cold}} = 0.7$, $\mathcal{T}_{\text{cold}} = 260\,\text{K}$, $\alpha_{\text{warm}} = 0.3$, $\mathcal{T}_{\text{warm}} = 295\,\text{K}$ and $\eta = 2$, we get $\lambda_{\alpha_p}(282\,\text{K}) = 2.90\,\text{W m}^{-2}\,\text{K}^{-1}$, $\lambda_{\alpha_p}(288\,\text{K}) = 1.56\,\text{W m}^{-2}\,\text{K}^{-1}$ and $\lambda_{\alpha_p}(293\,\text{K}) = 0.45\,\text{W m}^{-2}\,\text{K}^{-1}$.

The third feedback is the change in emissivity, which is the sum of several components:

$$\lambda_\varepsilon = \lambda_{\text{water vapor}} + \lambda_{\text{lapse rate}} + \lambda_{\text{cloud longwave}} + \cdots$$

The water vapor feedback, also known as the specific humidity feedback, is the most important positive and destabilizing feedback. It can be described as follows:

> "*[...] As the temperature increases, the amount of water vapor in saturated air increases. Since water vapor is the principal greenhouse gas, increasing water vapor content will increase the greenhouse effect of the atmosphere and raise the surface temperature even further.*" (Hartmann, 2016, page 297).

The mechanism behind water vapor feedback is related to the thermodynamics of moist (or humid) air in the atmosphere. Water molecules move freely between the liquid and vapor phases, and the equilibrium depends on the water vapor saturation e_s. The relationship between e_s and temperature is given by the Clausius-Clapeyron equation:

$$\frac{\mathrm{d}e_s}{\mathrm{d}\mathcal{T}_s} = \frac{1}{(\alpha_{\text{vapor}} - \alpha_{\text{liquid}})\,\mathcal{T}_s} = \frac{L_{\text{vapor}}e_s}{R_{\text{vapor}}\mathcal{T}_s^2}$$

where α_{vapor} is the specific volume of the vapor phase, α_{liquid} is the specific volume of the liquid phase, R_{vapor} is the gas constant of water vapor and L_{vapor} is the specific latent heat of evaporation of water. If we rearrange this equation, we obtain the following relationship:

$$\frac{\mathrm{d}e_s}{e_s} = \left(\frac{L_{\text{vapor}}}{R_{\text{vapor}}\mathcal{T}_s}\right)\frac{\mathrm{d}\mathcal{T}_s}{\mathcal{T}_s} = \varphi\frac{\mathrm{d}\mathcal{T}_s}{\mathcal{T}_s}$$

According to (Hartmann, 2016, page 298), $\varphi \approx 20$, which means that a 1% change in temperature is associated with a 20% change in saturation humidity. This means that a 20% change in saturation humidity causes a change in temperature of $1\% \times (273.15 + 15) \approx 2.9\,\text{K}$. According to IPCC AR6, the value $\lambda_{\text{water vapor}}$ of the water vapor feedback is assessed to be $1.85\,\text{W}\,\text{m}^{-2}\,\text{K}^{-1}$ (IPCC, 2021, Chapter 7, page 969). The lapse rate mechanism describes the relationship between temperature and altitude in the atmosphere (Colman and Soden, 2021, page 13):

$$\Gamma = -\frac{\partial \mathcal{T}}{\partial z} \in \left[4\,\text{K}\,\text{km}^{-1}, 10\,\text{K}\,\text{km}^{-1}\right]$$

where z is the altitude in kilometers. On average, the lapse rate is about 6.5°C per kilometer. This means that the temperature is not uniform throughout the atmosphere. According to IPCC (2021, Chapter 7, page 969), "*the warming is larger in the upper troposphere than in the lower troposphere [...] leading to a larger radiative emission to space and therefore a negative feedback*" and the average value of $\lambda_{\text{lapse rate}}$ is $-0.50\,\text{W}\,\text{m}^{-2}\,\text{K}^{-1}$. The case of cloud feedbacks is more complicated because it involves several mechanisms: (1) high cloud altitude, (2) tropical high cloud amount, (3) subtropical marine low cloud, (4) land cloud, (5) midlatitude cloud amount, (6) extratropical cloud optical depth, and (7) Arctic cloud. According to IPCC AR6, the value λ_{cloud} of the net cloud feedback is estimated to be $0.42\,\text{W}\,\text{m}^{-2}\,\text{K}^{-1}$ (IPCC, 2021, Chapter 7, page 974). One of the difficulties is to decompose the cloud feedback into shortwave and longwave feedbacks, since the global surface albedo feedback already includes shortwave cloud mechanisms. Assuming that 2/3 of the cloud feedback is longwave radiation, we get:

$$\lambda_\varepsilon \approx 1.85 - 0.50 + \frac{2}{3} \times 0.42 = 1.63\,\text{W}\,\text{m}^{-2}\,\text{K}^{-1}$$

Finally, we can compute the total feedback parameter by summing the Planck feedback, the surface albedo feedback, and the emissivity feedback:

$$
\begin{aligned}
\lambda &= \lambda_0 + \lambda_{\alpha_p} + \lambda_\varepsilon \\
&= -3.31 + 0.35 + 1.63 \\
&= -1.33 \, \mathrm{W \, m^{-2} \, K^{-1}}
\end{aligned}
$$

This value is obtained with a simple one-layer model with emissivity. It is relatively close to the average climate feedback parameter calculated with more complex models. However, using global climate models, we can assess the uncertainty in the climate feedback parameter. In this case, we consider the parameter λ as a Gaussian random variable $\tilde{\lambda} \sim \mathcal{N}\left(\mu_\lambda, \sigma_\lambda^2\right)$. As before, we can decompose the feedback as a sum of individual feedbacks:

$$
\tilde{\lambda} = \sum_{i=1}^{n} \tilde{\lambda}_i
$$

where $\tilde{\lambda}_i \sim \mathcal{N}\left(\mu_i, \sigma_i^2\right)$. Assuming that the individual feedbacks are independent, we have:

$$
\begin{cases}
\mu_\lambda = \sum_{i=1}^{n} \mu_i \\
\sigma_\lambda = \sqrt{\sum_{i=1}^{n} \sigma_i^2}
\end{cases}
$$

Table 8.10: Parameters μ_i and σ_i of feedback parameters

Feedback mechanism	μ_i	σ_i	Shortwave
High-cloud altitude	+0.20	0.10	
Tropical marine low cloud	+0.25	0.16	
Tropical anvil cloud area	−0.20	0.20	
Land cloud amount	+0.08	0.08	
Middle-latitude marine low-cloud amount	+0.12	0.12	
High-latitude low-cloud optical depth	+0.00	0.10	✓
Planck feedback	−3.20	0.10	✓
Water vapor + lapse rate	+1.15	0.15	
Surface albedo	+0.30	0.15	✓
Total cloud	+0.45	0.33	
Stratospheric	+0.00	0.10	
Atmospheric composition changes	+0.00	0.15	
Climate feedback parameter	−1.30	0.44	

Source: (Sherwood *et al.*, 2020, Table 1, page 18).

In Table 8.10, we report the estimated value of μ_i and σ_i for individual climate feedbacks calculated by Sherwood *et al.* (2020, Table 1, page 18). The estimation process is based on multiple lines of evidence, including historical observations, general circulation models, and theory. The first part of the table lists the parameters for the individual cloud feedbacks. Their corresponding probability density functions are shown in Figure 8.45. Using this modeling framework, we can compute the probability that the feedback is positive:

$$
\Pr\left\{\tilde{\lambda}_i \geq 0\right\} = 1 - \Phi\left(-\frac{\mu_i}{\sigma_i}\right)
$$

Figure 8.45: Probability density function of individual cloud feedbacks
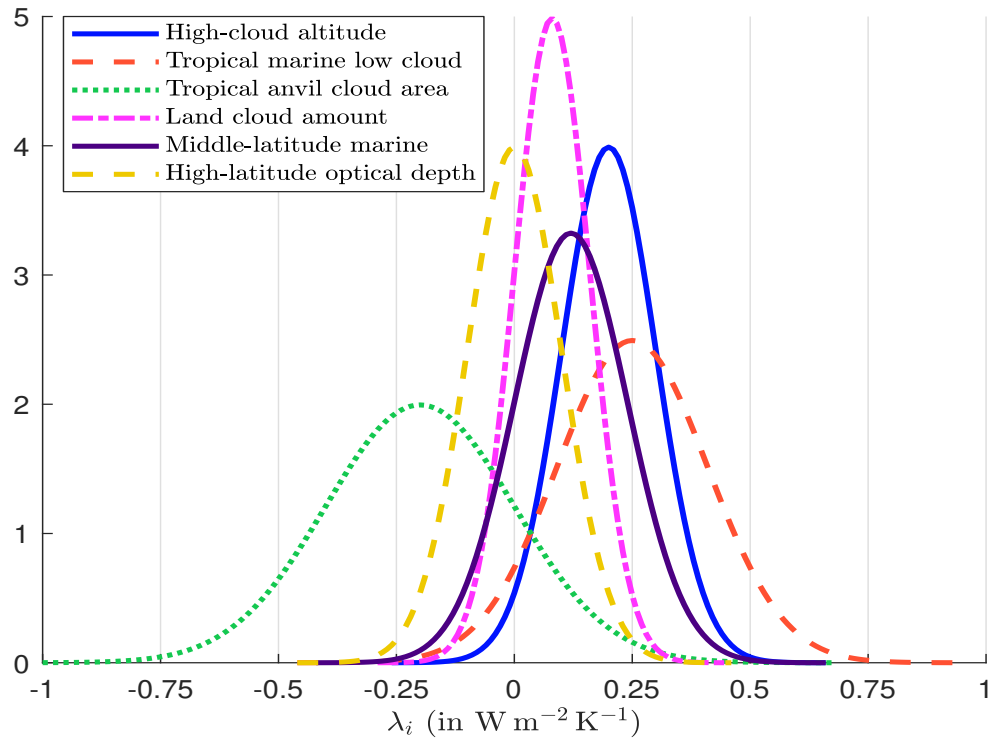
Figure 8.46: Probability density function of positive feedbacks
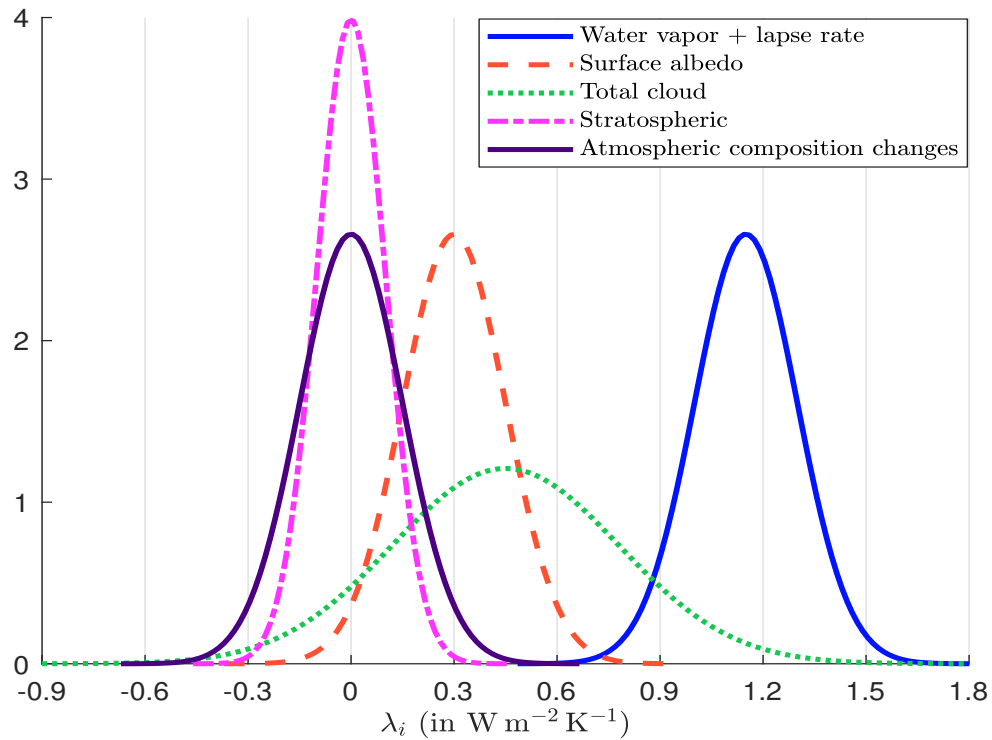
Figure 8.47: Comparison of $\tilde{\lambda}_{\text{Planck}}$ and $\tilde{\lambda}_{\text{positive}}$
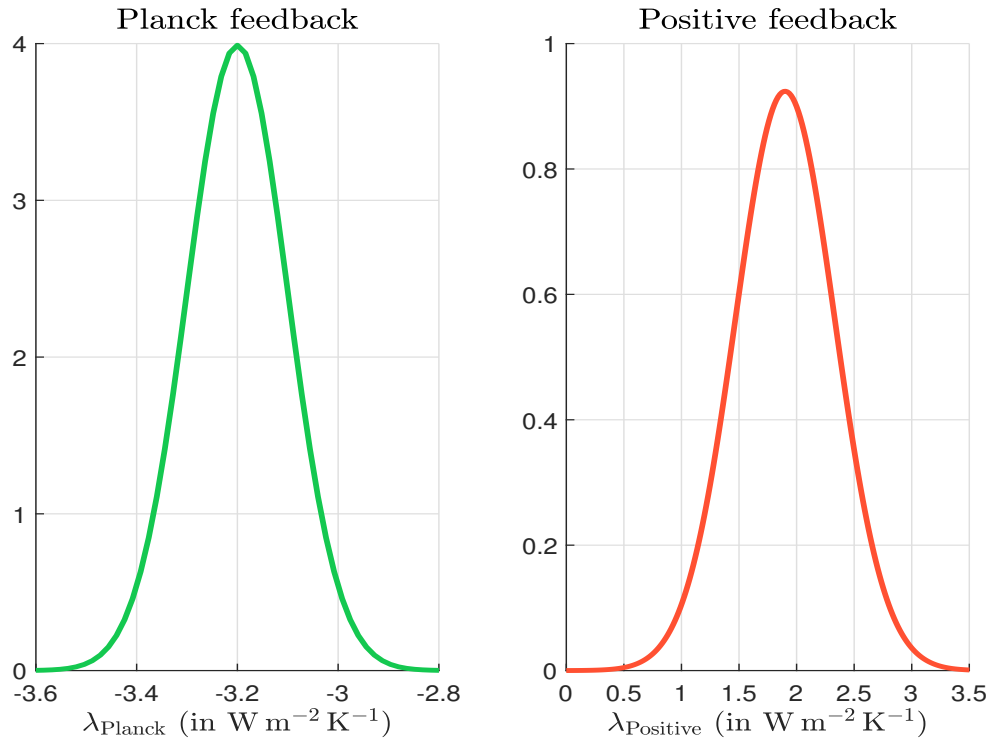
or the confidence interval at the α confidence level: $[\mu_i - c_\alpha \sigma_i, \mu_i + c_\alpha \sigma_i]$ where $c_\alpha = \Phi^{-1}((1+\alpha)/2)$. In IPCC AR6, the confidence intervals of $\tilde{\lambda}_i$ are called *likely* and *very likely* ranges and correspond to $\alpha = 66\%$ and $\alpha = 90\%$, respectively[51]. Using the parameters of the individual cloud feedbacks, we can infer the distribution of the total cloud feedback. We have:

$$\mu_{\text{total cloud}} = 0.20 + 0.25 - 0.20 + 0.08 + 0.12 + 0.00 = 0.45\,\text{W}\,\text{m}^{-2}\,\text{K}^{-1}$$

and:

$$\sigma_{\text{total cloud}} = \sqrt{0.10^2 + 0.16^2 + 0.20^2 + 0.08^2 + 0.12^2 + 0.10^2} = 0.3262\,\text{W}\,\text{m}^{-2}\,\text{K}^{-1}$$

These are the numbers obtained by Sherwood *et al.* (2020). We now plot the probability density functions of the five main positive feedback components in Figure 8.46 and aggregate them into a single feedback. We have:

$$\mu_{\text{positive}} = 1.15 + 0.30 + 0.45 + 0.00 + 0.00 = 1.90\,\text{W}\,\text{m}^{-2}\,\text{K}^{-1}$$

and:

$$\sigma_{\text{positive}} = \sqrt{0.15^2 + 0.15^2 + 0.33^2 + 0.10^2 + 0.15^2} = 0.4317\,\text{W}\,\text{m}^{-2}\,\text{K}^{-1}$$

The aggregated positive feedback is shown in Figure 8.47 and can be compared to the Planck feedback. Finally, Sherwood *et al.* (2020) aggregated all the positive and negative feedbacks to obtain the climate feedback parameter: $\tilde{\lambda} \sim \mathcal{N}(\mu_\lambda, \sigma_\lambda^2)$ where $\mu_\lambda = -1.30\,\text{W}\,\text{m}^{-2}\,\text{K}^{-1}$ and $\sigma_\lambda = 0.44\,\text{W}\,\text{m}^{-2}\,\text{K}^{-1}$. The probability density function of $\tilde{\lambda}$ is given in Figure 8.48. The probability that the climate feedback is positive is small[52], less than 1%.

---

[51] c_α is equal to 0.95 and 1.65.

[52] We have:

$$\Pr\left\{\tilde{\lambda} \geq 0\right\} = \Pr\left\{\frac{\tilde{\lambda} - \mu_\lambda}{\sigma_\lambda} \geq -\frac{\mu_\lambda}{\sigma_\lambda}\right\} = 1 - \Phi\left(-\frac{\mu_\lambda}{\sigma_\lambda}\right) = 1 - \Phi\left(\frac{1.30}{0.44}\right) = 0.16\%$$

Figure 8.48: Probability density function of the climate feedback parameter
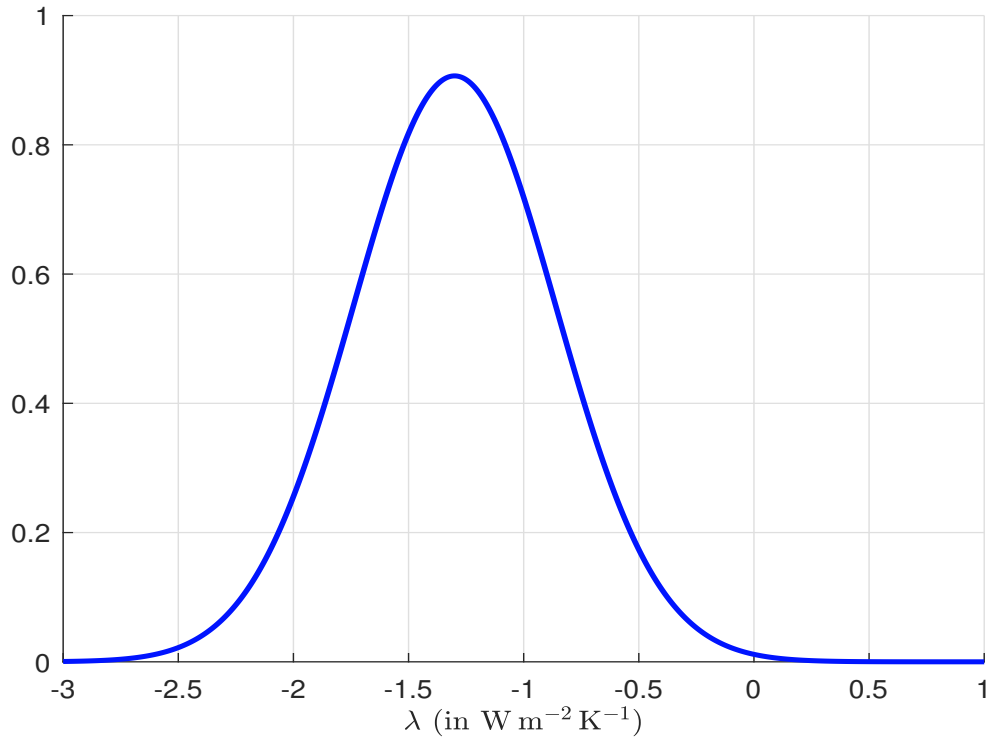
Equilibrium climate sensitivity (ECS)   We recall that the definition of equilibrium is given by Equation (8.13):

$$\Delta \mathcal{T}_s^\star = -\frac{\Delta F}{\lambda} = -\phi \Delta F$$

Therefore, we can calculate the effect of a specific value of radiative forcing ΔF on temperature change. In particular, Earth's equilibrium climate sensitivity (ECS) is the long-term global mean surface temperature change due to a specific value of radiative forcing corresponding to a doubling of CO_2 in the atmosphere (Jeevanjee, 2023). We have:

$$\text{ECS} := \Delta \mathcal{T}_{2\times CO_2} = -\frac{\Delta F_{2\times CO_2}}{\lambda}$$

where ECS (or $\Delta \mathcal{T}_{2\times CO_2}$) is the equilibrium climate sensitivity, λ is the climate sensitivity parameter and $\Delta F_{2\times CO_2}$ is the radiative forcing resulting from a doubling of the atmospheric carbon dioxide concentration.

Stochastic equilibrium temperature modeling   Let us assume that $\Delta F = 4\,\text{W/m}^2$ and $\lambda = -1.33\,\text{W}\,\text{m}^{-2}\,\text{K}^{-1}$. The equilibrium warming is then equal to 3°C:

$$\Delta \mathcal{T}_s^\star = -\frac{4\,\text{W/m}^2}{-1.33\,\text{W}\,\text{m}^{-2}\,\text{K}^{-1}} = 3.0075\,\text{K}$$

Remember that the dynamics of the temperature change is given by the ordinary differential equation:

$$c\frac{\mathrm{d}\Delta \mathcal{T}_s}{\mathrm{d}t} = \lambda \Delta \mathcal{T}_s + \Delta F$$

We assume that the solution has the following form:

$$\Delta \mathcal{T}_s(t) = e^{At} B + C$$

We deduce that:

$$\frac{\mathrm{d}\Delta \mathcal{T}_s}{\mathrm{d}t} = A e^{At} B = A\left(\Delta \mathcal{T}_s(t) - C\right)$$

The identification of the parameters results in:

$$\begin{cases} A = \dfrac{\lambda}{c} \\ -AC = \dfrac{\Delta F}{c} \\ B + C = \Delta \mathcal{T}_s(0) \end{cases}$$

The solutions are then $A = c^{-1}\lambda$,

$$C = -\frac{\Delta F}{Ac} = -\frac{\Delta F}{\lambda} = \Delta \mathcal{T}_s^{\star}$$

and:

$$B = \Delta \mathcal{T}_s(0) - C = \Delta \mathcal{T}_s(0) - \Delta \mathcal{T}_s^{\star}$$

Finally, we conclude that:

$$\Delta \mathcal{T}_s(t) = \exp\left(-\frac{t}{\tau}\right)\left(\Delta \mathcal{T}_s(0) - \Delta \mathcal{T}_s^{\star}\right) + \Delta \mathcal{T}_s^{\star} \qquad (8.14)$$

where:

$$\tau = -\frac{c}{\lambda} \qquad (8.15)$$

Again, we obtain that the equation for $\Delta \mathcal{T}_s(t)$ describes an exponential survival function with parameter τ^{-1}. Using a specific heat capacity of $c = 4 \times 10^8 \,\mathrm{J\,m^{-2}\,K^{-1}}$, we get[53]:

$$\tau = -\frac{c}{\lambda} = -\frac{4 \times 10^8 \,\mathrm{J\,m^{-2}\,K^{-1}}}{-1.33 \,\mathrm{W\,m^{-2}\,K^{-1}}} = \frac{4 \times 10^8 \,\mathrm{W\,m^{-2}\,K^{-1}\,s}}{1.33 \,\mathrm{W\,m^{-2}\,K^{-1}}} = 3.0075 \times 10^8 \,\mathrm{s}$$

The relaxation time τ of the climate system is then equal to $3\,480.92$ days or 9.53 years[54]. Due to the exponential distribution, τ is also the mean lifetime. The computation of the half-life[55] gives $t_{1/2} = \tau \ln(2) = 6.61$ years. Considering that the initial state is $\Delta \mathcal{T}_s(0) = 0$, the dynamics of $\Delta \mathcal{T}_s(t)$ when $\Delta F = 4 \,\mathrm{W/m^2}$ and $\lambda = -1.33 \,\mathrm{W\,m^{-2}\,K^{-1}}$ is shown in Figure 8.49. We check that the system converges to the equilibrium: $\lim_{t\to\infty} \Delta \mathcal{T}_s(t) = \Delta \mathcal{T}_s^{\star} = 3.0075°\mathrm{C}$.

The previous analysis assumes that the climate feedback parameter is certain. In fact, it is stochastic, which means that the equilibrium temperature and the temperature dynamics are stochastic. If $\tilde{\lambda} \sim \mathcal{N}\left(\mu_\lambda, \sigma_\lambda^2\right)$, the equilibrium temperature is equal to:

$$\Delta \widetilde{\mathcal{T}}_s^{\star} = -\frac{\Delta F}{\tilde{\lambda}} = \frac{1}{\xi}$$

---

[53] We have $\mathrm{J\,m^{-2}\,K^{-1}} = \mathrm{J\,s^{-1}\,m^{-2}\,K^{-1}\,s} = \mathrm{W\,m^{-2}\,K^{-1}\,s}$.

[54] We scale τ by $24 \times 3\,600$ seconds to get the relaxation time expressed in days and assume that one year is 362.25 days.

[55] Since we have:

$$\Delta \mathcal{T}_s(t) - \Delta \mathcal{T}_s^{\star} = \exp\left(-\frac{t}{\tau}\right)\left(\Delta \mathcal{T}_s(0) - \Delta \mathcal{T}_s^{\star}\right)$$

the half-life $t_{1/2}$ is defined as the solution of the equation $\exp\left(-\dfrac{t_{1/2}}{\tau}\right) = \dfrac{1}{2}$.

Figure 8.49: Surface temperature dynamics after a radiative forcing of $4\,\mathrm{W/m^2}$ ($\lambda = -1.33\,\mathrm{W\,m^{-2}\,K^{-1}}$)
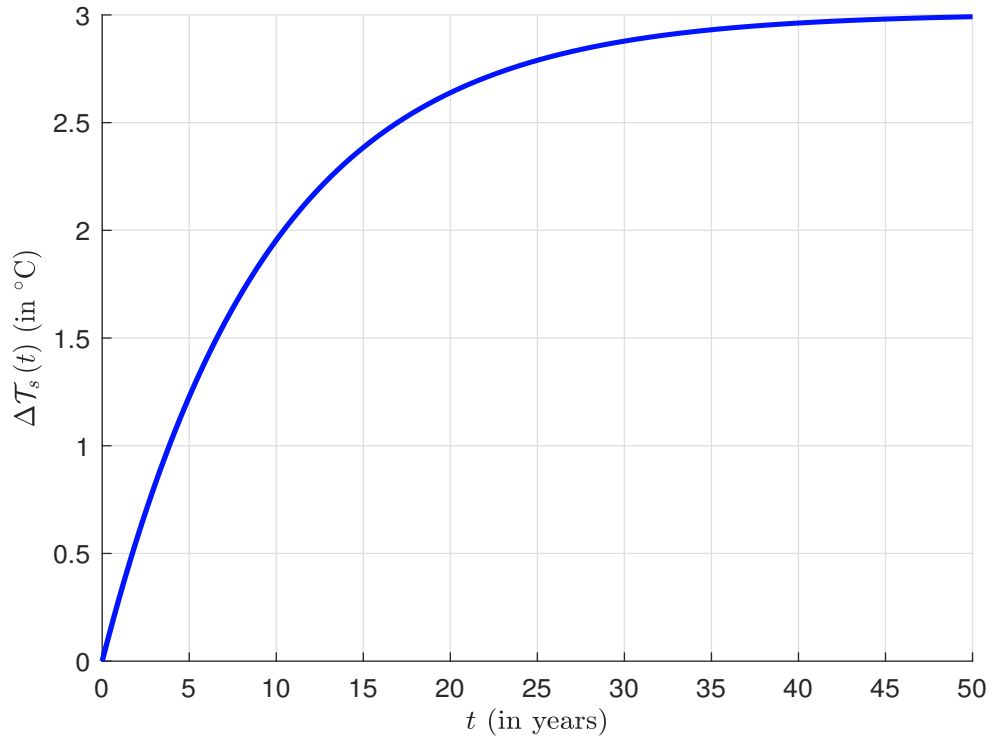
Figure 8.50: Probability density function of the equilibrium temperature ($\Delta F = 4\,\mathrm{W/m^2}$, $\mu_\lambda = -1.30\,\mathrm{W\,m^{-2}\,K^{-1}}$ and $\sigma_\lambda = 0.44\,\mathrm{W\,m^{-2}\,K^{-1}}$)
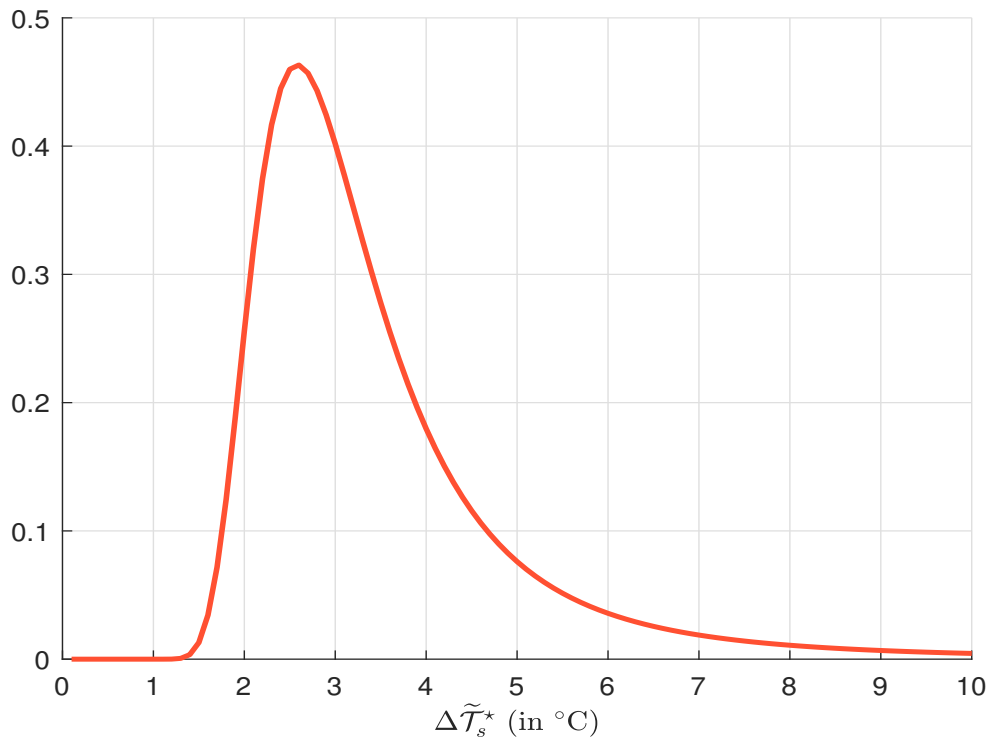
Figure 8.51: Probability density function of the relaxation time ($\mu_\lambda = -1.30 \, \mathrm{W \, m^{-2} \, K^{-1}}$ and $\sigma_\lambda = 0.44 \, \mathrm{W \, m^{-2} \, K^{-1}}$)
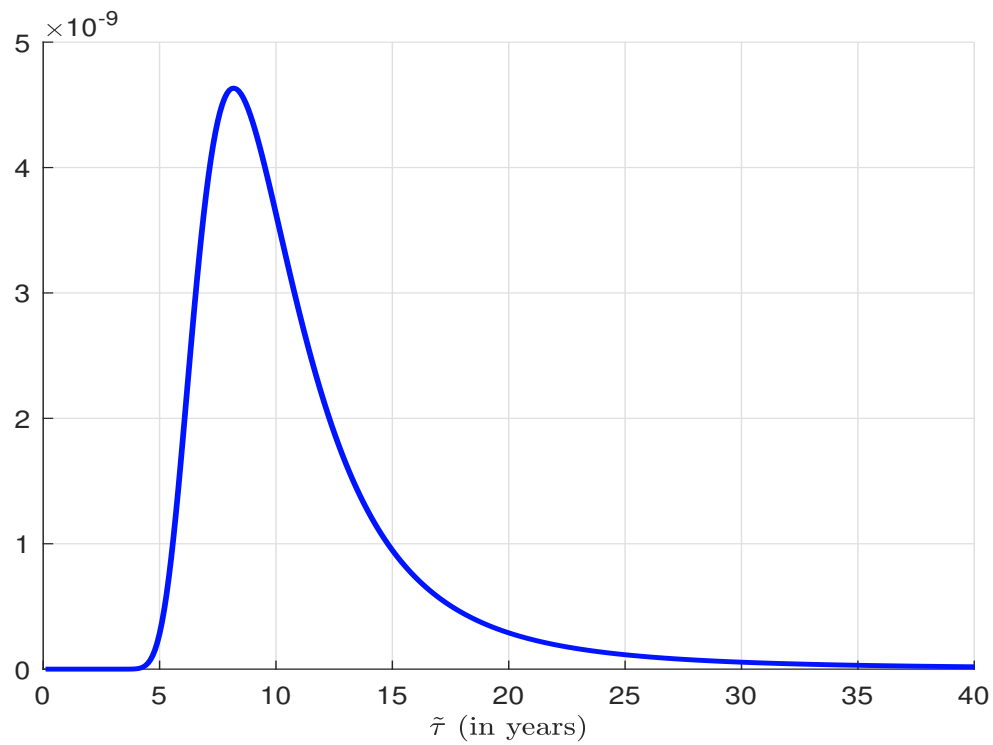
Figure 8.52: Monte Carlo simulation of the surface temperature dynamics after a radiative forcing of $4 \, \mathrm{W/m^2}$ ($\mu_\lambda = -1.30 \, \mathrm{W \, m^{-2} \, K^{-1}}$ and $\sigma_\lambda = 0.44 \, \mathrm{W \, m^{-2} \, K^{-1}}$)
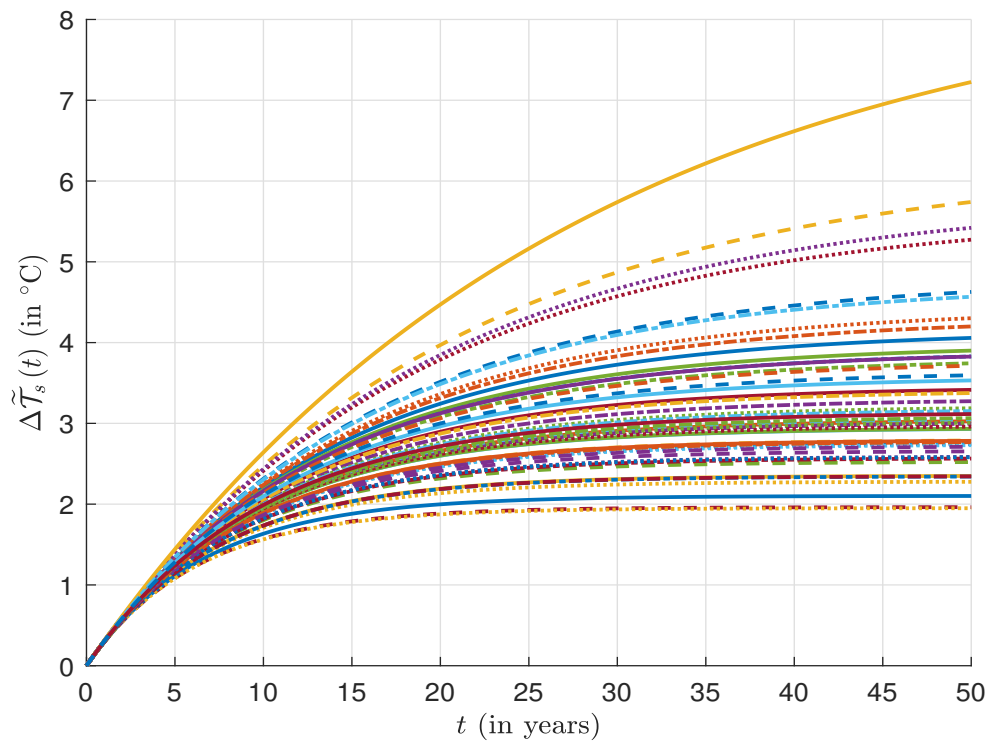
where:

$$\xi = -\frac{\tilde{\lambda}}{\Delta F} \sim \mathcal{N}\left(\mu_\xi, \sigma_\xi^2\right) \equiv \mathcal{N}\left(-\frac{\mu_\lambda}{\Delta F}, \frac{\sigma_\lambda^2}{\Delta F^2}\right)$$

We deduce that $\Delta \widetilde{\mathcal{T}}_s^\star$ follows a reciprocal normal distribution, and its probability density function is[56]:

$$f(x) = \frac{1}{\sigma_\xi x^2 \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x^{-1} - \mu_\xi}{\sigma_\xi}\right)^2}$$

Using the previous calibration of the climate feedback parameter $\tilde{\lambda} \sim \mathcal{N}\left(-1.30, 0.44^2\right)$, Figure 8.50 shows the probability density function of the equilibrium temperature when the radiative forcing is equal to $4\,\mathrm{W/m^2}$. The distribution of $\Delta \widetilde{\mathcal{T}}_s^\star$ is right skewed and has an excess of kurtosis. Using numerical integration, we compute the cumulative distribution function and the exceedance probability:

Temperature θ	2°C	3°C	4°C	5°C	7°C	10°C
$\Pr\left\{\Delta \widetilde{\mathcal{T}}_s^\star \geq \theta\right\}$	94.26%	52.86%	24.61%	12.63%	4.73%	1.88%

The probability of observing an equilibrium temperature greater than 4°C and 7°C is close to 25% and 5%, respectively. Therefore, there is a high uncertainty in $\Delta \mathcal{T}_s^\star$ when we introduce the random nature of the climate feedback. The uncertainty of λ also concerns the dynamics of $\Delta \mathcal{T}_s(t)$. Indeed, the mean lifetime becomes random and follows a reciprocal normal distribution:

$$\tilde{\tau} = -\frac{c}{\tilde{\lambda}} \sim \mathcal{RN}\left(-\frac{\mu_\lambda}{c}, \frac{\sigma_\lambda^2}{c}\right)$$

Figure 8.51 shows the probability distribution of $\tilde{\tau}$ when $\tilde{\lambda} \sim \mathcal{N}\left(-1.33, 0.44^2\right)$. Finally, we consider a Monte Carlo exercise and simulate 50 random paths of the surface temperature in Figure 8.51.

Equilibrium climate sensitivity estimation　According to IPCC (2021, Chapter 7, page 926), the total anthropogenic ERF[57] over the industrial era (1750-2019) is $2.72\,\mathrm{W/m^2}$, the contribution of anthropogenic greenhouse gas emissions is $3.84\,\mathrm{W/m^2}$ and the combined effect of all radiative feedbacks (including Planck feedback) is estimated to be $-1.16\,\mathrm{W/m^2}$. Thus, the best estimate of the ECS is 3°C. The likely range is 2.5°C to 4°C (with a probability of 66%), and the very likely range is 2°C to 5°C (with a probability of 90%).

Figure 8.53 reproduces the IPCC graph showing the change in effective radiative forcing (ERF) from 1750 to 2019 by contributing forcing agents. The total anthropogenic effective radiative forcing is divided into six categories: carbon dioxide, other well-mixed greenhouse gases (methane, nitrous oxide, and halogens), ozone, stratospheric water vapor, albedo, contrails & aviation-induced cirrus, and aerosols (aerosol-cloud and aerosol-radiation). The contributions of CO_2, CH_4 and N_2O are 79.41%, 19.9% and 17.3% of the total ERF, respectively. They are partially compensated by three negative contributors (aerosol-cloud, aerosol-radiation and land use). In Table 8.11, we report the evolution of these contributing forcings. In 2019, the total anthropogenic effective radiative forcing is $2.72\,\mathrm{W/m^2}$ and has increased by 29.3% over the last decade. The natural ERF due to volcanic activity and solar variations is $0.118\,\mathrm{W/m^2}$. We deduce that the total ERF, which is the sum of anthropogenic and natural ERF, is equal to $2.838\,\mathrm{W/m^2}$ in 2019. It has been multiplied by a factor of 4.85 in the last 70 years.

---

[56]We apply the theorem of variable change, which is explained on page 663.

[57]Effective radiative forcing (ERF) is a measure of the change in radiative flux at the top of the atmosphere and does not include the Planck feedback.

Figure 8.53: Anthropogenic effective radiative forcing (ERF) from 1750 to 2019 by contributing forcing agents

Source: (IPCC, 2021, Figure 7.6, Chapter 7, page 959) &
https://www.ipcc.ch/report/ar6/wg1/figures/chapter-7/figure-7-6.
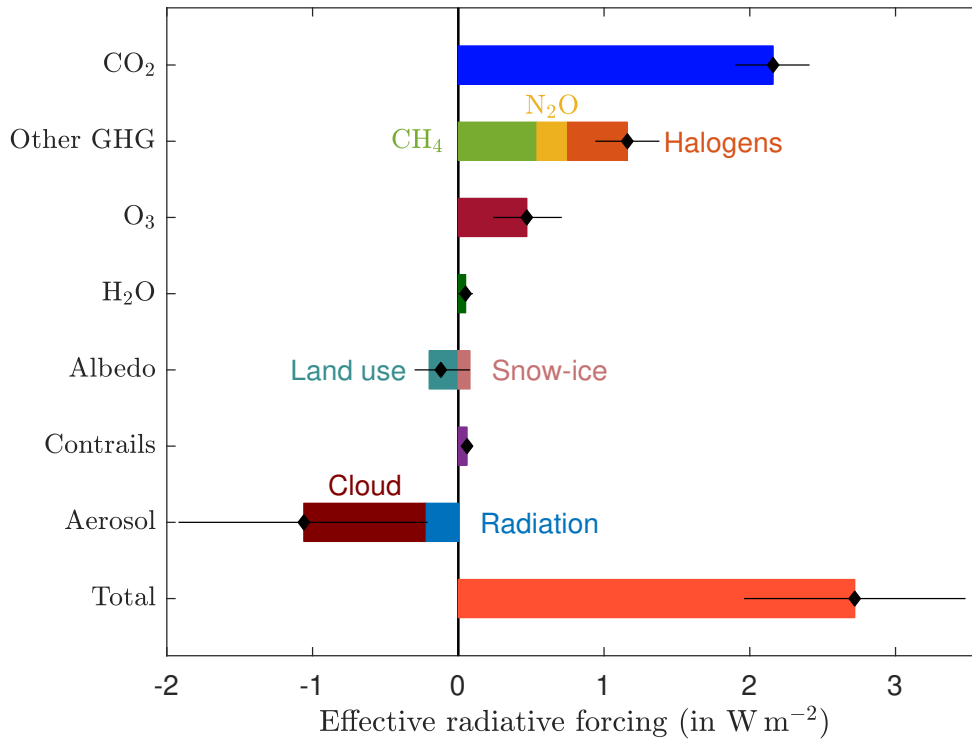
Table 8.11: Effective radiative forcing from 1750 to 2019

Forcing agent	1800	1850	1900	1950	2000	2010	2019
CO_2	0.070	0.140	0.346	0.648	1.561	1.854	2.156
CH_4	0.025	0.049	0.119	0.245	0.509	0.518	0.544
N_2O	0.004	0.007	0.032	0.069	0.157	0.181	0.208
Other GHG	0.000	0.000	0.000	0.010	0.375	0.392	0.408
O_3	0.015	0.030	0.081	0.167	0.399	0.443	0.474
H_2O (stratospheric)	0.002	0.005	0.011	0.022	0.047	0.048	0.050
Contrails	0.000	0.000	0.000	0.005	0.039	0.044	0.058
Aerosol	−0.018	−0.078	−0.346	−0.708	−1.221	−1.266	−1.058
Black carbon on snow	0.002	0.006	0.020	0.032	0.069	0.085	0.080
Land use	−0.011	−0.031	−0.084	−0.144	−0.194	−0.197	−0.200
Total (anthropogenic)	0.089	0.128	0.179	0.346	1.739	2.103	2.720
Volcanic	0.183	0.194	0.198	0.182	0.175	0.137	0.140
Solar	−0.043	0.008	−0.037	0.057	0.110	−0.008	−0.022
Total (natural)	0.140	0.202	0.160	0.239	0.285	0.129	0.118
Total	0.229	0.330	0.339	0.585	2.025	2.232	2.838

Source: (IPCC, 2021, Figure 7.6, Chapter 7, page 959) &
https://www.ipcc.ch/report/ar6/wg1/figures/chapter-7/figure-7-6.

Remark 53 *The previous decomposition based on radiative sources can be complemented by an attribution based on* GHG *emissions. For example, Panel (a) in Figure 8.54 shows the global and annual mean ERF attributed to emitted compounds over the period 1750-2019 based on AerChemMIP simulations (Thornhill et al., 2021), while the emissions-based contributions to* GSAT *change are estimated in Panel (b).*

Figure 8.54: Contribution to effective radiative forcing (ERF) (a) and global mean surface air temperature (GSAT) change (b) from component emissions between 1750 to 2019 based on CMIP6 models
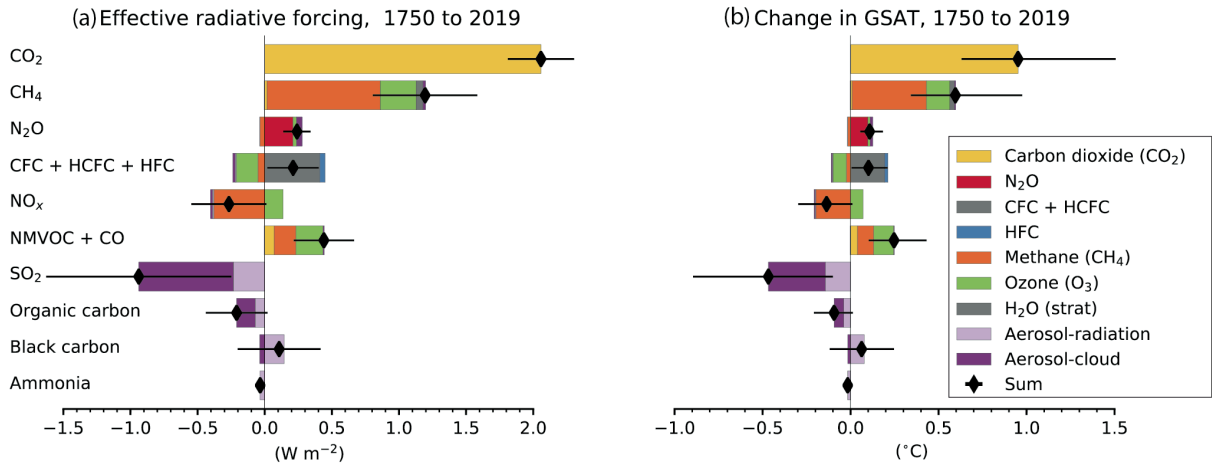
Source: (IPCC, 2021, Figure 6.12, Chapter 6, page 854) & https://www.ipcc.ch/report/ar6/wg1/figures/chapter-6/figure-6-12.

We now have all the information we need to calculate the equilibrium climate sensitivity. Using the parameters from Sherwood *et al.* (2020) $\lambda \sim \mathcal{N}\left(-1.30, 0.44^2\right)$ and $\Delta F_{2\times CO_2} \sim \mathcal{N}\left(4.00, 0.30^2\right)$, we get:

$$
\begin{aligned}
\text{ECS} &= -\frac{\Delta F_{2\times CO_2}}{\lambda} \\
&= -\frac{\mathcal{N}\left(4.00, 0.30^2\right)}{\mathcal{N}\left(-1.30, 0.44^2\right)} \\
&= \mathcal{NRD}\left(-4.00, 0.30^2, -1.30, 0.44^2\right)
\end{aligned}
$$

where $\mathcal{NRD}$ is the normal ratio distribution defined on page 664. The probability density function of ECS is shown in Figure 8.55 and the exceedance probability is given below:

Temperature θ	2°C	3°C	4°C	5°C	7°C	10°C
$\Pr\left\{\text{ECS} \geq \theta\right\}$	93.24%	52.79%	24.92%	12.85%	4.81%	1.91%

Therefore, we have a probability about 25% to observe a global warming greater than 4°C. In Figure 8.55, we compare the probability distribution when $\Delta F_{2\times CO_2}$ is constant and the probability distribution when $\Delta F_{2\times CO_2}$ is stochastic. This means comparing $\mathcal{RN}\left(\frac{1.30}{4.00}, \frac{0.44^2}{4.00^2}\right)$ and $\mathcal{NRD}\left(-4.00, 0.30^2, -1.30, 0.44^2\right)$. We conclude that the feedback uncertainty is the main contributor to the ECS uncertainty.

Figure 8.55: Probability density function of the equilibrium climate sensitivity ($\Delta F = +4.00 \, \text{W/m}^2$, $\sigma(\Delta F) = 0.30 \, \text{W/m}^2$, $\mu_\lambda = -1.30 \, \text{W} \, \text{m}^{-2} \, \text{K}^{-1}$ and $\sigma_\lambda = 0.44 \, \text{W} \, \text{m}^{-2} \, \text{K}^{-1}$)
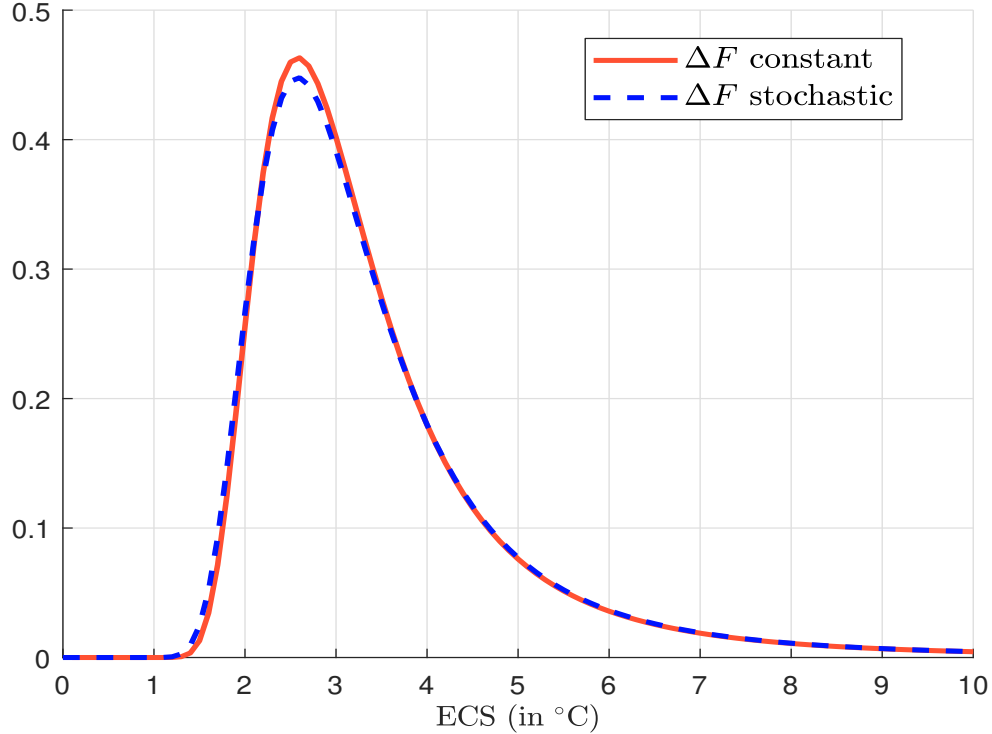
Instability, tipping point and chaos

An example of unstable equilibrium   We combine the Budyko model presented on page 319 with the one-layer model with emissivity. At equilibrium, we have:

$$\mathcal{E} = \frac{1}{4} \left(1 - \alpha_p(T_s) \right) S_0 - \left(\frac{2 - \varepsilon}{2} \right) \sigma \mathcal{T}_s^4 = 0$$

where[58]:

$$\alpha_p(\mathcal{T}_s) = \begin{cases} \alpha_{\text{cold}} & \text{if } \mathcal{T}_s \leq \mathcal{T}_{\text{cold}} \\ \alpha_{\text{warm}} + (\alpha_{\text{cold}} - \alpha_{\text{warm}}) \left(\dfrac{\mathcal{T}_{\text{warm}} - \mathcal{T}_s}{\mathcal{T}_{\text{warm}} - \mathcal{T}_{\text{cold}}} \right)^\eta & \text{if } \mathcal{T}_{\text{cold}} \leq \mathcal{T}_s \leq \mathcal{T}_{\text{warm}} \\ \alpha_{\text{warm}} & \text{if } \mathcal{T}_s \geq \mathcal{T}_{\text{warm}} \end{cases}$$

In the first panel of Figure 8.56 we plot the albedo function $\alpha_p(\mathcal{T}_s)$ when η is set to 3. The second panel shows the two forcing functions:

$$\begin{cases} F_{\text{solar}}(\mathcal{T}_s) = \dfrac{1}{4} \left(1 - \alpha_p((\mathcal{T}_s)) \right) S_0 \\ F_{\text{blackbody}}(\mathcal{T}_s) = \left(\dfrac{2 - \varepsilon}{2} \right) \sigma \mathcal{T}_s^4 \end{cases}$$

We then compare the net forcing function $F_{\text{solar}}(\mathcal{T}_s) - F_{\text{blackbody}}(\mathcal{T}_s)$ to the equation $\mathcal{E}(\mathcal{T}_s) = 0$

---

[58]We use the values $\alpha_{\text{cold}} = 0.7$, $\mathcal{T}_{\text{cold}} = 260 \, \text{K}$, $\alpha_{\text{warm}} = 0.3$ and $\mathcal{T}_{\text{warm}} = 295 \, \text{K}$.

Figure 8.56: Equilibrium states of the Budyko ice-albedo model
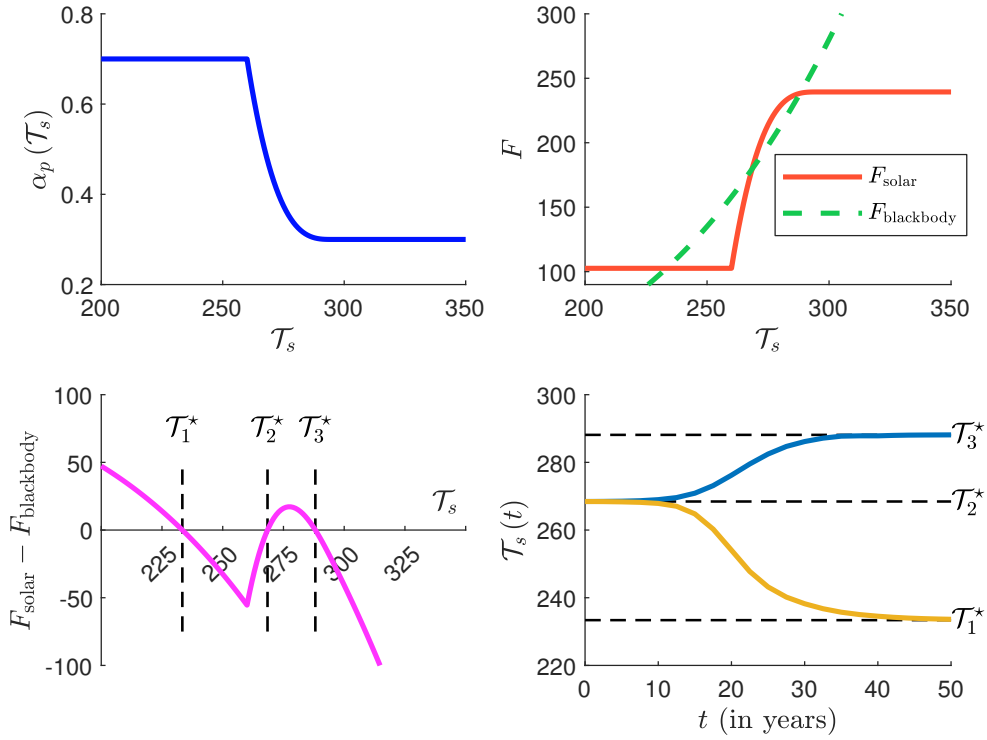
in the third panel. We can see graphically that there are three equilibrium states. By solving the equation $F_{\text{solar}}(\mathcal{T}_s) - F_{\text{blackbody}}(\mathcal{T}_s) = 0$ with a bisection algorithm, we find that the three equilibria are $\mathcal{T}_1^\star = 233.38\,\text{K}$, $\mathcal{T}_2^\star = 268.43\,\text{K}$, and $\mathcal{T}_3^\star = 288.13\,\text{K}$. To see if the equilibria are stable, we compute the total feedback:

$$\lambda(\mathcal{T}_s) = \lambda_0(\mathcal{T}_s) + \lambda_{\alpha_p}(\mathcal{T}_s)$$

where:

$$\lambda_0(\mathcal{T}_s) = -4\left(\frac{2-\varepsilon}{2}\right)\sigma\mathcal{T}_s^3$$

and:

$$\lambda_{\alpha_p}(\mathcal{T}_s) = \frac{1}{4}\eta S_0(\alpha_{\text{cold}} - \alpha_{\text{warm}})\frac{(\mathcal{T}_{\text{warm}} - \mathcal{T}_s)^{\eta-1}}{(\mathcal{T}_{\text{warm}} - \mathcal{T}_{\text{cold}})^\eta} \cdot \mathbb{1}\{\mathcal{T}_{\text{cold}} \leq \mathcal{T}_s \leq \mathcal{T}_{\text{warm}}\}$$

The results are shown below:

State	Temperature		$\lambda_0(\mathcal{T}_s)$	$\lambda_{\alpha_p}(\mathcal{T}_s)$	$\lambda(\mathcal{T}_s)$
$\mathcal{T}_1^\star$	$233.38\,\text{K}$	$-39.77°\text{C}$	-1.76	0.00	-1.76
$\mathcal{T}_2^\star$	$268.43\,\text{K}$	$-4.72°\text{C}$	-2.68	6.75	$+4.08$
$\mathcal{T}_3^\star$	$288.13\,\text{K}$	$14.98°\text{C}$	-3.31	0.45	-2.86

We conclude that $\mathcal{T}_1^\star$ and $\mathcal{T}_3^\star$ are two stable equilibria because λ is negative, but $\mathcal{T}_2^\star$ is an unstable equilibrium. This means that if the climate system is in equilibrium $\mathcal{T}_2^\star$, any perturbation will cause instability and the temperature will not return to that equilibrium. To illustrate this instability, we consider the dynamics of the temperature:

$$c\frac{\mathrm{d}\mathcal{T}_s}{\mathrm{d}t} = F_{\text{solar}}(\mathcal{T}_s) - F_{\text{blackbody}}(\mathcal{T}_s) + \Delta F$$

where $c = 4 \times 10^8 \, \mathrm{J\,m^{-2}\,K^{-1}}$ is the heat capacity and ΔF is the perturbation. We assume that the system is in equilibrium $\mathcal{T}_s(0) = \mathcal{T}_2^\star$. We introduce a small perturbation $\Delta F = +0.1 \, \mathrm{W/m^2}$ and solve the above ordinary differential equation numerically. The solution is reported in the fourth panel of Figure 8.56. In this case, we have:

$$\lim_{t \to \infty} \mathcal{T}_s(t) = \mathcal{T}_3^\star$$

If the perturbation $\Delta F = -0.1 \, \mathrm{W/m^2}$, we get:

$$\lim_{t \to \infty} \mathcal{T}_s(t) = \mathcal{T}_1^\star$$

We check that $\mathcal{T}_2^\star$ is an unstable equilibrium. The previous analysis concerns the sensitivity to the initial conditions. We can also consider the sensitivity to the parameters. For example, we can show that there is only one solution if $\eta < 1.9235$, while there are three solutions if $\eta > 1.9236$. Furthermore, there exists a parameter $\eta \in [1.9235, 1.9236]$ such that we get two solutions. This type of analysis is called bifurcation analysis.

Tipping point definition The concept of a tipping point emerged from stability analysis and bifurcation. The term was popularized in 2000 by Malcolm Gladwell in his book "*The Tipping Point: How Little Things Can Make a Big Difference*", which explored the concept in sociological change. In climate change, it was popularized in the late 2000s by Lenton *et al* (2008). In common parlance, a tipping point is a change. When applied to climate change, the term usually has a negative connotation and is generally associated with a change that could have devastating consequences for human society. However, when we look at scientific publications on climate change, a tipping point is a neutral mathematical concept. For example, here are the definitions adopted by Timothy Lenton and IPCC:

> "*The term tipping point commonly refers to a critical threshold at which a tiny perturbation can qualitatively alter the state or development of a system.*" (Lenton *et al*, 2008, page 1786).

> "*A climate tipping point occurs when a small change in forcing triggers a strongly nonlinear response in the internal dynamics of part of the climate system, qualitatively changing its future state.*" (Lenton, 2011, page 201).

> "*Tipping points refer to critical thresholds in a system that, when exceeded, can lead to a significant change in the state of the system, often with an understanding that the change is irreversible.*" (IPCC, 2018, page 262).

> In climate, a tipping point is "*a hypothesized critical threshold when global or regional climate changes from one stable state to another stable state. The tipping point event may be irreversible.*" (IPCC, 2021, page 1463).

In fact, these scientific definitions contrast with the content of these scientific publications, because a tipping point generally announces catastrophic events (Kemp *et al.*, 2022). For example, the term "*climate endgame*" refers to a hypothetical scenario in which climate change becomes catastrophic for human civilization and seems to be the Ten Plagues of Egypt with a series of super disasters: collapse of ecosystems, mass extinctions, global food shortages, water scarcity, etc.

Bifurcation theory   Lenton (2011, Figure 2, page 203) gives a heuristic interpretation of the global warning when approaching a bifurcation point with a ball in a landscape of valleys and mountains. In Figures 8.57-8.59 we use an analogous story where the valleys represent stable attractors and the ball represents the state of the system. Figure 8.57 shows a stable equilibrium corresponding to the green ball. If the ball moves away from the equilibrium, but not too far, it will return to the equilibrium. This situation corresponds to the two red balls. This illustrates a reversible behavior where the system can retrace its path and return to its initial state. Figure 8.58 shows a red ball representing a tipping point. In fact, this equilibrium is unstable and corresponds to a bifurcation point, since the red ball cannot remain in this position indefinitely. If the system is initially at the red ball and is then slightly perturbed by a small disturbance, it will move toward the green balls and settle there. However, it is unpredictable which direction the red ball will take. It depends on many parameters. The red arrows represent the transitions between equilibria and indicate the direction between the tipping point and a stable equilibrium. Figure 8.59 shows the new position of the ball after it has moved to the left valley. Again the ball is in equilibrium. However, when it leaves this equilibrium, it cannot return to its initial position as shown in 8.59. In fact, the system has changed because part of the mountain has filled the valley. Instead, there is a new valley, but it's higher than the old one. Therefore, when the green ball leaves its equilibrium, it can move to the position represented by the yellow ball, but it cannot reach the position represented by the grey ball because it no longer exists.

To present the bifurcation theory of one-dimensional dynamical systems, we use chapters 2 and 3 of the classic textbook on bifurcation theory and chaos by Strogatz (2015). We consider the dynamical system of the form:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = f(x, \mu)$$

where μ is the parameter set. We look at the mathematical properties of the system by analyzing the trajectories (or flows[59]) of $x(t)$. If $f(x, \mu) > 0$, then x increases with time, while if $f(x, \mu) < 0$, then x decreases with time. A fixed point $x^\star$ is then a value of x such that the system does not change with time, *i.e.* $f(x^\star, \mu) = 0$. Let $x = x^\star + \Delta x$ where Δx is small. We have:

$$
\begin{aligned}
\frac{\mathrm{d}x}{\mathrm{d}t} &= f(x^\star + \Delta x, \mu) \\
&= f(x^\star, \mu) + \frac{\partial f(x^\star, \mu)}{\partial x}\Delta x + O(\Delta x^2) \\
&= f(x^\star, \mu) + \lambda \Delta x + \mathcal{O}(\Delta x^2)
\end{aligned}
$$

where λ is the feedback of the system. Since we have:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{\mathrm{d}(x^\star + \Delta x)}{\mathrm{d}t} = \frac{\mathrm{d}\Delta x}{\mathrm{d}t}$$

and:

$$f(x^\star + \Delta x, \mu) \approx f(x^\star, \mu) + \lambda \Delta x = \lambda \Delta x$$

We deduce that:

$$\frac{\mathrm{d}\Delta x}{\mathrm{d}t} = \lambda \Delta x$$

If $\lambda > 0$, then any small perturbation of the fixed point grows exponentially, while if $\lambda < 0$, then any small perturbation of the fixed point decays exponentially. In the first case the fixed point is

---

[59]A flow is used here as a synonym for a trajectory, and corresponds to a vector field that gives the direction and magnitude of the trajectory.

Figure 8.57: Stable equilibria
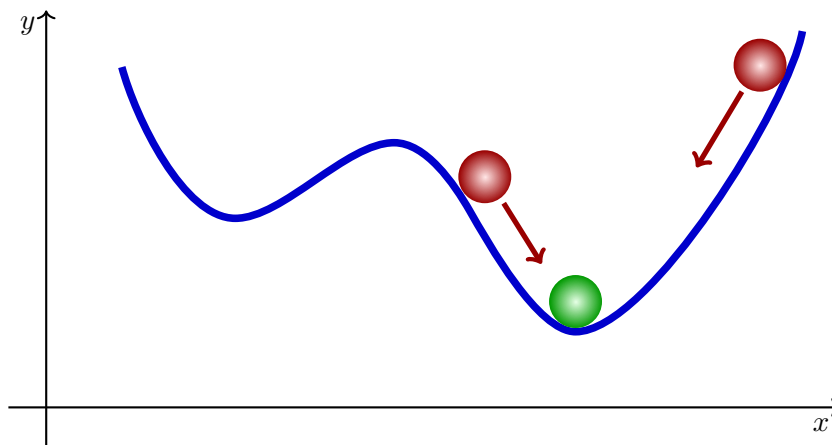
Figure 8.58: Unstable equilibrium
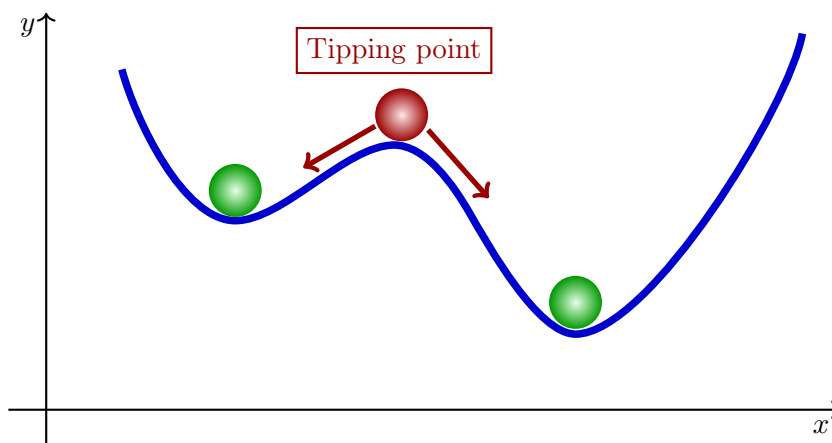
Tipping point

Figure 8.59: Irreversible tipping point
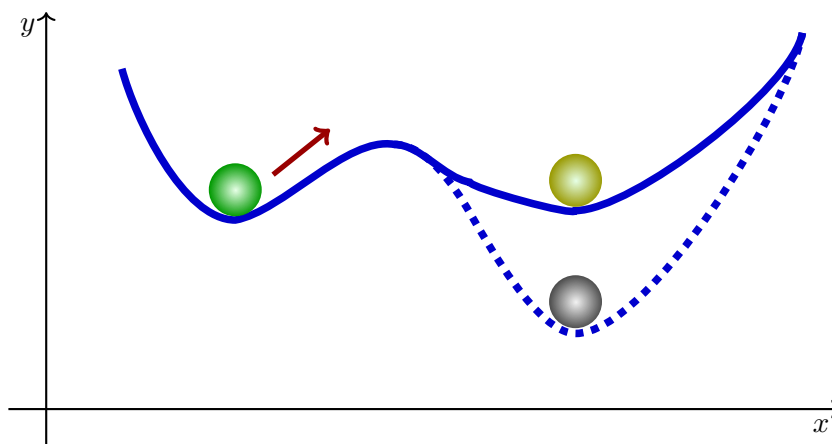
unstable, in the second case it is stable. When $\lambda = 0$, we need to study the second derivative of $f(x, \mu)$ to conclude. We also note that the relaxation timescale is equal to:

$$\tau = \frac{1}{|\partial_x f(x^\star, \mu)|}$$

Example 19 *Consider the dynamical system* $\dfrac{\mathrm{d}x}{\mathrm{d}t} = x^2 + \mu$*. We have* $\partial_x f(x, \mu) = 2x$*. The fixed points are solutions of the equation* $x^2 = -\mu$*. If* $\mu > 0$*, there are no fixed points. If* $\mu = 0$*, the fixed point is* $x^\star = 0$ *and is not stable. If* $\mu < 0$*, there are two fixed points.* $x_1^\star = -\sqrt{-\mu}$ *is stable while* $x_2^\star = \sqrt{-\mu}$ *is unstable.*

The stability analysis evaluates the behavior of $f(x^\star + \varepsilon, \mu)$ and checks if the point $x^\star + \Delta x$ converges to the point $x^\star$. We can consider a second approach to stability where we evaluate the behavior of $f(x^\star, \mu + \varepsilon)$, *i.e.* we apply the perturbation not directly to x but to the control parameter μ. In particular, we say that the value $\mu^\star$ is a bifurcation value if $f(x, \mu^\star)$ is not structurally stable. In the previous example, the dynamical system exhibits a bifurcation that occurs at $\mu = 0$. This type of bifurcation is called a saddle-node or fold bifurcation, because fixed points are created or destroyed. The first panel of Figure 8.60 shows the bifurcation diagram of the system. As μ increases, the distance between the stable and unstable fixed points decreases, and the equilibria are destroyed when μ is strictly positive. Following Strogatz (2015), other types of bifurcation are also common. A transcritical bifurcation occurs when fixed points exchange stability for a critical value of μ. For example, the system $\dfrac{\mathrm{d}x(t)}{\mathrm{d}t} = \mu x - x^2$ has a transcritical bifurcation at $\mu = 0$ (second panel of Figure 8.60). A pitchfork bifurcation occurs when the system is unchanged and exhibits symmetry when $x = -x$. There are two forms of pitchfork bifurcation:

1. In a supercritical pitchfork bifurcation, a stable fixed point becomes unstable at a critical value of μ. A canonical example is:
$$\frac{\mathrm{d}x}{\mathrm{d}t} = \mu x - x^3$$
   The corresponding bifurcation diagram is shown in the third panel of Figure 8.60. We observe that $\mu = 0$ is a supercritical pitchfork bifurcation.

2. In a subcritical pitchfork bifurcation, an unstable fixed point becomes stable at a critical value of μ. A canonical example is:
$$\frac{\mathrm{d}x}{\mathrm{d}t} = \mu x + x^3 - x^5$$
   The corresponding bifurcation diagram is shown in the fourth panel of Figure 8.60. Let $\mu^\star = -0.25$ be the first value of μ that exhibits fixed points. When $\mu < \mu^\star$, there is a stable fixed point at $x^\star = 0$ because $\partial_x f(x^\star, \mu) < 0$. When $\mu = \mu^\star$, two new equilibria are created. When $\mu^\star < \mu \leq 0$, there are five fixed points. Three of these are stable equilibria. When $\mu > 0$, the stable equilibrium $x^\star = 0$ is destroyed and jumps to the stable branches or becomes unstable.

In bifurcation theory, hysteresis refers to a phenomenon where the behavior of the system depends on the history of its past states. For example, consider the previous subcritical pitchfork bifurcation and suppose the system is at equilibrium $x^\star = 0$. If we increase μ, the system jumps to one of the stable branches. If we decrease μ, the system does not return to its past equilibrium, but stays on the stable branch. However, if we decrease μ even more, the system jumps to its past equilibrium $x^\star = 0$ when μ reaches $\mu^\star = -0.25$. Therefore, the path of the system has formed a hysteresis loop.

Figure 8.60: Bifurcation diagram
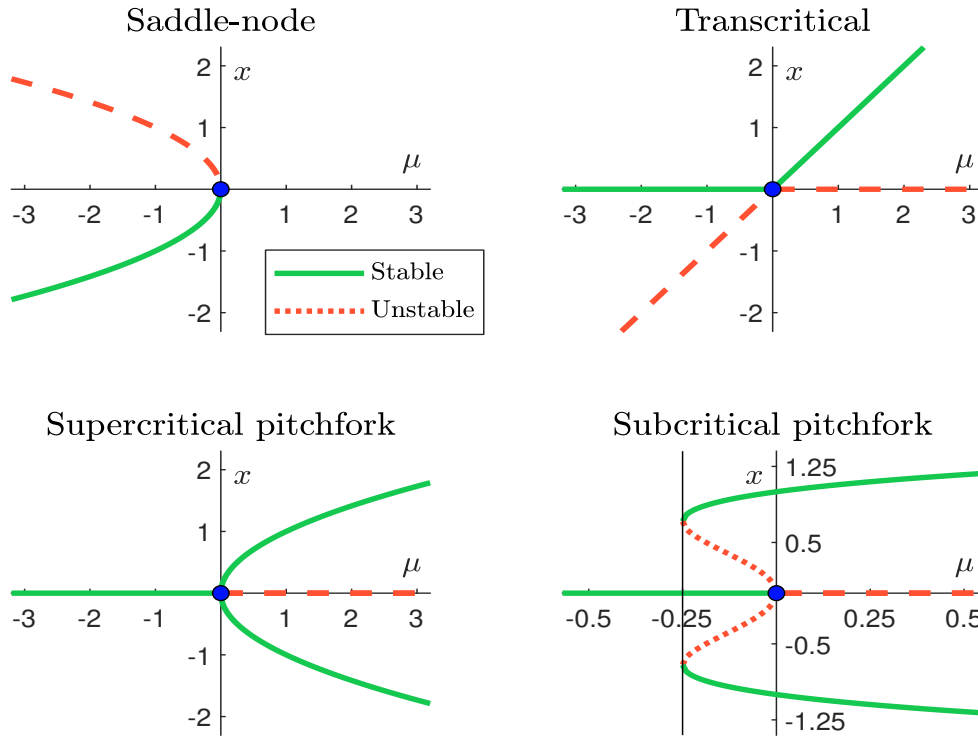
Remark 54 *Chaos theory is closely related to bifurcation theory because chaos theory is primarily concerned with the behavior of dynamical systems that are highly sensitive to initial conditions. Since chaos is synonymous with disorder, a deterministic chaotic system exhibits aperiodic long-term behavior that appears to be random. These include strange attractors, feedback loops, limit cycles, self-similarity, fractals, butterfly effects, irregular oscillations, etc. The most famous chaotic models are the Hénon map, the Lorenz system, the logistic map, and the Rössler attractor described in Box 8.10.*

Let us return to the Budyko ice-albedo model described on page 331. Using the default values $S_0 = 1\,368\,\text{W/m}^2$, $\varepsilon = 78\%$, $\eta = 3$, $\alpha_{\text{cold}} = 0.7$, $\mathcal{T}_{\text{cold}} = 260\,\text{K}$, $\alpha_{\text{warm}} = 0.3$ and $\mathcal{T}_{\text{warm}} = 295\,\text{K}$, we have obtained three fixed points: $\mathcal{T}_1^\star = 233.38\,\text{K}$, $\mathcal{T}_2^\star = 268.43\,\text{K}$, and $\mathcal{T}_3^\star = 288.13\,\text{K}$. $\mathcal{T}_1^\star$ and $\mathcal{T}_3^\star$ were stable, while $\mathcal{T}_3^\star$ was unstable. To perform the bifurcation analysis of the Budyko ice-albedo model, we consider the range $\mathcal{T}_s \in [200\,\text{K}, 300\,\text{K}]$ and divide the range into ten intervals. Using the bisection algorithm, we solve the equation for each interval of $\mathcal{T}_s$ and each value of the parameter of interest:

$$\frac{1}{4}\left(1 - \alpha_p\left(\left(\mathcal{T}_s\right)\right)\right)S_0 - \left(\frac{2 - \varepsilon}{2}\right)\sigma \mathcal{T}_s^4 = 0$$

We collect all fixed points $\{\mathcal{T}_1^\star, \mathcal{T}_2^\star, \ldots\}$. For each fixed point we calculate the feedback:

$$\lambda\left(\mathcal{T}_s\right) = \frac{1}{4}\eta S_0\left(\alpha_{\text{cold}} - \alpha_{\text{warm}}\right)\frac{\left(\mathcal{T}_{\text{warm}} - \mathcal{T}_s\right)^{\eta - 1}}{\left(\mathcal{T}_{\text{warm}} - \mathcal{T}_{\text{cold}}\right)^\eta} \cdot \mathbb{1}\left\{\mathcal{T}_{\text{cold}} \leq \mathcal{T}_s \leq \mathcal{T}_{\text{warm}}\right\} - 4\left(\frac{2 - \varepsilon}{2}\right)\sigma \mathcal{T}_s^3$$

We perform this analysis for a given range of the parameter of interest and plot the fixed points $\{\mathcal{T}_1^\star, \mathcal{T}_2^\star, \ldots\}$ against the parameter of interest. We also use a solid green line if the fixed point is stable or $\lambda\left(\mathcal{T}_s\right) < 0$, and a dashed red line otherwise. Results are shown in Figure 8.61, when the

Box 8.10: Rössler chaotic model

The Rössler model is a system of ordinary differential equations that exhibits chaotic dynamics (Rössler, 1976). It has become one of the most studied and widely used examples of a chaotic system. The Rössler equations are given by:

$$\begin{cases} \dfrac{dx(t)}{dt} = -(y(t) + z(t)) \\ \dfrac{dy(t)}{dt} = x(t) + ay(t) \\ \dfrac{dz(t)}{dt} = b + x(t)z(t) - cz(t) \end{cases}$$

where a, b and c are three parameters that control the behavior of the system. The Rössler attractor is a three-dimensional surface described by $(x(t), y(t), z(t))$. The attractor is very sensitive to the initial conditions $(x(0), y(0), z(0))$ and the set of parameters (a, b, c). Below we plot the attractor in the first panel and the bifurcation diagram of $x(t)$ with respect to each parameter in the other three panels. We assume that the initial condition is $(x(0), y(0), z(0)) = (1, 1, 0)$ and the parameters[a] are $a = 0.2$, $b = 0.2$ and $c = 5.7$. We check that the dynamic system $(x(t), y(t), z(t))$ does not necessarily converge to a fixed point and is highly sensitive to the parameters. While the Rössler model is deterministic, its behavior is unpredictable like a stochastic model.
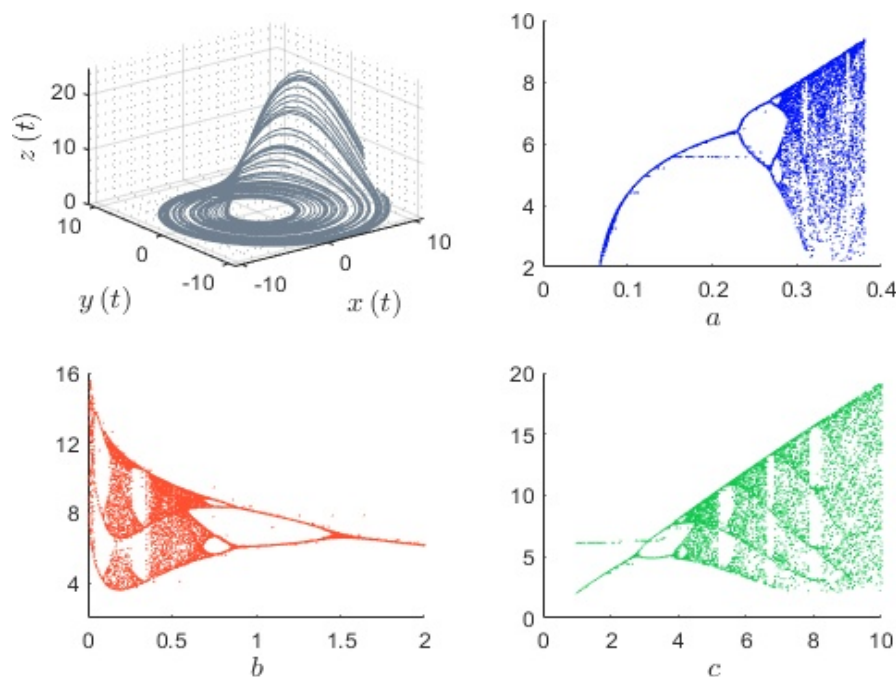
[a]For the bifurcation diagram of a, we set $b = 2$.

parameters of interest are η, ε, S_0 and α_{cold}. We can make several remarks. First, we note that we can have one, two, or three fixed points, meaning that there is at least one equilibrium and no more than three equilibria for a set of parameters. Second, one of these fixed points is stable and it always generates a saddle-node bifurcation since two new equilibria are created. Third, the unstable fixed point is always between the two stable fixed points. Figure 8.62 shows the relaxation timescale of the different equilibria:

$$\tau = \frac{c}{|\lambda(\mathcal{T}_s)|}$$

Figure 8.61: Bifurcation of the Budyko ice-albedo model
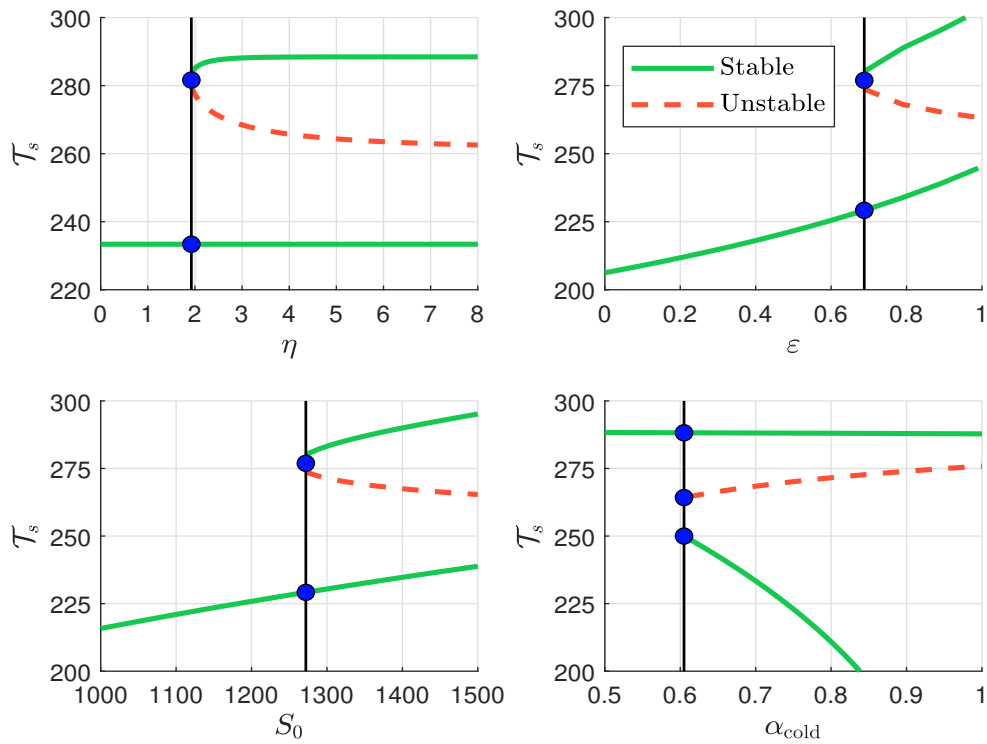
Figure 8.62: Relaxation timescale of the Budyko ice-albedo model (in years)
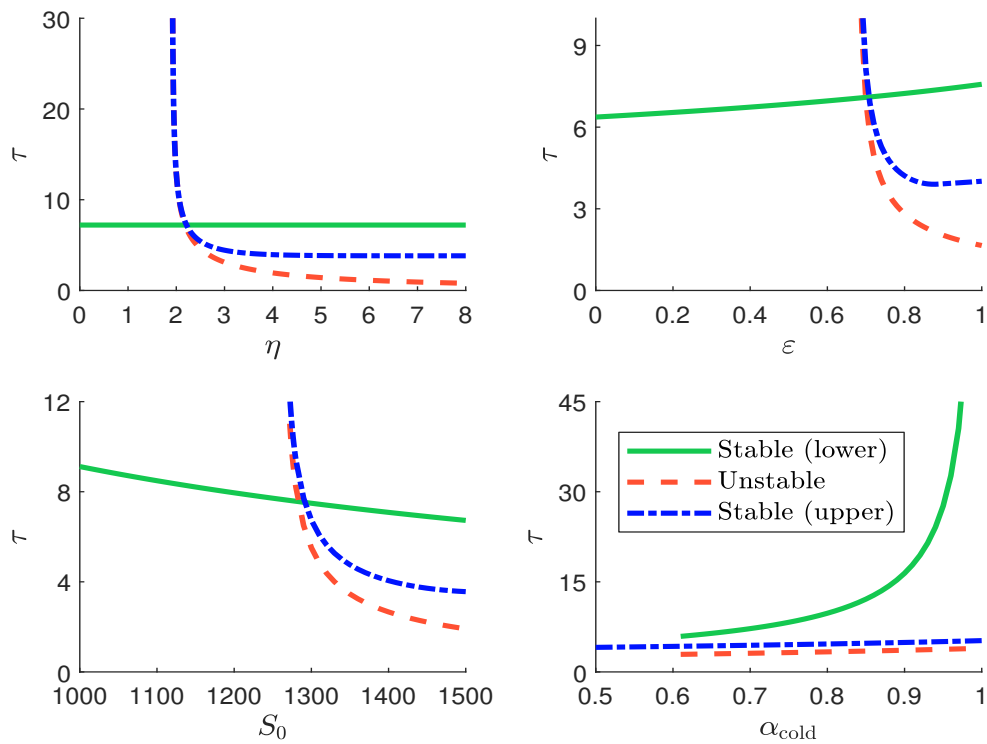
where $c = 4 \times 10^8 \, \text{J}\,\text{m}^{-2}\,\text{K}^{-1}$ is the heat capacity. τ is expressed in years. Note that τ is less than 50 years, which is not realistic. In fact, the framework we have used to model the ice-albedo feedback is very simplistic and far from reality.

Figure 8.63: Bifurcation and overshooting tipping points

(a) Slow/fast tipping onset elements   (b) AMOC overshooting
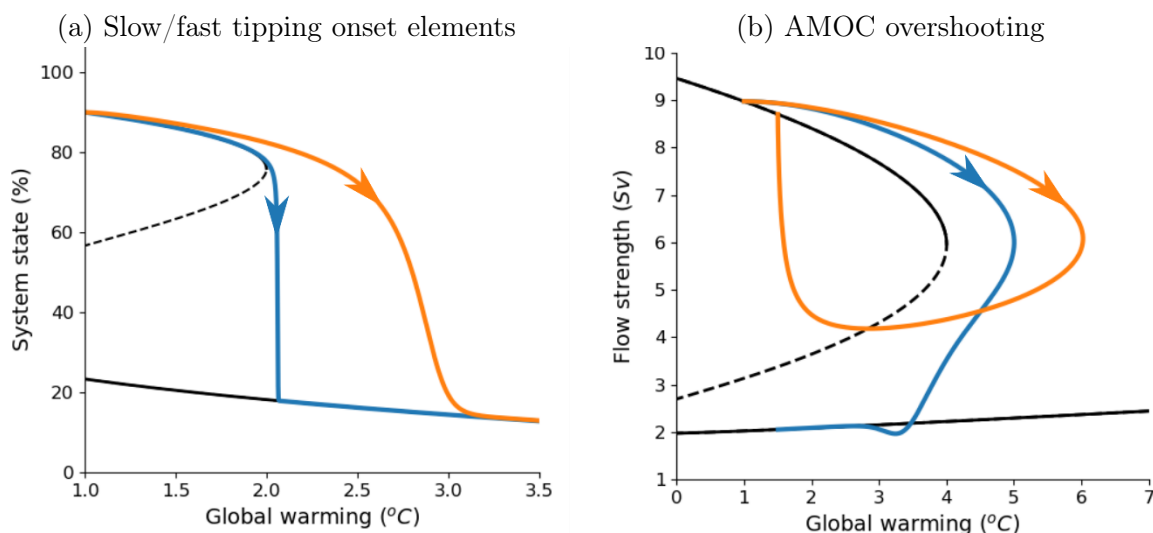
Source: Ritchie *et al.* (2021, Figures 1c & 3c, pages 518 & 520) &
https://www.nature.com/articles/s41586-021-03263-2.

The Budyko ice-albedo model produces a relatively classical bifurcation where equilibria are created or destroyed at a bifurcation point. To obtain a more complex bifurcation diagram, we need to use a higher dimensional model or nonlinear feedback relationships. In Figure 8.63 we show two examples calculated by Ritchie *et al.* (2021). The first panel compares slow and fast overshooting tipping points and the importance of the radiative timescale[60]. The blue curve represents a rapidly overshooting tipping point system, such as the dieback of the Amazon forest, while the orange curve shows that the transition can take a long time even after we have passed the tipping point, such as the collapse of the Atlantic meridional overturning circulation (AMOC). This means that some tipping points may have already been reached, but we will not see their effects until the next century or millennium. The second panel shows the behavior of the AMOC when implementing an overshoot[61]. Two different trajectories are considered, one with a small overshoot and long recovery time (blue curve) and one with a large overshoot and short recovery time (orange curve). Contrary to expectations, the small overshoot leads to a persistent collapse of the AMOC, while the large overshoot allows a recovery. The resilience of the AMOC in this second scenario can be explained by the rapid reversal of global warming.

---

[60]Here is the description of the figure provided by the authors: "*System state vs global warming for the fast onset tipping element (blue) and slow onset tipping element (orange). Both systems have a desired state, which is the upper solid black curve and represent contemporary conditions. An undesired stable state, given by the lower solid black curve, coexists with the desired state for warming levels below the $2°$C threshold, separated by an unstable state (black dashed curve). Above the threshold, only the undesired equilibrium state remains.*" (Ritchie *et al.*, 2021, page 518).

[61]This figure shows the flow strength of the AMOC against global warming for short long and large fast overshoots: "*The blue curve is a small and long overshoot, while the orange curve is a much larger yet quick overshoot. An AMOC on state (upper solid black curve) and an AMOC off state (lower solid black curve) both coexist for warming levels below the threshold of $4°$C and are separated by an unstable state (black dashed curve). Above the threshold only the AMOC off state remains.*" (Ritchie *et al.*, 2021, page 520).

Climate tipping elements Lenton *et al* (2008) introduced the term "*tipping element*" to describe large-scale components of the Earth system that may pass a tipping point. They identified 9 climate tipping elements: (1) Arctic sea ice, (2) Greenland ice sheet, (3) West Antarctic ice sheet, (4) Atlantic thermohaline circulation, (5) El Niño Southern Oscillation (ENSO), (6) Indian summer monsoon, (7) Sahara/Sahel and West African monsoons, (8) Amazon rainforest, and (9) boreal forest. They added to this list other candidate tipping elements that did not meet all the criteria for a tipping point: (10) Antarctic bottom water, (11) tundra, (12) permafrost, (13) marine methane hydrates, (14) ocean anoxia, and (15) Arctic ozone. The above short list of climate tipping elements has evolved over time (Lenton, 2011; Steffen *et al.*, 2018; Lenton *et al*, 2019). Today, there is a consensus on the most important tipping points, which has been published by Armstrong McKay *et al.* (2022). We report this short list in Table 8.12, while Figure 8.64 shows their location on the Earth.

Figure 8.64: Geographical distribution of global and regional tipping elements
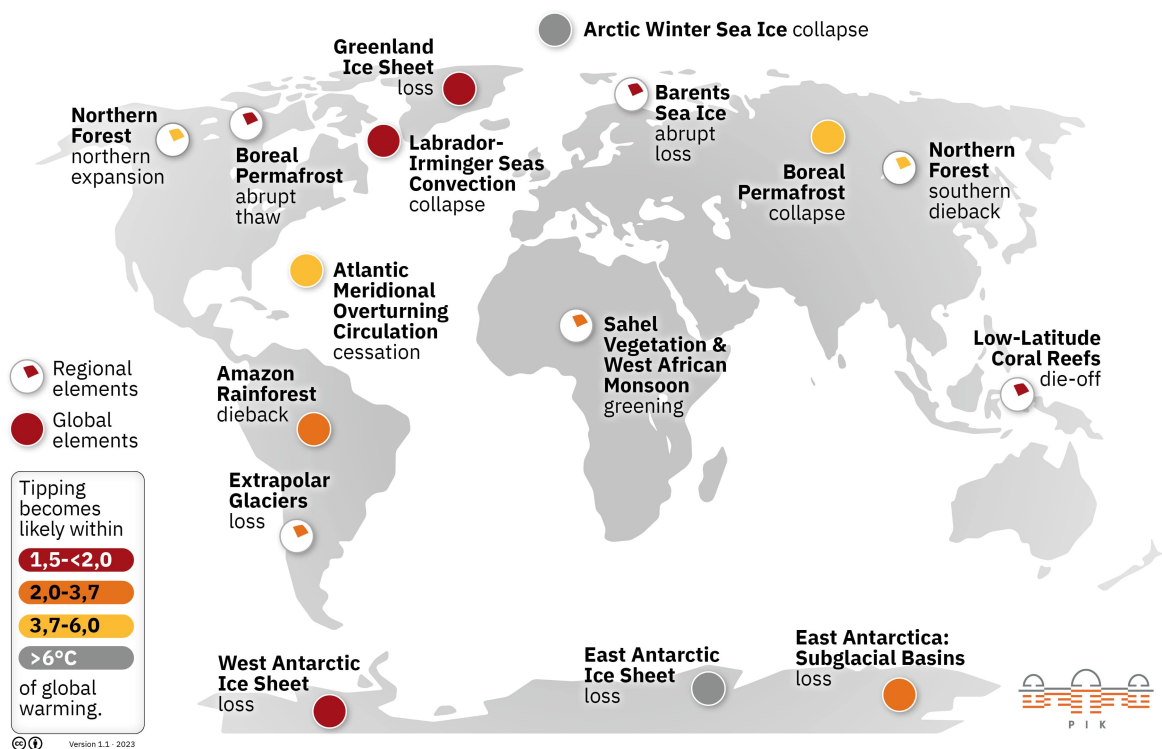
Armstrong McKay *et al.* (2022) distinguish 9 global core and 7 regional impact tipping elements that affect the cryosphere, ocean-atmosphere, and biosphere. For each climate tipping point, they estimate the threshold θ, the timescale τ and the maximum impact δ. This means that the tipping point is reached when the temperature anomaly is greater than θ. The result is a new equilibrium that induces a temperature change of δ after τ years:

$$\Delta \mathcal{T}_s(0) \geq \theta \Rightarrow \text{tipping point} \Rightarrow \Delta \mathcal{T}_s(\tau) = \delta$$

For example, if $\theta = +1°C$, $\tau = 50$ years and $\delta = 0.2°C$, the tipping point is crossed when the temperature increase is 1°C. The new equilibrium will be reached after 50 years with a supplementary temperature increase of 0.2°C. The plausible range of these estimates is given in parentheses.

Table 8.12: Threshold, timescale, and impact estimates for the global and regional tipping elements

Category	Climate tipping element	Tipping point	Threshold		Timescale		Maximum impact	
	Greenland ice sheet	collapse	1.5°C	(0.8-3.0)	10 kyr	(1-15)	0.13°C	(0.5-3.0)
	West Antarctic ice sheet	collapse	1.5°C	(1.0-3.0)	2 kyr	(0.5-13)	0.05°C	(1.0)
	Labrador-Irminger seas/SPG convection	collapse	1.8°C	(1.1-3.8)	10 yr	(5-50)	−0.50°C	(-3.0)
	East Antarctic subglacial basins	collapse	3.0°C	(2.0-6.0)	2 kyr	(0.5-10)	0.05°C	
Global	Amazon rainforest	dieback	3.5°C	(2.0-6.0)	100 yr	(50-200)	0.20°C	(0.4-2.0)
	Boreal permafrost	collapse	4.0°C	(3.0-6.0)	50 yr	(10-300)	0.40°C	
	AMOC	collapse	4.0°C	(1.4-8.0)	50 yr	(15-300)	−0.50°C	(-4/-10)
	Arctic winter sea ice	collapse	6.3°C	(4.5-8.7)	20 yr	(10-100)	0.60°C	(0.6-1.2)
	East Antarctic ice sheet	collapse	7.5°C	(5.0-10.0)	> 10 kyr		0.60°C	(2.0)
	Low-latitude coral reefs	die-off	1.5°C	(1.0-2.0)	10 yr			
	Boreal permafrost	abrupt thaw	1.5°C	(1.0-2.3)	200 yr	(100-300)	0.04°C	
	Barents sea ice	abrupt loss	1.6°C	(1.5-1.7)	25 yr			
Regional	Mountain glaciers	loss	2.0°C	(1.5-3.0)	200 yr	(50-1000)	0.08°C	
	Sahel and West African monsoon	greening	2.8°C	(2.0-3.5)	50 yr	(10-500)		
	Boreal forest (southern)	die-off	4.0°C	(1.4-5.0)	100 yr	(50+)	−0.18°C	(-0.5/-2)
	Boreal forest (northern)	expansion	4.0°C	(1.5-7.2)	100 yr	(40+)	0.14°C	(0.5-1.0)

Source: Armstrong McKay et al. (2022, Table 1, page 3).

Below is a list of the different tipping elements and some explanations of the mechanism behind them:

1. The Greenland ice sheet is the second largest ice sheet in the world. It covers 80% of the surface of Greenland. Its melting would increase sea level rise, possibly up to 7.4 meters, accelerate ocean acidification, and have a potentially positive feedback effect on climate change (Shepherd et al., 2020).

2. The West Antarctic ice sheet is a large ice sheet in Antarctica. It sits on a bedrock that is mostly below sea level and has formed a deep subglacial basin due to the weight of the ice sheet, which can be up to 4 kilometers thick in places. Its collapse could raise global sea levels, possibly up to 3 meters (Alley et al., 2015).

3. The Labrador and Irminger Seas are located in the subpolar North Atlantic, between Canada, Greenland and Iceland. These seas are characterized by cold and salty waters that generate deep convection. This deep convection and the regulation of ocean salinity influence the circulation of the Atlantic meridional overturning circulation (AMOC). The collapse of the deep convection system in the Labrador and Irminger Seas would affect the overall circulation in the North Subpolar Gyre (Chafik et al., 2022).

4. East Antarctic subglacial basins are large, ice-filled depressions in the bedrock adjacent to the East Antarctic ice sheet. They also serve as reservoirs for meltwater. Certain subglacial basins, such as Wilkes, Aurora, and Recovery, are more susceptible to a situation called marine ice sheet instability (MISI). As the ice shelf at the edge of the ice sheet retreats, warm ocean water can flow into the deeper basin, further destabilizing the ice shelf. This process can create a self-perpetuating cycle of ice melt and ice shelf retreat (Stokes et al., 2022).

5. The Amazon rainforest, also known as the Amazon jungle or Amazonia, is the largest rainforest in the world. It spans nine countries: Brazil, Peru, Colombia, Ecuador, Bolivia, Venezuela, Guyana, Suriname, and French Guiana. It contains the largest and most biodiverse area of tropical rainforest in the world. The Amazon rainforest acts as a massive carbon sink, absorbing and storing vast amounts of carbon dioxide through the process of photosynthesis. If the forest were to be subject to widespread deforestation or degradation, the stored carbon could be released back into the atmosphere, contributing to increased greenhouse gas concentrations. The Amazon also plays a critical role in the Earth's water cycle, influencing regional and global weather patterns. Its dense vegetation effectively captures rainwater and slowly releases it into streams and rivers, helping to maintain stable water levels, prevent flooding, and provide a steady source of fresh water (Gatti et al., 2021).

6. Boreal permafrost is a permanently frozen layer of soil and rock that underlies much of the world's boreal forest. It is found in Siberia, Alaska, northern Canada, and the Tibetan Plateau. Permafrost forms when the soil temperature remains below 0°C for at least two consecutive years. Boreal permafrost contains large amounts of organic carbon stored in the form of dead plant material that could not decompose due to the cold temperatures. It is also one of the largest reservoirs of methane. Rising temperatures may cause the boreal permafrost to thaw, releasing large amounts of carbon and methane into the atmosphere (Schuur et al., 2015).

7. The Atlantic meridional overturning circulation (AMOC) is a large, complex system of ocean currents that transports warm water from the tropics to the North Atlantic and cold water from the North Atlantic to the subtropics. It is also known as the Gulf Stream system. A weakening

of the AMOC could have complex and regionally specific effects on temperatures. On a global scale, it could result in less warm water reaching higher latitudes, leading to cooler sea surface temperatures in the North Atlantic and warmer temperatures in the Southern Hemisphere (Liu *et al.*, 2017).

8. Arctic winter sea ice is the maximum extent of sea ice that forms in the Arctic Ocean during the winter months. It helps to regulate global temperatures by reflecting sunlight back into space. This albedo reflection helps to cool the Arctic. As sea ice melts, more sunlight is absorbed by the ocean, causing a further warming trend[62]. However, the impact of the albedo effect remains controversial (Bathiany *et al.*, 2016).

9. The East Antarctic ice sheet is the largest and thickest ice sheet on Earth. A complete collapse would raise the global sea levels by 50 meters. However, the East Antarctic ice sheet is generally considered to be more stable than the West Antarctic ice sheet, due to its higher elevation and more remote location (Stokes *et al.*, 2022).

10. Low-latitude coral reefs occur in the Atlantic, Indian, and Pacific Oceans, most notably in the Philippines, Indonesia, and Australia. They require warm, sunny weather and unpolluted water. Therefore, coral reefs can be affected by climate change, although their impact on climate change is more limited (Descombes *et al.*, 2015).

11. We have already seen that the boreal permafrost is a global tipping element, but it is also a regional tipping element. In fact, an abrupt thaw of the boreal permafrost would have devastating consequences for the region, affecting infrastructure (roads, buildings, transportation), the environment (flooding, forests, vegetation), and living conditions and health (Hjort *et al.*, 2022).

12. Barents sea ice is found in the Barents Sea, an arm of the Arctic Ocean between Norway and Russia. The sea ice forms during the winter months and melts during the summer months. This regional tipping element is strongly related to two global tipping elements: Labrador-Irminger Seas and AMOC (Moore *et al.*, 2022).

13. Mountain glaciers are large masses of ice that form on mountains at high altitudes. They are formed from compacted snow that has accumulated over many years. The melting of mountain glaciers would have a major regional impact on human life (Carey, 2010).

14. The West African monsoon is a seasonal wind pattern that affects the Sahel, bringing moisture from the Atlantic Ocean during the rainy season and drying out the region during the dry season. It is responsible for the region's agriculture and supports the livelihoods of millions of people. Changes in rainfall can affect vegetation, agriculture and people (Pausata *et al.*, 2020).

15. The boreal forest, also known as the taiga, is a biome that surrounds the Arctic region. Countries with significant areas of boreal forest include Canada, Russia, Sweden, Norway and Finland. The southern edge of the boreal forest is the boundary between the boreal forest and temperate forests or grasslands. The risk could be an abrupt die-off (Venålåinen *et al.*, 2020).

16. The northern edge of the boreal forest is typically found at higher latitudes, closer to the Arctic Circle. The change could be an abrupt expansion into a tundra forest characterised by treeless landscapes and permafrost (Girona *et al.*, 2023).

---

[62]Moreover, the warming Arctic has the potential to release methane from permafrost.

Almost all the climate tipping elements induce a temperature increase, except Labrador-Irminger Seas/SPG convection, AMOC and the southern edge of the boreal forest. We also find some large differences in timescale (between 10 years and more than 10 000 years) and impact. According to Armstrong McKay *et al.* (2022), the Greenland and West Antarctic ice sheets have certainly passed their tipping points, and other tipping points such as the die-off of tropical coral reefs and boreal permafrost are close to crossing their threshold.

Cascading tipping points and climate domino effects

Mathematical framework of coupling tipping points   Klose *et al.* (2020) assumes that the climate tipping elements follow an ordinary differential equation with a double fold bifurcation:

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = f(x_i, \mu_i) = \frac{\alpha_i x_i - \beta_i x_i^3 + \mu_i}{\tau_i} \tag{8.16}$$

where x_i is the ith tipping element, τ_i is the timescale, $\alpha_i > 0$ and $\beta_i > 0$ are two fixed parameters[63] and μ_i is the bifurcation parameter. Note that the cubic equation $f(x_i, \mu_i) = 0$ has three roots (real and complex). The feedback is equal to:

$$\lambda = \frac{\partial f(x_i, \mu_i)}{\partial x_i} = \frac{\alpha_i - 3\beta_i x_i^2}{\tau_i} \tag{8.17}$$

The number of stable equilibria depends on the discriminant $\Delta = \frac{1}{4}(\beta_i \mu_i)^2 - \frac{1}{27}\beta_i \alpha_i^3$. If $\Delta > 0$, there is one stable equilibrium, while $\Delta < 0$ exhibits two stable equilibria. Therefore, the bifurcation occurs when $\Delta = 0$ or:

$$\mu_i = \pm\sqrt{\frac{4\alpha_i^3}{27\beta_i}}$$

Figure 8.65 shows the bifurcation diagram of the system when[64] $\alpha_i = \beta_i = 1$. If μ_i is strongly negative, there is only one stable equilibrium. A first bifurcation occurs at $\mu_i = -\sqrt{\frac{4}{27}}$ and two more equilibria are created, one stable and another one unstable. Then two equilibria are destroyed when μ_i reaches the value $\sqrt{\frac{4}{27}}$ and $x_i^\star = -\sqrt{\frac{1}{3}}$. They are the lower stable equilibrium and the middle unstable equilibrium. Only the upper equilibrium remains when μ_i is greater than $\sqrt{\frac{4}{27}}$.

Let $\alpha_i = \beta_i = 1$ and $\mu_i = 0.5$. We assume that the current state of the system is $x_i(0) = -1$. Figure 8.66 shows the convergence of $x_i(t)$ to equilibrium for three values of the timescale. Whatever the value of τ_i, we have $\lim_{t\to\infty} x_i(t) = 1.1915$. The equilibrium is reached in seven years when $\tau_i = 1$, while we need more than 70 years when $\tau_i = 10$. In the case $\tau_i = 10$, convergence to equilibrium takes between 300 and 400 years.

To introduce the cascading effects between the tipping elements, Klose *et al.* (2020) assume that x_i depends on the other tipping elements x_j with $j \neq i$:

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = f_i(x_i, \mu_i) + c_i(x_{-i}) = g_i(x_1, \ldots, x_n, \mu_i)$$

---

[63]According to Klose *et al.* (2020), α_i corresponds to the distance between the upper and lower layers of stable equilibria of the cusp, while β_i controls the strength of the nonlinearity in the system.

[64]Note that the bifurcation diagram does not depend on the timescale τ_i.

Figure 8.65: Double fold bifurcation
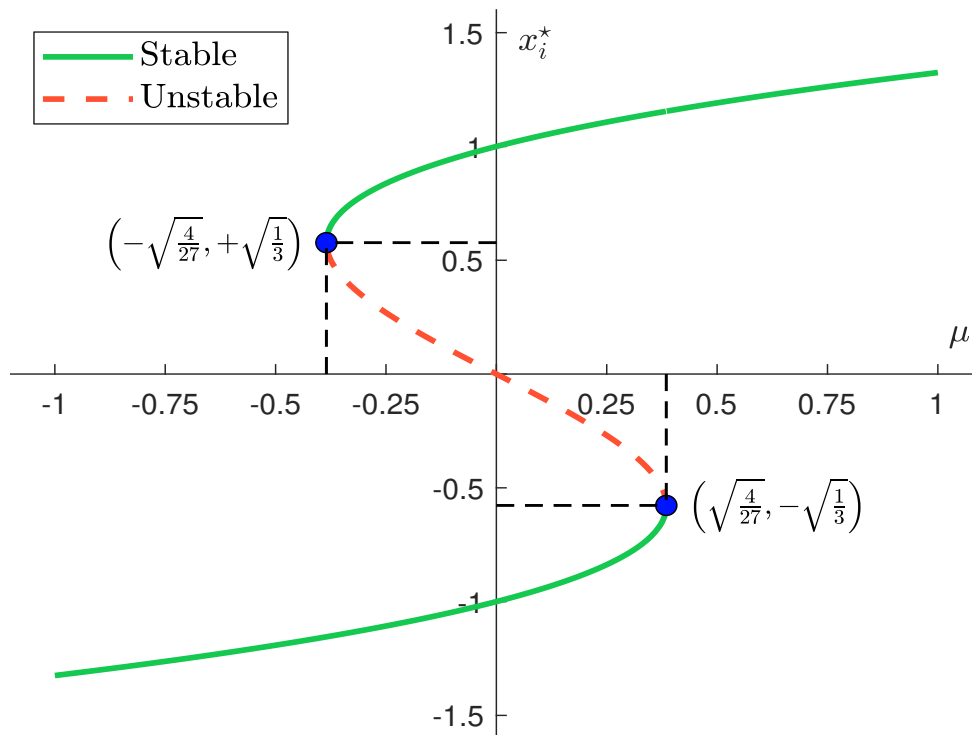
Source: Klose *et al.* (2020, Figure 1, page 3).

Figure 8.66: Convergence to the equilibrium
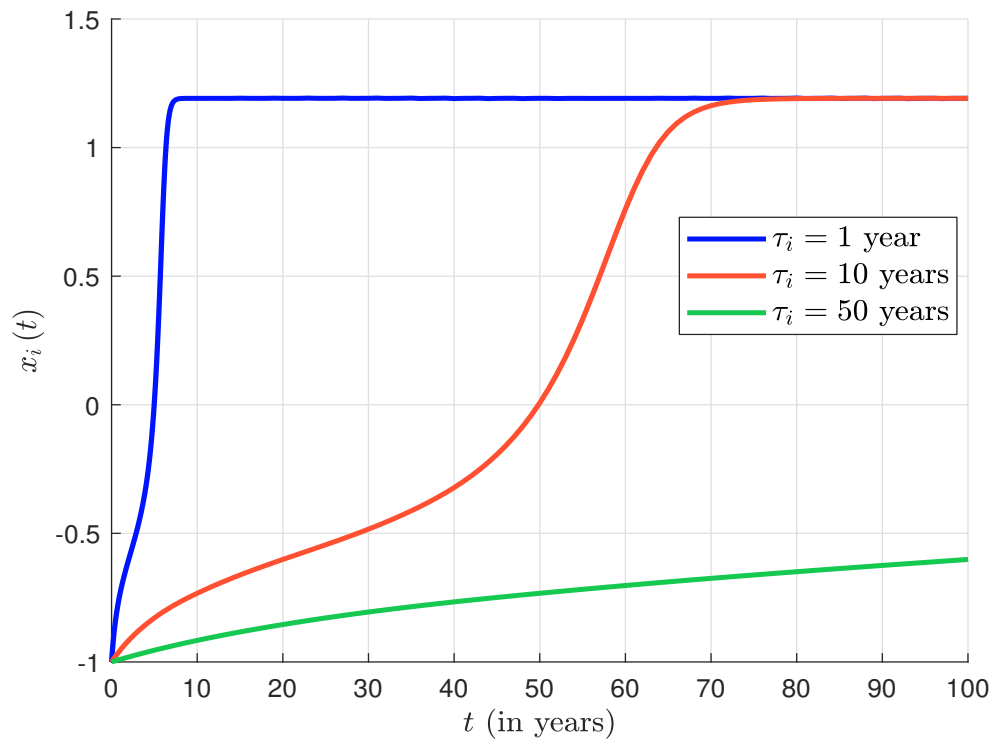
where $c_i(x_{-i})$ is the coupling function. In particular, they study the linear case where $c_i(x_{-i}) = \tau_i^{-1} \sum_{j \neq i} \gamma_{i,j} x_j$ and $\gamma_{i,j}$ measures the influence of subsystem x_j on subsystem x_i. Therefore, Equation (8.16) becomes:

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = \frac{\alpha_i x_i - \beta_i x_i^3 + \mu_i + \sum_{j \neq i} \gamma_{i,j} x_j}{\tau_i} \qquad (8.18)$$

We may wonder how the coupling function affects the tipping points. For example, in the two-dimensional case we have:

$$\begin{cases} \dfrac{\mathrm{d}x_1}{\mathrm{d}t} = \dfrac{\alpha_1 x_1 - \beta_1 x_1^3 + \mu_1 + \gamma_{1,2} x_2}{\tau_1} \\ \dfrac{\mathrm{d}x_2}{\mathrm{d}t} = \dfrac{\alpha_2 x_2 - \beta_2 x_2^3 + \mu_2 + \gamma_{2,1} x_1}{\tau_2} \end{cases}$$

Sed commodo posuere pede. Mauris ut est. Ut quis purus. Sed ac odio. Sed vehicula hendrerit sem. Duis non odio. Morbi ut dui. Sed accumsan risus eget odio. In hac habitasse platea dictumst. Pellentesque non elit. Fusce sed justo eu urna porta tincidunt. Mauris felis odio, sollicitudin sed, volutpat a, ornare ac, erat. Morbi quis dolor. Donec pellentesque, erat ac sagittis semper, nunc dui lobortis purus, quis congue purus metus ultricies tellus. Proin et quam. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent sapien turpis, fermentum vel, eleifend faucibus, vehicula eu, lacus.

An application with Greenland and West Antarctic ice sheets, AMOC and Amazon rainforest　　Using the previous framework developed by Klose *et al.* (2020), Wunderling *et al.* (2021) analyze the dynamics of four tipping climate elements, namely (1) Greenland ice sheet, (2) West Antarctic ice sheet, (3) the AMOC and (4) the Amazon rainforest. The dynamics of the tipping elements are defined as:

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = \frac{1}{\tau_i}\left(x_i - x_i^3 + \sqrt{\frac{4}{27}}\frac{\Delta\mathcal{T}_s}{\mathcal{T}_i^\star} + \frac{\gamma}{10}\sum_{j\neq i}s_{i,j}\left(x_j + 1\right)\right)$$

where $\Delta\mathcal{T}_s$ is the variation of the global mean surface temperature, $\mathcal{T}_i^\star$ is the critical temperature anomaly to reach a tipping point, γ is the overall interaction strength and $s_{i,j}$ is the strength coupling between x_i and x_j. Using the previous notations, we have:

$$\begin{cases} \alpha_i = 1 \\ \beta_i = 1 \\ \mu_i = \sqrt{\frac{4}{27}}\frac{\Delta T_s}{T_i^\star} + \frac{\gamma}{10}\sum_{j\neq i}s_{i,j} \\ \gamma_{i,j} = \frac{\gamma}{10}s_{i,j} \end{cases}$$

Figure 8.67: Interactions between climate tipping elements and their roles in tipping cascades
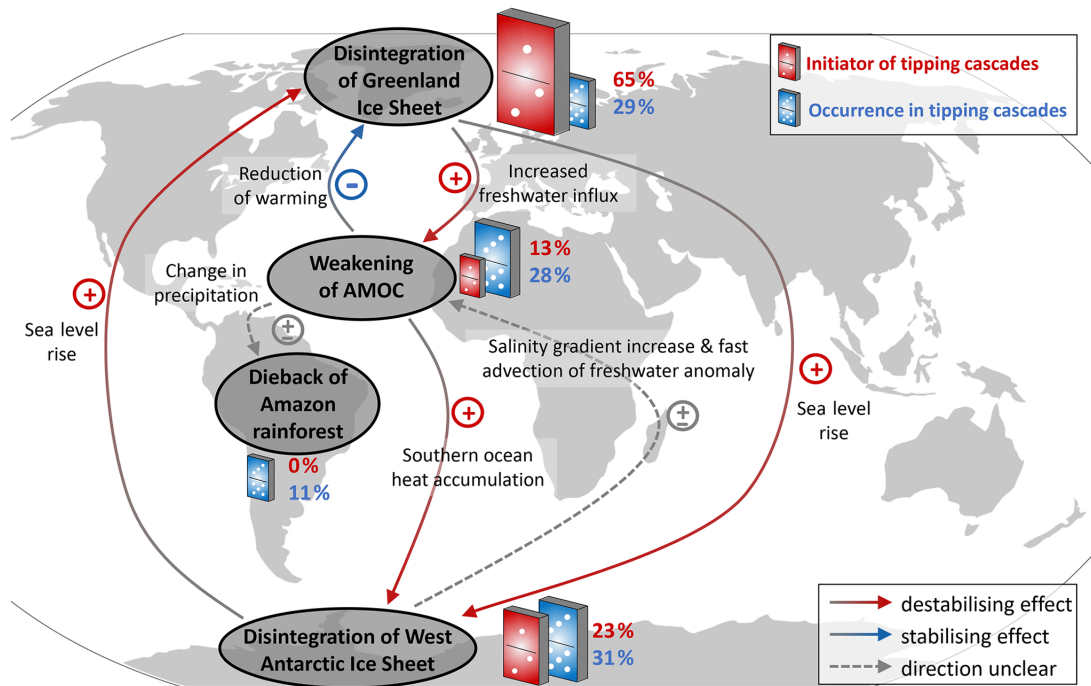
Etiam euismod. Fusce facilisis lacinia dui. Suspendisse potenti. In mi erat, cursus id, nonummy sed, ullamcorper eget, sapien. Praesent pretium, magna in eleifend egestas, pede pede pretium lorem, quis consectetuer tortor sapien facilisis magna. Mauris quis magna varius nulla scelerisque imperdiet. Aliquam non quam. Aliquam porttitor quam a lacus. Praesent vel arcu ut tortor cursus volutpat. In vitae pede quis diam bibendum placerat. Fusce elementum convallis neque. Sed dolor orci, scelerisque ac, dapibus nec, ultricies ut, mi. Duis nec dui quis leo sagittis commodo.

Remark 55 *This section on the physics of climate change provides a basic overview of the mechanisms behind climate change. It may be supplemented by Section 12.1 on Chapter 12 of Physical Risk Modeling, which is a comprehensive presentation of general circulation models.*

8.2   The ecosystem of climate change

8.2.1   Scientists

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula, urna sed ultricies auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur et nunc. Aliquam dolor odio, commodo pretium, ultricies non, pharetra in, velit. Integer arcu est, nonummy in, fermentum faucibus, egestas vel, odio.

8.2.2   Conferences of the parties

Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Donec odio elit, dictum in, hendrerit sit amet, egestas sed, leo. Praesent feugiat sapien aliquet odio. Integer vitae justo. Aliquam vestibulum fringilla lorem. Sed neque lectus, consectetuer at, consectetuer sed, eleifend ac, lectus. Nulla facilisi. Pellentesque eget lectus. Proin eu metus. Sed porttitor. In hac habitasse platea dictumst. Suspendisse eu lectus. Ut mi mi, lacinia sit amet, placerat et, mollis vitae, dui. Sed ante tellus, tristique ut, iaculis eu, malesuada ac, dui. Mauris nibh leo, facilisis non, adipiscing quis, ultrices a, dui.

8.2.3   Regulation policies

Morbi luctus, wisi viverra faucibus pretium, nibh est placerat odio, nec commodo wisi enim eget quam. Quisque libero justo, consectetuer a, feugiat vitae, porttitor eu, libero. Suspendisse sed mauris vitae elit sollicitudin malesuada. Maecenas ultricies eros sit amet ante. Ut venenatis velit. Maecenas sed mi eget dui varius euismod. Phasellus aliquet volutpat odio. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Pellentesque sit amet pede ac sem eleifend consectetuer. Nullam elementum, urna vel imperdiet sodales, elit ipsum pharetra ligula, ac pretium ante justo a nulla. Curabitur tristique arcu eu metus. Vestibulum lectus. Proin mauris. Proin eu nunc eu urna hendrerit faucibus. Aliquam auctor, pede consequat laoreet varius, eros tellus scelerisque quam, pellentesque hendrerit ipsum dolor sed augue. Nulla nec lacus.

8.3   Integrated assessment models

Integrated assessment models (IAMs) can be defined as approaches that link main features of society and economy with the biosphere and atmosphere into a common modeling framework. They generally couple a macroeconomic model with a climate risk model in order to simulate the economic impacts of climate change. IAMs are used by policy markers to analyze the economic cost of climate change and the impact of climate action. According to Nordhaus (2017b), the most important applications are:

1. Making projections of economic variables (*e.g.*, GDP) that take into account global warming;

2. Calculating the impacts of alternative assumptions;

3. Tracing the effects and estimating the costs and benefits of alternative climate policies;

4. Estimating the uncertainty of future economic pathways.

For instance, the US government uses three IAMs to estimate the social cost of greenhouse gases, that is the value of avoiding one tonne of GHG emissions (IWG, 2021). These models are Dynamic Integrated Climate and Economy (DICE), Climate Framework for Uncertainty, Negotiation, and Distribution (FUND) and Policy Analysis of the Greenhouse Gas Effect (PAGE). The number of IAMs has grown rapidly, especially these last years with the increasing awareness of scientists to take into account climate change into economic modeling. Already in the late 1990s, Kelly and Kolstad (1999) count 21 major integrated assessment models. Today, current IAMs may be very complex and integrate social dimensions (*e.g.*, inequality, education, health, food security), industry dimensions (*e.g.*, sectors, infrastructure, rare earth elements), biodiversity dimensions (*e.g.*, species, ecosystem, food), etc. In this section, we first present the DICE model, which is certainly one of the simplest IAMs, but also the most famous. Even if DICE is a highly stylized reference point, it is an excellent educational tool for understanding the economics and physics of IAMs.

8.3.1   The DICE model

There are many versions of DICE. William Nordhaus began to develop DICE with a simple energy/climate model in the 1970s (Nordhaus, 1977). The current format of the model can already be found in Nordhaus (1992). Since this publication in *Science*, William Nordhaus has multiplied the research projects on DICE. Some of them were published in academic journals, others in many books[65], but most of them were unpublished. In what follows, we use the presentation and the notations of Nordhaus and Sztorc (2013), which corresponds to the user's manual of the DICE 2013R software. We also extensively refer to the comprehensive survey of Le Guenedal (2019).

The 2013 model

DICE uses a standard neoclassical model of economic growth known as the Ramsey-Cass-Koopmans model. It is an extension of the Solow model when the saving rate is not constant but endogenous. The social planner maximizes then the welfare utility function and determines the optimal path of saving rates to increase global consumption in the future. Nordhaus introduces the impact of climate change as a negative externality that hurts the economy and the output. In order to mitigate the cost

---

[65]The most famous are *Managing the Global Commons: The Economics of Climate Change* (1994), *A Question of Balance: Weighing the Options on Global Warming Policies* (2008), *The Climate Casino: Risk, Uncertainty, and Economics for a Warming World* (2013) and *The Spirit of Green* (2021).

of climate physical risks, the social planner can increase climate investments via a control variable on climate transition risks. Therefore, the general equilibrium depends on two decision variables: the saving rate and the climate control variable.

Production and consumption functions Nordhaus and Sztorc (2013) assume that the gross production $Y(t)$ is given by the Cobb-Douglas function:

$$Y(t) = A(t) K(t)^\gamma L(t)^{1-\gamma}$$

where $A(t)$ is the total productivity factor (or technological progress), $K(t)$ is the capital input and $L(t)$ is the labor input. The parameter $\gamma \in]0,1[$ measures the elasticity of the capital factor:

$$\gamma = \frac{\partial \ln Y(t)}{\partial \ln K(t)} = \frac{\partial Y(t)}{\partial K(t)} \frac{K(t)}{Y(t)}$$

Traditional economic models do not make the distinction between the production $Y(t)$ and the net output $Q(t)$ because we have the identity $Y(t) = Q(t)$. Nevertheless, physical and transition climate risks generate losses:

$$Q(t) = \Omega_{\text{climate}}(t) Y(t) \leq Y(t)$$

where $\Omega_{\text{climate}}(t) \in]0,1[$ is the loss percentage of the production. $Q(t)$ is then the net output when taking into account negative externalities of climate change. The saving rate $s(t)$ is assumed to be time-dependent and is a control variable. Using the classical identities $Q(t) = C(t) + I(t)$ and $I(t) = s(t) Q(t)$ where $I(t)$ is the investment, the expression of the consumption $C(t)$ is then equal to:

$$\begin{aligned} C(t) &= (1 - s(t)) Q(t) \\ &= (1 - s(t)) \Omega_{\text{climate}}(t) A(t) K(t)^\gamma L(t)^{1-\gamma} \end{aligned} \tag{8.19}$$

In order to introduce the time dependence and complete the economic model, the authors assume that the dynamics of the state variables are:

$$\begin{cases} A(t) = (1 + g_A(t)) A(t-1) \\ K(t) = (1 - \delta_K) K(t-1) + I(t) \\ L(t) = (1 + g_L(t)) L(t-1) \end{cases} \tag{8.20}$$

where $g_A(t)$ is the growth rate of the technological progress, δ_K is the depreciation rate of the capital stock and $g_L(t)$ is the growth rate of the labor factor. The two growth rates $g_A(t)$ and $g_L(t)$ decline over time, implying that:

$$g_A(t) = \frac{1}{1 + \delta_A} g_A(t-1)$$

and:

$$g_L(t) = \frac{1}{1 + \delta_L} g_L(t-1)$$

where $\delta_A \geq 0$ and $\delta_L \geq 0$.

Example 20 *The world population was equal to 7.725 billion in 2019 and 7.805 billion in 2020. At the beginning of the 1970s, we estimate that the annual growth rate was equal to 2.045%. According to the United Nations, the global population could surpass 10 billion by 2100.*

In 2020, the annual growth rate was equal to:

$$g_L\left(2020\right) = \frac{L\left(2020\right)}{L\left(2019\right)} - 1 = \frac{7.805}{7.725} - 1 = 1.036\%$$

Since we have $g_L\left(t\right) = \left(\dfrac{1}{1+\delta_L}\right)^{t-t_0} g_L\left(t_0\right)$, we deduce that:

$$\delta_L = \left(\frac{g_L\left(t_0\right)}{g_L\left(t\right)}\right)^{1/(t-t_0)} - 1$$

An estimate of δ_L is then:

$$\delta_L = \left(\frac{g_L\left(1970\right)}{g_L\left(2020\right)}\right)^{1/30} - 1 = 2.292\%$$

In Figure 8.68, we report the dynamics of the labor input $L\left(t\right)$ by assuming that $L\left(2020\right) = 7.805$, $g_L\left(2020\right) = 1.036\%$ and $\delta_L = 2.292\%$. We estimate that the global population could reach 11.4 billion by 2100. In order to measure the sensitivity to the model parameters, we have also estimated the world population when δ_L is equal to 1.50% and 3.25%. The range is 10.5 and 12.6 billion at the end of this century.

Figure 8.68: Evolution of the labor input $L\left(t\right)$
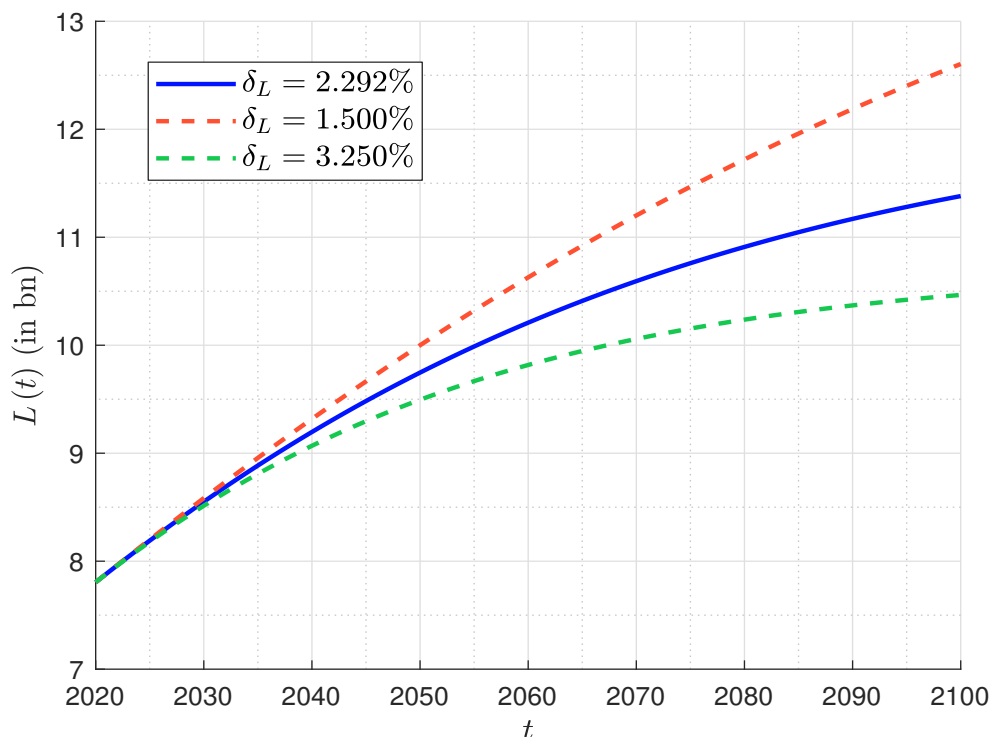
In Figure 8.69, we compare the different probabilistic projections, which are computed by the Department of Economic and Social Affairs (DESA) of the United Nations[66]. We notice that the previous calibration is close to the median estimation. Of course, we can use these data in order to better calibrate the DICE model. For instance, if we have a time series of $g_L\left(t\right)$, we can estimate δ_L

---

[66]The data source are available at `https://population.un.org/wpp`.

Figure 8.69: Projection of the world population
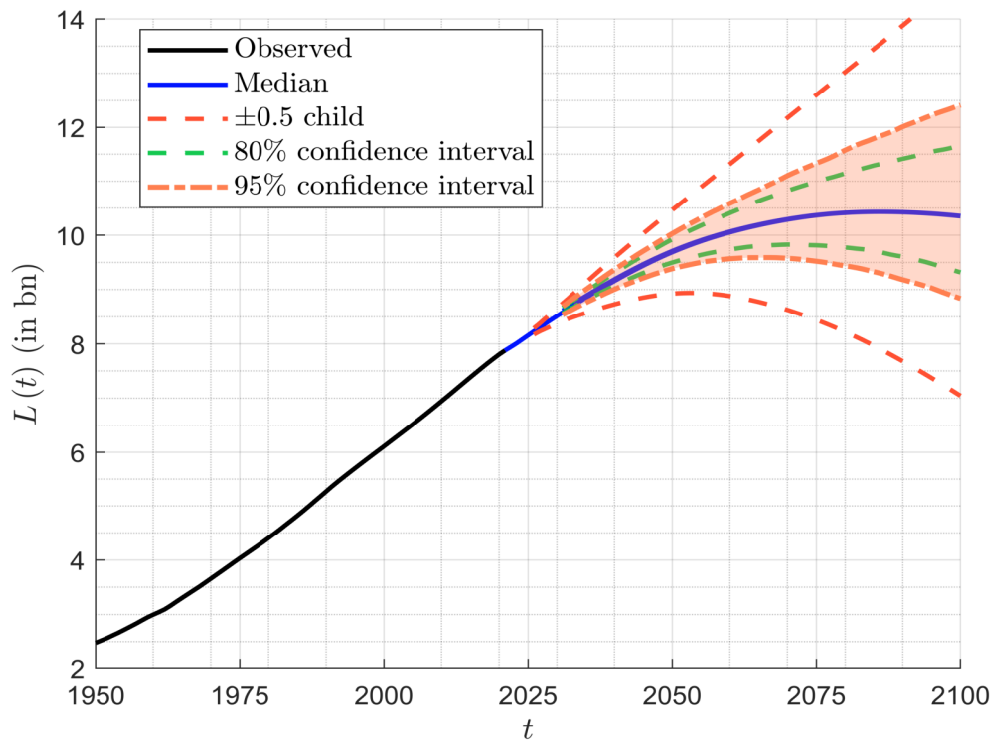
Figure 8.70: Population growth rate
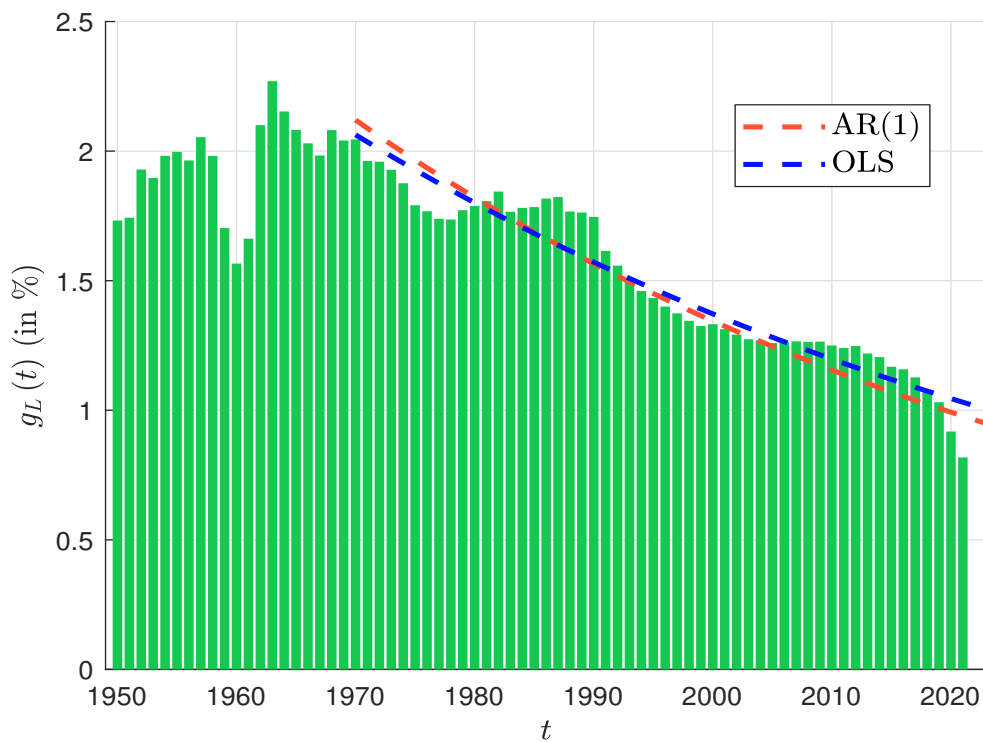
by considering the AR(1) process: $g_L(t) = \phi g_L(t-1) + \varepsilon(t)$. We have $\hat{\delta}_L = \left(1 - \hat{\phi}\right)/\hat{\phi}$. Another method consists in using the following relationship:

$$g_L(t) = \left(\frac{1}{1 + \delta_L}\right)^{t - t_0} g_L(t_0)$$

We deduce that $\ln g_L(t) = \ln g_L(t_0) - (t - t_0)\ln(1 + \delta_L)$. We can then estimate the parameter δ_L by considering the linear regression: $\ln g_L(t) = \beta_0 + \beta_1(t - t_0) + \varepsilon(t)$. It follows that $\hat{\delta}_L = e^{-\hat{\beta}_1} - 1$. Using the historical population growth rates between 1970 and 2021 provided by the United Nations (Figure 8.70), we obtain $\hat{\phi} = 0.985$, $\hat{\beta}_0 = 0.724$ and $\hat{\beta}_1 = -0.014$. The estimated value of δ_L is then equal to 1.529% for the AR(1) process and 1.369% for the log-linear model. In Figure 8.70, we observe a high decline of the population growth rate in the last years, since the last value $g_L(2021)$ was equal to 0.82%.

Table 8.13: Average productivity growth rate (in %)

Country	1960-1970	1970-1980	1980-1990	1990-2000	2000-2010	2010-2020
AUS	1.02	0.07	-0.23	1.02	0.36	0.13
BRA	2.39	2.05	-1.04	-1.12	-0.17	-1.63
CAN	2.18	0.38	-0.25	0.21	-0.21	0.40
CHN	-0.03	-0.06	-0.04	-0.41	2.24	-0.35
FRA	3.59	1.63	1.12	0.61	-0.11	0.02
DEU	2.33	1.63	0.75	1.52	0.01	0.74
IND	2.37	-1.22	1.06	1.04	0.70	1.89
ITA	3.71	1.66	-0.19	-0.20	-1.32	-0.34
JPN	4.05	0.77	1.09	-0.22	-0.15	0.69
ZAF	2.37	0.30	-0.84	-1.11	0.50	-1.20
GBR	0.50	0.72	0.75	0.42	0.12	0.08
USA	1.00	0.42	0.46	0.73	0.65	0.56

Source: Penn World Table 10.01 (Feenstra *et al.*, 2015) & Author's calculations.

Concerning the total factor productivity (TFP), also called the multifactor productivity (MFP), Nordhaus and Sztorc (2013) estimated that $g_A(2010) = 1.53\%$ and $\delta_A = 0.12\%$. If we would like to recalibrate these figures, there are currently four main databases that provide TFP/MFP statistics[67]: the OECD Productivity Statistics, the EUKLEMS–INTANProd database, the Conference Board Total Economy Database, and the Penn World Table (PWT), which is certainly the most known framework among economists (Feenstra *et al.*, 2015). We use this last database and consider a sample of twelve countries. The TFP index is reported in Figure 8.71. We notice a high discrepancy between countries and also between periods (see Table 8.13). Based on these statistics, $g_A(2020)$ is closer to 0.75% if we include developing countries. If we assume that the TFP growth rate has been divided by a factor d in n years, we have $\delta_A = \sqrt[n]{d} - 1$. Some examples of the growth rate path $g_A(t)$ are given in Figure 8.72.

For the investment $I(t)$, the capital stock $K(t)$ and the gross output $Y(t)$, we can use the previous databases such as the Penn World Table, or the investment and capital stock dataset (ICSD)

---

[67]We can download them at www.oecd.org/sdd/productivity-stats, https://euklems-intanprod-llee.luiss.it, www.conference-board.org/data/economydatabase/total-economy-database-productivity and www.rug.nl/ggdc/productivity/pwt.

Figure 8.71: Total factor productivity index (base 100 = 1960)

Figure 8.72: Dynamics of the TFP growth rate
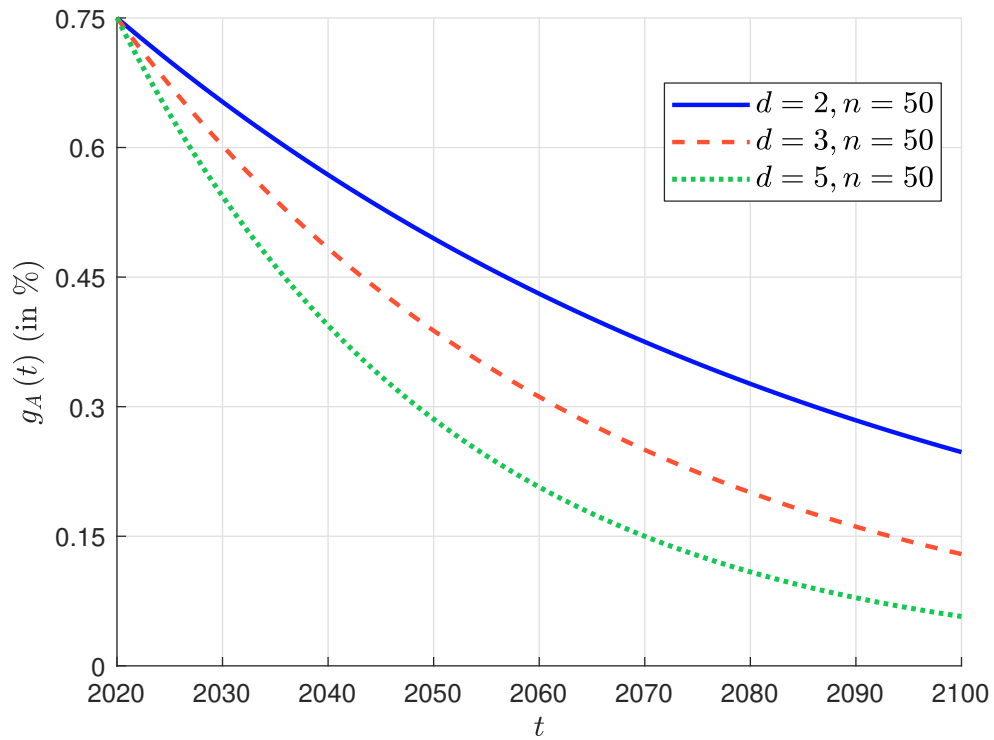
of the International Monetary Fund[68]. We have reported the evolution of the variables[69] in Figure 8.73. In 2019, we obtain $I(2019) = \$30.625$ tn, $K(2019) = \$318.773$ tn and $Y(2019) = \$124.418$ tn. We have also calculated the annual depreciation rate[70]:

$$\delta_K(t) = \frac{K(t-1) - K(t) + I(t)}{K(t-1)}$$

and we obtain $\delta_K(2019) = 6.25\%$. We have now all the elements to calibrate the initial value of $A(t)$. For that, we assume that γ takes the typical value of 0.30 and we obtain:

$$A(2019) = \frac{Y(t)}{K(t)^\gamma L(t)^{1-\gamma}} = \frac{124.418}{318.773^{0.30} \times 7.725^{0.70}} = 5.276$$

We can now solve the macroeconomic model in the absence of climate change effects. In this case, $\Omega_{\text{climate}}(t) = 1$ and $Q(t) = Y(t)$. By assuming that the saving rate $s(t)$ is constant[71], we obtain the simulations[72] in Figure 8.74. We recall that the variables $C(t)$, $I(t)$, $K(t)$ and $Y(t)$ are expressed in trillions of constant 2017 international dollars. If $s(t) = 25\%$, the GDP in 2100 would be equal to $\$507.47$ tn. It would be reduced to $\$408.65$ tn if $s(t) = 15\%$.

Cost function of climate change   We have seen that the net output is reduced because of climate change. The survival function is given by:

$$\Omega_{\text{climate}}(t) = \Omega_D(t)\,\Omega_\Lambda(t) = \frac{1}{1 + D(t)}(1 - \Lambda(t))$$

where $D(t) \geq 0$ is the climate damage function (physical risk) and $\Lambda(t) \geq 0$ is the mitigation or abatement cost (transition risk). The costs $D(t)$ result from natural disasters and climatic events, such as wildfires, storms, and floods, whereas the costs $\Lambda(t)$ are incurred from reducing GHG emissions and result from policy for financing the transition to a low-carbon economy. Nordhaus and Sztorc (2013) assume that $D(t)$ depends on the atmospheric temperature $\mathcal{T}_{\text{AT}}(t)$:

$$D(t) = \psi_1 \mathcal{T}_{\text{AT}}(t) + \psi_2 \mathcal{T}_{\text{AT}}(t)^2$$

where $\psi_1 \geq 0$ and $\psi_2 \geq 0$ are two parameters. $\mathcal{T}_{\text{AT}}(t)$ measures the global mean surface temperature and corresponds to the temperature increase in °C from 1900. Therefore, $\mathcal{L}_D(t) = 1 - \Omega_D(t) = 1 - (1 + D(t))^{-1}$ represents the fraction of net output that is lost because of the global warming. For the abatement cost function, the authors consider that it depends on the control variable $\mu(t)$:

$$\Lambda(t) = \theta_1(t)\,\mu(t)^{\theta_2}$$

where $\theta_1(t) \geq 0$ and $\theta_2 \geq 0$ are two parameters. For the emission-control rate, we have $\mu(t) \in [0,1]$. Finally, we deduce that the global impact of climate change is equal to:

$$\Omega_{\text{climate}}(t) = \frac{1 - \theta_1(t)\,\mu(t)^{\theta_2}}{1 + \psi_1 \mathcal{T}_{\text{AT}}(t) + \psi_2 \mathcal{T}_{\text{AT}}(t)^2} \tag{8.21}$$

---

[68]The database is available at https://data.imf.org.

[69]$I(t)$ is the sum of the general government investment (variable `igov_rppp`), the private investment (variable `ipriv_rppp`) and public-private partnership investment (variable `ippp_rppp`); $K(t)$ is the sum of general government capital stock (variable `kgov_rppp`), private capital stock (variable `kpriv_rppp`) and public-private partnership capital stock (variable `kppp_rppp`); $Y(t)$ corresponds to the gross domestic product (variable `gdp_rppp`). All the variables are expressed in trillions of constant 2017 international dollars.

[70]Since the trajectory of $\delta_K(t)$ is erratic, we have preferred to report the 3-year moving average of $\delta_K(t)$.

[71]We recall that $s(t)$ is a control variable.

[72]We use the following initial values: $L(2019) = 7.725$ bn, $g_L = 1\%$, $\delta_L = 1.5\%$, $I(2019) = \$30.625$ tn, $K(2019) = \$318.773$ tn, $\delta_K = 6.25\%$, $A(2019) = 5.276$, $g_A(2019) = 0.75\%$, $\delta_A = 0.12\%$, $\gamma = 0.30$ and $Y(2019) = \$124.418$ tn.

Figure 8.73: Historical estimates of $I(t)$, $K(t)$, $Y(t)$ and $\delta_K(t)$
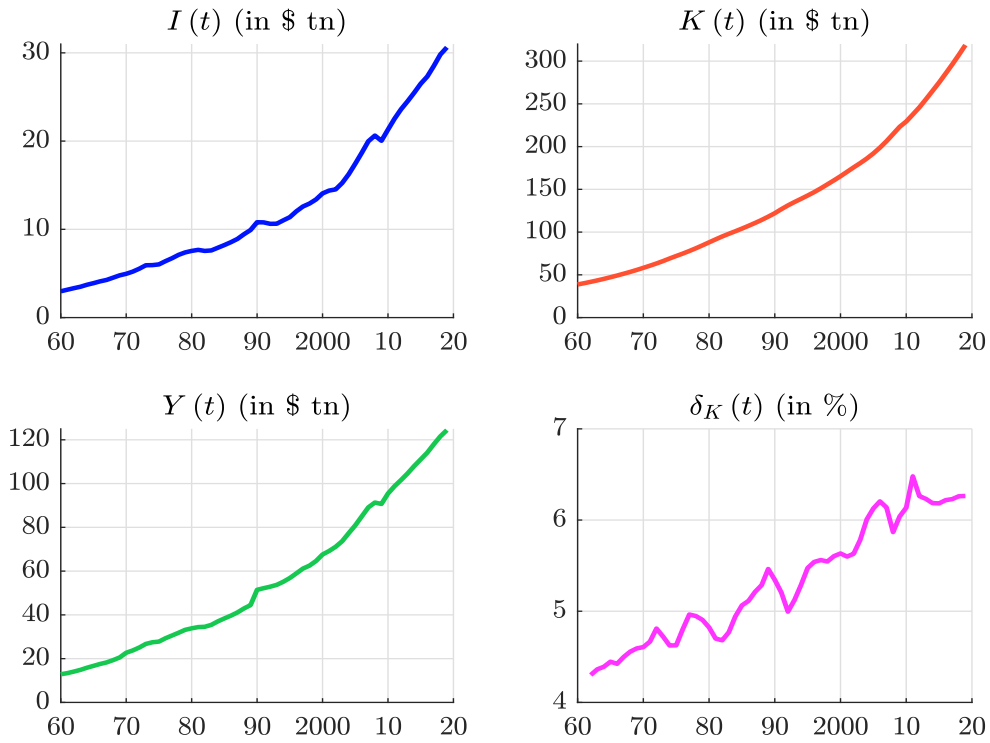
Source: IMF Investment and Capital Stock Dataset (2021) & Author's calculations.

Figure 8.74: Simulation of the DICE macroeconomic module
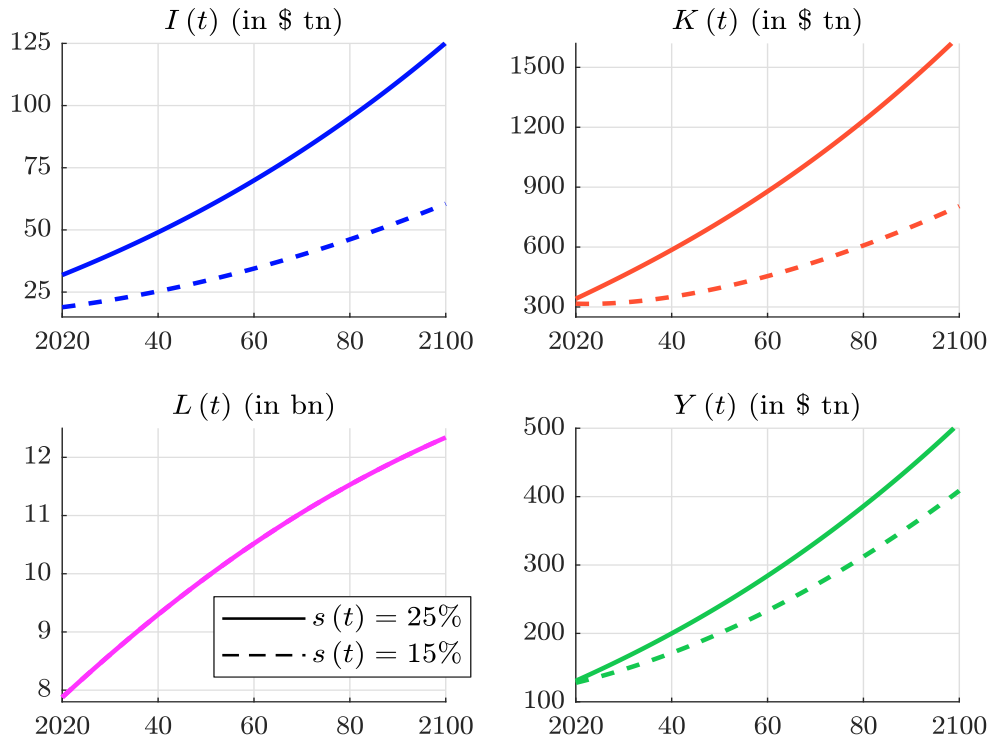
We notice that $\Omega_{\text{climate}}(t)$ depends on two variables: the atmospheric temperature $\mathcal{T}_{\text{AT}}(t)$ and the control variable $\mu(t)$. In the DICE model, $\mathcal{T}_{\text{AT}}(t)$ is a state variable while $\mu(t)$ is a decision variable. $\Omega_{\text{climate}}(t)$ is a then a decreasing function of $\mathcal{T}_{\text{AT}}(t)$ and $\mu(t)$ and we have the following properties:

$$
\begin{cases}
\Omega_{\text{climate}}(t) = 1 \Leftrightarrow \mathcal{T}_{\text{AT}}(t) = 0 \wedge \mu(t) = 0 \\
\lim_{\mathcal{T}_{\text{AT}}(t) \to \infty} \Omega_{\text{climate}}(t) = 0 \\
\lim_{\mu(t) \to 1} \Omega_{\text{climate}}(t) = \dfrac{1 - \theta_1(t)}{1 + \psi_1 \mathcal{T}_{\text{AT}}(t) + \psi_2 \mathcal{T}_{\text{AT}}(t)^2}
\end{cases}
$$

Nordhaus and Sztorc (2013) used $\psi_1 = 0$ and $\psi_2 = 2.67 \times 10^{-3}$ for the specification of $D(t)$. In Figure 8.75, we have represented the corresponding loss function $\mathcal{L}_D(\mathcal{T}_{\text{AT}}) := \mathcal{L}_D(t) = 1 - \Omega_D(t)$ as a function of the temperature increase $\mathcal{T}_{\text{AT}}$. For instance, we obtain $\mathcal{L}_D(1°\text{C}) = 0.27\%$, $\mathcal{L}_D(2°\text{C}) = 1.06\%$, $\mathcal{L}_D(5°\text{C}) = 6.26\%$ and $\mathcal{L}_D(10°\text{C}) = 21.07\%$ (dashed green line). We may think that these figures are underestimated in 2023. Already, fifteen years ago, the magnitude of these losses has created a scientific debate among economists. For instance, Hanemann (2008) found a higher loss function if the temperature reaches 2.5°C. Indeed, he estimated an annual loss of \$113 bn in the US versus \$28 bn for Nordhaus and Boyer (2000) (base year = 1990). Therefore, we can show that the parameter ψ_2 must take the value 1.135×10^{-3} to obtain the ratio 113/28:

$$
D(t) = 11.35 \times 10^{-3} \times \mathcal{T}_{\text{AT}}(t)^2
$$

In a series of research papers (Weitzman, 2009, 2010, 2012), Martin Weitzman investigated several functional forms of the climate damage function. Using a stochastic modeling of the temperature $\mathcal{T}_{\text{AT}}$, Weitzman (2012) estimated the following climate damage function:

$$
D(t) = \left(\frac{\mathcal{T}_{\text{AT}}(t)}{20.46}\right)^2 + \left(\frac{\mathcal{T}_{\text{AT}}(t)}{6.081}\right)^{6.754}
$$

The first term corresponds to the Nordhaus specification[73], whereas the second term corresponds to "*a tipping point where the damages function changes dramatically around the iconic global warming level of* 6°C" (Weitzman, 2012, page 235). If we fit the Hanemann damage function using the Weitzman function, we obtain an alternative specification with higher damage for low temperature increase:

$$
D(t) = 1.35 \times 10^{-2} \times \mathcal{T}_{\text{AT}}(t)^2 + 6.0287 \times 10^{-7} \times \mathcal{T}_{\text{AT}}(t)^{7.5}
$$

While the previous damage functions are based on the functional form $D(t) = \sum_{k=1}^{m} \psi_k \mathcal{T}_{\text{AT}}(t)^{\alpha_k}$, Weitzman (2009) suggested the exponential quadratic damage function[74]: $\Omega_D(t) = \exp\left(-\beta \mathcal{T}_{\text{AT}}(t)^2\right)$. Pindyck (2012) proposed then the following parameterization:

$$
\Omega_D(t) = e^{-\beta' \mathcal{T}_{\text{AT}}(t)^2 G}
$$

where $G \sim \mathbf{F}_G$ is a random variable that follows a displaced gamma distribution[75] $\mathcal{DG}(\alpha, \beta, \theta)$ where $\alpha = 4.5$, $\beta = 4.69 \times 10^{-5}$ and $\theta = -7.46 \times 10^{-5}$. Following Daniel *et al.* (2016), we set $\beta' = 13.97$ and $G = \mathbf{F}_G^{-1}(99.99\%)$. We have reported all these alternative loss functions[76] in Figure 8.75. We notice a high level of disagreement between these different research works (Howard and Sterner, 2017; Tol, 2022).

---

[73]We have $1/20.46^2 \approx 2.4 \times 10^{-3}$, which was the value used by Nordhaus (2008).

[74]See Pindyck (2013, page 867) and Wagner and Weitzman (2015).

[75]If $X \sim \mathcal{DG}(\alpha, \beta, \theta)$, then $X - \theta \sim \mathcal{G}(\alpha, \beta)$.

[76]We also include the piecewise function $D(t) = 0.011 \mathcal{T}_{\text{AT}}(t) + \mathbb{1}\{\mathcal{T}_{\text{AT}}(t) > 3°\text{C}\} \cdot (0.036 - 0.011)(\mathcal{T}_{\text{AT}}(t) - 3°\text{C})$, which has been estimated by Newbold and Marten (2014).

Figure 8.75: Loss function due to climate damage costs
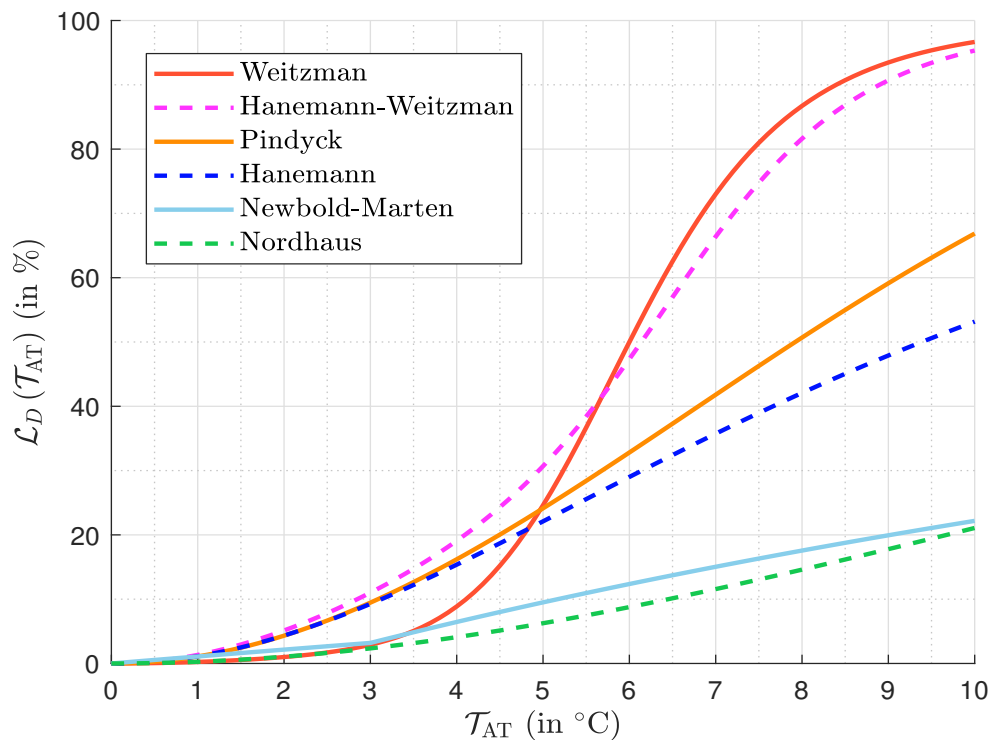
For the abatement cost function, Nordhaus (2018b) assumed that $\theta_1(t) = 0.0741 \times 0.98^{t-t_0}$ where t_0 is the base year of the abatement technology and $\theta_2 = 2.6$. The explanation is the following:

> "*The interpretation here is that at zero emissions for the first period* $(t = t_0)$*, abatement is 7.41% of output. That percentage declines at 2% per year. The abatement cost function is highly convex, reflecting the sharp diminishing returns to reducing emissions. [...] The model assumes the existence of a backstop technology, which is a technology that produces energy services with zero GHG emissions* ($\mu = 1$)*. The backstop price in 2020 is \$550 per ton of* CO_2*e, and the backstop cost declines at 0.5% per year. Additionally, it is assumed that there are no negative emissions technologies initially, but that negative emissions are available after 2150.*" (Nordhaus, 2018b, page 357).

Therefore, we can also define the cost of mitigation efforts as the following alternative form (Kellett et al., 2019):

$$\theta_1(t) \propto p_b(t_0)(1 - \delta_b)^{t-t_0} \sigma(t)$$

where $p_b(2020) = \$550/tCO_2e$, $\delta_b = 0.5\%$ and $\sigma(t)$ is the carbon intensity of economic activity, which is equal to the ratio of anthropogenic GHG emissions to economic output $Y(t)$. In Figure 8.76, we show the abatement cost function $\Lambda(t)$ for different values of the control rate $\mu(t)$, which is the second control variable with the saving rate $s(t)$.

Remark 56 *There is a large literature on abatement costs, especially about the concept of marginal abatement cost (MAC). Since this concept is at the core of the transition risk (Gillingham and Stock, 2018) and is highly related to the carbon tax (Daniel et al., 2019), it will be developed in Section 10.1.2 on page 544.*

Figure 8.76: Abatement cost function
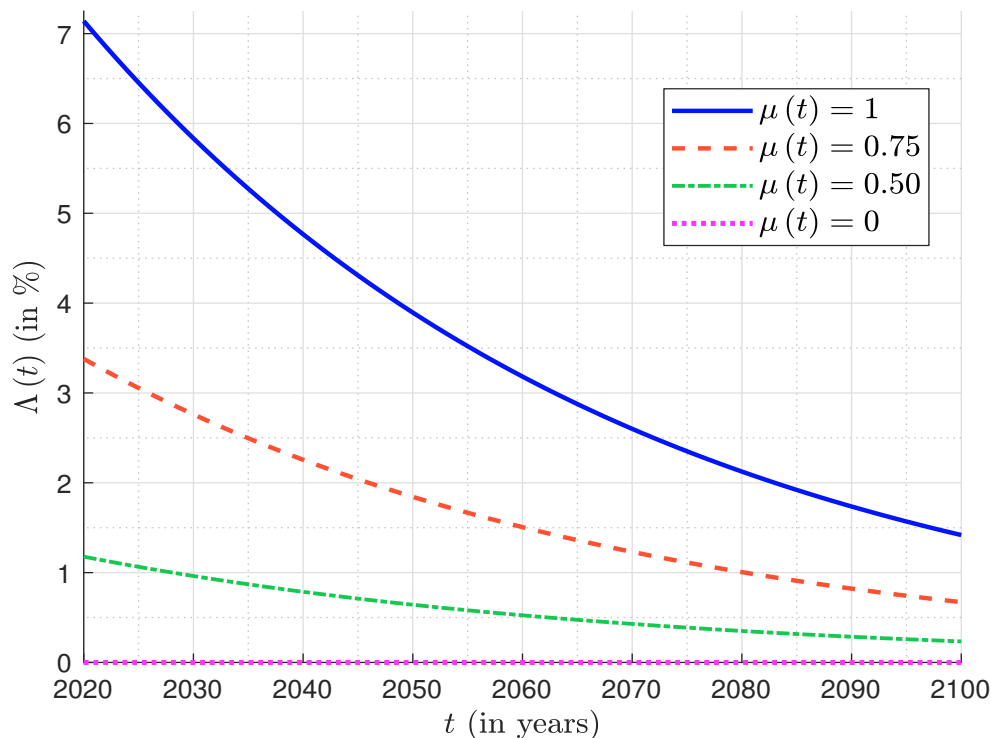
Temperature modeling   We now turn to the physics of the DICE model. As noticed by Nordhaus (2017b), IAMs are "*highly simplified representations of complex economic and geophysical realities*", and can not compete with general circulation models[77] (GCM). Nevertheless, these models are based on basic physics of temperature forecasts, produce plausible temperature trajectories and "*do a reasonable job capturing uncertainty about the equilibrium climate sensitivity*" (Calel and Stainforth, 2017a, page 1202).

The GHG emissions $\mathcal{CE}(t)$ depends on the production $Y(t)$ and the land-use emissions $\mathcal{CE}_{\text{Land}}(t)$:

$$
\begin{aligned}
\mathcal{CE}(t) &= \mathcal{CE}_{\text{Industry}}(t) + \mathcal{CE}_{\text{Land}}(t) \\
&= (1 - \mu(t))\sigma(t)\,Y(t) + \mathcal{CE}_{\text{Land}}(t)
\end{aligned}
\tag{8.22}
$$

where $\sigma(t)$ is the impact of the production on GHG emissions. $\mathcal{CE}_{\text{Industry}}(t)$ corresponds to the anthropogenic emissions due to industrial activities. We have We have $\mathcal{CE}_{\text{Land}}(t) = (1 - \delta_{\text{Land}})\mathcal{CE}_{\text{Land}}(t-1)$. The control variable $\mu(t)$ measures the impact of climate change mitigation policies:

$$
\underbrace{\mathcal{CE}_{\text{Land}}(t)}_{\mu(t)=1} \leq \mathcal{CE}(t) \leq \underbrace{\sigma(t)\,Y(t) + \mathcal{CE}_{\text{Land}}(t)}_{\mu(t)=0}
$$

$\mu(t) = 1$ corresponds to the case where mitigation policies have eliminated the effects of climate change, while $\mu(t) = 0$ indicates that no specific policy has been put in place. Since $\mu(t)$ is an endogenous variable, it must be viewed as an effort rate that the economy must bear to limit global

---

[77]They simulate general circulation of planetary atmosphere and oceans, using the Navier-Stokes equations on a rotating sphere and thermodynamic terms for energy sources (radiation, latent heat). They are used for weather forecasting, understanding the climate, and forecasting climate change.

warming. Concerning the ratio $\sigma(t)$, it measures the relationship between the carbon emissions due to industrial activities and the gross output in the absence of mitigation policies ($\mu(t) = 0$):

$$\sigma(t) = \frac{\mathcal{CE}(t) - \mathcal{CE}_{\text{Land}}(t)}{Y(t)} = \frac{\mathcal{CE}_{\text{Industry}}(t)}{Y(t)}$$

Therefore, it can be interpreted as an emission factor or the anthropogenic carbon intensity of the economy. This parameter is integrated into the model as follows:

$$\sigma(t) = (1 + g_\sigma(t))\sigma(t-1) \tag{8.23}$$

where:

$$g_\sigma(t) = \frac{1}{1 + \delta_\sigma} g_\sigma(t-1)$$

and δ_σ is the decline rate of the GHG emissions growth.

Figure 8.77: Physical carbon pump
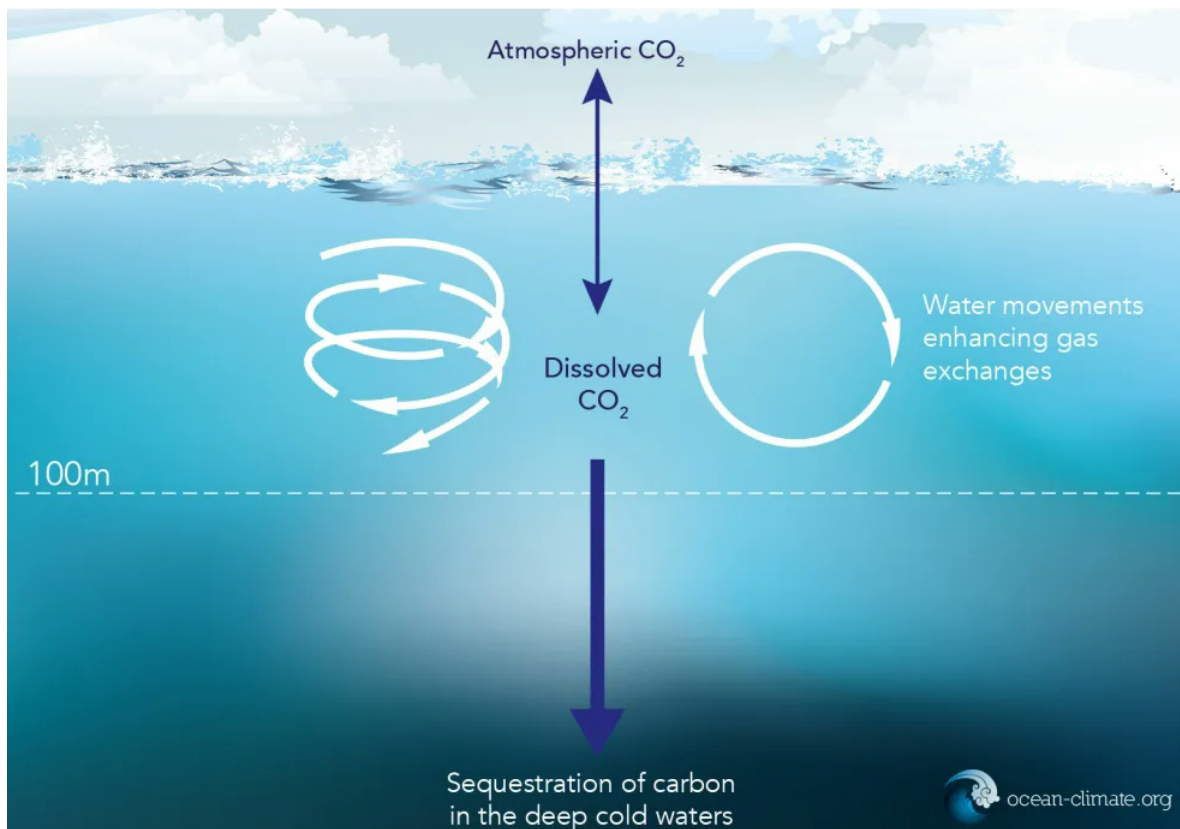
In order to relate the carbon emissions $\mathcal{CE}(t)$ and the atmospheric temperature $\mathcal{T}_{\text{AT}}(t)$, the DICE model uses a reduced form of a general circulation model describing the evolution of GHG concentrations in three carbon-sink reservoirs: the atmosphere AT, the upper ocean UP and the deep (or lower) ocean LO. As we will see on page 634, the ocean hold fifty times more CO_2 than the atmosphere. Indeed, a fraction of carbon from the atmosphere is absorbed by the upper ocean and transported by ocean circulation patterns. It is then stored in the deep ocean or re-radiated

from the upper ocean to the atmosphere influencing regional temperatures. The sequestration of carbon in the deep ocean depends on the temperature of cold waters, implying that a small part of the carbon stored in the deep ocean may return to the upper ocean. This physical carbon pump is illustrated in Figure 8.77. Taking into account this global heat transfer scheme, Nordhaus and Sztorc (2013) assume that the dynamics of carbon concentrations[78] are:

$$\begin{cases} \mathcal{CC}_{\mathrm{AT}}(t) = \phi_{1,1}\mathcal{CC}_{\mathrm{AT}}(t-1) + \phi_{1,2}\mathcal{CC}_{\mathrm{UP}}(t-1) + \phi_1 \boldsymbol{\mathcal{CE}}(t) \\ \mathcal{CC}_{\mathrm{UP}}(t) = \phi_{2,1}\mathcal{CC}_{\mathrm{AT}}(t-1) + \phi_{2,2}\mathcal{CC}_{\mathrm{UP}}(t-1) + \phi_{2,3}\mathcal{CC}_{\mathrm{LO}}(t-1) \\ \mathcal{CC}_{\mathrm{LO}}(t) = \phi_{3,2}\mathcal{CC}_{\mathrm{UP}}(t-1) + \phi_{3,3}\mathcal{CC}_{\mathrm{LO}}(t-1) \end{cases}$$

where $\phi_{i,j}$ represents the flow parameters between carbon-sink reservoirs[79], and ϕ_1 is the mass percentage of carbon in CO_2. Let $\mathcal{CC} = (\mathcal{CC}_{\mathrm{AT}}, \mathcal{CC}_{\mathrm{UP}}, \mathcal{CC}_{\mathrm{LO}})$ be the vector of the three-reservoir layers. The dynamics of $\mathcal{CC}$ is then a vector autoregressive process:

$$\mathcal{CC}(t) = \Phi_{\mathcal{CC}}\mathcal{CC}(t-1) + B_{\mathcal{CC}}\boldsymbol{\mathcal{CE}}(t) \tag{8.24}$$

where $B_{\mathcal{CC}} = (\phi_1, 0, 0)$ and:

$$\Phi_{\mathcal{CC}} = \begin{pmatrix} \phi_{1,1} & \phi_{1,2} & 0 \\ \phi_{2,1} & \phi_{2,2} & \phi_{3,2} \\ 0 & \phi_{3,2} & \phi_{3,3} \end{pmatrix}$$

The next step consists in linking accumulated carbon emissions in the atmosphere and global warming at the Earth's surface through increases in radiative forcing:

$$\mathcal{F}_{\mathrm{RAD}}(t) = \frac{\eta}{\ln 2} \ln\left(\frac{\mathcal{CC}_{\mathrm{AT}}(t)}{\mathcal{CC}_{\mathrm{AT}}(1750)}\right) + \mathcal{F}_{\mathrm{EX}}(t) \tag{8.25}$$

where $\mathcal{F}_{\mathrm{RAD}}(t)$ is the change in total radiative forcing of GHG emissions since 1750 (expressed in W/m^2), η is the temperature forcing parameter and $\mathcal{F}_{\mathrm{EX}}(t)$ is the exogenous forcing. Finally, the climate system for temperatures is characterized by a two-layer system:

$$\begin{cases} \mathcal{T}_{\mathrm{AT}}(t) = \mathcal{T}_{\mathrm{AT}}(t-1) + \xi_1\left(\mathcal{F}_{\mathrm{RAD}}(t) - \xi_2\mathcal{T}_{\mathrm{AT}}(t-1) - \xi_3\left(\mathcal{T}_{\mathrm{AT}}(t-1) - \mathcal{T}_{\mathrm{LO}}(t-1)\right)\right) \\ \mathcal{T}_{\mathrm{LO}}(t) = \mathcal{T}_{\mathrm{LO}}(t-1) + \xi_4\left(\mathcal{T}_{\mathrm{AT}}(t-1) - \mathcal{T}_{\mathrm{LO}}(t-1)\right) \end{cases}$$

where $\mathcal{T}_{\mathrm{AT}}(t)$ and $\mathcal{T}_{\mathrm{LO}}(t)$ are respectively the mean surface temperature and the temperature of the deep ocean. This system of equations means that "*higher radiative forcing warms the atmospheric layer, which then warms the upper ocean, gradually warming the deep ocean*" (Nordhaus and Sztorc, 2013, page 17). The authors simplify the relationships by only considering a two-box model: the atmosphere and the upper ocean. According to Calel and Stainforth (2017a), ξ_1 measures the speed of adjustment parameter for atmospheric temperature, ξ_2 is the ratio of increased forcing from CO_2 doubling to the climate sensitivity, ξ_3 is the heat loss coefficient from atmosphere to oceans and ξ_4 is the heat gain coefficient by deep oceans. Let $\mathcal{T} = (\mathcal{T}_{\mathrm{AT}}, \mathcal{T}_{\mathrm{LO}})$ be the temperature vector. We have:

$$\mathcal{T}(t) = \Xi_{\mathcal{T}}\mathcal{T}(t-1) + B_{\mathcal{T}}\mathcal{F}_{\mathrm{RAD}}(t) \tag{8.26}$$

where $B_{\mathcal{T}} = (\xi_1, 0)$ and:

$$\Xi_{\mathcal{T}} = \begin{pmatrix} 1 - \xi_1(\xi_2 + \xi_3) & \xi_1\xi_3 \\ \xi_4 & 1 - \xi_4 \end{pmatrix} = \begin{pmatrix} \xi'_{1,1} & \xi'_{1,2} \\ \xi'_{2,1} & \xi'_{2,2} \end{pmatrix}$$

---

[78]The carbon concentration is expressed in grams of carbon, and not in gCO$_2$e. The conversion is 12/44 or 0.2727 gC/g CO$_2$.

[79]By construction, we have $\sum_{i=1}^{3}\phi_{i,j} = 1$ for $j = 1, 2, 3$. $\phi_{i,i}$ is then the probability that a molecule remains in its carbon sink, while $\phi_{i,j}$ is the probability that a molecule stored in the carbon sink j goes to the carbon sink i.

Box 8.11: Correspondence between DICE and the two-box model

We have previously seen that the two-box model of climate change is based on the heat diffusion between the atmosphere and the deep ocean:

$$
\begin{cases}
c_{\mathrm{AT}} \dfrac{\mathrm{d}\mathcal{T}_{\mathrm{AT}}(t)}{\mathrm{d}t} &= \mathcal{F}_{\mathrm{RAD}}(t) - \lambda \mathcal{T}_{\mathrm{AT}}(t) - \beta\left(\mathcal{T}_{\mathrm{AT}}(t) - \mathcal{T}_{\mathrm{LO}}(t)\right) \\
c_{\mathrm{LO}} \dfrac{\mathcal{T}_{\mathrm{LO}}(t)}{\mathrm{d}t} &= \beta\left(\mathcal{T}_{\mathrm{AT}}(t) - \mathcal{T}_{\mathrm{LO}}(t)\right)
\end{cases}
$$

where c_{AT} is the effective heat capacity of the atmosphere (including the upper ocean and the land surface), c_{LO} is the effective heat capacity of the deep ocean, λ is the feedback parameter and β is the heat transfer coefficient between the upper and lower oceans. Using a forward Euler discretization, Calel and Stainforth (2017a) found that:

$$
\begin{cases}
\mathcal{T}_{\mathrm{AT}}(t) &= \mathcal{T}_{\mathrm{AT}}(t-1) + \dfrac{1}{c_{\mathrm{AT}}}\left(\mathcal{F}_{\mathrm{RAD}}(t-1) - \lambda\mathcal{T}_{\mathrm{AT}}(t-1) - \right.\\
 & \qquad\qquad\qquad\qquad \left. \beta\left(\mathcal{T}_{\mathrm{AT}}(t-1) - \mathcal{T}_{\mathrm{LO}}(t-1)\right)\right)\Delta t \\
\mathcal{T}_{\mathrm{LO}}(t) &= \mathcal{T}_{\mathrm{LO}}(t-1) + \dfrac{\beta}{c_{\mathrm{LO}}}\left(\mathcal{T}_{\mathrm{AT}}(t-1) - \mathcal{T}_{\mathrm{LO}}(t-1)\right)\Delta t
\end{cases}
$$

where Δt is the length of time step (expressed in seconds). Therefore, we have the following relationships[a]:

$$
\begin{cases}
\xi_1 = \dfrac{1}{c_{\mathrm{AT}}}\Delta t \\
\xi_2 = \lambda \\
\xi_3 = \beta \\
\xi_4 = \dfrac{\beta}{c_{\mathrm{LO}}}\Delta t
\end{cases}
\quad \text{and} \quad
\begin{cases}
\xi'_{1,1} = 1 - \dfrac{\lambda + \beta}{c_{\mathrm{AT}}}\Delta t \\
\xi'_{1,2} = \dfrac{\beta}{c_{\mathrm{AT}}}\Delta t \\
\xi'_{2,1} = \dfrac{\beta}{c_{\mathrm{LO}}}\Delta t \\
\xi'_{2,2} = 1 - \dfrac{\beta}{c_{\mathrm{LO}}}\Delta t
\end{cases}
$$

---

[a] Calel and Stainforth (2017a) noticed that the only difference is that the DICE model uses $\mathcal{F}_{\mathrm{RAD}}(t)$ instead of $\mathcal{F}_{\mathrm{RAD}}(t-1)$, which is not a problem if the time step is small.

Below, we report the parameter and initial values used by the 2013 version[80]:

- $t_0 = 2010$, $\Delta t = 5$ years;

- Equation (8.22): $\mathcal{CE}(t_0) = 36.91$ GtCO$_2$e, $\mu(t_0) = 3.9\%$, $\sigma(t_0) = 0.5491$ GtCO$_2$e/\$ tn, $Y(t_0) = \$63.69$ tn, $\mathcal{CE}_{\mathrm{Industry}}(t_0) = 33.61$ GtCO$_2$e, $\delta_{\mathrm{Land}} = 0.2$, and $\mathcal{CE}_{\mathrm{land}}(t_0) = 3.3$ GtCO$_2$e;

- Equation (8.23): $\sigma(t_0) = 0.5491$ GtCO$_2$e/\$ tn, $g_\sigma(t_0) = 1\%$ and $\delta_\sigma = 0.1\%$;

- Equation (8.24): $\mathcal{CC}_{\mathrm{AT}}(t_0) = 830.4$, $\mathcal{CC}_{\mathrm{UP}}(t_0) = 1\,527$, $\mathcal{CC}_{\mathrm{LO}}(t_0) = 10\,010$, $\phi_1 = 0.2727$ gC/gCO$_2$, $\mathcal{CE}(t_0) = 36.91$ GtCO$_2$e, and the carbon cycle diffusion matrix is equal to:

$$
\Phi_{\mathcal{CC}} = \begin{pmatrix} 91.20\% & 3.83\% & 0 \\ 8.80\% & 95.92\% & 0.03\% \\ 0 & 0.25\% & 99.97\% \end{pmatrix}
$$

---

[80] They can be found in Kellett *et al.* (2019, Table 2 & Appendices A and B).

- Equation (8.25): $\mathcal{F}_{\text{RAD}}(t_0) = 2.14$ W/m^2, $\eta = 3.8$ W/m^2, $\mathcal{CC}_{\text{AT}}(t_0) = 830.4$, $\mathcal{CC}_{\text{AT}}(1750) = 588$ GtC, $\mathcal{F}_{\text{EX}}(t)$ is the linear interpolation function between $\mathcal{F}_{\text{EX}}(t_0) = 0.25$ W/m^2 and $\mathcal{F}_{\text{EX}}(2100) = 0.70$ W/m^2 for $t \in [t_0, 2100]$, and is equal to $\mathcal{F}_{\text{EX}}(2100)$ for $t \geq 2100$;

- Equation (8.26): $\mathcal{T}_{\text{AT}}(t_0) = 0.8°$C, $\mathcal{T}_{\text{LO}}(t_0) = 0.0068°$C, $\xi_1 = 0.098$, $\mathcal{F}_{\text{RAD}}(t_0) = 2.14$ W/m^2, and the temperature diffusion matrix is given by[81]:

$$\Xi_{\mathcal{T}} = \begin{pmatrix} 86.30 & 0.8624 \\ 2.5 & 97.5 \end{pmatrix} \times 10^{-2}$$

From Equation (8.26), we can compute the steady-state temperatures:

$$\begin{aligned} \mathcal{T}(\infty) &= (I_2 - \Xi_{\mathcal{T}})^{-1} B_{\mathcal{T}} \mathcal{F}_{\text{RAD}}(\infty) \\ &= \frac{\mathcal{F}_{\text{RAD}}(\infty)}{\xi_2} \mathbf{1}_2 \\ &= \begin{pmatrix} 0.76316 \\ 0.76316 \end{pmatrix} \mathcal{F}_{\text{RAD}}(\infty) \end{aligned}$$

We deduce that:

$$\Delta \mathcal{T}_{\text{AT}}(\infty) = \frac{1}{\xi_2} \Delta \mathcal{F}_{\text{RAD}}(\infty) = 0.76316 \times \Delta \mathcal{F}_{\text{RAD}}(\infty)$$

A variation of ± 1 W/m^2 implies a variation of $\pm 1°$C of the atmospheric temperature. Therefore, in order to limit global warming, we need to reduce radiative forcing and the carbon concentration $\mathcal{CC}_{\text{AT}}(t)$ in the atmosphere (Equation 8.25). This is done by emitting lower carbon emissions (Equation 8.24). In Figure 8.78, we show the impulse response analysis[82] by considering a reduction of one GtCO$_2$e. We notice that the carbon concentration in the atmosphere is reduced because the carbon is stored in the upper ocean and then in the deep ocean. To achieve carbon emissions reduction, we have three choices (Equation 8.22):

1. We can reduce the production $Y(t)$;

2. We can reduce the carbon intensity $\sigma(t)$ of industrial activities;

3. We can increase the mitigation effort $\mu(t)$ and accelerate the transition to a low-carbon economy;

In order to illustrate the climate module, we run Equations (8.22–8.26) with the parameter and initial values used by the 2013 version. For that, we need to define $Y(t)$ and $\mu(t)$, which are the two external variables of the temperature modeling system. By assuming that $Y(t) = Y(t_0)$ and $\mu(t) = \mu(t_0)$, we obtain the results in Table 8.14. If we consider the following functions $Y(t) = (1 + g_Y)^{t-t_0} Y(t_0)$ and $\mu(t) = \min\left(e^{g_\mu(t-t_0)} \mu(t_0), 1\right)$, the trajectories of $Y(t)$, $\mu(t)$, $\mathcal{T}_{\text{AT}}(t)$ and $\mathcal{T}_{\text{LO}}(t)$ those given in Figure 8.79. We observe that we can limit global warming if we limit the economic growth g_Y or we dramatically increase the mitigation effort $\mu(t)$. The nightmare climate-economic scenario is obtained when the economic growth is high and there is no mitigation effort (Figure 8.80).

---

[81]We have $\xi_1 = 0.098$, $\xi_2 = 3.8/2.9$, $\xi_3 = 0.088$, and $\xi_4 = 0.025$ (Calel and Stainforth, 2017a, page 1203). The DICE model uses a 5-year time step. If we prefer to use $\Delta t = 1$ year, we have:

$$\Xi_{\mathcal{T}} = \begin{pmatrix} 97.2592 & 0.1725 \\ 0.5 & 99.5 \end{pmatrix} \times 10^{-2}$$

and $\xi_1 = 0.0196$ because $c_{\text{AT}} = 51.0204$, $c_{\text{LO}} = 17.6$, $\lambda = 1.3103$ and $\beta = 0.088$.

[82]Following Roncalli (2020a, page 646), the matrix of impulse responses after t years is $\Phi_{\mathcal{CC}}^t$, while the matrix of cumulated responses is $I_3 + \sum_{k=1}^{t} \Phi_{\mathcal{CC}}^k$.

Figure 8.78: Impulse response analysis on carbon AT, LO, and UP concentrations ($\Delta \mathcal{CE} = -1$ $GtCO_2e$)
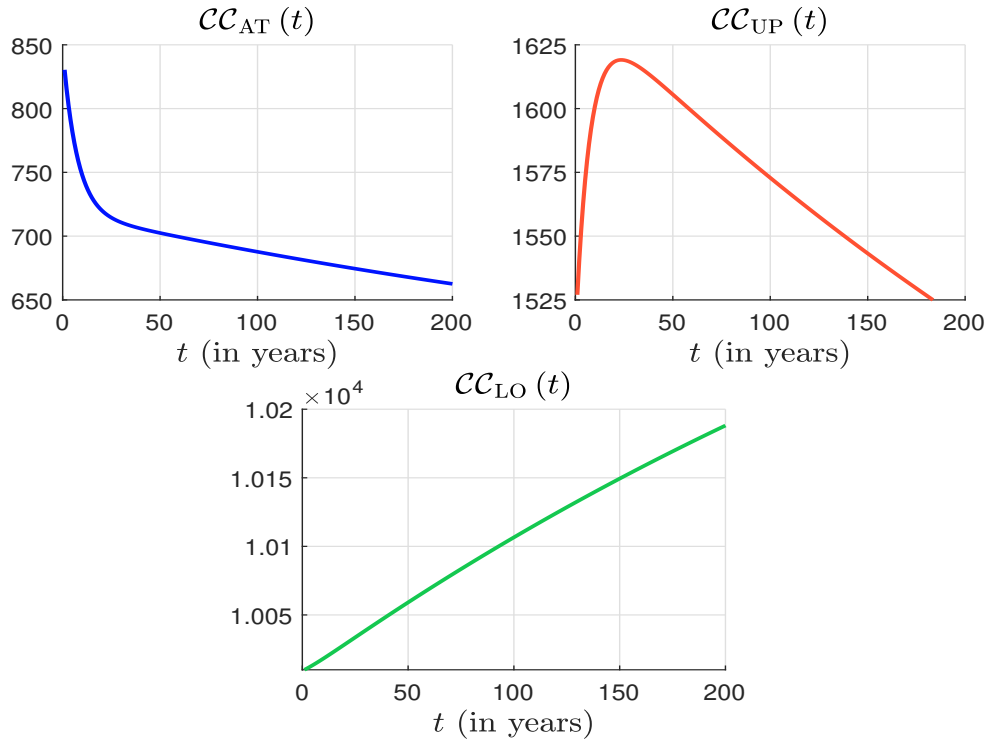
Table 8.14: Output of the DICE climate module ($Y(t) = Y(t_0)$, $\mu(t) = \mu(t_0)$)

t	$\mathcal{CE}(t)$	$\sigma(t)$	$\mathcal{CC}_{AT}(t)$	$\mathcal{F}_{RAD}(t)$	$\mathcal{T}_{AT}(t)$	$\mathcal{T}_{LO}(t)$
2010	36.91	0.55	830.4	2.14	0.800	0.007
2015	36.25	0.55	825.7	2.14	0.900	0.027
2020	36.06	0.56	821.9	2.14	0.986	0.048
2025	35.97	0.57	818.9	2.14	1.061	0.072
2030	35.98	0.57	816.6	2.15	1.127	0.097
2035	36.05	0.58	814.9	2.16	1.186	0.122
2040	36.18	0.58	813.9	2.18	1.238	0.149
2045	36.36	0.59	813.3	2.20	1.286	0.176
2050	36.58	0.59	813.3	2.23	1.329	0.204
2055	36.82	0.60	813.6	2.26	1.370	0.232
2060	37.09	0.61	814.4	2.29	1.408	0.261
2065	37.39	0.61	815.4	2.32	1.445	0.289
2070	37.70	0.62	816.8	2.35	1.480	0.318
2075	38.02	0.62	818.4	2.39	1.514	0.347
2080	38.36	0.63	820.3	2.43	1.547	0.376
2085	38.71	0.64	822.4	2.46	1.580	0.406
2090	39.06	0.64	824.7	2.50	1.612	0.435
2095	39.43	0.65	827.1	2.55	1.645	0.464
2100	39.80	0.66	829.7	2.59	1.677	0.494

Figure 8.79: Simulation of the DICE climate module
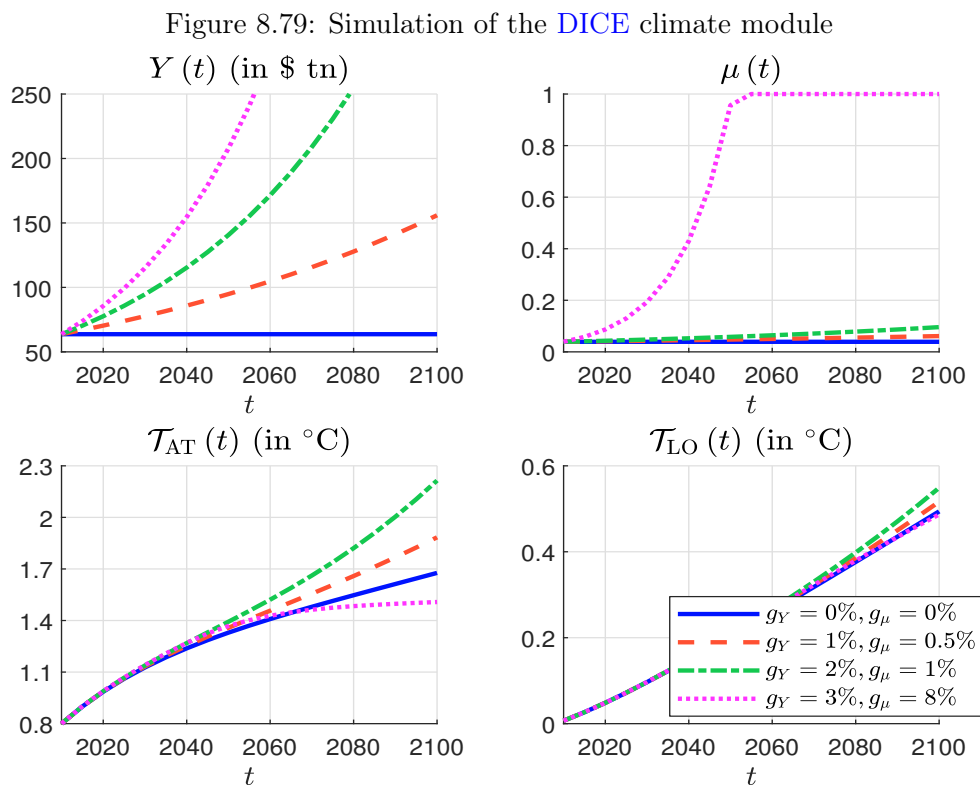
Figure 8.80: The nightmare climate-economic scenario ($g_Y = 0\%$, $\mu(t) = 0$)
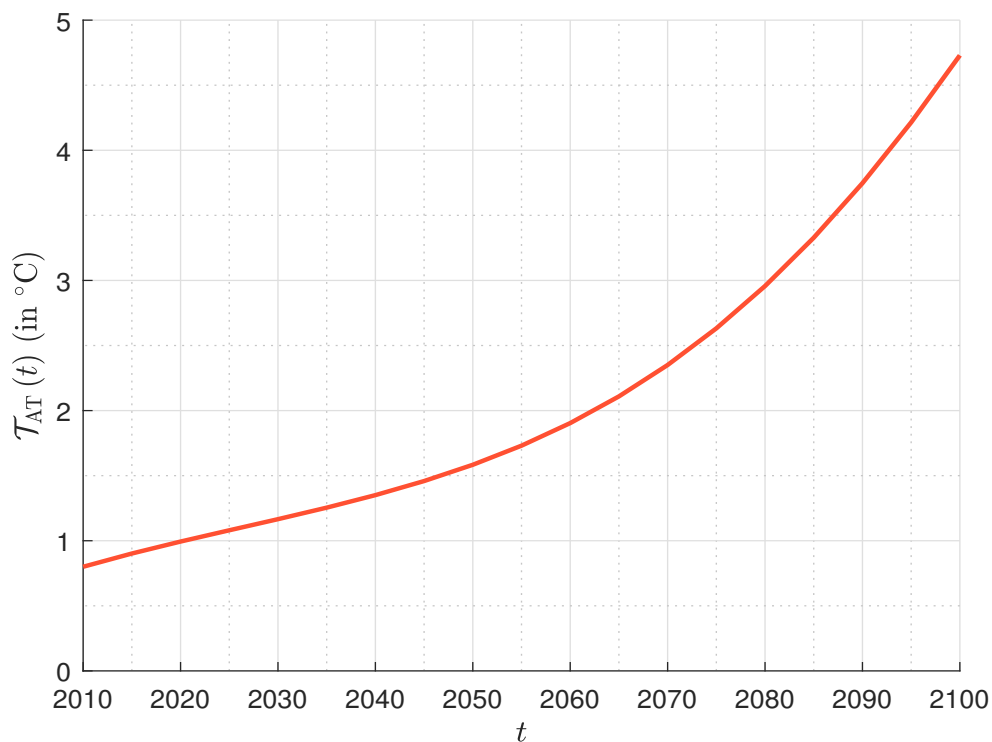
Social welfare maximization The last step of the DICE model is to define the social preferences and the objective function to optimize. The model assumes that the climate policy $\mu(t)$ is chosen to maximize the social welfare function W, which is the discounted sum of the generational utilities:

$$W\left(s\left(t\right),\mu\left(t\right)\right) = \sum_{t=t_0+1}^{T} \frac{L\left(t\right)\boldsymbol{\mathcal{U}}\left(c\left(t\right)\right)}{\left(1+\rho\right)^{t-t_0}}$$

where ρ is the (generational) discount rate and $c\left(t\right) = C\left(t\right)/L\left(t\right)$ is the consumption per capita. $\boldsymbol{\mathcal{U}}\left(c\right)$ is the CRRA utility function:

$$\boldsymbol{\mathcal{U}}\left(c\right) = \frac{c^{1-\alpha}-1}{1-\alpha}$$

where $\alpha \geq 0$ is the constant elasticity of marginal utility of consumption, which can be interpreted as the generational inequality aversion[83] in the context of the Ramsey-Cass-Koopmans model. When α is close to zero, there is low risk aversion to intergenerational inequality. When α is equal to one, the social welfare becomes egalitarian between generations. When $\alpha \to \infty$, we obtain an infinite total utility.

The optimal control problem is then given by:

$$\begin{aligned}
\left(s^{\star}\left(t\right),\mu^{\star}\left(t\right)\right) &= \arg\max W\left(s\left(t\right),\mu\left(t\right)\right) \\
&\text{s.t.} \quad \left\{ \begin{array}{l} \text{Equations (8.19–8.26)} \\ \mu\left(t\right) \in [0,1] \\ s\left(t\right) \in [0,1] \end{array} \right.
\end{aligned} \tag{8.27}$$

We reiterate that the optimization problem has two control variables: the saving rate $s\left(t\right)$ and the climate mitigation policy $\mu\left(t\right)$. The first control variable $s\left(t\right)$ is common to many economic models of optimal growth. It manages the substitution between the present consumption and the present investment, which determines the future consumption. The second control variable $\mu\left(t\right)$ is a new variable that manages the substitution between the present climate damages and the future climate damages. The two variables do not have the same status, because $s\left(t\right)$ is generally chosen by households, which is not the case of $\mu\left(t\right)$.

Remark 57 *The 2013 version of the DICE model assumes that $\alpha = 1.45$ and $\rho = 1.5\%$.*

Social cost of carbon

Above, we have presented the mathematical objective function of the DICE model. Nevertheless, the final objective of the DICE model is not to estimate the optimal pathway of $\mu\left(t\right)$, which is a conceptual measure of the mitigation effort. The main goal is to compute the social cost of carbon, which is the central pillar in the cost benefit analysis of climate change policies:

> "*The most important single economic concept in the economics of climate change is the social cost of carbon (SCC). This term designates the economic cost caused by an additional tonne of carbon dioxide emissions or its equivalent. In a more precise definition, it is the change in the discounted value of economic welfare from an additional unit of COtwo-equivalent emissions. The SCC has become a central tool used in climate change policy, particularly in the determination of regulatory policies that involve greenhouse gas emissions.*" (Nordhaus, 2017a, page 1518).

---

[83]When $\alpha \to 1$, $\mathcal{U}\left(c\right) = \ln c$.

From a mathematical viewpoint, the social cost of carbon is then defined as:

$$\text{SCC}(t) = \frac{\dfrac{\partial W(t)}{\partial \boldsymbol{\mathcal{CE}}(t)}}{\dfrac{\partial W(t)}{\partial C(t)}} = \frac{\partial C(t)}{\partial \boldsymbol{\mathcal{CE}}(t)}$$

It is expressed in $\$/tCO_2$. This measure can be extended to other greenhouse gases. For instance, IWG (2021) also estimated this figure for the Methane and the Nitrous Oxide. They found a magnitude order of \$200 for CO_2, \$6 000 for CH_4, and \$70 000 for N_2O.

The DICE model can be applied to different scenarios. For instance, we can consider the baseline scenario ($\mu(t) = \mu(t_0)$), the optimal scenario (social welfare maximization) or the 2°C scenario ($\mathcal{T}_{AT}(t) \leq 2$°C). In Figures 8.81–8.84, we report the results found by Le Guenedal (2019) when we use the 2013 and 2016 versions of the model[84]. The difference between the two versions is mainly explained by the update of the parameter values (Kellett *et al.*, 2019, Table 2). With the 2013 version, the control variable $\mu(t)$ remains below one when considering the optimal scenario, the temperature $\mathcal{T}_{AT}(t)$ crosses 2°C around 2050 and the social cost of carbon is less than \$150 (Figure 8.81). With the 2016 version, the optimal scenario is a little bit different since we need to make more efforts in terms of mitigation and the SCC in greater than \$150 after 2050 (Figure 8.83). The 2°C scenario implies higher values of $\mu(t)$ and SCC(t) for the 2013 version (Figure 8.82). Unfortunately, we observe that this scenario is no longer feasible with the 2016 version[85] (Figure 8.84).

Table 8.15: Global SCC under different scenario assumptions (in $\$/tCO_2$)

Scenario	2015	2020	2025	2030	2050	CAGR
Baseline	31.2	37.3	44.0	51.6	102.5	3.46%
Optimal	30.7	36.7	43.5	51.2	103.6	3.54%
2.5°C-max	184.4	229.1	284.1	351.0	1 006.2	4.97%
2.5°C-mean	106.7	133.1	165.1	203.7	543.3	4.76%

Source: Nordhaus (2017a, Table 1, page 1520).

In table 8.15, we report the SCC estimated for standard DICE models computed by Nordhaus (2017a). The baseline scenario corresponds to the current policy, while the optimal scenario is the given by the optimized control path (Equation 8.27). Under the baseline scenario assumption, the SCC value is $\$31.2/tCO_2$ in 2015 and reaches $\$102.5/tCO_2$ in 2050, implying a compound annual growth rate of 3.46%. The optimal scenario gives SCC figures that are very close to the baseline scenario. Nordhaus (2017a) also evaluated two alternative scenarios: the 2.5°C-max scenario constraints the temperature to be below 2.5°C, whereas the 2.5°C-mean imposes an average temperature of 2.5°C for the next 100 years. The impact of these two alternative scenarios is significant. In this case, the social cost of carbon can reach the value $\$1 000/tCO_2$ in 2050.

Remark 58 *The social cost of carbon is also known as the carbon tax, the carbon price or the shadow price of carbon emissions. Since it is one of the main policy tools to manage the transition risk, it will be extensively studied in Chapter 10.*

---

[84] These simulations are computed with the Matlab implementation of MPC-DICE, which replicates the functionality of the DICE-2013R/2016R model (Faulwasser *et al.*, 2018). The source code is available at https://github.com/cmkellett/MPC-DICE. These results can also be generated using the two other Maltab implementations provided by Kellett *et al.* (2016) and Lemoine (2020).

[85] See Nordhaus (2017b, page 17).

Figure 8.81: Optimal welfare scenario (DICE 2013R)

Figure 8.82: 2°C scenario (DICE 2013R)

Figure 8.83: Optimal welfare scenario (DICE 2016R)
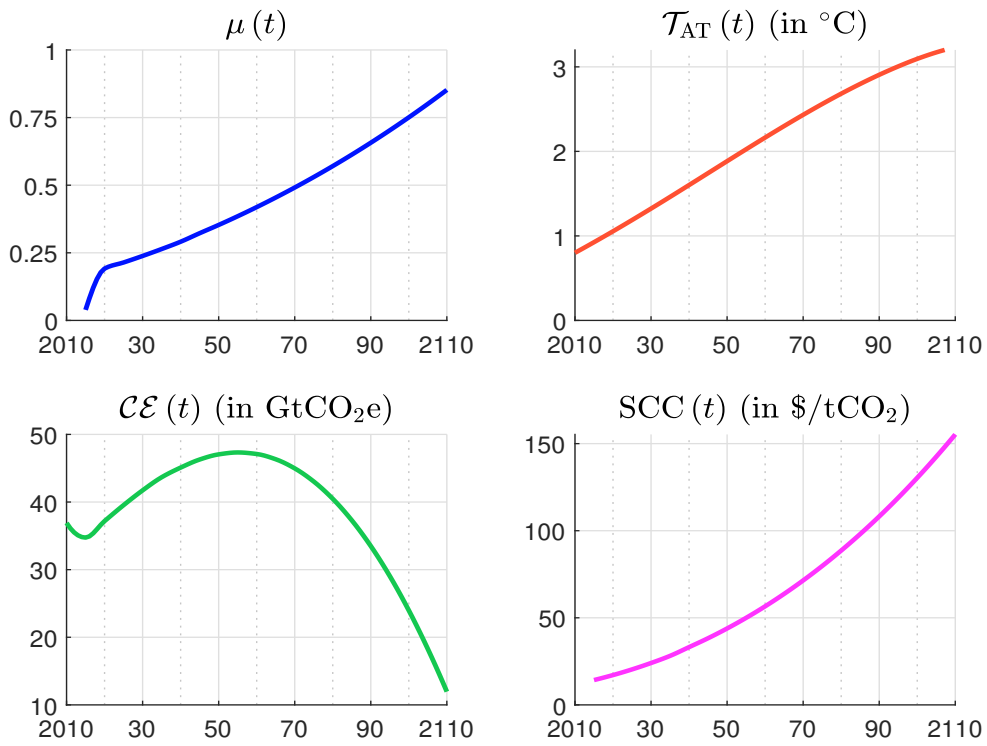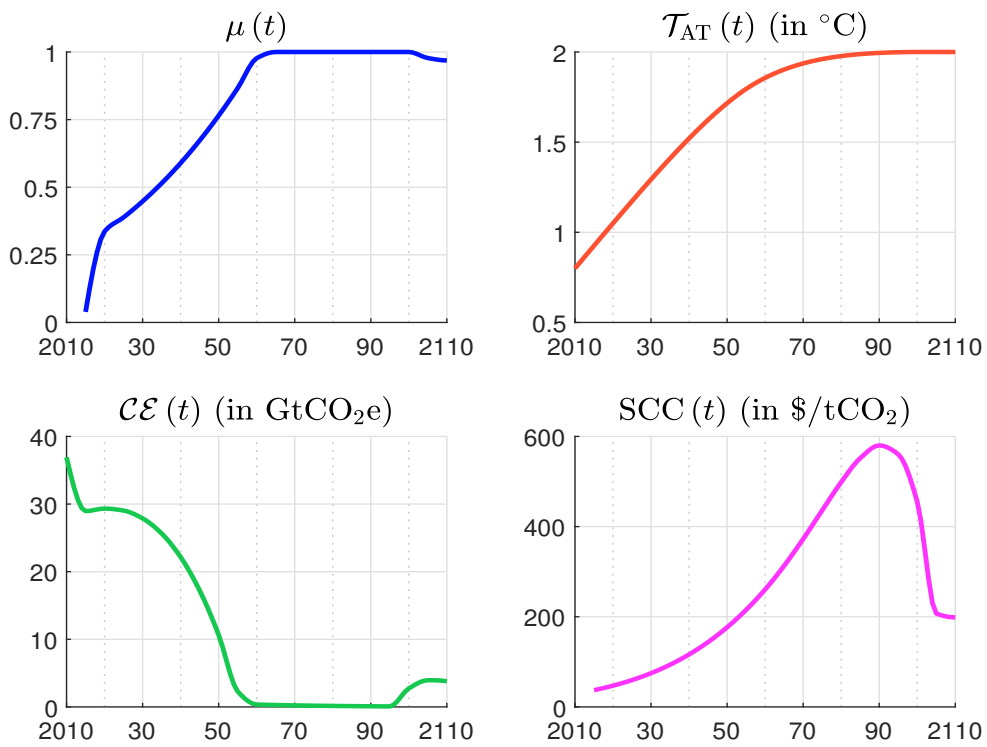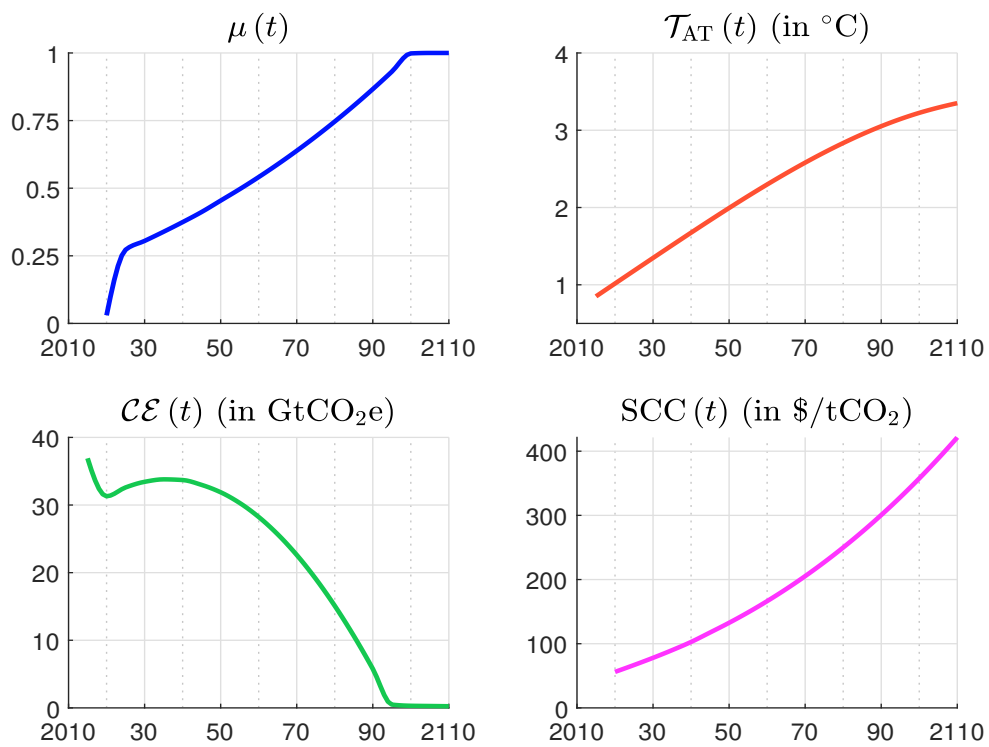
Source: Le Guenedal (2019).

Figure 8.84: 2°C scenario (DICE 2016R)
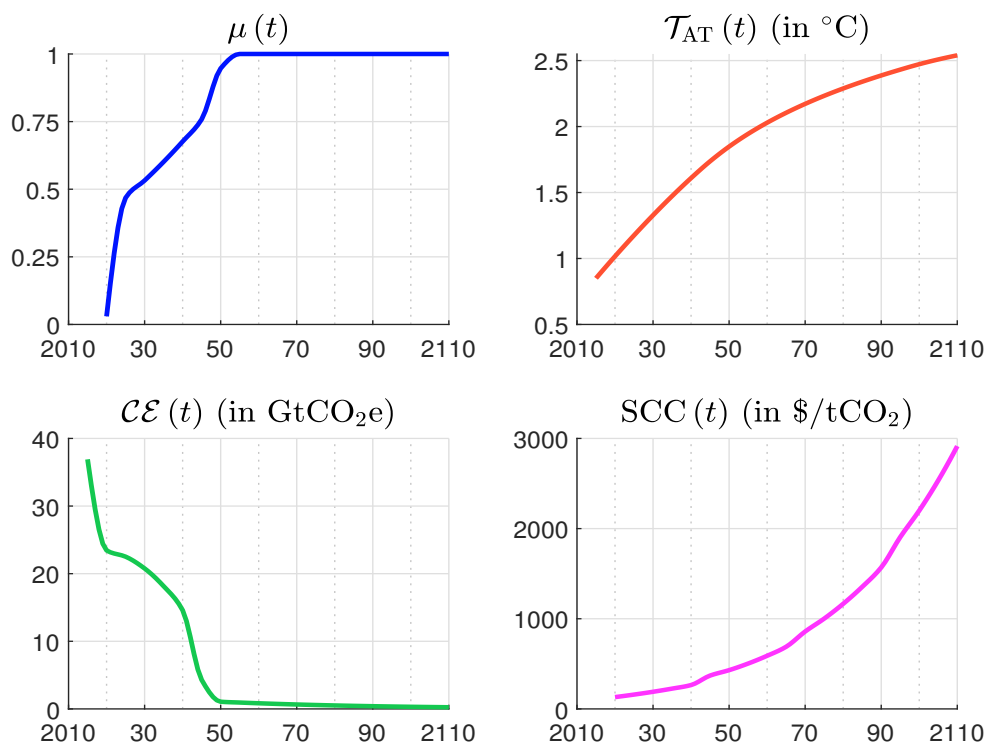
Source: Le Guenedal (2019).

The Stern-Nordhaus controversy

In 2007, Nicholas Stern published a report called *The Economics of Climate Change: The Stern Review,* that was commissioned by the British government (Stern, 2007). The Stern Review called for sharp and immediate action to stabilize greenhouse gases because "*the benefits of strong, early action on climate change outweighs the costs*". According to Weitzman (2007), the analysis supporting this conclusion is based on two main arguments. The first one concerns the high value of the discount rate, which is used by IAMs and the DICE model. The second argument is more ethical and related to the large uncertainty about the future of the world. Therefore, the Stern Review proposes to use $\rho = 0.10\%$. The publication of the book has created turmoil in the scientific community, and many economists have participated to the debate. For instance, we can find dozens of reviews and critiques of the Stern Review that have been published in academic journals (*Journal of Economic Literature, Science, Energy Policy, Climate Change*, etc.).

Figure 8.85: Discounted value of \$10
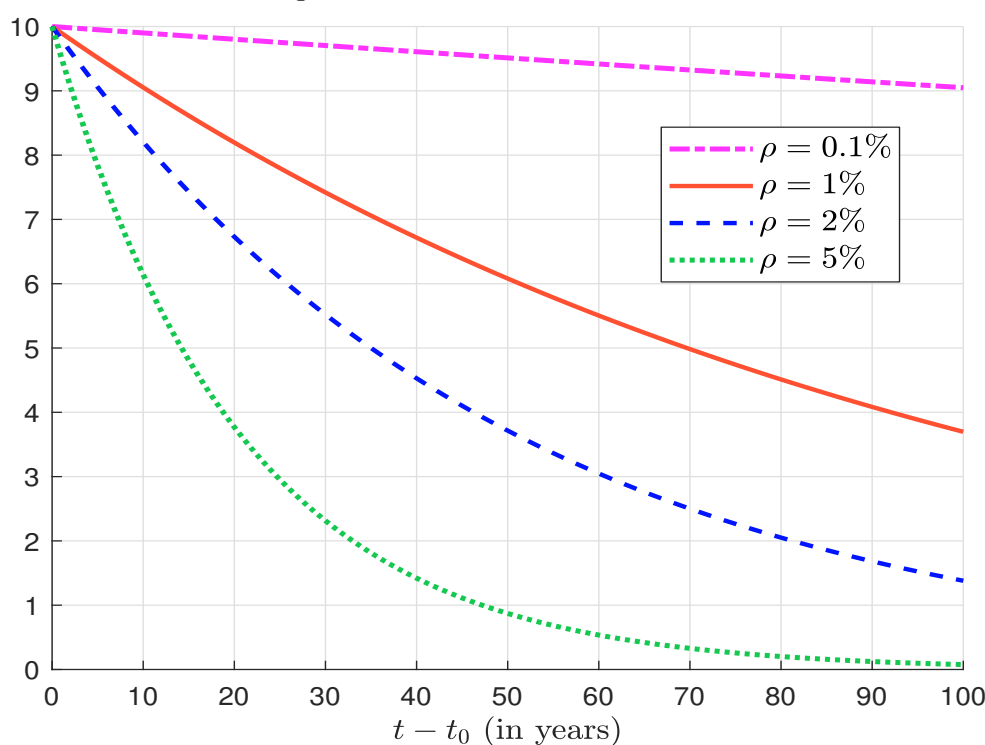
In order to illustrate the impact of ρ, we report the discounted value of \$100 in Figure 8.85. Since the time horizon of the economic analysis is relatively long (greater or equal to 100 years), we obtain very different discounted values. For instance, if ρ is equal to 2% (or 0.1%), the well-being of someone 100yr from now would be valued 86% (or 10%) less than the well-being of someone living today (Hänsel *et al.*, 2020). In this context, it is obvious that ρ is certainly the most important parameter when computing the social cost of carbon.

The time (or generational) discount rate ρ used in the DICE model is also called the pure rate of time preference and is related to the Ramsey rule[86]:

$$r = \rho + \alpha g_c$$

---

[86]See Box 8.12.

<div align="center">Box 8.12: The Ramsey rule</div>

We consider the optimization problem[a]:

$$v\left(k\left(t\right)\right) \quad = \quad \max_{c(t)} \int_0^\infty e^{-\rho t} \mathcal{U}\left(c\left(t\right)\right) \, \mathrm{d}t$$
$$\text{s.t.} \quad \mathrm{d}k\left(t\right) = rk\left(t\right) \, \mathrm{d}t - c\left(t\right) \, \mathrm{d}t$$

where $c\left(t\right) = C\left(t\right)/L\left(t\right)$ and $k\left(t\right) = K\left(t\right)/L\left(t\right)$ are the per capita consumption and capital, and r is the real interest rate. The Bellman equation is:

$$\max_{c(t)} \left\{ \mathcal{U}\left(c\left(t\right)\right) - \rho v\left(k\left(t\right)\right) + \frac{\partial v\left(k\left(t\right)\right)}{\partial k}\left(rk\left(t\right) - c\left(t\right)\right) \right\} = 0$$

or:

$$\rho v\left(k\left(t\right)\right) = \max_{c(t)} \left\{ \mathcal{U}\left(c\left(t\right)\right) - v'\left(k\left(t\right)\right) c\left(t\right) \right\} + rk\left(t\right) v'\left(k\left(t\right)\right)$$

If we assume that the utility function is CRRA — $\mathcal{U}\left(c\right) = c^{1-\alpha}/\left(1-\alpha\right)$, we obtain:

$$\max_c \left\{ \mathcal{U}\left(c\right) - v'\left(k\right) c \right\} \Leftrightarrow v'\left(k\right) = \mathcal{U}'\left(c\right) = c^{-\alpha}$$

implying that $c = v'\left(k\right)^{-1/\alpha}$ and:

$$\max_c \left\{ \mathcal{U}\left(c\right) - v'\left(k\right) c \right\} = \frac{c^{1-\alpha}}{1-\alpha} - c^{1-\alpha} = \frac{\alpha}{1-\alpha} c^{1-\alpha} = \frac{\alpha}{1-\alpha} v'\left(k\right)^{-(1-\alpha)/\alpha}$$

The Bellman equation becomes then:

$$\rho v\left(k\right) = \frac{\alpha}{1-\alpha} v'\left(k\right)^{-(1-\alpha)/\alpha} + rkv'\left(k\right)$$

We can verify that the optimal solution of the previous equation has the following expression:

$$v^\star\left(k\left(t\right)\right) = \left(\frac{\rho - r\left(1-\alpha\right)}{\alpha}\right)^{-\alpha} \frac{k\left(t\right)^{1-\alpha}}{1-\alpha}$$

The optimal consumption is then equal to:

$$c^\star\left(t\right) = \left(\frac{\partial v^\star\left(k\left(t\right)\right)}{\partial k}\right)^{-1/\alpha} = \frac{\rho - r\left(1-\alpha\right)}{\alpha} k\left(t\right)$$

It follows that:

$$\mathrm{d}k^\star\left(t\right) = rk^\star\left(t\right) \, \mathrm{d}t - c^\star\left(t\right) \, \mathrm{d}t = \left(r - \frac{\rho - r\left(1-\alpha\right)}{\alpha}\right) k^\star\left(t\right) \, \mathrm{d}t = \underbrace{\frac{r-\rho}{\alpha}}_{g_k} k^\star\left(t\right) \, \mathrm{d}t$$

In the Ramsey model, the optimal growth of capital is also equal to the optimal growth of consumption ($g_k = g_c$). Finally, we obtain the Ramsey rule:

$$g_c = \frac{\partial_t c\left(t\right)}{c\left(t\right)} = \frac{r-\rho}{\alpha}$$

---

[a]This derivation of the Ramsey rule has been suggested by Peter Tankov.

where r is the real interest rate, $g_c = \partial c(t)/c(t)$ is the growth rate of per capita consumption, and α is the consumption elasticity of the utility function (or the elasticity of substitution)[87]. The Ramsey rule can be interpreted as a condition for the optimality of intertemporal consumption choice. Therefore, r can be interpreted as the consumption discount rate, *the social discount rate* or *the discount rate on goods* (Nordhaus, 2019). There are then two ways to use the Ramsey rule. From the previous relationship, we can deduce the optimal value of ρ from the historical estimates of r and g_C: $\rho = r - \alpha g_c$. For example, if the real interest rate is equal to 4%, $g_c = 2\%$ and $\alpha = 1$, we obtain $\rho = 2\%$. Nordhaus (2019) justified the historical (or descriptive) approach as follows:

> "*This approach assumes that investments to slow climate change must compete with investments in other areas. The benchmark should therefore reflect the opportunity cost of investment. The descriptive approach yields a market rate of return in the neighborhood of 5% per year when risks are appropriately included.*"

According to Nordhaus (2007), the assumption $\rho = 0.10\%$ is not consistent with historical observations, because this implies a low goods discount rate, around 1% per year, while the descriptive approach yields a market rate of return in the neighborhood of 5% per year. The first approach contrasts with the normative (or prescriptive) approach, which consists in fixing directly ρ for ethical reasons. This is the position of Stern and Taylor (2007), who claims that there are "*many reasons for thinking that market rates and other approaches that illustrate observable market behavior cannot be seen as reflections of an ethical response to the issues at hand. There is no real economic market that reveals our ethical decisions on how we should act together on environmental issues in the very long term. Most long-term capital markets are very thin and imperfect.*". In fact, as shown by Gollier (2010, 2013), the Ramsey rule is not always valid when the uncertainty on economic growth is high. In particular, there is no reason to use the same rate to discount environmental impacts and monetary benefits.

Table 8.16: Global SCC under different discount rate assumptions

Discount rate	2015	2020	2025	2030	2050	CAGR
Stern	197.4	266.5	324.6	376.2	629.2	3.37%
Nordhaus	30.7	36.7	43.5	51.2	103.6	3.54%
2.5%	128.5	140.0	152.0	164.6	235.7	1.75%
3%	79.1	87.3	95.9	104.9	156.6	1.97%
4%	36.3	40.9	45.8	51.1	81.7	2.34%
5%	19.7	22.6	25.7	29.1	49.2	2.65%

Source: Nordhaus (2017a, Table 1, page 1520).

In Table 8.16, we report the estimated values of SCC, which was obtained by Nordhaus (2017a) under different discount rate assumptions. We better understand the call for sharp and immediate action by Nicholas Stern, since the social cost of carbon was equal to \$197.4/tCO$_2$ in 2015 when the discount rate is set to 0.1%!

---

[87]Hänsel *et al.* (2020, page 783) interpret the parameter α as measuring intertemporal inequality aversion:

> "*Due to diminishing marginal utility, the idea is that an additional \$1 is worth more to a poor person than to a rich one. In a growing economy, citizens in the future will be richer and their lower marginal utility motivates discounting. Suppose the economy grows at 2%. People living in 100yr will be seven times richer. If inequality aversion is the only reason for discounting, if $\alpha = 1$ (or 1.45), which corresponds to the values of the median expert (Nordhaus), the value of \$1 in 100yr is only 14 (or 6) cents.*"

8.3.2   Other models

In Table 8.17, we report a short list of IAMs. For instance, Grubb *et al.* (2021) counts 28 major climate economic models[88]. They make the distinction between stylized simple IAMs and complex IAMs. While the purpose of the first category is focused on the optimal path of the economy based on policy optimization, the second category generally aims to evaluate the impact of a climate scenario or a given policy on the economy.

Table 8.17: Main integrated assessment models

Model	Reference	Name
Stylized simple models		
DICE	Nordhaus and Sztorc (2013)	Dynamic Integrated Climate-Economy
FUND	Anthoff and Tol (2014)	Climate Framework for Uncertainty, Negotiation and Distribution
PAGE	Hope (2011)	Policy Analysis of the Greenhouse Effect
Complex models		
AIM/CGE	Fujimori *et al.* (2017)	Asia-Pacific Integrated Model/Computable General Equilibrium
GCAM	Calvin *et al.* (2019)	Global Change Assessment Model
GLOBIOM	Havlìk *et al.* (2018)	Global Biosphere Management Model
IMACLIM-R	Sassi *et al.* (2010)	Integrated Model to Assess Climate Change
IMAGE	Stehfest *et al.* (2014)	Integrated Model to Assess the Greenhouse Effect
MAGICC	Meinshausen *et al.* (2011)	Model for the Assessment of Greenhouse Gas Induced Climate Change
MAgPIE	Dietrich *et al.* (2019)	Model of Agricultural Production and its Impact on the Environment
MESSAGEix	Huppmann *et al.* (2019)	Model for Energy Supply Strategy Alternatives and their General Environmental Impact
REMIND	Aboumahboub *et al.* (2020)	REgional Model of INvestments and Development
WITCH	Bosetti *et al.* (2006)	World Induced Technical Change Hybrid

Source: Grubb *et al.* (2021, Table A1, pages 24-26) & Author's research.

The FUND model is available at www.fund-model.org. A `Julia` implementation can be found at https://github.com/fund-model/MimiFUND.jl. The PAGE model requires Microsoft Excel and the @RISK Excel plug-in, but Moore *et al.* (2018) propose a `Julia` open-source implementation at https://github.com/anthofflab/MimiPAGE2009.jl and http://anthofflab.berkeley.edu/MimiPAGE2009.jl/stable. The AIM/CGE is hosted at www-iam.nies.go.jp/aim/index.html. The GCAM model is available at http://jgcri.github.io/gcam-doc/index.html (documentation and open-source code). GLOBIOM has been developed using `GAMS`[89]. The source code is not available, but a graphical user interface and a ℝ interface is provided at https://iiasa.

---

[88]A more exhaustive list can be found in www.iamcdocumentation.eu, where a reference card is provided for each model.

[89]`GAMS` (General Algebraic Modeling System) is a high level modeling system for mathematical programming and optimization. It consists of a language compiler and a range of optimizer solvers, and requires a user license to be used (www.gams.com).

github.io/GLOBIOM. IMACLIM-R is not public. This is also the case of IMAGE, but we can download many data items, such as SSP scenario data, spatial data global land-use maps, scenario results and renewable energy data (www.pbl.nl/en/image/about-image). A compiled version of MAGGIC can be download at https://magicc.org. An online version is also available at https://live.magicc.org. The documentation of MAgPIE can be found at www.pik-potsdam.de/en/institute/departments/activities/land-use-modelling/magpie, while the link https://github.com/magpiemodel/magpie contains the model code and the ℝ implementation. For MESSAGEix, the documentation is available at https://docs.messageix.org/en/stable, while the 🐍 python programs can be found at https://github.com/iiasa/message_ix. The documentation of REMIND can be found at www.pik-potsdam.de/en/institute/departments/transformation-pathways/models/remind, while the link https://github.com/remindmodel/remind contains the ℝ implementation. Finally, the website www.witchmodel.org is dedicated to the WITCH model. The code is open-source (https://github.com/witch-team/witchmodel), but requires a GAMS license.

Stylized IAMs

Stylized integrated assessment models use the same structure as DICE (Figure 8.86). The FUND model has five main components: (1) population, production and income, (2) emissions, abatements and costs, (3) atmosphere and climate, (4) impacts (agriculture, forestry, water resources, energy consumption, sea level rise, ecosystems, human health, extreme weather, mortality), (5) optimization of the utility function, where the social welfare if a mixture of per capita income, damage of climate change and air pollution, and emission reduction costs Tol (1997). It also considers 16 regions. For its part, PAGE has two main modules (Hope, 2011; Moore *et al.*, 2018). Like the DICE model, the climate sensitivity is derived from a stylized GCM, but it explicitly introduces a sea level rise component and also considers CH_4 and N_2O forcing cycles, and not only the CO_2 cycle[90]. The economic module has four components: (1) population, production and income, (2) damage costs, (3) climate policy action and (4) consumption utility maximization. It is also possible to perform a regional analysis[91].

Figure 8.86: Stylized integrated assessment models
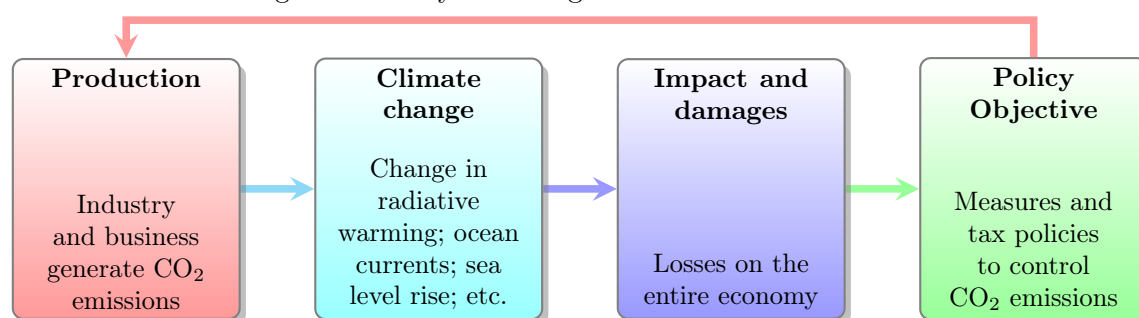
A comprehensive comparison of DICE, FUND and PAGE can be found in Diaz and Moore (2017a,b), and Rose *et al.* (2017a,b). It appears that parameter values and functional forms are not always the same (Diaz and Moore (2017a, Table 1, page 776), Rose *et al.* (2017b, Table S2, pages 2-5)), and damage estimates projected by the three models are dramatically different (Diaz

---

[90]We also find these additions in the FUND model
[91]The default regions are Africa, China, Europe, Latin America, other OECD countries, Russia, South-East Asia and the United States.

Figure 8.87: Histogram of the 150 000 US Government SCC estimates for 2020 with a 3% discount rate
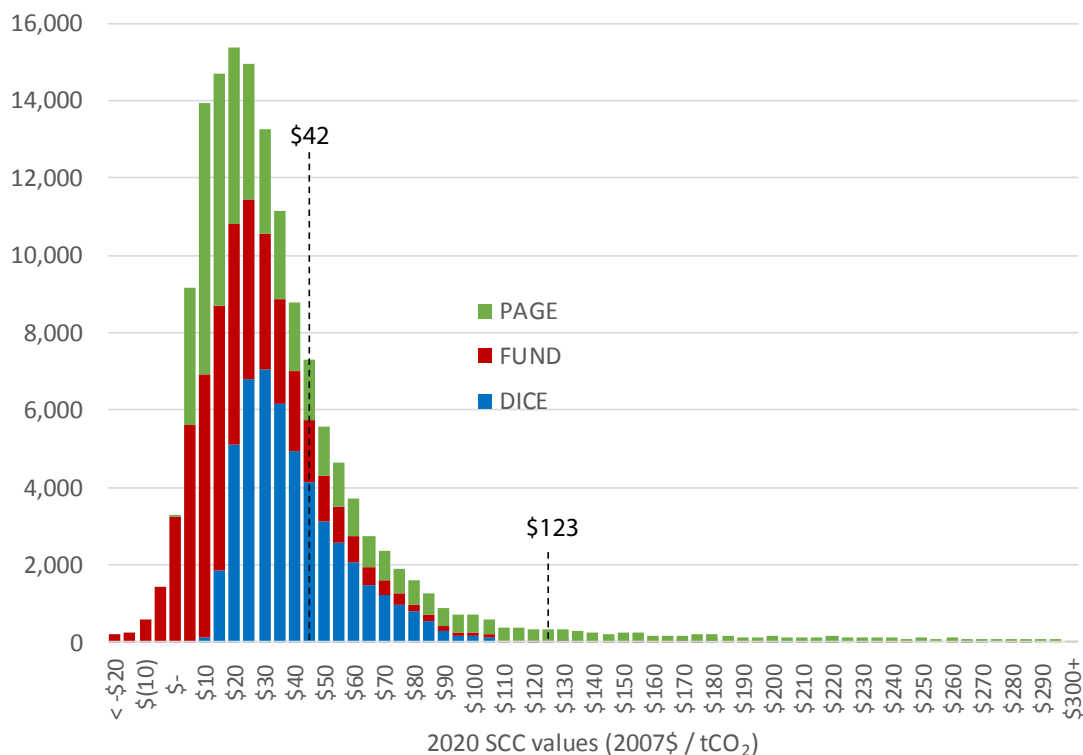
The figure combines the 50 000 2020 3% discount rate estimates from each of the three US Government models to illustrate their influence on the aggregate histogram that determines the official USG SCC for 2020 at 3%, which is equal to $42 (average) and $123 (95th percentile).

Source: (Rose *et al.*, 2017a, page 3).

and Moore, 2017a, Figure 2, page 777). In this context, the social cost of carbone reaches different levels depending on the model (Figure 8.87). In particular, Rose *et al.* (2017a) confirmed the ordering/raking PAGE ≻ DICE ≻ FUND that most academic studies have found when computing the social cost of carbon:

> "*We find significant variation in component-level behavior between models driven by model-specific structural and implementation elements, some resulting in artificial differences in results. These elements combine to produce model-specific tendencies in climate and damage responses that contribute to differences observed in SCC outcomes — producing PAGE SCC distributions with longer and fatter right tails and higher averages, followed by DICE with more compact distributions and lower averages, and FUND with distributions that include net benefits and the lowest averages.*" (Rose *et al.*, 2017a, page 1).

The previous stylized models have inspired many research projects and IAMs. This is especially true for DICE. For instance, the RESPONSE model of Pottier *et al.* (2015) and the ENTICE model of Popp (2004) are two variants of DICE. We can also mention the RICE model developed by Nordhaus and Yang (1996), which is a regional extension of the DICE framework with 12 regions. The RICE model introduces an important feature, which is the use of Negishi welfare weights with

multiple countries. Those weights are given by the relationship $\omega_k(t) \propto 1/\boldsymbol{\mathcal{U}}'(c_k(t))$ where k is the country index. In the case of the CRRA utility function $\boldsymbol{\mathcal{U}}(c) = c^{1-\alpha}/(1-\alpha)$, we obtain:

$$\omega_k(t) = \frac{c_k(t)^\alpha}{\sum_{k'=1}^{n_C} c_{k'}(t)^\alpha}$$

where n_C is the number of countries. Then, the social welfare function becomes:

$$W\left(s_1(t), \ldots, s_{n_C}(t), \mu_1(t), \ldots, \mu_{n_C}(t),\right) = \sum_{k=1}^{n_C} \sum_{t=t_0+1}^{T} \omega_k(t) \frac{L_k(t) \boldsymbol{\mathcal{U}}(c_k(t))}{(1+\rho_k)^{t-t_0}}$$

Negishi welfare weights are generally used to find a market equilibrium by optimising the weighted objective function (or Negishi social welfare function). However, these weights have been criticized to be inequitable, since the current wealth distribution would change. For instance, Stanton (2011) pointed out that the choice of weights are related to the regional inequality issue:

> "*In a global climate policy debate fraught with differing understandings of right and wrong, the importance of making transparent the ethical assumptions used in climate-economics models cannot be overestimated. [...] Negishi weights freeze the current distribution of income between world regions; without this constraint, IAMs that maximize global welfare would recommend an equalization of income across regions as part of their policy advice. With Negishi weights in place, these models instead recommend a course of action that would be optimal only in a world in which global income redistribution cannot and will not take place.*"

This question of regional inequality is central when considering climate change and the transition risk. For instance, Dennig *et al.* (2015) choose $\omega_k(t) = 1$ in their Nested Inequalities Climate-Economy (NICE) model, which is a direct extension of RICE, because they restrict consumption redistribution between regions.

Complex IAMs

As their name suggest, complex IAMs are complex. While stylized simple IAMs are generally developed by one or two academics, complex IAMs are built by an institution with a big research team and are continuously improved and updated[92]. For that reason, it seems a pointless exercise to present one of them in details. For instance, GCAM has been developed at Pacific Northwest National Laboratory for over 30 years. The source code is today about 120 Mo, there are 910 CSV data files for a total of 230 Mo. Therefore, it is better to understand how they work and how they can be used. In Figure 8.88, we have reported the global structure of GCAM. In this type of complex models, the economic component is a small part of the structure. Furthermore, these big complex models are generally split into elementary complex models. In Figure 8.89, we have reported the structure of GLOBIOM. This model is specialized in agriculture, bioenergy and forestry. This model can communicate with MESSAGEix and other models as shown in Figure 8.90. With GCAM and MESSAGE, REMIND is another key complex IAM (Figure 8.91). Again, we notice that REMIND

---

[92]AIM/CGE, IMAGE, GCAM, IMACLIM-R, MESSAGE-GLOBIOM, REMIND-MAgPIE, WITCH are respectively hosted by the National Institute for Environmental Studies (NIES, Japan), the PBL Environmental Assessment Agency (Netherlands), the Pacific Northwest National Laboratory (PNNL, US), the Centre international de recherche sur l'environnement et le développement (CIRED, France), the International Institute for Applied Systems Analysis (IIASA, Austria), the Potsdam Institute for Climate Impact Research (PIK, Germany) and the European Institute on Economics and the Environment (RFF-CMCC-EIEE, Italy).

Figure 8.88: Linkages between the major systems in GCAM
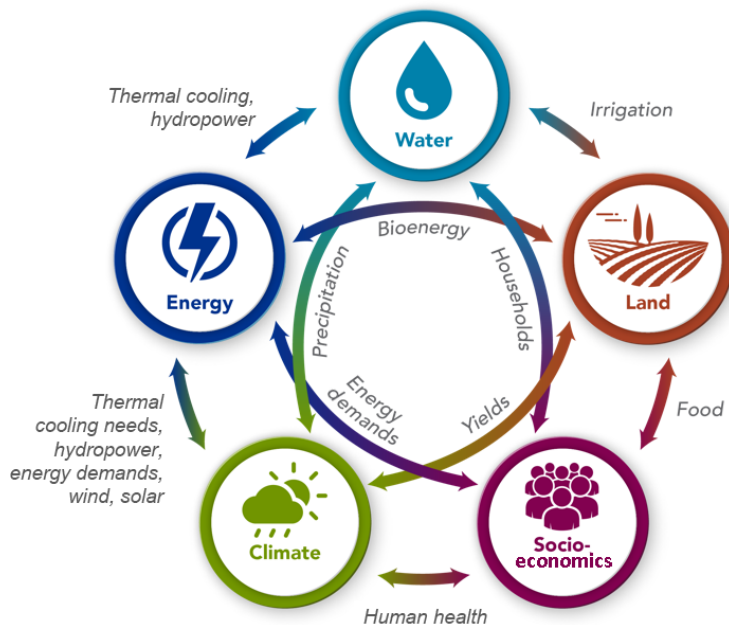
Figure 8.89: The main land-use sectors of GLOBIOM
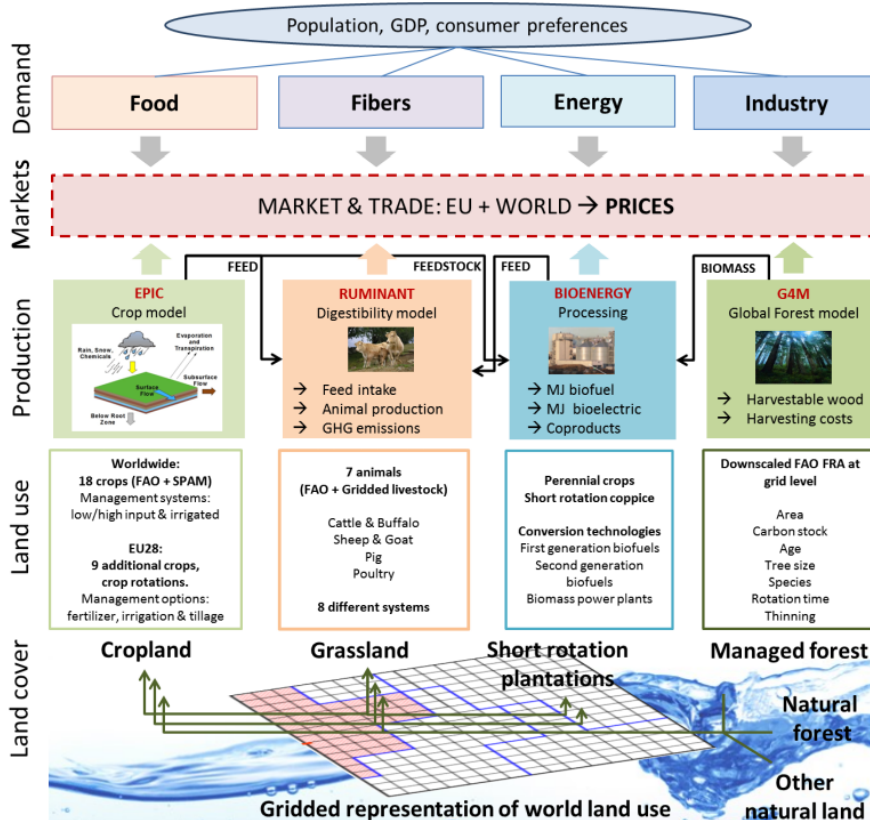
Figure 8.90: Overview of the IIASA IAM framework

Source: https://docs.messageix.org/projects/global/en/latest/overview/index.html.
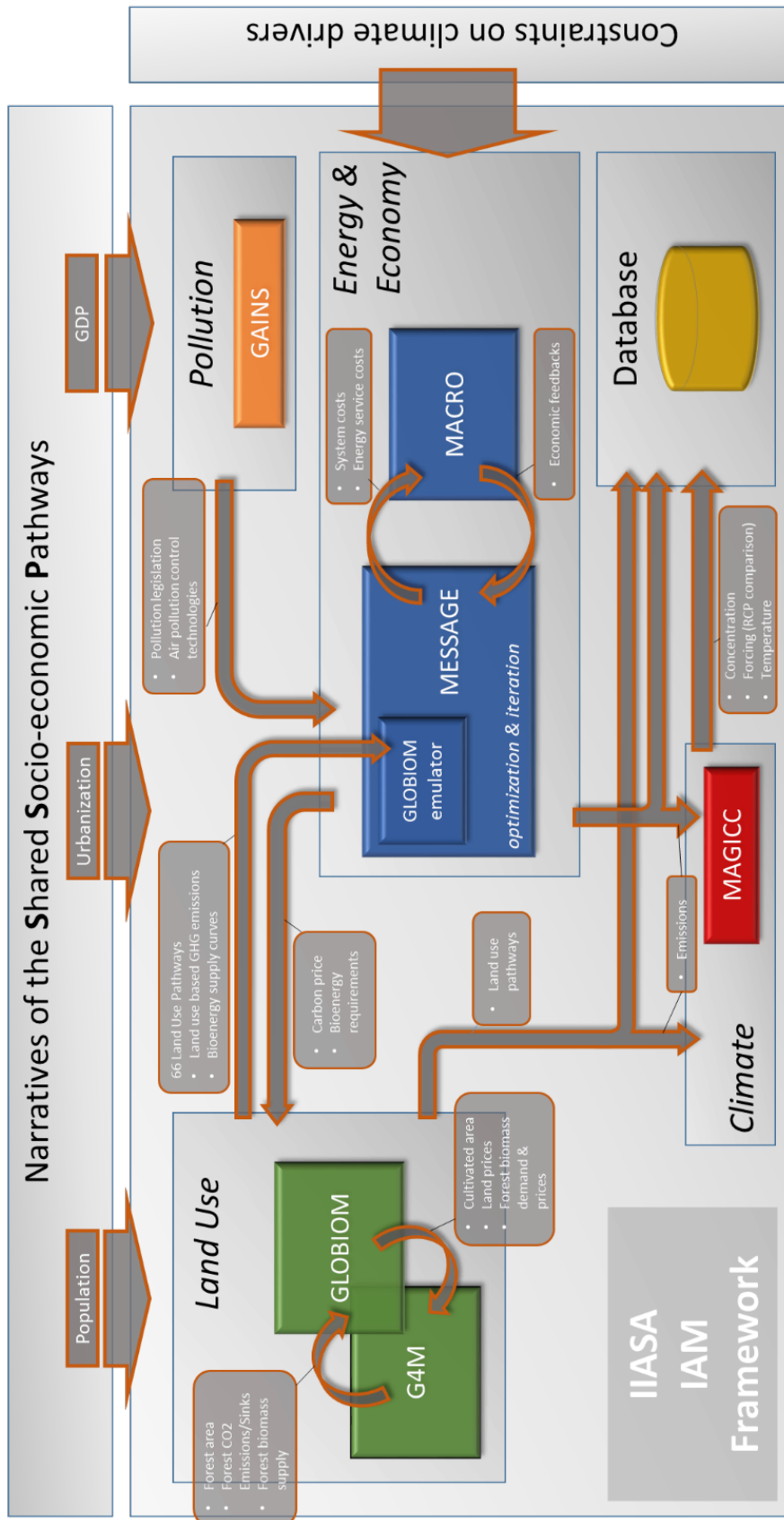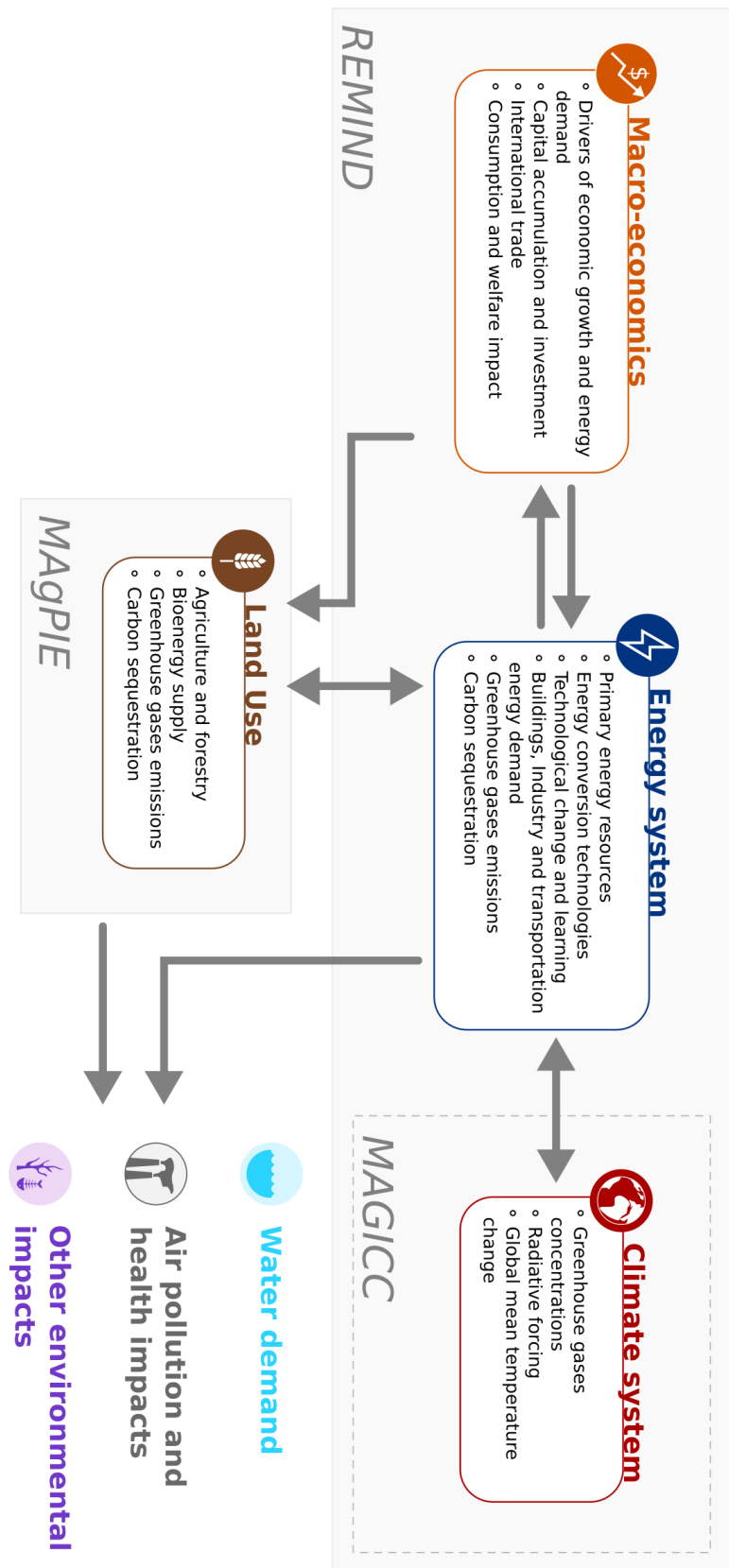
Figure 8.91: The Remind-MAgPIE framework

can be connected with MAgPIE, which is specialized in the land use. We also remark that it can take input data from MAGICC, exactly like MESSAGEix, and it can also feed MAGICC with output data.

Even if these models can be used in an optimization mode, whose objective is to compute optimal paths and perform a cost benefit analysis, they are mainly exploited in order to evaluate a climate policy or to simulate climate/economic trajectories with respect to a given scenario. Like any models, even if they are extensively used by policy markers and the scientific community, they have often been criticised.

Criticisms of integrated assessment models

According to Pindyck (2017), "*IAM-based analyses of climate policy create a perception of knowledge and precision that is illusory and can fool policymakers into thinking that the forecasts the models generate have some kind of scientific legitimacy*". In particular, Robert Pindyck considers that the most important flaws are the following:

- Certain inputs, such as the discount rate, are arbitrary, but they can have a big impact on the model outputs, such as the social cost of carbon;

- There is a lot of uncertainty about climate sensitivity and the temperature trajectory;

- Modeling damage functions is arbitrary, because we have little data to estimate the relationship between an increase in temperature and gross domestic product;

- IAMs are unable to consider tail risk, *i.e.* a tipping point or the likelihood/possible impact of a catastrophic climate outcome.

In Table 8.18, we report another list of criticisms found by Gambhir *et al.* (2019). We can multiply this type of criticism, and it is relatively essay to do it. We know that all models have flaws and are a simplification of reality. Nevertheless, to repeat the words of George box, "*all models are wrong, but some are useful*". For instance, we notice that Pindyck (2017) only referenced DICE, FUND and PAGE, while Gambhir *et al.* (2019) focused on complex IAMs. In both cases, criticisms are general and dot not concern a specific model. In this context, it is not always obvious to form an opinion on a model versus another one. For complex IAMs, it is almost impossible. As well, we notice that most comparative studies concern the stylized simple models.

8.3.3 Scenarios

As we have already explained, there are two ways to use IAMs: policy optimization and policy/scenario evaluation/simulation. In this section, we focus on the second approach, and more precisely on the scenario simulation. We distinguish two types of simulation outputs: shared socioeconomic pathways (SSPs) and macroeconomic pathways. Before presenting them, we briefly present climate scenarios, which are the input of the evaluation process.

Figure 8.92: Scenario evaluation
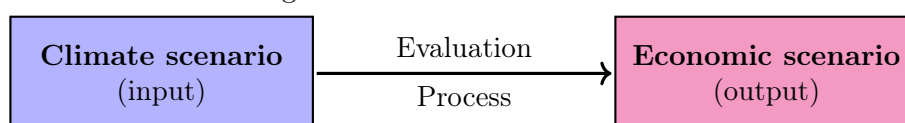
Table 8.18: Main integrated assessment models

Criticism category	Example of specific criticism
Lack of transparency	Lack of documentation
Inappropriate input assumptions	Low share of solar PV in mitigation modelling exercises
Lack of clarity on model inputs versus outputs	Energy demand reduction as a result of model choice or modeler input?
Reliance of mitigation costs on baseline assumptions	Significant differences in costs of achieving mitigation
Inadequate representation of innovation processes	Spillover between low-carbon technologies
Lack of representation of behavioural and economic systems	Customer behavior changes
Lack of assessment of real-world feasibility	Feasibility of pathways given full consideration of social, political and technical barriers
Lack of interaction with other policy goals	Lack of consideration of mitigation pathways in light of other policy goals such as energy security
Lack of representation of fine temporal and geographical scale	Geographical dispersion of electricity systems
Over-reliance on negative emissions technologies	Unrealistically high levels of negative emissions technologies

Source: Adapted from Gambhir *et al.* (2019, Table 1, page 10).

Climate scenarios

It is important to understand the different climate scenarios that are used to simulate economic scenarios. In what follows, we focus on four main scenario databases:

- The representative concentration pathways (RCPs) that can be found in the IPCC AR5 (IPCC, 2013);

- The decarbonization pathway scenarios developed by the IEA (IEA, 2017);

- The 1.5°C scenarios proposed by the IPCC SR15 (IPCC, 2018) and the IEA (IEA, 2021);

- The new scenarios for the future published in the IPCC AR6 (IPCC, 2022).

The RCP scenarios   A representative concentration pathway (RCP) is a greenhouse gas concentration trajectory adopted by IPCC (2013). The AR5 scenario database comprises 31 models and 1 184 scenarios. Among these different scenarios, four pathways have been selected to represent four reference scenarios:

1. RCP 2.6: GHG emissions start declining by 2020 and go to zero by 2100 (IMAGE);

2. RCP 4.5: GHG emissions peak around 2040, and then decline (MiniCAM);

3. RCP 6.0: GHG emissions peak around 2080, and then decline (AIM);

4. RCP 8.5: GHG emissions continue to rise throughout the 21st century (MESSAGE).

Figure 8.93: Total radiative forcing (in W/m^2)
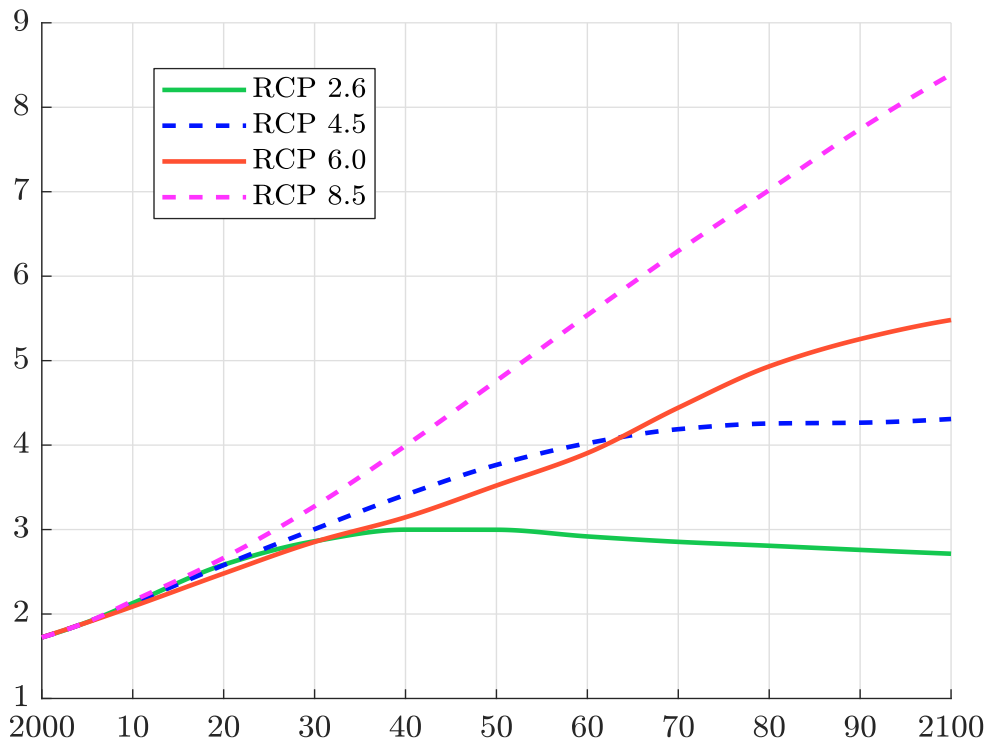
Figure 8.94: Greenhouse gas concentration trajectory
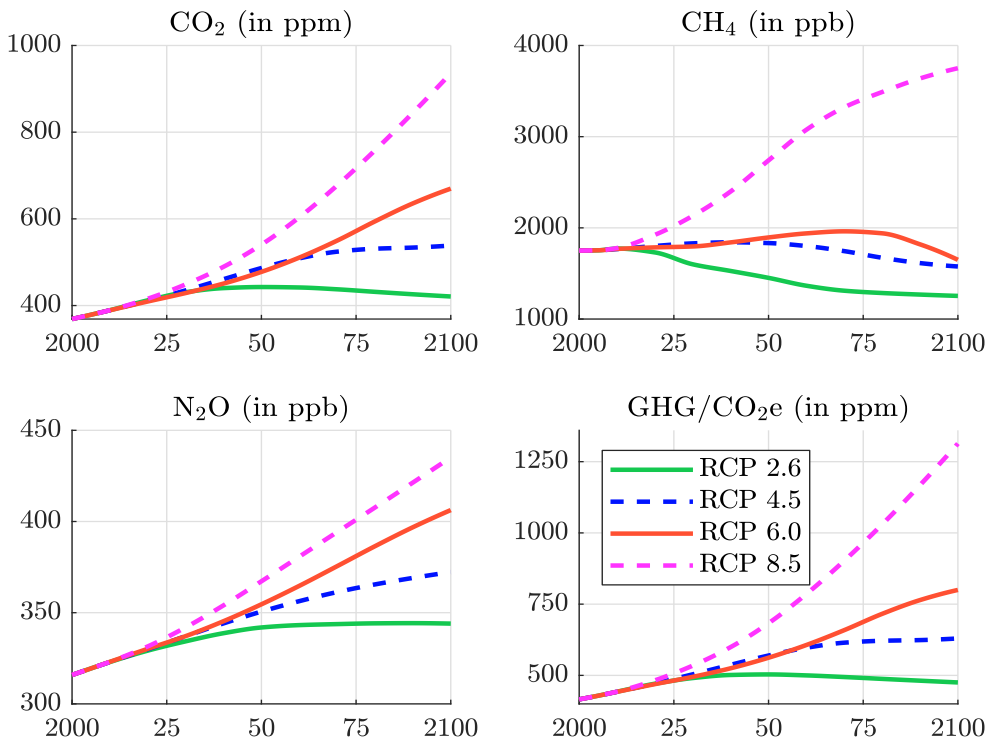
Figure 8.95: Greenhouse gas emissions trajectory

Source: https://tntcat.iiasa.ac.at/RcpDb.
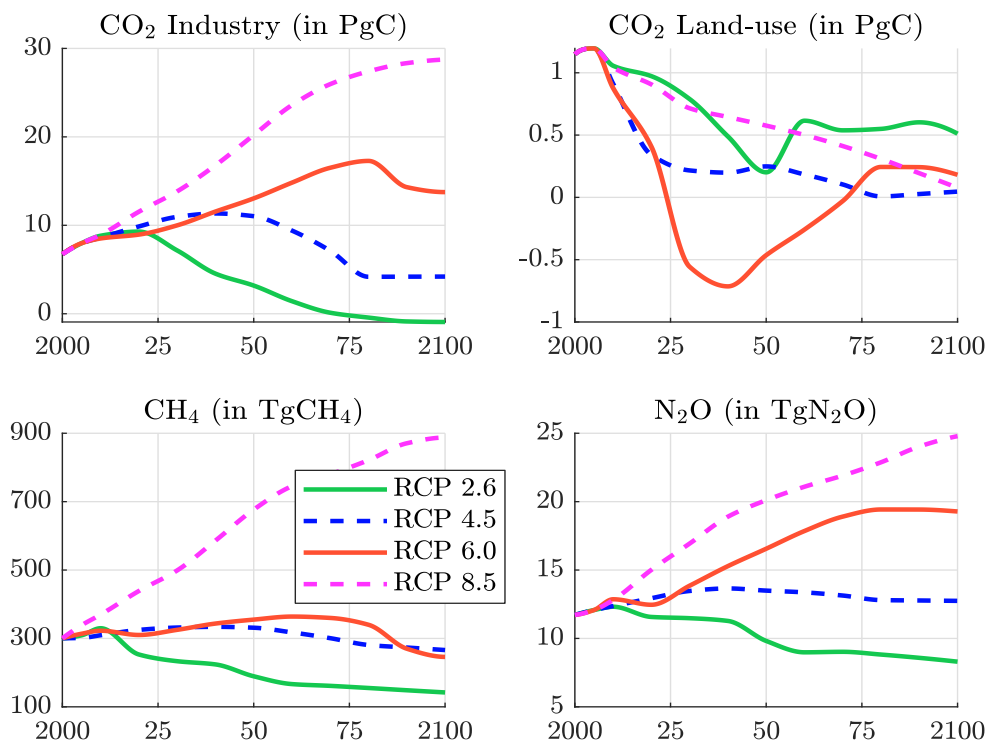
Figure 8.96: Total GHG emissions trajectory (in $GtCO_2e$)

Source: https://tntcat.iiasa.ac.at/RcpDb.
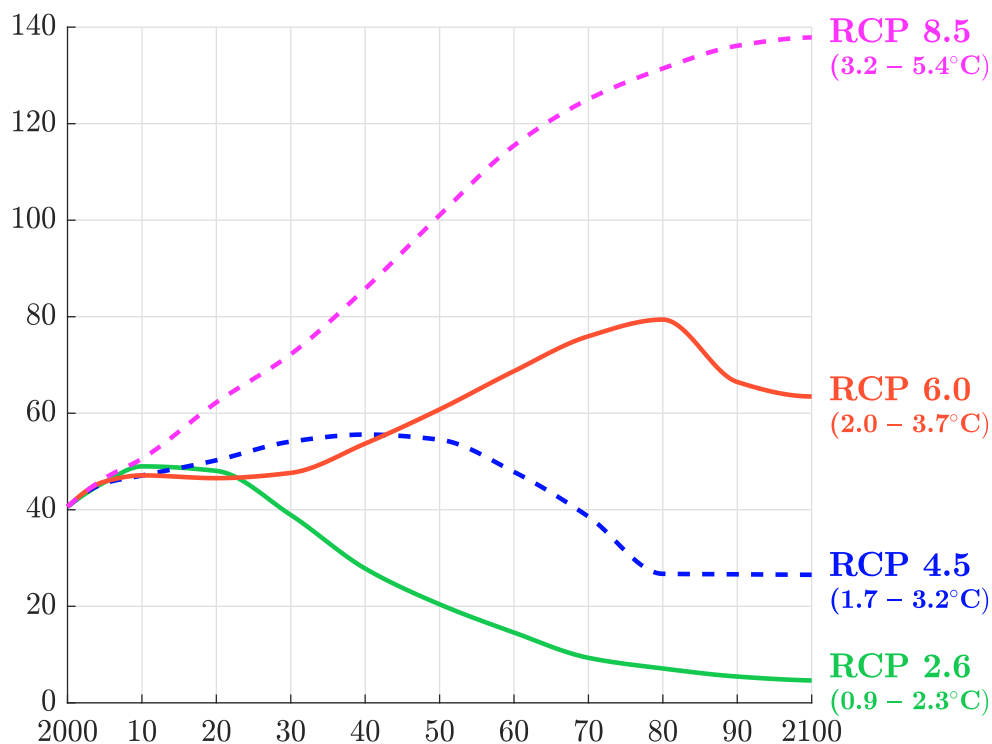
The name of each scenario refers to the value of the radiative forcing by 2050 (Figure 8.93). Figure 8.94 shows the GHG concentration trajectory for CO_2, CH_4 and N_2O gases. We also report the CO_2e concentration, which is the aggregation of the several GHG concentrations. In Figure 8.95, we show the emissions trajectories for the three main greenhouse gases. For CO_2, we distinguish between two categories: industry and land-use change. Finally, we give the total GHG emissions[93] trajectories in Figure 8.95. We notice that the RCP scenario database do not provide a temperature scenario, because the relationship between GHG emissions and temperature depends on several factors. This is why we can only provide a temperature range for each RCP scenario (Figure 8.96).

The IEA scenarios   As energy represents the largest portion of GHG emissions, the most common way to reduce global emissions is to impose a global shift in the energy supply or demand. To encourage this shifting, the international energy agency has developed energy pathways. These scenarios are not a forecast of the future but explore the different possibilities across the energy system. For instance, IEA (2017) presented three pathways for energy sector development to 2060. The reference technology scenario (RTS) was a baseline scenario that took into account existing climate-related commitments by countries, while the 2°C scenario (2DS) corresponded to a rapid decarbonisation pathway in line with international policy goals. The beyond 2°C scenario (B2DS) was a variant of the 2DS, where the energy sector reached carbon neutrality by 2060 to limit future temperature increases to 1.75°C by 2100. The previous three scenarios were replaced by two new scenarios in 2020. The sustainable development scenario (SDS) "*sets out the major changes that would be required to reach the key energy-related goals of the UN sustainable development agenda*", including a universal access to modern energy by 2030. This scenario is consistent with reaching net zero emissions by around 2070. The previous RTS scenario becomes the stated policies scenario (STEPS). Finally, the last edition of the IEA's Energy Technology Perspectives (ETP) only focuses on the NZE[94] and the announced pledges scenario (APS), which assumes that governments will meet all the climate-related commitments they have announced.

The international energy agency evaluates these different scenarios using its Global Energy and Climate (GEC) model[95]. This model covers 26 regions individually with dedicated bottom-up mod-elling for final energy demand, energy transformation and energy supply. Using assumptions on population, GDP, fossil fuel prices and ressources (crude oil, natural gas and steam coal), CO_2 prices, electricity and technology costs, the GEC model predicts primary energy demand and sup-ply by sources (oil, coal, natural gas, nuclear, biomass and waste, hydro, geothermal, wind, solar, hydrogen), electricity generation and capacity, CO_2 and CH_4 emissions, investment needs and costs, and materials and critical minerals demand. For instance, we report the CO_2 emissions estimated in ETP 2017 in Figure 8.97. We notice that we can not achieve the degC2 scenario when considering the country commitments (RTS).

---

[93]They have been calculated as follows:

$$\text{Total GHG emissions in GtCO}_2\text{e} \quad = \quad \begin{aligned} &\text{CO}_2 \text{ emissions in PgC} \times \frac{44}{12} + \\ &\text{CH}_4 \text{ emissions in Tg} \times \frac{28}{1000} + \\ &\text{N}_2\text{O emissions in Tg} \times \frac{265}{1000} \end{aligned}$$

This formula uses the AR5 GWP values (see Section 9.1.1 on page 470), the mass percentage of carbon in carbon dioxide, which is equal to $12/44 = 27.27\%$ (because carbon has an atomic mass of 12 and oxygen has an atomic mass of 16, implying that CO_2 has an atomic mass of 44), and the petagram/teragram conversion factor (1 Pg = 1000 Tg).
[94]See page 519.
[95]Its description is available at `www.iea.org/reports/global-energy-and-climate-model`.

Figure 8.97: Direct CO_2 emissions (in Gt)
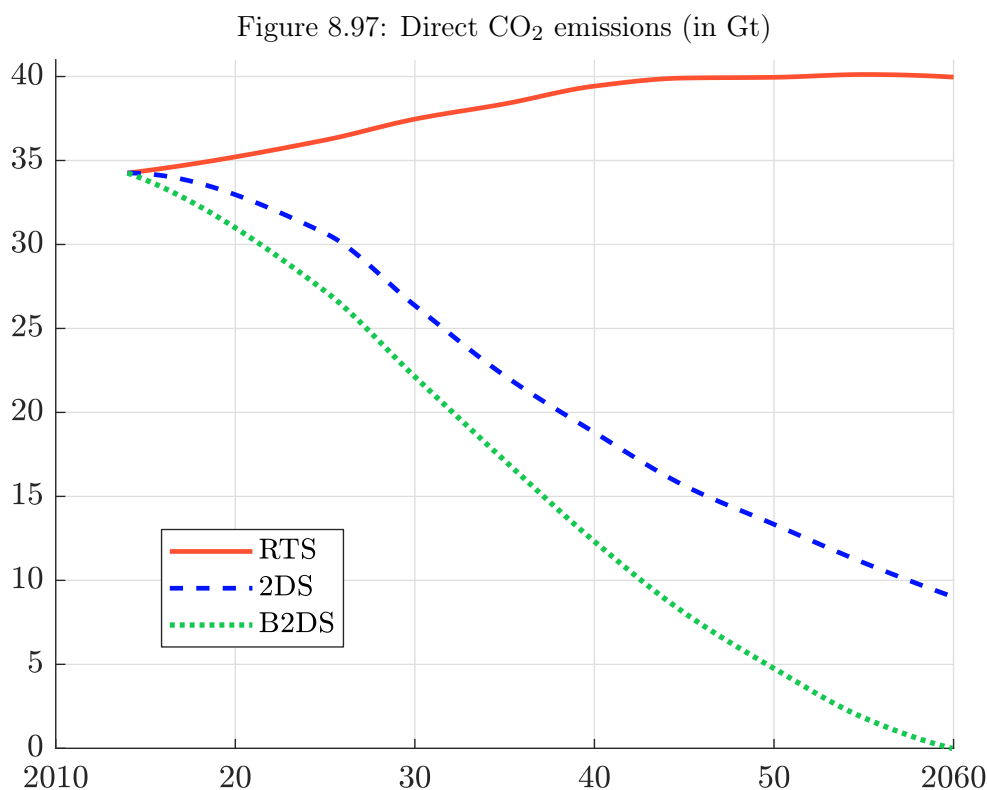
The 1.5°C **scenarios** In 2018, IPCC issued a special report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways. According to the report, "*limiting global warming to 1.5°C, compared with 2°C, could reduce the number of people both exposed to climate-related risks and susceptible to poverty by up to several hundred million by 2050*". While the goal of the climate Paris Agreement was to limit global temperature increase to well below 2°C, the IPCC reaffirms then that states must pursue efforts to limit the temperature increase to 1.5°C, because the cost-benefit analysis of the 1.5°C scenario versus 2°C scenario shows that we can reach some tipping points in the second scenario that will generate too much costs for the society, governments and people.

The IPCC SR15 database contains 177 scenarios computed with 24 models[96]. We filter the database in order to keep climate scenarios, whose explicit objective is 1.5°C or well below 2°C. We obtain 114 scenarios, whose CO2 emissions trajectory is reported in Figure 8.98. For each date, we estimate the mean and the standard deviation and compute the 50% and 956% confidence interval in Figure 8.99. We notice that the emissions are close to zero by 2050 on average. This is why the 11.5°C scenario is generally associated to the net zero emissions by 2050 (NZE) scenario. We must be careful because these figures correspond to expected values, meaning that there is a significant probability that the temperature increase can be greater than 1.5°C even if the scenario occurs. For instance, we have reported the global mean temperature of the 114 scenarios, which has been computed with the MAGICC 6 model (Figure 8.100). We observe that the temperature can be

---

[96]AIM/CGE (2.0 & 2.1), C-ROADS-5.005, GCAM 4.2, GENeSYS-MOD 1.0, IEA 2017 models (ETP & WEM), IMAGE (3.0.1 & 3.0.2), MERGE-ETL 6.0, MESSAGE (V.3, GLOBIOM & ix-GLOBIOM), POLES (ADVANCE, CD-LINKS & EMF33), REMIND (1.5 & 1.7), REMIND-MAgPIE (1.5 & 1.7-3.0), Shell World Energy Model 2018, and WITCH-GLOBIOM (3.1, 4.2 & 4.4).

Figure 8.98: IPCC 1.5°C scenarios of CO_2 emissions

Figure 8.99: Confidence interval of the average IPCC 1.5°C scenario
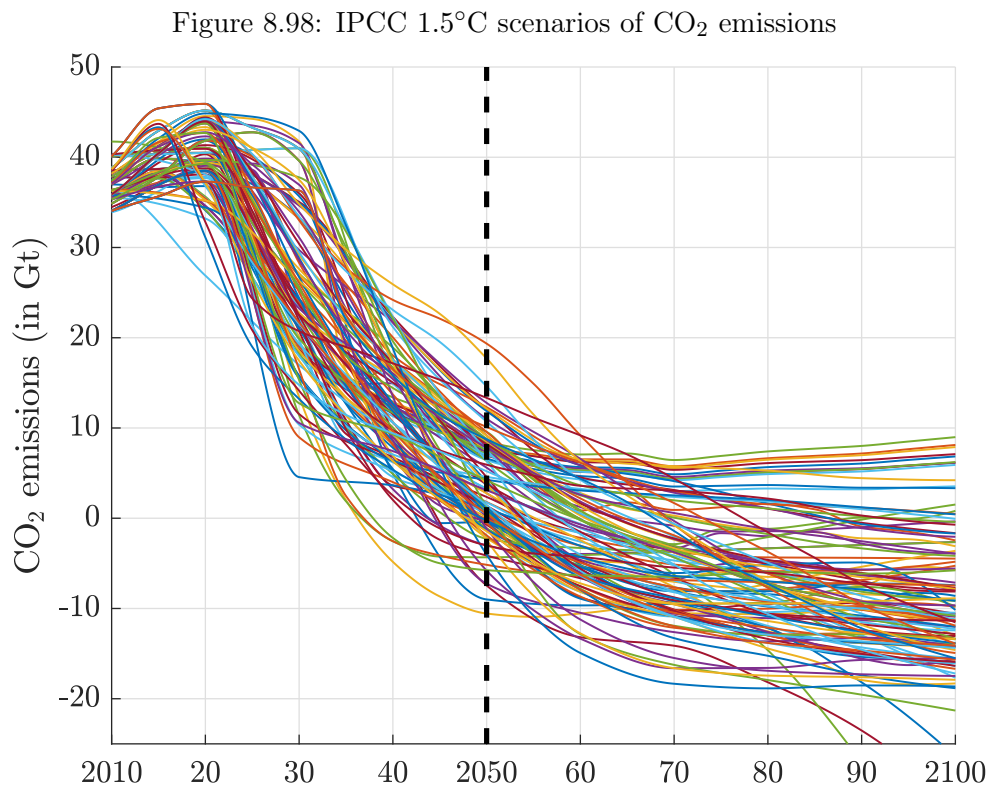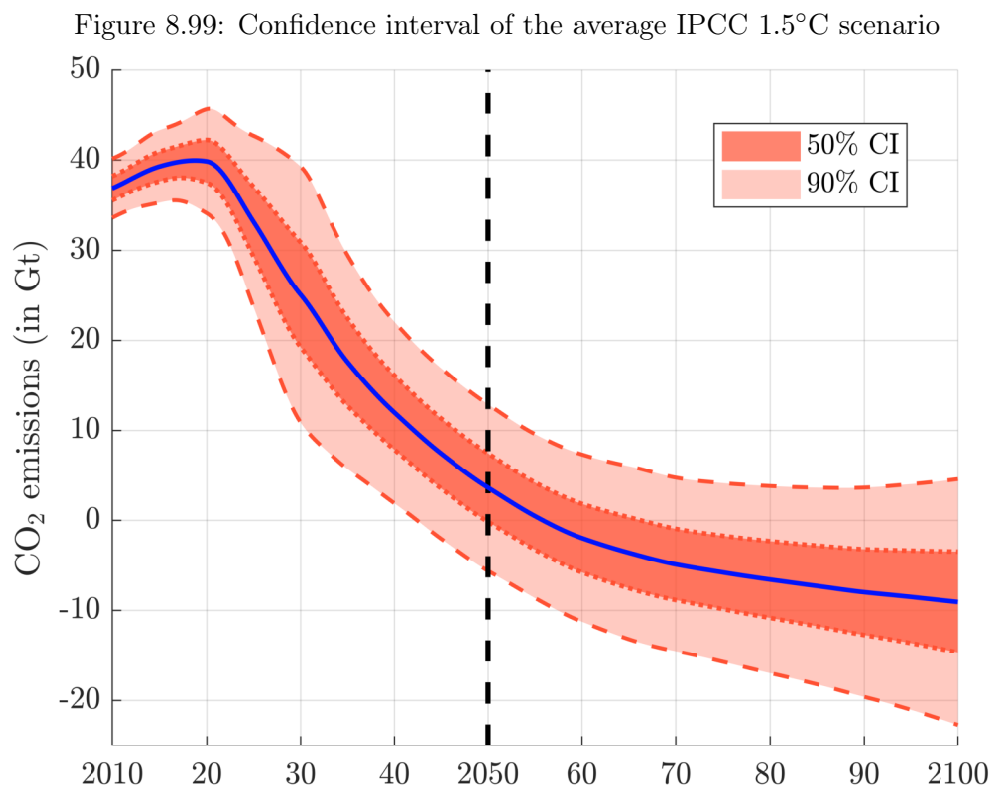
Figure 8.100: IPCC 1.5°C scenarios of the global mean temperature
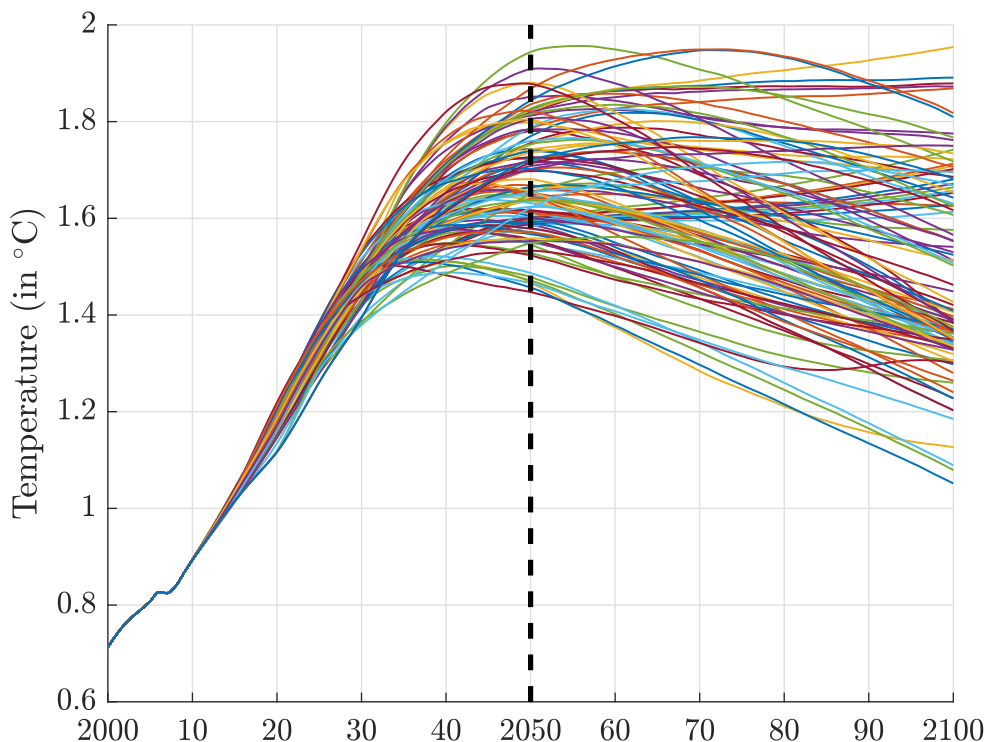
greater than 1.5°C by 2100, since only 56% of the 114 scenarios produce an expected temperature below this threshold. In Figures 8.101 and 8.102, we also give the evolution of the exceedance probabilities. We notice that the probability to observe a temperature greater than 1.5°C by 2050 is greater than 60%, while it is equal to only 15% if the threshold is 2°C.

The AR6 scenarios   IPCC (2022) has updated the RCP scenarios. The new dataset[97] contains 188 models, 1 389 scenarios, 244 countries and regions, and 1 791 variables, which can be split into six main categories:

- Agriculture: agricultural demand, crop, food, livestock, production, etc.

- Capital cost: coal, electricity, gas, hydro, hydrogen, nuclear, etc.

- Energy: capacity, efficiency, final energy, lifetime, OM cost, primary/secondary energy, etc.

- GHG impact: carbon sequestration, concentration, emissions, forcing, temperature, etc.

- Natural resources: biodiversity, land cover, water consumption, etc.

- Socio-economic variables: capital formation, capital stock, consumption, discount rate, employment, expenditure, export, food demand, GDP, Gini coefficient, import, inequality, interest rate, investment, labour supply, policy cost, population, prices, production, public debt, government revenue, taxes, trade, unemployment, value added, welfare, etc.

For example, we have reported the distribution of some output variables on page 390.

---

[97]The amount of data is huge (3 Go).

Figure 8.101: Confidence interval of the exceedance probability $\Pr\{\mathcal{T} > 1.5°C\}$ (MAGICC 6)

Source: https://data.ene.iiasa.ac.at/iamc-1.5c-explorer.
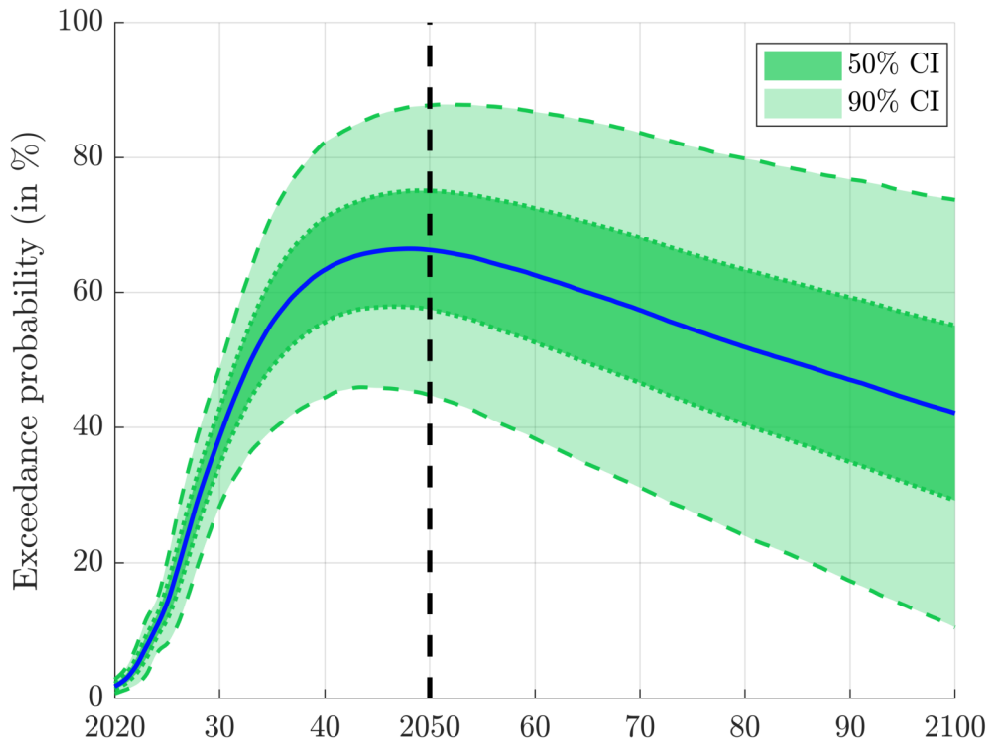
Figure 8.102: Confidence interval of the exceedance probability $\Pr\{\mathcal{T} > 2°C\}$ (MAGICC 6)

Source: https://data.ene.iiasa.ac.at/iamc-1.5c-explorer.
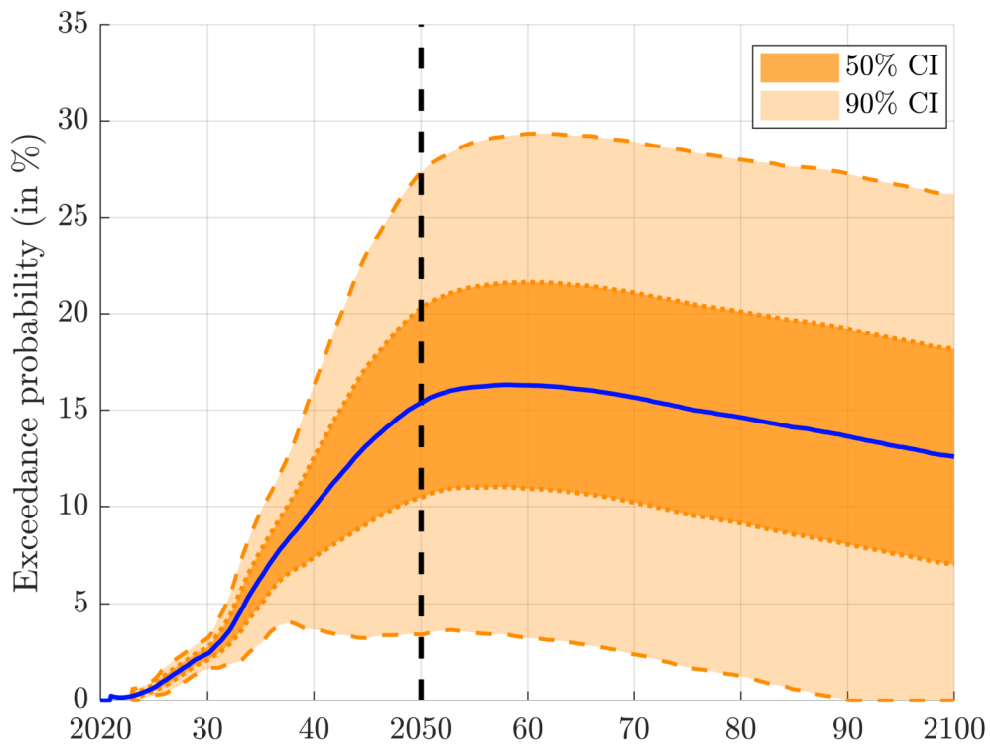
Figure 8.103: Histogram of some AR6 output variables by 2100

Source: https://data.ene.iiasa.ac.at/ar6.
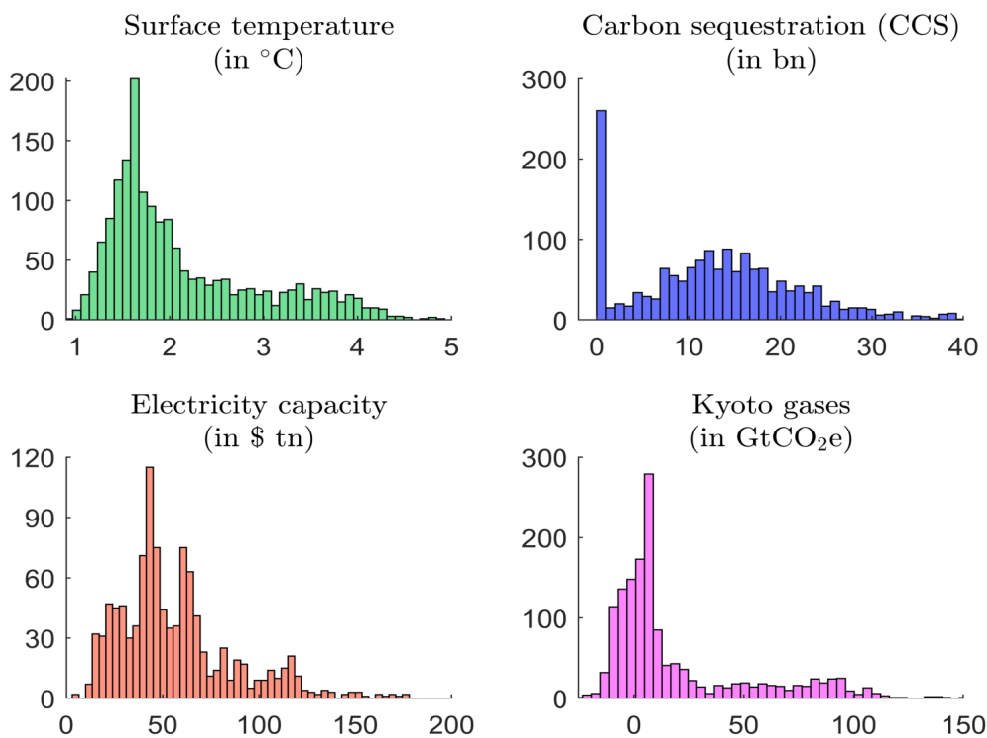
Figure 8.104: Histogram of some AR6 output variables by 2100

Source: https://data.ene.iiasa.ac.at/ar6.
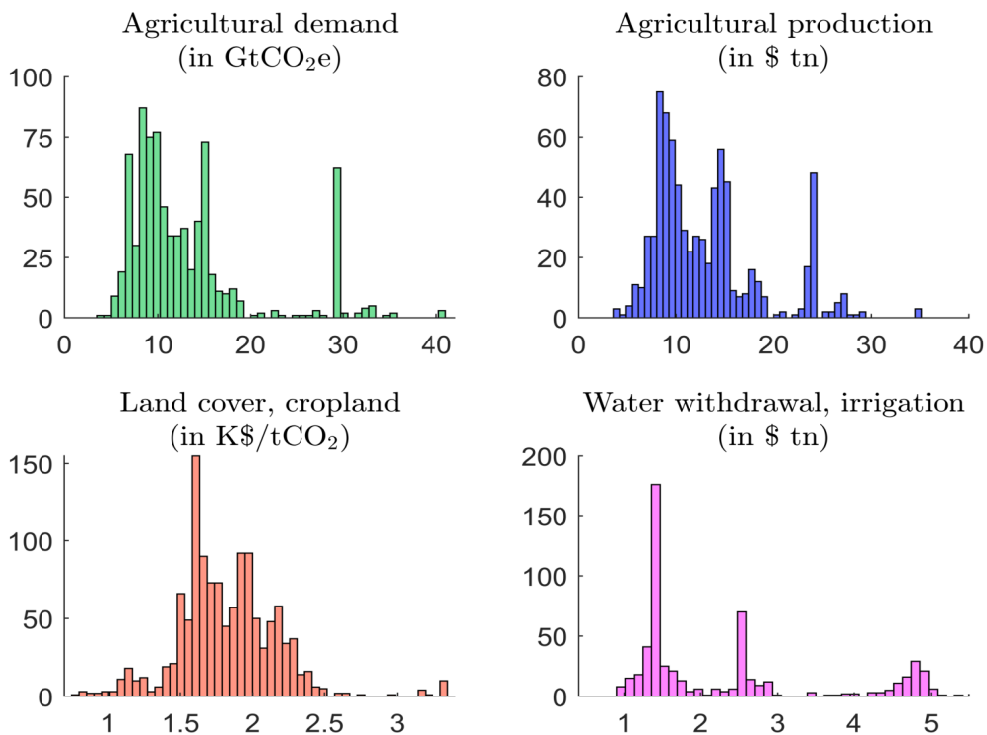
Shared socioeconomic pathways

Shared socioeconomic pathways are a set of five narrative scenarios proposed by O'Neill *et al.* (2014). These narratives describe alternative economic developments that are related to challenges for climate change adaptation and mitigation (Figure 8.105). According to (O'Neill *et al.*, 2014, page 389), "*the SSPs describe plausible alternative trends in the evolution of society and natural systems over the 21st century at the level of the world and large world regions; They consist of two elements: a narrative storyline and a set of quantified measures of development. SSPs are reference pathways in that they assume no climate change or climate impacts, and no new climate policies*".

Figure 8.105: The shared socioeconomic pathways
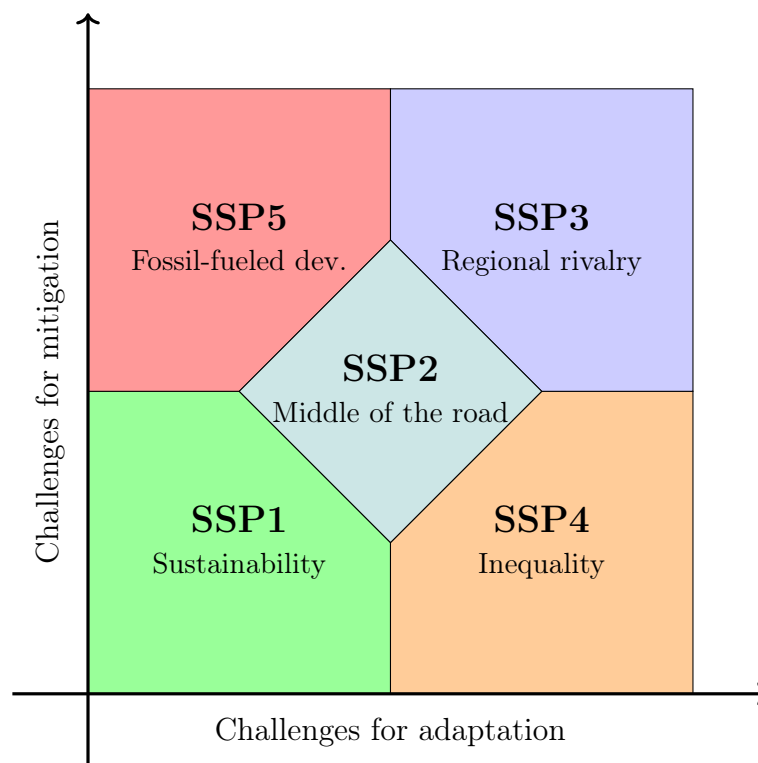
The concept of SSP is clarified in O'Neill *et al.* (2017):

> "*The SSP narratives [are] a set of five qualitative descriptions of future changes in demographics, human development, economy and lifestyle, policies and institutions, technology, and environment and natural resources. [...] Development of the narratives drew on expert opinion to (1) identify key determinants of the challenges [to mitigation and adaptation] that were essential to incorporate in the narratives and (2) combine these elements in the narratives in a manner consistent with scholarship on their interrelationships. The narratives are intended as a description of plausible future conditions at the level of large world regions that can serve as a basis for integrated scenarios of emissions and land use, as well as climate impact, adaptation and vulnerability analyses.*"

In particular, they described the five pathways and developed the narrative storyline for each of them (Figure 8.106). SSP1 and SSP2 corresponds to smooth scenarios, which are characterized by

Figure 8.106: The shared socioeconomic pathways

SSP1

Sustainability (Taking the Green Road)
Low challenges for both mitigation and adaptation, rapid development

SSP2

Middle of the Road
Moderate challenges for mitigation and adaptation

SSP3

Regional Rivalry (A Rocky Road)
High challenges for both mitigation and adaptation —
Concern about competitiveness, security and regional
conflict pushing countries to focus on regional issues

SSP4

Inequality (A Road Divided)
Low challenges for mitigation, high for adaptation — Unequal invest-
ment in human capital, concentration of power in a small business elite

SSP5

Fossil-fueled Development (Taking the Highway)
High challenges for mitigation, low for adaptation

Source: O'Neill *et al.* (2017).

a relative easy transition toward a sustainable world. SSP1 implies a shift from historical patterns, less resource-intensive lifestyles, an improvement of natural resource efficiency and a global green growth strategy. SSP2 may be interpreted as a continuation of the economic development that was observed in the twentieth century. Emerging economies become rapidly developed and middle-income economies, global population growth is under control and moderate, but inequalities remains. SSP3 describes an increasingly individualist world with more nationalism and regional conflicts. In this situation, countries mainly focus on regional issues and less on international cooperation. This fragmented world implies high inequalities within and between countries. Inequality and poverty is also the central theme of SSP4, which corresponds to a multi-level society with a lot of issues (social cohesion, access to water, sanitation and health ware, education, etc). SSP5 corresponds to a high resilience of the economy, and consequently nothing is done to reduce inequalities and negative externalities. We observe that the SSPs can be related to ESG dimensions:

(E) The mitigation/adaptation trade-off is obviously an environmental issue, but the SSPs encompass other environmental narratives, *e.g.*, land use, energy efficiency and green economy;

(S) The social dimension is the central theme of SSPs, and concerns demography, wealth, inequality & poverty, health, education, employment, and more generally the evolution of society. This explains that SSPs and SDGs are highly interconnected;

(G) Finally, the governance dimension is present though two major themes: international fragmentation or cooperation, and the political/economic system, including corruption, stability, rule of law, etc.

We notice that all these SSP themes can be found in a sovereign ESG framework[98].

While O'Neill *et al.* (2014) have introduced the concept of SSP and given a very broad definition, the analysis proposed by O'Neill *et al.* (2017) is purely qualitative and descriptive. In fact, the first broad quantification of SSPs is done by Riahi *et al.* (2017), who have generated the five narrative scenarios with the following IAMs:

- SSP1: IMAGE (PBL)

- SSP2: MESSAGE-GLOBIOM (IIASA)

- SSP3: AIM/CGE (NIES)

- SSP4: GCAM (PNNL)

- SSP5: REMIND-MAGPIE (PIK) and WITCH-GLOBIOM (FEEM)

Results obtained by these authors are summarized in Figures 8.107–8.110. We only focus on baseline scenarios and we group the indicators into four main categories. In Figure 8.107, the demography projection shows three main pathways: population decline (SSP1, SSP5), population stabilization (SSP2, SSP4) and population growth (SSP5). If we consider the economic growth projections, we obtain another kind of breakdown (Figure 8.108). SSP5 corresponds to the highest GDP projections. It is following by SSP1 and SSP2, while SSP4 and SSP3 are the two laggard scenarios. In the last panel, we have reported the GDP per primary energy ratio. It shows that SSP1 has the highest efficiency. The case of SSP3 is interesting since the improvement of the energy efficiency follows the economic growth beyond 2040.

Figures 8.109 and 8.110 give an overview of the environmental narratives. As expected, SSP1 and SSP5 are the two extreme scenarios if we focus on climate change (temperature and CO_2 emissions). If we consider the land use, the opposite scenario become SSP1 and SSP3. In SSP1, the forest increases while cropland and pasture decrease or are stable. In SSP3, the forest dramatically decreases and there is 50% more cropland by 2100. Another interesting SSP indicator is the urban share (first panel in Figure 8.110). The urbanization of the world continues, but at different paces. By the end of century, the projected rate of urbanization reaches 60% for SSP3, 80% for SSP3 and more than 90% for the other SSPs. O'Neill *et al.* (2014) explained that "*urbanization is constrained by slow economic growth and limited mobility across regions in SSP3, while urbanization is assumed to be rapid in both SSP1 and SSP5, which are associated with high income growth*".

To illustrate inequality issues, we can conduct the previous at a regional/country level. For example, Figure 8.111 shows demography and GDP/Capita growth projections in Africa and Asia. It is obvious that these two regions will not have the same economic development even if they follow the same shared socioeconomic pathway. This is particular true for the SSP4 scenario, for which we observe a population growth in Africa and a population decline in Asia. The impact of SSP4 on the GDP growth is also large in Africa with respect to the SSP2 scenario. The relationship between income inequalities and SSPs has been extensively studied by Rao *et al.* (2019). Using their dataset, we report in Figure 8.111 the boxplots of the Gini coefficient by 2100 for 184 countries, and compare these values with the 2020 estimates. We notice that income inequality within a country decreases for SSP1, SSP2 and SSP5 scenarios[99], while it increases for SSP3 and SSP4. These results illustrate that inequalities across/within countries are important to understand the social cohesion dimension.

---

[98]See Section 2.1.1 on page 48.

[99]Nevertheless, we also observe that there are more outlier countries for these scenarios.

Figure 8.107: SSP demography projections
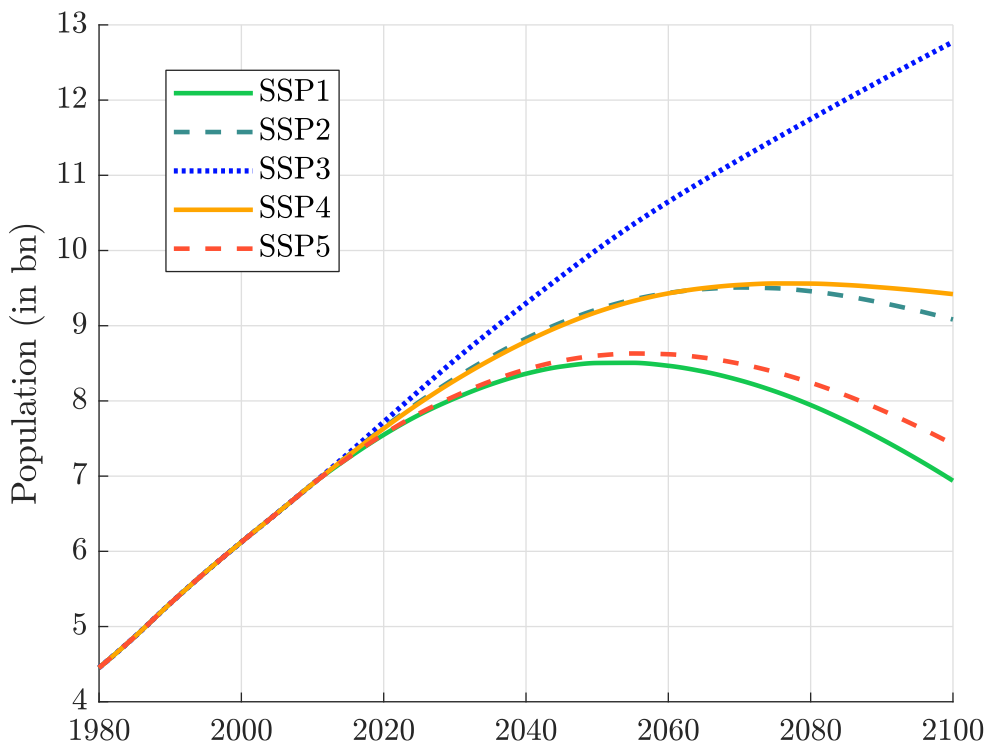
Figure 8.108: SSP economic projections
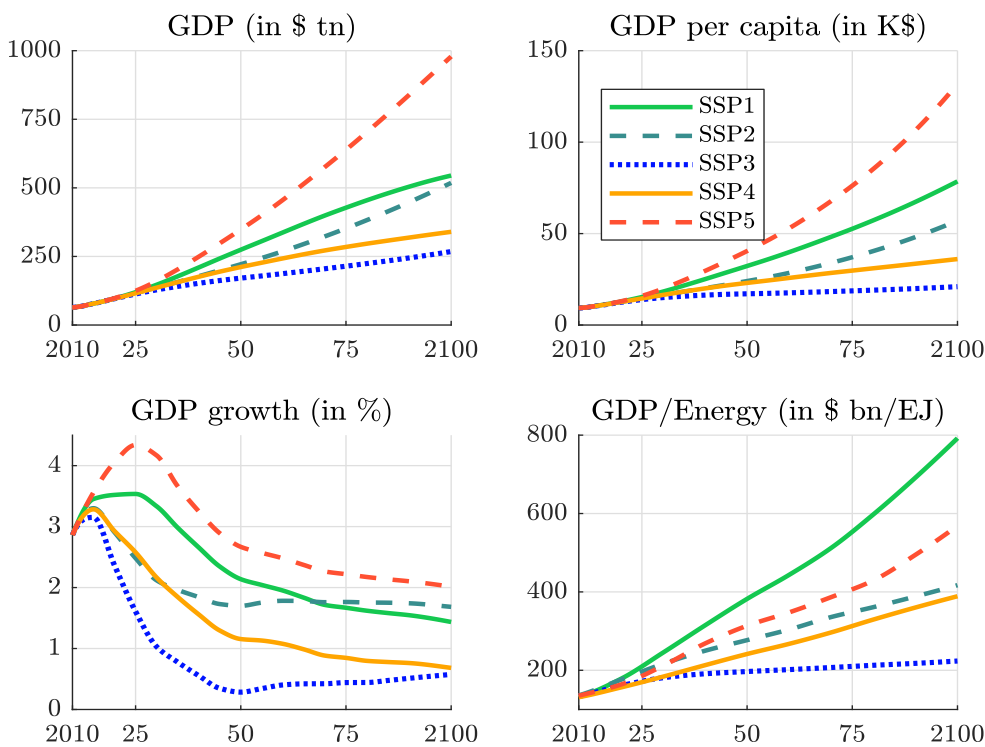
Figure 8.109: SSP environmental projections

Source: https://tntcat.iiasa.ac.at/SspDb.
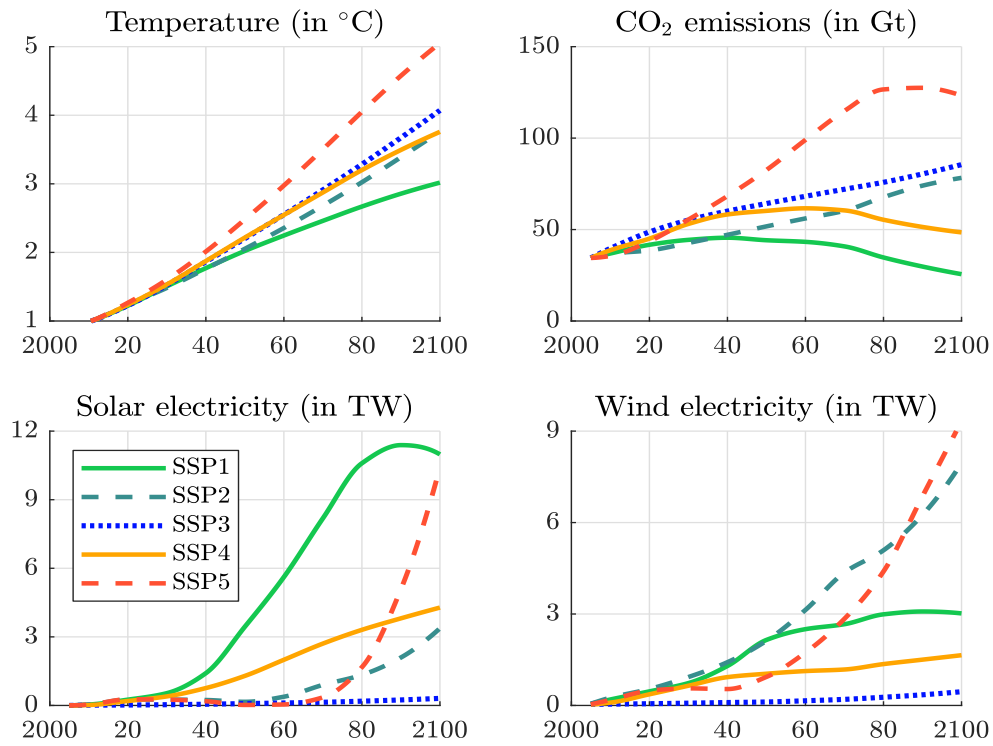
Figure 8.110: SSP land-use projections

Source: https://tntcat.iiasa.ac.at/SspDb.
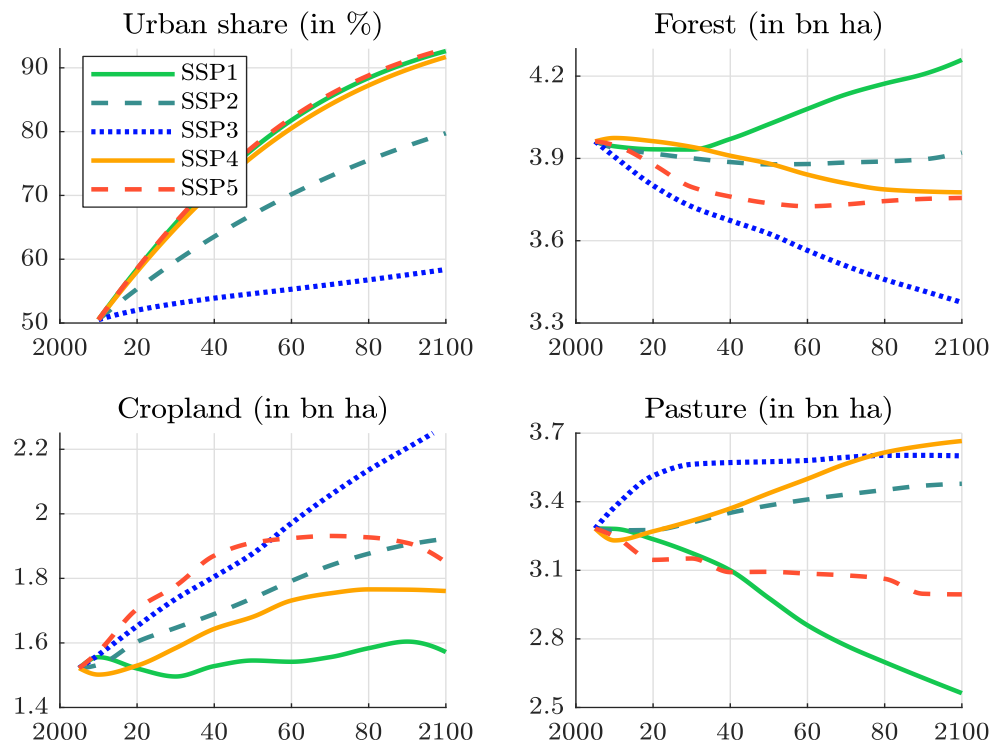
Figure 8.111: Example of SSP regional differences

Source: https://tntcat.iiasa.ac.at/SspDb.
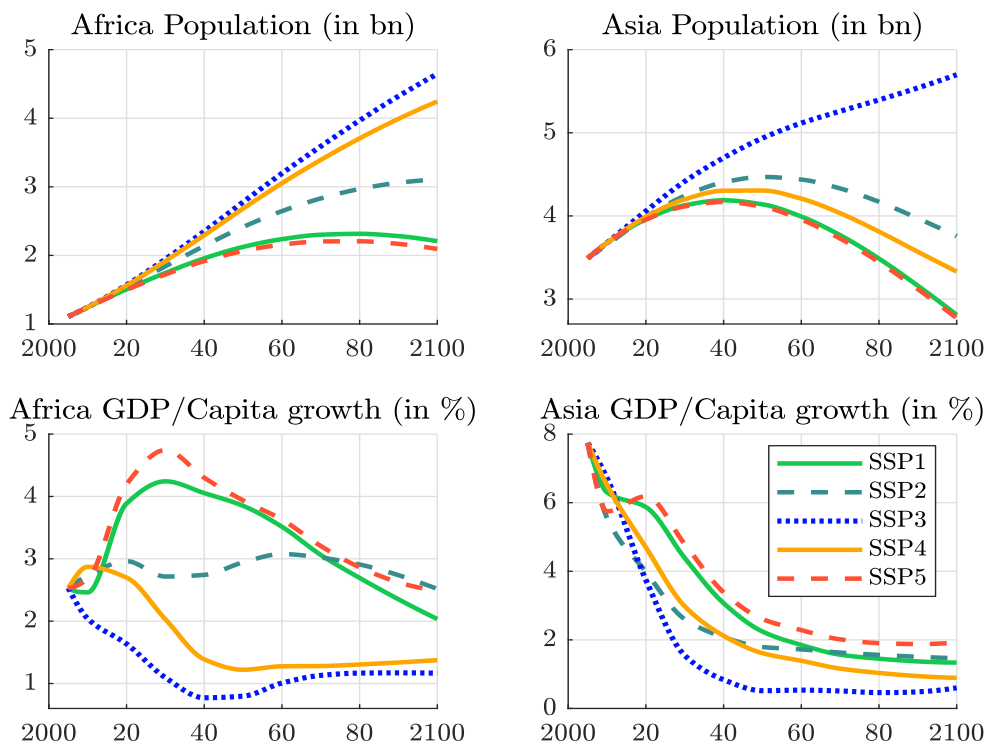
Figure 8.112: Gini coefficient projections by 2100

Source: https://tntcat.iiasa.ac.at/SspDb.
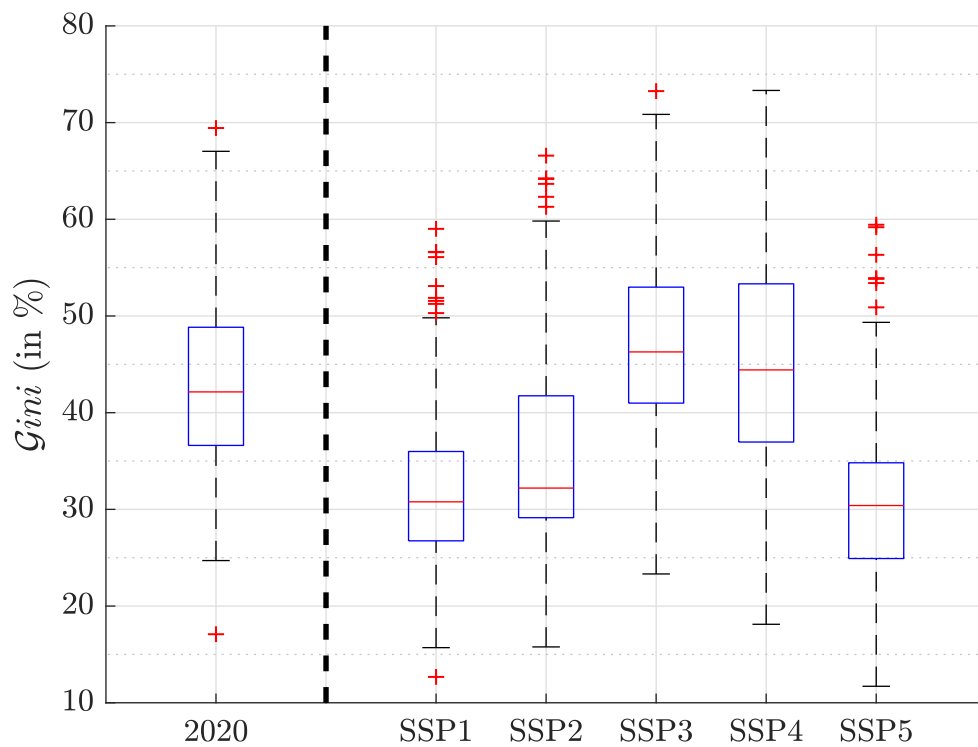
Macroeconomic scenarios

The main objective of IAMs is to produce macroeconomic scenarios. For instance, the stylized simple models generate population, capital stock and GDP pathways, even though they are generally used only to compute the social cost of carbon. Indeed, these models are too simple to describe realistic economic scenarios. This is why these latter are simulated using complex IAMs. In the previous section dedicated to SSPs, we have already seen some output variables that are generated. In what follows, we focus on the NGFS scenarios, which have recently become a common standard. Today, the NGFS framework is an official tool that is used and recognized by central banks and supervisors, and also a reference point for financial institutions (investors, asset managers, and banks).

NGFS scenarios framework   In 2020, the Network for Greening the Financial System (NGFS) has developed a set of scenarios for climate financial risk assessment[100]. The initial goal was to help central banks and supervisors exploring the possible impacts on the economy and the financial system. The NGFS scenarios framework consists in six scenarios located in the NGFS climate risk matrix given in Figure 8.113.

Figure 8.113: NGFS scenarios framework
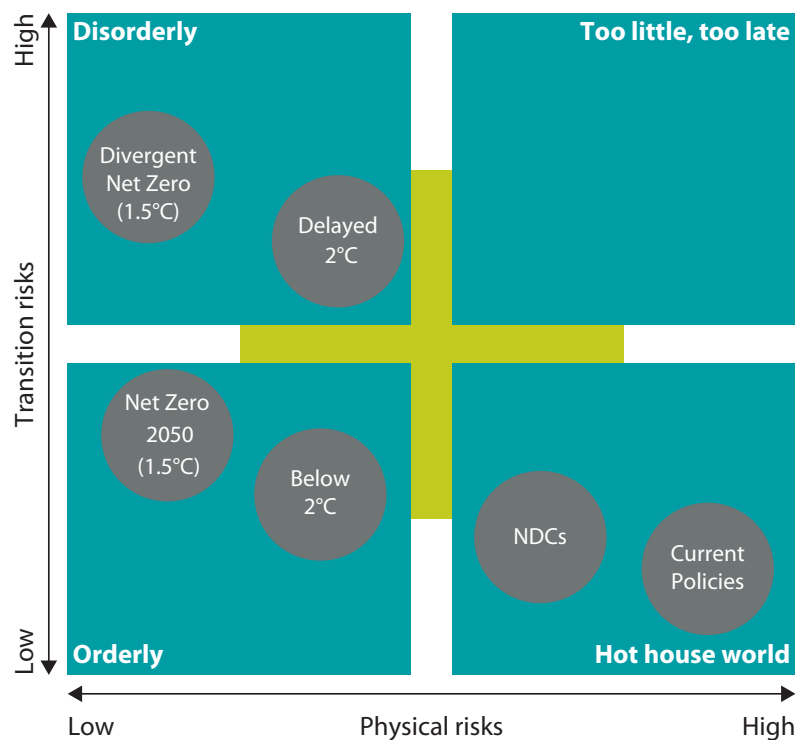
Below, we report the description of the six scenarios, which is done by NGFS (2022).

- Orderly scenarios assume climate policies are introduced early and become gradually more stringent. Both physical and transition risks are relatively subdued.

---

[100]The third version of these scenarios was released in September 2022.

#1 Net-zero 2050 limits global warming to 1.5°C through stringent climate policies and innovation, reaching global net zero CO_2 emissions around 2050. Some jurisdictions such as the US, EU, UK, Canada, Australia and Japan reach net zero for all GHGs.

#2 Below 2°C gradually increases the stringency of climate policies, giving a 67% chance of limiting global warming to below 2°C.

- Disorderly scenarios explore higher transition risk due to policies being delayed or divergent across countries and sectors. For example, carbon prices are typically higher for a given temperature outcome.

#3 Divergent net zero reaches net zero around 2050 but with higher costs due to divergent policies introduced across sectors leading to a quicker phase out of oil use.

#4 Delayed transition assumes annual emissions do not decrease until 2030. Strong policies are needed to limit warming to below 2°C. Negative emissions are limited.

- Hot house world scenarios assume that some climate policies are implemented in some jurisdictions, but globally efforts are insufficient to halt significant global warming. The scenarios result in severe physical risk including irreversible impacts like sea-level rise.

#5 Nationally determined contributions (NDCs) includes all pledged targets even if not yet backed up by implemented effective policies.

#6 Current policies assumes that only currently implemented policies are preserved, leading to high physical risks.

In what follows, we use the following acronyms NZ, B2D, DNZ, D2D, NDC, CP when we refer to these 6 scenarios. Table 8.19 shows the impact of climate risk on each scenario. For the physical risk, NGFS uses both chronic and acute impacts. Chronic risks[101] affect agriculture, labor productivity and natural capital. Acute risks or extreme weather events[102] implies destruction, economic losses, lower insurance cover and business disruption. Transition risk takes several forms: policy risk, technology risk and coordination risk. The climate transition pathways are generated by three IAMs: GCAM, MESSAGEix-GLOBIOM and REMIND-MAgPIE, while the economic variables are produced by the NiGEM macroeconomic model[103] in order to obtain harmonized metrics between scenarios and climate models. Moreover, a seventh scenario is produced by NiGEM without the IAM inputs and corresponds to the baseline (BSL) scenario, which does not take into account climate risk.

Description of the database   The NGFS database[104] is entirely described in Richters *et al.* (2022). It contains three EXCEL files. `IAM_data` contains the data generated by the three IAM models, `Downscaled_data` are the energy data produced by the previous models at the country level, while `NiGEM_data` corresponds to the data generated by the NiGEM model. In what follows, we only consider this last file. In order to select data, we first must specify the model and the scenario. The baseline scenario is obtained with the NiGEM model. For the six climate scenarios, we must choose one integrated assessment model (GCAM, MESSAGEix-GLOBIOM index[general]MESSAGE modelindex[general]GLOBIOM model or REMID-MAgPIE) and associate

---

[101]Such as increased temperatures, rising sea levels, ocean acidification and seasonal precipitation.

[102]Such as such as tropical cyclones, droughts, floods and wildfires.

[103]It is developed by the National Institute for Economic and Social Research (NIESR). A description can be found at www.niesr.ac.uk/nigem-macroeconomic-model.

[104]The economic scenarios can be downloaded at https://data.ene.iiasa.ac.at/ngfs.

Table 8.19: Physical and transition risk level of NGFS scenarios

Category	Scenario	Physical risk	Transition risk			
		Policy ambition	Policy reaction	Technology change	Carbon dioxide removal⁻	Regional policy variation⁺
Orderly	Net Zero 2050	1.4°C	Immediate and smooth	Fast change	Medium-high use	Medium variation
	Below 2°C	1.6°C	Immediate and smooth	Moderate change	Medium-high use	Low variation
Disorderly	Divergent Net Zero	1.4°C	Immediate but divergent across sectors	Fast change	Low-medium use	Medium variation
	Delayed Transition	1.6 °C	Delayed	Slow / Fast change	Low-medium use	High variation
Hot house world	Nationally Determined Contributions (NDCs)	2.6°C	NDCs	Slow change	Low-medium use	Medium variation
	Current Policies	3°C +	Non-currente policies	Slow change	Low use	Low variation

Lower risk, moderate risk and higher risk

Color coding indicates whether the characteristic makes the scenario more or less severe from a macro-financial risk perspective.

Source: NGFS (2022) & www.ngfs.net.

Table 8.20: NGFS database — Economic block

Scenarios

1. Net Zero 2050 (NZ)
2. Below 2°C (B2D)
3. Divergent Net Zero (DNZ)
4. Delayed Transition (DT)
5. Notionally Determined Contribution (NDC)
6. Current Policies (CP)
7. Baseline (BSL)

Models

- GCAM 5.3+ NGFS
- MESSAGEix-GLOBIOM 1.1-M-R12
- NiGEM NGFS v1.22
- REMIND-MAgPIE 3.0-4.4
- NiGEM NGFS v1.22 / Downscaling
  - GCAM 5.3+ NGFS
  - MESSAGEix-GLOBIOM 1.1-M-R12
  - REMIND-MAgPIE 3.0-4.4
- REMIND-MAgPIE 3.0-4.4
  - IntegratedPhysicalDamages (95th-high)
  - IntegratedPhysicalDamages (median)

Energy variables

- Coal price
- Gas price
- Oil price
- Quarterly consumption of coal
- Quarterly consumption of gas
- Quarterly consumption of non-carbon
- Quarterly consumption of oil
- Total energy consumption

Economic variables

- Central bank intervention rate
- Consumption (private sector)
- Domestic demand
- Effective exchange rate
- Equity prices
- Exchange rate
- Exports (goods and services excluding MTIC)
- Exports (goods and services)
- Government consumption
- Gross Domestic Product (GDP)
- Gross domestic income
- Gross operating surplus
- House prices (residential)
- Imports (goods and services excluding MTIC)
- Imports (goods and services)
- Inflation rate
- Investment (government)
- Investment (private sector)
- Long term interest rate
- Productivity (output per hour worked)
- Real personal disposable income
- Trend output for capacity utilisation
- Unemployment rate

it with the NiGEM model. Once the model and the scenario are chosen, we can filter the data by region and/or variable. We have reported the name of the variables in Table 8.20. They are of two types. Energy variables concern price and consumption of coal, gas, oil and non-carbon (or renewables/green energy). Economic variables cover production, consumption, income, investment, interest rates, inflation and unemployment.

Empirical results   We report the impact of NGFS scenarios on GDP in Tables 8.21, 8.22 and 8.23 when the time horizon is set to 2050 and the region corresponds to the world. The results depend on the selected model. For instance, if we consider the B2D scenario and the chronic physical risk, the GDP impact is equal to -3.09% with GCAM, -2.05% with MESSAGEix and -2.24% with REMIND. For each table, the impact is related to the selected scenario and the selected risk. We notice that the largest loss is obtained with the combined risk (physical and transition), but there is no individual risk that dominates the other. For example, the chronic physical risk has more impact for the B2D, DT, NDC and NZ scenarios while the transition risk dominates for the DNZ scenario. We also observe that the impacts of physical and transition risks are almost additive when they are combined. Another interesting remark concerns the magnitude of the GDP loss. If we consider GCAM and REMIND, there is no obvious ranking in terms of scenario severity. This is not the case of the MESSAGEix model, which assumes that DNZ and DT are the two most severe scenarios, followed by four other scenarios.

Table 8.21: Impact of climate change on the GDP loss by 2050 (GCAM)

Risk	B2D	CP	DNZ	DT	NDC	NZ
Chronic physical risk	−3.09	−5.64	−2.35	−3.28	−5.15	−2.56
Transition risk	−0.75		−3.66	−1.78	−0.89	−0.88
Combined risk	−3.84	−5.64	−6.00	−5.05	−6.03	−3.44
Combined + business confidence			−6.03	−5.09		

Source: https://data.ene.iiasa.ac.at/ngfs.

Table 8.22: Impact of climate change on the GDP loss by 2050 (MESSAGEix-GLOBIOM)

Risk	B2D	CP	DNZ	DT	NDC	NZ
Chronic physical risk	−2.05	−5.26	−1.55	−2.64	−4.78	−1.59
Transition risk	−1.46		−10.00	−10.77	−1.39	−3.26
Combined risk	−3.51	−5.26	−11.53	−13.37	−6.16	−4.84
Combined + business confidence			−11.57	−13.40		

Source: https://data.ene.iiasa.ac.at/ngfs.

Table 8.23: Impact of climate change on the GDP loss by 2050 (REMIND-MAgPIE)

Risk	B2D	CP	DNZ	DT	NDC	NZ
Chronic physical risk	−2.24	−6.05	−1.67	−2.65	−5.41	−1.76
Transition risk	−0.78		−3.01	−1.95	−0.33	−1.46
Combined risk	−3.02	−6.05	−4.68	−4.59	−5.73	−3.21
Combined + business confidence			−4.70	−4.63		

Source: https://data.ene.iiasa.ac.at/ngfs.

By using the MESSAGEix-GLOBIOM model and selecting the combined risk, we obtain the results per country and region in Table 8.24. On average, the GDP loss by 2050 is between −3.51% (current policies) and −13.37% (delayed transition). This large difference is also observed when we compare countries and region. For example, if we focus on the NZ scenario, the impact is negative and equal to −17.11% for Russia while it is positive for developing Europe, South Korea and Switzerland. Therefore, we conclude that there is a high heterogeneity between countries (see also Figures 8.114 and 8.115).

Table 8.24: Impact of climate change on the GDP loss by 2050 (MESSAGEix-GLOBIOM)

Region	B2D	CP	DNZ	DT	NDC	NZ
Africa	−13.58	−7.50	−27.35	−29.37	−11.78	−18.36
Asia	−1.50	−7.29	−5.44	−8.76	−6.78	−1.38
Australia	−4.11	−3.90	−11.03	−11.74	−5.77	−5.19
Brazil	−4.43	−5.92	−13.15	−15.90	−6.67	−6.65
Canada	−1.02	−2.37	−15.07	−18.12	−4.33	−4.87
China	−2.33	−4.97	−5.13	−6.73	−4.67	−2.76
Developing Europe	−0.28	−3.11	−0.56	−7.38	−2.73	0.39
Europe	−1.02	−2.84	−9.64	−11.02	−4.01	−1.62
France	−1.15	−2.80	−8.35	−9.48	−3.68	−1.56
Germany	−0.77	−2.38	−8.58	−9.38	−3.63	−1.21
India	−3.45	−8.61	−16.43	−17.74	−8.71	−3.86
Italy	−0.15	−3.69	−9.23	−12.88	−4.85	−0.89
Japan	−1.26	−4.14	−7.16	−10.05	−4.61	−1.40
Latam	−4.35	−6.10	−12.70	−14.58	−6.97	−5.74
Middle East	−9.97	−7.98	−22.03	−21.96	−10.28	−15.24
Russia	−12.18	−2.26	−23.46	−23.80	−7.54	−17.11
South Africa	−2.02	−5.06	−7.24	−9.16	−5.38	−3.04
South Korea	0.11	−3.49	−3.23	−7.57	−3.33	0.12
Spain	−2.41	−3.81	−12.49	−12.89	−5.41	−3.30
Switzerland	2.32	−2.25	−9.47	−10.35	−2.18	2.30
United Kingdom	−0.86	−1.90	−6.50	−8.05	−2.56	−1.33
United States	−2.67	−4.38	−15.37	−17.66	−6.31	−4.36
World	−3.51	−5.26	−11.53	−13.37	−6.16	−4.84

Source: https://data.ene.iiasa.ac.at/ngfs.

The previous analysis can be extended to other economic variables. In Figure 8.116, we show the impact on inflation, unemployment, private investment, productivity, equity prices and public investment in the case of China. With respect to the baseline scenario, the delayed transition may induce more inflation (up to 6%) and unemployment (up to 0.8%) between 2030 and 2040, and less public and private investments. It may also strongly impact the equity market and the financial performance of stocks. Finally, the economy may be less productive. We have reported a similar analysis in the case of United States, France and United Kingdom in Figures 8.117, 8.118 and 8.119. Again, we notice that the impact on the economy depends on the country or region. Even if there is no winner for some scenarios, some countries will loose less than others. Climate change will then create new inequalities between regions. In particular, two regions seems to be especially vulnerable: Africa and the Middle East.

Figure 8.114: GDP impact by 2050 (% change from baseline) — Delayed transition scenario
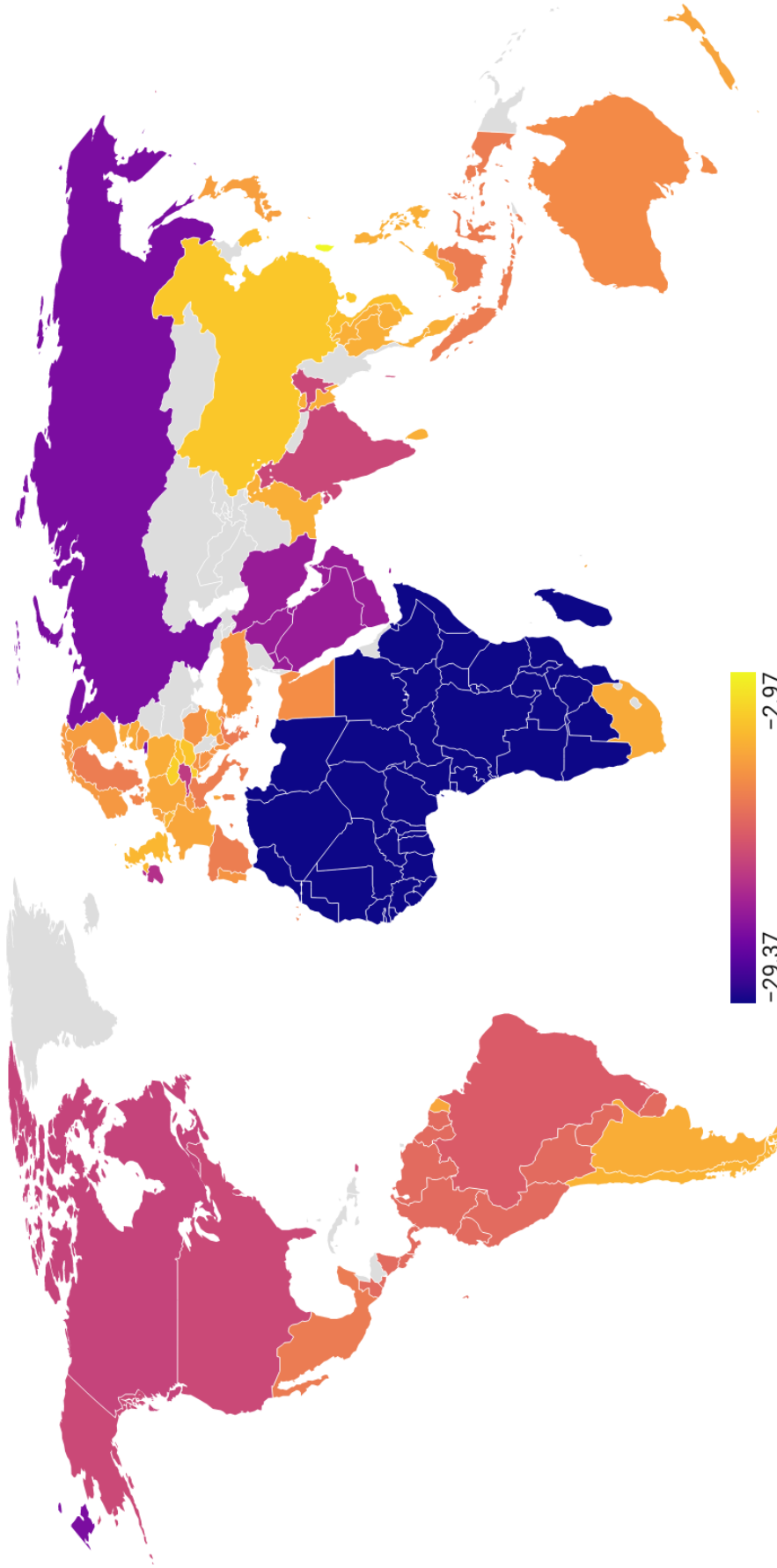
Source: https://data.ene.iiasa.ac.at/ngfs & Author's calculations (created by Datawrapper).
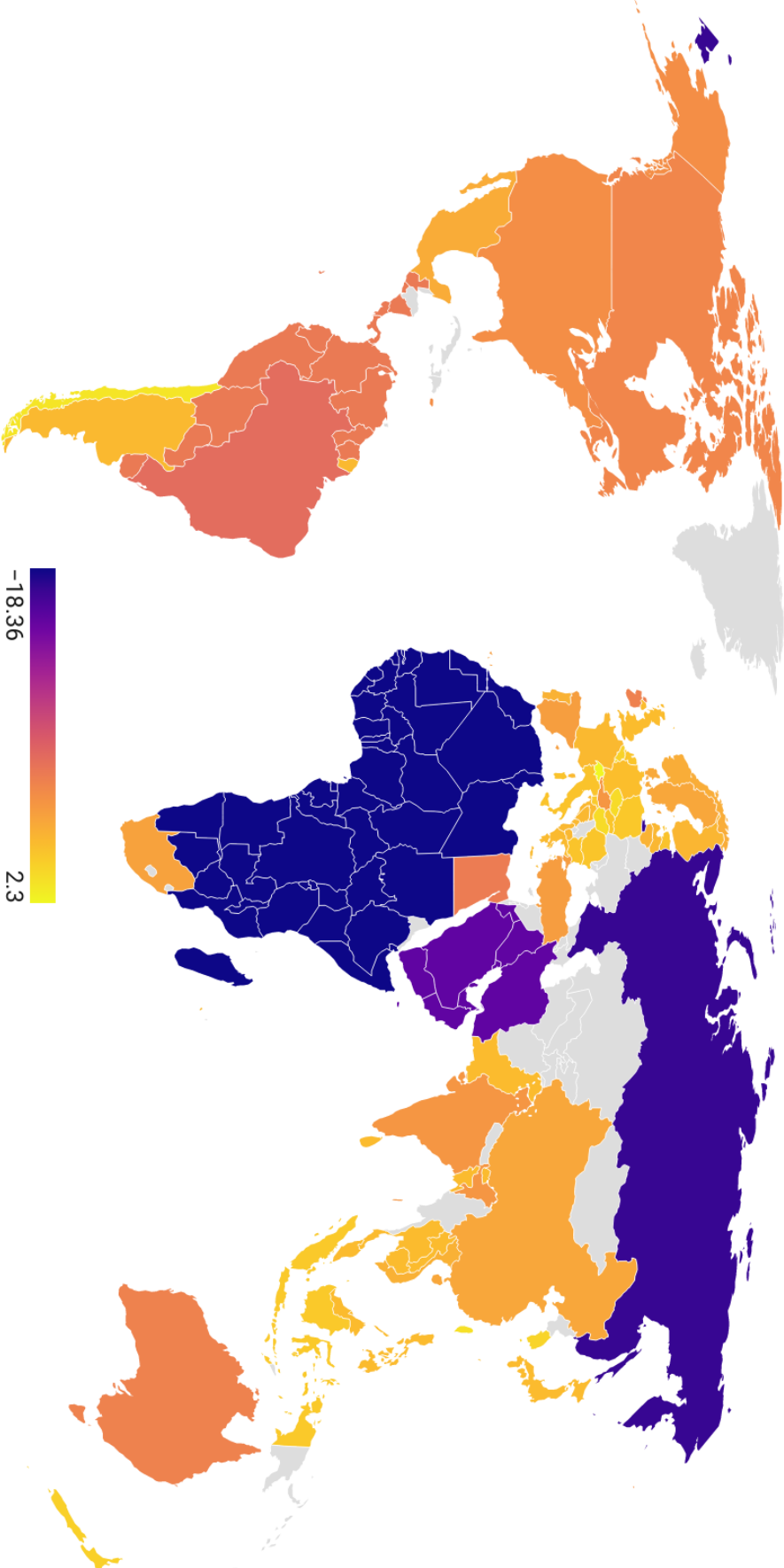
Figure 8.115: GDP impact by 2050 (% change from baseline) — Net-zero 2050 scenario

Source: https://data.ene.iiasa.ac.at/ngfs & Author's calculations (created by Datawrapper).

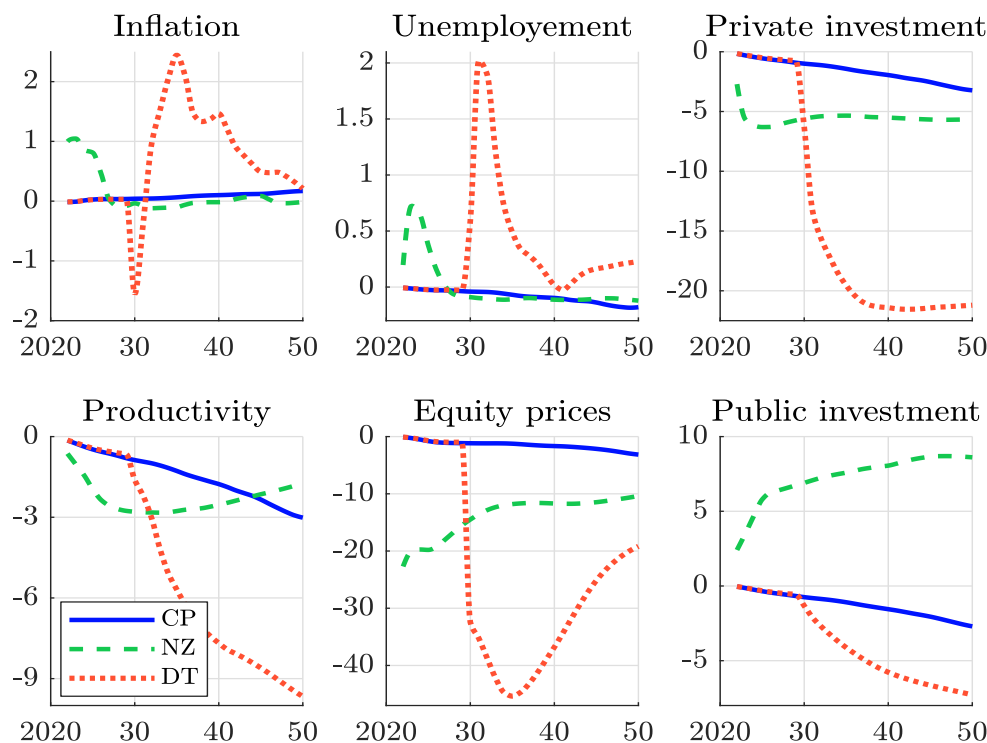Figure 8.116: Impact of climate scenarios on economics (% change from baseline) — China
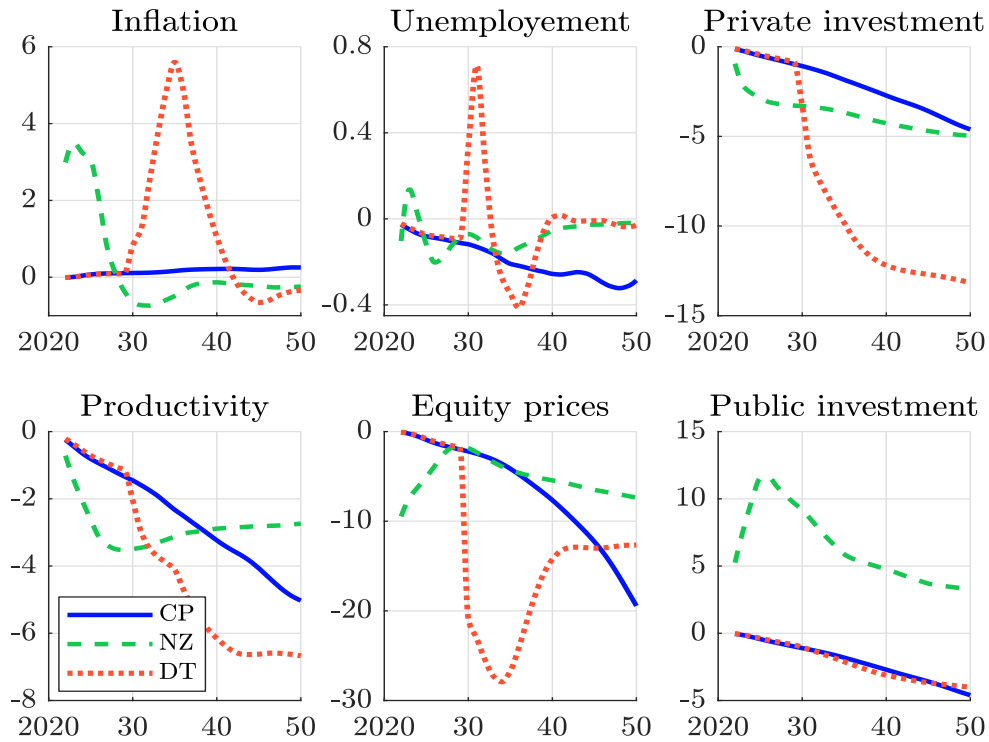
Source: https://data.ene.iiasa.ac.at/ngfs & Author's calculations.

Figure 8.117: Impact of climate scenarios on economics (% change from baseline) — United States
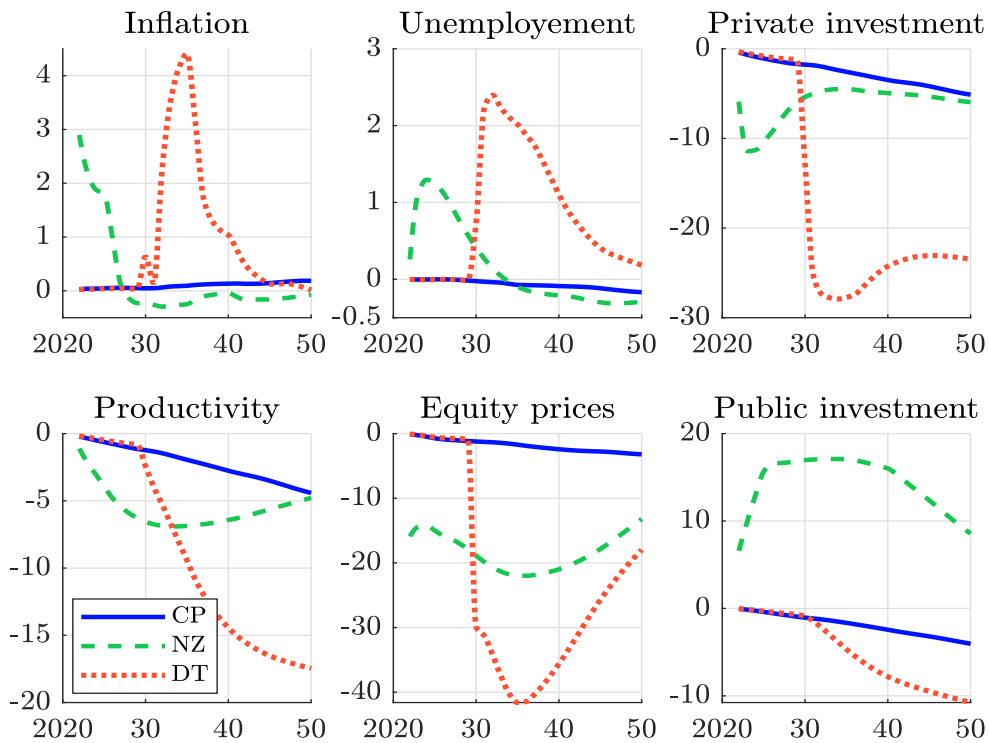
Source: https://data.ene.iiasa.ac.at/ngfs & Author's calculations.

Figure 8.118: Impact of climate scenarios on economics (% change from baseline) — France

Source: https://data.ene.iiasa.ac.at/ngfs & Author's calculations.

Figure 8.119: Impact of climate scenarios on economics (% change from baseline) — United Kingdom
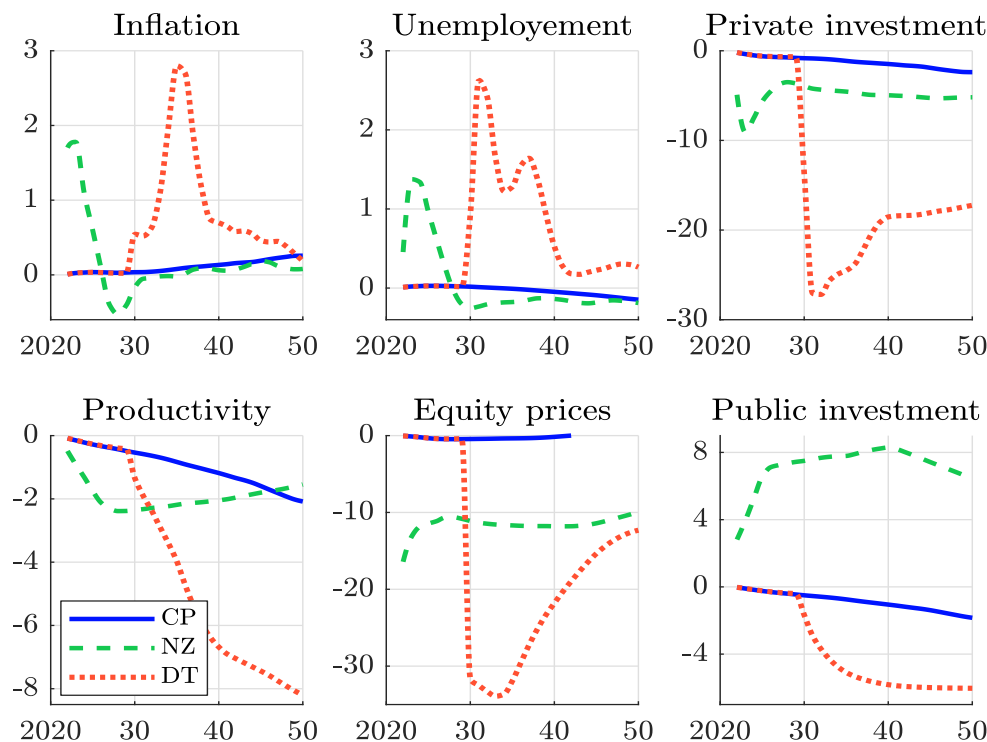
Source: https://data.ene.iiasa.ac.at/ngfs & Author's calculations.

8.4 Environmentally-extended input-output model

An input-output model is a mathematical tool that represents the macroeconomic relationships between different units or industries. It can be used to model the supply chain of a product, the sectoral structure of an economy, the production network of a country or the foreign exchange between regions. Most of these models are monetary in nature. When applied to economic transactions between different sectors, they can be used to calculate the contribution or value added of each sector to the final output of an economy. Environmentally-extended input-output (EEIO) analysis is an extension of the input-output framework to include environmental externalities such as pollution and greenhouse gas emissions. In particular, we can use EEIO models to estimate upstream Scope 3 emissions, define a stress testing framework or calculate the value-at-risk of a portfolio (Desnos *et al.*, 2023).

8.4.1 Input-output analysis

The input-output model was first introduced by Leontief (1936, 1941). It quantifies the interdependencies between different sectors in a single or multi-regional economy, based on the product flows between sectors (Miller and Blair, 2009). The underlying idea is to model the linkages between sectors and to describe the relationships from each of the producer/seller sectors to each of the purchaser/buyer sectors.

The demand-pull quantity model

Following Miller and Blair (2009), we consider n different sectors and we note $Z_{i,j}$ the value of transactions from sector i to sector j. We can interpret $Z_{i,j}$ in different ways:

1. It is the output that sector i sells to sector j;

2. It is the input of sector i required by sector j for its production (or output).

Let y_i be the final demand for products sold by sector i. This final demand is made up of the external sales to households, government purchases, and demand resulting from investment capacity and foreign trade. Then, the total production x_i of sector i is equal to:

$$\underbrace{x_i}_{\text{Supply}} = \underbrace{\sum\nolimits_{j=1}^n Z_{i,j} + y_i}_{\text{Demand}} \tag{8.28}$$

In this equation, x_i and $\sum_{j=1}^n Z_{i,j} + y_i$ are the supply and demand related to products of sector i, and $z_i = \sum_{j=1}^n Z_{i,j}$ represents intermediate demand. The interdependence relation between sectors is usually expressed as a ratio between $Z_{i,j}$ and x_j:

$$A_{i,j} = \frac{Z_{i,j}}{x_j}$$

Let $A = (A_{i,j}) = Z \operatorname{diag}(x)^{-1}$ be the input-output matrix of the technical coefficients $A_{i,j}$. In a matrix form, we have $x = Z\mathbf{1}_n + y$ and $Z \equiv A \operatorname{diag}(x) = A \odot x^\top$, and we deduce that:

$$x = Ax + y$$

where $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$. Assuming that final demand is exogenous, technical coefficients are fixed and output is endogenous, we obtain:

$$x = (I_n - A)^{-1} y \tag{8.29}$$

$\mathcal{L} = (I_n - A)^{-1}$ is known as the Leontief inverse (or multiplier) matrix and represents the amount of total output from sector i that is required by sector j to satisfy its final demand. Equation (8.29) describes a *demand-pull quantity* model.

Example 21 *We consider the basic economy given below:*

		To				Final	Total
		Energy	Materials	Industrials	Services	Demand	Output
			Z			y	x
From	Energy	500	800	1 600	1 250	850	5 000
	Materials	500	400	1 600	625	875	4 000
	Industrials	250	800	2 400	1 250	3 300	8 000
	Services	100	200	800	4 375	7 025	12 500

This basic economy has four sectors: energy, materials, industrials and services. In this economy, businesses in the energy sector buy $500 of goods and services from other businesses in the energy sector, $500 of goods and services from the materials sector, $250 of goods and services from the industrials sector, and $100 of goods and services from the services sector. The final demand for goods and services produced in the energy sector is equal to $850, while the total output of this sector is equal to $5 000.

We deduce that the matrix of technical coefficients is equal to:

$$A = Z \operatorname{diag}(x)^{-1} = \begin{pmatrix} 10\% & 20\% & 20\% & 10\% \\ 10\% & 10\% & 20\% & 5\% \\ 5\% & 20\% & 30\% & 10\% \\ 2\% & 5\% & 10\% & 35\% \end{pmatrix}$$

It follows that the multiplier matrix is equal to:

$$\mathcal{L} = (I_4 - A)^{-1} = \begin{pmatrix} 1.1881 & 0.3894 & 0.4919 & 0.2884 \\ 0.1678 & 1.2552 & 0.4336 & 0.1891 \\ 0.1430 & 0.4110 & 1.6303 & 0.3044 \\ 0.0715 & 0.1718 & 0.2993 & 1.6087 \end{pmatrix}$$

We verify that:

$$x = \mathcal{L}y = \begin{pmatrix} 1.1881 & 0.3894 & 0.4919 & 0.2884 \\ 0.1678 & 1.2552 & 0.4336 & 0.1891 \\ 0.1430 & 0.4110 & 1.6303 & 0.3044 \\ 0.0715 & 0.1718 & 0.2993 & 1.6087 \end{pmatrix} \begin{pmatrix} 850 \\ 875 \\ 3 300 \\ 7 025 \end{pmatrix} = \begin{pmatrix} 5 000 \\ 4 000 \\ 8 000 \\ 12 500 \end{pmatrix}$$

Suppose we have a variation in final demand. From Equation (8.29), we obtain $\Delta x = \mathcal{L}\Delta y$. For instance, an increase of $10 in the final demand for services implies:

$$\Delta x = \mathcal{L}\Delta y = \begin{pmatrix} 1.1881 & 0.3894 & 0.4919 & 0.2884 \\ 0.1678 & 1.2552 & 0.4336 & 0.1891 \\ 0.1430 & 0.4110 & 1.6303 & 0.3044 \\ 0.0715 & 0.1718 & 0.2993 & 1.6087 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 10 \end{pmatrix} = \begin{pmatrix} 2.8842 \\ 1.8907 \\ 3.0444 \\ 16.0872 \end{pmatrix}$$

This means that energy production increases by $2.88, materials production increases by $1.89, and so on.

The cost-push price model

Let m be number of primary inputs (*e.g.*, labor, capital, etc.). Let $V = (V_{k,j})$ be the value added matrix where $V_{k,j}$ represents the amount of primary input k required to produce the output of sector j. Since the total input of each sector is equal to its total output, we have $x_j = \sum_{i=1}^{n} Z_{i,j} + \sum_{k=1}^{m} V_{k,j}$. Therefore, $v_j = \sum_{k=1}^{m} V_{k,j} = x_j - \sum_{i=1}^{n} Z_{i,j}$ represents the other expenditure of sector j or the total primary inputs used in sector j. We have $v = (v_1, \ldots, v_n) = V^\top \mathbf{1}_m$. Let $p = (p_1, \ldots, p_n)$ and $\psi = (\psi_1, \ldots, \psi_m)$ be the vector of sector prices and primary inputs. p_j and ψ_k are then the prices per unit of sector j and primary input k. As in the quantity model, the interdependence relationship between primary inputs and sectors is expressed as the ratio between $V_{k,j}$ and x_j:

$$B_{k,j} = \frac{V_{k,j}}{x_j}$$

We denote the input-output matrix of the technical coefficients by $B = (B_{k,j}) \equiv V \operatorname{diag}(x)^{-1}$. Following Gutierrez (2008), the value of the output must be equal to the value of its inputs:

$$\underbrace{p_j x_j}_{\text{Value of the output}} = \underbrace{\sum_{i=1}^{n} Z_{i,j} p_i + \sum_{k=1}^{m} V_{k,j} \psi_k}_{\text{Value of the inputs}}$$

We deduce that:

$$
\begin{aligned}
p_j &= \sum_{i=1}^{n} \frac{Z_{i,j}}{x_j} p_i + \sum_{k=1}^{m} \frac{V_{k,j}}{x_j} \psi_k \\
&= \sum_{i=1}^{n} A_{i,j} p_i + \sum_{k=1}^{m} B_{k,j} \psi_k
\end{aligned}
$$

In a matrix form, we get $p = A^\top p + B^\top \psi$. $v = B^\top \psi$ is the vector of value added ratios. Finally, the output prices are equal to:

$$p = \left(I_n - A^\top\right)^{-1} v \tag{8.30}$$

$\tilde{\mathcal{L}} = \left(I_n - A^\top\right)^{-1}$ is known as the dual inverse matrix and represents the amount of costs from sector j that are passed on to sector i. Equation (8.30) describes a *cost-push price* model. By adding the income identity[105], Gutierrez (2008) proposed the following complete version of the full basic input-output model:

$$
\begin{cases}
x = (I_n - A)^{-1} y \\
v = V^\top \mathbf{1}_m \\
v = B^\top \psi \\
p = \left(I_n - A^\top\right)^{-1} v \\
x^\top v = y^\top p
\end{cases}
\tag{8.31}
$$

It mixes both the quantity and price models. In this system, A, B and V are the model parameters, ψ, v and y are the exogenous variables, and x and p are the endogenous variables. By changing the model parameters or the exogenous variables, we can measure the impacts Δy and Δv on the quantities and prices in the economy.

Remark 59 *The previous analysis was derived for physical input-output tables, where flows are expressed in product units. However, the analysis remains valid when monetary input-output tables*

---

[105]Since the input-output analysis assumes an equilibrium model, the total value of the revenues $y^\top p$ is equal to the total value of costs $x^\top v$.

are considered. The only difference is the calculation of the primary cost vector. In a monetary input-output analysis, ψ is set by construction to $\mathbf{1}_m$, which means that $\upsilon = B^{\top}\mathbf{1}_m = (\upsilon_1/x_1, \ldots, \upsilon_n/x_n)$ and $p = \mathbf{1}_n$ (Miller and Blair, 2009, Section 2.6.3, pages 43-44).

Example 22 *We consider the previous basic economy. We assume that the value added is made up of two items: labour and capital. We have:*

		To				Final	Total
		Energy	Materials	Industrials	Services	Demand	Output
				Z		y	x
	Energy	500	800	1 600	1 250	850	5 000
From	Materials	500	400	1 600	625	875	4 000
	Industrials	250	800	2 400	1 250	3 300	8 000
	Services	100	200	800	4 375	7 025	12 500
Value	Labour	3 000	800	1 000	3 000		
added	Capital	650	1 000	600	2 000		
	Income	5 000	4 000	8 000	12 000		

The energy sector has a labour consumption of \$3 000 and a total output of \$5 000. By construction, the income of the sector is equal to the output of the sector. We deduce that the capital item (capital interest and net profit) is equal to \$650.

We have:

$$V = \begin{pmatrix} 3\,000 & 800 & 1\,000 & 3\,000 \\ 650 & 1\,000 & 600 & 2\,000 \end{pmatrix}$$

and:

$$\upsilon = V^{\top}\mathbf{1}_2 = x - Z^{\top}\mathbf{1}_4 = \begin{pmatrix} 3\,650 \\ 1\,800 \\ 1\,600 \\ 5\,000 \end{pmatrix}$$

We deduce that:

$$B = \begin{pmatrix} 0.60 & 0.20 & 0.125 & 0.24 \\ 0.13 & 0.25 & 0.075 & 0.16 \end{pmatrix}$$

Since we have a monetary input-output table, the labour and capital costs are equal to the monetary unit, *i.e.* one dollar ($\psi_1 = \psi_2 = 1$). It follows that:

$$\upsilon = B^{\top}\mathbf{1}_2 = \begin{pmatrix} 0.73 \\ 0.45 \\ 0.20 \\ 0.40 \end{pmatrix}$$

The interpretation is as follows. For the energy sector, intermediate consumption is 27% and value added is 73%. For the other three sectors, the value added ratios are 45%, 20% and 40% respectively. Finally, we obtain:

$$\tilde{\mathcal{L}} = \left(I_n - A^{\top}\right)^{-1} = \begin{pmatrix} 1.1881 & 0.1678 & 0.1430 & 0.0715 \\ 0.3894 & 1.2552 & 0.4110 & 0.1718 \\ 0.4919 & 0.4336 & 1.6303 & 0.2993 \\ 0.2884 & 0.1891 & 0.3044 & 1.6087 \end{pmatrix}$$

and:

$$p = \tilde{\boldsymbol{\mathcal{L}}} v = \mathbf{1}_4$$

We check that the prices in a monetary input-output table are normalized to one dollar. In this basic economy, the total final demand $y^\top p$ is equal to \$12050, which is equal to the total value added $x^\top v$. Suppose we have a variation in the labour/capital costs. From Equation (8.31), we obtain $\Delta p = \tilde{\boldsymbol{\mathcal{L}}} \Delta v$. For example, a 10% increase in costs in the energy sector means that the price of energy increases by 11.88%, the price of materials by 3.89%, and so on:

$$\Delta p = \tilde{\boldsymbol{\mathcal{L}}} \Delta v = \begin{pmatrix} 1.1881 & 0.1678 & 0.1430 & 0.0715 \\ 0.3894 & 1.2552 & 0.4110 & 0.1718 \\ 0.4919 & 0.4336 & 1.6303 & 0.2993 \\ 0.2884 & 0.1891 & 0.3044 & 1.6087 \end{pmatrix} \begin{pmatrix} 0.10 \\ 0.00 \\ 0.00 \\ 0.00 \end{pmatrix} = \begin{pmatrix} 11.88\% \\ 3.89\% \\ 4.92\% \\ 2.88\% \end{pmatrix}$$

The definition of a price index is:

$$\mathcal{PI} = \sum_{j=1}^n \alpha_j p_j = \alpha^\top p$$

where $\alpha = (\alpha_1, \ldots, \alpha_n)$ is the weights of the basket of items. We deduce that the inflation rate between two dates t_0 and t_1 is:

$$\pi = \frac{\mathcal{PI}(t_1) - \mathcal{PI}(t_0)}{\mathcal{PI}(t_0)} = \frac{\alpha^\top \left(I_n - A^\top\right)^{-1} \Delta v(t_0, t_1)}{\alpha^\top \left(I_n - A^\top\right)^{-1} v(t_0)}$$

We can simplify this formula because $p(t_0) = \left(I_n - A^\top\right)^{-1} v(t_0) = \mathbf{1}_n$ and $\mathbf{1}_n^\top \alpha = 1$. Finally, we have:

$$\pi = \alpha^\top \left(I_n - A^\top\right)^{-1} \Delta v \tag{8.32}$$

In general, we define two price indices: the producer price index (PPI) where the basket weights are proportional to the output ($\alpha_j \propto x_j$) and the consumer price index (CPI) where the basket weights are proportional to the final demand ($\alpha_j \propto y_j$). In the latter case we obtain:

$$\text{CPI} = \frac{y^\top \left(I_n - A^\top\right)^{-1} \Delta v}{\mathbf{1}_n^\top y} = \frac{\Delta v^\top \left(I_n - A\right)^{-1} y}{\mathbf{1}_n^\top y} = \frac{\Delta v^\top x}{\mathbf{1}_n^\top y} = \frac{\sum_{j=1}^n x_j \Delta v_j}{\sum_{j=1}^n y_j}$$

Looking at the previous example, a 10% increase in energy costs will cause the producer price index to rise by 5.10% and the consumer price index by 4.15%.

Mathematical properties

Since we have $\left(I_n - A^\top\right)^{-1} = \left(\left(I_n - A\right)^\top\right)^{-1} = \left(\left(I_n - A\right)^{-1}\right)^\top$, we deduce that $\tilde{\boldsymbol{\mathcal{L}}} = \boldsymbol{\mathcal{L}}^\top$. So we can study $\boldsymbol{\mathcal{L}}$ or $\tilde{\boldsymbol{\mathcal{L}}}$ indifferently, because they share the same properties. The matrix $\boldsymbol{\mathcal{L}}$ admits the following Neumann series:

$$\begin{aligned} \boldsymbol{\mathcal{L}} &= (I_n - A)^{-1} \\ &= I_n + A + A^2 + A^3 + \ldots \\ &= \sum_{k=0}^\infty A^k \end{aligned} \tag{8.33}$$

Box 8.13: Neumann series

A Neumann series is a mathematical series of the form $S := \sum_{k=0}^{\infty} T^k$ where T is a bounded linear operator and $T^k = T^{k-1} \circ T = T \circ T^{k-1}$. If the Neumann series converges in the operator norm, then $\mathrm{Id} - T$ is invertible and its inverse is the Neumann series:

$$(\mathrm{Id} - T)^{-1} = S = \sum_{k=0}^{\infty} T^k$$

where Id is the identity operator. A sufficient condition is that the spectral radius of T is less than one. If A is an invertible matrix, we conclude that:

$$(I_n - A)^{-1} = \sum_{k=0}^{\infty} A^k$$

and $\lim_{k=0} \lVert A^k \rVert = 0$. This result generalizes the geometric series:

$$\frac{1}{1-x} = 1 + x + x^2 + \dots \qquad \text{where } |x| < 1$$

Since A is a doubly substochastic matrix[106], the eigendecomposition of A is $A = V\Lambda V^{-1}$ where V is the matrix of eigenvectors, $\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ and $|\lambda_i| \leq 1$. If we assume that A is irreducible, the Perron-Frobenius theorem states that the spectral radius $\varrho(A)$ is strictly lower than 1 and corresponds to the largest positive eigenvalue λ_1. This implies that all the eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ satisfies $|\lambda_i| < 1$. As we have $A^k = V\Lambda^k V^{-1}$, we deduce that $\lim_{k \to \infty} A^k = V\Lambda^k V^{-1} = \mathbf{0}_{n,n}$. The Neumann series $\sum_{k=0}^{\infty} A^k$ converges then to a finite matrix, which implies that the multiplier matrix $\boldsymbol{\mathcal{L}}$ is nonsingular.

From Equation (8.33), we deduce that $\boldsymbol{\mathcal{L}} \succeq I_n$ because $A^k \succeq \mathbf{0}_{n,n}$ for $k \geq 1$ (Property NN4 on page 656). We also get the following decomposition:

$$
\begin{aligned}
x &= \sum_{k=0}^{\infty} A^k y \\
&= y + Ay + A^2 y + \dots \\
&= \sum_{k=0}^{\infty} y_{(k)}
\end{aligned}
$$

where $y_{(0)} = y$ is the final demand (or zero-tier intermediate demand), $y_{(1)} = Ay$ is the first-tier intermediate demand, $y_{(2)} = A^2 y$ is the second-tier intermediate demand, and $y_{(k)} = A^k y$ is the kth-tier intermediate demand. Furthermore, we have:

$$\frac{\partial x}{\partial y} = (I_n - A)^{-1} \equiv \boldsymbol{\mathcal{L}} \succeq I_n$$

We can better understand why the matrix $\boldsymbol{\mathcal{L}}$ is also called the multiplier matrix because it is an analogy to Keynesian consumption theory and the effect of a change in aggregate demand on the output.

---

[106]See Appendix A.1.4 on page 656.

Table 8.25: Tier decomposition of the output (Example 21)

	k	0	1	2	3	4	5
	Energy	850.0	1 622.5	1 066.0	630.8	362.1	205.2
	Materials	875.0	1 183.8	805.1	486.3	282.1	160.7
$y_{(k)}$	Industrials	3 300.0	1 910.0	1 175.8	695.1	400.0	227.1
	Services	7 025.0	2 849.5	1 280.0	627.1	325.9	175.4
	Energy	850.0	2 472.5	3 538.5	4 169.2	4 531.3	4 736.5
	Materials	875.0	2 058.8	2 863.9	3 350.1	3 632.2	3 792.9
$y_{(0:k)}$	Industrials	3 300.0	5 210.0	6 385.8	7 080.9	7 480.9	7 708.0
	Services	7 025.0	9 874.5	11 154.5	11 781.6	12 107.5	12 283.0

In Tables 8.25 and 8.26, we report the values of intermediate demand $y_{(k)}$ in the case of Example 21. It can be seen that the first-tier intermediate demand is larger than final demand when we look at the energy and materials sector. For instance, we have $y_{(1),1} = 1622.5$ but $y_{(0),1} = y_1 = 850$. In the case of industrials and services, we obtain the opposite effect. We also check that the k-tier intermediate demand converges to zero. We have also reported the cumulative demand of the first k tiers: $y_{(0:k)} = \sum_{h=0}^{k} y_{(k)}$. This illustrates that $\lim_{k \to \infty} y_{(0:k)} = x$.

Table 8.26: Tier decomposition of the output (Example 21)

	k	10	20	25
	Energy	11.45	0.03	0.00
	Materials	9.01	0.03	0.00
$y_{(k)}$	Industrials	12.69	0.04	0.00
	Services	9.29	0.03	0.00
	Energy	4 985.42	4 999.96	5 000.00
	Materials	3 988.52	3 999.97	4 000.00
$y_{(0:k)}$	Industrials	7 983.83	7 999.95	8 000.00
	Services	12 488.18	12 499.96	12 500.00

Since $y_{(k)}$ converges to $\mathbf{0}_n$, we can ask whether this convergence is monotone. In particular, do we check that $y_{(k)} \succeq y_{(k+1)}$? The answer to this question is no. In fact, we can easily find counterexamples. The reason is that $A^k \not\succeq A^{k+1}$. For instance, we have seen that $y_{(1),1} \geq y_{(0),1}$ in Example 21. Nevertheless, we observe empirically that the relation $y_{(k)} \succeq y_{(k+1)}$ is satisfied for $k \geq k^\star$. This means that for sufficiently large k, the contribution of the kth tier decreases with respect to k. A sufficient (but not necessary) condition is that $A^k \succeq A^{k+1}$ for $k = k^\star$. Let us assume that $A^k \succeq A^{k+1}$ holds for $k = k^\star$. Let us note $B = A^k$, $C = A^{k+1}$ and $D = A$. Property NN1 (Appendix A.1.4 on page 656) implies that $BD \succeq CD$. This means that if $A^k \succeq A^{k+1}$, then $A^{k+1} \succeq A^{k+2}$. We conclude that if $A^k \succeq A^{k+1}$ for $k = k^\star$, then $A^k \succeq A^{k+1}$ for $k \geq k^\star$ and the relation $y_{(k)} \succeq y_{(k+1)}$ is satisfied for $k \geq k^\star$ if we assume that $y \succeq \mathbf{0}_n$.

Remark 60 *Theoretically, the partial monotonicity property does not hold if some elements of final demand are negative. Empirically, we observe that the property is always satisfied because $y_{(0:k)}$ converges to $\mathcal{L}y$ and the production is always positive: $\mathcal{L}y \equiv x \succeq \mathbf{0}_n$.*

Looking again at our numerical example, we check that the monotonicity property $y_{(k)} \succeq y_{(k+1)}$ is satisfied for $k \geq 1$.

Multi-regional input-output analysis

While an input-output table (IOT) usually covers only one region and tracks economic flows between sectors within that region, a multi-regional input-output table (MRIO) involves several regions. It includes national harmonized input-output tables and export and import data calculated within the same sectoral structure. The two best known MRIO databases are GTAP (global trade analysis project) and WIOD (world input-output database). They can be downloaded from `www.gtap.agecon.purdue.edu` and `www.rug.nl/ggdc/valuechain/wiod` respectively. The advantage of WIOD is that it is free, which is not the case with GTAP. However, the latest version is from November 2014. An alternative is to use the input-output tables provided by the OECD. The data can be found at `https://stats.oecd.org/Index.aspx?DataSetCode=IOTS`. Below we describe the WIOD table in order to understand the basic structure of an MRIO database.

We download the zipped file `WIOTS_in_EXCEL.zip` with the button `WIOT tables Excel` from `www.rug.nl/ggdc/valuechain/wiod/wiod-2016-release` and use the Excel binary spreadsheet `WIOT2014_Nov16_ROW.xlsb`, which is shown in Figure 8.120. This file contains the 2014 input-output table for 44 regions (28 EU countries, 15 other major countries and a global region corresponding to the rest-of-the-world aggregate) and 56 sectors. The list of countries and their ISO codes can be found in Table 8.28. ROW is the ISO code for the rest of the world. We also report the list of industries in Table 8.27. A comprehensive description of database can be found in Dietzenbacher *et al.* (2023) and Timmer *et al.* (2015). The structure of the database is:

	Z $(56 \times 44) \times (56 \times 44)$	y $(56 \times 44) \times (5 \times 44)$	x $(56 \times 44) \times 1$
Sum	$1 \times (56 \times 44)$	$1 \times (5 \times 44)$	$\mathbf{0}_{1,1}$
Value added	V $6 \times (56 \times 44)$	$6 \times (5 \times 44)$	$\mathbf{0}_{6,1}$
Output	w $1 \times (56 \times 44)$	$\mathbf{0}_{6,5 \times 44}$	$\mathbf{0}_{1,1}$

We have five main blocks of matrices: Z, y, x, V and w. First we have the matrix Z with 2 464 rows and 2 464 columns. It shows the flows between all sectors for all countries. For instance, the first 56 rows correspond to the 56 sectors of Australia (AUS), the first 56 columns correspond to the 56 sectors of Australia (AUS), the next 56 columns correspond to the 56 sectors of Austria (AUT). The structure of Z is then:

				To		
		AUS	AUT	$\cdots$	USA	ROW
	AUS	AUS $\longrightarrow$ AUS	AUS $\longrightarrow$ AUT	$\cdots$	AUS $\longrightarrow$ USA	AUS $\longrightarrow$ ROW
From	AUT	AUT $\longrightarrow$ AUS	AUT $\longrightarrow$ AUT		AUT $\longrightarrow$ USA	AUT $\longrightarrow$ ROW
	$\vdots$	$\vdots$				$\vdots$
	ROW	ROW $\longrightarrow$ AUS	ROW $\longrightarrow$ AUT	$\cdots$	ROW $\longrightarrow$ USA	ROW $\longrightarrow$ ROW

The dimension of each submatrix $\mathcal{C}_i \longrightarrow \mathcal{C}_j$ is equal to 56×56. All the items are expressed in millions of dollars. For example, we read in Figure 8.120 that $Z_{1,1} = \$12\,924$ mn and $Z_{1,2} = \$112$ mn. This means that the sector *crop and animal production, hunting and related service activities* in Australia sells $12.924 bn and $112 mn of goods and services to this sector and the sector *forestry and logging* in Australia. Second, the matrix y consists of five final demand items: (a) final consumption expenditure by households; (b) final consumption expenditure by non-profit organisations serving households (NPISH); (c) final consumption expenditure by government; (d)

Figure 8.120: World input-output database (November 2014)

Table 8.27: List of industries/sectors in the 2014 WIOD table

No.	Name
$\mathcal{S}_1$	Accommodation and food service activities
$\mathcal{S}_2$	Activities auxiliary to financial services and insurance activities
$\mathcal{S}_3$	Activities of extraterritorial organizations and bodies
$\mathcal{S}_4$	Activities of households as employers; undifferentiated goods- and services-producing activities of households for own use
$\mathcal{S}_5$	Administrative and support service activities
$\mathcal{S}_6$	Advertising and market research
$\mathcal{S}_7$	Air transport
$\mathcal{S}_8$	Architectural and engineering activities; technical testing and analysis
$\mathcal{S}_9$	Computer programming, consultancy and related activities; information service activities
$\mathcal{S}_{10}$	Construction
$\mathcal{S}_{11}$	Crop and animal production, hunting and related service activities
$\mathcal{S}_{12}$	Education
$\mathcal{S}_{13}$	Electricity, gas, steam and air conditioning supply
$\mathcal{S}_{14}$	Financial service activities, except insurance and pension funding
$\mathcal{S}_{15}$	Fishing and aquaculture
$\mathcal{S}_{16}$	Forestry and logging
$\mathcal{S}_{17}$	Human health and social work activities
$\mathcal{S}_{18}$	Insurance, reinsurance and pension funding, except compulsory social security
$\mathcal{S}_{19}$	Land transport and transport via pipelines
$\mathcal{S}_{20}$	Legal and accounting activities; activities of head offices; management consultancy activities
$\mathcal{S}_{21}$	Manufacture of basic metals
$\mathcal{S}_{22}$	Manufacture of basic pharmaceutical products and pharmaceutical preparations
$\mathcal{S}_{23}$	Manufacture of chemicals and chemical products
$\mathcal{S}_{24}$	Manufacture of coke and refined petroleum products
$\mathcal{S}_{25}$	Manufacture of computer, electronic and optical products
$\mathcal{S}_{26}$	Manufacture of electrical equipment
$\mathcal{S}_{27}$	Manufacture of fabricated metal products, except machinery and equipment
$\mathcal{S}_{28}$	Manufacture of food products, beverages and tobacco products
$\mathcal{S}_{29}$	Manufacture of furniture; other manufacturing
$\mathcal{S}_{30}$	Manufacture of machinery and equipment n.e.c.
$\mathcal{S}_{31}$	Manufacture of motor vehicles, trailers and semi-trailers
$\mathcal{S}_{32}$	Manufacture of other non-metallic mineral products
$\mathcal{S}_{33}$	Manufacture of other transport equipment
$\mathcal{S}_{34}$	Manufacture of paper and paper products
$\mathcal{S}_{35}$	Manufacture of rubber and plastic products
$\mathcal{S}_{36}$	Manufacture of textiles, wearing apparel and leather products
$\mathcal{S}_{37}$	Manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials
$\mathcal{S}_{38}$	Mining and quarrying
$\mathcal{S}_{39}$	Motion picture, video and television programme production, sound recording and music publishing activities; programming and broadcasting activities
$\mathcal{S}_{40}$	Other professional, scientific and technical activities; veterinary activities
$\mathcal{S}_{41}$	Other service activities
$\mathcal{S}_{42}$	Postal and courier activities
$\mathcal{S}_{43}$	Printing and reproduction of recorded media
$\mathcal{S}_{44}$	Public administration and defence; compulsory social security
$\mathcal{S}_{45}$	Publishing activities
$\mathcal{S}_{46}$	Real estate activities
$\mathcal{S}_{47}$	Repair and installation of machinery and equipment
$\mathcal{S}_{48}$	Retail trade, except of motor vehicles and motorcycles
$\mathcal{S}_{49}$	Scientific research and development
$\mathcal{S}_{50}$	Sewerage; waste collection, treatment and disposal activities; materials recovery; remediation activities and other waste management services
$\mathcal{S}_{51}$	Telecommunications
$\mathcal{S}_{52}$	Warehousing and support activities for transportation
$\mathcal{S}_{53}$	Water collection, treatment and supply
$\mathcal{S}_{54}$	Water transport
$\mathcal{S}_{55}$	Wholesale and retail trade and repair of motor vehicles and motorcycles
$\mathcal{S}_{56}$	Wholesale trade, except of motor vehicles and motorcycles

gross fixed capital formation; (e) changes in inventories and valuables. The sum of these five items is then the aggregate final demand. Third, the vector x corresponds to the output. For example, $x_1 = \$70\,292$ mn is the total production of the sector *crop and animal production, hunting and related service activities* in Australia. The fourth block matrix is V and has six rows: (a) taxes less subsidies on products; (b) cif/fob adjustments on exports; (c) direct purchases abroad by residents; (d) purchases on the domestic territory by non-residents; (e) value added at basic prices; (f) international transport margins. Again, we can aggregate these six items to calculate a total value added. Finally, the last block is the row vector w. By construction, we have $w_j = x_j$. For example, we check that w_1 is equal to \$70\,292 mn with the following decomposition: the total intermediate consumption is equal to \$39\,039 mn while the total value added is equal to $502 + 0 + 0 + 0 + 30\,489 + 262 = \$31\,253$ mn.

Table 8.28: List of countries/regions in the 2014 WIOD table

No.	ISO	Name	No.	ISO	Name
$\mathcal{C}_1$	AUS	Australia	$\mathcal{C}_2$	AUT	Austria
$\mathcal{C}_3$	BEL	Belgium	$\mathcal{C}_4$	BGR	Bulgaria
$\mathcal{C}_5$	BRA	Brazil	$\mathcal{C}_6$	CAN	Canada
$\mathcal{C}_7$	CHE	Switzerland	$\mathcal{C}_8$	CHN	China
$\mathcal{C}_9$	CYP	Cyprus	$\mathcal{C}_{10}$	CZE	Czech Republic
$\mathcal{C}_{11}$	DEU	Germany	$\mathcal{C}_{12}$	DNK	Denmark
$\mathcal{C}_{13}$	ESP	Spain	$\mathcal{C}_{14}$	EST	Estonia
$\mathcal{C}_{15}$	FIN	Finland	$\mathcal{C}_{16}$	FRA	France
$\mathcal{C}_{17}$	GBR	United Kingdom	$\mathcal{C}_{18}$	GRC	Greece
$\mathcal{C}_{19}$	HRV	Croatia	$\mathcal{C}_{20}$	HUN	Hungary
$\mathcal{C}_{21}$	IDN	Indonesia	$\mathcal{C}_{22}$	IND	India
$\mathcal{C}_{23}$	IRL	Ireland	$\mathcal{C}_{24}$	ITA	Italy
$\mathcal{C}_{25}$	JPN	Japan	$\mathcal{C}_{26}$	KOR	Republic of Korea
$\mathcal{C}_{27}$	LTU	Lithuania	$\mathcal{C}_{28}$	LUX	Luxembourg
$\mathcal{C}_{29}$	LVA	Latvia	$\mathcal{C}_{30}$	MEX	Mexico
$\mathcal{C}_{31}$	MLT	Malta	$\mathcal{C}_{32}$	NLD	Netherlands
$\mathcal{C}_{33}$	NOR	Norway	$\mathcal{C}_{34}$	POL	Poland
$\mathcal{C}_{35}$	PRT	Portugal	$\mathcal{C}_{36}$	ROU	Romania
$\mathcal{C}_{37}$	RUS	Russian Federation	$\mathcal{C}_{38}$	SVK	Slovakia
$\mathcal{C}_{39}$	SVN	Slovenia	$\mathcal{C}_{40}$	SWE	Sweden
$\mathcal{C}_{41}$	TUR	Turkey	$\mathcal{C}_{42}$	TWN	Taiwan
$\mathcal{C}_{43}$	USA	United States of America	$\mathcal{C}_{44}$	ROW	Rest-of-the-world

In Figure 8.121 we plot the sparsity pattern of the input-output matrix, and only the values of $A_{i,j}$ greater than 5% are colored. We see that the density of the matrix is mainly within the country submatrices. Outside of these intra-country matrices, the input-output table is sparse except for a few countries: China, Germany, Russia, USA and the rest-of-the-world region.

Let us now analyze the Leontief matrix $\mathcal{L} = (I - A)^{-1}$. In Figure 8.122, we perform the eigen-decomposition $A = V\Lambda V^{-1}$ and plot the spectrum of A. Figure 8.123 shows the Frobenius norm of the matrix A^k for $k = 0, \ldots, 10$ and the Leontief matrix $\mathcal{L}$. We find that the convergence of the Leontief matrix is achieved very fast since the Frobenius norm of A^k is less than 1 after the third tier and 0.1 after the seventh tier.

Figure 8.121: Sparsity pattern of the input-output matrix A
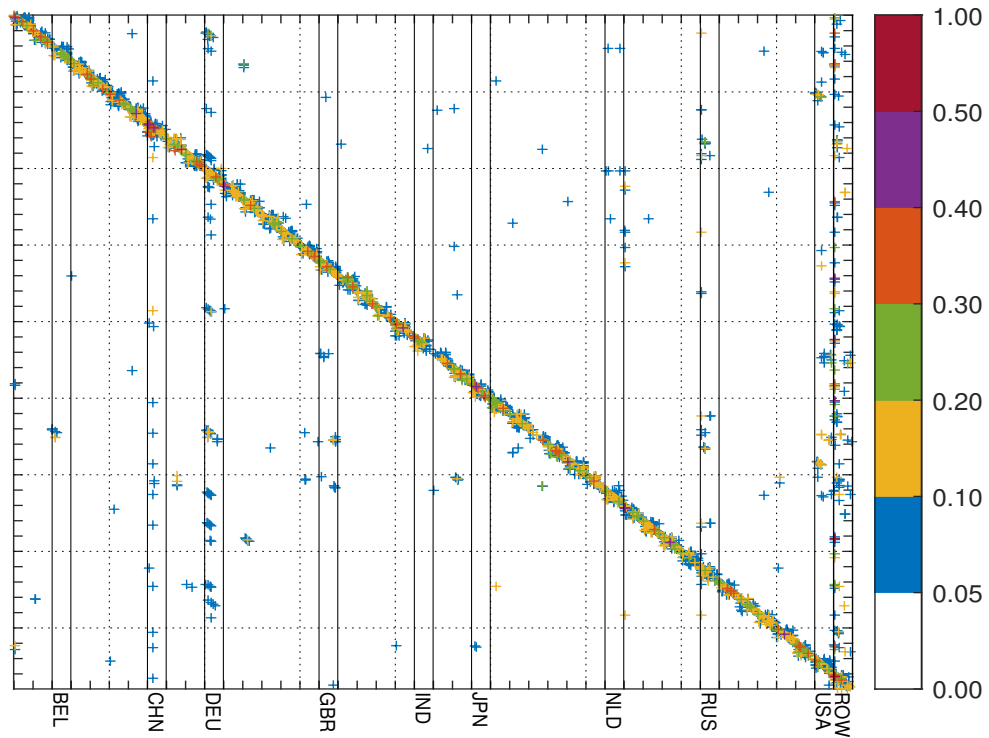
Figure 8.122: Spectrum of the matrix A
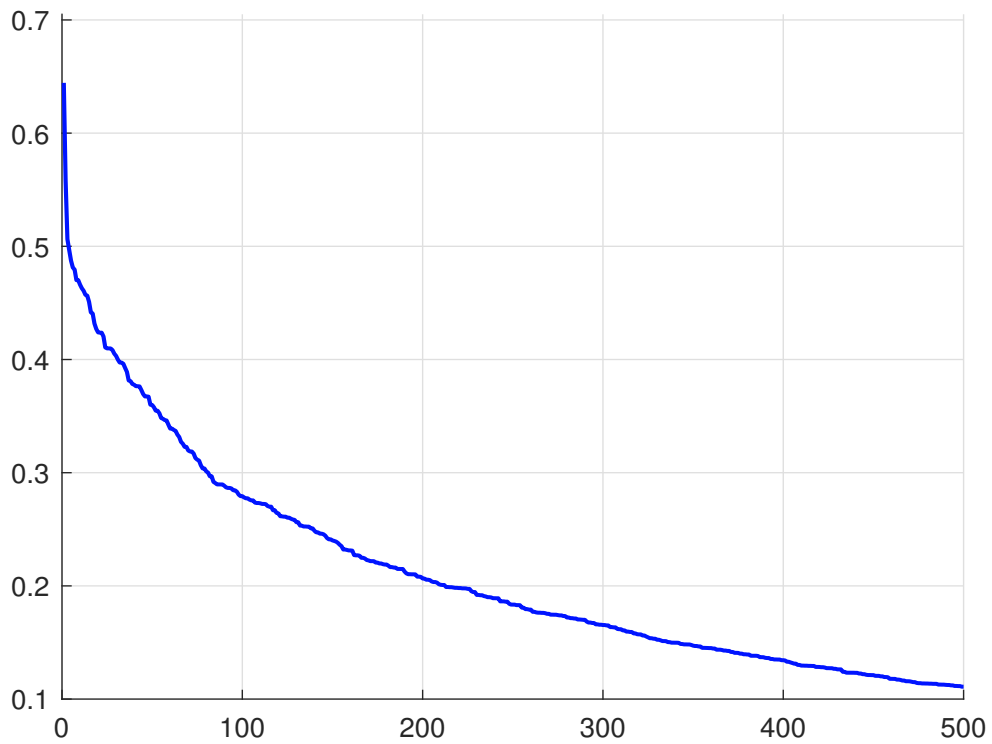
Figure 8.123: Frobenious norm of the matrix A^k
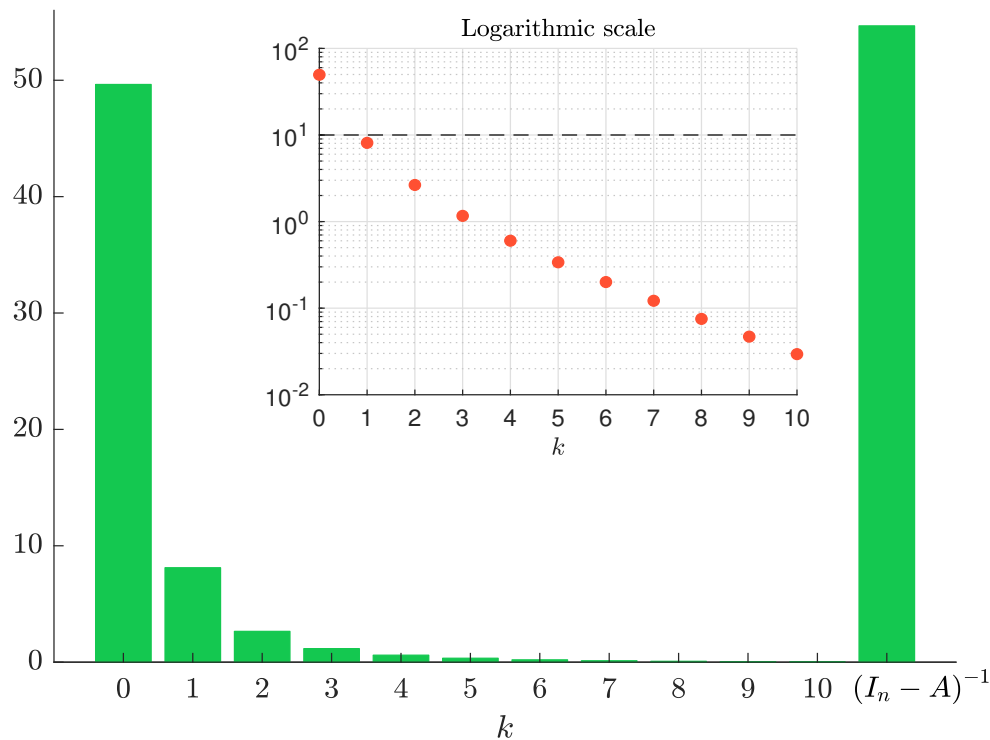
8.4.2   Application to environmental problems

In the late 1960s, several authors proposed linking economic and ecological systems using generalized input-output models. For example, Daly (1968) suggested extending the technical coefficients with additional rows/columns to reflect non-human sectors such as animals, plants, and bacteria, and non-living sectors such as the atmosphere, hydrosphere, and lithosphere. Leontief (1970) himself explained how externalities such as pollution could be incorporated into a basic input-output model. Since these first contributions, input-output analysis has been extended to many environmental problems[107].

Production-based vs. consumption-based inventory

To understand how input-output analysis can be used to measure carbon emissions, consider the mathematical problem of calculating the contribution of carbon emissions per product. Following Miller and Blair (2009), we denote by $C^{(x)} = \left(C_{g,j}^{(x)} \right)$ the pollution output matrix where $C_{g,j}^{(x)}$ is the total amount of the gth pollutant generated by the output of the jth sector. Similarly, we define $D^{(y)} = C^{(x)} \operatorname{diag}(x)^{-1} = \left(D_{g,j}^{(y)} \right)$ the matrix of direct impact coefficients where $D_{g,j}^{(y)} = c_{g,j}^{(y)}/x_j$ is the amount of the gth pollutant generated by 1\$ of the output of the jth sector. Let $\varpi = (\varpi_1, \ldots, \varpi_m)$ be the vector of pollution level. We have:

$$\varpi = D^{(x)}x = D^{(x)} \left(I_n - A \right)^{-1} y = D^{(y)}y$$

where $D^{(y)} = D^{(x)} \left(I_n - A \right)^{-1}$ is the pollutant multiplier matrix with respect to the final demand y. $D^{(y)}$ also measures the product carbon footprint (PCF). Since we have the following identity

---

[107]See Chapters 9 and 10 of Miller and Blair (2009).

$\varpi_g = \left(D^{(y)}y\right)_g = \sum_{j=1}^{n} D^{(y)}_{g,j} y_j$, we deduce that the total contribution of sector j to the gth pollutant is equal to:

$$C^{(x)}_{g,j} = \frac{\partial \varpi_g}{\partial y_j} y_j = D^{(y)}_{g,j} y_j$$

Again, we can decompose the pollutant level according to the kth tier. We have:

$$\varpi = D^{(y)}y = \sum_{k=0}^{\infty} D^{(x)} A^k y = \sum_{k=0}^{\infty} \varpi_{(k)}$$

where $\varpi_{(0)} = D^{(x)}y$ is the pollutant level due to the final demand (or the zero-tier pollutant level), $\varpi_{(1)} = D^{(x)}Ay$ is the pollutant level due to the first-tier supply chain, and $\varpi_{(k)} = D^{(x)}A^k y$ is the kth-tier pollutant level. The matrix $D^{(y)}_{(k)} = D^{(x)}A^k$ is called the kth-tier multiplier matrix and satisfies the identity $D^{(y)} \equiv \sum_{k=0}^{\infty} D^{(y)}_{(k)}$.

Example 23 *We consider three products, whose input-output table is given below:*

		To			Final demand	Total output
		P_1	P_2	P_3	y	x
From	P_1	100	300	100	500	1 000
	P_2	250	150	200	1 600	2 000
	P_3	25	200	75	200	500
	Value added	625	1 350	125		
	Total outlays	1 000	2 000	500		
GHG	CO_2	50	20	5	75	
	CH_4	3	1	0	4	

Intermediate production of \$100 of P_1, \$300 of P_2, and \$100 of P_3 is required to produce \$500 of P_1. This environmentally-extended input-output table has two additional rows corresponding to the GHG *emissions. For instance, the production of P_1 causes 50 kgCO_2 and 3 kgCH_4.*

The matrix of technical coefficients is equal to:

$$A = Z \operatorname{diag}(x)^{-1} = \begin{pmatrix} 10.0\% & 15.0\% & 20.0\% \\ 25.0\% & 7.5\% & 40.0\% \\ 2.5\% & 10.0\% & 15.0\% \end{pmatrix}$$

It follows that the matrix of multipliers is equal to:

$$\mathcal{L} = (I_3 - A)^{-1} = \begin{pmatrix} 1.1871 & 0.2346 & 0.3897 \\ 0.3539 & 1.2090 & 0.6522 \\ 0.0766 & 0.1491 & 1.2647 \end{pmatrix}$$

The direct impact matrix is equal to the GHG emissions divided by the output:

$$D^{(x)} = \begin{pmatrix} 50/1000 & 20/2000 & 5/500 \\ 3/1000 & 1/2000 & 0/500 \end{pmatrix} = \begin{pmatrix} 0.05 & 0.01 & 0.01 \\ 0.003 & 0.0005 & 0 \end{pmatrix}$$

The unit of $D^{(x)}$ is expressed in kilogram of the gas per dollar. For instance, the GHG intensities of the product P_1 are equal to $0.05 \text{ kgCO}_2/\$$ and $0.003 \text{ kgCH}_4/\$$. Finally, we obtain:

$$D^{(y)} = D^{(x)}\mathcal{L} = \left(\begin{array}{ccc} 0.0637 & 0.0253 & 0.0387 \\ 0.0037 & 0.0013 & 0.0015 \end{array} \right)$$

While $D^{(x)}$ corresponds to the production-based inventory, $D^{(y)}$ measures the carbon footprint from the perspective of the consumption-based inventory (Kitzes, 2013). This gives us the following decomposition:

$$C^{(y)} = \left(\begin{array}{ccc} 31.83 & 35.44 & 7.73 \\ 1.87 & 1.83 & 0.30 \end{array} \right) \neq \left(\begin{array}{ccc} 50 & 20 & 5 \\ 3 & 1 & 0 \end{array} \right) = C^{(x)}$$

We notice that the two contribution matrices are different. For instance, while the production of P_1 is responsible of 50 kgCO_2, the final consumption of P_1 is responsible of only 31.83 kgCO_2, meaning that 18.17 kgCO_2 are emitted by P_1 for the other two products. The difference between $C^{(x)}$ and $C^{(y)}$ depends on the structure of the matrix A and the vector y. Let us apply a perturbation to the previous example. We assume that there is no final demand of P_1: $y_1 = 0$. For consistency, we also assume that the GHG emissions of Product P_1 are divided by two because the production is reduced by 50%. We find that:

$$C^{(y)} = \left(\begin{array}{ccc} 0 & 40.96 & 9.04 \\ 0 & 2.13 & 0.37 \end{array} \right) \neq \left(\begin{array}{ccc} 25 & 20 & 5 \\ 1.5 & 1 & 0 \end{array} \right) = C^{(x)}$$

The consumption-based contribution of P_1 is null because P_1 is manufactured for the other two products.

Remark 61 *We can show that $C^{(x)} = C^{(y)}$ implies that A is a diagonal matrix. We conclude that the supply chain and interconnectedness between sectors can lead to a misperception of the sectoral carbon footprint.*

The previous framework can be applied to many problems involving the calculation of carbon footprints. Miller and Blair (2009) examined three categories of EEIO analysis: generalized input-output, economic-ecological, and commodity-by-industry models. An overview of generalized input-output models can be found in Minx *et al.* (2009) and Wiedmann (2009). These models are commonly used to calculate the carbon footprint of nations, sectors, supply chains, etc., and to analyze the impact of foreign trade. The use of economic-ecological models is less popular, since it involves building an input-output table for ecological sectors (species, plants, etc.). Commodity-by-industry models are more studied because it is easier to collect data for the commodity sector.

EEIO databases

The use of environmentally-extended input-output models requires credible database. Two EEIO databases dominate the market: Eora and Exiobase. According to https://worldmrio.com, the Eora global supply chain database uses more than $15\,000$ sectors across 190 countries, and contains about $2\,700$ environmental indicators covering GHG emissions, air pollution, energy use, water demand, land use, etc. Exiobase is a multi-regional environmentally-extended supply-use and input-output model with 44 countries, 163 industries, 200 products and 417 emission categories (www.exiobase.eu).

We consider the Exiobase data for the year 2022. To do this, we download the MRIO archive in industry by industry format (IOT_2022_ixi.zip) and the MRIO archive in product by product format (IOT_2022_pxp.zip) from the website: https://zenodo.org/record/5589597. Both archives

have the same structure. The economic input-output data are stored in the root of the archive, which contains four data files:

1. The Z matrix (`Z.txt` — flow/transactions matrix)

2. The A matrix (`A.txt` — direct requirements matrix)

3. The y matrix (`y.txt` — final demand)

4. The x vector (`x.txt` — gross/total output)

and four description files (`finaldemand.txt`, `industries.txt`, `products.txt` and `units.txt`). Exiobase provides also two set of extension data stored in the sub-folders `satellite` (uncharacterized stressors data — *e.g.*, CO_2 emissions, land use per category, etc.) and `impacts` (characterized stressors data — *e.g.*, total GWP100, total land use, etc.). The structure of the two sub-folders is the following:

- Factors of productions/stressors/impacts: `F.txt`

- Stressors/impacts of the final demand: `F_Y.txt`

- Direct stressor/impact coefficients: `S.txt`

- Stressor/impact coefficients of the final demand: `S_Y.txt`

- MRIO extension multipliers (total requirement factors of consumption): `M.txt`

- Consumption based accounts per sector: `D_cba.txt`

- Production based accounts per sector: `D_pba.txt`

- Consumption based accounts per region: `D_cba_reg.txt`

- Production based accounts per region: `D_pba_reg.txt`

- Import accounts per region: `D_imp_reg.txt`

- Export accounts per region: `D_exp_reg.txt`

- Absolute units of the stressor and impacts: `unit.txt`

Previously, we have given the detailed sectors in the WIOD database, which is a 56×56 industry by industry IOT. Exiobase is both a 163×163 industry by industry IOT and a 200×200 product by product IOT. The first fifteen products are listed here: paddy rice, wheat, cereal grains nec, vegetables, fruit, nuts, oil seeds, sugar cane and beet, plant-based fibers, crops nec, cattle, pigs, poultry, meat animals nec, animal products nec, raw milk, wool and silk-worm cocoons. In the impact sub-directory, we find measurement that concerns employment (in hour), GHG emissions (in GWP100), water consumption (in Mm^3), Nitrogen (in kg), land-use crop, forest, pasture (in km^2), etc. In the satellite sub-directory, we have similar information but in a disaggregated form. For instance, the employment item has seven categories: low-skilled male, low-skilled female, medium-skilled male, medium-skilled female, high-skilled male, high-skilled female, vulnerable.

8.4.3 Estimation of first-tier and indirect emissions

Basic formula

We assume that the carbon footprint is evaluated in CO_2e, which means that the input-output analysis considers only one pollutant, with all greenhouse gases being converted to the carbon based on their global warming potential. In this case, $D^{(x)}$ is a row vector of dimension n, and $D_j^{(x)}$ measures the direct emission intensity of sector j. We reiterate that the total emission intensities are equal to $D^{(y)} = D^{(x)}\mathcal{L} = D^{(x)}(I_n - A)^{-1}$. $D^{(y)}$ is a row vector of dimension n, and $D_j^{(y)}$ measures the direct and indirect emission intensity of sector j. Using the usual notation $\mathcal{CI}$ for the carbon intensity, we have[108]:

$$
\begin{aligned}
\mathcal{CI}_{\text{total}} &= \mathcal{L}^\top \mathcal{CI}_1 \\
&= \left(I_n - A^\top\right)^{-1}\mathcal{CI}_1 \\
&= \tilde{\mathcal{L}}\,\mathcal{CI}_1
\end{aligned}
\tag{8.34}
$$

where $\mathcal{CI}_1 = \mathcal{CI}_{\text{direct}}$ is the vector of direct carbon intensities and $\mathcal{CI}_{\text{total}}$ is the vector of direct plus indirect carbon intensities. It follows that the indirect carbon intensities are given by:

$$
\begin{aligned}
\mathcal{CI}_{\text{indirect}} &= \mathcal{CI}_{\text{total}} - \mathcal{CI}_1 \\
&= \left(\left(I_n - A^\top\right)^{-1} - I_n\right)\mathcal{CI}_{\text{direct}}
\end{aligned}
\tag{8.35}
$$

In particular, we can decompose $\mathcal{CI}_{\text{indirect}}$ using the Neumann series:

$$
\mathcal{CI}_{\text{indirect}} = \underbrace{A^\top \mathcal{CI}_1}_{\text{First-tier}} + \underbrace{\left(A^\top\right)^2 \mathcal{CI}_1}_{\text{Second-tier}} + \ldots + \underbrace{\left(A^\top\right)^k \mathcal{CI}_1}_{k^{th}\text{-tier}} + \ldots
\tag{8.36}
$$

and we have:

$$
\mathcal{CI}_{\text{total}} = \underbrace{\underbrace{\mathcal{CI}_1}_{\text{Scope 1}}}_{\text{Direct intensities}} + \underbrace{\underbrace{A^\top \mathcal{CI}_1}_{\text{First-tier}} + \underbrace{\left(A^\top\right)^2 \mathcal{CI}_1}_{\text{Second-tier}} + \ldots + \underbrace{\left(A^\top\right)^k \mathcal{CI}_1}_{k^{th}\text{-tier}} + \ldots}_{\text{Indirect intensities}}
\tag{8.37}
$$

Equations (8.34–8.37) are the core formulas of the consumption-based inventory approach and the calculation of indirect carbon intensities.

Illustration

Continuing with Example 22, let us assume that the carbon emissions, expressed in $ktCO_2e$, are as follows: 500 for the energy sector, 200 for the materials sector, 200 for the industrials sector and 125 for the services sector. We deduce that the vector of Scope 1 carbon intensities is equal to:

$$
\mathcal{CI}_1 = \text{diag}(x)^{-1}\mathcal{CE}_1 = \begin{pmatrix} 500/5\,000 \\ 200/4\,000 \\ 200/8\,000 \\ 125/12\,500 \end{pmatrix} \times 10^3 = \begin{pmatrix} 100 \\ 50 \\ 25 \\ 10 \end{pmatrix}
$$

---

[108]Because $D^{(x)} = \mathcal{CI}_1^\top$ and $D^{(y)} = \mathcal{CI}_{\text{total}}^\top$.

We multiply the carbon emissions by a factor of $1\,000$ to express carbon intensities[109] in $gCO_2e/\$$ or equivalently $tCO_2e/\$$ mn. Energy is the most polluting sector with 500 $ktCO_2e$, followed by materials and industrials with 200 $ktCO_2e$. Energy and services have respectively the highest and lowest carbon intensity, respectively 100 $tCO_2e/\$$ mn and 10 $tCO_2e/\$$ mn. We have:

$$
\begin{aligned}
\mathcal{CI}_{\text{total}} &= \tilde{\mathcal{L}}\,\mathcal{CI}_1 \\
&= \begin{pmatrix} 1.1881 & 0.1678 & 0.1430 & 0.0715 \\ 0.3894 & 1.2552 & 0.4110 & 0.1718 \\ 0.4919 & 0.4336 & 1.6303 & 0.2993 \\ 0.2884 & 0.1891 & 0.3044 & 1.6087 \end{pmatrix} \begin{pmatrix} 100 \\ 50 \\ 25 \\ 10 \end{pmatrix} \\
&= \begin{pmatrix} 131.49 \\ 113.69 \\ 114.62 \\ 61.99 \end{pmatrix}
\end{aligned}
$$

We can then decompose $\mathcal{CI}_{\text{total}}$ between $\mathcal{CI}_{\text{direct}}$ and $\mathcal{CI}_{\text{indirect}}$. Finally, we obtain we obtain the direct and indirect carbon intensities given in Table 8.29. While the Scope 1 carbon intensity of the energy sector is equal to 100 $tCO_2e/\$$ mn, its total carbon intensity is equal to 131.49 $tCO_2e/\$$ mn. The difference of 31.49 $tCO_2e/\$$ mn corresponds to the indirect emissions. In the case of the energy sector, direct and indirect emissions represent 76.05% and 23.95% of the total emissions, respectively. In fact, this sector has the lowest ratio of indirect carbon emissions. On the contrary, for the services sector, 83.87% of the total emissions are indirect.

Table 8.29: Direct and indirect carbon intensities

Sector	$\mathcal{CI}_1$	$\mathcal{CI}_{\text{total}}$	$\mathcal{CI}_{\text{direct}}$	$\mathcal{CI}_{\text{indirect}}$	$\mathcal{CI}_{\text{direct}}$	$\mathcal{CI}_{\text{indirect}}$	$\dfrac{\mathcal{CI}_{\text{total}}}{\mathcal{CI}_1}$
		(in $tCO_2e/\$$ mn)			(in %)		
Energy	100.00	131.49	100.00	31.49	76.05%	23.95%	1.31
Materials	50.00	113.69	50.00	63.69	43.98%	56.02%	2.27
Industrials	25.00	114.62	25.00	89.62	21.81%	78.19%	4.58
Services	10.00	61.99	10.00	51.99	16.13%	83.87%	6.20

We denote by $\mathcal{CI}_{(k)} = \left(A^{\top}\right)^k \mathcal{CI}_1$ the indirect carbon intensity when considering the kth tier, and $\mathcal{CI}_{(1-k)} = \sum_{h=1}^{k} \left(A^{\top}\right)^h \mathcal{CI}_1$ the cumulative indirect carbon intensity for the first k tiers. The results are shown in Table 8.30. For the services sector, the first and second tiers add 18.50 and 13.50 $tCO_2e/\$$ mn to the indirect carbon intensity, respectively. If we restrict the analysis to the first two tiers, the indirect carbon intensity is equal to 32. The tree presented in Figure 8.124 explains this computation[110]. To produce $\$1$ of services, we must purchase $0.10\$$ of energy, $\$0.05$ of materials, and so on. It follows that the first-tier indirect carbon intensities for the services sector is equal to:

$$
\begin{aligned}
\mathcal{CI}_{(1)}\left(\mathcal{S}_4\right) &= 0.10 \times 100 + 0.05 \times 50 + 0.10 \times 25 + 0.35 \times 10 \\
&= 10 + 2.5 + 2.5 + 3.5 \\
&= 18.50
\end{aligned}
$$

---

[109]Remember that the unit of carbon intensity is $tCO_2e/\$$ mn. When we divide by 1 million, we get an equivalent unit of $gCO_2e/\$$.

[110]To make the graph easier to read, we use the following correspondences: energy $\longleftarrow \mathcal{S}_1$, materials $\longleftarrow \mathcal{S}_2$, industrials $\longleftarrow \mathcal{S}_3$ and services $\longleftarrow \mathcal{S}_4$.

Figure 8.124: Upstream tree of the first- and second-tier rounds for the services sector
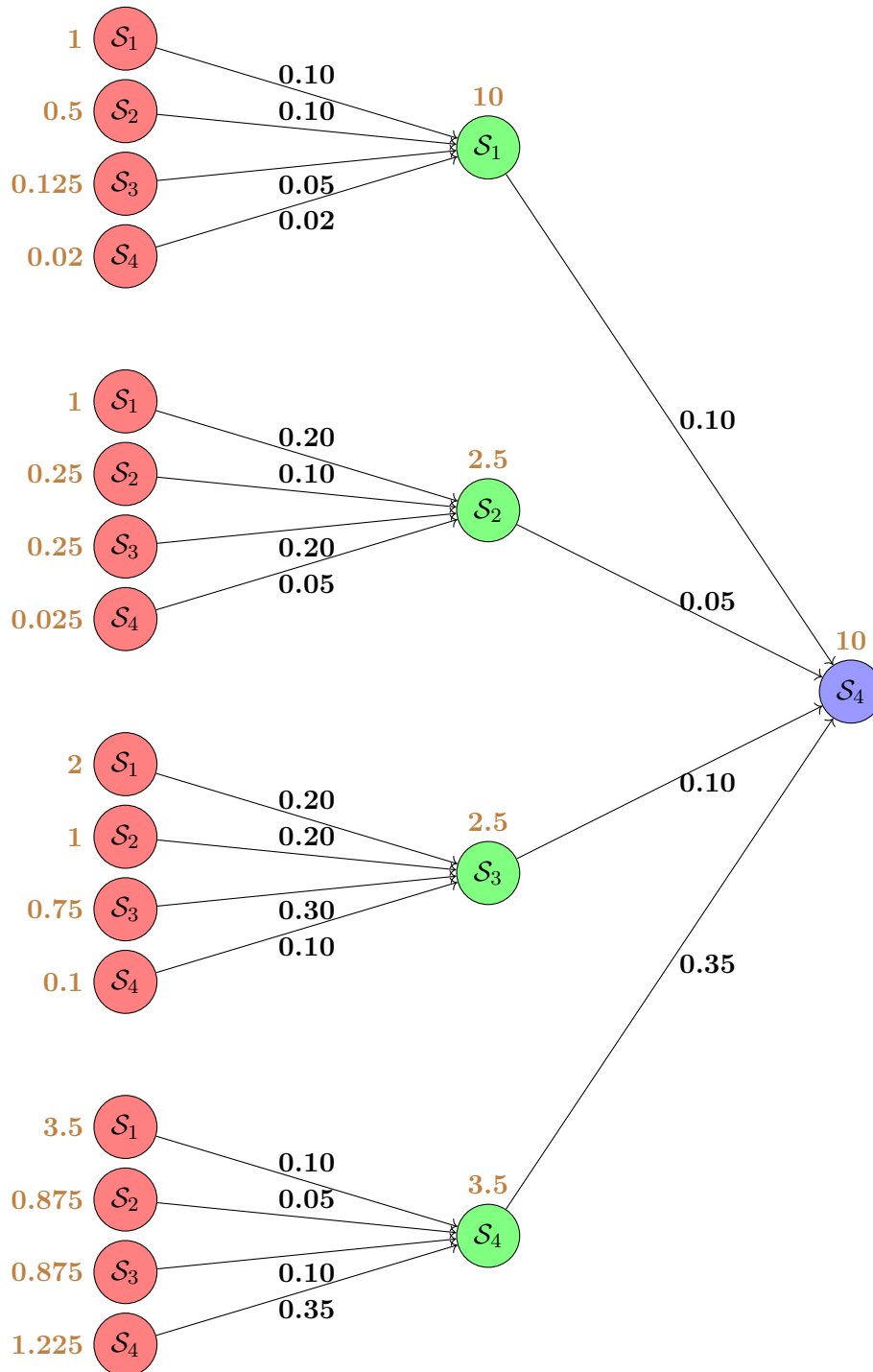
We can continue the analysis and consider the second tier. In fact, the companies involved in the first-tier round also purchase goods and services that generate new indirect emissions. We have:

$$
\begin{aligned}
\boldsymbol{CI}_{(2)}\left(\mathcal{S}_4\right) \;=\; & 0.10 \times \underbrace{(0.10 \times 100 + 0.10 \times 50 + 0.05 \times 25 + 0.02 \times 10)}_{\text{Indirect emissions from energy sector companies}} + \\
& 0.05 \times \underbrace{(0.20 \times 100 + 0.10 \times 50 + 0.20 \times 25 + 0.05 \times 10)}_{\text{Indirect emissions from materials sector companies}} + \\
& 0.10 \times \underbrace{(0.20 \times 100 + 0.20 \times 50 + 0.30 \times 25 + 0.10 \times 10)}_{\text{Indirect emissions from industrials sector companies}} + \\
& 0.35 \times \underbrace{(0.10 \times 100 + 0.05 \times 50 + 0.10 \times 25 + 0.35 \times 10)}_{\text{Indirect emissions from services sector companies}} \\
\;=\; & 13.495
\end{aligned}
$$

We can continue the analysis, and we verify that $\boldsymbol{CI}_{(3)}\left(\mathcal{S}_4\right) = 8.45$, $\boldsymbol{CI}_{(4)}\left(\mathcal{S}_4\right) = 4.98$, etc. Finally, the cumulative sum converges to $\boldsymbol{CI}_{(1-\infty)}\left(\mathcal{S}_4\right) = 51.99$.

Table 8.30: Tier decomposition of carbon intensities

	Sector	1	2	3	4	5	10	15	∞
	Energy	16.45	6.99	3.60	1.97	1.09	0.06	0.00	0.00
$\boldsymbol{CI}_{(k)}$	Materials	30.50	14.97	8.13	4.47	2.48	0.14	0.01	0.00
	Industrials	38.50	22.79	12.58	6.96	3.88	0.21	0.01	0.00
	Services	18.50	13.50	8.45	4.98	2.86	0.16	0.01	0.00
	Energy	16.45	23.44	27.04	29.02	30.11	31.41	31.48	31.49
$\boldsymbol{CI}_{(1-k)}$	Materials	30.50	45.47	53.59	58.06	60.55	63.52	63.68	63.69
	Industrials	38.50	61.29	73.87	80.83	84.71	89.35	89.61	89.62
	Services	18.50	32.00	40.44	45.43	48.29	51.79	51.98	51.99

The previous analysis concerns the carbon intensity. To estimate total emissions, we simply multiply by the output and we have the following identities:

$$
\frac{\boldsymbol{CE}_{\text{total}}}{\boldsymbol{CE}_1} = \frac{\boldsymbol{CI}_{\text{total}}}{\boldsymbol{CI}_1} \Leftrightarrow \boldsymbol{CE}_{\text{total}} = \boldsymbol{CI}_{\text{total}} \odot \frac{\boldsymbol{CE}_1}{\boldsymbol{CI}_1} = x \odot \boldsymbol{CI}_{\text{total}} \tag{8.38}
$$

Therefore, the indirect emissions are given by:

$$
\boldsymbol{CE}_{\text{indirect}} = \boldsymbol{CE}_{\text{total}} - \boldsymbol{CE}_{\text{direct}} = \left(\boldsymbol{CI}_{\text{total}} - \boldsymbol{CI}_1\right) \odot \frac{\boldsymbol{CE}_1}{\boldsymbol{CI}_1} \tag{8.39}
$$

The breakdown of the total carbon emissions is provided in Table 8.31. We notice that indirect carbon emissions are double counted. Indeed, the total direct carbon emissions are equal to $1\,025$ ktCO$_2$e and the indirect emissions add $1\,779$ ktCO$_2$e. Based on direct emissions, we have the following distribution: 49% for energy, 20% for materials, 20% for industrials and 12% for the sector of services. If we include the indirect emissions, we get a different picture. For instance, the services sector represents more than 25% of the total emissions because the direct emissions have been multiplied by a factor of 6.2, while the contribution of energy is now less than 25%.

Remark 62 *It would be wrong to directly diffuse the carbon emissions instead of the carbon intensities: $\boldsymbol{CE}_{\text{total}} = \left(I_n - A^\top\right)^{-1} \boldsymbol{CE}_1$. Indeed, carbon emissions are not comparable from one sector to another sector, because they are not normalized and monetary input-output tables give the technical coefficients for $1 of output from each sector.*

Table 8.31: Decomposition of carbon emissions

Sector	$\mathcal{CE}_{\text{direct}}$	$\mathcal{CE}_{\text{indirect}}$ (in ktCO$_2$e)	$\mathcal{CE}_{\text{total}}$	$\mathcal{CE}_{\text{direct}}$	$\mathcal{CE}_{\text{indirect}}$ (in %)	$\mathcal{CE}_{\text{total}}$
Energy	500	157.44	657.44	48.78	8.85	23.45
Materials	200	254.76	454.76	19.51	14.32	16.22
Industrials	200	716.97	916.97	19.51	40.30	32.70
Services	125	649.92	774.92	12.20	36.53	27.64
Total	1 025	1 779.10	2 804.10	100.00	100.00	100.00

Upstream vs. downstream analysis

The previous analysis is an output-based analysis. This is obvious if we look at Figure 8.124, which illustrates the impacts of requirement to produce \$1 in a sector. Once we have produced \$1 in a given sector, we may wonder how it is used by the value chain. In this case, we get an input-based analysis. In fact, instead of moving up the supply chain, we move down the value chain (Figure 8.125). Therefore, this approach is also known as downstream analysis while the output-based approach is known as upstream analysis.

To perform a downstream analysis, we must first define the technical coefficients for \$1 input (and not output):

$$\breve{A}_{i,j} = \frac{Z_{i,j}}{x_i}$$

$\breve{A}_{i,j}$ indicates the proportion of \$1 produced by sector i that is used by sector j. We denote the matrix of input impacts by $\breve{A} = \left(\breve{A}_{i,j}\right) = \text{diag}(x)^{-1} Z$. We note that:

$$\breve{A}_{i,j} = \frac{Z_{i,j}}{x_j} \cdot \frac{x_j}{x_i} = A_{i,j} \cdot T_{i,j}$$

In a matrix form, we have $\breve{A} = A \odot T$ where $T = (T_{i,j}) = \left(x_i^{-1} x_j\right)$. Using the same reasoning as in the previous paragraphs, we can show that:

$$\mathcal{CI}_{\text{total}}^{\text{down}} = \left(I_n - \breve{A}\right)^{-1} \mathcal{CI}_1 \tag{8.40}$$

where $\mathcal{CI}_{\text{total}}^{\text{down}}$ is the vector of downstream (direct plus indirect downstream) carbon intensities. It follows that the indirect downstream carbon intensities are given by:

$$\mathcal{CI}_{\text{indirect}}^{\text{down}} = \mathcal{CI}_{\text{total}}^{\text{down}} - \mathcal{CI}_1 = \left(\left(I_n - \breve{A}\right)^{-1} - I_n\right) \mathcal{CI}_{\text{direct}} \tag{8.41}$$

In particular, we can decompose $\mathcal{CI}_{\text{indirect}}^{\text{down}}$ as follows:

$$\mathcal{CI}_{\text{indirect}}^{\text{down}} = \underbrace{\breve{A}\mathcal{CI}_1}_{\text{First-tier}} + \underbrace{\breve{A}^2\mathcal{CI}_1}_{\text{Second-tier}} + \ldots + \underbrace{\breve{A}^k\mathcal{CI}_1}_{k^{th}\text{-tier}} + \ldots$$

and we have:

$$\mathcal{CI}_{\text{total}}^{\text{down}} = \underbrace{\underbrace{\mathcal{CI}_1}_{\text{Scope 1}}}_{\text{Direct downstream}} + \underbrace{\underbrace{\breve{A}\mathcal{CI}_1}_{\text{First-tier}} + \underbrace{\breve{A}^2\mathcal{CI}_1}_{\text{Second-tier}} + \ldots + \underbrace{\breve{A}^k\mathcal{CI}_1}_{k^{th}\text{-tier}} + \ldots}_{\text{Indirect downstream}}$$

Figure 8.125: Downstream tree of the first- and second-tier rounds for the services sector
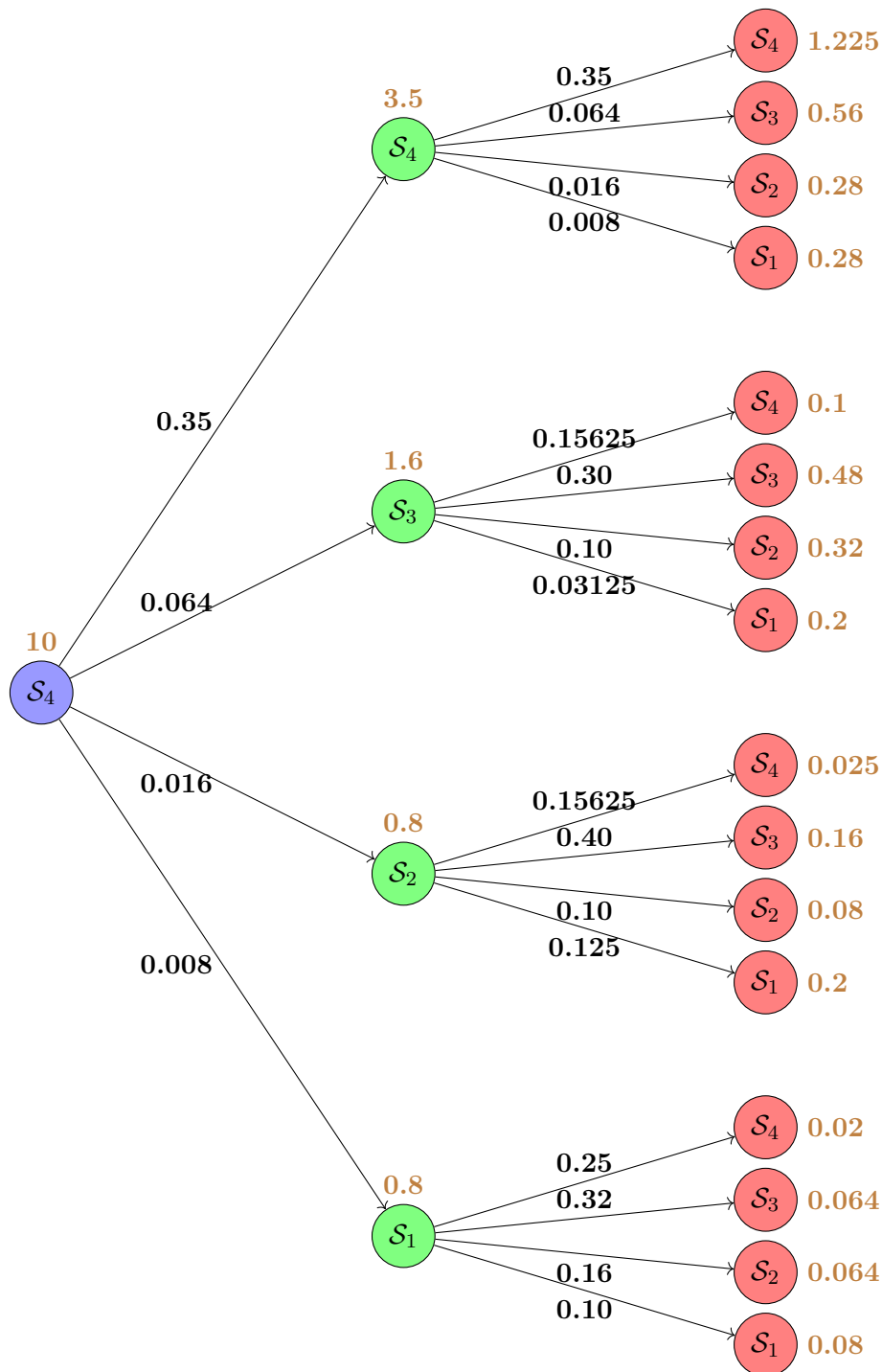
Again, we use the proportionality rule to calculate carbon emissions. We have:

$$\mathcal{CE}_{\text{total}}^{\text{down}} = \mathcal{CI}_{\text{total}}^{\text{down}} \odot \frac{\mathcal{CE}_1}{\mathcal{CI}_1}$$

and:

$$\mathcal{CE}_{\text{indirect}}^{\text{down}} = \mathcal{CE}_{\text{total}}^{\text{down}} - \mathcal{CE}_{\text{direct}}^{\text{down}} = \left(\mathcal{CI}_{\text{total}}^{\text{down}} - \mathcal{CI}_1 \right) \odot \frac{\mathcal{CE}_1}{\mathcal{CI}_1}$$

Using our previous example, the downstream matrix $\breve{A}$ is equal to:

$$
\begin{aligned}
\breve{A} &= \text{diag}\,(x)^{-1} Z \\
&= \begin{pmatrix}
0.10000 & 0.16000 & 0.32000 & 0.25000 \\
0.12500 & 0.10000 & 0.40000 & 0.15625 \\
0.03125 & 0.10000 & 0.30000 & 0.15625 \\
0.00800 & 0.01600 & 0.06400 & 0.35000
\end{pmatrix}
\end{aligned}
$$

We deduce the downstream multiplier matrix:

$$
\breve{\mathcal{L}} = \begin{pmatrix}
1.18811 & 0.31149 & 0.78705 & 0.72104 \\
0.20970 & 1.25525 & 0.86717 & 0.59085 \\
0.08938 & 0.20550 & 1.63035 & 0.47568 \\
0.02859 & 0.05497 & 0.19156 & 1.60872
\end{pmatrix}
$$

In Tables 8.32, 8.33 and 8.34, we report the downstream carbon intensities and emissions. These figures can be compared with those obtained in the case of the upstream analysis (Tables 8.29, 8.30 and 8.31).

Table 8.32: Direct and indirect downstream carbon intensities

Sector	$\mathcal{CI}_1$	$\mathcal{CI}_{\text{total}}^{\text{down}}$	$\mathcal{CI}_{\text{direct}}^{\text{down}}$	$\mathcal{CI}_{\text{indirect}}^{\text{down}}$	$\mathcal{CI}_{\text{direct}}^{\text{down}}$	$\mathcal{CI}_{\text{indirect}}^{\text{down}}$	$\dfrac{\mathcal{CI}_{\text{total}}^{\text{down}}}{\mathcal{CI}_1}$
		(in tCO$_2$e/\$ mn)			(in %)		
Energy	100.00	161.27	100.00	61.27	62.01%	37.99%	1.61
Materials	50.00	111.32	50.00	61.32	44.92%	55.08%	2.23
Industrials	25.00	64.73	25.00	39.73	38.62%	61.38%	2.59
Services	10.00	26.48	10.00	16.48	37.76%	62.24%	2.65

Table 8.33: Tier decomposition of downstream carbon intensities

	Sector	1	2	3	4	5	10	15	∞
$\mathcal{CI}_{(k)}^{\text{down}}$	Energy	28.50	14.68	8.00	4.45	2.48	0.14	0.00	0.00
	Materials	29.06	14.39	7.92	4.39	2.45	0.13	0.01	0.00
	Industrials	17.19	10.00	5.54	3.09	1.72	0.09	0.01	0.00
	Services	6.70	4.14	2.44	1.40	0.79	0.04	0.00	0.00
$\mathcal{CI}_{(1-k)}^{\text{down}}$	Energy	28.50	43.17	51.18	55.63	58.11	61.10	61.26	61.27
	Materials	29.06	43.45	51.37	55.76	58.21	61.15	61.31	61.32
	Industrials	17.19	27.19	32.73	35.82	37.54	39.61	39.72	39.73
	Services	6.70	10.84	13.27	14.67	15.46	16.43	16.48	16.48

We find that the results of the downstream analysis are different. While the energy and materials sectors have the lowest upstream indirect emissions, they have the highest downstream emissions.

This is due to the structure of the supply chain. Most of the output from the energy and materials sectors is destined to be used by the value chain to produce goods and services. In contrast, the industrials and services sectors require a lot of output from the value chain to produce goods and services directly. In this context, carbon emissions generally move downward for the energy and materials sectors, while they move upward for the industrials and services sectors. Finally, the indirect downstream emissions are equal to $1\,075$ ktCO$_2$e, while the indirect upstream emissions are equal to $1\,779$ ktCO$_2$e. To better understand the difference between the downstream and the upstream, we also plot the downstream tree of the first two tiers for the services sector in Figure 8.125. If we compare this tree with Figure 8.125, we see that the downstream trees grow to the right, while the upstream trees grow to the left.

Table 8.34: Decomposition of downstream carbon emissions

Sector	$\mathcal{CE}_{\text{direct}}^{\text{down}}$	$\mathcal{CE}_{\text{indirect}}^{\text{down}}$ (in ktCO$_2$e)	$\mathcal{CE}_{\text{total}}^{\text{down}}$	$\mathcal{CE}_{\text{direct}}^{\text{down}}$	$\mathcal{CE}_{\text{indirect}}^{\text{down}}$ (in %)	$\mathcal{CE}_{\text{total}}^{\text{down}}$
Energy	500	306.36	806.36	48.78	28.49	38.39
Materials	200	245.28	445.28	19.51	22.81	21.20
Industrials	200	317.83	517.83	19.51	29.55	24.65
Services	125	206.04	331.04	12.20	19.16	15.76
Total	$1\,025$	$1\,075.50$	$2\,100.50$	100.00	100.00	100.00

Remark 63 *As noted by Desnos et al. (2023), we must be careful with the upstream and downstream concepts of input-output analysis because they do not correspond to the upstream and downstream concepts of the GHG Protocol. There are several reasons for this. First, an input-output analysis does not distinguish between Scope 2 and Scope 3 emissions. Both are embedded in indirect emissions. Second, the GHG Protocol divides Scope 3 emissions into 8 upstream and 7 downstream categories. Downstream in the GHG Protocol refers to carbon emissions after goods and services have been produced. It includes their use by other sectors as well as final demand. Input-output analysis does not take into account downstream carbon emissions due to final demand (e.g., recycling or waste management). Thus, the downstream concept in input-output analysis is not consistent with the GHG Protocol definition. Third, we can see that there is a lot of double counting in the two analyses. For example, suppose that the matrix A is diagonal. In this case, we can show that $\mathcal{CE}_{\text{indirect}}^{\text{up}} = \mathcal{CE}_{\text{indirect}}^{\text{down}}$. In this particular case, the upstream and downstream analyses refer to the same carbon emissions, and we do not really know whether these emissions are upstream or downstream in the value chain.*

Mathematical properties

The mathematical properties derived in Section 8.4.1 on page 407 remain valid when we consider the upstream or downstream analysis of carbon intensities. Indeed, we reiterate that the properties require only that A is a doubly substochastic matrix. This is also the case for the matrices $A^{\top}$ (upstream multiplier matrix $\tilde{\mathcal{L}}$) and $\breve{A}$ (downstream multiplier matrix $\breve{\mathcal{L}}$). In particular, since the carbon intensities are not negative, the partial monotony property is satisfied. This means that there exists an index $k^{\star}$ such that $\mathcal{CI}_{(k)} \succeq \mathcal{CI}_{(k+1)}$ for the upstream analysis and $\mathcal{CI}_{(k)}^{\text{down}} \succeq \mathcal{CI}_{(k+1)}^{\text{down}}$ for the downstream analysis.

We recall that $\mathcal{CI}_{\text{total}} = \left(I_n - A^{\top}\right)^{-1}\mathcal{CI}_1 = \sum_{k=0}^{\infty}\left(A^{\top}\right)^k\mathcal{CI}_1 = \sum_{k=0}^{\infty}\mathcal{CI}_{(k)}$. Let $w_{(k)}$ be the relative contribution vector of the kth tier. We have:

$$w_{(k),j} = \frac{\mathcal{CI}_{(k),j}}{\sum_{h=0}^{\infty}\mathcal{CI}_{(k),j}}$$

Box 8.14: Calculating the upstreamness index

Let M be a square matrix of dimension $n \times n$. We have:

$$\frac{\partial \left(I_n - M\right)^{-1}}{\partial M} = \left(I_n - M\right)^{-1} \left(I_n - M\right)^{-1}$$

and:

$$\frac{\partial \sum_{k=0}^{\infty} M^k}{\partial M} = \sum_{k=0}^{\infty} k M^{k-1}$$

It follows that:

$$
\begin{aligned}
\sum_{k=0}^{\infty} k M^k &= M \sum_{k=0}^{\infty} k M^{k-1} \\
&= M \frac{\partial \sum_{k=0}^{\infty} M^k}{\partial M} \\
&= M \frac{\partial \left(I_n - M\right)^{-1}}{\partial M} \\
&= M \left(I_n - M\right)^{-1} \left(I_n - M\right)^{-1}
\end{aligned}
$$

Let z be a vector. Since we have $\left(I_n - M\right)^{-1} z = \sum_{k=0}^{\infty} M^k z = \sum_{k=0}^{\infty} z_{(k)}$ with $z_{(k)} = M^k z$, we deduce that:

$$\sum_{k=0}^{\infty} k \cdot z_{(k)} = \left(\sum_{k=0}^{\infty} k M^k\right) z = M \left(I_n - M\right)^{-1} \left(I_n - M\right)^{-1} z$$

The upstreamness index of sector j is then equal to:

$$\boldsymbol{\kappa}_j^{\mathrm{up}} := \frac{\left(\sum_{k=0}^{\infty} k \cdot z_{(k)}\right)_j}{\left(\sum_{k=0}^{\infty} z_{(k)}\right)_j} = \frac{\left(M \left(I_n - M\right)^{-1} \left(I_n - M\right)^{-1} z\right)_j}{\left(\left(I_n - M\right)^{-1} z\right)_j}$$

This expression is not exactly the formula proposed by Antràs *et al.* (2012), because they do not weight the tiers in the same way.

Following Antràs *et al.* (2012), we define the upstreamness index as the weighted average of the tiers with respect to their relative contributions:

$$\boldsymbol{\kappa}_j^{\mathrm{up}} = \sum_{k=0}^{\infty} k \cdot w_{(k),j} = 0 \times \frac{\boldsymbol{\mathcal{CI}}_{(0),j}}{\boldsymbol{\mathcal{CI}}_{\mathrm{total},j}} + 1 \times \frac{\boldsymbol{\mathcal{CI}}_{(1),j}}{\boldsymbol{\mathcal{CI}}_{\mathrm{total},j}} + 2 \times \frac{\boldsymbol{\mathcal{CI}}_{(2),j}}{\boldsymbol{\mathcal{CI}}_{\mathrm{total},j}} + \ldots = \frac{\left(\sum_{k=0}^{\infty} k \cdot \boldsymbol{\mathcal{CI}}_{(k)}\right)_j}{\left(\sum_{k=0}^{\infty} \boldsymbol{\mathcal{CI}}_{(k)}\right)_j}$$

In Box 8.14, we show that[111]:

$$\boldsymbol{\kappa}_j^{\mathrm{up}} = \frac{\left(A^\top \left(I_n - A^\top\right)^{-2} \boldsymbol{\mathcal{CI}}_1\right)_j}{\left(\left(I_n - A^\top\right)^{-1} \boldsymbol{\mathcal{CI}}_1\right)_j}$$

---

[111]We set $M = A^\top$ and $M = \breve{A}$, respectively.

and:

$$\kappa_j^{\text{down}} = \frac{\left(\breve{A} \left(I_n - \breve{A} \right)^{-2} \boldsymbol{\mathcal{CI}}_1 \right)_j}{\left(\left(I_n - \breve{A} \right)^{-1} \boldsymbol{\mathcal{CI}}_1 \right)_j}$$

If we consider our previous example, we obtain the following results:

		Energy	Materials	Industrials	Services
$M = A$	$z = y$	2.01	1.92	1.40	0.88
$M = A^\top$	$z = \boldsymbol{\mathcal{CI}}_1$	0.49	1.21	1.79	2.13
$M = \breve{A}$	$z = \boldsymbol{\mathcal{CI}}_1$	0.84	1.20	1.40	1.48

The upstreamness index of the energy sector is equal to 0.49, while its downstreamness index is equal to 0.84. In the case of the services sector, we obtain a higher upstreamness index ($\kappa_4^{\text{up}} = 2.13$). This indicates that the generation of carbon emissions in this sector affects higher tiers.

Comparison of upstream emissions between Exiobase, Trucost and WIOD

We report here some calculations made by Desnos *et al.* (2023). We recall that the total carbon intensity and emission vectors are equal to $\boldsymbol{\mathcal{CI}}_{\text{total}} = \left(I_n - A^\top \right)^{-1} \boldsymbol{\mathcal{CI}}_1$ and $\boldsymbol{\mathcal{CE}}_{\text{total}} = x \odot \boldsymbol{\mathcal{CI}}_{\text{total}}$. For the kth tier, the formulas become $\boldsymbol{\mathcal{CI}}_{(k)} = \left(A^\top \right)^k \boldsymbol{\mathcal{CI}}_1$ and $\boldsymbol{\mathcal{CE}}_{(k)} = x \odot \boldsymbol{\mathcal{CI}}_{(k)}$. The dimension of all these vectors is $n \times 1$, where n is the number of countries times the number of industries. For Trucost, carbon emissions and intensities are directly available. Direct emissions from input-output models can be compared to Scope 1 emissions of Trucost, while the total emissions correspond to Scope 1 plus Scope 2 plus upstream Scope 3 emissions from Trucost. We can also compare the direct plus first-tier indirect emissions of Trucost with the first-tier cumulative emissions $\boldsymbol{\mathcal{CE}}_{(0-1)}$ calculated from input-output models.

If we want to aggregate the results such that $i \in \Omega$, we have:

$$\boldsymbol{\mathcal{CE}}_{\text{total}} (\Omega) = \sum_{i \in \Omega} \boldsymbol{\mathcal{CE}}_{\text{total},i} = \omega^\top \boldsymbol{\mathcal{CE}}_{\text{total}}$$

where $\omega = (\omega_i)$ is a vector of dimension $n \times 1$ with $\omega_i = 1$ if $i \in \Omega$ and $\omega_i = 0$ otherwise. We deduce that the carbon intensity of Ω is equal to:

$$\boldsymbol{\mathcal{CI}}_{\text{total}} (\Omega) = \frac{\sum_{i \in \Omega} \boldsymbol{\mathcal{CE}}_{\text{total},i}}{\sum_{i \in \Omega} x_i} = \frac{\sum_{i \in \Omega} x_i \boldsymbol{\mathcal{CI}}_{\text{total},i}}{\sum_{i \in \Omega} x_i} = \sum_{i \in \Omega} w_i \boldsymbol{\mathcal{CI}}_{\text{total},i}$$

where $w_i = \left(\sum_{j \in \Omega} x_j \right)^{-1} x_i$ is the weight of item i in the set Ω. The carbon intensity of Ω is then equal to its weighted average carbon intensity (WACI). These calculations also apply to $\boldsymbol{\mathcal{CE}}_{(k)}$, $\boldsymbol{\mathcal{CE}}_{(0:k)}$, $\boldsymbol{\mathcal{CI}}_{(k)}$ and $\boldsymbol{\mathcal{CI}}_{(0:k)}$.

We perform a global analysis by setting $\omega = \mathbf{1}_n$. In Table 8.35, we report the multiplication coefficients $m_{(k)} = \boldsymbol{\mathcal{CE}}_{(k)} / \boldsymbol{\mathcal{CE}}_1$ and $m_{(0-k)} = \boldsymbol{\mathcal{CE}}_{(0-k)} / \boldsymbol{\mathcal{CE}}_1$, and we also compute the contribution ratio $c_{(0-k)} = \boldsymbol{\mathcal{CE}}_{(0-k)} / \boldsymbol{\mathcal{CE}}_{\text{total}}$. For the WIOD table, the direct plus indirect emissions are 3.14 times the Scope 1 emissions, which means that the indirect emissions are more than twice the direct emissions. In the case of the Exiobase tables, the ratio $m_{(0-\infty)}$ is equal to 2.76 in 2014 and 2.75 in 2022. We note that the convergence is rapid, as more than 90% of the total emissions are located in the first five tiers.

Table 8.35: Ratio of upstream carbon emissions (global analysis)

Tier	WIOD 2014			Exiobase 2014			Exiobase 2022		
	$m_{(k)}$	$m_{(0-k)}$	$c_{(0-k)}$	$m_{(k)}$	$m_{(0-k)}$	$c_{(0-k)}$	$m_{(k)}$	$m_{(0-k)}$	$c_{(0-k)}$
0	1.00	1.00	31.8%	1.00	1.00	36.2%	1.00	1.00	36.4%
1	0.77	1.76	56.1%	0.72	1.72	62.5%	0.73	1.73	62.9%
2	0.50	2.26	71.9%	0.43	2.15	78.0%	0.42	2.15	78.3%
3	0.32	2.58	82.1%	0.25	2.40	87.0%	0.25	2.40	87.3%
4	0.20	2.78	88.6%	0.15	2.55	92.3%	0.14	2.54	92.5%
5	0.13	2.91	92.7%	0.09	2.63	95.5%	0.08	2.62	95.5%
6	0.08	3.00	95.4%	0.05	2.69	97.3%	0.05	2.67	97.3%
7	0.05	3.05	97.0%	0.03	2.72	98.4%	0.03	2.70	98.4%
8	0.03	3.08	98.1%	0.02	2.73	99.0%	0.02	2.72	99.0%
9	0.02	3.11	98.8%	0.01	2.74	99.4%	0.01	2.73	99.4%
10	0.01	3.12	99.2%	0.01	2.75	99.7%	0.01	2.74	99.7%
∞	0.00	3.14	100.0%	0.00	2.76	100.0%	0.00	2.75	100.0%

Source: Desnos *et al.* (2023, Table 15, page 52).

Figure 8.126: Multiplication coefficient $m_{(0-1)}$ and $m_{(0-\infty)}$ (global analysis)
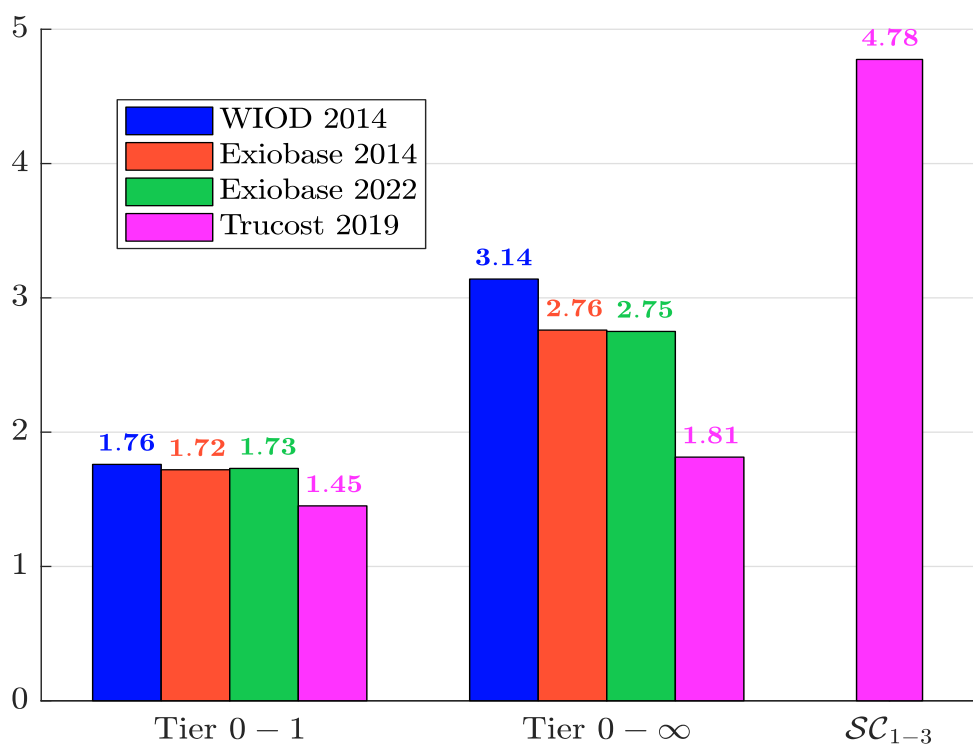
Source: Desnos *et al.* (2023, Figure 25, page 53).

In the case of Trucost, we can only compute $m_{(0-1)}$ and $m_{(0-\infty)}$. Results are shown in Figure 8.126. We notice that the multiplication coefficients obtained with Trucost are smaller than those computed with input-output models (1.45–1.54 versus 1.7 for the first tier). Nevertheless, the multiplication coefficient is very high when we integrate Scope 3 emissions since we obtain a value of 4.78 in 2019 and 6.75 in 2021.

Desnos *et al.* (2023) also suggested a country analysis[112]. They observed some large differences from one country to another. For instance, Figure 8.127 shows the multiplication coefficient $m_{(0-\infty)}$ of the different countries. The lowest value is obtained for the USA ($m_{(0-\infty)} = 2.19$), while the highest factor is observed for Switzerland ($m_{(0-\infty)} = 7.21$).

Figure 8.127: Multiplying coefficient $m_{(0-\infty)}$ (country analysis, WIOD 2014)

Desnos *et al.* (2023) applied the previous framework to estimate the upstream of the MSCI World index. They first estimated the total carbon intensity of all issuers in the portfolio:

$$\mathcal{CI}_{\text{total},i} = \mathcal{CI}_{1,i}^{\text{reported}} + \mathcal{CI}_{\text{indirect},i}^{\text{estimated}}$$

where $\mathcal{CI}_{1,i}^{\text{reported}}$ is the Scope 1 carbon intensity reported by the issuer i and $\mathcal{CI}_{\text{indirect},i}^{\text{estimated}}$ is the estimated indirect carbon intensity. In the case of Trucost, they used the values estimated by the data provider. For the input-output databases, they used the formula $\mathcal{CI}_{\text{indirect}}^{\text{estimated}} = \left(\left(I_n - A^\top\right)^{-1} - I_n\right)\mathcal{CI}_{\text{direct}}$ and considered the row corresponding to the sector and the country of the issuer i. They then aggregated the carbon intensity of the MSCI World Index at the GICS level 1 sectors. The results are shown in Table 8.36 and Figure 8.128. The direct plus indirect intensity of

---

[112]In this case, $\omega_i = 1$ if i belongs to the country and $\omega_i = 0$ otherwise.

the MSCI World Index is equal to 299 tCO$_2$e/\$ mn with Exiobase 2022, 281 tCO$_2$e/\$ mn with Trucost 2021 and 278 tCO$_2$e/\$ mn with WIOD 2014. The difference between the lowest and the highest value is then equal to 7.5%, which is a low figure. If we look at the GICS sectors, the differences are more significant, especially for consumer staples, energy and materials. For instance, the carbon intensity of the energy sector is equal to 757 tCO$_2$e/\$ mn with WIOD 2014 and 1 373 tCO$_2$e/\$ mn with Exiobase 2022. They also calculated the contribution of each sector to the carbon intensity of the MSCI World Index:

$$c_j(w) = \frac{\sum_{i \in \boldsymbol{S}ector_j} w_i \cdot \boldsymbol{C\mathcal{I}}_{\text{total},i}}{\sum_i w_i \cdot \boldsymbol{C\mathcal{I}}_{\text{total},i}}$$

where w is the vector of weights in the MSCI World Index and $c_j(w)$ is the contribution of the jth sector. In Table 8.37, we see some significant differences. This is true for the previously mentioned sectors (consumer staples, energy and materials), but also for consumer discretionary, health care and information technology.

Table 8.36: Direct + indirect carbon intensities of GICS sectors (MSCI World index, May 2023)

Sector	Exiobase 2022	Trucost 2021	WIOD 2014
Communication Services	66	78	102
Consumer Discretionary	168	209	219
Consumer Staples	437	387	277
Energy	1 373	796	757
Financials	83	55	83
Health Care	108	120	167
Industrials	276	277	307
Information Technology	110	138	131
Materials	791	973	747
Real Estate	128	134	138
Utilities	1 872	1 833	1 889
MSCI World	299	281	278

Source: Desnos *et al.* (2023, Table 57, page 178).

Table 8.37: Breakdown of the portfolio intensity by GICS sector (MSCI World Index, May 2023)

Sector	Exiobase 2022	Trucost 2021	WIOD 2014
Communication Services	1.5%	1.9%	2.5%
Consumer Discretionary	5.9%	7.8%	8.4%
Consumer Staples	11.6%	10.9%	7.9%
Energy	22.9%	14.1%	13.6%
Financials	4.2%	2.9%	4.5%
Health Care	4.8%	5.7%	8.0%
Industrials	10.2%	10.8%	12.1%
Information Technology	7.5%	10.0%	9.6%
Materials	11.7%	15.3%	11.9%
Real Estate	1.1%	1.2%	1.2%
Utilities	18.6%	19.4%	20.2%

Source: Desnos *et al.* (2023, Table 21, page 62).

Figure 8.128: Total carbon intensity $\mathcal{CI}_{\text{total}}$ by GICS sector (MSCI World Index, May 2023)
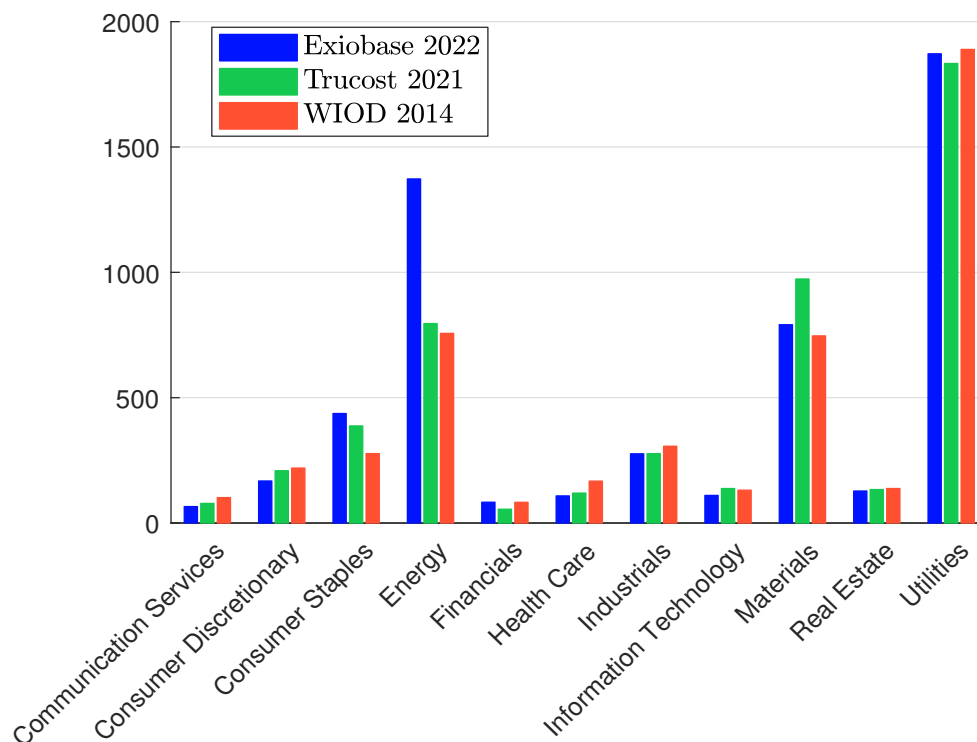
8.4.4 Imported and exported carbon emissions

On page 419, we have analyzed the difference in carbon footprint when we use a production-based or consumption-based inventory. We can extend this framework by considering multiple regions. In this case, we can adjust a country's carbon footprint for imports and exports. For instance, it may be unfair to attribute the CO_2 emissions of a good to a country if that country exports 100% of that good to the rest of the world. The concept of imported and exported carbon emissions is then central when we would like to measuring a country's carbon footprint, which reflects the consumption and lifestyle choices of its citizens (Turner *et al.*, 2007; Wiedmann *et al.*, 2007). The groundbreaking research paper by Peters and Hertwich (2008) has been followed by a large number of studies[113]. They all conclude that there is a large bias when focusing on territorial emissions due to production.

Calculation of balanced emissions

Let's learn how to calculate the CO_2 embodied in international trade by using the input-output framework. We have:

- $y_{j,r}$ is the final demand of the jth sector and the rth region.

- $y = (y_1, \dots, y_n)$ is the vector of final demand where $y_j = \sum_{r=1}^{p} y_{j,r}$ is the final demand of the jth sector.

---

[113]For example, we can cite the works of Davis and Caldeira (2010), Peters *et al.* (2011), Yamano and Guilhoto (2020), Lamb *et al.* (2021) and Friedlingstein *et al.* (2022).

- $C^{(x)} = \left(C_{g,j}^{(x)}\right)$ is the pollution output matrix where $C_{g,j}^{(x)}$ is the total amount of the gth pollutant generated by the output of the jth sector.

- $D^{(y)} = C^{(x)}\operatorname{diag}(x)^{-1} = \left(D_{g,j}^{(y)}\right)$ is the matrix of direct impact coefficients where $D_{g,j}^{(y)} = c_{g,j}^{(y)}/x_j$ is the amount of the gth pollutant generated by 1\$ of the output of the jth sector.

- $\varpi = (\varpi_1, \ldots, \varpi_m)$ is the vector of pollution level. We have:

$$\varpi = D^{(x)}x = D^{(x)}\left(I_n - A\right)^{-1}y = D^{(y)}y$$

where $D^{(y)} = D^{(x)}\left(I_n - A\right)^{-1}$ is the pollutant multiplier matrix with respect to the final demand y.

In the production-based approach, the total contribution of the rth to the gth pollutant is equal to:

$$C_g^{(x,r)} = \sum_{j \in r} C_{g,j}^{(x)} \tag{8.42}$$

In the consumption-based approach, we have the following identity:

$$\varpi_g = \left(D^{(y)}y\right)_g = \sum_{j=1}^n D_{g,j}^{(y)}y_j = \sum_{j=1}^n D_{g,j}^{(y)}\sum_{r=1}^p y_{j,r} = \sum_{r=1}^p \sum_{j=1}^n D_{g,j}^{(y)}y_{j,r}$$

we deduce that the total contribution of the rth region to the gth pollutant is equal to:

$$
\begin{aligned}
C_g^{(y,r)} &= \sum_{j=1}^n \frac{\partial\,\varpi_g}{\partial\,y_{j,r}}y_{j,r} \\
&= \sum_{j=1}^n D_{g,j}^{(y)}y_{j,r} \tag{8.43}
\end{aligned}
$$

Example 24 *We consider two sectors ($\mathcal{S}_1$ and $\mathcal{S}_2$) and three regions ($\mathcal{R}_1$, $\mathcal{R}_2$ and $\mathcal{R}_3$), whose input-output table is given below:*

		Z						y			x
		$\mathcal{R}_1$		$\mathcal{R}_2$		$\mathcal{R}_3$		$\mathcal{R}_1$	$\mathcal{R}_2$	$\mathcal{R}_3$	
		$\mathcal{S}_1$	$\mathcal{S}_2$	$\mathcal{S}_1$	$\mathcal{S}_2$	$\mathcal{S}_1$	$\mathcal{S}_2$				
$\mathcal{R}_1$	$\mathcal{S}_1$	100	300	10	10	20	0	500	200	25	1 165
	$\mathcal{S}_2$	250	150	20	0	10	0	800	100	17	1 347
$\mathcal{R}_2$	$\mathcal{S}_1$	10	10	110	310	0	0	20	200	15	675
	$\mathcal{S}_2$	20	20	80	25	15	20	0	200	5	385
$\mathcal{R}_3$	$\mathcal{S}_1$	10	5	8	3	40	7	5	25	50	153
	$\mathcal{S}_2$	5	2	8	8	12	35	3	50	50	173
V		770	860	439	29	56	111				
x		1 165	1 347	675	385	153	173				
GHG	CO_2	50 000	20 000	10 000	10 000	5 000	5 000	100 000			
	CH_4	5 000	3 000	0	0	1 000	1 000	10 000			

Z, y, x and V are expressed in \$ mn, while the GHG are calculated in tonne. Final demand is split between the three regions. For instance, x_1 is equal to \$1 165 mn, its intermediary demand is equal to $100+300+10+10+20+0$ or \$440 mn and its final demand is equal to \$725 mn with the following breakdown: 500 for the region $\mathcal{R}_1$, 200 for the region $\mathcal{R}_2$ and 25 for the region $\mathcal{R}_3$.

If we take a production-based approach, x_1 and x_2 are in region $\mathcal{R}_1$, which means that this region is responsible for $70\,000$ tCO$_2$ and $8\,000$ tCH$_4$. Finally, we obtain the following distribution among the three regions:

	Absolute			Relative		
	$\mathcal{R}_1$	$\mathcal{R}_2$	$\mathcal{R}_3$	$\mathcal{R}_1$	$\mathcal{R}_2$	$\mathcal{R}_3$
CO$_2$	70 000	20 000	10 000	70%	20%	10%
CH$_4$	8 000	0	2 000	80%	0%	20%

We have:

$$A = Z \operatorname{diag}(x)^{-1} = \begin{pmatrix} 0.0858 & 0.2227 & 0.0148 & 0.0260 & 0.1307 & 0.0000 \\ 0.2146 & 0.1114 & 0.0296 & 0.0000 & 0.0654 & 0.0000 \\ 0.0086 & 0.0074 & 0.1630 & 0.8052 & 0.0000 & 0.0000 \\ 0.0172 & 0.0148 & 0.1185 & 0.0649 & 0.0980 & 0.1156 \\ 0.0086 & 0.0037 & 0.0119 & 0.0078 & 0.2614 & 0.0405 \\ 0.0043 & 0.0015 & 0.0119 & 0.0208 & 0.0784 & 0.2023 \end{pmatrix}$$

and:

$$\mathcal{L} = (I_6 - A)^{-1} = \begin{pmatrix} 1.1674 & 0.2953 & 0.0454 & 0.0741 & 0.2450 & 0.0232 \\ 0.2848 & 1.1989 & 0.0586 & 0.0601 & 0.1663 & 0.0171 \\ 0.0482 & 0.0445 & 1.3705 & 1.1871 & 0.1893 & 0.1816 \\ 0.0351 & 0.0320 & 0.1817 & 1.2327 & 0.1927 & 0.1884 \\ 0.0167 & 0.0109 & 0.0263 & 0.0362 & 1.3706 & 0.0748 \\ 0.0101 & 0.0064 & 0.0280 & 0.0538 & 0.1442 & 1.2687 \end{pmatrix}$$

The direct impact matrix is equal to the GHG emissions divided by the output:

$$\begin{aligned} D^{(x)} &= \begin{pmatrix} 50\,000/1\,165 & 20\,000/1\,347 & 10\,000/675 & 10\,000/385 & 5\,000/153 & 5\,000/173 \\ 5\,000/1\,165 & 3\,000/1\,347 & 0/675 & 0/385 & 1\,000/153 & 1\,000/173 \end{pmatrix} \\ &= \begin{pmatrix} 42.9185 & 14.8478 & 14.8148 & 25.9740 & 32.6797 & 28.9017 \\ 4.2918 & 2.2272 & 0.0000 & 0.0000 & 6.5359 & 5.7803 \end{pmatrix} \end{aligned}$$

The unit of $D^{(x)}$ is expressed in tonnes of the gas per million dollars. The carbon intensity of the sector $\mathcal{S}_1$ in the region $\mathcal{R}_1$ is equal to 42.9185 tCO$_2$/\$ mn, while the methane intensity is equal to 4.2918 tCH$_4$/\$ mn. We have:

$$D^{(y)} = D^{(x)}\mathcal{L} = \begin{pmatrix} 56.7948 & 32.5033 & 29.5091 & 56.4128 & 69.7564 & 47.9460 \\ 5.8121 & 4.0453 & 0.6591 & 0.9993 & 11.2142 & 7.9600 \end{pmatrix}$$

Using Equation (8.43), we finally obtain the following decomposition:

	Absolute			Relative		
	$\mathcal{R}_1$	$\mathcal{R}_2$	$\mathcal{R}_3$	$\mathcal{R}_1$	$\mathcal{R}_2$	$\mathcal{R}_3$
CO$_2$	55 483	35 935	8 582	55.5%	35.9%	8.6%
CH$_4$	6 235	2 577	1 188	62.4%	25.8%	11.9%

In the consumption-based approach, the region $\mathcal{R}_1$ is responsible for 55.5% of carbon emissions, which is lower than the 70% figure obtained in the production-based approach. The second region sees a high increase in carbon emissions of 79.5%.

Let us now look at how the concept of imported and exported emissions is defined. In Table 8.38, we report the allocation of carbon emissions to intermediate and final uses. For intermediate consumption, the allocation to $Z_{i,j}$ is equal to:

$$\mathcal{CE}_i(Z_{i,j}) = \frac{Z_{i,j}}{x_i} \cdot \mathcal{CE}_i$$

while we have for the final consumption:

$$\mathcal{CE}_i\left(y_{i,r}\right) = \frac{y_{i,r}}{x_i} \cdot \mathcal{CE}_i$$

By construction, we verify the equality:

$$\sum_{j=1}^{n} \mathcal{CE}_i\left(Z_{i,j}\right) + \sum_{r=1}^{p} \mathcal{CE}_i\left(y_{i,r}\right) = \left(\sum_{j=1}^{n} \frac{Z_{i,j}}{x_i} + \sum_{r=1}^{p} \frac{y_{i,r}}{x_i}\right) \mathcal{CE}_i = \mathcal{CE}_i$$

Table 8.38: Allocation of carbon emissions to intermediate and final uses (production-based approach)

	Z						y			Total
	$\mathcal{R}_1$	$\mathcal{R}_1$	$\mathcal{R}_2$	$\mathcal{R}_2$	$\mathcal{R}_3$	$\mathcal{R}_3$	$\mathcal{R}_1$	$\mathcal{R}_2$	$\mathcal{R}_3$	
$(\mathcal{S}_1, \mathcal{R}_1)$	4 292	12 876	429	429	858	0	21 459	8 584	1 073	50 000
$(\mathcal{S}_2, \mathcal{R}_1)$	3 712	2 227	297	0	148	0	11 878	1 485	252	20 000
$(\mathcal{S}_1, \mathcal{R}_2)$	148	148	1 630	4 593	0	0	296	2 963	222	10 000
$(\mathcal{S}_2, \mathcal{R}_2)$	519	519	2 078	649	390	519	0	5 195	130	10 000
$(\mathcal{S}_1, \mathcal{R}_3)$	327	163	261	98	1 307	229	163	817	1 634	5 000
$(\mathcal{S}_2, \mathcal{R}_3)$	145	58	231	231	347	1 012	87	1 445	1 445	5 000
Total	9 143	15 992	4 926	6 000	3 050	1 760	33 884	20 488	4 757	100 000

From Table 8.38, we can aggregate the columns to obtain the regional carbon allocation (Table 8.39). For instance, the carbon allocation to sector $\mathcal{S}_1$ and region $\mathcal{R}_1$ is equal to:

$$4\,292 + 12\,876 + 21\,459 = 38\,627 \text{ tCO}_2$$

Table 8.39: Carbon allocation by region (production-based approach, column aggregation)

	$\mathcal{R}_1$	$\mathcal{R}_2$	$\mathcal{R}_3$	Total
$(\mathcal{S}_1, \mathcal{R}_1)$	38 627	9 442	1 931	50 000
$(\mathcal{S}_2, \mathcal{R}_1)$	17 817	1 782	401	20 000
$(\mathcal{S}_1, \mathcal{R}_2)$	593	9 185	222	10 000
$(\mathcal{S}_2, \mathcal{R}_2)$	1 039	7 922	1 039	10 000
$(\mathcal{S}_1, \mathcal{R}_3)$	654	1 176	3 170	5 000
$(\mathcal{S}_2, \mathcal{R}_3)$	289	1 908	2 803	5 000
Total	59 018	31 415	9 567	100 000

Table 8.40: Carbon allocation by region (production-based approach, row aggregation)

	$\mathcal{R}_1$	$\mathcal{R}_2$	$\mathcal{R}_3$	Total
$\mathcal{R}_1$	56 444	11 224	2 332	70 000
$\mathcal{R}_2$	1 632	17 107	1 261	20 000
$\mathcal{R}_3$	943	3 084	5 973	10 000
Total	59 018	31 415	9 567	100 000

By aggregating the rows, we finally obtain the allocation given in Table 8.40. For example, the carbon allocation of goods produced in region $\mathcal{R}_1$ and consumed in region $\mathcal{R}_1$ is equal to:

$$3\,8627 + 17\,817 = 56\,444 \text{ tCO}_2$$

The carbon allocation of goods produced in region $\mathcal{R}_1$ and consumed in region $\mathcal{R}_2$ is equal to:

$$9\,442 + 1\,782 = 11\,224 \text{ tCO}_2$$

The row shows the allocation split of the emissions produced in a given region. The column shows the allocation split of the emissions consumed in a given region. For example, among the $70\,000$ tCO$_2$ emitted in region $\mathcal{R}_1$, $11\,224$ and $2\,332$ are exported in regions $\mathcal{R}_2$ and $\mathcal{R}_3$. Therefore, $56\,444$ tCO$_2$ remain in region $\mathcal{R}_1$. Since $\mathcal{R}_1$ also imports $1\,632$ tCO$_2$ from region $\mathcal{R}_2$ and 943 tCO$_2$ from region $\mathcal{R}_3$, the total carbon emissions allocated to region $\mathcal{R}_1$ is equal to $59\,018$ tCO$_2$. A summary of the allocation process is given below:

	$\mathcal{R}_1$	$\mathcal{R}_2$	$\mathcal{R}_3$	Total
Produced emissions	70 000	20 000	10 000	100 000
− Exported emissions	13 556	2 893	4 027	20 475
= Intermediate emissions	56 444	17 107	5 973	79 525
+ Imported emissions	2 574	14 308	3 593	20 475
= Regional emissions	59 018	31 415	9 567	100 000

By construction, exported emissions are equal to imported emissions at the aggregate level, but not at the regional level. In the case of region $\mathcal{R}_1$, we have the following results:

$$\underbrace{70\,000}_{\text{produced}} - \underbrace{13\,556}_{\text{exported}} + \underbrace{2\,574}_{\text{imported}} = \underbrace{59\,018}_{\text{allocated}} \neq \underbrace{55\,483}_{\text{consumed}}$$

Note that we do not get the previous figure of $55\,483$ tCO$_2$ emissions. The reason for this is that the above analysis is based on the tier-one decomposition and does not include the full value chain relationships.

Table 8.41: Carbon allocation by region (consumption-based approach, column aggregation)

	$\mathcal{R}_1$	$\mathcal{R}_2$	$\mathcal{R}_3$	Total	Produced emissions	EEIO correction
$(\mathcal{S}_1, \mathcal{R}_1)$	28 397	11 359	1 420	41 176	50 000	−8 824
$(\mathcal{S}_2, \mathcal{R}_1)$	26 003	3 250	553	29 806	20 000	9 806
$(\mathcal{S}_1, \mathcal{R}_2)$	590	5 902	443	6 935	10 000	−3 065
$(\mathcal{S}_2, \mathcal{R}_2)$	0	11 283	282	11 565	10 000	1 565
$(\mathcal{S}_1, \mathcal{R}_3)$	349	1 744	3 488	5 581	5 000	581
$(\mathcal{S}_2, \mathcal{R}_3)$	144	2 397	2 397	4 938	5 000	−62
Total	55 483	35 935	8 582	100 000	100 000	0

Table 8.42: Carbon allocation by region (consumption-based approach, row aggregation)

	$\mathcal{R}_1$	$\mathcal{R}_2$	$\mathcal{R}_3$	Total
$\mathcal{R}_1$	54 400	14 609	1 972	70 982
$\mathcal{R}_2$	590	17 184	725	18 499
$\mathcal{R}_3$	493	4 141	5 885	10 519
Total	55 483	35 935	8 582	100 000

In order to take into account the full value chain, we proceed as previously by considering the consumption-based emissions $D_{g,j}^{(y)} y_{j,r}$ instead of the production-based emissions $C_{g,j}^{(x)}$. We obtain

the results shown in Table 8.41. We can then compute the EEIO correction $cr_{g,j}^{(y)}$, which is defined by the following equation:

$$\sum_{r=1}^{p} D_{g,j}^{(y)} y_{j,r} = C_{g,j}^{(x)} + cr_{g,j}^{(y)}$$

For example, it is equal to $-8\,824$ tCO_2 for $(\mathcal{S}_1, \mathcal{R}_1)$. We then perform the row aggregation (Table 8.42), and we finally obtain the following allocation results:

	$\mathcal{R}_1$	$\mathcal{R}_2$	$\mathcal{R}_3$	Total
Produced emissions	70 000	20 000	10 000	100 000
+ EEIO correction	982	−1 501	519	0
= Consumed emissions	70 982	18 499	10 519	100 000
− Exported emissions	16 582	1 315	4 634	22 530
= Intermediate emissions	54 400	17 184	5 885	77 470
+ Imported emissions	1 083	18 751	2 697	22 530
= Regional emissions	55 483	35 935	8 582	100 000
Net imported emissions	−14 517	15 935	−1 418	0

We can define the concept of net imported emissions, which takes into account the EEIO correction, imported and exported emissions. It is the difference between the emissions due to consumption and the emissions due to production. In our example, $\mathcal{R}_1$ and $\mathcal{R}_3$ are net exporting regions, while $\mathcal{R}_2$ is a net importing region.

Remark 64 *The previous formulas can be easily translated using the traditional notations. Let $\boldsymbol{CI}_1 = (\boldsymbol{CI}_{1,1}, \ldots, \boldsymbol{CI}_{n,1})$ be the vector of carbon intensities evaluated in CO_2e, where $\boldsymbol{CI}_{j,1}$ measures the direct emission intensity of sector j. The vector of consumption-based carbon intensities is equal to:*

$$\boldsymbol{CI}^{(y)} = \tilde{\boldsymbol{\mathcal{L}}} \boldsymbol{CI}_1 := \boldsymbol{CI}_{\text{total}} \tag{8.44}$$

Let $y = (y_{j,r})$ be the $n \times p$ matrix, where $y_{j,r}$ is the final demand of the jth sector and the rth region. We have:

$$\boldsymbol{C\mathcal{E}}^{(y)} = \boldsymbol{CI}^{(y)} \odot y \tag{8.45}$$

$\boldsymbol{C\mathcal{E}}^{(y)}$ is the $n \times p$ matrix of carbon emissions. The consumption-based carbon emissions of the rth region is then equal to:

$$\boldsymbol{C\mathcal{E}}^{(y,r)} = \mathbf{1}_n^{\top} \left(\boldsymbol{C\mathcal{E}}^{(y)} \mathbf{e}_r \right) \tag{8.46}$$

while the imported and exported carbon emissions of the rth region are:

$$\boldsymbol{C\mathcal{E}}_{\text{imported}}^{(y,r)} = \sum_{j \notin r} \left(\boldsymbol{C\mathcal{E}}^{(y)} \mathbf{e}_r \right)_j \tag{8.47}$$

and:

$$\boldsymbol{C\mathcal{E}}_{\text{exported}}^{(y,r)} = \sum_{j \in r} \sum_{k \neq r} \left(\boldsymbol{C\mathcal{E}}^{(y)} \mathbf{e}_r \right)_j \tag{8.48}$$

Equations (8.44–8.48) are the basic formulas for calculating imported and exported carbon emissions. We can also easily define consumed, intermediate and net imported emissions and the EEIO correction at both the sectoral and regional levels.

Stylized facts

Figure 8.129 shows a global picture of carbon emissions embedded in trade, which can be found at https://ourworldindata.org/consumption-based-co2. The data are taken from the famous *Global Carbon Budget* (2022) report (Friedlingstein *et al.*, 2022). The exported (blue color) or imported (red color) emissions are normalized to domestic production emissions. For example, a value of 100% indicates a country whose imported emissions are equal to its domestic production emissions. We notice that the European countries and the United States are colored red, which means that they are net importers of CO_2. On the contrary, the BRICS countries[114] are net exporters of CO2, as are Australia, Canada and the Middle East region.

Figure 8.129: CO_2 emissions embedded in trade, 2020
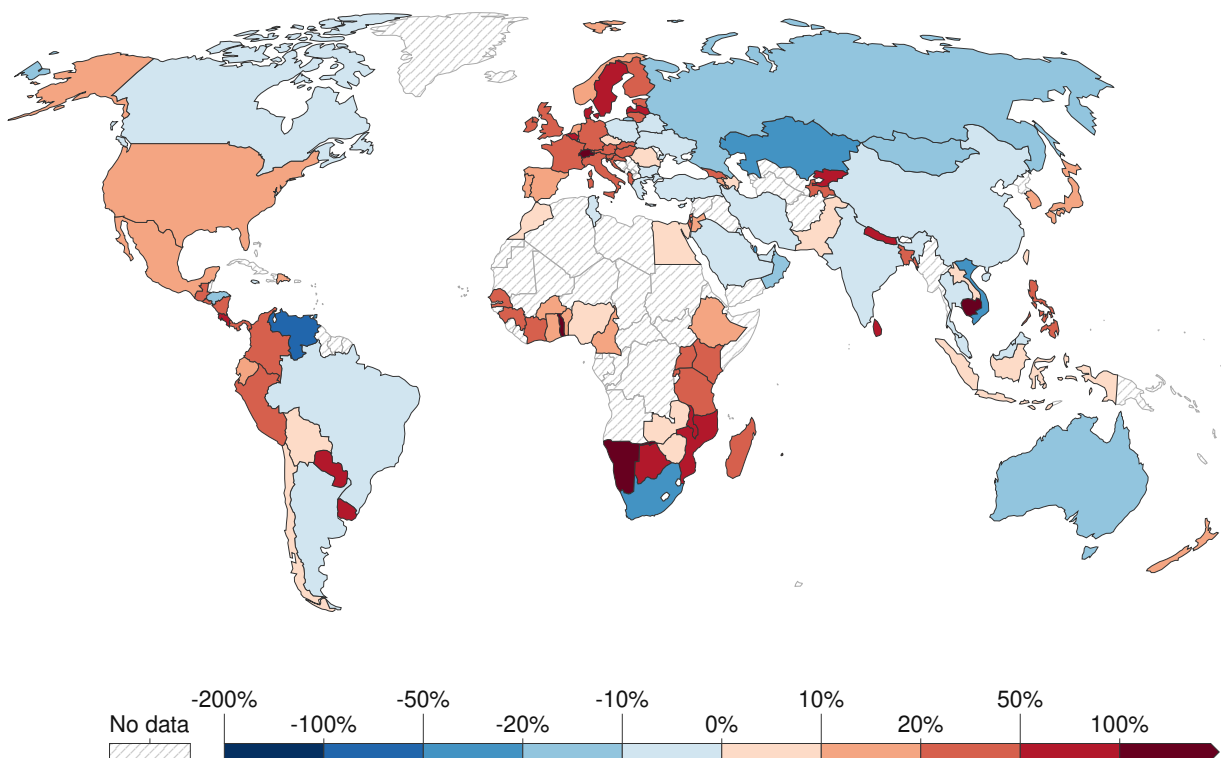
This is measured as emissions exported or imported as a percentage of domestic production emissions. Positive values (red) represent net importers of CO_2. Negative values (blue) represent net exporters of CO_2.

Source: https://ourworldindata.org/consumption-based-co2.

The OECD maintains a database of imported and exported CO_2 at www.oecd.org/industry/ind/carbondioxideemissionsembodiedininternationaltrade.htm. Calculations are detailed in Yamano and Guilhoto (2020). We use an updated version of these data, which can be obtained from the OECD statistics website[115]. In our case, we are interested in three variables:

- FD_CO2: CO2 emissions embodied in domestic final demand, by source country and industry;

- PROD_CO2: CO2 emissions based on production;

---

[114]Brazil, Russia, India, China and South Africa.

[115]The exact link is https://stats.oecd.org/Index.aspx?DataSetCode=IO_GHG_2021.

- BALCO2_FD: CO2 embodied in final demand, balance.

The first variable is the consumption-based emissions, the second variable is the production-based emissions, while the third variable is the difference between the previous two: BALCO2_FD = PROD_CO2 − FD_CO2. Using our notations, the GHG balance is equal to:

$$\mathcal{CE}^{(x-y,r)} = \mathcal{CE}^{(x,r)} - \mathcal{CE}^{(y,r)}$$

We can choose a specific industry or all activities (DTOTAL) and also the partner for the trade analysis. If the partner is WLD, we get the emissions of each country/region with respect to the world. If the partner is CHN, we obtain the emissions of each country/region with respect to China. In Table 8.43, we report the top importing and exporting countries. In 2018, the largest importer of carbon emissions is the United States, followed by Japan and Germany, while the largest exporter of carbon emissions is China, followed by Russia and South Africa. In the first case the GHG balance is negative, in the second case it is positive. In a sense, these rankings show a contrast between developed and emerging economies.

Table 8.43: Top importing and exporting countries by carbon emissions (in MtCO$_2$e, 2018)

		Top importers				Top exporters	
Rank	ISO	Country	Balance	Rank	ISO	Country	Balance
1	USA	United States	−752.10	1	CHN	China	895.45
2	JPN	Japan	−160.62	2	RUS	Russian Federation	343.48
3	DEU	Germany	−128.73	3	ZAF	South Africa	122.47
4	GBR	United Kingdom	−123.77	4	IND	India	106.10
5	FRA	France	−111.65	5	TWN	Chinese Taipei	77.03
6	ITA	Italy	−80.09	6	SGP	Singapore	62.19
7	HKG	Hong Kong, China	−70.14	7	KOR	Korea	54.35
8	CHE	Switzerland	−44.53	8	CAN	Canada	53.12
9	PHL	Philippines	−40.49	9	VNM	Viet Nam	52.31
10	SWE	Sweden	−29.67	10	MYS	Malaysia	46.52

Source: Yamano and Guilhoto (2020), https://stats.oecd.org & Author's calculations.

These rankings also show a divide between OECD and non-OECD countries. In Figure 8.130, we reproduce Figure 4-1 on page 24 of the OECD research report written by Yamano and Guilhoto (2020). The comparison between aggregate OECD and aggregate non-OECD production- and consumption-based emissions shows two different dynamics. For OECD countries, the long-term trend between 1995 and 2018 is quite stable. On the contrary, there is a general increase in emissions from non-OECD economies over the same period.

> "*These increases are in great part linked with the need of these countries to sustain their own development and to improve the quality of life of their population. A consequence being that many of these countries are important net exporters of CO$_2$ emissions as they develop a strong manufacturing base to meet the consumption needs of more developed nations. Despite increasing industrialisation, emissions per capita in non-OECD economies are still low compared to OECD countries*" (Yamano and Guilhoto, 2020, page 23).

Figure 8.130: Total production- and consumption-based CO_2 emitted by OECD and non-OECD countries (in GtCO$_2$e)
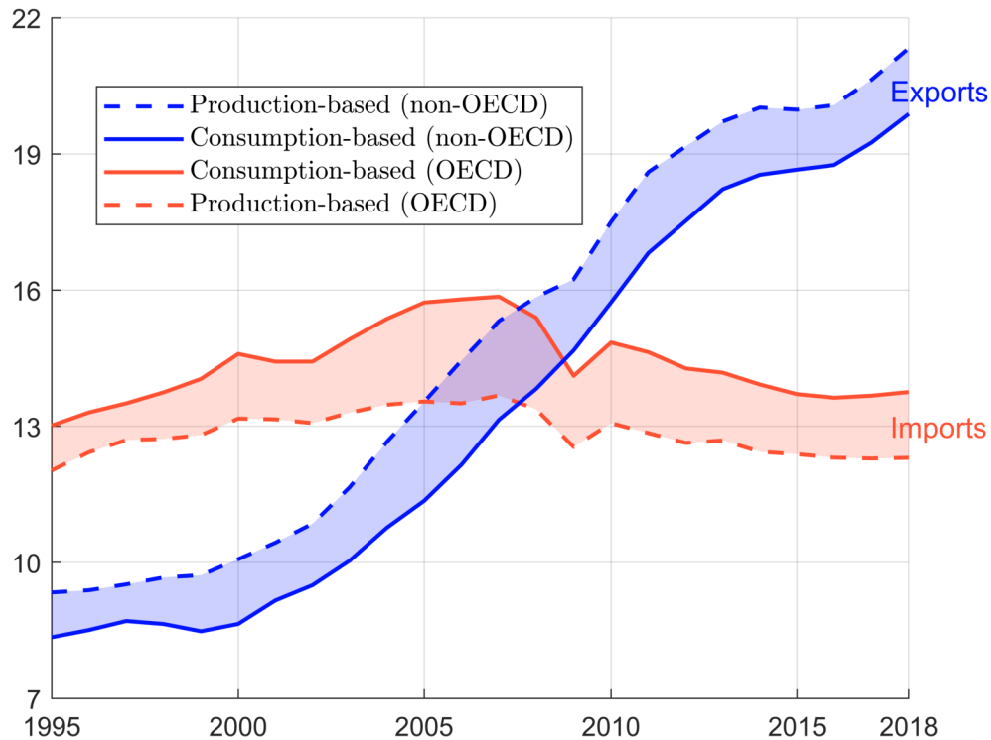
Figure 8.131: Decomposition of OECD imported emissions (in GtCO$_2$e)
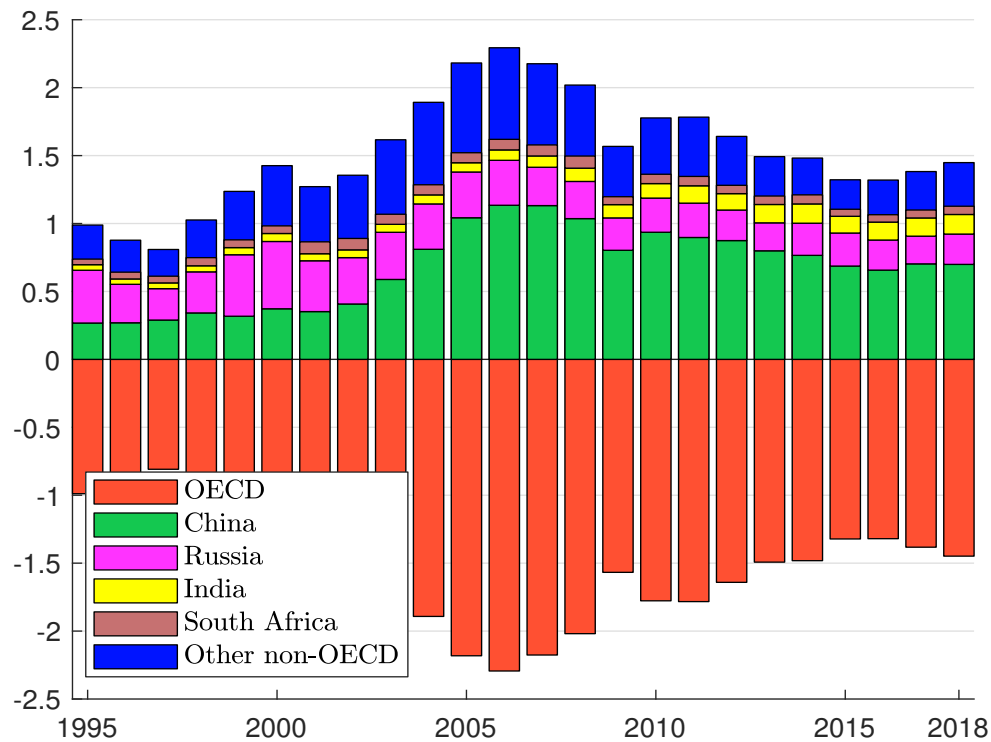
Figure 8.132: Net exported emissions (in MtCO$_2$e)
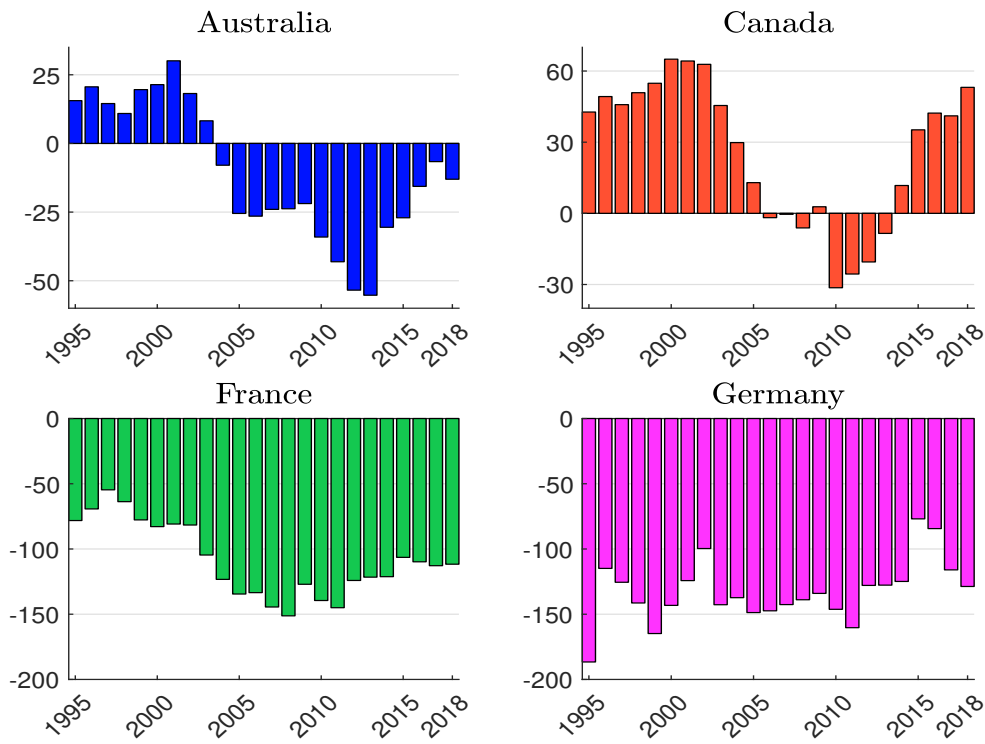
Source: Yamano and Guilhoto (2020), https://stats.oecd.org & Author's calculations.

Figure 8.133: Net exported emissions (in MtCO$_2$e)
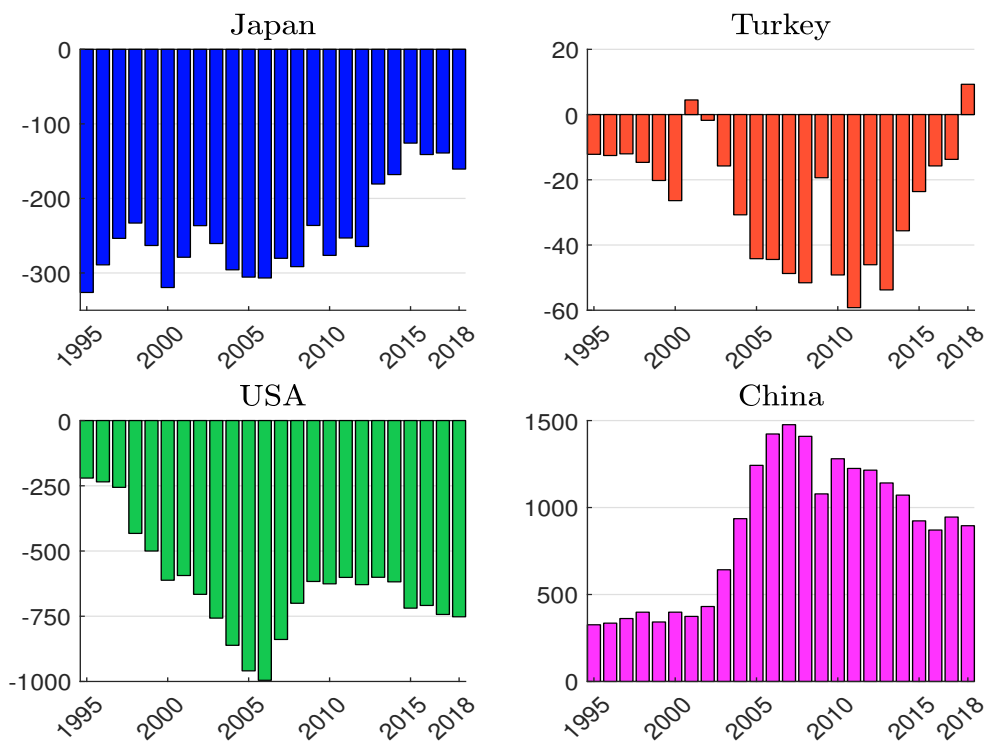
Source: Yamano and Guilhoto (2020), https://stats.oecd.org & Author's calculations.

Figures 8.132 and 8.133 show the evolution of net exported emissions for some countries. First, we see that they are highly time-varying. One reason for this is that the calculation is sensitive to the Leontief matrix of input-output analysis. Second, we observe different patterns across countries. For example, France and Germany are at similar levels in 2018, but this was not the case in 1995. Canada was a net exporter before 2005, a net importer between 2005 and 2015, and a net exporter again after 2015.

Remark 65 *The calculation of imported and exported emissions depends on many factors, especially when using input-output analysis. Depending on the data source, we may find different results. For instance, Australia is a net exporter according to Figure 8.129 and a net importer according to Figure 8.132. According to Commissariat Général au Développement Durable (2020), France's carbon footprint is made up of 45% domestic emissions and 55% imported emissions. The estimated value of imported emissions is then more than 350 $MtCO_2e$. According to the OECD calculation, this figure is close to 110 $MtCO_2e$. These two examples illustrate the large uncertainty in the calculation of GHG emissions embodied in trade.*

8.4.5   Taxation, pass-through and price dynamics

To study the impact of taxation on production costs, we need to diffuse the carbon tax in the input-output economic model to account for the cascading effects through the value chain. The diffusion of the carbon tax depends on the assumption of the reaction function of suppliers and pass-through mechanisms.

Value added approach

By construction, a carbon tax affects the income of producers, who may react in different ways. We first consider a flexible price model and assume that they want to maintain their level of value added.

Remark 66 *In the following, p^- is the price vector before the carbon tax, while p is the price vector including the tax effect.*

Impact on production prices   The absolute amount of the carbon tax for sector j is equal to:

$$T_{\text{direct},j} = \boldsymbol{\tau}_j \boldsymbol{\mathcal{CE}}_{1,j}$$

where $\boldsymbol{\tau}_j$ is the nominal carbon tax expressed in $\$/tCO_2e$ and $\boldsymbol{\mathcal{CE}}_{1,j}$ is the Scope 1 emissions of the sector. We deduce that the carbon tax rate is equal to:

$$t_{\text{direct},j} = \frac{T_{\text{direct},j}}{x_j} = \frac{\boldsymbol{\tau}_j \boldsymbol{\mathcal{CE}}_{1,j}}{x_j} = \boldsymbol{\tau}_j \boldsymbol{\mathcal{CI}}_{1,j}$$

Note that $t_{\text{direct},j}$ has no unit and is equal to the product of the tax and the Scope 1 carbon intensity. The input-output model implies that:

$$p_j x_j = \sum_{i=1}^{n} Z_{i,j} p_i + \sum_{k=1}^{m} V_{k,j} \psi_k + T_{\text{direct},j}$$

We deduce that:

$$p_j = \sum_{i=1}^{n} A_{i,j} p_i + \sum_{k=1}^{m} B_{k,j} \psi_k + t_{\text{direct},j} = \sum_{i=1}^{n} A_{i,j} p_i + \upsilon_j + t_{\text{direct},j}$$

It follows that:

$$p = \left(I_n - A^\top\right)^{-1} \left(\upsilon + t_{\text{direct}}\right)$$

where $t_{\text{direct}} = (t_{\text{direct},1}, \ldots, t_{\text{direct},n})$ is the vector of direct tax rates. We recover the cost-push price model, where the vector υ of value added ratios is replaced by $\upsilon + t_{\text{direct}}$. It follows that the vector of price changes due to the carbon tax is equal to:

$$\Delta p = \left(I_n - A^\top\right)^{-1} t_{\text{direct}} \tag{8.49}$$

This result is obvious since Equation (8.30) implies that $\Delta p = \left(I_n - A^\top\right)^{-1} \Delta \upsilon$ and $\Delta \upsilon$ corresponds to the vector t_{direct} of direct tax rates.

Impact on the price index The definition of a price index is:

$$\mathcal{PI} = \sum_{i=1}^{n} \alpha_i p_i = \alpha^\top p$$

where $\alpha = (\alpha_1, \ldots, \alpha_n)$ are the weights of the items in the basket. We deduce that the inflation rate is:

$$\pi = \frac{\Delta \mathcal{PI}}{\mathcal{PI}^-} = \frac{\mathcal{PI} - \mathcal{PI}^-}{\mathcal{PI}^-} = \frac{\alpha^\top \left(I_n - A^\top\right)^{-1} t_{\text{direct}}}{\alpha^\top \left(I_n - A^\top\right)^{-1} \upsilon}$$

We can simplify this formula because $p^- = \left(I_n - A^\top\right)^{-1} \upsilon = \mathbf{1}_n$ and $\mathbf{1}_n^\top \alpha = 1$. Finally, we have:

$$\pi = \alpha^\top \left(I_n - A^\top\right)^{-1} t_{\text{direct}} \tag{8.50}$$

Computation of the total tax amount The total tax cost is equal to:

$$T_{\text{total}} = x \odot \Delta p = x \odot \left(I_n - A^\top\right)^{-1} t_{\text{direct}} \tag{8.51}$$

while the direct tax cost is $T_{\text{direct}} = x \odot t_{\text{direct}}$. Since we have $x \succeq \mathbf{0}_n$ and $\left(I_n - A^\top\right)^{-1} \succeq I_n$ and using Hadarmard properties[116], then we conclude that the total tax cost is greater than the direct tax cost for all the sectors:

$$T_{\text{total},j} \geq T_{\text{direct},j}$$

Since the total cost to the economy is equal to $\mathcal{C}ost_{\text{total}} = \sum_{j=1}^{n} T_{\text{total},j} = x^\top \left(I_n - A^\top\right)^{-1} t_{\text{direct}}$, the tax incidence is then equal to:

$$\mathcal{TI} = \frac{\mathcal{C}ost_{\text{total}}}{\mathbf{1}_n^\top x} = \frac{x^\top \left(I_n - A^\top\right)^{-1} t_{\text{direct}}}{\mathbf{1}_n^\top x}$$

---

[116]Let A, B and C be three nonnegative matrices. If $B \preceq C$, then $A \odot B \preceq A \odot C$.

Common mistakes in calculating total tax costs In some research papers, we can find two formulas that seem to be intuitive:

$$T'_{\text{total}} = \left(I_n - A^\top\right)^{-1} T_{\text{direct}}$$

and:

$$T''_{\text{total}} = \boldsymbol{\tau} \odot \mathcal{CE}_{\text{total}}$$

The two previous equations are generally wrong because the Hadamard and matrix products are not associative: $A \odot (BC) \neq (A \odot B) C$.

Mathematical properties Let us denote by $f(\boldsymbol{\tau})$ the function f that depends on the vector $\boldsymbol{\tau} = (\boldsymbol{\tau}_1, \ldots, \boldsymbol{\tau}_n)$ of carbon taxes. Let $\lambda \geq 0$ be a positive scalar. The functions Δp, π, T_{total}, $\mathcal{C}ost_{\text{total}}$ and $\mathcal{TI}$ are homogeneous[117] and additive[118]. For instance, we have:

$$\Delta p\left(\lambda \boldsymbol{\tau}\right) = \left(I_n - A^\top\right)^{-1} t_{\text{direct}}\left(\lambda \boldsymbol{\tau}\right) = \lambda \left(I_n - A^\top\right)^{-1} t_{\text{direct}}\left(\boldsymbol{\tau}\right) = \lambda \Delta p\left(\boldsymbol{\tau}\right)$$

If the tax is uniform $\boldsymbol{\tau} = \tau \mathbf{1}_n$, the vector of total tax amount is the product of the tax by the total emissions:

$$T_{\text{total}}\left(\tau \mathbf{1}_n\right) = \tau \mathcal{CE}_{\text{total}}$$

The tax incidence for a given sector is then proportional to the direct plus indirect carbon emissions of the sector. At the global level, the tax incidence is equal to the carbon tax multiplied by the total carbon intensity of the world:

$$\mathcal{TI}\left(\tau \mathbf{1}_n\right) = \frac{\mathbf{1}_n^\top \tau \mathcal{CE}_{\text{total}}}{\mathbf{1}_n^\top x} = \tau \mathcal{CI}_{\text{total}}$$

Illustration Consider a variant of Example 21. Table 8.44 gives the values of $Z_{i,j}$, y_j, x_j and $V_{1,j}$ in \$ mn. The carbon emissions are expressed in $ktCO_2e$, while the carbon intensities are expressed in $tCO_2e/\$$ mn. We have:

$$A = Z \operatorname{diag}^{-1}(x) = \begin{pmatrix} 0.10 & 0.20 & 0.20 & 0.10 \\ 0.10 & 0.10 & 0.20 & 0.05 \\ 0.05 & 0.20 & 0.30 & 0.10 \\ 0.02 & 0.05 & 0.10 & 0.35 \end{pmatrix}$$

and:

$$\tilde{\boldsymbol{\mathcal{L}}} = \left(I_4 - A^\top\right)^{-1} = \begin{pmatrix} 1.1881 & 0.1678 & 0.1430 & 0.0715 \\ 0.3894 & 1.2552 & 0.4110 & 0.1718 \\ 0.4919 & 0.4336 & 1.6303 & 0.2993 \\ 0.2884 & 0.1891 & 0.3044 & 1.6087 \end{pmatrix}$$

Then, we calculate the vector υ of value added ratios:

$$\upsilon = \begin{pmatrix} 3\,650/5\,000 \\ 1\,800/4\,000 \\ 1\,600/8\,000 \\ 5\,000/12\,500 \end{pmatrix} = \begin{pmatrix} 0.73 \\ 0.45 \\ 0.20 \\ 0.40 \end{pmatrix}$$

---

[117]This means that $f(\lambda \boldsymbol{\tau}) = \lambda f(\boldsymbol{\tau})$.
[118]We have $f(\boldsymbol{\tau} + \boldsymbol{\tau}') = f(\boldsymbol{\tau}) + f(\boldsymbol{\tau}')$.

Table 8.44: Environmentally extended monetary input-output table

Sector	Z				y	x	$\mathcal{CE}_1$	$\mathcal{CI}_1$
Energy	500	800	1 600	1 250	850	5 000	500	100
Materials	500	400	1 600	625	875	4 000	200	50
Industrials	250	800	2 400	1 250	3 300	8 000	200	25
Services	100	200	800	4 375	7 025	12 500	125	10
Value added	3 650	1 800	1 600	5 000				
Income	5 000	4 000	8 000	12 500				

We check that $p^- = \tilde{\mathcal{L}} \upsilon = \mathbf{1}_4$. By construction, all the prices are standardized and equal to one in a monetary input-output model. We now introduce a differentiated carbon tax: $\boldsymbol{\tau}_1 = \$200/\text{tCO}_2\text{e}$ and $\boldsymbol{\tau}_2 = \boldsymbol{\tau}_3 = \boldsymbol{\tau}_4 = \$100/\text{tCO}_2\text{e}$. The direct tax costs are 100, 20, 20 and 12.5 million dollars for Energy, Materials, Industrials and Services respectively. We deduce that the vector of carbon tax rates is $t_{\text{direct}} = (2.00\%, 0.50\%, 0.25\%, 0.10\%)$. It follows that:

$$p = \left(I_n - A^\top \right)^{-1} (\upsilon + t_{\text{direct}}) = \begin{pmatrix} 1.0250 \\ 1.0153 \\ 1.0164 \\ 1.0091 \end{pmatrix}$$

If we assume that the basket of goods and services is $\alpha = (10\%, 20\%, 30\%, 40\%)$, the price index $\mathcal{PI}$ is 1.0141 while the inflation rate π is 1.410%. Finally, we calculate the total tax cost and obtain the results shown in Table 8.45. The direct tax cost is multiplied by a factor of 2.8 when we consider the diffusion of the carbon tax. We check that $T_{\text{total}} \neq T'_{\text{total}} \neq T''_{\text{total}}$. Services is the most affected sector, followed by Industrials, Materials and Energy with impact ratios $T_{\text{total}}/T_{\text{direct}}$ of 9.1, 6.6, 3.1 and 1.3 respectively. In Table 8.46, we consider a uniform tax of $\$100/\text{tCO}_2\text{e}$. We check that $T_{\text{total}} = T''_{\text{total}}$ but $T_{\text{total}} \neq T'_{\text{total}}$.

Table 8.45: Total carbon costs (in $ mn) (differentiated tax)

Sector	T_{direct}	T_{total}	T'_{total}	T''_{total}	$\mathcal{CE}_{\text{direct}}$	$\mathcal{CE}_{\text{total}}$
Energy	100.00	125.15	125.92	131.49	500.00	657.44
Materials	20.00	61.05	74.41	45.48	200.00	454.76
Industrials	20.00	131.05	94.21	91.70	200.00	916.97
Services	12.50	113.54	58.82	77.49	125.00	774.92
Sum	152.50	430.79	353.36	346.15	1 025.00	2 804.10

Table 8.46: Total carbon costs (in $ mn) (uniform taxation)

Sector	T_{direct}	T_{total}	T'_{total}	T''_{total}	$\mathcal{CE}_{\text{direct}}$	$\mathcal{CE}_{\text{total}}$
Energy	50.00	65.74	66.51	65.74	500.00	657.44
Materials	20.00	45.48	54.94	45.48	200.00	454.76
Industrials	20.00	91.70	69.62	91.70	200.00	916.97
Services	12.50	77.49	44.40	77.49	125.00	774.92
Sum	102.50	280.41	235.47	280.41	1 025.00	2 804.10

Mark-up pricing approach

Theoretical framework We consider a second approach proposed by Gemechu *et al.* (2014) and Mardones and Mena (2020). Mark-up pricing refers to a business strategy in which the suppliers determine the selling price by adding a fixed percentage to the cost of production. Let p_j^- be the price before the introduction of the carbon tax. We define ξ_j as the price factor induced by the carbon tax: $t_{\text{direct},j} = \xi_j p_j^-$. It follows that $p_j^- = \sum_{i=1}^n A_{i,j} p_i^- + v_j$ and[119]:

$$
\begin{aligned}
p_j &= \left(\sum_{i=1}^n A_{i,j} p_i + v_j \right) + t_{\text{direct},j} \\
&= \left(\sum_{i=1}^n A_{i,j} p_i + v_j \right) + \xi_j p_j^- \\
&= (1 + \xi_j) \left(\sum_{i=1}^n A_{i,j} p_i + v_j \right)
\end{aligned}
$$

We deduce that:

$$
\frac{p_j}{1 + \xi_j} = \sum_{i=1}^n A_{i,j} p_i + v_j
$$

and:

$$
p_j \left(1 - \frac{\xi_j}{1 + \xi_j} \right) = \sum_{i=1}^n A_{i,j} p_i + v_j
$$

It follows that:

$$
\begin{aligned}
p_j &= \sum_{i=1}^n A_{i,j} p_i + \frac{\xi_j}{1 + \xi_j} p_j + v_j \\
&= \sum_{i=1}^n A_{i,j} p_i + p_j \left(1 - \frac{1}{1 + \xi_j} \right) + v_j
\end{aligned}
$$

In a matrix form, we have:

$$
p = A^\top p + (I_n - D_\xi) p + v
$$

where:

$$
D_\xi = \text{diag} \left(\frac{1}{1 + \xi_1}, \dots, \frac{1}{1 + \xi_n} \right)
$$

Finally, we obtain:

$$
p = \left(I_n - A_\xi^\top \right)^{-1} v
$$

where $A_\xi = A + I_n - D_\xi$. Another expression is:

$$
p = \tilde{\boldsymbol{\mathcal{L}}}_m v = \left(D_\xi - A^\top \right)^{-1} v \tag{8.52}
$$

where $\tilde{\boldsymbol{\mathcal{L}}}_m = \left(D_\xi - A^\top \right)^{-1}$ is the mark-up inverse matrix. The vector of price variations is then:

$$
\Delta p = \left(\tilde{\boldsymbol{\mathcal{L}}}_m - \tilde{\boldsymbol{\mathcal{L}}} \right) v \tag{8.53}
$$

---

[119]We assume that $p_j \approx p_j^-$.

The expression of the price index is $\mathcal{PI} = \alpha^\top \left(D_\xi - A^\top \right)^{-1} \upsilon$ whereas the inflation rate is equal to $\pi = \alpha^\top \left(\tilde{\boldsymbol{\mathcal{L}}}_m - \tilde{\boldsymbol{\mathcal{L}}} \right) \upsilon$. From Equation (8.53), we also deduce the total tax cost:

$$T_{\text{total}} = x \odot \left(\tilde{\boldsymbol{\mathcal{L}}}_m - \tilde{\boldsymbol{\mathcal{L}}} \right) \upsilon \tag{8.54}$$

Note that the mark-up approach implies replacing the identity matrix I_n with the diagonal matrix D_ξ in the cost-push price model. Since we have $D_\xi \preceq I_n$, we deduce that $D_\xi^{-1} \succeq I_n$. Desnos *et al.* (2023) also showed that $\tilde{\boldsymbol{\mathcal{L}}}_m \succeq \tilde{\boldsymbol{\mathcal{L}}}$.

Illustration Considering the previous example, we have:

$$\tilde{\boldsymbol{\mathcal{L}}}_m = \left(D_\xi - A^\top \right)^{-1} = \begin{pmatrix} 1.2170 & 0.1730 & 0.1474 & 0.0735 \\ 0.4017 & 1.2650 & 0.4165 & 0.1740 \\ 0.5067 & 0.4398 & 1.6394 & 0.3021 \\ 0.2965 & 0.1919 & 0.3074 & 1.6121 \end{pmatrix}$$

In the case of differentiated taxation, we obtain:

$$p = \tilde{\boldsymbol{\mathcal{L}}}_m \upsilon = \begin{pmatrix} 1.0252 \\ 1.0154 \\ 1.0165 \\ 1.0091 \end{pmatrix}$$

The inflation rate π is equal to 1.421% and the total carbon costs (in \$ mn) are 125.82, 61.53, 132.10 and 114.33. The global cost is then \$433.78 mn compared to \$430.79 mn in the value added approach.

Pass-through integration

Definition According to RBB Economics (2014), "*cost pass-through describes what happens when a business changes the price of the production or services it sells following a change in the cost of producing them*". Therefore, a pass-through rate is closely related to the supply and demand elasticity. This concept of price adjustment is extremely common in many fields of economics: exchange rates, imperfect competition and Cournot-Bertrand equilibria, product taxation and retail prices, inflation regimes, etc. In other words, pass-through is the ability of a sector or a company to pass costs through its supply chain. In general, this parameter ranges from 0%, where the entire amount is supported by the agent, to 100%, where the entire amount is passed on to customers. As this parameter depends on several factors, such as supply and demand elasticity, international trade exposure, market concentration, product homogeneity, etc., its estimation is not easy, which implies a large uncertainty about the tax incidence in a transition risk framework.

Pass-through strongly depends on the market structure and the supply-demand equilibrium. In Figure 8.134, we show different demand curves whose slope depends on the consumer response to different price levels. If the slope of the curve is steep, it suggests that an increase in price would lead to a marginal decrease in sales. This scenario represents inelastic demand, where consumer demand is relatively unchanged when the price moves up or down. Conversely, if the demand curve is flatter, an increase in price will result in a significant reduction in the quantity demanded. This situation represents elastic demand, where consumers are highly responsive to price changes. If the demand curve is linear, there is no curvature, which means that the rate of decline in demand remains

Figure 8.134: Demand curvature

(a) Linear demand

(b) Concave demand

(c) Convex demand

(d) Extreme cases

constant as the price increases (top/left panel in Figure 8.134). In situations where demand falls more sharply as the price rises, this type of demand is classified as concave to the origin (top/right panel in Figure 8.134). As prices rise in this scenario, the demand curve becomes increasingly flatter, indicating increased price sensitivity or greater elasticity. In this scenario, firms should absorb part of the cost, implying a relatively low pass-through rate. Finally, if the rate of decline in demand slows with each price increase, this type of demand curve is said to be convex to the origin. In this case, as prices escalate, the remaining demand becomes less sensitive to these price fluctuations (bottom/left panel of the figure). Firms can then pass on the costs and set a relatively high pass-through rate.

From an economic point of view, the specification of the pass-through depends on several factors. In the case of competition, the general formula for the pass-through rate ϕ is:

$$\phi = \frac{\mathrm{d}\,p}{\mathrm{d}\,\tau} = \frac{\text{price sensitivity of supply}}{\text{price sensitivity of supply} - \text{price sensitivity of demand}}$$

We deduce that $\phi \in [0, 100\%]$. In a monopolistic situation, the previous formula becomes:

$$\phi = \frac{1}{2 + \text{elasticity of the slope of inverse demand}}$$

Since the slope elasticity of inverse demand is negative, $\phi \geq 50\%$. We get similar results in oligopolistic situations. In monopolistic and oligopolistic situations, it can also be greater than 100% if demand is highly convex. In Table 8.47, we report some estimates of path-through rates.

Table 8.47: Pass-through rates (in %) for intensive sectors

Sector	Rate
Electricity, gas and steam	100%
Petroleum refining	100%
Base metals	78%
Mining	78%
Waste/wastewater	78%
Land transport	78%
Fishery	75%
Non-metallic minerals	60%
Agriculture	50%
Chemicals	40%
Maritime transport	30%
Aviation	30%
Paper	10%

Source: (Sautel *et al.*, 2022, page 35).

Analytical formula We focus on the value added model, which is the most widely used approach in the academic literature. It is also the simplest model to introduce the pass-through mechanism. We have:

$$\Delta p = \tilde{\boldsymbol{\mathcal{L}}} \Delta \upsilon = \sum_{k=0}^{\infty} \left(A^{\top} \right)^k \Delta \upsilon = \sum_{k=0}^{\infty} \Delta p_{(k)}$$

where $\Delta p_{(k)} = \left(A^{\top} \right)^k \Delta \upsilon$ is the price impact at the kth tier. In fact, $\Delta p_{(k)}$ satisfies the following recurrence relation:

$$\begin{cases} \Delta p_{(k)} = A^{\top} \Delta p_{(k-1)} \\ \Delta p_{(0)} = \Delta \upsilon \end{cases}$$

If we consider the price p_j of sector j, we have $\Delta p_{(0),j} = \Delta \upsilon_j$ and:

$$\Delta p_{(k),j} = \sum_{i=1}^{n} A_{i,j} \Delta p_{(k-1),i}$$

This representation helps to better understand the cascading effect of the carbon tax. In the zeroth round, it induces an additional cost $\Delta \upsilon_j$, which is fully passed on to the price p_j of the sector. The new price is then $p_j + \Delta p_{(0),j} = p_j + \Delta \upsilon_j$. In the first round, sector j faces new additional costs due to the price increase of intermediate consumption. We have $\Delta p_{(1),j} = \sum_{i=1}^{n} A_{i,j} \Delta p_{(0),i} = \sum_{i=1}^{n} A_{i,j} \Delta \upsilon_i$. The iteration process continues and we have $\Delta p_{(2),j} = \sum_{i=1}^{n} A_{i,j} \Delta p_{(1),i} = \sum_{i=1}^{n} \sum_{k=1}^{n} A_{i,j} A_{k,i} \Delta \upsilon_k$ at the second round.

Now let us introduce the pass-through mechanism. By definition, we have $\Delta p_{(0),j} = \phi_j \Delta \upsilon_j$ where ϕ_j denotes the pass-through rate of sector j. In the first round, we have:

$$\Delta p_{(1),j} = \sum_{i=1}^{n} A_{i,j} \left(\phi_i \Delta p_{(0),i} \right) = \sum_{i=1}^{n} A_{i,j} \left(\phi_i \Delta \upsilon_i \right)$$

More generally, the recurrence relation is:

$$\Delta p_{(k),j} = \sum_{i=1}^{n} A_{i,j} \phi_i \Delta p_{(k-1),i}$$

Let $\phi = (\phi_1, \ldots, \phi_n)$ and $\Phi = \mathrm{diag}(\phi)$ be the pass-through vector and matrix. The recurrence matrix form is:

$$\begin{cases} \Delta p_{(k)} = A^{\top} \Phi \Delta p_{(k-1)} \\ \Delta p_{(0)} = \Phi \Delta v \end{cases}$$

We deduce that:

$$\begin{aligned} \Delta p &= \sum_{k=0}^{\infty} \left(A^{\top} \Phi \right)^k \Phi \Delta v \\ &= \left(I_n - A^{\top} \Phi \right)^{-1} \Phi \Delta v \\ &= \tilde{\mathcal{L}}(\phi) \Delta v \end{aligned} \qquad (8.55)$$

where $\tilde{\mathcal{L}}(\phi) = \left(I_n - A^{\top} \Phi \right)^{-1} \Phi$.

Since A is a substochastic matrix and Φ is a positive diagonal matrix, we verify that $\phi' \succeq \phi \Rightarrow \tilde{\mathcal{L}}(\phi') \succeq \tilde{\mathcal{L}}(\phi)$. The lower bound is then reached when $\phi = \mathbf{0}_n$ while the upper bound is reached when $\phi = \mathbf{1}_n$.

Application to the carbon tax Applying the previous analysis to the carbon tax, we have $\Delta v = t_{\mathrm{direct}}$. In this case, the concept of total tax cost must be redefined because part of the cost is paid by producers and part by consumers. By consumers, we must understand the downstream of the value chain. We have:

$$\begin{aligned} T_{\mathrm{producer}} &= x \odot (I_n - \Phi) t_{\mathrm{direct}} \\ &= x \odot (\mathbf{1}_n - \phi) \odot t_{\mathrm{direct}} \\ &= (\mathbf{1}_n - \phi) \odot T_{\mathrm{direct}} \end{aligned}$$

and:

$$T_{\mathrm{consumer}} = T_{\mathrm{downstream}} = x \odot \tilde{\mathcal{L}}(\phi) t_{\mathrm{direct}}$$

We deduce that:

$$\begin{aligned} T_{\mathrm{total}} &= T_{\mathrm{producer}} + T_{\mathrm{consumer}} \\ &= x \odot \left(I_n - \Phi + \tilde{\mathcal{L}}(\phi) \right) t_{\mathrm{direct}} \end{aligned}$$

If $\phi_j = 100\%$, we have $\tilde{\mathcal{L}}(\mathbf{1}_n) = \tilde{\mathcal{L}}$ and $\Delta p = \tilde{\mathcal{L}} t_{\mathrm{direct}}$. This is the original approach. If $\phi_j = 0\%$, we have $\tilde{\mathcal{L}}(\mathbf{0}_n) = \mathbf{0}_{n,n}$, $\Delta p = \mathbf{0}_n$, $T_{\mathrm{producer}} = T_{\mathrm{direct}}$ but $T_{\mathrm{consumer}} = \mathbf{0}_n$. The costs passed on to consumers (or the downstream of the value chain) are zero because the direct costs are initially absorbed by the producers.

Remark 67 *The functions Δp, π, T_{total}, $\mathcal{C}ost_{\mathrm{total}}$ and $\mathcal{TI}$ remain homogeneous and additive with respect to τ. We can also show that:*

$$\phi' \succeq \phi \Rightarrow T_{\mathrm{total}}(\tau, \phi') \succeq T_{\mathrm{total}}(\tau, \phi)$$

The effects of the tax is maximum when $\phi = \mathbf{1}_n$ and minimum when $\phi = \mathbf{0}_n$. If we consider a uniform pass-through, the total cost of the carbon tax is an increasing function of the pass-through rate.

Figure 8.135: Producer and consumer cost contributions (uniform pass-through rate)

Figure 8.136: Producer and consumer cost contributions ($\phi_2 = \phi_3 = \phi_4 = 0\%$)

Illustration   Continuing our example of differential taxation, let's assume that pass-through rates are uniform ($\phi_1 = \phi_2 = \phi_3 = \phi_4$). The evolution of the total cost is shown in Figure 8.135. When $\phi_j = 0\%$, T_{total} is equal to \$152.50 mn and is the lower bound. The upper bound is reached when $\phi_j = 100\%$ and we get $T_{\text{total}} = \$430.79$ mn. We have also shown the contribution of each sector by distinguishing between direct and indirect costs. Figure 8.136 corresponds to the case where the Energy sector passes on the direct costs to the other sectors.

Empirical results

There are many studies of carbon taxation and input-output models (Köppl and Schratzenstaller, 2023). In the following, we focus on two studies that have extensively examined the impact of pass-through rates on the economic cost of a carbon tax and its impact on inflation. We also produce new results by implementing a carbon tax of \$100/tCO$_2$e and differentiated pass-through rates.

The study of Desnos *et al.* (2023)   The authors analyze the impact of a uniform tax τ for all countries and a uniform pass-through rate ϕ for all sectors. The direct cost $\mathcal{C}ost_{\text{direct}} = \sum_{j=1}^{n} T_{\text{direct}}$ and the total cost $\mathcal{C}ost_{\text{total}} = \sum_{j=1}^{n} T_{\text{total}}$ are shown in Figure 8.137. If the carbon tax is set at \$100/tCO$_2$e, the direct cost is \$4.8 tn, while the total cost is \$6.1 tn if $\phi = 50\%$ and \$13.3 tn if $\phi = 100\%$. These correspond to 2.8%, 3.6% and 7.8% of the world GDP respectively. If we apply a carbon tax of \$500/tCO$_2$e, these costs become \$24.2, \$30.4 and \$66.4 tn respectively. They show that the relationship between total costs and the pass-through parameter is cubic. They suggest the following approximation:

$$\frac{\mathcal{C}ost_{\text{total}}\left(\boldsymbol{\tau}, \phi \mathbf{1}_n\right)}{\mathcal{C}ost_{\text{direct}}\left(\boldsymbol{\tau}, \phi \mathbf{1}_n\right)} \approx 1 + m_{(1-\infty)}\phi^3$$

Therefore, a small error in pass-through rate estimate can lead to a large error in the cost estimate.

Desnos *et al.* (2023) then analyze the impact of the carbon tax on the inflation. To do this, they define two price indices: the producer price index (PPI), where the basket weights are proportional to the output ($\alpha_j \propto x_j$) and the consumer price index (CPI), where the basket weights are proportional to the final demand ($\alpha_j \propto y_j$). Results are shown in Figure 8.138. Again, the inflation rate depends on the pass-through rate. For a carbon tax of \$500/tCO$_2$e and a pass-through rate of 100%, the PPI inflation rate is close to 40%, while the CPI inflation rate reaches 30%. These global figures are the result of a large discrepancy between country inflation rates. Figure 8.139 shows the world map of the country inflation rates for a uniform tax of \$100/tCO$_2$e. There are three factors (basket composition, value chain impact and direct carbon emissions of the country) that explain the dispersion of the inflation rates:

$$\pi = \underbrace{\alpha^{\top}}_{\text{Basket}} \cdot \underbrace{\tilde{\boldsymbol{\mathcal{L}}}\left(\boldsymbol{\phi}\right)}_{\text{Value chain}} \cdot \underbrace{t_{\text{direct}}}_{\text{Scope 1}}$$

Direct costs are the main contributor, followed by the impact of the downstream diffusion of the carbon tax. For example, Europe's low inflation rate is explained by its low direct emissions, but Europe is heavily penalized by its value chain. China is affected by both factors, while Russia's high inflation is mainly due to its direct emissions, as its value chain impact is among the lowest in the world.

Desnos *et al.* (2023) analyze a regional taxation scenario in which a carbon tax is imposed uniquely within a specific region of the world. This situation is likely to arise due to the lack of uniformity in carbon pricing. They first consider a uniform taxation across EU member states, which is certainly the most likely scenario. In this scenario, a \$500/tCO$_2$e carbon tax with a 100% pass-through would result in a global cost of \$4.5 tn, of which \$4 tn would be borne by EU countries

Figure 8.137: World economic cost in $ tn (global analysis, uniform tax, Exiobase 2022)

Figure 8.138: World inflation rate in % (global analysis, uniform tax, Exiobase 2022)

Figure 8.139: Production inflation rate in % (global analysis, uniform tax, $\tau = \$100/\text{tCO}_2\text{e}$, $\phi = 100\%$, Exiobase 2022)

and a \$0.5 tn by non-EU countries. If EU sectors absorb their increased costs by passing only 50% through the value chain, non-EU countries are less affected by carbon tax diffusion, and their costs fall from \$521 bn to \$54 bn. Moreover, the costs relative to GDP for EU sectors fall as they absorb the carbon tax, from 14% if they pass on the carbon tax in full to 8% if direct emitters bear half of the cost of the carbon tax. Among the EU countries, Germany is the most affected, followed by Poland and Italy (Figure 8.140).

Table 8.48: Domestic and foreign impacts (in \$ bn) of a regional tax (uniform taxation, $\phi = 100\%$, Exiobase 2022)

Carbon tax	Domestic impact			Foreign impact		
	EU	USA	China	EU	USA	China
\$100/tCO$_2$e	792	886	4 710	104	118	257
\$250/tCO$_2$e	1 979	2 215	11 774	261	296	643
\$500/tCO$_2$e	3 959	4 430	23 549	521	592	1 287

Table 8.48 shows the global foreign impact of a tax in the EU, US and China for three values of the carbon price. China always has the highest external impact, with a cost of \$1 287 bn for a \$500/tCO$_2$e carbon price, while the EU's external impact is only \$521 bn. To better understand the winners and the losers, we report the fifteen largest countries affected by a carbon tax in Table 8.49. In this scenario, the tax is set at \$100/tCO$_2$e and the pass-through parameter is set to 100%. When the tax is applied in the European Union, the rest-of-the-world region is most affected, accounting for 25.25% of the total costs supported by foreign countries. This is followed by China (23.62%),

Figure 8.140: Cost breakdown (EU, uniform tax, $\phi = 50\%$, Exiobase 2022)

Table 8.49: Fifteen most affected foreign countries (uniform tax, $\tau = \$100/tCO_2e$, $\phi = 100\%$, Exiobase 2022)

Rank	EU tax		US tax		Chinese tax	
1	ROW	25.25%	CHN	24.74%	ROW	36.89%
2	CHN	23.62%	ROW	18.60%	USA	12.95%
3	USA	11.45%	CAN	9.35%	KOR	8.87%
4	GBR	8.77%	MEX	8.51%	IND	6.91%
5	CHE	4.32%	KOR	6.89%	JPN	6.44%
6	KOR	4.05%	JPN	5.05%	DEU	3.61%
7	IND	3.67%	IND	4.28%	MEX	2.19%
8	JPN	3.31%	DEU	2.80%	FRA	1.88%
9	TUR	2.62%	BRA	2.51%	GBR	1.83%
10	TWN	2.08%	GBR	2.34%	BRA	1.75%
11	CAN	2.06%	FRA	1.63%	IDN	1.74%
12	RUS	1.96%	TWN	1.59%	CAN	1.62%
13	BRA	1.90%	IRL	1.47%	ITA	1.59%
14	MEX	1.70%	ITA	1.43%	AUS	1.31%
15	NOR	1.46%	NLD	1.23%	TUR	1.11%

the United States (11.45%) and the United Kingdom (8.77%). This would also be the case with a Chinese tax, but it now represents more than 36% of the foreign costs. This is followed by the United States (12.95%), South Korea (8.87%) and India (6.91%). In the case of a US tax, China would be the most affected country with 24.74% of the total impact, followed by the rest-of-the-world region (18.60%), Canada (9.35%) and Mexico (8.51%). It is important to note that the US has strong trade relations with China, but also with other countries in the Americas (Canada, Mexico, Brazil). These results highlight the trade links between countries, and consequently, the potential exposure to a carbon tax. For example, if we focus on Turkey, it is highly linked to the EU, as it would be the 9th country most affected by an EU carbon tax. It would also be affected by a Chinese tax, as it would be the 15th most affected country. The impact would be smaller in the case of a US carbon tax. Similarly, we see the importance of Germany in the Chinese and US supply chain.

The study of Sautel *et al.* (2022)   The analysis carried out by Sautel *et al.* (2022) uses the previous input-output framework. It focuses on France and considers a carbon tax of €250 per tonne of CO_2. The estimates are based on the WIOD database and default values of pass-through rates (see Table 8.47 on page 453). The conclusions are as follows:

> "*The total additional cost of introducing a price of €250 per tonne of CO_2 to be paid by French emitting installations is €57.6 billion, or about 2.5 points of GDP. Of this total, €7 billion corresponds to purchases by foreign operators and investments by French and foreign operators. [...] Of this €50.3 billion, French companies would ultimately bear 57% of the additional costs, or about €28.7 billion. The rest would be passed on to final demand, i.e. 21.6 billion euros.*" (Sautel *et al.*, 2022, page 39).

New results   It is common to assume that the pass-through rate follows a beta distribution, as it is a parameter between 0 and 1:

$$\phi \sim \mathcal{B}(\alpha, \beta)$$

Following Sautel *et al.* (2022), Desnos *et al.* (2023) considered four types of sectors with respect to price-demand elasticity (highly-elastic, high-elastic, medium-elastic and low-elastic), and used the expert-opinion values of the parameters α and β given in Table 8.50. We give the mean, the standard deviation and the 95% range, and show the associated probability density functions in Figure 8.141. The first type is right-skewed, while the fourth type is left-skewed. The second and third types are more symmetric. Moreover, these four distribution functions are ordered since they verify the first-order stochastic dominance principle (Figure 8.142).

Table 8.50: Probabilistic characterization of the four pass-through types

	Statistic	Highly-elastic	High-elastic	Medium-elastic	Low-elastic
Parameters	α	3.0	4.0	14.0	12.0
	β	12.0	6.0	6.0	0.6
Moments	μ_ϕ	20%	40%	70%	95%
	σ_ϕ	10%	15%	10%	6%
Range	$Q_\phi(2.5\%)$	5%	14%	49%	79%
	$Q_\phi(97.5\%)$	43%	70%	87%	100%

Desnos *et al.* (2023) defined a mapping between the WIOD sectors and the four types of pass-through mechanisms. They used this classification to build a Monte Carlo Value-at-Risk engine[120].

---

[120]See Section 13.3 on page 639.

Figure 8.141: Probability density function of pass-through rates

Figure 8.142: Probability density function of pass-through rates

We use their classification to measure the impact of a $100/tCO_2Eq$ carbon tax. We assume that the carbon tax is applied to a given region and compute the total cost vector as follows:

$$T_{\text{total}} = x \odot \left(I_n - diag\left(\boldsymbol{\phi}\right) + \tilde{\boldsymbol{\mathcal{L}}}\left(\boldsymbol{\phi}\right) \right) t_{\text{direct}}$$

We assume that the pass-through rate of the sector is constant and equal to the mean of the corresponding beta distribution:

$$\boldsymbol{\phi} = \mu_{\boldsymbol{\phi}}$$

We decompose the total costs into what is paid by the producer and what is paid by the downstream value chain, including the final consumer. We have the following decomposition:

$$\begin{cases} T_{\text{producer}} = x \odot \left(\mathbf{1}_n - \boldsymbol{\phi}\right) \odot t_{\text{direct}} \\ T_{\text{downstream}} = x \odot \tilde{\boldsymbol{\mathcal{L}}}\left(\boldsymbol{\phi}\right) t_{\text{direct}} \end{cases}$$

We also consider a second decomposition between the direct costs of the carbon tax and the indirect costs due to the pass-through mechanism:

$$\begin{cases} T_{\text{direct}} = x \odot t_{\text{direct}} \\ T_{\text{indirect}} = T_{\text{total}} - T_{\text{direct}} = x \odot \left(\tilde{\boldsymbol{\mathcal{L}}}\left(\boldsymbol{\phi}\right) - diag\left(\boldsymbol{\phi}\right) \right) t_{\text{direct}} \end{cases}$$

By definition, government revenue is equal to the direct cost of the carbon tax:

$$R_{\text{government}} = T_{\text{direct}} = x \odot t_{\text{direct}}$$

All of the previous measures can be aggregated at a global level or at a country/regional level. Inflation rates are calculated using the following formula:

$$\pi = \alpha^\top \Delta p = \alpha^\top \tilde{\boldsymbol{\mathcal{L}}}\left(\boldsymbol{\phi}\right) t_{\text{direct}}$$

We distinguish between PPI and CPI inflation. In the first case, the basket weights $\alpha = x/\left(\mathbf{1}_n^\top x\right)$ are defined with respect to the output vector, while in the second case we use the final demand vector: $\alpha_{\text{ppi}} = x/\left(\mathbf{1}_n^\top x\right)$ and $\alpha_{\text{cpi}} = y/\left(\mathbf{1}_n^\top y\right)$. If the analysis is conducted at the country/regional level, the weights are calculated using the country's output and final demand vectors.

In Tables 8.51 and 8.52, we report on the results when we consider a global carbon tax. Cost and revenue outcomes are expressed as a percentage of GDP[121], while inflation is expressed in %. For the world, the total cost is 4.2% of GDP, according to the WIOD 2014 MRIO table. These costs are relatively balanced between direct and indirect effects. This is not the case when looking at the producer/downstream distribution. Most of the costs are borne by the downstream value chain and the final consumers. Inflation is higher when looking at the producer price index than at the consumer price index: 3.7% versus 2.9%. India is the most affected country (9.6%), followed by Russia (8.6%) and China (7.2%). If we use the Exiobase 2022 MRIO table, we get similar global figures, but the results at the regional level can be different. Total costs are now equal to 5% of GDP and CPI inflation reaches 3.5%. Russia is the most affected country with total costs of 12.8% of GDP. It is followed by India (11.4%) and Indonesia (7.8%). China ranks fourth with total costs of 7.5%. If we look at an EU carbon tax, we see that the overall economic costs are mainly in the EU countries (Table 8.53). We get similar results with a carbon tax in the US or China (Table 8.54 and 8.55). This analysis shows that the introduction of a regional carbon tax creates a competitive distortion between the region introducing the tax and the rest of the world. Furthermore, implementing a carbon tax leads to inflation. This is because path-through mechanisms amplify the price dynamics of the costs passed on by producers. The real benefits of a carbon tax are then challenging and will be discussed in detail in Chapter 10.

---

[121]This means that we normalize these figures by the total output $\mathbf{1}_n^\top x$.

Table 8.51: Economic impact of a global carbon tax ($100/tCO$_2$e, WIOD 2014)

Region	Cost						Revenue	Inflation	
	T_{total}	T_{direct}	T_{indirect}	T_{producer}	$T_{\text{downstream}}$	T_{net}	$R_{\text{government}}$	π_{ppi}	π_{cpi}
World	4.17%	2.01%	2.16%	0.48%	3.68%	2.16%	2.01%	3.68%	2.92%
AUS	2.58%	1.34%	1.24%	0.27%	2.30%	1.24%	1.34%	2.30%	1.84%
AUT	1.25%	0.54%	0.71%	0.17%	1.08%	0.71%	0.54%	1.08%	1.02%
BEL	1.60%	0.65%	0.95%	0.23%	1.36%	0.95%	0.65%	1.36%	1.26%
BGR	6.49%	3.51%	2.98%	0.57%	5.92%	2.98%	3.51%	5.92%	3.78%
BRA	2.07%	1.21%	0.86%	0.38%	1.69%	0.86%	1.21%	1.69%	1.51%
CAN	2.37%	1.43%	0.93%	0.34%	2.03%	0.93%	1.43%	2.03%	1.55%
CHE	0.63%	0.19%	0.44%	0.07%	0.55%	0.44%	0.19%	0.55%	0.76%
CHN	7.23%	3.13%	4.09%	0.91%	6.32%	4.09%	3.13%	6.32%	5.25%
CYP	2.36%	1.38%	0.99%	0.45%	1.91%	0.99%	1.38%	1.91%	2.41%
CZE	3.34%	1.66%	1.68%	0.26%	3.08%	1.68%	1.66%	3.08%	2.19%
DEU	1.78%	0.96%	0.83%	0.23%	1.56%	0.83%	0.96%	1.56%	1.37%
DNK	1.62%	1.03%	0.58%	0.54%	1.07%	0.58%	1.03%	1.07%	1.12%
ESP	1.67%	0.81%	0.87%	0.24%	1.44%	0.87%	0.81%	1.44%	1.13%
FIN	1.83%	0.88%	0.95%	0.24%	1.59%	0.95%	0.88%	1.59%	1.36%
FRA	0.89%	0.46%	0.43%	0.17%	0.72%	0.43%	0.46%	0.72%	0.85%
GBR	1.30%	0.69%	0.61%	0.19%	1.10%	0.61%	0.69%	1.10%	1.08%
GRC	3.02%	1.71%	1.30%	0.26%	2.76%	1.30%	1.71%	2.76%	1.82%
HRV	2.38%	1.35%	1.03%	0.38%	2.00%	1.03%	1.35%	2.00%	11.31%
HUN	2.41%	1.22%	1.19%	0.32%	2.08%	1.19%	1.22%	2.08%	1.80%
IDN	4.73%	2.74%	1.99%	0.95%	3.78%	1.99%	2.74%	3.78%	3.08%
IND	9.56%	5.12%	4.44%	1.20%	8.36%	4.44%	5.12%	8.36%	6.90%
IRL	1.17%	0.64%	0.54%	0.24%	0.93%	0.54%	0.64%	0.93%	0.97%
ITA	1.41%	0.64%	0.77%	0.19%	1.22%	0.77%	0.64%	1.22%	1.25%
JPN	2.68%	1.29%	1.39%	0.29%	2.39%	1.39%	1.29%	2.39%	1.74%
KOR	4.20%	1.82%	2.38%	0.39%	3.81%	2.38%	1.82%	3.81%	4.03%
LTU	3.27%	1.77%	1.50%	0.43%	2.84%	1.50%	1.77%	2.84%	1.45%
LUX	0.79%	0.32%	0.47%	0.18%	0.61%	0.47%	0.32%	0.61%	0.79%
LVA	2.37%	1.05%	1.33%	0.31%	2.07%	1.33%	1.05%	2.07%	1.98%
MEX	3.03%	1.87%	1.15%	0.51%	2.51%	1.15%	1.87%	2.51%	2.12%
MLT	2.03%	1.18%	0.85%	0.62%	1.41%	0.85%	1.18%	1.41%	4.08%
NLD	2.04%	1.34%	0.70%	0.44%	1.61%	0.70%	1.34%	1.61%	1.84%
NOR	0.90%	0.55%	0.34%	0.17%	0.72%	0.34%	0.55%	0.72%	0.77%
POL	4.45%	2.44%	2.01%	0.50%	3.95%	2.01%	2.44%	3.95%	2.72%
PRT	2.00%	0.99%	1.00%	0.33%	1.67%	1.00%	0.99%	1.67%	1.05%
ROU	3.20%	1.65%	1.55%	0.40%	2.81%	1.55%	1.65%	2.81%	2.31%
RUS	8.63%	4.51%	4.12%	0.75%	7.88%	4.12%	4.51%	7.88%	3.64%
SVK	2.50%	1.24%	1.26%	0.25%	2.25%	1.26%	1.24%	2.25%	1.75%
SVN	2.40%	1.22%	1.18%	0.28%	2.13%	1.18%	1.22%	2.13%	1.96%
SWE	0.89%	0.41%	0.47%	0.15%	0.74%	0.47%	0.41%	0.74%	0.89%
TUR	3.78%	1.81%	1.97%	0.56%	3.22%	1.97%	1.81%	3.22%	2.18%
TWN	5.41%	2.41%	3.00%	0.63%	4.78%	3.00%	2.41%	4.78%	4.60%
USA	2.26%	1.40%	0.86%	0.28%	1.98%	0.86%	1.40%	1.98%	2.65%
ROW	6.17%	2.72%	3.44%	0.50%	5.67%	3.44%	2.72%	5.67%	4.03%

Table 8.52: Economic impact of a global carbon tax ($100/tCO$_2$e, Exiobase 2022)

Region	Cost						Revenue	Inflation	
	T_{total}	T_{direct}	T_{indirect}	T_{producer}	$T_{\text{downstream}}$	T_{net}	$R_{\text{government}}$	π_{ppi}	π_{cpi}
World	5.01%	2.82%	2.18%	0.93%	4.08%	2.18%	2.82%	4.08%	3.53%
AUS	4.63%	2.93%	1.70%	0.81%	3.82%	1.70%	2.93%	3.82%	2.65%
AUT	2.08%	0.91%	1.17%	0.30%	1.77%	1.17%	0.91%	1.77%	1.94%
BEL	1.77%	0.94%	0.82%	0.44%	1.33%	0.82%	0.94%	1.33%	1.63%
BGR	7.07%	3.94%	3.12%	0.89%	6.18%	3.12%	3.94%	6.18%	3.89%
BRA	5.22%	3.78%	1.44%	2.01%	3.21%	1.44%	3.78%	3.21%	2.38%
CAN	3.74%	2.25%	1.49%	0.57%	3.17%	1.49%	2.25%	3.17%	2.86%
CHE	0.75%	0.30%	0.45%	0.16%	0.59%	0.45%	0.30%	0.59%	0.74%
CHN	7.47%	3.44%	4.03%	1.21%	6.26%	4.03%	3.44%	6.26%	6.35%
CYP	5.05%	3.94%	1.11%	2.49%	2.56%	1.11%	3.94%	2.56%	3.57%
CZE	4.47%	2.13%	2.34%	0.44%	4.03%	2.34%	2.13%	4.03%	4.63%
DEU	1.99%	1.10%	0.89%	0.35%	1.64%	0.89%	1.10%	1.64%	1.60%
DNK	1.47%	0.98%	0.49%	0.54%	0.93%	0.49%	0.98%	0.93%	1.35%
ESP	2.25%	1.15%	1.11%	0.41%	1.84%	1.11%	1.15%	1.84%	1.71%
FIN	2.80%	1.36%	1.44%	0.36%	2.44%	1.44%	1.36%	2.44%	1.82%
FRA	1.39%	0.79%	0.60%	0.35%	1.04%	0.60%	0.79%	1.04%	6.29%
GBR	1.53%	0.88%	0.65%	0.33%	1.20%	0.65%	0.88%	1.20%	2.62%
GRC	6.39%	4.61%	1.78%	2.52%	3.87%	1.78%	4.61%	3.87%	4.35%
HRV	3.57%	2.18%	1.38%	0.89%	2.67%	1.38%	2.18%	2.67%	4.42%
HUN	3.41%	1.83%	1.58%	0.61%	2.80%	1.58%	1.83%	2.80%	3.17%
IDN	7.85%	5.53%	2.31%	2.08%	5.77%	2.31%	5.53%	5.77%	6.75%
IND	11.38%	6.83%	4.55%	2.28%	9.11%	4.55%	6.83%	9.11%	5.98%
IRL	1.47%	0.95%	0.52%	0.57%	0.89%	0.52%	0.95%	0.89%	1.97%
ITA	2.22%	0.93%	1.29%	0.28%	1.93%	1.29%	0.93%	1.93%	2.13%
JPN	2.85%	1.38%	1.47%	0.32%	2.53%	1.47%	1.38%	2.53%	3.36%
KOR	4.23%	1.61%	2.61%	0.38%	3.85%	2.61%	1.61%	3.85%	3.06%
LTU	4.06%	2.41%	1.65%	1.00%	3.06%	1.65%	2.41%	3.06%	2.52%
LUX	1.15%	0.51%	0.64%	0.35%	0.80%	0.64%	0.51%	0.80%	3.05%
LVA	3.43%	2.15%	1.28%	1.07%	2.36%	1.28%	2.15%	2.36%	2.00%
MEX	5.59%	3.60%	1.99%	1.02%	4.57%	1.99%	3.60%	4.57%	3.15%
MLT	1.82%	0.64%	1.18%	0.17%	1.65%	1.18%	0.64%	1.65%	3.38%
NLD	2.25%	1.14%	1.12%	0.51%	1.74%	1.12%	1.14%	1.74%	2.63%
NOR	1.81%	1.31%	0.51%	0.58%	1.23%	0.51%	1.31%	1.23%	1.04%
POL	5.84%	3.44%	2.40%	0.98%	4.86%	2.40%	3.44%	4.86%	4.14%
PRT	3.77%	2.13%	1.64%	0.70%	3.07%	1.64%	2.13%	3.07%	3.03%
ROU	4.10%	2.19%	1.91%	0.69%	3.42%	1.91%	2.19%	3.42%	2.38%
RUS	12.79%	8.55%	4.24%	1.44%	11.34%	4.24%	8.55%	11.34%	5.72%
SVK	3.29%	1.62%	1.66%	0.42%	2.87%	1.66%	1.62%	2.87%	3.36%
SVN	2.79%	1.51%	1.28%	0.47%	2.32%	1.28%	1.51%	2.32%	2.32%
SWE	1.21%	0.59%	0.62%	0.21%	1.00%	0.62%	0.59%	1.00%	1.14%
TUR	5.78%	3.73%	2.05%	1.39%	4.39%	2.05%	3.73%	4.39%	2.50%
TWN	5.16%	2.21%	2.95%	0.75%	4.41%	2.95%	2.21%	4.41%	3.73%
USA	2.17%	1.40%	0.78%	0.34%	1.83%	0.78%	1.40%	1.83%	1.30%
ROW	7.55%	5.14%	2.40%	1.87%	5.68%	2.40%	5.14%	5.68%	3.82%

Table 8.53: Economic impact of an EU carbon tax ($100/tCO$_2$e, Exiobase 2022)

Region	Cost						Revenue	Inflation	
	T_{total}	T_{direct}	T_{indirect}	T_{producer}	$T_{\text{downstream}}$	T_{net}	$R_{\text{government}}$	π_{ppi}	π_{cpi}
World	0.36%	0.22%	0.14%	0.07%	0.28%	0.14%	0.22%	0.28%	0.48%
AUT	1.80%	0.91%	0.89%	0.30%	1.50%	0.89%	0.91%	1.50%	1.54%
BEL	1.29%	0.94%	0.34%	0.44%	0.85%	0.34%	0.94%	0.85%	1.06%
BGR	6.30%	3.94%	2.35%	0.89%	5.41%	2.35%	3.94%	5.41%	3.16%
CYP	4.86%	3.94%	0.92%	2.49%	2.37%	0.92%	3.94%	2.37%	3.32%
CZE	3.90%	2.13%	1.76%	0.44%	3.46%	1.76%	2.13%	3.46%	4.07%
DEU	1.72%	1.10%	0.62%	0.35%	1.38%	0.62%	1.10%	1.38%	1.30%
DNK	1.28%	0.98%	0.30%	0.54%	0.74%	0.30%	0.98%	0.74%	1.18%
ESP	1.82%	1.15%	0.68%	0.41%	1.41%	0.68%	1.15%	1.41%	1.20%
FIN	2.27%	1.36%	0.91%	0.36%	1.91%	0.91%	1.36%	1.91%	1.32%
FRA	1.15%	0.79%	0.36%	0.35%	0.81%	0.36%	0.79%	0.81%	5.95%
GRC	5.64%	4.61%	1.03%	2.52%	3.12%	1.03%	4.61%	3.12%	3.59%
HRV	2.88%	2.18%	0.70%	0.89%	1.99%	0.70%	2.18%	1.99%	3.83%
HUN	2.70%	1.83%	0.87%	0.61%	2.08%	0.87%	1.83%	2.08%	2.20%
IRL	1.17%	0.95%	0.22%	0.57%	0.59%	0.22%	0.95%	0.59%	1.56%
ITA	1.57%	0.93%	0.65%	0.28%	1.29%	0.65%	0.93%	1.29%	1.49%
LTU	3.22%	2.41%	0.82%	1.00%	2.22%	0.82%	2.41%	2.22%	2.11%
LUX	0.71%	0.51%	0.20%	0.35%	0.36%	0.20%	0.51%	0.36%	2.30%
LVA	3.11%	2.15%	0.96%	1.07%	2.05%	0.96%	2.15%	2.05%	1.69%
MLT	1.30%	0.64%	0.66%	0.17%	1.13%	0.66%	0.64%	1.13%	2.69%
NLD	1.48%	1.14%	0.34%	0.51%	0.97%	0.34%	1.14%	0.97%	2.11%
POL	5.21%	3.44%	1.77%	0.98%	4.24%	1.77%	3.44%	4.24%	3.49%
PRT	3.28%	2.13%	1.15%	0.70%	2.58%	1.15%	2.13%	2.58%	2.58%
ROU	3.60%	2.19%	1.41%	0.69%	2.91%	1.41%	2.19%	2.91%	1.90%
SVK	2.72%	1.62%	1.09%	0.42%	2.30%	1.09%	1.62%	2.30%	2.80%
SVN	2.38%	1.51%	0.87%	0.47%	1.91%	0.87%	1.51%	1.91%	1.90%
SWE	0.93%	0.59%	0.34%	0.21%	0.72%	0.34%	0.59%	0.72%	0.74%
AUS	0.01%	0.00%	0.01%	0.00%	0.01%	0.01%	0.00%	0.01%	0.02%
BRA	0.02%	0.00%	0.02%	0.00%	0.02%	0.02%	0.00%	0.02%	0.03%
CAN	0.02%	0.00%	0.02%	0.00%	0.02%	0.02%	0.00%	0.02%	0.03%
CHE	0.10%	0.00%	0.10%	0.00%	0.10%	0.10%	0.00%	0.10%	0.12%
CHN	0.01%	0.00%	0.01%	0.00%	0.01%	0.01%	0.00%	0.01%	0.01%
GBR	0.06%	0.00%	0.06%	0.00%	0.06%	0.06%	0.00%	0.06%	0.10%
IDN	0.01%	0.00%	0.01%	0.00%	0.01%	0.01%	0.00%	0.01%	0.01%
IND	0.01%	0.00%	0.01%	0.00%	0.01%	0.01%	0.00%	0.01%	0.02%
JPN	0.01%	0.00%	0.01%	0.00%	0.01%	0.01%	0.00%	0.01%	0.01%
KOR	0.02%	0.00%	0.02%	0.00%	0.02%	0.02%	0.00%	0.02%	0.01%
MEX	0.02%	0.00%	0.02%	0.00%	0.02%	0.02%	0.00%	0.02%	0.03%
NOR	0.10%	0.00%	0.10%	0.00%	0.10%	0.10%	0.00%	0.10%	0.14%
RUS	0.02%	0.00%	0.02%	0.00%	0.02%	0.02%	0.00%	0.02%	0.03%
TUR	0.06%	0.00%	0.06%	0.00%	0.06%	0.06%	0.00%	0.06%	0.09%
TWN	0.03%	0.00%	0.03%	0.00%	0.03%	0.03%	0.00%	0.03%	0.02%
USA	0.01%	0.00%	0.01%	0.00%	0.01%	0.01%	0.00%	0.01%	0.01%
ROW	0.04%	0.00%	0.04%	0.00%	0.04%	0.04%	0.00%	0.04%	0.04%

Table 8.54: Economic impact of a US carbon tax ($100/tCO$_2$e, Exiobase 2022)

Region	Cost						Revenue	Inflation	
	T_{total}	T_{direct}	T_{indirect}	T_{producer}	$T_{\text{downstream}}$	T_{net}	$R_{\text{government}}$	π_{ppi}	π_{cpi}
World	0.44	0.29	0.14	0.07	0.37	0.14	0.29	0.37	0.27
BRA	0.05	0.00	0.05	0.00	0.05	0.05	0.00	0.05	0.04
CAN	0.18	0.00	0.18	0.00	0.18	0.18	0.00	0.18	0.16
GBR	0.03	0.00	0.03	0.00	0.03	0.03	0.00	0.03	0.04
IND	0.04	0.00	0.04	0.00	0.04	0.04	0.00	0.04	0.03
IRL	0.06	0.00	0.06	0.00	0.06	0.06	0.00	0.06	0.05
KOR	0.07	0.00	0.07	0.00	0.07	0.07	0.00	0.07	0.03
MEX	0.18	0.00	0.18	0.00	0.18	0.18	0.00	0.18	0.16
NLD	0.03	0.00	0.03	0.00	0.03	0.03	0.00	0.03	0.03
TWN	0.04	0.00	0.04	0.00	0.04	0.04	0.00	0.04	0.02
USA	1.96	1.40	0.57	0.34	1.62	0.57	1.40	1.62	1.06
ROW	0.04	0.00	0.04	0.00	0.04	0.04	0.00	0.04	0.04

Table 8.55: Economic impact of a carbon tax in China ($100/tCO$_2$e, Exiobase 2022)

Region	Cost						Revenue	Inflation	
	T_{total}	T_{direct}	T_{indirect}	T_{producer}	$T_{\text{downstream}}$	T_{net}	$R_{\text{government}}$	π_{ppi}	π_{cpi}
World	1.66	0.81	0.85	0.29	1.38	0.85	0.81	1.38	1.15
AUS	0.03	0.00	0.03	0.00	0.03	0.03	0.00	0.03	0.07
BEL	0.03	0.00	0.03	0.00	0.03	0.03	0.00	0.03	0.04
BRA	0.03	0.00	0.03	0.00	0.03	0.03	0.00	0.03	0.04
CAN	0.03	0.00	0.03	0.00	0.03	0.03	0.00	0.03	0.07
CHN	6.89	3.44	3.45	1.21	5.68	3.45	3.44	5.68	5.88
CZE	0.04	0.00	0.04	0.00	0.04	0.04	0.00	0.04	0.02
DEU	0.03	0.00	0.03	0.00	0.03	0.03	0.00	0.03	0.02
HUN	0.04	0.00	0.04	0.00	0.04	0.04	0.00	0.04	0.02
IDN	0.05	0.00	0.05	0.00	0.05	0.05	0.00	0.05	0.04
IND	0.05	0.00	0.05	0.00	0.05	0.05	0.00	0.05	0.07
JPN	0.04	0.00	0.04	0.00	0.04	0.04	0.00	0.04	0.04
KOR	0.12	0.00	0.12	0.00	0.12	0.12	0.00	0.12	0.08
MEX	0.06	0.00	0.06	0.00	0.06	0.06	0.00	0.06	0.06
POL	0.04	0.00	0.04	0.00	0.04	0.04	0.00	0.04	0.04
SVK	0.03	0.00	0.03	0.00	0.03	0.03	0.00	0.03	0.02
SVN	0.03	0.00	0.03	0.00	0.03	0.03	0.00	0.03	0.02
TUR	0.04	0.00	0.04	0.00	0.04	0.04	0.00	0.04	0.05
ROW	0.13	0.00	0.13	0.00	0.13	0.13	0.00	0.13	0.16

Table 8.56: Economic impact of a carbon tax in India ($100/tCO$_2$e, Exiobase 2022)

Region	Cost						Revenue	Inflation	
	T_{total}	T_{direct}	T_{indirect}	T_{producer}	$T_{\text{downstream}}$	T_{net}	$R_{\text{government}}$	π_{ppi}	π_{cpi}
World	0.40	0.25	0.15	0.08	0.31	0.15	0.25	0.31	0.23
BEL	0.03	0.00	0.03	0.00	0.03	0.03	0.00	0.03	0.05
IND	10.52	6.83	3.69	2.28	8.24	3.69	6.83	8.24	5.35
KOR	0.03	0.00	0.03	0.00	0.03	0.03	0.00	0.03	0.03
LUX	0.03	0.00	0.03	0.00	0.03	0.03	0.00	0.03	0.06
TUR	0.03	0.00	0.03	0.00	0.03	0.03	0.00	0.03	0.05
TWN	0.03	0.00	0.03	0.00	0.03	0.03	0.00	0.03	0.02
ROW	0.06	0.00	0.06	0.00	0.06	0.06	0.00	0.06	0.07

8.5 Exercises

Chapter 9

Climate Risk Measures

In this chapter, we list the various metrics that are helpful in assessing climate risk. We focus on traditional metrics related to the concept of carbon footprint. This concept seems easy to understand, but difficult to define precisely. In their seminal book on the ecological footprint, Wackernagel and Rees (1996) stated that "*the carbon footprint stands for a certain amount of gaseous emissions that are relevant to climate change and associated with human production or consumption activities*". Wiedmann and Minx (2008, pages 2-4) listed several definitions found in the scientific literature and the grey literature (NGOs, consultants, data providers, etc.) between 1960 and 2007. They also proposed a definition of their own:

> "*The carbon footprint is a measure of the exclusive total amount of carbon dioxide emissions that is directly and indirectly caused by an activity or is accumulated over the life stages of a product.*" (Wiedmann and Minx, 2008, page 5).

Note that this definition includes only carbon dioxide emissions. The authors excluded methane, because they felt that it should be reserved for another measure, that could be called climate footprint. Furthermore, they pointed out that the carbon footprint should be measured physically in a unit of mass (*e.g.*, kg) and not converted into a land area unit (*e.g.*, ha). This digression shows that the concepts of climate risk measurement are very recent[1] and not necessarily clear and stable. In the first section, we will then discuss how to measure carbon emissions. In particular, we will distinguish the three scopes, define the carbon dioxide equivalent (CO_2e) unit based on global warming potential (GWP) values, and present some examples of carbon footprint calculations. We will examine some specific sectors, such as transportation and agriculture, and the two main approaches: activity-based and energy-based methods. In the second section, we will analyze the concept of carbon intensity, which is a normalization of the carbon emissions measure in order to compare countries, companies or portfolios. In the case of physical intensity ratios, total emissions are divided by physical quantities. For instance, it is common to measure the carbon footprint of transportation in terms of $kgCO_2e$ per kilometer traveled, while the carbon footprint of a country is typically measured in terms of tCO_2e per capita. The second way to calculate carbon intensity is to use monetary quantities in the denominator. In the previous example, we can divide the total carbon emissions by the revenue generated by the transportation activity or the GDP of the country. Carbon intensity based on monetary units plays an important role in finance and especially in portfolio management. While carbon emissions and intensity are two static measures, in the third section we will develop some dynamic risk measures based on carbon budget and trend. Finally, the last section is dedicated to green intensity measures, whose objective is to assess the greenness of products, countries, companies or assets.

---

[1]The publication of this much cited research paper was only fifteen years ago.

9.1   Carbon emissions

Carbon footprint is a generic term used to define the total greenhouse gas (GHG) emissions caused by a given system, activity, company, country, or region. Greenhouse gases are made up of water vapor (H_2O), carbon dioxide (CO_2), methane (CH_4), nitrous oxide (N_2O), Ozone (O_3), etc. They absorb and emit radiation energy, causing the greenhouse effect. We remind that the greenhouse effect was a crucial factor for the development of human life on Earth. Indeed, without the greenhouse effect, the average temperature of Earth's surface would be about $-18°$C. With the greenhouse effect, the current temperature of Earth's surface is about $+15°$C. Nevertheless, the increasing concentration of some GHGs is an issue because it is a factor in global warming. It mainly concerns carbon dioxide, and to a lesser extent, methane and nitrous oxide.

9.1.1   Global warming potential

Carbon footprint is generally measured in carbon dioxide equivalent (CO_2e), which is a term for describing different GHGs in a common unit. In this framework, a quantity of GHG is expressed as CO_2e by multiplying the GHG amount by its global warming potential (GWP):

$$\text{equivalent mass of } CO_2 = \text{mass of the gas} \times \text{gwp of the gas}$$

where the GWP of a gas is the amount of CO_2 that would warm the earth equally. Since the mass of the gas is expressed in kilogram and the GWP has no unit, the mass of CO_2 equivalent is also expressed in kilogram. For instance, the IPCC's 5th assessment report has used the following rules (IPCC, 2014): 1 kg of methane corresponds to 28 kg of CO_2 and 1 kg of nitrous oxide corresponds to 265 kg of CO_2. The definition of a common unit allows two companies to be compared properly. To compute the carbon footprint of a system that is made up of several gases, we apply the weighted sum formula:

$$m = \sum_{i=1}^{n} m_i \cdot \text{gwp}_i$$

where m is the mass of CO_2 equivalent, m_i and gwp_i are the mass and the global warming potential of the ith gas, and n is the number of gases. m and m_i have the same mass unit (eg., kilogram or kg, tonne or t, kilotonne or kt, megatonne or Mt, gigatonne or Gt). However, m measures a mass of CO_2 equivalent. Therefore, it better to use the following units: $kgCO_2$e, tCO_2e, $ktCO_2$e, $MtCO_2$e and $GtCO_2$e.

Example 25 *We consider a company A that emits* 3 017 *tonnes of* CO_2, 10 *tonnes of* CH_4 *and* 1.8 *tonnes of* N_2O. *For the company B, the* GHG *emissions are respectively equal to* 2 302 *tonnes of* CO_2, 32 *tonnes of* CH_4 *and* 3.0 *tonnes of* N_2O.

The mass of CO_2 equivalent for companies A and B is equal to:

$$m_A = 3017 \times 1 + 10 \times 28 + 1.8 \times 265 = 3\,774 \text{ tCO}_2\text{e}$$

and:

$$m_B = 2302 \times 1 + 32 \times 28 + 3.0 \times 265 = 3\,993 \text{ tCO}_2\text{e}$$

We notice that company B emits more carbon emissions than company A when they are measured in CO_2 equivalent. We can also compute the mass contribution of each gas:

$$c_i = \frac{m_i \cdot \text{gwp}_i}{m}$$

The mass decomposition is reported below:

Company	Mass	Absolute contribution			Relative contribution		
A	3 774	3 017	280	477	79.94%	7.42%	12.64%
B	3 993	2 302	896	795	57.65%	22.44%	19.91%

The contribution of the carbon dioxide gas is equal to 79.94% for company A and 57.65% for company B. Concerning methane gas, its contribution is respectively equal to 7.42% and 22.44%.

Box 9.1: Estimation of the global warming potential

According to IPCC (2007), GWP is defined as "*the cumulative radiative forcing, both direct and indirect effects, over a specified time horizon resulting from the emission of a unit mass of gas related to some reference gas*". Since each gas differs in their capacity to absorb the energy (radiative efficiency) and how long it stays in the atmosphere (lifetime), its impact on global warming depends on these two factors. GWP is then a synthetic measure that combines radiative efficiency and lifetime.

The mathematical definition of the global warming potential is:

$$\text{gwp}_i(t) = \frac{\text{Agwp}_i(t)}{\text{Agwp}_0(t)} = \frac{\int_0^t RF_i(s)\,\mathrm{d}s}{\int_0^t RF_0(s)\,\mathrm{d}s} = \frac{\int_0^t A_i(s)\,\mathbf{S}_i(s)\,\mathrm{d}s}{\int_0^t A_0(s)\,\mathbf{S}_0(s)\,\mathrm{d}s} \tag{9.1}$$

where $A_i(t)$ is the radiative efficiency value of gas i (or the radiative forcing increase per unit mass increase of gas i in the atmosphere), $\mathbf{S}_i(t)$ is the decay function (or the fraction of gas i remaining in the atmosphere after t years following an incremental pulse of the gas) and $i = 0$ is the reference gas (e.g, CO_2). The radiative forcing $RF_i(t) = A_i(t)\,\mathbf{S}_i(t)$ is the product of the radiative efficiency and the decay function, whereas the absolute global warming potential $\text{Agwp}_i(t)$ is the cumulative radiative forcing of the gas i between 0 and t. GWP is then the ratio between the cumulative radiative forcing of the gas and this of the reference gas. We also notice that it depends on the time horizon t.

It is generally accepted to describe the decay function (or impulse response function) by exponential functions (Joos *et al.*, 2013):

$$\mathbf{S}_i(t) = \sum_{j=1}^m a_{i,j} e^{-\lambda_{i,j} t} \tag{9.2}$$

where $\sum_{j=1}^m a_{i,j} = 1$. Once we have defined the radiative efficiency function $A_i(t)$ and the set of parameters $\{(a_{i,j}, \lambda_{i,j}), j = 1, \ldots, m\}$ of the impulse function, we compute Equation (9.1) using numerical integration. In the case where $A_i(t)$ and $A_0(t)$ are constant, we obtain:

$$\text{gwp}_i(t) = \frac{A_i \sum_{j=1}^m a_{i,j} \lambda_{i,j}^{-1} \left(1 - e^{-\lambda_{i,j} t}\right)}{A_0 \sum_{j=1}^m a_{0,j} \lambda_{0,j}^{-1} \left(1 - e^{-\lambda_{0,j} t}\right)}$$

In Box 9.1, we explain how GWP is computed. We notice that the global warming potential value depends on the time horizon. For instance, the relative warming impact of one molecule of a greenhouse gas is not the same at 20 years than 100 years. We also notice that the estimation of GWP implies to make some assumptions about the impulse response (or decay) function and the radiative efficiency value. Let us see how the value for methane has been obtained. IPCC (2013)

assumed that the radiative intensity is constant and used the following values[2]: $A_{CO_2} = 1.76 \times 10^{-18}$ and $A_{CH_4} = 2.11 \times 10^{-16}$. The impulse response functions were estimated by least squares and they found the following approximated curve:

$$\mathbf{S}_{CO_2}(t) = 0.2173 + 0.2240 \cdot \exp\left(-\frac{t}{394.4}\right) + 0.2824 \cdot \exp\left(-\frac{t}{36.54}\right) + 0.2763 \cdot \exp\left(-\frac{t}{4.304}\right)$$

and:

$$\mathbf{S}_{CH_4}(t) = \exp\left(-\frac{t}{12.4}\right)$$

These two decay functions are reported in Figure 9.1. We can interpret them as survival functions[3], meaning that the density function can be computed as $f_i(t) = -\partial_t \mathbf{S}_i(t)$.

Figure 9.1: Fraction of gas remaining in the atmosphere

In the case of the exponential distribution $\mathcal{E}(\lambda)$, we have $\mathbf{S}_i(t) = e^{-\lambda t}$ and $f_i(t) = \lambda e^{-\lambda t}$ where λ is the rate parameter. Let τ_i be random time that the gas remains in the atmosphere. We have $\mathbb{E}[\tau_i] = 1/\lambda$ for the exponential random time. The survival function of the CH_4 gas is exponential with a mean time equal to 12.4 years ($\lambda = 1/12.4$). In the case of the general formula (9.2), the probability density function is equal to:

$$f_i(t) = -\partial_t \mathbf{S}_i(t) = \sum_{j=1}^{m} a_{i,j} \lambda_{i,j} e^{-\lambda_{i,j} t}$$

---

[2]The unit is $W \cdot m^{-2} g^{-1}$.
[3]This is why we use the notation $\mathbf{S}(t)$.

Figure 9.2: Probability density function of the random time

and the mean time $\mathcal{T}_i$ is given by:

$$
\begin{aligned}
\mathcal{T}_i := \mathbb{E}\left[\tau_i\right] &= \int_0^\infty s f_i\left(s\right) \, \mathrm{d}s \\
&= \sum_{j=1}^m a_{i,j} \int_0^\infty \lambda_{i,j} s e^{-\lambda_{i,j} s} \, \mathrm{d}s \\
&= \sum_{j=1}^m \frac{a_{i,j}}{\lambda_{i,j}}
\end{aligned}
$$

Another way to find this result is to notice that $f_i\left(t\right)$ is an exponential mixture distribution where m is the number of mixture components, $\mathcal{E}\left(\lambda_{i,j}\right)$ is the probability distribution associated with the jth component and $a_{i,j}$ is the mixture weight of the jth component. Therefore, we deduce that the mean time is equal to the weighted average of the mean times of the mixture components:

$$
\mathcal{T}_i = \mathbb{E}\left[\tau_i\right] = \sum_{j=1}^m a_{i,j} \mathbb{E}\left[\tau_{i,j}\right] = \sum_{j=1}^m a_{i,j} \mathcal{T}_{i,j}
$$

For the CO_2 gas, the exponential mixture distribution is defined by the following parameters:

j	1	2	3	4
$a_{i,j}$	0.2173	0.2240	0.2824	0.2763
$\lambda_{i,j}$ $(\times 10^3)$	0.00	2.535	27.367	232.342
$\mathcal{T}_{i,j}$ (in years)	∞	394.4	36.54	4.304

We can now explain why the carbon dioxide stays longer in the atmosphere than the methane. When we compare the density functions of the two gases (Figure 9.2), we observe that the disappearance

probabilities are located before 50 years, since the probability to stay in the atmosphere after 50 years is less than 2%. For the CO_2 gas, we have roughly 50% that the molecule disappears and 50% that the molecule stays. In fact, we notice that one mixture component corresponds to a permanent state ($\lambda_{i,1} = 0$) with a weight of 21.73%. This explains that the CO_2 molecule can stay in the atmosphere, and we have $\mathbf{S}_{CO_2}(\infty) = 21.73\%$.

We compute $Agwp_{CO_2}(t)$ and $Agwp_{CH_4}(t)$ and report their values in Figure 9.3. Even if the methane has a much shorter atmospheric lifetime than the carbon dioxide, it absorbs much more energy — because of the value of A_{CH_4} compared to the value of A_{CO_2}. Nevertheless, the absolute global warming potential is unbounded for COtwo because we have $Agwp_{CO_2}(\infty) = \infty$. For the methane, it reaches an upper bound, which corresponds to the ratio $A_{CH_4} \times \mathcal{T}_{CH_4} \propto 2.11 \times 12.4 = 26.164$. Therefore, $gwp_{CH_4}(t)$ is a decreasing function with respect to the time horizon (Figure 9.4). The instantaneous global warming potential of the methane is equal to:

$$gwp_{CH_4}(0) = \frac{A_{CH_4}}{A_{CO_2}} = \frac{2.11 \times 10^{-16}}{1.76 \times 10^{-18}} \approx 119.9$$

After 100 years, we obtain $gwp_{CH_4}(100) = 28.3853$, which is the value calculated by IPCC (2013, 2014). Because of the persistant regime of the carbon dioxyde, we have $gwp_{CH_4}(\infty) = 0$. In fact, the global warming potential of CO_2 becomes greater than this of CH_4 when $t \geq 6\,382$ years.

The previous analysis shows that the estimation of the global warming potential involves significant scientific uncertainty. First, the choice of a 100-year time horizon is arbitrary, and any other choice will change the GWP value. Second, the survival function $\mathbf{S}_i(t)$ is estimated and based on empirical experiments. Third, we have assumed that the radiative efficiency $A_i(t)$ is constant and equal to the initial value $A_i(0)$. In this context, we can consider that $gwp_i(t)$ is stochastic or cannot be observed without any error. The GHG protocol considers the six gases listed in the Kyoto Protocol: carbon dioxide, methane, nitrous oxide, sulphur hexafluoride, hydrofluorocarbons, and perfluorocarbons. Table 9.1 gives their GWP values according to the different IPCC reports. We notice that they have continuously changed.

Table 9.1: GWP values for 100-year time horizon

Name	Formula	AR2	AR4	AR5	AR6
Carbon dioxide	CO_2	1	1	1	1
Methane	CH_4	21	25	28	27.9
Nitrous oxide	N_2O	310	298	265	273
Sulphur hexafluoride	SF_6	23 900	22 800	23 500	25 200
Hydrofluorocarbons (HFC)	CHF_3	11 700	14 800	12 400	14 600
	CH_2F_2	650	675	677	771
	Etc.				
Perfluorocarbons (PFC)	CF_4	6 500	7 390	6 630	7 380
	C_2F_6	9 200	12 200	11 100	12 400
	Etc.				

Remark 68 *The GWP has been subjected to many criticisms because it does not directly measure the impact on the temperature and it is a single-pulse emission metric (Kleinberg, 2020). The alternative metric is the global temperature potential (GTP) proposed by Shine et al. (2005). While the GWP is a measure of the energy absorbed over a given time period, the GTP is a measure of the temperature change at the end of that time period. The adoption of the GTP is however extremely rare, because its calculation is complicated and it appeared after the Kyoto Protocol.*

Figure 9.3: Absolute global warming potential

Source: Kleinberg (2020) & Author's calculations.

Figure 9.4: Global warming potential for methane

$\text{gwp}_{\text{CH}_4}(100) = 28.39$

Source: Kleinberg (2020) & Author's calculations.

9.1.2   Consolidation accounting at the company level

Greenhouse gas accounting requires to define rules about what is reported and what is not reported. For instance, we have seen that the GHG protocol does not consider all greenhouse gases. In a similar way, the consolidation of GHG emissions (also named organizational boundary) follows specific principles, which are similar to those we can find in financial consolidation accounting. For corporate reporting, the GHG protocol distinguishes two approaches:

1. Equity share approach

2. Control approach

Under the first approach, a parent company must report carbon emissions of a subsidiary company according to its share of equity (or ownership ratio). This is the simplest accounting method. For example, if company A owns 25% of company B, company A have to take into account 25% of the company B's GHG emissions. Under the second approach, it can use either the financial control method or the operational control method. The company financially controls an operation if it bears the majority risks and rewards of this operation[4], whereas it has operational control if it has the full authority to implement the operation. The control approach is based on the all-or-none principle: Company A includes 100% of the company B's GHG emissions if A controls B, otherwise it includes 0%. In Table 9.2, we report the three accounting principles. By definition, the company has financial (and operational) control on group companies or subsidiaries. This explains that 100% of GHG emissions are consolidated. Associated and affiliated companies differ from the previous categories, because the company do not have the financial control. In this case, no GHG emissions are consolidated under the financial control method. Nevertheless, the company may have operational control, which explains that 100% of GHG emissions may be consolidated[5]. For joint ventures and partnerships, we apply the equity share principle for the financial control method, and we use the same rule than for associated and affiliated companies when we consider the financial control approach. Since the category fixed assets correspond to investments, the company receives dividends but has no control and GHG emissions are not consolidated. The same case applies to franchises because they are sperate legal entities and franchisers have no equity rights or control. In the opposite situation, the franchise is considered as a subsidiary.

Table 9.2: Percent of reported GHG emissions under each consolidation method

Accounting categories	GHG accouting based on		
	equity share	financial control	operational control
Wholly owned asset	100%	100%	100%
Group companies/subsidiaries	Ownership ratio	100%	100%
Associated/affiliated companies	Ownership ratio	0%	0%/100%
Joint ventures/partnerships	Ownership ratio	Ownership ratio	0%/100%
Fixed asset investments	0%	0%	0%
Franchises	0%	0%	0%
	Ownership ratio	100%	100%

Source: GHG Protocol (2004, Table 1, page 19).

Remark 69 *An illustration of the differences between the equity share and control approaches is given in GHG Protocol (2004) on pages 22 and 23. This concerns the Holland Industries Group.*

---

[4]For instance when the company has more than 50% voting rights.

[5]If the company has no operational control, GHG emissions are not consolidated.

Example 26 *We report the organizational structure of company A in Figure 9.5. This industrial group has several subsidiaries, partnerships and joint ventures. For instance, Companies B and C are integrated to company A, implying that the latter has financial and operational control on them. Company D is in the same situation even if company A has not the majority of capital. Indeed, company A treats company D as a subsidiary in its financial accounts, because the remaining capital is diluted and they control the management. The participation in company E is an investment. Company C has two joint ventures. For company F, it owns 75% of the capital, while company G is held in equal proportion by company A and another partner. Finally, company H is affiliated to company D, which has no financial or operational control.*

Figure 9.5: Defining the organizational boundary of company A

For each company, the brown number corresponds to the carbon emissions in tCO$_2$e. The three figures at the right or left of the node corresponds respectively to the equity share, the financial control and the operational control.

Computing the carbon emissions reported by company A will depend on the accounting method. In the case of the equity share approach, we exclude the investment in company H and obtain:

$$
\begin{aligned}
\mathcal{CE}_A &= 827 + 100\% \times 135 + 90\% \times 261 + 45\% \times 220 + 0\% \times 1\,385 + \\
&\quad 90\% \times 75\% \times 63 + 90\% \times 50\% \times 179 + 45\% \times 33\% \times 37 \\
&= 1\,424.4\text{tCO}_2\text{e}
\end{aligned}
$$

If we use the financial control approach, the reported carbon emissions become:

$$
\begin{aligned}
\mathcal{CE}_A &= 827 + 100\% \times 135 + 100\% \times 261 + 100\% \times 220 + 0\% \times 1\,385 + \\
&\quad 100\% \times 100\% \times 63 + 100\% \times 50\% \times 179 + 100\% \times 0\% \times 37 \\
&= 1\,595.50\text{tCO}_2\text{e}
\end{aligned}
$$

With the operational control approach, they are slightly different from above because of the treatment of company G:

$$
\begin{aligned}
\mathcal{CE}_A &= 827 + 100\% \times 135 + 100\% \times 261 + 100\% \times 220 + 0\% \times 1385 + \\
&\quad 100\% \times 100\% \times 63 + 100\% \times 0\% \times 179 + 100\% \times 0\% \times 37 \\
&= 1\,506.00\text{tCO}_2\text{e}
\end{aligned}
$$

Generally, an equity share of 50% or more induces a financial/operational control, which implies that 100% of carbon emissions are consolidated and reported by the parent company. By contrast, the consolidation factor may be equal to 0% if the parent company has no control, even in the case it has a significant equity share (*e.g.*, between 30% and 40%).

9.1.3   Scope 1, 2 and 3 emissions

The GHG Protocol corporate standard classifies a company's greenhouse gas emissions in three scopes[6]:

- Scope 1 denotes direct GHG emissions occurring from sources that are owned and controlled by the issuer.

- Scope 2 corresponds to the indirect GHG emissions from the consumption of purchased electricity, heat or steam.

- Scope 3 are other indirect emissions (not included in scope 2) of the entire value chain. They can be divided into two main categories[7]:

    - Upstream scope 3 emissions are defined as indirect carbon emissions related to purchased goods and services.

    - Downstream scope 3 emissions are defined as indirect carbon emissions related to sold goods and services.

Scope 1 emissions are also called direct emissions, whereas indirect emissions encompass both scope 2 and 3 GHG emissions. Unlike scope 1 and 2, scope 3 is an optional reporting category.

Remark 70 *The GHG protocol defines "the operational boundary as the scope of direct and indirect emissions for operations that fall within a company's established organizational boundary. The operational boundary (scope 1, scope 2, scope 3) is decided at the corporate level after setting the organizational boundary. The selected operational boundary is then uniformly applied to identify and categorize direct and indirect emissions at each operational level. The established organizational and operational boundaries together constitute a company's inventory boundary."*

Before explaining in fine detail the different scopes and their computation, we report six examples of carbon footprint reporting in Table 9.3. We consider the CDP database, since most of companies reporting to CDP use the GHG protocol framework. The CDP reporting framework is based on a questionnaire (see Box 9.2 on page 480). We first notice all the figures are not calculated. This is normal since scope 3 is not mandatory in the GHG protocol framework. When the sub-category is empty, we don't know whether it is equal to zero or a missing value, meaning that the company has not the capacity or implemented the method to compute it. In the case of Amazon, the sub-category end-of-life treatment of sold products is equal to zero and not an empty case. A second remark concerns the definition of the scope 2, because there are two approaches: location-based and market-based. How to read this table? If we consider the first company Amazon, its scope 1 emissions are equal to 9.62 MtCO$_2$e. For the scope 2 emissions, they are equal to 9.02 MtCO$_2$e when they are calculated with the location-based method or 5.27 MtCO$_2$e when they are calculated

---

[6]The latest version of corporate accounting and reporting standard can be found at `www.ghgprotocol.org/corporate-standard`.

[7]The upstream value chain includes all activities related to the suppliers whereas the downstream value chain refers to post-manufacturing activities.

Table 9.3: Examples of CDP reporting (carbon emissions in tCO$_2$e, reporting year 2020)

Scope	Category	Sub-category	Amazon	Danone	ENEL	Pfizer	Netflix	Walmart
1			9 623 138	668 354	45 255 000	654 460	30 883	7 236 499
2	Location-based (2a)		9 019 786	864 710	4 990 685	551 577	28 585	11 031 800
	Market-based (2b)		5 265 089	479 210	7 855 954	542 521	141	9 190 337
3	Upstream	Purchased goods and services	16 683 423	19 920 918		2 526 537	765 208	130 200 000
		Capital goods	13 202 065			191 894	116 366	645 328
		Fuel and energy related activities	1 248 847	283 764	1 061 268	203 093	12 287	3 327 874
		Upstream transportation and distribution	8 563 695	321 558	112 358	723 558	64 693	342 577
		Waste generated in operations	16 628	152 789	3 161	14 940		869 927
		Business travel	313 043			35 128	41 439	37 439
		Employee commuting	306 033			48 414	19 116	3 500 000
		Upstream leased assets	1 223 903			30 522	131	
	Downstream	Downstream transportation and distribution	2 785 676	1 627 090		7 295		5 099
		Processing of sold products						
		Use of sold products	1 426 543	1 885 548	46 524 860		952	32 211 000
		End-of-life treatment of sold products	0	782 649			349	130
		Downstream leased assets						130 000
		Franchises						
		Investments				36 839		
Total		Scope 1 + 2a	18 642 924	1 533 064	50 245 685	1 206 037	59 468	18 268 299
		Scope 1 + 2b	14 888 227	1 147 564	53 110 954	1 196 981	31 024	16 426 836
		Scope 3 upstream	41 557 637	20 679 029	1 176 787	3 774 086	1 019 240	138 923 145
		Scope 3 downstream	4 212 219	4 295 287	46 524 860	44 134	1 301	32 346 229
		Scope 3	45 769 856	24 974 316	47 701 647	3 818 220	1 020 541	171 269 374
		Scope 1 + 2a + 3	64 412 780	26 507 380	97 947 332	5 024 257	1 080 009	189 537 673
		Scope 1 + 2b + 3	60 658 083	26 121 880	100 812 601	5 015 201	1 051 565	187 696 210

Source: CDP database as of 01/07/2022 & Author's computation.

with the market-based method. The fifteen sub-categories of scope 3 emissions are not aggregated, and each item is filled separately. The last of part of the table (from Scope 1 + 2a to Scope 1 + 2a + 3 is not included in the CDP questionnaire. These figures are calculated by summing the different items.

<div align="center">Box 9.2: CDP questionnaire for corporates</div>

In order to report to CDP, companies must fill the CDP questionnaire on climate change, which is available at `www.cdp.net/en/guidance/guidance-for-companies` in HTML, Word and PDF formats. The full questionnaire has 129 pages and 16 sections. Emissions methodology corresponds to section C5, while emissions data are reported in section C6 with the following breakdown: scope 1 emissions (§C6.1), scope 2 emissions (§C6.3), scope 3 emissions (§C6.5) and emissions intensities (§C6.10).

Scope 1 emissions

According to GHG Protocol (2004, page 40)GHG-Protocol, once the inventory boundary has been established, "*companies generally calculate GHG emissions using the following steps: (1) Identify GHG emissions sources; (2) Select a GHG emissions calculation approach; (3) Collect activity data and choose emission factors; (4) Apply calculation tools and (5) Roll-up GHG emissions data to corporate level. The identification step helps to categorize the GHG emission sources*". The identification step consists in categorizing the GHG emissions in the four main source categories:

1. Stationary combustion: combustion of fuels in stationary equipment (*e.g.*, boilers, turbines, heaters, incinerators);

2. Mobile combustion: combustion of fuels in transportation devices (*e.g.*, automobiles, trucks, trains, airplanes, boats);

3. Process emissions: emissions from physical or chemical processes (*e.g.*, cement manufacturing, petrochemical processing, aluminum smelting);

4. Fugitive emissions: intentional and unintentional releases as well as fugitive emissions (*e.g.*, such as equipment leaks, coal piles, wastewater treatment, cooling towers).

Concretely, the company lists all the activities that result in a GHG emission, and allocates them to the three scopes. Then, we apply an emission factor to each activity and each gas:

$$E_{g,h} = A_h \cdot \boldsymbol{\mathcal{EF}}_{g,h}$$

where A_h is the hth activity rate (also called activity data) and $\boldsymbol{\mathcal{EF}}_{g,h}$ is the emission factor for the hth activity and the gth gas. A_h can be measured in volume, weight, distance, duration, surface, frequency, etc. Since $E_{g,h}$ is expressed in tonne, $\boldsymbol{\mathcal{EF}}_{g,h}$ is measured in tonne per activity unit. For instance, if A_h is measured in hectare, $EF_{g,h}$ is measured in tonne per hectare. In fact, the emission factor is a coefficient that attempts to quantify how much of a greenhouse gas is released into the atmosphere by an activity that releases that gas[8]. For each gas, we calculate the total emissions:

$$E_g = \sum_{h=1}^{n_A} E_{g,h} = \sum_{h=1}^{n_A} A_h \cdot \boldsymbol{\mathcal{EF}}_{g,h}$$

where n_A is the number of activities. Finally, we estimate the carbon emissions by applying the right GWP and summing up all the gases:

$$\boldsymbol{\mathcal{CE}} = \sum_{g=1}^{n_G} \text{gwp}_g \cdot E_g$$

where n_G is the number of gases[9]. Therefore, the compact formula is:

$$\boldsymbol{\mathcal{CE}} = \sum_{g=1}^{n_G} \text{gwp}_g \cdot \left(\sum_{h=1}^{n_A} A_h \cdot \boldsymbol{\mathcal{EF}}_{g,h} \right)$$

The carbon footprint of the company can be split into activities:

$$\boldsymbol{\mathcal{CE}} = \sum_{h=1}^{n_A} A_h \left(\sum_{g=1}^{n_G} \text{gwp}_g \cdot \boldsymbol{\mathcal{EF}}_{g,h} \right) = \sum_{h=1}^{n_A} A_h \cdot \boldsymbol{\mathcal{EF}}_h = \sum_{h=1}^{n_A} \boldsymbol{\mathcal{CE}}_h$$

where $\boldsymbol{\mathcal{EF}}_h$ and $\boldsymbol{\mathcal{CE}}_h$ are the global emission factor and carbon emissions related to the hth activity. We can also aggregate several activities:

$$\boldsymbol{\mathcal{CE}}_{\mathcal{A}} = \sum_{h \in \mathcal{A}} \boldsymbol{\mathcal{CE}}_h = \sum_{h \in \mathcal{A}} A_h \cdot \boldsymbol{\mathcal{EF}}_h$$

where $\mathcal{A}$ is the set of activities. It may happen that some emission factors are defined without a reference to a specific gas (*e.g.*, CO_2 or CH_4). In this case, the emission factor is a synthetic measure which already take into account the GWP of the gases:

$$\boldsymbol{\mathcal{EF}}_h = \sum_{g=1}^{n_G} \text{gwp}_g \cdot \boldsymbol{\mathcal{EF}}_{g,h}$$

---

[8]For example, how many kg of GHG are emitted by 1 kWh of natural gas?
[9]n_G is equal to six in the GHG Protocol.

The expression of the carbon footprint becomes:

$$\mathcal{CE} = \sum_{h \in \mathcal{A}_1} A_h \left(\sum_{g=1}^{n_G} \text{gwp}_g \cdot \mathcal{EF}_{g,h} \right) + \sum_{h \in \mathcal{A}_2} A_h \cdot \mathcal{EF}_h$$

where $\mathcal{A}_1$ and $\mathcal{A}_2$ are the sets of activities without and with synthetic emission factors.

The choice of data inputs is codified by IPCC (2006, 2019):

- Tier 1 methods use global default emission factors;

- Tier 2 methods use country-level or region-specific emission factors;

- Tier 3 methods use directly monitored or site-specific emission factors.

We can find emission factors in several sources: IPCC Emission Factor Database (Box 9.3), National Inventory Reports[10] (NIRs), country emission factor databases[11], international agencies or academic publications. In the US, the emission factors are calculated by the Environmental Protection Agency (US EPA). In the UK, this is the National Atmospheric Emissions Inventory (NAEI) agency, which is in charge to define the emission factors. In France, the database is managed by ADEME (Agence de l'Environnement et de la Maîtrise de l'Energie) and contains about 5 300 validated emission factors. Generally, we can download these emission factors in an Excel or PDF file.

Let us see an example. We consider the GHG inventory document[12] published by Enel. The scope 1 is based on the following activities: (1) combustion of fossil fuels in electricity generation activities; (2) combustion of fossil fuels in generators used for electricity generation and distribution activities; (3) combustion of fossil fuels in vehicles under the Company's control; (4) combustion of fuels for heating offices and canteens; (5) CH_4 leakage in gas-fired thermoelectric power plants; (6) SF_6 losses in electricity generation and distribution activities; (7) HFCs gas losses from cooling systems; (8) NF_3 losses from the production of solar panels; (9) Transportation of fuel (LNG and coal) on vessels under own operational control and (10) CH_4 emissions from the decomposition of organic matter in hydroelectric basins. For the calculations, they use the parameter values of IPCC (2006) for emission factors and the GWP figures of IPCC (2014). They obtained the following results in 2021 expressed in $ktCO_2e$:

	CO_2	CH_4	N_2O	NF_3	SF_6	HFCs	Total
Electricity power generation	50 643.54	385.25	98.14	0.014	31.15	10.22	51 168.32
Electricity distribution	208.33	0.24	0.45		111.62		320.64
Real estate	79.87	0.22	1.24				81.30
Total	50 931.72	385.71	99.83	0.014	142.77	10.22	51 750.26

The scope 1 emissions of Enel is then equal to 51.75 $MtCO_2e$. The contribution of CO_2 is the most important since it represents 98.4% of the total emissions, implying that the other gases have a small impact. In terms of activities, GHG emissions are mainly located in the electricity power generation. Buildings has a contribution of 0.2%.

---

[10]The NIR reports can be found at the UNFCCC website: https://unfccc.int/ghg-inventories-annex-i-parties/2021.

[11]Here are some websites: www.epa.gov/climateleadership/ghg-emission-factors-hub (US), https://naei.beis.gov.uk/data/ef-all (UK), https://bilans-ges.ademe.fr (France), www.dcceew.gov.au/climate-change/publications/national-greenhouse-accounts-factors-2021 (Australia), www.isprambiente.gov.it (Italy), https://publications.gc.ca/site/eng/9.911206/publication.html (Canada).

[12]Enel (2022). *Quantification and Reporting of Greenhouse Gas Emissions in Accordance with the Corporate GHG Protocol.* 12th April 2022, www.enel.com/investors/sustainability.

Box 9.3: IPCC emission factor database (EFDB)

The IPCC emission factor database[a] (EFDB) is a database on various parameters to be used in calculation of anthropogenic emissions by sources and removals by sinks of greenhouse gases. It contains the IPCC default data[b], and data from peer-reviewed journals and other publications including national inventory reports (NIRs). The database includes emission factors for five categories:

1. Energy (fuel combustion activities, fugitive emissions from fuels, carbon dioxide transport and storage);

2. Industrial processes and product use (mineral industry, chemical industry, metal industry, non-energy products from fuels and solvent use, electronics industry, product uses as substitutes for ozone depleting substances, other product manufacture and use, other);

3. Agriculture, forestry, and other land use (livestock, land, aggregate sources and non-CO2 emissions sources on land, other);

4. Waste (solid waste disposal, biological treatment of solid waste, incineration and open burning of waste, wastewater treatment and discharge, other)

5. Other (Indirect N_2O emissions from the atmospheric deposition of nitrogen in NO_x and NH_3, other)

Some figures of the library are given in Table 9.4. We notice that they may depend on the region and technical criteria. The unit is also important. For instance, if the emission factors are expressed in tCarbon/TeraJoule (tonne of carbon per terajoule energy), we multiply the emission factor $\mathcal{EF}$ (in tC/TJ) by the energy consumption C (in TJ) to obtain the carbon content (in tonnes of carbon). Since one tonne of CO_2 contains 0.2727 tonne of carbon, we then deduce that the CO_2 emissions are equal to $\dfrac{C \cdot \mathcal{EF}}{0.2727}$ tCO_2e. In this case, the activity data corresponds to the energy consumption.

---

[a]The website is www.ipcc-nggip.iges.or.jp/EFDB.

[b]Revised 1996 IPCC Guidelines, IPCC Good Practice Guidance and Uncertainty Management in National Greenhouse Gas Inventories, IPCC Good Practice Guidance for Land Use, Land-Use Change and Forestry, 2006 IPCC Guidelines for National Greenhouse Gas Inventories and 2013 Supplement to the 2006 IPCC Guidelines for National Greenhouse Gas Inventories: Wetlands.

Table 9.4: Examples of emission factors (EFDB, IPCC)

Category	Description	Gas	Region	Value	Unit
Iron and steel production	Integrated facility	CO_2	Canada	1.6	t/tonne
	Electrode consumption from steel produced in electric arc furnaces	CO_2	Global	5.0	kg/tonne
	Steel processing (rolling mills)	N_2O	Global	40	g/tonne
Manufacture of solid fuels	Metallurgical coke production	CO_2	Global	0.56	t/tonne
		CH_4	Global	0.1	g/tonne
Fuel combustion activities	Crude oil	CO_2	Global	20	tCarbon/TeraJoule
	Natural gas	CO_2	Global	15.3	tCarbon/TeraJoule
	Ethane	CO_2	Global	16.8	tCarbon/TeraJoule
Integrated circuit or semi-conductor	Semiconductor manufacturing (silicon)	CF_4	Global	0.9	kg/m²
Cement production	Cement production	CO_2	Global	0.4985	t/tonne
Horses	Enteric fermentation	CH_4	Global	18	kg/head/year
	Manure management (annual average temperature is less than 15oC)	CH_4	Developed countries	1.4	kg/head/year
	Manure management (annual average temperature is between 15oC and 25oc)	CH_4	Developed countries	2.1	kg/head/year
Buffalo	Enteric fermentation	CH_4	Global	55	kg/head/year
	Manure management (annual average temperature is less than 15oC)	CH_4	Developed countries	0.078	kg/head/year
Poultry	Manure management (annual average temperature is between 15oC and 25c)	CH_4	Developed countries	0.117	kg/head/year
	Manure management (annual average temperature greater than 25oC)	CH_4	Developed countries	0.157	kg/head/year
	Manure management (annual average temperature is greater than 25oC)	CH_4	Developing countries	0.023	kg/head/year

Source: EFDB, `www.ipcc-nggip.iges.or.jp/EFDB`.

Scope 2 emissions

Scope 2 is "*an indirect emission category that includes GHG emissions from the purchased or acquired electricity, steam, heat, or cooling consumed*" (GHG Protocol, 2015, page 34). There are then four forms of energy that are tracked in scope 2:

- Electricity
  People use electricity for operating machines, lighting, heating, cooling, electric vehicle charging, computers, electronics, public transportation systems, etc.

- Steam
  Industries use steam for mechanical work, heating, propulsion, driven turbines in electric power plants, etc.

- Heat
  Buildings use heat to control inside temperature and heat water, while the industrial sector uses heat for washing, cooking, sterilizing, drying, etc. Heat may be produced from electricity, solar heat processes or thermal combustion.

- Cooling
  It is produced from electricity or though the processes of forced air, conduction, convection, etc.

Scope 2 includes indirect emissions from generation only. For instance, the distribution of energy within a grid is tracked in scope 3. In Figures 9.6–9.9, we report the different cases that are illustrated in the GHG Protocol: if the consumed electricity comes from owned/operated equipment, no scope 2 emissions are reported (Figure 9.6); if the consumed electricity comes from a direct line transfer or the grid[13], the consumer of the energy reports the emissions in scope 2 (Figures 9.7 and (Figure 9.8); if some consumed electricity comes from the owned/operated equipment, and some is purchased from the grid, the operator (company A) has both scope 1 emissions from energy generation, and scope 3 emissions from energy purchased on the grid (Figure 9.9).

Figure 9.6: Energy production and consumption from owned/operated generation

Figure 9.7: Direct line energy transfer

Source: GHG Protocol (2015, Figures 5.1 and 5.2, pages 35-36).

---

[13]A grid is "*a system of power transmission and distribution lines under the control of a coordinating entity or grid operator, which transfers electrical energy generated by power plants to energy users — also called a power grid.*" (CDP, 2022, page 8).

Figure 9.8: Electricity production on a grid

Source: GHG Protocol (2015, Figure 5.4, page 38).

Figure 9.9: Facility consuming both energy generated on-site and purchased from the grid

Source: GHG Protocol (2015, Figure 5.3, page 37).

Scope 2 emissions are calculated using activity data and emission factors[14] expressed in MWh and tCO_2e/MWh:

$$\boldsymbol{\mathcal{CE}} = \sum_s A_s \cdot \boldsymbol{\mathcal{EF}}_s$$

where A_s is the amount of purchased electricity for the energy generation source s and $\boldsymbol{\mathcal{EF}}_s$ is the emission factor of the source s. The source can be an electricity supplier[15], a specific or country grid, a specific power station, etc.

Remark 71 *A Megawatt-hour is the common billing unit for electrical energy delivered to consumers. 1 000 MWh is equivalent to 3.6 TeraJoule (TJ). The TJ unit is used by IPCC (2006) (see Table 9.4 on page 484). A third energy unit is also used to defined emissions factors in North America (Canada and the US) and the United Kingdom: the British thermal unit or Btu (1 Btu is equivalent to 1.0551 KJ or 0.2931 Wh).*

Example 27 *We consider a company, whose electricity consumption is equal to 2 000 MWh per year. The electricity comes from two sources: 60% from a direct line with an electricity supplier (source S_1) and 40% from the country grid (source S_2). The emission factors are respectively equal to 200 and 350 gCO_2e/kWh.*

The electricity consumption from source S_1 is equal to $60\% \times 2\,000 = 1\,200$ MWh or $1\,200\,000$ kWh. We deduce that the carbon emissions from this source is:

$$\boldsymbol{\mathcal{CE}}\,(S_1) = \left(1.2 \times 10^6\right) \times 200 = 240 \times 10^6 \text{ gCO}_2\text{e} = 240 \text{ tCO}_2\text{e}$$

For the second source, we obtain:

$$\boldsymbol{\mathcal{CE}}\,(S_2) = \left(0.8 \times 10^6\right) \times 350 = 280 \times 10^6 \text{ gCO}_2\text{e} = 280 \text{ tCO}_2\text{e}$$

We deduce that the scope 2 carbon emissions of this company is equal to 520 tCO_2e.

Let us consider again the GHG inventory report of Enel (page 482). We remind that the scope 1 emissions of Enel is equal to 51 750 265 tCO_2e. In the same document, we learn that the ratio between scope 1 emissions and the total electricity production is equal to 227 gCO_2e/kWh (or 0.227 tCO_2e/MWh). We deduce that the 2021 electricity production of Enel is [16]:

$$A = \frac{51{,}750\,265}{0.227} = 227\,974\,735 \text{ MWh} = 228 \text{ TWh}$$

Two main methods are available for accounting scope 2 emissions:

- Location-based method
  In this approach, the company uses the average emission factor of the region or the country. For instance, if the electricity consumption is located in France, the company can use the emission intensity of the French energy mix;

---

[14]This approach is also known as the emission rate approach.

[15]The largest electricity companies are EDF, Enel, Engie, E.ON, Fortum, Marubeni, Siemens, State Grid Corporation of China, Tokyo Electric Power, and Uniper.

[16]In this example, we inverse the equation in order to estimate the activity data of an electricity supplier:

$$A = \frac{\boldsymbol{\mathcal{CE}}}{\boldsymbol{\mathcal{EF}}}$$

where $\boldsymbol{\mathcal{CE}}$ is the scope 1 emissions of the utility company.

- Market-based method
  This approach reflects the GHG emissions from the electricity that the company has chosen in
  the market. This means that the scope 2 carbon emissions will depend on the scope 1 carbon
  intensity of the electricity supplier.

Under the market-based method, an emission factor is associated to each electricity contract. To be
relevant, contacts must meet some quality criteria and concern some specific instruments: energy
attribute certificates [17] (*e.g.*, RECs, GOs), power purchase agreements with energy generators and
green electricity products.

Figure 9.10: Emission factor in gCO_2e/kWh of electricity generation (European Union, 1990–1992)

   The location-based method depends on the emission factor of the regional, subnational or na-
tional grid. Its value highly depends on the energy mix and the grid infrastructure. For instance,
we report the evolution of national emission factors of some European countries in Figure 9.10. We
notice that they tend to decrease since thirty years and they differ from one country to another.
On average, the emission factor is equal to 275 gCO_2e/kWh in the European Union in 2021. The
two extreme countries are Sweden (9 gCO_2e/kWh) and Poland (750 gCO_2e/kWh). The reason is
that most of Sweden's electricity supply comes from hydropower and nuclear, while Poland pro-
duces 83% of its electricity from fossil fuels (and 72% from coal). Emission factors for several region
and countries in the world are given in Table 9.5. We notice the high heterogeneity of the figures.
The continent with the lowest value is South America (204 gCO_2e/kWh) while Asia has the largest
emission factor (539 gCO_2e/kWh). Two countries which are geographically close may have different

---

[17]REC (or renewable energy certificate) is an energy attribute certificate used in Australia, Canada, India and the
US, while GO (guarantee of origin) is an energy attribute certificate used in Europe.

emission factors. This is the case of France and Germany (58 vs 354 gCO$_2$e/kWh), Canada and the US (128 vs 380 gCO$_2$e/kWh), etc.

Table 9.5: Emission factor in gCO$_2$e/kWh of electricity generation in the world

Region	$\mathcal{EF}$	Country	$\mathcal{EF}$	Country	$\mathcal{EF}$	Country	$\mathcal{EF}$
Africa	484	Australia	531	Germany	354	Portugal	183
Asia	539	Canada	128	India	637	Russia	360
Europe	280	China	544	Iran	492	Spain	169
North America	352	Costa Rica	33	Italy	226	Switzerland	47
South America	204	Cuba	575	Japan	479	United Kingdom	270
World	442	France	58	Norway	26	United States	380

Source: https://ourworldindata.org/grapher/carbon-intensity-electricity

Example 28 *We consider a French bank, whose activities are mainly located in France and the Western Europe. Below, we report the energy consumption (in MWh) by country:*

Belgium	125 807	France	1 132 261	Germany	71 890	Ireland	125 807
Italy	197 696	Luxembourg	33 069	Netherlands	18 152	Portugal	12 581
Spain	61 106	Switzerland	73 148	UK	124 010	World	37 742

If we consider a Tier 1 approach, we can estimate scope 2 emissions of the bank by computing the total activity data and multiplying by the global emission factor. Since we have twelve sources, we obtain:

$$A = \sum_{s=1}^{12} A_s = 125\,807 + 1\,132\,261 + \ldots + 37\,742 = 2\,013\,269 \text{ MWh}$$

and:

$$\begin{aligned} \mathcal{CE} &= A \cdot \mathcal{EF}_{World} \\ &= \left(2\,013{,}269 \times 10^3\right) \times 442 \\ &= 889\,864\,898\,000 \text{ gCO}_2\text{e} \\ &= 889.86 \text{ ktCO}_2\text{e} \end{aligned}$$

Another Tier 1 approach is to consider the emission factor of the European Union, because the rest of the world represents less than 2% of the electricity consumption. Using $\mathcal{EF}_{EU} = 275$, we obtain $\mathcal{CE} = 553.65$ ktCO$_2$e. The third approach uses a Tier 2 method by considering the emission factor of each country. In this case, we have to collect the data. We use figures in Table 9.5 and the following emission factors: Belgium (143); Ireland (402); Luxembourg (68) and Netherlands (331). It follows that:

$$\begin{aligned} \mathcal{CE} &= \sum_{s=1}^{12} A_s \cdot \mathcal{EF}_s \\ &= (125\,807 \times 143 + 1\,132\,261 \times 58 + \ldots + 124\,010 \times 270 + 37\,742 \times 442) \times \frac{10^3}{10^9} \\ &= 278.85 \text{ ktCO}_2\text{e} \end{aligned}$$

We notice that the estimated scope 2 emissions of this bank are sensitive to the chosen approach.

The market-based accounting approach requires to track the electricity of each supplier. For that, the company can use reliable tracking systems (North American REC, European Energy Certificate System GO, International REC standard, TIGR registry) and supplier-based contactual instruments. This Tier 1 approach is based on contract-specific emission factors. Nevertheless, when they are not available, the company can use supplier-based average and residual mix emission factors.

Example 29 *We consider a Norwegian company, whose current electricity consumption is equal to* 1 351 *Mwh. 60% of the electricity comes from the Norwegian hydroelectricity and the GO system guarantees that this green electricity emits* 1 gCO$_2$e/kWh.

If we assume that the remaining 40% of the electricity consumption comes from the Norwegian grid[18], the market-based scope 2 emissions of this company are equal to:

$$
\begin{aligned}
\mathcal{CE} &= \frac{10^6 \times 60\% \times 1 + 10^6 \times 40\% \times 26}{10^6} \\
&= 11 \text{ ktCO}_2\text{e}
\end{aligned}
$$

The market-based approach may reduce the scope 2 emissions when the company purchases green electricity. For instance, the emission factors in France are the following: 6 for nuclear, 418 for natural gas, 730 for fuel oil and 1 058 for coal. In Table 9.6, we have reported the life cycle emission factors for several technologies. Even if these figures depend on many parameters (vintage, country, etc.) and the ranges are relatively wide, we clearly observe an ordering. Wind, nuclear, hydro and solar electricity generates less GHG emissions than gas, fuel oil and coal.

Table 9.6: Emission factor in gCO$_2$e/KWh from electricity supply technologies (IPCC, 2014; UN-ECE, 2022)

Technology	Characteristic	IPCC Mean	IPCC Min–Max	UNECE Mean	UNECE Min–Max
Wind	Onshore	11	7–56	12	8–16
	Offshore	12	8–35	13	13–23
Nuclear		12	3–110	6	
Hydro power		24	1–2200	11	6–147
Solar power	CSP	27	9–63	32	14–122
	Rooftop (PV)	41	26–60	22	9–83
	Utility/Ground (PV)	48	18–180	20	8–82
Geothermal		38	6–79		
Biomass	Dedicated	230	130–420		
Gas	CCUS	169	90-370	130	92–221
	Combined cycle	490	410–650	430	403–513
Fuel oil			510–1170		
Coal	CCUS	161	70–290	350	190–470
	PC	820	740–650	1 000	912–1095

CSP: concentrated solar power; PV: photovoltaic power; CCUS: carbon capture, use, and storage; PC: pulverized coal.

---

[18]The emission factor for Norway is 26 gCO$_2$e/kWh.

Let us consider again the example of Enel (page 482). They obtained the following figures (expressed in $ktCO_2e$):

	Electricity purchased from the grid	Losses on the distribution grid	Total
Location-based	1 336.67	2 966.52	4 303.18
Market-based	2 351.00	4 763.15	7 114.15

The first category derives from the generation of electricity purchased and consumed by Enel (electricity consumption taken from the network for civil use or for energy generation in thermoelectric and hydroelectric plants). The second category includes indirect emissions due to dissipated energy emissions from technical losses from Enel's distribution network and from the transmission system. We notice that this second category represented 67% and 69% of scope 2 emissions. Curiously, the market-based figure is greater than the location-based approach: 7.11 vs. 4.30 $MtCO_2e$.

We consider the CDP database and compare the location-based and market-based values for the year 2020. Statistics are reported in Table 9.7. Less than 1% of issuers have declared zero scope 2 carbon emissions with the location-based approach. This figure becomes 8.78% when we consider the market-based approach. 70% of issuers have greater location-based emissions than market-based emissions. About 10% have the same value, meaning that these issuers have certainly used the mix residual approach to compute the scope 2 emissions with the market-based approach. The mean variation ratio[19] is equal to +26.59%. This result is explained by the frequency asymmetry, but also by the fact that the variation is higher for issuers that have greater location-based emissions than market-based emissions ($+43.29\%$ vs. -22.04%).

Table 9.7: Statistics of CDP scope 2 emissions (2020)

	$\mathcal{CE}_{\mathrm{loc}} = 0$	$\mathcal{CE}_{\mathrm{loc}} = \mathcal{CE}_{\mathrm{mkt}} = 0$	$\mathcal{CE}_{\mathrm{mkt}} = 0$
Frequency	0.89%	0.39%	8.78%
	$\mathcal{CE}_{\mathrm{loc}} > \mathcal{CE}_{\mathrm{mkt}}$	$\mathcal{CE}_{\mathrm{loc}} = \mathcal{CE}_{mkt}$	$\mathcal{CE}_{\mathrm{loc}} < \mathcal{CE}_{\mathrm{mkt}}$
Frequency	70.43%	9.48%	20.09%
Mean variation ratio	+43.89%	0.00%	−22.04%

Source: CDP database as of 01/07/2022 & Author's computation.

Scope 3 emissions

Scope 3 emissions are all the indirect emissions in the company's value chain, apart from indirect emissions which are reported in scope 2. They are divided into fifteen categories of emissions: eight upstream categories and seven downstream categories (Table 9.8). We report below their description as it appears in GHG Protocol (2011, Table 5.4, pages 34-37):

1. Purchased goods and services (not included in categories 2-8)
   Extraction, production, and transportation of goods and services purchased or acquired by the company;

2. Capital goods
   Extraction, production, and transportation of capital goods purchased or acquired by the company;

---

[19]The variation ratio is equal to $\dfrac{CE_{\mathrm{loc}} - CE_{\mathrm{mkt}}}{\max\left(CE_{\mathrm{loc}}, CE_{\mathrm{mkt}}\right)}$.

Table 9.8: The scope 3 carbon emissions categories

Upstream

1. Purchased goods and services
2. Capital goods
3. Fuel and energy related activities
4. Upstream transportation and distribution
5. Waste generated in operations
6. Business travel
7. Employee commuting
8. Upstream leased assets
9. Other upstream

Downstream

1. Downstream transportation and distribution
2. Processing of sold products
3. Use of sold products
4. End-of-life treatment of sold products
5. Downstream leased assets
6. Franchises
7. Investments
8. Other downstream

3. Fuel- and energy-related activities (not included in scope 1 or 2)
   Extraction, production, and transportation of fuels and energy purchased or acquired by the company;

4. Upstream transportation and distribution
   Transportation and distribution of products purchased by the company between the company's tier 1 suppliers and its own operations; Transportation and distribution services purchased by the company, including inbound logistics, outbound logistics (*e.g.*, sold products), and transportation and distribution between the company's own facilities;

5. Waste generated in operations
   Disposal and treatment of waste generated in the company's operations;

6. Business travel
   Transportation of employees for business-related activities;

7. Employee commuting
   Transportation of employees between their homes and their work sites;

8. Upstream leased assets
   Operation of assets leased by the company (lessee);

9. Downstream transportation and distribution
   Transportation and distribution of products sold by the company between the company's operations and the end consumer (if not paid for by the company);

10. Processing of sold products
    Processing of intermediate products sold by downstream companies (*e.g.*, manufacturers);

11. Use of sold products
    End use of goods and services sold by the company;

12. End-of-life treatment of sold products
Waste disposal and treatment of products sold by the company at the end of their life;

13. Downstream leased assets
Operation of assets owned by the company (lessor) and leased to other entities;

14. Franchises
Operation of franchises reported by franchisor;

15. Investments
Operation of investments (including equity and debt investments and project finance).

All these categories share the principle that there is no double counting of emissions between the scopes. For instance, the transport categories do not concern vehicles and facilities owned, controlled or operated by the company, because their GHG emissions are already reported in scope 1 and 2. This means that the transport of employees with a company's vehicle is reported in scope 1 and 2, but not in scope 3. On the contrary, the public transport of employees is reported in scope 3.

Table 9.9: Scope 3 emission factors for business travel and employee commuting (United States)

Vehicle type	CO_2 (kg/unit)	CH_4 (g/unit)	N_2O (g/unit)	Unit
Passenger car	0.332	0.0070	0.0070	vehicle-mile
Light-duty truck	0.454	0.0120	0.0090	vehicle-mile
Motorcycle	0.183	0.0700	0.0070	vehicle-mile
Intercity rail (northeast corridor)	0.058	0.0055	0.0007	passenger-mile
Intercity rail (other routes)	0.150	0.0117	0.0038	passenger-mile
Intercity rail (national average)	0.113	0.0092	0.0026	passenger-mile
Commuter rail	0.139	0.0112	0.0028	passenger-mile
Transit rail (subway, tram)	0.099	0.0084	0.0012	passenger-mile
Bus	0.056	0.0210	0.0009	passenger-mile
Air travel (short haul, < 300 miles)	0.207	0.0064	0.0066	passenger-mile
Air travel (medium haul, 300-2300 miles)	0.129	0.0006	0.0041	passenger-mile
Air travel (long haul, > 2300 miles)	0.163	0.0006	0.0052	passenger-mile

These factors are intended for use in the distance-based method defined in the scope 3 calculation guidance. If fuel data are available, then the fuel-based method should be used.

Source: US EPA (2020), Table 10, www.epa.gov, ghg-emission-factors-hub.xlsx.

The computation of scope 3 emissions requires specific emission factors. For example, Table 9.9 gives their values for business travel (category 6) and employee commuting (category 7) in the US. In the same document, we can find other scope 3 emissions factors (categories 4, 5, 9 and 12). Collecting data is not an easy task since there is no available comprehensive database at the global level. Nevertheless, we can find documented databases at the sector level. For instance, AGRIBALYSE provides references data on the environmental impacts of agricultural and food products through a database built according to the life cycle analysis (LCA) methodology[20]. Other databases can be found in the GHG Protocol website (https://ghgprotocol.org/life-cycle-databases). The

---

[20]The web site is https://doc.agribalyse.fr.

GHG protocol has also developed several calculation tools (cross-sector, country-specific, sector-specific and cities). With Quantis, they also provide scope 3 evaluator (S3E), which is a free web-based tool[21].

Since it may be sometimes difficult to manipulate physical units, the organizations have also developed monetary emission factors, which are expressed in $kgCO_2e/k\$$ or $kgCO_2e/k€$. Some figures are reported in Table 9.10. For example, a business air travel, whose cost is equal to \$1 000, induces a scope 3 emissions of $1\,970$ $kgCO_2e$ according to the scope 3 evaluator tool.

Table 9.10: Examples of monetary scope 3 emission factors

Category	S3E	ADEME	Category	S3E	ADEME
Agriculture	2 500	2 300	Air transport	1 970	1 190
Construction	810	360	Education	310	120
Financial intermediation	140	110	Health and Social Work	300	500
Hotels and restaurants	560	320	Rubber and plastics	1 270	800
Telecommunications	300	170	Textiles	1 100	600

Source: Scope 3 Evaluator (S3E), `https://quantis-suite.com/Scope-3-Evaluator`
& ADEME, `https://bilans-ges.ademe.fr`.

Ducoulombier (2021) highlights the importance of scope 3 emissions, but also the lack of data robustness. Since the reporting of these indirect emissions remains voluntary, we observe heterogenous data in the CDP database with scope 3 items that are partially or not calculated. In this context, most of ESG data providers estimate scope 3 upstream and downstream values using statistical model or environmentally-extended input-output (EEIO) framework[22]. This means that the reported scope 3 emissions are rarely used.

Remark 72 *In order to distinguish the different scopes, we use the following notations: $\mathcal{SC}_1$ for scope 1 emissions, $\mathcal{SC}_2$ for scope 2 emissions and $\mathcal{SC}_3 = \mathcal{SC}_3^{\mathrm{up}} + \mathcal{SC}_3^{\mathrm{down}}$ for scope 3 emissions, where $\mathcal{SC}_3^{\mathrm{up}}$ and $\mathcal{SC}_3^{\mathrm{down}}$ refer to upstream and downstream scope 3 emissions. The cumulative emissions are then denoted by $\mathcal{SC}_{1-2} = \mathcal{SC}_1 + \mathcal{SC}_2$, $\mathcal{SC}_{1-3}^{\mathrm{up}} = \mathcal{SC}_1 + \mathcal{SC}_2 + \mathcal{SC}_3^{\mathrm{up}}$ and $\mathcal{SC}_{1-3} = \mathcal{SC}_1 + \mathcal{SC}_2 + \mathcal{SC}_3$.*

9.1.4   Carbon emissions of investment portfolios

There are two main methods for measuring the carbon footprint of an investment portfolio. The first method if the financed emissions approach. In this case, the investor calculates the carbon emissions that are financed across both equity and debt. Generally, we use EVIC to estimate the value of the enterprise. It is "*the sum of the market capitalization of ordinary and preferred shares at fiscal year end and the book values of total debt and minorities interests*" (TEG, 2019b). Let W be the wealth invested in the company, the financed emissions are equal to:

$$\mathcal{CE}\left(W\right) = \frac{W}{\mathrm{EVIC}} \cdot \mathcal{CE}$$

In the case of a portfolio $(W_1, \ldots, W_n)$ where W_i is the wealth invested in company i, we have:

$$\mathcal{CE}\left(W\right) = \sum_{i=1}^{n} \mathcal{CE}_i\left(W_i\right) = \sum_{i=1}^{n} \frac{W_i}{\mathrm{EVIC}_i} \cdot \mathcal{CE}_i \tag{9.3}$$

---

[21]The tool is available at `https://quantis-suite.com/Scope-3-Evaluator`.
[22]This model is studied in Section 8.4 on page 407.

where EVIC_i and $\boldsymbol{\mathcal{CE}}_i$ are the enterprise value and carbon emissions of company i. It follows that $\boldsymbol{\mathcal{CE}}(W)$ is expressed in tCO_2e.

A second method is to use the ownership approach (Le Guenedal and Roncalli, 2022). In this case, we break down the carbon emissions between the stockholders of the company. Equation (9.3) becomes:

$$\boldsymbol{\mathcal{CE}}(W) = \sum_{i=1}^{n} \frac{W_i}{\text{MV}_i} \cdot \boldsymbol{\mathcal{CE}}_i = \sum_{i=1}^{n} \varpi_i \cdot \boldsymbol{\mathcal{CE}}_i \tag{9.4}$$

where MV_i is the market value of company i and ϖ_i is the ownership ratio of the investor. Let $W = \sum_{i=1}^{n} W_i$ be the portfolio value. The portfolio weight of asset i is given by:

$$w_i = \frac{W_i}{W}$$

We deduce that:

$$\varpi_i = \frac{W_i}{\text{MV}_i} = \frac{w_i \cdot W}{\text{MV}_i}$$

and:

$$\boldsymbol{\mathcal{CE}}(W) = \sum_{i=1}^{n} \frac{w_i \cdot W}{\text{MV}_i} CE_i = W \left(\sum_{i=1}^{n} w_i \cdot \frac{\boldsymbol{\mathcal{CE}}_i}{\text{MV}_i} \right) = W \left(\sum_{i=1}^{n} w_i \cdot \boldsymbol{\mathcal{CI}}_i^{\text{MV}} \right)$$

where $\boldsymbol{\mathcal{CI}}_i^{\text{MV}}$ is the market value-based carbon intensity of company i:

$$\boldsymbol{\mathcal{CI}}_i^{\text{MV}} = \frac{\boldsymbol{\mathcal{CE}}_i}{\text{MV}_i}$$

Since $\boldsymbol{\mathcal{CE}}(W)$ is a linear function of W, the carbon footprint of the portfolio is generally computed with $W = \$1$ mn and is expressed in tCO_2e (per $ mn invested).

Remark 73 *The second approach is valid only for equity portfolios. To compute the market value (or the total market capitalization), we use the following approximation:*

$$\text{MV} = \frac{\text{MC}}{\mathcal{FP}}$$

where MC *and* $\mathcal{FP}$ *are the free float market capitalisation and percentage of the company.*

Example 30 *We consider a $100 mn investment portfolio with the following composition: $63.1 mn in company A, $16.9 mn in company B and $20.0 mn in company C. The data are the following:*

Issuer	Market capitalization (in $ bn)			Debt	$\mathcal{FP}$	$\boldsymbol{\mathcal{SC}}_{1-2}$
	31/12/2021	31/12/2022	31/01/2023	(in $ bn)	(in %)	(in ktCO_2e)
A	12.886	10.356	10.625	1.112	99.8	756.144
B	7.005	6.735	6.823	0.000	39.3	23.112
C	3.271	3.287	3.474	0.458	96.7	454.460

As of 31 January 2023, the EVIC value for company A is equal to:

$$\text{EVIC}_A = \frac{10\,356}{0.998} + 1\,112 = \$11\,489 \text{ mn}$$

We deduce that the financed emissions are equal to:

$$\boldsymbol{\mathcal{CE}}_A(\$63.1 \text{ mn}) = \frac{63.1}{11\,489} \times 756.144 = 4.153 \text{ ktCO}_2\text{e}$$

If we assume that the investor has no bond in the portfolio, we can use the ownership approach:

$$\varpi_A = \frac{63.1}{(10\,625/0.998)} = 59.2695 \text{ bps}$$

The carbon emissions of the investment in company A is then equal to:

$$\mathcal{CE}_A\,(\$63.1 \text{ mn}) = 59.2695 \times 10^{-4} \times 756.144 = 4.482 \text{ ktCO}_2\text{e}$$

Finally, we obtain the following results[23]:

	Financed emissions	Carbon emissions
Company A	4.153	4.482
Company B	0.023	0.022
Company C	2.356	2.530
Portfolio	6.532	7.034

9.1.5 Statistics

In what follows, we use the analysis done by Barahhou *et al.* (2022). We consider the Trucost dataset of carbon emissions as of $01/06/2022$ and analyze the distribution of carbon emissions in 2019 for around $15\,000$ companies. We prefer to use the year 2019 instead of the year 2020, because the covid-19 crisis had a significant impact on the carbon footprint. In Figure 9.11, we have reported the scope 1 and 2 carbon emissions per GICS sector. We notice that including scope 2 has a limited impact, except for some low-carbon sectors such as Consumer Services, Information Technology and Real Estate. In Table 9.11, we have calculated the breakdown of carbon emissions. Scope 1 and 2 emissions represent 17.6 GtCO$_2$e, and the most important sector contributors are Utilities (34.4%), Materials (31.4%), Energy (14.0%) and Industrials (10.0%). This means that these 4 strategic sectors explain about 90% of scope 1 and 2 carbon emissions.

Table 9.11: Breakdown (in %) of carbon emissions in 2019

Sector	$\mathcal{SC}_1$	$\mathcal{SC}_2$	$\mathcal{SC}_{1-2}$	$\mathcal{SC}_3^{\text{up}}$	$\mathcal{SC}_3^{\text{down}}$	$\mathcal{SC}_3$	$\mathcal{SC}_{1-3}$
Communication Services	0.1	5.1	0.8	1.5	0.2	0.4	0.5
Consumer Discretionary	1.7	9.7	2.9	14.1	10.2	10.8	9.1
Consumer Staples	2.3	6.7	2.9	18.6	1.6	4.4	4.1
Energy	15.0	8.5	14.0	14.1	40.1	36.0	31.2
Financials	0.7	1.8	0.9	2.6	1.8	2.0	1.7
Health Care	0.3	1.7	0.5	2.6	0.2	0.6	0.6
Industrials	10.2	8.9	10.0	15.6	24.2	22.8	20.0
Information Technology	0.6	6.8	1.5	4.9	2.3	2.7	2.5
Materials	29.8	40.7	31.4	20.2	13.5	14.6	18.2
Real Estate	0.3	2.8	0.6	1.1	1.0	1.0	0.9
Utilities	39.0	7.3	34.4	4.7	4.8	4.8	11.2
Total (in GtCO$_2$e)	15.1	2.6	17.6	10.3	53.7	64.0	81.6

Source: Trucost (2022) & Barahhou *et al.* (2022).

---

[23]For the financed emissions, we use the data as of 31 December 2022 while the ownership ratio is based on the current data (as of 31 January 2023). In this example, the data as of 31 December 2021 are never used.

Figure 9.11: 2019 carbon emissions per GICS sector in GtCO$_2$e (scope 1 & 2)
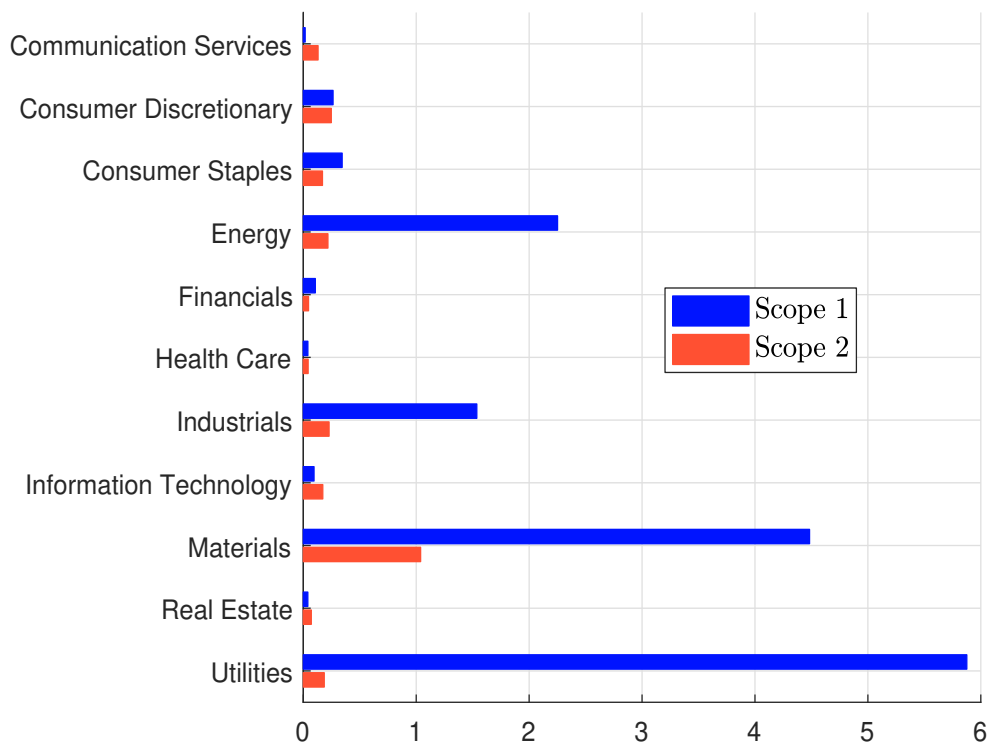
Figure 9.12: 2019 carbon emissions per GICS sector in GtCO$_2$e (scope 1, 2 & 3 upstream)
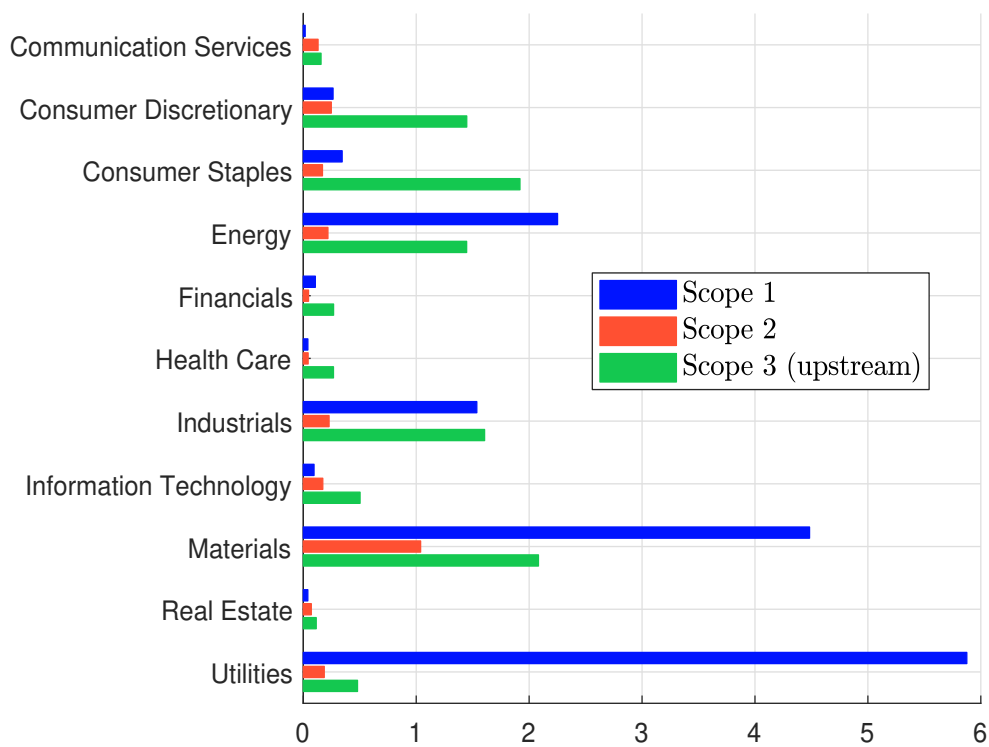
Figure 9.13: 2019 carbon emissions per GICS sector in GtCO$_2$e (scope 1, 2 & 3)
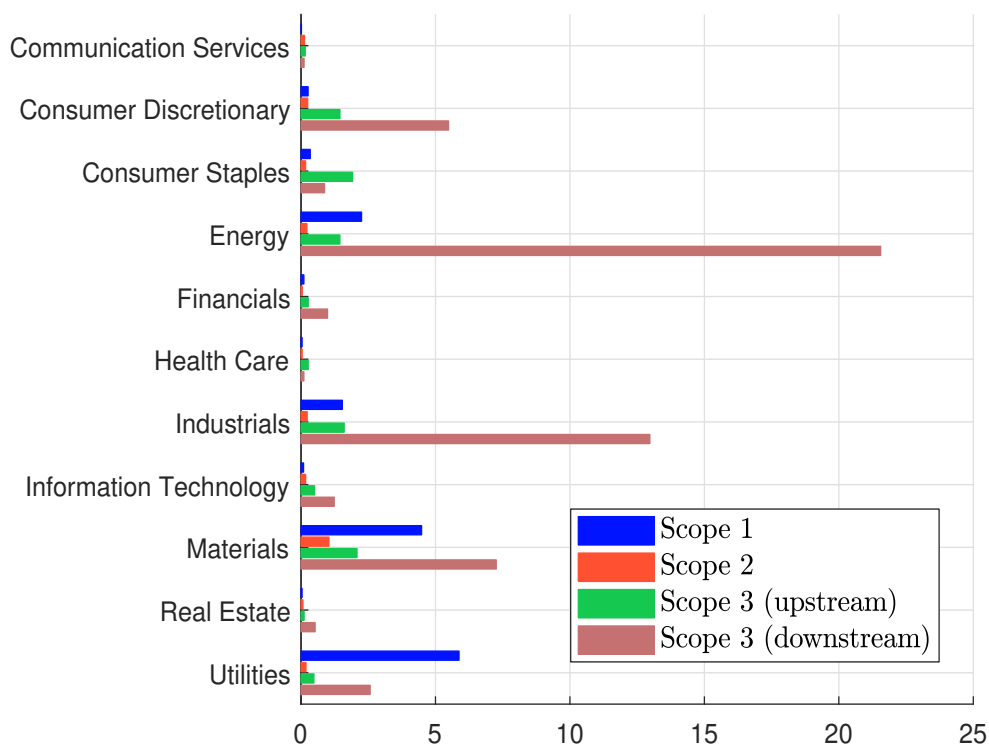
In Figure 9.12, we observe that some sectors are highly impacted by the upstream scope 3 emissions. For instance, the ratio $\dfrac{\mathcal{SC}_3^{\mathrm{up}}}{\mathcal{SC}_{1-2}}$ is greater than 2.5 for Consumer Discretionary, Consumer Staples and Health Care, and is close to 2 for Information Technology. Among the strategic sectors, Energy and Industrials are the most penalized whereas the upstream scope 3 emissions of Utilities is relatively small compared to its scope 1 emissions.

While the impact of the upstream scope 3 is significant, the impact of the downstream scope 3 is huge as demonstrated in Figure 9.13. Four sectors have very large downstream carbon emissions: Consumer Discretionary, Energy, Industrials and Materials. While Utilities has the most important contribution in terms of scope 1 and 2 since it represents 34.4% of carbon emissions, its contribution to scope 3 is relatively modest and is equal to 4.8%. Including or not scope 3, in particular the downstream carbon emissions, changes the whole picture of the breakdown between the sectors. Figure 9.14 is a visualisation of the sector contribution by considering the addition of several scopes. At each step, the contribution of Materials and Utilities decreases whereas it increases for Consumer Discretionary, Energy, Industrials and Information Technology. Among the most significant sectors[24], the behavior of Consumer Staples is singular since its contribution increases when adding scope 2 and upstream scope 3, but decreases when considering downstream scope 3.

Remark 74 *When considering scope 3 emissions, double counting is a real issue. According to Table 9.11, the total carbon emissions is* 17.6 GtCO$_2$e *for scope* 1 + 2, *and* 81.6 GtCO$_2$e *for scope* 1 + 2 + 3, *while we estimate that the world emits about* 36 GtCO$_2$e *per year.*

---

[24]They correspond to sectors that have a contribution greater than 2%.

Figure 9.14: Sector contribution in %
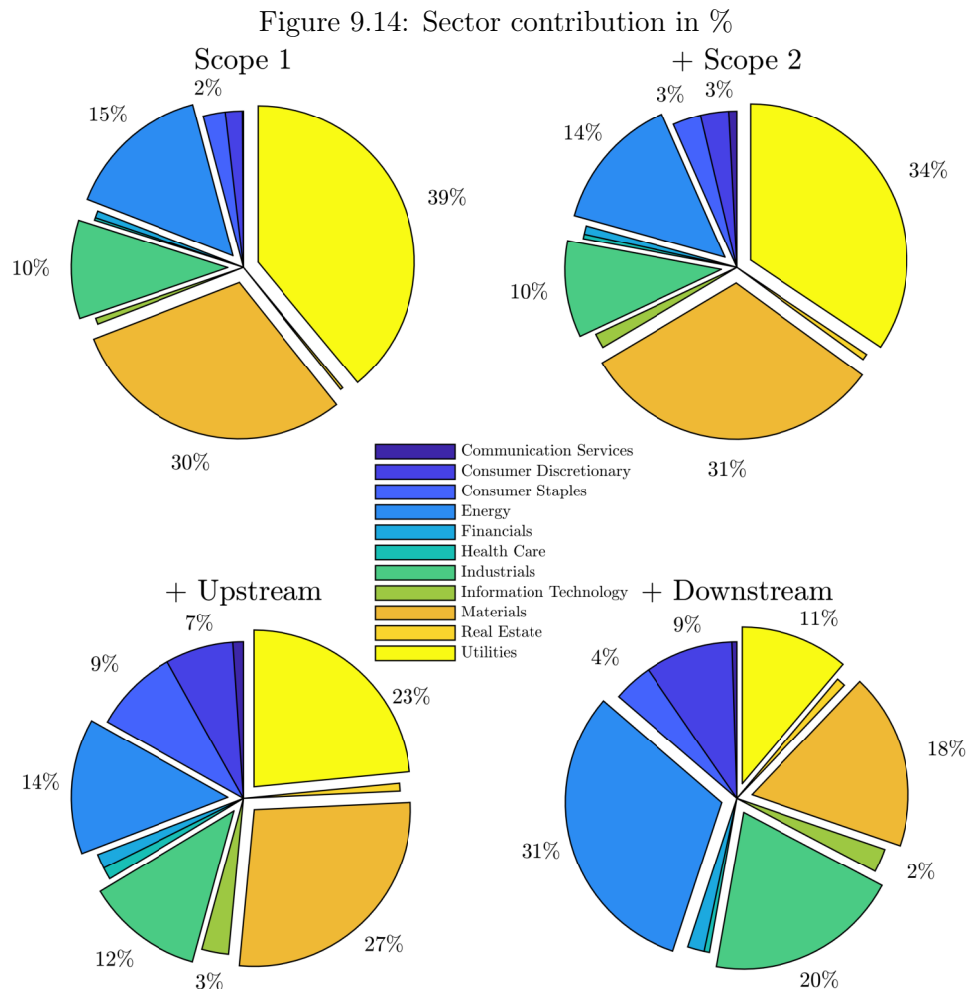
In Figure 9.15, we draw the histogram of carbon emissions and indicate the 5% and 95% percentile values. We need to use a logarithmic scale, because the range is between some tonnes of CO_2e to several dozen tonnes of CO_2e. This graph shows that it is difficult to compute the carbon footprint of a portfolio based on carbon emissions, because this metric is not homogeneous to the company size. This is why the carbon intensity metric is preferred in financial markets.

9.1.6 Negative emissions, avoided emissions, and carbon offsetting

Negative emissions, also known as carbon dioxide removal or CDR, is the process of removing CO_2 from the atmosphere. There are two main categories of negative emissions:

1. Natural climate solutions
   Examples include forest restoration and afforestation[25], reducing soil disturbance[26], etc.

---

[25]Afforestation is the process of creating a new forest (planting trees in an area where there was no forest in the past), while reforestation is the process of planting trees in areas where there was forest before.

[26]This is the practice of minimizing disturbance to the soil surface and structure, such as using minimum tillage or planting certain crops that protect the soil.

Figure 9.15: Histogram of 2019 carbon emissions (logarithmic scale, tCO_2e)
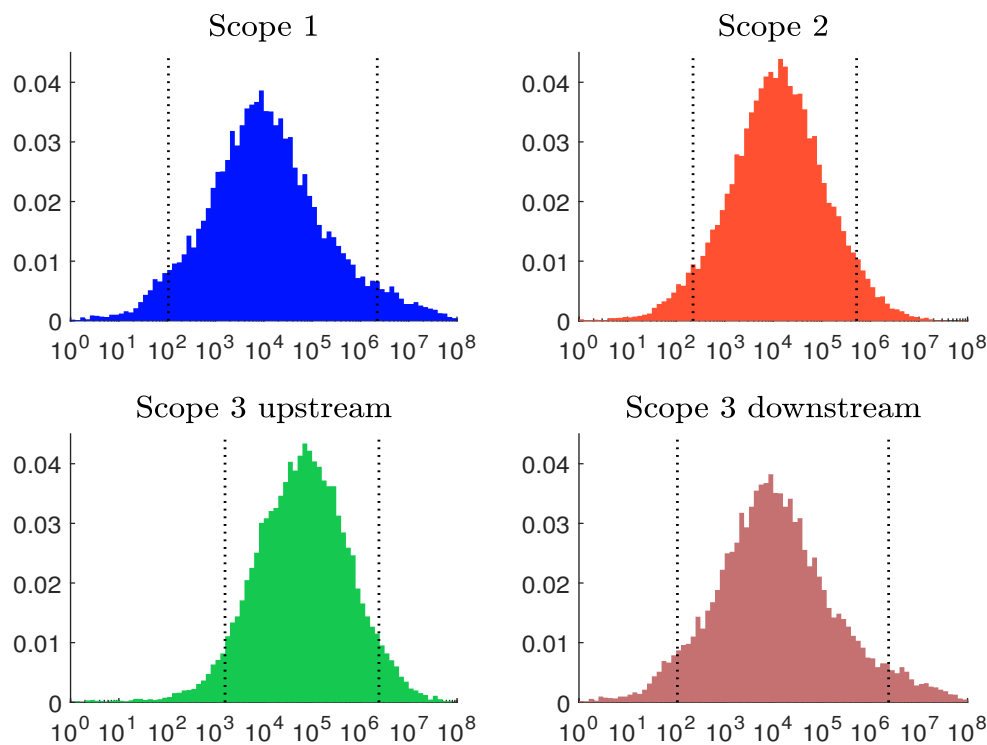
2. Negative emission technologies
   Examples are direct air capture with carbon storage[27] (DACCS), bioenergy with carbon capture and storage[28] (BECCS), enhanced weathering[29], ocean fertilization[30], etc.

Tanzer and Ramírez (2019) gives a more formal definition of negative emissions by considering four minimum criteria for determining whether a technology induces negative emissions:

> "[...] (1) Physical greenhouse gases are removed from the atmosphere. (2) The removed gases are stored out of the atmosphere in a manner intended to be permanent. (3) Upstream and downstream greenhouse gas emissions associated with the removal and storage process, such as biomass origin, energy use, gas fate, and co-product fate, are comprehensively estimated and included in the emission balance. (4) The total quantity of atmospheric greenhouse gases removed and permanently stored is greater than the total quantity of greenhouse gases emitted to the atmosphere." (Tanzer and Ramírez, 2019, page 1216)

In a series of three review papers, Jan Minx and his co-authors provided a comprehensive overview of negative emissions (Minx *et al.*, 2018; Fuss *et al.*, 2018; Nemet *et al.*, 2018). They emphasized

---

[27]This technology uses special filters to capture CO_2 directly from the air, while the captured CO_2 is then stored underground or used in other applications.

[28]This process involves capturing and storing the CO_2 emissions from burning biomass, such as wood or grasses.

[29]This process involves the application of finely ground minerals, such as olivine or basalt, to land surfaces. When these minerals react with atmospheric CO_2, they form harmless minerals and carbonates, trapping the carbon in a stable mineral form. The goal is to accelerate the natural process of weathering.

[30]This technology involves adding nutrients to the ocean, which can stimulate the growth of phytoplankton in the ocean, which then absorbs CO_2 through photosynthesis.

that the efficiency, capacity, and cost of different technologies vary widely. A typical example is direct air capture technology[31]. The rationale for the technology is presented in the book published by National Academies of Sciences, Engineering, and Medicine (2019). There are two general types of DAC processes: DAC with liquid solvents (L-DAC) and DAC with solid sorbents (S-DAC). In an L-DAC process, there are four stages: absorption, regeneration, purification and separation. Each phase involves a chemical reaction:

$$
\begin{aligned}
2\,KOH + CO_2 &\longrightarrow H_2O + K_2CO_3 \\
CaO + H_2O &\longrightarrow Ca\,(OH)_2 \\
K_2CO_3 + Ca\,(OH)_2 &\longrightarrow 2\,KOH + CaCO_3 \\
CaCO_3 &\longrightarrow CaO + CO_2
\end{aligned}
$$
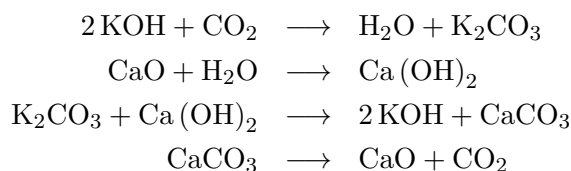
The goal is to use the liquid solvent KOH to react with atmospheric carbon dioxide CO_2 to produce pure CO_2 and calcium oxide CaO. In an S-DAC process, solid materials or sorbents, such as porous polymers or metal-organic frameworks, are used to adsorb CO_2. The costs associated with DAC technology include the initial investment to build the DAC system (*e.g.*, air contractor, causticizer, calciner, and slaker), the price of solvents and sorbents, the electricity needs to perform the chemical reactions, and the cost of storage. The current price of removing a tonne of CO_2 is around \$1 000, which is high compared to the price of carbon traded on CO_2 markets. Another factor in assessing the relevance of DAC technologies is the measurement of carbon efficiency, which depends on the amount and carbon intensity of electricity used to remove atmospheric CO_2. Today, the carbon efficiency of the best DAC plans is less than 70%. This is, of course, the current situation and many improvements are expected in the coming years. For instance, IEA (2022) estimated that DAC costs could fall below \$100/tCO_2 by 2030.

Box 9.4: An example of DAC companies: Climeworks

Climeworks (https://climeworks.com) is a Swiss company founded in 2009 as a spin-off from ETH Zurich. It specializes in DAC technology and has established itself as a pioneer in this field with two other companies: Carbon Engineering (Canada) and Global Thermostat (USA). In September 2021, Climeworks inaugurates the world's first large-scale direct air capture and storage plant "*Orca*" in Iceland, with a capacity to capture 4 000 tonnes of CO_2 per year. The storage of CO_2 is carried out by the company Carbfix, which injects it deep underground, where it mineralizes and turns into stone. In June 2022, Climeworks announces a second, newest and largest direct air capture and storage facility, "*Mammoth*", also in Iceland. It will have a nominal CO_2 capture capacity of up to 36 000 tonnes per year when fully operational.

A related concept to negative emissions is avoided emissions, often incorrectly referred to as Scope 4 emissions. According to Russell (2023), "*comparative impacts are estimated as the difference between the total, attributional, life-cycle GHG inventories of a company's product (the assessed product) and an alternative (or reference) product that provides an equivalent function*":

$$
\mathcal{AE} = \mathcal{CE}\,(\text{reference product}) - \mathcal{CE}\,(\text{assessed product})
$$

Avoided emissions can be positive ($\mathcal{AE} \geq 0$) or negative ($\mathcal{AE} < 0$). For example, an electric car emits CO_2, especially when we consider the life cycle of the batteries, but electric cars do not emit

---

[31]DAC and DACCS are two interchangeable terms because carbon storage is implicit in all carbon dioxide capture, use, and storage (CCUS) technologies.

greenhouse gases from burning gasoline. In this example, the reference product is the gasoline-powered car and the assessed product is the electric car, and we expect the avoided emissions to be positive. However, there are two issues in calculating avoided emissions. First, which car should we choose to represent the gasoline car or the reference product? Second, what is the use of the electric car? In fact, the avoided emissions depend on many factors, such as the carbon intensity of the electricity, recycling assumptions, etc.

In addition to negative emissions and avoided emissions, carbon offsetting includes a third concept: carbon credits. Carbon credits are transferable financial instruments that represent one tonne of carbon dioxide or another greenhouse gas. They are traded on carbon markets where companies, governments and individuals can buy and sell credits to meet their emission reduction targets. The price of carbon credits can vary depending on supply and demand, as well as the type of project and the region in which it is located. There are two main types of carbon credit systems:

- Cap-and-trade systems
  These systems place a limit on the total amount of GHG emissions that can be released from a given region or industry. Companies are allocated a certain number of carbon credits (emission allowances) and can buy or sell credits to meet their emissions targets. These government-regulated schemes make up the compliance carbon market.

- Voluntary carbon markets
  These markets are not regulated by the government, and companies can voluntarily buy carbon credits to offset their emissions. Voluntary carbon markets are often used to offset emissions from activities not covered by cap-and-trade systems. In this case, the avoided emissions from a carbon offset (e.g., through the use of negative emission technologies) must be counted on the balance sheet of the buyer, not the seller, who is the developer of the project.

We can now give a precise definition of carbon offsetting. Carbon offsetting is when a company offsets its own carbon emissions by providing emission reductions outside of its own operations. This means that the company purchases a verified carbon credit in a voluntary carbon market that funds a negative emission project. Carbon offsetting does not involve avoided emissions because they concerns the company's own operations and are associated with a change in the company's business strategy. Carbon offsets also do not include carbon credits purchased in a cap-and-trade system because these carbon credits do not necessarily result in negative emissions. Because carbon offsetting involves reducing a company's carbon footprint, it is commonly associated with the race to net-zero emissions. However, we need to make a clear distinction between the two concepts. It is now accepted that some activities will continue to emit GHGs in 2050 due to a lack of carbon-free alternatives, even in the most stringent net-zero scenario. In such situations, carbon offsets must be used primarily by companies exposed to these hard-to-abate sectors, such as cement or airlines. In Figure 9.16 we reproduce the taxonomy of carbon offsets proposed by Allen *et al.* (2020). Based on our definition, only categories IV and V fall under the strict definition of carbon offsetting.

Allen *et al.* (2020) proposed a framework for assessing the relevance of carbon offsets to ensure that they contribute to a net-zero economy. The Oxford principles for net-zero aligned carbon offsetting are:

1. Cut emissions, use high quality offsets, and regularly revise offsetting strategy as best practice evolves
   Companies' first priority is to reduce their own emissions, not to purchase carbon offsets. If they do, they need to buy offsets that ensure environmental integrity, high standards and certification in line with accounting practices. The largest GHG offset programs are the Verified

Figure 9.16: Taxonomy of carbon offsets
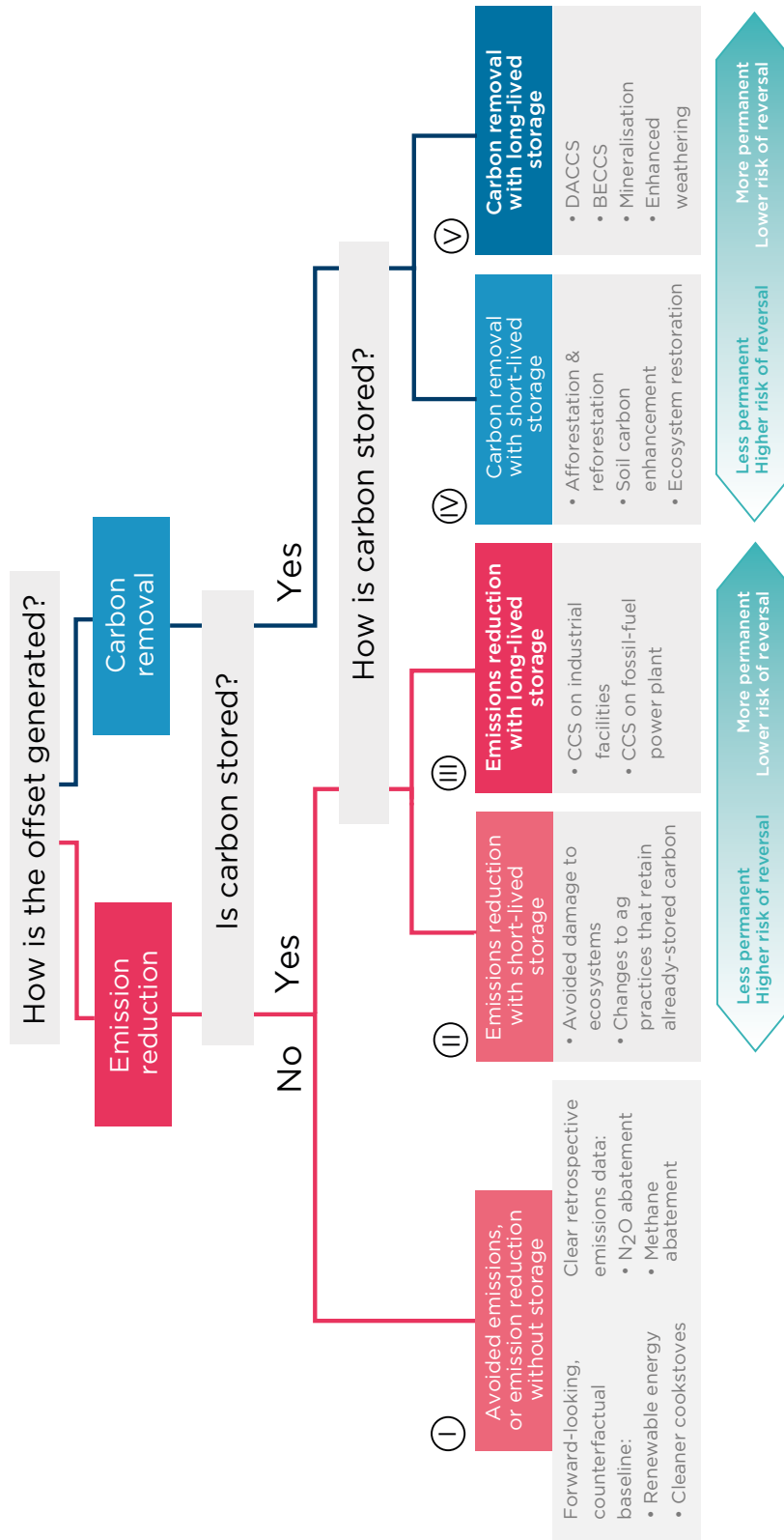
Carbon Standard (VCS) and the Gold Standard (GS). The status of projects that meet these standards can be tracked through official registries whose goal is to certify the ownership of each negative emissions project. Specifically, ownership is transferred to the buyer of the carbon credits and then canceled when the credits are sold. The purpose of registries is to ensure that negative emissions are not counted twice.

2. Shift to carbon removal offsetting
   There is clearly an imbalance between the supply of certified negative emissions projects and the projects needed to achieve net zero in the long term. Creating demand for carbon removal offsets today will send the necessary market signal to increase supply. Figure 9.17 shows the size of the voluntary carbon market (VCM). It has been multiplied by two after 2019. According to Ecosystem Marketplace (2023b), the cumulative volume has reached $2.3\,\mathrm{GtCO_2e}$ with a value of \$10 billion. This implies an average price of \$4.35 per tonne of CO_2. From 2021, the average price is more likely to be between \$7 and \$8 per tonne of CO_2. The market is largely dominated by renewable energy projects and forestry & land use. Since 2020, projects on household & community devices have also been promoted. Although developing, the voluntary carbon market remains relatively small and immature, with many intermediaries and few end users[32]. For example, the energy sector is the main buyer of voluntary carbon credits, accounting for more than 50% of the market. However, the market is expected to reach between \$10 billion and \$40 billion by 2030, up from a record \$2.1 billion in 2021 (BCG, 2023).

Figure 9.17: Voluntary carbon market size by volume of traded carbon credits

---
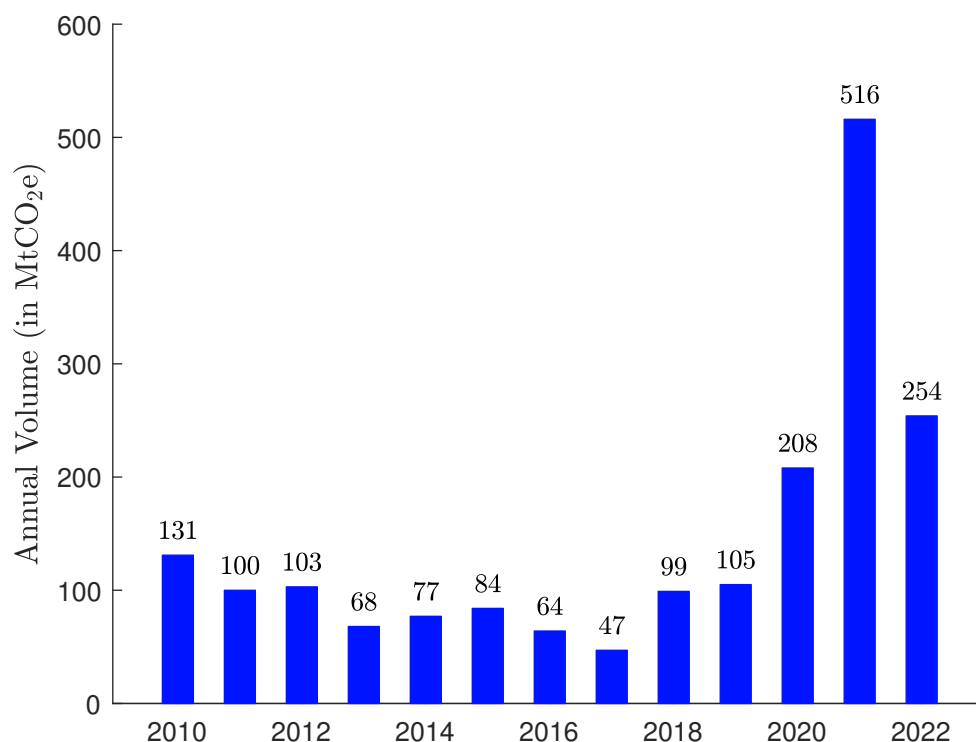
[32]In fact, it is concentrated in a few companies. According to Ecosystem Marketplace (2023a), the top 10 buyers in 2021 were Delta Air Lines, TotalEnergies, Shell, Volkswagen, Takeda Pharmaceuticals, Comcast, Diamondback Energy, La Poste, Telstra and Eni.

3. Shift to long-lived storage
   The issue of CO_2 storage and sequestration is an important one. As noted by Allen *et al.* (2020), "*short-lived storage involves methods that have a higher risk of being reversed over decades. Long-lived storage refers to methods of storing carbon that have a low risk of reversal over centuries to millennia, such as storing CO2 in geological reservoirs or mineralising carbon into stable forms. Short-lived storage offsets help buy time to reduce emissions and invest in long-lived storage, but they are not a long-term solution for achieving balance between sinks and sources.*" Measuring the efficiency of a technology is not straightforward and is highly dependent on the lifetime of the project (Terlouw *et al.*, 2021) and the system boundary. Figure 9.18 shows an example taken from Tanzer and Ramírez (2019). Chiquier *et al.* (2022) proposed to evaluate the efficiency of carbon dioxide removal by considering the amount of CO_2 stored (or removed) and the amount of CO_2 leaked (or emitted) over the supply chain:

$$\eta\left(t\right) = \frac{CO_2^{stored}\left(t\right) - CO_2^{leaked}\left(t\right)}{CO_2^{stored}\left(t\right)}$$

The metric $\eta\left(t\right)$ depends on the lifetime t expressed in years. In general, it is a decreasing function of time t, which means that the efficiency is maximum at the beginning of the project. In the case of an afforestation/reforestation project implemented in 2020 in the UK, Chiquier *et al.* (2022) estimates $\eta\left(10\right) = 87.1\%$, $\eta\left(30\right) = 98.8\%$, $\eta\left(100\right) = 98.9\%$, and $\eta\left(1000\right) = 61.9\%$. Here, the CDR efficiency increases in the beginning because the forest establishment emits CO_2 and the trees are young. Then the trees grow and the efficiency is close to 100% between 30 and 100 years. In the long term, the efficiency decreases due to the risk of forest fires. A summary of key features for each CDR pathway is provided in Table 9.12.

Table 9.12: Summary of key features for each CDR pathway

CDR	$\eta\left(100\right)$	$\eta\left(1000\right)$	Timing	Permanence
Afforestation	63 to 99%	31 to 95%	Decades	Very low
Reforestation	63 to 99%	31 to 95%	Decades	Very low
BECCS	52 to 87%	78 to 87%	Immediate to decades	High/very high
Biochar	20 to 39%	-3 to 5%	Immediate	Low/very low
DACCS	-5 to 90%	-5 to 90%	Immediate	Very high
Enhanced weathering	17 to 92%	51 to 92%	Immediate to decades	High/very high

Source: Chiquier *et al.* (2022, Table 1, page 4400).

4. Support the development of net-zero aligned offsetting
   The fourth principle is to promote carbon offsetting. To develop this market, companies can enter into long-term agreements, form sector-specific alliances, support the restoration and protection of natural and semi-natural ecosystems, and incorporate these principles into regulation.

Carbon accounting for negative emission technologies is still an open question (Brander *et al.*, 2021; Kaplan *et al.*, 2023). It is closely related to the issue of certification and credibility of CDR projects. Accounting for carbon offsets also challenges the economic incentives of these projects from the perspective of buyers of carbon credits. Microsoft's experience described in *Nature* is an interesting testimony from an end user and provides some insights to improve the ecosystem of negative emission technologies and carbon offsetting (Joppa *et al.*, 2021).

Figure 9.18: Perceived CO_2 emissions of a simplified steel production system when viewed from different system boundaries
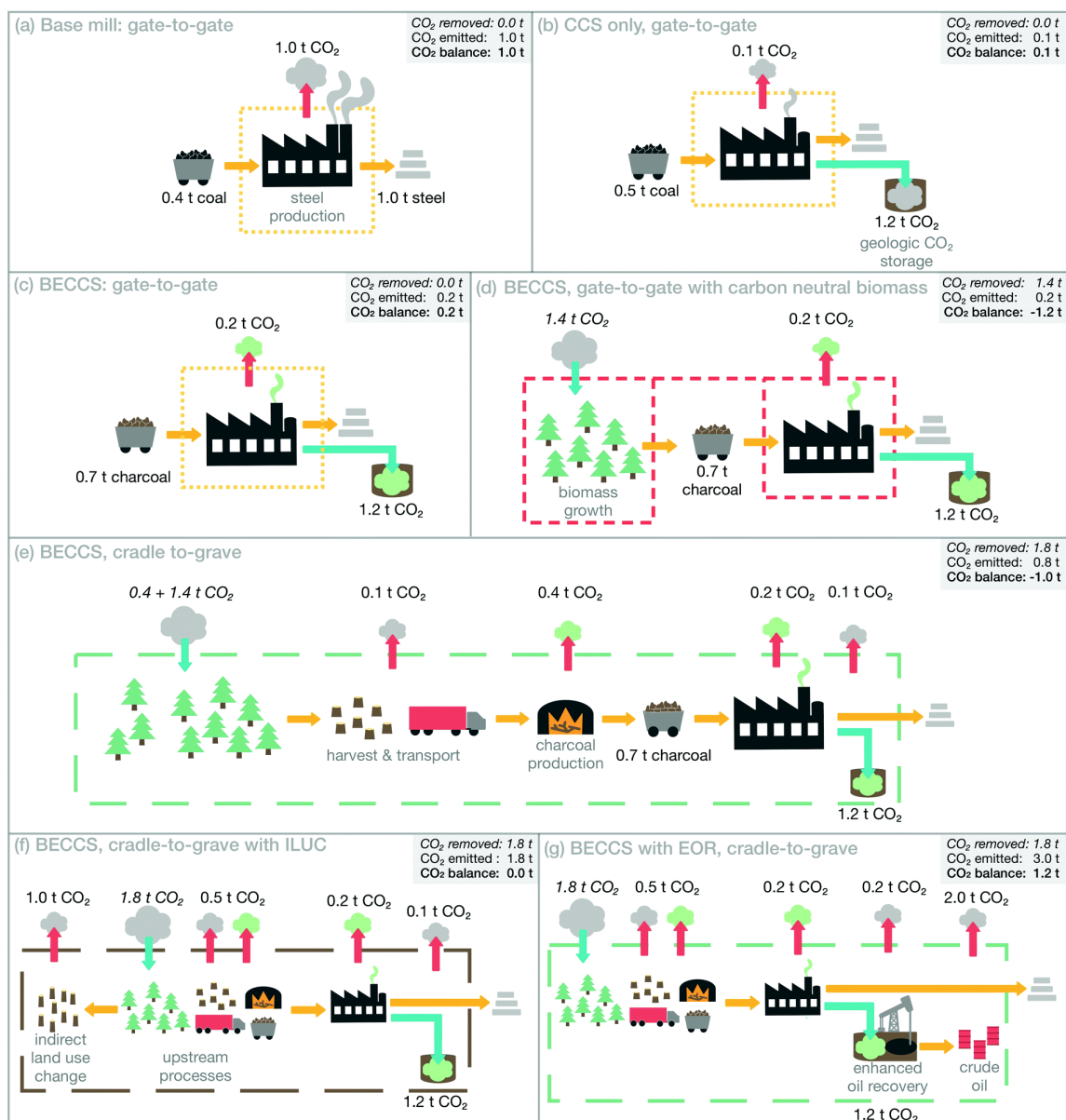
The dashed line in each sub-figure represents the system boundaries used to estimate the total CO_2 emissions in the upper right corner of each figure. The system design and numbers used are greatly simplified for illustrative purposes. (a-c) show the gate-to-gate CO_2 emissions of a steel mill, considering only the CO_2 produced at the mill itself for normal production (a), with the use of carbon capture and storage (b), and with the use of bioenergy with carbon capture and storage (c). (d) extends the system boundaries to include photosynthetic uptake of the exact amount of CO_2 released by combustion, assuming the charcoal is carbon neutral. (e) is a simplified cradle-to-grave system that includes in its boundaries the CO_2 absorbed by the wood that is lost in the charcoal production process, the CO_2 emissions from biomass harvesting and transportation, the CO_2 emissions from charcoal production, and the CO_2 emissions from CO_2 storage. (f) is a variant where biomass production has significant indirect land use change emissions. (g) is a variant where geological storage of CO_2 results in the production and combustion of fossil fuels whose CO_2 emissions exceed the CO_2 stored.

Source: Tanzer and Ramírez (2019, Figure 2, page 1214).

9.2 Carbon intensity

While carbon emissions measure the carbon footprint in an absolute value, the carbon intensity is a relative metric of the carbon footprint. The underlying idea is to normalize the carbon emissions by a size or activity unit. For instance, we can measure the carbon footprint of countries by tCO_2e per capita, watching television by CO_2e emissions per viewer-hour, washing machines by $kgCO_2e$ per wash, cars by $kgCO_2e$ per kilometer driven, companies by $ktCO_2e$ per \$1 mn revenue, etc. We distinguish two types of carbon intensity: carbon intensities whose activity units are physical and carbon intensities whose activity units are monetary.

9.2.1 Physical intensity ratios

The product carbon footprint (PCF) measures the relative carbon emissions of a product throughout its life cycle. This approach, which is called life cycle assessment (LCA), distinguishes two methods:

- Cradle-to-gate refers to the carbon footprint of a product from the moment it is produced (including the extraction of raw materials) to the moment it enters the store;

- In contrast, cradle-to-grave covers the entire life cycle of a product, including the use-phase and recycling.

Below, we report some examples of product carbon footprint computed by ADEME.

Table 9.13: Examples of product carbon footprint (in $kgCO_2e$ per unit)

Product	Category	Cradle-to-gate	Cradle-to-grave
Screen	21.5 inches	222	236
	23.8 inches	248	265
Computer	Laptop	156	169
	Desktop	169	189
	High performance	295	394
Smartphone	Classical	16	16
	5 inches	33	32
Oven	Built-in electric	187	319
	Professional (combi steamer)	734	12 676
Washing machine	Capacity 5kg	248	468
	Capacity 7kg	275	539
Shirt	Coton	10	13
	Viscose	9	12
Balloon	Football	3.4	5.1
	Basket-ball	3.6	5.9

Source: Lhotellier *et al.* (2018, Annex 4, pages 212-215).

The previous analysis can be extended to corporate carbon footprint (CCF). For instance, we can measure the CCF of a cement manufacturer by the amount of GHG emissions per tonne of cement. In the airline sector, the main traffic metric is the revenue passenger kilometers (RPK), which is calculated by multiplying the number of paying passengers by the distance traveled. Therefore, the CCF of airlines can be measured by the amount of GHG emissions per RPK (Table 9.14).

Table 9.14: Physical carbon intensity per production unit

Sector	Unit	Description
Transport sector (aviation)	CO_2e/RPK	Revenue passenger kilometers
Transport sector (shipping)	CO_2e/RTK	Revenue tonne kilometers
Industry (cement)	CO_2e/t cement	Tonne of cement
Industry (steel)	CO_2e/t steel	Tonne of steel
Electricity	CO_2e/MWh	Megawatt hour
Buildings	CO_2e/SQM	Square meter

9.2.2   Monetary intensity ratios

From a financial point of view, it does not make sense to compare and aggregate the carbon emissions of a large cap company with the carbon emissions of a small cap company. Carbon intensity is then a more relevant metric. ESG analysts can then compare companies that belong to the same activity sector by using physical intensity ratios. For example, they can compare all the cement manufacturers, because they can normalize the carbon emissions by the volume of cement production. In a similar way, they can compare all the airline companies, because they can normalize the carbon emissions by the RPK metric. Nevertheless, the physical intensity ratios are not relevant when we consider a portfolio that is invested in several sectors. How to compare a cement-based carbon intensity with a RPK-based carbon intensity? How to aggregate the two metrics? Until now, nobody has the answer.

Therefore, portfolio managers will use monetary intensity ratios, which are defined as:

$$\mathcal{CI} = \frac{\mathcal{CE}}{Y}$$

where $\mathcal{CE}$ is the company's carbon emissions and Y is a monetary variable measuring its activity. For instance, we can use revenues, sales, etc. to normalize carbon emissions:

- Revenue:

$$\mathcal{CI}^{\text{Revenue}} = \frac{\mathcal{CE}}{\text{Revenue}}$$

- Sales:

$$\mathcal{CI}^{\text{Sales}} = \frac{\mathcal{CE}}{\text{Sales}}$$

- Enterprise value including cash:

$$\mathcal{CI}^{\text{EVIC}} = \frac{\mathcal{CE}}{\text{EVIC}}$$

- Market value:

$$\mathcal{CI}^{\text{MV}} = \frac{\mathcal{CE}}{\text{MV}}$$

Even the previous carbon emission metrics based on EVIC and market value can be viewed as carbon intensity metrics.

If we consider the EVIC-based approach, the carbon intensity of the portfolio is given by:

$$
\begin{aligned}
\mathcal{CI}^{\mathrm{EVIC}}(w) &= \frac{\mathcal{CE}^{\mathrm{EVIC}}(W)}{W} \\
&= \frac{1}{W} \sum_{i=1}^{n} \frac{W_i}{\mathrm{EVIC}_i} \cdot \mathcal{CE}_i \\
&= \sum_{i=1}^{n} \frac{W_i}{W} \cdot \frac{\mathcal{CE}_i}{\mathrm{EVIC}_i} \\
&= \sum_{i=1}^{n} w_i \cdot \mathcal{CI}_i^{\mathrm{EVIC}}
\end{aligned}
$$

where $w = (w_1, \ldots, w_n)$ is the vector of portfolio weights. We notice that the carbon intensity satisfies the additivity property. In a similar way, we obtain:

$$
\mathcal{CI}^{\mathrm{MV}}(w) = \sum_{i=1}^{n} w_i \cdot \mathcal{CI}_i^{\mathrm{MV}}
$$

Let us now consider the revenue-based carbon intensity (also called the economic carbon intensity). We denote by Y_i the revenue of issuer i. The carbon intensity of the portfolio becomes:

$$
\mathcal{CI}^{\mathrm{Revenue}}(w) = \frac{\mathcal{CE}(w)}{Y(w)}
$$

where $\mathcal{CE}(w)$ measures the carbon emissions of the portfolio:

$$
\mathcal{CE}(w) = \sum_{i=1}^{n} W_i \cdot \frac{\mathcal{CE}_i}{\mathrm{MV}_i} = W \sum_{i=1}^{n} \frac{w_i}{\mathrm{MV}_i} \cdot \mathcal{CE}_i
$$

and $Y(w)$ is the total revenue of the portfolio:

$$
Y(w) = \sum_{i=1}^{n} W_i \cdot \frac{Y_i}{\mathrm{MV}_i} = W \sum_{i=1}^{n} \frac{w_i}{\mathrm{MV}_i} \cdot Y_i
$$

We deduce that:

$$
\begin{aligned}
\mathcal{CI}^{\mathrm{Revenue}}(w) &= \frac{\sum_{i=1}^{n} \dfrac{w_i}{\mathrm{MV}_i} \cdot \mathcal{CE}_i}{\sum_{i=1}^{n} \dfrac{w_i}{\mathrm{MV}_i} \cdot Y_i} \\
&= \sum_{i=1}^{n} w_i \cdot \omega_i \cdot \mathcal{CI}_i^{\mathrm{Revenue}}
\end{aligned}
$$

where ω_i is the ratio between the revenue per market value of company i and the weighted average revenue per market value of the portfolio:

$$
\omega_i = \frac{\dfrac{Y_i}{\mathrm{MV}_i}}{\sum_{k=1}^{n} w_k \cdot \dfrac{Y_k}{\mathrm{MV}_k}}
$$

Except when all the companies have the same revenue per market value ratio, we deduce that the revenue-based carbon intensity does not satisfy the additivity property since we have $\boldsymbol{\mathcal{CI}}^{\text{Revenue}}(w) \neq \sum_{i=1}^{n} w_i \cdot \boldsymbol{\mathcal{CI}}_i^{\text{Revenue}}$. In order to avoid this problem, we generally use the weighted average carbon intensity (WACI) of the portfolio:

$$\boldsymbol{\mathcal{CI}}^{\text{Revenue}}(w) = \sum_{i=1}^{n} w_i \cdot \boldsymbol{\mathcal{CI}}_i^{\text{Revenue}} \tag{9.5}$$

This method is the standard approach in portfolio management.

Remark 75 *Carbon intensity is additive when we consider a given issuer:*

$$
\begin{aligned}
\boldsymbol{\mathcal{CI}}_i(\boldsymbol{\mathcal{SC}}_{1-3}) &= \frac{\boldsymbol{\mathcal{CE}}_i(\boldsymbol{\mathcal{SC}}_1) + \boldsymbol{\mathcal{CE}}_i(\boldsymbol{\mathcal{SC}}_2) + \boldsymbol{\mathcal{CE}}_i(\boldsymbol{\mathcal{SC}}_3)}{Y_i} \\
&= \boldsymbol{\mathcal{CI}}_i(\boldsymbol{\mathcal{SC}}_1) + \boldsymbol{\mathcal{CI}}_i(\boldsymbol{\mathcal{SC}}_2) + \boldsymbol{\mathcal{CI}}_i(\boldsymbol{\mathcal{SC}}_3)
\end{aligned}
$$

Example 31 *We assume that $\boldsymbol{\mathcal{CE}}_1 = 5 \times 10^6$ CO$_2$e, $Y_1 = \$0.2 \times 10^6$, $\text{MV}_1 = \$10 \times 10^6$, $\boldsymbol{\mathcal{CE}}_2 = 50 \times 10^6$ CO$_2$e, $Y_2 = \$4 \times 10^6$ and $\text{MV}_2 = \$10 \times 10^6$. We invest $W = \$10$ mn.*

We deduce that:

$$\boldsymbol{\mathcal{CI}}_1 = \frac{5 \times 10^6}{0.2 \times 10^6} = 25.0 \text{ tCO}_2\text{e/\$ mn}$$

and $\boldsymbol{\mathcal{CI}}_2 = 12.5$ tCO$_2$e/\$ mn. Since we have:

$$
\begin{cases}
\boldsymbol{\mathcal{CE}}(w) = W\left(w_1 \dfrac{\boldsymbol{\mathcal{CE}}_1}{\text{MV}_1} + w_2 \dfrac{\boldsymbol{\mathcal{CE}}_2}{\text{MV}_2}\right) \\
Y(w) = W\left(w_1 \dfrac{Y_1}{\text{MV}_1} + w_2 \dfrac{Y_2}{\text{MV}_2}\right) \\
\boldsymbol{\mathcal{CI}}(w) = w_1 \boldsymbol{\mathcal{CI}}_1 + w_2 \boldsymbol{\mathcal{CI}}_2
\end{cases}
$$

We obtain the following results:

w_1	w_2	$\boldsymbol{\mathcal{CE}}(w)$ $(\times 10^6$ CO$_2$e$)$	$Y(w)$ $(\times \$10^6)$	$\dfrac{\boldsymbol{\mathcal{CE}}(w)}{Y(w)}$	$\boldsymbol{\mathcal{CI}}(w)$
0%	100%	50.00	4.00	12.50	12.50
10%	90%	45.50	3.62	12.57	13.75
20%	80%	41.00	3.24	12.65	15.00
30%	70%	36.50	2.86	12.76	16.25
50%	50%	27.50	2.10	13.10	18.75
70%	30%	18.50	1.34	13.81	21.25
80%	20%	14.00	0.96	14.58	22.50
90%	10%	9.50	0.58	16.38	23.75
100%	0%	5.00	0.20	25.00	25.00

We notice that the weighted average carbon intensity can be very different than the economic carbon intensity. Let us assume that we buy the two companies, implying that $W = \$20$ mn, $w_1 = 50\%$ and $w_2 = 50\%$. In this case, we obtain $\boldsymbol{\mathcal{CE}}(w) = 55 \times 10^6$ and $Y(w) = \$4 \times 10^6$. The economic carbon intensity is then equal to $55/4 = 13.10$ while the WACI is 18.75.

Remark 76 *For sovereign issuers, the economic carbon intensity is measured in mega-tonnes of CO$_2$e per million dollars of GDP while the physical carbon intensity unit is tCO$_2$e per capita.*

9.2.3 Statistics

Some CCF values are provided in Table 9.15. These figures illustrate some issues in the computation of the carbon footprint at the issuer level. First, it is obvious that it is important to take into account scope 3 to have the real picture of the carbon footprint of an issuer. Indeed, we notice that some issuers have a low scope 1, because they have more or less outsourced the manufacturing of their products. Since a part of the production is located in upstream scope 3, we can not make a fair comparison between issuers if we only consider scope 1 and 2. We face a similar issue with the distribution of the products. The magnitude of some scope 3 carbon intensities raises also the question of their computation. Indeed, while scope 1 and 2 are mandatory to report, there is no obligation for a company to report its scope 3. Moreover, while there is one unique figure for scope 1 and 2 emissions in the CDP reporting files, scope 3 emissions are split into 15 categories, and it is extremely rare that a company reports all scope 3 categories. This explains that the frequency of estimated values is larger for scope 3.

Figure 9.19: Histogram of 2019 carbon intensities (logarithmic scale, $tCO_2e/\$$ mn)

Source: Trucost (2022) & Barahhou *et al.* (2022).

In Figure 9.19, we show the distribution of carbon intensities. Since the range may be very large (from zero to several thousand), we use a logarithmic scale. Moreover, the dotted vertical lines indicate the 5th and 95th percentiles. We observe that the distribution support is very large for scope 1, 2 and 3 downstream. In this case, there are many extreme points with very low and very high carbon intensities. Therefore, we will see that it is relatively easy to reduce the carbon footprint of a portfolio. Now, if we focus on upstream scope 3, we obtain another story, because the range is not so large. Indeed, we do not have issuers with very low carbon intensity. Therefore, incorporating upstream scope 3 will change the nature of portfolio decarbonization, which will become more difficult.

Table 9.15: Examples of 2019 carbon emissions and intensities

Company	Carbon emissions (in tCO$_2$e)				Revenue (in $ mn)	Intensity (in tCO$_2$e/$ mn)			
	SC_1	SC_2	SC_3^{up}	SC_3^{down}		SC_1	SC_2	SC_3^{up}	SC_3^{down}
Airbus	576 705	386 674	12 284 183	23 661 432	78 899	7.3	4.9	155.7	299.9
Allianz	46 745	224 315	3 449 234	3 904 000	135 279	0.3	1.7	25.5	28.9
Alphabet	111 283	5 118 152	7 142 566		161 857	0.7	31.6	44.1	
Amazon	5 760 000	5 500 000	20 054 722	10 438 551	280 522	20.5	19.6	71.5	37.2
Apple	50 549	862 127	27 624 282	5 470 771	260 174	0.2	3.3	106.2	21.0
BNP Paribas	64 829	280 789	1 923 307	1 884	78 244	0.8	3.6	24.6	0.0
Boeing	611 001	871 000	9 878 431	22 959 719	76 559	8.0	11.4	129.0	299.9
BP	49 199 999	5 200 000	103 840 194	582 639 687	276 850	177.7	18.8	375.1	2 104.5
Caterpillar	905 000	926 000	15 197 607	401 993 744	53 800	16.8	17.2	282.5	7 472.0
Danone	722 122	944 877	28 969 780	4 464 773	28 308	25.5	33.4	1 023.4	157.7
Enel	69 981 891	5 365 386	8 726 973	53 774 821	86 610	808.0	61.9	100.8	620.9
Exxon	111 000 000	9 000 000	107 282 831	594 131 943	255 583	434.3	35.2	419.8	2 324.6
JPMorgan Chase	81 655	692 299	3 101 582	15 448 469	115 627	0.7	6.0	26.8	133.6
Juventus	6 665	15 739	35 842	77 114	709	9.4	22.2	50.6	108.8
LVMH	67 613	262 609	11 853 749	942 520	60 083	1.1	4.4	197.3	15.7
Microsoft	113 414	3 556 553	5 977 488	4 003 770	125 843	0.9	28.3	47.5	31.8
Nestle	3 291 303	3 206 495	61 262 078	33 900 606	93 153	35.3	34.4	657.6	363.9
Netflix	38 481	145 443	1 900 283	2 192 255	20 156	1.9	7.2	94.3	108.8
NVIDIA	2 767	65 048	2 756 353	1 184 981	11 716	0.2	5.6	235.3	101.1
PepsiCo	3 552 415	1 556 523	32 598 029	14 229 956	67 161	52.9	23.2	485.4	211.9
Pfizer	734 638	762 840	4 667 225	133 468	51 750	14.2	14.7	90.2	2.6
Roche	288 157	329 541	5 812 735	347 437	64 154	4.5	5.1	90.6	5.4
Samsung Electronics	5 067 000	10 998 000	33 554 245	60 978 947	197 733	25.6	55.6	169.7	308.4
TotalEnergies	40 909 135	3 596 127	49 817 293	456 993 576	200 316	204.2	18.0	248.7	2 280.0
Toyota	2 522 987	5 227 844	66 148 020	330 714 268	272 608	9.3	19.2	242.6	1 213.2
Volkswagen	4 494 066	5 973 894	65 335 372	354 913 446	282 817	15.9	21.1	231.0	1 254.9
Walmart	6 101 641	13 057 352	40 651 079	32 346 229	514 405	11.9	25.4	79.0	62.9

Source: Trucost (2022) & Barahhou et al. (2022).

Table 9.16: Carbon intensity in tCO$_2$e/\$ mn per GICS sector and sector contribution in % (MSCI World, June 2022)

Sector	b_i (in %)	Carbon intensity				Risk contribution			
		$\mathcal{SC}_1$	$\mathcal{SC}_{1-2}$	$\mathcal{SC}_{1-3}^{up}$	$\mathcal{SC}_{1-3}$	$\mathcal{SC}_1$	$\mathcal{SC}_{1-2}$	$\mathcal{SC}_{1-3}^{up}$	$\mathcal{SC}_{1-3}$
Communication Services	7.58	2	28	134	172	0.14	1.31	3.30	1.31
Consumer Discretionary	10.56	23	65	206	590	1.87	4.17	6.92	6.21
Consumer Staples	7.80	28	55	401	929	1.68	2.66	10.16	7.38
Energy	4.99	632	698	1 006	6 823	24.49	21.53	16.33	34.37
Financials	13.56	13	19	52	244	1.33	1.58	2.28	3.34
Health Care	14.15	10	22	120	146	1.12	1.92	5.54	2.12
Industrials	9.90	111	130	298	1 662	8.38	7.83	9.43	16.38
Information Technology	21.08	7	23	112	239	1.13	3.03	7.57	5.06
Materials	4.28	478	702	1 113	2 957	15.89	18.57	15.48	12.93
Real Estate	2.90	22	101	167	571	0.48	1.81	1.57	1.65
Utilities	3.21	1 744	1 794	2 053	2 840	43.47	35.59	21.41	9.24
MSCI World		130	163	310	992				
MSCI World EW		168	211	391	1 155				

Source: MSCI (2022), Trucost (2022) & Barahhou *et al.* (2022).

Let $b = (b_1, \ldots, b_n)$ be the weights of the assets that belong to a benchmark. Its weighted average carbon intensity is given by $\mathcal{CI}(b) = \sum_{i=1}^{n} b_i \cdot \mathcal{CI}_i$ where $\mathcal{CI}_i$ is the carbon intensity of asset i. If we focus on the carbon intensity for a given sector, we use the following formula:

$$\mathcal{CI}(\boldsymbol{S}ector_j) = \frac{\sum_{i \in \boldsymbol{S}ector_j} b_i \cdot \mathcal{CI}_i}{\sum_{i \in \boldsymbol{S}ector_j} b_i}$$

In Table 9.16, we report the carbon intensity of the MSCI World index and its sectors. We also compute the risk contribution of each sector as follows: $\mathcal{RC}(\boldsymbol{S}ector_j) = \left(\sum_{i \in \boldsymbol{S}ector_j} b_i \cdot \mathcal{CI}_i\right)/\mathcal{CI}(b)$. We obtain 130 tCO$_2$e/\$ mn for scope 1, 163 tCO$_2$e/\$ mn if we include scope 2, 310 tCO$_2$e/\$ mn if we add upstream scope 3, and finally 992 tCO$_2$e/\$ mn if we consider the full scope 3. We notice a large cap bias because the MSCI World equally-weighted portfolio shows higher figures. We also observe a high discrepancy between sectors. Low-carbon sectors are Communication Services, Financials, Health Care and Information Technology, whereas high-carbon sectors are Energy, Materials and Utilities. In terms of risk contribution, Consumer Services represents 7.58% of the nominal allocation, but only 0.14% of the carbon allocation if we consider scope 1. If we focus on the first two scopes, Utilities is the main contributor, followed by Energy and Materials. By including upstream scope 3 emissions, the contribution of Consumer Staples becomes significant. We also notice that the Utilities contribution has strongly been reduced whereas the Industrials contribution increases when we consider the three scopes.

Remark 77 *The question of double-counting is less important when we consider carbon intensities, especially monetary measures. Indeed, the carbon intensity can be seen as a scoring system, and portfolio managers generally use carbon intensity in a relative way, and not in an absolute way. For instance, they do not target a given carbon intensity. Their goal is more reducing the carbon intensity relatively to a benchmark, without analyzing the absolute value of the benchmark itself. Moreover, the aggregation at the portfolio level is generally done thanks to the WACI measure, which indicates that the carbon intensity is more viewed as a score than a physical measure.*

9.3   Dynamic risk measures

In this section, we present the basics for building dynamic carbon metrics that are very useful when defining net-zero investment portfolios and assessing the decarbonization policy of issuers. The main tools are the carbon budget, the carbon trend and the carbon target. By combining these tools, we will be able to present the $\mathcal{PAC}$ framework which is the cornerstone of implied temperature ratings (ITR). It measures the participation, the ambition and the credibility of a company to reduce its carbon emissions.

9.3.1   Carbon budget

Definition

The carbon budget defines the amount of GHG emissions that a country, a company or an organization produces over the time period $[t_0, t]$. From a mathematical point of view, it corresponds to the signed area of the region bounded by the function $\mathcal{CE}(t)$:

$$\mathcal{CB}(t_0, t) = \int_{t_0}^{t} \mathcal{CE}(s) \, \mathrm{d}s$$

The carbon budget can be computed with other functions than the carbon emissions. For instance, if the reference level is equal to $\mathcal{CE}^{\star}(t)$ at time t, we obtain:

$$\mathcal{CB}^{\star}(t_0, t) = \int_{t_0}^{t} \mathcal{CE}^{\star}(s) \, \mathrm{d}s$$

Therefore, we can easily compute the excess (or net) carbon budget since we have:

$$\int_{t_0}^{t} (\mathcal{CE}(s) - \mathcal{CE}^{\star}(s)) \, \mathrm{d}s = \mathcal{CB}(t_0, t) - \mathcal{CB}^{\star}(t_0, t)$$

If the reference level is constant — $\mathcal{CE}^{\star}(t) = \mathcal{CE}^{\star}$, the previous formula becomes:

$$\int_{t_0}^{t} (\mathcal{CE}(s) - \mathcal{CE}^{\star}) \, \mathrm{d}s = \mathcal{CB}(t_0, t) - \mathcal{CE}^{\star}(t - t_0)$$

Example 32 *In Table 9.17, we report the historical data of carbon emissions from 2010 to 2020. Moreover, the company has announced his carbon targets for the years until 2050.*

Table 9.17: Carbon emissions in $\mathrm{MtCO_2e}$

t	2010	2011	2012	2013	2014	2015	2016	2017
$\mathcal{CE}(t)$	4.800	4.950	5.100	5.175	5.175	5.175	5.175	5.100
t	2018	2019	2020	2025*	2030*	2035*	2040*	2050*
$\mathcal{CE}(t)$	5.025	4.950	4.875	4.200	3.300	1.500	0.750	0.150

The asterisk * indicates that the company has announced a carbon target for this year.

We consider the carbon pathway given in Example 32 and report different carbon budgets in Figure 9.20. The first panel (top/left) corresponds to the carbon budget that was spent by the company from 2010 to 2020. The second panel (top/right) is the targeted carbon budget that is estimated or planned by the company for the period between 2020 and 2035. The last two panels (bottom/left and bottom/right) considers a constant reference level, which may be for example the

average target of the industry. If we assume that the reference level $\mathcal{CE}^\star$ is equal to 3 MtCO$_2$e (bottom/left panel), we notice that the net carbon budget is the difference between two areas. From January 2020 to October 2030, the carbon emissions are greater than $\mathcal{CE}^\star$ and this period has a positive contribution to the carbon budget. On the contrary, the period from November 2030 to December 2035 has a negative contribution. On average, the net carbon budget is positive. In the case where the reference level $\mathcal{CE}^\star$ is equal to 4 MtCO$_2$e (bottom/right panel), the excess carbon budget is negative.

Figure 9.20: Past, expected and net carbon budgets (Example 32)

Computation of the carbon budget (numerical solution)

We consider the equally-spaced partition $\{[t_0, t_0 + \Delta t], \ldots, [t - \Delta t, t]\}$ of $[t_0, t]$. Let $m = \dfrac{t - t_0}{\Delta t}$ be the number of intervals. We set $\mathcal{CE}_k = \mathcal{CE}(t_0 + k\Delta t)$. The right Riemann approximation is:

$$\mathcal{CB}(t_0, t) = \int_{t_0}^{t} \mathcal{CE}(s) \, \mathrm{d}s \approx \sum_{k=1}^{m} \mathcal{CE}(t_0 + k\Delta t) \Delta t = \Delta t \sum_{k=1}^{m} \mathcal{CE}_k$$

If we use the left Riemann sum, we obtain:

$$\mathcal{CB}(t_0, t) \approx \Delta t \sum_{k=0}^{m-1} \mathcal{CE}_k$$

Finally, the midpoint rule is given by:

$$\mathcal{CB}(t_0, t) \approx \Delta t \sum_{k=1}^{m} \mathcal{CE}\left(t_0 + \frac{k}{2}\Delta t\right)$$

In the case of a yearly partition, the previous formulas are simplified since we have $\Delta t = 1$. For instance, the left Riemann sum becomes:

$$\mathcal{CB}(t_0, t) = \sum_{k=0}^{m-1} \mathcal{CE}_k = \mathcal{CE}(t_0) + \ldots + \mathcal{CE}(t-1)$$

If we consider Example 32, the carbon budget from 1st January 2010 to 1st January 2020 is equal to:

$$\begin{aligned}
\mathcal{CB}(2010, 2020) &= 4.8 + 4.95 + 5.1 + 5.175 + 5.175 + 5.175 + 5.175 + 5.1 + 5.025 + 4.95 \\
&= 50.625 \text{ MtCO}_2\text{e}
\end{aligned}$$

Remark 78 *Instead of Riemann sums, we can use more sophisticated methods such as trapezoidal and Simpon's rules (Roncalli, 2020a, Section A.1.2.3, pages 1037-1041). We can also interpolate the carbon emissions with spline functions and then implement a Gaussian quadrature.*

Computation of the carbon budget (analytical solution)

Constant reduction rate   If we use a constant linear reduction rate $\mathcal{R}(t_0, t) = \mathcal{R}(t - t_0)$, we obtain the following analytical expression:

$$\mathcal{CB}(t_0, t) = \int_{t_0}^{t} (\mathcal{CE}(t_0) - \mathcal{R}(s - t_0)) \, \mathrm{d}s = (t - t_0)\mathcal{CE}(t_0) - \frac{(t - t_0)^2}{2}\mathcal{R} \tag{9.6}$$

In the case of a constant compound reduction rate:

$$\mathcal{CE}(t) = (1 - \mathcal{R})^{(t - t_0)} \mathcal{CE}(t_0)$$

we obtain:

$$\begin{aligned}
\mathcal{CB}(t_0, t) &= \mathcal{CE}(t_0) \int_{t_0}^{t} (1 - \mathcal{R})^{(s - t_0)} \, \mathrm{d}s \\
&= \mathcal{CE}(t_0) \left[\frac{(1 - \mathcal{R})^{(s - t_0)}}{\ln(1 - \mathcal{R})} \right]_{t_0}^{t} \\
&= \frac{(1 - \mathcal{R})^{(t - t_0)} - 1}{\ln(1 - \mathcal{R})} \mathcal{CE}(t_0)
\end{aligned} \tag{9.7}$$

If we assume that $\mathcal{CE}(t) = e^{-\mathcal{R}(t - t_0)}\mathcal{CE}(t_0)$, we have:

$$\mathcal{CB}(t_0, t) = \mathcal{CE}(t_0) \left[-\frac{e^{-\mathcal{R}(s - t_0)}}{\mathcal{R}} \right]_{t_0}^{t} = \mathcal{CE}(t_0) \frac{\left(1 - e^{-\mathcal{R}(t - t_0)}\right)}{\mathcal{R}} \tag{9.8}$$

Remark 79 *If the carbon emissions increase at a positive growth rate g, we set $\mathcal{R} = -g$.*

According to IPCC (2018), the probability that the temperature $\mathcal{T}$ remains below 1.5°C by 2050 depends on a carbon budget. They estimated that the remaining carbon budget $\mathcal{CB}(2019, t)$ is 580 GtCO$_2$e for a 50% probability of limiting warming to 1.5°C, 420 GtCO$_2$e for a 66% probability and 300 GtCO$_2$e for a 83% probability. In Figure 9.21, we have computed $\mathcal{CB}(2019, t)$ by setting $\mathcal{CE}(2019) = 36$ GtCO$_2$e and assuming a constant compound reduction rate $\mathcal{R}$. If nothing is done, the probability to reach 1.5°C by 2035 is close to 50% since we obtain $\mathcal{CB}(2019, 2035) = 571.41$ GtCO$_2$e. With a reduction rate of 7%, we have a probability of 65% to limit global warming to 1.5°C by 2050. This explains that many reduction targets are calibrated on this figure, for example the Paris aligned benchmarks (CTB and PAB).

Figure 9.21: Probability to reach 1.5°C

Linear function If we assume that $\mathcal{CE}(t) = \beta_0 + \beta_1 t$, we deduce that:

$$
\begin{aligned}
\mathcal{CB}(t_0, t) &= \int_{t_0}^{t} (\beta_0 + \beta_1 s) \, \mathrm{d}s \\
&= \left[\beta_0 s + \frac{1}{2} \beta_1 s^2 \right]_{t_0}^{t} \\
&= \beta_0 (t - t_0) + \frac{1}{2} \beta_1 \left(t^2 - t_0^2 \right)
\end{aligned}
\tag{9.9}
$$

We can extend this formula to a piecewise linear function. We assume that $\mathcal{CE}(t)$ is known for $t \in \{t_0, t_1, \ldots, t_m\}$ and $\mathcal{CE}(t)$ is linear between two consecutive dates:

$$
\mathcal{CE}(t) = \mathcal{CE}(t_{k-1}) + \frac{\mathcal{CE}(t_k) - \mathcal{CE}(t_{k-1})}{t_k - t_{k-1}} (t - t_{k-1}) \qquad \text{if } t \in [t_{k-1}, t_k]
$$

We notice that this equation can be written as:

$$
\mathcal{CE}(t) = \underbrace{\frac{t_k}{t_k - t_{k-1}} \mathcal{CE}(t_{k-1}) - \frac{t_{k-1}}{t_k - t_{k-1}} \mathcal{CE}(t_k)}_{\beta_{0,k}} + \underbrace{\frac{\mathcal{CE}(t_k) - \mathcal{CE}(t_{k-1})}{t_k - t_{k-1}}}_{\beta_{1,k}} t
$$

We deduce that:

$$
\mathcal{CB}(t_0, t) = \sum_{k=1}^{k(t)} \int_{t_{k-1}}^{t_k} \mathcal{CE}(s) \, \mathrm{d}s + \int_{t_{k(t)}}^{t} \mathcal{CE}(s) \, \mathrm{d}s
$$

where $k(t) = \{\max k : t_k \leq t\}$. Using Equation (9.9), we conclude that[33]:

$$\mathcal{CB}(t_0, t) = \sum_{k=1}^{k(t)} \beta_{0,k}^{\star}(t_k - t_{k-1}) + \frac{1}{2}\sum_{k=1}^{k(t)} \beta_{1,k}\left(t_k^2 - t_{k-1}^2\right) + \beta_{0,k(t)+1}\left(t - t_{k(t)}\right) + \frac{1}{2}\beta_{1,k(t)+1}\left(t^2 - t_{k(t)}^2\right)$$
$$(9.10)$$

If we consider Example 32 and assume that the carbon emissions are linear between two consecutive years, the carbon budget from 1st January 2010 to 1st January 2020 is equal to 50.662 MtCO$_2$e, which is close to the value 50.625 MtCO$_2$e obtained previously. When we consider the carbon targets, the Riemann sums are not appropriate because the targets are measured every five or ten years. Since the objective is that the company reduces continuously its carbon emissions, it is better to compute the carbon budget by assuming a piecewise linear function. In our example, we obtain $\mathcal{CB}(2020, 2035) = 53.437$ MtCO$_2$e with the following decomposition[34]: $\mathcal{CB}(2020, 2025) = 22.687$ MtCO$_2$e (42.46%), $\mathcal{CB}(2025, 2030) = 18.750$ MtCO$_2$e (35.09%) and $\mathcal{CB}(2030, 2035) = 12.000$ MtCO$_2$e (22.46%). We also have $\mathcal{CB}(2020, 2050) = 63.562$ MtCO$_2$e, implying that the first period 2020–2035 represents 84.07% of the carbon emissions.

Table 9.18: IEA NZE scenario (in GtCO$_2$e)

Sector	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Electricity	12.4	13	13.3	13.5	13.6	13.3	13.3	13.5	14	13.8
Buildings	2.89	2.81	2.78	2.9	2.84	2.87	2.91	2.95	2.98	3.01
Transport	7.01	7.13	7.18	7.37	7.5	7.72	7.88	8.08	8.25	8.29
Industry	8.06	8.47	8.57	8.71	8.78	8.71	8.56	8.52	8.72	8.9
Other	1.87	1.89	1.91	1.96	1.87	1.89	1.89	1.92	1.92	1.91
Gross emissions	32.2	33.3	33.7	34.4	34.5	34.5	34.5	35	35.9	35.9
BECCS/DACCS	0	0	0	0	0	0	0	0	0	0
Net emissions	32.2	33.3	33.7	34.4	34.5	34.5	34.5	35	35.9	35.9

Sector	2020	2025	2030	2035	2040	2045	2050
Electricity	13.5	10.8	5.82	2.12	-0.08	-0.31	-0.37
Buildings	2.86	2.43	1.81	1.21	0.69	0.32	0.12
Transport	7.15	7.23	5.72	4.11	2.69	1.5	0.69
Industry	8.48	8.14	6.89	5.25	3.48	1.8	0.52
Other	1.91	1.66	0.91	0.09	-0.46	-0.82	-0.96
Gross emissions	33.9	30.3	21.5	13.7	7.77	4.3	1.94
BECCS/DACCS	0	-0.06	-0.32	-0.96	-1.46	-1.8	-1.94
Net emissions	33.9	30.2	21.1	12.8	6.32	2.5	0.00

Source: IEA (2021, Figure 2.3, page 55).

---

[33]When t belongs to the set $\{t_0, t_1, \ldots, t_m\}$, we can simplify this expression as follows:

$$\mathcal{CB}(t_0, t) = \sum_{k=1}^{k(t)}\left(\mathcal{CE}(t_{k-1})\,t_k - \mathcal{CE}(t_k)\,t_{k-1}\right) + \frac{1}{2}\sum_{k=1}^{k(t)}\left(\mathcal{CE}(t_k) - \mathcal{CE}(t_{k-1})\right)(t_k + t_{k-1})$$

[34]We use the Chasles property of the Riemann integral:

$$\mathcal{CB}(t_0, t_2) = \int_{t_0}^{t_2}\mathcal{CE}(s)\,\mathrm{d}s = \int_{t_0}^{t_1}\mathcal{CE}(s)\,\mathrm{d}s + \int_{t_1}^{t_2}\mathcal{CE}(s)\,\mathrm{d}s = \mathcal{CB}(t_0, t_1) + \mathcal{CB}(t_1, t_2)$$

Example 33 *We consider the net-zero emissions (NZE) scenario provided by the International Energy Agency (IEA, 2021). We remind that it is a normative scenario that shows a pathway for the global energy sector to achieve net-zero CO_2e emissions by 2050. For each important sector, IEA gives the past trajectory of carbon emissions and the decarbonization pathway that could be achievable (Table 9.18). This net-zero scenario has been calibrated with a carbon budget of approximatively* 500 $GtCO_2e$.

Figure 9.22: CO_2 emissions by sector in the IEA NZE scenario (in $GtCO_2e$)

Source: IEA (2021) & Author's calculations.

In Figure 9.22, we show the decarbonization pathway of each sector and the global economy. We notice the importance of the technologies bioenergy with carbon capture and storage (BECCS) and direct air carbon capture with carbon storage (DACCS). Indeed, they will help to compensate the remaining 2 $GtCO_2e$ of carbon emissions. In Table 9.19, we have computed the carbon budget $\mathcal{CB}(2019, t)$ for each sector and the global system. We notice that the sectors Electricity, Industry and Transport represent about 85% of the global budget.

Table 9.19: Carbon budget in the IEA NZE scenario (in $GtCO_2e$)

t	Electricity	Buildings	Transport	Industry	Other	Gross emissions
2025	74.4	50.2	43.7	16.2	10.8	195.4
2030	115.9	87.8	76.0	26.8	17.3	324.9
2040	140.9	140.0	117.6	39.1	18.8	466.6
2045	139.9	153.2	128.1	41.6	15.6	496.8
2050	138.2	159.0	133.6	42.7	11.2	512.4

Source: IEA (2021) & Author's calculations.

9.3.2   Carbon trend

Linear trend model

Le Guenedal *et al.* (2022) defined the carbon trend by considering the linear trend model:

$$\mathcal{CE}\left(t\right) = \beta_0 + \beta_1 t + u\left(t\right) \tag{9.11}$$

where $u\left(t\right) \sim \mathcal{N}\left(0, \sigma_u^2\right)$. We estimate the parameters β_0 and β_1 with the least squares method and a sample of observations. Therefore, the projected carbon trajectory is given by:

$$\mathcal{CE}^{\mathcal{T}rend}\left(t\right) = \widehat{\mathcal{CE}}\left(t\right) = \hat{\beta}_0 + \hat{\beta}_1 t \tag{9.12}$$

Let $\{t_{\mathcal{F}irst}, t_{\mathcal{F}irst} + 1, \ldots, t_{\mathcal{L}ast}\}$ be the set of observation dates. We interpret $\mathcal{CE}^{\mathcal{T}rend}\left(t\right)$ as follows:

- If $t < t_{\mathcal{F}irst}$, $\mathcal{CE}^{\mathcal{T}rend}\left(t\right)$ is the *back-calculated* value of carbon emissions before the first observation date (retropolation);

- If $t_{\mathcal{F}irst} \leq t \leq t_{\mathcal{L}ast}$, $\mathcal{CE}^{\mathcal{T}rend}\left(t\right)$ is the *predicted* value of carbon emissions and can be compared with the *observed* value of carbon emissions (out-of-sample estimation);

- If $t > t_{\mathcal{L}ast}$, $\mathcal{CE}^{\mathcal{T}rend}\left(t\right)$ is the *forecast* value of carbon emissions and could be compared with the *future* value of carbon emissions when this later will be available (out-of-sample estimation);

This model is very simple and the underlying idea is to extrapolate the past trajectory. Nevertheless, we can derive several metrics that are useful to compare the existing track record of the issuer with its willingness to really reduce its carbon emissions.

Equation (9.12) is not easy to interpret, because the intercept $\hat{\beta}_0$ corresponds to the estimated value $\widehat{\mathcal{CE}}\left(0\right)$ at time $t = 0$. Then, it is convenient to use another base year t_0, implying that Equation (9.11) becomes:

$$\mathcal{CE}\left(t\right) = \beta_0' + \beta_1'\left(t - t_0\right) + u\left(t\right) \tag{9.13}$$

In this case, the carbon trend is given by:

$$\mathcal{CE}^{\mathcal{T}rend}\left(t\right) = \hat{\beta}_0' + \hat{\beta}_1'\left(t - t_0\right) \tag{9.14}$$

We can show that the two models (9.12) and (9.14) are equivalent and give the same value $\widehat{\mathcal{CE}}\left(t\right)$. Indeed, we have the following relationships:

$$\begin{cases} \beta_0' = \beta_0 + \beta_1 t_0 \\ \beta_1' = \beta_1 \end{cases}$$

The new parameterization does not change the slope of the trend, but only the constant $\hat{\beta}_0'$ which is now equal to $\widehat{\mathcal{CE}}\left(t_0\right)$.

Remark 80 *The previous approach can be extended to the carbon intensity measure $\mathcal{CI}\left(t\right)$.*

Example 34 *In Table 9.20, we report the evolution of scope $1+2$ carbon emissions for company A.*

Table 9.20: Carbon emissions in $MtCO_2e$ (company A)

Year	2007	2008	2009	2010	2011	2012	2013
$\mathcal{CE}(t)$	57.8	58.4	57.9	55.1	51.6	48.3	47.1
Year	2014	2015	2016	2017	2018	2019	2020
$\mathcal{CE}(t)$	46.1	44.4	42.7	41.4	40.2	41.9	45.0

Using the carbon emissions given in Table 9.20, we obtain the following estimates[35]: $\hat{\beta}_0 = 2\,970.43$ and $\hat{\beta}_1 = -1.4512$. If we consider the regression model (9.13), the results become $\hat{\beta}'_0 = 57.85$ and $\hat{\beta}'_1 = -1.4512$ if the base year t_0 is set to 2007, and $\hat{\beta}'_0 = 38.99$ and $\hat{\beta}'_1 = -1.4512$ if the base year t_0 is set to 2020. We verify that all the figures are coherent:

$$
\begin{aligned}
\mathcal{CE}^{\mathcal{T}rend}(t) &= 38.99 - 1.4512 \times (t - 2020) \\
&= 2\,970.43 - 1.4512 \times t
\end{aligned}
$$

We notice that the trend model is more intuitive if we use the base year $t_0 = 2020$. The estimated carbon emissions is equal to 38.99 $MtCO_2e$ in 2020 and we observe a reduction of 1.4512 $MtCO_2e$ every year. For instance, the forecast value for the year 2025 is:

$$
\mathcal{CE}^{\mathcal{T}rend}(2025) = 38.99 - 1.4512 \times 5 = 31.73 \text{ } MtCO_2e
$$

We have reported the in-sample estimated values and out-of-sample forecast values in Figure 9.23. When t is set to the year 2020, we observe that there is a gap between the observed value $\mathcal{CE}(t)$ and the estimated value $\widehat{\mathcal{CE}}(t)$ because $\mathcal{CE}(2020) = 45.0 \gg \widehat{\mathcal{CE}}(2020) = 38.99$. We deduce that the current carbon emissions are greater than the figure given by the trend, meaning that the company has made less effort in recent years compared to the past history. We can then rescale the trend model by imposing that the last value $\mathcal{CE}(t_{\mathcal{L}ast})$ is equal to the estimated value $\widehat{\mathcal{CE}}(t_{\mathcal{L}ast})$. We deduce that:

$$
\hat{\beta}'_0 + \hat{\beta}'_1(t_{\mathcal{L}ast} - t_0) = \mathcal{CE}(t_{\mathcal{L}ast}) \Leftrightarrow \hat{\beta}'_0 = \mathcal{CE}(t_{\mathcal{L}ast}) - \hat{\beta}'_1(t_{\mathcal{L}ast} - t_0)
$$

If $t_{\mathcal{L}ast} = t_0$, we obtain $\hat{\beta}'_0 = \mathcal{CE}(t_{\mathcal{L}ast})$. In our example, the rescaled model has the following expression:

$$
\mathcal{CE}^{\mathcal{T}rend}(t) = 45 - 1.4512 \times (t - 2020)
$$

In Figure 9.23, we verify that the rescaled model has the same slope as previously, but it is now coherent with the last observation of carbon emissions. In the sequel, we will always consider rescaled trend models, because they are more relevant.

Log-linear trend model

Instead of a linear model, we can use a log-linear trend model:

$$
\ln \mathcal{CE}(t) = \gamma_0 + \gamma_1(t - t_0) + v(t) \tag{9.15}
$$

where $v(t) \sim \mathcal{N}(0, \sigma_v^2)$. Again, we estimate the parameters γ_0 and γ_1 by ordinary least squares. Let $Y(t) = \ln \mathcal{CE}(t)$ be the logarithmic transform of the carbon emissions. We have:

$$
\hat{Y}(t) = \hat{\gamma}_0 + \hat{\gamma}_1(t - t_0)
$$

---

[35]$\hat{\sigma}_u$ is equal to 2.5844.

Figure 9.23: Linear carbon trend (Example 34)

and:

$$
\begin{aligned}
\widehat{\mathcal{CE}}\,(t) &= \exp\left(\hat{Y}\,(t)\right) \\
&= \exp\left(\hat{\gamma}_0 + \hat{\gamma}_1\,(t - t_0)\right) \\
&= \widehat{\mathcal{CE}}\,(t_0)\exp\left(\hat{\gamma}_1\,(t - t_0)\right)
\end{aligned}
$$

where $\widehat{\mathcal{CE}}\,(t_0) = \exp\left(\hat{\gamma}_0\right)$. The estimator (9.15) does not take into account the variance bias of log-linear models. Indeed, the correct value of the mathematical expectation is equal to[36]:

$$
\begin{aligned}
\mathbb{E}\left[\mathcal{CE}\,(t)\right] &= \mathbb{E}\left[e^{Y(t)}\right] \\
&= \mathbb{E}\left[\mathcal{LN}\left(\gamma_0 + \gamma_1\,(t - t_0)\,, \sigma_v^2\right)\right] \\
&= \exp\left(\gamma_0 + \gamma_1\,(t - t_0) + \frac{1}{2}\sigma_v^2\right)
\end{aligned}
$$

Therefore, we obtain:

$$
\begin{aligned}
\widehat{\mathcal{CE}}\,(t) &= \exp\left(\hat{\gamma}_0 + \hat{\gamma}_1\,(t - t_0) + \frac{1}{2}\hat{\sigma}_v^2\right) \\
&= \widehat{\mathcal{CE}}\,(t_0)\exp\left(\hat{\gamma}_1\,(t - t_0)\right)
\end{aligned}
$$

where $\widehat{\mathcal{CE}}\,(t_0) = \exp\left(\hat{\gamma}_0 + \frac{1}{2}\hat{\sigma}_v^2\right)$. Again, we can rescale the trend model such that $\widehat{\mathcal{CE}}\,(t_{\mathcal{L}ast}) = \mathcal{CE}\,(t_{\mathcal{L}ast})$. It follows that:

$$
\hat{\gamma}_0 + \frac{1}{2}\hat{\sigma}_v^2 = \ln\mathcal{CE}\,(t_{\mathcal{L}ast}) - \hat{\gamma}_1\,(t_{\mathcal{L}ast} - t_0)
$$

---

[36]We remind that $\mathbb{E}\left[X\right] = \exp\left(\mu + \frac{1}{2}\sigma^2\right)$ if $X \sim \mathcal{LN}\left(\mu, \sigma^2\right)$ (see Section A.2.1 on page 660).

If the base year t_0 is equal to the last year $t_{\mathcal{L}ast}$, the forecast value is equal to:

$$\boldsymbol{\mathcal{CE}}^{\mathcal{T}rend}(t) = \boldsymbol{\mathcal{CE}}(t_0) \exp\left(\hat{\gamma}_1(t - t_0)\right) \tag{9.16}$$

Remark 81 *While the slope of the trend is measured in* CO_2e *in the linear trend model, it is measured in % in the log-linear model. In fact, we estimate an absolute trend in the former model and a relative trend in the later model. From Equation (9.15), we have:*

$$\frac{\partial \boldsymbol{\mathcal{CE}}(t)}{\partial t} = \frac{\exp\left(\gamma_0 + \gamma_1(t - t_0) + v(t)\right)}{\partial t} = \gamma_1 \boldsymbol{\mathcal{CE}}(t)$$

We verify that the slope γ_1 *is the relative variation of carbon emissions:*

$$\frac{\dfrac{\partial \boldsymbol{\mathcal{CE}}(t)}{\partial t}}{\boldsymbol{\mathcal{CE}}(t)} = \frac{\partial \ln \boldsymbol{\mathcal{CE}}(t)}{\partial t} = \gamma_1$$

Using Example 34 and the 2020 base year, we obtain the following results: $\hat{\gamma}_0 = 3.6800$, $\hat{\gamma}_1 = -2.95\%$ and $\hat{\sigma}_v = 0.0520$. It follows that $\widehat{\boldsymbol{\mathcal{CE}}}(2020)$ takes the value 39.65 MtCO$_2$e without the correction of the variance bias and 39.70 MtCO$_2$e with the correction of the variance bias. Using the parameterization (9.16), we compare the estimated log-linear trend with the estimated linear trend in Figure 9.24. We notice that the log-linear trend is convex and the future reduction rate of carbon emissions are less important than those obtained with the linear model.

Figure 9.24: Log-linear carbon trend (Example 34)

Example 35 *We consider several historical trajectories of scope* 1 *carbon emissions:*

#	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
1	10.0	11.1	10.5	12.5	13.0	14.8	16.0	16.5	17.0	17.5	19.8
2	10.0	10.2	10.5	11.0	10.8	10.8	13.0	12.5	13.5	13.6	13.6
3	10.0	9.9	9.5	9.0	9.3	8.8	8.7	8.5	9.0	8.0	8.2
4	10.0	9.5	9.0	9.0	8.3	8.1	7.7	6.5	7.0	6.1	6.2

For each historical trajectory, we estimate the two models and report the estimated trends in Figure 9.25. When the slope is positive, the log-linear trend is systematically above the linear trend in the long run ($t \to \infty$). When the slope is negative, we observe the oppositive phenomenon. In this last case, the linear trend $\widehat{\mathcal{CE}}(\infty)$ tends to $-\infty$ while the log-linear trend $\widehat{\mathcal{CE}}(\infty)$ tends to 0. The fact that carbon emissions are negative may be disturbing. From a theoretical viewpoint, it is not impossible because of the impact of negative emissions (due to negative emissions, carbon credits or carbon removal methods for instance). Nevertheless, it is extremely rare especially if we take into account scope 3 emissions. From a modeling viewpoint, it is then better to impose that carbon emissions are positive.

Figure 9.25: Log-linear vs. linear carbon trend (Example 35)

It seems then that the log-linear model is more relevant when the trend is negative. If we consider companies with a positive trend, the log-linear model may produce exploding carbon emissions. Let us compute the ratio of expected growth rates between t_0 and t. For the linear model, the expected growth rate is:

$$\frac{\widehat{\mathcal{CE}}(t) - \mathcal{CE}(t_0)}{\mathcal{CE}(t_0)} = \frac{\mathcal{CE}(t_0) + \hat{\beta}_1(t - t_0) - \mathcal{CE}(t_0)}{\mathcal{CE}(t_0)} = \frac{\hat{\beta}_1(t - t_0)}{\mathcal{CE}(t_0)}$$

while the log-linear model gives:

$$\frac{\widehat{\mathcal{CE}}(t) - \mathcal{CE}(t_0)}{\mathcal{CE}(t_0)} = \frac{\mathcal{CE}(t_0)\exp(\hat{\gamma}_1(t - t_0)) - \mathcal{CE}(t_0)}{\mathcal{CE}(t_0)} = \exp(\hat{\gamma}_1(t - t_0)) - 1$$

Therefore, the ratio of expected growth rates is proportional to:

$$\begin{aligned}
\frac{\exp(\hat{\gamma}_1(t - t_0)) - 1}{\hat{\beta}_1(t - t_0)} &\approx \frac{1}{\hat{\beta}_1(t - t_0)} \sum_{n=1}^{\infty} \frac{\hat{\gamma}_1^n(t - t_0)^n}{n!} \\
&= \frac{\hat{\gamma}_1}{\hat{\beta}_1} \sum_{n=0}^{\infty} \frac{\hat{\gamma}_1^n(t - t_0)^n}{(n+1)!} \\
&\approx \frac{\hat{\gamma}_1}{\hat{\beta}_1} \left(1 + \frac{1}{2}\hat{\gamma}_1(t - t_0) + \frac{1}{6}\hat{\gamma}_1^2(t - t_0)^2 + \frac{1}{24}\hat{\gamma}_1^3(t - t_0)^3 + \ldots\right)
\end{aligned}$$

We deduce that the exploding effect cannot be avoided. This is why we must be very careful when we consider the log-linear model. One possible solution is to use the linear model when $\hat{\gamma}_1$ is greater than a threshold and the log-linear model otherwise. In Table 9.21, we compute the multiplication factor $M(\gamma_1, t) = e^{\gamma t}$ for different values of the growth rate γ_1 and different time horizon. For instance, if we target the year 2050, we have $t \approx 30$ years, we notice that carbon emissions are multiplied by a factor greater than 10 if the growth rate is 8%. Since there are many uncertainties about data collection and data computation, an historical growth rate of 8% may be explained by several factors:

- The company has really increased its carbon emissions by 8%;

- The company has underestimated its carbon emissions in the past and is more conservative today;

- The company has changed the reporting perimeter;

- Etc.

Applying a factor of 10 with a 30-year time horizon may then be not realistic. This is why we must be careful when estimating the carbon trend and analyze the outlier companies.

Table 9.21: Multiplication factor $M(\gamma_1, t)$

t (in years)	γ_1									
	1%	2%	3%	4%	5%	6%	7%	8%	9%	10%
1	1.01	1.02	1.03	1.04	1.05	1.06	1.07	1.08	1.09	1.11
5	1.05	1.11	1.16	1.22	1.28	1.35	1.42	1.49	1.57	1.65
10	1.11	1.22	1.35	1.49	1.65	1.82	2.01	2.23	2.46	2.72
15	1.16	1.35	1.57	1.82	2.12	2.46	2.86	3.32	3.86	4.48
20	1.22	1.49	1.82	2.23	2.72	3.32	4.06	4.95	6.05	7.39
25	1.28	1.65	2.12	2.72	3.49	4.48	5.75	7.39	9.49	12.18
30	1.35	1.82	2.46	3.32	4.48	6.05	8.17	11.02	14.88	20.09

Remark 82 *The expected growth rate is related to the concept of carbon momentum that will be presented later on page 527.*

Stochastic trend model

Following Roncalli (2020a), the linear trend model can be written as:

$$\begin{cases} y(t) = \mu(t) + u(t) \\ \mu(t) = \mu(t-1) + \beta_1 \end{cases}$$

where $u(t) \sim \mathcal{N}\left(0, \sigma_u^2\right)$. In this case, we have $y(t) = \beta_0 + \beta_1 t + u(t)$ where $\beta_0 = \mu(t_0) - \beta_1 t_0$. A way to introduce a stochastic trend is to add a noise $\eta(t)$ in the trend equation: $\mu(t) = \mu(t-1) + \beta_1 + \eta(t)$ where $\eta(t) \sim \mathcal{N}\left(0, \sigma_\eta^2\right)$. Let us now assume that the slope of the trend is also stochastic:

$$\begin{cases} y(t) = \mu(t) + u(t) \\ \mu(t) = \mu(t-1) + \beta_1(t-1) + \eta(t) \\ \beta_1(t) = \beta_1(t-1) + \zeta(t) \end{cases}$$

where $\zeta(t) \sim \mathcal{N}\left(0, \sigma_\zeta^2\right)$. This model is called the local linear trend (LLT) model (Roncalli, 2020a, page 653). Using the Kalman filter (KF), we can estimate both the stochastic trend $\mu(t)$ and the stochastic slope $\beta_1(t)$.

Let us come back to Example 34. We estimate the parameters $(\sigma_u, \sigma_\eta, \sigma_\zeta)$ by maximizing the Whittle log-likelihood function (Roncalli, 2020a, pages 686-687). We obtain $\hat{\sigma}_u = 0.7022$, $\hat{\sigma}_\eta = 0.7019$ and $\hat{\sigma}_\zeta = 0.8350$. We deduce that the standard deviation of the stochastic slope variation $\beta_1(t) - \beta_1(t-1)$ is equal to 0.8350 $MtCO_2e$. This indicates that there is a high uncertainty in the trend computation. Then, we run the Kalman filter to estimate $\hat{\mu}(t)$ and $\hat{\beta}_1(t)$. In Table 9.22, we report these values and compare them with the estimates using the rolling least squares (RLS). We notice that the time-varying slope produced by the Kalman filter may be very different from the one produced by the method of least squares[37]. In particular, the magnitude of the variability is not the same. This is normal since the rolling least squares estimate a global slope from the beginning of the sample to time t whereas the Kalman filter estimates a local slope for the period $[t-1, t]$. Therefore, $\hat{\beta}_i(t)$ is an estimator of the average slope in the case of the linear trend model and an estimator of the marginal slope in the case of the local linear trend model.

Remark 83 *With the stochastic trend model, Le Guenedal et al. (2022) introduce the concept of carbon velocity, which measures the normalized slope change between $t - h$ and t:*

$$\boldsymbol{v}^{(h)}(t) = \frac{\hat{\beta}_1(t) - \hat{\beta}_1(t-h)}{h}$$

The rationale for this measure is the following. A commitment to reduce carbon emissions implies a negative trend: $\hat{\beta}_1(t) < 0$. Nevertheless, it can take many years for a company to change the sign of the trend slope if it has a bad track record. Therefore, we can use the velocity to verify that the company is making significant efforts in the recent period. In this case, we must have $\boldsymbol{v}^{(h)}(t) < 0$ for low values[38] of h. In the case of the local linear trend model, we notice that the one-step velocity is equal to the innovation of the slope:

$$\boldsymbol{v}^{(1)}(t) = \hat{\beta}_1(t) - \hat{\beta}_1(t-1) = \hat{\zeta}(t)$$

---

[37]For instance, we have $\hat{\beta}_1(2020) = -1.4512$ $MtCO_2e$ with the least squares method and $\hat{\beta}_1(2020) = +1.7701$ $MtCO_2e$ with the Kalman filter.

[38]Generally, h is equal to 1, 2 or 3 years.

Table 9.22: Kalman filter estimation of the stochastic trend (Example 34)

t	$\mathcal{CE}(t)$	$\hat{\beta}_1(t)$ (RLS)	$\beta_1(t)$ (KF)	$\mu(t)$ (KF)
2007	57.80		0.0000	57.80
2008	58.40		0.2168	58.25
2009	57.90	0.0500	-0.0441	58.00
2010	55.10	-0.8600	-1.3941	55.56
2011	51.60	-1.5700	-2.6080	52.01
2012	48.30	-2.0200	-3.1288	48.47
2013	47.10	-2.0929	-2.2977	46.82
2014	46.10	-2.0321	-1.5508	45.85
2015	44.40	-1.9817	-1.5029	44.38
2016	42.70	-1.9406	-1.5887	42.73
2017	41.40	-1.8891	-1.4655	41.36
2018	40.20	-1.8329	-1.3202	40.15
2019	41.90	-1.6824	0.1339	41.41
2020	45.00	-1.4512	1.7701	44.45

Carbon momentum

Le Guenedal *et al.* (2022) define the long-term carbon momentum as the growth rate of carbon emissions. In the case of the linear trend model, we have:

$$\mathcal{CM}^{\mathcal{L}ong}(t) = \frac{\hat{\beta}_1(t)}{\mathcal{CE}(t)}$$

while it is directly equal to $\hat{\gamma}_1(t)$ in the case of the log-linear trend model:

$$\mathcal{CM}^{\mathcal{L}ong}(t) = \hat{\gamma}_1(t)$$

Le Guenedal *et al.* (2022) also define the short-term carbon momentum as the one-year carbon velocity:

$$\mathcal{CM}^{\mathcal{S}hort}(t) = \frac{\boldsymbol{v}^{(1)}(t)}{\mathcal{CE}(t)}$$

If we apply this concept to the log-linear model, we obtain $\mathcal{CM}^{\mathcal{S}hort}(t) = \boldsymbol{v}^{(1)}(t)$ where:

$$\boldsymbol{v}^{(h)}(t) = \frac{\hat{\gamma}_1(t) - \hat{\gamma}_1(t-h)}{h}$$

In the case of the stochastic trend model, we have $\mathcal{CM}^{\mathcal{S}hort}(t) = \hat{\zeta}(t)$.

Remark 84 *Carbon momentum plays a key role when we will define net-zero investment portfolios, because it is highly related to the concept of self-decarbonization*[39].

---

[39]See Section 11.3 on page 591.

Application

Table 9.23 and 9.24 gives some statistics about carbon momentum. It reproduces the results obtained by Barahhou *et al.* (2022) by considering the issuers of the MSCI World index. Since it is difficult to obtain at least 5-year historical data, we focus on the scopes $\mathcal{SC}_1$, $\mathcal{SC}_{1-2}$ and $\mathcal{SC}_{1-3}^{\mathrm{up}}$, and we do not consider the scope $\mathcal{SC}_{1-3}$. If we use the linear trend model, the median value of $\mathcal{CM}^{\mathcal{L}ong}(t)$ is equal to 0% for scope 1, 1.6% when we include the scope 2, and 2.3% when we add the upstream scope 3. The carbon momentum is negative for only 29.4% of issuers when we consider $\mathcal{SC}_{1-3}^{\mathrm{up}}$. This means that a majority of issuers have a positive carbon trend. For instance, about 10% of issuers have a carbon momentum greater than 10%! If we consider carbon intensity instead of carbon emission, we obtain another story. Indeed, issuers with a negative trend dominate issuers with a positive trend. Therefore, it is easier to build a self-decarbonized portfolio when we consider the carbon intensity measure. If we estimate the carbon momentum with the log-linear trend model, results are slightly different. For instance, 19.2% of issuers have a carbon momentum $\mathcal{SC}_{1-3}^{\mathrm{up}}$ greater than 10% versus 8.0% with the linear trend model.

Table 9.23: Statistics (in %) of carbon momentum $\mathcal{CM}^{\mathcal{L}ong}(t)$ (MSCI World index, $1995 - 2021$, linear trend)

Statistics	Carbon emissions			Carbon intensity		
	$\mathcal{SC}_1$	$\mathcal{SC}_{1-2}$	$\mathcal{SC}_{1-3}^{\mathrm{up}}$	$\mathcal{SC}_1$	$\mathcal{SC}_{1-2}$	$\mathcal{SC}_{1-3}^{\mathrm{up}}$
Median	0.0	1.6	2.3	-4.8	-2.4	-1.3
Negative	49.9	41.1	29.4	76.0	69.6	75.6
Positive	50.1	58.9	70.6	24.0	30.4	24.4
$< -10\%$	23.4	15.8	5.8	36.0	25.0	5.7
$< -5\%$	32.1	22.2	10.6	48.6	36.7	13.4
$> +5\%$	22.9	27.5	23.6	6.2	7.3	2.7
$> +10\%$	9.2	9.5	8.0	2.3	2.6	1.0

Source: Trucost database (2022) & Authors' calculations.

Table 9.24: Statistics (in %) of carbon momentum $\mathcal{CM}^{\mathcal{L}ong}(t)$ (MSCI World index, $1995 - 2021$, log-linear trend)

Statistics	Carbon emissions			Carbon intensity		
	$\mathcal{SC}_1$	$\mathcal{SC}_{1-2}$	$\mathcal{SC}_{1-3}^{\mathrm{up}}$	$\mathcal{SC}_1$	$\mathcal{SC}_{1-2}$	$\mathcal{SC}_{1-3}^{\mathrm{up}}$
Median	-0.1	1.7	2.8	-3.6	-1.9	-1.2
Negative	50.6	40.3	29.0	76.3	69.0	75.8
Positive	49.4	59.7	71.0	23.7	31.0	24.2
$< -10\%$	13.6	8.0	2.8	20.8	12.3	2.1
$< -5\%$	26.6	16.9	7.5	42.3	29.0	8.4
$> +5\%$	29.8	35.9	37.1	9.0	10.1	4.0
$> +10\%$	16.9	19.4	19.2	4.0	4.1	1.6

Source: Trucost database (2022) & Authors' calculations.

9.3.3 Participation, ambition and credibility for an alignment strategy

In this section, we define the three pillars that help to evaluate a company's alignment strategy with respect to a given climate scenario, *e.g.*, the net-zero emissions scenario. These three pillars are participation, ambition and credibility. They form the $\mathcal{PAC}$ framework, and they can be quantified using the tool of carbon budget.

Carbon target and decarbonization scenario

In addition to the historical pathway of carbon emissions, the $\mathcal{PAC}$ framework requires two other time series:

- The reduction targets announced by the company;

- The market-based sector scenario associated to the company that defines the decarbonization pathway.

Carbon reduction targets are defined by companies at a scope emissions level with different time horizons[40]. For instance, the issuer can commit to reduce its scope 1 emissions by 50% over a period of 20 years and its scope 3 emissions by 30% over a period of 10 years. Even if the time frame of carbon reduction targets goes to 60 years, most of reduction targets concern the next twenty years. In the CDP database, we observe that most targets are underway or new, and a large proportion of companies set targets to reduce emissions by less than 50% from their base year. We also notice that some targets are reported over multiple scopes[41] and we can have multiple release dates. Therefore, it is important to transform these heterogenous figures into a unique reduction pathway with one base year t_0:

$$\mathbb{CT} = \left\{ \boldsymbol{\mathcal{R}}^{\mathcal{T}arget} \left(t_0, t_k \right), k = 1, \ldots, n_T \right\}$$

where n_T is the number of targets and $\boldsymbol{\mathcal{R}}^{\mathcal{T}arget} \left(t_0, t_k \right)$ is the reduction rate between t_0 and t_k for the kth target.

Concerning the market-based scenario, we generally use sector scenarios provided by IPCC, IEA or IIASA. In some circumstances, we can take global scenarios, but only when we do not have the choice because there is no appropriate scenario for the sector. Again, the decarbonization scenario is defined as a set of reduction rates:

$$\mathbb{CS} = \left\{ \boldsymbol{\mathcal{R}}^{\mathcal{S}cenario} \left(t_0, t_k \right), k = 1, \ldots, n_S \right\}$$

where n_S is the number of scenario data points. The reduction rate is calculated as follows:

$$\boldsymbol{\mathcal{R}}^{\mathcal{S}cenario} \left(t_0, t_k \right) = 1 - \frac{\boldsymbol{\mathcal{CE}}^{\mathcal{S}cenario} \left(t_k \right)}{\boldsymbol{\mathcal{CE}}^{\mathcal{S}cenario} \left(t_0 \right)}$$

where t_0 is the base year and $\boldsymbol{\mathcal{CE}}^S \left(t_k \right)$ is the value of carbon emissions at time t_k in the market-based scenario. For instance, if we consider the IEA NZE scenario (see Table 9.18 on page 518), we obtain the results given in Table 9.25. Carbon emissions have been floored at zero in order to verify that the reduction rate is always less than or equal to 100%. We notice that the Electricity sector must decarbonize very quickly: -20% in 2025, -57% in 2030 and -84% in 2035. The carbon emissions reduction of the Industry and Transport sectors is delayed and really begins after 2025. If we consider the global scenario, the reduction rate is set to 10% in 2025 and increases by 5% every year until 2035.

---

[40]Carbon reduction targets can be found in the CDP database.
[41]For instance, the target can concern only one scope, scope $1 + 2$ or all scopes.

Table 9.25: Reduction rates of the IEA NZE scenario (base year = 2020)

Year	Electricity	Industry	Transport	Buildings	Other	Global
2025	20.0	4.0	−1.1	15.0	13.1	10.6
2030	56.9	18.8	20.0	36.7	52.4	36.6
2035	84.3	38.1	42.5	57.7	95.3	59.6
2040	100.0	59.0	62.4	75.9	100.0	77.1
2045	100.0	78.8	79.0	88.8	100.0	87.3
2050	100.0	93.9	90.3	95.8	100.0	94.3

Source: IEA (2021) & Author's calculations.

Definition of the $\mathcal{PAC}$ framework

The $\mathcal{PAC}$ framework has been introduced by Le Guenedal *et al.* (2022) and is based on the relative positioning of three carbon trajectories: (1) the historical trajectory and its trend, (2) the carbon targets and (3) the decarbonization scenario. It helps to answer several operational questions.

1. First, is the trend of the issuer in line with the scenario? In this case, we would like to know if the company has already reduced its carbon emissions. While the reduction targets correspond to future intentions, the carbon trend measures the past efforts of the company.

2. Is the commitment of the issuer to fight climate change ambitious? In particular, we would like to know if the target trajectory is above, below or in line with the market-based scenario, which is appropriate for the sector of the issuer. This is an important topic, because achieving the net-zero emissions scenario can only be possible if there are no free riders.

3. Finally, a third question is critical and certainly the most important issue. Is the target setting of the company relevant and robust? Indeed, we may wonder if the target trajectory is a too ambitious promise and a form of greenwashing or, on the contrary, a plausible decarbonization pathway.

Therefore, the assessment of the company's targets has three dimensions or pillars: (historical) participation, ambition and credibility. They form the $\mathcal{PAC}$ framework.

Example 36 *We consider again Example 34. Company A has announced the following targets:* $\mathcal{R}^{Target}(2020, 2025) = 40\%$, $\mathcal{R}^{Target}(2020, 2030) = 50\%$, $\mathcal{R}^{Target}(2020, 2035) = 75\%$, $\mathcal{R}^{Target}(2020, 2040) = 80\%$ *and* $\mathcal{R}^{Target}(2020, 2050) = 90\%$. *Since company A is an utility corporation, we propose to use the IEA NZE scenario for the sector Electricity.*

We have reported the different pathways of company A in Figure 9.26. We notice that the announced targets are below the carbon trend except in 2050. A comparison between the targets and the global scenario indicates that company A is more ambitious than the average firm. Nevertheless, the comparison is less favorable when we consider the decarbonization scenario of the corresponding sector. In order to quantify the relative the relative positioning of these trajectories, we compute the carbon budgets with the different pathways. If we consider the time horizon 2035, the carbon budget of the targets is slightly lower than the carbon budget of the decarbonization scenario (388 vs. 407 MtCO$_2$e). This indicates a true ambition to reduce its carbon emissions in line with what the market expects. Nevertheless, we observe a high carbon budget based on the trend model (512 MtCO$_2$e). This questions the credibility of the targets, even if the company has done some efforts in the past.

Table 9.26: Comparison of carbon budgets (base year $= 2020$, Example 36)

Year	Trend (linear)	Trend (log-linear)	Target	Scenario (global)	Scenario (electricity)
2025	207	209	180	213	203
2030	377	390	304	385	341
2035	512	546	388	502	407
2040	610	680	439	573	425
2045	671	796	478	613	425
2050	697	896	506	634	425

Assessment of the $\mathcal{PAC}$ pillars

These three pillars depend on the carbon trajectories $\mathcal{CE}(t)$, $\mathcal{CE}^{\mathcal{T}rend}(t)$, $\mathcal{CE}^{\mathcal{T}arget}(t)$ and $\mathcal{CE}^{\mathcal{S}cenario}(t)$, where $\mathcal{CE}(t)$ is the time series of historical carbon emissions, $\mathcal{CE}^{\mathcal{T}rend}(t)$ and $\mathcal{CE}^{\mathcal{T}arget}(t)$ are the estimated carbon emissions deduced from the trend model and the targets, and $\mathcal{CE}^{\mathcal{S}cenario}(t)$ is the market-based decarbonization scenario. Generally, the participation only depends on the past observations and corresponds to the track record analysis of historical carbon emissions. The ambition compares the target trajectory on one side and the scenario or the trend on the other side. Indeed, we measure to what extent companies are willing to reverse their current carbon emissions and have objectives that match the scenario. Finally, we can measure the credibility of the targets by comparing the current trend of carbon emissions and the reduction targets or by analyzing the recent dynamics of the track record.

We note $t_{\mathcal{F}irst}$ as the first date, $t_{\mathcal{L}ast}$ as the last reporting date and $t_{\mathcal{S}cenario}$ as the target date of the decarbonization scenario. In Figure 9.27, we illustrate the underlying ideas of the $\mathcal{PAC}$ pillars. Let us consider the first three panels. They show the historical carbon emissions of different companies. It is obvious that the company in the top-left panel has a positive participation to slow global warming, whereas the participation of the company in the top-center panel is negative. In the top-right panel, we give three examples that are mixed. In this case, we do not observe a clear pattern: downward or upward trend of carbon emissions. Therefore, the company's participation can be measured by the metrics that are related to the carbon trend. The next three panels in Figure 9.27 illustrate the ambition pillar. In this case, we directly compare the carbon targets of the company and the market-based risk scenario. The companies belong to the same sector, implying that the decarbonization scenario is the same for the middle-left, middle-center and middle-right panels. The middle-left panel shows an ambitious company since its carbon targets are lower than the market-based scenario. In other words, the company has announced that it will make a greater effort than is expected by the market. On the contrary, the company in the middle-center panel is less ambitious, because it plans to reduce its carbon emissions at a slower pace. Finally, the middle-right panel presents two mixed situations. The first one concerns a company that has high ambitions at the beginning of the period $[t_{\mathcal{L}ast}, t_{\mathcal{S}cenario}]$ but it has not disclosed its ambitions for the end of the period. The company's ambition in the short term is then counterweighted by the absence of ambition in the long run. The second example is about a company that concentrates its ambition in the long run. These two examples question the true willingness of these companies to substantially reduce their carbon emissions. Finally, the credibility pillar is illustrated in the last three panels in Figure 9.27. In this case, we compare the carbon emissions trend and the targets communicated by the company. The bottom-left panel corresponds to a credible company, since it has announced more or less a reduction trajectory that is in line with what it has done in the

Figure 9.26: Carbon trend, targets and NZE scenario of company A

past. This is not the case of the company in the bottom-center panel. Clearly, it has announced a reduction of its carbon emissions, but it has continuously increased them in the past. Again, the bottom-right panel presents a mixed situation. The company has announced a reduction trajectory that is not very far from the past trend, but there are two issues. The first one is that it has increased its carbon emissions in the short term, implying that we can have some doubts about the downward trend. The second issue is that it accelerates its objective of carbon emissions reduction at the end of the period $[t_{\mathcal{L}ast}, t_{\mathcal{S}cenario}]$ in order to meet the requirements of the market-based scenario, but its efforts are not very substantial in the short term.

Temperature scoring system

Le Guenedal *et al.* (2022) derive many metrics to measure the three dimensions. They can be classified into four main families. The gap metrics measure the differences between two trajectories or carbon budgets[42]. The duration metrics calculate the time which is necessary to achieve a given objective[43]. The velocity metrics assess the short-term dynamics of the carbon emissions. Finally,

---

[42]An example is the difference between the trend budget and the scenario budget:

$$\mathcal{G}ap\left(t, t^{\mathcal{S}cenario}\right) = \mathcal{CB}^{\mathcal{T}rend}\left(t, t^{\mathcal{S}cenario}\right) - \mathcal{CB}^{\mathcal{S}cenario}\left(t, t^{\mathcal{S}cenario}\right)$$

[43]For instance, the trend duration is the time horizon for achieving zero carbon emissions:

$$\boldsymbol{\mathcal{D}}^{\mathcal{T}rend} = \inf\left\{t : \boldsymbol{\mathcal{CE}}^{\mathcal{T}rend}(t) \leq 0\right\}$$

Figure 9.27: Illustration of the participation, ambition and credibility pillars

the last category computes the short-term reduction that the company must implement to satisfy the objective imposed by the market-based scenario[44]. Each pillar can then be studied using different metrics. Like the ESG risk, we can use a scoring system in order to analyze the $\mathcal{PAC}$ pillars and build three scores: the participation score $\mathcal{PS}_i$, the ambition score $\mathcal{AS}_i$ and the credibility score $\mathcal{CS}_i$. In Figure 9.28, we have represented several configurations of the $\mathcal{PAC}$ scoring system. If the three scores $\mathcal{PS}_i$, $\mathcal{AS}_i$ and $\mathcal{CS}_i$ are high and greater than 0.5 (which is the median value of a q-score), the company is both ambitious and credible and has already made some efforts to reduce its carbon emissions (Panel (a)). On the contrary, in Panel (b), we have a company, whose three scores are below the median. These two extreme cases are very frequent. Nevertheless, we can also obtain a more balanced scoring. For instance, Panel (c) corresponds to a company that has substantially reduced its past emissions but has announced weak reduction targets. Therefore, its ambition score is low, but its credibility score is high. It may be a company that does not talk a lot about its climate change policy, but its track record has demonstrated that it is committed. Finally, Panel (d) represents the scoring of a company with very high ambition, but it has continuously increased

---

[44]The burn-out scenario refers to a sudden and violent reduction of carbon emissions such that the gap is equal to zero.

its carbon emissions in the past. Therefore, we can suspect a type of greenwashing. These examples show that the three dimensions are correlated. For instance, we can assume a positive correlation between participation and credibility, and a negative correlation between ambition and credibility. Indeed, high credibility can only be obtained if participation is high or ambition is weak. Similarly, low credibility can be associated with excessively high ambition or weak participation, implying that the correlation between participation and credibility is unclear.

Figure 9.28: The $\mathcal{PAC}$ scoring system

Source: Le Guenedal *et al.* (2022).

Remark 85 *While the $\mathcal{PAC}$ framework is the backbone to analyze the decarbonization commitment of the company's climate strategy (like the ESG pillars for analyzing the extra-financial risks of the company), the $\mathcal{PAC}$ scoring system is similar to the ESG scoring system. Most of implied temperature ratings are based on the participation, ambition and credibility pillars[45].*

9.3.4   Illustration

We consider the analysis done by Le Guenedal *et al.* (2022). The base year is 2013 and they rebase the trajectories by the carbon emissions $\mathcal{CE}$ (2013). The market-based scenario is the IEA NZE scenario. Company B is a US based multinational technology conglomerate. Carbon emissions and targets are reported in Figure 9.29. Company B is a particularly relevant example of the expectations from investors in the NZE context. Indeed, participation switched favorably after 2018 with a significant reduction in the scope 3 emissions from the use of sold products. The credibility is confirmed with the 2020 data point as the duration $\mathcal{D}^{\mathcal{T}rend}$ drops under the duration of the NZE scenario. Figure 9.30 illustrates Company C which is a major US airline. Its participation switched favorably in 2020.

---

[45]See Section 10.4.1 on page 545.

The credibility has not switched even with the significant drop in 2020, since the duration $\mathcal{D}^{\mathcal{T}rend}$ remains larger than the NZE time horizon. In fact, the reduction of $\mathcal{CE}$ (2020) sourced from both scope 1 and scope 3 emissions is related to the drop in activity due to the Covid-19 crisis. Company D is a European multinational company which supplies industrial resources and services to various industries (Figure 9.31). The company has a clear ambition and has embraced the NZE context. However, the metrics indicate that in terms of participation, the trend has not been negative and has deteriorated in previous years. We stress here that although Company D pays attention to its carbon intensity policy, it has not been active on the absolute carbon emissions level.

Figure 9.29: Carbon emissions, trend, targets and NZE scenario (Company B)

On a global basis, we observe an increase of carbon emissions between 2013 and 2019, and a plateau in 2020, which is probably due to emissions reduction related to the Covid-19 crisis (Figure 9.32). Carbon targets are in line with the NZE scenario until 2025. After this date, we clearly see that the targeted reduction rates are lower than the NZE required reduction rates. We notice that the reduction targets are more or less in line with the NZE scenario. However, and more strikingly, there is a huge gap between the upward trend between 2013 and 2020 and what has been announced by the companies. Figure 9.32 perfectly illustrates the interest of the $\mathcal{PAC}$ framework, since we observe inconsistencies between the ambition of these issuers on one side, and their participation and credibility on another side. The sector analysis is interesting, because we observe some large differences between the sectors in Figure 9.33. First, the trajectory of carbon emissions is highly dependent on the sector. Electricity is the sector that has been making the greatest effort whereas we observe a large increase in carbon emissions for the Industry sector. The impact of the Covid-19 pandemic on the transport sector is particularly striking. Second, there are small differences between the reduction targets, except for the Transport sector that is slightly less ambitious.

Figure 9.30: Carbon emissions, trend, targets and NZE scenario (Company C)

Figure 9.31: Carbon emissions, trend, targets and NZE scenario (Company D)

Figure 9.32: Carbon emissions, trend, targets and NZE scenario (median analysis, global universe)

Source: CDP database (2021), IEA (2021) & Le Guenedal *et al.* (2022).

Figure 9.33: Carbon emissions, trend, targets and NZE scenario (median analysis, sector universe)

Source: CDP database (2021), IEA (2021) & Le Guenedal *et al.* (2022).

9.4   Greenness measures

Until now, we have focused on carbon metrics to measure the "*brownness*" of organizations or products. We now consider "*greenness*" measures, whose objective is to assess the positive contribution to limit global warming. In some sense, brownness and greenness measures are related. Nevertheless, we cannot deduce one measure from another one. Let us define the brown and green intensities of a company as the proportion of brown and green activities. We note them $\mathcal{BI}$ and $\mathcal{GI}$. By construction, we have $\mathcal{BI} \in [0,1]$, $\mathcal{GI} \in [0,1]$ and $0 \leq \mathcal{BI} + \mathcal{GI} \leq 1$. Most of the time, we have $\mathcal{BI} + \mathcal{GI} \neq 1$, meaning that we cannot deduce the green intensity from the brown intensity. While the carbon footprint is a well-defined concept, greenness is then more difficult to assess. In fact, it is a multi-faceted concept. For instance, if one company changes its business model so that its new products are carbon efficient, we can measure the company's greenness based on the avoided emissions generated by the change of the business model. For other companies, the greenness can be evaluated by estimating the R&D amount dedicated to green projects. Therefore, we observe a big difference between carbon and greenness metrics. Indeed, while it makes sense to compute the carbon footprint of all companies, the greenness may be indefinite for some companies, because they have no vocation to participate in the transition to a low-carbon economy. Therefore, Figure 9.34 illustrates that we cannot classifies all activities into these two categories, since there are many activities that are neither brown nor green. Some companies are then neutral and are not exposed to the green business. These remarks argue in favor of considering simple and homogeneous measures of greenness.

Figure 9.34: Brown and green activities at the company level

(a) Green activities

(b) Brown activities

(c) All activities

9.4.1   Green taxonomy

The purpose of a green taxonomy is to define what is green, and its objective is to inform investors about the greenness of their investments. Therefore, they can evaluate whether these levels satisfy or not their expectations. A green taxonomy is all the more important as we observe a strong development of green sentiment among investors (Brière and Ramelli, 2021). In this context, the investor may want to assess the proportion of his portfolio that is invested in environmentally sustainable assets. Therefore, a green taxonomy is necessary for both asset owners and managers.

The most famous example is the European green taxonomy, which has been already presented on page 30. We remind that the EU taxonomy for sustainable activities is "*a classification system, establishing a list of environmentally sustainable economic activities.*". These economic activities must have a substantive contribution to at least one of the following six environmental objectives: (1) climate change mitigation; (2) climate change adaptation; (3) sustainable use and protection of water and marine resources; (4) transition to a circular economy; (5) pollution prevention and control; (6) protection and restoration of biodiversity and ecosystem. Moreover, a business activity must also meet two other criteria to qualify as sustainable. First, the activity must do no significant

harm to the other environmental objectives (DNSH constraint) and second, it must comply with minimum social safeguards (MS constraint). Figure 9.35 summarizes the different steps.

Figure 9.35: EU taxonomy for sustainable activities

Remark 86 *The EU taxonomy is not finalized and only concerns the first two objectives as of today (January 2023).*

9.4.2 Green revenue share

Relationship between the green intensity and the green revenue share

There are several ways to compute the green intensity. This is why we observe some significant differences between data providers. One method is to translate the 3-step approach of the EU taxonomy into the following equation:

$$\boldsymbol{\mathcal{GI}} = \frac{\mathcal{GR}}{\mathcal{TR}} \cdot (1 - \mathcal{P}) \cdot \mathbb{1}\left\{\boldsymbol{\mathcal{S}} \geq \boldsymbol{\mathcal{S}}^\star\right\}$$

where $\mathcal{GR}$ is the green revenue deduced from the six environmentally sustainable objectives, $\mathcal{TR}$ is the total revenue, $\mathcal{P}$ is the penalty coefficient reflecting the DNSH constraint, $\boldsymbol{\mathcal{S}}$ is the minimum safeguard score and $\boldsymbol{\mathcal{S}}^\star$ is the threshold. The first term is a proxy of the turnover KPI and corresponds to the green revenue share:

$$\boldsymbol{\mathcal{GRS}} = \frac{\mathcal{GR}}{\mathcal{TR}}$$

By construction, we have $0 \leq \boldsymbol{\mathcal{GRS}} \leq 1$. This measure is then impacted by the DNSH coefficient. If the penalty coefficient is equal to zero, the green activities of the issuer do not significantly harm the other objectives and we have $\boldsymbol{\mathcal{GI}} = \boldsymbol{\mathcal{GRS}}$. Otherwise, the green intensity satisfies $0 \leq \boldsymbol{\mathcal{GI}} = \boldsymbol{\mathcal{GRS}} \cdot (1 - \mathcal{P}) \leq \boldsymbol{\mathcal{GRS}}$. Finally, the indicator function $\mathbb{1}\left\{\boldsymbol{\mathcal{S}} \geq \boldsymbol{\mathcal{S}}^\star\right\}$ is a binary all-or-nothing variable. It is equal to one if the firm complies with minimum social safeguards. Otherwise, the green intensity is equal to zero if the firm doesn't pass this materiality test. It follows that an upper bound of the green intensity is the green revenue share since we have $\boldsymbol{\mathcal{GI}} \leq \boldsymbol{\mathcal{GRS}}$.

Box 9.5: EU green taxonomy

The EU taxonomy is described in the Delegated Act on the climate objectives of 4 June 2021. For each activity, three items are provided: the description of the activity, the technical screening criteria, and the DNSH compliance. Let us consider Transport of CO_2 (page 100 of the Delegated Act). We have the following information:

1. Description of the activity
   This concerns the transport of captured CO_2 via all modes; the construction and operation of CO_2 pipelines and retrofit of gas networks where the main purpose is the integration of captured CO_2. The economic activities in this category could be associated with several NACE codes, in particular F42.21 and H49.50. An economic activity in this category is an enabling activity.

2. Technical screening criteria
   This activity has a substantial contribution to climate change mitigation:

   (a) The CO_2 transported from the installation where it is captured to the injection point does not lead to CO_2 leakages above 0.5% of the mass of CO_2 transported.

   (b) The CO_2 is delivered to a permanent CO_2 storage site that meets the criteria for underground geological storage of CO_2 set out in Section 5.12 of this Annex; or to other transport modalities, which lead to permanent CO_2 storage site that meet those criteria.

   (c) Appropriate leak detection systems are applied and a monitoring plan is in place, with the report verified by an independent third party.

   (d) The activity may include the installation of assets that increase the flexibility and improve the management of an existing network.

3. Do no significant harm
   Three out of five categories are concerned: (2) climate change adaptation: the activity complies with the criteria set out in Appendix A to this Annex; (3) sustainable use and protection of water and marine resources: The activity complies with the criteria set out in Appendix B to this Annex. (4) transition to a circular economy: N/A; (5) pollution prevention and control: N/A; (6) protection and restoration of biodiversity and ecosystems: the activity complies with the criteria set out in Appendix D to this Annex.

This example shows that the revenues generated by the transport of CO_2 are not necessarily green, because the technical screening criteria imply that CO_2 leakages must be below 0.5% of the mass of CO_2 transported. We also observe that some criteria are generic while others are specific to an activity. For instance, the life cycle GHG emissions from the generation of electricity (whatever the electricity source) must be lower than 100 gCO_2e per kWh. Some criteria also concern the activity efficiency. For example, the power density of the electricity generation facility must be above 5 Watt per m^2 for the hydropower sector.

Concerning the DNSH compliance criteria, they may for instance imply that the activity does use persistent organic pollutants, ether, mercury, substances that deplete the ozone layer, certain hazardous substances, etc. In a similar way, an environmental impact assessment (EIA) must be conducted for sites located near biodiversity-sensitive areas.

Example 37 *We consider a company in the hydropower sector which has five production sites. Below, we indicate the power density efficiency, the GHG emissions, the DNSH compliance with respect to the biodiversity and the corresponding revenue:*

Site	#1	#2	#3	#4	#5
Efficiency (in Watt per m^2)	3.2	3.5	3.3	5.6	4.2
GHG emissions (in gCO$_2$e per kWh)	35	103	45	12	36
Biodiversity DNSH compliance	✓	✓	✓	✓	
Revenue (in \$ mn)	103	256	89	174	218

The total revenue is equal to:

$$\mathcal{TR} = 103 + 256 + 89 + 174 + 218 = \$840 \text{ mn}$$

We notice that the fourth site does not pass the technical screening, because the power density is above 5 Watt per m^2. The second site does not also comply because it has a GHG emissions greater than 100 gCO$_2$e per kWh. We deduce that the green revenue is equal to:

$$\mathcal{GR} = 103 + 89 + 218 = \$410 \text{ mn}$$

We conclude that the green revenue share is equal to 48.8%. According to the EU green taxonomy, the green intensity is lower because the last site is close to a biodiversity area and has a negative impact. Therefore, we have:

$$\mathcal{GI} = \frac{103 + 89}{840} = 22.9\%$$

Statistics

In Table 9.27, we report the descriptive statistics of green revenue share calculated by Barahhou *et al.* (2022) with the MSCI database. For each category[46], they have computed:

- The frequency $\mathbf{F}(x) = \Pr\{\mathcal{GRS} > x\}$;

- The statistical quantile $\mathbf{Q}(\alpha) = \inf\{x : \Pr\{\mathcal{GRS} \leq x\} \geq \alpha\}$;

- The arithmetic average $n^{-1}\sum_{i=1}^{n}\mathcal{GRS}_i$ and the weighted mean $\mathcal{GRS}(b) = \sum_{i=1}^{n} b_i \mathcal{GRS}_i$ where b_i is the weight of issuer i in the MSCI ACWI IMI benchmark.

For instance, 9.82% of issuers have a green revenue share that concerns alternative energy. This figure becomes less than 1% if we consider a green revenue share greater than 50%. The average value is equal to 1.36% whereas the weighted value is equal to 0.77%. This indicates a small cap bias. For energy efficiency, the average is lower than the weighted mean, implying a bias towards big companies. If we consider the total green revenue share, 27.85% have a positive figure and only 3.17% have a figure greater than 50%. The 90% quintile is equal to 11.82%. Therefore, we notice a high positive skewness for the distribution. The green revenue share is then located in a small number of companies.

Remark 87 *Barahhou et al. (2022) estimated that the green revenue share of the MSCI World index and the Bloomberg Global Investment Grade Corporate Bond index are respectively equal to 5.24% and 3.49% in June 2022. This is not a high figure, because the economy is today far to be green. These results are confirmed by Alessi and Battiston (2022), who estimated "a greenness of about 2.8% for EU financial markets" according to the existing EU green taxonomy[47].*

---

[46]The MSCI taxonomy uses 6 categories: (1) alternative energy, (2) energy efficiency, (3) green building, (4) pollution prevention and control, (5) sustainable agriculture and (6) sustainable water.

[47]This concerns the first two categories, which are the most important.

Table 9.27: Statistics in % of green revenue share (MSCI ACWI IMI, June 2022)

Category	Frequency $\mathbf{F}(x)$				Quantile $\mathbf{Q}(\alpha)$				Mean	
	0	25%	50%	75%	75%	90%	95%	Max	Avg	Wgt
(1)	9.82	1.47	0.96	0.75	0.00	0.00	2.85	100.00	1.36	0.77
(2)	14.10	1.45	0.65	0.31	0.00	1.25	6.12	100.00	1.39	3.50
(3)	4.84	1.68	1.02	0.31	0.00	0.00	0.00	100.00	1.16	0.51
(4)	4.79	0.30	0.10	0.06	0.00	0.00	0.00	99.69	0.32	0.22
(5)	1.00	0.39	0.20	0.09	0.00	0.00	0.00	98.47	0.26	0.10
(6)	4.75	0.28	0.11	0.05	0.00	0.00	0.00	99.98	0.29	0.14
Total	27.85	5.82	3.17	1.68	0.42	11.82	30.36	100.00	4.78	5.24

Source: MSCI (2022) & Barahhou *et al.* (2022).

9.4.3   Green capex

Curabitur tellus magna, porttitor a, commodo a, commodo in, tortor. Donec interdum. Praesent scelerisque. Maecenas posuere sodales odio. Vivamus metus lacus, varius quis, imperdiet quis, rhoncus a, turpis. Etiam ligula arcu, elementum a, venenatis quis, sollicitudin sed, metus. Donec nunc pede, tincidunt in, venenatis vitae, faucibus vel, nibh. Pellentesque wisi. Nullam malesuada. Morbi ut tellus ut pede tincidunt porta. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam congue neque id dolor.

9.4.4   Green-to-brown ratio

Donec et nisl at wisi luctus bibendum. Nam interdum tellus ac libero. Sed sem justo, laoreet vitae, fringilla at, adipiscing ut, nibh. Maecenas non sem quis tortor eleifend fermentum. Etiam id tortor ac mauris porta vulputate. Integer porta neque vitae massa. Maecenas tempus libero a libero posuere dictum. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Aenean quis mauris sed elit commodo placerat. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Vivamus rhoncus tincidunt libero. Etiam elementum pretium justo. Vivamus est. Morbi a tellus eget pede tristique commodo. Nulla nisl. Vestibulum sed nisl eu sapien cursus rutrum.

9.4.5   Other metrics

Nulla non mauris vitae wisi posuere convallis. Sed eu nulla nec eros scelerisque pharetra. Nullam varius. Etiam dignissim elementum metus. Vestibulum faucibus, metus sit amet mattis rhoncus, sapien dui laoreet odio, nec ultricies nibh augue a enim. Fusce in ligula. Quisque at magna et nulla commodo consequat. Proin accumsan imperdiet sem. Nunc porta. Donec feugiat mi at justo. Phasellus facilisis ipsum quis ante. In ac elit eget ipsum pharetra faucibus. Maecenas viverra nulla in massa.

9.5   Exercises

9.5.1   Stochastic modeling of global warming potentials

9.5.2   Calculation of global temperature potentials

Chapter 10

Transition Risk Modeling

10.1 Carbon tax

10.1.1 Mathematics of carbon tax

10.1.2 Abatement cost

10.2 Stranded assets

10.3   Decarbonization pathway

10.3.1   Global analysis

10.3.2   Sector analysis

Power and electricity

Hydrogen

Buildings

Mobility and transport

Materials

Industry

Water management

Waste management and circular economy

10.4   Transition risk measures

10.4.1   Temperature rating modeling

10.5   Exercises

Chapter 11

Climate Portfolio Construction

With the 2015 Paris Agreement, the development of ESG investing, and the emergence of net-zero investment policies, climate risk is undoubtedly the most important issue and challenge for asset owners and managers today and in the coming years. Building portfolios to manage climate risk began in 2014, when asset owners AP4 and FRR worked with asset manager Amundi and index provider MSCI to define an investment strategy to help hedge climate risk (Andersson *et al.*, 2016). Together, they defined the concept of a low-carbon portfolio. The 2015-2020 period corresponds to the growth of ESG investing and the adoption of sustainable finance by many asset owners and managers. During this period, climate investing is part of ESG investing, and climate portfolio construction remains relatively marginal, essentially involving passive management. On the contrary, ESG investing has seen a major development with the adoption of ESG scores in active management. Since 2020, we have seen a new trend. The line between ESG investing and climate investing is becoming increasingly blurred, and the two issues are now separate. One of the reasons for this is the emergence of net-zero investment policies, which have profoundly changed the investment decisions of asset owners. This separation has accelerated with COP26. The proliferation of net zero alliances (GFANZ, NZAOA, NZAM, NZBA, etc.), the commitments made by financial institutions (asset managers, banks, pension funds, insurance companies, etc.), and the push for regulation[1] are all contributing to the shift from ESG investing to climate investing.

This chapter is dedicated to portfolio construction when we integrate climate risk measures. It is therefore closely related to Chapter 9, as we use carbon footprint and green footprint metrics. It is also related to Chapter 10 on transition risk, as the goal of climate investing is to reduce the transition risk of investment portfolios. Integrating physical risk is more complicated today, because we do not have the right metrics to assess physical risk at the corporate or security level. Finally, it is related to Chapter 2, which is dedicated to the impact of ESG investing on asset prices and portfolio returns, because we use the same tools and methodologies. We will therefore make extensive use of portfolio optimization. A comprehensive review of portfolio optimization is presented in the first section. We distinguish between allocations to equity and fixed income portfolios because they require two different approaches. The reason is that we generally measure equity risk in terms of volatility risk, while bond risk is multi-dimensional and must at least integrate duration and credit risk. The second section is a guide to building a low-carbon portfolio. While there are several approaches, the choice of Scope emissions is certainly the most important decision and has a major impact on asset allocation. Finally, the last section focuses on net-zero investing and lists many challenges to defining an investment portfolio that is aligned with a net-zero emissions scenario.

---

[1]Examples include the Net-Zero Industry Act of the European Commission, the work of the NGFS on climate scenarios, or the climate stress tests organized by the ECB.

11.1   Portfolio optimization in practice

Before studying portfolio allocation in the context of climate risk, we first begin to remind some basics about portfolio optimization. As mentioned by Perrin and Roncalli (2020), the success of mean-variance optimization is due to the appealing properties of the quadratic utility function, and it is easy to solve numerically quadratic programming problems[2]. This is why most of the portfolio allocation problems that we will encounter in this chapter will be cast into a QP problem, whose standard formulation is[3]:

$$
\begin{aligned}
x^\star \quad &= \quad \arg\min \frac{1}{2} x^\top Q x - x^\top R \\
&\text{s.t.} \quad \begin{cases} Ax = B \\ Cx \leq D \\ x^- \leq x \leq x^+ \end{cases}
\end{aligned}
\tag{11.1}
$$

where x is a $n \times 1$ vector, Q is a $n \times n$ matrix, R is a $n \times 1$ vector, A is a $n_A \times n$ matrix, B is a $n_A \times 1$ vector, C is a $n_C \times n$ matrix, D is a $n_C \times 1$ vector, and x^- and x^+ are two $n \times 1$ vectors. If $n_A = 0$, there is no equality constraints. Similarly, there is no inequality constraints if $n_C = 0$. If there is no lower bounds or/and upper bounds, this implies that $x^- = -\infty \cdot \mathbf{1}_n$ and $x^+ = \infty \cdot \mathbf{1}_n$. From a numerical viewpoint, QP solvers generally replace these bounds by $x^- = -c \cdot \mathbf{1}_n$ and $x^+ = c \cdot \mathbf{1}_n$ where c is a large floating-point number[4] (e.g., $c = 10^{200}$).

11.1.1   Equity portfolios

Basic optimization problems

We consider a universe of n assets. We note w the vector of portfolio weights. Let μ and Σ be the vector of expected returns and the covariance matrix of asset returns. The long-only mean-variance optimization problem is given by:

$$
\begin{aligned}
w^\star \quad &= \quad \arg\min \frac{1}{2} w^\top \Sigma w - \gamma w^\top \mu \\
&\text{s.t.} \quad \begin{cases} \mathbf{1}_n^\top w = 1 \\ \mathbf{0}_n \leq w \leq \mathbf{1}_n \end{cases}
\end{aligned}
$$

where γ is the risk-tolerance coefficient, the equality constraint is the budget constraint ($\sum_{i=1}^n w_i = 1$) and the bounds correspond to the no short-selling restriction ($w_i \geq 0$). We recognize a QP problem where $Q = \Sigma$, $R = \gamma\mu$, $A = \mathbf{1}_n^\top$, $B = 1$, $w^- = \mathbf{0}_n$ and $w^+ = \mathbf{1}$. In this problem, we have one equality constraint ($n_A = 1$) and zero inequality constraint[5] ($n_C = 1$).

In many problems, we will have to manage the portfolio with respect to a benchmark. In this case, the objective function depends on the tracking error risk variance $\sigma^2(w \mid b) = (w - b)^\top \Sigma (w - b)$ where b is the vector of benchmark weights. In Section 3.1.1 on page 142, we have seen that the tracking error optimization problem is defined as:

$$
\begin{aligned}
w^\star \quad &= \quad \arg\min \frac{1}{2} w^\top \Sigma w - w^\top (\gamma\mu + \Sigma b) \\
&\text{s.t.} \quad \begin{cases} \mathbf{1}_n^\top w = 1 \\ \mathbf{0}_n \leq w \leq \mathbf{1}_n \end{cases}
\end{aligned}
$$

---

[2] See Appendix A.1.2 on page 652.

[3] The objective function is a quadratic form and is noted $\mathcal{QF}(x; Q, R, \mathbf{0}_n)$.

[4] The largest finite floating-point number in IEEE[®] double precision is equal to $(2 - 2^{-52}) \times 2^{1023}$.

[5] We do not take into account the bounds.

This is exactly the same QP problem as previously except that $R = \gamma\mu + \Sigma b$. If the objective of the portfolio manager is to minimize the tracking error risk, we obtain $R = \Sigma b$.

Specification of the constraints

We can extend the previous framework by considering more constraints. For instance, we consider a sector weight constraint:

$$s_j^- \leq \sum_{i \in \boldsymbol{S}ector_j} w_i \leq s_j^+$$

We notice that:

$$\sum_{i \in \boldsymbol{S}ector_j} w_i = \boldsymbol{s}_j^\top w$$

where $\boldsymbol{s}_j$ is the $n \times 1$ sector-mapping vector whose elements are $\boldsymbol{s}_{i,j} = \mathbb{1}\{i \in \boldsymbol{S}ector_j\}$. We deduce that the sector constraint can be written as:

$$s_j^- \leq \sum_{i \in \boldsymbol{S}ector_j} w_i \leq s_j^+ \Leftrightarrow \left\{ \begin{array}{l} s_j^- \leq \boldsymbol{s}_j^\top w \\ \boldsymbol{s}_j^\top w \leq s_j^+ \end{array} \right. \Leftrightarrow \left\{ \begin{array}{l} -\boldsymbol{s}_j^\top w \leq -s_j^- \\ \boldsymbol{s}_j^\top w \leq s_j^+ \end{array} \right.$$

It follows that the inequality constraint $Cw \leq D$ is defined by the following system:

$$\underbrace{\begin{pmatrix} -\boldsymbol{s}_j^\top \\ \boldsymbol{s}_j^\top \end{pmatrix}}_{C} w \leq \underbrace{\begin{pmatrix} -s_j^- \\ s_j^+ \end{pmatrix}}_{D}$$

In this case, C is a $2 \times n$ matrix and D is a 2×1 vector. The previous analysis can be extended when there are many sectors.

We denote by $\boldsymbol{S}$ a vector of scores (e.g., ESG scores) and we would like to impose that the (linear) score of the portfolio is greater than a threshold $\boldsymbol{S}^\star$:

$$\sum_{i=1}^n w_i \boldsymbol{S}_i \geq \boldsymbol{S}^\star$$

The QP form of this constraint is:

$$-\boldsymbol{S}^\top w \leq -\boldsymbol{S}^\star$$

Let us now assume that we would like to apply this constraint to a sector. In this case, we have:

$$\begin{aligned} \sum_{i \in \boldsymbol{S}ector_j} w_i \boldsymbol{S}_i \geq \boldsymbol{S}_j^\star \quad &\Leftrightarrow \quad \sum_{i=1}^n \mathbb{1}\{i \in \boldsymbol{S}ector_j\} w_i \boldsymbol{S}_i \geq \boldsymbol{S}_j^\star \\ &\Leftrightarrow \quad \sum_{i=1}^n \boldsymbol{s}_{i,j} w_i \boldsymbol{S}_i \geq \boldsymbol{S}_j^\star \\ &\Leftrightarrow \quad \sum_{i=1}^n w_i (\boldsymbol{s}_{i,j} \boldsymbol{S}_i) \geq \boldsymbol{S}_j^\star \\ &\Leftrightarrow \quad (\boldsymbol{s}_j \circ \boldsymbol{S})^\top w \geq \boldsymbol{S}_j^\star \end{aligned}$$

where $a \circ b$ is the Hadamard product: $(a \circ b)_i = a_i b_i$. The QP form of the sector-specific score constraint is defined by $C = -(\boldsymbol{s}_j \circ \boldsymbol{S})^\top$ and $D = -\boldsymbol{S}_j^\star$.

Example 38 *We consider a capitalization-weighted equity index, which is composed of 8 stocks. The weights are equal to 23%, 19%, 17%, 13%, 9%, 8%, 6% and 5%. We assume that the stock volatilities are equal to 22%, 20%, 25%, 18%, 35%, 23%, 13% and 29%. The correlation matrix is given by:*

$$\mathbb{C} = \begin{pmatrix} 100\% \\ 80\% & 100\% \\ 70\% & 75\% & 100\% \\ 60\% & 65\% & 80\% & 100\% \\ 70\% & 50\% & 70\% & 85\% & 100\% \\ 50\% & 60\% & 70\% & 80\% & 60\% & 100\% \\ 70\% & 50\% & 70\% & 75\% & 80\% & 50\% & 100\% \\ 60\% & 65\% & 70\% & 75\% & 65\% & 70\% & 80\% & 100\% \end{pmatrix}$$

The ESG score, carbon intensity and sector of the eight stocks are the following:

Stock	#1	#2	#3	#4	#5	#6	#7	#8
$\mathcal{S}$	−1.20	0.80	2.75	1.60	−2.75	−1.30	0.90	−1.70
$\mathcal{CI}$	125	75	254	822	109	17	341	741
$\mathcal{S}ector$	1	1	2	2	1	2	1	2

The objective function is minimizing the tracking error risk. We deduce that the equivalent QP problem is:

$$w^\star \quad = \quad \arg\min \frac{1}{2} w^\top Q w - w^\top R$$
$$\text{s.t.} \quad \begin{cases} Aw = B \\ Cw \leq D \\ w^- \leq w \leq w^+ \end{cases}$$

where[6]:

$$Q = \Sigma = \begin{pmatrix} 484.00 & 352.00 & 385.00 & 237.60 & 539.00 & 253.00 & 200.20 & 382.80 \\ 352.00 & 400.00 & 375.00 & 234.00 & 350.00 & 276.00 & 130.00 & 377.00 \\ 385.00 & 375.00 & 625.00 & 360.00 & 612.50 & 402.50 & 227.50 & 507.50 \\ 237.60 & 234.00 & 360.00 & 324.00 & 535.50 & 331.20 & 175.50 & 391.50 \\ 539.00 & 350.00 & 612.50 & 535.50 & 1225.00 & 483.00 & 364.00 & 659.75 \\ 253.00 & 276.00 & 402.50 & 331.20 & 483.00 & 529.00 & 149.50 & 466.90 \\ 200.20 & 130.00 & 227.50 & 175.50 & 364.00 & 149.50 & 169.00 & 301.60 \\ 382.80 & 377.00 & 507.50 & 391.50 & 659.75 & 466.90 & 301.60 & 841.00 \end{pmatrix} \times 10^{-4}$$

and:

$$R = \Sigma b = \begin{pmatrix} 3.74 \\ 3.31 \\ 4.39 \\ 3.07 \\ 5.68 \\ 3.40 \\ 2.02 \\ 4.54 \end{pmatrix} \times 10^{-2}$$

---

[6]We have $\Sigma_{i,j} = \mathbb{C}_{i,j} \sigma_i \sigma_j$.

We assume that the portfolio is long-only. It follows that $w^- = \mathbf{0}_8$ and $w^+ = \mathbf{1}_8$. To satisfy the budget constraint $\sum_{i=1}^{8} w_i = 1$, we have a first linear equation $A_0 w = B_0$ where $A_0 = \mathbf{1}_8^\top$ and $B_0 = 1$. We consider three type of constraints:

- We impose a relative reduction of the benchmark carbon intensity:

$$\boldsymbol{CI}(w) \leq (1 - \boldsymbol{\mathcal{R}})\,\boldsymbol{CI}(b)$$

where $\boldsymbol{\mathcal{R}}$ is the reduction rate. Since $\boldsymbol{CI}(w) = \boldsymbol{CI}^\top w$, we deduce the following inequality constraint $C_1 w \leq D_1$ where $C_1 = \boldsymbol{CI}^\top$ and $D_1 = (1 - \boldsymbol{\mathcal{R}})\,\boldsymbol{CI}(b)$.

- We impose an absolute increase of the benchmark ESG score:

$$\boldsymbol{S}(w) \geq \boldsymbol{S}(b) + \Delta\boldsymbol{S}^\star$$

Since $\boldsymbol{S}(w) = \boldsymbol{S}^\top w$, we deduce the following inequality constraint $C_2 w \leq D_2$ where $C_2 = -\boldsymbol{S}^\top$ and $D_2 = -\left(\boldsymbol{S}(b) + \Delta\boldsymbol{S}^\star\right)$.

- We impose the sector neutrality of the portfolio meaning that:

$$\sum_{i \in \boldsymbol{Sector}_j} w_i = \sum_{i \in \boldsymbol{Sector}_j} b_i$$

We note:

$$A_1 = \boldsymbol{s}_1^\top = \begin{pmatrix} 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 \end{pmatrix}$$

and:

$$A_2 = \boldsymbol{s}_2^\top = \begin{pmatrix} 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}$$

We compute $B_1 = \boldsymbol{s}_1^\top b = \sum_{i \in \boldsymbol{Sector}_1} b_i$ and $B_2 = \boldsymbol{s}_2^\top b = \sum_{i \in \boldsymbol{Sector}_2} b_i$. The sector neutrality constraint can be written as:

$$\begin{pmatrix} A_1 \\ A_2 \end{pmatrix} w = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}$$

Let us now combine the different constraints. For that, we use the block matrix notation, which is particularly convenient when manipulating nested QP problems. The set #1 of constraint corresponds to the reduction of the carbon intensity, the set #2 corresponds to the ESG score improvement, the set #3 combines the two constraints and we add the sector neutrality in the set #4 of constraints.

Set of constraints	Carbon intensity	ESG score	Sector neutrality	A	B	C	D
#1	✓			A_0	B_0	C_1	D_1
#2		✓		A_0	B_0	C_2	D_2
#3	✓	✓		A_0	B_0	$\begin{bmatrix} C_1 \\ C_2 \end{bmatrix}$	$\begin{bmatrix} D_1 \\ D_2 \end{bmatrix}$
#4	✓	✓	✓	$\begin{bmatrix} A_0 \\ A_1 \\ A_2 \end{bmatrix}$	$\begin{bmatrix} B_0 \\ B_1 \\ B_2 \end{bmatrix}$	$\begin{bmatrix} C_1 \\ C_2 \end{bmatrix}$	$\begin{bmatrix} D_1 \\ D_2 \end{bmatrix}$

We can now solve the QP problem with the right specification of matrices A, B, C and D for each set of constraints. Results are reported below for the following parameters[7]: $\mathcal{R} = 30\%$ and $\Delta\mathcal{S}^{\star} = 0.50$. We indicate the optimal weights $w^{\star}$, the tracking error volatility $\sigma\left(w^{\star} \mid b\right)$, the carbon intensity $\mathcal{CI}\left(w^{\star}\right)$, the effective reduction rate[8] $\mathcal{R}\left(w^{\star} \mid b\right)$, the ESG score $\mathcal{S}\left(w^{\star}\right)$, the ESG score variation $\mathcal{S}\left(w^{\star}\right) - \mathcal{S}\left(b\right)$, the weights[9] $w^{\star}\left(\mathcal{S}ector_1\right)$ and $w^{\star}\left(\mathcal{S}ector_2\right)$ of sectors 1 and 2.

		Benchmark	Set #1	Set #2	Set #3	Set #4
Weights (in %)	$w_1^{\star}$	23.00	18.17	25.03	8.64	12.04
	$w_2^{\star}$	19.00	24.25	14.25	29.27	23.76
	$w_3^{\star}$	17.00	16.92	21.95	26.80	30.55
	$w_4^{\star}$	13.00	2.70	27.30	1.48	2.25
	$w_5^{\star}$	9.00	12.31	3.72	10.63	8.51
	$w_6^{\star}$	8.00	11.23	1.34	6.30	10.20
	$w_7^{\star}$	6.00	11.28	1.68	16.87	12.69
	$w_8^{\star}$	5.00	3.15	4.74	0.00	0.00
Statistics	$\sigma\left(w^{\star} \mid b\right)$ (in %)	0.00	0.50	1.18	1.90	2.12
	$\mathcal{CI}\left(w^{\star}\right)$	261.72	183.20	367.25	183.20	183.20
	$\mathcal{R}\left(w^{\star} \mid b\right)$ (in %)		30.00	-40.32	30.00	30.00
	$\mathcal{S}\left(w^{\star}\right)$	0.17	0.05	0.67	0.67	0.67
	$\mathcal{S}\left(w^{\star}\right) - \mathcal{S}\left(b\right)$		-0.12	0.50	0.50	0.50
	$w^{\star}\left(\mathcal{S}ector_1\right)$ (in %)	57.00	66.00	44.67	65.41	57.00
	$w^{\star}\left(\mathcal{S}ector_2\right)$ (in %)	43.00	34.00	55.33	34.59	43.00

If we only reduce the carbon intensity, the tracking error cost is 50 bps. Nevertheless, the ESG score of the optimized portfolio is below the ESG score of the benchmark (0.05 vs. 0.17). If we improve the ESG score by 0.50, the tracking error cost is 118 bps. This optimal portfolio has a carbon footprint which is higher than this of the benchmark (367.25 vs. 261.72). If we combine the two constraints, the tracking error cost is close to 2%. We observe a reallocation between the sectors. For instance, sector 1 represents 65.41% of the portfolio while its weight is equal to 57% in the benchmark. Finally, the tracking error volatility of the optimal portfolio is equal to 2.12% when we add the sector neutrality constraint.

Dealing with constraints on relative weights

Let us assume that we would like to reduce the carbon footprint at the sector level. In this case, we denote by $\mathcal{CI}\left(w; \mathcal{S}ector_j\right)$ the carbon intensity of the jth sector within the portfolio w:

$$\mathcal{CI}\left(w; \mathcal{S}ector_j\right) = \sum_{i \in \mathcal{S}ector_j} \tilde{w}_i \mathcal{CI}_i$$

where $\tilde{w}_i$ is the normalized weight in the sector bucket:

$$\tilde{w}_i = \frac{w_i}{\sum_{k \in \mathcal{S}ector_j} w_k}$$

---

[7]We have $B_1 = 0.57$, $B_2 = 0.43$, $D_1 = 183.2040$ and $D_2 = -0.6690$.

[8]We have:

$$\mathcal{R}\left(w \mid b\right) = \frac{\mathcal{CI}\left(b\right) - \mathcal{CI}\left(w\right)}{\mathcal{CI}\left(b\right)}$$

[9]We have:

$$w\left(\mathcal{S}ector_j\right) = \sum_{i \in \mathcal{S}ector_j} w_i = \boldsymbol{s}_j^{\top} w$$

Another expression of $\boldsymbol{CI}(w; \boldsymbol{S}ector_j)$ is:

$$\boldsymbol{CI}(w; \boldsymbol{S}ector_j) = \frac{\sum_{i \in \boldsymbol{S}ector_j} w_i \boldsymbol{CI}_i}{\sum_{i \in \boldsymbol{S}ector_j} w_i} = \frac{(\boldsymbol{s}_j \circ \boldsymbol{CI})^\top w}{\boldsymbol{s}_j^\top w}$$

If we consider the constraint $\boldsymbol{CI}(w; \boldsymbol{S}ector_j) \leq \boldsymbol{CI}_j^\star$, we obtain:

$$
\begin{aligned}
(*) \quad &\Leftrightarrow \quad \boldsymbol{CI}(w; \boldsymbol{S}ector_j) \leq \boldsymbol{CI}_j^\star \\
&\Leftrightarrow \quad (\boldsymbol{s}_j \circ \boldsymbol{CI})^\top w \leq \boldsymbol{CI}_j^\star \left(\boldsymbol{s}_j^\top w \right) \\
&\Leftrightarrow \quad \left((\boldsymbol{s}_j \circ \boldsymbol{CI}) - \boldsymbol{CI}_j^\star \boldsymbol{s}_j \right)^\top w \leq 0 \\
&\Leftrightarrow \quad \left(\boldsymbol{s}_j \circ (\boldsymbol{CI} - \boldsymbol{CI}_j^\star) \right)^\top w \leq 0
\end{aligned}
\tag{11.2}
$$

The QP form is then $C = \left(\boldsymbol{s}_j \circ (\boldsymbol{CI} - \boldsymbol{CI}_j^\star) \right)^\top$ and $D = 0$. We remark that we obtain a classical reduction constraint where the vector of carbon intensities is replaced by the vector of deviations $\boldsymbol{CI} - \boldsymbol{CI}_j^\star$.

Remark 88 *We have defined an absolute threshold $\boldsymbol{CI}_j^\star$, but we can implement a relative threshold by setting $\boldsymbol{CI}_j^\star = (1 - \boldsymbol{R}_j) \boldsymbol{CI}(b; \boldsymbol{S}ector_j)$.*

Let us consider again Example 38 on page 550. We would like to reduce the carbon footprint of the benchmark by 30% and impose the sector neutrality. In this case, the linear equality constraint $Aw = B$ is defined by:

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}$$

and:

$$B = \begin{pmatrix} 100\% \\ 57\% \\ 43\% \end{pmatrix}$$

For the linear inequality constraint $Cw \leq D$, we obtain:

$$C = \begin{pmatrix} 125 & 75 & 254 & 822 & 109 & 17 & 341 & 741 \end{pmatrix}$$

and:

$$D = 183.2040$$

The optimal solution is $w^\star = (21.54\%, 18.50\%, 21.15\%, 3.31\%, 10.02\%, 15.26\%, 6.94\%, 3.27\%)$. The tracking error volatility $\sigma(w^\star \mid b)$ is equal to 112 bps. We verify that the carbon intensity of $w^\star$ is equal to 183.20, which corresponds to a reduction rate of 30% compared to the carbon intensity of the benchmark — $\boldsymbol{CI}(b) = 261.72$. Nevertheless, we observe that the sector breakdown of the carbon footprint is the following:

$$\begin{cases} \boldsymbol{CI}(w^\star; \boldsymbol{S}ector_1) = 132.25 \\ \boldsymbol{CI}(w^\star; \boldsymbol{S}ector_2) = 250.74 \end{cases} \quad \text{versus} \quad \begin{cases} \boldsymbol{CI}(b; \boldsymbol{S}ector_1) = 128.54 \\ \boldsymbol{CI}(b; \boldsymbol{S}ector_2) = 438.26 \end{cases}$$

It follows that the 30% global reduction is explained by an increase of 2.89% of the carbon footprint for the first sector and a decrease of 42.79% of the carbon footprint for the second sector. We deduce that the decarbonization of this portfolio is only supported by the second sector. This is

why we impose that the first sector must reduce its carbon footprint by $\mathcal{R}_1 = 20\%$. In this case, the inequality system becomes:

$$
\begin{aligned}
C &= \begin{pmatrix} \mathcal{CI}^\top \\ \left(\boldsymbol{s}_1 \circ \left(\mathcal{CI} - (1 - \mathcal{R}_1)\,\mathcal{CI}\,(b; \boldsymbol{S}ector_1)\right)\right)^\top \end{pmatrix} \\
&= \begin{pmatrix} 125 & 75 & 254 & 822 & 109 & 17 & 341 & 741 \\ 22.1649 & -27.8351 & 0 & 0 & 6.1649 & 0 & 238.1649 & 0 \end{pmatrix}
\end{aligned}
$$

and:

$$
D = \begin{pmatrix} 183.2040 \\ 0 \end{pmatrix}
$$

Solving this new QP problem gives the following optimal portfolio:

$$
w^\star = (22.70\%, 22.67\%, 19.23\%, 5.67\%, 11.39\%, 14.50\%, 0.24\%, 3.61\%)
$$

The tracking error volatility $\sigma\,(w^\star \mid b)$ becomes 144 bps. We verify that the carbon intensity of $w^\star$ is equal to 183.20. Moreover, we obtain $\mathcal{CI}\,(w^\star; \boldsymbol{S}ector_1) = 102.84$ and $\mathcal{CI}\,(w^\star; \boldsymbol{S}ector_2) = 289.74$, i.e. a reduction of 20% and 33.89% respectively.

11.1.2 Bond portfolios

Risk measure of a bond portfolio

There is a large consensus to use the historical volatility as a risk measure for an equity portfolio, even if it is not perfect and has many drawbacks. For instance, skewness and kurtosis risks are not taken into account, correlations are time-varying, we generally observe jumps in the time series of stock returns, volatility is heteroscedastic, etc. In the case of fixed-income portfolios, the historical volatility is not a relevant risk measure, because the bond volatility depends on duration, credit spread, liquidity, etc. Let us consider a zero-coupon bond, whose price and maturity date are $B\,(t, T)$ and T. Following Roncalli (2013, page 225), we have:

$$
B_t\,(t, T) = e^{-(r(t)+s(t))(T-t)+L(t)}
$$

where $r\,(t)$, $s\,(t)$ and $L\,(t)$ are the interest rate, the credit spread and the liquidity premium. We deduce that:

$$
\begin{aligned}
\mathrm{d}\ln B\,(t, T) &= -(T-t)\,\mathrm{d}r\,(t) - (T-t)\,\mathrm{d}s\,(t) + \mathrm{d}L\,(t) \\
&= -D\,\mathrm{d}r\,(t) - (D\,s\,(t))\frac{\mathrm{d}s\,(t)}{s\,(t)} + \mathrm{d}L\,(t) \\
&= -D\,\mathrm{d}r\,(t) - \mathrm{DTS}\,(t)\frac{\mathrm{d}s\,(t)}{s\,(t)} + \mathrm{d}L\,(t)
\end{aligned}
$$

where $D = T - t$ is the remaining maturity (or duration) and $\mathrm{DTS}\,(t)$ is the duration-times-spread factor. If we assume that $r\,(t)$, $s\,(t)$ and $L\,(t)$ are independent, the risk of the defaultable bond is equal to:

$$
\sigma^2\,(\mathrm{d}\ln B\,(t, T)) = D^2\sigma^2\,(\mathrm{d}r\,(t)) + \mathrm{DTS}\,(t)^2\,\sigma^2\left(\frac{\mathrm{d}s\,(t)}{s\,(t)}\right) + \sigma^2\,(\mathrm{d}L\,(t))
$$

It can then be decomposed into three components: an interest-rate risk component, a credit risk component and a liquidity risk component. In the case where the three volatility risks (interest rate, credit spread and liquidity) are constant, we obtain:

$$
\sigma^2\,(\mathrm{d}\ln B\,(t, T)) = D^2\sigma_r^2 + \mathrm{DTS}\,(t)^2\,\sigma_s^2 + \sigma_L^2
$$

We see that the bond risk depends not only on interest rate, credit and liquidity volatilities but it varies with respect to the duration and the level of the credit spread. Therefore, we confirm that the historical volatility of a bond price is not a relevant risk measure.

Remark 89 *In the sequel, we do not consider the liquidity risk, since it is generally measured by proprietary liquidity scores (Ben Slimane and De Jong, 2017).*

Basic optimization problems

In practice, the duration risk is measured using the modified duration (MD) metric. We have:

$$\text{MD}\left(w\right) = \sum_{i=1}^{n} w_i \,\text{MD}_i$$

The portfolio DTS is equal to:

$$\text{DTS}\left(w\right) = \sum_{i=1}^{n} w_i \,\text{DTS}_i$$

Generally, the construction of a bond portfolio uses the clustering approach. The bond universe is divided into clusters based on the currency, the sector, the credit quality and the maturity band. For instance, the cluster (EUR, Financials, AAA to A−, 1Y-3Y) corresponds to euro-denominated bonds from the financials sector, whose rating is greater than A− and the duration is between one and three years. A cluster is then a generalization of the sector concept. Nevertheless, we will continue to use the term sector instead of the term cluster in the sequel when we will build bond portfolios. Let $\text{MD}_j\left(w\right)$ and $\text{DTS}_j\left(w\right)$ be the MD and DTS contributions[10] of the jth sector. We have $\text{MD}_j\left(w\right) = \sum_{i \in \boldsymbol{S}ector_j} w_i \,\text{MD}_i$ and $\text{DTS}_j\left(w\right) = \sum_{i \in \boldsymbol{S}ector_j} w_i \,\text{DTS}_i$. Bond portfolio optimization consists in maximizing the expected carry of the portfolio under a set of numerous constraints:

$$\Omega = \left\{ j = 1, \ldots, n_{\boldsymbol{S}ector} : \text{MD}_j\left(w\right) \approx \text{MD}_j^{\star}, \text{DTS}_j\left(w\right) \approx \text{DTS}_j^{\star} \right\}$$

Therefore, the objective function of the optimization problem has the following form:

$$w^{\star} = \arg\min \frac{\varphi_{\text{MD}}}{2} \sum_{j=1}^{n_{\boldsymbol{S}ector}} \left(\text{MD}_j\left(w\right) - \text{MD}_j^{\star}\right)^2 + \frac{\varphi_{\text{DTS}}}{2} \sum_{j=1}^{n_{\boldsymbol{S}ector}} \left(\text{DTS}_j\left(w\right) - \text{DTS}_j^{\star}\right)^2 - \gamma \sum_{i=1}^{n} w_i \mathcal{C}_i \tag{11.3}$$

where $\varphi_{\text{MD}} \geq 0$ and $\varphi_{\text{DTS}} \geq 0$ indicate the relative weight of each risk component, $\mathcal{C}_i$ is the expected carry[11] of bond i and γ is the risk-tolerance coefficient. In order to solve this optimization problem, the goal is to transform Equation (11.3) into a QP problem:

$$
\begin{aligned}
w^{\star} \quad &= \quad \arg\min \mathcal{QF}\left(w; Q, R, c\right) \\
\text{s.t.} \quad &\left\{ \begin{array}{l} \mathbf{1}_n^{\top} w = 1 \\ \mathbf{0}_n \leq w \leq \mathbf{1}_n \end{array} \right.
\end{aligned} \tag{11.4}
$$

where $\mathcal{QF}\left(w; Q, R, c\right)$ is the quadratic form of the objective function.

---

[10]Be careful, they do not measure the MD and DTS of the sector, which are measured with relative weights within the sector.

[11]The expected carry of a bond corresponds to the carry of the bond plus its expected mark-to-market return minus the expected credit loss.

We use the properties of quadratic forms presented in Section A.1.3 on page 654. We have[12]:

$$
\begin{aligned}
\frac{1}{2}\left(\mathrm{MD}_j\left(w\right)-\mathrm{MD}_j^\star\right)^2 &= \frac{1}{2}\left(\sum_{i\in\boldsymbol{S}ector_j} w_i\,\mathrm{MD}_i - \mathrm{MD}_j^\star\right)^2 \\
&= \frac{1}{2}\left(\sum_{i=1}^n \boldsymbol{s}_{i,j} w_i\,\mathrm{MD}_i - \mathrm{MD}_j^\star\right)^2 \\
&= \frac{1}{2}\left(\sum_{i=1}^n \boldsymbol{s}_{i,j}\,\mathrm{MD}_i\,w_i\right)^2 - w^\top\left(\boldsymbol{s}_j\circ\mathrm{MD}\right)\mathrm{MD}_j^\star + \frac{1}{2}\mathrm{MD}_j^{\star 2} \\
&= \mathcal{QF}\left(w;\mathcal{T}\left(\boldsymbol{s}_j\circ\mathrm{MD}\right),\left(\boldsymbol{s}_j\circ\mathrm{MD}\right)\mathrm{MD}_j^\star,\frac{1}{2}\mathrm{MD}_j^{\star 2}\right)
\end{aligned}
$$

where $\mathrm{MD}=\left(\mathrm{MD}_1,\ldots,\mathrm{MD}_n\right)$ is the vector of modified durations. It follows that:

$$
\frac{1}{2}\sum_{j=1}^{n_{\boldsymbol{S}ector}}\left(\mathrm{MD}_j\left(w\right)-\mathrm{MD}_j^\star\right)^2 = \mathcal{QF}\left(w;Q_{\mathrm{MD}},R_{\mathrm{MD}},c_{\mathrm{MD}}\right)
$$

where:

$$
\begin{cases}
Q_{\mathrm{MD}} = \displaystyle\sum_{j=1}^{n_{\boldsymbol{S}ector}}\mathcal{T}\left(\boldsymbol{s}_j\circ\mathrm{MD}\right) \\
R_{\mathrm{MD}} = \displaystyle\sum_{j=1}^{n_{\boldsymbol{S}ector}}\left(\boldsymbol{s}_j\circ\mathrm{MD}\right)\mathrm{MD}_j^\star \\
c_{\mathrm{MD}} = \dfrac{1}{2}\displaystyle\sum_{j=1}^{n_{\boldsymbol{S}ector}}\mathrm{MD}_j^{\star 2}
\end{cases}
$$

In a similar way, we have:

$$
\frac{1}{2}\sum_{j=1}^{n_{\boldsymbol{S}ector}}\left(\mathrm{DTS}_j\left(w\right)-\mathrm{DTS}_j^\star\right)^2 = \mathcal{QF}\left(w;Q_{\mathrm{DTS}},R_{\mathrm{DTS}},c_{\mathrm{DTS}}\right)
$$

where:

$$
\begin{cases}
Q_{\mathrm{DTS}} = \displaystyle\sum_{j=1}^{n_{\boldsymbol{S}ector}}\mathcal{T}\left(\boldsymbol{s}_j\circ\mathrm{DTS}\right) \\
R_{\mathrm{MD}} = \displaystyle\sum_{j=1}^{n_{\boldsymbol{S}ector}}\left(\boldsymbol{s}_j\circ\mathrm{DTS}\right)\mathrm{DTS}_j^\star \\
c_{\mathrm{DTS}} = \dfrac{1}{2}\displaystyle\sum_{j=1}^{n_{\boldsymbol{S}ector}}\mathrm{DTS}_j^{\star 2}
\end{cases}
$$

where $\mathrm{DTS}=\left(\mathrm{DTS}_1,\ldots,\mathrm{DTS}_n\right)$ is the vector of DTS values. Since we have $-\gamma\sum_{i=1}^n w_i\mathcal{C}_i = \mathcal{QF}\left(w;\mathbf{0}_{n,n},\gamma\mathcal{C},0\right)$ where $\mathcal{C}=\left(\mathcal{C}_1,\ldots,\mathcal{C}_n\right)$ is the vector of expected carry values, we conclude that the function to optimize is:

$$
\begin{aligned}
\mathcal{QF}\left(w;Q,R,c\right) = {} & \varphi_{\mathrm{MD}}\mathcal{QF}\left(w;Q_{\mathrm{MD}},R_{\mathrm{MD}},c_{\mathrm{MD}}\right)+\varphi_{\mathrm{DTS}}\mathcal{QF}\left(w;Q_{\mathrm{DTS}},R_{\mathrm{DTS}},c_{\mathrm{DTS}}\right)+ \\
& \mathcal{QF}\left(w;\mathbf{0}_{n,n},\gamma\mathcal{C},0\right)
\end{aligned}
$$

---

[12]We remind that $\mathcal{T}\left(u\right)=uu^\top$.

where:

$$\begin{cases} Q = \varphi_{\mathrm{MD}}Q_{\mathrm{MD}} + \varphi_{\mathrm{DTS}}Q_{\mathrm{DTS}} \\ R = \gamma\mathcal{C} + \varphi_{\mathrm{MD}}R_{\mathrm{MD}} + \varphi_{\mathrm{DTS}}R_{\mathrm{DTS}} \\ c = \varphi_{\mathrm{MD}}c_{\mathrm{MD}} + \varphi_{\mathrm{DTS}}c_{\mathrm{DTS}} \end{cases}$$

We can extend the previous analysis when there is a benchmark. In this case, the threshold values $\mathrm{MD}_j^\star$ and $\mathrm{DTS}_j^\star$ are equal to the MD and DTS contributions of the benchmark: $\mathrm{MD}_j^\star = \sum_{i \in \boldsymbol{Sector}_j} b_i \, \mathrm{MD}_i$ and $\mathrm{DTS}_j^\star = \sum_{i \in \boldsymbol{Sector}_j} b_i \, \mathrm{DTS}_i$. The MD- and DTS-based tracking error variances are then equal to:

$$\mathcal{R}_{\mathrm{MD}}\left(w \mid b\right) = \sigma_{\mathrm{MD}}^2\left(w \mid b\right) = \sum_{j=1}^{n_{\boldsymbol{Sector}}} \left(\sum_{i \in \boldsymbol{Sector}_j} \left(w_i - b_i\right) \mathrm{MD}_i \right)^2$$

and:

$$\mathcal{R}_{\mathrm{DTS}}\left(w \mid b\right) = \sigma_{\mathrm{DTS}}^2\left(w \mid b\right) = \sum_{j=1}^{n_{\boldsymbol{Sector}}} \left(\sum_{i \in \boldsymbol{Sector}_j} \left(w_i - b_i\right) \mathrm{DTS}_i \right)^2$$

Generally, we also consider a third active risk component called the active share risk[13]:

$$\mathcal{R}_{\mathrm{AS}}\left(w \mid b\right) = \sigma_{\mathrm{AS}}^2\left(w \mid b\right) = \sum_{i=1}^{n} \left(w_i - b_i\right)^2$$

The basic optimization problem becomes:

$$\begin{aligned} w^\star \quad &= \quad \arg\min \frac{1}{2}\mathcal{R}\left(w \mid b\right) - \gamma \sum_{i=1}^{n} \left(w_i - b_i\right)\mathcal{C}_i \qquad (11.5) \\ \text{s.t.} \quad &\begin{cases} \mathbf{1}_n^\top w = 1 \\ \mathbf{0}_n \le w \le \mathbf{1}_n \end{cases} \end{aligned}$$

where the synthetic risk measure is equal to:

$$\mathcal{R}\left(w \mid b\right) = \varphi_{\mathrm{AS}}\mathcal{R}_{\mathrm{AS}}\left(w \mid b\right) + \varphi_{\mathrm{MD}}\mathcal{R}_{\mathrm{MD}}\left(w \mid b\right) + \varphi_{\mathrm{DTS}}\mathcal{R}_{\mathrm{DTS}}\left(w \mid b\right)$$

By using the properties $\mathcal{QF}\left(x - y; Q, R, c\right) = \mathcal{QF}\left(x; Q, R + Qy, \frac{1}{2}y^\top Qy + y^\top R + c\right)$, $\mathcal{T}\left(\boldsymbol{s}_j \circ \mathrm{MD}\right)b = \left(\boldsymbol{s}_j \circ \mathrm{MD}\right)\left(\boldsymbol{s}_j \circ \mathrm{MD}\right)^\top b = \left(\boldsymbol{s}_j \circ \mathrm{MD}\right)\mathrm{MD}_j^\star$, $\mathcal{T}\left(\boldsymbol{s}_j \circ \mathrm{DTS}\right)b = \left(\boldsymbol{s}_j \circ \mathrm{DTS}\right)\mathrm{DTS}_j^\star$, $b^\top \mathcal{T}\left(\boldsymbol{s}_j \circ \mathrm{MD}\right)b = \mathrm{MD}_j^{\star 2}$ and $b^\top \mathcal{T}\left(\boldsymbol{s}_j \circ \mathrm{DTS}\right)b = \mathrm{DTS}_j^{\star 2}$, we have:

$$\begin{cases} \dfrac{1}{2}\mathcal{R}_{\mathrm{AS}}\left(w \mid b\right) = \mathcal{QF}\left(w; Q_{\mathrm{AS}}\left(b\right), R_{\mathrm{AS}}\left(b\right), c_{\mathrm{AS}}\left(b\right)\right) \\ \dfrac{1}{2}\mathcal{R}_{\mathrm{MD}}\left(w \mid b\right) = \mathcal{QF}\left(w; Q_{\mathrm{MD}}\left(b\right), R_{\mathrm{MD}}\left(b\right), c_{\mathrm{MD}}\left(b\right)\right) \\ \dfrac{1}{2}\mathcal{R}_{\mathrm{DTS}}\left(w \mid b\right) = \mathcal{QF}\left(w; Q_{\mathrm{DTS}}\left(b\right), R_{\mathrm{DTS}}\left(b\right), c_{\mathrm{DTS}}\left(b\right)\right) \end{cases}$$

---

[13]The traditional definition of the active share is:

$$\mathrm{AS}\left(w \mid b\right) = \frac{1}{2}\sum_{i=1}^{n} \left|w_i - b_i\right|$$

We prefer to use the $\mathcal{L}_2$-norm which is more tractable from a numerical viewpoint.

where $Q_{AS}(b) = I_n$, $R_{AS}(b) = b$, $c_{AS}(b) = \frac{1}{2}b^\top b$, $Q_{MD}(b) = Q_{MD}$, $R_{MD}(b) = Q_{MD}b = R_{MD}$, $c_{MD}(b) = \frac{1}{2}b^\top Q_{MD}b = c_{MD}$, $Q_{DTS}(b) = Q_{DTS}$, $R_{DTS}(b) = Q_{DTS}b = R_{DTS}$, and $c_{DTS}(b) = \frac{1}{2}b^\top Q_{DTS}b = c_{DTS}$. We conclude that Equation (11.5) becomes:

$$
\begin{aligned}
w^\star &= \arg\min \mathcal{QF}(w; Q(b), R(b), c(b)) \tag{11.6}\\
&\text{s.t.} \quad \begin{cases} \mathbf{1}_n^\top w = 1 \\ \mathbf{0}_n \le w \le \mathbf{1}_n \end{cases}
\end{aligned}
$$

where[14]:

$$
\begin{cases}
Q(b) = \varphi_{AS}Q_{AS}(b) + \varphi_{MD}Q_{MD}(b) + \varphi_{DTS}Q_{DTS}(b) \\
R(b) = \gamma\mathcal{C} + \varphi_{AS}R_{AS}(b) + \varphi_{MD}R_{MD}(b) + \varphi_{DTS}R_{DTS}(b) \\
c(b) = \gamma b^\top \mathcal{C} + \varphi_{AS}c_{AS}(b) + \varphi_{MD}c_{MD}(b) + \varphi_{DTS}c_{DTS}(b)
\end{cases}
$$

Remark 90 *We notice that the optimization problem (11.4) is a special case of the optimization problem (11.6) when φ_{AS} and b are set to 0 and $\mathbf{0}_n$.*

Example 39 *We consider an investment universe of five bonds with the following characteristics: the benchmark is $b = (30\%, 25\%, 20\%, 15\%, 10\%)$; the modified durations (in years) are $MD = (3, 7, 6, 12, 2)$; the DTS values (in bps) are $DTS = (100, 250, 70, 400, 150)$; the carry vector (in bps) is $\mathcal{C} = (200, 300, 150, 250, 600)$; the first, third and fourth bonds belong to the first sector while the second and fourth bonds belong to the second sector. We assume that $\varphi_{AS} = 1$, $\varphi_{MD} = 1$, $\varphi_{DTS} = 0.001$ and $\gamma = 20\%$.*

We proceed by steps in order to obtain the final quadratic form. First, we first compute the modified duration and the duration-times-spread factor of the benchmark. We obtain $MD(b) = 5.85$ years and $DTS(b) = 181.50$ bps. The sector contributions are $MD_1(b) = 3.90$ and $DTS_1(b) = 104.00$ for sector $\mathcal{S}ector_1$, and $MD_2(b) = 1.95$ and $DTS_2(b) = 77.50$ for sector $\mathcal{S}ector_2$. The target values are then $MD_1^\star = 3.90$, $MD_2^\star = 1.95$, $DTS_1^\star = 104.00$ and $DTS_2^\star = 77.50$. Second, we compute the matrices Q_{MD}, R_{MD}, c_{MD}, Q_{DTS}, R_{DTS}, and c_{DTS}. Third, using the benchmark composition b, we deduce the matrices $Q_{AS}(b)$, $R_{AS}(b)$, $c_{AS}(b)$, $Q_{MD}(b)$, $R_{MD}(b)$, $c_{MD}(b)$, $Q_{DTS}(b)$, $R_{DTS}(b)$ and $c_{DTS}(b)$. Finally, we calculate the matrices $Q(b)$, $R(b)$ and $c(b)$. All these matrices are reported in Table 11.1 on page 559. If we compute $\mathcal{QF}(w_{ew}; Q(b), R(b), c(b))$ where w_{ew} is the equally-weighted portfolio, we obtain -7.37812. Another way to check that we obtain the good results is to verify that $\mathcal{QF}(b; Q(b), R(b), c(b)) = 0$.

Illustration

Example 40 *We consider an investment universe of nine corporate bonds with the following characteristics:*

Issuer	#1	#2	#3	#4	#5	#6	#7	#8	#9
b_i (in %)	21	19	16	12	11	8	6	4	3
$\mathcal{CI}_i$ (in tCO$_2$e/\$ mn)	111	52	369	157	18	415	17	253	900
MD_i (in years)	3.16	6.48	3.54	9.23	6.40	2.30	8.12	7.96	5.48
DTS_i (in bps)	107	255	75	996	289	45	620	285	125
$\mathcal{S}ector$	1	1	1	2	2	2	3	3	3

We assume that the portfolio weights can not deviate too far from the benchmark and impose the following constraint: $0.25 \times b_i \le w_i \le 4 \times b_i$. We have $\varphi_{AS} = 100$, $\varphi_{MD} = 25$ and $\varphi_{DTS} = 0.001$.

Table 11.1: Quadratic form (Example 39)

$Q_{AS}(b) = Q_{AS}$					R_{AS}	$R_{AS}(b)$	c_{AS}	$c_{AS}(b)$
1	0	0	0	0	0.00	0.30		
0	1	0	0	0	0.00	0.25		
0	0	1	0	0	0.00	0.20	0.0000	0.1125
0	0	0	1	0	0.00	0.15		
0	0	0	0	1	0.00	0.10		
$Q_{MD}(b) = Q_{MD}$					R_{MD}	$R_{MD}(b)$	c_{MD}	$c_{MD}(b)$
9	0	18	36	0	11.70	11.70		
0	49	0	0	14	13.65	13.65		
18	0	36	72	0	23.40	23.40	9.5062	9.5062
36	0	72	144	0	46.80	46.80		
0	14	0	0	4	3.90	3.90		
$Q_{DTS}(b) = Q_{DTS}$					R_{DTS}	$R_{DTS}(b)$	c_{DTS}	$c_{DTS}(b)$
10000	0	7000	40000	0	10400	10400		
0	62500	0	0	37500	19375	19375		
7000	0	4900	28000	0	7280	7280	8411.1250	8411.1250
40000	0	28000	160000	0	41600	41600		
0	37500	0	0	22500	11625	11625		
$Q(b)$					$R(b)$		$c(b)$	
20.0	0.0	25.0	76.0	0.0	62.400			
0.0	112.5	0.0	0.0	51.5	93.275			
25.0	0.0	41.9	100.0	0.0	60.880		70.530	
76.0	0.0	100.0	305.0	0.0	138.550			
0.0	51.5	0.0	0.0	27.5	135.625			

The objective is to reduce the carbon footprint of the benchmark and minimizing the tracking error risk. In this case, the carry vector $\mathcal{C}$ is set to the zero vector and the optimization problem is defined as:

$$w^\star(\mathcal{R}) = \arg\min \frac{1}{2} w^\top Q(b) w - w^\top R(b)$$

$$\text{s.t.} \begin{cases} \mathbf{1}_9^\top w = 1 \\ \mathcal{CI}^\top w \leq (1 - \mathcal{R}) \mathcal{CI}(b) \\ \frac{b}{4} \leq w \leq 4b \end{cases}$$

where $\mathcal{R}$ is the reduction rate. Since the bonds are ordering by sectors, $Q(b)$ is a block diagonal matrix:

$$Q(b) = \begin{pmatrix} Q_1 & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} \\ \mathbf{0}_{3\times3} & Q_2 & \mathbf{0}_{3\times3} \\ \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & Q_3 \end{pmatrix} \times 10^3$$

where:

$$Q_1 = \begin{pmatrix} 0.3611 & 0.5392 & 0.2877 \\ 0.5392 & 1.2148 & 0.5926 \\ 0.2877 & 0.5926 & 0.4189 \end{pmatrix}, \qquad Q_2 = \begin{pmatrix} 3.2218 & 1.7646 & 0.5755 \\ 1.7646 & 1.2075 & 0.3810 \\ 0.5755 & 0.3810 & 0.2343 \end{pmatrix}$$

---

[14]We have $-\gamma \sum_{i=1}^n (w_i - b_i) \mathcal{C}_i = \mathcal{QF}(w; \mathbf{0}_{n,n}, \gamma\mathcal{C}, \gamma b^\top \mathcal{C})$.

Figure 11.1: Relationship between the reduction rate and the tracking risk (Example 40)

and:

$$Q_3 = \begin{pmatrix} 2.1328 & 1.7926 & 1.1899 \\ 1.7926 & 1.7653 & 1.1261 \\ 1.1899 & 1.1261 & 0.8664 \end{pmatrix}$$

The vector $R(b)$ is equal to $(2.2431, 4.3886, 2.4004, 6.2678, 3.7506, 1.2972, 2.3537, 2.1195, 1.4243) \times 10^2$. In Tables 11.2 and 11.3, we report the optimal weights and the risk statistics. Increasing the reduction rate implies to take more active risk in terms of active share, modified duration and duration-times-spread factor. For instance, if we target $\mathcal{R} = 50\%$, the MD-based tracking risk is 2.4 months while the DTS-based tracking risk is 30.11 bps. The asset- and sector-based[15] active shares are respectively equal to 21.21% and 28.31%. In Figure 11.1, we show the relationship between the reduction rate and these different risk statistics. We notice that there is no solution to the optimization problem when $\mathcal{R} \geq 65.73\%$.

Table 11.2: Weights in % of optimized bond portfolios (Example 40)

Portfolio	#1	#2	#3	#4	#5	#6	#7	#8	#9
b	21.00	19.00	16.00	12.00	11.00	8.00	6.00	4.00	3.00
$w^\star(10\%)$	21.92	19.01	15.53	11.72	11.68	7.82	6.68	4.71	0.94
$w^\star(30\%)$	26.29	20.24	10.90	10.24	16.13	3.74	9.21	2.50	0.75
$w^\star(50\%)$	27.48	23.97	4.00	6.94	22.70	2.00	11.15	1.00	0.75

---

[15]The sector-based active share is defined as $\mathrm{AS}_{\boldsymbol{\mathcal{S}}ector}(w \mid b) = \frac{1}{2}\sum_{j=1}^{n_{\boldsymbol{\mathcal{S}}ector}}\left|\sum_{i \in Sector_j}(w_i - b_i)\right|$.

Table 11.3: Risk statistics of optimized bond portfolios (Example 40)

Portfolio	$\text{AS}_{\mathcal{S}ector}$ (in %)	$\text{MD}(w)$ (in years)	$\text{DTS}(w)$ (in bps)	$\sigma_{\text{AS}}(w \mid b)$ (in %)	$\sigma_{\text{MD}}(w \mid b)$ (in years)	$\sigma_{\text{DTS}}(w \mid b)$ (in bps)	$\mathcal{CI}(w)$ $\text{gCO}_2\text{e}/\$$
b	0.00	5.43	290.18	0.00	0.00	0.00	184.39
$w^\star(10\%)$	3.00	5.45	293.53	2.62	0.02	3.80	165.95
$w^\star(30\%)$	14.87	5.58	303.36	10.98	0.10	14.49	129.07
$w^\star(50\%)$	28.31	5.73	302.14	21.21	0.19	30.11	92.19

11.1.3 Advanced optimization problems

In this section, we go beyond the previous approaches in order to solve more complex portfolio optimization problems. First, we consider the case of large bond universes when there are more than 10 000 bonds in the index. In this case, we generally use linear programming (LP) algorithms. Second, we show how to incorporate some standard nonlinear constraints in the QP/LP framework. Finally, we consider mathematical problems when the variable to optimize is not a vector of weights, but the numbers of shares. In particular, this last problem occurs when we would like to build an investible fixed-income portfolio.

Large bond universe

Quadratic problem algorithms are efficient when the dimension of the problem is relative small, say, when $n \leq 5\,000$. The issue is not the convergence of the algorithm, but more the manipulation of the Hessian matrix Q of the quadratic form. Indeed, since Q is a $n \times n$ matrix, we have to store n^2 floating-point numbers. For instance, Q has 25 millions of elements when n is equal to $5\,000$. In the case of bond portfolio optimization, we have seen that Q is a block matrix with the following structure[16]:

$$Q = \begin{pmatrix} Q_1 & \mathbf{0}_{n_{(1)} \times n_{(2)}} & & \mathbf{0} \\ \mathbf{0}_{n_{(2)} \times n_{(1)}} & Q_2 & & \\ & & \ddots & \\ \mathbf{0} & & & Q_{n_{\mathcal{S}ector}} \end{pmatrix}$$

where Q_j is the $n_{(j)} \times n_{(j)}$ matrix associated to the jth sector and $n_{(j)}$ is the number of assets that belong to the jth sector. By construction, we have $\sum_{j=1}^{n_{\mathcal{S}ector}} n_{(j)} = n$. Therefore, the number of non-zero entries is equal to:

$$\eta(Q) = \sum_{j=1}^{n_{\mathcal{S}ector}} n_{(j)}^2 \ll n^2$$

Let us assume that the number of assets are equally distributed among the sectors; $n_{(j)} = n/n_{\mathcal{S}ector}$. We obtain:

$$\eta(Q) = \sum_{j=1}^{n_{\mathcal{S}ector}} \left(\frac{n}{n_{\mathcal{S}ector}}\right)^2 = \frac{n^2}{n_{\mathcal{S}ector}}$$

We deduce that Q is a highly sparse matrix since the number of non-zero elements is a fraction of the full number of elements. If $n = 5\,000$ and $n_{\mathcal{S}ector} = 10$, the number of non-zero elements is equal to 2.5 millions of elements. We reiterate that the bucket concept in bond portfolio optimization generalizes the sector concept. For instance, if we consider ten sectors, three currencies and five

---

[16]We assume that the bonds are ordered according to the sectors.

maturity bands, the number of buckets is equal to $10 \times 3 \times 5 = 150$. If $n = 5\,000$, the number of non-zero elements is approximatively equal to $167\,000$ of elements. This is equivalent to manipulate a dense matrix of dimension 409×409. It follows that portfolio optimization with large bond universe remains feasible by using sparse quadratic programming solvers.

As of 31/01/2023, the Bloomberg Global Aggregate Total Return Index has $28\,799$ securities, while the ICE BOFA Global Broad Market Index has $33\,575$ securities. For these two bond universes, we reach the limits of QP optimization from a numerical viewpoint. For instance, the number of non-zero entries is equal to more than 8 million floating-point numbers for the Hessian Q matrix when n is equal to $35\,000$ and we have 150 equally-distributed buckets. In the case where we only consider sectors as buckets, we have more than 120 million floating-point numbers for 10 equally-distributed sectors. Moreover, the number of bonds are generally concentrated in some sectors, e.g. Financials. This explains that bond portfolio managers prefer to use another approach than the QP framework when performing portfolio optimization with such large investment universes.

Figure 11.2: Relationship between $\mathcal{L}_1$-norm $\mathcal{L}_2$-norm
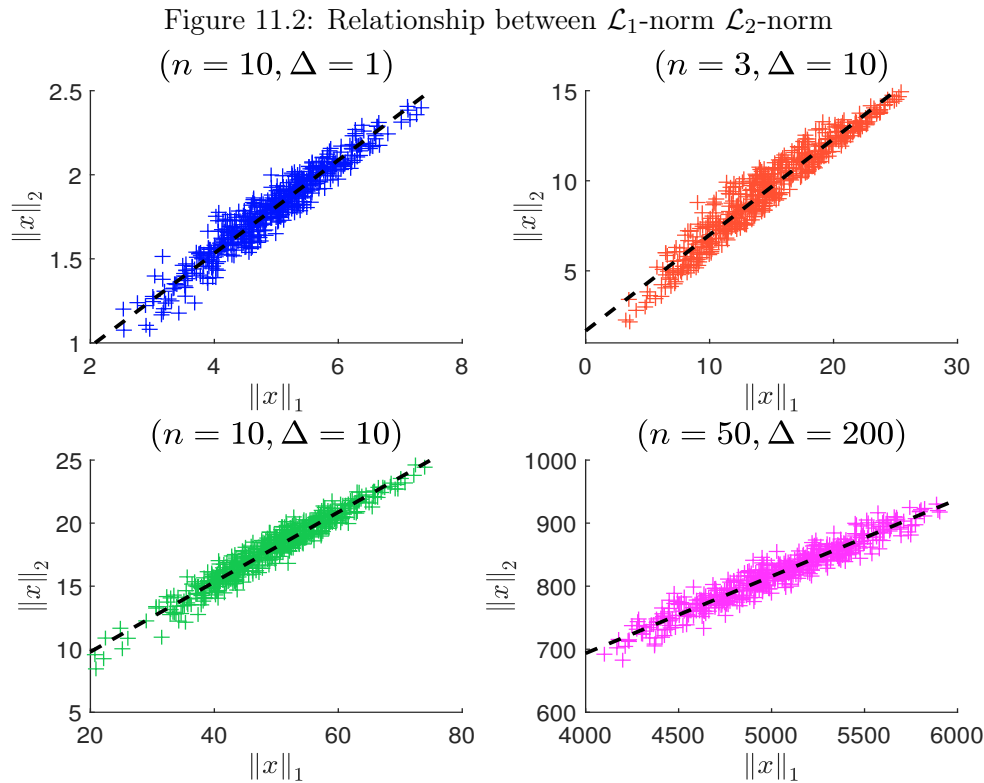
The underlying idea is to replace $\mathcal{L}_2$-norm risk measures by $\mathcal{L}_1$-norm risk measures. We recall that $f(x) = \sqrt{x^2} = |x|$ when x is a scalar. In the general case where x is a $n \times 1$ vector, we have $\|x\|_1 = \sum_{i=1}^n |x_i|$ and $\|x\|_2 = \sqrt{\sum_{i=1}^n x_i^2} = \sqrt{x^\top x}$. We notice that $\|x\|_2^2 \le \sum_{i=1}^n |x_i| \sum_{i=1}^n |x_i| = \|x\|_1^2$ and $\|x\|_1^2 = \mathbf{1}_n^\top |x| \le \|\mathbf{1}_n\|_2 \cdot \|x\|_2 \le \sqrt{n} \|x\|_2$, implying that $\|x\|_2 \le \|x\|_1 \le \sqrt{n} \|x\|_2$. This means that there is not a simple relationship between the two norms. Nevertheless, if the vector x is "*homogenous*", we can use the following approximation $\|x\|_2 \approx \beta_0 + \beta_1 \|x\|_1$ where $0 < \beta_1 \le 1$. In Figure 11.2, we report the scatter plot of $(\|x\|_1, \|x\|_2)$ and the least square fitting model $\|x\|_2 = \hat{\beta}_0 + \hat{\beta}_1 \|x\|_1$ when x_i is the random variate, whose probability distribution is $\mathbf{F}(x_i) = \Pr\{\mathcal{U}_{[0,1]} \le \Delta^{-1} x_i\}$. The estimated values $\left(\hat{\beta}_0, \hat{\beta}_1\right)$ depend then on the dimension n of the vector x and the parameter Δ.

We exploit the previous relationship and replace the synthetic risk measure by:

$$\mathcal{D}\left(w \mid b\right) = \varphi'_{\mathrm{AS}}\mathcal{D}_{\mathrm{AS}}\left(w \mid b\right) + \varphi'_{\mathrm{MD}}\mathcal{D}_{\mathrm{MD}}\left(w \mid b\right) + \varphi'_{\mathrm{DTS}}\mathcal{D}_{\mathrm{DTS}}\left(w \mid b\right)$$

where:

$$\mathcal{D}_{\mathrm{AS}}\left(w \mid b\right) = \frac{1}{2}\sum_{i=1}^{n}|w_i - b_i| = \frac{1}{2}\left\|w - b\right\|_1$$

$$\mathcal{D}_{\mathrm{MD}}\left(w \mid b\right) = \sum_{j=1}^{n_{\boldsymbol{Sector}}}\left|\sum_{i\in\boldsymbol{Sector}_j}\left(w_i - b_i\right)\mathrm{MD}_i\right| = \sum_{j=1}^{n_{\boldsymbol{Sector}}}\left\|\boldsymbol{s}_j \circ (w - b) \circ \mathrm{MD}\right\|_1$$

$$\mathcal{D}_{\mathrm{DTS}}\left(w \mid b\right) = \sum_{j=1}^{n_{\boldsymbol{Sector}}}\left|\sum_{i\in\boldsymbol{Sector}_j}\left(w_i - b_i\right)\mathrm{DTS}_i\right| = \sum_{j=1}^{n_{\boldsymbol{Sector}}}\left\|\boldsymbol{s}_j \circ (w - b) \circ \mathrm{DTS}\right\|_1$$

The optimization problem becomes then:

$$w^\star = \arg\min \mathcal{D}\left(w \mid b\right) - \gamma\sum_{i=1}^{n}\left(w_i - b_i\right)\mathcal{C}_i \tag{11.7}$$

$$\text{s.t.} \quad \left\{ \begin{array}{l} \mathbf{1}_n^\top w = 1 \\ \mathbf{0}_n \leq w \leq \mathbf{1}_n \end{array} \right.$$

Using the absolute value trick, we transform the nonlinear objective function into a linear function:

$$w^\star = \arg\min \frac{1}{2}\varphi'_{\mathrm{AS}}\sum_{i=1}^{n}\tau_{i,w} + \varphi'_{\mathrm{MD}}\sum_{j=1}^{n_{\boldsymbol{Sector}}}\tau_{j,\mathrm{MD}} + \varphi'_{\mathrm{DTS}}\sum_{j=1}^{n_{\boldsymbol{Sector}}}\tau_{j,\mathrm{DTS}} - \gamma\sum_{i=1}^{n}\left(w_i - b_i\right)\mathcal{C}_i$$

$$\text{s.t.} \quad \left\{ \begin{array}{l} \mathbf{1}_n^\top w = 1 \\ \mathbf{0}_n \leq w \leq \mathbf{1}_n \\ |w_i - b_i| \leq \tau_{i,w} \\ \left|\sum_{i\in\boldsymbol{Sector}_j}\left(w_i - b_i\right)\mathrm{MD}_i\right| \leq \tau_{j,\mathrm{MD}} \\ \left|\sum_{i\in\boldsymbol{Sector}_j}\left(w_i - b_i\right)\mathrm{DTS}_i\right| \leq \tau_{j,\mathrm{DTS}} \\ \tau_{i,w} \geq 0, \tau_{j,\mathrm{MD}} \geq 0, \tau_{j,\mathrm{DTS}} \geq 0 \end{array} \right.$$

We notice that:

$$|w_i - b_i| \leq \tau_{i,w} \Leftrightarrow \left\{ \begin{array}{l} w_i - \tau_{i,w} \leq b_i \\ -w_i - \tau_{i,w} \leq -b_i \end{array} \right.$$

and:

$$(*) \quad \Leftrightarrow \quad \left|\sum_{i\in\boldsymbol{Sector}_j}\left(w_i - b_i\right)\mathrm{MD}_i\right| \leq \tau_{j,\mathrm{MD}}$$

$$\Leftrightarrow \quad -\tau_{j,\mathrm{MD}} \leq \sum_{i\in\boldsymbol{Sector}_j}\left(w_i - b_i\right)\mathrm{MD}_i \leq \tau_{j,\mathrm{MD}}$$

$$\Leftrightarrow \quad -\tau_{j,\mathrm{MD}} + \sum_{i\in\boldsymbol{Sector}_j}b_i\,\mathrm{MD}_i \leq \sum_{i\in\boldsymbol{Sector}_j}w_i\,\mathrm{MD}_i \leq \tau_{j,\mathrm{MD}} + \sum_{i\in\boldsymbol{Sector}_j}b_i\,\mathrm{MD}_i$$

$$\Leftrightarrow \quad -\tau_{j,\mathrm{MD}} + \mathrm{MD}_j^\star \leq \left(\boldsymbol{s}_j \circ \mathrm{MD}\right)^\top w \leq \tau_{j,\mathrm{MD}} + \mathrm{MD}_j^\star$$

$$\Leftrightarrow \quad \left\{ \begin{array}{l} \left(\boldsymbol{s}_j \circ \mathrm{MD}\right)^\top w - \tau_{j,\mathrm{MD}} \leq \mathrm{MD}_j^\star \\ -\left(\boldsymbol{s}_j \circ \mathrm{MD}\right)^\top w - \tau_{j,\mathrm{MD}} \leq -\mathrm{MD}_j^\star \end{array} \right.$$

We can now formulate problem (11.7) as a standard linear programming (LP) problem:

$$x^\star \quad = \quad \arg\min c^\top x \qquad\qquad (11.8)$$

$$\text{s.t.} \quad \begin{cases} Ax = B \\ Cx \le D \\ x^- \le x \le x^+ \end{cases}$$

where:

$$x = \begin{pmatrix} w \\ \tau_w \\ \tau_{\mathrm{MD}} \\ \tau_{\mathrm{DTS}} \end{pmatrix}$$

is a vector of dimension $n_x = 2 \times (n + n_{\boldsymbol{Sector}})$. The vector c is equal to:

$$c = \begin{pmatrix} -\gamma \mathcal{C} \\ \dfrac{1}{2}\varphi'_{\mathrm{AS}}\mathbf{1}_n \\ \varphi'_{\mathrm{MD}}\mathbf{1}_{n_{\boldsymbol{Sector}}} \\ \varphi'_{\mathrm{DTS}}\mathbf{1}_{n_{\boldsymbol{Sector}}} \end{pmatrix}$$

The equality constraint is defined by $A = \begin{pmatrix} \mathbf{1}_n^\top & \mathbf{0}_n^\top & \mathbf{0}_{n_{\boldsymbol{Sector}}}^\top & \mathbf{0}_{n_{\boldsymbol{Sector}}}^\top \end{pmatrix}$ and $B = 1$. We also have:

$$Cx \le D \Leftrightarrow \begin{pmatrix} I_n & -I_n & \mathbf{0}_{n,n_{\boldsymbol{Sector}}} & \mathbf{0}_{n,n_{\boldsymbol{Sector}}} \\ -I_n & -I_n & \mathbf{0}_{n,n_{\boldsymbol{Sector}}} & \mathbf{0}_{n,n_{\boldsymbol{Sector}}} \\ C_{\mathrm{MD}} & \mathbf{0}_{n_{\boldsymbol{Sector}},n} & -I_{n_{\boldsymbol{Sector}}} & \mathbf{0}_{n_{\boldsymbol{Sector}},n_{\boldsymbol{Sector}}} \\ -C_{\mathrm{MD}} & \mathbf{0}_{n_{\boldsymbol{Sector}},n} & -I_{n_{\boldsymbol{Sector}}} & \mathbf{0}_{n_{\boldsymbol{Sector}},n_{\boldsymbol{Sector}}} \\ C_{\mathrm{DTS}} & \mathbf{0}_{n_{\boldsymbol{Sector}},n} & \mathbf{0}_{n_{\boldsymbol{Sector}},n_{\boldsymbol{Sector}}} & -I_{n_{\boldsymbol{Sector}}} \\ -C_{\mathrm{DTS}} & \mathbf{0}_{n_{\boldsymbol{Sector}},n} & \mathbf{0}_{n_{\boldsymbol{Sector}},n_{\boldsymbol{Sector}}} & -I_{n_{\boldsymbol{Sector}}} \end{pmatrix} x \le \begin{pmatrix} b \\ -b \\ \mathrm{MD}^\star \\ -\mathrm{MD}^\star \\ \mathrm{DTS}^\star \\ -\mathrm{DTS}^\star \end{pmatrix}$$

where C_{MD} and C_{DTS} are two $n_{\boldsymbol{Sector}} \times n$ matrices, whose (j,i)-elements are $\boldsymbol{s}_{i,j}\,\mathrm{MD}_i$ and $\boldsymbol{s}_{i,j}\,\mathrm{DTS}_i$, $\mathrm{MD}^\star = \left(\mathrm{MD}_1^\star, \ldots, \mathrm{MD}_{n_{\boldsymbol{Sector}}}^\star\right)$ and $\mathrm{DTS}^\star = \left(\mathrm{DTS}_1^\star, \ldots, \mathrm{DTS}_{n_{\boldsymbol{Sector}}}^\star\right)$. Finally, the bounds are $x^- = \mathbf{0}_{n_x}$ and[17] $x^+ = \infty \cdot \mathbf{1}_{n_x}$.

Remark 91 *In the case where there are additional constraints, they are cast into LP constraints using the following equivalence:*

$$\begin{cases} Aw = B \\ Cw \le D \end{cases} \Leftrightarrow \begin{cases} \begin{pmatrix} A & \mathbf{0}_{n_A,n_x-n} \end{pmatrix} x = B \\ \begin{pmatrix} C & \mathbf{0}_{n_A,n_x-n} \end{pmatrix} x \le D \end{cases}$$

Remark 92 *While the dimension of the QP problem was n, the dimension of the LP problem is $2 \times (n + n_{\boldsymbol{Sector}})$. This is not an issue since LP algorithms are efficient to solve optimization problems with million of decision variables.*

Example 41 *We consider a toy example with four corporate bonds:*

Issuer	#1	#2	#3	#4
b_i (in %)	35	15	20	30
$\mathcal{CI}_i$ (in tCO$_2$e/\$ mn)	117	284	162.5	359
MD_i (in years)	3.0	5.0	2.0	6.0
DTS_i (in bps)	100	150	200	250
$\boldsymbol{Sector}$	1	1	2	2

We would like to reduce the carbon footprint by 20%, and we set $\varphi'_{\mathrm{AS}} = 100$, $\varphi'_{\mathrm{MD}} = 25$ and $\varphi'_{\mathrm{DTS}} = 1$.

---

[17]We remark that the constraint $w \le \mathbf{1}_n$ is already embedded into the linear equality constraint.

We have $n = 4$, $n_{\boldsymbol{\mathcal{S}}ector} = 2$ and:

$$x = (\underbrace{w_1, w_2, w_3, w_4}_{w}, \underbrace{\tau_{w_1}, \tau_{w_2}, \tau_{w_3}, \tau_{w_4}}_{\tau_w}, \underbrace{\tau_{\text{MD}_1}, \tau_{\text{MD}_2}}_{\tau_{\text{MD}}}, \underbrace{\tau_{\text{DTS}_1}, \tau_{\text{DTS}_2}}_{\tau_{\text{DTS}}})$$

Since the vector $\mathcal{C}$ is equal to $\mathbf{0}_4$, we obtain:

$$c = (0, 0, 0, 0, 50, 50, 50, 50, 25, 25, 1, 1)$$

The equality system $Ax = B$ is defined by:

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

and $B = 1$, while the inequality system $Cx \leq D$ is given by:

$$C = \begin{pmatrix}
 & I_4 & & & -I_4 & & \mathbf{0}_{4,4} & \\
 & -I_4 & & & -I_4 & & \mathbf{0}_{4,4} & \\
3 & 5 & 0 & 0 & -1 & 0 & 0 & 0 \\
0 & 0 & 2 & 6 & 0 & -1 & 0 & 0 \\
-3 & -5 & 0 & 0 \quad \mathbf{0}_{4,4} & -1 & 0 & 0 & 0 \\
0 & 0 & -2 & -6 & 0 & -1 & 0 & 0 \\
100 & 150 & 0 & 0 & 0 & 0 & -1 & 0 \\
0 & 0 & 200 & 250 & 0 & 0 & 0 & -1 \\
-100 & -150 & 0 & 0 \quad \mathbf{0}_{4,4} & 0 & 0 & -1 & 0 \\
0 & 0 & -200 & -250 & 0 & 0 & 0 & -1 \\
117 & 284 & 162.5 & 359 \quad \mathbf{0}_{1,4} & 0 & 0 & 0 & 0
\end{pmatrix}$$

and:

$$D = (0.35, 0.15, 0.2, 0.3, -0.35, -0.15, -0.2, -0.3, 1.8, 2.2, -1.8, -2.2, 57.5, 115, -57.5, -115, 179)$$

The last row of $Cx \leq D$ corresponds to the carbon footprint constraint. Indeed, we have $\boldsymbol{\mathcal{CI}}(b) = 223.75$ tCO$_2$e/\$ mn and $(1 - \boldsymbol{\mathcal{R}})\boldsymbol{\mathcal{CI}}(b) = 0.80 \times 223.75 = 179.00$ tCO$_2$e/\$ mn. We can now solve the LP program, and we obtain the following solution: $w^\star = (47.34\%, 0\%, 33.3\%, 19.36\%)$, $\tau_w^\star = (12.34\%, 15\%, 13.3\%, 10.64\%)$, $\tau_{\text{MD}}^\star = (0.3798, 0.3725)$ and $\tau_{\text{DTS}}^\star = (10.1604, 0)$.

We consider Example 40 on page 558. By assuming that $\varphi'_{\text{AS}} = \varphi_{\text{AS}} = 100$, $\varphi'_{\text{MD}} = \varphi_{\text{MD}} = 25$ and $\varphi'_{\text{DTS}} = \varphi'_{\text{DTS}} = 0.001$, we obtain the results given in Table 11.4 and 11.5. By construction, these results differ from those obtained in 11.2 and 11.3 on page 560. However, we observe a consistency between the two solutions. In particular, the overweighting and underweighting directions are respected by the QP and LP algorithms.

Table 11.4: Weights in % of optimized bond portfolios (Example 40)

Portfolio	#1	#2	#3	#4	#5	#6	#7	#8	#9
b	21.00	19.00	16.00	12.00	11.00	8.00	6.00	4.00	3.00
$w^\star$ (10%)	21.70	19.00	16.00	12.00	11.00	8.00	7.46	4.00	0.84
$w^\star$ (30%)	34.44	19.00	4.00	11.65	11.98	6.65	7.52	4.00	0.75
$w^\star$ (50%)	33.69	19.37	4.00	3.91	24.82	2.00	10.46	1.00	0.75

Remark 93 *The linear programming approach is used by Barahhou et al. (2022) and Ben Slimane et al. (2023a) to implement portfolio alignment and net-zero investment policies[18]. The reason is that the investment universe of corporate bonds is too large to be manipulated by the QP algorithm.*

---

[18]See Section 11.2 on page 567 for using linear programming for decarbonized bond portfolios and Section 11.3 on page 591 for net-zero bond portfolios.

Table 11.5: Risk statistics of optimized bond portfolios (Example 40)

Portfolio	$\text{AS}_{\boldsymbol{\mathcal{S}}ector}$ (in %)	$\text{MD}(w)$ (in years)	$\text{DTS}(w)$ (in bps)	$\sigma_{\text{AS}}(w \mid b)$ (in %)	$\sigma_{\text{MD}}(w \mid b)$ (in years)	$\sigma_{\text{DTS}}(w \mid b)$ (in bps)	$\boldsymbol{\mathcal{CI}}(w)$ $\text{gCO}_2\text{e}/\$$
b	0.00	5.43	290.18	0.00	0.00	0.00	184.39
$w^\star(10\%)$	2.16	5.45	297.28	2.16	0.02	7.10	165.95
$w^\star(30\%)$	15.95	5.43	300.96	15.95	0.00	13.20	129.07
$w^\star(50\%)$	31.34	5.43	268.66	31.34	0.00	65.12	92.19

Nonlinear constraints

A first idea to deal with nonlinear constraints is to use a general optimization solver, such as the sequential quadratic programming algorithm.However, this algorithm is not efficient when the dimension of the problem is large. In most cases, nonlinear constraints are standard in portfolio optimization. They concern the turnover, the transaction costs, the leverage limit, etc. Therefore, we can use the quadratic property of the objective function to solve the nonlinear constrained optimization problem. The underlying idea is to preserve the QP or LP structure of the problem by considering sequential optimization. Among the different algorithms, the most famous is the alternating direction method of multipliers[19] (ADMM), which can be used to solve many nonlinear portfolio optimization problems (Perrin and Roncalli, 2020). In particular, Lezmi *et al.* (2022) used it to solve multi-period portfolio allocation and some classes of net-zero optimization problems.

Investible portfolios

The previous optimization problems are defined with respect to the unknown variable $w = (w_1, \ldots, w_n)$, where w_i is the portfolio weight of Security i. In this approach, $w \in \mathbb{R}^n$ and the budget constraint is $\sum_{i=1}^n w_i = 1$. This type of problem falls under the umbrella of continuous portfolio optimization. In some cases, it does not make sense to define the portfolio allocation by considering continuous weights because the solution is far from investible. Therefore, the unknown variable becomes $q = (q_1, \ldots, q_n)$, where q_i is the number of shares of Security i. Let P_i be the price and $Q_i = q_i P_i$ be the tradable amount. In this approach, $q \in \mathbb{N}^n$ and the budget constraint is $\sum_{i=1}^n q_i P_i = \sum_{i=1}^n Q_i = \mathcal{A}$ where $\mathcal{A}$ is the investment amount. This type of problem falls under the umbrella of discrete portfolio optimization and can be solved using genetic algorithms (Ben Slimane, 2021) or mixed integer programming algorithms (Ben Slimane and Menchaoui, 2023).

Given an optimal solution $w^\star$, the amount invested in Security i is equal to $Q_i = w_i^\star \mathcal{A}$. For some markets, Q_i may be far from the optimal investible solution, which is equal to $q_i^\star P_i$. This is for example the case of corporate bonds[20]. Let us introduce some definitions. The minimum tradable lot size is the minimum order size that can be bought or sold for a bond, while the incremental lot size is the minimum multiple that must be added to the minimum tradable lot to buy or sell a bond. For example, if the minimum tradable lot size is $10\,000$ and the incremental lot size is $1\,000$, the order sizes are $Q = 10\,000 + 1\,000 \cdot k$ where $k \in \mathbb{N}$. For EUR and USD corporate bonds, the minimum tradable size is typically €$100\,000$ and \$$2\,000$, respectively. Suppose we want to invest 5% of a million euros in a security. If the lot size of the security is €540, we have the choice of buying 92 or 93 shares, that is €$49\,680$ or €$50\,220$. Both solutions are close to the desired amount of €$50\,000$. If the minimum tradeable amount of the security is €$100\,000$, we have the choice of investing 0 or €$100\,000$. Both solutions are far from the desired amount of €$50\,000$.

---

[19]The ADMM algorithm is described on page 653.

[20]For equity and sovereign bond portfolios, it is common practice to use continuous portfolio optimization.

11.2 Portfolio decarbonization

The objective of portfolio decarbonization is to construct a portfolio w with a low carbon metric, which is generally the carbon intensity. Most of the time, the fund manager needs a reference portfolio b in order to measure the reduction $\mathcal{R}(w \mid b)$. This reference portfolio can be the capitalization-weighted benchmark of the investment universe, a model portfolio, the managed portfolio one year ago, etc. When an explicit objective is defined, portfolio decarbonization can be viewed as a portfolio management constraint:

$$\mathcal{CI}(w) \leq (1 - \mathcal{R})\,\mathcal{CI}(b)$$

In this case, the decarbonization issue is solved by using constrained portfolio optimization. This framework can be extended when the global decarbonization constraint is replaced by a set of sector-specific decarbonization constraints. At first sight, portfolio decarbonization is then an application of basic portfolio optimization. Nevertheless, there are many implementation differences that make the exercise more difficult than we can imagine. In particular, portfolio decarbonization is ultimately an exclusion process and can be put in the category of exclusion ESG strategies. In some sense, portfolio decarbonization is a disguised form of divestment, which impacts the worst-in-class issuers that have the higher carbon intensities.

11.2.1 Global reduction of the carbon footprint

Equity portfolios

Let b be the reference portfolio (e.g., a benchmark). The optimization problem is defined as:

$$
\begin{aligned}
w^{\star} &= \arg\min \frac{1}{2}(w - b)^{\top}\Sigma(w - b) && (11.9)\\[4pt]
&\text{s.t.} \quad \begin{cases}
\mathbf{1}_n^{\top} w = 1 \\
w \in \Omega \\
\mathbf{0}_n \leq w \leq \mathbf{1}_n \\
\mathcal{CI}(w) \leq (1 - \mathcal{R})\,\mathcal{CI}(b)
\end{cases}
\end{aligned}
$$

where $\mathcal{R}$ is the reduction rate and $w \in \Omega$ is a set of additional weight constraints (e.g., weight or sector deviation). Therefore, we minimize the tracking error variance of the investment portfolio w with respect to the reference portfolio b by imposing a long-only constraint and reducing the carbon intensity of the benchmark. Since we both impose a constraint and minimize the tracking error risk, the portfolio w has fewer stocks than the benchmark b. In fact, the number of stocks depends on several parameters: the reduction rate $\mathcal{R}$, the number n of stocks in the benchmark and the covariance matrix. This implies that portfolio w is less diversified than benchmark b. In order to explicitly control the number of removed stocks, Andersson *et al.* (2016) proposed a second portfolio decarbonization approach by eliminating the m worst performing issuers in terms of carbon intensity.

Let $\mathcal{CI}_{i:n}$ be the order statistics of $(\mathcal{CI}_1, \ldots, \mathcal{CI}_n)$ such that:

$$\min \mathcal{CI}_i = \mathcal{CI}_{1:n} \leq \mathcal{CI}_{2:n} \leq \cdots \leq \mathcal{CI}_{i:n} \leq \cdots \leq \mathcal{CI}_{n:n} = \max \mathcal{CI}_i$$

The carbon intensity bound $\mathcal{CI}^{(m,n)}$ is defined as $\mathcal{CI}^{(m,n)} = \mathcal{CI}_{n-m+1:n}$ where $\mathcal{CI}_{n-m+1:n}$ is the $(n - m + 1)$-th order statistic of $(\mathcal{CI}_1, \ldots, \mathcal{CI}_n)$. Eliminating the m worst performing assets is equivalent to imposing the following constraint: $\mathcal{CI}_i \geq \mathcal{CI}^{(m,n)} \Rightarrow w_i = 0$. We then obtain the

following optimization problem:

$$w^\star \quad = \quad \arg\min \frac{1}{2}(w-b)^\top \Sigma (w-b) \tag{11.10}$$

$$\text{s.t.} \quad \begin{cases} \mathbf{1}_n^\top w = 1 \\ w \in \Omega \\ \mathbf{0}_n \le w \le \mathbb{1}\left\{ \boldsymbol{\mathcal{CI}} < \boldsymbol{\mathcal{CI}}^{(m,n)} \right\} \end{cases}$$

Finally, a third method consists in re-weighting the remaining assets:

$$w_i^\star = \frac{\mathbb{1}\left\{ \boldsymbol{\mathcal{CI}}_i < \boldsymbol{\mathcal{CI}}^{(m,n)} \right\} \cdot b_i}{\sum_{k=1}^n \mathbb{1}\left\{ \boldsymbol{\mathcal{CI}}_k < \boldsymbol{\mathcal{CI}}^{(m,n)} \right\} \cdot b_k} \tag{11.11}$$

Remark 94 *Problem (11.9) is called the "threshold" approach, whereas problems (11.10) and (11.11) are known as the "order-statistic" and "naive" approach.*

The QP form of problem (11.9) is: $Q = \Sigma$, $R = \Sigma b$, $A = \mathbf{1}_n^\top$, $B = 1$, $C = \boldsymbol{\mathcal{CI}}^\top$, $D = (1 - \boldsymbol{\mathcal{R}})\boldsymbol{\mathcal{CI}}(b)$, $w^- = \mathbf{0}_n$ and $w^+ = \mathbf{1}_n$. For problem (11.10), the matrices C and D vanishes and the upper bound becomes $w^+ = \mathbb{1}\left\{ \boldsymbol{\mathcal{CI}} < \boldsymbol{\mathcal{CI}}^{(m,n)} \right\}$, which is a vector of zeros and ones.

Example 42 *We consider a capitalization-weighted equity index, which is composed of eight stocks. Their weights are equal to 20%, 19%, 17%, 13%, 12%, 8%, 6% and 5%. The carbon intensities (expressed in* tCO$_2$e/\$ mn*) are respectively equal to 100.5, 97.2, 250.4, 352.3, 27.1, 54.2, 78.6 and 426.7. To evaluate the risk of the portfolio, we use the market one-factor model: the beta β_i of each stock is equal to 0.30, 1.80, 0.85, 0.83, 1.47, 0.94, 1.67 and 1.08, the idiosyncratic volatilities $\tilde{\sigma}_i$ are respectively equal to 10%, 5%, 6%, 12%, 15%, 4%, 8% and 7%, and the estimated market volatility σ_m is 18%.*

In order to solve the different optimization problems, we need to compute the covariance matrix:

$$\Sigma = \beta\beta^\top \sigma_m^2 + D$$

where β is the vector of beta coefficients, σ_m^2 is the variance of the market portfolio and $D = diag\left(\tilde{\sigma}_1^2, \dots, \tilde{\sigma}_n^2\right)$ is the diagonal matrix, whose elements are the idiosyncratic variances. The specification of the D matrix in the optimization problem (11.9) also requires to compute the carbon intensity of the benchmark and we obtain $\boldsymbol{\mathcal{CI}}(b) = 160.57$ tCO$_2$e/\$ mn. We report the different results in Tables 11.6, 11.7 and 11.8. With the threshold approach, we can target a explicit value of the reduction rate. For instance, if $\boldsymbol{\mathcal{R}}$ is set to 10%, the tracking error volatility $\sigma(w^\star \mid b)$ is equal to 30 bps. In the case of the order-statistic approach, it is not possible to have an optimized portfolio, whose tracking error volatility is less than 37 bps. This figure is obtained when we exclude the eight stock, which has the largest carbon intensity. In this case, the reduction rate is equal to 9.62%. If we exclude the two assets that have the largest carbon footprint (asset #4 and #8), the tracking error volatility and the reduction rate are equal to 1.68% and 29.33%. If we use the order-statistic approach, it is then not possible to target an intermediary reduction rate, for instance 20% in our example. For a given value of m, the naive implies a lower reduction rate and a higher tracking risk compared to the order-statistic approach. For example, if m is equal to 3, their values are respectively equal to 2.25% and 54.05% versus 3.04% and 51.26%. In Figure 11.3, we show the efficient decarbonization frontier which indicates the relationship between the reduction rate and

Table 11.6: Optimal decarbonization portfolios (Example 42, threshold approach)

$\mathcal{R}$	0	10	20	30	40	50	$\mathcal{CI}_i$
$w_1^\star$	20.00	20.54	21.14	21.86	22.58	22.96	100.5
$w_2^\star$	19.00	19.33	19.29	18.70	18.11	17.23	97.2
$w_3^\star$	17.00	15.67	12.91	8.06	3.22	0.00	250.4
$w_4^\star$	13.00	12.28	10.95	8.74	6.53	3.36	352.3
$w_5^\star$	12.00	12.26	12.60	13.07	13.53	14.08	27.1
$w_6^\star$	8.00	11.71	16.42	22.57	28.73	34.77	54.2
$w_7^\star$	6.00	6.36	6.69	7.00	7.30	7.59	78.6
$w_8^\star$	5.00	1.86	0.00	0.00	0.00	0.00	426.7
$\sigma\left(w^\star \mid b\right)$	0.00	30.01	61.90	104.10	149.65	196.87	
$\mathcal{CI}\left(w\right)$	160.57	144.52	128.46	112.40	96.34	80.29	
$\mathcal{R}\left(w \mid b\right)$	0.00	10.00	20.00	30.00	40.00	50.00	

The reduction rate and weights are expressed in %, while the tracking error volatility is measured in bps.

Table 11.7: Optimal decarbonization portfolios (Example 42, order-statistic approach)

m	0	1	2	3	4	5	6	7	$\mathcal{CI}_i$
$w_1^\star$	20.00	20.40	22.35	26.46	0.00	0.00	0.00	0.00	100.5
$w_2^\star$	19.00	19.90	20.07	20.83	7.57	0.00	0.00	0.00	97.2
$w_3^\star$	17.00	17.94	21.41	0.00	0.00	0.00	0.00	0.00	250.4
$w_4^\star$	13.00	13.24	0.00	0.00	0.00	0.00	0.00	0.00	352.3
$w_5^\star$	12.00	12.12	12.32	12.79	13.04	14.26	18.78	100.00	27.1
$w_6^\star$	8.00	10.04	17.14	32.38	74.66	75.12	81.22	0.00	54.2
$w_7^\star$	6.00	6.37	6.70	7.53	4.73	10.62	0.00	0.00	78.6
$w_8^\star$	5.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	426.7
$\sigma\left(w^\star \mid b\right)$	0.00	0.37	1.68	2.25	3.98	4.04	4.30	15.41	
$\mathcal{CI}\left(w\right)$	160.57	145.12	113.48	73.78	55.08	52.93	49.11	27.10	
$\mathcal{R}\left(w \mid b\right)$	0.00	9.62	29.33	54.05	65.70	67.04	69.42	83.12	

The reduction rate, weights, and tracking error volatility are expressed in %.

Table 11.8: Optimal decarbonization portfolios (Example 42, naive approach)

m	0	1	2	3	4	5	6	7	$\mathcal{CI}_i$
$w_1^\star$	20.00	21.05	24.39	30.77	0.00	0.00	0.00	0.00	100.5
$w_2^\star$	19.00	20.00	23.17	29.23	42.22	0.00	0.00	0.00	97.2
$w_3^\star$	17.00	17.89	20.73	0.00	0.00	0.00	0.00	0.00	250.4
$w_4^\star$	13.00	13.68	0.00	0.00	0.00	0.00	0.00	0.00	352.3
$w_5^\star$	12.00	12.63	14.63	18.46	26.67	46.15	60.00	100.00	27.1
$w_6^\star$	8.00	8.42	9.76	12.31	17.78	30.77	40.00	0.00	54.2
$w_7^\star$	6.00	6.32	7.32	9.23	13.33	23.08	0.00	0.00	78.6
$w_8^\star$	5.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	426.7
$\sigma\left(w^\star \mid b\right)$	0.00	0.39	1.85	3.04	9.46	8.08	8.65	15.41	
$\mathcal{CI}\left(w\right)$	160.57	146.57	113.95	78.26	68.38	47.32	37.94	27.10	
$\mathcal{R}\left(w \mid b\right)$	0.00	8.72	29.04	51.26	57.41	70.53	76.37	83.12	

The reduction rate, weights, and tracking error volatility are expressed in %.

Figure 11.3: Efficient decarbonization frontier (Example 42)
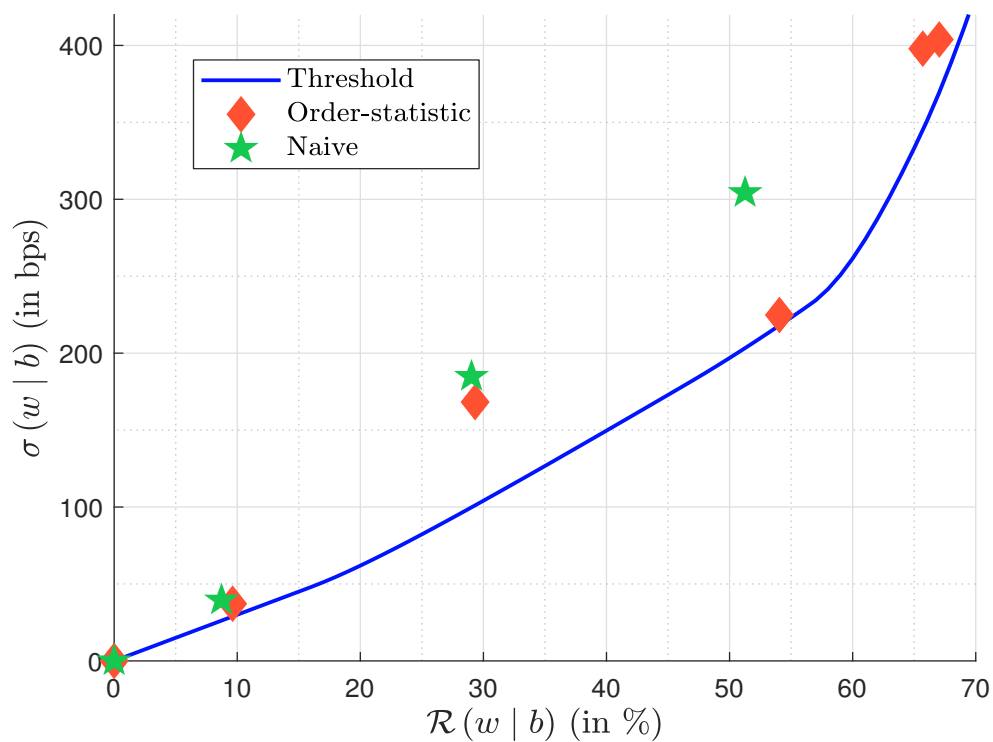
Figure 11.4: Efficient decarbonization frontier of the interpolated naive approach (Example 42)
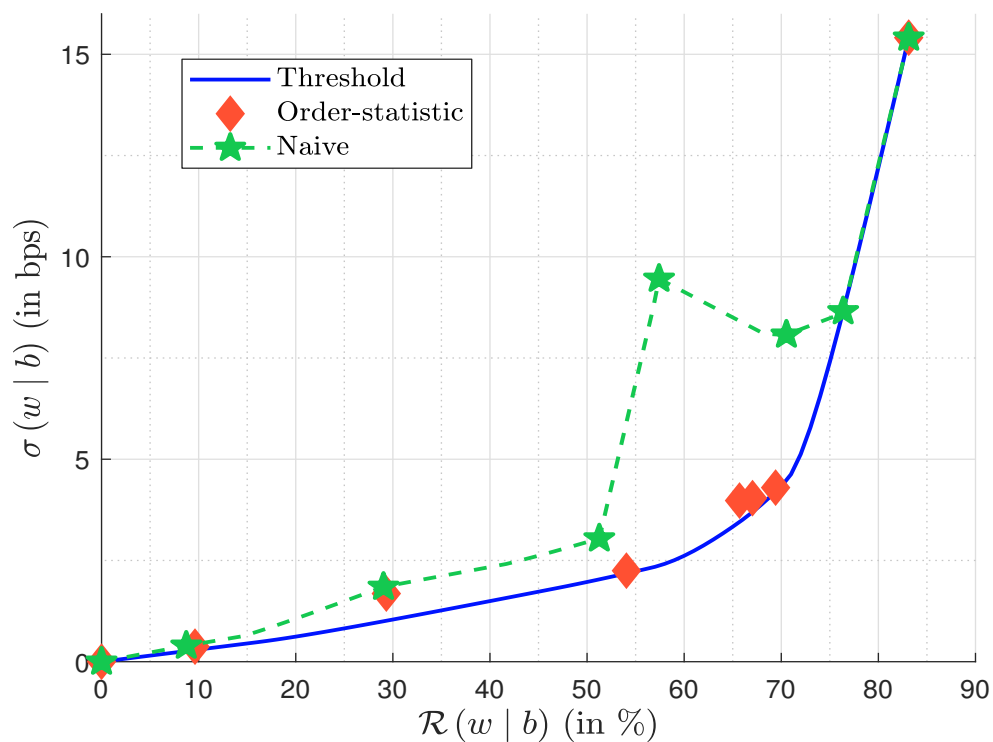
the tracking error volatility. In the case of the threshold approach, it is a continuous curve, which is not the case for the two other approaches.

In the case of the naive approach, we can extend the method to a given reduction rate $\mathcal{R}$. Let us denote by $w^\star_{(m)}$ the optimal portfolio when we exclude m assets. We consider the index m such that[21]:

$$\mathcal{R}\left(w^\star_{(m-1)} \mid b\right) \leq \mathcal{R} \leq \mathcal{R}\left(w^\star_{(m)} \mid b\right)$$

We define the portfolio $w^\star$ as the linear interpolation between $w^\star_{(m-1)}$ and $w^\star_{(m)}$:

$$w^\star = \alpha w^\star_{(m-1)} + (1 - \alpha)\, w^\star_{(m)}$$

where:

$$\alpha = \frac{\mathcal{R}\left(w^\star_{(m)} \mid b\right) - \mathcal{R}}{\mathcal{R}\left(w^\star_{(m)} \mid b\right) - \mathcal{R}\left(w^\star_{(m-1)} \mid b\right)}$$

We deduce that:

$$
\begin{aligned}
\mathcal{R}\left(w^\star \mid b\right) &= 1 - \frac{w^{\star\top}\boldsymbol{CI}}{b^\top\boldsymbol{CI}} \\
&= 1 - \frac{\left(\alpha w^\star_{(m-1)} + (1 - \alpha)\, w^\star_{(m)}\right)^\top \boldsymbol{CI}}{b^\top\boldsymbol{CI}} \\
&= (\alpha + (1 - \alpha)) - \left(\alpha \frac{w^{\star\top}_{(m-1)}\boldsymbol{CI}}{b^\top\boldsymbol{CI}} + (1 - \alpha)\frac{w^{\star\top}_{(m)}\boldsymbol{CI}}{b^\top\boldsymbol{CI}}\right) \\
&= \alpha\mathcal{R}\left(w^\star_{(m-1)} \mid b\right) + (1 - \alpha)\mathcal{R}\left(w^\star_{(m)} \mid b\right) \\
&= \mathcal{R}
\end{aligned}
$$

In Figure 11.4, we report the efficient decarbonization frontier when we apply the interpolated naive method to Exercise. We notice that the tracking error volatility is not necessarily an increasing function of the reduction rate. This is certainly the main drawback of the naive approach. The monotonicity property is also not always satisfied when we consider the order-statistic approach[22].

Bond portfolios

In the case of bond portfolios, we can use the same framework as equity portfolios except that we use the synthetic risk measure $\mathcal{R}\left(w \mid b\right)$ instead of the tracking error variance. We can implement the threshold or order-statistic approach, but the naive approach is not really appropriate since it may induce high deformation of the MD and DTS profiles.

Example 43 *We consider a debt-weighted bond index, which is composed of eight bonds. Their weights are equal to* 20%, 19%, 17%, 13%, 12%, 8%, 6% *and* 5%. *The carbon intensities (expressed in* $tCO_2e/\$$ mn*) are respectively equal to* 100.5, 97.2, 250.4, 352.3, 27.1, 54.2, 78.6 *and* 426.7. *To evaluate the risk of the portfolio, we use the modified duration which is respectively equal to* 3.1, 6.6, 7.2, 5, 4.7, 2.1, 8.1 *and* 2.6 *years, and the duration-times-spread factor, which is respectively equal to* 100, 155, 575, 436, 159, 145, 804 *and* 365 *bps. There are two sectors. Bonds #1, #3, #4 and #8 belong to* $\boldsymbol{S}ector_1$ *while Bonds #2, #5, #6 and #7 belong to* $\boldsymbol{S}ector_2$.

Results are given in Tables 11.9 and 11.10. We see the impact of the decarbonization constraint on the modified duration or the DTS factor on the portfolio.

---

[21]We have $\mathcal{R}\left(w^\star_{(0)}\right) = 0$ and $w^\star_{(0)} = b$.

[22]See Exercise 11.4.2 on page 632.

Table 11.9: Optimal decarbonization portfolios (Example 43, threshold approach)

$\mathcal{R}$	0	10	20	30	40	50	$\mathcal{CI}_i$
$w_1^\star$	20.00	21.62	23.93	26.72	30.08	33.44	100.5
$w_2^\star$	19.00	18.18	16.98	14.18	7.88	1.58	97.2
$w_3^\star$	17.00	18.92	21.94	22.65	16.82	11.00	250.4
$w_4^\star$	13.00	11.34	5.35	0.00	0.00	0.00	352.3
$w_5^\star$	12.00	13.72	16.14	21.63	33.89	46.14	27.1
$w_6^\star$	8.00	9.60	10.47	10.06	7.21	4.36	54.2
$w_7^\star$	6.00	5.56	5.19	4.75	4.11	3.48	78.6
$w_8^\star$	5.00	1.05	0.00	0.00	0.00	0.00	426.7
$\mathrm{AS}_{\boldsymbol{\mathcal{S}}ector}$	0.00	6.87	15.49	24.07	31.97	47.58	
$\mathrm{MD}(w)$	5.48	5.49	5.45	5.29	4.90	4.51	
$\mathrm{DTS}(w)$	301.05	292.34	282.28	266.12	236.45	206.78	
$\sigma_{\mathrm{AS}}(w \mid b)$	0.00	5.57	12.31	19.82	30.04	43.58	
$\sigma_{\mathrm{MD}}(w \mid b)$	0.00	0.01	0.04	0.17	0.49	0.81	
$\sigma_{\mathrm{DTS}}(w \mid b)$	0.00	8.99	19.29	35.74	65.88	96.01	
$\mathcal{CI}(w)$	160.57	144.52	128.46	112.40	96.34	80.29	
$\mathcal{R}(w \mid b)$	0.00	10.00	20.00	30.00	40.00	50.00	

The reduction rate, weights, and active share metrics are expressed in %, the MD metrics are measured in years, and the DTS metrics are computed in bps.

Table 11.10: Optimal decarbonization portfolios (Example 43, order-statistic approach)

m	0	1	2	3	4	5	6	7	$\mathcal{CI}_i$
$w_1^\star$	20.00	20.83	24.62	64.64	0.00	0.00	0.00	0.00	100.5
$w_2^\star$	19.00	18.60	18.13	21.32	3.32	0.00	0.00	0.00	97.2
$w_3^\star$	17.00	17.79	26.30	0.00	0.00	0.00	0.00	0.00	250.4
$w_4^\star$	13.00	14.53	0.00	0.00	0.00	0.00	0.00	0.00	352.3
$w_5^\star$	12.00	12.89	13.96	6.00	36.57	41.27	41.27	100.00	27.1
$w_6^\star$	8.00	9.74	11.85	0.00	60.11	58.73	58.73	0.00	54.2
$w_7^\star$	6.00	5.62	5.15	8.03	0.00	0.00	0.00	0.00	78.6
$w_8^\star$	5.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	426.7
$\mathrm{AS}_{\boldsymbol{\mathcal{S}}ector}$	0.00	5.78	19.72	49.00	76.68	80.00	80.00	88.00	
$\mathrm{MD}(w)$	5.48	5.52	5.54	4.77	3.27	3.17	3.17	4.70	
$\mathrm{DTS}(w)$	301.05	295.08	284.71	171.82	150.45	150.78	150.78	159.00	
$\sigma_{\mathrm{AS}}(w \mid b)$	0.00	5.73	17.94	50.85	66.96	68.63	68.63	95.33	
$\sigma_{\mathrm{MD}}(w \mid b)$	0.00	0.03	0.04	0.63	2.66	2.64	2.64	3.21	
$\sigma_{\mathrm{DTS}}(w \mid b)$	0.00	6.21	16.87	128.04	197.22	197.29	197.29	199.22	
$\mathcal{CI}(w)$	160.57	147.94	122.46	93.63	45.72	43.02	43.02	27.10	
$\mathcal{R}(w \mid b)$	0.00	7.87	23.74	41.69	71.53	73.21	73.21	83.12	

The reduction rate, weights and active share metrics are expressed in %, the MD metrics are measured in years, and the DTS metrics are computed in bps.

11.2.2 Sector-specific constraints

Sectors are very important in portfolio construction and allocation. For instance, performance attribution is generally done by distinguishing three components: sector allocation, asset selection and interaction effect. Since the goal of portfolio decarbonization is not to take active bets, this explains that most of allocation constraints concern sectors. Therefore, it is necessary to understand sector classification systems in asset management. In Box 11.1, we describe the GICS system, which is the most known and used sector taxonomy.

Sector scenario

The previous approach can be extended by considering one decarbonization scenario per sector instead of a global reduction rate:

$$\mathcal{CI}\left(w; \boldsymbol{Sector}_j\right) \leq \left(1 - \mathcal{R}_j\right)\mathcal{CI}\left(b; \boldsymbol{Sector}_j\right)$$

for $j = 1, \ldots, n_{\boldsymbol{Sector}}$. On page 553, we have seen that these constraints can be written as:

$$\left(\boldsymbol{s}_j \circ \left(\mathcal{CI} - \mathcal{CI}_j^{\star}\right)\right)^{\top} w \leq 0$$

where $\mathcal{CI}_j^{\star} = \left(1 - \mathcal{R}_j\right)\mathcal{CI}\left(b; \boldsymbol{Sector}_j\right)$. The QP form $Cw \leq D$ is then:

$$C = \begin{pmatrix} \left(\boldsymbol{s}_1 \circ \left(\mathcal{CI} - \mathcal{CI}_1^{\star}\right)\right)^{\top} \\ \vdots \\ \left(\boldsymbol{s}_j \circ \left(\mathcal{CI} - \mathcal{CI}_j^{\star}\right)\right)^{\top} \\ \vdots \\ \left(\boldsymbol{s}_{n_{\boldsymbol{Sector}}} \circ \left(\mathcal{CI} - \mathcal{CI}_{n_{\boldsymbol{Sector}}}^{\star}\right)\right)^{\top} \end{pmatrix}$$

and:

$$D = \begin{pmatrix} \left(1 - \mathcal{R}_1\right)\mathcal{CI}\left(b; \boldsymbol{Sector}_1\right) \\ \vdots \\ \left(1 - \mathcal{R}_j\right)\mathcal{CI}\left(b; \boldsymbol{Sector}_j\right) \\ \vdots \\ \left(1 - \mathcal{R}_{n_{\boldsymbol{Sector}}}\right)\mathcal{CI}\left(b; \boldsymbol{Sector}_{n_{\boldsymbol{Sector}}}\right) \end{pmatrix}$$

We can also apply these constraints on a subset of sectors and combine them with a global reduction constraint.

In Table 9.25 on page 530, we have seen that the reduction rates of the IEA NZE scenario depend on the sector. In particular, the Electricity sector must decarbonize itself faster than the other sectors. For 2030, using 2020 as the base year, the reduction rates are 56.9% for the Electricity sector, 18.8% for the Industry sector, 20.0% for the Transport sector, 36.7% for the Buildings sector and 52.4% for the other sectors. The global reduction rate is set to 36.6%. In Table 9.16 on page 513, we have reported the carbon intensity of the MSCI World index. It is equal to 163 tCO$_2$e/\$ mn if we consider the scope $1 + 2$ emissions. By 2030, we can then use the following global threshold:

$$\mathcal{CI}^{\star} = \left(1 - \mathcal{R}\right)\mathcal{CI}\left(b\right) = \left(1 - 0.366\right) \times 163 = 103.3 \text{ tCO}_2\text{e/\$ mn}$$

If we use the sector scenarios, we first have to map the IEA sectors on to the GICS sectors. For instance, the Utilities sector corresponds to the Electricity sector. Since the carbon intensity of this sector is equal to 1 794 tCO$_2$e/\$ mn, we can use the following specific threshold:

$$\mathcal{CI}_j^{\star} = \left(1 - \mathcal{R}_j\right)\mathcal{CI}\left(b; \boldsymbol{Sector}_j\right) = \left(1 - 0.569\right) \times 1\,794 = 773.2 \text{ tCO}_2\text{e/\$ mn}$$

Box 11.1: Global Industry Classification Standard (GICS®)

The Global Industry Classification Standard (GICS) is an industry taxonomy developed in 1999 by MSCI and Standard & Poor's (S&P) and based on a 4-level structure:

Level	1	2	3	4
Name	Sector	Industry group	Industry	Sub-industry
Code	AA	AABB	AABBCC	AABBCCDD

GICS is used as a basis for S&P and MSCI market indexes. The level codes are generated using 2-, 4-, 6- and 8-number classifications, whose root (AA, BB, CC and DD) is a multiple of ten $(10, 20, \ldots, 90)$. For instance, the Energy sector has the following structure:

- Energy (1010)
  - Energy (1010)
    * Energy equipment & services (101010)
      · Oil & gas drilling (10101010): drilling contractors or owners of drilling rigs that contract their services for drilling wells;
      · Oil & gas equipment & services (10101020): manufacturers of equipment, including drilling rigs and equipment, and providers of supplies and services to companies involved in the drilling, evaluation and completion of oil and gas wells;
    * Oil, gas & consumable fuels (101020)
      · Integrated oil & gas (10102010): integrated oil companies engaged in the exploration & production of oil and gas;
      · Oil & gas exploration & production (10102020): companies engaged in the exploration and production of oil and gas not classified elsewhere;
      · Oil & gas refining & marketing (10102030): companies engaged in the refining and marketing of oil, gas and/or refined products;
      · Oil & gas storage & transportation (10102040): companies engaged in the storage and/or transportation of oil, gas and/or refined products;
      · Coal & consumable fuels (10102050): companies primarily involved in the production and mining of coal, related products and other consumable fuels related to the generation of energy.

The Energy sector has then one industry group, two industries and seven sub-industries. Here is the number of items per sector:

Name	Sector	Industry Group	Industry	Sub-industry
Energy	1	1	2	7
Materials	1	1	5	17
Industrials	1	3	14	27
Consumer Discretionary	1	4	10	27
Consumer Staples	1	3	6	12
Health Care	1	2	6	10
Financials	1	3	6	18
Information Technology	1	3	6	12
Communication Services	1	2	5	10
Utilities	1	1	5	6
Real Estate	1	2	10	17
Total	11	25	75	163

Finally, we report the level 3 classification in Tables 11.11 and 11.12.

Table 11.11: GICS level 3 classification (March 2023)

Sector		Industry Group		Industry	
10	Energy	1010	Energy	101010	Energy Equipment & Services
				101020	Oil, Gas & Consumable Fuels
15	Materials	1510	Materials	151010	Chemicals
				151020	Construction Materials
				151030	Containers & Packaging
				151040	Metals & Mining
				151050	Paper & Forest Products
20	Industrials	2010	Capital Goods	201010	Aerospace & Defense
				201020	Building Products
				201030	Construction & Engineering
				201040	Electrical Equipment
				201050	Industrial Conglomerates
				201060	Machinery
				201070	Trading Companies & Distributors
		2020	Commercial & Professional Services	202010	Commercial Services & Supplies
				202020	Professional Services
		2030	Transportation	203010	Air Freight & Logistics
				203020	Passenger Airlines
				203030	Marine Transportation
				203040	Ground Transportation
				203050	Transportation Infrastructure
25	Consumer Discretionary	2510	Automobiles & Components	251010	Automobile Components
				251020	Automobiles
		2520	Consumer Durables & Apparel	252010	Household Durables
				252020	Leisure Products
				252030	Textiles, Apparel & Luxury Goods
		2530	Consumer Services	253010	Hotels, Restaurants & Leisure
				253020	Diversified Consumer Services
		2550	Consumer Discretionary Distribution & Retail	255010	Distributors
				255030	Broadline Retail
				255040	Specialty Retail
30	Consumer Staples	3010	Consumer Staples Dist. & Retail	301010	Consumer Staples Distribution & Retail
		3020	Food, Beverage & Tobacco	302010	Beverages
				302020	Food Products
				302030	Tobacco
		3030	Household & Personal Products	303010	Household Products
				303020	Personal Care Products

Table 11.12: GICS level 3 classification (March 2023)

Sector		Industry Group		Industry	
35	Health Care	3510	Health Care Equipment & Services	351010	Health Care Equipment & Supplies
				351020	Health Care Providers & Services
				351030	Health Care Technology
		3520	Pharmaceuticals, Biotechnology & Life Sciences	352010	Biotechnology
				352020	Pharmaceuticals
				352030	Life Sciences Tools & Services
40	Financials	4010	Banks	401010	Banks
		4020	Financial Services	402010	Financial Services
				402020	Consumer Finance
				402030	Capital Markets
				402040	Mortgage Real Estate Investment Trusts (REITs)
		4030	Insurance	403010	Insurance
45	Information Technology	4510	Software & Services	451020	IT Services
				451030	Software
		4520	Technology Hardware & Equipment	452010	Communications Equipment
				452020	Technology Hardware, Storage & Peripherals
				452030	Electronic Equipment, Instruments & Components
		4530	Semiconductors & SC. Equip.	453010	Semiconductors & Semiconductor Equipment
50	Communication Services	5010	Telecommunication Services	501010	Diversified Telecommunication Services
				501020	Wireless Telecommunication Services
		5020	Media & Entertainment	502010	Media
				502020	Entertainment
				502030	Interactive Media & Services
55	Utilities	5510	Utilities	551010	Electric Utilities
				551020	Gas Utilities
				551030	Multi-Utilities
				551040	Water Utilities
				551050	Independent Power and Renewable Electricity Producers
60	Real Estate	6010	Equity Real Estate Investment Trusts (REITs)	601010	Diversified REITs
				601025	Industrial REITs
				601030	Hotel & Resort REITs
				601040	Office REITs
				601050	Health Care REITs
				601060	Residential REITs
				601070	Retail REITs
				601080	Specialized REITs
		6020	Real Estate Management & Dev.	602010	Real Estate Management & Development

Table 11.13 shows the different sector thresholds for the different scope emissions. For that, we have assumed the following mapping between the GICS sectors and the IEA sectors: Communication Services → Other, Consumer Discretionary → Other, Consumer Staples → Other, Energy → Electricity, Financials → Other, Health Care → Other, Industrials → Industry, Information Technology → Other, Materials → Buildings, Real Estate → Buildings and Utilities → Electricity. This mapping is simplistic since the Industrials GICS sector is composed of three industry groups: capital goods, commercial & professional services and transportation. Therefore, we can obtain a more accurate mapping if we use the GICS level II classification.

Table 11.13: Carbon intensity and threshold in tCO$_2$e/\$ mn by GICS sector (MSCI World, 2030)

Sector	$\mathcal{CI}\,(b; \mathcal{S}ector_j)$				$\mathcal{R}_j$	$\mathcal{CI}_j^\star$			
	$\mathcal{SC}_1$	$\mathcal{SC}_{1-2}$	$\mathcal{SC}_{1-3}^{\mathrm{up}}$	$\mathcal{SC}_{1-3}$	(in %)	$\mathcal{SC}_1$	$\mathcal{SC}_{1-2}$	$\mathcal{SC}_{1-3}^{\mathrm{up}}$	$\mathcal{SC}_{1-3}$
Communication Services	2	28	134	172	52.4	1	13	64	82
Consumer Discretionary	23	65	206	590	52.4	11	31	98	281
Consumer Staples	28	55	401	929	52.4	13	26	191	442
Energy	632	698	1 006	6 823	56.9	272	301	434	2 941
Financials	13	19	52	244	52.4	6	9	25	116
Health Care	10	22	120	146	52.4	5	10	57	70
Industrials	111	130	298	1 662	18.8	90	106	242	1 350
Information Technology	7	23	112	239	52.4	3	11	53	114
Materials	478	702	1 113	2 957	36.7	303	445	704	1 872
Real Estate	22	101	167	571	36.7	14	64	106	361
Utilities	1 744	1 794	2 053	2 840	56.9	752	773	885	1 224
MSCI World	130	163	310	992	36.6	82	103	196	629

Source: MSCI (2022), Trucost (2022) & Author's calculations.

Remark 95 *The mapping between the different classifications is not always obvious. Indeed, asset owners and managers generally use the Global Industry Classification System (GICS), while the European Union and the United Nations have their own classification systems: the Statistical Classification of Economic Activities in the European Community (NACE) and the International Standard Industrial Classification (ISIC). Le Guenedal et al. (2022) perform a mapping between the four classification systems: GICS, NACE, ISIC and IEA. At the more granular level, they obtain a perfect match only in 70% of cases.*

Sector and weight deviation constraints

Equity portfolio If we consider problem (11.9), we notice that the set of weight constraints are $\Omega_0 \cap \Omega$. The first set of constraint $\Omega_0 = \left\{ w : \mathbf{1}_n^\top w = 1, \mathbf{0}_n \leq w \leq \mathbf{1}_n \right\}$ is necessary to obtain a long-only portfolio, but is not sufficient to define a realistic portfolio because it can be concentrated in few number of assets. Therefore, portfolio managers and index providers generally impose some additional constraints to obtain a diversified portfolio:

1. To control the weight deviation between portfolio w and benchmark b, we can use:

$$\Omega := \mathcal{C}_1\left(m_w^-, m_w^+\right) = \left\{ w : m_w^- b \leq w \leq m_w^+ b \right\}$$

   where $m_w^- \in [0, 1[$ and $m_w^+ \in [1, \infty[$. In this case, the portfolio's weight w_i can only deviate from the benchmark's weight b_i by lower and upper ratios m_w^- and m_w^+. Typical figures are $m_w^- = 1/2$ and $m_w^+ = 2$.

2. Another approach consists in controlling the sector deviations. In this case, we can use a relative deviation allowance[23]:

$$\Omega := \mathcal{C}_2\left(m_s^-, m_s^+\right) = \left\{\forall j : m_s^- \sum_{i \in \boldsymbol{S}ector_j} b_i \leq \sum_{i \in \boldsymbol{S}ector_j} w_i \leq m_s^+ \sum_{i \in \boldsymbol{S}ector_j} b_i\right\}$$

where $m_s^- \in [0, 1[$ and $m_s^+ \in [1, \infty[$. Typical figures are $m_s^- = 1/2$, $m_s^+ = 2$ and $\delta_s^+ = 5\%$.

In the sequel, we define 4 sets of constraints: $\mathcal{C}_0$ imposes long-only constraints, $\mathcal{C}_1\left(m_w^-, m_w^+\right)$ adds stock weight constraints, $\mathcal{C}_2\left(m_s\right)$ adds sector relative allocation constraints with $m_s^- = 1/m_s$ and $m_s^+ = m_s$, and $\mathcal{C}_3\left(m_w^-, m_w^+, m_s\right) = \mathcal{C}_1\left(m_w^-, m_w^+\right) \cap \mathcal{C}_2\left(m_s\right)$ combines $\mathcal{C}_1$ and $\mathcal{C}_2$.

Bond portfolio In the case of bond portfolio optimization, the current problem is already highly constrained at the sector level, because the objective function controls MD and DTS tracking risks at the sector level. Therefore, we need to use another constraints $\Omega_0 \cap \Omega$, where the first set $\Omega_0 := \mathcal{C}_0 = \left\{w : \mathbf{1}_n^\top w = 1, \mathbf{0}_n \leq w \leq \mathbf{1}_n\right\}$ defines a long-only portfolio and the second set Ω controls the deviation of risk metrics between portfolio w and benchmark b:

1. We can neutralize the modified duration at the portfolio level:

$$\Omega := \mathcal{C}_1' = \{w : \text{MD}(w) = \text{MD}(b)\} = \left\{w : \sum_{i=1}^n (x_i - b_i)\,\text{MD}_i = 0\right\}$$

2. We can neutralize the duration-times-spread factor at the portfolio level:

$$\Omega := \mathcal{C}_2' = \{w : \text{DTS}(w) = \text{DTS}(b)\} = \left\{w : \sum_{i=1}^n (x_i - b_i)\,\text{DTS}_i = 0\right\}$$

3. We can constraint the portfolio to have the same weight per maturity bucket as the benchmark:

$$\Omega := \mathcal{C}_3' = \left\{w : \sum_{i \in \boldsymbol{B}ucket_j} (x_i - b_i) = 0\right\}$$

where $\boldsymbol{B}ucket_j$ is the jth maturity bucket. Typical maturity buckets (expressed in years) are: 0–1, 1–3, 3–5, 5–7, 7–10 and 10+.

4. Instead of maturity buckets, we can consider rating categories:

$$\Omega := \mathcal{C}_4' = \left\{w : \sum_{i \in \mathcal{R}_j} (x_i - b_i) = 0\right\}$$

where $\mathcal{R}_j$ is the jth rating category. For investment grade bonds, typical rating categories are: AAA–AA (AAA, AA+, AA and AA−), A (A+, A and A−) and BBB (BBB+, BBB, BBB−).

---

[23]An alternative approach is to use an absolute deviation allowance:

$$\Omega = \left\{\forall j : \left|\sum_{i \in \boldsymbol{S}ector_j} (w_i - b_i)\right| \leq \delta_s^+\right\}$$

where $\delta_s^+ \in [0, 1]$.

HCIS constraint

The CTB and PAB labels require that the exposure to sectors highly exposed to climate change is at least equal to the exposure in the investment universe. TEG (2019a) distinguishes two types of sectors:

1. High climate impact sectors (HCIS);

2. Low climate impact sectors (LCIS).

According to TEG (2019a), the first category is made up of "*sectors that are key to the low-carbon transition*". They correspond to the following NACE classes: A. Agriculture, Forestry, and Fishing; B. Mining and Quarrying; C. Manufacturing; D. Electricity, Gas, Steam, and Air Conditioning Supply; E. Water Supply; Sewerage, Waste Management, and Remediation Activities; F. Construction; G. Wholesale and Retail Trade; Repair of Motor Vehicles and Motorcycles; H. Transportation and Storage; L. Real Estate Activities. Let $\mathcal{HCIS}(w) = \sum_{i \in \text{HCIS}} w_i$ be the HCIS weight of portfolio w. In order to comply with CTB and PAB labels, we must verify that:

$$\mathcal{HCIS}(w) \geq \mathcal{HCIS}(b) \tag{11.12}$$

TEG (2019b, Appendix B, pages 26-170) has published a mapping between the NACE classes and several sector classification structures: BICS (Bloomberg), GICS (MSCI and S&P), ICB (FTSE) and TRBC (Refinitiv). In the case of the GICS taxonomy, about 70% of sub-industries are classified as high climate impact sectors. The HCIS constraint has been criticized because this figure is very high. This would mean that almost all activities are critical for building a low-carbon economy. Therefore, only two sectors are classified in low climate impact sectors (Communication Services and Financials), but more than half of the Health Care and Information Technology sub-industries are viewed as high climate impact sectors. The original idea of the HCIS constraint was to continue financing the sectors that are essential for reaching a low-carbon economy (e.g., Energy and Utilities) and at the same time promoting investments in green issuers instead of brown issuers in these sectors. Nevertheless, the constraint (11.12) is not very restrictive with this broad HCIS measure. Moreover, this constraint encourages substitutions between sectors or industries and not substitutions between issuers within a same sector. Therefore, the trade-off is not necessarily between green electricity and brown electricity, but for example between electricity generation and health care equipment.

In Table 11.14, we report the weight and the carbon intensity of each sector when applying the HCIS constraint to the MSCI World index. The carbon intensity of the MSCI World index is equal to 992 tCO$_2$e/\$ mn when we consider the scope $\mathcal{SC}_{1-3}$ and 1 498 after the HCIS filter. We verify that the HCIS investment universe has a higher carbon intensity than the reference universe. Nevertheless, the HCIS universe represents 59.79% of the MSCI World index, which is a very high figure. This means that more than half of stocks have a high climate impact. This explains that the HCIS constraint is not really used in portfolio allocation. Indeed, it is difficult to justify that 70% of the Health Care sector, 50% of the Information Technology sector but 0% of the Financials sector are classified as high climate impact sectors.

11.2.3 Empirical results

We apply the previous framework to the MSCI World index and the ICE BofA Global Corporate index. The numerical results are those obtained by Barahhou *et al.* (2022).

Table 11.14: Weight and carbon intensity when applying the HCIS filter (MSCI World, June 2022)

Sector	Index b_j	HCIS b'_j	$\mathcal{SC}_1$ $\mathcal{CI}$	$\mathcal{CI}'$	$\mathcal{SC}_{1-2}$ $\mathcal{CI}$	$\mathcal{CI}'$	$\mathcal{SC}^{\mathrm{up}}_{1-3}$ $\mathcal{CI}$	$\mathcal{CI}'$	$\mathcal{SC}_{1-3}$ $\mathcal{CI}$	$\mathcal{CI}'$
Communication Services	7.58	0.00	2		28		134		172	
Consumer Discretionary	10.56	8.01	23	14	65	31	206	189	590	462
Consumer Staples	7.80	7.80	28	28	55	55	401	401	929	929
Energy	4.99	4.99	632	632	698	698	1 006	1 006	6 823	6 823
Financials	13.56	0.00	13		19		52		244	
Health Care	14.15	9.98	10	13	22	26	120	141	146	177
Industrials	9.90	7.96	111	132	130	151	298	332	1 662	1 921
Information Technology	21.08	10.67	7	12	23	30	112	165	239	390
Materials	4.28	4.28	478	478	702	702	1 113	1 113	2 957	2 957
Real Estate	2.90	2.90	22	22	101	101	167	167	571	571
Utilities	3.21	3.21	1 744	1 744	1 794	1 794	2 053	2 053	2 840	2 840
MSCI World	100.00	59.79	130	210	163	252	310	458	992	1 498

The weights are expressed in %. Column b_j represents the sector weight in the benchmark, while column b'_j indicates the sector weight when applying the HCIS filter. The carbon intensities are measured in tCO$_2$e/\$ mn. Column $\mathcal{CI}$ is the WACI measure of the sector, while column $\mathcal{CI}'$ takes into account the HCIS filter.

Source: MSCI (2022), Trucost (2022) & Author's calculations.

Equity portfolios

In Table 11.13 on page 577, we have reported the carbon intensity of the MSCI World index and its sectors. We obtain 130 tCO$_2$e/\$ mn for scope $\mathcal{SC}_1$, 163 tCO$_2$e/\$ mn if we include scope 2, 310 tCO$_2$e/\$ mn if we add the upstream scope 3, and finally 992tCO$_2$e/\$ mn if we consider the full scope $\mathcal{SC}_{1-3}$. We observe a high discrepancy between sectors. Low-carbon sectors are Communication Services, Financials, Health Care and Information Technology, whereas high-carbon sectors are Energy, Materials and Utilities. We foresee that decarbonizing a portfolio implies reducing the exposure to high-carbon sectors and increasing the exposure to low-carbon sectors. When imposing sector constraints, for instance sector neutrality, the trade-off will be between issuers of the same sector. In Figures 11.5 and 11.6, we show the empirical distribution of carbon intensities for each sector (logarithmic scale). Portfolio decarbonization consists in underweighting the issuers with high carbon intensities and overweigthing the issuers with high carbon intensities. The number of outliers or extreme cases depends on the scope. Therefore, portfolio decarbonization will be more or less difficult depending on the definition of scope emissions.

Barahhou *et al.* (2022) consider the basic optimization problem:

$$
\begin{aligned}
w^\star &= \arg\min \frac{1}{2}(w-b)^\top \Sigma (w-b) \\
\text{s.t.} &\quad \begin{cases} \mathcal{CI}(w) \le (1-\mathcal{R})\mathcal{CI}(b) \\ w \in \Omega_0 \cap \Omega \end{cases}
\end{aligned}
$$

and measure the impact of constraints $\Omega_0 \cap \Omega$ on the tracking error volatility. Figure 11.7 corresponds to the $\mathcal{C}_0$ constraint. The tracking risk increases when we include scope 2 or upstream scope 3, whereas downstream scope 3 reduces it because of its large dispersion. Sector allocation is given in Table 11.15. We observe that portfolio decarbonization is a strategy that is long on Financials sector and short on Energy, Materials and Utilities sectors.

Figure 11.5: Boxplot of carbon intensity per sector (MSCI World, June 2022, scope $\mathcal{SC}_{1-2}$)

Figure 11.6: Boxplot of carbon intensity per sector (MSCI World, June 2022, scope $\mathcal{SC}_{1-3}$)

Figure 11.7: Impact of the carbon scope on the tracking error volatility (MSCI World, June 2022, $\mathcal{C}_0$ constraint)

Source: MSCI (2022), Trucost (2022) & Barahhou *et al.* (2022).

Table 11.15: Sector allocation in % (MSCI World, June 2022, $\mathcal{C}_0$ constraint, scope $\boldsymbol{\mathcal{SC}}_{1-3}$)

Sector	Index	Reduction rate $\mathcal{R}$						
		30%	40%	50%	60%	70%	80%	90%
Communication Services	7.58	7.95	8.15	8.42	8.78	9.34	10.13	12.27
Consumer Discretionary	10.56	10.69	10.69	10.65	10.52	10.23	9.62	6.74
Consumer Staples	7.80	7.80	7.69	7.48	7.11	6.35	5.03	1.77
Energy	4.99	4.14	3.65	3.10	2.45	1.50	0.49	0.00
Financials	13.56	14.53	15.17	15.94	16.90	18.39	20.55	28.62
Health Care	14.15	14.74	15.09	15.50	16.00	16.78	17.77	17.69
Industrials	9.90	9.28	9.01	8.71	8.36	7.79	7.21	6.03
Information Technology	21.08	21.68	22.03	22.39	22.88	23.51	24.12	24.02
Materials	4.28	3.78	3.46	3.06	2.56	1.85	1.14	0.24
Real Estate	2.90	3.12	3.27	3.41	3.57	3.72	3.71	2.51
Utilities	3.21	2.28	1.79	1.36	0.90	0.54	0.24	0.12

Source: MSCI (2022), Trucost (2022) & Barahhou *et al.* (2022).

If we impose the classical weight constraint $\mathcal{C}_1 (1/3, 3)$, which is very popular in index portfolio management, we observe a high increase in the tracking error volatility (Panel 1, Figure 11.8). Moreover, we generally have no solution when $\mathcal{R} > 60\%$. The issue comes from the lower bound, which is way to narrow. Indeed, portfolio decarbonization is an exclusion process. By imposing a lower bound, we then limit portfolio decarbonization. For instance, we obtain similar results between constraint $\mathcal{C}_1 (0, 3)$ and constraint $\mathcal{C}_0$ (Panel 2, Figure 11.8). Nevertheless, we must be careful when choosing m_w^+, because a low value can lead to infeasible solutions. For instance, this is the case of constraint $\mathcal{C}_1 (0, 1.25)$, as shown in Panel 4, Figure 11.8.

The impact of sector constraints is less important than the impact of weight constraints. For instance, constraint $\mathcal{C}_2 (2)$ does not increase the tracking error volatility with respect to constraint $\mathcal{C}_0$ (Panel 1, Figure 11.9). If we impose sector neutrality (constraint $\mathcal{C}_2 (1)$), we observe a small increase of the tracking risk. Indeed, for low reduction rates (less than 50%), the tracking error volatility remains below 40 bps (Panel 2, Figure 11.9). At first sight, it may be surprising that weight constraints are more binding than sector constraints. Indeed, we generally consider that the sector contribution is greater than the idiosyncratic contribution. Therefore, we expect that the inter-class dispersion largely dominates the intra-class variance. Nevertheless, this viewpoint is biased because it considers homogeneous sectors. In our case, we use level 1 of the GICS classification. The concept of sector is then very heterogeneous. Within a particular sector, we can have low-carbon and high-carbon issuers as we have seen in Figures 11.5 and 11.6.

The combination of weight and sectoral constraints is a more difficult exercise as shown in the bottom panels in Figure 11.9. This is why portfolio managers generally use less restrictive rules. Thus, $\mathcal{C}_3 (0, 10, 2)$ constraint has became the standard approach in the ETF market when building climate benchmarks. As shown in Figure 11.10, the impact of this constraint is relatively low.

Figure 11.8: Impact of $\mathcal{C}_1$ constraint on the tracking error volatility (MSCI World, June 2022)

Figure 11.9: Impact of $\mathcal{C}_2$ and $\mathcal{C}_3$ constraints (MSCI World, June 2022)

Source: MSCI (2022), Trucost (2022) & Barahhou *et al.* (2022).

Figure 11.10: Tracking error volatility with $\mathcal{C}_3\,(0, 10, 2)$ constraint (MSCI World, June 2022)

Source: MSCI (2022), Trucost (2022) & Barahhou *et al.* (2022).

The efficiency of the order-statistic and naive approaches depends on the distribution of carbon intensities. For instance, if the carbon intensity of the index is concentrated in one stock, then removing this stock reduces dramatically the carbon footprint of the portfolio. In Figure 11.11, we have reported the relative carbon footprint contribution of the m worst performing assets:

$$\mathcal{CFC}^{(m,n)} = \frac{\sum_{i=1}^{n} \mathbb{1}\left\{\mathcal{CI}_i \geq \mathcal{CI}^{(m,n)}\right\} \cdot b_i \mathcal{CI}_i}{\mathcal{CI}(b)}$$

where $\mathcal{CI}^{(m,n)} = \mathcal{CI}_{n-m+1:n}$ is the $(n-m+1)$-th order statistic. By construction, $\mathcal{CFC}^{(m,n)}$ is an increasing function with respect to m and we have the following properties: $\mathcal{CFC}^{(0,n)} = 0$ and $\mathcal{CFC}^{(n,n)} = 1$. The fifty worst performing stocks[24] represent respectively about 50% and 40% of $\boldsymbol{\mathcal{SC}}_1$ and $\boldsymbol{\mathcal{SC}}_{1-3}$ carbon intensities of the MSCI World index. The previous analysis can be extended by considering another metric to perform the rank ordering, for instance the absolute contribution of stock i:

$$\mathcal{CFC}^{(m,n)} = \frac{\sum_{i=1}^{n} \mathbb{1}\left\{\mathcal{CIC}_i \geq \mathcal{CIC}^{(m,n)}\right\} \cdot b_i \mathcal{CI}_i}{\mathcal{CI}(b)}$$

where $\mathcal{CIC}_i = b_i \mathcal{CI}_i$ and $\mathcal{CIC}^{(m,n)} = \mathcal{CIC}_{n-m+1:n}$. In this approach, we do not consider $\mathcal{CI}_i$ but the product $b_i \mathcal{CI}_i$ to define the exclusion process. Results are given in Figure 11.12. The 50 worst performing stocks represent then 70% and 55% of $\boldsymbol{\mathcal{SC}}_1$ and $\boldsymbol{\mathcal{SC}}_{1-3}$ carbon intensities.

Figure 11.11: Carbon footprint contribution $\mathcal{CFC}^{(m,n)}$ in % (MSCI World, June 2022, first approach)

Source: MSCI (2022), Trucost (2022) & Author's calculations.

---

[24]The MSCI World index was composed of $1\,513$ stocks in June 2022.

Figure 11.12: Carbon footprint contribution $\mathcal{CFC}^{(m,n)}$ in % (MSCI World, June 2022, second approach)

Source: MSCI (2022), Trucost (2022) & Author's calculations.

Table 11.16: Carbon footprint contribution $\mathcal{CFC}^{(m,n)}$ in % (MSCI World, June 2022, second approach, $\boldsymbol{\mathcal{SC}}_{1-3}$)

Sector	m							
	1	5	10	25	50	75	100	200
Communication Services						0.44	0.44	0.73
Consumer Discretionary				0.78	1.37	2.44	2.93	4.28
Consumer Staples		2.46	2.46	2.46	3.75	4.44	4.92	5.62
Energy		9.61	17.35	23.78	29.56	31.78	33.02	33.89
Financials						0.72	1.53	1.88
Health Care							0.21	0.37
Industrials			2.16	5.59	7.13	8.70	9.48	13.05
Information Technology				0.98	1.58	1.94	2.15	3.30
Materials	4.08	4.08	4.08	5.81	7.31	8.81	9.59	10.75
Real Estate					0.77	0.77	0.77	0.85
Utilities				0.81	3.20	3.89	5.24	7.98
Total	4.08	16.15	26.06	40.21	54.66	63.94	70.29	82.70

Source: MSCI (2022), Trucost (2022) & Author's calculations.

Table 11.17: Weight contribution $\mathcal{WC}^{(m,n)}$ in % (MSCI World, June 2022, second approach, $\boldsymbol{\mathcal{SC}}_{1-3}$)

Sector	b_j (in %)	m							
		1	5	10	25	50	75	100	200
Communication Services	7.58						0.08	0.08	3.03
Consumer Discretionary	10.56				0.58	1.79	2.44	4.51	5.89
Consumer Staples	7.80		0.70	0.70	0.70	1.90	2.50	2.84	3.84
Energy	4.99		1.71	2.25	2.96	3.62	3.99	4.33	4.65
Financials	13.56						0.74	1.17	2.33
Health Care	14.15							0.95	1.34
Industrials	9.90			0.06	0.32	0.70	0.96	1.20	4.12
Information Technology	21.08				0.16	4.70	8.42	8.78	11.62
Materials	4.28	0.29	0.29	0.29	0.47	0.88	1.10	1.40	1.87
Real Estate	2.90					0.05	0.05	0.05	0.23
Utilities	3.21				0.31	0.86	1.04	1.31	2.33
Total		0.29	2.71	3.30	5.49	14.50	21.32	26.63	41.24

Source: MSCI (2022), Trucost (2022) & Author's calculations.

In Table 11.16, we have reported the relative carbon intensity contribution $\mathcal{CFC}^{(m,n)}$ for several values of m. For example, one stock contributes to 4.08% of the $\boldsymbol{\mathcal{SC}}_{1-3}$ carbon intensity[25], while the fifty worst performing stocks have a contribution of 54.66%. Table 11.16 also shows the sector allocation of $\mathcal{CFC}^{(m,n)}$. We verify that the Energy sector is the most important contributor and represents about 50% on average when $m \leq 100$. In Table 11.17, we calculate the weight contribution:

$$\mathcal{WC}^{(m,n)} = \sum_{i=1}^{n} \mathbb{1} \left\{ \mathcal{CIC}_i \geq \mathcal{CIC}^{(m,n)} \right\} \cdot b_i$$

We notice that the fifty worst performing stocks represent then 14.50% of the MSCI World index in terms of allocation, but 54.66% of its carbon footprint[26].

Remark 96 *The previous figures highly depend on the scope definition and the ordering approach. For instance, if we consider scope $\boldsymbol{\mathcal{SC}}_{1-2}$, the fifty worst performing stocks represent then 12.65% of the MSCI World index in terms of allocation, but 60.58% of its carbon footprint. If we consider the first ordering approach based on the carbon intensity $\boldsymbol{\mathcal{CI}}_i$ and not the absolute contribution $\mathcal{CIC}_i$, the figures become respectively 2.57% and 45.34%.*

We implement the order-statistic optimization problem:

$$\begin{aligned} w^{\star} &= \arg \min \frac{1}{2} (w - b)^{\top} \Sigma (w - b) \\ &\text{s.t.} \quad \begin{cases} \mathbf{1}_n^{\top} w = 1 \\ \mathbf{0}_n \leq w \leq w^{(m,n)} \end{cases} \end{aligned}$$

where the upper bound $w^{(m,n)}$ is equal to $\mathbb{1} \left\{ \boldsymbol{\mathcal{CI}} < \boldsymbol{\mathcal{CI}}^{(m,n)} \right\}$ for the first ordering approach and $\mathbb{1} \left\{ \mathcal{CIC} < \mathcal{CIC}^{(m,n)} \right\}$ for the second ordering approach[27]. We compare the optimization method

---

[25]We remind that the $\boldsymbol{\mathcal{SC}}_{1-3}$ carbon intensity of the MSCI World index is equal to 992 tCO$_2$e/\$ mn.

[26]If we consider the first approach, 3.21% of the allocation explains 39.15% of the carbon footprint.

[27]We have $\mathcal{CIC} = (\mathcal{CIC}_1, \ldots, \mathcal{CIC}_n)$ where $\mathcal{CIC}_i = b_i \boldsymbol{\mathcal{CI}}_i$.

Figure 11.13: Tracking error volatility (MSCI World, June 2022, $\mathcal{SC}_{1-3}$, first ordering method)

Figure 11.14: Tracking error volatility (MSCI World, June 2022, $\mathcal{SC}_{1-3}$, second ordering method)

with the naive method:

$$w_i^\star = \frac{e_i b_i}{\sum_{k=1}^n e_k b_k}$$

where e_i is defined as $\mathbb{1}\left\{\mathcal{CI}_i < \mathcal{CI}^{(m,n)}\right\}$ for the first ordering approach and $\mathbb{1}\left\{\mathcal{CIC}_i < \mathcal{CIC}^{(m,n)}\right\}$ for the second ordering approach. Results are given in Figures 11.13 and 11.14. Compared to the threshold method, the order-statistic and naive solutions are less efficient with a higher tracking error volatility. We also observe that the second ordering approach is not robust, because it may remove stocks, whose carbon intensity contribution is mainly explained by their weights.

Bond portfolios

In the case of corporate bonds, Barahhou *et al.* (2022) solve the following optimization problem:

$$
w^\star = \arg\min \frac{1}{2} \sum_{i=1}^n |w_i - b_i| + 50 \sum_{j=1}^{n_{\boldsymbol{Sector}}} \left| \sum_{i \in \boldsymbol{Sector}_j} (w_i - b_i) \, \mathrm{DTS}_i \right|
$$

$$
\text{s.t.} \quad \begin{cases} \mathcal{CI}(w) \le (1 - \mathcal{R}) \mathcal{CI}(b) \\ w \in \mathcal{C}_0 \cap \mathcal{C}_1' \cap \mathcal{C}_3' \cap \mathcal{C}_4' \end{cases}
$$

We remind that $\mathcal{C}_1'$ constraint neutralizes the modified duration at the portfolio level, whereas $\mathcal{C}_3'$ and $\mathcal{C}_4'$ constraints requires the portfolio to have the same weights as the benchmark per maturity bucket and rating category. The tracking risk measures $\mathcal{D}_{\mathrm{AS}}(w \mid b)$ and $\mathcal{D}_{\mathrm{DTS}}(w \mid b)$ are reported in Figures 11.15 and 11.16. We observe that the tracking risk is low when we consider the DTS component, whereas it is significant when we focus on the active share component.

Table 11.18: Sector allocation in % (ICE Global Corp., June 2022, scope $\boldsymbol{SC}_{1-3}$)

Sector	Index	Reduction rate $\mathcal{R}$						
		30%	40%	50%	60%	70%	80%	90%
Communication Services	7.34	7.35	7.34	7.37	7.43	7.43	7.31	7.30
Consumer Discretionary	5.97	5.97	5.96	5.94	5.93	5.46	4.48	3.55
Consumer Staples	6.04	6.04	6.04	6.04	6.04	6.02	5.39	4.06
Energy	6.49	5.49	4.42	3.84	3.69	3.23	2.58	2.52
Financials	33.91	34.64	35.66	35.96	36.09	37.36	38.86	39.00
Health Care	7.50	7.50	7.50	7.50	7.50	7.50	7.52	7.48
Industrials	8.92	9.38	9.62	10.19	11.34	12.07	13.55	18.13
Information Technology	5.57	5.57	5.59	5.59	5.60	5.60	5.52	5.27
Materials	3.44	3.43	3.31	3.18	3.12	2.64	2.25	1.86
Real Estate	4.76	4.74	4.74	4.74	4.74	4.66	4.61	3.93
Utilities	10.06	9.89	9.82	9.64	8.52	8.04	7.92	6.88

Source: ICE (2022), Trucost (2022) & Barahhou *et al.* (2022).

Table 11.18 shows the sector allocation when considering scope $\boldsymbol{SC}_{1-3}$. Like for equities, the decarbonization process is a strategy that is long on the Financials sector and short on Materials and Utilities sectors. Health care, Communication Services, Consumer Discretionary, and Information Technology weights are very close to their benchmark's. The case of the Industrials sector may be disturbing, but the deviation highly depends on the scope.

Figure 11.15: Impact of the carbon scope on the active share in % (ICE Global Corp., June 2022)

Figure 11.16: Impact of the carbon scope on the DTS risk in bps (ICE Global Corp., June 2022)

11.3 Net-zero investing

The emergence of net-zero emissions policies is one of the hottest topics in finance today. In particular, net-zero emissions policies have gained significant traction in recent years with the proliferation of net-zero alliances (GFANZ, NZAOA, NZAM, NZBA, etc.) and their commitments. This issue is significantly changing portfolio allocation and investment frameworks for both passive and active investors. Indeed, it implies that investors need to implement dynamic decarbonization pathways with continuous reference to business-as-usual benchmarks. In addition, a net-zero portfolio must be considered to finance the transition. The greenness or green intensity of the portfolio is therefore important. A net-zero investment policy therefore has two main dimensions: decarbonizing the portfolio and financing the transition. This means that net-zero investing is not only a carbon footprint issue, but also a green footprint issue.

According to Ben Slimane *et al.* (2023b), there are two approaches to implementing a net-zero investment policy. The integrated approach combines the decarbonization and financing dimensions in an allocation process that considers both carbon intensity for the decarbonization dimension and green intensity for the financing dimension. In the core-satellite strategy, the decarbonization dimension is managed within the core portfolio, while the objective of the satellite strategy is to finance the transition to a low-carbon economy.

11.3.1 Integrated approach

Choice of the decarbonization scenario

The carbon emissions/intensity approach A decarbonization scenario is defined as a function that relates a decarbonization rate to a time index t:

$$f \quad : \quad \mathbb{R}^+ \longrightarrow [0,1]$$
$$t \longmapsto \boldsymbol{\mathcal{R}}\left(t_0, t\right)$$

where t_0 is the base year and $\boldsymbol{\mathcal{R}}\left(t_0, t_0^-\right) = 0$. In general, we assume that $\boldsymbol{\mathcal{R}}\left(t_0, t\right)$ is a nondecreasing function of time t. When considering a decarbonization pathway, we need to distinguish between two different concepts: economic decarbonization and financial decarbonization. In the first case, the variable of interest is the level of carbon emissions, while in the second case we use carbon intensity. Figure 11.17 shows the Net-Zero Emissions by 2050 (NZE) scenario provided by the International Energy Agency[28]. This is a normative scenario based on a number of assumptions about the global energy sector. From this scenario, we can calculate the decarbonization path of the real economy and the different sectors. Figure 11.18 compares these with those used by the CTB and PAB benchmarks. It is clear that we are not comparing apples to apples. In fact, in the case of the real economy, carbon emissions $\boldsymbol{\mathcal{CE}}\left(t\right)$ are assumed to follow the following trajectory:

$$\boldsymbol{\mathcal{CE}}\left(t\right) = \left(1 - \boldsymbol{\mathcal{R}}\left(t_0, t\right)\right)\boldsymbol{\mathcal{CE}}\left(t_0\right)$$

while we have for the PAB and CTB pathways:

$$\boldsymbol{\mathcal{CI}}\left(t\right) = \left(1 - \Delta\boldsymbol{\mathcal{R}}\right)^{t-t_0}\left(1 - \boldsymbol{\mathcal{R}}^-\right)\boldsymbol{\mathcal{CI}}\left(t_0\right)$$

where $\Delta\boldsymbol{\mathcal{R}} = 7\%$ and $\boldsymbol{\mathcal{R}}^-$ takes the values 30% (CTB) and 50% (PAB) respectively (TEG, 2019a). By construction, the reduction path expressed in terms of carbon intensity must be lower than the

---

[28]This scenario has already been explored in Chapter 9 (see Figure 9.22 on page 519).

Figure 11.17: CO_2 emissions by sector in the IEA NZE scenario (in $GtCO_2e$)

Figure 11.18: IEA, NZAOA, CTB and PAB decarbonization pathways

reduction path expressed in terms of carbon emissions. This observation raises the question of the magnitude of the reduction rate. Let us assume that the base date is 2020. The Paris aligned benchmarks imply a reduction rate of 65% by 2025 and 75% by 2030 (Table 11.19). This is much higher than the reduction rates proposed by the International Energy Agency, which are about 15% and 40% by 2025 and 2030 respectively. By comparison, the net-zero frameworks for asset owners propose a reduction rate of about 30% by 2025 and 50% by 2030. The NZAOA curve corresponds to this average asset owner trajectory.

Table 11.19: IEA, NZAOA, CTB and PAB decarbonization rates (baseline = 2020)

Year	CTB	PAB	NZE	NZAOA
$\mathcal{R}^-$	30%	50%	IEA	Average
$\Delta\mathcal{R}$	7%	7%	Scenario	Scenario
2020	30.0%	50.0%	0.0%	10.0%
2021	34.9%	53.5%	1.7%	14.0%
2022	39.5%	56.8%	3.9%	18.0%
2023	43.7%	59.8%	6.7%	22.0%
2024	47.6%	62.6%	9.9%	26.0%
2025	51.3%	65.2%	13.6%	30.0%
2026	54.7%	67.7%	17.8%	34.0%
2027	57.9%	69.9%	22.3%	38.0%
2028	60.8%	72.0%	27.2%	42.0%
2029	63.6%	74.0%	32.1%	46.0%
2030	66.1%	75.8%	37.1%	50.0%
2035	76.4%	83.2%	60.2%	70.3%
2040	83.6%	88.3%	77.2%	89.6%
2045	88.6%	91.9%	87.6%	95.2%
2050	92.1%	94.3%	94.6%	100.0%

Source: Ben Slimane *et al.* (2023b).

The previous analysis considered a global path for the entire economy. However, Figure 11.17 shows that not all sectors are the same. In particular, three major sectors are affected (buildings, electricity and transportation), while some sectors are "*hard-to-abate*" such as materials, steel, cement, petrochemicals, etc. Therefore, a net-zero investment policy must focus on these sectors, which means that we must not spend too much effort on some sectors, such as health care or communication services. There is also a sequencing of decarbonization across sectors as shown in Figure 11.19. The order is as follows:

$$\text{Electricity} \succ \text{Buildings} \succ \text{Transport} \succ \text{Industry}$$

The carbon budget approach   A different approach has been suggested by Bolton *et al.* (2022). The underlying idea is to consider the definition of a net-zero emissions scenario. In fact, a NZE scenario can be defined by a decarbonization pathway that satisfies the following constraints:

$$\begin{cases} \mathcal{CB}\left(t_0, 2050\right) \leq \mathcal{CB}^+ \text{ GtCO}_2\text{e} \\ \mathcal{CE}\left(2050\right) \approx 0 \text{ GtCO}_2\text{e} \end{cases}$$

where t_0 is the base date and $\mathcal{CB}^+$ is the maximum carbon budget. If we look at the SR15 results from IPCC (2018), we can set $t_0 = 2019$ and $\mathcal{CB}^+ = 580\,\text{GtCO}_2\text{e}$, meaning that there is a 50%

Figure 11.19: Sectoral decarbonization pathways

probability of limiting the global warning to 1.5°C. If we want to increase this probability, we can replace the maximum carbon budget $\mathcal{CB}^+$ with a lower number, *e.g.*, $420\,\text{GtCO}_2\text{e}$ for a 66% probability and $300\,\text{GtCO}_2\text{e}$ for an 83% probability. Over the years, the budget constraint moves, especially if the economy's decarbonization path is not met.

Let us assume that carbon dioxide emissions follow the pathways of the CTB/PAB:

$$\mathcal{CE}(t) = (1 - \Delta\mathcal{R})^{t-t_0}\left(1 - \mathcal{R}^-\right)\mathcal{CE}(t_0)$$

Using Equation (9.7) on page 516, we obtain:

$$\mathcal{CB}(t_0, t) = \left(\frac{(1 - \Delta\mathcal{R})^{t-t_0} - 1}{\ln(1 - \Delta\mathcal{R})}\right)\left(1 - \mathcal{R}^-\right)\mathcal{CE}(t_0) \tag{11.13}$$

By considering several values of $\mathcal{R}^-$ and $\Delta\mathcal{R}$, and assuming that $\mathcal{CE}(2020) = 36\,\text{GtCO}_2\text{e}$, we obtain the figures given in Table 11.20. For instance, the carbon budget $\mathcal{CB}(2020, 2050)$ is equal to $308\,\text{GtCO}_2\text{e}$ if $\mathcal{R}^- = 30\%$ and $\Delta\mathcal{R} = 7\%$.

Recall that carbon intensity $\mathcal{CI}(t)$ is defined as the ratio of carbon emissions $\mathcal{CE}(t)$ and the normalization variable $Y(t)$:

$$\mathcal{CI}(t) = \frac{\mathcal{CE}(t)}{Y(t)}$$

Let $\mathcal{R}_{\mathcal{CI}}(t_0, t)$ and $\mathcal{R}_{\mathcal{CE}}(t_0, t)$ be the reduction rates of carbon intensity and emissions between t_0 and t. We have the following relationship:

$$\mathcal{R}_{\mathcal{CI}}(t_0, t) = \frac{\mathcal{CI}(t_0) - \mathcal{CI}(t)}{\mathcal{CI}(t_0)} = \frac{g_Y(t_0, t) + \mathcal{R}_{\mathcal{CE}}(t_0, t)}{1 + g_Y(t_0, t)}$$

Table 11.20: Carbon budget $\mathcal{CB}(2020, 2050)$ (in $GtCO_2e$) when defining the decarbonization pathway of carbon emissions

	$\mathcal{R}^-$	0%	10%	20%	30%	50%	75%
	5%	551	496	441	386	276	138
	6%	491	442	393	344	245	123
$\Delta\mathcal{R}$	7%	440	396	352	308	220	110
	8%	396	357	317	277	198	99
	9%	359	323	287	251	180	90
	10%	327	294	262	229	164	82

where $g_Y(t_0, t)$ is the growth rate of the normalization variable. Assuming that $g_Y(t_0, t) \geq 0$, we can show that the rate of reduction in carbon intensity is always greater than the rate of reduction in carbon emissions. In most cases, we assume that the annual growth rate of the normalization variable is constant: $Y(t) = (1 + g_Y) Y(t - 1)$. We deduce that the compound growth rate is equal to $g_Y(t_0, t) = (1 + g_Y)^{t-t_0} - 1$.

Since we have $\mathcal{CE}(t) = Y(t) \mathcal{CI}(t)$, we obtain:

$$
\begin{aligned}
\mathcal{CB}(t_0, t) &= \mathcal{CE}(t_0) \int_{t_0}^{t} (1 + g_Y(t_0, s)) (1 - \mathcal{R}_{\mathcal{CI}}(t_0, s)) \, ds \\
&= \underbrace{(t - t_0) \mathcal{CE}(t_0)}_{\mathcal{CB}_1(t_0, t)} + \underbrace{\mathcal{CE}(t_0) \int_{t_0}^{t} g_Y(t_0, s) \, ds}_{\mathcal{CB}_2(t_0, t)} - \\
&\quad \underbrace{\mathcal{CE}(t_0) \int_{t_0}^{t} (1 + g_Y(t_0, s)) \mathcal{R}_{\mathcal{CI}}(t_0, s) \, ds}_{\mathcal{CB}_3(t_0, t)}
\end{aligned}
$$

We can divide the carbon budget into three components. The first component $\mathcal{CB}_1(t_0, t)$ corresponds to the total carbon emissions if nothing is done. The second component $\mathcal{CB}_2(t_0, t)$ corresponds to the additional carbon budget if the carbon intensity remains unchanged. The third component $\mathcal{CB}_3(t_0, t)$ is the carbon budget removed by the intensity reduction. If we assume that the annual growth rate of $Y(t)$ is constant and use the PAB/CTB formula for the intensity decarbonization pathway, we get:

$$
\mathcal{CB}(t_0, t) = \frac{(1 + g_Y)^{t-t_0} (1 - \Delta\mathcal{R}_{\mathcal{CI}})^{t-t_0} - 1}{\ln(1 + g_Y) + \ln(1 - \Delta\mathcal{R}_{\mathcal{CI}})} (1 - \mathcal{R}_{\mathcal{CI}}^-) \mathcal{CE}(t_0) \tag{11.14}
$$

Table 11.21 shows the effect of g_Y, $\mathcal{R}_{\mathcal{CI}}^-$ and $\Delta\mathcal{R}_{\mathcal{CI}}$ on the calculation of $\mathcal{CB}(2020, 2050)$ when $\mathcal{CE}(2020) = 36\,GtCO_2e$. If $g_Y = 0\%$, we get the previous results given in Table 11.20. When $g_Y > 0$, the estimated carbon budget is higher. For example, if $\mathcal{R}_{\mathcal{CI}}^- = 30\%$ and $\Delta\mathcal{R}_{\mathcal{CI}} = 7\%$, the carbon budget is equal to $540\,GtCO_2e$ if $g_Y = 5\%$ while it is equal to $308\,GtCO_2e$ if $g_Y = 0\%$. Therefore, we need to use a more aggressive decarbonization pathway for carbon intensity than for carbon emissions.

To satisfy a given carbon budget $\mathcal{CB}(t_0, t)$, we can calibrate a decarbonization pathway in terms of carbon emissions or in terms of carbon intensity. In the first case, we can approximate the integral by the finite sum:

$$
\mathcal{CB}(t_0, t) \approx \sum_{t_i = t_0 + 1}^{t} \mathcal{CE}(t_i)
$$

Table 11.21: Carbon budget $\mathcal{CB}(2020, 2050)$ (in GtCO$_2$e) when defining the decarbonization pathway of carbon intensity

$\mathcal{R}_{\mathcal{CI}}^{-}$	0%	10%	20%	30%	50%	75%	0%	10%	20%	30%	50%	75%
			$g_Y = 1\%$						$g_Y = 3\%$			
5%	619	557	495	433	309	155	793	714	635	555	397	198
6%	547	493	438	383	274	137	691	622	553	484	346	173
7%	487	438	390	341	244	122	607	546	485	425	303	152
$\Delta\mathcal{R}_{\mathcal{CI}}$ 8%	436	392	349	305	218	109	536	482	429	375	268	134
9%	393	353	314	275	196	98	476	429	381	333	238	119
10%	356	320	285	249	178	89	426	383	341	298	213	107
			$g_Y = 5\%$						$g_Y = 10\%$			
5%	1040	936	832	728	520	260	2245	2021	1796	1572	1123	561
6%	893	804	715	625	447	223	1859	1673	1487	1301	930	465
7%	772	695	618	540	386	193	1549	1394	1239	1084	774	387
$\Delta\mathcal{R}_{\mathcal{CI}}$ 8%	672	605	538	470	336	168	1299	1169	1039	909	649	325
9%	589	530	471	412	295	147	1096	987	877	767	548	274
10%	520	468	416	364	260	130	932	839	746	653	466	233

and choose a path $\{\mathcal{CE}(t_0 + 1), \ldots, \mathcal{CE}(t)\}$ that satisfies the carbon budget. We can also assume that $\mathcal{CE}(t) = (1 - \Delta\mathcal{R}_{\mathcal{CE}})^{t-t_0}(1 - \mathcal{R}_{\mathcal{CE}}^{-})\mathcal{CE}(t_0)$ and calibrate the parameters $(\mathcal{R}_{\mathcal{CE}}^{-}, \Delta\mathcal{R}_{\mathcal{CE}})$ such that they satisfy the inequality:

$$\frac{(1 - \Delta\mathcal{R}_{\mathcal{CE}})^{t-t_0} - 1}{\ln(1 - \Delta\mathcal{R}_{\mathcal{CE}})}(1 - \mathcal{R}_{\mathcal{CE}}^{-})\mathcal{CE}(t_0) \leq \mathcal{CB}(t_0, t)$$

We remark that:

$$\frac{(1 - \Delta\mathcal{R}_{\mathcal{CE}})^{t-t_0} - 1}{\ln(1 - \Delta\mathcal{R}_{\mathcal{CE}})} \leq \frac{1}{1 - \mathcal{R}_{\mathcal{CE}}^{-}}\frac{\mathcal{CB}(t_0, t)}{\mathcal{CE}(t_0)} \tag{11.15}$$

This means that the solution does not depend on the absolute value of the carbon budget, but on the ratio of the carbon budget to current emissions. In the second case, we assume that g_Y is given and calibrate the parameters $(\mathcal{R}_{\mathcal{CI}}^{-}, \Delta\mathcal{R}_{\mathcal{CI}})$ in the same way:

$$\frac{(1 - \Delta\mathcal{R}_{\mathcal{CI}})^{t-t_0} - 1}{\ln(1 + g_Y) + \ln(1 - \Delta\mathcal{R}_{\mathcal{CI}})} \leq \frac{1}{(1 + g_Y)^{t-t_0}(1 - \mathcal{R}_{\mathcal{CI}}^{-})}\frac{\mathcal{CB}(t_0, t)}{\mathcal{CE}(t_0)} \tag{11.16}$$

Example 44 *We want to reduce the carbon emissions by* 400% *from* $t_0 = 2020$ *to* $t = 2050$. *We assume that* $\mathcal{R}_{\mathcal{CE}}^{-} = \mathcal{R}_{\mathcal{CI}}^{-} = 0$ *and* $g_Y = 3\%$.

We have:

$$\frac{(t - t_0)\mathcal{CE}(t_0)}{\mathcal{CB}(t_0, t)} = 400\% \Leftrightarrow \frac{\mathcal{CB}(t_0, t)}{\mathcal{CE}(t_0)} = \frac{30}{4} = 7.5$$

because $t - t_0$ is equal to 30 years. The numerical solution to Equation (11.15) is $\Delta\mathcal{R}_{\mathcal{CE}} = 5.88\%$. It follows that:

$$(1 - \Delta\mathcal{R}_{\mathcal{CE}}) = (1 + g_Y)(1 - \Delta\mathcal{R}_{\mathcal{CI}}) \Leftrightarrow \Delta\mathcal{R}_{\mathcal{CI}} = 1 - \left(\frac{1 - \Delta\mathcal{R}_{\mathcal{CE}}}{1 + g_Y}\right)$$

$$\Leftrightarrow \Delta\mathcal{R}_{\mathcal{CI}} = 1 - \left(\frac{1 - 5.88\%}{1 + 3\%}\right) = 8.62\%$$

We check that this is the solution of Equation (11.16).

Dynamic decarbonization and portfolio alignment

General framework While the decarbonization problem finds an optimal portfolio $w^\star(\mathcal{R})$ with respect to a given reduction rate $\mathcal{R}$, the alignment problem defines an optimal portfolio $w^\star(t)$ with respect to a given date t. Therefore, this second problem can be seen as a special case of the first problem, where we use the mapping function between the date t and the reduction rate $\mathcal{R}$. The inequality constraint $\mathcal{CI}(w) \leq (1 - \mathcal{R})\mathcal{CI}(b)$ becomes:

$$\mathcal{CI}(t, w) \leq (1 - \mathcal{R}(t_0, t))\mathcal{CI}(t_0, b(t_0)) \tag{11.17}$$

where t_0 is the base year, $\mathcal{R}(t_0, t)$ is the decarbonization pathway of the NZE scenario and $\mathcal{CI}(t_0, b(t_0))$ is the carbon intensity of the benchmark at time t_0. Portfolio alignment is the process of aligning the decarbonization rate of the portfolio $w^\star(t)$ with respect to the decarbonization pathway of the net-zero scenario. We have the following properties:

- Decarbonizing the aligned portfolio becomes easier over time as the benchmark decarbonizes itself:
$$\mathcal{CI}(t, b(t)) \ll \mathcal{CI}(t_0, b(t_0)) \qquad \text{for } t > t_0$$

- Decarbonizing the aligned portfolio becomes more difficult over time as the benchmark carbonizes itself:
$$\mathcal{CI}(t, b(t)) \gg \mathcal{CI}(t_0, b(t_0)) \qquad \text{for } t > t_0$$

- The aligned portfolio matches the benchmark portfolio if the benchmark is sufficiently decarbonized:
$$\mathcal{CI}(t, b(t)) \leq (1 - \mathcal{R}(t_0, t))\mathcal{CI}(t_0, b(t_0))$$

Since we have $\mathcal{CI}(t, b(t)) = \sum_{i=1}^{n} \mathcal{CI}_i(t)b_i(t)$, the decarbonization of a net-zero investment process is strongly influenced by two factors: changes in benchmark weights and changes in the carbon intensity of assets. In fact, we can imagine that the decarbonization process will become easier over time because the market capitalization of green assets will grow faster than the market capitalization of brown assets and/or because the global decarbonization of the world is well established and on the right track.

Remark 97 *If we consider a carbon budget/emissions approach, we change the inequality constraint (11.17) by $\mathcal{CE}(t, w) \leq (1 - \mathcal{R}(t_0, t))\mathcal{CE}(t_0, b(t_0))$ or $\mathcal{CB}(t_0, t, w) \leq \mathcal{CB}^\star(t_0, t)$ where $\mathcal{CB}^\star(t_0, t)$ is the carbon budget target at time t. In addition, we saw in the previous section that we can calibrate an intensity-based scenario from an emissions- or budget-based scenario.*

Equity portfolios Suppose the objective is to replicate a benchmark and to align the portfolio with a given scenario. The optimization problem becomes:

$$
\begin{aligned}
w^\star(t) &= \arg\min \frac{1}{2}(w - b(t))^\top \Sigma(t)(w - b(t)) \\
&\text{s.t.} \begin{cases} \mathcal{CI}(t, w) \leq (1 - \mathcal{R}(t_0, t))\mathcal{CI}(t_0, b(t_0)) \\ w \in \Omega_0 \cap \Omega \end{cases}
\end{aligned}
\tag{11.18}
$$

where $\Omega_0 = \mathcal{C}_0 = \{w : \mathbf{1}_n^\top w = 1, \mathbf{0}_n \leq w \leq \mathbf{1}_n\}$ defines the long-only constraint and Ω is the set of additional constraints. Note that the benchmark $b(t)$, the covariance matrix $\Sigma(t)$, and the carbon intensity $\mathcal{CI}(t, x)$ are functions of time t. This means that the data is updated each time we rebalance the portfolio[29].

---

[29]For example, at time $t + 1$, the optimization problem depends on the data available at that time, not the data available at time t in the past.

Example 45 *We consider Example 42. We want to align the portfolio with respect to the* CTB
scenario. To compute the optimal portfolio $w^{\star}(t)$ *where* $t = t_0 + h$ *and* $h = 0, 1, 2, ...$ *years, we assume
that the benchmark* $b(t)$, *the covariance matrix* $\Sigma(t)$, *and the vector* $\boldsymbol{CI}(t)$ *of carbon intensities do
not change over time.*

First, we compute the mapping function between the time t and the decarbonization rate
$\boldsymbol{\mathcal{R}}(t_0, t)$:
$$\boldsymbol{\mathcal{R}}(t_0, t) = 1 - (1 - 30\%) \times (1 - 7\%)^h$$

We get $\boldsymbol{\mathcal{R}}(t_0, t_0) = 30\%$, $\boldsymbol{\mathcal{R}}(t_0, t_0 + 1) = 34.90\%$, $\boldsymbol{\mathcal{R}}(t_0, t_0 + 2) = 39.46\%$, and so on. Second, we
solve the optimization problem (11.18) for the different values of time t. The results are shown in
Table 11.22.

Table 11.22: Equity portfolio alignment (Example 45)

t	$b(t_0)$	t_0	$t_0 + 1$	$t_0 + 2$	$t_0 + 3$	$t_0 + 4$	$t_0 + 5$	$t_0 + 10$
$w_1^{\star}$	20.00	21.86	22.21	22.54	22.84	23.02	22.92	8.81
$w_2^{\star}$	19.00	18.70	18.41	18.15	17.90	17.58	17.04	0.00
$w_3^{\star}$	17.00	8.06	5.69	3.48	1.43	0.00	0.00	0.00
$w_4^{\star}$	13.00	8.74	7.66	6.65	5.72	4.56	2.70	0.00
$w_5^{\star}$	12.00	13.07	13.29	13.51	13.70	13.91	14.18	21.22
$w_6^{\star}$	8.00	22.57	25.59	28.39	31.00	33.39	35.54	62.31
$w_7^{\star}$	6.00	7.00	7.15	7.29	7.42	7.53	7.63	7.66
$w_8^{\star}$	5.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$\sigma(w^{\star} \mid b(t))$	0.01	104.10	126.22	147.14	166.79	185.24	203.51	352.42
$\boldsymbol{CI}(t, w)$	160.57	112.40	104.53	97.22	90.41	84.08	78.20	54.40
$\boldsymbol{\mathcal{R}}(w \mid b(t_0))$	0.00	30.00	34.90	39.46	43.70	47.64	51.30	66.12

The reduction rate and weights are expressed in %, while the tracking error volatility is measured in bps.

Instead of using the threshold approach, we can implement the order statistic or the naive
approach when performing portfolio alignment. These approaches make sense because net-zero in-
vesting is an exclusion process, as shown by Barahhou *et al.* (2022). Examples of portfolio alignment
based on order-statistic and naive approaches can be found in Jondeau *et al.* (2021), Bolton *et al.*
(2022), and Ben Slimane *et al.* (2023b).

Bond portfolios   For bonds, the tracking error volatility is replaced by the active risk function
proposed by Barahhou *et al.* (2022):

$$\mathcal{D}(w \mid b) = \varphi \underbrace{\sum_{s=1}^{n_{\mathcal{S}ector}} \left| \sum_{i \in s} (w_i - b_i) \operatorname{DTS}_i \right|}_{\text{DTS component}} + \underbrace{\frac{1}{2} \sum_{i \in b} |w_i - b_i|}_{\text{AS component}} + \underbrace{\mathbb{1}_{\Omega_{\text{MD}}}(w)}_{\text{MD component}} \qquad (11.19)$$

where DTS_i and MD_i are the duration-times-spread and modified duration factors, $\Omega_{\text{MD}} = \{w : \sum_{i=1}^{n} (w_i - b_i) \operatorname{MD}_i = 0\}$ and $\mathbb{1}_{\Omega}(w)$ is the convex indicator function. The optimization prob-
lem becomes then:

$$
\begin{aligned}
w^{\star}(t) &= \arg\min \mathcal{D}(w \mid b(t)) \\
&\text{s.t.} \quad \begin{cases} \boldsymbol{CI}(t, w) \leq (1 - \boldsymbol{\mathcal{R}}(t_0, t)) \boldsymbol{CI}(t_0, b(t_0)) \\ w \in \Omega_0 \cap \Omega \end{cases}
\end{aligned} \qquad (11.20)
$$

Again, the benchmark $b(t)$ and the risk factors (modified duration, DTS, etc.) are functions of time t. To solve this mathematical problem, we transform it into a linear programming problem and use standard LP algorithms.

Example 46 *We consider Example 43. We want to align the portfolio with respect to the CTB scenario. To compute the optimal portfolio $w^\star(t)$ where $t = t_0 + h$ and $h = 0, 1, 2, ...$ years, we assume that the benchmark, the modified duration and the duration-times-spread factors do not change over time.*

The corresponding LP problem is[30]:

$$
\begin{aligned}
x^\star &= \arg\min c^\top x \\
&\text{s.t.} \quad \begin{cases} Ax = B \\ Cx \leq D \\ x^- \leq x \leq x^+ \end{cases}
\end{aligned}
\tag{11.21}
$$

where $x = (w, \tau_w, \tau_{\text{DTS}})$ is a 18×1 vector. The 18×1 vector c is equal to $\left(\mathbf{0}_8, \dfrac{1}{2}\mathbf{1}_8, \varphi\mathbf{1}_2\right)$. The equality constraint includes the convex indicator function $\mathbb{1}_{\Omega_{\text{MD}}}(w)$ and is defined by:

$$
Ax = B \Leftrightarrow \begin{pmatrix} \mathbf{1}_8^\top & \mathbf{0}_8^\top & \mathbf{0}_2^\top \\ \text{MD}^\top & \mathbf{0}_8^\top & \mathbf{0}_2^\top \end{pmatrix} x = \begin{pmatrix} 1 \\ 5.476 \end{pmatrix}
$$

The inequality constraints are:

$$
Cx \leq D \Leftrightarrow \begin{pmatrix} I_8 & -I_8 & \mathbf{0}_{8,2} \\ -I_8 & -I_8 & \mathbf{0}_{8,2} \\ C_{\text{DTS}} & \mathbf{0}_{2,8} & -I_2 \\ -C_{\text{DTS}} & \mathbf{0}_{2,8} & -I_2 \\ \boldsymbol{\mathcal{CI}}(t)^\top & \mathbf{0}_{1,8} & 0 \end{pmatrix} x \leq \begin{pmatrix} b \\ -b \\ 192.68 \\ 108.37 \\ -192.68 \\ -108.37 \\ 160.574 \times (1 - \boldsymbol{\mathcal{R}}(t_0, t)) \end{pmatrix}
$$

where:

$$
C_{\text{DTS}} = \begin{pmatrix} 100 & 0 & 575 & 436 & 0 & 0 & 0 & 365 \\ 0 & 155 & 0 & 0 & 159 & 145 & 804 & 0 \end{pmatrix}
$$

Finally, the bounds are $x^- = \mathbf{0}_{18}$ and $x^+ = \infty \cdot \mathbf{1}_{18}$. The solutions are shown in Table 11.23.

Defining a net-zero investment policy

General framework As explained by Barahhou *et al.* (2022) and Ben Slimane *et al.* (2023b), net-zero investment policies must address two dimensions: portfolio alignment and financing the transition. Therefore, the set of constraints to be applied must include this second dimension:

$$
\Omega = \Omega_{\text{alignment}} \cap \Omega_{\text{transition}}
$$

where:

$$
\Omega_{\text{alignment}} = \{w : \boldsymbol{\mathcal{CI}}(t, w) \leq (1 - \boldsymbol{\mathcal{R}}(t_0, t))\boldsymbol{\mathcal{CI}}(t_0, b(t_0))\}
$$

---

[30]We use the LP framework and the transformation problem (11.8) defined on page 564.

Table 11.23: Bond portfolio alignment (Example 46)

t	$b(t_0)$	t_0	t_0+1	t_0+2	t_0+3	t_0+4	t_0+5	t_0+10
$w_1^\star$	20.00	20.00	20.00	20.00	13.98	17.64	16.02	5.02
$w_2^\star$	19.00	13.99	17.79	19.00	19.00	19.00	19.00	19.00
$w_3^\star$	17.00	25.43	20.96	17.78	17.00	13.64	11.65	4.61
$w_4^\star$	13.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$w_5^\star$	12.00	28.97	30.71	35.84	43.52	48.80	53.33	71.37
$w_6^\star$	8.00	8.00	8.00	5.67	6.46	0.92	0.00	0.00
$w_7^\star$	6.00	3.61	2.53	1.70	0.04	0.00	0.00	0.00
$w_8^\star$	5.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$\mathrm{AS}(w)$	0.00	25.40	22.68	24.62	31.52	36.80	41.33	59.37
$\mathrm{MD}(w)$	5.48	5.48	5.48	5.48	5.48	5.48	5.48	5.48
$\mathrm{DTS}(w)$	301.05	274.61	248.91	230.60	220.10	204.46	197.26	174.46
$\mathcal{D}(w\mid b)$	0.00	0.39	0.49	0.60	0.72	0.85	0.99	1.57
$\mathcal{CI}(w)$	160.57	112.40	104.53	97.22	90.41	84.08	78.20	54.40
$\mathcal{R}(w\mid b)$	0.00	30.00	34.90	39.46	43.70	47.64	51.30	66.12

The reduction rate, weights, and active share metrics are expressed in %, the MD metrics are measured in years, and the DTS metrics are calculated in bps.

While specifying $\Omega_{\text{alignment}}$ is straightforward, specifying $\Omega_{\text{transition}}$ is more complex. Indeed, the goal of a net-zero investment policy is to participate in the transformation to a low-carbon economy. As explained in Chapter 15 of IPCC (2022), this transformation involves large programs of capital reallocation for the low-carbon transition. However, ensuring efficient capital allocation in line with climate targets is not an easy task. Therefore, the number of factors to define net-zero investments can be very large. Barahhou *et al.* (2022) considers three of them, which are:

1. The self-decarbonization of the portfolio;

2. The greenness of the portfolio;

3. The exclusion of net-zero "*enemies*" (free riders).

Therefore, we have:

$$\Omega_{\text{transition}} = \Omega_{\text{self-decarbonization}} \cap \Omega_{\text{greenness}} \cap \Omega_{\text{exclusion}}$$

Self-decarbonization and endogeneity of the decarbonization pathway   In the context of a net-zero scenario, portfolio alignment is a dynamic approach to portfolio decarbonization. Most investors have solved this problem by considering a time-varying rate of carbon footprint reduction. In this case, the portfolio is periodically rebalanced to match the decarbonization pathway. This is what we have done in the previous paragraph. However, the decarbonization dimension of net-zero investing cannot be summarized by a sequence of decarbonization rates or a sequence of portfolio rebalancing. In fact, if net-zero investing consists of building successive independent portfolios, there is no mechanism that respects the endogenous aspect of the decarbonization pathway. In particular, if the time-varying decarbonization is due only to the rebalancing process, it is clear that the portfolio cannot claim to be net-zero. In fact, the endogenous aspect of the decarbonization pathway implies a self-decarbonization of the portfolio. Consider an example to illustrate the concept of self-decarbonization. Suppose the decarbonization rate at the beginning of year t is 30%. For

the next year $t + 1$, the goal is to achieve a decarbonization rate of 35%. Two extreme cases are considered below. In the first panel on the left, effective decarbonization of the portfolio at the end of the year is 25%, which means that the carbon footprint of the portfolio has increased. In this case, we need to rebalance the portfolio to reach the 35% level at the beginning of year $t + 1$. This is the bad case because the self-decarbonization of the portfolio is zero. In the third panel on the right, the decarbonization rate of the portfolio is higher than 35% at the end of the year, which means that we do not need to rebalance the portfolio. This is the good case because the portfolio has decarbonized itself, meaning that the carbon footprint of the issuers in the portfolio is following a trajectory consistent with the net-zero scenario.

	Bad case	Mixed case	Good case
Effective decarbonization			
at the beginning of the year t	30%	30%	30%
at the end of the year t	25%	33%	36%
Self-decarbonization	0%	3%	6%
Relabancing requirement	10%	2%	0%

The previous example illustrates that we can always follow a decarbonization path by rebalancing the portfolio if it is composed of liquid assets, but this does not mean that the investment process is a net-zero investment policy. In particular, there are some financial businesses where it is difficult to rebalance the portfolio because the assets are not liquid, such as a portfolio of private equities, a portfolio of car insurance policies, or a portfolio of loans.

We therefore need to introduce an incentive mechanism to achieve a minimum level of self-decarbonization. The objective of the temperature ratings is precisely to assess an issuer's ability to adapt to a carbon emissions scenario. Implied temperature ratings can be seen as a synthetic scoring system based on the $\mathcal{PAC}$ framework (Le Guenedal *et al.*, 2022), which measures the (past) participation, ambition and credibility of the issuer[31]. As a temperature rating system is often perceived as a black box, we may consider a simplified approach that is more transparent. For example, we can use net-zero targets that have been approved and validated by a third party. Using a linear interpolation model, we can calculate the annual self-decarbonization rate of issuers and derive the self-decarbonization level of portfolios. This simple approach is limited for two reasons. First, the data is not homogeneous, as target dates and scopes may differ. Second, self-decarbonization cannot be calculated for issuers without a net-zero commitment or validation. Another approach is to focus on the first pillar, which is participation. In fact, participation is a technical term used to identify past self-decarbonization. This explains why carbon trends and carbon momentum measures are very important metrics for a net-zero investor. This is a way to introduce a dynamic approach to carbon footprints and move beyond current levels, which are a poor estimate of the issuer's finish line and an even poorer estimate of how quickly the issuer will get there.

Following Barahhou *et al.* (2022), we can specify the self-decarbonization constraint as follows:

$$\Omega_{\text{self-decarbonization}} = \{w : \mathcal{CM}(t, w) \leq \mathcal{CM}^{\star}(t)\}$$

where $\mathcal{CM}(t, w)$ is the carbon momentum of the portfolio w at time t and $\mathcal{CM}^{\star}(t)$ is the self-decarbonization minimum threshold. For example, if $\mathcal{CM}^{\star}(t) = -3\%$, we expect to reduce the carbon footprint of the portfolio by 3% next year if the observed trend continues. By construction, carbon momentum is inherently a backward-looking approach to self-decarbonization that can be complemented by more forward-looking measures of self-decarbonization.

---

[31]See Section 9.3.2 on page 520.

Green footprint   The second factor in the transition dimension is the greenness of the portfolio. In fact, to achieve a low-carbon economy by 2050, we need to reduce the carbon footprint, but we also need to improve the green intensity of the economy. This means that net-zero investing is not just a carbon footprint issue, it is also a green footprint issue, and these two concepts are different. The goal of the second factor is then to finance the transition to a low-carbon economy, or in other words, to reallocate capital investment to green activities. The greenness constraint can be written as follows:

$$\Omega_{\text{greenness}} = \{w : \boldsymbol{\mathcal{GI}}(t,w) \geq \boldsymbol{\mathcal{GI}}^\star(t)\}$$

where $\boldsymbol{\mathcal{GI}}(t,w)$ is the green intensity of the portfolio w at time t and $\boldsymbol{\mathcal{GI}}^\star(t)$ is the minimum threshold. In general, the absolute measure $\boldsymbol{\mathcal{GI}}^\star(t)$ is expressed as a relative value with respect to the benchmark:

$$\boldsymbol{\mathcal{GI}}^\star(t) = (1 + \boldsymbol{\mathcal{G}})\,\boldsymbol{\mathcal{GI}}(t, b(t))$$

where $\boldsymbol{\mathcal{G}}$ is the minimum growth value. For example, if $\boldsymbol{\mathcal{G}} = 100\%$, we want to improve the green footprint of the benchmark so that the green intensity of the portfolio is at least twice the green intensity of the benchmark[32].

Remark 98 *Another approach is to use a greenness pathway* $\boldsymbol{\mathcal{GI}}^\star(t) = (1 + \boldsymbol{\mathcal{G}}(t_0, t))\,\boldsymbol{\mathcal{GI}}(t, b(t_0))$. *In this case,* $\Omega_{greenness}$ *and* $\Omega_{alignment}$ *are two symmetric constraints.*

The choice of green intensity is critical. In theory, we want to improve the future green footprint of the economy. Therefore, an appropriate measure would be a forward-looking metric. The ideal candidate is green capex because it measures current green investment. It is therefore a proxy for future green revenues. However, the scarcity and current robustness of green capex data is an obstacle. As a result, most investors prefer to use green revenue share.

Net-zero exclusion policy   Any net-zero investment portfolio must include an exclusion policy, which may include sectors and/or issuers. Most asset owners and managers have a coal exclusion policy when implementing net-zero. Some may also include the fossil fuel sector. These sector exclusions are related to ESG exclusion strategies. As with ESG investing, investors also create an exclusion list of issuers. In some cases, the net-zero exclusion list is the same as the ESG exclusion list, but in most cases, the net-zero exclusion list is specific. The criteria can be the issuer's carbon emissions or intensity relative to the sector, but the most popular metric is the temperature score. For example, an issuer will often be excluded if its temperature score is above a cap threshold, such as 4 or 5 degrees Celsius. The rationale is to exclude issuers that are not willing to participate in the net-zero journey. In the same spirit as the temperature score, Barahhou *et al.* (2022) suggests excluding issuers whose carbon momentum is greater than a threshold $\boldsymbol{\mathcal{CM}}^+$:

$$\Omega_{\text{exclusion}} = \left\{w : \boldsymbol{\mathcal{CM}}_i \geq \boldsymbol{\mathcal{CM}}^+ \Rightarrow w_i = 0\right\}$$

Again, the underlying rationale is to measure the willingness of issuers to follow a net-zero scenario, and to exclude issuers who have not played the game in the past. For example, if $\boldsymbol{\mathcal{CM}}^+ = 10\%$, we exclude issuers that have recently increased their carbon footprint by more than 10%. On page 528 we have seen that the proportion of excluded issuers can be high, especially if we consider carbon emissions.

Remark 99 *To manage the self-decarbonization of the portfolio or to define the exclusion list, it is better to use the long-term carbon momentum than the short-term carbon momentum.*

---

[32]$\boldsymbol{\mathcal{G}} = 100\%$ is the standard measure in the ETF market.

Equity portfolios The optimization problem (11.18) on page 597 becomes:

$$w^\star(t) = \arg\min \frac{1}{2}(w - b(t))^\top \Sigma(t)(w - b(t)) \tag{11.22}$$

$$\text{s.t.} \begin{cases} \boldsymbol{\mathcal{CI}}(t, w) \leq (1 - \boldsymbol{\mathcal{R}}(t_0, t))\boldsymbol{\mathcal{CI}}(t_0, b(t_0)) & \leftarrow \text{Alignment} \\ \boldsymbol{\mathcal{CM}}(t, w) \leq \boldsymbol{\mathcal{CM}}^\star(t) & \leftarrow \text{Self-decarbonization} \\ \boldsymbol{\mathcal{GI}}(t, w) \geq (1 + \boldsymbol{\mathcal{G}})\boldsymbol{\mathcal{GI}}(t, b(t)) & \leftarrow \text{Greenness} \\ 0 \leq w_i \leq \mathbb{1}\left\{\boldsymbol{\mathcal{CM}}_i(t) \leq \boldsymbol{\mathcal{CM}}^+\right\} & \leftarrow \text{Exclusion} \\ w \in \Omega_0 \cap \Omega & \leftarrow \text{Other constraints} \end{cases}$$

We deduce that the quadratic form is $Q = \Sigma(t)$, $R = \Sigma(t)b(t)$, $A = \mathbf{1}_n^\top$, $B = 1$, $w^- = \mathbf{0}_n$, $w^+ = \mathbf{1}\left\{\boldsymbol{\mathcal{CM}}(t) \leq \boldsymbol{\mathcal{CM}}^+\right\}$ and:

$$Cw \leq D \Leftrightarrow \begin{pmatrix} \boldsymbol{\mathcal{CI}}(t)^\top \\ \boldsymbol{\mathcal{CM}}(t)^\top \\ -\boldsymbol{\mathcal{GI}}(t)^\top \end{pmatrix} w \leq \begin{pmatrix} (1 - \boldsymbol{\mathcal{R}}(t_0, t))\boldsymbol{\mathcal{CI}}(t_0, b(t_0)) \\ \boldsymbol{\mathcal{CM}}^\star(t) \\ -(1 + \boldsymbol{\mathcal{G}})\boldsymbol{\mathcal{GI}}(t, b(t)) \end{pmatrix}$$

Here, we assume that the carbon momentum function is a linear function:

$$\boldsymbol{\mathcal{CM}}(t, w) = w^\top\boldsymbol{\mathcal{CM}}(t) = \sum_{i=1}^n w_i\boldsymbol{\mathcal{CM}}_i(t)$$

where $\boldsymbol{\mathcal{CM}}(t) = (\boldsymbol{\mathcal{CM}}_1(t), \ldots, \boldsymbol{\mathcal{CM}}_n(t))$ is the carbon momentum vector.

Remark 100 *In the previous optimization problem, the carbon momentum constraint can be replaced by a temperature constraint $\boldsymbol{\mathcal{TS}}(t, w) \leq \boldsymbol{\mathcal{TS}}^\star(t)$ where $\boldsymbol{\mathcal{TS}}(t, w) = w^\top\boldsymbol{\mathcal{TS}}(t) = \sum_{i=1}^n w_i\boldsymbol{\mathcal{TS}}_i(t)$ and $\boldsymbol{\mathcal{TS}}_i(t)$ is the temperature score of issuer i.*

Note that the aggregation of the carbon momentum or temperature score at the portfolio level uses the weighted average approach, as does the WACI formulation for carbon intensity. If we use long-term carbon momentum estimated with a linear trend model, we can perform an exact calculation of the carbon momentum at the portfolio level. Using the results in Box 11.2, we have:

$$\begin{aligned} \boldsymbol{\mathcal{CM}}(t, w) \leq \boldsymbol{\mathcal{CM}}^\star(t) &\Leftrightarrow \sum_{i=1}^n \tilde{w}_i\boldsymbol{\mathcal{CM}}_i(t) \leq \boldsymbol{\mathcal{CM}}^\star(t) \\ &\Leftrightarrow \sum_{i=1}^n \frac{w_i\boldsymbol{\mathcal{CI}}_i(t)}{\sum_{j=1}^n w_j\boldsymbol{\mathcal{CI}}_j(t)}\boldsymbol{\mathcal{CM}}_i(t) \leq \boldsymbol{\mathcal{CM}}^\star(t) \\ &\Leftrightarrow \sum_{i=1}^n w_i\boldsymbol{\mathcal{CI}}_i(t)\boldsymbol{\mathcal{CM}}_i(t) \leq \sum_{j=1}^n w_j\boldsymbol{\mathcal{CI}}_j(t)\boldsymbol{\mathcal{CM}}^\star(t) \\ &\Leftrightarrow \sum_{i=1}^n w_i\boldsymbol{\mathcal{CI}}_i(t)(\boldsymbol{\mathcal{CM}}_i(t) - \boldsymbol{\mathcal{CM}}^\star(t)) \leq 0 \\ &\Leftrightarrow \sum_{i=1}^n w_i\zeta_i \leq 0 \end{aligned}$$

where:

$$\zeta_i = \boldsymbol{\mathcal{CI}}_i(t)(\boldsymbol{\mathcal{CM}}_i(t) - \boldsymbol{\mathcal{CM}}^\star(t))$$

Box 11.2: Carbon momentum aggregation at the portfolio level

We recall that $\mathcal{CM}_i^{\mathcal{L}ong}(t) = \dfrac{\hat{\beta}_{i,1}(t)}{\mathcal{CI}_i(t)}$ where i is the issuer, $\mathcal{CI}_i(t)$ is the carbon intensity, and $\hat{\beta}_{i,1}(t)$ is the slope of the linear trend model:

$$\widehat{\mathcal{CI}}_i(t) = \hat{\beta}_{i,0}(t) + \hat{\beta}_{i,1}(t)(t - t_0)$$

The carbon intensity of the portfolio is given by its weighted average: $\mathcal{CI}(t,w) = \sum_{i=1}^{n} w_i \mathcal{CI}_i(t)$. This follows:

$$\begin{aligned}
\widehat{\mathcal{CI}}(t,w) &= \sum_{i=1}^{n} w_i \widehat{\mathcal{CI}}_i(t) \\
&= \sum_{i=1}^{n} w_i \hat{\beta}_{i,0}(t) + \underbrace{\sum_{i=1}^{n} w_i \hat{\beta}_{i,1}(t)}_{\hat{\beta}_1(t,w)}(t - t_0)
\end{aligned}$$

where $\hat{\beta}_1(t,w) = \sum_{i=1}^{n} w_i \hat{\beta}_{i,1}(t)$. We deduce that:

$$\begin{aligned}
\mathcal{CM}^{\mathcal{L}ong}(t,w) &= \frac{\hat{\beta}_1(t,w)}{\mathcal{CI}(t,w)} \\
&= \frac{\sum_{i=1}^{n} w_i \hat{\beta}_{i,1}(t)}{\sum_{i=1}^{n} w_i \mathcal{CI}_i(t)} \\
&= \frac{\sum_{i=1}^{n} w_i \mathcal{CI}_i(t) \mathcal{CM}_i^{\mathcal{L}ong}(t)}{\sum_{i=1}^{n} w_i \mathcal{CI}_i(t)} \\
&= \sum_{i=1}^{n} \tilde{w}_i \mathcal{CM}_i^{\mathcal{L}ong}(t)
\end{aligned}$$

where the adjusted weight $\tilde{w}_i$ is equal to:

$$\tilde{w}_i = \frac{w_i \mathcal{CI}_i(t)}{\sum_{j=1}^{n} w_j \mathcal{CI}_j(t)}$$

We see that $\mathcal{CM}^{\mathcal{L}ong}(t,w) \neq \sum_{i=1}^{n} w_i \mathcal{CM}_i^{\mathcal{L}ong}(t)$. This aggregation method is also valid at the sector level by using the weight of each issuer in its respective sector.

The formulation of the QP problem remains the same, but the inequality constraints are changed:

$$Cw \leq D \Leftrightarrow \begin{pmatrix} \mathcal{CI}(t)^{\top} \\ \zeta^{\top} \\ -\mathcal{GI}(t)^{\top} \end{pmatrix} w \leq \begin{pmatrix} (1 - \mathcal{R}(t_0,t))\, \mathcal{CI}(t_0, b(t_0)) \\ 0 \\ -(1 + \mathcal{G})\, \mathcal{GI}(t, b(t)) \end{pmatrix}$$

where $\zeta = (\zeta_1, \ldots, \zeta_n)$.

If we define the carbon momentum by considering the log-linear trend model, we get $\mathcal{CM}_i^{\mathcal{L}ong}(t) = \hat{\gamma}_{i,1}(t)$. We cannot find an analytically exact formula $\hat{\gamma}_1(t,w)$ for the portfolio. Therefore, we use the weighted average approach:

$$\mathcal{CM}(t,w) = \sum_{i=1}^{n} w_i \hat{\gamma}_{i,1}(t)$$

Example 47 *We consider Example 45. The carbon momentum values are equal to* -3.1%, -1.2%, -5.8%, -1.4%, $+7.4\%$, -2.6%, $+1.2\%$, *and* -8.0%. *We measure the green intensity by the green revenue share. Its values are equal to* 10.2%, 45.3%, 7.5%, 0%, 0%, 35.6%, 17.8% *and* 3.0%. *The net-zero investment policy imposes to follow the* CTB *decarbonization pathway with a self-decarbonization of* 3%, *and to improve the green intensity of the benchmark by* 100%.

The solutions of the QP optimization problem are shown in Table 11.24. The carbon momentum of the benchmark is $\boldsymbol{\mathcal{CM}}(t_0, b) = -1.66\%$, which is not enough to ensure a self-decarbonization of 3%. The green intensity of the benchmark is $\boldsymbol{\mathcal{GI}}(t_0, b) = 15.99\%$, which means that the green intensity target is 31.98%. If we consider the base date t_0, the solution has a lower carbon intensity than the target defined by the CTB decarbonization pathway, because the self-decarbonization and green intensity constraints imply a greater reduction in carbon footprint. We also note that there is no solution for the years $t_0 + 5$ and $t_0 + 10$.

Table 11.24: Net-zero equity portfolio (Example 47)

t	$b(t_0)$	t_0	t_0+1	t_0+2	t_0+3	t_0+4	t_0+5	t_0+10
$w_1^\star$	20.00	5.26	3.51	1.49	0.00	0.02		
$w_2^\star$	19.00	20.96	17.27	13.00	8.82	4.16		
$w_3^\star$	17.00	3.35	7.27	11.82	15.02	14.32		
$w_4^\star$	13.00	0.00	0.00	0.00	0.00	0.00	No feasible	
$w_5^\star$	12.00	0.00	0.00	0.00	0.00	0.00	solution	
$w_6^\star$	8.00	60.06	64.69	70.05	75.37	81.51		
$w_7^\star$	6.00	0.00	0.00	0.00	0.00	0.00		
$w_8^\star$	5.00	10.37	7.25	3.64	0.79	0.00		
$\sigma(w^\star \mid b(t))$	0.00	370.16	376.38	398.30	430.94	472.44		
$\boldsymbol{\mathcal{CI}}(t,w)$	160.57	110.85	104.53	97.22	90.41	84.08		
$\boldsymbol{\mathcal{R}}(w \mid b(t_0))$	0.00	30.96	34.90	39.46	43.70	47.64		
$\boldsymbol{\mathcal{CM}}(t,w)$	-1.66	-3.00	-3.00	-3.00	-3.00	-3.00		
$\boldsymbol{\mathcal{GI}}(t,w)$	15.99	31.98	31.98	31.98	31.98	31.98		

The reduction rate, weights, carbon momentum and green intensity are expressed in %, while the tracking error volatility is measured in bps.

Remark 101 *This example shows that there is not always a mathematical solution to the net-zero optimization problem when we stack many constraints. Moreover, even if there is a mathematical solution, it may not be investable due to poor liquidity or poor diversification. In this example, the net-zero portfolio is highly concentrated in the sixth asset. Of course, our example has only eight assets, so the net-zero portfolio can become concentrated very quickly. However, this drawback may be present if we also consider large equity universes such as the MSCI World. Therefore, there is a trade-off between the existence of the solution and the investability of the solution.*

Bond portfolios By introducing the transition constraints $\Omega_{\text{transition}}$, the optimization problem (11.20) becomes:

$$w^\star(t) = \arg\min \mathcal{D}(w \mid b(t)) \tag{11.23}$$

$$\text{s.t.} \begin{cases} \boldsymbol{\mathcal{CI}}(t,w) \leq (1 - \boldsymbol{\mathcal{R}}(t_0,t))\, \boldsymbol{\mathcal{CI}}(t_0, b(t_0)) & \leftarrow \text{Alignment} \\ \boldsymbol{\mathcal{CM}}(t,w) \leq \boldsymbol{\mathcal{CM}}^\star(t) & \leftarrow \text{Self-decarbonization} \\ \boldsymbol{\mathcal{GI}}(t,w) \geq (1 + \mathcal{G})\, \boldsymbol{\mathcal{GI}}(t, b(t)) & \leftarrow \text{Greenness} \\ 0 \leq w_i \leq \mathbb{1}\left\{\boldsymbol{\mathcal{CM}}_i(t) \leq \boldsymbol{\mathcal{CM}}^+\right\} & \leftarrow \text{Exclusion} \\ w \in \Omega_0 \cap \Omega & \leftarrow \text{Other constraints} \end{cases}$$

We get the same LP form except for the set of inequality constraints:

$$Cx \leq D \Leftrightarrow \begin{pmatrix} I_n & -I_n & \mathbf{0}_{n,n_{\boldsymbol{S}ector}} \\ -I_n & -I_n & \mathbf{0}_{n,n_{\boldsymbol{S}ector}} \\ C_{\mathrm{DTS}} & \mathbf{0}_{n_{\boldsymbol{S}ector},n} & -I_{n_{\boldsymbol{S}ector}} \\ -C_{\mathrm{DTS}} & \mathbf{0}_{n_{\boldsymbol{S}ector},n} & -I_{n_{\boldsymbol{S}ector}} \\ \boldsymbol{\mathcal{CI}}(t)^{\top} & \mathbf{0}_{1,n} & \mathbf{0}_{1,n_{\boldsymbol{S}ector}} \\ \boldsymbol{\mathcal{CM}}(t)^{\top} & \mathbf{0}_{1,n} & \mathbf{0}_{1,n_{\boldsymbol{S}ector}} \\ -\boldsymbol{\mathcal{GI}}(t)^{\top} & \mathbf{0}_{1,n} & \mathbf{0}_{1,n_{\boldsymbol{S}ector}} \end{pmatrix} x \leq \begin{pmatrix} b \\ -b \\ \mathrm{DTS}^{\star} \\ -\mathrm{DTS}^{\star} \\ (1 - \boldsymbol{\mathcal{R}}(t_0,t))\,\boldsymbol{\mathcal{CI}}(t_0, b(t_0)) \\ \boldsymbol{\mathcal{CM}}^{\star}(t) \\ -(1 + \boldsymbol{\mathcal{G}})\,\boldsymbol{\mathcal{GI}}(t, b(t)) \end{pmatrix}$$

and the upper bound: $x^{+} = \left(\mathbb{1}\left\{\boldsymbol{\mathcal{CM}}(t) \leq \boldsymbol{\mathcal{CM}}^{+}\right\}, \infty \cdot \mathbf{1}_n, \infty \cdot \mathbf{1}_{n_{\boldsymbol{S}ector}}\right)$.

Example 48 *We consider Example 46. The carbon momentum values are equal to -3.1%, -1.2%, -5.8%, -1.4%, $+7.4\%$, -2.6%, $+1.2\%$, and -8.0%. We measure the green intensity by the green revenue share. Its values are equal to 10.2%, 45.3%, 7.5%, 0%, 0%, 35.6%, 17.8% and 3.0%. The net-zero investment policy imposes to follow the CTB decarbonization pathway with a self-decarbonization of 2%, and to improve the green intensity of the benchmark by 100%.*

The solutions of the LP optimization problem are shown in Table 11.25. Note that again there is no solution for some dates. In addition, compared to the net-zero equity portfolio exercise, we have changed the self-decarbonization target $\boldsymbol{\mathcal{CM}}^{\star}(t)$, which is equal to -2%. Indeed, if we set $\boldsymbol{\mathcal{CM}}^{\star}(t) = -3\%$, we have no solution to the optimization problem even for the base date t_0.

Table 11.25: Net-zero bond portfolio (Example 48)

t	$b(t_0)$	t_0	$t_0 + 1$	$t_0 + 2$	$t_0 + 3$	$t_0 + 4$	$t_0 + 5$	$t_0 + 10$
$w_1^{\star}$	20.00	4.28	13.80	20.48	26.34	19.02		
$w_2^{\star}$	19.00	34.78	38.94	42.72	46.23	49.01		
$w_3^{\star}$	17.00	21.03	13.86	7.73	2.11	0.00		
$w_4^{\star}$	13.00	0.00	0.00	0.00	0.00	0.00	No feasible	
$w_5^{\star}$	12.00	0.00	0.00	0.00	0.00	0.00	solution	
$w_6^{\star}$	8.00	39.91	33.40	29.07	25.32	31.97		
$w_7^{\star}$	6.00	0.00	0.00	0.00	0.00	0.00		
$w_8^{\star}$	5.00	0.00	0.00	0.00	0.00	0.00		
AS(w)	0.00	51.72	45.34	45.27	50.89	53.98		
MD(w)	5.48	5.48	5.48	5.48	5.48	5.48		
DTS(w)	301.05	236.99	202.30	173.29	146.83	141.34		
$\mathcal{D}(w \mid b)$	0.00	0.87	0.95	1.09	1.28	1.48		
$\boldsymbol{\mathcal{CI}}(w)$	160.57	112.40	104.53	97.22	90.41	84.08		
$\boldsymbol{\mathcal{R}}(w \mid b)$	0.00	30.00	34.90	39.46	43.70	47.64		
$\boldsymbol{\mathcal{CM}}(t, w)$	-1.66	-2.81	-2.57	-2.35	-2.15	-2.01		
$\boldsymbol{\mathcal{GI}}(t, w)$	15.99	31.98	31.98	32.37	32.80	35.52		

The reduction rate, weights, carbon momentum, green intensity and active share metrics are expressed in %, the MD metrics are measured in years, and the DTS metrics are calculated in bps.

Empirical results

The following empirical results are taken from Barahhou *et al.* (2022) and Ben Slimane *et al.* (2023b). First, we examine the dynamic decarbonization of equity portfolios. In Figure 11.20, we show the relationship between time and tracking error volatility with respect to Scope emissions when

Figure 11.20: Tracking error volatility of dynamic decarbonized portfolios (MSCI World, June 2022, $\mathcal{C}_0$ constraint)

Figure 11.21: Tracking error volatility of dynamic decarbonized portfolios (MSCI World, June 2022, $\mathcal{C}_3(0, 10, 2)$ constraint)

Source: MSCI (2022), Trucost (2022) & Barahhou *et al.* (2022).

considering the CTB and PAB decarbonization pathways and the MSCI World universe. Including Scope 3 has a significant impact on the tracking risk, especially when upstream Scope 3 emissions are considered. On average, the inclusion of Scope 3 results in a multiplication of the tracking error volatility by a factor of three. Barahhou *et al.* (2022) reported results considering the $\mathcal{C}_3(0,2,1)$ constraint, which imposes sector neutrality. They showed that the solution may not exist even before 2030 for the PAB decarbonization pathway. In order to have acceptable solutions, the ETF industry generally uses the $\mathcal{C}_3(0,10,2)$ constraint (Figure 11.21).

The previous analysis deals only with the decarbonization dimension. Barahhou *et al.* (2022) then introduced the transition dimension and solved the following optimization problem:

$$
\begin{aligned}
w^\star(t) &= \arg\min \frac{1}{2}(w - b(t))^\top \Sigma(t)(w - b(t)) \\
&\text{s.t.} \begin{cases} \mathcal{CI}(t,w) \le (1 - \mathcal{R}(t_0,t))\,\mathcal{CI}(t_0,b(t_0)) \\ \mathcal{CM}(t,w) \le \mathcal{CM}^\star(t) \\ \mathcal{GI}(t,w) \ge (1 + \mathcal{G})\,\mathcal{GI}(t,b(t)) \\ w \in \mathcal{C}_0 \cap \mathcal{C}_3(0,10,2) \end{cases}
\end{aligned}
$$
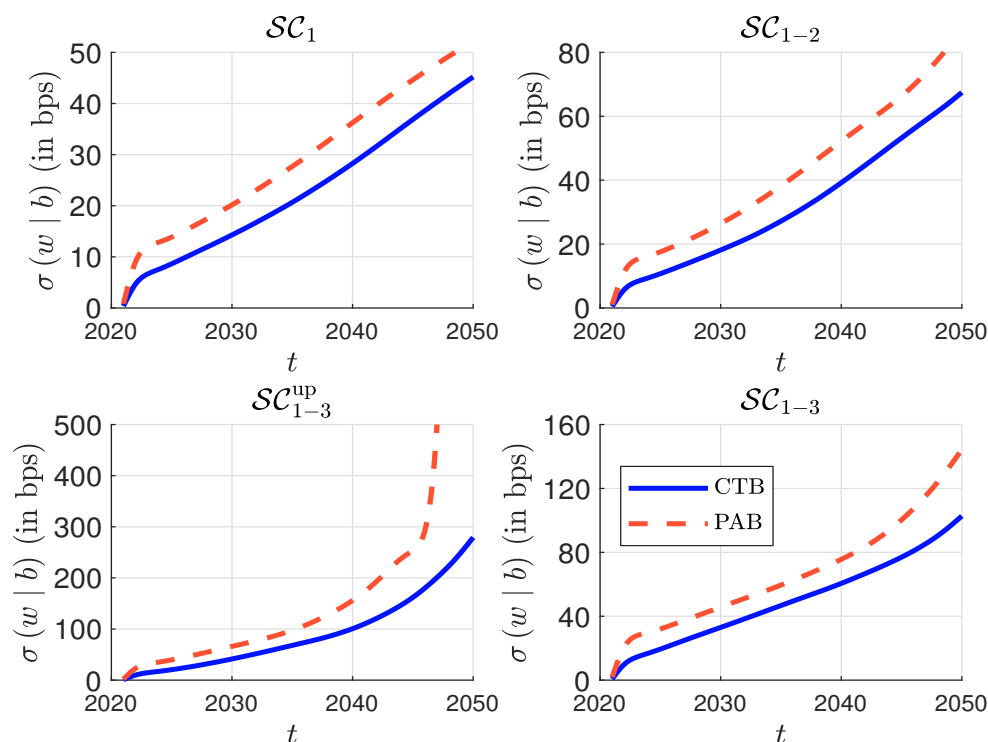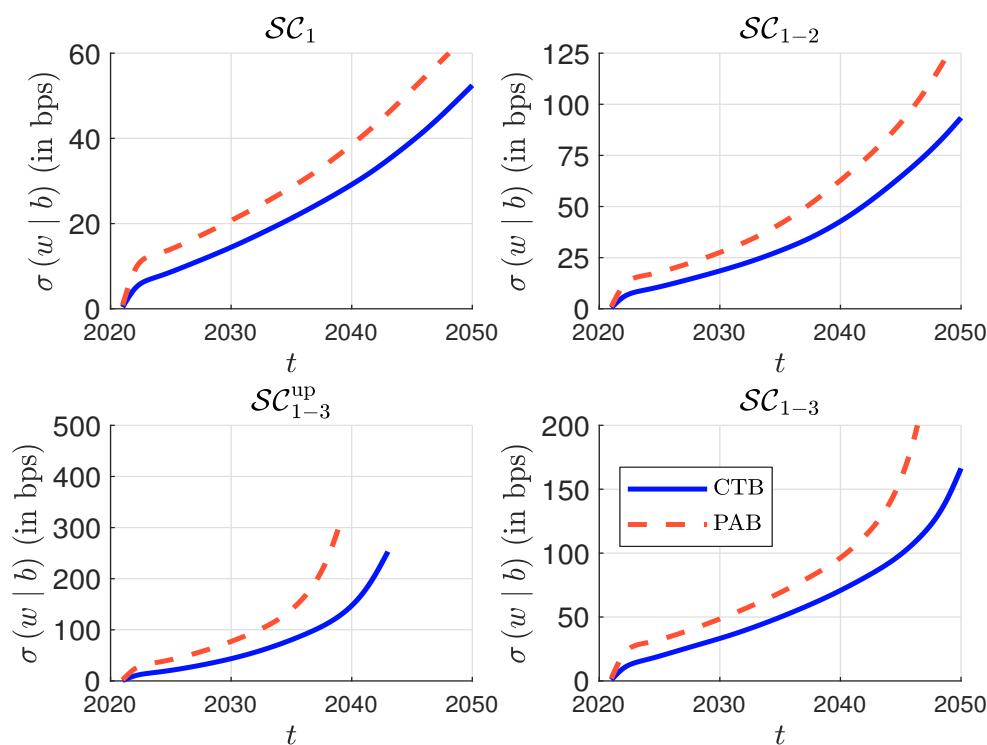
where $\mathcal{CM}^\star(t) = -5\%$ and $\mathcal{G} = 100\%$. In Figure 11.23, we plot the relationship between time t and the tracking error volatility $\sigma(w^\star(t) \mid b(t))$, measured in bps, when considering the PAB decarbonization pathway[33]. We also report the decomposition between the decarbonization and transition dimensions. The results of these simulations clearly show that the transition dimension induces significant and additional costs. On average, we observe that the additional cost of the tracking error for the years 2022-2030 is 27, 25, 21 and 19 bps for Scopes $\mathcal{SC}_1$, $\mathcal{SC}_{1-2}$, $\mathcal{SC}_{1-3}^{\mathrm{up}}$ and $\mathcal{SC}_{1-3}$, respectively. Moreover, there may be no solution to the optimization problem by 2050, especially if the carbon footprint is based on upstream/downstream Scope 3 emissions. Of course, all these results are very sensitive to the choice of the green multiplier $\mathcal{G}$ and the carbon threshold $\mathcal{CM}^\star$.

The previous results are valid for the MSCI World index, which is a large investment universe with more than 1 500 stocks. Let us focus on smaller investment universes by considering the MSCI EMU and USA indexes. The results are shown in Figures 11.24 and 11.25. The tracking error volatilities for smaller universes become larger in fewer years than for the MSCI World index, and we also fail to find solutions sooner. We could separate these results by putting the Scope 1 and 2 alignment on one side and Scope 3 on the other. Looking at Scopes 1 and 2, we see that in both universes the aligned portfolio breaks earlier than the MSCI World. However, even though the MSCI EMU universe is smaller than the MSCI USA universe, we can find solutions for a longer period of time. This is due to the distribution of green revenues and carbon dynamics, which are easier to reconcile with the intensity reduction constraint for the EMU. The inclusion of Scope 3 intensities paints a different picture. Although the EMU net-zero portfolios have lower tracking errors than the USA net-zero portfolios, larger universes tend to provide longer solutions. The fact that we are unable to align the EMU portfolio in terms of Scope 3 carbon intensities after 2040 therefore highlights the difficulty of portfolio alignment for a relatively small investment universe.

In practice, many constraints can be used to construct net-zero portfolios. We have seen above that the cost of tracking error can be significant and that the solution may not exist for long time horizons. Because the net-zero portfolio excludes some assets, it may be more concentrated than the benchmark. Therefore, we may face not only diversification risk, but also liquidity risk. These risks will be reduced as the economy decarbonizes in the coming years. However, we are not immune to

---

[33]To assess the impact of the weight/sector constraint $\mathcal{C}_3(0,10,2)$, we report the results based on the $\mathcal{C}_0$ constraint only in Figure 11.22.

Figure 11.22: Tracking error volatility of net-zero portfolios (MSCI World, June 2022, $\mathcal{C}_0$ constraint, $\mathcal{G} = 100\%$, $\mathcal{CM}^\star = -5\%$, PAB)
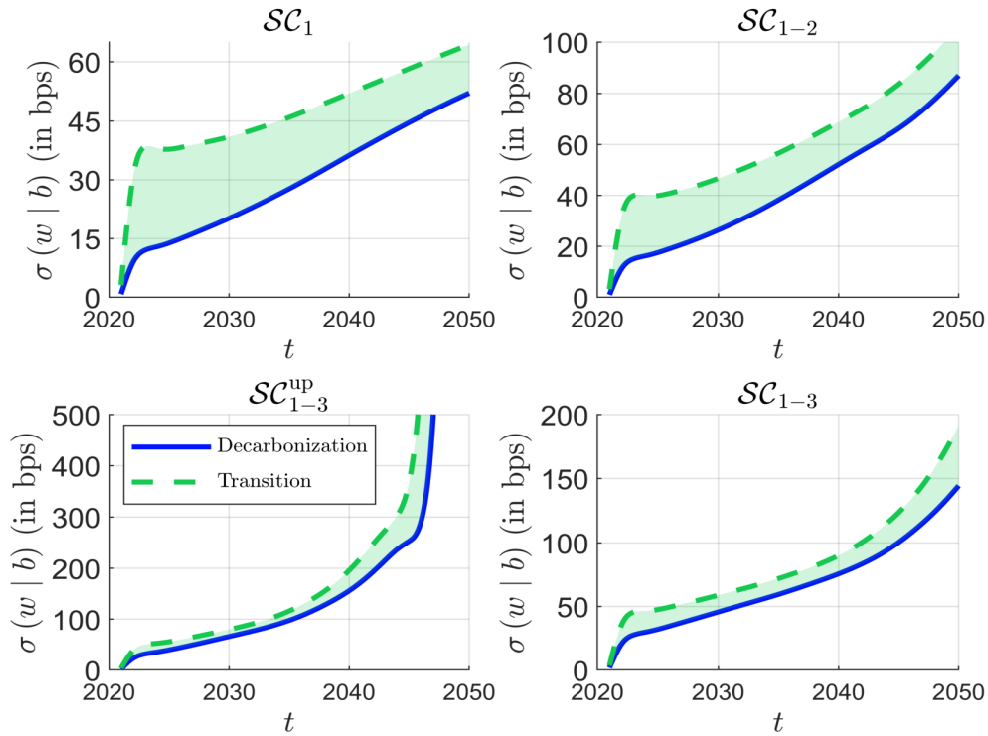
Figure 11.23: Tracking error volatility of net-zero portfolios (MSCI World, June 2022, $\mathcal{C}_3\,(0, 10, 2)$ constraint, $\mathcal{G} = 100\%$, $\mathcal{CM}^\star = -5\%$, PAB)
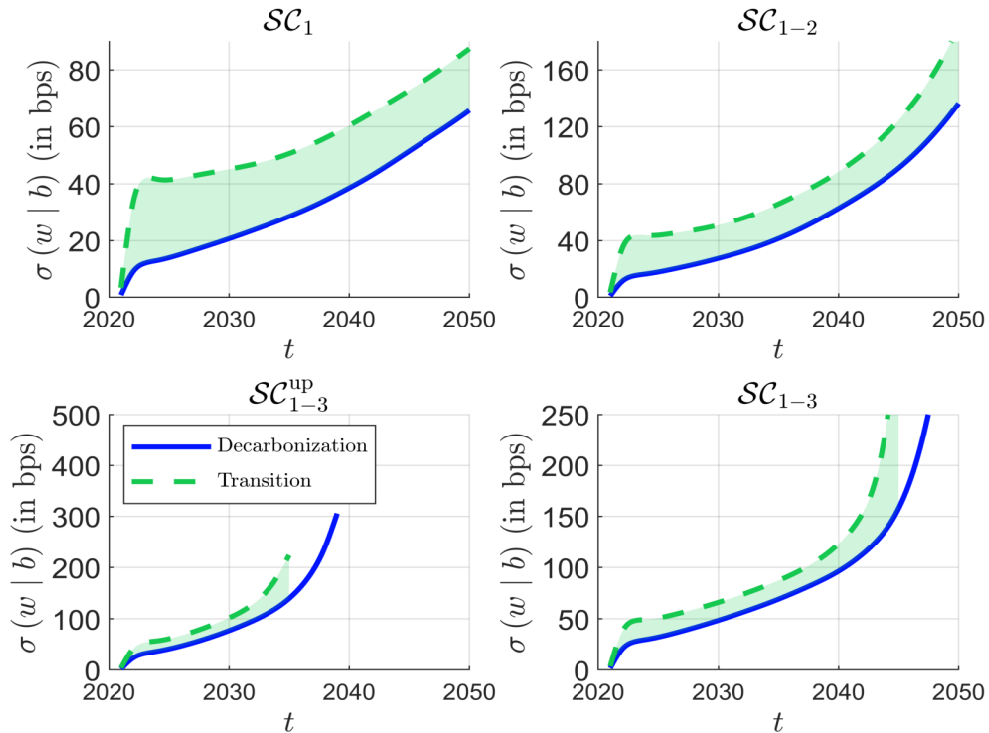
Figure 11.24: Tracking error volatility of net-zero portfolios (MSCI EMU, June 2022, $\mathcal{C}_3\,(0, 10, 2)$ constraint, $\boldsymbol{\mathcal{G}} = 100\%$, $\boldsymbol{\mathcal{CM}}^\star = -5\%$, PAB)
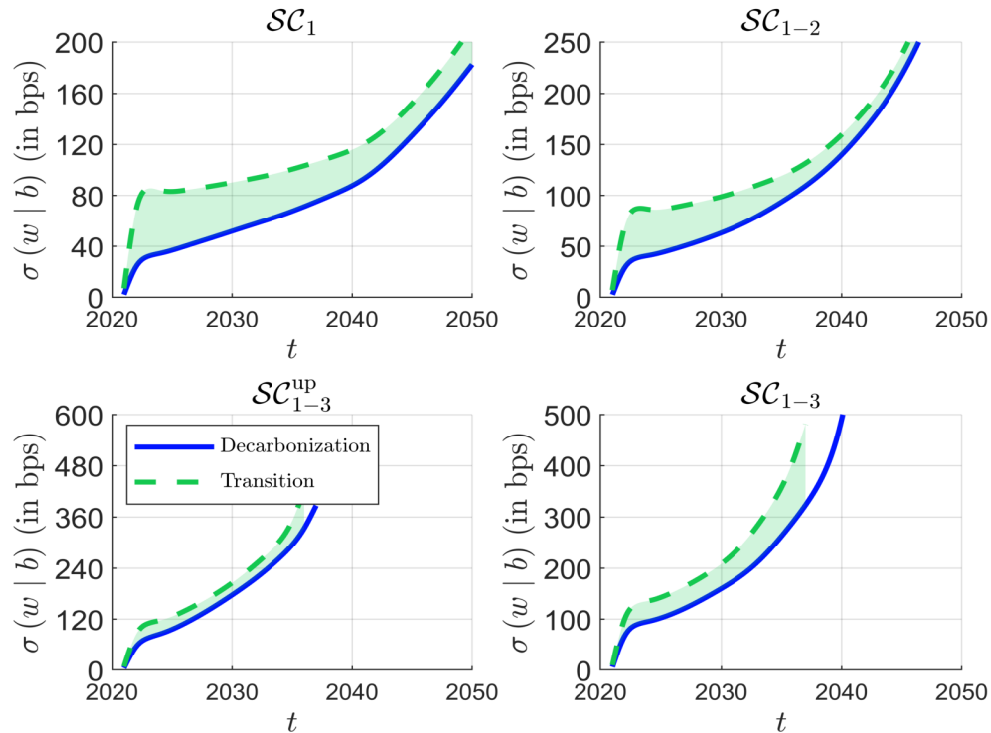
Figure 11.25: Tracking error volatility of net-zero portfolios (MSCI USA, Jun. 2022, $\mathcal{C}_3\,(0, 10, 2)$ constraint, $\boldsymbol{\mathcal{G}} = 100\%$, $\boldsymbol{\mathcal{CM}}^\star = -5\%$, PAB)
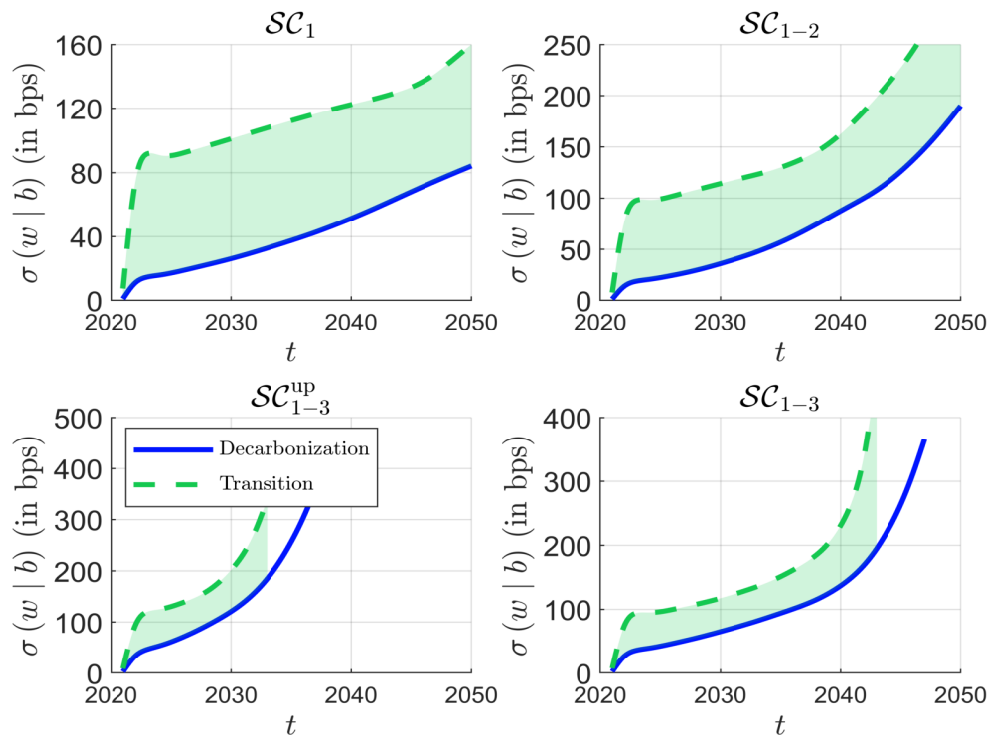
Source: MSCI (2022), Trucost (2022) & Barahhou *et al.* (2022).

the possibility that carbon emissions will continue to rise in the short term. In this case, solutions will be very sensitive to the gap between the carbon target of net-zero portfolios and the carbon footprint of the economy. To illustrate the shrinkage risk of the investment universe, we calculate the number of stocks selected per sector for each optimized portfolio and divide these figures by the corresponding number of stocks in the index[34]. In the case of Scope $\mathcal{SC}_{1-3}$, the radar chart of these frequencies is shown in Figure 11.26. We observe that the investment universe is shrunk at the first date. The green area represents the removed part by 2030. With the exception of the communication services, financials, health care, information technology and real estate sectors, the investment in the other sectors is concentrated on few stocks. This shrinkage effect is also observed for small investment universes. By construction, the shrinkage of the investment universe worsens if we add other constraints. For instance, the impact of the momentum exclusion constraint is illustrated in Figure 11.27. In this case, we complete the set of constraints by the exclusion constraint $\{\mathcal{CM}_i(t) \geq 0 \Rightarrow w_i = 0\}$, meaning that we exclude issuers with a positive carbon trend. We notice that the investment universe is highly reduced even from the first year. These results show that we cannot reduce the cost of net-zero investing to the cost of tracking risk. As seen above, there is also a cost of diversification risk. There is also a liquidity risk, as illustrated by Barahhou *et al.* (2022). Indeed, they showed that the repartition between large, mid and small caps changes. In some particular cases, they observed that the allocation to small- and micro-cap buckets increases over time.

Figure 11.26: Radar chart of investment universe shrinkage (MSCI World, June 2022, $\mathcal{C}_3(0, 10, 2)$ constraint, $\mathcal{G} = 100\%$, $\mathcal{CM}^\star = -5\%$, PAB, Scope $\mathcal{SC}_{1-3}$)
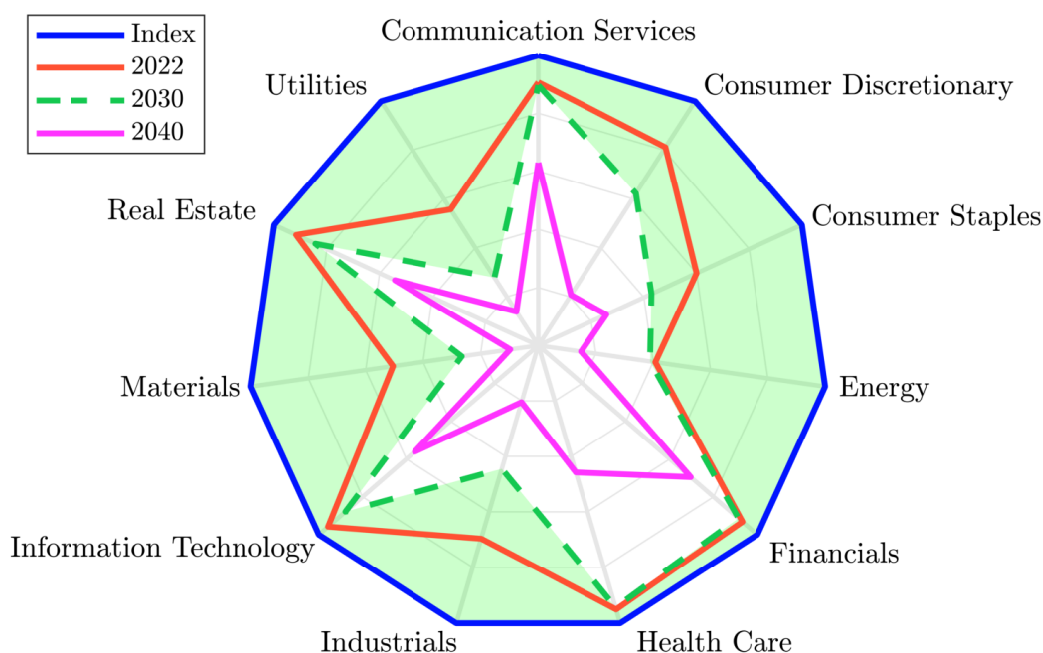
Source: MSCI (2022), Trucost (2022) & Barahhou *et al.* (2022).

The results obtained by Barahhou *et al.* (2022) are based on simulations and do not take into account all the investment constraints that may be encountered in a net-zero portfolio, such as ESG

---

[34]For example, if the frequency for the energy sector is 25%, this means that the optimized portfolio selected 25% of the energy stocks and removed 75% of the energy investment universe.

constraints and the exclusion of some activities. Therefore, the estimated cost of tracking error risk can be viewed as a lower bound on the cost of real net-zero investing.

Figure 11.27: Impact of momentum exclusion on universe shrinkage (MSCI World, June 2022, $\mathcal{C}_3(0,10,2)$ constraint, $\mathcal{G} = 100\%$, $\mathcal{CM}^{\star} = -5\%$, PAB, Scope $\mathcal{SC}_{1-3}$, $\mathcal{CM}^{+} = 0\%$)
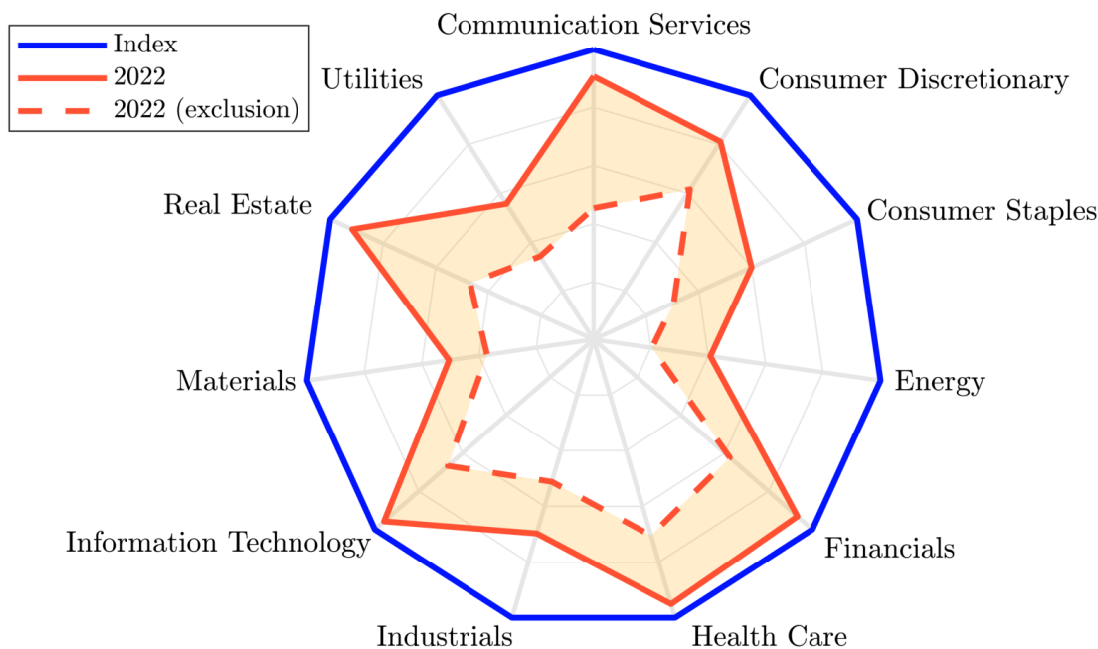
Source: MSCI (2022), Trucost (2022) & Barahhou *et al.* (2022).

We now turn to the case of corporate bonds. Using the CTB and PAB decarbonization scenarios, we get the results in Figures 11.28 and 11.29. The DTS risk is not significant and is less than 6 bps until 2030. This is not the case for the active share risk, which can reach 20% in 2030 for the PAB decarbonization pathway. When we include the transition constraints, the additional cost seems relatively low compared to what we have observed for equity investment universes. For example, Barahhou *et al.* (2022) found that the DTS tracking risk and active share increase by less than 1 bp and 1%, respectively, when $\mathcal{G}$ is set to 100%. However, they found that the exclusion constraint can significantly increase costs. For example, requiring issuers to have a negative carbon momentum increases the tracking risk by 20% on average.

The case of government bonds is studied by Barahhou *et al.* (2023). As with corporate bonds, the integrated approach for sovereign bonds consists of several steps: (1) we need to define the decarbonization scenario at the country level; (2) we can assess the self-decarbonization of a country by considering the government's credible commitments and decarbonization plans towards a low-carbon economy; (3) a specific green intensity measure needs to measure the country's contribution to the climate transition and its greenness. Unlike corporate bonds, there are many options for choosing the net-zero scenario. We have already discussed some of them in Section 8.3.3 on page 381. We can choose between the IPCC scenarios, the NGFS scenarios, the IEA scenarios, etc. Barahhou *et al.* (2023) explained that the stated policies scenario (STEPS) and the announced pledges scenario (APS) cannot be used because they are not consistent with limiting global warming to 1.5°C. It is better to use the NZE scenarios from NGFS or IEA. Figure 11.30 shows the decarbonization pathway derived from the IEA NZE scenario (IEA, 2021, Figure 2.2, page 53). We distinguish between advanced economies, and emerging market and developing economies. Using 2020 as a

Figure 11.28: Duration-times-spread cost of dynamically decarbonized portfolios (Global Corporate, June 2022)
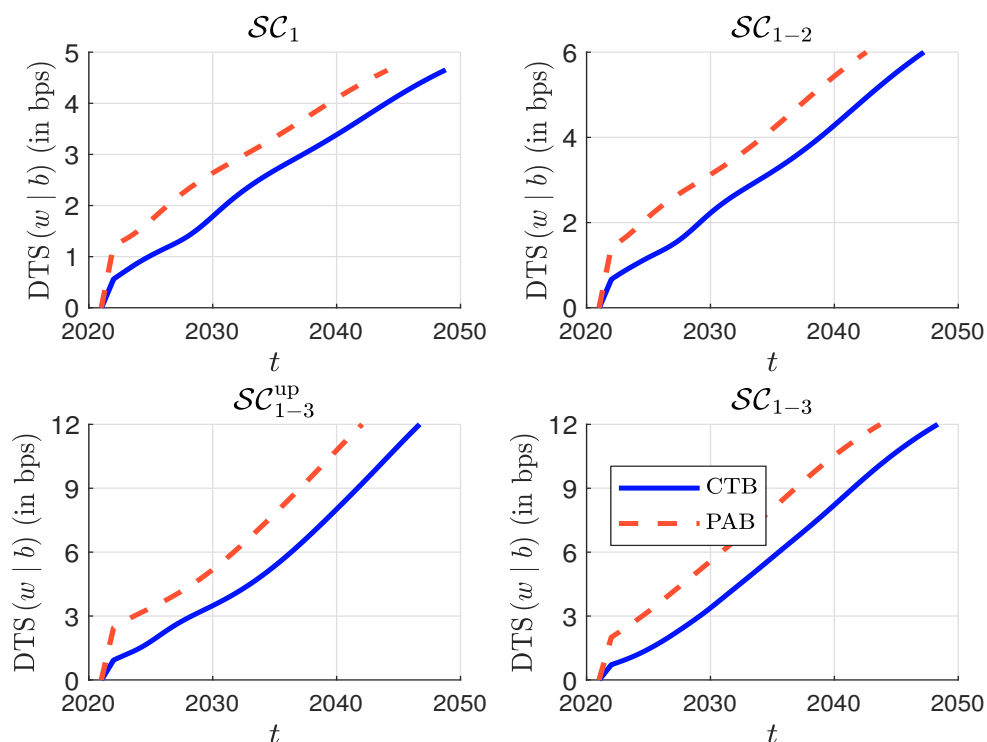
Figure 11.29: Active share of dynamically decarbonized portfolios (Global Corporate, June 2022)
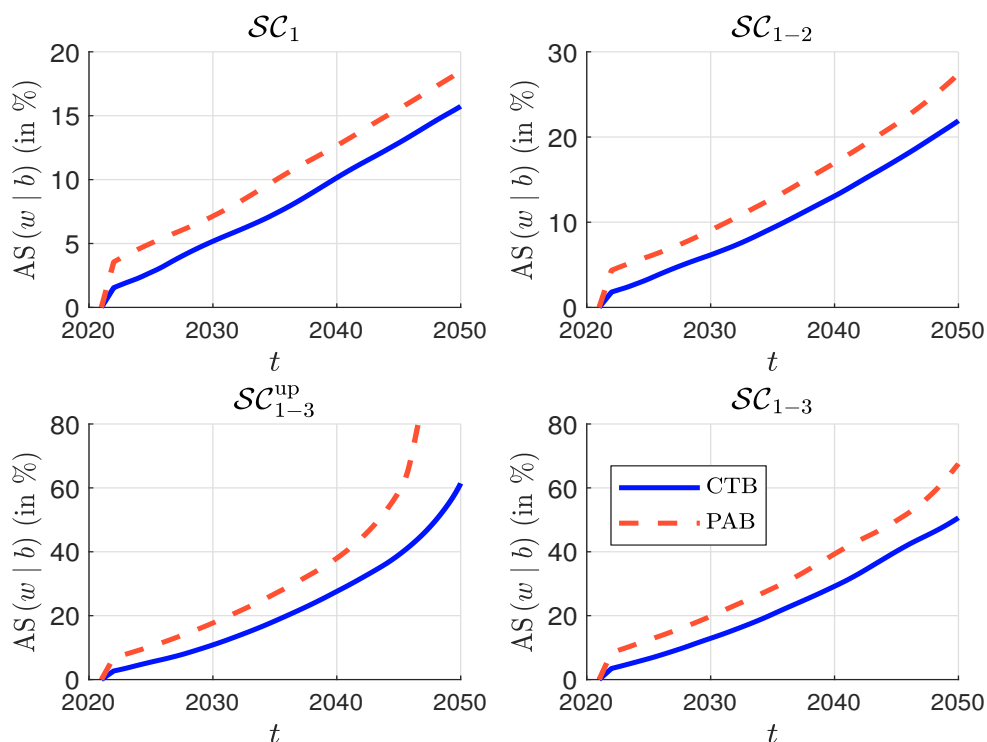
Source: ICE (2022), Trucost (2022) & Barahhou *et al.* (2022).

baseline, the decarbonization pathways in terms of carbon emissions are similar for developed and developing countries. This is not the case when looking at carbon intensity, which corresponds to CO_2 emissions per capita.

Figure 11.30: IEA decarbonization pathways
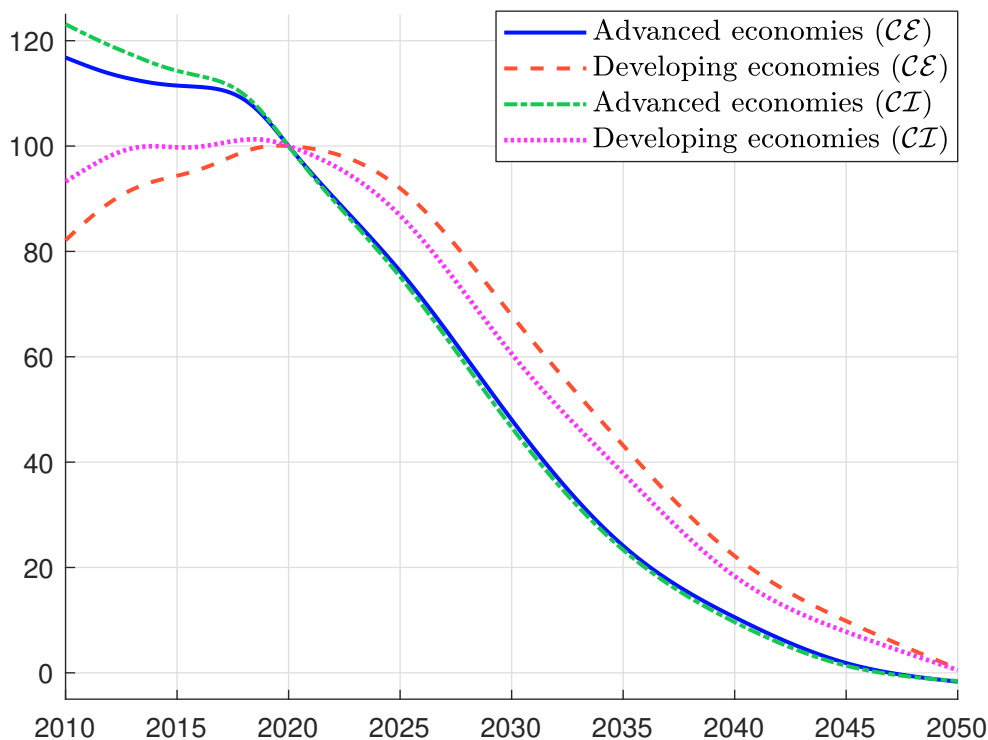
Choosing the scope of emissions and how to measure the carbon footprint is another important issue. As seen in Section 8.4.4 on page 436, we can define carbon emissions at the country level by considering a production-based or consumption-based inventory. If we prefer a carbon intensity measure, the normalization variable can be population, GDP, or public debt. Barahhou *et al.* (2023) showed that the distribution of carbon intensity metrics across countries is very different for these three measures. For the green footprint, they listed several metrics: government spending data[35] from the IEA, government spending on environmental protection[36] from the IMF's Climate Change Dashboard, the amount of green bonds issued, and environmental taxes[37] by country from the IMF's Climate Change Dashboard. Using data on carbon emissions and commitments, they also derived four types of forward-looking metrics: (1) carbon trend, (2) nationally determined contribution (NDC), (3) NDC ambition, and (4) NDC fulfillment. In particular, Barahhou *et al.* (2023) defined two simple criteria:

---

[35]The website is https://www.iea.org/reports/government-energy-spending-tracker-2.

[36]This database contains the following seven time series: (1) Environmental protection expenditure, (2) Biodiversity and landscape protection expenditure, (3) Expenditure on environmental protection n.e.c., (4) Environmental protection R&D expenditure, (5) Pollution abatement expenditure, (6) Waste management expenditure and (7) Waste water management expenditure. The website for downloading the data is https://climatedata.imf.org/datasets/d22a6decd9b147fd9040f793082b219b_0/explore.

[37]This database contains the following five time series: (1) Environmental taxes, (2) Taxes on energy (including fuels for transport), (3) Taxes on pollution, (4) Taxes on resources and (5) Taxes on transport (excluding fuels for transport). The website for downloading the data is https://climatedata.imf.org/datasets/3fb1ed30d3394574b3145246846023b1_0/explore.

- Commitments aligned with the NZE scenario (CAS) imply that countries have NDC target emissions that fall within a certain range around the value projected by the NZE scenario;

- Emissions on track with commitments (EOTC) indicate countries with historical trends in line with their NDC commitments.

These two metrics measure the ambition and credibility of the $\mathcal{PAC}$ framework. In Table 11.29, we report the CAS and EOTC statistics found by Barahhou *et al.* (2023). Only 6.6% of countries have sufficient ambition and credibility consistent with the NZE scenario, while about 70% of countries do not meet these two criteria.

Table 11.26: CAS and EOTC statistics on Bloomberg Global Aggregate Treasuries

CAS	✓	✓		No criteria
EOTC	✓		✓	met
Frequency	6.6%	23.3%	0.7%	69.5%

Source: Barahhou *et al.* (2023, Figure 14, page 23).

Very quickly, there is no solution to the optimization problem after 2030, which means that the current ambitions of countries are not sufficient to build a net-zero sovereign bond portfolio in the long run. In Table 11.27, we report the authors' estimated first year of country exit for a given set of parameters and the GHG/GDP intensity metric. Some countries are removed from the portfolio in 2024 such as Canada, Indonesia, New Zealand and South Korea, but most of developed countries exit the portfolio in 2029. Changing the green constraint could delay the first year of exit by a year or two, but the problem is the constraint of the decarbonization pathway, which is impossible to manage after 2032.

Table 11.27: First year of country exit from the NZE investment portfolio (GHG/GDP intensity metric)

Australia	2025	Finland	2029	Lithuania	2025	Romania	2029
Austria	2029	France	2029	Luxembourg	2029	Singapore	2029
Belgium	2028	Germany	2029	Mexico	2029	Slovakia	2025
Canada	2024	Hong Kong	2029	Malaysia	2028	Slovenia	2028
Chile	2029	Hungary	2029	Malta	2029	South Korea	2024
China	2028	Indonesia	2024	Netherlands	2029	Spain	2028
Colombia	2029	Ireland	2029	Norway	2029	Switzerland	2029
Cyprus	2029	Israel	2029	New Zealand	2024	Sweden	2029
Czechia	2024	Italy	2029	Peru	2029	Thailand	2025
Denmark	2029	Japan	2029	Poland	2029	United Kingdom	2029
Estonia	2025	Latvia	2028	Portugal	2028	United States	2028

Source: Barahhou *et al.* (2023, Table 9, page 26).

The effect of the intensity metric is also significant. Table 11.28 shows when the country is removed from the NZE investment portfolio. We observe significant differences when we use GHG per GDP or GHG per capita. In addition, the use of a production-based or a consumption-based inventory also affects the results[38]. However, regardless of the metrics used, we generally found that a net-zero sovereign bond portfolio tended to overweight European countries.

[38]In this case, estimated metrics are generally available for CO_2 emissions than for all GHG emissions because

Table 11.28: Country exclusion year by intensity metric

Metric	GHG / GDP	GHG / Population	CO$_2$ (production) / GDP	CO$_2$ (consumption) / Population
China	2028	2031	2027	2031
France	2029	2032	2027	2031
Indonesia	2024	2032	2024	2031
Ireland	2029	2030	2027	2030
Japan	2029	2032	2027	2031
United States	2028	2030	2026	2029
United Kingdom	2029	2032	2027	2031
Sweden	2029	2032	2027	2031

Source: Barahhou *et al.* (2023, Table 14, page 31).

The previous empirical results for equities, corporate bonds, and sovereign bonds suggest the following lessons. First, the solution is parameter and data sensitive. In particular, we need to be careful in choosing the carbon scope metric to assess the decarbonization rate. A net-zero investment policy only makes sense for a closed system. Therefore, Scope 3 emissions need to be taken into account to align a portfolio with a net-zero scenario. The problem is that we see a lack of data reliability on Scope 3 emissions today. Similarly, the solution is highly dependent on the green intensity target and the level of self-decarbonization we want to achieve. Then we have to be careful because there may be no solution to the optimization problem in the medium term. The question of no solution depends on the relative speed of the portfolio's decarbonization path relative to the economy's decarbonization path and the initial starting point. The second key finding is that portfolio alignment (or decarbonization) and net-zero construction lead to different solutions. In particular, decarbonizing a portfolio is easier than constructing a net-zero portfolio. We find that decarbonizing along CTB or PAB pathways never leads to exploding tracking errors by 2030. In fact, the real problem with decarbonization is the diversification and liquidity risk that an investor may face. These results are amplified when we add the transition dimension to the optimization program. In addition to higher tracking risk, there is no guarantee that there will always be a solution. Moreover, the introduction of the transition pillar highlights the difficulty of choosing an appropriate set of constraints for net-zero portfolios, as some metrics may be negatively correlated with others. Portfolio decarbonization is systematically a strategy that is long financial issuers and short energy, materials and utilities issuers. Therefore, we have a situation where the transition dimension of a decarbonized portfolio is weaker than that of the benchmark portfolio, as green solutions are also located in carbon-intensive sectors. It is therefore crucial to distinguish between issuers with a high carbon footprint that will not participate in the transition and those that will reduce their carbon emissions and find low-carbon solutions. The third key finding is that portfolio decarbonization and alignment are two processes of exclusion. This means that it is quite impossible to achieve net-zero alignment without allowing the algorithm to exclude companies (or countries) from the benchmark. For example, the optimization program will generally not find a solution if it imposes non-zero lower bounds. As a result, some key players in the transition, such as energy and utility companies, unfortunately disappear. Furthermore, imposing sector neutrality can lead

---

consumption-based estimates are obtained with input-output analysis. In addition, it is traditionally accepted that the intensity measure uses GDP as the normalization variable for the production-based inventory, because it depends on national production, and the population for the consumption-based inventory, because it depends on population size.

to similar problems in finding a solution. The final lesson is that it is easier to implement net-zero in bonds than in equities. At first glance, this result may seem surprising, since there is no reason why net-zero should affect the equity and bond markets differently. In fact, there are two possible explanations. First, the structure of equity and bond indices is different, with the latter having a more balanced allocation across sectors and a high exposure to financial issuers. Second, bond indices are strongly influenced by new fresh capital, while equity indices are sticky to the stock of existing capital. This is because the primary bond market is very active, which implies a significant impact on the secondary market. Indeed, bonds mature and are replaced by new, greener bonds. The primary market then helps to achieve net-zero in bonds. This is not the case in the stock market, where IPOs and capital increases are only a small part of the secondary market. This means that portfolio holdings change faster for bond indices than for equity indices. Therefore, the greenness of bond indices increases faster than the greenness of equity indices. All of these factors suggests that the cost of implementing net-zero investments relative to traditional investments will be higher for equity portfolios than for bond portfolios, and that the bond market will benefit more quickly from the transition to a low-carbon economy.

11.3.2 Core-satellite approach

We have seen that the comprehensive integrated approach can sometimes be difficult to implement because today, on average, carbon intensities are positively correlated with green intensities. This means that the greenness of the economy is not necessarily found in companies with low carbon footprints. Therefore, a second approach has emerged that is easier to implement. It consists of adopting a core-satellite strategy, where decarbonization is applied to the core portfolio, while the objective of the satellite portfolio is to finance the transition to a low-carbon economy. In the financial literature, the core portfolio is called the net-zero decarbonization portfolio, while the satellite portfolio is called the net-zero contribution portfolio, but other terms are used, as shown in Table 11.29.

Table 11.29: The two building block approach

Decarbonizing the portfolio	Financing the transition
• Net-zero decarbonization portfolio	• Net-zero contribution portfolio
• Net-zero transition portfolio	• Net-zero funding portfolio
• Dynamic low-carbon portfolio	• Net-zero transformation portfolio

This is equivalent to splitting the problem into two sub-problems. The goal of the first sub-problem is to decarbonize and manage the carbon footprint of the investment. The goal of the second sub-problem is to contribute to increasing the green footprint of the economy. The two sub-problems are summarized in Table 11.30. The core portfolio is more of a top-down allocation process and exclusion strategy, where the central climate risk metric is carbon intensity. The satellite portfolio is more of a bottom-up allocation process and asset selection strategy, where the central climate risk metric is green intensity. This approach also has the advantage of making the allocation between the two net-zero strategies clear. Of course, the allocation $\alpha(t)$ to the satellite can be dynamic and change over time as the world and economy progresses towards net-zero.

Table 11.30: The core-satellite approach

Decarbonized portfolio	Contribution portfolio
• Carbon intensity	• Green intensity
• Decarbonization pathway	• Financing the transition
• Top-down approach	• Bottom-up approach
• Portfolio construction	• Security selection
• Net-zero carbon metrics	• Net-zero transition metrics

$$\mathbf{1-\alpha(t)} \qquad\qquad\qquad \mathbf{\alpha(t)}$$
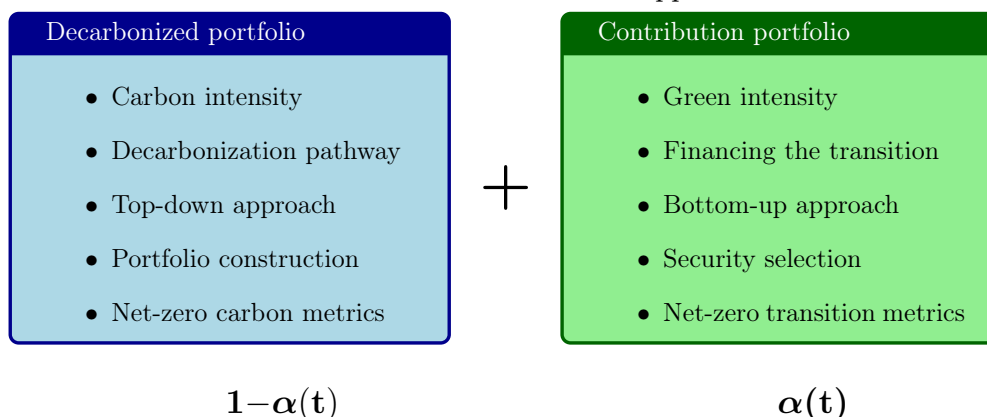
Core portfolio

The core portfolio is the largest part of the net-zero investment portfolio and the objective is to manage the decarbonization dimension. Therefore, only carbon metrics are used to assess the carbon footprint of the portfolio and its trajectory. If we write the alignment process in terms of portfolio optimization, a typical program for the equity bucket looks like this:

$$
\begin{aligned}
w^\star(t) \quad=\quad & \arg\min \frac{1}{2}\left(w-b\left(t\right)\right)^\top \Sigma\left(t\right)\left(w-b\left(t\right)\right) \qquad\qquad (11.24)\\[4pt]
& \text{s.t.}
\begin{cases}
\mathcal{CI}\left(t,w\right) \leq \left(1-\mathcal{R}\left(t_0,t\right)\right)\mathcal{CI}\left(t_0,b\left(t_0\right)\right)\\
\mathcal{CM}\left(t,w\right) \leq \mathcal{CM}^\star\left(t\right)\\
0 \leq w_i \leq \mathbb{1}\left\{\mathcal{CM}_i\left(t\right) \leq \mathcal{CM}^+\right\}\\
w \in \Omega_0 \cap \Omega
\end{cases}
\end{aligned}
$$

This is exactly the optimization problem defined in Equation (11.22), except that the green intensity constraints are removed because the green footprint is managed through the satellite portfolio. For the bond bucket, we get a similar optimization problem:

$$
\begin{aligned}
w^\star(t) \quad=\quad & \arg\min \mathcal{D}\left(w \mid b\left(t\right)\right) \qquad\qquad\qquad\qquad\qquad (11.25)\\[4pt]
& \text{s.t.}
\begin{cases}
\mathcal{CI}\left(t,w\right) \leq \left(1-\mathcal{R}\left(t_0,t\right)\right)\mathcal{CI}\left(t_0,b\left(t_0\right)\right)\\
\mathcal{CM}\left(t,w\right) \leq \mathcal{CM}^\star\left(t\right)\\
0 \leq w_i \leq \mathbb{1}\left\{\mathcal{CM}_i\left(t\right) \leq \mathcal{CM}^+\right\}\\
w \in \Omega_0 \cap \Omega
\end{cases}
\end{aligned}
$$

In both cases, the implementation of the portfolio alignment process considers a global decarbonization scenario constraint, a self-decarbonization constraint, and an exclusion constraint. Ben Slimane *et al.* (2023b) also suggested that a number of specific sectoral decarbonization pathways can be considered, consistent with the International Energy Agency's sequencing principles. In particular, the issue of electricity is central to the net zero scenario. Electricity has to be green by 2035. From a financial perspective, this means that we will have to finance the transition of the power sector on a massive scale. However, if we only implement a global decarbonization scenario, Ben Slimane *et al.* (2023b) noted that utility issuers are quickly excluded from the net-zero optimization program. The core net-zero portfolio invests in pure players that produce green power and doesn't give others time to transform their business models. This is not consistent with the goal of net-zero investing. Ben Slimane *et al.* (2023b) then proposed to constrain the optimization problem to follow the NZE

scenario for the electricity sector. Using the results on page 552, the constraint to meet a reduction rate for a given sector $\boldsymbol{Sector}_j$ can be expressed as:

$$\frac{\sum_{i=1}^{n} \mathbb{1} \{i \in \boldsymbol{Sector}_j\} w_i \boldsymbol{CI}_i}{\sum_{i=1}^{n} \mathbb{1} \{i \in \boldsymbol{Sector}_j\} w_i} = \boldsymbol{CI} (\boldsymbol{Sector}_j, \boldsymbol{R}_j)$$

where $\boldsymbol{CI} (\boldsymbol{Sector}_j, \boldsymbol{R}_j)$ is the carbon intensity target for the given sector:

$$\boldsymbol{CI} (\boldsymbol{Sector}_j, \boldsymbol{R}_j) = (1 - \boldsymbol{R}_j) \frac{\sum_{i=1}^{n} \mathbb{1} \{i \in \boldsymbol{Sector}_j\} b_i \boldsymbol{CI}_i}{\sum_{i=1}^{n} \mathbb{1} \{i \in \boldsymbol{Sector}_j\} b_i}$$

We deduce that:

$$\sum_{i=1}^{n} \mathbb{1} \{i \in \boldsymbol{Sector}_j\} w_i \boldsymbol{CI}_i = \boldsymbol{CI} (\boldsymbol{Sector}_j, \boldsymbol{R}_j) \sum_{i=1}^{n} \mathbb{1} \{i \in \boldsymbol{Sector}_j\} w_i$$

which is equivalent to the following constraint:

$$\sum_{i=1}^{n} \mathbb{1} \{i \in \boldsymbol{Sector}_j\} w_i (\boldsymbol{CI}_i - \boldsymbol{CI} (\boldsymbol{Sector}_j, \boldsymbol{R}_j)) = 0 \Leftrightarrow \left(\boldsymbol{s}_j \circ \left(\boldsymbol{CI}_i - \boldsymbol{CI}_j^\star\right)\right)^\top w = 0$$

where $\boldsymbol{CI}_j^\star = \boldsymbol{CI} (\boldsymbol{Sector}_j, \boldsymbol{R}_j)$. Since this is a linear equation, the previous optimization problems with this constraint remain a QP or LP problem and can be solved easily. In Figures 11.31 and 11.32, we reproduce the tracking error volatility $\sigma (w \mid b)$ of equity portfolios and the active risk $\mathcal{D} (w \mid b)$ of bond portfolios when we impose the IEA NZE scenario for the electricity sector, $\boldsymbol{CM}^\star = -3.5\%$ and $\boldsymbol{CM}^+ = 10\%$, and we consider different NZE decarbonization scenarios (IEA, NZAOA, CTB and PAB). In these simulations, Ben Slimane *et al.* (2023b) found that if we allocate the tracking risk between the different constraints, the two most important contributors are the global decarbonisation pathway and the exclusion constraints, while the costs of self-decarbonisation and the specific electricity decarbonisation pathway are relatively low.

Satellite portfolio

While the core portfolio aims to implement decarbonization policies, the satellite portfolio aims to finance the transition to a low-carbon economy and monitor green intensity. As the core-satellite approach is implemented in strategic asset allocation or multi-asset portfolios, the investment universe is diversified and typically consists of the following asset classes:

- Green, sustainability and sustainability-linked bonds

- Green stocks

- Green infrastructure

- Sustainable real estate

The net-zero transition analysis on page 545 shows that the list of major sectors to be financed is relatively small. The mapping to level 4 of the GICS classification is shown in Table 11.31. In fact, we identify 29 sub-industries out of the 163 included in the GICS classification. Looking at the major sub-industries that can be included in the net-zero satellite portfolio, four of the eleven GICS sectors are over-represented, two are included, and the other five are excluded. The four over-represented

Figure 11.31: Tracking error volatility of decarbonized portfolios (MSCI World, December 2021, $\mathcal{CM}^\star = -3.5\%$, $\mathcal{CM}^+ = 10\%$, IEA NZE electricity sector scenario)
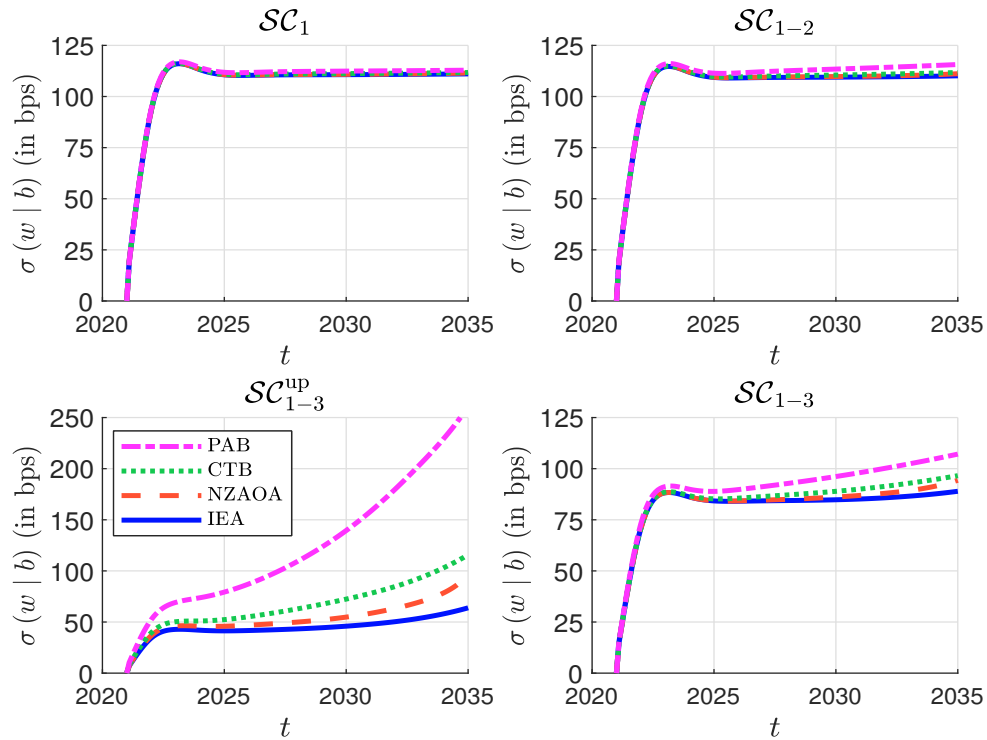
Figure 11.32: Active risk of decarbonized portfolios (Global Corporate, December 2021, $\mathcal{CM}^\star = -3.5\%$, $\mathcal{CM}^+ = 10\%$, IEA NZE electricity sector scenario)
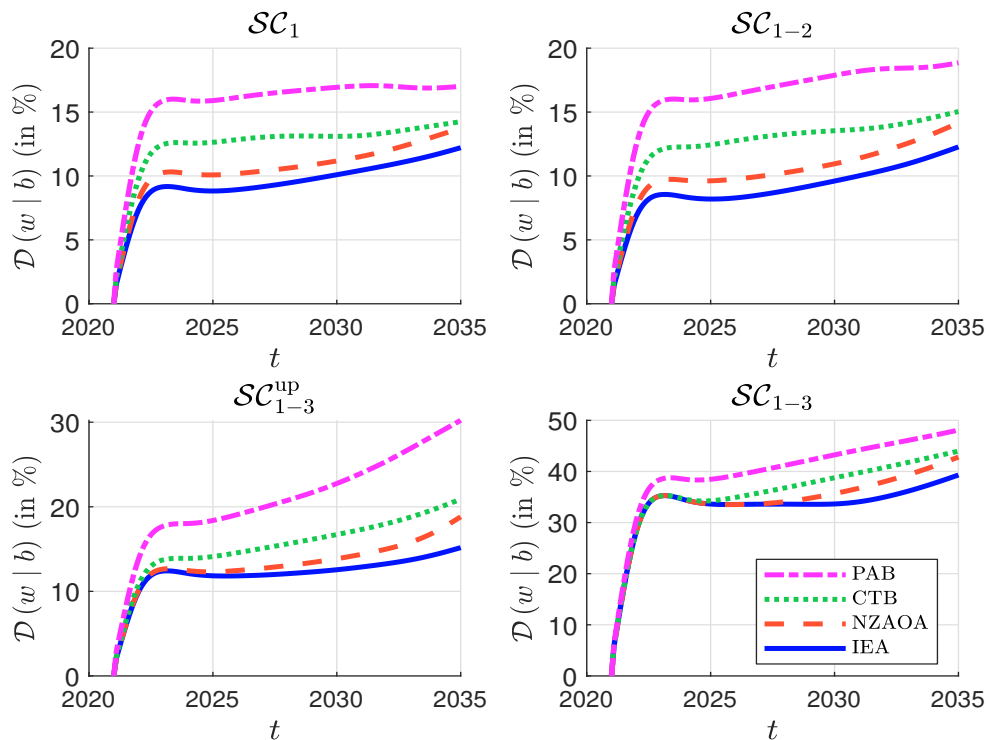
Source: Ben Slimane et al. (2023b).

Table 11.31: Main sub-industries of the net-zero satellite portfolio (GICS level 4)

Code	Level 4	Level 2	Level 1
15102010	Construction Materials	Materials	Materials
15104010	Aluminium		
15104020	Diversified Metals & Mining		
15104025	Copper		
15104040	Precious Metals & Minerals		
15104045	Silver		
15104050	Steel		
20102010	Building Products	Capital Goods	Industrials
20103010	Construction & Engineering		
20104010	Electrical Components & Equipment		
20104020	Heavy Electrical Equipment		
20106010	Construction Machinery & Heavy Transportation Eqpt.		
20106015	Agricultural & Farm Machinery		
20201050	Environmental & Facilities Services	Commercial & Professional Services	
20304010	Rail Transportation	Transportation	
20305010	Airport Services		
20305020	Highways & Railtracks		
20305030	Marine Ports & Services		
25101010	Automotive Parts & Equipment	Automobiles & Components	Consumer Discretionary
25102010	Automobile Manufacturers		
25201010	Consumer Electronics	Consumer Durables & Apparel	
25201030	Homebuilding		
25201040	Household Appliances		
30202010	Agricultural Products	Services Food, Beverage & Tobacco	Consumer Staples
55101010	Electric Utilities	Utilities	Utilities
55103010	Multi-Utilities		
55104010	Water Utilities		
55105020	Renewable Electricity		
60201030	Real Estate Development	Real Estate Management & Development	Real Estate

Source: Ben Slimane *et al.* (2023b, Table 14, page 44).

sectors are Industrials, Materials, Consumer Discretionary and Utilities. The Industrials sector is divided into three industry groups: Capital Goods, which includes sub-industries related to machinery, equipment and construction; Transportation, where we look at how to improve public transportation systems and their infrastructure; and finally, Commercial & Professional Services, which includes the Environmental & Facilities Services sub-industry, which mainly covers waste management and pollution control. The Materials sector is characterized by the various materials used in the energy transition, such as aluminium, copper, steel, etc. The Consumer Discretionary sector is divided into two industry groups, Automobiles & Components, where we find auto manufacturers and auto parts, and Consumer Durables & Apparel, related to housing and appliances. Finally, the Utilities sectors will show different types of utilities needed for the transition, such as electric utilities, water utilities, or even renewable energy. Then we include two sub-industries that belong to the Consumer Staples and Real Estate sectors, respectively: agricultural products and real estate development. In summary, this means that not all sectors are represented. Figure 11.33 shows the four levels of the GICS classification[39] and indicates which sub-industry falls within the definition of the satellite investment universe.

Remark 102 *It is clear that the GICS classification is not relevant when considering a net-zero investing framework. For example, there is no sector such as electricity storage, hydrogen storage, photovoltaic electricity generation, wind electricity generation, nuclear electricity generation in existing plants, etc. The NACE classification is more appropriate, and has the advantage to be in line with the EU green taxonomy. Nevertheless, GICS is the classification used by investors.*

Figure 11.33: Narrow specification of the satellite investment universe

Source: Ben Slimane *et al.* (2023b, Figure 31, page 45).

---

[39]It is described in Box 11.1 on page 574.

Green bonds We look at the Bloomberg database of the GSS+ investment universe. For each bond, Bloomberg indicates whether it is a green, social, sustainability, sustainability-linked or conventional bond. Issue amounts are shown in Table 11.32. In 2022, 1 784 green bonds were issued for a total of \$531.6 bn. This represents 15% of the net-zero financing needs[40]. For the other categories, the amount to be issued in 2022 is equal to \$152.8 bn for social bonds, \$174.8 bn for sustainability bonds and \$144.3 bn for sustainability-linked bonds. As explained by Ben Slimane *et al.* (2023a), social bonds are not net-zero transition instruments, but more conventional bonds to finance social debt and social infrastructure. Therefore, if we look at a broad definition of the net-zero fixed income universe (green, sustainability and sustainability-linked bonds), we get a total of \$850.7 bn, which can be seen as the upper bound of current investment opportunities. This is less than 25% of the \$3.5 tn previously required to achieve net-zero.

Table 11.32: GSS+ bond issuance

Year	Green		Social		Sustainability		SLB	
	#	\$ bn	#	\$ bn	#	\$ bn	#	\$ bn
2022	1 784	531.6	542	152.8	614	174.8	382	144.3
2021	1 971	686.1	554	242.1	646	233.2	343	161.5
2020	1 076	291.2	273	172.0	308	154.8	47	16.5
2019	877	268.0	99	22.2	333	85.2	18	8.9
2018	582	165.3	48	16.5	52	22.1	1	2.2
2017	472	160.9	46	11.8	17	9.2	1	0.2
2016	285	99.7	14	2.2	16	6.6	0	0.0

Source: Bloomberg (2023), GSS+ Instrument Indicator & Author's calculations.

In Figure 11.34 we compare the performance of the Bloomberg Global Green Bond index with the performance of the Bloomberg Global Aggregate index. We see that there is a high tracking risk. Between January 2014 and September 2023, the volatility of the tracking error is 2.7%. Several factors explain this high figure: sector allocation, duration, credit risk, etc. Investors must therefore accept a higher active risk for the satellite portfolio than for the core portfolio.

Green stocks We can invest in green stocks through thematic funds, which focus on a specific theme, or through an equity basket. In general, the second approach is used by equity managers, while the first is preferred by other fund managers, especially multi-asset managers. The emergence of economic, social and technological megatrends is strongly influencing the expansion of thematic funds. Some funds address the environmental challenge with a broader range of climate policies and solutions, investing only in companies with the best environmental practices and targets. These funds do not necessarily focus on net-zero transition. However, they generally include renewable energy stocks. Some more specific thematic funds have been developed to address net-zero issues such as clean energy, hydrogen, water management and future mobility. Figure 11.35 shows the performance of four thematic equity indices: Bloomberg BioEnergy, Bloomberg Hydrogen, MSCI Future Mobility and MSCI New Energy. These indices have a high risk of tracking error relative to the MSCI World index. On average, the tracking error volatility of the satellite equity portfolio is expected to be around 20%. The fund manager can also develop a stock-picking process related to the net-zero theme. In this case, he can build a screening based on green revenue share, green capex or green opex measures. More generally, the idea behind building a satellite equity portfolio with a basket of stocks is to select stocks according to a green intensity measure that can be aligned with the EU taxonomy.

---

[40]On page 10.3.1 we saw that the net-zero funding requirement is equivalent to \$3.5 tn per year.

Figure 11.34: Performance and duration of the Bloomberg Global Green Bond and Aggregate indices
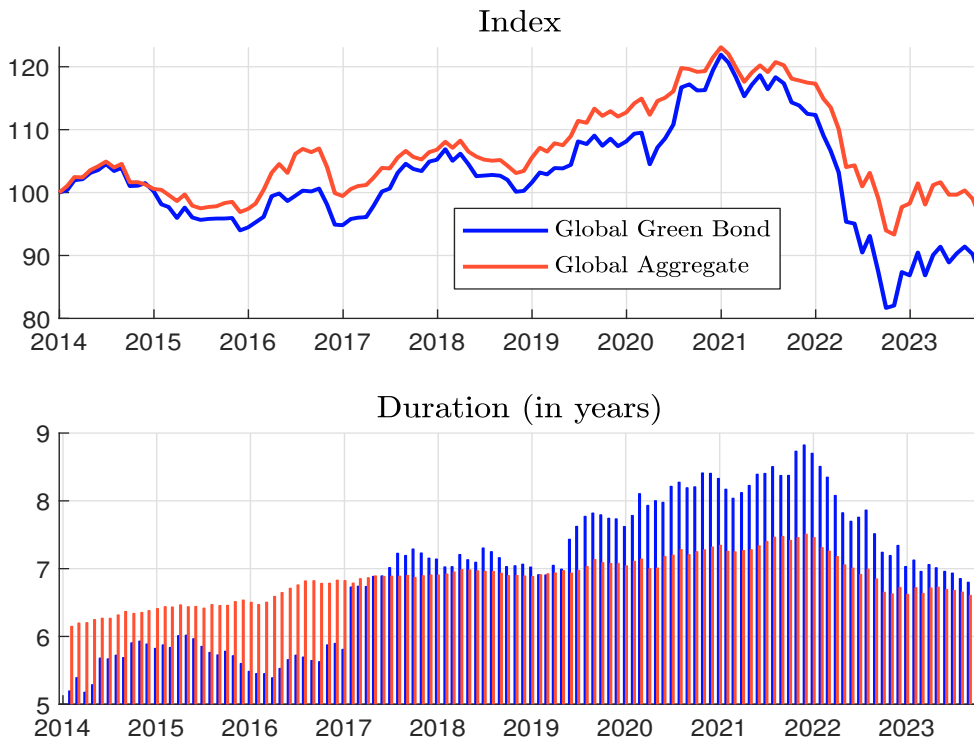
Source: Bloomberg (2023) & Author's calculations.

Figure 11.35: Performance and tracking error volatility of thematic equity indices
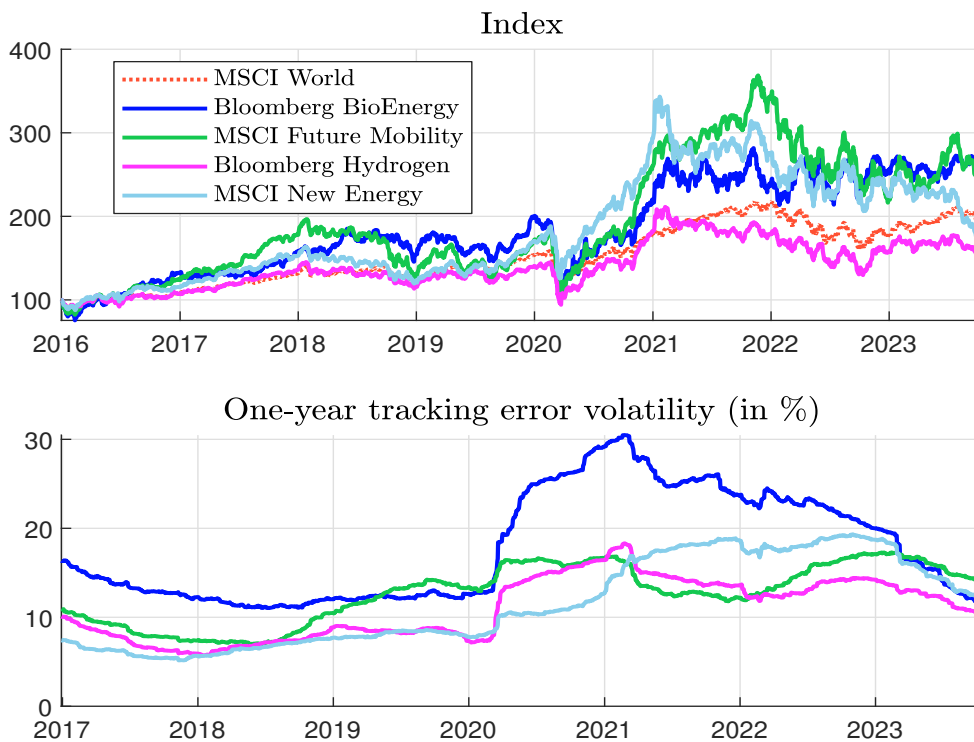
Source: Bloomberg (2023), MSCI (2023) & Author's calculations.

Green infrastructure The European Commission[41] defines green infrastructure as "*a strategically planned network of natural and semi-natural areas with other environmental features, designed and managed to deliver a wide range of ecosystem services, while also enhancing biodiversity*". Green infrastructure is implemented in a variety of sectors, from energy through energy transmission infrastructure, water through natural water retention measures or sustainable urban drainage systems, to the urban landscape with street trees to help sequester carbon or green roofs to help regulate the temperature of buildings. The cost of implementing green infrastructure is in the identification, mapping, planning and creation of the infrastructure, but the environmental, economic and social benefits make it worthwhile. Funds that assess infrastructure needs are emerging in the market and typically invest in owners of sustainable infrastructure assets as well as companies that are leaders in infrastructure investment. In addition to infrastructure funds, investors are also considering direct investments such as green car parks, water infrastructure and flood defences.

Sustainable real estate The real estate sector emits significant amounts of CO_2 through building operations, building materials, and construction. Action is needed in the construction of new buildings, but also in the renovation of existing buildings. For existing buildings, it is very important to reduce energy consumption, eliminate emissions from energy and refrigerants, and reduce or eliminate the use of fossil fuels. This is done by improving equipment such as insulation, ventilation, and the use of renewable energy, as well as optimizing operations by installing GHG monitors or adjusting temperature settings. New construction must be energy and carbon efficient, taking into account new and clean technologies. Sustainable real estate funds have entered the market, typically targeting multiple sectors and countries with a specific allocation to achieve net-zero by 2050. They mostly follow the CRREM (carbon risk real estate monitor) pathway, targeting 1.5°C/2°C using a Paris-aligned decarbonization pathway per country and building type, ranging from office buildings to retail stores and hotels.

Allocation process

In the following, we assume that the core-satellite portfolio is invested in stocks and bonds, but the inclusion of green infrastructure and sustainable real estate is straightforward.

The stock/bond mix allocation Let α_{equity} and α_{bond} be the proportions of stocks and bonds in the multi-asset portfolio. Let $\alpha^{\text{satellite}}$ be the weight of the satellite portfolio. The core allocation is given by the vector $\left(\alpha_{\text{equity}}^{\text{core}}, \alpha_{\text{bond}}^{\text{core}}\right)$, while the satellite allocation is defined by $\left(\alpha_{\text{equity}}^{\text{satellite}}, \alpha_{\text{bond}}^{\text{satellite}}\right)$. We have the following identities:

$$\begin{cases} \alpha_{\text{equity}} = \left(1 - \alpha^{\text{satellite}}\right) \alpha_{\text{equity}}^{\text{core}} + \alpha^{\text{satellite}} \alpha_{\text{equity}}^{\text{satellite}} \\ \alpha_{\text{bond}} = \left(1 - \alpha^{\text{satellite}}\right) \alpha_{\text{bond}}^{\text{core}} + \left(1 - \alpha^{\text{satellite}}\right) \alpha_{\text{bond}}^{\text{satellite}} \end{cases} \tag{11.26}$$

In general, the fund manager targets a strategic asset allocation at the portfolio level, *i.e.* the proportions α_{equity} and α_{bond} are given. For example, a defensive portfolio corresponds to a 20/80 constant mix strategy, while the 50/50 allocation is known as a balanced portfolio. Another famous allocation rule is the 60/40 portfolio, which is 60% in stocks and 40% in bonds. A first solution to Equation (11.26) is to maintain the same proportion of stocks and bonds in the core and satellite portfolios:

$$\begin{cases} \alpha_{\text{equity}}^{\text{core}} = \alpha_{\text{equity}}^{\text{satellite}} = \alpha_{\text{equity}} \\ \alpha_{\text{bond}}^{\text{core}} = \alpha_{\text{bond}}^{\text{satellite}} = \alpha_{\text{bond}} \end{cases}$$

---

[41]https://environment.ec.europa.eu/topics/nature-and-biodiversity/green-infrastructure_en.

However, this solution is not always satisfactory. In fact, the satellite portfolio generally has more bonds than the overall portfolio because the investment universe of green bonds is larger than the investment universe of green stocks. As a result, bonds are overweighted in the satellite portfolio. A second solution to Equation (11.26) is to calculate the proportion of bonds in the core portfolio relative to the proportion of bonds in the satellite portfolio:

$$\alpha_{\text{bond}}^{\text{core}} = \frac{\alpha_{\text{bond}} - \alpha^{\text{satellite}} \alpha_{\text{bond}}^{\text{satellite}}}{1 - \alpha^{\text{satellite}}}$$

Example 49 *We consider a 60/40 constant mix strategy. The satellite portfolio represents 10% of the net zero investments. We assume that the satellite portfolio has 70% exposure to green bonds.*

We have $\alpha_{\text{equity}} = 60\%$, $\alpha_{\text{bond}} = 40\%$, $\alpha^{\text{core}} = 90\%$, $\alpha^{\text{satellite}} = 10\%$ and $\alpha_{\text{bond}}^{\text{satellite}} = 70\%$. We deduce that:

$$\alpha_{\text{bond}}^{\text{core}} = \frac{0.40 - 0.10 \times 0.70}{1 - 0.10} = \frac{33}{90} = 36.67\%$$

The core allocation is then $(63.33\%, 36.67\%)$, while the satellite allocation is $(30\%, 70\%)$. We check that:

$$\begin{cases} \alpha_{\text{equity}} = 0.90 \times \left(1 - \dfrac{33}{90}\right) + 0.10 \times 0.30 = 60\% \\ \alpha_{\text{bond}} = 0.90 \times \dfrac{33}{90} + 0.10 \times 0.70 = 40\% \end{cases}$$

In Table 11.33, we report the values taken by $\alpha_{\text{bond}}^{\text{core}}$ when we target several constant mix strategies (60/40, 50/50, and 20/80). We assume that $\alpha_{\text{bond}}^{\text{satellite}}$ is set to 70% or 90% and consider several values of $\alpha^{\text{satellite}}$. For example, if green bonds account for 70% of the satellite portfolio's allocation and the satellite has a weight of 25%, the core portfolio's bond allocation must be set to 30% to achieve a 60/40 constant mix strategy.

Table 11.33: Calculating the bond allocation in the core portfolio ($\alpha_{\text{bond}}^{\text{core}}$ in %)

Strategy		60/40			50/50			20/80		
$\alpha_{\text{bond}}^{\text{satellite}}$		70.0	80.0	90.0	70.0	80.0	90.0	70.0	80.0	90.0
	0%	40.0	40.0	40.0	50.0	50.0	50.0	80.0	80.0	80.0
	1%	39.7	39.6	39.5	49.8	49.7	49.6	80.1	80.0	79.9
	5%	38.4	37.9	37.4	48.9	48.4	47.9	80.5	80.0	79.5
$\alpha^{\text{satellite}}$	10%	36.7	35.6	34.4	47.8	46.7	45.6	81.1	80.0	78.9
	15%	34.7	32.9	31.2	46.5	44.7	42.9	81.8	80.0	78.2
	20%	32.5	30.0	27.5	45.0	42.5	40.0	82.5	80.0	77.5
	25%	30.0	26.7	23.3	43.3	40.0	36.7	83.3	80.0	76.7

Tracking error risk of the core-satellite portfolio   Let w, w^{core}, $w^{\text{satellite}}$, and b be the core-satellite, core, satellite and benchmark portfolios, respectively. The return of the core-satellite portfolio is equal to:

$$\begin{aligned} R(w) = & \left(1 - \alpha^{\text{satellite}}\right) \underbrace{\left(\alpha_{\text{equity}}^{\text{core}} R\left(w_{\text{equity}}^{\text{core}}\right) + \alpha_{\text{bond}}^{\text{core}} R\left(w_{\text{bond}}^{\text{core}}\right)\right)}_{\text{Core portfolio's return}} + \\ & \alpha^{\text{satellite}} \underbrace{\left(\alpha_{\text{equity}}^{\text{satellite}} R\left(w_{\text{equity}}^{\text{satellite}}\right) + \alpha_{\text{bond}}^{\text{satellite}} R\left(w_{\text{bond}}^{\text{satellite}}\right)\right)}_{\text{Satellite portfolio's return}} \end{aligned}$$

By construction, we have $\alpha_{\text{equity}}^{\text{core}} + \alpha_{\text{bond}}^{\text{core}} = 1$ and $\alpha_{\text{equity}}^{\text{satellite}} + \alpha_{\text{bond}}^{\text{satellite}} = 1$. The proportion invested in equities is equal to $\alpha_{\text{equity}} = \left(1 - \alpha^{\text{satellite}}\right) \alpha_{\text{equity}}^{\text{core}} + \alpha^{\text{satellite}} \alpha_{\text{equity}}^{\text{satellite}}$ whereas the proportion invested in bonds is the complementary part $(\alpha_{\text{bond}} = 1 - \alpha_{\text{equity}})$. We deduce that:

$$R(w) = \tilde{\boldsymbol{\alpha}}^\top \tilde{\boldsymbol{R}}(w)$$

where $\tilde{\boldsymbol{R}}(w) = \left(R\left(w_{\text{equity}}^{\text{core}}\right), R\left(w_{\text{bond}}^{\text{core}}\right), R\left(w_{\text{equity}}^{\text{satellite}}\right), R\left(w_{\text{bond}}^{\text{satellite}}\right)\right)$ and:

$$\tilde{\boldsymbol{\alpha}} = \begin{pmatrix} \left(1 - \alpha^{\text{satellite}}\right) \alpha_{\text{equity}}^{\text{core}} \\ \left(1 - \alpha^{\text{satellite}}\right) \alpha_{\text{bond}}^{\text{core}} \\ \alpha^{\text{satellite}} \alpha_{\text{equity}}^{\text{satellite}} \\ \alpha^{\text{satellite}} \alpha_{\text{bond}}^{\text{satellite}} \end{pmatrix}$$

The benchmark portfolio's return is given by:

$$R(b) = \alpha_{\text{equity}} R(b_{\text{equity}}) + (1 - \alpha_{\text{equity}}) R(b_{\text{bond}})$$

Ben Slimane *et al.* (2023b) showed that:

$$\begin{aligned} R(b) &= \left(1 - \alpha^{\text{satellite}}\right) \left(\alpha_{\text{equity}}^{\text{core}} R(b_{\text{equity}}) + \alpha_{\text{bond}}^{\text{core}} R(b_{\text{bond}})\right) + \\ &\quad \alpha^{\text{satellite}} \left(\alpha_{\text{equity}}^{\text{satellite}} R(b_{\text{equity}}) + \alpha_{\text{bond}}^{\text{satellite}} R(b_{\text{bond}})\right) \end{aligned}$$

We deduce that:

$$R(b) = \tilde{\boldsymbol{\alpha}}^\top \tilde{\boldsymbol{R}}(b)$$

where $\tilde{\boldsymbol{R}}(b) = (R(b_{\text{Equity}}), R(b_{\text{bond}}), R(b_{\text{equity}}), R(b_{\text{bond}}))$. The tracking error is defined as:

$$e = R(w) - R(b) = \tilde{\boldsymbol{\alpha}}^\top \left(\tilde{\boldsymbol{R}}(w) - \tilde{\boldsymbol{R}}(b)\right)$$

Let $\tilde{\boldsymbol{\Sigma}}(w \mid b)$ be the 4×4 covariance matrix of $\tilde{\boldsymbol{R}}(w) - \tilde{\boldsymbol{R}}(b)$. We conclude that the tracking error volatility of the core-satellite portfolio has the following expression:

$$\sigma(w \mid b) = \sqrt{\tilde{\boldsymbol{\alpha}}^\top \tilde{\boldsymbol{\Sigma}}(w \mid b) \tilde{\boldsymbol{\alpha}}} = \sqrt{(\tilde{\boldsymbol{\alpha}} \circ \tilde{\boldsymbol{\sigma}}(w \mid b))^\top \tilde{\boldsymbol{\rho}}(w \mid b)(\tilde{\boldsymbol{\alpha}} \circ \tilde{\boldsymbol{\sigma}}(w \mid b))}$$

where $\tilde{\boldsymbol{\rho}}(w \mid b)$ is the correlation matrix of $R(w) - R(b)$ and $\tilde{\boldsymbol{\sigma}}(w \mid b)$ is the vector of tracking error volatilities:

$$\tilde{\boldsymbol{\sigma}}(w \mid b) = \begin{pmatrix} \sigma\left(w_{\text{equity}}^{\text{core}} \mid b_{\text{equity}}\right) \\ \sigma\left(w_{\text{bond}}^{\text{core}} \mid b_{\text{bond}}\right) \\ \sigma\left(w_{\text{equity}}^{\text{satellite}} \mid b_{\text{equity}}\right) \\ \sigma\left(w_{\text{bond}}^{\text{satellite}} \mid b_{\text{bond}}\right) \end{pmatrix}$$

Example 50 *The tracking error volatilities are 2% for the core equity portfolio, 25 bps for the core bond portfolio, 20% for the satellite equity portfolio, and 3% for the satellite bond portfolio. To define the correlation matrix $\tilde{\boldsymbol{\rho}}(w \mid b)$, we assume an 80% correlation between the two equity baskets, a 50% correlation between the two bond baskets, and a 0% correlation between the equity and bond baskets. We consider a 60/40 constant mix strategy. The satellite portfolio represents 10% of the net zero portfolio and has 70% exposure to green bonds.*

We compute the tracking error covariance matrix $\tilde{\boldsymbol{\Sigma}}(w \mid b)$ as follows: the tracking error variance for the core equity portfolio is $\tilde{\boldsymbol{\Sigma}}_{1,1}(w \mid b) = 0.02^2$, the tracking error variance for the satellite equity portfolio is $\tilde{\boldsymbol{\Sigma}}_{3,3}(w \mid b) = 0.20^2$, the tracking error covariance for the two core portfolios is $\tilde{\boldsymbol{\Sigma}}_{1,2}(w \mid b) = 0 \times 0.02 \times 0.0025$, the tracking error covariance for the core equity portfolio and the satellite equity portfolio is $\tilde{\boldsymbol{\Sigma}}_{1,3}(w \mid b) = 0.80 \times 0.02 \times 0.20$, and so on. Finally, we get:

$$\tilde{\boldsymbol{\Sigma}}(w \mid b) = \begin{pmatrix} 4 & 0 & 32 & 0 \\ 0 & 0.0625 & 0 & 0.375 \\ 32 & 0 & 400 & 0 \\ 0 & 0.375 & 0 & 9 \end{pmatrix} \times 10^{-4}$$

Since we have $\tilde{\boldsymbol{\alpha}} = (57\%, 33\%, 3\%, 7\%)$, we deduce that $\sigma(w \mid b) = 1.68\%$.

We consider several constant mix strategies with the same allocation within the core and satellite portfolios: $\alpha_{\text{equity}}^{\text{core}} = \alpha_{\text{equity}}^{\text{satellite}} = \alpha_{\text{equity}}$ and $\alpha_{\text{bond}}^{\text{core}} = \alpha_{\text{bond}}^{\text{satellite}} = \alpha_{\text{bond}}$. We assume that the tracking error volatilities are 2% for the core equity portfolio, 25 bps for the core bond portfolio, 20% for the satellite equity portfolio, and 3% for the satellite bond portfolio. We consider a lower and an upper bound on the correlation matrix $\tilde{\boldsymbol{\rho}}(w \mid b)$. For the lower bound, we set all cross-correlations to zero. For the upper bound, we assume an 80% correlation between the two equity baskets, an 80% correlation between the two bond baskets, and a 0% correlation between the equity and bond baskets. The results[42] are shown in Table 11.34 for three different values of $\alpha^{\text{satellite}}$. When the satellite weight is set to 20%, the tracking error volatility for the pure bond allocation is between 63 and 80 bps. For the 60/40 constant mix allocation, we obtain $\sigma(w \mid b) \in [2.60\%, 3.68\%]$, while for the pure equity allocation we have $\sigma(w \mid b) \in [4.31\%, 5.60\%]$. Ben Slimane *et al.* (2023b) estimated that the tracking error volatility of a net-zero core-satellite portfolio is currently around 3% for a 60/40 constant mix strategy.

Table 11.34: Estimation of the tracking error volatility of the core-satellite portfolio (in %)

	$\alpha^{\text{satellite}}$	Bond	Defensive	Balanced	60/40	Dynamic	Equity
	10%	0.38	0.62	1.36	1.62	2.15	2.69
Lower bound	20%	0.63	1.00	2.18	2.60	3.45	4.31
	30%	0.92	1.43	3.11	3.71	4.93	6.16
	10%	0.53	1.18	2.16	2.49	3.15	3.80
Upper bound	20%	0.80	1.76	3.20	3.68	4.64	5.60
	30%	1.07	2.34	4.24	4.87	6.13	7.40

Dynamic allocation   The allocation between the core and satellite portfolios depends on the investment opportunities for the satellite portfolio. Today, we see a large imbalance between the supply of green assets and what is needed to finance the transition. Therefore, we can expect the allocation in the satellite portfolio to be small, but to increase in the future:

$$\frac{\partial \alpha^{\text{satellite}}(t)}{\partial t} \geq 0$$

The allocation $\alpha^{\text{satellite}}$ is then time-varying and must be revised periodically, *e.g.* once a year.

---

[42]Defensive, balanced and dynamic portfolios correspond to 20/80, 50/50 and 80/20 constant mix strategies, respectively.

11.4 Exercise

11.4.1 Equity and bond portfolio optimization with green preferences

We consider an investment universe of 8 issuers. In the table below, we report the carbon emissions $\mathcal{CE}_{i,j}$ (in ktCO$_2$e) of these companies and their revenues Y_i (in \$ bn), and we indicate in the last row whether the company belongs to sector $\boldsymbol{Sector}_1$ or $\boldsymbol{Sector}_2$:

Issuer	#1	#2	#3	#4	#5	#6	#7	#8
$\mathcal{CE}_{i,1}$	75	5 000	720	50	2 500	25	30 000	5
$\mathcal{CE}_{i,2}$	75	5 000	1 030	350	4 500	5	2000	64
$\mathcal{CE}_{i,3}$	24 000	15 000	1 210	550	500	187	30 000	199
Y_i	300	328	125	100	200	102	107	25
$\boldsymbol{Sector}$	1	2	1	1	2	1	2	2

The benchmark b of this investment universe is defined as:

$$b = (22\%, 19\%, 17\%, 13\%, 11\%, 8\%, 6\%, 4\%)$$

In what follows, we consider long-only portfolios.

1. We want to compute the carbon intensity of the benchmark.

   (a) Compute the carbon intensities $\mathcal{CI}_{i,j}$ of each company i for the scopes 1, 2 and 3.

   (b) Deduce the carbon intensities $\mathcal{CI}_{i,j}$ of each company i for the scopes $1+2$ and $1+2+3$.

   (c) Deduce the weighted average carbon intensity (WACI) of the benchmark if we consider the scope $1+2+3$.

   (d) We assume that the market capitalization of the benchmark portfolio is equal to \$10 tn and we invest \$1 bn.

      i. Deduce the market capitalization of each company (expressed in \$ bn).

      ii. Compute the ownership ratio for each asset (expressed in bps).

      iii. Compute the carbon emissions of the benchmark portfolio[43] if we invest \$1 bn and we consider the scope $1+2+3$.

      iv. Compare the (exact) carbon intensity of the benchmark portfolio with the WACI value obtained in Question 1.(c).

2. We want to manage an equity portfolio with respect to the previous investment universe and reduce the weighted average carbon intensity of the benchmark by the rate $\mathcal{R}$. We assume that the volatility of the stocks is respectively equal to 22%, 20%, 25%, 18%, 40%, 23%, 13% and 29%. The correlation matrix between these stocks is given by:

$$\rho = \begin{pmatrix} 100\% \\ 80\% & 100\% \\ 70\% & 75\% & 100\% \\ 60\% & 65\% & 80\% & 100\% \\ 70\% & 50\% & 70\% & 85\% & 100\% \\ 50\% & 60\% & 70\% & 80\% & 60\% & 100\% \\ 70\% & 50\% & 70\% & 75\% & 80\% & 50\% & 100\% \\ 60\% & 65\% & 70\% & 75\% & 65\% & 70\% & 60\% & 100\% \end{pmatrix}$$

[43]We assume that the float percentage is equal to 100% for all the 8 companies.

(a) Compute the covariance matrix Σ.

(b) Write the optimization problem if the objective function is to minimize the tracking error risk under the constraint of carbon intensity reduction.

(c) Give the QP formulation of the optimization problem.

(d) $\mathcal{R}$ is equal to 20%. Find the optimal portfolio if we target scope $1+2$. What is the value of the tracking error volatility?

(e) Same question if $\mathcal{R}$ is equal to 30%, 50%, and 70%.

(f) We target scope $1+2+3$. Find the optimal portfolio if $\mathcal{R}$ is equal to $20\%, 30\%, 50\%$ and 70%. Give the value of the tracking error volatility for each optimized portfolio.

(g) Compare the optimal solutions obtained in Questions 2.(e) and 2.(f).

3. We want to manage a bond portfolio with respect to the previous investment universe and reduce the weighted average carbon intensity of the benchmark by the rate $\mathcal{R}$. We use the scope $1+2+3$. In the table below, we report the modified duration MD_i and the duration-times-spread factor DTS_i of each corporate bond i:

Asset	#1	#2	#3	#4	#5	#6	#7	#8
MD_i (in years)	3.56	7.48	6.54	10.23	2.40	2.30	9.12	7.96
DTS_i (in bps)	103	155	75	796	89	45	320	245
$\boldsymbol{S}ector$	1	2	1	1	2	1	2	2

We remind that the active risk can be calculated using three functions. For the active share, we have:

$$\mathcal{R}_{\mathrm{AS}}\left(w \mid b\right) = \sigma_{\mathrm{AS}}^2\left(w \mid b\right) = \sum_{i=1}^{n}\left(w_i - b_i\right)^2$$

We also consider the MD-based tracking error risk:

$$\mathcal{R}_{\mathrm{MD}}\left(w \mid b\right) = \sigma_{\mathrm{MD}}^2\left(w \mid b\right) = \sum_{j=1}^{n_{\boldsymbol{S}ector}}\left(\sum_{i \in \boldsymbol{S}ector_j}\left(w_i - b_i\right)\mathrm{MD}_i\right)^2$$

and the DTS-based tracking error risk:

$$\mathcal{R}_{\mathrm{DTS}}\left(w \mid b\right) = \sigma_{\mathrm{DTS}}^2\left(w \mid b\right) = \sum_{j=1}^{n_{\boldsymbol{S}ector}}\left(\sum_{i \in \boldsymbol{S}ector_j}\left(w_i - b_i\right)\mathrm{DTS}_i\right)^2$$

Finally, we define the synthetic risk measure as a combination of AS, MD and DTS active risks:

$$\mathcal{R}\left(w \mid b\right) = \varphi_{\mathrm{AS}}\mathcal{R}_{\mathrm{AS}}\left(w \mid b\right) + \varphi_{\mathrm{MD}}\mathcal{R}_{\mathrm{MD}}\left(w \mid b\right) + \varphi_{\mathrm{DTS}}\mathcal{R}_{\mathrm{DTS}}\left(w \mid b\right)$$

where $\varphi_{\mathrm{AS}} \geq 0$, $\varphi_{\mathrm{MD}} \geq 0$ and $\varphi_{\mathrm{DTS}} \geq 0$ indicate the weight of each risk. In what follows, we use the following numerical values: $\varphi_{\mathrm{AS}} = 100$, $\varphi_{\mathrm{MD}} = 25$ and $\varphi_{\mathrm{DTS}} = 1$. The reduction rate $\mathcal{R}$ of the weighted average carbon intensity is set to 50% for the scope $1+2+3$.

(a) Compute the modified duration $\mathrm{MD}\left(b\right)$ and the duration-times-spread factor $\mathrm{DTS}\left(b\right)$ of the benchmark.

(b) Let w_{ew} be the equally-weighted portfolio. Compute[44] $\mathrm{MD}\left(w_{\mathrm{ew}}\right)$, $\mathrm{DTS}\left(w_{\mathrm{ew}}\right)$, $\sigma_{\mathrm{AS}}\left(w_{\mathrm{ew}} \mid b\right)$, $\sigma_{\mathrm{MD}}\left(w_{\mathrm{ew}} \mid b\right)$ and $\sigma_{\mathrm{DTS}}\left(w_{\mathrm{ew}} \mid b\right)$.

---

[44]Precise the corresponding unit (years, bps or %) for each metric.

(c) We consider the following optimization problem:

$$w^\star = \arg\min \frac{1}{2}\mathcal{R}_{\mathrm{AS}}\left(w \mid b\right)$$

$$\text{s.t.} \quad \begin{cases} \sum_{i=1}^{n} w_i = 1 \\ \mathrm{MD}\left(w\right) = \mathrm{MD}\left(b\right) \\ \mathrm{DTS}\left(w\right) = \mathrm{DTS}\left(b\right) \\ \boldsymbol{CI}\left(w\right) \leq \left(1 - \boldsymbol{\mathcal{R}}\right)\boldsymbol{CI}\left(b\right) \\ 0 \leq w_i \leq 1 \end{cases}$$

Give the analytical value of the objective function. Find the optimal portfolio $w^\star$. Compute $\mathrm{MD}\left(w^\star\right)$, $\mathrm{DTS}\left(w^\star\right)$, $\sigma_{\mathrm{AS}}\left(w^\star \mid b\right)$, $\sigma_{\mathrm{MD}}\left(w^\star \mid b\right)$ and $\sigma_{\mathrm{DTS}}\left(w^\star \mid b\right)$.

(d) We consider the following optimization problem:

$$w^\star = \arg\min \frac{\varphi_{\mathrm{AS}}}{2}\mathcal{R}_{\mathrm{AS}}\left(w \mid b\right) + \frac{\varphi_{\mathrm{MD}}}{2}\mathcal{R}_{\mathrm{MD}}\left(w \mid b\right)$$

$$\text{s.t.} \quad \begin{cases} \sum_{i=1}^{n} w_i = 1 \\ \mathrm{DTS}\left(w\right) = \mathrm{DTS}\left(b\right) \\ \boldsymbol{CI}\left(w\right) \leq \left(1 - \boldsymbol{\mathcal{R}}\right)\boldsymbol{CI}\left(b\right) \\ 0 \leq w_i \leq 1 \end{cases}$$

Give the analytical value of the objective function. Find the optimal portfolio $w^\star$. Compute $\mathrm{MD}\left(w^\star\right)$, $\mathrm{DTS}\left(w^\star\right)$, $\sigma_{\mathrm{AS}}\left(w^\star \mid b\right)$, $\sigma_{\mathrm{MD}}\left(w^\star \mid b\right)$ and $\sigma_{\mathrm{DTS}}\left(w^\star \mid b\right)$.

(e) We consider the following optimization problem:

$$w^\star = \arg\min \frac{1}{2}\mathcal{R}\left(w \mid b\right)$$

$$\text{s.t.} \quad \begin{cases} \sum_{i=1}^{n} w_i = 1 \\ \boldsymbol{CI}\left(w\right) \leq \left(1 - \boldsymbol{\mathcal{R}}\right)\boldsymbol{CI}\left(b\right) \\ 0 \leq w_i \leq 1 \end{cases}$$

Give the analytical value of the objective function. Find the optimal portfolio $w^\star$. Compute $\mathrm{MD}\left(w^\star\right)$, $\mathrm{DTS}\left(w^\star\right)$, $\sigma_{\mathrm{AS}}\left(w^\star \mid b\right)$, $\sigma_{\mathrm{MD}}\left(w^\star \mid b\right)$ and $\sigma_{\mathrm{DTS}}\left(w^\star \mid b\right)$.

(f) Comment on the results obtained in Questions 3.(c), 3.(d) and 3.(e).

(g) How to find the previous solution of Question 3.(e) using a QP solver?

4. We consider a variant of Question 3 and assume that the synthetic risk measure is:

$$\mathcal{D}\left(w \mid b\right) = \varphi_{\mathrm{AS}}\mathcal{D}_{\mathrm{AS}}\left(w \mid b\right) + \varphi_{\mathrm{MD}}\mathcal{D}_{\mathrm{MD}}\left(w \mid b\right) + \varphi_{\mathrm{DTS}}\mathcal{D}_{\mathrm{DTS}}\left(w \mid b\right)$$

where:

$$\mathcal{D}_{\mathrm{AS}}\left(w \mid b\right) = \frac{1}{2}\sum_{i=1}^{n}\left|w_i - b_i\right|$$

$$\mathcal{D}_{\mathrm{MD}}\left(w \mid b\right) = \sum_{j=1}^{n_{\boldsymbol{Sector}}}\left|\sum_{i\in\boldsymbol{Sector}_j}\left(w_i - b_i\right)\mathrm{MD}_i\right|$$

$$\mathcal{D}_{\mathrm{DTS}}\left(w \mid b\right) = \sum_{j=1}^{n_{\boldsymbol{Sector}}}\left|\sum_{i\in\boldsymbol{Sector}_j}\left(w_i - b_i\right)\mathrm{DTS}_i\right|$$

(a) Define the corresponding optimization problem when the objective is to minimize the active risk and reduce the carbon intensity of the benchmark by $\mathcal{R}$.

(b) Give the LP formulation of the optimization problem.

(c) Find the optimal portfolio when $\mathcal{R}$ is set to 50%. Compare the solution with this obtained in Question 3.(e).

11.4.2   Monotonicity property of the order-statistic and naive approaches

Chapter 12

Physical Risk Modeling

12.1 General circulation model

12.1.1 Carbon cycle

12.1.2 Greenhouse gas chemistry

12.1.3 Horizontal and vertical heat transport

12.1.4 Oceans

12.1.5 Carbon sinks

12.2 Statistical modeling of climate hazards

12.2.1 Chronic risk

12.2.2 Acute risk

12.3 Geolocation

12.3.1 Climate hazard location

12.3.2 Asset location

12.4 Applications

12.4.1 Cyclones and hurricanes

12.4.2 Drought

12.4.3 Floods

12.4.4 Extreme heat

12.4.5 Water stress

12.4.6 Wildfire

12.5 Exercises

Chapter 13

Climate Stress Testing and Risk Management

13.1   Transmission channels

13.1.1   Direct and indirect transmission

13.1.2   Credit transmission channel

13.1.3   Market transmission channel

13.1.4   Systemic risk

13.2 Climate risk hedging

13.3 Climate value-at-risk

13.4 Climate stress testing

13.5 Exercises

Conclusion

Appendices

Appendix A

Technical Appendix

A.1 Mathematical tools

A.1.1 Linear algebra

Eigendecomposition

The value λ is an eigenvalue of the $n \times n$ matrix A if there exists a non-zero eigenvector v such that we have $Av = \lambda v$. We denote V the matrix composed of the n eigenvectors. We have $AV = V\Lambda$ where $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ is the diagonal matrix of eigenvalues. We finally obtain the eigendecomposition of the matrix A:

$$A = V\Lambda V^{-1}$$

If A is an hermitian matrix[1], then the matrix V of eigenvectors is unitary. It follows that:

$$A = V\Lambda V^*$$

In particular, if A is a symmetric real matrix, we obtain the following relationship[2]:

$$A = V\Lambda V^\top$$

Schur decomposition and matrix function

The Schur decomposition of the $n \times n$ matrix A is equal to:

$$A = QTQ^* \tag{A.1}$$

where Q is a unitary matrix and T is an upper triangular matrix[3]. This decomposition is useful to calculate matrix functions.

Let us consider the matrix function in the space $\mathbb{M}$ of square matrices:

$$f : \quad \mathbb{M} \longrightarrow \mathbb{M}$$
$$A \longmapsto B = f(A)$$

For instance, if $f(x) = \sqrt{x}$ and A is positive, we can define the matrix B such that:

$$BB^* = B^*B = A$$

---

[1]The square matrix A is hermitian if it is equal to its own conjugate transpose A^*, implying that we have $A_{i,j} = \operatorname{conj} A_{j,i}$;

[2]We have $A^\top = \left(V\Lambda V^{-1}\right)^\top = \left(V^{-1}\right)^\top \Lambda V^\top$. We deduce that $V^{-1} = V^\top$.

[3]Q and T are also called the transformation matrix and the Schur form of A.

B is called the square root of A and we note $B = A^{1/2}$. This matrix function generalizes the scalar-valued function to the set of matrices. Let us consider the following Taylor expansion:

$$f(x) = f(x_0) + (x - x_0) f'(x_0) + \frac{(x - x_0)^2}{2!} f''(x_0) + \dots$$

We can show that if the series converge for $|x - x_0| < \alpha$, then the matrix $f(A)$ defined by the following expression:

$$f(A) = f(x_0) + (A - x_0 I) f'(x_0) + \frac{(A - x_0 I)^2}{2!} f''(x_0) + \dots$$

converges to the matrix B if $|A - x_0 I| < \alpha$ and we note $B = f(A)$. In the case of the exponential function, we have:

$$f(x) = e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

We deduce that the exponential of the matrix A is equal to:

$$B = e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!}$$

In a similar way, the logarithm of A is the matrix B such that $e^B = A$ and we note $B = \ln A$.

Let A and B be two $n \times n$ square matrices. Using the Taylor expansion, Golub and Van Loan (2013) showed that $f(A^\top) = f(A)^\top$, $A f(A) = f(A) A$ and $f(B^{-1} A B) = B^{-1} f(A) B$. It follows that:

$$e^{A^\top} = (e^A)^\top$$

and:

$$e^{B^{-1} A B} = B^{-1} e^A B$$

If $AB = BA$, we can also prove that $A e^B = e^B A$ and $e^{A+B} = e^A e^B = e^B e^A$.

Remark 103 *There are different ways to compute numerically $f(A)$. For transcendental functions, we have:*

$$f(A) = Q f(T) Q^*$$

where $A = Q T Q^*$ is the Schur decomposition of A. Because T is an upper diagonal matrix, $f(T)$ is also a diagonal matrix whose elements can be calculated with Algorithm 9.1.1 of Golub and Van Loan (2013). This algorithm is reproduced below[4].

Sherman-Morrison-Woodbury formula

Let u and v be two $n \times 1$ vectors and A be an invertible $n \times n$ matrix. We can show that (Golub and Van Loan, 2013):

$$\left(A + uv^\top\right)^{-1} = A^{-1} - \frac{1}{1 + v^\top A^{-1} u} A^{-1} uv^\top A^{-1}$$

---

[4]For the exponential matrix, we may prefer to use the Pade approximation method, which is described in Algorithm 9.3.1 (scaling and squaring) of Golub and Van Loan (2013). See also the survey of Moler and Van Loan (2003).

Algorithm 2 Schur-Parlett matrix function $f(A)$

Compute the Schur decomposition $A = QTQ^*$
Initialize F to the matrix $\mathbf{0}_{n \times n}$
for $i = 1 : n$ **do**
    $f_{i,i} \leftarrow f(t_{i,i})$
end for
for $p = 1 : n - 1$ **do**
  for $i = 1 : n - p$ **do**
    $j \leftarrow i + p$
    $s \leftarrow t_{i,j}(f_{j,j} - f_{i,i})$
    for $k = i + 1 : j - 1$ **do**
      $s \leftarrow s + t_{i,k}f_{k,j} - f_{i,k}t_{k,j}$
    end for
    $f_{i,j} \leftarrow s/(t_{j,j} - t_{i,i})$
  end for
end for
$B \leftarrow QFQ^*$
return B

Source: Golub and Van Loan (2013), page 519.

Batista and Karawia (2009) extended the SMW formula when the outer product is a sum:

$$\left(A + \sum_{k=1}^{m} u_k v_k^\top\right)^{-1} = A^{-1} - A^{-1}US^{-1}V^\top A^{-1}$$

where $U = \begin{pmatrix} u_1 & \cdots & u_m \end{pmatrix}$ and $V = \begin{pmatrix} v_1 & \cdots & v_m \end{pmatrix}$ are two $n \times m$ matrices, and $S = I_m + T$ and $T = (T_{i,j})$ are two $m \times m$ matrices where $T_{i,j} = v_i^\top A^{-1} u_j$. In the case $m = 2$, Roncalli *et al.* (2020) showed that the SMW formula becomes:

$$\left(A + u_1 v_1^\top + u_2 v_2^\top\right)^{-1} = A^{-1} - A^{-1}CA^{-1}$$

where:

$$C = \frac{\left(I_n - u_2 v_2^\top A^{-1}\right) u_1 v_1^\top + \left(I_n - u_1 v_1^\top A^{-1}\right) u_2 v_2^\top + u_1 v_2^\top A^{-1} u_2 v_1^\top + u_2 v_1^\top A^{-1} u_1 v_2^\top}{1 + \left(v_1^\top A^{-1} u_1\right)\left(v_2^\top A^{-1} u_2\right) + v_1^\top A^{-1} u_1 + v_2^\top A^{-1} u_2 - \left(v_2^\top A^{-1} u_1\right)\left(v_1^\top A^{-1} u_2\right)}$$

A.1.2   Optimization

Bisection algorithm

The simplest algorithm to find the roots of the equation $f(x) = 0$ is the bisection algorithm. We assume that the function f is continuous on the interval $[a, b]$, and $f(a)$ and $f(b)$ have oppositive signs: $f(a)f(b) < 0$. The underlying idea is to reduce the bracket by computing the midpoint $c = (a + b)/2$. If $f(a)f(c) < 0$, then the interval becomes $[a, c]$, otherwise it is equal to $[c, b]$. This step is repeated until the interval is sufficiently small. The absolute error is divided by 2 at each step so the method converges linearly. Let $c^\star$ be a root such that $f(c^\star) = 0$ and $c^\star \in [a, b]$. Let $c_{(k)}$ be the midpoint at the kth iteration. Then the difference between $c_{(k)}$ and $c^\star$ is bounded by $\left|c_{(k)} - c^\star\right| \leq 2^{-k}(b - a)$. For instance, after 10 iterations, we are sure that the interval has been divided by $2^{10} = 1024$. After 20 iterations, the interval is divided by 10^6.

Linear programming

The canonical form of a linear program is:

$$x^\star \quad = \quad \arg\max c^\top x$$
$$\text{s.t.} \quad \begin{cases} Sx \le T \\ x \ge \mathbf{0}_n \end{cases}$$

where x is a $n \times 1$ vector, c is a $n \times 1$ vector, S is a $n_S \times n$ matrix, and T is a $n_S \times 1$ vector. A minimization linear programming (LP) problem can be cast into a maximization LP problem because $\min c^\top x \equiv \max -c^\top x$. A maximization LP problem is in standard form if it is written as:

$$x^\star \quad = \quad \arg\max c^\top x$$
$$\text{s.t.} \quad \begin{cases} Sx = T \\ x \ge \mathbf{0}_n \end{cases}$$

We notice that a canonical problem can always be put into a standard form by introducing the slack variables ξ:

$$Sx \le T \Leftrightarrow \{\xi \ge \mathbf{0}_{n_S} : Sx + I_{n_S}\xi = T\}$$

We obtain the following LP problem:

$$y^\star \quad = \quad \arg\max d^\top y$$
$$\text{s.t.} \quad \begin{cases} My = N \\ y \ge \mathbf{0}_n \end{cases}$$

where $d = (c, \mathbf{0}_{n_S})$, $y = (x, s)$, $M = \begin{pmatrix} S & I_{n_S} \end{pmatrix}$ and $N = T$. The technique of slack variables[5] can also be used when the contraint is $Sx \ge T$ and we have $Sx - \xi = T$. If we introduce lower and upper bounds ($x^- \le x \le x^+$), we can use the change of variable $y = x - x^-$. It follows that $y \ge \mathbf{0}_n$ and $I_n y \le x^+ - x^-$. Since we have $x = x^- + y$, we obtain $c^\top x = c^\top x^- + c^\top y$ and $Sx \le T \Leftrightarrow Sy \le T - Sx^-$. Finally, we deduce that:

$$\begin{cases} x^\star & = & \arg\max c^\top x \\ & \text{s.t.} & \begin{cases} Sx \le T \\ x \ge \mathbf{0}_n \end{cases} \end{cases} \quad \Longleftrightarrow \quad \begin{cases} y^\star & = & \arg\max c^\top y \\ & \text{s.t.} & \begin{cases} \begin{pmatrix} S \\ I_n \end{pmatrix} y \le \begin{pmatrix} T - Sx^- \\ x^+ - x^- \end{pmatrix} \\ y \ge \mathbf{0}_n \end{cases} \end{cases}$$

If some variables $x_i \in \mathbb{R}$, we can write $x_i = y_i - z_i$ where $y_i \ge 0$ and $z_i \ge 0$.

The previous analysis shows that we can always transform a general LP problem into a canonical or standard form. This is why most numerical packages consider the following general formulation:

$$x^\star \quad = \quad \arg\min c^\top x \qquad\qquad (\text{A.2})$$
$$\text{s.t.} \quad \begin{cases} Ax = B \\ Cx \le D \\ x^- \le x \le x^+ \end{cases}$$

where x is a $n \times 1$ vector, c is a $n \times 1$ vector, A is a $n_A \times n$ matrix, B is a $n_A \times 1$ vector, C is a $n_C \times n$ matrix, D is a $n_C \times 1$ vector, and x^- and x^+ are two $n \times 1$ vectors. If $n_A = 0$, there is no equality constraints. Similarly, there is no inequality constraints if $n_C = 0$. If there is no lower

---

[5]When we rewrite an inequality by substracting new variables, these last ones are also called surplus variables.

bounds or/and upper bounds, this implies that $x^- = -\infty \cdot \mathbf{1}_n$ and $x^+ = \infty \cdot \mathbf{1}_n$. A solution x is said to be feasible if it satisfies all the constraints of the linear program (A.2). The set of feasible solutions is called the feasible space Ω:

$$\Omega = \left\{ x \in \mathbb{R}^n : Ax = B, Cx \le D, x^- \le x \le x^+ \right\}$$

Since $Ax = B$ and $Cx \le D$ define two convex sets (a hyperplane and a closed half-space), Ω is a convex set. If Ω is not empty and $c^\top x$ is bounded below, then there is an optimal solution.

Example 51 *We consider the following optimization problem:*

$$(x_1^\star, x_2^\star, x_3^\star) \quad = \quad \arg\min x_1 + x_2 + x_3$$

$$s.t. \quad \begin{cases} x_1 + x_2 + x_3 = 1 \\ 3x_1 + x_2 + 5x_3 \ge 2 \\ 2x_1 + x_2 + 3x_3 \le s \\ 0 \le x_i \le 1 \end{cases}$$

If $s \ge 1.5$, the optimal solution is $(0, 0.75, 0.25)$. If $s < 1.5$, $\Omega = \emptyset$ and there is no solution.

Let us consider the standard program: $x^\star = \arg\max f(x) = c^\top x$ subject to $Ax = b$. The Lagrange function is:

$$\mathcal{L}(x; \lambda) = c^\top x - \lambda^\top (Ax - b)$$

The first-order condition is:

$$\begin{cases} \partial_x \mathcal{L}(x; \lambda) = c - A^\top \lambda = \mathbf{0}_n \\ \partial_\lambda \mathcal{L}(x; \lambda) = Ax - b = \mathbf{0}_{n_A} \end{cases}$$

The optimal solution $(x^\star, \lambda^\star)$ satisfies the following system of equations: $Ax^\star = b$ and $A^\top \lambda^\star = c$. We also have $f(x^\star) = \mathcal{L}(x^\star, \lambda^\star) = c^\top x^\star - \lambda^{\star\top}(Ax^\star - b)$ and:

$$\mathcal{L}(x^\star, \lambda^\star) = \sum_{i=1}^n c_i x_i^\star - \sum_{j=1}^{n_A} \lambda_j^\star \sum_{i=1}^n A_{j,i} x_i^\star + \sum_{j=1}^{n_A} \lambda_j^\star b_j$$

Let us assume that we increase b by a vector $\varepsilon \ge \mathbf{0}_{n_A}$. The Lagrange function becomes:

$$\begin{aligned} \mathcal{L}_\varepsilon(x^\star, \lambda^\star) &= \sum_{i=1}^n c_i x_i^\star - \sum_{j=1}^{n_A} \lambda_j^\star \sum_{i=1}^n A_{j,i} x_i^\star + \sum_{j=1}^{n_A} \lambda_j^\star (b_j + \varepsilon_j) \\ &= f(x^\star) + \sum_{j=1}^{n_A} \lambda_j^\star \varepsilon_j \end{aligned}$$

If we relax the jth constraint by a small value of ε_j, we change the objective function by $\lambda_j^\star \varepsilon_j$. If $\lambda_j^\star = 0$, then modifying locally the constraint has no impact. If $\lambda_j^\star < 0$, we have $f_\varepsilon(x^\star) < f(x^\star)$, implying that we have no incentive to change the value b_j of the constraint, otherwise we will reduce the objective function. If $\lambda_j^\star > 0$, we have $f_\varepsilon(x^\star) > f(x^\star)$, meaning that we can improve the objective function. From an economic point of view, a linear program is usually a profit maximization, a production maximization or a cost minimization. In this case, we can ask ourselves what price p_j we are willing to pay to increase the right-hand side value b_j by one unit? By taking $\varepsilon_j = 1$, we obtain $p_j \le \lambda_j^\star$ and $\lambda_j^\star$ is called the shadow or dual price. This explains why duality plays an important role in economics.

Example 52 *We consider the following optimization problem:*

$$
\begin{aligned}
x^{\star} \quad &= \quad \arg\max 2x_2 + 3x_3 - x_1 \\
\text{s.t.} \quad &\left\{
\begin{array}{l}
x_1 + x_2 + x_3 = 1 \\
2x_1 + 5x_2 + 6x_3 = 0.2
\end{array}
\right.
\end{aligned}
$$

The optimal solution is $x^{\star} = (1.45, 0, -0.45)$ and $\lambda^{\star} = (3, -1)$, and the objective function is equal to $f(x^{\star}) = 2.80$. The shadow prices are then equal to $+3$ for the first constraint and -1 for the second constraint. We do not want to increase the second constraint by ε_2 because it reduces the objective function by ε_2. Indeed, if $\varepsilon_2 = 0.04$, the new solution is $x^{*} = (1.44, 0, -0.44)$ and we have $f(x^{*}) = 2.76$. On the contrary, if we increase the first constraint by ε_1, we improve the objective function by $3\varepsilon_1$. For instance, if $\varepsilon_1 = 0.1$, the new solution is $x^{*} = (1.6, 0, -0.5)$ and we have $f(x^{*}) = 3.1$. Let us combine the two effects. If $(\varepsilon_1, \varepsilon_2) = (0.1, 0.04)$, we have $x^{*} = (1.59, 0, -0.49)$ and $f(x^{*}) = 3.06$. We verify that:

$$
\begin{aligned}
f(x^{\star}) + \lambda_1^{\star}\varepsilon_1 + \lambda_2^{\star}\varepsilon_2 \quad &= \quad 2.80 + 3 \times 0.1 - 1 \times 0.04 \\
&= \quad 3.06 \\
&= \quad f(x^{*})
\end{aligned}
$$

If the primal problem is maximizing $c^{\top}x$ subject to $Ax \leq b$ and $x \geq \mathbf{0}_n$, then the dual problem is defined as:

$$
\begin{aligned}
x^{\star} \quad &= \quad \arg\min b^{\top}y \\
\text{s.t.} \quad &\left\{
\begin{array}{l}
A^{\top}y \geq c \\
y \geq \mathbf{0}_m
\end{array}
\right.
\end{aligned} \tag{A.3}
$$

We notice that:

- The dual problem has a linear programming form;

- The maximization problem and the $\leq$ inequality type are changed into a minimization problem and a $\geq$ inequality type;

- Each dual variable y_j may be assigned to a corresponding primal constraint $\sum_{i=1}^{n} A_{j,i}x_i \leq b_j$;

- The roles of the vectors c and b are switched: c becomes the right-hand side of the constraints and b the coefficients of the objective function;

- The matrix A of the inequality constraints is transposed.

Let $\Omega_x = \{x : Ax \leq b, x \geq \mathbf{0}_n\}$ and $\Omega_y = \{y : A^{\top}y \geq c, y \geq \mathbf{0}_m\}$ be the primal and dual feasible spaces. The weak duality theorem states that[6]:

$$
\forall x \in \Omega_x, y \in \Omega_y : c^{\top}x \leq b^{\top}y
$$

This means that the objective value of the dual problem is an upper bound on the objective value of the primal problem, and vice versa. Using this property, we can also formulate the strong duality theorem: A primal feasible solution $x^{\star}$ is optimal if and only if there exists a dual feasible solution $y^{\star}$ such that $c^{\top}x^{\star} = b^{\top}y^{\star}$. We also deduce that $y^{\star}$ is the optimal solution of the dual LP problem.

---

[6]If $x \geq \mathbf{0}_n$ and $\alpha \geq \beta$, we have $\alpha^{\top}x \geq \beta^{\top}x$ because $\alpha_i x_i \geq \beta_i x_i$ and $\sum_{i=1}^{n} \alpha_i x_i \geq \sum_{i=1}^{n} \beta_i x_i$. Using the constraints $Ax \leq b$ and $A^{\top}y \geq c$, we deduce that $c^{\top}x = x^{\top}c \leq x^{\top}A^{\top}y = (Ax)^{\top}y \leq b^{\top}y$.

Example 53 *We consider a classical ressource allocation problem. We would like to produce 3 goods, whose market prices are \$3, \$5 and \$7. The production of these goods requires some ressources, such as capital, labor, raw materials and land. The available ressources are respectively equal to 100, 50, 80 and 200. For the first good, the production of one unit needs 5 units of capital, 3 units of labor, 1 unit of raw materials and 1 unit of land. For the second good, the intermediary consumptions are 10, 2, 10 and 2, while they are equal to 4, 7, 8 and 4 for the third good.*

The problem is to maximize the revenue under allocation constraints. By definition, the revenue is the product of sold quantities and market prices:

$$\mathcal{R}evenue = 3x_2 + 5x_3 + 7x_1$$

If we consider the first ressource constraint, we have 100 units of capital that we can allocate between the three goods in the following way. We need 5 units of capital to produce one unit of the first good, 10 units of capital to produce one unit of the second good and 4 units of capital to produce one unit of the third good. The optimization problem is then defined as follows:

$$
\begin{aligned}
x^\star \quad = \quad & \arg\max 3x_2 + 5x_3 + 7x_1 \\
\text{s.t.} \quad & \begin{cases}
5x_1 + 10x_2 + 4x_3 \leq 100 \\
3x_1 + 2x_2 + 7x_3 \leq 50 \\
x_1 + 10x_2 + 8x_3 \leq 80 \\
x_1 + 2x_2 + 4x_3 \leq 200 \\
x_1 \geq 0, x_2 \geq 0, x_3 \geq 0
\end{cases}
\end{aligned}
$$

The optimal solution is $x_1^\star = 7.4390$, $x_2^\star = 5.3049$ and $x_3^\star = 2.4390$. The optimal value of the revenue is \$65.91. The computation of the Lagrange multipliers gives $\lambda_1^\star = 0.1555$, $\lambda_2^\star = 0.6707$, $\lambda_3^\star = 0.2104$ and $\lambda_4^\star = 0$. While the primal LP was a resource allocation problem, the dual LP is a resource valuation problem from an economic viewpoint. Indeed, y_i is the implicit cost per unit associated with each resource i and the objective is then to minimize the total cost of the resources:

$$\mathcal{C}ost = 100y_1 + 50y_2 + 80y_3 + 200y_4$$

Moreover, if we consider the first good, its market price p_1 is \$3, and its internal value to produce it is equal to $v_1 = 5y_1 + 3y_2 + y_3 + y_4$. Therefore, we have the following constraint $v_1 \geq p_1$. Finally, we deduce that the resource valuation problem is to minimize the total cost under the constraints $v_1 \geq p_1$, $v_2 \geq p_2$ and $v_3 \geq p_3$. Therefore, the dual problem is:

$$
\begin{aligned}
y^\star \quad = \quad & \arg\min 100y_1 + 50y_2 + 80y_3 + 200y_4 \\
\text{s.t.} \quad & \begin{cases}
5y_1 + 3y_2 + y_3 + y_4 \geq 3 \\
10y_1 + 2y_2 + 10y_3 + 2y_4 \geq 5 \\
4y_1 + 7y_2 + 8y_3 + 4y_4 \geq 7 \\
y_1 \geq 0, y_2 \geq 0, y_3 \geq 0, y_3 \geq 0
\end{cases}
\end{aligned}
$$

The optimal solution is $y_1^\star = 0.1555$, $y_2^\star = 0.6707$, $y_3^\star = 0.2104$ and $y_4^\star = 0$. The optimal value of the dual objective function is \$65.91. The computation of the Lagrange multipliers gives $\vartheta_1^\star = 7.4390$, $\vartheta_2^\star = 5.3049$ and $\vartheta_3^\star = 2.4390$. We can make the following observations. First, the strong duality property is satisfied because $\mathcal{R}evenue(x^\star) = \mathcal{C}ost(y^\star) = \$65.91$. Second, we verify that $y^\star = \lambda^\star$ and $x^\star = \vartheta^\star$. The optimal solution of one problem is equal to the Lagrange coefficients of the second problem. Therefore, we interpret $y^\star$ as the shadow prices of the resources. Third, a shadow price $y_j^\star$ is equal to zero when the jth constraint is non-binding. Otherwise, it is strictly positive. If we

consider the previous example, the shadow price of the fourth resource is equal to zero because this resource is not scarce. In particular, we do not need to buy supplementary land to increase the revenue. Therefore, we are ready to pay zero for additional land. On the contrary, we can pay the supplementary unit up to $0.1555 to buy more capital. If the price of capital is greater than $0.1555, the increase of revenue will be lower than the increase of cost and we have no interest to increase the production capacity.

Quadratic programming

A quadratic programming (QP) problem is an optimization problem with a quadratic objective function and linear inequality constraints:

$$
\begin{aligned}
x^\star \quad = \quad & \arg\min \frac{1}{2} x^\top Q x - x^\top R \\
\text{s.t.} \quad & Sx \leq T
\end{aligned}
\tag{A.4}
$$

where x is a $n \times 1$ vector, Q is a $n \times n$ matrix and R is a $n \times 1$ vector. We note that the system of constraints $Sx \leq T$ allows specifying linear equality constraints[7] $Ax = B$ or weight constraints $x^- \leq x \leq x^+$. Most numerical packages then consider the following formulation:

$$
\begin{aligned}
x^\star \quad = \quad & \arg\min \frac{1}{2} x^\top Q x - x^\top R \\
\text{s.t.} \quad & \begin{cases} Ax = B \\ Cx \leq D \\ x^- \leq x \leq x^+ \end{cases}
\end{aligned}
\tag{A.5}
$$

because the problem (A.5) is equivalent to the canonical problem (A.4) with the following system of linear inequalities:

$$
\begin{bmatrix} -A \\ A \\ C \\ -I_n \\ I_n \end{bmatrix} x \leq \begin{bmatrix} -B \\ B \\ D \\ -x^- \\ x^+ \end{bmatrix}
$$

If the space Ω defined by $Sx \leq T$ is non-empty and if Q is a symmetric positive definite matrix, the solution exists because the function $f(x) = \frac{1}{2} x^\top Q x - x^\top R$ is convex. In the general case where Q is a square matrix, the solution may not exist.

The Lagrange function is also:

$$
\mathcal{L}(x; \lambda) = \frac{1}{2} x^\top Q x - x^\top R + \lambda^\top (Sx - T)
$$

We deduce that the dual problem is defined by:

$$
\begin{aligned}
\lambda^\star \quad = \quad & \arg\max \left\{ \inf_x \mathcal{L}(x; \lambda) \right\} \\
\text{s.t.} \quad & \lambda \geq 0
\end{aligned}
$$

---

[7]This is equivalent to imposing that $Ax \geq B$ and $Ax \leq B$.

We note that $\partial_x \mathcal{L}(x; \lambda) = Qx - R + S^\top \lambda$. The solution to the problem $\partial_x \mathcal{L}(x; \lambda) = 0$ is then $x = Q^{-1}\left(R - S^\top \lambda\right)$. We obtain:

$$
\begin{aligned}
\inf_x \mathcal{L}(x; \lambda) &= \frac{1}{2}\left(R^\top - \lambda^\top S\right) Q^{-1}\left(R - S^\top \lambda\right) - \left(R^\top - \lambda^\top S\right) Q^{-1} R + \\
&\quad \lambda^\top \left(SQ^{-1}\left(R - S^\top \lambda\right) - T\right) \\
&= \frac{1}{2}R^\top Q^{-1} R - \lambda^\top SQ^{-1} R + \frac{1}{2}\lambda^\top SQ^{-1} S^\top \lambda - R^\top Q^{-1} R + \\
&\quad 2\lambda^\top SQ^{-1} R - \lambda^\top SQ^{-1} S^\top \lambda - \lambda^\top T \\
&= -\frac{1}{2}\lambda^\top SQ^{-1} S^\top \lambda + \lambda^\top \left(SQ^{-1} R - T\right) - \frac{1}{2}R^\top Q^{-1} R
\end{aligned}
$$

The dual program is another quadratic program:

$$
\begin{aligned}
\lambda^\star &= \arg\min \frac{1}{2}\lambda^\top \bar{Q}\lambda - \lambda^\top \bar{R} \\
&\text{s.t.} \quad \lambda \geq 0
\end{aligned}
\tag{A.6}
$$

where $\bar{Q} = SQ^{-1}S^\top$ and $\bar{R} = SQ^{-1}R - T$.

Alternating direction method of multipliers

The alternating direction method of multipliers (ADMM) is an algorithm introduced by Gabay and Mercier (1976) to solve optimization problems that can be expressed as:

$$
\begin{aligned}
\{x^\star, y^\star\} &= \arg\min_{(x,y)} f_x(x) + f_y(y) \\
&\text{s.t.} \quad Ax + By = c
\end{aligned}
\tag{A.7}
$$

where $A \in \mathbb{R}^{p \times n}$, $B \in \mathbb{R}^{p \times m}$, $c \in \mathbb{R}^p$, and the functions $f_x : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ and $f_y : \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$ are proper closed convex functions. Boyd *et al.* (2011) showed that the ADMM algorithm consists of the following three steps:

1. The x-update is:

$$
x^{(k+1)} = \arg\min_x \left\{ f_x^{(k+1)}(x) := f_x(x) + \frac{\varphi}{2}\left\| Ax + By^{(k)} - c + u^{(k)} \right\|_2^2 \right\}
\tag{A.8}
$$

2. The y-update is:

$$
y^{(k+1)} = \arg\min_y \left\{ f_y^{(k+1)}(y) := f_y(y) + \frac{\varphi}{2}\left\| Ax^{(k+1)} + By - c + u^{(k)} \right\|_2^2 \right\}
\tag{A.9}
$$

3. The u-update is:

$$
u^{(k+1)} = u^{(k)} + \left(Ax^{(k+1)} + By^{(k+1)} - c \right)
\tag{A.10}
$$

In this approach, $u^{(k)}$ is the dual variable of the primal residual $r = Ax + By - c$ and φ is the $\mathcal{L}_2$-norm penalty variable. The parameter φ may be constant or it may change at each iteration. The ADMM algorithm benefits from the dual ascent principle and the multipliers method. The difference with the latter is that the x- and y-updates are performed alternately. This makes it more flexible, since the updates are equivalent to computing proximal operators for f_x and f_y independently. In practice, ADMM may be slow to converge at high accuracy, but is fast to converge at modest accuracy. Therefore, ADMM is a good candidate for solving large-scale machine learning problems where high accuracy does not necessarily lead to a better solution.

A.1.3   Quadratic form

Definition

In mathematics, a quadratic form is a polynomial with terms all of degree two:

$$\mathcal{QF}(x_1, \ldots, x_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{i,j} x_i x_j = x^\top A x$$

where $A = (a_{i,j})$ is a $n \times n$ matrix. Since $x^\top A x$ is a scalar, we have $\left(x^\top A x\right)^\top = x^\top A^\top x$ and:

$$\begin{aligned}
\mathcal{QF}(x_1, \ldots, x_n) &= \frac{1}{2}\left(x^\top A x + x^\top A^\top x\right) \\
&= \frac{1}{2} x^\top \left(A + A^\top\right) x \\
&= \frac{1}{2} x^\top Q x
\end{aligned}$$

where $Q = A + A^\top$ is a symmetric matrix. We use the notation $\mathcal{QF}(x; Q) = \dfrac{1}{2} x^\top Q x$ to define a canonical quadratic form. A generalized quadratic form is a polynomial including terms with degrees one and zero[8]:

$$\mathcal{QF}(x; Q, R, c) = \frac{1}{2} x^\top Q x - x^\top R + c$$

Main properties

We list here some properties that are helpful when considering portfolio optimization:

- The multiplication of a quadratic form by a scalar $\varphi \geq 0$ remains a quadratic form:

$$\varphi \cdot \mathcal{QF}(w; Q, R, c) = \mathcal{QF}(w; \varphi Q, \varphi R, \varphi c)$$

- The sum of two quadratic forms is a quadratic form:

$$\mathcal{QF}(x; Q_1, R_1, c_1) + \mathcal{QF}(x; Q_2, R_2, c_2) = \mathcal{QF}(x; Q_1 + Q_2, R_1 + R_2, c_1 + c_2)$$

- A quadratic form applied to the difference vector $x - y$ is a quadratic form in x:

$$\begin{aligned}
\mathcal{QF}(x - y; Q, R, c) &= \frac{1}{2}(x - y)^\top Q (x - y) - (x - y)^\top R + c \\
&= \frac{1}{2}\left(x^\top Q x - 2 x^\top Q y + y^\top Q y\right) - x^\top R + y^\top R + c \\
&= \frac{1}{2} x^\top Q x - x^\top (R + Q y) + \frac{1}{2} y^\top Q y + y^\top R + c \\
&= \mathcal{QF}\left(x; Q, R + Q y, \frac{1}{2} y^\top Q y + y^\top R + c\right)
\end{aligned}$$

and a quadratic form in y:

$$\begin{aligned}
\mathcal{QF}(x - y; Q, R, c) &= \frac{1}{2} y^\top Q y - y^\top (Q x - R) + \frac{1}{2} x^\top Q x - x^\top R + c \\
&= \mathcal{QF}\left(y; Q, Q x - R, \frac{1}{2} x^\top Q x - x^\top R + c\right)
\end{aligned}$$

---

[8]We add the one-degree polynomial $-x^\top R$ instead of $+x^\top R$ in order to match the canonical form of a QP problem.

- When Q is a diagonal matrix, the quadratic form reduces to:

$$\mathcal{QF}\left(x;Q,R,c\right) = \frac{1}{2}\sum_{i=1}^{n}q_{i,i}x_i^2 - \sum_{i=1}^{n}r_ix_i + c$$

  We deduce that:

$$\frac{1}{2}\sum_{i=1}^{n}q_ix_i^2 = \mathcal{QF}\left(x;\mathcal{D}\left(q\right),\mathbf{0}_n,0\right)$$

  where $q = (q_1,\ldots,q_n)$ is a $n \times 1$ vector and $\mathcal{D}\left(q\right) = \mathrm{diag}\left(q\right)$.

- Using the previous properties, we have:

$$
\begin{aligned}
\frac{1}{2}\sum_{i=1}^{n}q_i\left(x_i - y_i\right)^2 &= \mathcal{QF}\left(x - y;\mathcal{D}\left(q\right),\mathbf{0}_n,0\right) \\
&= \mathcal{QF}\left(x;\mathcal{D}\left(q\right),\mathcal{D}\left(q\right)y,\frac{1}{2}y^{\top}\mathcal{D}\left(q\right)y\right)
\end{aligned}
$$

- The square of the weighted sum is also a quadratic form:

$$\frac{1}{2}\left(\sum_{i=1}^{n}q_ix_i\right)^2 = \frac{1}{2}\left(x^{\top}q\right)^2 = \frac{1}{2}\left(x^{\top}q\right)\left(x^{\top}q\right)^{\top} = \frac{1}{2}x^{\top}qq^{\top}x^{\top} = \mathcal{QF}\left(x;\mathcal{T}\left(q\right),\mathbf{0}_n,0\right)$$

  where $\mathcal{T}\left(q\right) = qq^{\top}$.

- We deduce that:

$$
\begin{aligned}
\frac{1}{2}\left(\sum_{i=1}^{n}q_i\left(x_i - y_i\right)\right)^2 &= \mathcal{QF}\left(x - y;\mathcal{T}\left(q\right),\mathbf{0}_n,0\right) \\
&= \mathcal{QF}\left(x;\mathcal{T}\left(q\right),\mathcal{T}\left(q\right)y,\frac{1}{2}y^{\top}\mathcal{T}\left(q\right)y\right)
\end{aligned}
$$

- The previous results can be extended when we consider partial sums:

$$\sum_{i\in\Omega}q_ix_i = \sum_{i=1}^{n}\mathbb{1}\left\{i \in \Omega\right\} \cdot q_ix_i = \sum_{i=1}^{n}\left(\omega_iq_i\right)x_i = \sum_{i=1}^{n}\tilde{q}_ix_i$$

  where $\omega_i = \mathbb{1}\left\{i \in \Omega\right\}$ and $\tilde{q}_i = \omega_iq_i$. In a matrix form, we have:

$$\sum_{i\in\Omega}q_ix_i = \left(\omega \circ q\right)^{\top}x$$

  where $\omega = (\omega_1,\ldots,\omega_n)$. We deduce that:

$$
\begin{cases}
\frac{1}{2}\sum_{i\in\Omega}q_ix_i^2 = \mathcal{QF}\left(x;\mathcal{D}\left(\omega \circ q\right),\mathbf{0}_n,0\right) \\
\frac{1}{2}\sum_{i\in\Omega}q_i\left(x_i - y_i\right)^2 = \mathcal{QF}\left(x;\mathcal{D}\left(\omega \circ q\right),\mathcal{D}\left(\omega \circ q\right)y,\frac{1}{2}y^{\top}\mathcal{D}\left(\omega \circ q\right)y\right) \\
\frac{1}{2}\left(\sum_{i\in\Omega}q_ix_i\right)^2 = \mathcal{QF}\left(x;\mathcal{T}\left(\omega \circ q\right),\mathbf{0}_n,0\right) \\
\frac{1}{2}\left(\sum_{i\in\Omega}q_i\left(x_i - y_i\right)\right)^2 = \mathcal{QF}\left(x;\mathcal{T}\left(\omega \circ q\right),\mathcal{T}\left(\omega \circ q\right)y,\frac{1}{2}y^{\top}\mathcal{T}\left(\omega \circ q\right)y\right)
\end{cases}
$$

  We notice that $\mathcal{D}\left(\omega \circ q\right) = \mathrm{diag}\left(\omega \circ q\right) = \mathcal{D}\left(\omega\right)\mathcal{D}\left(q\right)$ and $\mathcal{T}\left(\omega \circ q\right) = \left(\omega \circ q\right)\left(\omega \circ q\right)^{\top} = \left(\omega\omega^{\top}\right) \circ qq^{\top} = \mathcal{T}\left(\omega\right) \circ \mathcal{T}\left(q\right)$.

A.1.4   Nonnegative matrix

Definition

$A = (a_{i,j})$ is a nonnegative matrix if all the elements are greater than or equal to zero: $a_{i,j} \geq 0$. Here are some special cases:

- A positive matrix is one where all elements are strictly greater than zero.

- A stochastic matrix is a square nonnegative matrix whose rows or columns are probability vectors[9].

  - A right/row stochastic matrix is a stochastic matrix where each row sums to 1: $\sum_{j=1}^{n} a_{i,j} = 1$.
  - A left/column stochastic matrix is a stochastic matrix where each column sums to 1: $\sum_{i=1}^{n} a_{i,j} = 1$.
  - A doubly stochastic matrix is a right and left stochastic matrix.

- A substochastic matrix is a square nonnegative matrix whose rows or columns add up to at most 1.

  - A is a right/row substochastic matrix if $\sum_{j=1}^{n} a_{i,j} \leq 1$.
  - A is a left/column substochastic matrix if $\sum_{i=1}^{n} a_{i,j} \leq 1$.
  - A is a doubly substochastic matrix if it is a right and left substochastic matrix.

Ordering properties

A is greater than or equal to B if and only if $A_{i,j} \geq B_{i,j}$ for all i, j. We use the notation $A \succeq B$. By definition, a square nonnegative matrix A satisfies $A \succeq \mathbf{0}_{n,n}$. If A is a stochastic matrix, we have $\mathbf{0}_{n,n} \preceq A \preceq \mathbf{1}_{n,n}$. Let A, B, C and D be square nonnegative matrices. We can show that:

(NN1) $A \succeq B \Rightarrow AC \succeq BC$;

(NN2) $A \succeq B \wedge C \succeq D \Rightarrow AC \succeq BD$;

(NN3) $A \succeq B \wedge C \succeq D \Rightarrow A + C \succeq B + D$;

(NN4) $A \succeq B \wedge k \geq 1 \Rightarrow A^k \succeq B^k$;

The proofs can be found in Desnos *et al.* (2023).

---

[9]This means that each entry is a real number between 0 and 1 and the sum of entries is equal to 1.

A.2 Statistical and probability analysis

A.2.1 Probability distributions

The Bernoulli distribution

The Bernoulli random variable X takes the value 1 with success probability of p and the value 0 with failure probability of $q = 1 - p$. We note $X \sim \mathcal{B}(p)$. The probability mass function may also be expressed as follows:

$$\Pr\{X = k\} = p^k (1 - p)^{1-k} \qquad \text{with } k = 0, 1$$

We have $\mathbb{E}[X] = p$ and $\operatorname{var}(X) = p(1 - p)$.

The binomial distribution

The binomial random variable X is the sum of n independent Bernoulli random variables with the same probability of success p:

$$X = \sum_{i=1}^{n} \mathcal{B}_i(p)$$

We note $X \sim \mathcal{B}(n, p)$. The probability mass function is equal to:

$$\Pr\{X = k\} = \binom{n}{k} p^k (1 - p)^{n-k} \qquad \text{with } k = 0, 1, \ldots, n$$

We have $\mathbb{E}[X] = np$ and $\operatorname{var}(X) = np(1 - p)$.

The geometric distribution

The geometric random variable X is the number of Bernoulli trials needed to get one success. We note $X \sim \mathcal{G}(p)$. The probability mass function is equal to:

$$\Pr\{X = k\} = (1 - p)^{k-1} p \qquad \text{with } k \in \mathbb{N}^\star$$

We have $\mathbb{E}[X] = 1/p$ and $\operatorname{var}(X) = (1 - p)/p^2$.

Remark 104 *If we define X as the number of failures before the first success, we have $\Pr\{X = k\} = (1 - p)^k p$ with $k \in \mathbb{N}$, $\mathbb{E}[X] = (1 - p)/p$ and $\operatorname{var}(X) = (1 - p)/p^2$.*

The Poisson distribution

The Poisson random variable X is the number of times an event occurs in the unit interval of time. We note $X \sim \mathcal{P}(\lambda)$ where λ is the parameter of the Poisson distribution. The probability mass function is equal to:

$$\Pr\{X = k\} = \frac{\lambda^k e^{-\lambda}}{k!} \qquad \text{with } k \in \mathbb{N}$$

We have $\mathbb{E}[X] = \operatorname{var}(X) = \lambda$. The parameter λ is then the expected number of events occurring in the unit interval of time.

The 1- and n-diversity distributions

Let $\pi_n \in \mathbb{R}_+^n$ such that $\mathbf{1}^\top \pi_n = 1$. π_n is then a probability distribution. The probability distribution π_n^- is perfectly concentrated if there exists one observation i_0 such that $\pi_{n,i_0}^- = 1$ and $\pi_{n,i}^- = 0$ if $i \neq i_0$. When n tends to $+\infty$, the limit distribution is noted π_∞^-. On the opposite, the probability distribution π_n^+ such that $\pi_{n,i}^+ = 1/n$ for all $i = 1, \ldots, n$ has no concentration. π_n^- and π_n^+ are respectively called 1- and n-diversity distributions.

The uniform distribution

X is a uniform random variable $\mathcal{U}_{[a,b]}$ if the density function is $f(x) = (b-a)^{-1}$ for $x \in [a, b]$. We deduce that the cumulative density function is equal to:

$$\mathbf{F}(x) = \begin{cases} 0 & \text{if } x \leq a \\ \dfrac{x-a}{b-a} & \text{if } x \in [a, b] \\ 1 & \text{if } x \geq b \end{cases}$$

We have:

$$\mathbb{E}[X] = \frac{a+b}{2}$$

and:

$$\text{var}(X) = \frac{(b-a)^2}{12}$$

The standard uniform distribution $\mathcal{U}_{[0,1]}$ is obtained with $a = 0$ and $b = 1$.

Remark 105 *The uniform distribution is related to the probability integral transform. We assume that $X \sim \mathbf{F}$ is a continuous random variable. Let $U = \mathbf{F}(X)$ be the integral transform of X. Its cumulative distribution function $\mathbf{G}$ is equal to:*

$$\begin{aligned} \mathbf{G}(u) &= \Pr\{U \leq u\} \\ &= \Pr\{\mathbf{F}(X) \leq u\} \\ &= \Pr\{X \leq \mathbf{F}^{-1}(u)\} \\ &= \mathbf{F}(\mathbf{F}^{-1}(u)) \\ &= u \end{aligned}$$

where $\mathbf{G}(0) = 0$ and $\mathbf{G}(1) = 1$. We deduce that U is a standard uniform random variable.

The gamma distribution

The gamma distribution is a two-parameter family of continuous probability distributions, whose support is $[0, \infty)$. We note $X \sim \mathcal{G}(\alpha, \beta)$ where $\alpha > 0$ and $\beta > 0$. α and β are called the shape parameter and the rate parameter. The probability density function is equal to:

$$f(x) = \frac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)}$$

where $\Gamma(\alpha)$ is the gamma function defined as:

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} \, \mathrm{d}t$$

The cumulative distribution function is the regularized gamma function:

$$\mathbf{F}(x) = \frac{\gamma(\alpha, \beta x)}{\Gamma(\alpha)}$$

where $\gamma(\alpha, x)$ is the lower incomplete gamma function defined as:

$$\gamma(\alpha, x) = \int_0^x t^{\alpha-1} e^{-t} \, \mathrm{d}t$$

We have $\mathbb{E}[X] = \alpha/\beta$ and $\mathrm{var}(X) = \alpha/\beta^2$. We verify the following properties:

- $\mathcal{G}(1, \beta) \sim \mathcal{E}(\beta)$;

- if $X \sim \mathcal{G}(\alpha, \beta)$, then $cX \sim \mathcal{G}(\alpha, \beta/c)$ when $c > 0$;

- $\sum_{i=1}^n \mathcal{G}(\alpha_i, \beta) \sim \mathcal{G}(\sum_{i=1}^n \alpha_i, \beta)$.

Remark 106 *The standard gamma distribution corresponds to $\mathcal{G}(\alpha, 1)$ and is denoted by $\mathcal{G}(\alpha)$.*

The beta distribution

The beta distribution is a two-parameter family of continuous probability distributions defined on the interval $[0, 1]$. We note $X \sim \mathcal{B}(\alpha, \beta)$ where $\alpha > 0$ and $\beta > 0$. The probability density function is equal to:

$$f(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{\mathfrak{B}(\alpha, \beta)}$$

where $\mathfrak{B}(\alpha, \beta)$ is the beta function defined as:

$$\mathfrak{B}(\alpha, \beta) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1} \, \mathrm{d}t = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

The cumulative distribution function is the regularized incomplete beta function:

$$\mathbf{F}(x) = \mathcal{IB}(x; \alpha, \beta) = \frac{\mathfrak{B}(x; \alpha, \beta)}{\mathfrak{B}(\alpha, \beta)}$$

where $\mathfrak{B}(x; \alpha, \beta)$ is the incomplete beta function defined as:

$$\mathfrak{B}(x; \alpha, \beta) = \int_0^x t^{\alpha-1}(1-t)^{\beta-1} \, \mathrm{d}t$$

We have $\mathbb{E}[X] = \alpha/(\alpha+\beta)$ and:

$$\mathrm{var}(X) = \frac{\alpha\beta}{(\alpha+\beta)^2 (\alpha+\beta+1)}$$

The exponential distribution

X is an exponential random variable $\mathcal{E}(\lambda)$ if the density function is $f(x) = \lambda e^{-\lambda x}$ for $x \geq 0$. We deduce that $\mathbf{F}(x) = 1 - e^{-\lambda x}$. We have $\mathbb{E}[X] = 1/\lambda$ and $\mathrm{var}(X) = 1/\lambda^2$. More generally, we can show that $\mathbb{E}[X^n] = n!/\lambda^n$. This distribution verifies the lack of memory property:

$$\Pr\{X \geq s + t \mid X \geq s\} = \Pr\{X \geq t\}$$

for all $s \geq 0$ and $t \geq 0$.

The normal distribution

Let $\mathbb{C}$ be a correlation matrix. We consider the standardized Gaussian random vector $X \sim \mathcal{N}\left(\mathbf{0}, \mathbb{C}\right)$ of dimension n. We note $\phi_n\left(x; \mathbb{C}\right)$ the associated density function defined as:

$$\phi_n\left(x; \mathbb{C}\right) = (2\pi)^{-n/2} \left|\mathbb{C}\right|^{-1/2} \exp\left(-\frac{1}{2} x^\top \mathbb{C}^{-1} x\right)$$

We deduce that the expression of cumulative distribution function is:

$$\Phi_n\left(x; \mathbb{C}\right) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_2} \phi_n\left(u; \mathbb{C}\right) \, \mathrm{d}u$$

By construction, we have $\mathbb{E}\left[X\right] = \mathbf{0}$ and $\mathrm{cov}\left(x\right) = \mathbb{C}$. In the bivariate case, we use the notations $\phi_2\left(x_1, x_2; \rho\right) = \phi_2\left(x; \mathbb{C}\right)$ and $\Phi_2\left(x_1, x_2; \rho\right) = \Phi_2\left(x; \mathbb{C}\right)$ where $\rho = \mathbb{C}_{1,2}$ is the correlation between the components X_1 and X_2. In the univariate case, we also consider the alternative notations $\phi\left(x\right) = \phi_1\left(x; 1\right)$ and $\Phi\left(x\right) = \Phi_1\left(x; 1\right)$. The density function reduces then to:

$$\phi\left(x\right) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} x^2\right)$$

Concerning the moments, we have $\mu\left(X\right) = 0$, $\sigma\left(X\right) = 1$, $\gamma_1\left(X\right) = 0$ and $\gamma_2\left(X\right) = 0$.

Adding a mean vector μ and a covariance matrix Σ is equivalent to apply the linear transformation to X:

$$Y = \mu + \sigma X$$

where $\sigma = \mathrm{diag}^{1/2}\left(\Sigma\right)$.

The log-normal distribution

Let $Z \sim \mathcal{N}\left(\mu, \sigma^2\right)$ be a normal-distributed random variable. $X = e^Z$ is a log-normal random variable and we note $X \sim \mathcal{LN}\left(\mu, \sigma^2\right)$. The probability distribution function is equal to:

$$f\left(x\right) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

whereas the cumulative distribution function has the following expression:

$$\mathbf{F}\left(x\right) = \Phi\left(\frac{\ln x - \mu}{\sigma}\right)$$

We have:

$$\mathbb{E}\left[X\right] = e^{\mu + \frac{1}{2}\sigma^2}$$

and:

$$\mathrm{var}\left(X\right) = e^{2\mu + \sigma^2}\left(e^{\sigma^2} - 1\right)$$

The inverse Gaussian distribution

The inverse Gaussian distribution is a two-parameter family of continuous probability distribution, whose support is $[0, \infty)$. We note $X \sim \mathcal{IG}(\mu, \lambda)$ where $\mu > 0$ and $\lambda > 0$, μ and λ are called the mean parameter and the shape parameter. The probability density function is equal to:

$$f(x) = \sqrt{\frac{\lambda}{2\pi x^3}} \exp\left(-\frac{\lambda}{2\mu^2 x}(x - \mu)^2\right)$$

while the cumulative density function has the following expression:

$$\mathbf{F}(x) = \Phi\left(\sqrt{\frac{\lambda}{x}}\left(\frac{x - \mu}{\mu}\right)\right) + \exp\left(2\frac{\lambda}{\mu}\right)\Phi\left(-\sqrt{\frac{\lambda}{x}}\left(\frac{x + \mu}{\mu}\right)\right)$$

We have $\mathbb{E}[X] = \mu$ and $\operatorname{var}(X) = \mu^3/\lambda$. We verify the following properties:

- If $X \sim \mathcal{IG}(\mu, \lambda)$, then $cX \sim \mathcal{IG}(c\mu, c\lambda)$ when $c > 0$;

- If $X_i \sim \mathcal{IG}(\mu, \lambda)$, then $\sum_{i=1}^{n} X_i \sim \mathcal{IG}(n\mu, n^2\lambda)$;

- If $X \sim \mathcal{IG}(\mu, \lambda)$, then $\mathbb{E}[X^{-k}] = \mu^{-(2k+1)}\mathbb{E}(X^{k+1})$ for $k \in \mathbb{N}^*$.

The Pareto distribution

The Pareto distribution is denoted by $\mathcal{P}(\alpha, x_-)$. We have:

$$f(x) = \frac{\alpha}{x}\left(\frac{x}{x_-}\right)^{-\alpha}$$

and:

$$\mathbf{F}(x) = 1 - \left(\frac{x}{x_-}\right)^{-\alpha}$$

where $x \geq x_-$, $\alpha > 0$ and $x_- > 0$. Concerning the first two moments, we obtain:

$$\mathbb{E}[X] = \frac{\alpha x_-}{\alpha - 1}$$

if $\alpha > 1$ and:

$$\operatorname{var}(X) = \frac{\alpha x_-^2}{(\alpha - 1)^2(\alpha - 2)}$$

if $\alpha > 2$.

Remark 107 *The Pareto distribution: can be parameterized as follows;*

$$\mathbf{F}(x) = 1 - \left(\frac{\theta + x}{\theta}\right)^{-\alpha}$$

where $x \geq 0$, $\alpha > 0$ and $\theta > 0$. In this case, it is denoted by $\mathcal{P}(\alpha, \theta)$.

The generalized extreme value distribution

The generalized extreme value distribution is denoted by $\mathcal{GEV}\left(\mu,\sigma,\xi\right)$. We have:

$$f\left(x\right) = \frac{1}{\sigma}\left(1+\xi\left(\frac{x-\mu}{\sigma}\right)\right)^{-(1+1/\xi)}\exp\left(-\left(1+\xi\left(\frac{x-\mu}{\sigma}\right)\right)^{-1/\xi}\right)$$

and:

$$\mathbf{F}\left(x\right) = \exp\left(-\left(1+\xi\left(\frac{x-\mu}{\sigma}\right)\right)^{-1/\xi}\right)$$

where $x > \mu - \sigma/\xi$, $\sigma > 0$ and $\xi > 0$. Concerning the first two moments, we obtain:

$$\mathbb{E}\left[X\right] = \mu + \frac{\sigma}{\xi}\left(\Gamma\left(1-\xi\right)-1\right)$$

if $\xi < 1$ and:

$$\mathrm{var}\left(X\right) = \frac{\sigma^2}{\xi^2}\left(\Gamma\left(1-2\xi\right)-\Gamma^2\left(1-\xi\right)\right)$$

if $\xi < 1/2$.

The generalized Pareto distribution

The generalized Pareto distribution is denoted by $\mathcal{GPD}\left(\sigma,\xi\right)$. We have:

$$f\left(x\right) = \frac{1}{\sigma}\left(1+\frac{\xi x}{\sigma}\right)^{-1/\xi-1}$$

and:

$$\mathbf{F}\left(x\right) = 1 - \left(1+\frac{\xi x}{\sigma}\right)^{-1/\xi}$$

where $x \geq 0$, $\sigma > 0$ and $\xi > 0$. Concerning the first two moments, we obtain:

$$\mathbb{E}\left[X\right] = \frac{\sigma}{1-\xi}$$

if $\xi < 1$ and:

$$\mathrm{var}\left(X\right) = \frac{\sigma^2}{\left(1-\xi\right)^2\left(1-2\xi\right)}$$

if $\xi < 1/2$.

The Bates distribution

Let $U_1,\ldots,U_n$ be a sequence of *iid* uniform random variables $\mathcal{U}_{[0,1]}$. The mean $\bar{U}_n$ follows the Bates distribution:

$$X := \bar{U}_n = \frac{1}{n}\sum_{k=1}^{n} U_k \sim \mathfrak{B}ates\left(n\right)$$

The support is $[0,1]$. The probability density function is equal to[10]:

$$f\left(x\right) = \frac{n}{2\left(n-1\right)!}\sum_{k=0}^{n}\left(-1\right)^k\binom{n}{k}\left(nx-k\right)^{n-1}\mathrm{sgn}\left(nx-k\right)$$

---

[10]If $k = nx$, $\mathrm{sgn}\left(nx-k\right) = 0$.

whereas the cumulative distribution function has the following expression:

$$\mathbf{F}(x) = \frac{1}{n!} \sum_{k=0}^{n} \mathbb{1}\{nx > k\}(-1)^k \binom{n}{k}(nx - k)^n$$

We have:

$$\mathbb{E}[X] = \frac{1}{2}$$

and

$$\text{var}(X) = \frac{1}{12n}$$

Conditional probability distribution in the Gaussian case

Let us consider a Gaussian random vector defined as follows:

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}, \begin{pmatrix} \Sigma_{x,x} & \Sigma_{x,y} \\ \Sigma_{y,x} & \Sigma_{y,y} \end{pmatrix} \right)$$

The conditional probability distribution of Y given $X = x$ is a multivariate normal distribution. We have:

$$\mu_{y|x} = \mathbb{E}[Y \mid X = x] = \mu_y + \Sigma_{y,x}\Sigma_{x,x}^{-1}(x - \mu_x)$$

and:

$$\Sigma_{y,y|x} = \sigma^2[Y \mid X = x] = \Sigma_{y,y} - \Sigma_{y,x}\Sigma_{x,x}^{-1}\Sigma_{x,y}$$

We deduce that:

$$Y = \mu_y + \Sigma_{y,x}\Sigma_{x,x}^{-1}(x - \mu_x) + u$$

where u is a centered Gaussian random variable with variance $\sigma^2 = \Sigma_{y,y|x}$. It follows that:

$$Y = \underbrace{\left(\mu_y - \Sigma_{y,x}\Sigma_{x,x}^{-1}\mu_x\right)}_{\beta_0} + \underbrace{\Sigma_{y,x}\Sigma_{x,x}^{-1}}_{\beta^\top}x + u$$

We recognize the linear regression of Y on a constant and a set of exogenous variables X:

$$Y = \beta_0 + \beta^\top X + u$$

Moreover, we have:

$$\mathfrak{R}^2 = 1 - \frac{\sigma^2}{\Sigma_{y,y}} = \frac{\Sigma_{y,x}\Sigma_{x,x}^{-1}\Sigma_{x,y}}{\Sigma_{y,y}}$$

Change of variables

Let X be a random variable whose probability density function is $f(x)$. We consider the change of the variable $Y = \varphi(X)$. If the function φ is monotone, then the probability density function $g(y)$ of Y is equal to:

$$g(y) = f(x)\left|\frac{\mathrm{d}x}{\mathrm{d}y}\right|$$

In the multivariate case, let $(X_1, \ldots, X_n)$ be the random vector with density function $f(x_1, \ldots, x_n)$. If the function φ is bijective, we can show that the probability density function of $(Y_1, \ldots, Y_n) = \varphi(X_1, \ldots, X_n)$ is equal to:

$$g(y_1, \ldots, y_n) = f(x_1, \ldots, x_n)\left|\frac{1}{\det J_\varphi}\right|$$

where J_φ is the Jacobian associated with the variable change.

Ratio distribution

Given two independent random variables X and Y, the distribution of the random variable $Z = Y/X$ is a ratio distribution. We have:

$$
\begin{aligned}
\mathbf{F}_z\left(z\right) &= \Pr\left\{Z \le z\right\} \\
&= \Pr\left\{\frac{X}{Y} \le z\right\} \\
&= \Pr\left\{X \ge zY, Y \le 0\right\} + \Pr\left\{X \le zY, Y > 0\right\} \\
&= \int_{-\infty}^{0}\int_{zy}^{\infty} f_{x,y}\left(x,y\right)\,\mathrm{d}x\,\mathrm{d}y + \int_{0}^{\infty}\int_{-\infty}^{zy} f_{x,y}\left(x,y\right)\,\mathrm{d}x\,\mathrm{d}y \\
&= \int_{-\infty}^{0}\int_{zy}^{\infty} f_{x}\left(x\right) f_{y}\left(y\right)\,\mathrm{d}x\,\mathrm{d}y + \int_{0}^{\infty}\int_{-\infty}^{zy} f_{x}\left(x\right) f_{y}\left(y\right)\,\mathrm{d}x\,\mathrm{d}y \\
&= \int_{-\infty}^{0}\left(\int_{zy}^{\infty} f_{x}\left(x\right)\,\mathrm{d}x\right) f_{y}\left(y\right)\,\mathrm{d}y + \int_{0}^{\infty}\left(\int_{-\infty}^{zy} f_{x}\left(x\right)\,\mathrm{d}x\right) f_{y}\left(y\right)\,\mathrm{d}y \\
&= \int_{-\infty}^{0}\left(1 - \mathbf{F}_x\left(zy\right)\right) f_{y}\left(y\right)\,\mathrm{d}y + \int_{0}^{\infty}\mathbf{F}_x\left(zy\right) f_{y}\left(y\right)\,\mathrm{d}y
\end{aligned}
$$

and:

$$
\begin{aligned}
f_z\left(z\right) &= \partial_z \Pr\left\{Z \le z\right\} \\
&= -\int_{-\infty}^{0} y f_{x}\left(zy\right) f_{y}\left(y\right)\,\mathrm{d}y + \int_{0}^{\infty} y f_{x}\left(zy\right) f_{y}\left(y\right)\,\mathrm{d}y \\
&= \int_{-\infty}^{0} \left|y\right| f_{x}\left(zy\right) f_{y}\left(y\right)\,\mathrm{d}y + \int_{0}^{\infty} \left|y\right| f_{x}\left(zy\right) f_{y}\left(y\right)\,\mathrm{d}y \\
&= \int_{-\infty}^{\infty} \left|y\right| f_{x}\left(zy\right) f_{y}\left(y\right)\,\mathrm{d}y
\end{aligned}
$$

In the case where $X \sim \mathcal{N}\left(\mu_x, \sigma_x^2\right)$ and $Y \sim \mathcal{N}\left(\mu_y, \sigma_y^2\right)$, we note $Z \sim \mathcal{NRD}\left(\mu_x, \sigma_x^2, \mu_y, \sigma_y^2\right)$, and Hinkley (1969) showed that:

$$
\mathbf{F}_z\left(z\right) = \Phi_2\left(-\frac{\mu_y z - \mu_x}{\sigma_x \sigma_y a\left(z\right)}, -\frac{\mu_y}{\sigma_y}, \rho_z\right) + \Phi_2\left(\frac{\mu_y z - \mu_x}{\sigma_x \sigma_y a\left(z\right)}, \frac{\mu_y}{\sigma_y}, \rho_z\right)
$$

and:

$$
f_z\left(z\right) = \frac{b\left(z\right)}{\sigma_x \sigma_y \sqrt{2\pi} a^3\left(z\right)} \left(2\Phi\left(\frac{b\left(z\right)}{a\left(z\right)}\right) - 1\right) \exp\left(\frac{b^2\left(z\right) - c a^2\left(z\right)}{2a^2\left(z\right)}\right) + \frac{1}{\sigma_x \sigma_y a^2\left(z\right) \pi}\exp\left(-\frac{c}{2}\right)
$$

where $a\left(z\right) = \sqrt{\dfrac{1}{\sigma_x^2}z^2 + \dfrac{1}{\sigma_y^2}}$, $b\left(z\right) = \dfrac{\mu_x}{\sigma_x^2}z + \dfrac{\mu_y}{\sigma_y^2}$, $c = \dfrac{\mu_x^2}{\sigma_x^2} + \dfrac{\mu_y^2}{\sigma_y^2}$ and $\rho_z = \dfrac{z}{\sigma_x a\left(z\right)}$. Depending on the parameter values, Z has one or two modes. The moments of Z do not exist. In the case where $\mu_x = \mu_y = 0$, $b\left(z\right) = 0$ and $c = 0$, which implies that Z follows a Cauchy distribution:

$$
f_z\left(z\right) = \frac{\sigma_x \sigma_y}{\pi\left(\sigma_y^2 z^2 + \sigma_x^2\right)}
$$

The previous result can be easily extended to the correlated case, because $\mathbf{F}_z\left(z\right)$ has exactly the same expression with $a\left(z\right) = \sqrt{\dfrac{1}{\sigma_x^2}z^2 - 2\dfrac{\rho_{x,y}}{\sigma_x \sigma_y}z + \dfrac{1}{\sigma_y^2}}$ and $\rho_z = \dfrac{\sigma_y z - \rho_{x,y}\sigma_x}{\sigma_x \sigma_y a\left(z\right)}$.

A.2.2   Copula functions

The concept of copula has been introduced by Sklar in 1959. During a long time, only a small number of people have used copula functions, more in the field of mathematics than this of statistics. The publication of Genest and MacKay (1986b) in the *American Statistician* marks a breakdown and opens areas of study in empirical modeling, statistics and econometrics. In what follows, we intensively use the materials developed in the books of Joe (1997), Nelsen (2006) and Roncalli (2020a).

Definition and main properties

Nelsen (2006) defines a bi-dimensional copula (or a 2-copula) as a function $\mathbf{C}$ which satisfies the following properties:

1. Dom $\mathbf{C} = [0, 1] \times [0, 1]$;

2. $\mathbf{C}(0, u) = \mathbf{C}(u, 0) = 0$ and $\mathbf{C}(1, u) = \mathbf{C}(u, 1) = u$ for all u in $[0, 1]$;

3. $\mathbf{C}$ is 2-increasing:

$$\mathbf{C}(v_1, v_2) - \mathbf{C}(v_1, u_2) - \mathbf{C}(u_1, v_2) + \mathbf{C}(u_1, u_2) \geq 0$$

for all $(u_1, u_2) \in [0, 1]^2$, $(v_1, v_2) \in [0, 1]^2$ such that $0 \leq u_1 \leq v_1 \leq 1$ and $0 \leq u_2 \leq v_2 \leq 1$.

This definition means that $\mathbf{C}$ is a cumulative distribution function with uniform margins:

$$\mathbf{C}(u_1, u_2) = \Pr\{U_1 \leq u_1, U_2 \leq u_2\}$$

where U_1 and U_2 are two uniform random variables.

Let us consider the function $\mathbf{C}^{\perp}(u_1, u_2) = u_1 u_2$. We have $\mathbf{C}^{\perp}(0, u) = \mathbf{C}^{\perp}(u, 0) = 0$ and $\mathbf{C}^{\perp}(1, u) = \mathbf{C}^{\perp}(u, 1) = u$. Since we have $v_2 - u_2 \geq 0$ and $v_1 \geq u_1$, it follows that $v_1(v_2 - u_2) \geq u_1(v_2 - u_2)$ and $v_1 v_2 + u_1 u_2 - u_1 v_2 - v_1 u_2 \geq 0$. We deduce that $\mathbf{C}^{\perp}$ is a copula function. It is called the product copula.

Let $\mathbf{F}_1$ and $\mathbf{F}_2$ be any two univariate distributions. It is obvious that $\mathbf{F}(x_1, x_2) = \mathbf{C}(\mathbf{F}_1(x_1), \mathbf{F}_2(x_2))$ is a probability distribution with margins $\mathbf{F}_1$ and $\mathbf{F}_2$. Indeed, $u_i = \mathbf{F}_i(x_i)$ defines a uniform transformation ($u_i \in [0, 1]$). Moreover, we verify that $\mathbf{C}(\mathbf{F}_1(x_1), \mathbf{F}_2(\infty)) = \mathbf{C}(\mathbf{F}_1(x_1), 1) = \mathbf{F}_1(x_1)$. Copulas are then a powerful tool to build a multivariate probability distribution when the margins are given. Conversely, Sklar (1959) proves that any bivariate distribution $\mathbf{F}$ admits such a representation:

$$\mathbf{F}(x_1, x_2) = \mathbf{C}(\mathbf{F}_1(x_1), \mathbf{F}_2(x_2)) \tag{A.11}$$

and that the copula $\mathbf{C}$ is unique provided the margins are continuous. This result is important, because we can associate to each bivariate distribution a copula function. We then obtain a canonical representation of a bivariate probability distribution: on one side, we have the margins or the univariate directions $\mathbf{F}_1$ and $\mathbf{F}_2$; on the other side, we have the copula $\mathbf{C}$ that links these margins and gives the dependence between the unidimensional directions.

If the joint distribution function $\mathbf{F}(x_1, x_2)$ is absolutely continuous, we obtain:

$$\begin{aligned} f(x_1, x_2) &= \partial_{1,2} \mathbf{F}(x_1, x_2) \\ &= \partial_{1,2} \mathbf{C}(\mathbf{F}_1(x_1), \mathbf{F}_2(x_2)) \\ &= c(\mathbf{F}_1(x_1), \mathbf{F}_2(x_2)) \cdot f_1(x_1) \cdot f_2(x_2) \end{aligned} \tag{A.12}$$

where $f(x_1, x_2)$ is the joint probability density function, f_1 and f_2 are the marginal densities and c is the copula density:

$$c(u_1, u_2) = \partial_{1,2} \mathbf{C}(u_1, u_2)$$

We notice that the condition $\mathbf{C}(v_1, v_2) - \mathbf{C}(v_1, u_2) - \mathbf{C}(u_1, v_2) + \mathbf{C}(u_1, u_2) \geq 0$ is then equivalent to $\partial_{1,2} \mathbf{C}(u_1, u_2) \geq 0$ when the copula density exists. From Equation (A.12), we deduce that:

$$c(u_1, u_2) = \frac{f\left(\mathbf{F}_1^{-1}(u_1), \mathbf{F}_2^{-1}(u_2)\right)}{f_1\left(\mathbf{F}_1^{-1}(u_1)\right) \cdot f_2\left(\mathbf{F}_2^{-1}(u_2)\right)} \tag{A.13}$$

We obtain a second canonical representation based on density functions. For some copulas, there is no explicit analytical formula. This is the case of the Normal copula, which is equal to $\mathbf{C}(u_1, u_2; \rho) = \Phi\left(\Phi^{-1}(u_1), \Phi^{-1}(u_2); \rho\right)$. Using Equation (A.13), we can however characterize its density function:

$$
\begin{aligned}
c(u_1, u_2; \rho) &= \frac{2\pi\left(1 - \rho^2\right)^{-1/2} \exp\left(-\frac{1}{2(1-\rho^2)}\left(x_1^2 + x_2^2 - 2\rho x_1 x_2\right)\right)}{(2\pi)^{-1/2} \exp\left(-\frac{1}{2}x_1^2\right) \cdot (2\pi)^{-1/2} \exp\left(-\frac{1}{2}x_2^2\right)} \\
&= \frac{1}{\sqrt{1 - \rho^2}} \exp\left(-\frac{1}{2}\frac{\left(x_1^2 + x_2^2 - 2\rho x_1 x_2\right)}{(1 - \rho^2)} + \frac{1}{2}\left(x_1^2 + x_2^2\right)\right)
\end{aligned}
$$

where $x_1 = \mathbf{F}_1^{-1}(u_1)$ and $x_2 = \mathbf{F}_2^{-1}(u_2)$. It is then easy to generate bivariate non-normal distributions.

Example 54 *In Figure A.1, we have built a bivariate probability distribution by considering that the margins are a Student's distribution and a beta distribution. The copula function corresponds to the Normal copula such that its Kendall's tau is equal to* 60%.

Fréchet classes and concordance ordering

The goal of Fréchet classes is to study the structure of the class of distributions with given margins. Let us first consider the bivariate case. The distribution function $\mathbf{F}$ belongs to the Fréchet class $(\mathbf{F}_1, \mathbf{F}_2)$ and we note $\mathbf{F} \in \mathcal{F}(\mathbf{F}_1, \mathbf{F}_2)$ if an only if the margins of $\mathbf{F}$ are $\mathbf{F}_1$ and $\mathbf{F}_2$, meaning that $\mathbf{F}(x_1, \infty) = \mathbf{F}_1(x_1)$ and $\mathbf{F}(\infty, x_2) = \mathbf{F}_2(x_2)$. Characterizing the Fréchet class $\mathcal{F}(\mathbf{F}_1, \mathbf{F}_2)$ is then equivalent to find the set $\mathcal{C}$ of copula functions:

$$\mathcal{F}(\mathbf{F}_1, \mathbf{F}_2) = \{\mathbf{F} : \mathbf{F}(x_1, x_2) = \mathbf{C}(\mathbf{F}_1(x_1), \mathbf{F}_2(x_2)), \mathbf{C} \in \mathcal{C}\}$$

Therefore this problem does not depend on the margins $\mathbf{F}_1$ and $\mathbf{F}_2$. We can show that the extremal distribution functions $\mathbf{F}^-$ and $\mathbf{F}^+$ of the Fréchet class $\mathcal{F}(\mathbf{F}_1, \mathbf{F}_2)$ are $\mathbf{F}^-(x_1, x_2) = \max(\mathbf{F}_1(x_1) + \mathbf{F}_2(x_2) - 1, 0)$ and $\mathbf{F}^+(x_1, x_2) = \min(\mathbf{F}_1(x_1), \mathbf{F}_2(x_2))$. $\mathbf{F}^-$ and $\mathbf{F}^+$ are called the Fréchet lower and upper bounds. We deduce that the corresponding copula functions are:

$$\mathbf{C}^-(u_1, u_2) = \max(u_1 + u_2 - 1, 0)$$

and:

$$\mathbf{C}^+(u_1, u_2) = \min(u_1, u_2)$$

Figure A.1: Example of a bivariate probability distribution with given margins
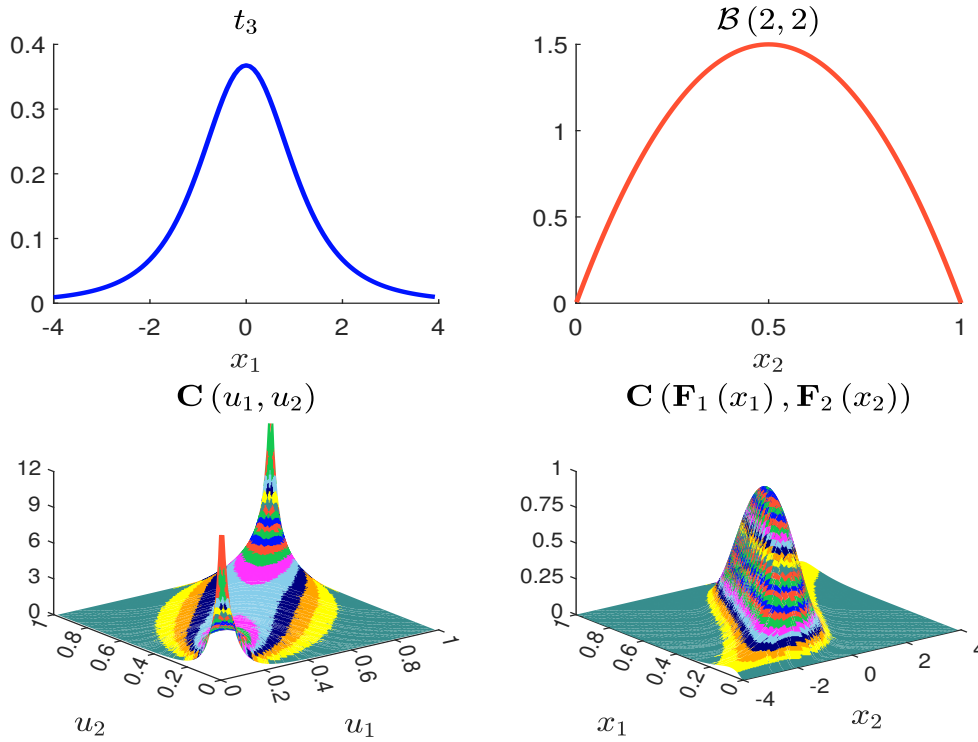
The extension of bivariate copulas to multivariate copulas is straightforward. Thus, the canonical decomposition of a multivariate distribution function is:

$$\mathbf{F}(x_1, \ldots, x_n) = \mathbf{C}(\mathbf{F}_1(x_1), \ldots, \mathbf{F}_n(x_n))$$

We note $\mathbf{C}_{\mathcal{E}}$ the sub-copula of $\mathbf{C}$ such that arguments that are not in the set $\mathcal{E}$ are equal to 1. For instance, with a dimension of 4, we have $\mathbf{C}_{12}(u, v) = \mathbf{C}(u, v, 1, 1)$ and $\mathbf{C}_{124}(u, v, w) = \mathbf{C}(u, v, 1, w)$. Let us consider the 2-copulas $\mathbf{C}_1$ and $\mathbf{C}_2$. It seems logical to build a copula of higher dimension with copulas of lower dimensions. In fact, the function $\mathbf{C}_1(u_1, \mathbf{C}_2(u_2, u_3))$ is not a copula in most cases. In the multivariate case, we define:

$$\mathbf{C}^-(u_1, \ldots, u_n) = \max\left(\sum_{i=1}^n u_i - n + 1, 0\right)$$

and:

$$\mathbf{C}^+(u_1, \ldots, u_n) = \min(u_1, \ldots, u_n)$$

We can show that $\mathbf{C}^+$ is a copula, but $\mathbf{C}^-$ does not belong to the set $\mathcal{C}$. Nevertheless, $\mathbf{C}^-$ is the best-possible bound, meaning that for all $(u_1, \ldots, u_n) \in [0, 1]^n$, there is a copula that coincide with $\mathbf{C}^-$ (Nelsen, 2006). This implies that $\mathcal{F}(\mathbf{F}_1, \ldots, \mathbf{F}_n)$ has a minimal distribution function if and only if $\max\left(\sum_{i=1}^n \mathbf{F}_i(x_i) - n + 1, 0\right)$ is a probability distribution.

We now introduce a stochastic ordering on copulas. Let $\mathbf{C}_1$ and $\mathbf{C}_2$ be two copula functions. We say that the copula $\mathbf{C}_1$ is smaller than the copula $\mathbf{C}_2$ and we note $\mathbf{C}_1 \prec \mathbf{C}_2$ if we verify that $\mathbf{C}_1(u_1, u_2) \leq \mathbf{C}_2(u_1, u_2)$ for all $(u_1, u_2) \in [0, 1]^2$. This stochastic ordering is called the concordance ordering and may be viewed as the first order of the stochastic dominance on probability

distributions. Using the previous results on Fréchet classes, we deduce that:

$$\mathbf{C}^- \prec \mathbf{C} \prec \mathbf{C}^+$$

for all $\mathbf{C} \in \mathcal{C}$. It follows that $\mathbf{C}^- \prec \mathbf{C}^\perp \prec \mathbf{C}^+$. A copula $\mathbf{C}$ has a positive quadrant dependence (PQD) if it satisfies the inequality $\mathbf{C}^\perp \prec \mathbf{C} \prec \mathbf{C}^+$. In a similar way, $\mathbf{C}$ has a negative quadrant dependence (NQD) if it satisfies the inequality $\mathbf{C}^- \prec \mathbf{C} \prec \mathbf{C}^\perp$. As it is a partial ordering, there exist copula functions $\mathbf{C}$ such that $\mathbf{C} \not\succ \mathbf{C}^\perp$ and $\mathbf{C} \not\prec \mathbf{C}^\perp$. A copula function may then have a dependence structure that is neither positive or negative. In Figure A.2, we report the cumulative distribution function (above panel) and its contour lines (below panel) of the three copula functions $\mathbf{C}^-$, $\mathbf{C}^\perp$ and $\mathbf{C}^+$, which plays an important role to understand the dependance between unidimensional margins.

Figure A.2: The three copula functions $\mathbf{C}^-$, $\mathbf{C}^\perp$ and $\mathbf{C}^+$
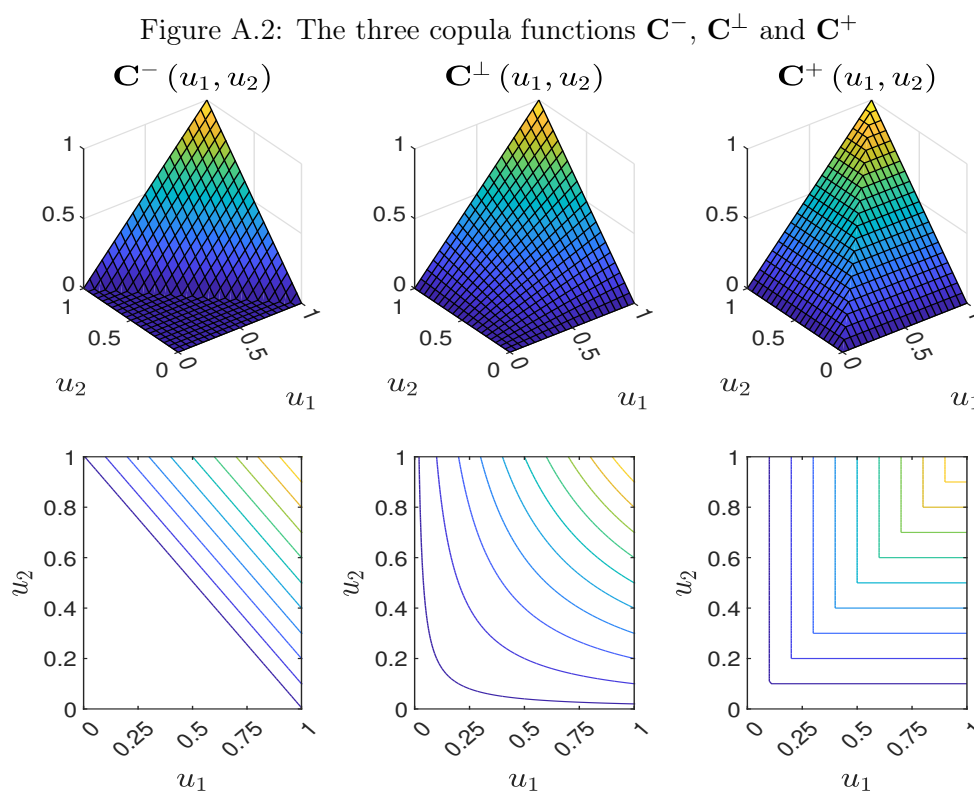
Let $\mathbf{C}_\theta(u_1, u_2) = \mathbf{C}(u_1, u_2; \theta)$ be a family of copula functions that depends on the parameter θ. The copula family $\{\mathbf{C}_\theta\}$ is totally ordered if, for all $\theta_2 \geq \theta_1$, $\mathbf{C}_{\theta_2} \succ \mathbf{C}_{\theta_1}$ (positively ordered) or $\mathbf{C}_{\theta_2} \prec \mathbf{C}_{\theta_1}$ (negatively ordered). For instance, the Frank copula defined by:

$$\mathbf{C}(u_1, u_2; \theta) = -\frac{1}{\theta} \ln \left(1 + \frac{\left(e^{-\theta u_1} - 1\right)\left(e^{-\theta u_2} - 1\right)}{e^{-\theta} - 1} \right)$$

where $\theta \in \mathbb{R}$ is a positively ordered family. An illustration is showed in Figure A.3). For a given value $\alpha \in [0, 1]$, we also verify that the level curves of $\mathbf{C}$ are in the triangle defined as follows:

$$\{(u_1, u_2) : \max(u_1 + u_2 - 1, 0) \leq \alpha, \min(u_1, u_2) \geq \alpha\}$$

Figure A.3: Concordance ordering of the Frank copula
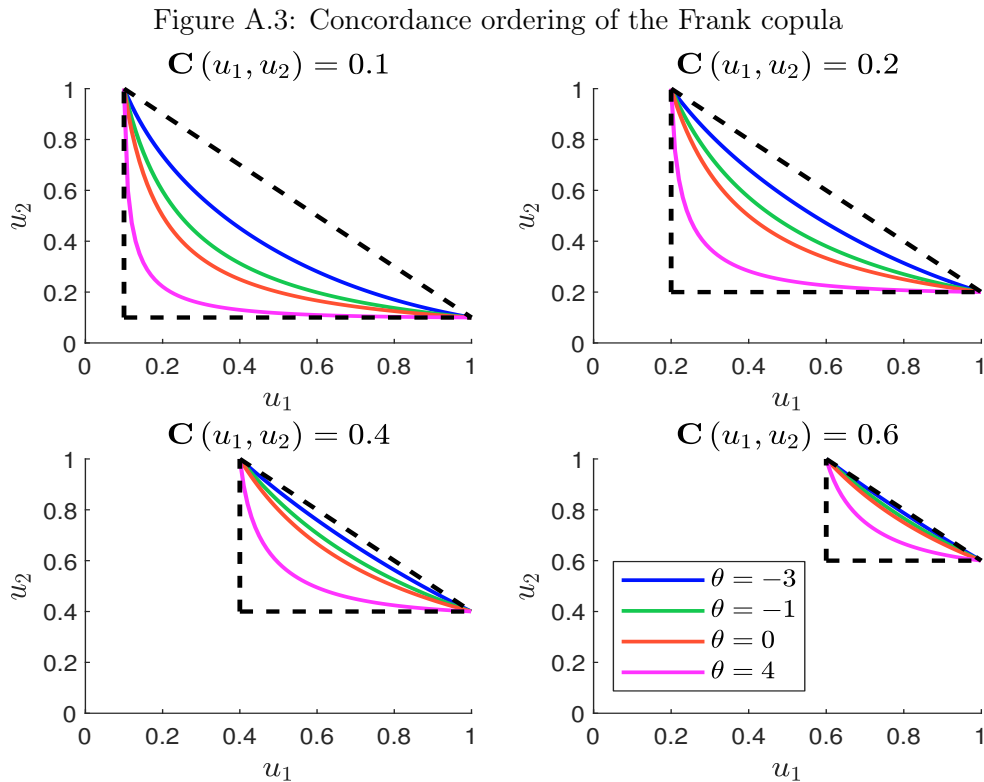
Copula function and random vector

Let $X = (X_1, X_2)$ be a random vector with distribution $\mathbf{F}$. We define the copula of (X_1, X_2) by the copula of $\mathbf{F}$:

$$\mathbf{F}(x_1, x_2) = \mathbf{C} \langle X_1, X_2 \rangle (\mathbf{F}_1(x_1), \mathbf{F}_2(x_2))$$

In what follows, we give the main results on the dependence of the random vector X found in Nelsen (2006). We first consider the probabilistic interpretation of the three copula functions $\mathbf{C}^-$, $\mathbf{C}^\perp$ and $\mathbf{C}^+$:

- X_1 and X_2 are countermonotonic — or $\mathbf{C} \langle X_1, X_2 \rangle = \mathbf{C}^-$ — if there exists a random variable X such that $X_1 = f_1(X)$ and $X_2 = f_2(X)$ where f_1 and f_2 are respectively decreasing and increasing functions[11];

- X_1 and X_2 are independent if the dependence function is the product copula $\mathbf{C}^\perp$;

- X_1 are X_2 are comonotonic — or $\mathbf{C} \langle X_1, X_2 \rangle = \mathbf{C}^+$ — if there exists a random variable X such that $X_1 = f_1(X)$ and $X_2 = f_2(X)$ where f_1 and f_2 are both increasing functions[12].

Let us consider a uniform random vector (U_1, U_2). We have $U_2 = 1 - U_1$ when $\mathbf{C} \langle X_1, X_2 \rangle = \mathbf{C}^-$ and $U_2 = U_1$ when $\mathbf{C} \langle X_1, X_2 \rangle = \mathbf{C}^+$. In the case of a standardized Gaussian random vector, we obtain $X_2 = -X_1$ when $\mathbf{C} \langle X_1, X_2 \rangle = \mathbf{C}^-$ and $X_2 = X_1$ when $\mathbf{C} \langle X_1, X_2 \rangle = \mathbf{C}^+$. If the marginals are log-normal, it follows that $X_2 = X_1^{-1}$ when $\mathbf{C} \langle X_1, X_2 \rangle = \mathbf{C}^-$ and $X_2 = X_1$ when $\mathbf{C} \langle X_1, X_2 \rangle = \mathbf{C}^+$. For these three examples, we verify that X_2 is a decreasing (resp. increasing) function of X_1 if

---

[11]We also have $X_2 = f(X_1)$ where $f = f_2 \circ f_1^{-1}$ is a decreasing function.

[12]In this case, $X_2 = f(X_1)$ where $f = f_2 \circ f_1^{-1}$ is an increasing function.

the copula function $\mathbf{C}\langle X_1, X_2 \rangle$ is $\mathbf{C}^-$ (resp. $\mathbf{C}^+$). The concepts of counter- and comonotonicity concepts generalize the cases where the linear correlation of a Gaussian vector is equal to -1 or $+1$. Indeed, $\mathbf{C}^-$ and $\mathbf{C}^+$ define respectively perfect negative and positive dependence.

Let (X_1, X_2) be a random vectors, whose copula is $\mathbf{C}\langle X_1, X_2 \rangle$. If h_1 and h_2 are two increasing functions on $\operatorname{Im} X_1$ and $\operatorname{Im} X_2$, then we have:

$$\mathbf{C}\langle h_1(X_1), h_2(X_2) \rangle = \mathbf{C}\langle X_1, X_2 \rangle$$

This means that copula functions are invariant under strictly increasing transformations of the random variables. To prove this theorem, we note $\mathbf{F}$ and $\mathbf{G}$ the probability distributions of the random vectors (X_1, X_2) and $(Y_1, Y_2) = (h_1(X_1), h_2(X_2))$. The margins of $\mathbf{G}$ are:

$$
\begin{aligned}
\mathbf{G}_1(y_1) &= \Pr\{Y_1 \leq y_1\} \\
&= \Pr\{h_1(X_1) \leq y_1\} \\
&= \Pr\{X_1 \leq h_1^{-1}(y_1)\} \quad \text{(because h_1 is strictly increasing)} \\
&= \mathbf{F}_1\left(h_1^{-1}(y_1)\right)
\end{aligned}
$$

and $\mathbf{G}_2(y_2) = \mathbf{F}_2\left(h_2^{-1}(y_2)\right)$. We deduce that $\mathbf{G}_1^{-1}(u_1) = h_1\left(\mathbf{F}_1^{-1}(u_1)\right)$ and $\mathbf{G}_2^{-1}(u_2) = h_2\left(\mathbf{F}_2^{-1}(u_2)\right)$. By definition, we have:

$$\mathbf{C}\langle Y_1, Y_2 \rangle (u_1, u_2) = \mathbf{G}\left(\mathbf{G}_1^{-1}(u_1), \mathbf{G}_2^{-1}(u_2)\right)$$

Moreover, it follows that:

$$
\begin{aligned}
\mathbf{G}\left(\mathbf{G}_1^{-1}(u_1), \mathbf{G}_2^{-1}(u_2)\right) &= \Pr\{Y_1 \leq \mathbf{G}_1^{-1}(u_1), Y_2 \leq \mathbf{G}_2^{-1}(u_2)\} \\
&= \Pr\{h_1(X_1) \leq \mathbf{G}_1^{-1}(u_1), h_2(X_2) \leq \mathbf{G}_2^{-1}(u_2)\} \\
&= \Pr\{X_1 \leq h_1^{-1}\left(\mathbf{G}_1^{-1}(u_1)\right), X_2 \leq h_2^{-1}\left(\mathbf{G}_2^{-1}(u_2)\right)\} \\
&= \Pr\{X_1 \leq \mathbf{F}_1^{-1}(u_1), X_2 \leq \mathbf{F}_2^{-1}(u_2)\} \\
&= \mathbf{F}\left(\mathbf{F}_1^{-1}(u_1), \mathbf{F}_2^{-1}(u_2)\right)
\end{aligned}
$$

Because we have $\mathbf{C}\langle X_1, X_2 \rangle (u_1, u_2) = \mathbf{F}\left(\mathbf{F}_1^{-1}(u_1), \mathbf{F}_2^{-1}(u_2)\right)$, we deduce that $\mathbf{C}\langle Y_1, Y_2 \rangle = \mathbf{C}\langle X_1, X_2 \rangle$.

We can interpret the copula function $\mathbf{C}\langle X_1, X_2 \rangle$ as a standardization of the joint distribution after eliminating the effects of margins. Indeed, it is a comprehensive statistic of the dependence function between X_1 and X_2. Therefore, a non-comprehensive statistic will be a dependence measure if it can be expressed using $\mathbf{C}\langle X_1, X_2 \rangle$. Following Nelsen (2006), a numeric measure m of association between X_1 and X_2 is a measure of concordance if it satisfies the following properties:

1. $-1 = m\langle X, -X \rangle \leq m\langle \mathbf{C} \rangle \leq m\langle X, X \rangle = 1$;

2. $m\langle \mathbf{C}^\perp \rangle = 0$;

3. $m\langle -X_1, X_2 \rangle = m\langle X_1, -X_2 \rangle = -m\langle X_1, X_2 \rangle$;

4. if $\mathbf{C}_1 \prec \mathbf{C}_2$, then $m\langle \mathbf{C}_1 \rangle \leq m\langle \mathbf{C}_2 \rangle$;

Using this last property, we have: $\mathbf{C} \prec \mathbf{C}^\perp \Rightarrow m\langle \mathbf{C} \rangle < 0$ and $\mathbf{C} \succ \mathbf{C}^\perp \Rightarrow m\langle \mathbf{C} \rangle > 0$. The concordance measure can then be viewed as a generalization of the linear correlation when the dependence function is not normal. Indeed, a positive quadrant dependence copula will have a positive concordance measure whereas a negative quadrant dependence copula will have a negative

concordance measure. Moreover, the bounds -1 and $+1$ are reached when the copula function is countermonotonic and comonotonic.

Among the several concordance measures, we find Kendall's tau and Spearman's rho, which play an important role in non-parametric statistics. Let us consider a sample of n observations $\{(x_1, y_1), \ldots, (x_n, y_n)\}$ of the random vector (X, Y). Kendall's tau is the probability of concordance — $(X_i - X_j) \cdot (Y_i - Y_j) > 0$ — minus the probability of discordance — $(X_i - X_j) \cdot (Y_i - Y_j) < 0$:

$$\tau = \Pr\left\{(X_i - X_j) \cdot (Y_i - Y_j) > 0\right\} - \Pr\left\{(X_i - X_j) \cdot (Y_i - Y_j) < 0\right\}$$

Spearman's rho is the linear correlation of the rank statistics $(X_{i:n}, Y_{i:n})$. We can also show that Spearman's rho has the following expression:

$$\varrho = \frac{\mathrm{cov}\left(\mathbf{F}_X\left(X\right), \mathbf{F}_Y\left(Y\right)\right)}{\sigma\left(\mathbf{F}_X\left(X\right)\right) \cdot \sigma\left(\mathbf{F}_Y\left(Y\right)\right)}$$

Schweizer and Wolff (1981) showed that Kendall's tau and Spearman's rho are concordance measures and have the following expressions:

$$\tau = 4 \iint_{[0,1]^2} \mathbf{C}\left(u_1, u_2\right) \, \mathrm{d}\mathbf{C}\left(u_1, u_2\right) - 1$$

$$\varrho = 12 \iint_{[0,1]^2} u_1 u_2 \, \mathrm{d}\mathbf{C}\left(u_1, u_2\right) - 3$$

From a numerical point of view, the following formulas should be preferred (Nelsen, 2006):

$$\tau = 1 - 4 \iint_{[0,1]^2} \partial_{u_1} \mathbf{C}\left(u_1, u_2\right) \, \partial_{u_2} \mathbf{C}\left(u_1, u_2\right) \, \mathrm{d}u_1 \, \mathrm{d}u_2$$

$$\varrho = 12 \iint_{[0,1]^2} \mathbf{C}\left(u_1, u_2\right) \, \mathrm{d}u_1 \, \mathrm{d}u_2 - 3$$

For some copulas, we have analytical formulas. For instance, we have:

Copula	ϱ	τ
Normal	$6\pi^{-1} \arcsin\left(\rho/2\right)$	$2\pi^{-1} \arcsin\left(\rho\right)$
Gumbel	$\checkmark$	$(\theta - 1)/\theta$
FGM	$\theta/3$	$2\theta/9$
Frank	$1 - 12\theta^{-1}\left(\mathbf{D}_1\left(\theta\right) - \mathbf{D}_2\left(\theta\right)\right)$	$1 - 4\theta^{-1}\left(1 - \mathbf{D}_1\left(\theta\right)\right)$

where $\mathbf{D}_k\left(x\right)$ is the Debye function. The Gumbel (or Gumbel-Hougaard) copula is equal to:

$$\mathbf{C}\left(u_1, u_2; \theta\right) = \exp\left(-\left[\left(-\ln u_1\right)^\theta + \left(-\ln u_2\right)^\theta\right]^{1/\theta}\right)$$

for $\theta \geq 1$, whereas the expression of the Farlie-Gumbel-Morgenstern (or FGM) copula is:

$$\mathbf{C}\left(u_1, u_2; \theta\right) = u_1 u_2 \left(1 + \theta\left(1 - u_1\right)\left(1 - u_2\right)\right)$$

for $-1 \leq \theta \leq 1$.

For illustration, we report in Figures A.4, A.5 and A.6 the level curves of several density functions built with Normal, Frank and Gumbel copulas. In order to compare them, the parameter of each copula is calibrated such that Kendall's tau is equal to 50%. This means that these 12 distributions functions have the same dependence with respect to Kendall's tau. However, the dependence is different from one figure to another, because their copula function is not the same. This is why Kendall's tau is not an exhaustive statistic of the dependence between two random variables.

Figure A.4: Contour lines of bivariate densities (Normal copula)
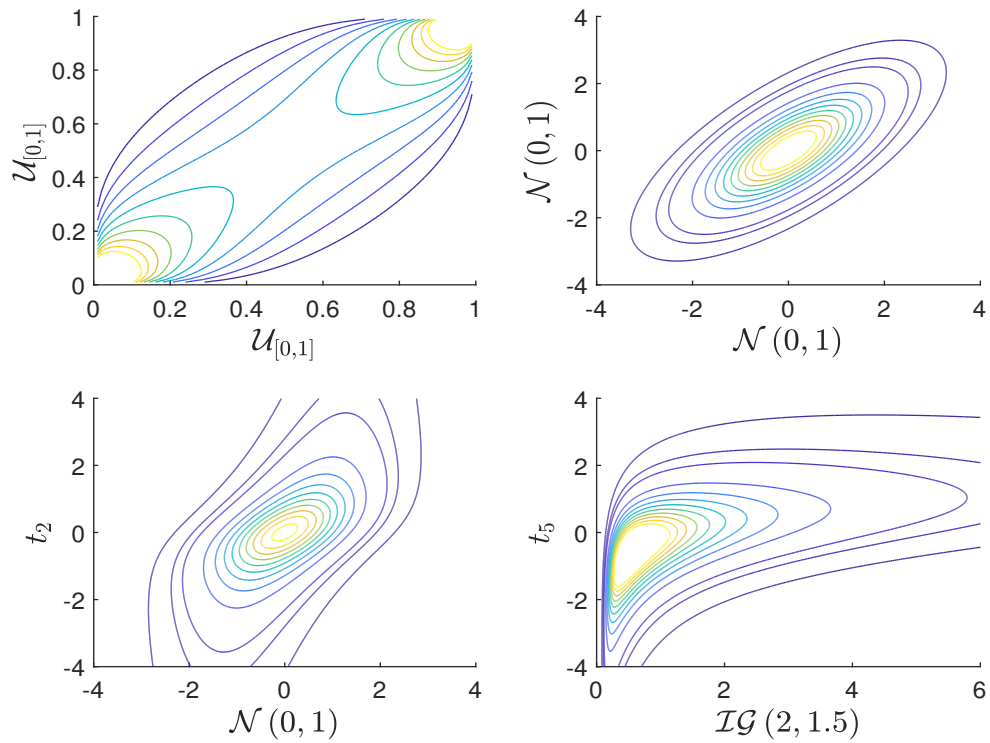
Figure A.5: Contour lines of bivariate densities (Frank copula)
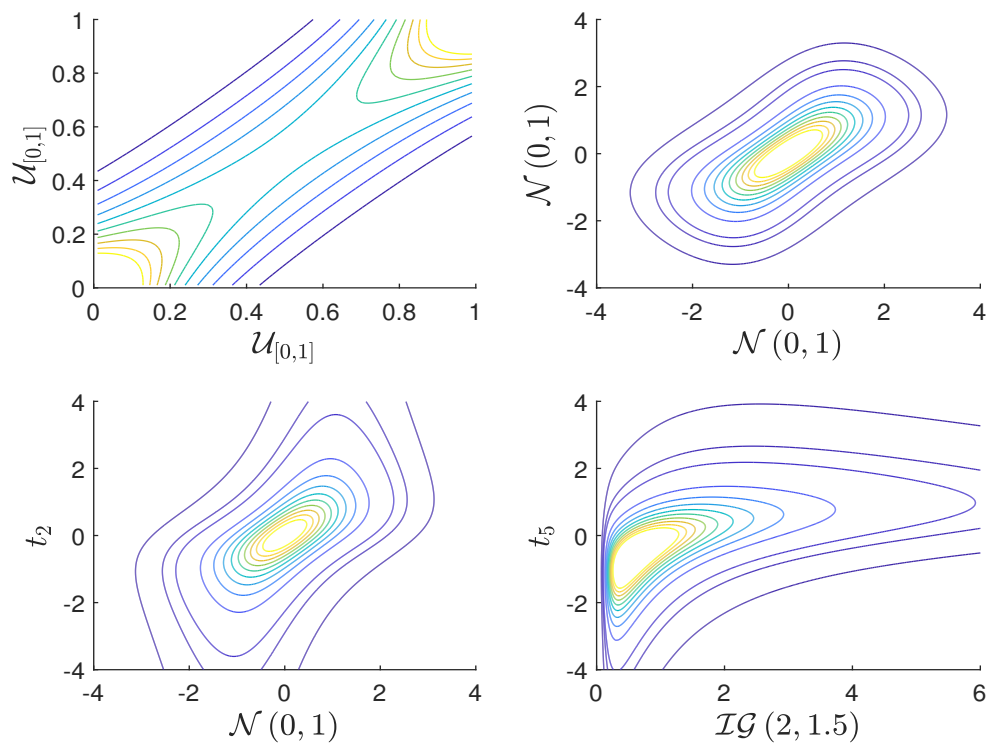
Figure A.6: Contour lines of bivariate densities (Gumbel copula)
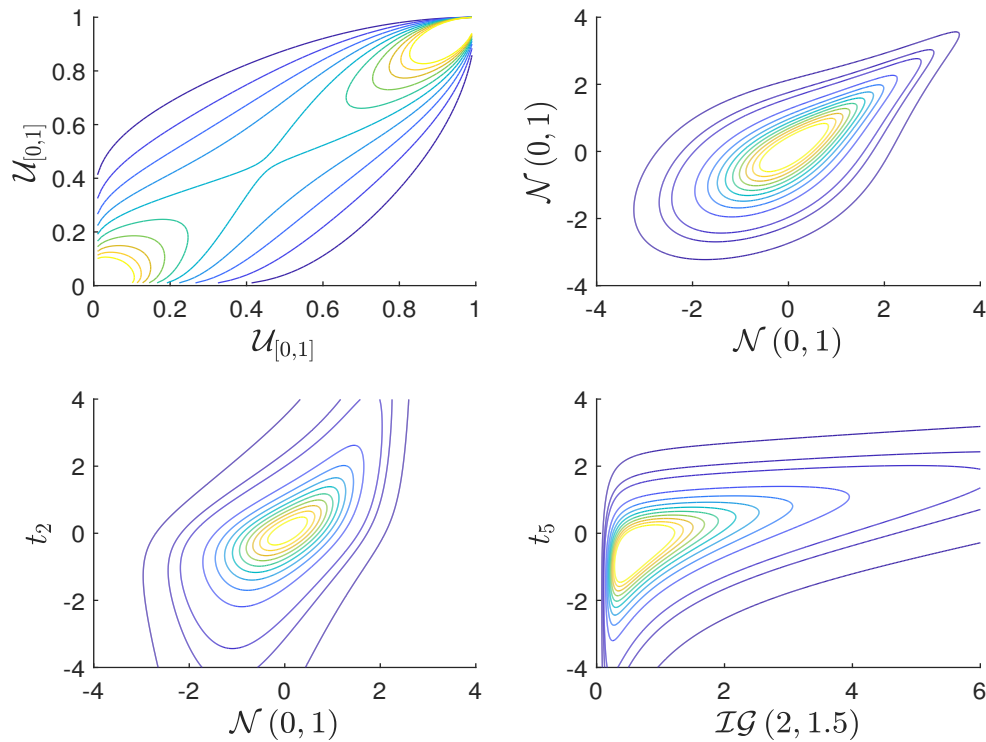
Remark 108 *We can show that the linear correlation (or Pearson's correlation) is not a concordance measure ([Roncalli, 2020a](), pages 727-729). In particular, the lower and upper bounds $\rho \langle \mathbf{C}^- \rangle$ and $\rho \langle \mathbf{C}^+ \rangle$ are not necessarily equal to -1 and $+1$. While the copula function is an invariant measure by increasing transformations, the Pearson's correlation is an invariant measure by increasing linear transformations. This is why the correlation is called a linear dependence measure.*

Parametric copula functions

Archimedean copulas   [Genest and MacKay]() ([1986b]()) define Archimedean copulas as follows:

$$\mathbf{C}(u_1, u_2) = \begin{cases} \varphi^{-1}(\varphi(u_1) + \varphi(u_2)) & \text{if } \varphi(u_1) + \varphi(u_2) \leq \varphi(0) \\ 0 & \text{otherwise} \end{cases}$$

where φ a C^2 is a function which satisfies $\varphi(1) = 0$, $\varphi'(u) < 0$ and $\varphi''(u) > 0$ for all $u \in [0, 1]$. $\varphi(u)$ is called the generator of the copula function. If $\varphi(0) = \infty$, the generator is said to be strict. [Genest and MacKay]() ([1986a]())MacKay, Jock link the construction of Archimedean copulas to the independence of random variables. Indeed, by considering the multiplicative generator $\lambda(u) = \exp(-\varphi(u))$, the authors show that:

$$\mathbf{C}(u_1, u_2) = \lambda^{-1}(\lambda(u_1)\lambda(u_2))$$

This means that:

$$\lambda(\Pr\{U_1 \leq u_1, U_2 \leq u_2\}) = \lambda(\Pr\{U_1 \leq u_1\}) \times \lambda(\Pr\{U_2 \leq u_2\})$$

In this case, the random variables (U_1, U_2) become independent when the scale of probabilities has been transformed.

The product copula $\mathbf{C}^{\perp}$ is Archimedean and the associated generator is $\varphi(u) = -\ln u$. Concerning Fréchet copulas, only $\mathbf{C}^{-}$ is Archimedean with $\varphi(u) = 1 - u$. In Table A.1, we provide another examples of Archimedean copulas[13].

Table A.1: Archimedean copula functions

Copula	$\varphi(u)$	$\mathbf{C}(u_1, u_2)$
Clayton	$u^{-\theta} - 1$	$\left(u_1^{-\theta} + u_2^{-\theta} - 1\right)^{-1/\theta}$
Frank	$-\ln \dfrac{e^{-\theta u} - 1}{e^{-\theta} - 1}$	$-\dfrac{1}{\theta}\ln\left(1 + \dfrac{\left(e^{-\theta u_1} - 1\right)\left(e^{-\theta u_2} - 1\right)}{e^{-\theta} - 1}\right)$
Gumbel	$(-\ln u)^{\theta}$	$\exp\left(-\left(\tilde{u}_1^{\theta} + \tilde{u}_2^{\theta}\right)^{1/\theta}\right)$
Joe	$-\ln\left(1 - (1-u)^{\theta}\right)$	$1 - \left(\bar{u}_1^{\theta} + \bar{u}_2^{\theta} - \bar{u}_1^{\theta}\bar{u}_2^{\theta}\right)^{1/\theta}$

Nelsen (2006) showed that if $\varphi(t)$ is a strict generator, then we can build two-parameter Archimedean copulas by considering the following generator:

$$\varphi_{\alpha,\beta}(t) = (\varphi(t^{\alpha}))^{\beta}$$

where $\alpha > 0$ and $\beta > 1$. For instance, if $\varphi(t) = t^{-1} - 1$, the two-parameter generator is $\varphi_{\alpha,\beta}(t) = (t^{-\alpha} - 1)^{\beta}$. Therefore, the corresponding copula function is defined by:

$$\mathbf{C}(u_1, u_2) = \left(\left[\left(u_1^{-\alpha} - 1\right)^{\beta} + \left(u_2^{-\alpha} - 1\right)^{\beta}\right]^{1/\beta} + 1\right)^{-1/\alpha}$$

This is a generalization of the Clayton copula, which is obtained when the parameter β is equal to 1.

We can build multivariate Archimedean copulas in the following way:

$$\mathbf{C}(u_1, \ldots, u_n) = \varphi^{-1}(\varphi(u_1) + \ldots + \varphi(u_n))$$

However, $\mathbf{C}$ is a copula function if and only if the function $\varphi^{-1}(u)$ is completely monotone (Nelsen, 2006):

$$(-1)^k \frac{\mathrm{d}^k}{\mathrm{d}u^k}\varphi^{-1}(u) \geq 0 \qquad \forall\, k \geq 1$$

For instance, the multivariate GumbelCopula function!Gumbel-Hougaard copula is defined by:

$$\mathbf{C}(u_1, \ldots, u_n) = \exp\left(-\left((-\ln u_1)^{\theta} + \ldots + (-\ln u_n)^{\theta}\right)^{1/\theta}\right)$$

Normal copula   The Normal copula is the dependence function of the multivariate normal distribution with a correlation matrix ρ:

$$\mathbf{C}(u_1, \ldots, u_n; \rho) = \Phi_n\left(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_n); \rho\right)$$

By using the canonical decomposition of the multivariate density function:

$$f(x_1, \ldots, x_n) = c(\mathbf{F}_1(x_1), \ldots, \mathbf{F}_n(x_n))\prod_{i=1}^{n} f_i(x_i)$$

---

[13]We use the notations $\bar{u} = 1 - u$ and $\tilde{u} = -\ln u$.

we deduce that the probability density function of the Normal copula is:

$$c\left(u_1, \ldots, u_n, ; \rho\right) = \frac{1}{|\rho|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}x^\top \left(\rho^{-1} - I_n\right) x\right)$$

where $x_i = \Phi^{-1}\left(u_i\right)$. In the bivariate case, we obtain[14]:

$$c\left(u_1, u_2; \rho\right) = \frac{1}{\sqrt{1 - \rho^2}} \exp\left(-\frac{x_1^2 + x_2^2 - 2\rho x_1 x_2}{2\left(1 - \rho^2\right)} + \frac{x_1^2 + x_2^2}{2}\right)$$

It follows that the expression of the bivariate Normal copula function is also equal to:

$$\mathbf{C}\left(u_1, u_2; \rho\right) = \int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} \phi_2\left(x_1, x_2; \rho\right) \, \mathrm{d}x_1 \, \mathrm{d}x_2$$

where $\phi_2\left(x_1, x_2; \rho\right)$ is the bivariate normal density:

$$\phi_2\left(x_1, x_2; \rho\right) = \frac{1}{2\pi\sqrt{1 - \rho^2}} \exp\left(-\frac{x_1^2 + x_2^2 - 2\rho x_1 x_2}{2\left(1 - \rho^2\right)}\right)$$

Another expression of the bivariate Normal copula density is:

$$\mathbf{C}\left(u_1, u_2; \rho\right) = \int_0^{u_1} \Phi\left(\frac{\Phi^{-1}\left(u_2\right) - \rho\Phi^{-1}\left(u\right)}{\sqrt{1 - \rho^2}}\right) \, \mathrm{d}u$$

Student's t copula   In a similar way, the Student's t copula is the dependence function associated with the multivariate Student's t probability distribution:

$$\mathbf{C}\left(u_1, \ldots, u_n; \rho, \nu\right) = \mathbf{T}_n\left(\mathbf{T}_\nu^{-1}\left(u_1\right), \ldots, \mathbf{T}_\nu^{-1}\left(u_n\right); \rho, \nu\right)$$

By using the definition of the cumulative distribution function:

$$\mathbf{T}_n\left(x_1, \ldots, x_n; \rho, \nu\right) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} \frac{\Gamma\left(\frac{\nu+n}{2}\right)|\rho|^{-\frac{1}{2}}}{\Gamma\left(\frac{\nu}{2}\right)(\nu\pi)^{\frac{n}{2}}} \left(1 + \frac{1}{\nu}x^\top\rho^{-1}x\right)^{-\frac{\nu+n}{2}} \mathrm{d}x$$

we can show that the copula density is then:

$$c\left(u_1, \ldots, u_n, ; \rho, \nu\right) = |\rho|^{-\frac{1}{2}} \frac{\Gamma\left(\frac{\nu+n}{2}\right)\left[\Gamma\left(\frac{\nu}{2}\right)\right]^n}{\left[\Gamma\left(\frac{\nu+1}{2}\right)\right]^n \Gamma\left(\frac{\nu}{2}\right)} \frac{\left(1 + \frac{1}{\nu}x^\top\rho^{-1}x\right)^{-\frac{\nu+n}{2}}}{\prod_{i=1}^n \left(1 + \frac{x_i^2}{\nu}\right)^{-\frac{\nu+1}{2}}}$$

where $x_i = \mathbf{T}_\nu^{-1}\left(u_i\right)$. In the bivariate case, we deduce that the t copula has the following expression:

$$\mathbf{C}\left(u_1, u_2; \rho, \nu\right) = \int_{-\infty}^{\mathbf{T}_\nu^{-1}(u_1)} \int_{-\infty}^{\mathbf{T}_\nu^{-1}(u_2)} \frac{1}{2\pi\sqrt{1 - \rho^2}} \left(1 + \frac{x_1^2 + x_2^2 - 2\rho x_1 x_2}{\nu\left(1 - \rho^2\right)}\right)^{-\frac{\nu+2}{2}} \mathrm{d}x_1 \, \mathrm{d}x_2$$

Like the Normal copula, we can obtain another expression, which is easier to manipulate. Let $\left(X_1, X_2\right)$ be a random vector whose probability distribution is $\mathbf{T}_2\left(x_1, x_2; \rho, \nu\right)$. Conditionally to $X_1 = x_1$, we have:

$$\left(\frac{\nu + 1}{\nu + x_1^2}\right)^{1/2} \frac{X_2 - \rho x_1}{\sqrt{1 - \rho^2}} \sim \mathbf{T}_{\nu+1}$$

---

[14]In the bivariate case, the parameter ρ is the cross-correlation between X_1 and X_2, that is the element $(1, 2)$ of the correlation matrix.

The conditional distribution $\mathbf{C}_{2|1}(u_1, u_2)$ is then equal to:

$$\mathbf{C}_{2|1}(u_1, u_2; \rho, \nu) = \mathbf{T}_{\nu+1}\left(\left(\frac{\nu+1}{\nu + \left[\mathbf{T}_\nu^{-1}(u_1)\right]^2}\right)^{1/2} \frac{\mathbf{T}_\nu^{-1}(u_2) - \rho\mathbf{T}_\nu^{-1}(u_1)}{\sqrt{1-\rho^2}}\right)$$

We deduce that:

$$\mathbf{C}(u_1, u_2; \rho, \nu) = \int_0^{u_1} \mathbf{C}_{2|1}(u, u_2; \rho, \nu)\, \mathrm{d}u$$

A.2.3 Estimation methods

Linear regression

Let Y and X be two random vectors. We consider the conditional expectation problem:

$$y = \mathbb{E}\left[Y \mid X = x\right] = m\left(x\right)$$

The underlying idea is to find an estimate $\hat{m}\left(x\right)$ of the function $m\left(x\right)$. In the general case, this problem is extremely difficult to solve. However, if (Y, X) is a Gaussian random vector, the function $m\left(x\right)$ can then be determined by considering the Gaussian linear model:

$$Y = \beta^\top X + u$$

where $u \sim \mathcal{N}\left(0, \sigma^2\right)$. Most of the time, the joint distribution of (Y, X) is unknown. In this case, the linear model is estimated by applying least squares techniques[15] to a given sample $(\mathbf{Y}, \mathbf{X})$:

$$\mathbf{Y} = \mathbf{X}\beta + \mathbf{U}$$

Derivation of the OLS estimator We consider a training set of n *iid* samples (y_i, x_i). For the ith observation, we have:

$$y_i = \sum_{k=1}^{K} \beta_k x_{i,k} + u_i \tag{A.14}$$

The least squares estimate of the parameter vector β is defined as follows:

$$\hat{\beta} = \arg\min \sum_{i=1}^{n} u_i^2$$

We introduce the following matrix notations: $\mathbf{Y}$ is the $n \times 1$ vector with elements $\mathbf{Y}_i = y_i$, $\mathbf{X}$ is the $n \times K$ matrix defined as follows:

$$\mathbf{X} = \begin{pmatrix} x_{1,1} & & x_{1,K} \\ & \ddots & \\ x_{n,1} & & x_{n,K} \end{pmatrix}$$

and $\mathbf{U}$ is the $n \times 1$ vector with elements $\mathbf{U}_i = u_i$. In this case, the system of equations (A.14) becomes:

$$\mathbf{Y} = \mathbf{X}\beta + \mathbf{U}$$

Let $\mathrm{RSS}\left(\beta\right)$ be the residual sum of squares. We have:

$$\begin{aligned} \mathrm{RSS}\left(\beta\right) &= \sum_{i=1}^{n} u_i^2 \\ &= \mathbf{U}^\top \mathbf{U} \\ &= \mathbf{Y}^\top \mathbf{Y} - 2\beta^\top \mathbf{X}^\top \mathbf{Y} + \beta^\top \mathbf{X}^\top \mathbf{X}\beta \end{aligned}$$

---

[15]In order to distinguish random variables and observations, we write matrices and vectors that are related to observations in bold style.

The least squares estimator verifies the set of normal equations $\partial_\beta \mathbf{U}^\top \mathbf{U} = \mathbf{0}$ and we deduce that $-2\mathbf{X}^\top \mathbf{Y} + 2\mathbf{X}^\top \mathbf{X}\hat{\beta} = \mathbf{0}$. The expression of the least squares estimator is then:

$$\hat{\beta} = \left(\mathbf{X}^\top \mathbf{X}\right)^{-1} \mathbf{X}^\top \mathbf{Y} \tag{A.15}$$

To obtain the expression of $\hat{\beta}$, we only need the assumption that the rank of the matrix $\mathbf{X}$ is K. In this case, $\hat{\beta}$ is the solution of the least squares problem. To go further, we assume that (Y, X) is a Gaussian random vector. The solution of the conditional expectation problem $\mathbb{E}[Y \mid X = x] = m(x)$ is then:

$$\hat{m}(x) = x^\top \hat{\beta} = x^\top \left(\mathbf{X}^\top \mathbf{X}\right)^{-1} \mathbf{X}^\top \mathbf{Y}$$

It means that the prediction of Y given that $X = x$ is equal to $\hat{y} = x^\top \hat{\beta}$. If we consider the training data $\mathbf{X}$, we obtain:

$$\hat{Y} = \hat{m}(\mathbf{X}) = \mathbf{X} \left(\mathbf{X}^\top \mathbf{X}\right)^{-1} \mathbf{X}^\top \mathbf{Y} = \mathbf{H}\mathbf{Y}$$

where $\mathbf{H} = \mathbf{X} \left(\mathbf{X}^\top \mathbf{X}\right)^{-1} \mathbf{X}^\top$ is called the '*hat*' matrix[16]. We notice that $\hat{m}(\mathbf{X})$ is a linear predictor of $\mathbf{Y}$.

Statistical inference   Because (Y, X) is a Gaussian random vector, it implies that $u = Y - \beta^\top X$ is a Gaussian random variable. We notice that:

$$\hat{\beta} = \left(\mathbf{X}^\top \mathbf{X}\right)^{-1} \mathbf{X}^\top \mathbf{Y} = \beta + \left(\mathbf{X}^\top \mathbf{X}\right)^{-1} \mathbf{X}^\top \mathbf{U}$$

By assuming the exogeneity of the variables X — meaning that $\mathbb{E}[u \mid X = x] = 0$ — we deduce that $\hat{\beta}$ is an unbiased estimator:

$$\mathbb{E}\left[\hat{\beta}\right] = \beta + \left(\mathbf{X}^\top \mathbf{X}\right)^{-1} \mathbb{E}\left[\mathbf{X}^\top \mathbf{U}\right] = \beta$$

We recall that $\mathbf{U} \sim \mathcal{N}\left(0, \sigma^2 I_n\right)$. It follows that:

$$\begin{aligned}
\mathrm{var}\left(\hat{\beta}\right) &= \mathbb{E}\left[\left(\hat{\beta} - \beta\right)\left(\hat{\beta} - \beta\right)^\top\right] \\
&= \mathbb{E}\left[\left(\mathbf{X}^\top \mathbf{X}\right)^{-1} \mathbf{X}^\top \mathbf{U}\mathbf{U}^\top \mathbf{X} \left(\mathbf{X}^\top \mathbf{X}\right)^{-1}\right] \\
&= \left(\mathbf{X}^\top \mathbf{X}\right)^{-1} \mathbf{X}^\top \mathbb{E}\left[\mathbf{U}\mathbf{U}^\top\right] \mathbf{X} \left(\mathbf{X}^\top \mathbf{X}\right)^{-1} \\
&= \left(\mathbf{X}^\top \mathbf{X}\right)^{-1} \mathbf{X}^\top \left(\sigma^2 I_n\right) \mathbf{X} \left(\mathbf{X}^\top \mathbf{X}\right)^{-1} \\
&= \sigma^2 \left(\mathbf{X}^\top \mathbf{X}\right)^{-1}
\end{aligned}$$

We conclude that:

$$\hat{\beta} \sim \mathcal{N}\left(\beta, \sigma^2 \left(\mathbf{X}^\top \mathbf{X}\right)^{-1}\right)$$

---

[16]We interpret $\mathbf{H}$ as the orthogonal projection matrix generated by $\mathbf{X}$ implying that $\mathbf{H}$ is idempotent, that is $\mathbf{H}\mathbf{H} = \mathbf{H}$. Indeed, we have:

$$\mathbf{H}\mathbf{H} = \mathbf{X} \left(\mathbf{X}^\top \mathbf{X}\right)^{-1} \mathbf{X}^\top \mathbf{X} \left(\mathbf{X}^\top \mathbf{X}\right)^{-1} \mathbf{X}^\top = \mathbf{X} \left(\mathbf{X}^\top \mathbf{X}\right)^{-1} \mathbf{X}^\top = \mathbf{H}$$

In most cases, σ^2 is unknown and we have to estimate it. The vector of residuals is:

$$\hat{\mathbf{U}} = \mathbf{Y} - \hat{\mathbf{Y}} = \mathbf{Y} - \mathbf{X}\hat{\beta}$$

We notice that $\mathbb{E}\left[\hat{\mathbf{U}}\right] = \mathbf{0}$ and $\text{var}\left(\hat{\mathbf{U}}\right) = \sigma^2 \left(I_n - \mathbf{H}\right)$. Because $\text{RSS}\left(\hat{\beta}\right) = \hat{\mathbf{U}}^\top \left(I_n - \mathbf{H}\right) \hat{\mathbf{U}}$ is a quadratic form, we can show that:

$$\hat{\sigma}^2 = \frac{\text{RSS}\left(\hat{\beta}\right)}{n - K}$$

is an unbiased estimator of σ^2 and $\hat{\sigma}^2/\sigma^2 \sim \chi^2_{n-K}$. In order to measure the model quality, we consider the coefficient of determination or R_c^2. It is defined as follows:

$$R_c^2 = 1 - \frac{\text{RSS}\left(\hat{\beta}\right)}{\text{TSS}}$$

where $\text{TSS} = \sum_{i=1}^{n} (y_i - \bar{y})$ is the total sum of squares. We have $R_c^2 \leq 1$. A high (resp. low) level indicates a good (resp. bad) goodness-of-fit of the regression model.

Lasso regression

The lasso method consists in adding a $\mathcal{L}_1$-norm penalty function to the optimization function in order to obtain a sparse parameter vector θ:

$$L_1(\theta) = \|\theta\|_1 = \sum_{k=1}^{K} |\theta_k|$$

For example, the lasso regression model is specified as follows (Tibshirani, 1996):

$$y_i = \sum_{k=1}^{K} \beta_k x_{i,k} + u_i \quad \text{s.t.} \quad \sum_{k=1}^{K} |\beta_k| \leq \tau$$

where τ is a scalar to control the sparsity. Using the notations introduced on page 677, we have:

$$\hat{\beta}(\tau) = \arg\min (\mathbf{Y} - \mathbf{X}\beta)^\top (\mathbf{Y} - \mathbf{X}\beta) \tag{A.16}$$
$$\text{s.t.} \quad \|\beta\|_1 \leq \tau$$

This problem is equivalent to the Lagrange optimization program $\hat{\beta}(\lambda) = \arg\min \mathcal{L}(\beta; \lambda)$ where[17]:

$$\mathcal{L}(\beta; \lambda) = \frac{1}{2}(\mathbf{Y} - \mathbf{X}\beta)^\top (\mathbf{Y} - \mathbf{X}\beta) + \lambda \|\beta\|_1$$
$$\propto \frac{1}{2}\beta^\top \left(\mathbf{X}^\top \mathbf{X}\right) \beta - \beta^\top \left(\mathbf{X}^\top \mathbf{Y}\right) + \lambda \|\beta\|_1$$

The solution $\hat{\beta}(\lambda)$ can be found by solving the augmented QP program where $\beta = \beta^+ - \beta^-$ under the constraints $\beta^+ \geq \mathbf{0}$ and $\beta^- \geq \mathbf{0}$. We deduce that:

$$\|\beta\|_1 = \sum_{k=1}^{K} |\beta_k^+ - \beta_k^-|$$
$$= \sum_{k=1}^{K} |\beta_k^+| + \sum_{k=1}^{K} |\beta_k^-|$$
$$= \mathbf{1}^\top \beta^+ + \mathbf{1}^\top \beta^-$$

---

[17]τ and λ are related by the relationship $\tau = \left\|\hat{\beta}(\lambda)\right\|_1$.

Since we have:

$$\beta = \left(\begin{array}{cc} I_K & -I_K \end{array}\right) \left(\begin{array}{c} \beta^+ \\ \beta^- \end{array}\right)$$

the augmented QP program is specified as follows:

$$\hat{\theta} = \arg\min \frac{1}{2}\theta^\top Q\theta - \theta^\top R$$
$$\text{s.t.} \quad \theta \geq \mathbf{0}$$

where $\theta = (\beta^+, \beta^-)$, $\tilde{\mathbf{X}} = \left(\begin{array}{cc} \mathbf{X} & -\mathbf{X} \end{array}\right)$, $Q = \tilde{\mathbf{X}}^\top \tilde{\mathbf{X}}$ and $R = \tilde{\mathbf{X}}^\top \mathbf{Y} - \lambda \cdot \mathbf{1}$. If we denote $A = \left(\begin{array}{cc} I_K & -I_K \end{array}\right)$, we obtain $\hat{\beta}(\lambda) = A\hat{\theta}$.

Remark 109 *If we consider Problem (A.16), we can also solve it using another augmented QP program:*

$$\hat{\theta} = \arg\min \frac{1}{2}\theta^\top Q\theta - \theta^\top R$$
$$\text{s.t.} \quad \left\{ \begin{array}{l} C\theta \geq D \\ \theta \geq \mathbf{0} \end{array} \right.$$

where $Q = \tilde{\mathbf{X}}^\top \tilde{\mathbf{X}}$, $R = \tilde{\mathbf{X}}^\top \mathbf{Y}$, $C = -\mathbf{1}^\top$ and $D = -\tau$. We again have $\hat{\beta}(\tau) = A\hat{\theta}$.

We have:

$$
\begin{aligned}
\text{RSS}(\beta) &= (\mathbf{Y} - \mathbf{X}\beta)^\top (\mathbf{Y} - \mathbf{X}\beta) \\
&= \left(\mathbf{Y} - \mathbf{X}\left(\hat{\beta}^{\text{ols}} + \beta - \hat{\beta}^{\text{ols}}\right)\right)^\top \left(\mathbf{Y} - \mathbf{X}\left(\hat{\beta}^{\text{ols}} + \beta - \hat{\beta}^{\text{ols}}\right)\right) \\
&= \left(\mathbf{Y} - \mathbf{X}\hat{\beta}^{\text{ols}}\right)^\top \left(\mathbf{Y} - \mathbf{X}\hat{\beta}^{\text{ols}}\right) + 2\left(\mathbf{Y} - \mathbf{X}\hat{\beta}^{\text{ols}}\right)^\top \mathbf{X}\left(\beta - \hat{\beta}^{\text{ols}}\right) + \\
&\quad \left(\beta - \hat{\beta}^{\text{ols}}\right)^\top \mathbf{X}^\top \mathbf{X}\left(\beta - \hat{\beta}^{\text{ols}}\right)
\end{aligned}
$$

We notice that:

$$
\begin{aligned}
(*) &= \left(\mathbf{Y} - \mathbf{X}\hat{\beta}^{\text{ols}}\right)^\top \mathbf{X}\left(\beta - \hat{\beta}^{\text{ols}}\right) \\
&= \left(\mathbf{Y}^\top - \left(\hat{\beta}^{\text{ols}}\right)^\top \mathbf{X}^\top\right) \mathbf{X}\left(\beta - \hat{\beta}^{\text{ols}}\right) \\
&= \left(\mathbf{Y}^\top - \left(\left(\mathbf{X}^\top \mathbf{X}\right)^{-1} \mathbf{X}^\top \mathbf{Y}\right)^\top \mathbf{X}^\top\right) \mathbf{X}\left(\beta - \hat{\beta}^{\text{ols}}\right) \\
&= \left(\mathbf{Y}^\top \mathbf{X} - \left(\left(\mathbf{X}^\top \mathbf{X}\right)^{-1} \mathbf{X}^\top \mathbf{Y}\right)^\top \mathbf{X}^\top \mathbf{X}\right)\left(\beta - \hat{\beta}^{\text{ols}}\right) \\
&= \left(\mathbf{Y}^\top \mathbf{X} - \mathbf{Y}^\top \mathbf{X}\right)\left(\beta - \hat{\beta}^{\text{ols}}\right) \\
&= 0
\end{aligned}
$$

Finally, we obtain:

$$\text{RSS}(\beta) = \text{RSS}\left(\hat{\beta}^{\text{ols}}\right) + \left(\beta - \hat{\beta}^{\text{ols}}\right)^\top \mathbf{X}^\top \mathbf{X}\left(\beta - \hat{\beta}^{\text{ols}}\right)$$

If we consider the equation $\text{RSS}(\beta) = c$, we distinguish three cases:

1. if $c < \text{RSS}\left(\hat{\beta}^{\text{ols}}\right)$, there is no solution;

2. if $c = \text{RSS}\left(\hat{\beta}^{\text{ols}}\right)$, there is one solution $\beta^{\star} = \hat{\beta}^{\text{ols}}$;

3. if $c > \text{RSS}\left(\hat{\beta}^{\text{ols}}\right)$, we have:

$$\left(\beta - \hat{\beta}^{\text{ols}}\right)^{\top} A \left(\beta - \hat{\beta}^{\text{ols}}\right) = 1$$

where:

$$A = \frac{\mathbf{X}^{\top}\mathbf{X}}{c - \text{RSS}\left(\hat{\beta}^{\text{ols}}\right)}$$

The solution $\beta^{\star}$ is an ellipsoid, whose center is $\hat{\beta}^{\text{ols}}$ and principal axes are the eigenvectors of the matrix A.

If we add the lasso constraint $\sum_{k=1}^{K} |\beta_k| \leq \tau$, the lasso estimator $\hat{\beta}(\tau)$ corresponds to the tangency between the diamond shaped region and the ellipsoid that corresponds to the possible maximum value of c. The diamond shape region due to the lasso constraint ensures that the lasso estimator is sparse:

$$\exists \eta > 0 : \forall \tau < \eta, \; \min\left(\hat{\beta}_1(\tau), \ldots, \hat{\beta}_K(\tau)\right) = 0$$

For example, the two-dimensional case is represented in Figure A.7. We notice that $\hat{\beta}_1(\tau)$ is equal to zero if $\tau < \eta$. This sparsity property is central for understanding the variable selection procedure.

Figure A.7: Interpretation of the lasso regression
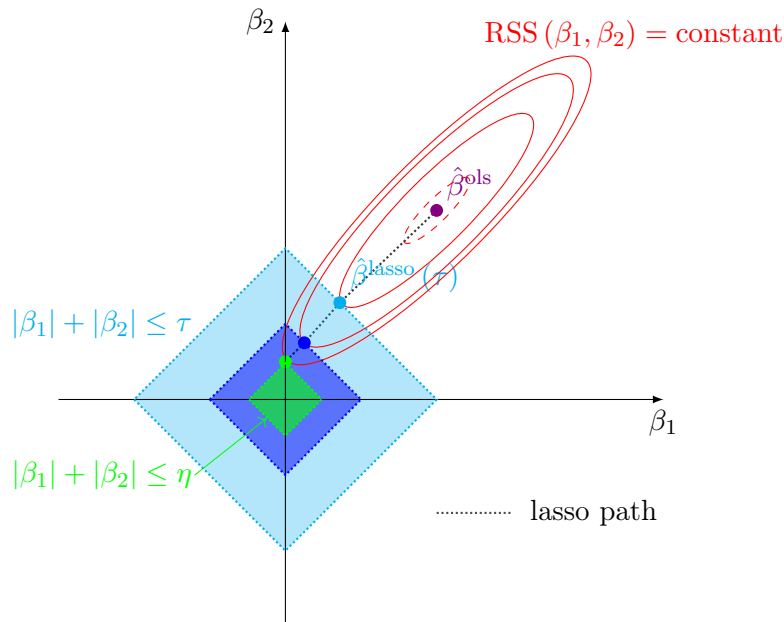
Example 55 *Using the data given in Table A.2, we consider the linear regression model:*

$$y_i = \beta_0' + \sum_{k=1}^{5} \beta_k' x_{i,k} + u_i \tag{A.17}$$

The objective is to determine the importance of each variable.

Table A.2: Data of the lasso regression problem

i	y	x_1	x_2	x_3	x_4	x_5
1	3.1	2.8	4.3	0.3	2.2	3.5
2	24.9	5.9	3.6	3.2	0.7	6.4
3	27.3	6.0	9.6	7.6	9.5	0.9
4	25.4	8.4	5.4	1.8	1.0	7.1
5	46.1	5.2	7.6	8.3	0.6	4.5
6	45.7	6.0	7.0	9.6	0.6	0.6
7	47.4	6.1	1.0	8.5	9.6	8.6
8	−1.8	1.2	9.6	2.7	4.8	5.8
9	20.8	3.2	5.0	4.2	2.7	3.6
10	6.8	0.5	9.2	6.9	9.3	0.7
11	12.9	7.9	9.1	1.0	5.9	5.4
12	37.0	1.8	1.3	9.2	6.1	8.3
13	14.7	7.4	5.6	0.9	5.6	3.9
14	−3.2	2.3	6.6	0.0	3.6	6.4
15	44.3	7.7	2.2	6.5	1.3	0.7

The lasso method can be used for ranking the variables. For that, we consider the following linear regression:

$$\tilde{y}_i = \sum_{k=1}^{5} \beta_k \tilde{x}_{i,k} + u_i$$

where $\tilde{y}_i$ and $\tilde{x}_{i,k}$ are the standardized data[18]:

$$\frac{y_i - \bar{y}}{s_y} = \sum_{k=1}^{5} \beta_k \left(\frac{x_{i,k} - \bar{x}_k}{s_{x_k}} \right) + u_i \tag{A.18}$$

Linear regressions (A.17) and (A.18) are related by the following equation:

$$y_i = \left(\bar{y} - \sum_{k=1}^{5} \frac{s_y \beta_k}{s_{x_k}} \bar{x}_k \right) + \sum_{k=1}^{5} \frac{s_y \beta_k}{s_{x_k}} x_{i,k} + s_y u_i$$

We deduce that $\beta_0' = \bar{y} - \sum_{k=1}^{5} (s_y/s_{x_k}) \beta_k \bar{x}_k$ and $\beta_k' = (s_y/s_{x_k}) \beta_k$. When performing lasso regression, we always standardize the data in order to obtain comparable beta's. Otherwise, the penalty function $\|\beta\|_1$ does not make a lot of sense. In Table A.3, we have estimated the lasso coefficients $\beta_k(\lambda)$ for different values of the shrinkage parameter λ. When $\lambda = 0$, we obtain the OLS estimate, and the lasso regression selects all the available variables. When $\lambda \to \infty$, the solution is $\hat{\beta}(\infty) = \mathbf{0}$, and the lasso regression selects no explanatory variables. In Table A.3, we verify that the number of selected variables is a decreasing function of λ. For instance, the lasso regression selects respectively four and three variables when λ is equal to 0.9 and 2.5. It follows that the most important variable is the third one, followed by the first, second, fourth and fifth variables.

---

[18]The notations $\bar{x}_k$ and s_{x_k} represent the mean and the standard deviation of the data $\{x_{i,k}, i = 1, \ldots, n\}$.

Table A.3: Results of the lasso regression

λ	0.0	0.9	2.5	5.5	7.5
$\hat{\beta}_1\left(\lambda\right)$	0.4586	0.4022	0.3163	0.1130	
$\hat{\beta}_2\left(\lambda\right)$	−0.1849	−0.2005	−0.1411		
$\hat{\beta}_3\left(\lambda\right)$	0.8336	0.7265	0.5953	0.3951	0.2462
$\hat{\beta}_4\left(\lambda\right)$	−0.1893	−0.1102			
$\hat{\beta}_5\left(\lambda\right)$	0.0931				
$\left\Vert\hat{\beta}\left(\lambda\right)\right\Vert_1$	1.7595	1.4395	1.0527	0.5081	0.2462
$\mathrm{RSS}\left(\hat{\beta}\left(\lambda\right)\right)$	0.0118	0.0304	0.1180	0.4076	0.6306
R_c^2	0.9874	0.9674	0.8735	0.5633	0.3244
$\mathrm{df}^{(\mathrm{model})}$	5	4	3	2	1

In Figure A.8, we have reported the path of the lasso estimate $\hat{\beta}\left(\lambda\right)$ with respect to the scaling factor $\tau^\star \in [0,1]$, which is defined as follows:

$$\tau^\star = \frac{\tau}{\tau_{\max}} = \frac{\left\Vert\hat{\beta}\left(\lambda\right)\right\Vert_1}{\left\Vert\hat{\beta}\left(0\right)\right\Vert_1}$$

$\tau^\star$ is equal to zero when $\lambda \to \infty$ (no selected variable) and one when $\lambda = 0$, which corresponds to the OLS case. From this path, we verify the lasso ordering $x_3 \succ x_1 \succ x_2 \succ x_4 \succ x_5$.

Figure A.8: Variable selection with the lasso regression
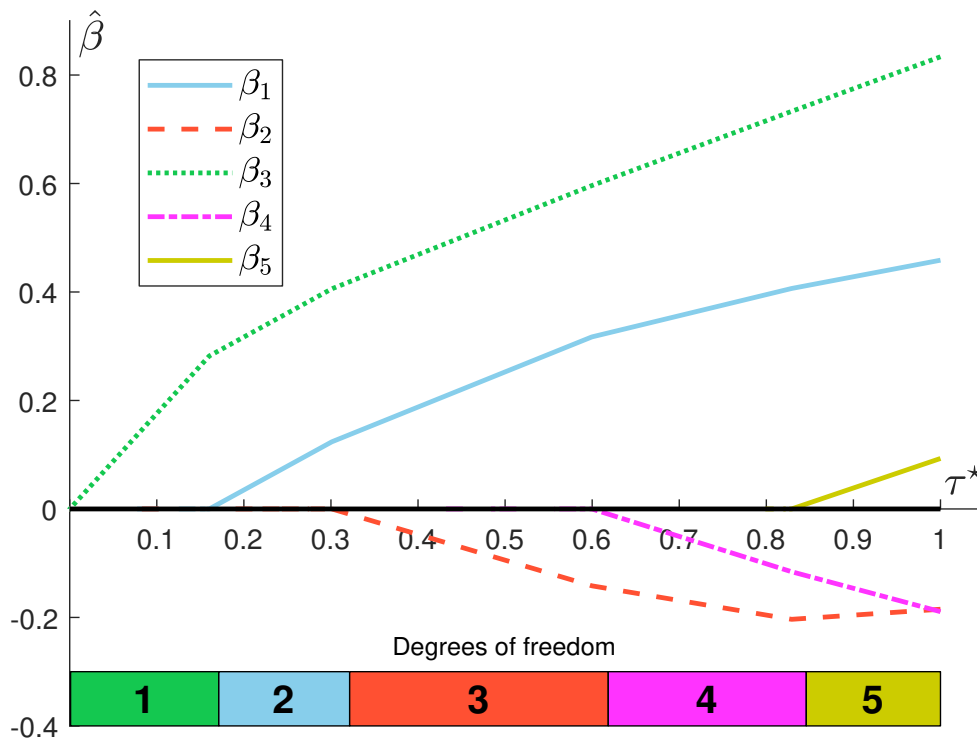
State space models

A state space model (SSM) includes a measurement equation and a transition equation. In the measurement equation, we define the relationship between an observable system and state variables, whereas the transition equation describes the dynamics of state variables. Generally, the state vector α_t is generated by a Markov linear process[19]:

$$\alpha_t = T_t \alpha_{t-1} + c_t + R_t \eta_t$$

where α_t is a $m \times 1$ vector, T_t is a $m \times m$ matrix, c_t is a $m \times 1$ vector and R_t is a $m \times p$ matrix. In the case of a linear SSM, the measurement equation is given by:

$$y_t = Z_t \alpha_t + d_t + \epsilon_t$$

where y_t is a n-dimensional time series, Z_t is a $n \times m$ matrix, d_t is a $n \times 1$ vector. We also assume that η_t and ϵ_t are two independent white noise processes of dimension p and n with covariance matrices Q_t and H_t.

Kalman filtering   In the state space model, the variable y_t is observable, but it is generally not the case of the state vector α_t. The Kalman filter is a statistical tool to estimate the distribution function of α_t. Let $\alpha_0 \sim \mathcal{N}(\hat{\alpha}_0, P_0)$ the initial position of the state vector. We note $\hat{\alpha}_{t|t}$ (or $\hat{\alpha}_t$) and $\hat{\alpha}_{t|t-1}$ the optimal estimators of α_t given the available information until time t and $t-1$:

$$
\begin{aligned}
\hat{\alpha}_{t|t} &= \mathbb{E}\left[\alpha_t \mid \mathcal{F}_t\right] \\
\hat{\alpha}_{t|t-1} &= \mathbb{E}\left[\alpha_t \mid \mathcal{F}_{t-1}\right]
\end{aligned}
$$

$P_{t|t}$ (or P_t) and $P_{t|t-1}$ are the covariance matrices associated to $\hat{\alpha}_{t|t}$ and $\hat{\alpha}_{t|t-1}$:

$$
\begin{aligned}
P_{t|t} &= \mathbb{E}\left[\left(\hat{\alpha}_{t|t} - \alpha_t\right)\left(\hat{\alpha}_{t|t} - \alpha_t\right)^{\top}\right] \\
P_{t|t-1} &= \mathbb{E}\left[\left(\hat{\alpha}_{t|t-1} - \alpha_t\right)\left(\hat{\alpha}_{t|t-1} - \alpha_t\right)^{\top}\right]
\end{aligned}
$$

These different quantities are calculated thanks to the Kalman filter, which consists in a recursive algorithm[20] (Harvey, 1990):

$$
\begin{cases}
\hat{\alpha}_{t|t-1} = T_t \hat{\alpha}_{t-1|t-1} + c_t \\
P_{t|t-1} = T_t P_{t-1|t-1} T_t^{\top} + R_t Q_t R_t^{\top} \\
\hat{y}_{t|t-1} = Z_t \hat{\alpha}_{t|t-1} + d_t \\
v_t = y_t - \hat{y}_{t|t-1} \\
F_t = Z_t P_{t|t-1} Z_t^{\top} + H_t \\
\hat{\alpha}_{t|t} = \hat{\alpha}_{t|t-1} + P_{t|t-1} Z_t^{\top} F_t^{-1} v_t \\
P_{t|t} = \left(I_m - P_{t|t-1} Z_t^{\top} F_t^{-1} Z_t\right) P_{t|t-1}
\end{cases}
$$

where $\hat{y}_{t|t-1} = \mathbb{E}\left[y_t \mid \mathcal{F}_{t-1}\right]$ is the best estimator of y_t given the available information until time $t-1$, v_t is the innovation process and F_t is the associated covariance matrix.

Remark 110 *Harvey (1990) showed that we can directly calculate $\hat{\alpha}_{t+1|t}$ from $\hat{\alpha}_{t|t-1}$:*

$$\hat{\alpha}_{t+1|t} = \left(T_{t+1} - K_t Z_t\right)\hat{\alpha}_{t|t-1} + K_t y_t + c_{t+1} - K_t d_t$$

---

[19]The presentation is based on the book of Harvey (1990).
[20]The algorithm is initialized with values $\hat{\alpha}_{0|0} = \hat{\alpha}_0$ and $P_{0|0} = P_0$.

where $K_t = T_{t+1} P_{t|t-1} Z_t^\top F_t^{-1}$ is the gain matrix. It follows that:

$$\hat{\alpha}_{t+1|t} = T_{t+1}\hat{\alpha}_{t|t-1} + c_{t+1} + K_t \left(y_t - Z_t\hat{\alpha}_{t|t-1} - d_t\right)$$

By recognizing the innovation process v_t, we obtain the following innovation representation:

$$\begin{cases} y_t = Z_t\hat{\alpha}_{t|t-1} + d_t + v_t \\ \hat{\alpha}_{t+1|t} = T_{t+1}\hat{\alpha}_{t|t-1} + c_{t+1} + K_t v_t \end{cases}$$

Estimation of unknown parameters   In many cases, the state space model depends on certain parameters that are unknown. Given a set θ of values for these unknown parameters, the Kalman filter may be applied to estimate the state vector α_t. We have:

$$v_t \sim \mathcal{N}\left(\mathbf{0}, F_t\right)$$

where $v_t = y_t - \hat{y}_{t|t-1}$ is the innovation at time t and $F_t = Z_t P_{t|t-1} Z_t^\top + H_t$ is the covariance matrix. If we change θ and we run the Kalman filter, we will obtain other values of v_t and F_t, meaning that v_t and F_t depend on θ. This is why we can write $v_t(\theta)$ and $F_t(\theta)$. We deduce that the likelihood function of the sample $\{y_1, \ldots, y_T\}$ is equal to:

$$\boldsymbol{\ell}(\theta) = -\frac{nT}{2}\ln\left(2\pi\right) - \frac{1}{2}\sum_{t=1}^{T}\left(\ln|F_t(\theta)| + v_t(\theta)^\top F_t(\theta)^{-1} v_t(\theta)\right)$$

We can then estimate the vector θ of unknown parameters by the method of maximum likelihood:

$$\hat{\theta} = \arg\max \boldsymbol{\ell}(\theta)$$

Once the ML estimate $\hat{\theta}$ is found, we can run again[21] the Kalman filter to estimate the other quantities $\hat{\alpha}_{t|t-1}$, $\hat{\alpha}_{t|t}$, $P_{t|t-1}$ and $P_{t|t}$.

Time-invariant state space model   We consider the time-invariant model:

$$\begin{cases} y_t = Z\alpha_t + d + \epsilon_t \\ \alpha_t = T\alpha_{t-1} + c + R\eta_t \end{cases}$$

where $\epsilon_t \sim \mathcal{N}\left(\mathbf{0}, H\right)$ and $\eta_t \sim \mathcal{N}\left(\mathbf{0}, Q\right)$. If the state space model converges to a steady state, the estimators $(\hat{\alpha}_\infty, P_\infty)$ must satisfy the following equations:

$$\begin{cases} \hat{\alpha}_\infty = T\hat{\alpha}_\infty + c \\ P_\infty = TP_\infty T^\top + RQR^\top \end{cases}$$

It follows that the solution is:

$$\begin{cases} \hat{\alpha}_\infty = (I_m - T)^{-1} c \\ \mathrm{vec}(P_\infty) = (I_{m^2} - T \otimes T)^{-1} \mathrm{vec}\left(RQR^\top\right) \end{cases}$$

where $\hat{\alpha}_\infty$ and P_∞ are the unconditional mean and covariance matrix of α_t. Without any knowledge of the initial position α_0, the best way to define $\hat{\alpha}_0$ and P_0 is then to use the steady state:

$$\begin{cases} \hat{\alpha}_0 = \hat{\alpha}_\infty \\ P_0 = P_\infty \end{cases}$$

---

[21]We say again, because computing the log-likelihood function requires one Kalman filter run, implying that many Kalman filter runs are used for maximizing the log-likelihood function.

In many state space models, the matrices T, c, R and Q depend on unknown parameters θ, implying that $\hat{\alpha}_\infty$ and P_∞ also depend on θ. This means that when maximizing the log-likelihood function, the Kalman filter is initialized by values of $\hat{\alpha}_0$ and P_0 that depend on θ. This is the main difference with time-varying state space model since the Kalman filter is initialized by fixed values of $\hat{\alpha}_0$ and P_0.

A.3   Stochastic analysis

A.3.1   Stochastic optimal control

A.3.2   Jump-diffusion processes

A.4   Spatial data

A.4.1   Spherical coordinates

A.4.2   Geographic coordinate systems

A.4.3   Network common data form

Appendix B

Solutions to the Tutorial Exercices

B.1 Exercises related to ESG risk

B.1.1 Score normalization when the features are independent (Exercise 2.4.1)

B.1.2 Score normalization when the features are correlated (Exercise 2.4.2)

B.1.3 Construction of a sovereign ESG score (Exercise 2.4.3)

B.1.4 Probability distribution of ESG scores (Exercise 2.4.4)

1. (a) We have:
$$\boldsymbol{\mathcal{S}}_i^{(\text{ESG})} = 0.4 \times \boldsymbol{\mathcal{S}}_i^{(\text{E})} + 0.4 \times \boldsymbol{\mathcal{S}}_i^{(\text{S})} + 0.2 \times \boldsymbol{\mathcal{S}}_i^{(\text{G})}$$

We deduce the following results:

Issuer	#1	#2	#3	#4	#5	#6	#7	#8
$\boldsymbol{\mathcal{S}}_i^{(\text{E})}$	-2.80	-1.80	-1.75	0.60	0.75	1.30	1.90	2.70
$\boldsymbol{\mathcal{S}}_i^{(\text{S})}$	-1.70	-1.90	0.75	-1.60	1.85	1.05	0.90	0.70
$\boldsymbol{\mathcal{S}}_i^{(\text{G})}$	0.30	-0.70	-2.75	2.60	0.45	2.35	2.20	1.70
$\boldsymbol{\mathcal{S}}_i^{(\text{ESG})}$	-1.74	-1.62	-0.95	0.12	1.13	1.41	1.56	1.70

(b) We obtain:
$$\boldsymbol{\mathcal{S}}^{(\text{ESG})}(x_{\text{ew}}) = \sum_{i=1}^{8} x_{\text{ew},i} \times \boldsymbol{\mathcal{S}}_i^{(\text{ESG})} = 0.2013$$

2. (a) We have:
$$\boldsymbol{\mathcal{S}}\left(x_{\text{ew}}^{(n)}\right) = \sum_{i=1}^{n} x_{\text{ew},i}^{(n)} \times \boldsymbol{\mathcal{S}}_i = \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{\mathcal{S}}_i$$

We deduce that $\boldsymbol{\mathcal{S}}\left(x_{\text{ew}}^{(n)}\right)$ follows a Gaussian distribution. Its mean is equal to:
$$\mathbb{E}\left[\boldsymbol{\mathcal{S}}\left(x_{\text{ew}}^{(n)}\right)\right] = \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}\left[\boldsymbol{\mathcal{S}}_i\right] = 0$$

whereas its standard deviation is equal to:
$$\sigma\left(\boldsymbol{\mathcal{S}}\left(x_{\text{ew}}^{(n)}\right)\right) = \sqrt{\frac{1}{n^2} \sum_{i=1}^{n} \sigma^2\left(\boldsymbol{\mathcal{S}}_i\right)} = \frac{1}{\sqrt{n}}$$

Finally, we deduce that:

$$\mathcal{S}\left(x_{\text{ew}}^{(n)}\right) \sim \mathcal{N}\left(0, \frac{1}{n}\right)$$

(b) The behavior of a well-diversified portfolio is close to an equally-weighted portfolio with n sufficiently large. Therefore, the ESG score is close to zero because we have:

$$\lim_{n\to\infty} \mathcal{S}\left(x_{\text{ew}}^{(n)}\right) = 0$$

(c) We have:

$$f_\alpha(t) = \alpha t^{\alpha-1}$$

The probability density function $f_\alpha(t)$ is reported in Figure B.1. We notice that the function $f_\alpha(t)$ tends to the dirac delta function when α tends to infinity:

$$\lim_{\alpha\to\infty} f_\alpha(t) = \delta_1(t) = \begin{cases} 0 & \text{if } t \neq 1 \\ +\infty & \text{if } t = 1 \end{cases}$$

Figure B.1: Probability density function $f_\alpha(t)$
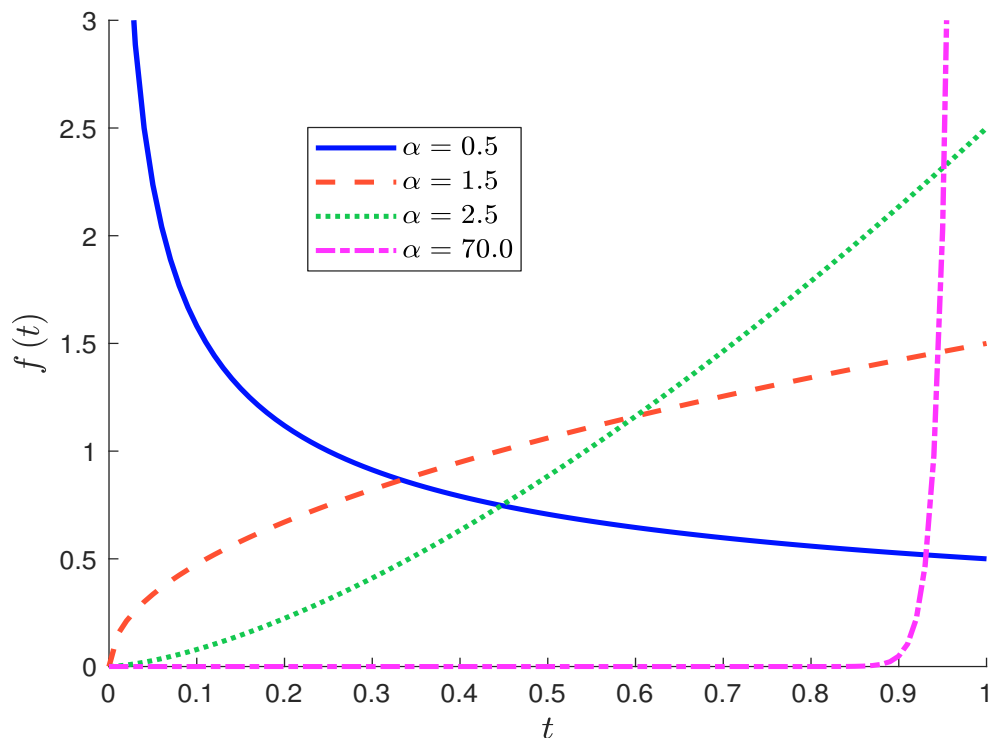
(d) To simulate T_i, we use the property of the probability integral transform: $U_i = \mathbf{F}_\alpha(T_i) \sim \mathcal{U}_{[0,1]}$. We deduce that:

$$T_i = \mathbf{F}_\alpha^{-1}(U_i) = U_i^{1/\alpha}$$

The algorithm for simulating the portfolio x is then the following:

   i. We simulate n independent uniform random numbers $(u_1, \ldots, u_n)$;

ii. We compute the random variates $(t_1, \ldots, t_n)$ where:

$$t_i = u_i^{1/\alpha}$$

iii. We calculate the normalization constant:

$$c = \left(\sum_{i=1}^{n} t_i \right)^{-1} = \left(\sum_{i=1}^{n} u_i^{1/\alpha} \right)^{-1}$$

iv. We deduce the portfolio weights $x = (x_1, \ldots, x_n)$:

$$x_i = c \cdot t_i = c \cdot u_i^{1/\alpha} = \frac{u_i^{1/\alpha}}{\sum_{j=1}^{n} u_j^{1/\alpha}}$$

In Figure B.2, we have represented the composition of the portfolio x for the 4 values of α. The weights are ranked in descending order. We deduce that the portfolio x is uniform when $\alpha \to \infty$. The parameter α controls the concentration of the portfolio. Indeed, when α is small, the portfolio is highly concentrated. It follows that the Herfindahl index $\mathcal{H}_\alpha(x)$ of the portfolio weights is a decreasing function of the parameter α.

Figure B.2: Repartition of the portfolio weights in descending order
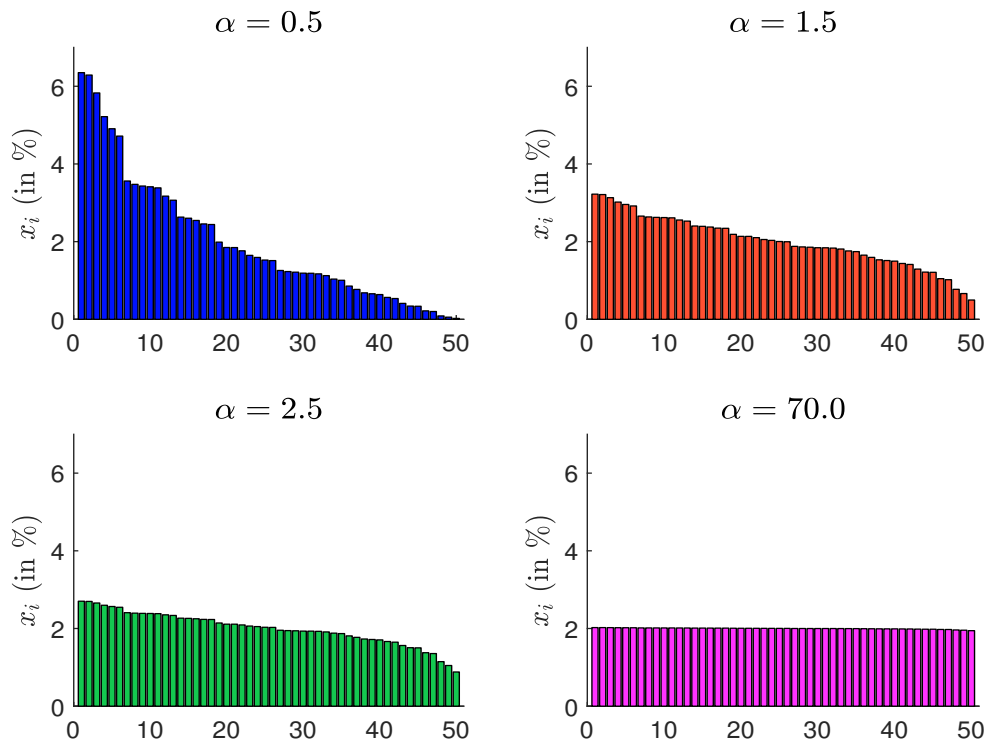
(e) We simulate $x = (x_1, \ldots, x_n)$ using the previous algorithm. The vector of ESG scores $\boldsymbol{\mathcal{S}} = (\boldsymbol{\mathcal{S}}_1, \ldots, \boldsymbol{\mathcal{S}}_n)$ is generated with normally-distributed random variables since we have $\boldsymbol{\mathcal{S}}_i \sim \mathcal{N}(0, 1)$. We deduce that the simulated value of the portfolio ESG score $\boldsymbol{\mathcal{S}}(x)$ is equal to:

$$\boldsymbol{\mathcal{S}}(x) = \sum_{i=1}^{n} x_i \cdot \boldsymbol{\mathcal{S}}_i$$

We replicate the simulation of $\boldsymbol{\mathcal{S}}(x)$ 50 000 times and draw the corresponding histogram in Figure B.3. We also report the fitted Gaussian distribution. We observe that the portfolio's ESG score $\boldsymbol{\mathcal{S}}(x)$ is equal to zero on average, and its variance is an increasing function of the portfolio concentration.

Figure B.3: Histogram of the portfolio ESG score $\boldsymbol{\mathcal{S}}(x)$
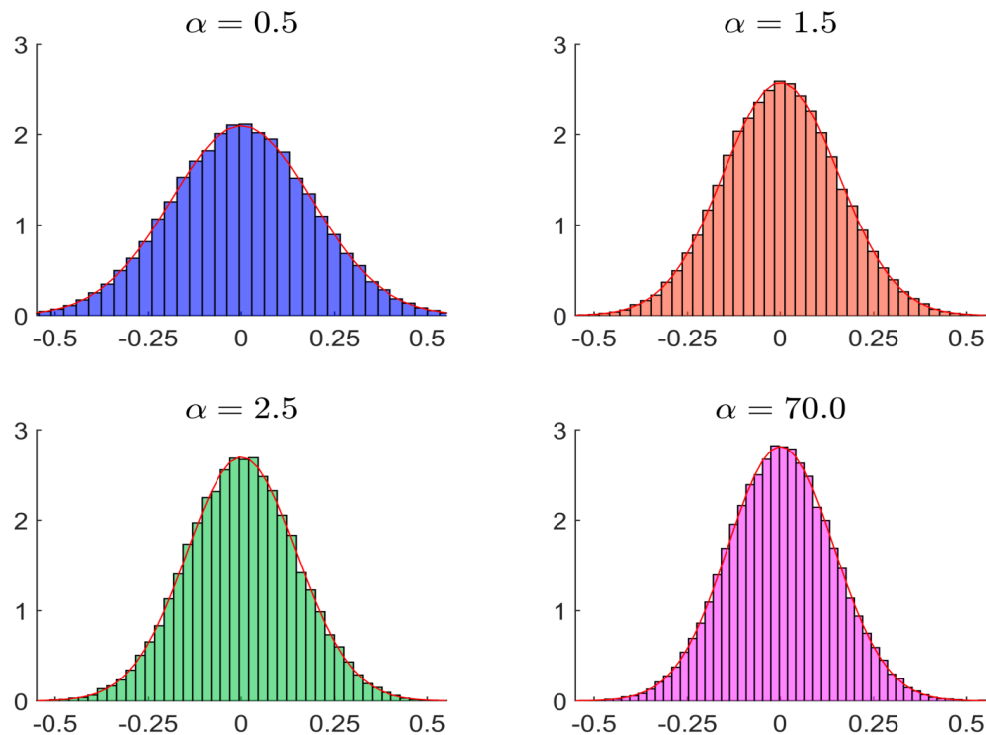
(f) Since $x_i \sim cT_i$, x_i is an increasing function of T_i. We deduce that the copula function of $(T_i, \boldsymbol{\mathcal{S}}_i)$ is the same as the copula function of $(x_i, \boldsymbol{\mathcal{S}}_i)$. To simulate the Normal copula function $\mathbf{C}(u, v)$, we use the transformation algorithm based on the Cholesky decomposition:

$$\begin{cases} u_i = \Phi(g_i') \\ v_i = \Phi\left(\rho g_i' + \sqrt{1-\rho^2} g_i''\right) \end{cases}$$

where g_i' and g_i'' are two independent random numbers from the probability distribution $\mathcal{N}(0,1)$. Here is the algorithm to simulate the portfolios's ESG score $\boldsymbol{\mathcal{S}}(x)$:

    i. We simulate n independent normally-distributed random numbers g_i' and g_i'' and compute (u_i, v_i):

$$\begin{cases} u_i = \Phi(g_i') \\ v_i = \Phi\left(\rho g_i' + \sqrt{1-\rho^2} g_i''\right) \end{cases}$$

    ii. We compute the random variates $(t_1, \ldots, t_n)$ where $t_i = u_i^{1/\alpha}$;

    iii. We deduce the vector of weights $x = (x_1, \ldots, x_n)$:

$$x_i = t_i \left/ \sum_{j=1}^{n} t_j \right.$$

iv. We simulate the vector of scores $\boldsymbol{\mathcal{S}} = (\boldsymbol{\mathcal{S}}_1, \ldots, \boldsymbol{\mathcal{S}}_n)$:

$$\boldsymbol{\mathcal{S}}_i = \Phi^{-1}(v_i) = \rho g'_i + \sqrt{1 - \rho^2} g''_i$$

v. We calculate the portfolio score:

$$\boldsymbol{\mathcal{S}}(x) = \sum_{i=1}^n x_i \cdot \boldsymbol{\mathcal{S}}_i$$

Figure B.4: Histogram of the portfolio ESG score $\boldsymbol{\mathcal{S}}(x)$ ($\rho = 50\%$)
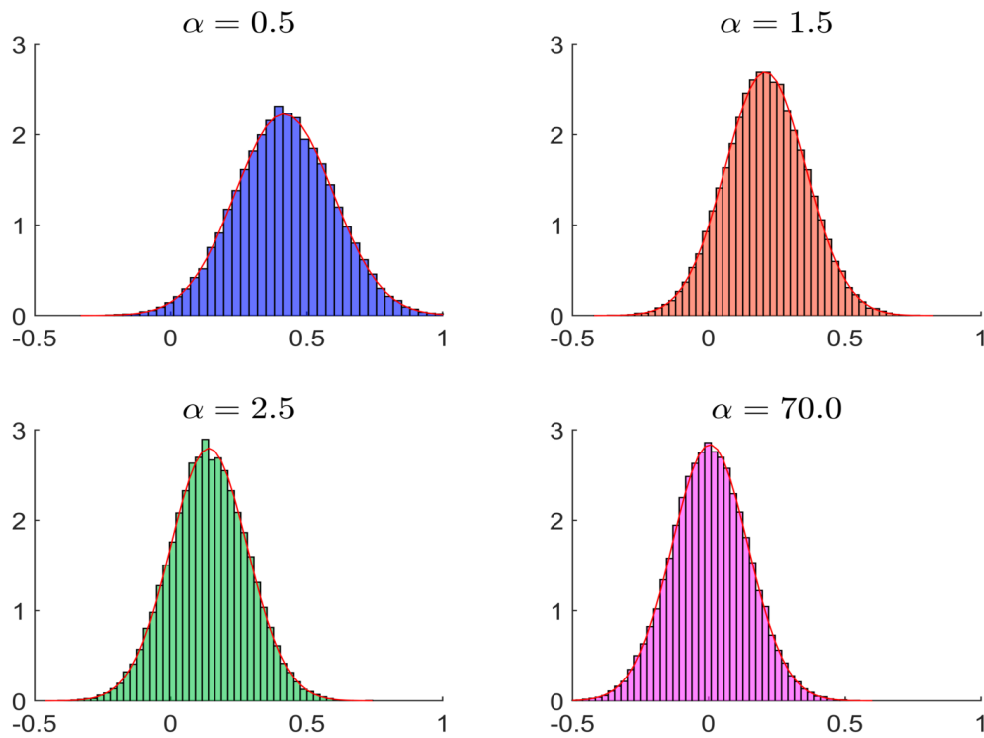
(g) In the independent case, we found that $\mathbb{E}[\boldsymbol{\mathcal{S}}(x)] = 0$. In Figure B.4, we notice that $\mathbb{E}[\boldsymbol{\mathcal{S}}(x)] \neq 0$ when ρ is equal to 50%. Indeed, we obtain:

$$\mathbb{E}[\boldsymbol{\mathcal{S}}(x)] = \begin{cases} 0.418 & \text{if } \alpha = 0.5 \\ 0.210 & \text{if } \alpha = 1.5 \\ 0.142 & \text{if } \alpha = 2.5 \\ 0.006 & \text{if } \alpha = 70.0 \end{cases}$$

(h) In Figure B.5, we notice that there is a positive relationship between ρ and $\mathbb{E}[\boldsymbol{\mathcal{S}}(x)]$ and the slope increases with the concentration of the portfolio.

(i) Big cap companies have more (financial and human) resources to develop an ESG policy than small cap companies. Therefore, we observe a positive correlation between the market capitalization and the ESG score of an issuer. It follows that ESG portfolios have generally a size bias. For instance, we generally observe that cap-weighted indexes have an ESG score which is greater than the average of ESG scores. In the previous questions, we verify that $\mathbb{E}[\boldsymbol{\mathcal{S}}(x)] \geq \mathbb{E}[\boldsymbol{\mathcal{S}}]$ when the Herfindahl index of the portfolio x is high and the correlation between x_i and $\boldsymbol{\mathcal{S}}_i$ is positive.

Figure B.5: Relationship between ρ and $\mathbb{E}\left[\boldsymbol{S}\left(x\right)\right]$
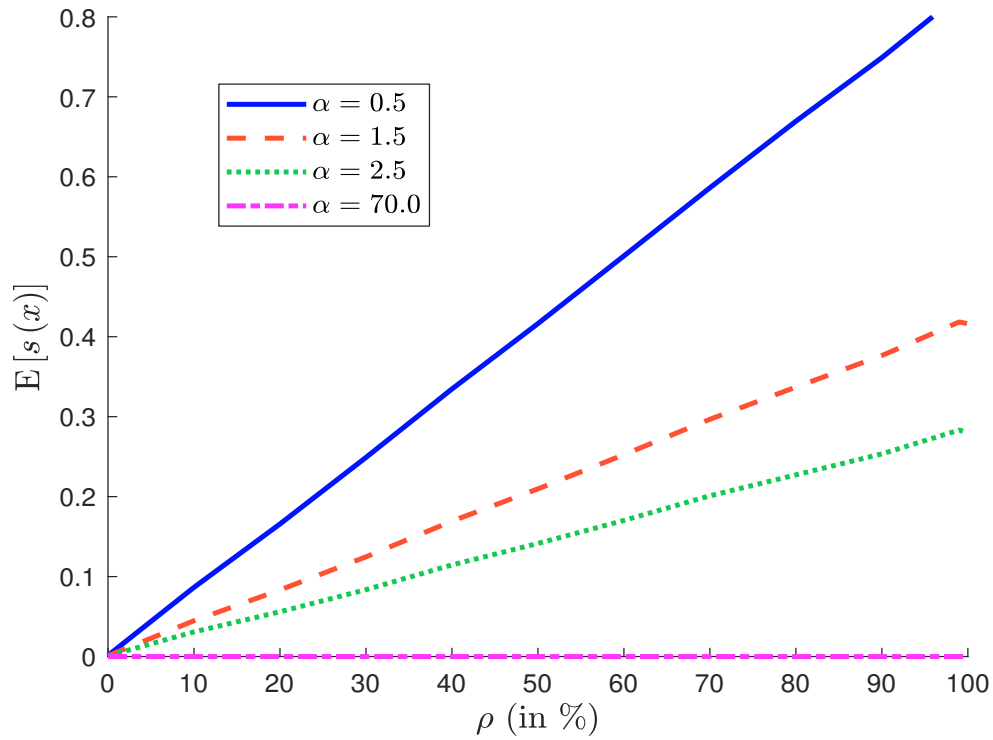
3. (a) We have:

$$
\begin{aligned}
\Pr\left\{\mathcal{R}=\mathbf{A}\right\} &= \Pr\left\{\boldsymbol{S}\geq 1.5\right\} \\
&= 1-\Phi\left(1.5\right) \\
&= 6.68\%
\end{aligned}
$$

and:

$$
\begin{aligned}
\Pr\left\{\mathcal{R}=\mathbf{B}\right\} &= \Pr\left\{0\leq\boldsymbol{S}<1.5\right\} \\
&= \Phi\left(1.5\right)-\Phi\left(0\right) \\
&= 43.32\%
\end{aligned}
$$

Since the Gaussian distribution is symmetric around 0, we also have:

$$
\Pr\left\{\mathcal{R}=\mathbf{C}\right\}=\Pr\left\{\mathcal{R}=\mathbf{B}\right\}=43.32\%
$$

and:

$$
\Pr\left\{\mathcal{R}=\mathbf{D}\right\}=\Pr\left\{\mathcal{R}=\mathbf{A}\right\}=6.68\%
$$

The mapping function is then equal to:

$$
\mathcal{M}_{\text{appring}}\left(\boldsymbol{S}\right)=\begin{cases}
\mathbf{A} & \text{if } \boldsymbol{S}<-1.5 \\
\mathbf{B} & \text{if } -1.5\leq\boldsymbol{S}<0 \\
\mathbf{C} & \text{if } 0\leq\boldsymbol{S}<1.5 \\
\mathbf{D} & \text{if } \boldsymbol{S}\geq 1.5
\end{cases}
$$

(b) Since we have:

$$\Pr\{\mathcal{R}(t) = \mathbf{A}\} = \Pr\{\mathcal{R}(t) = \mathbf{B}\} = \Pr\{\mathcal{R}(t) = \mathbf{C}\} = \Pr\{\mathcal{R}(t) = \mathbf{D}\}$$

and:

$$\Pr\{\mathcal{R}(t) = \mathbf{A}\} + \Pr\{\mathcal{R}(t) = \mathbf{B}\} + \Pr\{\mathcal{R}(t) = \mathbf{C}\} + \Pr\{\mathcal{R}(t) = \mathbf{D}\} = 1$$

we deduce that:

$$\Pr\{\mathcal{R}(t) = \mathbf{A}\} = \frac{1}{4} = 25\%$$

and $\Pr\{\mathcal{R}(t) = \mathbf{B}\} = \Pr\{\mathcal{R}(t) = \mathbf{C}\} = \Pr\{\mathcal{R}(t) = \mathbf{D}\} = 25\%$. We want to find the breakpoints (s_1, s_2, s_3) such that:

$$\begin{cases} \Pr\{\mathcal{S} < s_1\} = 25\% \\ \Pr\{s_1 \leq \mathcal{S} < s_2\} = 25\% \\ \Pr\{s_2 \leq \mathcal{S} < s_3\} = 25\% \\ \Pr\{\mathcal{S} \geq s_3\} = 25\% \end{cases}$$

We deduce that:

$$\begin{cases} s_1 = \Phi^{-1}(0.25) = -0.6745 \\ s_2 = \Phi^{-1}(0.50) = 0 \\ s_3 = \Phi^{-1}(0.75) = +0.6745 \end{cases}$$

The mapping function is then given by:

$$\mathcal{M}_{\text{appring}}(\mathcal{S}) = \begin{cases} \mathbf{A} & \text{if } \mathcal{S} < -0.6745 \\ \mathbf{B} & \text{if } -0.6745 \leq \mathcal{S} < 0 \\ \mathbf{C} & \text{if } 0 \leq \mathcal{S} < 0.6745 \\ \mathbf{D} & \text{if } \mathcal{S} \geq 0.6745 \end{cases}$$

(c) We have:

$$\begin{cases} s_1 = \Phi^{-1}(0.10) = -1.2816 \\ s_2 = \Phi^{-1}(0.50) = 0 \\ s_3 = \Phi^{-1}(0.90) = +1.2816 \end{cases}$$

We deduce that the mapping function is equal to:

$$\mathcal{M}_{\text{appring}}(\mathcal{S}) = \begin{cases} \mathbf{A} & \text{if } \mathcal{S} < -1.2816 \\ \mathbf{B} & \text{if } -1.2816 \leq \mathcal{S} < 0 \\ \mathbf{C} & \text{if } 0 \leq \mathcal{S} < 1.2816 \\ \mathbf{D} & \text{if } \mathcal{S} \geq 1.2816 \end{cases}$$

4. (a) The joint distribution of $(\mathcal{S}(t-1), \Delta\mathcal{S}(t))$ is:

$$\begin{pmatrix} \mathcal{S}(t-1) \\ \Delta\mathcal{S}(t) \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & \sigma^2 \end{pmatrix} \right)$$

(b) Since we have:

$$\mathcal{S}(t) = \mathcal{S}(t-1) + \Delta\mathcal{S}(t)$$

we deduce that:

$$\begin{pmatrix} \mathcal{S}(t-1) \\ \mathcal{S}(t) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \mathcal{S}(t-1) \\ \Delta\mathcal{S}(t) \end{pmatrix}$$

We conclude that $(\boldsymbol{S}(t-1), \boldsymbol{S}(t))$ is a Gaussian random vector. We have:

$$\text{var}(\boldsymbol{S}(t)) = 1 + \sigma^2$$

and:

$$
\begin{aligned}
\text{cov}(\boldsymbol{S}(t-1), \boldsymbol{S}(t)) &= \mathbb{E}[\boldsymbol{S}(t-1) \cdot \boldsymbol{S}(t)] \\
&= \mathbb{E}[\boldsymbol{S}^2(t-1) + \boldsymbol{S}(t-1) \cdot \Delta\boldsymbol{S}(t)] \\
&= 1
\end{aligned}
$$

It follows that:

$$\begin{pmatrix} \boldsymbol{S}(t-1) \\ \boldsymbol{S}(t) \end{pmatrix} \sim \mathcal{N}(\mathbf{0}_2, \Sigma_\sigma)$$

where Σ_σ is the covariance matrix:

$$\Sigma_\sigma = \begin{pmatrix} 1 & 1 \\ 1 & 1+\sigma^2 \end{pmatrix}$$

(c) We have:

$$\Pr\{\mathcal{R}(t-1) = \mathcal{R}_k\} = \Pr\{s_{k-1} \le \boldsymbol{S}(t-1) < s_k\} = \Phi(s_k) - \Phi(s_{k-1})$$

(d) We have:

$$
\begin{aligned}
(*) &= \Pr\{\mathcal{R}(t) = \mathcal{R}_k, \mathcal{R}(t-1) = \mathcal{R}_j\} \\
&= \Pr\{s_{k-1} \le \boldsymbol{S}(t) < s_k, s_{j-1} \le \boldsymbol{S}(t-1) < s_j\} \\
&= \Phi_2(s_j, s_k; \Sigma_\sigma) - \Phi_2(s_{j-1}, s_k; \Sigma_\sigma) - \Phi_2(s_j, s_{k-1}; \Sigma_\sigma) + \Phi_2(s_{j-1}, s_{k-1}; \Sigma_\sigma)
\end{aligned}
$$

where $\Phi_2(x, y; \Sigma_\sigma)$ is the bivariate Normal cdf with covariance matrix Σ_σ.

(e) We have:

$$
\begin{aligned}
p_{j,k} &= \Pr\{\mathcal{R}(t) = \mathcal{R}_k \mid \mathcal{R}(t-1) = \mathcal{R}_j\} \\
&= \frac{\Pr\{\mathcal{R}(t) = \mathcal{R}_k, \mathcal{R}(t-1) = \mathcal{R}_j\}}{\Pr\{\mathcal{R}(t-1) = \mathcal{R}_j\}} \\
&= \frac{\Phi_2(s_j, s_k; \Sigma_\sigma) + \Phi_2(s_{j-1}, s_{k-1}; \Sigma_\sigma)}{\Phi(s_j) - \Phi(s_{j-1})} - \frac{\Phi_2(s_{j-1}, s_k; \Sigma_\sigma) + \Phi_2(s_j, s_{k-1}; \Sigma_\sigma)}{\Phi(s_j) - \Phi(s_{j-1})}
\end{aligned}
$$

(f) We have:

$$
\begin{aligned}
\mathcal{T}(\mathcal{R}_k) &= \Pr\{\mathcal{R}(t) \ne \mathcal{R}_k \mid \mathcal{R}(t-1) = \mathcal{R}_k\} \\
&= 1 - \Pr\{\mathcal{R}(t) = \mathcal{R}_k \mid \mathcal{R}(t-1) = \mathcal{R}_k\} \\
&= 1 - p_{k,k}
\end{aligned}
$$

(g) We have:

$$
\begin{aligned}
\mathcal{T}(\mathcal{R}_1, \ldots, \mathcal{R}_K) &= \sum_{k=1}^{K} \Pr\{\mathcal{R}(t-1) = \mathcal{R}_k\} \cdot \mathcal{T}(\mathcal{R}_k) \\
&= \sum_{k=1}^{K} \Pr\{\mathcal{R}(t) \ne \mathcal{R}_k, \mathcal{R}(t-1) = \mathcal{R}_k\}
\end{aligned}
$$

Table B.1: ESG migration matrix (Question 3.a)

Rating	s_k	p_k	Transition probability $p_{j,k}$				$\mathcal{T}(\mathcal{R}_k)$
D	-1.50	6.68%	92.96%	7.04%	0.00%	0.00%	7.04%
C	0.00	43.32%	1.31%	95.03%	3.66%	0.00%	4.97%
B	1.50	43.32%	0.00%	3.66%	95.03%	1.31%	4.97%
A		6.68%	0.00%	0.00%	7.04%	92.96%	7.04%
$\mathcal{T}(\mathcal{R}_1,\ldots,\mathcal{R}_K)$							5.25%

Table B.2: ESG migration matrix (Question 3.b)

Rating	s_k	p_k	Transition probability $p_{j,k}$				$\mathcal{T}(\mathcal{R}_k)$
D	-0.67	25.00%	95.15%	4.85%	0.00%	0.00%	4.85%
C	0.00	25.00%	5.27%	88.38%	6.35%	0.00%	11.62%
B	0.67	25.00%	0.00%	6.35%	88.38%	5.27%	11.62%
A		25.00%	0.00%	0.00%	4.85%	95.15%	4.85%
$\mathcal{T}(\mathcal{R}_1,\ldots,\mathcal{R}_K)$							8.23%

Table B.3: ESG migration matrix (Question 3.c)

Rating	s_k	p_k	Transition probability $p_{j,k}$				$\mathcal{T}(\mathcal{R}_k)$
D	-1.28	10.00%	93.54%	6.46%	0.00%	0.00%	6.46%
C	0.00	40.00%	1.89%	94.14%	3.97%	0.00%	5.86%
B	1.28	40.00%	0.00%	3.97%	94.14%	1.89%	5.86%
A		10.00%	0.00%	0.00%	6.46%	93.54%	6.46%
$\mathcal{T}(\mathcal{R}_1,\ldots,\mathcal{R}_K)$							5.98%

(h) The ESG migration matrices are given in Tables B.1, B.2 and B.3. We deduce that the ESG rating system defined in Question 3.a is the best rating system if we would like to reduce the monthly turnover of ESG ratings.

(i) The relationship between the parameter σ and the turnover $\mathcal{T}(\mathcal{R}_1,\ldots,\mathcal{R}_K)$ is given in Figure B.6.

(j) The relationship between the number of notches K and the turnover $\mathcal{T}(\mathcal{R}_1,\ldots,\mathcal{R}_K)$ is given in Figure B.7.

(k) An ESG rating system is mainly quantitative and highly depends on the mapping function. This is not the case of a credit rating system, which is mainly qualitative and discretionary. This explains that the turnover of an ESG rating system is higher than the turnover of a credit rating system. The stabilization of the ESG rating system implies to reduce the turnover $\mathcal{T}(\mathcal{R}_1,\ldots,\mathcal{R}_K)$, which depends on three factors: (1) the number of notches[1] K; (2) the volatility σ of score changes; (3) the design of the ESG rating system $(s_1,\ldots,s_{K-1})$. The turnover $\mathcal{T}(\mathcal{R}_1,\ldots,\mathcal{R}_K)$ has a big impact on an ESG exclusion (or negative screening) policy, because it creates noisy short-term entry/exit positions that do not necessarily correspond to a decrease or increase of the long-term ESG risks.

---

[1]This is why ESG rating systems have less notches than credit rating systems.

Figure B.6: Relationship between σ and $\mathcal{T}\left(\mathcal{R}_1,\ldots,\mathcal{R}_K\right)$
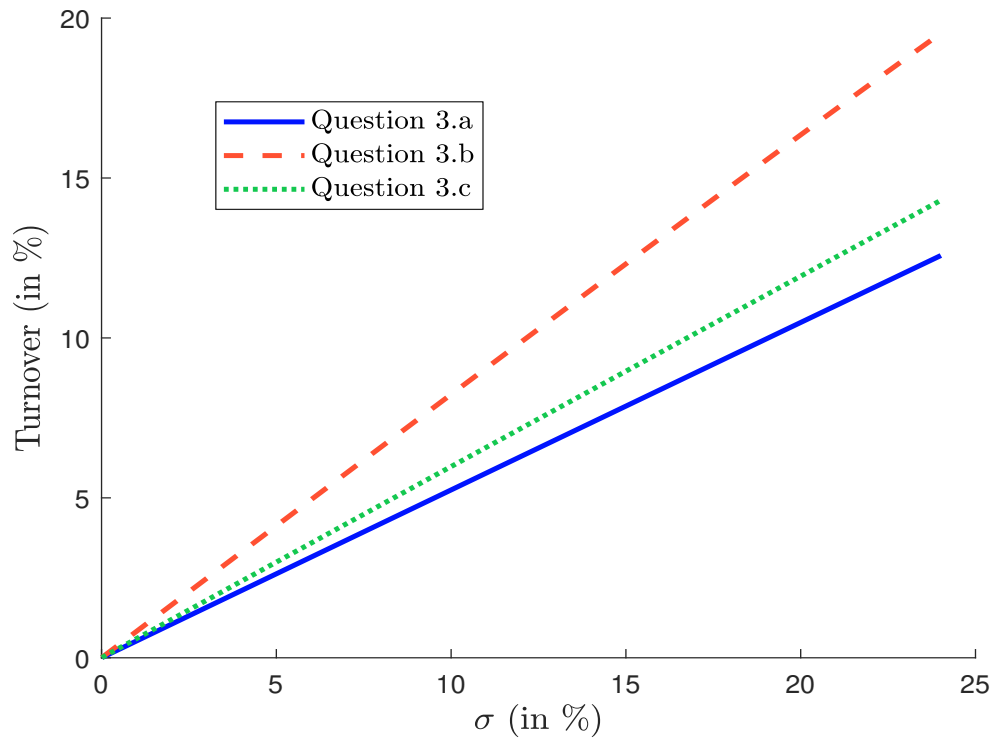
Figure B.7: Relationship between K and $\mathcal{T}\left(\mathcal{R}_1,\ldots,\mathcal{R}_K\right)$
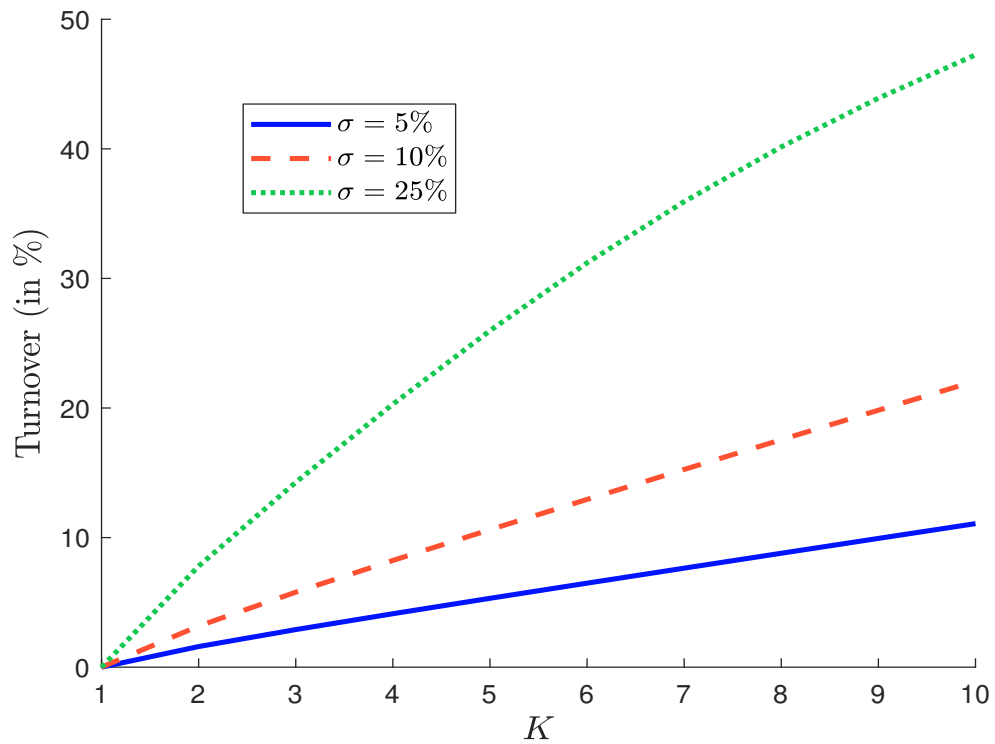
B.1.5   Markov generator of ESG migration matrix (Exercise 2.4.5)

B.1.6   Properties of Markov chains (Exercise 2.4.6)

B.1.7   Equity Portfolio optimization with ESG scores (Exercise 3.5.1)

1. We assume that the CAPM is valid.

   (a) Using the CAPM, we have:
   $$\mu_i = r + \beta_i (\mu_m - r)$$

   For instance, we have $\mu_1 = 1\% + 0.10 \times (6\% - 1\%) = 1.5\%$, $\mu_2 = 1\% + 0.30 \times 5\% = 2.5\%$. Finally, we obtain $\mu = (1.5\%, 2.5\%, 3.5\%, 5.5\%, 7.5\%, 11\%)$.

   (b) We have:
   $$\Sigma = \sigma_m^2 \beta \beta^\top + D$$

   where $D = \text{diag}\left(\tilde{\sigma}_1^2, \ldots, \tilde{\sigma}_6^2\right)$. The numerical value of Σ is:

   $$\Sigma = \begin{pmatrix} 293 & & & & & \\ 12 & 325 & & & & \\ 20 & 60 & 356 & & & \\ 36 & 108 & 180 & 424 & & \\ 52 & 156 & 260 & 468 & 797 & \\ 80 & 240 & 400 & 720 & 1\,040 & 1\,744 \end{pmatrix} \times 10^{-4}$$

   The computation of the asset volatility is equal to $\sigma_i = \sqrt{\Sigma_{i,i}}$. We deduce that $\sigma = (17.12\%, 18.03\%, 18.87\%, 20.59\%, 28.23\%, 41.76\%)$. The formula of the correlation is $\rho_{i,j} = (\sigma_i \sigma_j)^{-1} \Sigma_{i,j}$. We obtain the following correlation matrix expressed in %:

   $$\mathbb{C} = \begin{pmatrix} 100.00 & & & & & \\ 3.89 & 100.00 & & & & \\ 6.19 & 17.64 & 100.00 & & & \\ 10.21 & 29.09 & 46.33 & 100.00 & & \\ 10.76 & 30.65 & 48.81 & 80.51 & 100.00 & \\ 11.19 & 31.88 & 50.76 & 83.73 & 88.21 & 100.00 \end{pmatrix}$$

   (c) We have:

   $$w^* = \frac{\Sigma^{-1}(\mu - r\mathbf{1})}{\mathbf{1}^\top \Sigma^{-1}(\mu - r\mathbf{1})} = \begin{pmatrix} 0.94\% \\ 2.81\% \\ 5.28\% \\ 24.34\% \\ 29.06\% \\ 37.57\% \end{pmatrix}$$

   We deduce that $\mu(w^*) = w^{*\top}\mu = 7.9201\%$ and $\sigma(w^*) = \sqrt{w^{*\top}\Sigma w^*} = 28.3487\%$. The Sharpe ratio is then equal to:

   $$\text{SR}(w^* \mid r) = \frac{7.9201\% - 1\%}{28.3487\%} = 0.2441$$

   Finally, the ESG score of the tangency portfolio is:

   $$\boldsymbol{S}(w^*) = \sum_{i=1}^{6} w_i^* \boldsymbol{S}_i = -2.0347$$

(d) We have:
$$\beta_i\left(w^*\right) = \frac{\mathbf{e}_i^\top \Sigma w^*}{\sigma^2\left(w^*\right)}$$

We obtain:
$$\beta\left(w^*\right) = \begin{pmatrix} 0.0723 \\ 0.2168 \\ 0.3613 \\ 0.6503 \\ 0.9393 \\ 1.4451 \end{pmatrix}$$

The computation of the implied expected return $\tilde{\mu}_i = r + \beta_i\left(w^*\right)\left(\mu\left(w^*\right) - r\right)$ gives:
$$\tilde{\mu} = \begin{pmatrix} 1.50\% \\ 2.50\% \\ 3.50\% \\ 5.50\% \\ 7.50\% \\ 11.00\% \end{pmatrix}$$

(e) We notice that $\beta_i\left(w^*\right) \neq \beta_i\left(w_m\right)$ but the risk premia deduced from the tangency portfolio are the same as those computed with the market portfolio. In fact, the market portfolio cannot coincide with the tangency portfolio because of the cash component. Let us assume that the allocation of w_m is equal to α of the tangency portfolio w^* and $1 - \alpha$ of the risk-free asset. We deduce that:
$$\beta\left(w_m\right) = \frac{\Sigma w_m}{\sigma^2\left(w_m\right)} = \frac{\alpha \Sigma w^*}{\alpha^2 \sigma^2\left(w^*\right)} = \frac{1}{\alpha}\beta\left(w^*\right)$$

and:
$$\alpha = \frac{\beta_i\left(w^*\right)}{\beta_i\left(w_m\right)}$$

The computation gives $\alpha = 72.25\%$. The market portfolio w_m is equal to 72.25% of the tangency portfolio w^* and 27.75% of the risk-free asset. We have:
$$\begin{aligned} \mu\left(w_m\right) &= r + \alpha\left(\mu\left(w^*\right) - r\right) \\ &= 1\% + 72.25\% \times (7.9201\% - 1\%) \\ &= 6\% \end{aligned}$$

and:
$$\begin{aligned} \sigma\left(w_m\right) &= \alpha\,\sigma\left(w^*\right) \\ &= 72.25\% \times 28.3487\% \\ &= 20.48\% \end{aligned}$$

We deduce that:
$$\mathrm{SR}\left(w_m \mid r\right) = \frac{6\% - 1\%}{20.48\%} = 0.2441$$

Curiously, we do not obtain the true value of the Sharpe ratio:
$$\mathrm{SR}\left(w_m \mid r\right) = \frac{6\% - 1\%}{20\%} = 0.25$$

because the volatility $\sigma\left(w_m\right)$ is not equal to 20%. The reason is that the market portfolio computed with the tangency portfolio has an idiosyncratic risk. Indeed, we have:

$$\sqrt{w_m^\top \left(\sigma_m^2 \beta\beta^\top\right) w^\top} = 20\% < \sigma\left(w_m\right) = 20.48\%$$

Therefore, w_m is not fully diversified because the number of assets is small ($n = 6$).

2. We consider long-only portfolios and we also impose a minimum threshold $\boldsymbol{\mathcal{S}}^\star$ for the portfolio ESG score:

$$\boldsymbol{\mathcal{S}}\left(w\right) = \sum_{i=1}^n w_i \boldsymbol{\mathcal{S}}_i \geq \boldsymbol{\mathcal{S}}^\star$$

   (a) We have:

$$
\begin{aligned}
w^\star \quad &= \quad \arg\min \frac{1}{2}w^\top \Sigma w - \gamma w^\top \mu \\
&\text{s.t.} \quad \begin{cases} \mathbf{1}_6^\top w = 1 \\ w^\top \boldsymbol{\mathcal{S}} \geq \boldsymbol{\mathcal{S}}^\star \\ \mathbf{0}_6 \leq w \leq \mathbf{1}_6 \end{cases}
\end{aligned}
$$

Figure B.8: Impact of the minimum ESG score on the efficient frontier
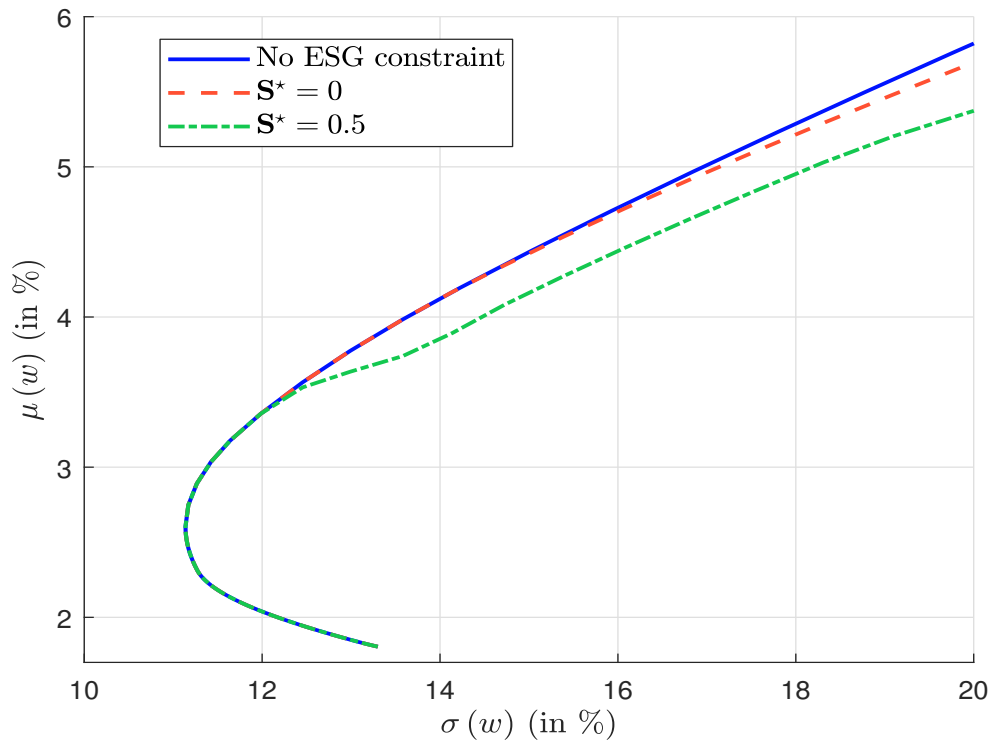
   (b) We remind that the matrix form of the QP problem is:

$$
\begin{aligned}
w^\star \quad &= \quad \arg\min \frac{1}{2}w^\top Q w - w^\top R \\
&\text{s.t.} \quad \begin{cases} Aw = B \\ Cw \leq D \\ w^- \leq w \leq w^+ \end{cases}
\end{aligned}
$$

We deduce that $Q = \Sigma$, $R = \gamma\mu$, $A = \mathbf{1}_6^\top$, $B = 1$, $C = -\boldsymbol{\mathcal{S}}^\top$, $D = -\boldsymbol{\mathcal{S}}^\star$, $w^- = \mathbf{0}_6$ and $w^+ = \mathbf{1}_6$.

(c) To compute the efficient frontier, we consider several value of $\gamma \in [-1, 2]$. For each value of γ, we compute the optimal portfolio $w^\star$ and deduce its expected return $\mu(w^\star)$ and its volatility $\sigma(w^\star)$. In Figure B.8, we compare the results for the three cases: $\boldsymbol{\mathcal{S}}^\star = -\infty$, $\boldsymbol{\mathcal{S}}^\star = 0$ and $\boldsymbol{\mathcal{S}}^\star = 0.5$. We notice that imposing a positive ESG score has little impact, especially for low-volatility portfolios when $\sigma(w^\star) \leq 16\%$. This is not the case when $\boldsymbol{\mathcal{S}}^\star = 0.5$.

(d) Let $w^\star(\gamma)$ be the MVO portfolio when the risk tolerance is equal to γ. If we use a fine grid of γ values, we can find the optimal value γ^* by solving numerically the following optimization problem with the brute force algorithm:

$$\gamma^* = \arg\max \frac{\mu(w^\star(\gamma)) - r}{\sigma(w^\star(\gamma))} \qquad \text{for } \gamma \in [0, 2]$$

Then we deduce the tangency portfolio $w^* = w^\star(\gamma^*)$. We also compute the following statistics: $\mu(w^*) = w^{*\top}\mu$, $\sigma(w^*) = \sqrt{w^{*\top}\Sigma w^*}$, $\text{SR}(w^* \mid r) = \sigma(w^*)^{-1}(\mu(w^*) - r)$ and $\boldsymbol{\mathcal{S}}(w^*) = w^{*\top}\boldsymbol{\mathcal{S}}$. Results are reported in Table B.4. In the case $\boldsymbol{\mathcal{S}}^\star = -\infty$, we retrieve the tangency portfolio, which has been found in Question 1(c). Moreover, we notice that the Sharpe ratio decreases when the minimum ESG score $\boldsymbol{\mathcal{S}}^\star$ increases.

Table B.4: Impact of the minimum ESG score on the efficient frontier

$\boldsymbol{\mathcal{S}}^\star$	$-\infty$	0	0.5
γ^*	1.1613	0.8500	0.8500
	0.9360	9.7432	9.1481
	2.8079	16.3317	19.0206
	5.2830	31.0176	40.3500
w^* (in %)	24.3441	5.1414	0.0000
	29.0609	11.6028	3.8248
	37.5681	26.1633	27.6565
$\mu(w^*)$ (in %)	7.9201	5.6710	5.3541
$\sigma(w^*)$ (in %)	28.3487	19.8979	19.2112
$\text{SR}(w^* \mid r)$	0.2441	0.2347	0.2266
$\boldsymbol{\mathcal{S}}(w^*)$	-2.0347	0.0000	0.5000

(e) We perform the same analysis as previously for several values $\boldsymbol{\mathcal{S}}^\star \in [-2.5, 2.5]$. Results are reported in Figure B.9. We verify that the Sharpe ratio is a decreasing function of $\boldsymbol{\mathcal{S}}^\star$.

(f) The market portfolio w_m is then equal to:

$$w_m = \begin{pmatrix} 9.15\% \\ 19.02\% \\ 40.35\% \\ 0.00\% \\ 3.82\% \\ 27.66\% \end{pmatrix}$$

We deduce that $\mu(w_m) = 5.3541\%$, $\sigma(w_m) = 19.2112\%$, $\text{SR}(w_m \mid r) = 0.2266$ and $\boldsymbol{\mathcal{S}}(w_m) = 0.5$.

Figure B.9: Relationship between the minimum ESG score $\boldsymbol{\mathcal{S}}^{\star}$ and the Sharpe ratio $\text{SR}(w^* \mid r)$ of the tangency portfolio
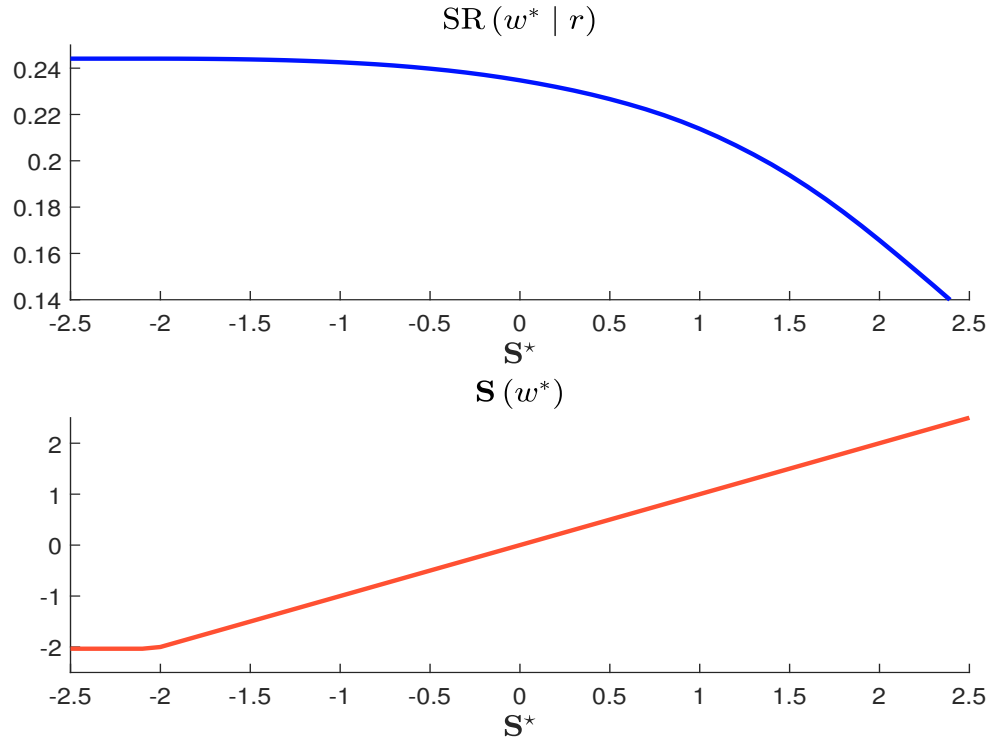
(f).i We have:

$$\beta_i(w_m) = \frac{\mathbf{e}_i^{\top} \Sigma w_m}{\sigma^2(w_m)}$$

and:

$$\tilde{\mu}_i(w_m) = r + \beta_i(w_m)(\mu(w_m) - r)$$

We deduce that the alpha return is equal to:

$$\begin{aligned} \alpha_i &= \mu_i - \tilde{\mu}_i(w_m) \\ &= (\mu_i - r) - \beta_i(w_m)(\mu(w_m) - r) \end{aligned}$$

Results are reported in Table B.5. We notice that $\alpha_i < 0$ for the first three assets and $\alpha_i > 0$ for the last three assets, implying that:

$$\begin{cases} \boldsymbol{\mathcal{S}}_i > 0 \Rightarrow \alpha_i < 0 \\ \boldsymbol{\mathcal{S}}_i < 0 \Rightarrow \alpha_i > 0 \end{cases}$$

(f).ii We have:

$$\beta(w_{\text{ew}} \mid w_m) = \frac{w_{\text{ew}}^{\top} \Sigma w_m}{\sigma^2(w_m)} = 0.9057$$

and:

$$\begin{aligned} \tilde{\mu}(w_{\text{ew}}) &= r + \beta(w_{\text{ew}} \mid w_m)(\mu(w_m) - r) \\ &= 1\% + 0.9057 \times (5.3541\% - 1\%) \\ &= 4.9435\% \end{aligned}$$

Table B.5: Computation of the alpha return due to the ESG constraint

Asset	$\beta_i(w_m)$	$\tilde{\mu}_i(w_m)$ (in %)	$\tilde{\mu}_i(w_m) - r$ (in %)	α_i (in bps)
1	0.1660	1.7228	0.7228	-22.28
2	0.4321	2.8813	1.8813	-38.13
3	0.7518	4.2733	3.2733	-77.33
4	0.8494	4.6984	3.6984	80.16
5	1.2395	6.3967	5.3967	110.33
6	1.9955	9.6885	8.6885	131.15

We deduce that:

$$
\begin{aligned}
\alpha(w_{\text{ew}}) &= \mu(w_{\text{ew}}) - \tilde{\mu}(w_{\text{ew}}) \\
&= 5.25\% - 4.9435\% \\
&= 30.65 \text{ bps}
\end{aligned}
$$

We verify that:

$$
\alpha(w_{\text{ew}}) = \sum_{i=1}^{6} w_{\text{ew},i}\alpha_i = \frac{\sum_{i=1}^{6}\alpha_i}{6} = 30.65 \text{ bps}
$$

The equally-weighted portfolio has a positive alpha. The main reason is that its ESG score is lower than the ESG score of the market portfolio:

$$
\boldsymbol{S}(w_{\text{ew}}) = -0.33 \ll \boldsymbol{S}(w_m) = 0.50
$$

3. The objective of the investor is twice. He would like to manage the tracking error risk of his portfolio with respect to the benchmark $b = (15\%, 20\%, 19\%, 14\%, 15\%, 17\%)$ and have a better ESG score than the benchmark. Nevertheless, this investor faces a long-only constraint because he cannot leverage his portfolio and he cannot also be short on the assets.

   (a) We have:

   $$
   \boldsymbol{S}(b) = \sum_{i=1}^{6} b_i \boldsymbol{S}_i = -0.1620
   $$

   (b) We have:

   $$
   \begin{cases}
   \boldsymbol{S}(w \mid b) = (w-b)^\top \boldsymbol{S} = 0.0470 \\
   \mu(w \mid b) = (w-b)^\top \mu = -0.5 \text{ bps} \\
   \sigma(w \mid b) = \sqrt{(w-b)^\top \Sigma (w-b)} = 2.8423\% \\
   \text{IR}(w \mid b) = \dfrac{\mu(w \mid b)}{\sigma(w \mid b)} = -0.0018
   \end{cases}
   $$

   The portfolio w is not optimal since it improves the ESG score of the benchmark, but its information ratio is negative. Nevertheless, the expected excess return is close to zero (less than -1 bps).

   (c) We have $\boldsymbol{S}(w \mid b) = 0.1305$, $\mu(w \mid b) = 29.5$ bps, $\sigma(w \mid b) = 2.4949\%$ and $\text{IR}(w \mid b) = 0.1182$. The portfolio w has then a positive expected excess return of 29.5 bps and a tracking error volatility of 2.4949%. Moreover, it has a better ESG score than the benchmark since its excess ESG score is equal to 0.1305.

(d) The optimization problem is:

$$w^\star = \arg\min \frac{1}{2}\sigma^2\left(w\mid b\right) - \gamma\boldsymbol{S}\left(w\mid b\right)$$

$$\text{s.t.} \quad \begin{cases} \mathbf{1}_6^\top w = 1 \\ \mathbf{0}_6 \leq w \leq \mathbf{1}_6 \end{cases}$$

(e) The objective function is equal to:

$$\begin{aligned} (*) &= \frac{1}{2}\sigma^2\left(w\mid b\right) - \gamma\boldsymbol{S}\left(w\mid b\right) \\ &= \frac{1}{2}\left(w - b\right)^\top \Sigma\left(w - b\right) - \gamma\left(w - b\right)^\top \boldsymbol{S} \\ &= \frac{1}{2}w^\top\Sigma w - w^\top\left(\Sigma b + \gamma\boldsymbol{S}\right) + \underbrace{\left(\gamma b^\top\boldsymbol{S} + \frac{1}{2}b^\top\Sigma b\right)}_{\text{does not depend on } w} \end{aligned}$$

We remind that the form of the QP problem is:

$$w^\star = \arg\min \frac{1}{2}w^\top Q w - w^\top R$$

$$\text{s.t.} \quad \begin{cases} Aw = B \\ Cw \leq D \\ w^- \leq w \leq w^+ \end{cases}$$

We deduce that $Q = \Sigma$, $R = \Sigma b + \gamma\boldsymbol{S}$, $A = \mathbf{1}_6^\top$, $B = 1$, $w^- = \mathbf{0}_6$ and $w^+ = \mathbf{1}_6$.

(f) We solve the QP problem for several values of $\gamma \in [0, 5\%]$ and obtain Figure B.10.

(g) Using the QP numerical algorithm, we compte the optimal value $\sigma\left(w\mid b\right)$ for $\gamma = 0$ and $\gamma = 5\%$. Then, we apply the bisection algorithm to find the optimal value $\gamma^\star$ such that:

$$\sigma\left(w\mid b\right) = \sigma^\star$$

We obtain the results given in Table B.6.

Table B.6: Solution of the σ-problem

Target $\sigma^\star$	0	1%	2%	3%	4%
$\gamma^\star$ (in bps)	0.000	4.338	8.677	13.015	18.524
	15.000	15.175	15.350	15.525	14.921
	20.000	21.446	22.892	24.338	25.385
$w^\star$ (in %)	19.000	23.084	27.167	31.251	35.589
	14.000	9.588	5.176	0.763	0.000
	15.000	12.656	10.311	7.967	3.555
	17.000	18.052	19.104	20.156	20.550
$\boldsymbol{S}\left(w^\star\mid b\right)$	0.000	0.230	0.461	0.691	0.915

(h) Using the QP numerical algorithm, we compte the optimal value $\boldsymbol{S}\left(w\mid b\right)$ for $\gamma = 0$ and $\gamma = 5\%$. Then, we apply the bisection algorithm to find the optimal value $\gamma^\star$ such that:

$$\boldsymbol{S}\left(w\mid b\right) = \boldsymbol{S}^\star$$

Figure B.10: Efficient frontier of tracking a benchmark with an ESG score objective
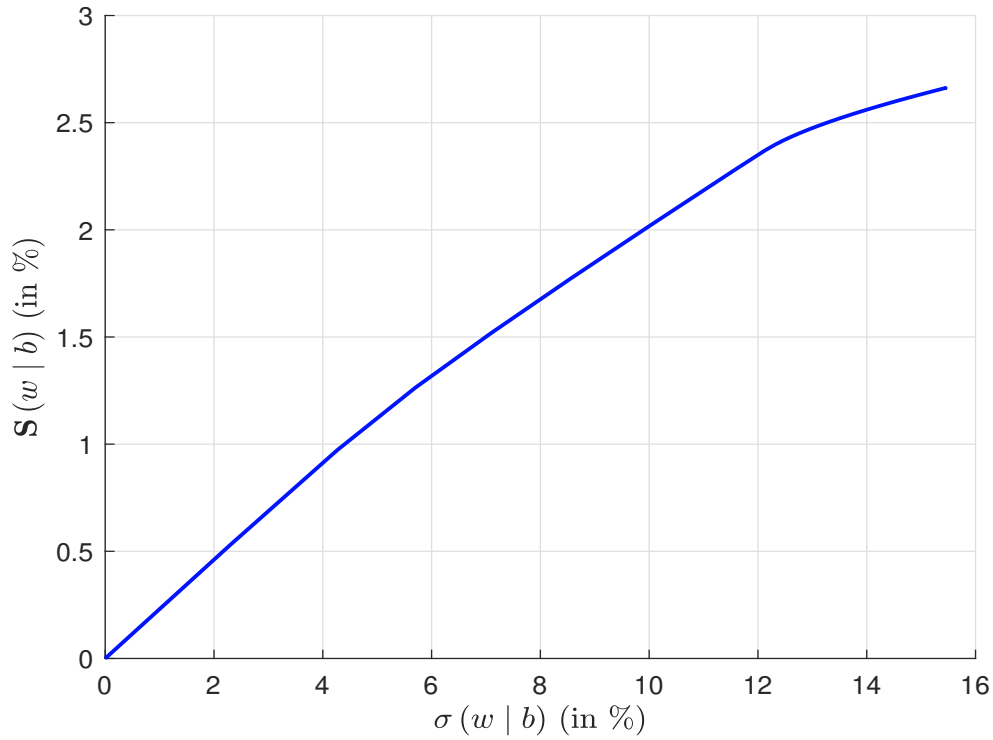
An alternative approach consists in solving the following optimization problem:

$$w^\star = \arg\min \frac{1}{2}\sigma^2(w \mid b)$$

$$\text{s.t.} \begin{cases} \mathbf{1}_6^\top w = 1 \\ \boldsymbol{S}(w \mid b) = \boldsymbol{S}^\star \\ \mathbf{0}_6 \le w \le \mathbf{1}_6 \end{cases}$$

It is easy to show that the QP problem is given by $Q = \Sigma$, $R = \Sigma b$, $A = \begin{pmatrix} \mathbf{1}_6^\top \\ \boldsymbol{S}^\top \end{pmatrix}$,

$B = \begin{pmatrix} 1 \\ \boldsymbol{S}^\star + \boldsymbol{S}^\top b \end{pmatrix}$, $w^- = \mathbf{0}_6$ and $w^+ = \mathbf{1}_6$. Finally, we obtain the results given in Table B.7.

Table B.7: Solution of the $\boldsymbol{S}$-problem

Target $\boldsymbol{S}^\star$	0	0.1	0.2	0.3	0.4
$\gamma^\star$ (in bps)	0.000	1.882	3.764	5.646	7.528
	15.000	15.076	15.152	15.228	15.304
	20.000	20.627	21.255	21.882	22.509
$w^\star$ (in %)	19.000	20.772	22.544	24.315	26.087
	14.000	12.086	10.171	8.257	6.343
	15.000	13.983	12.966	11.949	10.932
	17.000	17.456	17.913	18.369	18.825
$\sigma(w^\star \mid b)$ (in %)	0.000	0.434	0.868	1.301	1.735

(i) For the best-in-class strategy, the optimization problem becomes:

$$
w^\star = \arg\min \frac{1}{2}\sigma^2(w\mid b) - \gamma \boldsymbol{S}(w\mid b)
$$

$$
\text{s.t.} \quad \begin{cases} \mathbf{1}_6^\top w = 1 \\ w_4 = w_5 = w_6 = 0 \\ \mathbf{0}_6 \le w \le \mathbf{1}_6 \end{cases}
$$

The QP form is defined by $Q = \Sigma$, $R = \Sigma b + \gamma \boldsymbol{S}$, $A = \mathbf{1}_6^\top$, $B = 1$, $w^- = \mathbf{0}_6$ and $w^+ = \begin{pmatrix} \mathbf{1}_3 \\ \mathbf{0}_3 \end{pmatrix}$. For the worst-in-class strategy, the optimization problem becomes:

$$
w^\star = \arg\min \frac{1}{2}\sigma^2(w\mid b) - \gamma \boldsymbol{S}(w\mid b)
$$

$$
\text{s.t.} \quad \begin{cases} \mathbf{1}_6^\top w = 1 \\ w_6 = 0 \\ \mathbf{0}_6 \le w \le \mathbf{1}_6 \end{cases}
$$

The QP form is defined by $Q = \Sigma$, $R = \Sigma b + \gamma \boldsymbol{S}$, $A = \mathbf{1}_6^\top$, $B = 1$, $w^- = \mathbf{0}_6$ and $w^+ = \begin{pmatrix} \mathbf{1}_5 \\ 0 \end{pmatrix}$. The efficient frontiers are reported in Figure B.11. The exclusion strategy has less impact than the selection strategy. This last one implies a high tracking error risk.

Figure B.11: Comparison of the efficient frontiers (ESG integration, best-in-class selection and worst-in-class exclusion)
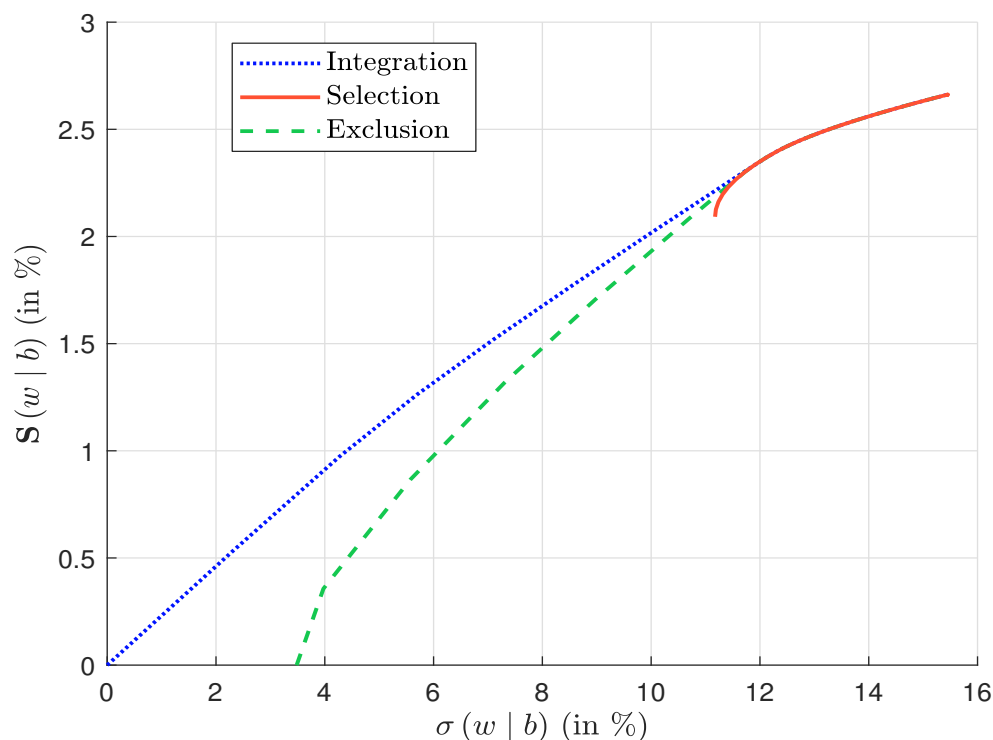
(j) We solve the first problem of Question 3(i) with $\gamma = 0$ and obtain $\sigma(w \mid b) \geq 11.17\%$. When $\sigma(w^\star \mid b) = 11.17\%$, the corresponding optimal portfolio is:

$$w^\star = \begin{pmatrix} 16.31\% \\ 34.17\% \\ 49.52\% \\ 0\% \\ 0\% \\ 0\% \end{pmatrix}$$

Remark 111 *The impact of ESG scores on optimized portfolios depends on their relationship with expected returns, volatilities, correlations, beta coefficients, etc. In the previous exercise, the results are explained because the best-in-class assets are those with the lowest expected returns and beta coefficients while the worst-in-class assets are those with the highest expected returns and beta coefficients. For instance, we obtain a high tracking error risk for the best-in-class selection strategy, because the best-in-class assets have low volatilities and correlations with respect to worst-in-class assets, implying that it is difficult to replicate these last assets with the other assets.*

B.1.8 Bond portfolio optimization with ESG scores (Exercise 3.5.2)

B.1.9 Minimum variance portfolio with climate risk (Exercise 3.5.3)

B.1.10 Cost of capital and green sentiment (Exercise 3.5.4)

B.1.11 Strategic asset allocation with ESG preferences (Exercise 3.5.5)

B.1.12 Computation of the greenium computing (Exercise 4.6.1)

B.1.13 Dependence modeling of ESG and credit ratings (Exercise 4.6.2)

B.2 Exercises related to climate risk

B.2.1 Computing the carbon risk contribution in input-output matrix models (Exercise 8.4.1)

B.2.2 Computing the carbon tax in a two-period model (Exercise 8.4.2)

B.2.3 Probability distribution of carbon momentum (Exercise 9.5.1)

B.2.4 Carbon trajectory denoising (Exercise 9.5.2)

B.2.5 Computing the optimal carbon price in ETS (Exercise 10.4.1)

B.2.6   Equity and bond portfolio optimization with green preferences (Exercise 11.3.1)

1. (a) We have:

$$\mathcal{CI}_{i,j} = \frac{\mathcal{CE}_{i,j}}{Y_i}$$

For instance, if we consider the 8th issuer, we have[2]:

$$\mathcal{CI}_{8,1} = \frac{\mathcal{CE}_{8,1}}{Y_8} = \frac{5}{25} = 0.20 \text{ tCO}_2\text{e/\$ mn}$$

$$\mathcal{CI}_{8,2} = \frac{\mathcal{CE}_{8,2}}{Y_8} = \frac{64}{25} = 2.56 \text{ tCO}_2\text{e/\$ mn}$$

$$\mathcal{CI}_{8,3} = \frac{\mathcal{CE}_{8,3}}{Y_8} = \frac{199}{25} = 7.96 \text{ tCO}_2\text{e/\$ mn}$$

Since we have:

Issuer	#1	#2	#3	#4	#5	#6	#7	#8
$\mathcal{CE}_{i,1}$	75	5 000	720	50	2 500	25	30 000	5
$\mathcal{CE}_{i,2}$	75	5 000	1 030	350	4 500	5	2000	64
$\mathcal{CE}_{i,3}$	24 000	15 000	1 210	550	500	187	30 000	199
Y_i	300	328	125	100	200	102	107	25

we obtain:

Issuer	#1	#2	#3	#4	#5	#6	#7	#8
$\mathcal{CI}_{i,1}$	0.25	15.24	5.76	0.50	12.50	0.25	280.37	0.20
$\mathcal{CI}_{i,2}$	0.25	15.24	8.24	3.50	22.50	0.05	18.69	2.56
$\mathcal{CI}_{i,3}$	80.00	45.73	9.68	5.50	2.50	1.83	280.37	7.96

(b) We have:

$$\mathcal{CI}_{i,1-2} = \frac{\mathcal{CE}_{i,1} + \mathcal{CE}_{i,2}}{Y_i} = \mathcal{CI}_{i,1} + \mathcal{CI}_{i,2}$$

and:

$$\mathcal{CI}_{i,1-3} = \mathcal{CI}_{i,1} + \mathcal{CI}_{i,2} + \mathcal{CI}_{i,3}$$

We deduce that:

Issuer	#1	#2	#3	#4	#5	#6	#7	#8
$\mathcal{CI}_{i,1}$	0.25	15.24	5.76	0.50	12.50	0.25	280.37	0.20
$\mathcal{CI}_{i,1-2}$	0.50	30.49	14.00	4.00	35.00	0.29	299.07	2.76
$\mathcal{CI}_{i,1-3}$	80.50	76.22	23.68	9.50	37.50	2.12	579.44	10.72

(c) We have:

$$
\begin{aligned}
\mathcal{CI}(b) &= \sum_{i=1}^{8} b_i \mathcal{CI}_i \\
&= 0.22 \times 80.50 + 0.19 \times 76.2195 + 0.17 \times 23.68 + 0.13 \times 9.50 + \\
&\quad 0.11 \times 37.50 + 0.08 \times 2.1275 + 0.06 \times 579.4393 + 0.04 \times 10.72 \\
&= 76.9427 \text{ tCO}_2\text{e/\$ mn}
\end{aligned}
$$

---

[2]Because 1 ktCO$_2$e/\$ bn = 1 tCO$_2$e/\$ mn.

(d) i. We have:
$$b_i = \frac{MC_i}{\sum_{k=1}^{8} MC_k}$$

and $\sum_{k=1}^{8} MC_k = \$10$ tn. We deduce that:

$$MC_i = 10 \times b_i$$

We obtain the following values of market capitalization expressed in \$ bn:

Issuer	#1	#2	#3	#4	#5	#6	#7	#8
MC_i	2 200	1 900	1 700	1 300	1 100	800	600	400

ii. Let W be the wealth invested in the benchmark portfolio b. The wealth invested in asset i is equal to $b_i W$. We deduce that the ownership ratio is equal to:

$$\varpi_i = \frac{b_i W}{MC_i} = \frac{b_i W}{b_i \sum_{k=1}^{n} MC_k} = \frac{W}{\sum_{k=1}^{n} MC_k}$$

When we invest in a capitalization-weighted portfolio, the ownership ratio is the same for all the assets. In our case, we have:

$$\varpi_i = \frac{1}{10 \times 1000} = 0.01\%$$

The ownership ratio is equal to 1 basis point.

iii. Using the financed emissions approach, the carbon emissions of our investment is equal to:

$$
\begin{aligned}
\mathcal{CE}\ (\$1\ \text{bn}) \ &= \ 0.01\% \times (75 + 75 + 24\,000) + \\
&\quad\ 0.01\% \times (5\,000 + 5\,000 + 15\,000) + \\
&\quad\ \ldots + \\
&\quad\ 0.01\% \times (5 + 64 + 199) \\
&= \ 12.3045\ \text{ktCO}_2\text{e}
\end{aligned}
$$

iv. We compute the revenues of our investment:

$$Y\ (\$1\ \text{bn}) = 0.01\% \sum_{i=1}^{8} Y_i = \$0.1287\ \text{bn}$$

We deduce that the exact carbon intensity is equal to:

$$\mathcal{CI}\ (\$1\ \text{bn}) = \frac{\mathcal{CE}\ (\$1\ \text{bn})}{Y\ (\$1\ \text{bn})} = \frac{12.3045}{0.1287} = 95.6061\ \text{tCO}_2\text{e}/\$\ \text{mn}$$

We notice that the WACI of the benchmark underestimates the exact carbon intensity of our investment by 19.5%:

$$76.9427 < 95.6061$$

2. (a) The covariance matrix $\Sigma = (\Sigma_{i,j})$ is defined by:

$$\Sigma_{i,j} = \rho_{i,j}\sigma_i\sigma_j$$

We obtain the following numerical values (expressed in bps):

$$\Sigma = \begin{pmatrix} 484.0 & 352.0 & 385.0 & 237.6 & 616.0 & 253.0 & 200.2 & 382.8 \\ 352.0 & 400.0 & 375.0 & 234.0 & 400.0 & 276.0 & 130.0 & 377.0 \\ 385.0 & 375.0 & 625.0 & 360.0 & 700.0 & 402.5 & 227.5 & 507.5 \\ 237.6 & 234.0 & 360.0 & 324.0 & 612.0 & 331.2 & 175.5 & 391.5 \\ 616.0 & 400.0 & 700.0 & 612.0 & 1600.0 & 552.0 & 416.0 & 754.0 \\ 253.0 & 276.0 & 402.5 & 331.2 & 552.0 & 529.0 & 149.5 & 466.9 \\ 200.2 & 130.0 & 227.5 & 175.5 & 416.0 & 149.5 & 169.0 & 226.2 \\ 382.8 & 377.0 & 507.5 & 391.5 & 754.0 & 466.9 & 226.2 & 841.0 \end{pmatrix}$$

(b) The tracking error variance of portfolio w with respect to benchmark b is equal to:

$$\sigma^2(w \mid b) = (w - b)^\top \Sigma (w - b)$$

The carbon intensity constraint has the following expression:

$$\sum_{i=1}^{8} w_i \mathcal{CI}_i \leq (1 - \mathcal{R}) \mathcal{CI}(b)$$

where $\mathcal{R}$ is the reduction rate and $\mathcal{CI}(b)$ is the carbon intensity of the benchmark. Let $\mathcal{CI}^\star = (1 - \mathcal{R})\mathcal{CI}(b)$ be the target value of the carbon footprint. The optimization problem is then:

$$w^\star = \arg\min \frac{1}{2}\sigma^2(w \mid b)$$
$$\text{s.t.} \quad \begin{cases} \sum_{i=1}^{8} w_i \mathcal{CI}_i \leq \mathcal{CI}^\star \\ \sum_{i=1}^{8} w_i = 1 \\ 0 \leq w_i \leq 1 \end{cases}$$

We add the second and third constraints in order to obtain a long-only portfolio.

(c) The objective function is equal to:

$$f(w) = \frac{1}{2}\sigma^2(w \mid b) = \frac{1}{2}(w - b)^\top \Sigma (w - b) = \frac{1}{2}w^\top \Sigma w - w^\top \Sigma b + \frac{1}{2}b^\top \Sigma b$$

while the matrix form of the carbon intensity constraint is:

$$\mathcal{CI}^\top w \leq \mathcal{CI}^\star$$

where $\mathcal{CI} = (\mathcal{CI}_1, \ldots, \mathcal{CI}_8)$ is the column vector of carbon intensities. Since $b^\top \Sigma b$ is a constant and does not depend on w, we can cast the previous optimization problem into a QP problem:

$$w^\star = \arg\min \frac{1}{2}w^\top Q w - w^\top R$$
$$\text{s.t.} \quad \begin{cases} Aw = B \\ Cw \leq D \\ w^- \leq w \leq w^+ \end{cases}$$

We have $Q = \Sigma$, $R = \Sigma b$, $A = \mathbf{1}_8^\top$, $B = 1$, $C = \mathcal{CI}^\top$, $D = \mathcal{CI}^\star$, $w^- = \mathbf{0}_8$ and $w^+ = \mathbf{1}_8$.

(d) We have:

$$\mathcal{CI}(b) = 0.22 \times 0.50 + 0.19 \times 30.4878 + \ldots + 0.04 \times 2.76$$
$$= 30.7305 \text{ tCO}_2\text{e/\$ mn}$$

We deduce that:

$$\mathcal{CI}^\star = (1 - \mathcal{R})\mathcal{CI}(b) = 0.80 \times 30.7305 = 24.5844 \text{ tCO}_2\text{e/\$ mn}$$

Therefore, the inequality constraint of the QP problem is:

$$\begin{pmatrix} 0.50 & 30.49 & 14.00 & 4.00 & 35.00 & 0.29 & 299.07 & 2.76 \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_7 \\ w_8 \end{pmatrix} \leq 24.5844$$

We obtain the following optimal solution:

$$w^\star = \begin{pmatrix} 23.4961\% \\ 17.8129\% \\ 17.1278\% \\ 15.4643\% \\ 10.4037\% \\ 7.5903\% \\ 4.0946\% \\ 4.0104\% \end{pmatrix}$$

The minimum tracking error volatility $\sigma(w^\star \mid b)$ is equal to 15.37 bps.

(e) In Table B.8, we report the optimal solution $w^\star$ (expressed in %) of the optimization problem for different values of $\mathcal{R}$. We also indicate the carbon intensity of the portfolio (in tCO$_2$e/\$ mn) and the tracking error volatility (in bps). For instance, if $\mathcal{R}$ is set to 50%, the weights of assets #1, #3, #4 and #8 increase whereas the weights of assets #2, #5, #6 and #7 decrease. The carbon intensity of this portfolio is equal to 15.3653 tCO$_2$e/\$ mn. The tracking error volatility is below 40 bps, which is relatively low.

(f) In this case, the inequality constraint $Cw \leq D$ is defined by:

$$C = \mathcal{CI}_{1-3}^\top = \begin{pmatrix} 80.5000 \\ 76.2195 \\ 23.6800 \\ 9.5000 \\ 37.5000 \\ 2.1275 \\ 579.4393 \\ 10.7200 \end{pmatrix}^\top$$

and:

$$D = (1 - \mathcal{R}) \times 76.9427$$

We obtain the results given in Table B.9.

Table B.8: Solution of the equity optimization problem (scope $\mathcal{SC}_{1-2}$)

$\mathcal{R}$	0%	20%	30%	50%	70%
w_1	22.0000	23.4961	24.2441	25.7402	30.4117
w_2	19.0000	17.8129	17.2194	16.0323	9.8310
w_3	17.0000	17.1278	17.1917	17.3194	17.8348
w_4	13.0000	15.4643	16.6964	19.1606	23.3934
w_5	11.0000	10.4037	10.1055	9.5091	7.1088
w_6	8.0000	7.5903	7.3854	6.9757	6.7329
w_7	6.0000	4.0946	3.1418	1.2364	0.0000
w_8	4.0000	4.0104	4.0157	4.0261	4.6874
$\mathcal{CI}(w)$	30.7305	24.5844	21.5114	15.3653	9.2192
$\sigma(w \mid b)$	0.00	15.37	23.05	38.42	72.45

Table B.9: Solution of the equity optimization problem (scope $\mathcal{SC}_{1-3}$)

$\mathcal{R}$	0%	20%	30%	50%	70%
w_1	22.0000	23.9666	24.9499	26.4870	13.6749
w_2	19.0000	17.4410	16.6615	8.8001	0.0000
w_3	17.0000	17.1988	17.2981	19.4253	24.1464
w_4	13.0000	16.5034	18.2552	25.8926	41.0535
w_5	11.0000	10.2049	9.8073	7.1330	3.5676
w_6	8.0000	7.4169	7.1254	7.0659	8.8851
w_7	6.0000	3.2641	1.8961	0.0000	0.0000
w_8	4.0000	4.0043	4.0065	5.1961	8.6725
$\mathcal{CI}(w)$	76.9427	61.5541	53.8599	38.4713	23.0828
$\sigma(w \mid b)$	0.00	21.99	32.99	104.81	414.48

(g) In Figure B.12, we report the relationship between the reduction rate $\mathcal{R}$ and the tracking error volatility $\sigma(w \mid b)$. The choice of the scope has little impact when $\mathcal{R} \leq 45\%$. Then, we notice a high increase when we consider the scope $1+2+3$. The portfolio's weights are given in Figure B.13. For assets #1 and #3, the behavior is divergent when we compare scopes $1+2$ and $1+2+3$.

3. (a) We have:

$$
\begin{aligned}
\mathrm{MD}(b) &= \sum_{i=1}^{n} b_i \, \mathrm{MD}_i \\
&= 0.22 \times 3.56 + 0.19 \times 7.48 + \ldots + 0.04 \times 7.96 \\
&= 5.96 \text{ years}
\end{aligned}
$$

and:

$$
\begin{aligned}
\mathrm{DTS}(b) &= \sum_{i=1}^{n} b_i \, \mathrm{DTS}_i \\
&= 0.22 \times 103 + 0.19 \times 155 + \ldots + 0.04 \times 155 \\
&= 210.73 \text{ bps}
\end{aligned}
$$

Figure B.12: Impact of the scope on the tracking error volatility
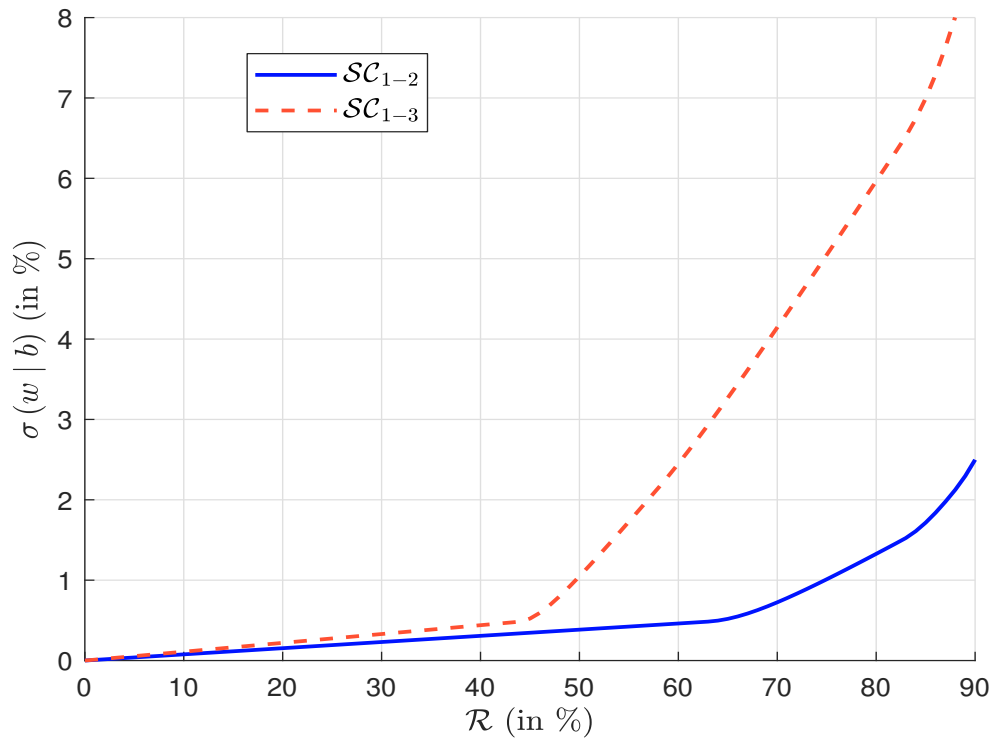
Figure B.13: Impact of the scope on the portfolio allocation (in %)
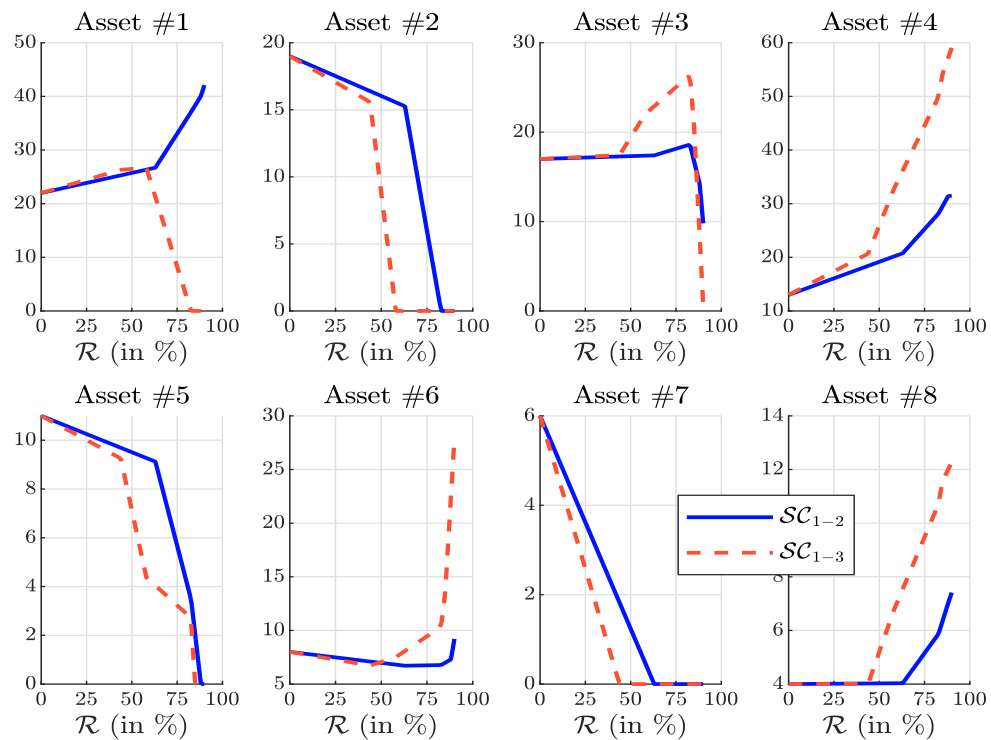
(b) We have:

$$\begin{cases} \mathrm{MD}\,(w_{\mathrm{ew}}) = 6.20 \text{ years} \\ \mathrm{DTS}\,(w_{\mathrm{ew}}) = 228.50 \text{ bps} \\ \sigma_{\mathrm{AS}}\,(w_{\mathrm{ew}} \mid b) = 17.03\% \\ \sigma_{\mathrm{MD}}\,(w_{\mathrm{ew}} \mid b) = 1.00 \text{ years} \\ \sigma_{\mathrm{DTS}}\,(w_{\mathrm{ew}} \mid b) = 36.19 \text{ bps} \end{cases}$$

(c) We have:

$$\begin{aligned} \mathcal{R}_{\mathrm{AS}}\,(w \mid b) \quad = \quad & (w_1 - 0.22)^2 + (w_2 - 0.19)^2 + (w_3 - 0.17)^2 + (w_4 - 0.13)^2 + \\ & (w_5 - 0.11)^2 + (w_6 - 0.08)^2 + (w_7 - 0.06)^2 + (w_8 - 0.04)^2 \end{aligned}$$

The objective function is then:

$$f\,(w) = \frac{1}{2}\mathcal{R}_{\mathrm{AS}}\,(w \mid b)$$

The optimal solution is equal to:

$$w^\star = (17.30\%, 17.41\%, 20.95\%, 14.41\%, 10.02\%, 11.09\%, 0\%, 8.81\%)$$

The risk metrics are:

$$\begin{cases} \mathrm{MD}\,(w^\star) = 5.96 \text{ years} \\ \mathrm{DTS}\,(w^\star) = 210.73 \text{ bps} \\ \sigma_{\mathrm{AS}}\,(w^\star \mid b) = 10.57\% \\ \sigma_{\mathrm{MD}}\,(w^\star \mid b) = 0.43 \text{ years} \\ \sigma_{\mathrm{DTS}}\,(w^\star \mid b) = 15.21 \text{ bps} \end{cases}$$

(d) We have[3]:

$$\begin{aligned} \mathcal{R}_{\mathrm{MD}}\,(w \mid b) \quad = \quad & \left(\sum_{i=1,3,4,6} (w_i - b_i)\,\mathrm{MD}_i \right)^2 + \left(\sum_{i=2,5,7,8} (w_i - b_i)\,\mathrm{MD}_i \right)^2 \\ = \quad & \left(\sum_{i=1,3,4,6} w_i\,\mathrm{MD}_i - \mathrm{MD}_1^\star \right)^2 + \left(\sum_{i=2,5,7,8} w_i\,\mathrm{MD}_i - \mathrm{MD}_2^\star \right)^2 \\ = \quad & (3.56w_1 + 6.54w_3 + 10.23w_4 + 2.30w_6 - 3.4089)^2 + \\ & (7.48w_2 + 2.40w_5 + 9.12w_7 + 7.96w_8 - 2.5508)^2 \end{aligned}$$

The objective function is then:

$$f\,(w) = \frac{\varphi_{\mathrm{AS}}}{2}\mathcal{R}_{\mathrm{AS}}\,(w \mid b) + \frac{\varphi_{\mathrm{MD}}}{2}\mathcal{R}_{\mathrm{MD}}\,(w \mid b)$$

The optimal solution is equal to:

$$w^\star = (16.31\%, 18.44\%, 17.70\%, 13.82\%, 11.67\%, 11.18\%, 0\%, 10.88\%)$$

The risk metrics are:

$$\begin{cases} \mathrm{MD}\,(w^\star) = 5.93 \text{ years} \\ \mathrm{DTS}\,(w^\star) = 210.73 \text{ bps} \\ \sigma_{\mathrm{AS}}\,(w^\star \mid b) = 11.30\% \\ \sigma_{\mathrm{MD}}\,(w^\star \mid b) = 0.03 \text{ years} \\ \sigma_{\mathrm{DTS}}\,(w^\star \mid b) = 3.70 \text{ bps} \end{cases}$$

---

[3]We verify that $3.4089 + 2.5508 = 5.9597$ years.

(e) We have[4]:

$$\mathcal{R}_{\mathrm{DTS}}\left(w \mid b\right) = \left(\sum_{i=1,3,4,6}\left(w_i - b_i\right)\mathrm{DTS}_i\right)^2 + \left(\sum_{i=2,5,7,8}\left(w_i - b_i\right)\mathrm{DTS}_i\right)^2$$

$$= \left(103w_1 + 75w_3 + 796w_4 + 45w_6 - 142.49\right)^2 +$$
$$\left(155w_2 + 89w_5 + 320w_7 + 245w_8 - 68.24\right)^2$$

The objective function is then:

$$f\left(w\right) = \frac{\varphi_{\mathrm{AS}}}{2}\mathcal{R}_{\mathrm{AS}}\left(w \mid b\right) + \frac{\varphi_{\mathrm{MD}}}{2}\mathcal{R}_{\mathrm{MD}}\left(w \mid b\right) + \frac{\varphi_{\mathrm{DTS}}}{2}\mathcal{R}_{\mathrm{DTS}}\left(w \mid b\right)$$

The optimal solution is equal to:

$$w^{\star} = \left(16.98\%, 17.21\%, 18.26\%, 13.45\%, 12.10\%, 9.46\%, 0\%, 12.55\%\right)$$

The risk metrics are:

$$\begin{cases} \mathrm{MD}\left(w^{\star}\right) = 5.97 \text{ years} \\ \mathrm{DTS}\left(w^{\star}\right) = 210.68 \text{ bps} \\ \sigma_{\mathrm{AS}}\left(w^{\star} \mid b\right) = 11.94\% \\ \sigma_{\mathrm{MD}}\left(w^{\star} \mid b\right) = 0.03 \text{ years} \\ \sigma_{\mathrm{DTS}}\left(w^{\star} \mid b\right) = 0.06 \text{ bps} \end{cases}$$

(f) We summarize the results in Table B.10.

Table B.10: Solution of the bond optimization problem (scope $\mathcal{SC}_{1-3}$)

Problem	Benchmark	3.(c)	3.(d)	3.(e)
w_1	22.0000	17.3049	16.3102	16.9797
w_2	19.0000	17.4119	18.4420	17.2101
w_3	17.0000	20.9523	17.6993	18.2582
w_4	13.0000	14.4113	13.8195	13.4494
w_5	11.0000	10.0239	11.6729	12.1008
w_6	8.0000	11.0881	11.1792	9.4553
w_7	6.0000	0.0000	0.0000	0.0000
w_8	4.0000	8.8075	10.8769	12.5464
$\mathrm{MD}\left(w\right)$	5.9597	5.9597	5.9344	5.9683
$\mathrm{DTS}\left(w\right)$	210.7300	210.7300	210.7300	210.6791
$\sigma_{\mathrm{AS}}\left(w \mid b\right)$	0.0000	10.5726	11.3004	11.9400
$\sigma_{\mathrm{MD}}\left(w \mid b\right)$	0.0000	0.4338	0.0254	0.0308
$\sigma_{\mathrm{DTS}}\left(w \mid b\right)$	0.0000	15.2056	3.7018	0.0561
$\mathcal{CI}\left(w\right)$	76.9427	38.4713	38.4713	38.4713

(g) The goal is to write the objective function into a quadratic function:

$$f\left(w\right) = \frac{\varphi_{\mathrm{AS}}}{2}\mathcal{R}_{\mathrm{AS}}\left(w \mid b\right) + \frac{\varphi_{\mathrm{MD}}}{2}\mathcal{R}_{\mathrm{MD}}\left(w \mid b\right) + \frac{\varphi_{\mathrm{DTS}}}{2}\mathcal{R}_{\mathrm{DTS}}\left(w \mid b\right)$$

$$= \frac{1}{2}w^{\top}Q\left(b\right)w - w^{\top}R\left(b\right) + c\left(b\right)$$

---

[4]We verify that $142.49 + 68.24 = 210.73$ bps.

where:

$$
\begin{aligned}
\mathcal{R}_{\mathrm{AS}}\left(w \mid b\right) &= (w_1 - 0.22)^2 + (w_2 - 0.19)^2 + (w_3 - 0.17)^2 + (w_4 - 0.13)^2 + \\
& \quad (w_5 - 0.11)^2 + (w_6 - 0.08)^2 + (w_7 - 0.06)^2 + (w_8 - 0.04)^2 \\
\mathcal{R}_{\mathrm{MD}}\left(w \mid b\right) &= (3.56w_1 + 6.54w_3 + 10.23w_4 + 2.30w_6 - 3.4089)^2 + \\
& \quad (7.48w_2 + 2.40w_5 + 9.12w_7 + 7.96w_8 - 2.5508)^2 \\
\mathcal{R}_{\mathrm{DTS}}\left(w \mid b\right) &= (103w_1 + 75w_3 + 796w_4 + 45w_6 - 142.49)^2 + \\
& \quad (155w_2 + 89w_5 + 320w_7 + 245w_8 - 68.24)^2
\end{aligned}
$$

We use the analytical approach which is described in Section 11.1.3 on pages 561-566. Moreover, we rearrange the universe such that the first fourth assets belong to the first sector and the last fourth assets belong to the second sector. In this case, we have:

$$
w = \Big(\underbrace{w_1, w_3, w_4, w_6}_{\boldsymbol{S}ector_1}, \underbrace{w_2, w_5, w_7, w_8}_{\boldsymbol{S}ector_2} \Big)
$$

The matrix $Q\left(b\right)$ is block-diagonal:

$$
Q\left(b\right) = \left(\begin{array}{cc} Q_1 & \mathbf{0}_{4,4} \\ \mathbf{0}_{4,4} & Q_2 \end{array} \right)
$$

where the matrices Q_1 and Q_2 are equal to:

$$
Q_1 = \left(\begin{array}{cccc}
11\,025.8400 & 8\,307.0600 & 82\,898.4700 & 4\,839.7000 \\
8\,307.0600 & 6\,794.2900 & 61\,372.6050 & 3\,751.0500 \\
82\,898.4700 & 61\,372.6050 & 636\,332.3225 & 36\,408.2250 \\
4\,839.7000 & 3\,751.0500 & 36\,408.2250 & 2\,257.2500
\end{array} \right)
$$

and:

$$
Q_2 = \left(\begin{array}{cccc}
25\,523.7600 & 14\,243.8000 & 51\,305.4400 & 39\,463.5200 \\
14\,243.8000 & 8\,165.0000 & 29\,027.2000 & 22\,282.6000 \\
51\,305.4400 & 29\,027.2000 & 104\,579.3600 & 80\,214.8800 \\
39\,463.5200 & 22\,282.6000 & 80\,214.8800 & 61\,709.0400
\end{array} \right)
$$

The vector $R\left(b\right)$ is defined as follows:

$$
R\left(b\right) = \left(\begin{array}{c}
15\,001.8621 \\
11\,261.1051 \\
114\,306.8662 \\
6\,616.0617 \\
11\,073.1996 \\
6\,237.4080 \\
22\,424.3824 \\
17\,230.4092
\end{array} \right)
$$

Finally, the value of $c\left(b\right)$ is equal to:

$$
c\left(b\right) = 12\,714.3386
$$

Using a QP solver, we obtain the following numerical solution:

$$
\begin{pmatrix} w_1 \\ w_3 \\ w_4 \\ w_6 \\ w_2 \\ w_5 \\ w_7 \\ w_8 \end{pmatrix} = \begin{pmatrix} 16.9796 \\ 18.2582 \\ 13.4494 \\ 9.4553 \\ 17.2102 \\ 12.1009 \\ 0.0000 \\ 12.5464 \end{pmatrix} \times 10^{-2}
$$

We observe some small differences (after the fifth digit) because the QP solver is more efficient than a traditional nonlinear solver.

4. (a) The optimization problem is:

$$
\begin{aligned}
w^\star &= \arg\min \mathcal{D}\left(w \mid b\right) \\
&\text{s.t.} \quad \begin{cases} \mathbf{1}_8^\top w = 1 \\ \mathcal{CI}^\top w \leq (1 - \mathcal{R})\,\mathcal{CI}\,(b) \\ \mathbf{0}_8 \leq w \leq \mathbf{1}_8 \end{cases}
\end{aligned}
$$

(b) We use the absolute value trick and obtain the following optimization problem:

$$
\begin{aligned}
w^\star &= \arg\min \frac{1}{2}\varphi_{\mathrm{AS}} \sum_{i=1}^{8} \tau_{i,w} + \varphi_{\mathrm{MD}} \sum_{j=1}^{2} \tau_{j,\mathrm{MD}} + \varphi_{\mathrm{DTS}} \sum_{j=1}^{2} \tau_{j,\mathrm{DTS}} \\
&\text{s.t.} \quad \begin{cases} \mathbf{1}_8^\top w = 1 \\ \mathbf{0}_8 \leq w \leq \mathbf{1}_8 \\ \mathcal{CI}^\top w \leq (1 - \mathcal{R})\,\mathcal{CI}\,(b) \\ |w_i - b_i| \leq \tau_{i,w} \\ \left| \sum_{i \in \boldsymbol{Sector}_j} (w_i - b_i)\,\mathrm{MD}_i \right| \leq \tau_{j,\mathrm{MD}} \\ \left| \sum_{i \in \boldsymbol{Sector}_j} (w_i - b_i)\,\mathrm{DTS}_i \right| \leq \tau_{j,\mathrm{DTS}} \\ \tau_{i,w} \geq 0, \tau_{j,\mathrm{MD}} \geq 0, \tau_{j,\mathrm{DTS}} \geq 0 \end{cases}
\end{aligned}
$$

We can now formulate this problem as a standard LP problem:

$$
\begin{aligned}
x^\star &= \arg\min c^\top x \\
&\text{s.t.} \quad \begin{cases} Ax = B \\ Cx \leq D \\ x^- \leq x \leq x^+ \end{cases}
\end{aligned}
$$

where x is the 20×1 vector defined as follows:

$$
x = \begin{pmatrix} w \\ \tau_w \\ \tau_{\mathrm{MD}} \\ \tau_{\mathrm{DTS}} \end{pmatrix}
$$

The 20×1 vector c is equal to:

$$
c = \begin{pmatrix} \mathbf{0}_8 \\ \frac{1}{2}\varphi_{\mathrm{AS}}\mathbf{1}_8 \\ \varphi_{\mathrm{MD}}\mathbf{1}_2 \\ \varphi_{\mathrm{DTS}}\mathbf{1}_2 \end{pmatrix}
$$

The equality constraint is defined by $A = \begin{pmatrix} \mathbf{1}_8^\top & \mathbf{0}_8^\top & \mathbf{0}_2^\top & \mathbf{0}_2^\top \end{pmatrix}$ and $B = 1$. The bounds are $x^- = \mathbf{0}_{20}$ and $x^+ = \infty \cdot \mathbf{1}_{20}$. For the inequality constraint, we have[5]:

$$Cx \leq D \Leftrightarrow \begin{pmatrix} I_8 & -I_8 & \mathbf{0}_{8,2} & \mathbf{0}_{8,2} \\ -I_8 & -I_8 & \mathbf{0}_{8,2} & \mathbf{0}_{8,2} \\ C_{\mathrm{MD}} & \mathbf{0}_{2,8} & -I_2 & \mathbf{0}_{2,2} \\ -C_{\mathrm{MD}} & \mathbf{0}_{2,8} & -I_2 & \mathbf{0}_{2,2} \\ C_{\mathrm{DTS}} & \mathbf{0}_{2,8} & \mathbf{0}_{2,2} & -I_2 \\ -C_{\mathrm{DTS}} & \mathbf{0}_{2,8} & \mathbf{0}_{2,2} & -I_2 \\ \boldsymbol{CI}^\top & \mathbf{0}_{1,8} & 0 & 0 \end{pmatrix} x \leq \begin{pmatrix} b \\ -b \\ \mathrm{MD}^\star \\ -\mathrm{MD}^\star \\ \mathrm{DTS}^\star \\ -\mathrm{DTS}^\star \\ (1 - \boldsymbol{\mathcal{R}})\,\boldsymbol{CI}\,(b) \end{pmatrix}$$

where:

$$C_{\mathrm{MD}} = \begin{pmatrix} 3.56 & 0.00 & 6.54 & 10.23 & 0.00 & 2.30 & 0.00 & 0.00 \\ 0.00 & 7.48 & 0.00 & 0.00 & 2.40 & 0.00 & 9.12 & 7.96 \end{pmatrix}$$

and:

$$C_{\mathrm{DTS}} = \begin{pmatrix} 103 & 0 & 75 & 796 & 0 & 45 & 0 & 0 \\ 0 & 155 & 0 & 0 & 89 & 0 & 320 & 245 \end{pmatrix}$$

The 2×1 vectors $\mathrm{MD}^\star$ and $\mathrm{DTS}^\star$ are respectively equal to $(3.4089, 2.5508)$ and $(142.49, 68.24)$.

(c) We obtain the following solution:

$$\begin{aligned} w^\star &= (18.7360, 15.8657, 17.8575, 13.2589, 11, 9.4622, 0, 13.8196) \times 10^{-2} \\ \tau_w^\star &= (3.2640, 3.1343, 0.8575, 0.2589, 0, 1.4622, 6, 9.8196) \times 10^{-2} \\ \tau_{\mathrm{MD}} &= (0, 0) \\ \tau_{\mathrm{DTS}} &= (0, 0) \end{aligned}$$

In Table B.11, we compare the two solutions[6]. They are very close. In fact, we notice that the LP solution matches perfectly the MD and DTS constraints, but has a higher AS risk $\sigma_{\mathrm{AS}}(w \mid b)$. If we note the two solutions $w^\star(\mathcal{L}_1)$ and $w^\star(\mathcal{L}_2)$, we have:

$$\begin{cases} \mathcal{R}(w^\star(\mathcal{L}_2) \mid b) = 1.4524 < \mathcal{R}(w^\star(\mathcal{L}_1) \mid b) = 1.5584 \\ \mathcal{D}(w^\star(\mathcal{L}_2) \mid b) = 13.9366 > \mathcal{D}(w^\star(\mathcal{L}_1) \mid b) = 12.3982 \end{cases}$$

There is a trade-off between the $\mathcal{L}_1$- and $\mathcal{L}_2$-norm risk measures. This is why we cannot say that one solution dominates the other.

---

[5]C is a 25×8 matrix and D is a 25×1 vector.

[6]The units are the following: % for the weights w_i, and the active share metrics $\sigma_{\mathrm{AS}}(w \mid b)$ and $\mathcal{D}_{\mathrm{AS}}(w \mid b)$; years for the modified duration metrics $\mathrm{MD}(w)$, $\sigma_{\mathrm{MD}}(w \mid b)$ and $\mathcal{D}_{\mathrm{MD}}(w \mid b)$; bps for the duration-times-spread metrics $\mathrm{DTS}(w)$, $\sigma_{\mathrm{DTS}}(w \mid b)$ and $\mathcal{D}_{\mathrm{DTS}}(w \mid b)$; $\mathrm{tCO_2e/\$}$ mn for the carbon intensity $\mathrm{DTS}(w)$.

Table B.11: Solution of the bond optimization problem (scope $\boldsymbol{SC}_{1-3}$)

Problem	Benchmark	3.(e)	4.(c)
w_1	22.0000	16.9796	18.7360
w_2	19.0000	17.2102	15.8657
w_3	17.0000	18.2582	17.8575
w_4	13.0000	13.4494	13.2589
w_5	11.0000	12.1009	11.0000
w_6	8.0000	9.4553	9.4622
w_7	6.0000	0.0000	0.0000
w_8	4.0000	12.5464	13.8196
MD (w)	5.9597	5.9683	5.9597
DTS (w)	210.7300	210.6791	210.7300
$\sigma_{\mathrm{AS}}(w \mid b)$	0.0000	11.9400	12.4837
$\sigma_{\mathrm{MD}}(w \mid b)$	0.0000	0.0308	0.0000
$\sigma_{\mathrm{DTS}}(w \mid b)$	0.0000	0.0561	0.0000
$\mathcal{D}_{\mathrm{AS}}(w \mid b)$	0.0000	25.6203	24.7964
$\mathcal{D}_{\mathrm{MD}}(w \mid b)$	0.0000	0.0426	0.0000
$\mathcal{D}_{\mathrm{DTS}}(w \mid b)$	0.0000	0.0608	0.0000
$\boldsymbol{\mathcal{CI}}(w)$	76.9427	38.4713	38.4713

Bibliography

ABOUMAHBOUB T, AUER C, BAUER N, ..., and UECKERDT F (2020). *REMIND — REgional Model of INvestments and Development*. Technical Documentation, Version 2.1.0.

ADMATI, A. R., and PFLEIDERER, P. (2009). The "Wall Street Walk" and Shareholder Activism: Exit as a Form of Voice. *Review of Financial Studies*, 22(7), pp. 2645-2685.

ALESSI, L., and BATTISTON, S. (2022). Two Sides of the Same Coin: Green Taxonomy Alignment versus Transition Risk in Financial Portfolios. *International Review of Financial Analysis*, 102319.

ALLEN, M., AXELSSON, K., CALDECOTT, B., HALE, T., HEPBURN, C., HICKEY, C., MITCHELL-LARSON, E., MALHI, Y., OTTO, F., SEDDON, N., and SMITH, S. (2020). The Oxford Principles for Net Zero Aligned Carbon Offsetting. *Report*, September, 13 pages.

ALLEY, R. B., ANANDAKRISHNAN, S., CHRISTIANSON, K., HORGAN, H. J., MUTO, A., PARIZEK, B. R., POLLARD, D., and WALKER, R. T. (2015). Oceanic Forcing of Ice-Sheet Retreat: West Antarctica and More. *Annual Review of Earth and Planetary Sciences*, 43, pp. 207-231.

ALMOND, R. E. A., GROOTEN, M., JUFFE BIGNOLI, D., and PETERSEN, T. (Eds) (2022). Living Planet Report 2022 — Building a Nature-positive Society. *Report*, WWF.

ALTMAN, E.I. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *Journal of Finance*, 23(4), pp. 589-609.

ANDERSSON, M., BOLTON, P., and SAMAMA, F. (2016). Hedging Climate Risk. *Financial Analysts Journal*, 72(3), pp. 13-32.

ANTHOFF, D., and TOL, R. S. J. (2014). *The Climate Framework for Uncertainty, Negotiation and Distribution (Fund)*. Technical Description, Version 3.9.

ANTRÀS, P., CHOR, D., FALLY, T., and HILLBERRY, R. (2012). Measuring the Upstreamness of Production and Trade Flows. *American Economic Review*, 102(3), pp. 412-416.

APPEL, I. R., GORMLEY, T. A., and KEIM, D. B. (2016). Passive Investors, Not Passive Owners. *Journal of Financial Economics*, 121(1), pp. 111-141.

ARCHER, D., and PIERREHUMBERT, R. (Eds) (2011). *The Warming Papers: The Scientific Foundation for the Climate Change Forecast*. John Wiley & Sons.

ARMSTRONG MCKAY, D. I., STAAL, A., ABRAMS, J. F., ..., and LENTON, T. M. (2022). Exceeding 1.5°C Global Warming Could Trigger Multiple Climate Tipping Points. *Science*, 377(6611), eabn7950, 11 pages.

ARRHENIUS, S. (1896). On the Influence of Carbonic Acid in the Air upon the Temperature of the Ground. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 41(251), pp. 237-276.

ARRHENIUS, S. (1908). *Worlds in the Making: The Evolution of the Universe*. Harper.

ATZ, U., VAN HOLT, T., LIU, Z. Z., and BRUNO, C. C. (2022). Does Sustainability Generate Better Financial Performance? Review, Meta-analysis, and Propositions. *Journal of Sustainable Finance & Investment*, 13(1), pp. 802-825.

Autorité de Contrôle Prudentiel et de Résolution (2021). A First Assessment of Financial Risks Stemming from Climate Change: The Main Results of the 2020 Climate Pilot Exercise. *Analyses et synthèses*, 122-2021, 4 May 2021.

AVRAMOV, D., CHENG, S., LIOUI, A., and TARELLI, A. (2022). Sustainable Investing with ESG Rating Uncertainty. *Journal of Financial Economics*, 145(2), pp. 642-664.

BAKER, M., BERGSTRESSER, D., SERAFEIM, G., and WURGLER, J. (2022). The Pricing and Ownership of US Green Bonds, *Annual Review of Financial Economics*, 14, pp. 415-437.

BANDEIRA, B., JAMET, J.-L., JAMET, D., and GINOUX, J.-M. (2013). Mathematical Convergences of Biodiversity Indices. *Ecological Indicators*, 29, pp. 522-528.

Bank of England (2022). Results of the 2021 Climate Biennial Exploratory Scenario (CBES). *Report*, 24 May 2022.

BARAHHOU, I., BEN SLIMANE, M., OULID AZOUZ, N., and RONCALLI, T. (2022). Net Zero Investment Portfolios — Part 1. The Comprehensive Integrated Approach. *SSRN*, 4283998.

BARAHHOU, I., FERREIRA, P., and MAALEJ, Y. (2023). A Framework to Align Sovereign Bond Portfolios with Net Zero Trajectories. *SSRN*, 4515462.

BARBER, B. M., MORSE, A., and YASUDA, A. (2021). Impact Investing. *Journal of Financial Economics*, 139(1), pp. 162-185.

BARKO, T., CREMERS, M., and RENNEBOOG, L. (2022). Shareholder Engagement on Environmental, Social, and Governance Performance. *Journal of Business Ethics*, 180(2), pp. 777-812.

BARNETT, M. L., and SALOMON, R. M. (2006). Beyond Dichotomy: The Curvilinear Relationship between Social Responsibility and Financial Performance. *Strategic Management Journal*, 27(11), pp. 1101-1122.

Basel Committee on Banking Supervision (2022). Principles for the Effective Management and Supervision of Climate-related Financial Risks. *Guidelines*, 15 June 2022.

BATHIANY, S., NOTZ, D., MAURITSEN, T., RAEDEL, G., and BROVKIN, V. (2016). On the Potential for Abrupt Arctic Winter Sea Ice Loss. *Journal of Climate*, 29(7), pp. 2703-2719.

BATISTA, M., and KARAWIA, A. (2009). The Use of the Sherman-Morrison-Woodbury Formula to Solve Cyclic Block Tri-diagonal and Cyclic Block Penta-diagonal Linear Systems of Equations. *Applied Mathematics and Computation*, 210(2), pp. 558-563.

BAUER, R., KOEDIJK, K., and OTTEN, R. (2005). International Evidence on Ethical Mutual Fund Performance and Investment Style. *Journal of Banking & Finance*, 29(7), pp. 1751-1767.

BCG (2023). The Voluntary Carbon Market: 2022 Insights and Trends. *Report*, January, 21 pages.

Beabout, G. R., and Schmiesing, K. E. (2003). Socially Responsible Investing: An Application of Catholic Social Thought. *Logos: A Journal of Catholic Thought and Culture*, 6(1), pp. 63-99.

Bekjarovski, F., and Brière, M. (2018). Shareholder Activism: Why Should Investors Care?. *Amundi Discussion Paper*, 30.

Ben Dor, A., Guan, J., Kelleher, A., Lauretig, A., Preclaw, R., and Zeng, X. (2022). ESG and Alternative Data: Capturing Corporates' Sustainability-Related Activities with Job Postings. *Journal of Financial Data Science*, 4(1), pp. 130-144.

Ben Slimane, M. (2021). Bond Index Tracking with Genetic Algorithms. *ResearchGate*, 350756323.

Ben Slimane, M., and De Jong, M. (2017). Bond Liquidity Scores. *Journal of Fixed Income*, 27(1), pp. 77-82.

Ben Slimane, M., Brard, E., Le Guenedal, T., Roncalli, T., and Sekine, T. (2019a). ESG Investing in Fixed Income: It's Time to Cross the Rubicon. *SSRN*, 3683477.

Ben Slimane, M., Le Guenedal, T., Roncalli, T., and Sekine, T. (2019b). ESG Investing in Corporate Bonds: Mind the Gap. *SSRN*, 3683472.

Ben Slimane, M., Da Fonseca, D., and Mahtani, V. (2020). Facts and Fantasies about the Green Bond Premium. *ResearchGate*, 348650560.

Ben Slimane, M., Roncalli, T., and Semet, R. (2023a). Green vs. Social Bond Premium. *SSRN*, 4448651.

Ben Slimane, M., Lucius, D., Roncalli, T., and Xu, J. (2023b). Net Zero Investment Portfolios — Part 2. The Core-Satellite Approach. *SSRN*, 4611418.

Ben Slimane, M., and Menchaoui, G. (2023c). Bond Portfolio Optimisation and Mixed Integer Programming. *ResearchGate*, 376687891.

Bennani, L., Le Guenedal, T., Lepetit, F., Ly, L., Mortier, V., Roncalli, T., and Sekine, T. (2018). How ESG Investing Has Impacted the Asset Pricing in the Equity Market. *SSRN*, 3316862.

Berg, F., Kölbel, J. F., and Rigobon, R. (2022). Aggregate Confusion: The Divergence of ESG Ratings. *Review of Finance*, 26(6), pp. 1315-1344.

Berrada, T., Engelhardt, L., Gibson, R., and Krueger, P. (2022). The Economics of Sustainability Linked Bonds. *Swiss Finance Institute Research Paper*, 22-26.

Billio, M., Costola, M., Hristova, I., Latino, C., and Pelizzon, L. (2021). Inside the ESG Ratings: (Dis)agreement and Performance. *Corporate Social Responsibility and Environmental Management*, 28(5), pp. 1426-1445.

Bingler, J. A., Kraus, M., Leippold, M., and Webersinke, N. (2022). Cheap Talk and Cherry-Picking: What ClimateBert has to Say on Corporate Climate Risk Disclosures. *Finance Research Letters*, 47(B), 102776, June.

BLACK, F. (1972). Capital Market Equilibrium with Restricted Borrowing. *Journal of Business*, 45(3), pp. 444-455.

BLACK, F. and LITTERMAN, R. B. (1991). Asset Allocation: Combining Investor Views with Market Equilibrium. *Journal of Fixed Income*, 1(2), pp. 7-18.

BLACK, F. and LITTERMAN, R. B. (1992). Global Portfolio Optimization. *Financial Analysts Journal*, 48(5), pp. 28-43.

Board of Governors of the Federal Reserve System (2011). Supervisory Guidance of Model Risk Management. Joint Publication with the Office of the Comptroller of the Currency. *Supervision and Regulation Letters*, SR 11-7, April.

Board of Governors of the Federal Reserve System (2021). Interagency Statement on Model Risk Management for Bank Systems Supporting Bank Secrecy Act/Anti-Money Laundering Compliance. Joint Publication with the Federal Deposit Insurance Corporation andthe Office of the Comptroller of the Currency. *Interagency Statement*, April.

BOFFO, R., and PATALANO, R. (2020). ESG Investing: Practices, Progress and Challenges, *Report*, OECD Paris.

BOLIN, B., and ERIKSSON, E. (1958). Changes in the Carbon Dioxide Content of the Atmosphere and Sea due to Fossil Fuel Combustion. In Bolin, B. (Ed.), *The Atmosphere and the Sea in Motion: Scientific Contributions to the Rossby Memorial Volume*, Rockefeller Institute Press, pp. 130-142.

BOLTON, P., and KACPERCZYK, M. (2021). Do Investors Care about Carbon Risk?. *Journal of Financial Economics*, 142(2), pp. 517-549.

BOLTON, P., KACPERCZYK, M., and SAMAMA, F. (2022). Net-zero Carbon Portfolio Alignment. *Financial Analysts Journal*, 78(2), pp. 19-33.

BOSETTI, V., CARRARO, C., GALEOTTI, M., MASSETTI, E., and TAVONI, M. (2006). A World Induced Technical Change Hybrid Model. *Energy Journal*, 27, Special Issue #2, pp. 13-37.

BOUYÉ, E., and MENVILLE, D. (2021). The Convergence of Sovereign Environmental, Social and Governance Ratings. *Policy Research Working Paper*, World Bank, 9583.

BOWEN, H. R. (1953). *Social Responsibilities of the Businessman*. Harper & Row.

BOWKER, G. C. (2000). Biodiversity Datadiversity. *Social Studies of Science*, 30(5), pp. 643-683.

BOYD, S., PARIKH, N., CHU, E., PELEATO, B., and ECKSTEIN, J. (2010). Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers. *Foundations and Trends® in Machine learning*, 3(1), pp. 1-122.

BRANDER, M., ASCUI, F., SCOTT, V., and TETT, S. (2021). Carbon Accounting for Negative Emissions Technologies. *Climate Policy*, 21(5), pp. 699-717.

BRIÈRE, M., KEIP, M., and LE BERTHE, T. (2022). Artificial Intelligence for Sustainable Finance: Why it may Help. *Amundi Institute Research Paper*, October 2022.

BRIÈRE, M., POUGET, S., and URECHE-RANGAU, L. (2020). BlackRock vs Norway Fund at Shareholder Meetings: Institutional Investors' Votes on Corporate Externalities. in Bank for International Settlements (Ed.), *Evolving Practices in Public Investment Management*, Proceedings of the Seventh Public Investors Conference, pp. 81-108.

BRIÈRE, M., and RAMELLI, S. (2021). Green Sentiment, Stock Returns, and Corporate Behavior. *SSRN*, 3850923.

BROCCARDO, E., HART, O. D., and ZINGALES, L. (2022). Exit versus Voice. *Journal of Political Economy*, 130(12), pp. 3101-3145.

BROECKER, W. S. (1975). Climatic Change: Are we on the Brink of a Pronounced Global Warming?. *Science*, 189(4201), pp. 460-463.

BUDYKO, M. I. (1969). The Effect of Solar Radiation Variations on the Climate of the Earth. *Tellus*, 21(5), pp. 611-619.

BURKE, M., DRISCOLL, A., LOBELL, D. B., and ERMON, S. (2021). Using Satellite Imagery to Understand and Promote Sustainable Development. *Science*, 371(6535), March.

BUTCHART, S. H. M., RESIT AKÇAKAYA, H., CHANSON, J., ..., and HILTON-TAYLOR, C. (2007). Improvements to the Red List Index. *PloS One*, 2(1), e140.

CALEL, R., and STAINFORTH, D. A. (2017a). On the Physics of Three Integrated Assessment Models. *Bulletin of the American Meteorological Society*, 98(6), pp. 1199-1216.

CALEL, R., and STAINFORTH, D. A. (2017b). On the Physics of Three Integrated Assessment Models: Supplement. *Internet Supplementary Materials*, https://doi.org/10.1175/BAMS-D-16-0034.1.

CALLENDAR, G. S. (1938). The Artificial Production of Carbon Dioxide and its Influence on Temperature. *Quarterly Journal of the Royal Meteorological Society*, 64(275), pp. 223-240.

CALVIN, K., PATEL, P., CLARKE, L., ..., and WISE, M. (2019). GCAM v5.1: Representing the Linkages between Energy, Water, Land, Climate, and Economic Systems. *Geoscientific Model Development*, 12, 677-698.

CAPELLE-BLANCARD, G., and MONJON, S. (2014). The Performance of Socially Responsible Funds: Does the Screening Process Matter?. *European Financial Management*, 20(3), pp. 494-520.

CARDINALE, B. J., DUFFY, J. E., GONZALEZ, A., ..., and NAEEM, S. (2012). Biodiversity Loss and its Impact on Humanity. *Nature*, 486(7401), pp. 59-67.

CAREY, M. (2010). *In the Shadow of Melting Glaciers: Climate Change and Andean Society*. Oxford University Press.

CARNEY, M. (2015). Breaking the Tragedy of the Horizon — Climate Change and Financial Stability. *Speech given at Lloyd's of London*, 29 September 2015.

CARNEY, M. (2019). Fifty Shades of Green. *Finance and Development*, 56(4), pp. 12-15.

CARROLL, A. B. (1999). Corporate Social Responsibility: Evolution of a Definitional Construct. *Business & Society*, 38(3), pp. 268-295.

CARROLL, A. B., and SHABANA, K. M. (2010). The Business Case for Corporate Social Responsibility: A Review of Concepts, Research and Practice. *International Journal of Management Reviews*, 12(1), pp. 85-105.

CASEAU, C., and GROLLEAU, G. (2020). Impact Investing: Killing two Birds with one Stone?. *Financial Analysts Journal*, 76(4), pp. 40-52.

CATLING, D. C., and ZAHNLE, K. J. (2020). The Archean Atmosphere. *Science Advances*, 6(9):eaax1420, pp. 1-16.

CDP (2022). Accounting of Scope 2 Emissions — CDP Climate Change Questionnaire. *Technical Note*, Version 9.0, March.

CHAFIK, L., HOLLIDAY, N. P., BACON, S., and ROSSBY, T. (2022). Irminger Sea is the Center of Action for Subpolar AMOC Variability. *Geophysical Research Letters*, 49(17), e2022GL099133, 11 pages.

CHAMBERS, D., DIMSON, E., and QUIGLEY, E. (2020). To Divest or to Engage? A Case Study of Investor Responses to Climate Activism. *Journal of Investing*, 29(2), pp. 10-20.

CHARNAY, B., WOLF, E. T., MARTY, B., and FORGET, F. (2020). Is the Faint Young Sun Problem for Earth Solved?. *Space Science Reviews*, 216(90), pp. 1-29.

CHARNEY, J. G., ARAKAWA, A., BAKER, D. J., BOLIN, B., DICKINSON, R. E., GOODY, R. M., LEITH, C. E., STOMMEL, H. M., and WUNSCH, C. I. (1979). *Carbon Dioxide and Climate: A Scientific Assessment*. US National Academy of Sciences, 22 pages.

CHARNEY, J. G., FJÖRTOFT, R., and VON NEUMANN, J. (1950). Numerical Integration of the Barotropic Vorticity Equation. *Tellus*, 2(4), pp. 237-254.

CHATTERJI, A. K., DURAND, R., LEVINE, D. I., and TOUBOUL, S. (2016). Do Ratings of Firms Converge? Implications for Managers, Investors and Strategy Researchers. *Strategic Management Journal*, 37(8), pp. 1597-1614.

CHEN, T., DONG, H., and LIN, C. (2020). Institutional Shareholders and Corporate Social Responsibility. *Journal of Financial Economics*, 135(2), pp. 483-504.

CHEN, Z., VAN KHOA, L. D., TEOH, E. N., NAZIR, A., KARUPPIAH, E. K., and LAM, K. S. (2018). Machine Learning Techniques for Anti-money Laundering (AML) Solutions in Suspicious Transaction Detection: A Review. *Knowledge and Information Systems*, 57(2), pp. 245-285.

CHERTOW, M. R. (2000). The IPAT Equation and its Variants. *Journal of Industrial Ecology*, 4(4), pp. 13-29.

CHIQUIER, S., PATRIZIO, P., BUI, M., SUNNY, N., and MAC DOWELL, N. (2022). A Comparative Analysis of the Efficiency, Timing, and Permanence of CO_2 Removal Pathways. *Energy & Environmental Science*, 15(10), pp. 4389-4403.

Climate Bonds Initiative (2019). Climate Bonds Standard Version 3.0. *Guidelines*, December.

Climate Bonds Initiative (2022a). Sustainable Debt — Global State of the Market 2021. *Report*, April.

Climate Bonds Initiative (2022b). Sustainable Debt Market — Summary H1 2022. *Report*, August.

Climate Bonds Initiative (2022c). Climate Bonds Standard Version 4.0. *Draft for Public Consultation*, September.

Climate Disclosure Standards Board (2021a). Application Guidance for Water-related Disclosures. *CDSB Framework*, August 2021.

Climate Disclosure Standards Board (2021b). Application Guidance for Biodiversity-related Disclosures. *CDSB Framework*, November 2021.

Climate Disclosure Standards Board (2021c). TCFD Good Practice Handbook. *Report*, November 2021.

CLINE, W. R. (1992). *The Economics of Global Warming*. Institute for International Economics.

COHEN, A. J., BRAUER, M., BURNETT, R., ..., and FOROUZANFAR, M. H. (2017). Estimates and 25-year Trends of the Global Burden of Disease Attributable to Ambient Air Pollution: An Analysis of Data from the Global Burden of Diseases Study 2015. *The Lancet*, 389(10082), pp. 1907-1918.

COLMAN, R., and SODEN, B. J. (2021). Water Vapor and Lapse Rate Feedbacks in the Climate System. *Reviews of Modern Physics*, 93(4), 045002, 90 pages.

Commissariat Général au Développement Durable (2020). LâĂŹempreinte carbone des Français reste stable. *Report*, 4 pages, January.

COQUERET, G. (2022). *Perspectives in Sustainable Equity Investing*. CRC Press.

CORTEZ, M. C., SILVA, F., and AREAL, N. (2009). The Performance of European Socially Responsible Funds. *Journal of Business Ethics*, 87(4), pp. 573-588.

COSTANZA, R., D'ARGE, R., DE GROOT, R., ..., and VAN DEN BELT, M. (1997). The Value of the World's Ecosystem Services and Natural Capital. *Nature*, 387(6630), pp. 253-260.

COX, G. M., HALVERSON, G. P., STEVENSON, R. K., ..., and MACDONALD, F. A. (2016). Continental Flood Basalt Weathering as a Trigger for Neoproterozoic Snowball Earth. *Earth and Planetary Science Letters*, 446, pp. 89-99.

CRAIG, H. (1961). Isotopic Variations in Meteoric Waters. *Science*, 133(3465), pp. 1702-1703.

CROWTHER, D., and ARAS, G. (2008). *Corporate Social Responsibility*. Bookboon.

CRUTZEN, P. J. and STOERMER, E. F. (2000). The Anthropocene. IGBP Global Change News, 41, pp. 17-18.

DALY, H. E. (1968). On Economics as a Life Science. *Journal of Political Economy*, 76(3), pp. 392-406.

DANIEL, K. D., LITTERMAN, R. B., and WAGNER, G. (2016). Applying Asset Pricing Theory to Calibrate the Price of Climate Risk. *NBER*, 22795.

DANIEL, K. D., LITTERMAN, R. B., and WAGNER, G. (2019). Declining CO_2 Price Paths. *Proceedings of the National Academy of Sciences*, 116(42), pp. 20886-20891.

DANSGAARD, W. (1964). Stable Isotopes in Precipitation. *Tellus*, 16(4), pp. 436-468.

DAVIES, S. W., and VAN WESEP, E. D. (2018). The Unintended Consequences of Divestment. *Journal of Financial Economics*, 128(3), pp. 558-575.

DAUBANES, J. X., MITALI, S. F., and ROCHET, J-C. (2021). Why Do Firms Issue Green Bonds?. *Swiss Finance Institute Research Paper*, 21-97.

DAVIS, S. J., and CALDEIRA, K. (2010). Consumption-based Accounting of CO_2 Emissions. *Proceedings of the National Academy of Sciences*, 107(12), pp. 5687-5692.

DELMAS, R. J., ASCENCIO, J. M., and LEGRAND, M. (1980). Polar Ice Evidence that Atmospheric CO_2 20,000 yr BP was 50% of Present. *Nature*, 284(5752), pp. 155-157.

DEMARTINI, A. (2020). Provision of Non-financial Data: Mapping of Stakeholders, Products and Services. *AMF Risks and Trend Mapping*, December.

DENNIG, F., BUDOLFSON, M. B., FLEURBAEY, M., SIEBERT, A., and SOCOLOW, R. H. (2015). Inequality, Climate Impacts on the Future Poor, and Carbon Prices. *Proceedings of the National Academy of Sciences*, 112(52), pp. 15827-15832.

DERWALL, J., and KOEDIJK, K. (2009). Socially Responsible Fixed-income Funds. *Journal of Business Finance & Accounting*, 36(1-2), pp. 210-229.

DESCOMBES, P., WISZ, M. S., LEPRIEUR, F., ..., and PELLISSIER, L. (2015). Forecasted Coral Reef Decline in Marine Biodiversity Hotspots under Climate Change. *Global Change Biology*, 21(7), pp. 2479-2487.

DESNOS, B., LE GUENEDAL, T., MORAIS, P., and RONCALLI, T. (2023). From Climate Stress Testing to Climate Value-at-Risk: A Stochastic Approach. *SSRN*, 4497124.

DIAZ, D. B., and MOORE, F. C. (2017a). Quantifying the Economic Risks of Climate Change. *Nature Climate Change*, 7(11), pp. 774-782.

DIAZ, D. B., and MOORE, F. C. (2017b). Quantifying the Economic Risks of Climate Change: Supplementary Information. Available at www.nature.com/articles/nclimate3411.

DIETRICH, J. P., BODIRSKY, B. L., HUMPENÖDER, F., ..., and POPP, A. (2019). MAgPIE 4 — A Modular Open-source Framework for Modeling Global Land Systems. *Geoscientific Model Development*, 12(4), pp. 1299-1317.

DIETZENBACHER, E., LOS, B., STEHRER, R., TIMMER, M., and DE VRIES, G. (2013). The Construction of World Input-Output Tables in the WIOD Project. *Economic Systems Research*, 25(1), pp. 71-98.

DIMSON, E., KARAKAŞ, O., and LI, X. (2015). Active Ownership. *Review of Financial Studies*, 28(12), pp. 3225-3268.

DIMSON, E., KARAKAŞ, O., and LI, X. (2021). Coordinated Engagements. *SSRN*, 3209072.

DIMSON, E., KREUTZER, I., LAKE, R., SJO, H., and STARKS, L. (2013). Responsible Investment and the Norwegian Government Pension Fund Global. *Strategy Council 2013*, Report, 34 pages, 11 November 2013.

DOMMENGET, D. (2022). An Introduction to Climate Dynamics. *Lectures Notes for the Undergraduate Course in Climate Dynamics*, 408 pages.

DREI, A., LE GUENEDAL, T., LEPETIT, F., MORTIER, V., RONCALLI, T. and SEKINE, T. (2019). ESG Investing in Recent Years: New Insights from Old Challenges. *SSRN*, 3683469.

DRUCKER, P. F. (1954). *The Practice of Management*. Harper & Row.

DRUCKER, P. F. (1984). The New Meaning of Corporate Social Responsibility. *California Management Review*, 26(2), pp. 53-63.

DUCOULOMBIER, F. (2021). Understanding the Importance of Scope 3 Emissions and the Implications of Data Limitations. *Journal of Impact and ESG Investing*, 1(4), pp. 63-71.

ECCLES, R. G., and STROEHLE, J. C. (2018). Exploring Social Origins in the Construction of Environmental, Social and Governance Measures. *SSRN*, 3212685.

ECCLES, R. G., LEE, L. E., and STROEHLE, J. C. (2020). The Social Origins of ESG: An Analysis of Innovest and KLD. *Organization & Environment*, 33(4), pp. 575-596.

Ecosystem Marketplace (2023a). All in on Climate: The Role of Carbon Credits in Corporate Climate Strategies. *Report*, October, 36 pages.

Ecosystem Marketplace (2023b). State of the Voluntary Carbon Markets 2023. *Insights Report*, November, 28 pages.

EDMANS, A. (2009). Blockholder Trading, Market Efficiency, and Managerial Myopia. *Journal of Finance*, 64(6), pp. 2481-2513.

EDMANS, A. (2011). Does the Stock Market fully Value Intangibles? Employee Satisfaction and Equity Prices. *Journal of Financial Economics*, 101(3), pp. 621-640.

EDMANS, A. (2014). Blockholders and Corporate Governance. *Annual Review of Financial Economics*, 6, pp. 23-50.

EDMANS, A., LEVIT, D., and SCHNEEMEIER, J. (2022). Socially Responsible Divestment. European Corporate Governance Institute, Finance Working Paper, 823, April.

EHLERS, T., and PACKER, F. (2017). Green Bond Finance and Certification. *BIS Quarterly Review*, September, pp. 89-104.

EL GHOUL, S., GUEDHAMI, O., KWOK, C. C., and MISHRA, D. R. (2011). Does Corporate Social Responsibility Affect the Cost of Capital?. *Journal of Banking & Finance*, 35(9), pp. 2388-2406.

EPICA community members (2004). Eight Glacial Cycles from an Antarctic Ice Core. *Nature*, 429(6992), pp. 623-628.

European Banking Authority (2014). Regulatory Technical Standards on the Conditions for Assessing the Materiality of Extensions and Changes of Internal Approaches for Credit, Market and Operational Risk. *Technical Standards, Guidelines & Recommendations*, July.

European Central Bank (2022). 2022 Climate Risk Stress Test. *Report*, July 2022.

European Commission (2022a). Targeted Consultation on the Functioning of the ESG Ratings Market in the European Union and on the Consideration of ESG Factors in Credit Ratings. *Consultation Document*, April 2022.

European Commission (2022b). Targeted Consultation on the Functioning of the ESG Ratings Market in the European Union and on the Consideration of ESG Factors in Credit Ratings. *Summary Report*, August 2022.

European Financial Reporting Advisory Group (2022). EFRAG Public consultation on ESRS Exposure Drafts (Appendix II — CSRD Requirements for Standard Setting and their Coverage by the ESRS Exposure Drafts). *Cover Note for Public Consultation*, April.

European Securities and Markets Authority (2022). ESMA's Technical Advice to the European Commission on Integrating Sustainability Risks and Factors in MiFID II. *Final Report*, September 2022.

European Sustainable Investment Forum (2018). *2018 European SRI Study*. November.

Fabozzi, F. J., Ma, K. C., and Oliphant, B. J. (2008). Sin Stock Returns. *Journal of Portfolio Management*, 35(1), pp. 82-94.

Fama E. F., and French, K. R. (1993). Common Risk Factors in the Returns on Stocks and Bonds. *Journal of Financial Economics*, 33(1), pp. 3-56.

Fatica, S., Panzica, R., and Rancan, M. (2021). The Pricing of Green Bonds: Are Financial Institutions Special?. *Journal of Financial Stability*, 54, 100873.

Faulwasser, T. , Kellett, C. M., and Weller, S. R. (2018). *MPC-DICE: Model Predictive Control — Dynamic Integrated model of Climate and Economy*. Code source available at https://github.com/cmkellett/MPC-DICE.

Fawcett, T. (2006). An Introduction to ROC Analysis, *Pattern Recognition Letters*, 27(8), pp. 861-874.

Feenstra, R. C., Inklaar, R., and Timmer, M. P. (2015). The Next Generation of the Penn World Table. *American Economic Review*, 105(10), pp. 3150-3182.

Fender, I., McMorrow, M., Sahakyan, V., and Zulaica, O. (2019). Green Bonds: The Reserve Management Perspective. *BIS Quarterly Review*, September, pp. 49-53.

Feulner, G. (2012). The Faint Young Sun Problem. *Reviews of Geophysics*, 50(2), RG2006, 29 pages.

Financial Stability Board (2015). Financial Reforms — Progress on the Work Plan for the Antalya Summit. *Chair's Letter to G20 Finance Ministers and Central Bank Governors*.

Finster, M., Wood, M., and Raja, S. N. (2005). The Apgar Score has Survived the Test of Time. *Journal of the American Society of Anesthesiologists*, 102(4), pp. 855-857.

Flammer, C. (2021). Corporate Green Bonds. *Journal of Financial Economics*, 142(2), pp. 499-516.

Foster, G. L., Royer, D. L., and Lunt, D. J. (2017). Future Climate Forcing Potentially without Precedent in the last 420 Million Years. *Nature Communications*, 8, 14845, 8 pages.

Fourier, J. (1824). Remarques générales sur les températures du globe terrestre et des espaces planétaires. *Annales de Chimie et de Physique*, 27, pp. 136-167.

Frankignoul, C., and Hasselmann, K. (1977). Stochastic Climate Models, Part II Application to Sea-surface Temperature Anomalies and Thermocline Variability. *Tellus*, 29(4), pp. 289-305.

Frazzini, A., and Pedersen, L. H. (2014). Betting Against Beta. *Journal of Financial Economics*, 111(1), pp. 1-25.

Freeman, R. E. (2004). *Strategic Management: A Stakeholder Approach*. Pitman.

Freshfields Bruckhaus Deringer (2005). A Legal Framework for the Integration of Environmental, Social and Governance Issues into Institutional Investment. *UNEP Report*, October.

Friede, G., Busch, T., and Bassen, A. (2015). ESG and Financial Performance: Aggregated Evidence from more than 2000 Empirical Studies. *Journal of Sustainable Finance & Investment*, 5(4), pp. 210-233.

Friederich, D., Kaack, L. H., Luccioni, A., and Steffen, B. (2021). Automated Identification of Climate Risk Disclosures in Annual Corporate Reports. *arXiv*, 2108.01415.

Friedlingstein, P., O'Sullivan, M., Jones, M. W., ..., and Zheng, B. (2022). Global Carbon Budget 2022. *Earth System Science Data*, 14, pp. 4811-49001.

Friedman, M. (1962). *Capitalism and Freedom*. University of Chicago Press.

Friedman, M. (1970). The Social Responsibility of Business Is to Increase Its Profits. *New York Times Magazine*, September 13, Section SM, pp. 17 (reproduced in Zimmerli, W. Ch., Holzinger, M., and Richter, K. (Eds). (2007), *Corporate Ethics and Corporate Governance*, Springer, pp. 173-178).

Friedrich, O., Norris, R. D., and Erbacher, J. (2012). Evolution of Middle to Late Cretaceous Oceans — A 55 my Record of Earth's Temperature and Carbon Cycle. *Geology*, 40(2), pp. 107-110.

Fujimori, S., Hasegawa, T., and Masui, T. (2017). AIM/CGE V2.0: Basic Feature of the Model. In Fujimori, S., Kainuma, M., and Masui, T. (Eds), *Post-2020 Climate Action: Global and Asian Perspectives*, Springer, pp. 305-328.

Fuss, S., Lamb, W. F., Callaghan, ..., and Minx, J. C. (2018). Negative Emissions — Part 2: Costs, Potentials and Side Effects. *Environmental Research Letters*, 13(6), 063002, 47 pages.

Gabay, D., and Mercier, B. (1976). A Dual Algorithm for the Solution of Nonlinear Variational Problems via Finite Element Approximation. *Computers & Mathematics with Applications*, 2(1), pp. 17-40.

Gambhir, A., Butnar, I., Li, P. H., Smith, P., and Strachan, N. (2019). A Review of Criticisms of Integrated Assessment Models and Proposed Approaches to Address These, Through the Lens of BECCS. *Energies*, 12(9), 1747.

Gantchev, N. (2013). The Costs of Shareholder Activism: Evidence from a Sequential Decision Model. *Journal of Financial Economics*, 107(3), pp. 610-631.

Gantchev, N., Gredil, O. R., and Jotikasthira, C. (2019). Governance under the Gun: Spillover Effects of Hedge Fund Activism. *Review of Finance*, 23(6), pp. 1031-1068.

Gatti, L. V., Basso, L. S., Miller, J. B., ..., and Neves, R. A. (2021). Amazonia as a Carbon Source Linked to Deforestation and Climate Change. *Nature*, 595(7867), pp. 388-393.

Geczy, C., Jeffers, J. S., Musto, D. K., and Tucker, A. M. (2021). Contracts with (Social) Benefits: The Implementation of Impact Investing. *Journal of Financial Economics*, 142(2), pp. 697-718.

GEMECHU, E. D., BUTNAR, I., LLOP, M., and CASTELLS, F. (2014). Economic and Environmental effects of CO$_2$ Taxation: An Input-output Analysis for Spain. *Journal of Environmental Planning and Management*, 57(5), pp. 751-768.

GENEST, C., and MACKAY, J. (1986a). Copules Archimédiennes et Familles de Lois Bidimensionnelles dont les Marges sont données. *Canadian Journal of Statistics*, 14(2), pp. 145-159.

GENEST, C., and MACKAY, J. (1986b), The Joy of Copulas: Bivariate Distributions with Uniform Marginals, *American Statistician*, 40(4), pp. 280-283.

GERARD, B. (2019). ESG and Socially Responsible Investment: A Critical Review. *Beta*, 33(1), pp. 61-83.

GHG Protocol (2004). The Greenhouse Gas Protocol: A Corporate Accounting and Reporting Standard. *Standards*.

GHG Protocol (2011). The Value Chain (Scope 3) Standard. *Standards*.

GHG Protocol (2013). Technical Guidance for Calculating Scope 3 Emissions: Supplement to the Corporate Value Chain (Scope 3) Accounting & Reporting Standard. *Standards*.

GHG Protocol (2015). GHG Protocol Scope 2 Guidance: An Amendment to the GHG Protocol Corporate Standard. *Standards*.

GIANFRATE, G., and PERI, M. (2019). The Green Advantage: Exploring the Convenience of Issuing Green Bonds. *Journal of Cleaner Production*, 219, pp. 127-135.

GIBBARD, P., WALKER, M., BAUER, A., ..., and RUDDIMAN, W. (2022). The Anthropocene as an Event, not an Epoch. *Journal of Quaternary Science*, 37(3), pp. 395-399.

GILLAN, S. L., and STARKS, L. T. (2000). Corporate Governance Proposals and Shareholder Activism: The Role of Institutional Investors. *Journal of Financial Economics*, 57(2), pp. 275-305.

GILLINGHAM, K., and STOCK, J. H. (2018). The Cost of Reducing Greenhouse Gas Emissions. *Journal of Economic Perspectives*, 32(4), pp. 53-72.

GINI, C. (1921). Measurement of Inequality of Incomes. *Economic Journal*, 31(121), pp. 124-126.

GIRONA, M. M., MORIN, H., GAUTHIER, S., and BERGERON, Y. (Eds) (2023). *Boreal Forests in the Face of Climate Change: Sustainable Management*. Springer Nature, Advances in Global Change Research, 14, 837 pages.

Global Sustainable Investment Alliance (2013). *2012 Global Sustainable Investment Review*. January.

Global Sustainable Investment Alliance (2015). *2014 Global Sustainable Investment Review*. February.

Global Sustainable Investment Alliance (2017). *2016 Global Sustainable Investment Review*. March.

Global Sustainable Investment Alliance (2019). *2018 Global Sustainable Investment Review*. April.

Global Sustainable Investment Alliance (2021). *2020 Global Sustainable Investment Review*. July.

GOLDREYER, E. F., AHMED, P., and DILTZ, J. D. (1999). The Performance of Socially Responsible Mutual Funds: Incorporating Sociopolitical Information in Portfolio Selection. *Managerial Finance*, 25(1), pp. 23-36.

GOLLIER, C. (2010). Ecological Discounting. *Journal of Economic Theory*, 145(2), pp. 812-829.

GOLLIER, C. (2013). *Pricing the Planet's Future: The Economics of Discounting in an Uncertain World*. Princeton University Press.

GOLUB, G. H., and VAN LOAN, C. F. (2013). *Matrix Computations*. Fourth edition, Johns Hopkins University Press.

GOMPERS, P., ISHII, J., and METRICK, A. (2003). Corporate Governance and Equity Prices. *Quarterly Journal of Economics*, 118(1), pp. 107-156.

GOOSSE, H., KAY, J. E., ARMOUR, K. C., ..., and VANCOPPENOLLE, M. (2018). Quantifying Climate Feedbacks in Polar Regions. *Nature Communications*, 9(1), 1919, 13 pages.

GOURIÉROUX, C. (1992). Courbes de Performance, de Sélection et de Discrimination. *Annales d'Économie et de Statistique*, 28(4), pp. 107-123.

GOURIÉROUX, C., and JASIAK, J. (2007). *The Econometrics of Individual Risk: Credit, Insurance, and Marketing*. Princeton University Press.

GRATCHEVA, E. M., EMERY, T., and WANG, D. (2020). Demystifying Sovereign ESG. *Report*, World Bank, Equitable Growth, Finance and Institutions Insight, December.

GREWAL, J., SERAFEIM, G., and YOON, A. (2016). Shareholder Activism on Sustainability Issues. *SSRN*, 2805512.

GROSSMAN, B. R., and SHARPE, W. F. (1986). Financial Implications of South African Divestment. *Financial Analysts Journal*, 42(4), pp. 15-29.

GRUBB, M., WIENERS, C., and YANG, P. (2021). Modeling Myths: On DICE and Dynamic Realism in Integrated Assessment Models of Climate Change Mitigation. *WIREs Climate Change*, 12(3), e698.

GUTIERREZ, M. S. (2008). *Economic Activity and Atmospheric Pollution in Spain: An Input-output Approach*. PhD Thesis, Universitat de Barcelona.

HANEMANN, W. M. (2008). What is the Economic Cost of Climate Change?. *Working Paper*, University of California, Berkeley, 1071.

HÄNSEL, M. C., DRUPP, M. A., JOHANSSON, D. J., ..., and STERNER, T. (2020). Climate Economics Support for the UN Climate Targets. *Nature Climate Change*, 10(8), pp. 781-789.

HANSEN, J., LACIS, A., RIND, D., RUSSELL, G., STONE, P., FUNG, I., RUEDY, R., and LERNER, J. (1984). Climate Sensitivity: Analysis of Feedback Mechanisms. In Hansen, J. E. and Takahashi, T. (Eds), *Climate Processes and Climate Sensitivity*, 29, pp. 130-163.

HANSEN, J., SATO, M., RUSSELL, G., and KHARECHA, P. (2013). Climate Sensitivity, Sea Level and Atmospheric Carbon Dioxide. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 371(2001), 20120294, 31 pages.

HASSELMANN, K. (1976). Stochastic Climate Models, Part I. Theory. *Tellus*, 28(6), pp. 473-485.

HARTMANN, D. L. (2016). *Global Physical Climatology*. Second edition, Elsevier Science.

HARVEY, A. C. (1990). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press.

HASTIE, T., TIBSHIRANI, R., and FRIEDMAN, J.H. (2009). *The Elements of Statistical Learning*. Second edition, Springer.

HAVLÌK, P., VALIN, H., MOSNIER, A., ..., and HASEGAWA, T. (2018). *GLOBIOM Documentation*. International Institute for Applied Systems Analysis.

HAYS, J. D., IMBRIE, J., and SHACKLETON, N. J. (1976). Variations in the Earth's Orbit: Pacemaker of the Ice Ages. *Science*, 194(4270), pp. 1121-1132.

HERSBACH, H., BELL, B., BERRISFORD, P., ..., and THẼᴌPAUT, J. N. (2020). The ERA5 Global Reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730), pp. 1999-2049.

High-Level Expert Group on Sustainable Finance (2018). Financing a Sustainable European Economy. *Final Report*, January.

HJORT, J., STRELETSKIY, D., DORÉ, G., WU, Q., BJELLA, K., and LUOTO, M. (2022). Impacts of Permafrost Degradation on Infrastructure. *Nature Reviews Earth & Environment*, 3(1), pp. 24-38.

HINKLEY, D. V. (1969). On the Ratio of Two Correlated Normal Random Variables. *Biometrika*, 56(3), pp. 635-639.

HIRSCHMAN, A. O. (1970). *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States*. Harvard University Press.

HOFFMAN, P. F., KAUFMAN, A. J., HALVERSON, G. P., and SCHRAG, D. P. (1998). A Neoproterozoic Snowball Earth. *Science*, 281(5381), pp. 1342-1346.

HOFFMAN, P. F., and SCHRAG, D. P. (2002). The Snowball Earth Hypothesis: Testing the Limits of Global Change. *Terra Nova*, 14(3), pp. 129-155.

HONG, H., and KACPERCZYK, M. (2009). The Price of Sin: The Effects of Social Norms on Markets. *Journal of Financial Economics*, 93(1), pp. 15-36.

HOPE, C. (2011). The PAGE09 Integrated Assessment Model: A Technical Description. *Cambridge Judge Business School Working Paper*, 4.

HOWARD, P. H., and STERNER, T. (2017). Few and not so Far Between: A Meta-analysis of Climate Damage Estimates. *Environmental and Resource Economics*, 68(1), pp. 197-225.

HUPPMANN, D., GIDDEN, M., FRICKO, ..., and KREY, V. (2019). The MESSAGEix Integrated Assessment Model and the ix Modeling Platform (ixmp): An Open Framework for Integrated and Cross-cutting Analysis of Energy, Climate, the Environment, and Sustainable Development. *Environmental Modelling & Software*, pp. 112, 143-156.

Interagency Working Group on Social Cost of Greenhouse Gases (2021). *Technical Support Document: Social Cost of Carbon, Methane, and Nitrous Oxide, Interim Estimates under Executive Order 13990*. United States Government, 2021.

Intergovernmental Panel on Climate Change (2006). 2006 IPCC Guidelines for National Greenhouse Gas Inventories. Prepared by the National Greenhouse Gas Inventories Programme, Eggleston H. S., Buendia L., Miwa K., Ngara T. and Tanabe K. (Eds).

Intergovernmental Panel on Climate Change (2007). Climate Change 2007: The Physical Science Basis — Contribution of Working Group I to the Fourth Assessment Report of the IPCC. *Report.*

Intergovernmental Panel on Climate Change (2013). Climate Change 2013: The Physical Science Basis — Contribution of Working Group I to the Fifth Assessment Report of the IPCC. *Report.*

Intergovernmental Panel on Climate Change (2014). Climate Change 2014: Mitigation of Climate Change — Contribution of Working Group III to the Fifth Assessment Report of the IPCC. *Report.*

Intergovernmental Panel on Climate Change (2018). Global Warming of 1.5°C. *Special Report.*

Intergovernmental Panel on Climate Change (2019). *2019 Refinement to the 2006 IPCC Guidelines for National Greenhouse Gas Inventories.* Prepared by Calvo Buendia, E., Tanabe, K., Kranjc, A., Baasansuren, J., Fukuda, M., Ngarize, S., Osako, A., Pyrozhenko, Y., Shermanau, P. and Federici, S. (Eds).

Intergovernmental Panel on Climate Change (2021). Climate Change 2021: The Physical Science Basis — Contribution of Working Group I to the Sixth Assessment Report of the IPCC. *6th Assesment Report.*

Intergovernmental Panel on Climate Change (2022). Climate Change 2022: Mitigation of Climate Change — Contribution of Working Group III to the Sixth Assessment Report of the IPCC. *Report.*

International Association of Insurance Supervisors (2021). Supervision of Climate-related Risks in the Insurance Sector. *Application Paper*, May 2021.

International Capital Market Association (2021a). Green Bond Principles — Voluntary Process Guidelines for Issuing Green Bonds. *Guidelines*, June.

International Capital Market Association (2021b). Social Bond Principles — Voluntary Process Guidelines for Issuing Social Bonds. *Guidelines*, June.

International Capital Market Association (2022). Green, Social and Sustainability Bonds: A High-Level Mapping to the Sustainable Development Goals. *Report*, June 2022.

International Energy Agency (2017). Energy Technology Perspectives 2017: Catalysing Energy Technology Transformations. *Flagship Report*, June.

International Energy Agency (2021). Net Zero by 2050: A Roadmap for the Global Energy Sector. *Report*, July.

International Energy Agency (2022). Direct Air Capture: A Key Technology for Net Zero. *Report*, April, 76 pages.

International Finance Corporation (2020). Green Bond Handbook: A Step-by-step Guide to Issuing a Green Bond. *Guidelines.*

International Finance Corporation (2022). Blue Finance — Guidance for Financing the Blue Economy, Building on the Green Bond Principles and the Green Loan Principles. *Guidelines*, January.

International Organization of Securities Commissions (2021). Environmental, Social and Governance (ESG) Ratings and Data Products Providers. *Report*, November 2021.

International Sustainability Standards Board (2022a). IFRS S1 General Requirements for Disclosure of Sustainability-related Financial Information. *Exposure Draft*, March 2022.

International Sustainability Standards Board (2022b). IFRS S2 Climate-related Disclosures. *Exposure Draft*, March 2022.

ISRAEL, R.B., ROSENTHAL, J.S., and WEI, J.Z. (2001). Finding Generators for Markov Chains via Empirical Transition Matrices, with Applications to Credit Ratings. *Mathematical Finance*, 11(2), pp. 245-265.

ISS Governance (2022). Climate & Voting: 2021 Review and Global Trends. *Insights Report*, May.

IUCN. 2022. The IUCN Red List of Threatened Species. *Version 2022-1*, https://www.iucnredlist.org.

JAHNKE, P. (2019). Ownership Concentration and Institutional Investors' Governance through Voice and Exit. *Business and Politics*, 21(3), pp. 327-350.

JARROW, R.A., LANDO, D., and TURNBULL, S.M. (1997). A Markov Model for the Term Structure of Credit Risk Spreads. *Review of Financial Studies*, 10(2), pp. 481-523.

JENSEN, M. C. (1968). The Performance of Mutual Funds in the Period 1945-1964. *Journal of Finance*, 23(2), pp. 389-416.

JENSEN, M. C. (1993). The Modern Industrial Revolution, Exit, and the Failure of Internal Control Systems. *Journal of Finance*, 48(3), pp. 831-880.

JENSEN, M. C., and MECKLING, W. H. (1976). Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure. *Journal of Financial Economics*, 3(4), pp. 305-360.

JEEVANJEE, N. (2023). Climate Sensitivity from Radiative-Convective Equilibrium: A Chalkboard Approach. *American Journal of Physics*, 91(9), pp. 731-745.

JHA, A., and COX, J. (2015). Corporate Social Responsibility and Social Capital. *Journal of Banking & Finance*, 60, pp. 252-270.

JOE, H. (1997). *Multivariate Models and Dependence Concepts*. Monographs on Statistics and Applied Probability, 73, Chapman & Hall.

JOHNSON, S. A., MOORMAN, T. C., and SORESCU, S. (2009). A Reexamination of Corporate Governance and Equity Prices. *Review of Financial Studies*, 22(11), pp. 4753-4786.

JONDEAU, E., MOJON, B., and PEREIRA DA SILVA, L. A. (2021). Building Benchmark Portfolios with Decreasing Carbon Footprints. *Swiss Finance Institute Research Paper*, 21-91.

JONES, M. W., PETERS, G. P., GASSER, T., ..., and LE QUÉRÉ, C. (2023). National Contributions to Climate Change due to Historical Emissions of Carbon Dioxide, Methane, and Nitrous Oxide since 1850. *Scientific Data*, 10(1), 23 pages.

JONES, T. M. (1980). Corporate Social Responsibility Revisited, Redefined. *California Management Review*, 22(3), pp. 59-67.

JOOS, F., ROTH, R., FUGLESTVEDT, J. S., ..., and WEAVER, A. J. (2013). Carbon Dioxide and Climate Impulse Response Functions for the Computation of Greenhouse Gas Metrics: A Multimodel Analysis. *Atmospheric Chemistry and Physics*, 13(5), pp. 2793-2825.

JOPPA, L., LUERS, A., WILLMOTT, E., FRIEDMANN, S. J., HAMBURG, S. P., and BROZE, R. (2021). Microsoft's Million-Tonne CO2-Removal Purchase — Lessons for Net Zero. *Nature*, 597, pp. 629-632.

JORION, P. (2003). Portfolio Optimization with Tracking-Error Constraints. *Financial Analysts Journal*, 59(5), pp. 70-82.

JOUZEL, J. (2013). A Brief History of Ice Core Science over the last 50 yr. *Climate of the Past*, 9(6), pp. 2525-2547.

JOUZEL, J., MASSON-DELMOTTE, V., CATTANI, O., ..., and WOLFF, E. W. (2007). Orbital and Millennial Antarctic Climate Variability over the Past 800,000 Years. *Science*, 317(5839), pp. 793-796.

KAPLAN, R. S., RAMANNA, K., and ROSTON, M. (2023). Accounting for Carbon Offsets — Establishing the Foundation for Carbon-trading Markets. *Harvard Business School Working Paper*, 23-050, 16 pages.

KAPRAUN, J., LATINO, C., SCHEINS, C., and SCHLAG, C. (2021). (In)-credibly Green: Which Bonds Trade at a Green Bond Premium?. *SSRN*, 3347337.

KAUFMANN, D., KRAAY, A., and MASTRUZZI, M. (2010). The Worldwide Governance Indicators: Methodology and Analytical Issues. *World Bank Policy Research Working Paper*, 5430 (SSRN, 1682130).

KAVVATHAS, D. (2001). Estimating Credit Rating Transition Probabilities for Corporate Bonds. *SSRN*, 248421.

KAYA, Y., and YOKOBORI, K. (Eds) (1997). *Environment, Energy, and Economy: Strategies for Sustainability*. United Nations University Press, 393 pages.

KEELING, C. D. (1960). The Concentration and Isotopic Abundances of Carbon Dioxide in the Atmosphere. *Tellus*, 12(2), pp. 200-203.

KEELING, C. D. (1978). The Influence of Mauna Loa Observatory on the Development of Atmospheric CO_2 Research. In Miller, J (Ed.), *Mauna Loa Observatory: A 20th Anniversary Report*, National Oceanic and Atmospheric Administration Special Report, pp. 36-54.

KEELING, C. D., PIPER, S. C., BACASTOW, R. B., WAHLEN, M., WHORF, T. P., HEIMANN, M., and MEIJER, H. A. (2001). Exchanges of Atmospheric CO_2 and $^{13}CO_2$ with the Terrestrial Biosphere and Oceans from 1978 to 2000. I. Global aspects. *Scripps Institution of Oceanography, San Diego, SIO Reference Series*, 01-06, 28 pages.

KELLETT, C. M., FAULWASSER, T., and WELLER, S. R. (2016). *DICE-2013R-mc — A Matlab/CasADi Implementation of Vanilla DICE-2013R*. Code source available at https://github.com/cmkellett/DICE2013R-mc.

KELLETT, C. M., WELLER, S. R., FAULWASSER, T., GRÜNE, L., and SEMMLER, W. (2019). Feedback, Dynamics, and Optimal Control in Climate Economics. *Annual Reviews in Control*, 47, pp. 7-20.

KELLY, D. L., and KOLSTAD, C. D. (1999). Integrated Assessment Models for Climate Change Control. In Folmer, H. and Tietenberg, T. (Eds), *International Yearbook of Environmental and Resource Economics 1999/2000*, Edward Elgar, pp. 171-197.

KEMP, L., XU, C., DEPLEDGE, J., ..., and LENTON, T. M. (2022). Climate Endgame: Exploring Catastrophic Climate Change Scenarios. *Proceedings of the National Academy of Sciences*, 119(34), e2108146119, 9 pages.

KIEHL, J. T., and TRENBERTH, K. E. (1997). Earth's Annual Global Mean Energy Budget. *Bulletin of the American Meteorological Society*, 78(2), pp. 197-208.

KITZES, J. (2013). An Introduction to Environmentally-extended Input-output Analysis. *Resources*, 2(4), pp. 489-503.

KLEINBERG, R. L. (2020). The Global Warming Potential Misrepresents the Physics of Global Warming Thereby Misleading Policy Makers. *EarthArXiv*, November.

KLOSE, A. K., KARLE, V., WINKELMANN, R., and DONGES, J. F. (2020). Emergence of Cascading Dynamics in Interacting Tipping Elements of Ecology and Climate. *Royal Society Open Science*, 7(6), 200599, 18 pages.

KÖLBEL, J. F., LEIPPOLD, M., RILLAERTS, J., and WANG, Q. (2022). Ask BERT: How Regulatory Disclosure of Transition and Physical Climate Risks Affects the CDS Term Structure. *Swiss Finance Institute Research Paper*, 21-19.

KÖPPL, A., and SCHRATZENSTALLER, M. (2023). Carbon Taxation: A Review of the Empirical Literature. *Journal of Economic Surveys*, 37(4), pp. 1353-1388.

KOSTANT, P. C. (1999). Exit, Voice and Loyalty in the Course of Corporate Governance and Counsel's Changing Role. *Journal of Socio-Economics*, 28(3), pp. 203-246.

KRUEGER, P., SAUTNER, Z., and STARKS, L. T. (2020). The Importance of Climate Risks for Institutional Investors. *Review of Financial Studies*, 33(3), pp. 1067-1111.

KVENVOLDEN, K. A. (1988). Methane Hydrates and Global Climate. *Global Biogeochemical Cycles*, 2(3), pp. 221-229.

LAMB, W. F., WIEDMANN, T., PONGRATZ, J., ..., and MINX, J. (2021). A Review of Trends and Drivers of Greenhouse Gas Emissions by Sector from 1990 to 2018. *Environmental Research Letters*, 16(7), 073005.

LANGWAY, C. C. (2008). The History of Early Polar Ice Cores. *Cold Regions Science and Technology*, 52(2), pp. 101-117.

LARCKER, D. F., and WATTS, E. M. (2020). Where's the Greenium?. *Journal of Accounting and Economics*, 69(2-3), 101312.

LE GUENEDAL, T. (2019). Economic Modeling of Climate Risks. *SSRN*, 3693661.

LE GUENEDAL, T., DROBINSKI, P., and TANKOV, P. (2021). Measuring and Pricing Cyclone-Related Physical Risk Under Changing Climate. *SSRN*, 3850673.

LE GUENEDAL, T., DROBINSKI, P., and TANKOV, P. (2022). Cyclone Generation Algorithm including a Thermodynamic Module for Integrated National Damage Assessment (CATHERINA 1.0) compatible with Coupled Model Intercomparison Project (CMIP) climate data. *Geoscientific Model Development*, 15(21), pp. 8001-8039.

LE GUENEDAL, T., LOMBARD, F, RONCALLI, T., and SEKINE, T. (2022). Net Zero Carbon Metrics. *SSRN*, 4033686.

LE GUENEDAL, T., and RONCALLI, T. (2022). Portfolio Construction with Climate Risk Measures. In Jurczenko, E. (Ed.), *Climate Investing: New Strategies and Implementation Challenges*, Wiley, December 2022, pp. 49-86.

LEITE, P., and CORTEZ, M. C. (2016). The Performance of European Socially Responsible Fixed-Income Funds. *SSRN*, 2726094.

LEMOINE, D. (2020). *DICE-2016R-Matlab*. Code source available at https://github.com/dlemoine1/DICE-2016R-Matlab.

LENSSEN, N. J., SCHMIDT, G. A., HANSEN, J. E., MENNE, M. J., PERSIN, A., RUEDY, R., and ZYSS, D. (2019). Improvements in the GISTEMP Uncertainty Model. *Journal of Geophysical Research: Atmospheres*, 124(12), pp. 6307-6326.

LENTON, T. M. (2011). Early Warning of Climate Tipping Points. *Nature Climate Change*, 1(4), pp. 201-209.

LENTON, T. M., HELD, H., KRIEGLER, E., HALL, J. W., LUCHT, W., RAHMSTORF, S., and SCHELLNHUBER, H. J. (2008). Tipping Elements in the Earth's Climate System. *Proceedings of the National Academy of Sciences*, 105(6), pp. 1786-1793.

LENTON, T. M., ROCKSTRÖM, J., GAFFNEY, O., RAHMSTORF, S., RICHARDSON, K., STEFFEN, W., and SCHELLNHUBER, H. J. (2019). Climate Tipping Points — Too Risky to Bet Against. *Nature*, 575(7784), pp. 592-595.

LEONTIEF, W. W. (1936). Quantitative Input and Output Relations in the Economic Systems of the United States. *Review of Economics and Statistics*, 18(3), pp. 105-125.

LEONTIEF, W. W. (1941). *Structure of American Economy, 1919-1929: An Empirical Application of Equilibrium Analysis*. Harvard University Press.

LEONTIEF, W. (1970). Environmental Repercussions and the Economic Structure: An Input-output Approach. *Review of Economics and Statistics*, 52(3), pp. 262-271.

LEVIT, D. (2019). Soft Shareholder Activism. *Review of Financial Studies*, 32(7), pp. 2775-2808.

LEWIS, S. L., and MASLIN, M. A. (2015). Defining the Anthropocene. *Nature*, 519(7542), pp. 171-180.

LEZMI, E., RONCALLI, T, and XU, J. (2022). Multi-Period Portfolio Optimization. *SSRN*, 4078043.

LHOTELLIER J., LESS, E. BOSSANNE, E., and PESNEL, S. (2017). Modélisation et évaluation du poids carbone de produits de consommation et biens d'équipements. *ADEME Report*, September.

LIANG, S., WANG, D., HE, T., and YU, Y. (2019). Remote Sensing of Earth's Energy Budget: Synthesis and Review. *International Journal of Digital Earth*, 12(7), pp. 737-780.

LIN, K., KABEL, A., PARKER, S., and JOYE, C. (2019). Are ESG Alpha and Beta Benefits in Corporate Bonds a Mirage?. *SSRN*, 3352950.

LISIECKI, L. E., and RAYMO, M. E. (2005). A Pliocene-Pleistocene Stack of 57 Globally Distributed benthic $\delta^{18}O$ Records, *Paleoceanography*, 20(1), 17 pages.

LIU, W., XIE, S. P., LIU, Z., and ZHU, J. (2017). Overlooked Possibility of a Collapsed Atlantic Meridional Overturning Circulation in Warming Climate. *Science Advances*, 3(1), e1601666, 7 pages.

LÖFFLER, K. U., PETRESKI, A., and STEPHAN, A. (2021). Drivers of Green Bond Issuance and New Evidence on the "Greenium". *Eurasian Economic Review*, 11(1), pp. 1-24.

LORENZ, M. O. (1905). Methods of Measuring the Concentration of Wealth. *Journal of the American Statistical Association*, 9(70), pp. 209-219.

MALTAIS, A., and NYKVIST, B. (2020). Understanding the Role of Green Bonds in Advancing Sustainability. *Journal of Sustainable Finance & Investment*, pp. 1-20.

MANABE, S., and WETHERALD, R. T. (1967). Thermal Equilibrium of the Atmosphere with a Given Distribution of Relative Humidity. *Journal of the Atmospheric Sciences*, 24(3), pp. 241-259.

MARDONES, C., and MENA, C. (2020). Economic, Environmental and Distributive Analysis of the Taxes to Global and Local Air Pollutants in Chile. *Journal of Cleaner Production*, 259, 120893.

MARGOLIS, J. D., ELFENBEIN, H. A., and WALSH, J. P. (2009). Does it Pay to be Good... and does it Matter? A Meta-analysis of the Relationship between Corporate Social and Financial Performance. *SSRN*, 1866371.

MARKOWITZ H. M. (1952), Portfolio Selection. *Journal of Finance*, 7(1), pp. 77-91.

MARKOWITZ H. M. (1956). The Optimization of a Quadratic Function Subject to Linear Constraints. *Naval Research Logistics Quarterly*, 3(1-2), pp. 111-133.

MATOS, P. (2020). *ESG and Responsible Institutional Investing around the World: A Critical Review.* CFA Institute Research Foundation.

MATSUSAKA, J. G., OZBAS, O., and YI, I. (2021). Can Shareholder Proposals Hurt Shareholders? Evidence from Securities and Exchange Commission No-action-letter Decisions. *Journal of Law and Economics*, 64(1), pp. 107-152.

MCCAHERY, J. A., SAUTNER, Z., and STARKS, L. T. (2016). Behind the Scenes: The Corporate Governance Preferences of Institutional Investors. *Journal of Finance*, 71(6), pp. 2905-2932.

McKinsey (2022). *The Net-zero Transition — What it Would Cost, What it Could Bring.* January 2022, 224 pages.

MEINSHAUSEN, M., RAPER, S. C. B., and WIGLEY, T. M. L. (2011). Emulating Coupled Atmosphere-Ocean and Carbon Cycle Models with a Simpler Model, MAGICC6 — Part 1: Model Description and Calibration. *Atmospheric Chemistry and Physics*, 11(4), pp. 1417-1456.

MENZ, K. M. (2010). Corporate Social Responsibility: Is it Rewarded by the Corporate Bond Market? A Critical Note. *Journal of Business Ethics*, 96(1), pp. 117-134.

MERLIVAT, L., and JOUZEL, J. (1979). Global Climatic Interpretation of the Deuterium-Oxygen 18 Relationship for Precipitation. *Journal of Geophysical Research: Oceans*, 84(C8), pp. 5029-5033.

MILLER, R. E., and BLAIR, P. D. (2009). *Input-output Analysis: Foundations and Extensions*. Second edition, Cambridge University Press.

MINX, J. C., WIEDMANN, T., WOOD, R., ..., and ACKERMAN, F. (2009). Input-output Analysis and Carbon Footprinting: An Overview of Applications. *Economic Systems Research*, 21(3), pp. 187-216.

MINTZBERG, H. (1983). The Case for Corporate Social Responsibility. *Journal of Business Strategy*, 4(2), pp. 3-13.

MINX, J. C., LAMB, W. F., CALLAGHAN, M. W., ..., and DEL MAR ZAMORA DOMINGUEZ, M. (2018). Negative Emissions — Part 1: Research Landscape and Synthesis. *Environmental Research Letters*, 13(6), 063001, 29 pages.

MITCHELL, R. K., AGLE, B. R., and WOOD, D. J. (1997). Toward a Theory of Stakeholder Identification and Salience: Defining the Principle of Who and What Really Counts. *Academy of Management Review*, 22(4), pp. 853-886.

MOLER, C., and VAN LOAN, C.F. (2003). Nineteen Dubious Ways to Compute the Exponential of a Matrix, Twenty-Five Years Later. *SIAM Review*, 45(1), pp. 3-49.

Moodys' (2022). Sovereign Default and Recovery Rates, 1983-2021. *Report*, April.

MOORE, F. C., RISING, J., LOLLO, N., ..., and ANTHOFF, D. (2018). Mimi-PAGE, An Open-source Implementation of the PAGE09 Integrated Assessment Model. *Scientific Data*, 5, 180187.

MOORE, G. W. K., VÅGE, K., RENFREW, I. A., and PICKART, R. S. (2022). Sea-ice Retreat Suggests Re-organization of Water Mass Transformation in the Nordic and Barents Seas. *Nature Communications*, 13(67), 8 pages.

MORICE, C. P., KENNEDY, J. J., RAYNER, N. A., ..., and SIMPSON, I. R. (2021). An Updated Assessment of Near-surface Temperature Change from 1850: The HadCRUT5 Data Set. *Journal of Geophysical Research: Atmospheres*, 126(3), 2019JD032361, 28 pages.

MOSKOWITZ, M. (1972). Choosing Socially Responsible Stocks. *Business and Society Review*, 1(1), pp. 71-75.

MSCI (2020). MSCI ESG Ratings Methodology. *Methodology Document*, April.

MSCI (2022). MSCI ESG Ratings Methodology. *Executive Summary*, June.

MURPHY, K. J. (1985). Corporate Performance and Managerial Remuneration: An Empirical Analysis. *Journal of Accounting and Economics*, 7(1-3), pp. 11-42.

National Academies of Sciences, Engineering, and Medicine (2019). *Negative Emissions Technologies and Reliable Sequestration: A Research Agenda*. The National Academy Press, 510 pages.

NBIM (2020). Shareholder Voting Process — Asset Management Perspective. *Report*, March.

NELSEN, R. B. (2006). *An Introduction to Copulas.* Second edition, Springer.

NEMET, G. F., CALLAGHAN, M. W., CREUTZIG, F., ..., and SMITH, P. (2018). Negative Emissions — Part 3: Innovation and Upscaling. *Environmental Research Letters*, 13(6), 063003, 31 pages.

NEWBOLD, S. C., and MARTEN, A. L. (2014). The Value of Information for Integrated Assessment Models of Climate Change. *Journal of Environmental Economics and Management*, 68(1), pp. 111-123.

NGFS (2022). NGFS Scenarios for Central Banks and Supervisors. *Report*, September.

NORDHAUS, W. D. (1977). Economic Growth and Climate: the Carbon Dioxide Problem. *American Economic Review*, 67(1), pp. 341-346.

NORDHAUS, W. D. (1992). An Optimal Transition Path for Controlling Greenhouse Gases. *Science*, 258(5086), pp. 1315-1319.

NORDHAUS, W. D. (2007). A Review of the Stern Review on the Economics of Climate Change. *Journal of Economic Literature*, 45(3), pp. 686-702.

NORDHAUS, W. D. (2008). *A Question of Balance: Weighing the Options on Global Warming Policies.* Yale University Press.

NORDHAUS, W. D. (2017a). Revisiting the Social Cost of Carbon. *Proceedings of the National Academy of Sciences*, 114(7), pp. 1518-1523.

NORDHAUS, W. D. (2017b). Integrated Assessment Models of Climate Change. *NBER Reporter*, 3, September, pp. 16-20.

NORDHAUS, W. D. (2018a). Evolution of Modeling of the Economics of Global Warming: Changes in the DICE Model, 1992-2017. *Climatic Change*, 148(4), pp. 623-640.

NORDHAUS, W. D. (2018b). Projections and Uncertainties about Climate Change in An Era of Minimal Climate Policies. *American Economic Journal: Economic Policy*, 10(3), pp. 333-360.

NORDHAUS, W. D. (2019). *Climate Change: The Ultimate Challenge for Economics. American Economic Review*, 109(6), pp. 1991-2014.

NORDHAUS, W. D., and BOYER, J. (2000). *Warming the World: Economic Models of Global Warming.* MIT Press.

NORDHAUS, W. D., and SZTORC, P. (2013). *DICE 2013R: Introduction and User's Manual.* Yale University, Second Edition, October.

NORDHAUS, W. D., and YANG, Z. (1996). A Regional Dynamic General-equilibrium Model of Alternative Climate-change Strategies. *American Economic Review*, 86(4), pp. 741-765.

NORRIS, J. R. (1997). *Markov Chains.* Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press.

North Greenland Ice Core Project members (2004). High-resolution Record of Northern Hemisphere Climate Extending into the Last Interglacial Period. *Nature*, 431(7005), pp. 147-151.

O'NEILL, B. C., KRIEGLER, E., RIAHI, K., EBI, K. L., HALLEGATTE, S., CARTER, T. R., MATHUR, R., and VAN VUUREN, D. P. (2014). A New Scenario Framework for Climate Change Research: The Concept of Shared Socioeconomic Pathways. *Climatic Change*, 122, pp. 387-400.

O'NEILL, B. C., KRIEGLER, E., EBI, K. L., ..., and SOLECKI, W. (2017). The Roads Ahead: Narratives for Shared Socioeconomic Pathways Describing World Futures in the 21st Century. *Global Environmental Change*, 42, pp. 169-180.

Observatoire de la Responsabilité Sociétale des Entreprises (2001). Guide des Organismes d'Analyse Sociétale et Environnementale. *Report*, October.

Observatoire de la Responsabilité Sociétale des Entreprises (2007). Guide to Sustainability Analysis Organisations — Profiles. *Report*, December.

OHLMANN, M., MIELE, V., DRAY, S., CHALMANDRIER, L., O'CONNOR, L., and THUILLER, W. (2019). Diversity Indices for Ecological Networks: A Unifying Framework using Hill Numbers. *Ecology letters*, 22(4), pp. 737-747.

OIKONOMOU, I., BROOKS, C., and PAVELIN, S. (2014). The Effects of Corporate Social Performance on the Cost of Corporate Debt and Credit Ratings. *Financial Review*, 49(1), pp. 49-75.

ORLITZKY, M., SCHMIDT, F. L., and RYNES, S. L. (2003). Corporate Social and Financial Performance: A Meta-analysis. *Organization Studies*, 24(3), pp. 403-441.

ORTIZ, J. D., and JACKSON, R. (2022). Understanding Eunice Foote's 1856 Experiments: Heat Absorption by Atmospheric Gases. *Notes and Records*, 76(1), pp. 67-84.

PÁSTOR, L., STAMBAUGH, R. F., and TAYLOR, L. A. (2021). Sustainable Investing in Equilibrium. *Journal of Financial Economics*, 142(2), pp. 550-571.

PÁSTOR, L., STAMBAUGH, R. F., and TAYLOR, L. A. (2022). Dissecting Green Returns. *Journal of Financial Economics*, 146(2), pp. 403-424.

PAUSATA, F. S. R., GAETANI, M., MESSORI, G., BERG, A., DE SOUZA, D. M., SAGE, R. F., and DEMENOCAL, P. B. (2020). The Greening of the Sahara: Past Changes and Future Implications. *One Earth*, 2(3), pp. 235-250.

PEDERSEN, L. H., FITZGIBBONS, S., and POMORSKI, L. (2021). Responsible Investing: The ESG-Efficient Frontier. *Journal of Financial Economics*, 142(2), pp. 572-597.

PEREIRA, P., CORTEZ, M. C., and SILVA, F. (2019). Socially Responsible Investing and the Performance of Eurozone Corporate Bond Portfolios. *Corporate Social Responsibility and Environmental Management*, 26(6), pp. 1407-1422.

PARRENIN, F., BARNOLA, J. M., BEER, J., ..., and WOLFF, E. (2007). The EDC3 Chronology for the EPICA Dome C Ice Core. *Climate of the Past*, 3(3), pp. 485-497.

PERRIN, S., and RONCALLI, T. (2020). Machine Learning Optimization Algorithms & Portfolio Allocation. In Jurczenko, E. (Ed.), *Machine Learning for Asset Management: New Developments and Financial Applications*, Chapter 8, Wiley, pp. 261-328.

PETERS, G. P., and HERTWICH, E. G. (2008). CO_2 Embodied in International Trade with Implications for Global Climate Policy. *Environmental Science & Technology*, 42(5), pp. 1401-1407.

PETERS, G. P., MINX, J. C., WEBER, C. L., and EDENHOFER, O. (2011). Growth in Emission Transfers via International Trade from 1990 to 2008. *Proceedings of the National Academy of Sciences*, 108(21), pp. 8903-8908.

PETIT, J. R., JOUZEL, J., RAYNAUD, D., ..., and STIEVENARD, M. (1999). Climate and Atmospheric History of the Past 420,000 Years from the Vostok Ice Core, Antarctica. *Nature*, 399(6735), pp. 429-436.

PHILLIPS, N. A. (1956). The General Circulation of the Atmosphere: A Numerical Experiment. *Quarterly Journal of the Royal Meteorological Society*, 82(352), pp. 123-164.

PINDYCK, R. S. (2012). Uncertain Outcomes and Climate Change Policy. *Journal of Environmental Economics and Management*, 63(3), pp. 289-303.

PINDYCK, R. S. (2013). Climate Change Policy: What Do the Models Tell Us?. *Journal of Economic Literature*, 51(3), pp. 860-872.

PINDYCK, R. S. (2017). The Use and Misuse of Models for Climate Policy. *Review of Environmental Economics and Policy*, 11(1), pp. 100-114.

PLASS, G. N. (1956). The Carbon Dioxide Theory of Climatic Change. *Tellus*, 8(2), pp. 140-154.

POLBENNIKOV, S., DESCLÉE, A., DYNKIN, L., and MAITRA, A. (2016). ESG Ratings and Performance of Corporate Bonds. *Journal of Fixed Income*, 26(1), pp. 21-41.

POPP, D. (2004). ENTICE: Endogenous Technological Change in the DICE Model of Global Warming. *Journal of Environmental Economics and Management*, 48(1), pp. 742-768.

POTTIER, A., ESPAGNE, E., PERRISSIN FABERT, B., and DUMAS, P. (2015). The Comparative Impact of Integrated Assessment Models' Structures on Optimal Mitigation Policies. *Environmental Modeling & Assessment*, 20, pp. 453-473.

POUILLET, C. (1838). *Mémoire sur la chaleur solaire, sur les pouvoirs rayonnants et absorbants de l'air atmosphérique et sur la température de l'espace.* Bachelier, Paris.

POWERS, D. M. (2011). Evaluation: From Precision, Recall and F-measure to ROC, Informedness, Markedness and Correlation. *Journal of Machine Learning Technologies*, 2(1), pp 37-63.

Principles for Responsible Investment (2019a). A Practical Guide to ESG integration in Sovereign Debt. *Report*, August.

Principles for Responsible Investment (2019b). Active Owenership 2.0: The Evolution Stewardship Urgently Needs. *Report*, November.

Principles for Responsible Investment (2020). Screening. *An Introduction to Responsible Investment*, May.

Principles for Responsible Investment (2021a). Stewardship. *An Introduction to Responsible Investment*, February.

Principles for Responsible Investment (2021b). Making Voting Count. *Report*, March.

Principles for Responsible Investment (2022a). Discussing Divestment: Developing an Approach when Pursuing Sustainability Outcomes in Listed Equities. *Report*, April.

Principles for Responsible Investment (2022b). PRI's Regulation Database. *Database*, Q1 2022.

RAMANATHAN, V. (1975). Greenhouse Effect Due to Chlorofluorocarbons: Climatic Implications. *Science*, 190(4209), pp. 50-52.

RAMANATHAN, V., CICERONE, R. J., SINGH, H. B., and KIEHL, J. T. (1985). Trace Gas Trends and Their Potential Role in Climate Change. *Journal of Geophysical Research: Atmospheres*, 90(D3), pp. 5547-5566.

RAO, N. D., SAUER, P., GIDDEN, M., and RIAHI, K. (2019). Income Inequality Projections for the Shared Socioeconomic Pathways (SSPs). *Futures*, 105, pp. 27-39.

RAUPACH, M. R., MARLAND, G., CIAIS, P., LE QUÉRÉ, C., CANADELL, J. G., KLEPPER, G., and FIELD, C. B. (2007). Global and Regional Drivers of Accelerating CO_2 Emissions. *Proceedings of the National Academy of Sciences*, 104(24), pp. 10288-10293.

RBB Economics (2014). Cost Pass-through: Theory, Measurement, and Potential Policy Implications. *A Report prepared for the Office of Fair Trading*, February.

Refinitiv (2022). Environmental, Social and Governance Scores from Refinitiv. *Report*, May.

RepRisk (2022). RepRisk Methodology Overview. *Report*, July.

REVELLE, R., and SUESS, H. E. (1957). Carbon Dioxide Exchange between Atmosphere and Ocean and the Question of an Increase of Atmospheric CO_2 during the Past Decades. *Tellus*, 9(1), pp. 18-27.

RIAHI, K., VAN VUUREN, D. P., KRIEGLER, E., ..., and TAVONI, M. (2017). The Shared Socioeconomic Pathways and Their Energy, Land Use, and Greenhouse Gas Emissions Implications: An Overview. *Global Environmental Change*, 42, pp. 153-168.

RICHARDSON, K., STEFFEN, W., LUCHT, W., ..., and ROCKSTRÖM, J. (2023). Earth Beyond Six of Nine Planetary Boundaries. *Science Advances*, 9(37), eadh2458, 16 pages.

RICHARDSON, L. F. (1922). *Weather Prediction by Numerical Process*. Cambridge University Press.

RICHTERS, O., BERTRAM, C., KRIEGLER, E., ..., and ZWERLING, M. (2022). NGFS Climate Scenario Database. *Technical Documentation*, V3.1, November.

RITCHIE, P. D., CLARKE, J. J., COX, P. M., and HUNTINGFORD, C. (2021). Overshooting Tipping Point Thresholds in A Changing Climate. *Nature*, 592(7855), pp. 517-523.

ROCKSTRÖM, J., STEFFEN, W., NOONE, K., ..., and FOLEY, J. (2009). A Safe Operating Space for Humanity. *Nature*, 461(7263), pp. 472-475.

ROHDE, R. A., and HAUSFATHER, Z. (2020). The Berkeley Earth Land/Ocean Temperature Record. *Earth System Science Data*, 12(4), pp. 3469-3479.

ROLL, R. (1992). A Mean/Variance Analysis of Tracking Error. *Journal of Portfolio Management*, 18(4), pp. 13-22.

RONCALLI, T. (2013). *Introduction to Risk Parity and Budgeting*. Chapman and Hall/CRC Financial Mathematics Series.

RONCALLI, T. (2017). Alternative Risk Premia: What Do We Know?. in Jurczenko, E. (Ed.), *Factor Investing: From Traditional to Alternative Risk Premia*, Elsevier.

RONCALLI, T. (2020a). *Handbook of Financial Risk Management*. Chapman and Hall/CRC Financial Mathematics Series.

RONCALLI, T. (2020b). ESG & Factor Investing: A New Stage has been Reached. *Amundi Viewpoint*, May.

RONCALLI, T., LE GUENEDAL, T., LEPETIT, F., RONCALLI, T., and SEKINE, T. (2020). Measuring and Managing Carbon Risk in Investment Portfolios. *arXiv*, 2008.13198.

ROSATI, B., MOOTE, K., KUMAR, R., and MAIOLO, M. (2022). A Look Back at the 2022 Proxy Season. *Report*, Georgeson LLC.

ROSE, S. K., DIAZ, D. B., and BLANFORD, G. J. (2017a). Understanding the Social Cost of Carbon: A Model Diagnostic and Inter-comparison Study. *Climate Change Economics*, 8(02), 1750009.

ROSE, S. K., DIAZ, D. B., and BLANFORD, G. J. (2017b). Understanding the Social Cost of Carbon: A Model Diagnostic and Inter-comparison Study: Supplementary Material. Available at https://doi.org/10.1142/S2010007817500099.

ROSS, S. (1976). The Arbitrage Theory of Capital Asset Pricing. *Journal of Economic Theory*, 13(3), pp. 341-360.

RÖSSLER, O. E. (1976). An Equation for Continuous Chaos. *Physics Letters A*, 57(5), pp. 397-398.

ROYER, D. L., BERNER, R. A., MONTAÃŚEZ, I. P., TABOR, N. J., and BEERLING, D. J. (2004). CO_2 as a Primary Driver of Phanerozoic Climate. *GSA Today*, 14(3), pp. 4-10.

RUDD, A. (1979). Divestment of South African Equities: How Risky?. *Journal of Portfolio Management*, 5(3), pp. 5-10.

RUSSELL, S. (2023). Estimating and Reporting the Comparative Emissions Impacts of Products. World Resources Institute, *Working Paper*, 26 pages.

SACHS, J. D., LAFORTUNE, G., KROLL, G., FULLER, G. and WOELM, F. (2022). From Crisis to Sustainable Development: the SDGs as Roadmap to 2030 and Beyond. *Sustainable Development Report 2022*, Cambride University Press.

SÆTRA, H. S. (2021). A Framework for Evaluating and Disclosing the ESG Related Impacts of AI with the SDGs. *Sustainability*, 13(15), 8503.

SAGAN, C. (1986). *Dragons of Eden: Speculations on the Evolution of Human Intelligence*. Ballantine Books.

SASSI, O., CRASSOUS, R., HOURCADE, J-C., GITZ, V., WAISMAN, H-D., and GUIVARCH, C. (2010). IMACLIM-R: A Modelling Framework to Simulate Sustainable Development Pathways. *International Journal of Global Environmental Issues*, 10 (1-2), pp. 5-24.

SAUTEL, O., MINI, C., BAILLY, H., and DIEYE, R. (2022). *La tarification du carbone et ses répercussions. Exposition sectorielle au surcoût carbone*. Les Notes de La Fabrique, Presses des Mines.

SCHUUR, E. A. G., MCGUIRE, A. D., SCHÄDEL, C., ..., and VONK, J. E. (2015). Climate Change and the Permafrost Carbon Feedback. *Nature*, 520(7546), pp. 171-179.

SCHWEIZER, B., and WOLFF, E. F. (1981). On Nonparametric Measures of Dependence for Random Variables. *Annals of Statistics*, 9(4), pp. 879-885.

SELLERS, W. D. (1969). A Global Climatic Model based on the Energy Balance of the Earth-atmosphere System. *Journal of Applied Meteorology and Climatology*, 8(3), pp. 392-400.

SEMET, R., RONCALLI, T., and STAGNOL, L. (2021). ESG and Sovereign Risk: What is Priced in by the Bond Market and Credit Rating Agencies?. *SSRN*, 3940945.

SENKOWSKI, C. K., and MCKENNEY, M. G. (1999). Trauma Scoring Systems: A Review. *Journal of the American College of Surgeons*, 189(5), pp. 491-503.

ShareAction (2019). Voting Matters — Are Asset Managers using their Proxy Votes for Climate Action?. *Rankings& Surveys*, November.

ShareAction (2020). Voting Matters 2020 — Are Asset Managers using their Proxy Votes for Action on Climate and Social Issues?. *Rankings & Surveys*, December.

ShareAction (2021). Voting Matters 2021 — Are Asset Managers using their Proxy Votes for Action on Climate and Social Issues?. *Rankings & Surveys*, December.

ShareAction (2023). Voting Matters 2022. *Rankings & Surveys*, January.

SHARPE, W. F. (1964). Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *Journal of Finance*, 19(3), pp. 425-442.

SHEPHERD, A., IVINS, E., RIGNOT, E., and the IMBIE team (2020). Mass Balance of the Greenland Ice Sheet from 1992 to 2018. *Nature*, 579(7798), pp. 233-239.

SHERWOOD, S. C., WEBB, M. J., ANNAN, J. D., ..., and ZELINKA, M. D. (2020). An Assessment of Earth's Climate Sensitivity using Multiple Lines of Evidence. *Reviews of Geophysics*, 58(4), 2019RG000678, 92 pages.

SHIELDS, G., and VEIZER, J. (2002). Precambrian Marine Carbonate Isotope Database: Version 1.1. G^3: Geochemistry, Geophysics, Geosystems, 3(6), pp. 1-12.

SHINE, K. P., FUGLESTVEDT, J. S., HAILEMARIAM, K., and STUBER, N. (2005). Alternatives to the Global Warming Potential for Comparing Climate Impacts of Emissions of Greenhouse Gases. *Climatic Change*, 68(3), pp. 281-302.

SKLAR, A. (1959). Fonctions de Répartition à n Dimensions et leurs Marges. *Publications de l'Institut de Statistique de l'Université de Paris*, 8(1), pp. 229-231.

S&P Dow Jones Indices (2022). S&P DJI ESG Score. *Methodology*, May.

STANTON, E. A. (2011). Negishi Welfare Weights in Integrated Assessment Models: The Mathematics of Global Inequality. *Climatic Change*, 107(3-4), pp. 417-432.

STEFFEN, W., RICHARDSON, K., ROCKSTRÖM, J., ..., and SÖRLIN, S. (2015). Planetary Boundaries: Guiding Human Development on a Changing Planet. *Science*, 347(6223), pp. 736 & 1259855 (10 pages).

STEFFEN, W., ROCKSTRÖM, J., RICHARDSON, K., ..., and SCHELLNHUBER, H. J. (2018). Trajectories of the Earth System in the Anthropocene. *Proceedings of the National Academy of Sciences*, 115(33), pp. 8252-8259.

STEHFEST, E., VAN VUUREN, D. P., KRAM, T., and BOUWMAN, L. (2014), *Integrated Assessment of Global Environmental Change with IMAGE 3.0: Model Description and Policy Applications.* Netherlands Environmental Assessment Agency.

STEPHENS, G. L., O'BRIEN, D., WEBSTER, P. J., PILEWSKI, P., KATO, S., & LI, J.-L. (2015). The Albedo of Earth. *Reviews of Geophysics*, 53(1), pp. 141-163.

STEPHENSON, P. J., and STENGEL, C. (2020). An Inventory of Biodiversity Data Sources for Conservation Monitoring. *PloS One*, 15(12), e0242923.

STERN, N. (2007). *The Economics of Climate Change: The Stern Review.* Cambridge University Press.

STERN, N., and TAYLOR, C. (2007). Climate Change: Risk, Ethics, and the Stern Review. *Science*, 317(5835), pp. 203-204.

STOKES, C. R., ABRAM, N. J., BENTLEY, M. J., ..., and WHITEHOUSE, P. L. (2022). Response of the East Antarctic Ice Sheet to Past and Future Climate Change. *Nature*, 608(7922), pp. 275-286.

STROGATZ, S. H. (2015). *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology. Chemistry, and Engineering.* Westview Press, Second Edition, 528 pages.

TANG, D. Y., and ZHANG, Y. (2020). Do Shareholders Benefit from Green Bonds?. *Journal of Corporate Finance*, 61, 101427.

TANZER, S. E., and RAMÍREZ, A. (2019). When Are Negative Emissions Negative Emissions?. *Energy & Environmental Science*, 12(4), pp. 1210-1218.

Task Force on Climate Related Financial Disclosures (2017). Recommendations of the Task Force on Climate-related Financial Disclosures. *Final Report*, June.

Task Force on Climate Related Financial Disclosures (2021a). Implementing the Recommendations of the Task Force on Climate-related Financial Disclosures. *Implementation Guidance*, October.

Task Force on Climate Related Financial Disclosures (2021b). Guidance on Metrics, Targets, and Transition Plans. *Implementation Guidance*, October.

Task Force on Climate Related Financial Disclosures (2022). 2022 Status Report. *Status Report*, October.

Technical Expert Group on Sustainable Finance (2019a). TEG Final Report on Climate Benchmarks and Benchmarks' ESG Disclosures. *Report*, September.

Technical Expert Group on Sustainable Finance (2019b). Handbook of Climate Transition Benchmarks, Paris-Aligned Benchmarks and Benchmarks' ESG Disclosures. *Report*, December.

Technical Expert Group on Sustainable Finance (2020). TEG Final Report on the EU Taxonomy. *Report*, March.

TERLOUW, T., BAUER, C., ROSA, L., and MAZZOTTI, M. (2021). Life Cycle Assessment of Carbon Dioxide Removal Technologies: A Critical Review. *Energy & Environmental Science*, 14(4), pp. 1701-1721.

THORNHILL, G. D., COLLINS, W. J., KRAMER, R. J., ..., and ZHANG, J. (2021). Effective Radiative Forcing from Emissions of Reactive Gases and Aerosols — A Multi-model Comparison. *Atmospheric Chemistry and Physics*, 21(2), pp. 853-874.

TIBSHIRANI, R. (1996). Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society B*, 58(1), pp. 267-288.

TIMMER, M. P., DIETZENBACHER, E., LOS, B., STEHRER, R., and DE VRIES, G. J. (2015). An Illustrated User Guide to the World Input-Output Database: The Case of Global Automotive Production. *Review of International Economics*, 23(3), pp. 575-605.

TOBIN J. (1958). Liquidity Preference as Behavior Towards Risk. *Review of Economic Studies*, 25(2), pp. 65-86.

TOL, R. S. J. (1997). On the Optimal Control of Carbon Dioxide Emissions: An Application of FUND. *Environmental Modeling & Assessment*, 2, pp. 151-163.

TOL, R. S. J. (2022). A Meta-analysis of the Total Economic Impact of Climate Change. *arXiv*, 2207.12199.

TONELLO, M. (2022). Shareholder Voting Trends (2018-2022). *Report*, The Conference Board.

TOWNSEND, B. (2020). From SRI to ESG: The Origins of Socially Responsible and Sustainable Investing. *Journal of Impact and ESG Investing*, 1(1), pp. 10-25.

TRENBERTH, K. E., FASULLO, J. T., and KIEHL, J. T. (2009). Earth's Global Energy Budget. *Bulletin of the American Meteorological Society*, 90(3), pp. 311-324.

TURNER, K., LENZEN, M., WIEDMANN, T., and BARRETT, J. (2007). Examining the Global Environmental Impact of Regional Consumption Activities — Part 1: A Technical Note on Combining Input-output and Ecological Footprint Analysis. *Ecological Economics*, 62(1), pp. 37-44.

TYNDALL, J. (1861). On the Absorption and Radiation of Heat by Gases and Vapours, and on the Physical Connexion of Radiation, Absorption, and Conduction — The Bakerian Lecture. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 22(146), pp. 169-194.

United Nations Economic Commission for Europe (2022). Carbon Neutrality in the UNECE Region: Integrated Life-cycle Assessment of Electricity Sources. *Report*, April.

United Nations Environment Program (2022). Advancing Delivery on Decarbonisation Targets. *Second Progress Report*, September.

VANCE, S. C. (1975). Are Socially Responsible Corporations Good Investment Risks. *Management Review*, 64(8), pp. 19-24.

VAN LOAN, C. F. (1978). Computing Integrals Involving the Matrix Exponential. *IEEE Transactions on Automatic Control*, 23(3), pp. 395-404.

VEIZER, J., ALA, D., AZMY, K., ..., and STRAUSS, H. (1999). $^{87}Sr/^{86}Sr$, $\delta^{13}C$ and $\delta^{18}O$ Evolution of Phanerozoic Seawater. *Chemical Geology*, 161(1-3), pp. 59-88.

VEIZER, J., GODDERIS, Y., and FRANÃĞOIS, L. M. (2000). Evidence for Decoupling of Atmospheric CO2 and Global Climate during the Phanerozoic Eon. *Nature*, 408(6813), pp. 698-701.

VENÅLÅINEN, A., LEHTONEN, I., LAAPAS, M., RUOSTEENOJA, K., TIKKANEN, O. P., VIIRI, H., IKONEN, V. P., and PELTOLA, H. (2020). Climate Change Induces Multiple Risks to Boreal Forests and Forestry in Finland: A Literature Review. *Global Change Biology*, 26(8), pp. 4178-4196.

VOSE, R. S., HUANG, B., YIN, X., ..., and ZHANG, H. M. (2021). Implementing Full Spatial Coverage in NOAA's Global Temperature Analysis. *Geophysical Research Letters*, 48(4), 2020GL090873.

WACKERNAGEL, M., and REES, W. (1996). *Our Ecological Footprint: Reducing Human Impact on the Earth*. New Society Publishers.

WAGNER, G., and WEITZMAN, M. L. (2015). *Climate Shock: The Economic Consequences of a Hotter Planet*. Princeton University Press.

WATERS, C. N., ZALASIEWICZ, J., SUMMERHAYES, ..., and WOLFE, A. P. (2016). The Anthropocene is Functionally and Stratigraphically Distinct from the Holocene. *Science*, 351(6269), page 137.

WEART, S. (2023). *The Discovery of Global Warming*. American Institute of Physics, https://history.aip.org/climate/index.htm.

WEITZMAN, M. L. (2007). A Review of the Stern Review on the Economics of Climate Change. *Journal of Economic Literature*, 45(3), pp. 703-724.

WEITZMAN, M. L. (2009). On Modeling and Interpreting the Economics of Catastrophic Climate Change. *Review of Economics and Statistics*, 91(1), pp. 1-19.

WEITZMAN, M. L. (2010). What is the "Damages Function" for Global Warming — And What Difference Might it Make?. *Climate Change Economics*, 1(1), pp. 57-69.

WEITZMAN, M. L. (2012). GHG Targets as Insurance against Catastrophic Climate Damages. *Journal of Public Economic Theory*, 14(2), pp. 221-244.

WESTVEER, J., FREEMAN, R., MCRAE, L., MARCONI, V., ALMOND, R. E. A., and GROOTEN, M. (2022). A Deep Dive into the Living Planet Index. *Technical Report*, WWF.

Who Cares Wins (2004). Connecting Financial Markets to a Changing World. *Conference Report*, December.

Who Cares Wins (2005). Investing for Long-Term Value — Integrating Environmental, Social and Governance Value Drivers in Asset Management and Financial Research. *Conference Report*, October.

WIEDMANN, T. (2009). A Review of Recent Multi-region Input-output Models used for Consumption-based Emission and Resource Accounting. *Ecological Economics*, 69(2), pp. 211-222.

WIEDMANN, T., LENZEN, M., TURNER, K., and BARRETT, J. (2007). Examining the Global Environmental Impact of Regional Consumption Activities — Part 2: Review of Input-output Models for the Assessment of Environmental Impacts Embodied in Trade. *Ecological Economics*, 61(1), pp. 15-26.

WIEDMANN, T., and MINX, J. (2008). A Definition of 'Carbon Footprint'. In Pertsova, C. C., *Ecological Economics Research Trends*, Chapter 1, pp. 1-11, Nova Science Publishers.

WILLIAMSON, O. E. (1970). *Corporate Control and Business Behavior*. Prentice-Hall.

WOOD, D. J. (1991). Corporate Social Performance Revisited. *Academy of Management Review*, 16(4), pp. 691-718.

World Meteorological Organization (1982). Meeting of Experts on Potential Climatic Effects of Ozone and Other Minor Trace Gases (13-17 September 1982). *Report (GORMP)*, 14, 35 pages.

WUNDERLING, N., DONGES, J. F., KURTHS, J., and WINKELMANN, R. (2021). Interacting Tipping Elements Increase Risk of Climate Domino Effects under Global Warming. *Earth System Dynamics*, 12(2), pp. 601-619.

YAMANO, N., and GUILHOTO, J. J. M. (2020). CO_2 Emissions Embodied in International Trade and Domestic Final Demand: Methodology and Results using the OECD Inter-Country Input-Output Database. *OECD Working Paper*, 2020/11, 55 pages.

YUAN, Y., LU, L. Y., TIAN, G., and YU, Y. (2020). Business Strategy and Corporate Social Responsibility. *Journal of Business Ethics*, 162(2), pp. 359-377.

ZACHOS, J. C., DICKENS, G. R., and ZEEBE, R. E. (2008). An Early Cenozoic Perspective on Greenhouse Warming and Carbon-cycle Dynamics. *Nature*, 451(7176), pp. 279-283.

ZACHOS, J., PAGANI, M., SLOAN, L., THOMAS, E., and BILLUPS, K. (2001). Trends, Rhythms, and Aberrations in Global Climate 65 Ma to Present. *Science*, 292(5517), pp. 686-693.

ZERBIB, O. D. (2019). The Effect of Pro-environmental Preferences on Bond Prices: Evidence from Green Bonds. *Journal of Banking & Finance*, 98, pp. 39-60.

ZERBIB, O. D. (2022). A Sustainable Capital Asset Pricing Model (S-CAPM): Evidence from Environmental Integration and Sin Stock Exclusion. *Review of Finance*, 26(6), pp. 1345-1388.

Subject Index

Author Index