Why Do Firms Choose to Greenwash: An Evolutionary Analysis of Greenwashing Incentives and Deterrents

Ille, Sebastian and Carrera, Edgar J. Sanchez

The American University in Cairo, University of Florence

3 August 2025

Why Do Firms Choose to Greenwash: An Evolutionary Analysis of Greenwashing Incentives and Deterrents

Sebastian Ille[a,*], Edgar J. Sanchez Carrera[b]

[a] *Onsi Sawiris School of Business, The American University in Cairo. E-mail:sebastian.ille@aucegypt.edu ORCID:*

[b] *Department of Economics and Management at the University of Florence, Italy. Research Fellow at the Applied Mathematics Research Center CIMA UAdeC, Mexico. E-mail: edgar.sanchez@unifi.it ORCID: 0000-0002-1191-8859*

Abstract

With the increasing demand for sustainable products, greenwashing has become more prevalent and sophisticated over the past decade. To better understand the incentives for firms to greenwash, we develop an evolutionary game-theoretic model in which firms may choose to mimic green behavior without having to bear the cost linked to green investment and production. We provide the conditions for the different evolutionarily stable equilibria. In a second step, we extend the model using agent-based simulations to incorporate path-dependent investment/production costs, history-dependent mimicry effectiveness, peer effects, and localized firm interactions. We show that the simpler model with random matching offers good approximations of the equilibrium conditions in more complex setups, but market segmentation supports green investment and production in contrast to higher penalties. While curtailing opportunities to pretend green behaviour boosts green production, we also find that increasing cost efficiencies encourage firms to engage in green production, even in the face of increasingly sophisticated deceptive strategies. Based on our results, we suggest trio-targeted policies that reduce the (initial) costs of green investment/production, curtail opportunities to mimic green behavior, and support segmentation.

Keywords: Climate change; Non-linear Macroeconomic Models; Greenwashing; Corporate Sustainability.
JEL classification: C27, G14, G32, Q56, M14

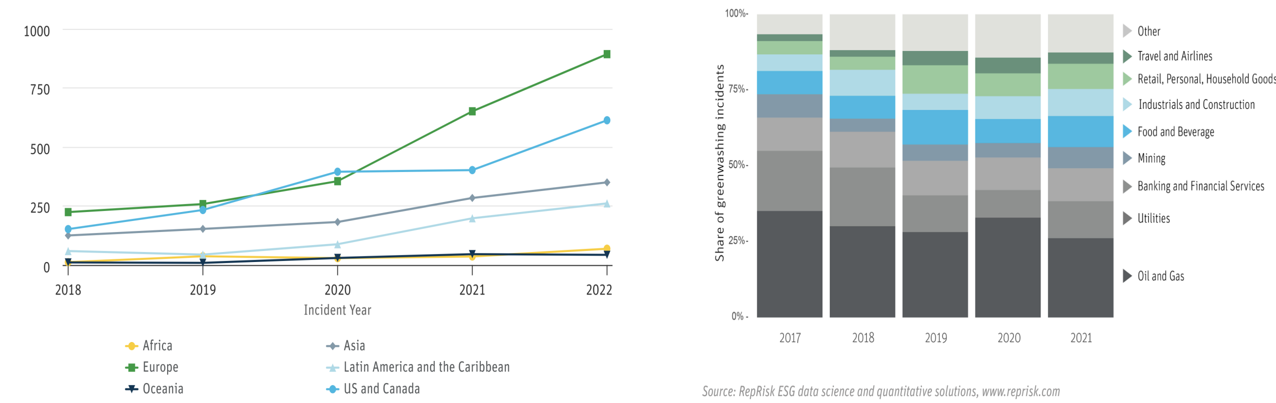
*Corresponding author. Email: sebastian.ille@aucegypt.edu

1. Introduction

The rise of corporate social responsibility (CSR) in recent years is part of a growing trend by firms to reaffirm their responsibility towards the communities they serve. At the same time, the rising threat of climate change has shifted the public focus to the contributing role of firms to climate change, excess waste, and other environmental impacts of provision. Consequently, ESG (environmental, social, and governance) investing has become a major principle for global investment strategies. Investors seek environmentally responsible firms that pledge to reduce harmful practices, regulate their energy consumption, and neutralize the negative environmental externalities of their business. While CSR increases the profit opportunities of firms, it dictates a balanced approach to providing socio-economic and environmental benefits, sustainability, and profit. Firms discovered CSR as an opportunity to boost their reputation and brand image and encourage investment and customer retention. Some firms, however, have abandoned a balanced approach and use it to maximize profits by making false claims about the social responsibility of their practices. Within the context of environmental responsibility, such actions have been termed greenwashing (or green sheen).

Empirical evidence is concerning: greenwashing activities carried out by firms are increasingly common. RepRisk database (`https://www.reprisk.com`) reports, apart from a recent mild decrease, a constant increase in the frequency of greenwashing linking one in four climate-related ESG risk incidents globally related to the former. In addition, greenwashing is linked to its social counterpart, social washing. RepRisk's data shows that nearly one in three public companies linked to greenwashing also engage in social washing. The data shows a significant increase in greenwashing over the past years since the signing of the Paris Agreement in 2015, especially in the US, Canada, and Europe (see Figure 1) and, in particular, in the banking and financial services sectors (RepRisk Report, 2022). With greater public attention to greenwashing and, eventually, broader criticism in Europe and America, firms have developed more sophisticated means of greenwashing. In addition to directly misleading consumers, firms are now using pledges, certifications, and future commitments that do not require immediate actions and render greenwashing less visible (Fürer, 2023).

Figure 1: Left: Unique entities with at least one ESG incident linked to both environmental footprint and misleading communication in a given year (Fürer, 2023). Right: breakdown of sectors linked to climate greenwashing (RepRisk Report, 2022).



With the growing empirical evidence of greenwashing and the rising sophistication of measures taken by firms to deceive investors and customers, this paper supports and extends the existing, predominantly empirical, and sector-specific literature by providing a dynamic model that considers the limited strategic and path-dependent aspects of greenwashing. Our model captures coevolutionary factors that influence the adoption of green and greenwashing firm behavior that are not captured by existing static models. In a first step, we develop an evolutionary game-theoretic model in which firms are confronted with the choice between three strategies: using a green or brown production technology or greenwashing. In this model, we capture several interesting strategic characteristics of greenwashing: To minimize discovery, greenwashing firms greenscam consumers and, in doing so, engage in the form of aggressive mimicry. The latter describes an evolutionary strategy in which predators convince their prey of their harmlessness by superficially resembling an innocuous third party. Similarly, greenwashing firms attempt to pose as sustainable firms to attract customers and investors. The success of greenwashing critically depends on the cost of mimicry, the risk of discovery, and the costs attributed to such a case in relation to the benefits of successful greenwashing. Greenwashing is further subject to peer dynamics and externalities. On the one hand, frequent exposure to firms' greenwashing practices erodes the public's trust in these industries, increasing scrutiny and the risk of discovery.[1] On the other hand, strong competition encourages firms to adopt practices that maximize profit. Since greenwashing allows firms to cut costs, a higher prevalence of successfully greenwashing firms encourages other firms to follow the same strategy. This creates a strategic coupling between firms, where a higher pervasiveness of greenwashing in an industry increases the likelihood of detection but also makes firms more inclined to utilize this strategy. Thus, greenwashing is both a strategic substitute and a strategic complement.

We study the dynamics for two different regimes, depending on whether the profit premium attributed to green production exceeds the cost of green production, and for each, provide the explicit condition which determines the type of the evolutionarily stable equilibria. In a second step, we incorporate more realistic dynamics, such as localized firm interactions and peer effects, as well as history-dependent costs of green investment, profits, and mimicry efficiency. To manage the increased complexity of this extended model, we simulate the interactions using an agent-based model. Comparing the results between both models reveals that the closed-form conditions obtained in the simpler version offer a good approximation for the dynamics of the extended model. However, market segmentation and history dependencies of both costs and mimicry efficiency encourage green production and discourage greenwashing beyond the initially predicted levels. In addition, using the flexibility of the agent-based approach, we further examine how different random network topologies influence the strategy distribution of the firm population in this context. Although the network structure affects the rate of convergence towards an equilibrium, the long-term strategy distributions are robust across these various network types.

Based on these results, we argue in favor of sets of trio-targeted policies that simultaneously reduce the (initial) costs of green investment/production, curtail opportunities to

[1] Increased public distrust can even lead green firms to hide their sustainability goals to avoid being falsely labeled greenwashers. This strategy is known as greenhushing, which we will not study as part of the simple model in this paper.

mimic green behavior, and support segmentation. At the end of this paper, we provide policy suggestions and support for existing policies that encourage these three dimensions.

The remainder of the paper proceeds as follows. After a short literature review in the following section 2, we develop the simple evolutionary model and determine the attractors of the system in section 3. Section 4 extends the model to an agent-based model, and we present the simulation results. Section 5 discusses the results and their implications for policymaking, and section 6 concludes.

2. Literature review

While the investment in green technology entails costs (Eyraud et al., 2013), investors and other stakeholders look favorably on firms which disclose green investments leading to significant gains in firm value (Martin and Moser, 2016; Yadav et al., 2016). Greenwashing then arises as an element of concern when firms fail to internalize the costs of brown productions and climate change, and perceive green investments as a vehicle to promote their products and attract investors.[2]

Greenwashing further implies less effective climate policies and is likely to influence attitudes not only in the financial sector but also those of consumers.[3] False or misleading information about a firm's environmental claims is difficult to spot for consumers (Juliana Fernandes and Leopold, 2020). However, if discovered, the practice damages corporate reputation and stock performance, while negatively impacting consumer purchases, creating a "lemon market" problem (Ioannou et al., 2023). Current research on the management of corporate greenwashing as a critical environmental policy issue has attracted substantial

[2]Coined back in the 1980s, the term characterizes the practice of promoting environmentally friendly programs to deflect attention from an organization's environmental unfriendly activities. This practice has become increasingly relevant in the context of financial investment where green bonds are perceived as an environmentally friendly signal (see Flammer, 2021). Greenwashing causes these bonds to no longer reflect a firm's real resilience to the climate-transition risk.

[3]Quantitative methods and models developed using Mean-Field-Games (MFGs) and Mean-Field Control (MFC) represent an interesting and innovative framework for studying this type of problem (see Carmona and Delarue, 2018). Both family of models are flexible enough to incorporate climate-adjusted credit risk models and mimic dynamic incentive schemes effects.

academic attention. Two broad streams of literature are relevant for our study:

- Ecological behavior games: This literature studies topics related to the management of green supply chains, examining how chains can provide green products that meet environmental and sustainability needs (Cai et al., 2023). In this context, game theoretic models are increasingly used to simulate collaborative evolutionary stability trends in emissions reductions (Liu et al., 2022, 2023), and explore emerging green corporate behaviors (Chen et al., 2021). Some scholars have studied green supply chain issues under a dynamic framework using differential game models (Liu et al., 2022).

- The motivation for greenwashing: Assuming that companies are self-interested profit seekers who minimize costs, greenwashing is a rational corporate strategy influenced by competitive pressures and opportunities (He et al., 2022). Environmental policy requirements (Zhang, 2022b) and consumer demand (Wang and Hou, 2020) have become important drivers of greenwashing. Limited consumer knowledge of the green market causes information asymmetry which firms exploit (Hameed et al., 2021) and imperfect regulations enable greenwashing through loopholes in certification (Demirel et al., 2018). Factors influencing the decision to greenwash can be external market factors (consumers, policies, etc.) and internal non-market factors (Delmas and Burbano, 2011a; Lyon and Maxwell, 2011a; Zhang, 2022a).

Those aforementioned studies, along with the literature on climate sentiments (Besley and Persson, 2023; Dunz et al., 2021) provide the basis for our analysis. However, by providing a dynamic model that accounts for the limited strategic as well as path-dependent aspects of greenwashing, our approach both supports and extends the predominantly empirical and sector-specific literature in the field of greenwashing (Huang et al., 2025; Gupta and Singh, 2024; Bernini et al., 2024, see also Delmas and Burbano, 2011b; Laufer, 2003; Parguel et al., 2011; Wu and Shen, 2013; Marquis et al., 2016; Wei Xu, 2023 for examples). Our evolutionary analysis has advantages over static models (e.g., Agi et al., 2021; Leng and Parlar, 2005;

Huang et al., 2025 and for a good example, see Lyon and Maxwell, 2011b) since we capture coevolutionary factors of greenwashing that include elements of strategic complements (the engagement in greenwashing of one firm incentivizes other firms) and strategic substitutes (a higher frequency of greenwashing increases public awareness). The extended model includes factors that have been shown to affect green production and greenwashing, such as path-dependent profits Gaio and Raposo (2011); Al Ani and Chavali (2023); Hirsch et al. (2021) and costs (Kesidou and Demirel, 2012; Zhou and Gu, 2019; Faber et al., 2022), and relates the ability to mimic green production to the previous history of successful and unsuccessful mimicry. It also adds localized peer effects (see, for example, the literature on strategic groups, Fiegenbaum and Thomas, 1995, three-dimensional institutional equivalence, Gao et al., 2024, and strategic change, Massini et al., 2005; Ma et al., 2024) that influence a firm's decision based on the historical prevalence of greenwashing in its neighborhood. In doing so, our paper goes beyond more classical approaches to modeling deceptive signaling.[4]

3. The Dynamical Model with Random Matching

Each firm i has to choose between three strategies $s_i = \{G, B, W\}$, where G indicates an investment in green technology, B in brown technology, and W denotes the decision to deceive investors and customers by greenwashing their production and products. For simplicity, we do not describe a production process nor a production function in this simplified model.

We assume that each firm receives a baseline profit of ϕ_i and each strategy is connected to additional revenues and costs independent of the quantity of output. If playing $s_i = G$, firm i is able to attract additional revenues given by a profit premium σ_i and an additional individual cost c_i as a share of the baseline profit. The latter is caused by additional investments in green technology and the production of green output. A brown firm, on the other hand, does neither accrue a profit premium nor does it bear the additional cost of investment. A greenwashing firm i is able to attract the profit premium σ_i but only bears a cost of mimicry

[4] See also the argument in Casey et al., 2021)

μ_i less than c_i and σ_i as a share of the baseline profit, which is paid for pretending being green (i.e. the cost of mimicry). If, however, a firm is identified as a greenwasher, it does not receive a profit premium, but to the contrary, bears a reduction of its profit ρ as a share of the baseline profit. Since the premium and costs are fractions of the baseline profit ϕ_i, we further have $\{\sigma_i, c_i, \mu_i, \rho_i\} \in (0, 1)$. Further assume that each firm i is subject to profit variations in each period t defined by $\epsilon_i \in (-\varepsilon, \varepsilon)$ with an expected value of 0. We then obtain the four payoff functions:[5]

$$\pi_i(G) = (1 + \sigma_i - c_i + \epsilon_i)\,\phi_i \tag{1a}$$

$$\pi_i(B) = (1 + \epsilon_i)\phi_i \tag{1b}$$

$$\pi_i^{nd}(W) = (1 + \sigma_i - \mu_i + \epsilon_i)\,\phi_i \qquad \text{if not detected} \tag{1c}$$

$$\pi_i^{d}(W) = (1 - \rho_i + \epsilon_i)\,\phi_i \qquad \text{if detected} \tag{1d}$$

The expected profit of a greenwashing firm then depends on the probability θ_i of being detected by the public. Probability θ_i is partially defined by the public awareness of greenwashing in the form of a level of suspicion α and the (individual) effectiveness of mimicry ν_i, which for the moment we consider to be constant but less than the level of suspicion. Let the share of G-players be x, the share of W-players be y, and thus, the share of B-players be z. We assume the simple relation:

$$\alpha = \frac{y^\beta}{x^\beta + y^\beta} \tag{2a}$$

$$\theta_i = \max\{\alpha - \nu_i, 0\} \tag{2b}$$

Parameter $\beta \geq 0$ defines the public inertia.[6] The larger β is, the more the public will need

[5]Note that since we do not model a demand and production function and therefore, we do not include a process of output optimization. We therefore assume that the parameters in the payoff functions are the result of such an optimization process and that the marginal profits of all firms are normalized relative to the profit of brown firms.

[6]Equation (2a) models a baseline probability without mimicry and is frequently used in the literature to

to be exposed to greenwashing firms before investors and consumers become suspicious. In turn, they react more strongly once suspicious. The expected payoff of a greenwashing firm is then a convex combination of (1c) and (1d) with weight (2b) given by:

$$\pi_i(W) = \phi_i \left(1 - \rho_i + \epsilon_i + \frac{(\sigma_i - \mu_i + \rho_i)\left(\nu_i\left(x^\beta + y^\beta\right) + x^\beta\right)}{x^\beta + y^\beta}\right) \tag{3}$$

Firms are expected to regularly update their strategy based on a payoff comparison with firms with which they interact. Assuming random matching, we can represent the system's dynamics using the following replicator dynamics,

$$\dot{x} = x\left(\pi_i(G) - \Delta\right) \tag{4a}$$

$$\dot{y} = y\left(\pi_i(W) - \Delta\right) \tag{4b}$$

where $\dot{x} = \partial x/\partial t$ and $\dot{y} = \partial y/\partial t$ define the corresponding time derivatives and $\Delta = \pi_i(B)z + \pi_i(G)x + \pi_i(W)y$ defines the average payoff of the population. It is sufficient to define the replicator for frequencies x and y, since $z = 1 - x - y$ and the structure of the replicator dynamics ensures that the strategy frequencies remain within the unit interval. A strategy then increases whenever firms following such a strategy outperform the market average and decreases whenever such firms perform less well than the market average (Ille, 2022).

For $\beta = 1$, we can solve the system analytically. In addition, we drop the indices since the replicator dynamics operate on population averages. Further note that a fixed point is obtained whenever the strategy distribution does not change. Solving for $\dot{x} = 0$ and $\dot{y} = 0$

model biased cultural transmission. Parameter β can then be interpreted as defining the degree of conformism in a society.

simultaneously and noting that[7]

$$\lim_{y\to 0, x\to 0} (\dot{x}, \dot{y}) = (0, 0) \tag{5}$$

we obtain the following proposition by identifying the fixed points.

Proposition 1. *The dynamical system (4a-b) is characterized by the following stationary states and their corresponding eigenvalues:*

$$(x = 0, y = 0): \ \lambda_{001} = \phi\left(\sigma - c\right), \ \lambda_{002} = \phi\left(\nu\left(\sigma - \mu + \rho\right) - \rho\right) \tag{6a}$$

$$\begin{aligned}(x = 0, y = 1): \ & \lambda_{011} = \phi\left(\rho - \nu\left(\sigma - \mu + \rho\right)\right), \\ & \lambda_{012} = \phi\left(\sigma - c + \rho - \nu\left(\sigma - \mu + \rho\right)\right)\end{aligned} \tag{6b}$$

$$(x = 1, y = 0): \ \lambda_{101} = -\phi\left(\sigma - c\right), \ \lambda_{102} = \phi\left(\nu\left(\sigma - \mu + \rho\right) + c - \mu\right) \tag{6c}$$

$$\begin{aligned}(x = x_0, y = y_0): \ & \lambda_{xy1} = -\phi\left(\sigma - c\right), \\ & \lambda_{xy2} = \frac{\phi\left(-\sigma + c - \rho + \nu\left(\sigma - \mu + \rho\right)\right)\left(c - \mu + \nu\left(\sigma - \mu + \rho\right)\right)}{\sigma - \mu + \rho}\end{aligned} \tag{6d}$$

where

$$x_0 = -\nu + \frac{\sigma - c + \rho}{\sigma - \mu + \rho} \tag{7a}$$

$$y_0 = 1 - x_0 \tag{7b}$$

define the interior fixed point if $x_0, y_0 \in (0, 1)$.

We can see that the interior equilibrium is in fact a corner solution at which $z = 0$. Further, we obtain the sensible result that the equilibrium share of greenwashing firms increases with a higher effectiveness of mimicry, while the share of green firms decreases at (x_0, y_0). Since $\mu < c, \sigma$, fixed point $(1, 0)$ is unstable. Assuming, $\sigma > c$, fixed point $(0, 1)$ requires that the effectiveness of mimicry is sufficiently high, i.e.

$$\nu > \frac{\sigma - c + \rho}{\sigma - \mu + \rho} \tag{8}$$

[7]Note that the eigenvalues at fixed point $(0, 0)$ depend on the sequence at which the arguments approach the limit. For $\lim_{x\to 0, y\to 0}(\dot{x}, \dot{y}) = (0, 0)$, we obtain the eigenvalues $\lambda_{001} = \phi\left(\sigma - c\right)$ and $\lambda_{002} = \phi\left(\nu\left(\sigma - \mu + \rho\right) + \sigma - \mu\right)$. In this case, the dynamics of the system are not fully described by the eigenvalues.

while fixed point (x_0, y_0) requires the opposite to be stable, which also ensures that the interior equilibrium is situated within the unit interval, i.e.

$$\nu < \frac{\sigma - c + \rho}{\sigma - \mu + \rho} \tag{9}$$

Assuming $\sigma < c$, fixed point $(0, 1)$ requires the slightly modified condition

$$\nu > \frac{\rho}{\sigma - \mu + \rho} \tag{10}$$

and the stability of $(0, 0)$ requires the inverse condition

$$\nu < \frac{\rho}{\sigma - \mu + \rho} \tag{11}$$

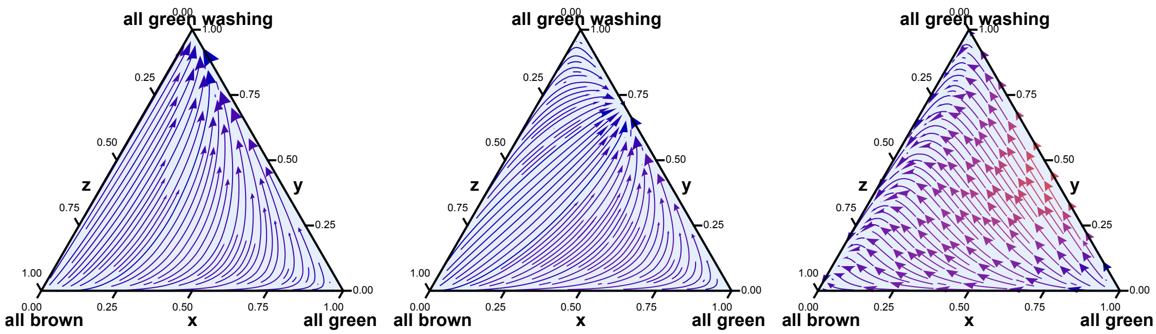
leading to the following qualitative result of the dynamical system with random matching (4a-b).

Corollary 1. *The dynamical system defined by payoffs* (1a), (1b), *and* (3) *and the underlying replicator dynamics* (4) *is subject to two regimes, depending on whether the premium from green investment exceeds its cost. In each regime, the firm population converges to a globally stable equilibrium determined by the efficiency of mimicry.*

- *In regime 1: $\sigma > c$, the system can sustain an equilibrium in which all firms engage in greenwashing or an equilibrium in which some firms greenwash while the remaining share of firms legitimately invests in green technologies. The stability of the former equilibrium requires that the effectiveness of mimicry exceeds the net benefit of green investment over being a detected greenwasher relative to the net benefit of undetected greenwashing over being a detected greenwasher. Otherwise, we observe the latter equilibrium.*

- *In regime 2: $\sigma < c$, the same system can sustain a pure population composed entirely of greenwashing firms or a pure population with only brown firms.The former occurs if the effectiveness of mimicry exceeds the cost of detection relative to the net benefit of undetected greenwashing over being a detected greenwasher (condition* (10)*). The latter equilibrium occurs in the opposite case* (11)*.*

Notice that the largest share of green firms occurs for $\lim_{\mu \to c} x_0 = 1 - \nu$ leading to the intuitive result that only if the cost-benefit of greenwashing converges to zero and mimicry is ineffective, all firms will invest in green technology.[8] The following simplexes (Fig. 2) represent a projection of the strategy space indicating the system's dynamics.



(a) $\nu_i = 0.7, \sigma_i = 0.8; y^* = 1$ (b) $\nu_i = 0.3, \sigma_i = 0.8; y^* = 0.7$ (c) $\nu_i = 0.3, \sigma_i = 0.5, z^* = 1$

Figure 2: $\beta = 1, c_i = 0.6, \mu_i = 0.4, \rho_i = 0.1, \phi_i = 100$

This simple dynamical model with random matching cannot explain the persistence of a completely mixed equilibrium in which a positive share of firms chooses each strategy, nor a stable equilibrium with only green firms. It ignores a number of critical factors. Random matching does not reflect how firms interact since any firm i has an equal probability of receiving profit information from any other firm $j \neq i$. Not only are some details relevant to a firm's profit (e.g., off-balance sheet items) not publicly available. In addition, firms may not compete under the same market conditions but are segmented across markets. An interaction with a fixed structure, in which any firm i has only partial access to information and, thus, interactions are constrained to a subset of firms, is a more realistic setting.

To obtain a closed-form solution, we made some further simplifying assumptions. The cost of investment is considered constant. However, while the initial costs of investment in green technology can be very high, leading to $\sigma_i < c_i$, these costs diminish over time (ignoring straight-line amortization, see also Karásek and Pavlica, 2016). Thus, even at initially high investment costs, firms may support green technology as long as the costs diminish sufficiently quickly.[9] Similarly, if found to be greenwashing, firms find it difficult to

[8]The condition further ensures that given (2), $\alpha - \nu$ is strictly non-negative, allowing us to simplify the probability of detection to $\theta_i = \alpha - \nu_i$.

[9]Although (Pekovic et al., 2018) indicate that the relationship between cost and degree of investment are

raise investment or to sell their goods and are unable to establish a green image convincingly. Yet, these perpetrating firms can switch to a green strategy once redeemed at the end of their punishment period. On the other hand, greenwashing firms that remain undetected for longer periods improve their mimicry and can better avoid detection. Thus, both green and greenwashing firms exhibit path dependencies. Similarly, investors look at a firm's performance history before an investment decision. The same can be assumed for firms when adopting the strategy of another firm. In addition, strategic change will not occur if profits are expected to increase only insignificantly. Thus, firms are likely to switch strategies if another firm's average profit over the past periods exceeds its own average profit by a certain margin.

4. Extended Model: An Agent-Based Approach with Local Interactions

To account for these simplifying assumptions, we extend the simple model using an agent-based model to test the robustness of the previous results.[10] This allows us not only to generate a more realistic model but to study the evolution of the strategy distribution and patterns over time.

4.1. Assumptions and Definitions

The simulations are initialized with a particular spatial structure which will be discussed further below. Firms are represented by agents, are initialized with a random strategy and seeded with an initial profit history equal to m time periods. The profit history is then given by $\pi_i^h = \{\pi_i(-m), \pi_i(-m+1), \ldots, \pi_i(0)$. Each $\pi_i(t) \in (\phi - \varepsilon, \phi + \varepsilon)$ for $t \leq 0$. Each agent interacts only with the set of neighbors to which they are connected. Each interaction period t is composed of a sequence of the following four stages: Profit generation → Strategy

more sophisticated than in our simple model.

[10]Ohtsuki and Nowak (2006) extend the replicator equations in (4) to account for non-random matching. However the approach requires a regular network (i.e. all firms have the same number of peers) as well as strategy related payoffs that show how well a strategy does against itself and each of its competing strategies, which does not apply in this context.

update → Detection and punishment → Update memory and costs.

Profit generation. Consistent with Gaio and Raposo (2011); Al Ani and Chavali (2023); Hirsch et al. (2021), we assume that external investments and profits are correlated with past profits. The profit of each firm i is based on its average profit of the previous m periods relative to the lowest and highest of the average profits in the population of firms of size n. Let $\bar{\pi}_i(t) = \sum_{r=1}^{m} \pi_i(t-r)/m$ be a firm i's average profit over the last m periods and $\hat{\pi}_i^*(t) = \max_{i \in n}\{\bar{\pi}_1(t), \bar{\pi}_2(t), \ldots, \bar{\pi}_n(t)\}$ be the maximum average profit in this period in the firm population. Equivalently, let $\check{\pi}_i^*(t) = \min_{i \in n}\{\bar{\pi}_1(t), \bar{\pi}_2(t), \ldots, \bar{\pi}_n(t)\}$ be the minimum average profit across all firms in t. We define the individual gross profit of firm i as

$$\phi_{i,t} = \varphi \frac{\bar{\pi}_i(t)}{\frac{\hat{\pi}_i^*(t)+\check{\pi}_i^*(t)}{2}} \tag{12}$$

Furthermore, we assume that the cost of green investment c_i depends on the earlier strategy of a firm, since the initial costs of green investment can decrease over time (Kesidou and Demirel, 2012; Zhou and Gu, 2019; Faber et al., 2022). Let the initial cost premium be constant $\hat{c}$ and let l_i be the number of periods that firm i engaged in investment in green technology, and k_i the number of periods that the firm used brown technology. We then define the cost of green investment as

$$c_{i,t} = \sup\{\inf\{\hat{c} - \delta\left(l_i - k_i\right), \hat{c}\}, \frac{\hat{c}}{2}\} \tag{13}$$

where δ defines the impact of each period which we set to 10^{-2} in the simulations. In other words, the cost decreases with each period firm i invests in green technology and increases in each period it invests in brown technology, but the individual cost does not exceed $\hat{c}$ or is less than $\hat{c}/2$. We then redefine the profit equations (1) to the following:

$$\pi_i(G,t) = (1 + \sigma - c_{i,t} + \epsilon_{i,t})\,\phi_{i,t} \tag{14a}$$

$$\pi_i(B,t) = (1 + \epsilon_{i,t})\phi_{i,t} \tag{14b}$$

$$\pi_i^{nd}(W,t) = (1 + \sigma - \mu + \epsilon_{it})\,\phi_{i,t} \qquad \text{if not detected} \tag{14c}$$

$$\pi_i^{d}(W,t) = (1 - \rho + \epsilon_{i,t})\,\phi_{i,t} \qquad \text{if detected} \tag{14d}$$

Any parameter without a subscript is a constant and does not change over time. The profit variation given by ϵ_{it} with mean equal to zero is randomly assigned to each firm in each period. Note that the population averages of the payoffs defined in (14) are equal to those defined in (1).

Strategy update. We assume that instead of being radmly matched with any other firm in the population, firms interact with a strict subset of other firms (Fiegenbaum and Thomas, 1995; Gao et al., 2024; Massini et al., 2005; Ma et al., 2024). Let N be the set of all firms, and firm j be a neighbor of firm i, denoted by $i \sim j$, if they are situated at adjacent cells in a regular two-dimensional lattice g_l wrapped around a torus or are connected via a link in a network g_n and let the set of neighbors of firm i be defined by $\mathfrak{N}_i = \{j \in N : j \sim i\}$.[11] Furthermore, let $\mathfrak{N}_i^+ = \{j \in N : j \sim i, s_j = G,\}$ and $\mathfrak{N}_i^- = \{j \in N : j \sim i, s_j = W\}$ be the set of neighbors of i that are green or greenwashing, respectively. The number of neighbors who are greenwashing or are green, respectively given by $|\mathfrak{N}_i^-|$ and $|\mathfrak{N}_i^+|$. The highest payoff in firm i's neighborhood is then given by $\bar{\pi}_i^*(t) = \max_{j \in \mathfrak{N}_i \cup \{i\}}\{\bar{\pi}_j(t)\}$ and the associated strategy is $s_{\mathfrak{N}}^*(t) = \{s_j | j \in \mathfrak{N} \cup \{i\}, \pi_j(s_j,t) = \bar{\pi}_i^*(t)\}$.[12]

Similarly, the average payoff of greenwashing neighbors is $\bar{\pi}_i^-(t) = \sum_{j \in \mathfrak{N}_i^-} \pi_j(t)/|\mathfrak{N}_i^-|$ and of green neighbors is $\bar{\pi}_i^+(t) = \sum_{j \in \mathfrak{N}_i^+} \pi_j(t)/|\mathfrak{N}_i^+|$. We then assume that each firm recalls the

[11] The torus avoids differences at the edges and corners by ensuring that each cell is surrounded by eight other cells

[12] In the lattice configuration, we assume that the adjacent cells and therefore neighbors are defined by the Moore neighborhood. Each firm can therefore have 8 neighbors or less if not all lattice positions are filled.

composition and average payoffs in its neighborhood in the past k periods, given by memory $\kappa_i = \{\xi_i(t), \xi_i(t-1), \ldots, \xi_i(t-k)\}$. For $s \leq 0$, we assume that the period's neighborhood corruption value is $\xi_i(s) = 0$, and for all other periods,

$$\xi_i(t) = \frac{|\mathfrak{N}_i^-|}{|\mathfrak{N}_i^+|} \frac{\bar{\pi}_i^-(t)}{\bar{\pi}_i^+(t)} \tag{15}$$

The corruption value is then the relative number of green to greenwashing neighbors weighted by their respective average payoff.[13] In each period, firm i calculates an average neighborhood corruption value across the k last memories, giving $\bar{\kappa}_{i,t} = \sum_{l=0}^{k} \xi_i(t-l)/k$. Furthermore, let $\omega_{i,t} = 1 - \bar{\pi}_i(t)/\bar{\pi}_i^*(t)$ be the baseline probability. We have the following imitation rules:

If firm i has been playing $s_i(t) = G$ in the current period:

$$s_i(t+1) = \begin{cases} W, & \text{if } s_{\mathfrak{N}}^*(t) = \{W\} \text{ and } \omega_{i,t} + \bar{\kappa}_{i,t} - \zeta > \eta_{i,t} \\ B, & \text{if } s_{\mathfrak{N}}^*(t) = \{B\} \text{ and } \omega_{i,t} - \zeta > \eta_{i,t} \\ G, & \text{otherwise} \end{cases} \tag{16}$$

The parameter ζ defines an inertia to ensure that a strategy switch only occurs if payoff differences are sufficiently large. The random variable $\eta_{i,t} \in (0, 1)$ is calculated for each firm in each period. In other words, a green firm switches to greenwashing with a probability equal to the baseline probability, which depends on its payoffs relative to the highest greenwashing payoff, and the average neighborhood corruption value minus the inertia. Thus, firms are more likely to greenwash if their neighborhood has a history of greenwashing.

[13] Equation (15) models pro-environmental behavior on the basis of peer effect and monetary incentives in line with empirical literature (see, for example, Chen et al., 2023; Gu and Shen, 2022; Maris et al., 2022).

If firm i has been playing $s_i(t) = W$ in the current period:

$$s_i(t+1) = \begin{cases} G, & \text{if } s^*_{\mathfrak{N}}(t) = \{G\} \text{ and } \omega_{i,t} - \zeta > \eta_{i,t} \\ B, & \text{if } s^*_{\mathfrak{N}}(t) = \{B\} \text{ and } \omega_{i,t} - \zeta > \eta_{i,t} \\ W, & \text{otherwise} \end{cases} \tag{17}$$

and lastly, if firm i has been playing $s_i(t) = B$ in the current period:

$$s_i(t+1) = \begin{cases} G, & \text{if } s^*_{\mathfrak{N}}(t) = \{G\} \text{ and } \omega_{i,t} - \zeta > \eta_{i,t} \\ W, & \text{if } s^*_{\mathfrak{N}}(t) = \{W\} \text{ and } \omega_{i,t} - \zeta > \eta_{i,t} \\ B, & \text{otherwise} \end{cases} \tag{18}$$

Detection and Punishment. While greenwashing implies reduced costs compared to firms investing in green technology, the former firms run a risk of their deceptive strategy being detected by investors and consumers. The likelihood of detection depends on public awareness. We may assume that awareness is affected by exogenous factors, such as labels or improved transparency rules. For simplicity and consistency, however, we assume that the probability of detection depends on the strategy composition of the entire firm population as well as the individual history of the firm.[14] The likelihood of detection and hence the expected profit of a greenwashing firm thus depend both on a public and private signal. As before, let

$$\hat{\alpha}_t = \frac{y(t)^\beta}{x(t)^\beta + y(t)^\beta} \tag{19}$$

In addition, we assume that firms, which have previously been reported as greenwashing, are

[14] We may argue that the introduction of transparency rules themselves depend on the prevalence of greenwashing and are therefore covered by our assumptions.

more likely to be detected again since they will be under tighter scrutiny. Let

$$\tilde{\alpha}_{it} = \gamma \, \frac{r_i}{t} \tag{20}$$

where parameter γ is a report impact multiplier and r_i defines the number of previous reports on the firm, i.e. the times the firms has been identified as a greenwasher in the past history of interactions. In contrast, the ability to mimic a green firm increases with the time the firm remains undetected. We choose the following simple relation

$$\nu_{it} = \inf\{t - t_{id}, 100\}/100 \tag{21}$$

where t_{id} denotes the last period that firm i was detected. Mimicry efficiency increases linearly in each period up to 100 interaction periods at which point, $\nu_{it} = 1$. With each detection, the mimicry efficiency is reset. Let $\alpha_{it} = \hat{\alpha}_t + \tilde{\alpha}_{it}$ and the probability of detection

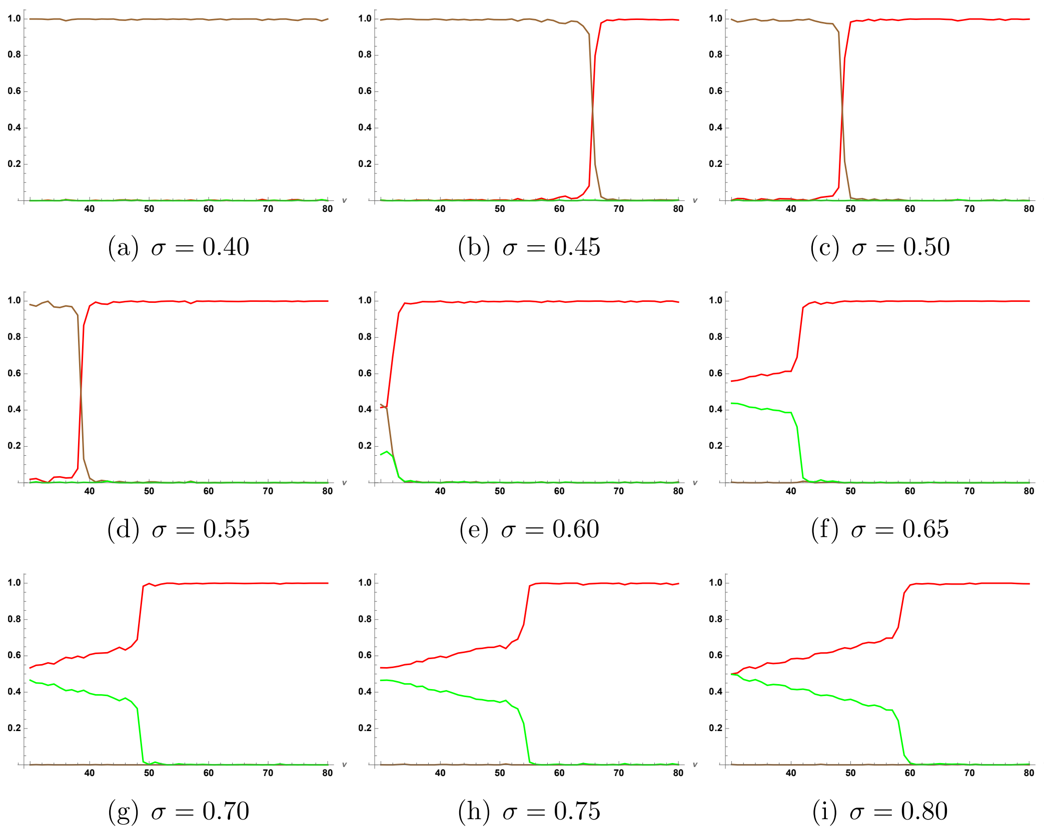
$$\theta_{it} = \sup\{\inf\{\alpha_{it} - \nu_{it}, 1\}, 0\} \tag{22}$$

Update memory and costs. Each individual firm's profit is then calculated based on the strategy chosen according to (16)-(18) and given (22), while the firm's payoff is given by (14). Each firm records the strategy distribution and payoffs in its neighborhood. If a firm i has been detected as a greenwashing firm, it will receive payoff $\pi^d_{i,t}(W)$ for l^p periods.

4.2. Simulation Results

To better understand how the segmented interaction on a lattice and the other assumptions affect previous results, we assume for the moment that firms do not benefit from history-dependent efficiencies and both cost $c_{i,t}$ and v_{it} are constant. We further set $l^p = 1$.[15]

[15] Note that we did not fully test the robustness of our results. While results hold across the large variety of parameter combinations for which we ran simulations, we did not conduct a systematic test for all combinations due to computational limitations.



(a) $\sigma = 0.40$ (b) $\sigma = 0.45$ (c) $\sigma = 0.50$

(d) $\sigma = 0.55$ (e) $\sigma = 0.60$ (f) $\sigma = 0.65$

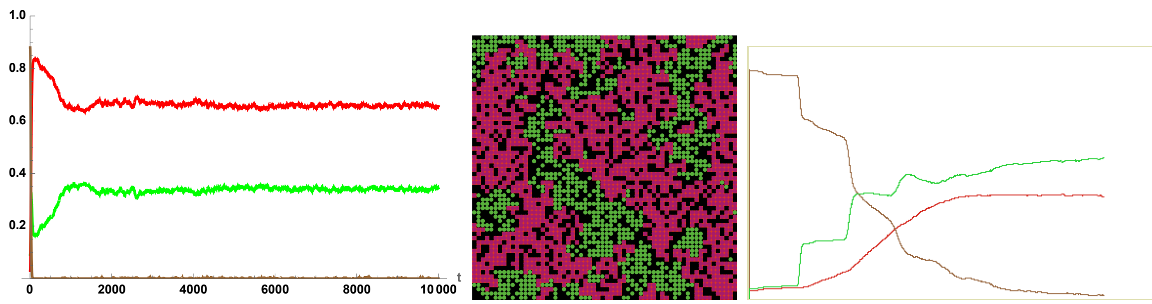
(g) $\sigma = 0.70$ (h) $\sigma = 0.75$ (i) $\sigma = 0.80$

Figure 3: Equilibrium distributions for different profit premia and effectiveness of mimicry $\nu \in [30, 80; 1]$. Initialization with $\varphi = 100$, $z(0) = 0.93$, $x(0) = 0.04$, $y(0) = 0.03$, $c = 0.60$, $\mu = 0.4$, $\rho = 0.1$, $\varepsilon = 0.05$.

Constant costs and mimicry efficiency. Figure 3 presents the simulation results after 10 000 interaction periods between 2690 firms to ensure that the firm population settled into a quasi-stable distribution. In all graphs, brown indicates the brown strategy, red the greenwashing strategy, and green defines the green strategy. The system is initialized with a majority of brown firms and only 4% green firms and 3% greenwashing firms.[16] The reason is to start with a baseline scenario in which firms predominantly had no environmental concerns. Firms interact with their Moore neighborhood, are positioned on a lattice wrapped around a torus, and therefore each firm has eight adjacent cells. To introduce additional randomness in the number of neighbors, only 70% of all cells are occupied. Since population averages of the payoff functions in (14) are equal to (1), the main differences between the closed-form

[16] The length of the profit history is set to 8 in all simulations. Longer periods have only an insignificant impact on the results. Figure present simulation averages.

conditions (8)-(11) and the simulation results in Figure 3 are caused by the differences in the updating algorithm. The system behaves as predicted by the closed-form solution for $\sigma \leq 60$ and the points of transition between regimes are as defined (10)-(11), respectively. While the transition points are still correctly predicted by (8)-(9) for larger values of the baseline profit, we can see that the share of green (greenwashing) firms exceeds (falls behind) the predicted value in (7). The reason is presented by the sample run in Figure 4(a)-(b).

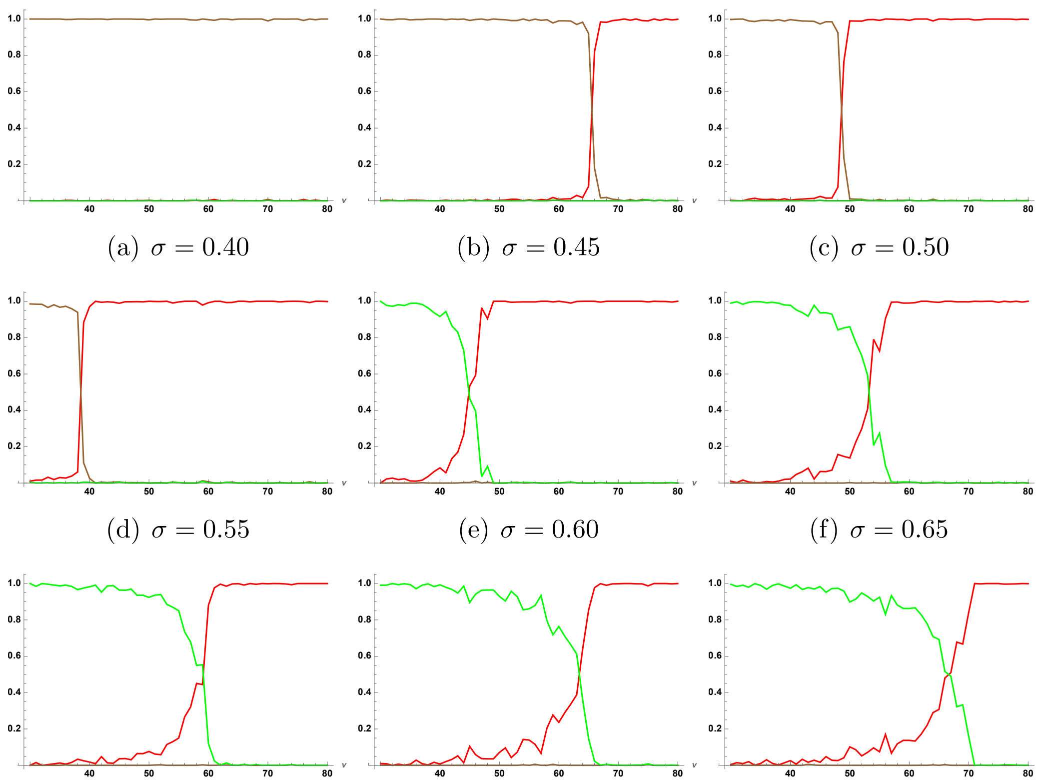


(a) Distribution $t \in (0, 10\,000)$ (b) Final distribution (c) Distribution $t \in (0, 3500)$

Figure 4: Strategy distribution for $\sigma = 75$ and $\nu = 0.50$

Initially, firms quickly adopt a greenwashing strategy and the level of public suspicion $\hat{\alpha}$ spikes at 0.84 in period 130 limiting profits for greenwashing firms (not shown). During this time, small clusters of green firms can sustain themselves and benefit from the high levels of public suspicion. They increase in size, boosting the number of green firms from 15% to roughly 33% (the predicted share is 5.5%) reducing public suspicion levels to approximately 66%. At this point, the firm population enters only a quasi-stable state. While the strategy distributions remain roughly constant, the spatial distribution slowly but gradually shifts as the clusters of green firms move across the lattice on the torus. We can see that segmentation supports the evolution of green investment by slowing down the adoption of other strategies and allowing for the evolution of successful green firms.

History-dependent costs and mimicry efficiency. Figure 5 shows the results if we relax the assumption that costs are constant but instead, are defined by (13). Results are unaffected for $\sigma < 60 = \hat{c}$ but higher payoff premia change both the transition points and the strategy distributions. Since costs are history dependent, green firms experience higher



(a) $\sigma = 0.40$ (b) $\sigma = 0.45$ (c) $\sigma = 0.50$

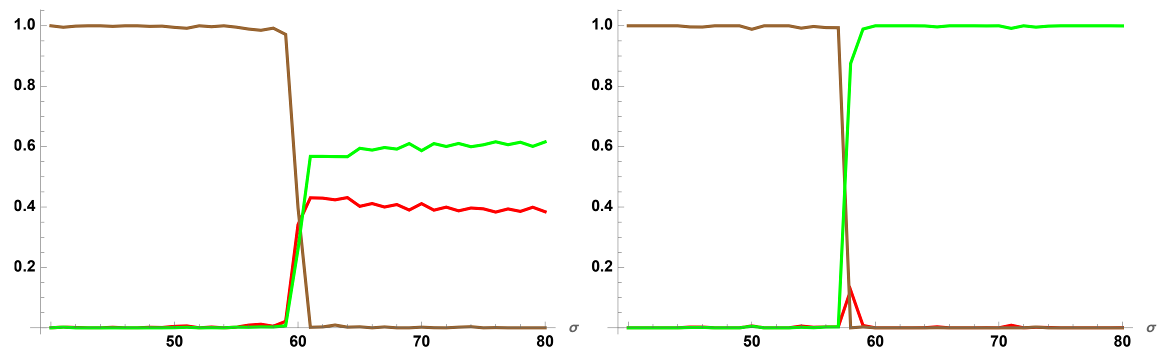
(d) $\sigma = 0.55$ (e) $\sigma = 0.60$ (f) $\sigma = 0.65$

(g) $\sigma = 0.70$ (h) $\sigma = 0.75$ (i) $\sigma_i = 0.80$

Figure 5: Dynamics for parameter values as in Figure 3 and dynamic costs according to (13)

profits in later periods which significantly boosts the share of green firms given a fixed effectiveness of mimicry. Figure 4(c) shows the convergence to one of the mixed equilibria in which green firms benefit from dynamic costs and exceed the number of greenwashing firms. The figure shows that the clustering effect can create intermittent shifts during which brown firms rapidly adopt green technology followed by a gradual but slower rise of greenwashing firms. Figure 6 then studies the effect of dynamic mimicry as well as the joint impact of both dynamic mimicry and dynamic costs. Since mimicry is endogenously defined by (21), results only vary based on differences in the payoff premia. A regime transition occurs at around $\sigma = 0.60$ as before.[17] Dynamic mimicry fails to adequately boost the share of greenwashing firms, since the latter only has an incrementally increasing impact on firm payoff. During the

[17] Since firms are not forward looking in this model, firms only consider current investment costs and green investment requires that $\sigma \geq \hat{c}$.



(a) Dynamic Mimicry (21) (b) Dynamic Mimicry and Cost

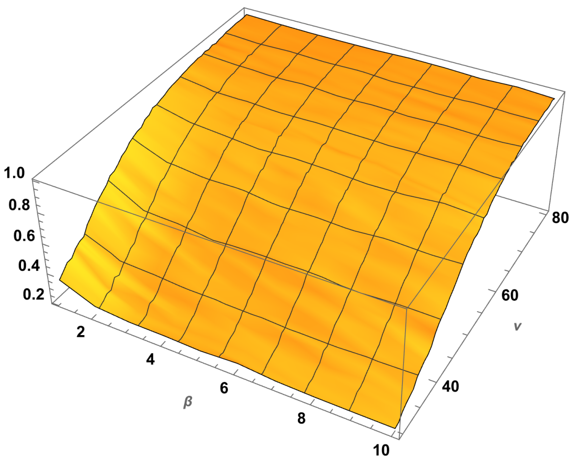
Figure 6: Population frequencies based on different payoff premia using the same parameter values as before.

initial periods, green firms have sufficient time to consolidate their profit while segmentation and the history of lower profits make it harder for greenwashing firms to catch up. The combined effect of dynamic mimicry and dynamic cost creates two simple regimes. For $\sigma < \hat{c}$, all firms are brown while for $\sigma \geq \hat{c}$, all firms eventually invest in green technology.

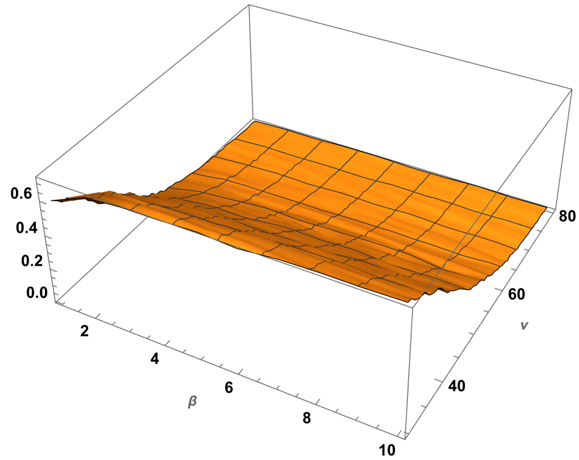
Public inertia and longer punishment. To obtain the results in the closed-form solution, we assumed that $\beta = 1$. Since results are independent of $\sigma \in [30, 80]$, Figures 7(a) and 7(b) show the impact of higher public inertiae for a single value $\sigma = 0.60$. We can see that the impact of inertia is insignificant for $\beta \geq 2$ and slightly increases the share of greenwashers and decreases the share of brown companies when moving from $\beta = 1$ to $\beta \geq 2$. At low payoff premia, a longer punishment period l^p reduces the share of greenwashing firms, but only leads to an increase in brown firms and therefore fails to encourage green investment (see Figures 7(c) and 7(d)). However, longer punishment periods become ineffective at larger payoff premia. At $\sigma \geq 0.65$, increasing the punishment period has only a negligible impact on the strategy distribution in the firm population (not shown).

4.3. Network Effects under Dynamic Mimicry and Dynamic Costs

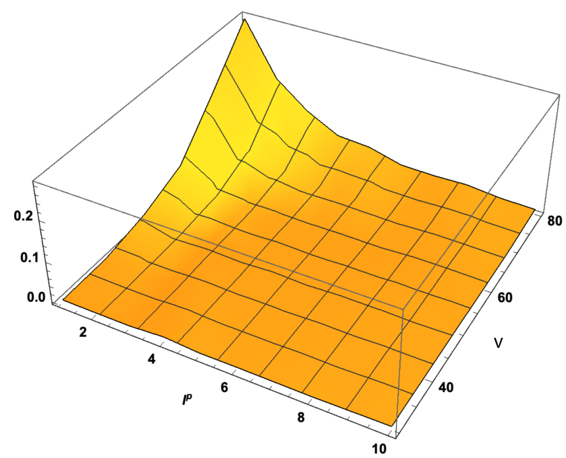
While an in-depth analysis goes beyond the scope of this paper, we tested the robustness of our results for different random networks. These random networks exhibit properties,



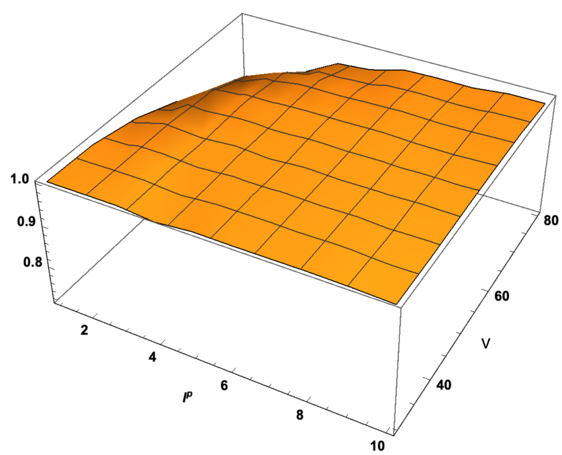
(a) Share of greenwashers y



(b) Share of brown firms z



(c) Share of greenwashers y



(d) Share of brown firms z

Figure 7: Figures 7(a) and 7(b) show the impact of changes to inertia β, given profit premium $\sigma = 0.60$ showing only small initial effect at $\beta = 2$. Figures 7(c) and 7(d) show the impact of changes to punishment duration, given profit premium $\sigma = 0.45$

which can be found in firm alliance networks and in internationalizing firms.[18]

These networks are generated using the dynamics similar to those in social networks and therefore, provide a more realistic representation than the lattice topology. The networks are created as follows:

[18] These networks exhibit small world properties (small length of shortest path between two firms, while the network demonstrates a low density and a high clustering coefficient), see Baum et al., 2010; Gaonkar and Mele, 2023, and the degree distribution follows a power-law, see Rezende et al., 2022. For a study of corporate networks in Malaysia, see Abdul Wahab et al., 2020 and for a study of these networks in Kuwait, see Alebrahim and Alnahedh, 2024.

- Type 1: Each firm is asked to create a link with a randomly chosen other firm returning a regular network of size 1.
- Type 2: A randomly chosen firm is asked to create a link with another randomly chosen firm. The process is repeated $n = 500$ times where n equals the the population size.
- Type 3: Similar to Type 2, but the process is repeated a fixed amount of times. In the simulations, this number was set to $m = 2400$. This process creates a classic Erdős-Rényi random network.
- Type 4: Firms are ordered in a line and each firm randomly connects to each of the firms further to its right each with a random probability of 5%.
- Type 5: In each period, a new firm is added to the network and connects to an incumbent firm with a probability proportional to its degree (number of links) creating a network of preferential attachment (see Yule, 1925; Barabási and Albert, 1999 and Ille, 2022, Ch. 6).[19]

We only studied the impact of the different types of random networks under dynamic mimicry and under both dynamic mimicry and dynamic costs and compared the results to those in Figure 6. Simulations ran for 500 000 periods to allow for enough time to settle on an equilibrium. In case of dynamic mimicry but fixed costs, all firms remain brown for profit premia of $\sigma < 60$. The only minor exception are Type 5 networks (preferential attachment) where we observe a more gradual regime transition. Instead of the sudden transition at around $\sigma = 60$, the share of brown firms starts declining at $\sigma = 50$ and reaches a value close to zero at $\sigma = 65$. After the regime transition, results are identical to 6(a) for Type 2 and 3 networks. In case of a Type 1, 4 or 5 network, the strategy distribution does not entirely settle on an equilibrium but fluctuates close to $(x = 0.5, y = 0.5)$. In the case in which firms benefit both from increasing efficiencies in mimicry and decreasing costs, we obtain the same

[19] This type of network creates a Matthew effect leading to a power law distribution of the degrees.

results as in Figure 6(b) with the only difference that regime transition occurs at around $\sigma \approx 48$ instead of $\sigma = 58$ as in the other scenarios. Consequently, different network topologies seem to have a very limited impact on the long-term strategy distribution, supporting our results in 6. However, these networks encourage green production for profit premia below the cost of green production (i.e. in the vicinity close to the critical value of the regime transition) if firms benefit from dynamic costs and dynamic mimicry and further, slow down the rates of convergence to an equilibrium.

5. Interpretation and Implication of the Results

Both the evolutionary game theoretic model as well as the extended model generate a number of insightful results. The conditions of the closed-form solution show the critical role of investment costs, which must be lower than the payoff premium linked to green investment (Regime 1: $\sigma > c$).[20] When interpreting this result, two important points need to be acknowledged: in our simple model, agents are not forward-looking. Investor support depends only on the history of past profits. In addition, we deliberately kept the meaning of the payoff premium σ vague. While the latter can be interpreted as immediate gains in profits, it is more realistic to understand these as the present value of future profit premia. In addition, premia should not be seen strictly as pecuniary in nature, but should also include reputation benefits that facilitate future investment and sales. However, condition (9) shows that the ability to mimic green production must be simultaneously curtailed and that at low cost of green production, increasing the obstacles or cost of mimicry can encourage the adoption of green production more strongly than increasing the profit premium of green production or the penalty after greenwashing is discovered.[21] In addition, conditions (10)-(11) indicate that if the cost of green production is high (Regime 2: $\sigma < c$), increasing the

[20] The conclusion is straightforward and the literature on reducing the cost relative to the benefits of green investment is ample, especially in relation to access and cost of capital (e.g.,Karpf and Mandel, 2018; Schnabel, 2023; Bacchiocchi et al., 2024).

[21] For $\nu_1^* = (\sigma - c + \rho)/(\sigma - \mu + \rho)$, we have $\partial\nu_1^*/\partial\sigma = \partial\nu_1^*/\partial\rho = (c - \mu)/(\sigma - \mu + \rho)^2$, $\partial\nu_1^*/\partial c = -1/(\sigma - \mu + \rho)$, $\partial\nu_1^*/\partial\mu = (\sigma - c + \rho)/(\sigma - \mu + \rho)^2$.

profit premium may only encourage greenwashing.[22]

The simulations of the extended agent-based model provide more nuanced results. Segmented interactions tend to support green investment, especially in conjunction with path-dependent cost efficiencies (see Figure 4(a)-(b) and Figure 5(e)-(i) compared to Figure 3(e)-(i)). The segmentation shields sectors of green firms during the initial phase of production, which enables the former to benefit from lower costs over time. We can see that combining measures to increase cost efficiency and delay the effectiveness of mimicry has the strongest effect (Figure 6). This result holds across various network topologies and is thus independent of the degree and type of segmentation of the market in which firms interact (see discussion in 4.3). Consequently, measures that affect the barriers to greenwashing at an early stage can be effective, even if firms are eventually able to evade these hurdles and become more effective in mimicking green behavior. The simulations also demonstrate that longer penalties have a limited effect on the firm's decision (Figure 7(c)-(d)). Consequently, curtailing the effectiveness of mimicry (i.e. improving transparency) can prove more impactful than longer or stronger penalties, since the latter may fail to encourage green investment in the first place.

Based on these results, we suggest several recommendations and demonstrate which sets of policies are especially effective in a) curtailing greenwashing, b) encouraging green investment and production by reducing costs, and c) supporting segmentation and peer effects that encourage green production and discourage greenwashing. It is straightforward to see that only focusing on increasing the profit premium of green production can inadvertently incentivize greenwashing if mimicry is not curtailed. We further observed that while heavier penalties can deter greenwashing, in contrast to peer effects and market segmentation, they may fail to encourage green investment. These results suggest a trio-targeted policy approach that implements financial opportunities for green firms to reduce the cost of green production and investment, in addition to imposing measures that increase the cost and

[22]For $\nu_2^* = \rho/(\sigma - \mu + \rho)$, we have $\partial \nu_2^*/\partial \sigma = -\partial \nu_2^*/\partial \mu = \rho/(\sigma - \mu + \rho)^2$, $\partial \nu_2^*/\partial \rho = (\sigma - \mu)/(\sigma - \mu + \rho)^2$.

curtail opportunities to mimic green behavior, and strategies supporting segmentation and peer effects.

Measures to reduce the overall costs are tax incentives and subsidies, especially if the latter are applied during the early-adoption stage, which helps alleviate the burden of high upfront costs. R&D tax credits help reward firms for incremental improvements in green technology efficiency and reduce production costs over time, strongly supporting a green economy. These measures counteract the initially high costs and help reinforce green production through path dependency. In addition, they can be adjusted as green technology matures (e.g., policymakers can phase out subsidies gradually to avoid dependency while maintaining anti-mimicry measures) in the form of tiered subsidies that are linked to specific green targets.

Since our model has shown that improving transparency and exposure in conjunction with cost-reducing measures can prove highly efficient in reducing greenwashing viability (see also Nieto and Papathanassiou, 2025 for a similar result), policies at the larger institutional level, such as the UK's Green Claims Code (CMA, 2021) or the EU's recent changes to consumer law, transparency and verification requirements (EP, 2024b,a) are steps in the right direction. At the firm level, policymakers should emphasize transparency measures that require firms to disclose audited data on their environmental impact (such as the EU's 'Green claims' directive[23]) along with replacing self-certified labels with government-audited certificates based on standardized guidelines over penalties (see again EP, 2024a). On the consumer side, educating stakeholders via prebunking and ESG campaigns will increase public scrutiny and pressure. Detection of greenwashing can be further improved by using machine learning tools that uncover inconsistencies in ESG claims, help predict greenwashing behavior (for an example, see Zeng et al., 2025) and generate more effective ESG rating systems (for an overview, see Huang et al., 2024).

[23] https://www.europarl.europa.eu/RegData/etudes/BRIE/2023/753958/EPRS_BRI(2023)753958_EN.pdf

In our model, the segmented interaction between firms supports green production via two channels: first, it localizes competition and thereby reduces competitive pressure on a firm by diminishing the number of firms with which it compares profits. Second, a firm's decision is also affected by the prevalence of greenwashing among its neighbors. Consequently, green investment and production can benefit from sector-specific initiatives that promote industry coalitions of green companies (e.g., green supply chain partnerships) which will strengthen peer pressure against greenwashing (e.g., Walmart's Project Gigatron, Unilever's Sustainable Sourcing Policy, Redwood Hyperion, Snowkap's ESG Hub). These initiatives can be supported by geographically concentrated initiatives, which help leverage network effects and thus, simultaneously help reduce the cost of green production while increasing peer scrutiny (e.g., RegioGreenTex Hubs, STRING Megaregion, Greater Copenhagen's green initiatives).

While the results of the dynamic models do not advocate for a new set of policies, we show the need for a more holistic approach to curtail greenwashing. Policies are efficient in reducing the incentives for greenwashing and mitigating the barriers to a green economy if they improve the benefit-to-cost ratio of green production over time while simultaneously curtailing the ability to mimic green investment. They can be additionally supported by measures that strengthen peer-pressure and market segmentation. Therefore, similar to Jaumotte et al. (2024), we advocate for an approach that combines early temporary subsidies with better regulatory and screening frameworks that focus on transparency, coupled with measures that promote industry coalitions and network effects.

6. Conclusion

In our simple dynamic model, firms choose between investing in green technology, continuing brown production, or only mimicking investment in green technology and production. We derived the equilibrium conditions for the model with random matching and showed which conditions can completely stifle green production either by discouraging any such production in the first place (i.e., most firms remain brown) or by encouraging spurious green investment in the form of greenwashing. We then extended the simpler model to fit a more realistic interaction environment with dynamic costs and efficiencies in mimicry. While localized interactions encourage the evolution of green firms, indicating that market segmentation can support green investment, we show that the simpler model with random matching offers good approximations of the equilibrium conditions in more complex setups. In addition, we show that increasing cost efficiencies can outperform increasing efficiencies in mimicry and create path dependencies that support green production.

Our study has some limitations. Firms are not forward-looking and base their strategies on a payoff comparison with other firms. Similarly, consumers are only aware of greenwashing based on the relative prevalence of green firms. In the future, a more comprehensive agent-based model can allow for less myopic firm interactions and explicitly model the reaction of consumers. In addition, we only briefly analyze the impact of networks and limit our study to a set of random networks. Future research can model the interaction using a broader range of networks based on empirical data across different sectors. We also ignore problems related to firms fearing being mislabeled as greenwashers and therefore, do not include the strategy of greenhushing, which can be added to the model in future iterations.

Acknowledgements. We thank the editor and the anonymous reviewers for their constructive comments, which helped us improve the manuscript. This paper benefited from discussions and comments from seminar participants at the: NED 2025 - Nonlinear Economic Dynamics Workshop held in University of Napoli "Federico II", Italy, 3-5 June, 2025; Renaissance in Economics, 2nd International Conference, Manifesto for New Economy, held in the Department of Economics and Management – University of Florence (Italy) 5-6 June 2025; The Research Seminars at DISEI University of Florence. This paper benefited from the PRIN 2022 under the Italian Ministry of University and Research (MUR) Prot. 2022YMLS4T - "TEC - Tax Evasion and Corruption: theoretical models and empirical studies. A quantitative based approach for the Italian case."

Declarations. The authors did not receive support from any organization for the submitted work. The authors have no relevant financial or non-financial interests to disclose. The authors have no conflicts of interest to declare that are relevant to the content of this article.

References

Abdul Wahab, E. A., M. F. Jamaludin, D. Agustia, and I. Harymawan (2020). Director networks, political connections, and earnings quality in malaysia. *Management and Organization Review 16*(3), 687–724.

Agi, M. A. N., S. Faramarzi-Oghani, and Öncü Hazır and (2021). Game theory-based models in green supply chain management: a review of the literature. *International Journal of Production Research 59*(15), 4736–4755.

Al Ani, M. K. and K. Chavali (2023). The relationship between investment intensity and profitability measures from the perspective of foreign investors. *Humanities and Social Sciences Communications 10*(76).

Alebrahim, E. and M. Alnahedh (2024). Corporate network formation in kuwait: board interlocks network analysis using exponential random graph models. *Corporate Governance: The International Journal of Business in Society 24*(7), 1638–1662.

Bacchiocchi, A., S. Ille, and G. Giombini (2024, October). The effects of a green monetary policy on firms financing cost. *Journal of Economic Interaction and Coordination 19*(4), 727–757.

Barabási, A.-L. and R. Albert (1999). Emergence of scaling in random networks. *Science 286*(5439), 509–512.

Baum, J. A. C., R. Cowan, and N. Jonard (2010). Network-independent partner selection and the evolution of innovation networks. *Management Science 56*(11), 2094–2110.

Bernini, F., M. Giuliani, and F. La Rosa (2024). Measuring greenwashing: A systematic methodological literature review. *Business Ethics, the Environment & Responsibility 33*(4), 649–667.

Besley, T. and T. Persson (2023, 01). The Political Economics of Green Transitions*. *The Quarterly Journal of Economics 138*(3), 1863–1906.

Cai, J., H. Sun, J. Shang, and G. G. Hegde (2023). Information structure selection in a green supply chain: Impacts of wholesale price and greenness level. *European Journal of Operational Research 306*(1), 34–46.

Carmona, R. and F. Delarue (2018). *Learning by Examples: What Is a Mean Field Game?*, pp. 3–65. Cham: Springer International Publishing.

Casey, W., S. E. Massey, and B. Mishra (2021). How signalling games explain mimicry at many levels: from viral epidemiology to human sociology. *Journal of the Royal Society, Interface 18*(175).

Chen, C., D. Jiang, and W. Li (2023). Keeping up with the csr joneses: The impact of industry peers on focal firms' csr performance. *Humanities and Social Sciences Communications 10*(106).

Chen, M.-H., H. Wei, M. Wei, H. Huang, and C.-H. J. Su (2021). Modeling a green supply chain in the hotel industry: An evolutionary game theory approach. *International Journal of Hospitality Management 92*, 102716.

CMA (2021). Making environmental claims on goods and services. *Competition and Markets Authority, UK Government*.

Delmas, M. A. and V. C. Burbano (2011a). The drivers of greenwashing. *California Management Review 54*(1), 64–87.

Delmas, M. A. and V. C. Burbano (2011b). The drivers of greenwashing. *California Management Review 54*(1), 64–87.

Demirel, P., K. Iatridis, and E. Kesidou (2018). The impact of regulatory complexity upon self-regulation: Evidence from the adoption and certification of environmental management systems. *Journal of Environmental Management 207*, 80–91.

Dunz, N., A. Naqvi, and I. Monasterolo (2021). Climate sentiments, transition risk, and financial stability in a stock-flow consistent model. *Journal of Financial Stability 54*(C).

EP (2024a). Meps adopt new law banning greenwashing and misleading product information. *European Parliament News, Press Release*.

EP (2024b). Parliament wants to improve consumer protection against misleading claims. *European Parliament News, Press Release*.

Eyraud, L., B. Clements, and A. Wane (2013). Green investment: Trends and determinants. *Energy Policy 60*, 852–865.

Faber, G., A. Ruttinger, T. Strunge, T. Langhorst, A. Zimmermann, M. van der Hulst, F. Bensebaa, S. Moni, and L. Tao (2022). Adapting technology learning curves for prospective techno-economic and life cycle assessments of emerging carbon capture and utilization pathways. *Frontiers in Climate Volume 4 - 2022*.

Fiegenbaum, A. and H. Thomas (1995). Strategic groups as reference groups: Theory, modeling and empirical examination of industry and competitive strategy. *Strategic Management Journal 16*(6), 461–476.

Flammer, C. (2021). Corporate green bonds. *Journal of Financial Economics 142*(2), 499–516.

Fürer, M. (2023). Reprisk data suggest increase in greenwashing and rise of social washing. `https://www.reprisk.com/research-insights/news-and-media-coverage/reprisk-data-shows-increase-in-greenwashing-with-one-in-three-greenwashing-public-com ~:text=While%20greenwashing%20incidents%20have%20accelerated,banking%20and%20financial%20services%20sectors`. Accessed: 2024-09-30.

Gaio, C. and C. Raposo (2011). Earnings quality and firm valuation: international evidence. *Accounting & Finance 51*(2), 467–499.

Gao, J., X. Gu, and X. Yang (2024). Three-dimensional institutional equivalence: how industry, community, and network peers influence corporate innovation quality in china. *European Journal of Innovation Management ahead of print*(ahead of print).

Gaonkar, S. and A. Mele (2023). A model of inter-organizational network formation. *Journal of Economic Behavior & Organization 214*, 82–104.

Gu, C. and M. Shen (2022). Peer effect of environmental information disclosure behavior in china:empirical test based on "real green" and "fake green". *Heliyon 10*(23).

Gupta, K. and N. Singh (2024, May). A Systematic Review of Literature on Greenwashing. In *Sustainability Development through Green Economics*, Volume 114 of *Contemporary Studies in Economic and Financial Analysis*, pp. 75–90. Emerald Group Publishing Limited.

Hameed, I., Z. Hyder, M. Imran, and K. Shafiq (2021). Greenwash and green purchase behavior: an environmentally sustainable perspective. *Environment, Development and Sustainability*, 1–22.

He, Q., Z. Wang, G. Wang, J. Xie, and Z. Chen (2022). The dark side of environmental sustainability in projects: Unraveling greenwashing behaviors. *Project Management Journal 53*(4), 349–366.

Hirsch, S., D. Lanter, and R. Finger (2021). Profitability and profit persistence in eu food retailing: Differences between top competitors and fringe firms. *Agribusiness 37*(2), 235–263.

Huang, S.-M. H. S.-M., S.-H. C. S.-M. Huang, C.-F. H. S.-H. Chung, and K.-Q. Z. C.-F.

Huang (2024). Discovering greenwashing behavior patterns: A systematic literature review approach. *International Journal of Computer Auditing*.

Huang, Z., Y. Shi, and M. Jia (2025). Greenwashing: A systematic literature review. *Accounting & Finance 65*(1), 857–882.

Ille, S. (2022). *Models of Society and Complex Systems*. London: Routledge.

Ioannou, I., G. Kassinis, and G. Papagiannakis (2023, June). The Impact of Perceived Greenwashing on Customer Satisfaction and the Contingent Role of Capability Reputation. *Journal of Business Ethics 185*(2), 333–347.

Jaumotte, F., J. Kim, S. Pienknagura, and G. Schwerhoff (2024). Policies to foster green fdi: Best practices for emerging market and developing economies. *Staff Climate Notes 2024*(004), A001.

Juliana Fernandes, S. S. and J. K. Leopold (2020). When consumers learn to spot deception in advertising: testing a literacy intervention to combat greenwashing. *International Journal of Advertising 39*(7), 1115–1149.

Karpf, A. and A. Mandel (2018). The changing value of the 'green' label on the us municipal bond market. *Nature Climate Change 8*, 161–165.

Karásek, J. and J. Pavlica (2016). Green investment scheme: Experience and results in the czech republic. *Energy Policy 90*, 121–130.

Kesidou, E. and P. Demirel (2012). On the drivers of eco-innovations: Empirical evidence from the uk. *Research Policy 41*(5), 862–870.

Laufer, W. S. (2003). Social accountability and corporate greenwashing. *Journal of Business Ethics 43*(3), 253–261.

Leng, M. and M. Parlar (2005). Game theoretic applications in supply chain management: A review. *INFOR: Information Systems and Operational Research 43*(3), 187–220.

Liu, C., Y. Song, W. Wang, and X. Shi (2023). The governance of manufacturers' greenwashing behaviors: A tripartite evolutionary game analysis of electric vehicles. *Applied Energy 333*, 120498.

Liu, Z., Y.-Q. Huang, W.-L. Shang, Y.-J. Zhao, Z.-L. Yang, and Z. Zhao (2022). Precooling energy and carbon emission reduction technology investment model in a fresh food cold chain based on a differential game. *Applied Energy 326*, 119945.

Liu, Z., Q. Qian, B. Hu, W.-L. Shang, L. Li, Y. Zhao, Z. Zhao, and C. Han (2022). Government regulation to promote coordinated emission reduction among enterprises in the green supply chain based on evolutionary game analysis. *Resources, Conservation and Recycling 182*, 106290.

Lyon, T. P. and J. W. Maxwell (2011a, March). Greenwash: Corporate Environmental Disclosure under Threat of Audit. *Journal of Economics & Management Strategy 20*(1), 3–41.

Lyon, T. P. and J. W. Maxwell (2011b). Greenwash: Corporate environmental disclosure under threat of audit. *Journal of Economics & Management Strategy 20*(1), 3–41.

Ma, P., L. Cui, and D. Xu (2024). Social comparison inside business groups and strategic change: Evidence from group-affiliated chinese firms. *Journal of Management Studies 61*(8), 3533–3565.

Maris, R., Z. Dorner, and F. Carlsson (2022). Nudges and monetary incentives: A green partnership? *Working Paper in Economics No. 842 - University of Gothenburg*.

Marquis, C., M. W. Toffel, and Y. Zhou (2016). Scrutiny, norms, and selective disclosure: A global study of greenwashing. *Organization Science 27*(2), 483–504.

Martin, P. R. and D. V. Moser (2016). Managers' green investment disclosures and investors' reaction. *Journal of Accounting and Economics 61*(1), 239–254.

Massini, S., A. Y. Lewin, and H. R. Greve (2005). Innovators and imitators: Organizational reference groups and adoption of organizational routines. *Research Policy 34*(10), 1550–1569.

Nieto, M. J. and C. Papathanassiou (2025). Different shades of green: Eu corporate disclosure rules and their effectiveness in limiting "greenwashing". *Occasional Paper Series, European Central Bank* (370).

Ohtsuki, H. and M. A. Nowak (2006). The replicator equation on graphs. *Journal of Theoretical Biology 243*(1), 86–97.

Parguel, B., F. Benoît-Moreau, and F. Larceneux (2011). How Sustainability Ratings Might Deter "Greenwashing": A Closer Look at Ethical Corporate Communication. *Journal of Business Ethics 102*(1), 15–28.

Pekovic, S., G. Grolleau, and N. Mzoughi (2018). Environmental investments: Too much of a good thing? *International Journal of Production Economics 197*, 297–302.

RepRisk Report, J. . (2022). Spotting greenwashing with esg data. `https://www.reprisk.com/research-insights/reports/spotting-greenwashing-with-esg-data`. Accessed: 2024-09-30.

Rezende, S., A. Alves, D. Bicalho, and Ângela Versiani (2022). The network evolution of the internationalizing firm. *Brazilian Business Review 19*, 116–132.

Schnabel, I. (2023). Monetary policy tightening and the green transition.

Wang, Y. and G. Hou (2020). A duopoly game with heterogeneous green supply chains in optimal price and market stability with consumer green preference. *Journal of Cleaner Production 255*, 120161.

Wei Xu, Mingzhu Li, S. X. (2023). Unveiling the "veil" of information disclosure: Sustainability reporting "greenwashing" and "shared value". *PLOSOne 18*(1).

Wu, M.-W. and C.-H. Shen (2013). Corporate social responsibility in the banking industry: Motives and financial performance. *Journal of Banking & Finance 37*(9), 3529–3547.

Yadav, P. L., S. H. Han, and J. J. Rho (2016). Impact of environmental performance on firm value for sustainable investment: Evidence from large us firms. *Business Strategy and the Environment 25*(6), 402–420.

Yule, G. U. (1925). Ii.—a mathematical theory of evolution, based on the conclusions of dr. j. c. willis, f. r. s. *Philosophical Transactions of the Royal Society of London. Series B, Containing Papers of a Biological Character 213*(402-410), 21–87.

Zeng, F., J. Wang, and C. Zeng (2025, 03). An optimized machine learning framework for predicting and interpreting corporate esg greenwashing behavior. *PLOS ONE 20*(3), 1–25.

Zhang, D. (2022a). Environmental regulation, green innovation, and export product quality: What is the role of greenwashing? *International Review of Financial Analysis 83*, 102311.

Zhang, D. (2022b). Green financial system regulation shock and greenwashing behaviors: Evidence from chinese firms. *Energy Economics 111*, 106064.

Zhou, Y. and A. Gu (2019). Learning curve analysis of wind power and photovoltaics technology in us: Cost reduction and the importance of research, development and demonstration. *Sustainability 11*(8).