Returns to scale with a Cobb-Douglas production function for a small italian mechanical firm

Osti, Davide

20 November 2021

Returns to scale with a Cobb-Douglas production function for a small italian firm

Davide Osti[1*]

[1*]Economic Research, Studio Osti - Sgarzi, Via della Zecca 1, Bologna, 40121, Emilia Romagna, Italy.

Corresponding author(s). E-mail(s): devidosti@gmail.com;

Abstract

with this piece of evidence, I try to she light upon the effects of fixed and variable costs on revenues for a firm operating in the sector of leathing and milling in the neighbourhood of Bologna, on the Tuscan - Emilian Appennines, through the estimation of a linear bivariate simultaneous equation model where variable and fixed costs explain revenues; with a sample of eleven years of annual data, I find that a marginal increase in variable costs of 1 euro, keeping the fixed costs constants, leads to higher revenues up to 1.155 euro; I further estimate a cobb douglas production function, in order to find out whether the returns to scale are increasing, constant or decreasing; I find support for the hypothesis of slightly increasing returns to scale with the baseline cobb douglas transformed in logarithms (with capital and labour only), while multiplicatively including an additional regressor for raw materials purchases, I find evidence for slightly decreasing returns to scale

Keywords: firm behaviour, production functions, returns to scale, cobb douglas, stochastic frontier model, non linear least squares, production sets, convex cones

JEL Classification: C01 , C51 , C80 , C81 , C87 , C88 , D01

1 The problem of estimation: ordinary least squares

the following sections on the theory of econometric estimation have largely been borrowed by [2].

in order to represent the evolution and the causal relationship holding together variable and fixed costs on the one side, and revenues on the other side, we can set up the following simple bivariate linear simultaneous equation model: $\mathbf{Y} = \mathbf{X}'\boldsymbol{\beta} + \boldsymbol{\varepsilon} \leftrightarrow \mathbf{Y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}$, where $\underset{(T\times 1)}{\mathbf{Y}} = \begin{bmatrix} y_1 \\ \vdots \\ y_T \end{bmatrix}$; $\underset{(T\times 1)}{\boldsymbol{\varepsilon}} = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_T \end{bmatrix}$; $\underset{T\times k}{\mathbf{X}} = \begin{bmatrix} x_{11} & x_{12} \\ \vdots & \vdots \\ x_{T1} & x_{T2} \end{bmatrix}$; $k = 2$ is the number of parameters to estimate,, that is, $\underset{(2\times 1)}{\boldsymbol{\beta}} = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$ in our case in which $\mathbf{X}_1$ is a column vector with time index $t = 2011, \ldots, 2021$ containing the balance sheet value of variable costs, in which each row is the sum of all the variable costs of a single fiscal year; while $\mathbf{X}_2$ contains the values of fixed costs for each year of the time - series at hand; and, finally, $\mathbf{Y}$ is made up of the revenues of each of the 11 years of the sample.

the aim of the estimation exercise is to attribute a numerical value to the vector of parameters $\boldsymbol{\beta}$; the method of ordinary least squares chooses the numerical values for the elements of the unknown parameters' vector such as to minimize the sum of squared residuals of the regression; we define the following value:

$\underset{(T\times 1)}{\mathbf{e}(\boldsymbol{\beta})} = \underset{(T\times 1)}{\mathbf{Y}} - \underset{(T\times k)}{\mathbf{X}}\underset{(k\times 1)}{\boldsymbol{\beta}}$; if $\mathbf{X}\boldsymbol{\beta}$ can be considered a predictor of $\mathbf{Y}$, $\mathbf{e}$ is the corresponding prediction or forecast error; the sum of the squares of the residuals is given by: $\mathbf{S}(\boldsymbol{\beta}) = \mathbf{e}(\boldsymbol{\beta})'\mathbf{e}(\boldsymbol{\beta})$; the method of ordinary least squares produces an estimate of $\boldsymbol{\beta}$, with $\hat{\boldsymbol{\beta}}$ such that: $\mathbf{S}(\hat{\boldsymbol{\beta}}) = \min_{\boldsymbol{\beta}} \mathbf{S}(\boldsymbol{\beta})$

we denote the corresponding estimator of $\boldsymbol{\varepsilon}$ with $\hat{\boldsymbol{\varepsilon}} = \mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}$, such that $\mathbf{S}(\hat{\boldsymbol{\beta}}) = \hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}}$. $\hat{\boldsymbol{\varepsilon}}$ is defined as the vector of the resioduals of the ordinary least squares estimator. at this point, we can obtain the estimator, considering that the necessary and sufficient conditions allowing the existence of a vector $\hat{\boldsymbol{\beta}}$ defining a unique minumum S are the following:

i) $\mathbf{X}'\hat{\boldsymbol{\varepsilon}} = 0$;

ii) $rank(\mathbf{X}) = k$.

the first condition imposes the orthogonality between the OLS residuals and the variable included in the right hand side of the model (commonly known as regressors);

the second condition imposes that the columns of the matrix $\mathbf{X}$ are linearly independent among each others, in other words that none of the explanatory variables can be expressed as a linear combination of each others; we note that condition *i)* guarantees that the residuals of the OLS sum up to zero, and that thus have null mean, whether a constant is included within the regressors;

from *i)*, we can derive an expression for the OLS estimator:

$\mathbf{X}\hat{\varepsilon} = \mathbf{X}'(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}) = \mathbf{X}'\mathbf{Y} - \mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} = 0$
$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$

1.1 properties of the OLS estimator

we derived the OLS estimator imposing a minimum set of hypotheses; we now focus on its properties under a particular set of additional statistical hypotheses; to this end, we shall review a few concepts of mean and variance in vectors of statistical variables; for a vector of variables

$\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 & \dots & \mathbf{x}_T \end{bmatrix}'$

we define the mean vector, expressed as $\mathbb{E}(\mathbf{x})$, as well as the mean matrix of external products $\mathbb{E}(\mathbf{x}'\mathbf{x})$ as follows:

$\mathbb{E}(\mathbf{x}) = \begin{bmatrix} \mathbb{E}(\mathbf{x}_1) & \dots & \mathbb{E}(\mathbf{x}_T) \end{bmatrix}'$

where the symbol ' stands for transposed, that is, a column vector rotated of 90 degrees towards a row vector;

$$\mathbb{E}(\mathbf{x}'\mathbf{x}) = \mathbb{E}\begin{bmatrix} \mathbf{x}_1^2 & \mathbf{x}_1\mathbf{x}_2 & \mathbf{x}_1\mathbf{x}_T \\ \vdots & \vdots & \vdots \\ \mathbf{x}_T\mathbf{x}_1 & \dots & \mathbf{x}_T^2 \end{bmatrix} = \begin{bmatrix} \mathbb{E}(\mathbf{x}_1^2) & \mathbb{E}(\mathbf{x}_1\mathbf{x}_2) & \dots & \mathbb{E}(\mathbf{x}_1\mathbf{x}_T) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{E}(\mathbf{x}_T\mathbf{x}_1) & E(\mathbf{x}_T\mathbf{x}_2) & \dots & \mathbb{E}(\mathbf{x}_t^2) \end{bmatrix}$$

the variance-covariance matrix of $\mathbf{x}$ is thus defined as follows:

$var(\mathbf{x}) = \mathbb{E}(\mathbf{x} - \mathbb{E}(\mathbf{x}))(\mathbf{x} - \mathbb{E}(\mathbf{x}))' = \mathbb{E}(\mathbf{x}\mathbf{x}') - \mathbb{E}(\mathbf{x})\mathbb{E}(\mathbf{x})'$

the variance-covariance matrix is symmetric and positive definite, by construction, indeed for an arbitrary vector of dimensions T, $\mathbf{A}$, we have:

$var(\mathbf{A}'\mathbf{x}) = \mathbf{A}'var(\mathbf{x})\mathbf{A}$

the former hypothesis to derive the statistical properties of the estimator is that the various components of the sample at hand y_t, $\mathbf{x}_t'$ are extracted independently within each others: no observation can help to predict the other observations; in such a case, the hypothesis $\mathbb{E}(y_t|\mathbf{x}_t) = \mathbf{x}_t'\boldsymbol{\beta}$ becomes equivalent to:

$\mathbb{E}(y_t|\mathbf{x}_1, \dots, \mathbf{x}_t, \dots, \mathbf{x}_T) = \mathbf{x}_t'\boldsymbol{\beta}$, for $t = 1, \dots, T$, or, in vector notation,

[A.1] $\mathbb{E}(\mathbf{y}|\mathbf{X}) = \mathbf{X}\boldsymbol{\beta}$,

we note that hypothesis [A.1] is very restrictive and it applies in our empirical issue at hand only as far as the supply side shocks hitting the various competing firms are specific to these lasts (sectoral shocks) and are not correlated with the contemporaneously observed explanatory variables, for both leads and lags; if this hypothesis is applicable to the sample at hand, it will hardly be applicable to a sample of time - series; as a matter of facts, time - series are characterized by the interdependence of the observations taken in different points in time (lack of independence → autocorrelation); we should thus use a sample of cross - sectional data (different variables for each firm at a given point in time) to introduce the econometric methodology.

the second hypothesis we put forth, partially follows the first one and strengthens it, imposing a constant variance of the shocks:

[A.2] $\mathbb{E}(\varepsilon\varepsilon'|\mathbf{X}) = \sigma^2\mathbf{I}$, where σ^2 is a constant independent of $\mathbf{X}$.

[A.3] $rank(\mathbf{X}) = k$;

under hypotheses [A.1]-[A.3] we can now derive the properties of the OLS estimator.

property 1 $\longrightarrow$ the estimator $\exists$, in fact [A.3] guarantees that $(\mathbf{X}'\mathbf{X})$ is invertible.

property 2 $\longrightarrow$ the mean of the estimatore, conditioned with respect to $\mathbf{X}$, is $\boldsymbol{\beta}$; the OLS estimator is unbiased. indeed, we have:

$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon})$

$= \boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}$

$\rightarrow \mathbb{E}(\hat{\boldsymbol{\beta}}|\mathbf{X}) = \boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \underbrace{\mathbb{E}(\boldsymbol{\varepsilon}|\mathbf{X})}_{=0}$

$= \boldsymbol{\beta}$, by hypothesis [A.1].

we point out that the result is valid for each $\mathbf{X}$, thus, even the unconditional mean of the OLS estimator does not coincide with the vector of parameters to be estimated;

property 3 $\rightarrow$ the variance of the OLS estimator, conditioned on $\mathbf{X}$, is $\sigma^2(\mathbf{X}'\mathbf{X})^{-1}$:

$var(\hat{\boldsymbol{\beta}}|\mathbf{X}) = \mathbb{E}\big[(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})'|\mathbf{X}\big]$

$= \mathbb{E}\big[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}|\mathbf{X}\big]$

$= \big[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbb{E}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'|\mathbf{X})\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\big]$

$= \big[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\sigma^2\mathbf{I})\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\big]$

$= \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$

we note that the expression for the conditional variance depends on $\mathbf{X}$, therefore it does not coincide with the non conditional variance.

property 4 (GAUSS-MARKOV THEOREM) $\rightarrow$ the OLS estimator is the estimator with minimal conditional variance within the class of linear unbiased estimators (UMVUE and BLUE).

this property is important for it shows the optimality of the OLS estimator with respect to a well defined criterion; we saw that the estimator is unbiased, it is thus natural to think of the optimality with respect to the variance of the estimator; the estimator with minimum variance is the most *efficient*, in the sense that is employs the information contained in the data in the most efficient way.

let us consider the class of linear estimators:

$\boldsymbol{\beta}_L = \mathbf{L}\mathbf{y}$

such a class is defined within the set of matrices $\mathbf{L}$ of dimensions $k \times T$, that are fixed when conditioning upon $\mathbf{X}$; $\mathbf{L}$ can be constant, can depend on $\mathbf{X}$, but cannot be function of $\mathbf{X}$; therefore:

$var(\boldsymbol{\beta}_L|\mathbf{X}) = \mathbb{E}(\mathbf{L}\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\mathbf{L}'|\mathbf{X}) = \sigma^2\mathbf{L}\mathbf{L}'$

we notice that, because $\mathbf{L}$ can depend on $\mathbf{X}$, the expression for the non conditional varianc does not generally coincide with that for the conditional variance.

at this point, we are ready to show that the OLS estimator is the most efficient within the class of linear unbiased estimators, by showing that the conditional variance of the OLS estimator differs from that of any other estimator within the class for a positive semi-definite matrix[1].

let us define $\mathbf{D} = \mathbf{L} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$; $\mathbf{LX} = \mathbf{I}$; we wish that $\mathbf{DX} = 0$:

$\mathbf{LL}' = ((\mathbf{X}'\mathbf{X})^{-1} + \mathbf{D})(\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} + \mathbf{D}')$

$= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{D}' + \mathbf{DX}(\mathbf{X}'\mathbf{X})^{-1} + \mathbf{DD}'$

$= (\mathbf{X}'\mathbf{X}^{-1}) + \mathbf{DD}'$

from which it follows that:

$var(\boldsymbol{\beta}_L|\mathbf{X}) = var(\hat{\boldsymbol{\beta}}|\mathbf{X}) + \sigma^2\mathbf{DD}'$

which shows that the symmetric matrix $\mathbf{DD}'$ is positive semi-definite, for every matrix $\mathbf{D}$, not necessarily squared

1.2 Analysis of the residuals

in order to produce the analysis of the residuals, let us consider the following representation:

$\hat{\boldsymbol{\varepsilon}} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{y} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \mathbf{My}$

where $\mathbf{M} = \mathbf{I}_T - \mathbf{Q}$ and $\mathbf{Q} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$; the matrices $T \times T$ $\mathbf{Q}$ and $\mathbf{M}$ have the following properties:

i) are symmetric, $\mathbf{Q} = \mathbf{Q}'$

i) are idempotent, $\mathbf{QQ} = \mathbf{Q}$, $\mathbf{M} = \mathbf{M}'$;

iii) $\mathbf{M}$ is orthogonal to $\mathbf{X}$ ($\mathbf{MX} = 0$), $\mathbf{M}$ is orthogonal to $\mathbf{Q}$ ($\mathbf{MQ} = 0$), $\mathbf{QX} = \mathbf{X}$.

notice that the OLS forecast for $\mathbf{y}$ can be written as $\mathbf{Qy}$, $\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{Qy}$, notice also that $\hat{\boldsymbol{\varepsilon}} = \mathbf{My}$, from which the result follows, already commented, of orthogonality between the OLS residuals and the OLS predictors; we further have that $\mathbf{My} = \mathbf{MX}\boldsymbol{\beta} + \mathbf{M}\boldsymbol{\varepsilon}$, since $\mathbf{MX} = 0$. there exists a well precise relation between the OLS residuals and the errors in the econometric model, that cannot however be used to get the errors given the residuals, as far as the matrix $\mathbf{M}$ is singular and thus not invertible.

we have:

$S(\hat{\boldsymbol{\beta}}) = \hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}} = \boldsymbol{\varepsilon}'\mathbf{M}'\mathbf{M}\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}'\mathbf{M}\boldsymbol{\varepsilon}$

we may get an estimator of σ^2 da $S(\hat{\boldsymbol{\beta}})$; for the derivation of the estimator, it is necessary to introduce the concept of trace; the *trace* of a square matrix is the sum of the elements on the main diagonal;

the trace enjoys some relevant properties, namely:

i) for every dyad of square matrices $\mathbf{A}$ and $\mathbf{B}$, $tr(\mathbf{A}+\mathbf{B}) = tr(\mathbf{A}) + tr(\mathbf{B})$;

ii) for every couple of matrices $\mathbf{A}$ and $\mathbf{B}$, $tr\mathbf{AB} = tr(\mathbf{BA})$, if both the products are defined (the equality applied also to rectangular matrics, in such case the matrices $\mathbf{AB}$ and $\mathbf{BA}$ are not of the same order[2]);

iii) the rank of an idempotent matrix is equal to its trace.

---

[1] a matrix is positive semi-definite if and only if the principal minors are all of the same sign (greater than or equal to zero), or, analogously, if and only if all the eigenvalues are greater than or equal to zero.

[2] the order of a matrix is its dimension: number of rows times number of columns.

using the property *ii)* as well as the fact that a scalar is a matrix of dimension 1×1 and it coincides with its tracewe have that:

$\boldsymbol{\varepsilon}'\mathbf{M}\boldsymbol{\varepsilon} = tr\boldsymbol{\varepsilon}'\mathbf{M}\boldsymbol{\varepsilon} = tr\mathbf{M}\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'$

we have seen that the expected value of a matrix is the matrix of the expected values, thus the expected value of a trace is the trace of expected values:

$\mathbb{E}(S(\hat{\boldsymbol{\beta}})|\mathbf{X}) = \mathbb{E}(tr\mathbf{M}\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'|\mathbf{X}) = tr\mathbb{E}(\mathbf{M}\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'|\mathbf{X}) = tr\mathbf{M}\mathbb{E}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'|\mathbf{X}) = \sigma^2 tr\mathbf{M} = T - k$

which proves that the rank of $\mathbf{M}$ is $T - k$ and that an unbiased estimator of the variance σ^2 is given by the following expression:

$s^2 = \frac{\hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}}}{T-k}$

such a result resolves the first issues[3], showing how the OLS estimation residuals can be used to construct an unbiased estimator of the variance of the residuals themselves; once obtained an expression for the variance of the residuals, it is possible to reconstruct the (estimate of) the variance-covariance matrix of the estimated parameters with the least squares method.

the analysis of the residuals allows also to shed light upon some aspect connected to the second issue[4];

given the orthogonality between the OLS forecasts and the residuals, we may write:

$var(\mathbf{y}) = var(\hat{\mathbf{y}}) + var(\hat{\boldsymbol{\epsilon}})$

from which we can construct the following measure of goodness of fit of the regression line to the data, which is defined in terms of the relationship between the variance of $\mathbf{y}$ and that of the estimated $\mathbf{y}$ values:

$R^2 = \frac{var(\hat{\mathbf{y}})}{var(\mathbf{y})} = 1 - \frac{var(\hat{\boldsymbol{\varepsilon}})}{var(\mathbf{y})}$, with $0 < R^2 < 1$, where $R^2 = 1$ if there is a perfect fit of the regression line or the lines to the data, $R^2 = 0$ if there is no fit at all of the regression line(s) to the data.

in our empirical application to the turning and milling firm in the neighbourhood of Bologna, the R^2 are all quite close to 1, suggesting a good fit of the model to the data, despite the residuals are very high for the estimated model with the variables in levels, e.g. $Y = AK^{\alpha}L^{\beta}$.

to the information coming from the R^2 is associated the one contained in σ^2, known as standard error of the regression, which represents the square root of the variance estimate of the error term defined above;

we note that in a model specified in logarithms, the standard error of the regession is a measure interpretable independently on the units of measure in which the variables are expressed and the standard error in the regression can be interpreted as the standard deviation of the forecast error.

[3] we have derived an expression for the OLS estimator which is a function of the sole observations on the vectorial variable $\mathbf{Y}$ and on the matrix $\mathbf{X}$; we have also derived an expression for the variance of the OLS estimator that is function of both the observables and of the error term.

[4] we are evaluating the empirical results, comparing the estimated parameters with the model forecasts; it could be the case to consider the residuals of the estimated model; in fact, what is omitted from the estimated model contributes to form the residual and analyzing the residuals seems like a very natural way of evaluating the goodness of fit of the chosen econometric specification.

1.3 elements of theory of the distributions

let us consider the distribution of an n-dimensional $\mathbf{x}$ vector together with the distribution derived from vector $\mathbf{y} = \mathbf{g}(\mathbf{x})$, a vector of invertible, continuous functions, with inverse $\mathbf{x} = \mathbf{h}(\mathbf{y})$; $\mathbf{h} = \mathbf{g}^{-1}$.

$\Pr(x_1 < x < x2) = \int_{x_2}^{x_1} f(x)dx$ and $\Pr(y_1 < y < y_2) = \int_{y_2}^{y_1} f^*(y)dy$, therefore

$f^*(\mathbf{y}) = f(\mathbf{h}(\mathbf{y}))\mathbf{J}$

where the jacobian matrix is

$$\underset{n\times n}{\mathbf{J}} = \begin{vmatrix} \frac{\partial h_1}{\partial y_1} & \cdots & \frac{\partial h_n}{\partial y_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial h_1}{\partial y_n} & \cdots & \frac{\partial h_n}{\partial y_N} \end{vmatrix} = \left|\frac{\partial \mathbf{h}}{\partial \mathbf{y}'}\right|$$

1.3.1 the normal distribution

the standard normal univariate distribution has the following probability density function:

$$f(z) = \frac{1}{\sqrt{2\pi}} \exp\Big\{ -\frac{1}{2}z^2\Big\}$$

$\mathbb{E}(z) = 0$, $var(z) = 1$

considering the transformation $x = \boldsymbol{\sigma} z + \boldsymbol{\mu}$, let us derive the univariate normal distribution:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\Big\{ -\frac{(x-\mu)^2}{2\sigma^2}\Big\}$$

$\mathbb{E}(x) = \mu$, $var(x) = \sigma^2$

let us consider the vector $\mathbf{z} = (z_1, \ldots, z_n)$, such that

$$f(\mathbf{z}) = \prod_{i=1}^{n} f(z_i) = (2\pi)^{-n/2} \exp\Big\{ -\frac{1}{2}\mathbf{z}'\mathbf{z}\Big\}$$

$\mathbf{z}$, by construction, is a vector of mutually independent normal variables, with zero mean and variance-covariance matrix equal to the identity matrix; the conventional notation is $\mathbf{z} \sim N(0, \mathbf{I}_n)$.

let's now consider a linear transformation of the following form

$$\mathbf{x} = \mathbf{A}\mathbf{z} + \boldsymbol{\mu}$$

where $\mathbf{A}$ is an $n \times n$ non-singular matrix; the proposed transformation is an invertible continuous function with inverse $\mathbf{z} = \mathbf{A}^{-1}(\mathbf{x} - \boldsymbol{\mu})$, and jacobian $\mathbf{J} = |\mathbf{A}^{-1}| = 1/|\mathbf{A}|$; applying the formula for the change of variable, we get:

$$f(x) = (2\pi)^{-n/2}|\mathbf{A}|\exp\Big\{ -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})(\mathbf{A}^{-1})'(\mathbf{A}^{-1})(\mathbf{x}-\boldsymbol{\mu})\Big\}$$

if we define the positive definite matrix $\boldsymbol{\Sigma} = \mathbf{A}\mathbf{A}'$, we may rewrite the density as follows:

$$f(x) = (2\pi)^{-n/2}|\mathbf{\Sigma}^{-1/2} \exp\Big\{ -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\Big\}$$

by convention, we denote the multivariate normal as $\mathbf{x} \sim N(\boldsymbol{\mu}, \mathbf{\Sigma})$.

upon the multivariate normal distribution, it is useful to remind the following result:

THEOREM 1 $\rightarrow$ any linear function of normal variables is normally distributed.

$\underset{n\times 1}{\mathbf{x}} \sim N(\boldsymbol{\mu}, \mathbf{\Sigma})$, given a generic matrix $\underset{m\times n}{\mathbf{B}}$ and a vector $\underset{m\times 1}{d}$, if $\mathbf{y} = \mathbf{B}\mathbf{x}+\mathbf{d}$, then

$$\mathbf{y} \sim N(\mathbf{B}\boldsymbol{\mu} + \mathbf{d}, \mathbf{B}\mathbf{\Sigma}\mathbf{B}')$$

applying the above defined theorem, we can show that, partitioning a normally distributed vector $n \times 1$ in two vectors of dimensions $n_1 \times 1$ and $n_2 \times 1$, where $n_1 + n_2 = n$, in the following way:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \sim N\left[\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}\right]$$

,

the following results hold:

(i) $x_1 \sim N(\mu_1, \Sigma_{11})$ attainable by applying the theorem with $\mathbf{B} = \begin{pmatrix} \mathbf{I}_{n_1} & \mathbf{0} \end{pmatrix}$ and $\mathbf{d} = 0$;

(ii) $x_1|x_2 \sim N(\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2), \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})$, attainable by applying the theorem with $\mathbf{B} = \begin{pmatrix} I_{n_1} & -\Sigma_{12}\Sigma_{22}^{-1} \end{pmatrix}, \mathbf{d} = \Sigma_{12}\Sigma_{22}^{-1}x_2$.

property *(ii)* illustrates how the non correlation within the field of the multivariate normal implies independence; such results, non always generally valid, does not surprise us, given that the normal distribution is entirely described by its first two moments.

1.3.2 distributions derived from the normal

let's consider $\underset{n\times 1}{\mathbf{z}} \sim N(=, \mathbf{I})$; the distribution of $\boldsymbol{\omega} = \mathbf{z}'\mathbf{z}$ is defined chi - squared with n degrees of freedom; such a distribution is tabulated for various values of n; the first two moments are respectively equal to n and to $2n$.

let us consider two vectors $\underset{n_1\times 1}{\mathbf{z}_1}$ and $\underset{n_2\times 1}{\mathbf{z}_2}$ for which it holds that:

$$\begin{pmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \end{pmatrix} \sim \left[0, \begin{pmatrix} I_{n_1} & 0 \\ 0 & I_{n_2} \end{pmatrix}\right]$$

we notice that $\boldsymbol{\omega}_1 = \mathbf{z}_1'\mathbf{z}_1 \sim \chi^2_{n_1}$, $\boldsymbol{\omega}_2 = \mathbf{z}_2'\mathbf{z}_2 \sim \chi^2_{n2}$, $\mathbf{z}_1'\mathbf{z}_1 + \mathbf{z}_2'\mathbf{z}_2 \sim \chi^2_{n_1+n_2}$; $\boldsymbol{\omega}_1$ and $\boldsymbol{\omega}_2$ are independent if and only if the elements of $\mathbf{z}_1$ and $\mathbf{z}_2$ are independent; the distribution has the following properties:

THEOREM 2 $\rightarrow$ the sum of independent chi-squared is distributed as a chi-squared with a number of degrees of freedom equal to the sum of the degrees of freedom of the two distributions;

from our discussion on the multivariate normal, it follows:

THEOREM 3 $\rightarrow$ if $\underset{n\times 1}{\mathbf{x}} \sim N(\boldsymbol{\mu}, \boldsymbol{\sigma}^2) \longrightarrow (\mathbf{x}-\boldsymbol{\mu})'\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}) \sim \chi^2_n$;

a connected result is that:

THEOREM 4 $\rightarrow$ if $\underset{n\times 1}{\mathbf{z}} \sim N(0, \mathbf{I})$ and $\mathbf{M}$ is an $n \times n$ symmetric and idempotent matrix of rank r, then $\mathbf{z}'\mathbf{M}\mathbf{z} \sim \chi^2_r$.

another family of distributions tabulated from the normal is the F distribution; it is obtained as the ration between two chi-squared distributions, independent between themselves, each divided by the number of its degrees of freedom; for example, given

$\boldsymbol{\omega}_1 \sim \chi^2_{n_1}$, $\boldsymbol{\omega}_2 \sim \chi^2_{n_2}$, reciprocally independent, we have that:

$$\frac{\boldsymbol{\omega}_1/n_1}{\boldsymbol{\omega}_2/n_2} \sim F_{n_1,n_2}$$

a distribution very much linked to the F is the t of student (attributed to W. S. Gosset) with n degrees of freedom, defined as:

$$t_n = \sqrt{F_{1,n}}$$

the most important application of the F distribution to our aims relies upon the following result:

TEOREM 5 $\rightarrow$ due idempotent quadratic forms in the standard normal vector $\mathbf{z}$, $\mathbf{z}'\mathbf{M}\mathbf{z}$ and $\mathbf{z}'\mathbf{Q}\mathbf{z}$, are between themselves idempotent if $\mathbf{MQ} = 0$.

combining theorems 4 and 5, we obtain the fundamental result for the application of statistical inference to the linear model:

THEOREM 6 $\rightarrow$ if $\underset{n\times 1}{\mathbf{z}} \sim N(=, \mathbf{I})$, and $\mathbf{M}$ and $\mathbf{Q}$ are symmetric, idempotent matrices of rank respectively r and s and $\mathbf{MQ} = 0$, then we have:

$$\frac{\mathbf{z}'\mathbf{Q}\mathbf{z}}{\mathbf{z}'\mathbf{M}\mathbf{z}}\frac{r}{s} \sim F(s, r)$$

1.3.3 Inference in the linear regression model

having introduced the basic elements for the statistical analysis of the linear model, we return to our first model and introduce another hypothesis: the distribution of $\mathbf{y}$ conditioned with respect to $\mathbf{X}$ is an independent normal

$$\mathbf{y}|\mathbf{X} \sim N(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$$

or, equivalently,

[A.4]

$$\mathbf{u}|\mathbf{X} \sim N(0, \sigma^2\mathbf{I})$$

the first indication of hypothesis [A.4] regards the distribution of $\hat{\boldsymbol{\beta}}|\mathbf{X}$, which, being a linear function of $\mathbf{u}$, is normal as well:

$$\hat{\boldsymbol{\beta}}|\mathbf{X} \sim N(\boldsymbol{\beta}, \sigma^2(\mathbf{XX})^{-1})$$

it shall be noticed how the conditional distribution of $\mathbf{u}$ does not depend on $\mathbf{X}$, thus it coincides with the non conditional distribution, while the conditional distribution of $\hat{\boldsymbol{\beta}}$ depends on $\mathbf{X}$.

hypothesis [A.4] forms the basis for the construction of confidence intervals and for the running of hypotheses tests on $\boldsymbol{\beta}$; consider first the following expression:

$$\frac{(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})'\mathbf{X}'\mathbf{X}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})}{\boldsymbol{\sigma}^2}=\frac{\mathbf{u}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}}{\boldsymbol{\sigma}^2}=\frac{\mathbf{u}'\mathbf{Q}\mathbf{u}}{\boldsymbol{\sigma}^2}$$

now, $\frac{\mathbf{u}'\mathbf{Q}\mathbf{u}}{\boldsymbol{\sigma}^2}|\mathbf{X}\sim\chi^2(k)\ \blacksquare$

based on theorem 4, because $\mathbf{Q}$ is an idempotent matrix, fixed when conditioning upon $\mathbf{X}$, and that, given **[A.4]**,

$$\frac{1}{\boldsymbol{\sigma}}\mathbf{u}|\mathbf{X}\sim N(0,\mathbf{I})$$

the above result is not applicable to the general case of the known variance, nevertheless, applying the same arguments to derive $\blacksquare$, we have:

$$S\Big(\frac{\hat{\boldsymbol{\beta}}}{\boldsymbol{\sigma}^2}\Big)=\frac{\mathbf{u}'\mathbf{M}\mathbf{u}}{\boldsymbol{\sigma}^2}\sim\chi^2_{T-k}\blacktriangle$$

we know that the two quadratic forms are independent among themselves, because $\mathbf{MQ}=0$; furthermore, both the $\blacksquare$ and the $\blacktriangle$ are proportional to the known variance, which disappears if we take the ration between these two entities; we therefore reach the following result:

$$\frac{(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})'\mathbf{X}'\mathbf{X}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})}{ks^2}=\frac{(\mathbf{u}'\mathbf{Q}\mathbf{u})/k\sigma^2}{(\mathbf{u}'\mathbf{M}\mathbf{u})/(T-k)\sigma^2}\sim F_{k,T-k}\bullet$$

the result $\bullet$ can be used obtaining from the tabulated distribution F the critical value $F^*_\alpha(k,T-k)$ such that

$$prob\{(k,T-k)>F^*_{\alpha,k,T-k}\}=\alpha$$

$0<\alpha<1$.

for different values of α, we are able to evaluate exactly inequalities of the form

$$prob[(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})'\mathbf{X}'\mathbf{X}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})]\leq ks^2F^*_\alpha(k,T-k)]=1-\alpha$$

which define the confidence intervals (geometrically some confidence ellipsoids) for $\boldsymbol{\beta}$, centered on $\hat{\boldsymbol{\beta}}$.

the hypotheses testing is strictly linked to the estimation of confidence intervals, with the difference that a decision should be taken, based on the sample evidence, whether to refuse or not the validity of specific restrictions imposed on the basic model;

in such a context, the hypotheses [A.1] - [A.4] are identified as maintained hypotheses and the reduced form of the model is identified with the null hypothesis H_0; in the hypotheses testing approach proposed by neyman and pearson, we derive a statistic with known distribution under the null hypothesis; letting the decision depend upon the absolute value of the statistic, it is therefore possible to fix the probability of making errors of the first type (refusing H_0 when H_0 is instead "true") at the level α; for example, a test of level α of the null hypothesis $\boldsymbol{\beta} = \boldsymbol{\beta}_0$, based on the F statistic, is given when we do not refuse H_0 if $\boldsymbol{\beta}_0$ lies within the confidence interval with associated probability $1 - \alpha$;

in practice, this way of verifying hypotheses is not very useful because the hypotheses of interest for the economist are rarely so complete as to specify a number of restrictions equal to the number of estimated parameters;

in general, the case of interest for the economist is the test of r restrictions on the vector of coefficients, where $\boldsymbol{\beta}_0$.

1.3.4 an application: the significance tests

let's consider a partitioned model of $\boldsymbol{\beta}$ in $\begin{bmatrix} \boldsymbol{\beta}_1 & \boldsymbol{\beta}_2 \end{bmatrix}$:

$\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}$

consider the partition of the "normal equations" $\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{y}$ in

$\begin{bmatrix} \mathbf{X}_1' \\ \mathbf{X}_2' \end{bmatrix} \begin{bmatrix} \mathbf{X}_1 & \mathbf{X}_2 \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\beta}}_1 \\ \hat{\boldsymbol{\beta}}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1' \\ \mathbf{X}_2' \end{bmatrix} \mathbf{y}$

or

$\begin{bmatrix} \mathbf{X_1'X_1} & \mathbf{X_1'X_2} \\ \mathbf{X_2'X_1} & \mathbf{X_2'X_2} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\beta}}_1 \\ \hat{\boldsymbol{\beta}}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{X_1'y} \\ \mathbf{X_2'y} \end{bmatrix}$

or

$\mathbf{X_1'X_1}\hat{\boldsymbol{\beta}}_1 + \mathbf{X_1'X_2}\hat{\boldsymbol{\beta}}_2 = \mathbf{X_1'y}$

$\mathbf{X_2'X_1}\hat{\boldsymbol{\beta}}_1 + \mathbf{X_2'X_2}\hat{\boldsymbol{\beta}}_2 = \mathbf{X_2'y}$

such a system in two blocks of equations can be solved in two steps; first derive $\hat{\boldsymbol{\beta}}_2$ from the second equation:

$\hat{\boldsymbol{\beta}}_2 = (\mathbf{X_2'X_2})^{-1}(\mathbf{X_2'y} - \mathbf{X_2'X_1}\hat{\boldsymbol{\beta}}_1)$

then substitute this in the first of the two equations of the system:

$\mathbf{X_1'X_1}\hat{\boldsymbol{\beta}}_1 + \mathbf{X_1'X_2}(\mathbf{X_2'X_2})^{-1}(\mathbf{X_2'y} - \mathbf{X_2'X_1}\hat{\boldsymbol{\beta}}_\mathbf{1}) = \mathbf{X_1'y}$

from which

$\hat{\boldsymbol{\beta}}_1 = [\mathbf{X_1'X_1} - \mathbf{X_1'X_2}(\mathbf{X_2'X_2})^{-1}\mathbf{X_2'X_1}]^{-1}[\mathbf{X_1'y} - \mathbf{X_1'X_2}(\mathbf{X_2'X_2})^{-1}\mathbf{X_2'y}]$

$\hat{\boldsymbol{\beta}}_1 = [\mathbf{X_1'M_2X_1}]^{-1}[\mathbf{X_1'M_2y}]$, with $\mathbf{M}_2 = \mathbf{I} - \mathbf{X_2}(\mathbf{X_2'X_2})^{\mathbf{-1}}\mathbf{X_2'}$

notice that[5], thanks to the idempotence of $\mathbf{M}_2$, we can write

$\hat{\boldsymbol{\beta}}_1 = [\mathbf{X_1'M_2'M_2X_1}]^{-1}[\mathbf{X_1'M_2'M_2y}]$

thus $\hat{\beta}_1$ can be seen as the result of the regression of y on $\mathbf{M_2X_1}$, that is, the regression of y on the matrix of the residuals of the regression of $\mathbf{X}_1$ on $\mathbf{X}_2$; the result by which the coefficients of a multiple regression can be calculated in a two step procedure known as the frisch - waugh theorem.

before returning to the hypotheses testing, consider the residuals of the partitioned model:

$\hat{\boldsymbol{\varepsilon}} = \mathbf{y} - \mathbf{X}_1\hat{\boldsymbol{\beta}}_1 - \mathbf{X}_2\hat{\boldsymbol{\beta}}_2$

$\boldsymbol{\varepsilon} = \mathbf{y} - \mathbf{X}_1\hat{\boldsymbol{\beta}}_1 - \mathbf{X_2}(\mathbf{X_2'X_2})^{\mathbf{-1}}(\mathbf{X_2'y} - \mathbf{X_2'X_1}\hat{\boldsymbol{\beta}}_\mathbf{1})$

---

[5]an alternative way to reach this result is to adopt the well known formula of the inverse of a partitioned matrix: $\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{E} & \mathbf{-EBD^{-1}} \\ \mathbf{-D^{-1}CE} & \mathbf{D^{-1} + D^{-1}CEBD^{-1}} \end{pmatrix}$, where $\mathbf{E} = (\mathbf{A} - \mathbf{BD^{-1}C})^{\mathbf{-1}}$

$\hat{\varepsilon} = \mathbf{M_2 y} - \mathbf{M_2 X_1}\hat{\boldsymbol{\beta}}_{\mathbf{1}} = \mathbf{M_2 y} - \mathbf{M_2 X_1}(\mathbf{X}_1'\mathbf{M_2 X_1})^{-1}(\mathbf{X}_2'\mathbf{M_2 y})$
$\hat{\varepsilon} = (\mathbf{M_2} - \mathbf{M_2 X_1}(\mathbf{X}_1'\mathbf{M_2 X_1})^{-1}(\mathbf{X}_1'\mathbf{M_2}))\mathbf{y}$
from what already seen, we know that $M = (\mathbf{M_2} - \mathbf{M_2 X_1}(\mathbf{X}_1'\mathbf{M_2 X_1})^{-1}(\mathbf{X}_1'\mathbf{M_2}))$;
this result is very useful to derive the form of test statistic for the hypothesis testing of significance to which we will return;
the specific hypothesis states that $\mathbf{X}_1$ does not have any additional explanatory power for $\mathbf{y}$, once considering $\mathbf{X}_2$; put differently,
$H_0 : \mathbf{y} = \mathbf{X_2}\hat{\boldsymbol{\beta}}_{\mathbf{2}} + \boldsymbol{\varepsilon}$
$(\boldsymbol{\varepsilon}|\mathbf{X_1}, \mathbf{X_2}) \sim N(0, \sigma^2\mathbf{I})$
notice that the statement
$\mathbf{y} = \mathbf{X_2}\boldsymbol{\gamma}_{\mathbf{2}} + \boldsymbol{\varepsilon}$
$(\boldsymbol{\varepsilon}|\mathbf{X_2}) \sim N(0, \sigma^2\mathbf{I})$
is true in the realm of the hypotheses maintained in **[A.1] - [A.4]**, even though $\boldsymbol{\gamma}_{\mathbf{2}} \neq \boldsymbol{\beta}_{\mathbf{2}}$, unless the null hypothesis holds ; in such a case, the matrix $\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}$ of dimensions $r \times r$ is nothing but the north-western submatrix of $(\mathbf{X}'\mathbf{X})^{-1}$ that, using the formula of the partitioned inverse introduced in footnote 5, we know being equal to $(\mathbf{X}_1'\mathbf{M_2 X_1})^{-1}$; thus the test statistic takes the form:
$\frac{\hat{\boldsymbol{\beta}}_1'(\mathbf{X}_1'\hat{\mathbf{M}}_2'\mathbf{X_1})\hat{\boldsymbol{\beta}}_{\mathbf{1}}}{rs^2} = \frac{\mathbf{y M_2 X_1}(\mathbf{X}_1'\mathbf{M_2 X_1})^{-1}\mathbf{X}_1'\mathbf{M_2 y}}{\mathbf{y}'\mathbf{M y}}\frac{(T-k)}{r} \sim F(T-k, r)$
given that $M = (\mathbf{M_2} - \mathbf{M_2 X_1}(\mathbf{X}_1'\mathbf{M_2 X_1})^{-1}(\mathbf{X}_1'\mathbf{M_2}))$, we can rewrite the statistic as
$\frac{\mathbf{y}'\mathbf{My} - \mathbf{y}'\mathbf{My}}{\mathbf{y}'\mathbf{My}}\frac{(T-k)}{r}$
where r is the number of restrictions on the coefficients deriving from economic theory, and the denominator is made up of the sum of the squared residuals of the regression without the imposition of the null hypothesis, while the numerator is made of the difference between the sum of the squared residuals of the regression under the null hypothesis and the sum of the squared residuals of the regression under the alternative;
our derivation establishes that the numerator is always positive;
let us consider the specific case in which $r = 1$ and $\boldsymbol{\beta}_{\mathbf{1}}$ is a scalar; for the reordering of the variables, the parameter considered can represent a whatever element of the vector $\boldsymbol{\beta}$;
the model is rewritten as

$$\underset{(T\times1)}{\mathbf{y}} = \underset{(T\times1)}{\mathbf{X}_1}\underset{(1\times1)}{\boldsymbol{\beta}_1} + \underset{(T\times1)}{\mathbf{X}_2}\underset{(1\times1)}{\boldsymbol{\beta}_2} + \underset{(T\times1)}{\boldsymbol{\varepsilon}}$$

in this case, the formula for the F statistic takes the following peculiar form:
$\frac{\hat{\boldsymbol{\beta}}_1^2}{s^2(\mathbf{x}_1'\mathbf{M_2 x_1})^{-1}} \sim F(1, T-k)$ under $H_0 : \beta_1 = 0$, where $(\mathbf{x}_1'\mathbf{M_2 x_1})^{-1}$ is the (1,1) element of the matrix $(\mathbf{X}'\mathbf{X})^{-1}$.
let us recall that the $F(1, T-k)$ distribution coincides with the t^2_{T-k}, where t_{T-k} is the student t distribution with $T-k$ degrees of freedom; rewriting the statistic under the null hypothesis as

$$\frac{\hat{\boldsymbol{\beta}}_1}{s(\mathbf{x_1'M_2x_1})^{-1/2}} \sim t_{T-k}$$

we have that the t - statistic in the form of the ratio between the coefficient and the associated standard error, exactly coinciding with the results reported in the third column of the regression; the associated level of probability to these values allows to largely refuse the hypothesis of equality to zero of the estimated coefficients at the conventional significance level of 5%;

we have thus developed the tools to establish the significance of the estimated coefficient in the economic model of interest (relationship between revenues and fixed/variable costs);

it also arises the problem of the correct specification of the econometric model to be estimated and the characteristics of the data generating process, from us not completely observed.

2 Effects of fixed and variable costs on revenues

after having reviewed the properties of the OLS estimator, we apply such a simple estimation method and apply it to the model with fixed and variable costs and revenues, defined as follows:

$\mathbf{Y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}$

where $\mathbf{X}_1$ represents variable costs; $\mathbf{X}_2$ stands for fixed costs; $\mathbf{Y}$ is revenues for years $t = 2011, \ldots, 2021$; notice that the balance sheet from which the data have been retrieved are the analytical financial statements of end of the solar year, for all but 2021, for which we have data up to the end of the third quarter (september 30); therefore, up to now, the length of the time series is $T = 11$ years with yearly observation and $k = 2$, the number of parameters to be estimated;

in particular, we considered as components of **variable costs** $\longrightarrow$ the following voices of the sequence of income statements of the firm:

`purchases` $\longrightarrow$ *purchase of raw materials, namely iron, aluminium, brass, inox; purchase of finite goods from both italy and abroad*; `production costs` $\longrightarrow$ *external processing, industrial lubricants, equipment and small parts, cleaning and garbage collection, compressor maintenance, petrol/diesel trucks, consumables, treatments, car fuel, truck fuel, truck insurance*; `sales costs` $\longrightarrow$ *transport for sales, travel and transfers, packaging for sales, commercial expenses, passive commissions*; `general expenses` $\longrightarrow$ *postal and telegraph expenses, telephone expenses, revenue stamps, bank expenses, administrative services, mobile phone expenses, various rentals*;

as part of the **fixed costs**, we included the following components of the income statements: `cost of productive labour` $\longrightarrow$ *gross workers' salaries, INPS and INAIL social security contributions for workers, severance indemnity*; `production costs` $\longrightarrow$ *electricity, maintenance and repairs, heating, water consumption, insurance, car insurance, computer rental fees, truck insurance, sylos system maintenance and repairs, truck maintenance and* *repairs, forklift truck repairs, heating system maintenance, electrical system maintenance, washing machine maintenance and repair*; `general expenses` $\longrightarrow$ *stationery and printed matter, legal and notary consultancy*[6], *administrative consultancy, directors' fees, computer programming assistance services, contribution of 10% for self-employed workers, compliance with law 626*[7], *ISO 9002 compliance, board of statutory auditors compensation*; `cost of administrative labour` $\longrightarrow$ *administrative salaries, INPS and INAIL social security contributions for employees, severance pay for employees*;

finally, as part of the **revenues**, we choose the following: `miscellaneous revenues and income` $\longrightarrow$ *sales of production in Italy, exports, sale of scrap and various scraps, recovery of expenses and other indemnities, bank interest income, interest income (coupons), contingent assets, capital gains.*

we report in analytical and graphical form the data employed in the analysis, arising from the illustrated aggregations.

Table 1 data on revenues (Y), variable costs (X_1), fixed costs (X_2), capital (K), the cost of productive labour (L_{prod}), and of administrative labour (L_{adm})

t	Y	X_1	X_2	K	L_{prod}	L_{adm}	L
2011	863,768	2,299,513	1,240,489	211,217	733,086.7	294,931.7	998,017.4
2012	2,815,571	1,361,284	1,107,425	1,161,120	543,928.8	271,331.5	815,260.3
2013	2,180,815	1,100,591	1,059,961	1,151,168	519,649	266,106.8	785,755.9
2014	2,239,722	1,095,824	1,035,184	1,088,539	507,263	296,094.9	776,357.9
2015	2,265,955	1,238,506	1,043,427	1,129,489	496,639.9	281,116.4	764,756.3
2016	2,271,333	967,703	1,064,599	1,041,229	447,950	260,831.2	708,781.2
2017	2,383,581	1,195,590	1,071,085	1,091,207	551,940.8	235,179.5	787,120.3
2018	768,506.3	302,473.2	290,349.2	1,488,349	593,765.3	259,711.7	853,476.8
2019	2,159,723	984,404	1,079,349	1,103,705	577,075.3	261,118.2	838,193.5
2020	1,814,393	740,421.8	1,025,192	2,014,860	853,476.8	241,305.5	724,987.1
2021	1,846,034	861,581.5	845,405.9	1,631,942	378,328.1	202,075	580,403.1

Table 2 descriptive statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
$\mathbf{y}$	11	2,429,347	579,151.1	1,814,393	3,863,768
$\mathbf{X}_1$	11	1,182,414	423,947.8	705,503.7	2,265,687
$\mathbf{X}_2$	11	1,110,932	116,349.9	845,405.9	1,288,783

we hereby report the various specifications of the models which we estimate with STATA 13.0 SE, that depart from the basic version of the bivariate equation with the variables in contemporaneous time, allowing for some lagged independent variables to appear on the right side of the equation:

$$\mathbf{y}_t = \mathbf{X}_{1t}\boldsymbol{\beta}_1 + \mathbf{X}_{2t}\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}_t \tag{1}$$

[6] due to their occasional occurrence.
[7] safety on the job for the workers.

$$\mathbf{y}_t = \mathbf{X}_{1t-1}\boldsymbol{\beta}_1 + \mathbf{X}_{2t-1}\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}_t \tag{2}$$

$$\mathbf{y}_t = \mathbf{X}_{1t-1}\boldsymbol{\beta}_{11} + \mathbf{X}_{1t}\boldsymbol{\beta}_{12} + \mathbf{X}_{2t-1}\boldsymbol{\beta}_{21} + \mathbf{X}_{2t}\boldsymbol{\beta}_{22} + \boldsymbol{\varepsilon}_t \tag{3}$$

$$\mathbf{y}_t = X_{1t-2}\boldsymbol{\beta}_{11} + \mathbf{X}_{1t-1}\boldsymbol{\beta}_{12} + \mathbf{X}_{2t-2}\beta_{21} + \mathbf{X}_{2t-1}\boldsymbol{\beta}_{22} + \boldsymbol{\varepsilon}_t \tag{4}$$

$$\mathbf{y} = \mathbf{X}_{1t-2}\boldsymbol{\beta}_{11} + \mathbf{X}_{1t-1}\boldsymbol{\beta}_{12} + \mathbf{X}_{1t}\boldsymbol{\beta}_{13} + \mathbf{X}_{2t-2}\boldsymbol{\beta}_{21} + \mathbf{X}_{2t-1}\boldsymbol{\beta}_{22} + \mathbf{X}_{2t}\boldsymbol{\beta}_{23} + \boldsymbol{\varepsilon}_t \tag{5}$$

the idea is that, especially fixed costs, may take time to produce effects on the revenues, since they are connected with fixed assets, which have a relatively long economic life; therefore, we consider equations with one to two lags, combining various layers of complication; below are reported the results of the regressions, where the number of the columns corresponds to the number of the equations:

Table 3 exploratory regressions of revenues on variable and fixed costs, with and without time lags in the independent variables

	dependent	**variable**	$\longrightarrow$	**revenues**	**of the year**
	{1}	{2}	{3}	{4}	{5}
$\mathbf{X}_{1t}$	1.155***		0.999***		0.805*
	(0.0987)		(0.223)		(0.272)
$\mathbf{X}_{2t}$	0.956***		0.852		1.341
	(0.110)		(0.508)		(0.682)
$\mathbf{X}_{1t-1}$		0.264	0.165	0.748	-0.0457
		(0.308)	(0.104)	(1.251)	(0.369)
$\mathbf{X}_{2t-1}$		1.721***	0.0705	0.268	0.135
		(0.346)	(0.407)	(3.697)	(0.957)
$\mathbf{X}_{1t-2}$				-0.504	-0.0650
				(0.706)	(0.199)
$\mathbf{X}_{2t-2}$				1.517	-0.0849
				(3.267)	(0.935)
observations	11	10	10	9	9
R-squared	0.998	0.983	0.999	0.982	0.999

standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

equation (1) is the most relevant one, as far as it allows to best exploit the information at hand, regressing the current values of revenues on those of two categories of costs within the same year (i.e. contemporaneous effect of variable and fixed costs on revenues); while in equation (2) we consider a case where the regression is of revenues each year on the fixed and variable costs of the year before; in equation (3) we see the effect of variable costs of a year on the revenues of the same year and of the subsequent year as well; equation

(4) we consider the independent variables lagged of two and of one year; and, finally, in equation (5) we insert a lag of two years, one year, and no lag;

despite the lagged structure of fixed costs on output, the most significant estimated coefficients seem to be those of the first regression, which points to the direction that marginal increases in variable costs lead to higher revenues, while, paradoxically, increases in fixed costs tend to reduce revenues; in particular, a marginal increase in variable costs of 1% should lead to an increase in revenues of about 1.55%; while an increase in fixed costs seems to be negatively related with revenues, being the associated coefficient point estimate slightly lower than 1; both the estimates are highly statistically significant;

the trend seems to invert when we move on to consider lagged variable and fixed costs on revenues: as a matter of facts, lagged fixed costs have a positive and statistically significant effect on revenues, raising them, while keeping variable costs constant, of 1.72% after a 1% increase in their entity; on the other hand, a marginal increase in variable costs keeping the fixed constant, does not seem to have a statistically significant effect on revenues of the subsequent year;

in the third model, with the regressors both contemporaneous to the dependent variable and lagged of one period, it appears that the variable costs reduce revenues and that the fixed cost raise revenues, instead; here, only the coefficient estimate associated with contemporaneous variable costs seems to be statistically significant, though, raising doubts on the validity of such a specification;

similar considerations hold for the fourth and fifth model as well; the firth regression presents similar results to the third one; we feel like the bivariate regression most likely to capture the relationships between fixed and variable costs, and revenues is the first one, with the condition that assumptions [A.1], [A.2], [A.3] for the OLS estimator seen in the previous sections hold; only if such a case happens to be holding, these econometric estimates could have a causal interpretation.

if that happened to be the case, it would be worthy for the firm in consideration to raise variable costs with the reasonable expectation of producing a higher share of output and thus having higher revenues, especially on the voices of cost most related with raw materials purchases.

the stochastic frontier model with the same contemporaneous variables as before, as of the STATA syntax, happens to produce similar results than the ordinary least squares estimation, especially in the case without a constant term in the regression; a marginal increase in one unit of variable costs, keeping fixed costs constant, seems to cause an increase in revenues of 1.55 units; while an increase in fixed costs of one unit, keeping the variables constant, seems to reduce revenues by about 0.05 units; considering an intercept, we obtain a slightly higher coefficient associated with X_1 and a slightly lower coefficient associated with X_2, slightly polarizing the effects of the no-intercept case.

Table 4 stochastic frontier model with and without an intercept

variables	**Y** [1a]	$\ln \sigma^2 v$ [1b]	$\ln \sigma^2 u$ [1c]	**Y** [2a]	$\ln \sigma^2 v$ [2b]	$\ln \sigma^2 u$ [2c]
$\mathbf{X}_{1t}$	1.155***			1.185***		
	(0.0893)			(0.107)		
$\mathbf{X}_{2t}$	0.956***			0.763*		
	(0.0997)			(0.408)		
constant		22.98***	-12.15	181,638	22.96***	-5.481
		(0.426)	(41,250)	(443,927)	(0.426)	(9.442e+06)
Observations	11	11	11	11	11	11

standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

3 Properties of production sets

PRODUCTION SETS[8] $\longrightarrow$ in an economy with L commodities, a production vector is a vector $\mathbf{y} = (y_1, \ldots, y_L) \in \mathbb{R}^L$, that describes the net output of the L commodities from a production process[9];

$Y \subset \mathbb{R}^L \longrightarrow$ production set; any $y \in Y$ is possible, any $y \notin Y$ isn't. a production set is a primitive datum of the theory; technological constraints $\rightarrow$ legal restrictions or prior contractual commitment $\rightarrow F(.)$: transformation function, $Y = \{y \in \mathbb{R}^L : F(y) \leq 0\}$ and $F(0) = 0$ if and only if y is an element of the boundary of Y;

$\{y \in \mathbb{R}^L : F(y) = 0\} \rightarrow$ boundary $\equiv$ transformation frontier;

$MRT_{lk}(\bar{y}) = \frac{\frac{\partial F(\bar{y})}{\partial y_l}}{\frac{\partial F(\bar{y})}{\partial y_k}}$, $\forall l, k$, $l \neq k$ goods, marginal rate of transformation of good l for good k at $\bar{y}$;

$\frac{\partial F(\bar{y})}{\partial y_k} \cdot dy_k + \frac{\partial F(\bar{y})}{\partial y_l} \cdot dy_l = 0$

$\mathbf{z} = (z_1, \ldots, z_{L-M}) \geq 0 \rightarrow$ firm's $L - M$ inputs;

$\mathbf{q} = (q_1, \ldots, q_M) \geq 0 \rightarrow$ outputs;

single output technology $\rightarrow f(z) \rightarrow$ max. amount of q that can be produced using input amounts $\mathbf{z} = (z_1, \ldots, z_{L-1}) \geq 0$; if the output is good $L \rightarrow Y = \{(z_1, \ldots, z_{L-1}, q) : q - f(z_1, \ldots, z_{L-1}) \leq 0$ and $(z_1, \ldots, z_{L-1}) \geq 0\}$;

$MRTS_{lk}(\bar{z}) = \frac{\frac{\partial f(\bar{z})}{\partial z_l}}{\frac{\partial f(\bar{z})}{\partial z_k}} \rightarrow$ marginal rate of technical substitution;

the cobb - douglas production function can be expressed as $f(z_1, z_2) = z_1^{\alpha} z_2^{\beta}$, and, if $\alpha + \beta = 1 \rightarrow f(z_1, z_2) = z_1^{\alpha} z_2^{1-\alpha}$

(i) Y **is non-empty**[10]$\rightarrow$ the firm has something to plan to do;

(ii) Y **is closed** $\longrightarrow$ the set Y includes its boundary, the limit of a sequence of feasible input - output vectors is also feasible; $y^n \rightarrow y$ and $y^n \in Y \rightarrow y \in Y$;

[8] see also [1] and [3].
[9] this section is largely based upon chapter 5 of [4], pp. 128 - 136.
[10] the production set.

(iii) **no free lunch** $\longrightarrow y \in Y$ and $y \geq 0$ so that y doesn't use any inputs; this property $\longrightarrow$ this production vector cannot produce output either; $Y \cap \mathbb{R}_+^L \subset \{0\}$

it's not possible to produce something out of nothing;

(iv) **possibility of inaction** $\longrightarrow 0 \in Y$; the point in time at which production possibilities are being analyzed is often important for the validity of this assumption; if we see a firm that could access a set of technological possibilities but hasn't yet been organized $\longrightarrow$ inaction is clearly possible; but otherwise (decisions already taken or irrevocable contracts signed), inaction isn't possible $\rightarrow$ sunk costs;

the firm is already committed to use at least $-\bar{y}_1$ units of good 1;

$\searrow$ the set is a restricted production set, reflecting the firm's remaining choices from some original production set Y like the ones in the previous graphs;

v. free disposal $\rightarrow$ holds if the absorption of any additional amount of inputs without any $\searrow$ in output is always possible, if $y \subset Y$ and $y' \leq y$ (so that y' produces at most the same amount of outputs using at least the same amount of inputs) $\rightarrow y' \in Y$; $Y - \mathbb{R}_+^L \subset Y \leftrightarrow$ the extra amounts of inputs (or outputs) can be disposed of or eliminated at no cost;

vi. irreversibility $\rightarrow y \in Y$ and $y \neq 0$;

$-y \notin Y$; it's impossible to reverse a technologically possible production vector to transform an amount of output into the same amount of input that was used to generate it;

drawing 5 — drawing 6 — drawing 7

vii. non $\nearrow$ returns to scale $\rightarrow$ the production technology Y exhibits non $\nearrow$ returns to scale if for any $y \in Y$, we've $\alpha y \in Y$ for all scalars $\alpha \in [0, 1]$;

any feasible input - output vector can be scaled down;

viii. non $\searrow$ returns to scale $\rightarrow$ if $\forall y \in Y \rightarrow \alpha y \in Y$ for any scale $\alpha \geq 1$.

any feasible input - output vector can be scaled up;

ix. constant returns to scale $\rightarrow$ the production set Y exhibits constant returns to scale if $y \in Y \rightarrow \alpha y \in Y$, for any scalar $\alpha \geq 0$. Y is a cone;

for single output technologies $\rightarrow$ properties of the production set translate into properties of the production function, $f(.)$; Y satisfies constant returns to scale if and only if $f(.)$ is homogeneous of degree 1. $f(2z_1, 2z_2) = 2^{\alpha+\beta} z_1^{\alpha} z_2^{\beta} = 2^{\alpha+\beta} f(z_1, z_2)$;

when $\alpha + \beta < 1 \rightarrow \searrow$ returns to scale;

when $\alpha + \beta = 1 \rightarrow$ constant returns to scale;

when $\alpha + \beta > 1 \rightarrow \nearrow$ returns to scale; ■

x. additivity $\rightarrow$ or free entry $\rightarrow y \in Y$ and $y' \in Y \rightarrow y + y' \in Y \leftrightarrow Y + Y \subset Y \rightarrow$ e.g. $ky \in Y$, $\forall k \in \mathbf{N}_+$; output here is available only in integer amounts. perhaps because of indivisibilities, the economic interpretation is that both y and y' are possible $\rightarrow$ one can set up two plants that don't interfere with each other and carry out production plans y and y' independently. the result is the production vector $y + y'$;

additivity $\rightarrow$ entry: if a firm produces $y \in Y$ $\rightarrow$ net result $\rightarrow$ $y + y'$ $\rightarrow$ the aggregate production set $\rightarrow$ must satisfy additivity when ever unrestricted entry or free entry is possible;

xi. convexity $\rightarrow$ one of the fundamental assumptions of micro-economics $\rightarrow$ production set Y is convex $\rightarrow$ if $y, y' \in Y$ and $\alpha \in [0,1]$, $\alpha y + (1-\alpha)y' \in Y$ $\rightarrow$ non $\nearrow$ returns, if inaction is possible, i.e. if $0 \in Y$ $\rightarrow$ convexity $\rightarrow$ Y has non increasing returns to scale; hence if any $\alpha \in [0,1] \rightarrow \alpha y = \alpha y + 0(1-\alpha)$, if $y \in Y$ and $0 \in Y$ $\rightarrow$ $\alpha y \in Y$, by convexity;

"unbalanced" inputs combinations aren't more productive than balanced ones;

if production plans y and y' produce exactly the same amount of output but use $\neq$ input combinations $\rightarrow$ a production vector that uses a level of each input that's the average of the levels used in these two plans can do at least as well as either y or y'.

ex. 5.B.3: Y is convex if $f(z)$ is concave. suppose Y is convex; $z, z' \in \mathbf{R}_+^{L-1}$ and $\alpha \in [0,1] \rightarrow (-z, f(z)) \in Y$ and $(-z', f(z')) \in Y$. by convexity

$\{-[\alpha z + (1-\alpha)z], \alpha f(z) + (1-\alpha)f(z)\} \in Y$

by convexity $\alpha f(z) + (1-\alpha)f(z) \leq f[\alpha z + (1-\alpha)z] \rightarrow f(z)$ is concave

suppose $f(z)$ is concave.

$(q, -z) \in Y$, $(q', -z') \in Y$, $\alpha \in [0,1]$ $q \leq f(z)$ and $q' \leq f(z) \rightarrow \alpha q + (1-\alpha)q' \leq \underbrace{\alpha f(z) + (1-\alpha)f(z')}$

by concavity

$\rightarrow \underbrace{\alpha f(z) + (1-\alpha)f(z')} \leq f[\alpha z + (1-\alpha)z']$ $\alpha q + (1-\alpha)q' \leq f[\alpha z + (1-\alpha z')]$

$\rightarrow \{-[\alpha z + (1-\alpha)z'], \alpha q + (1-\alpha)q\} = \alpha(-z, q) + (1-\alpha)(-z', q') \in Y$

$\rightarrow Y$ is convex. ■

xii. Y is a convex cone $\rightarrow$ convexity $\cap$ CRS. if for any production vector $y, y' \in Y$ and constants $\alpha \geq 0$ and $\beta \geq 0 \rightarrow \alpha y + \beta y' \in Y$.

proposition 5.B.1 the production set Y is additive and satisfies the non $\nearrow$ returns condition iff it's a convex cone.

proof $\alpha y + \beta y' \in Y$; $k > \max\{\alpha, \beta\}$,
$ky \in Y$, $ky' \in Y$; $\frac{\alpha}{k} < 1$ and $\alpha y = \frac{\alpha}{k} ky \rightarrow \alpha y \in Y$, similarly for β.

feasible input - output combination can be scaled down, and simultaneous operation of several technologies w/out mutual interference is possibile $\rightarrow$ convexity! production set $\rightarrow$ technology. $\searrow$ returns reflect the scarcity of an underlying, unlisted input of production.

proposition 5:B.2: for any convex production set $Y \subset \mathbb{R}^L$ with $0 \in Y$, there is a constant return convex production set $Y' \subset \mathbb{R}^{L+1}$ such that $Y = \{y \in \mathbb{R}^L : (y, -1) \in Y\}$

additional input $\rightarrow$ entrepreneurial factor - whose return's precisely the firm's profit. $Y' = \{y' \in \mathbb{R}^{L+1} : y' = \alpha(y, -1)$ for some $y' \in Y$ and $\alpha \geq 0\}$.

4 test of the exponents of a cobb-douglas production function

we set up an empirical estimation of a cobb douglas production function of the form

$$Y = AK^{\alpha}L^{\beta} \tag{6}$$

where the variables Y, K and L represent, respectively, revenues, capital and labour cost; the first one, a flow variable, is as of the ones resulting from the income statements; the second, is here defined as the balance sheet value of net fixed assets; while the third one is intended as the sum of both productive and administrative labour cost, inclusive of the social security contributions; the data are composed of observations spanning the period t = 2011,. . . ,2021, as in the previous exercise;

A is the so called technological augmenting factor, namely a measure of the technological intensity of the productive process of the firm at hand;

our main point is connected with exercise 3.B.1 of the [4] handbook: we attempt to see whether the sum of the estimated exponents α and β is $\gtreqless 1$, in order to conclude whether the returns to scale are respectively increasing, constant or decreasing.

in the appendix, we display some additional three dimensional graphs made with MATLAB 2022a, which may give an indication of the geometric properties of the production set of the firm which we are analyzing;

we report the results of a regression which estimates equation (6):

Table 5 cobb douglas in levels

	$\hat{A}$	$\hat{\alpha}$	$\hat{\beta}$
estimates	0.000176	0.0161	1.703***
	(0.00104)	(0.227)	(0.323)
observations	11	11	11
R-squared	0.988	0.988	0.988

standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

here, I obtain two estimated exponents, $\hat{\alpha}$ and $\hat{\beta}$, whose sum is higher than 1, finding support for increasing returns to scale, as of point ix. of the properties of production sets in mas colell et al. 1995; however, the only significant coefficient estimate of the three $\hat{A}$, $\hat{\alpha}$ and $\hat{\beta}$ is $\hat{\beta}$, and equal to about 1.7, significantly higher than one, seemingly indicating that with the basic cobb - douglas production function estimated in levels, there could be increasing returns to scale almost entirely driven by the labour share of the input of production;

we log - transform the same function of above, in order to make it linear, considering the three variables Y, K and L expressed in natural logarithms of their levels,

$$\ln Y = \ln A + \alpha \ln K + \beta \ln L \tag{7}$$

Table 6 cobb douglas in logarithms

	$\ln(\hat{A})$	$\hat{\alpha}$	$\hat{\beta}$
estimates	-2.332	-0.0471	1.303***
	(6.433)	(0.219)	(0.344)
observations	11	11	11
R-quadrato	0.682	0.682	0.682

standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

here, again the only significant coefficient is the one associated with labour, which is also slightly higher than 1, about 1.3, while the coefficient associated with capital is slightly negative but statistically not significantly different that zero; similarly for the technological augmenting factor, which is in magnitude considerably negative; these results would point again in favour of the increasing returns to scale hypothesis;

however, the firm in question is relying heavily on the supply of raw materials, given that it transforms them into bolts and small mechanical components for automobiles engines; this leads us to consider a version of the cobb douglas production function which also includes a term for the purchase of raw materials, variable extrapolated as well from the analytical income statements of the firm[11], such that:

$$Y = AK^{\alpha} L^{\beta} M^{\gamma} \tag{8}$$

finding the following parameter estimates, with the values of the variables in levels:

Table 7 cobb douglas in levels with raw materials

	$\hat{A}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$
estimates	27.19	-0.0405	0.298	0.583***
	(120.5)	(0.141)	(0.373)	(0.143)
observations	11	11	11	11
R-squared	0.997	0.997	0.997	0.997

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

here, surprisingly, the entity of the technological augmenting factor is much higher than in the baseline version of the function, though not statistically different than zero, perhaps for a lack of more data; in addition to that, $\hat{\alpha} =$

[11] borowski, borwein 1989.

Table 8 cobb douglas in logarithms with raw materials

	$\widehat{\ln(A)}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$
estimates	4.841	-0.0962	0.261	0.564***
	(4.260)	(0.132)	(0.336)	(0.144)
observations	11	11	11	11
R-squared	0.901	0.901	0.901	0.901

standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

$-.0405$, $\hat{\beta} = .2982$, and $\hat{\gamma} = .5825$, but only this last is significant, reversing the previously accepted hypothesis; the only significant factor of production is purchases of raw materials, rather than cost of labour and capital; if we also take into account the non significant estimate, we get $\hat{\alpha} + \hat{\beta} + \hat{\gamma} = .8402 < 1$, which would indicate the existence of decreasing returns to scale; which appear to be even lower if we just consider the statistically significant estimate of $\hat{\gamma} = .5825$.

as a last experiment, we log transform equation (8), to get:

$$\ln Y = \ln A + \alpha \ln K + \beta \ln L + \gamma \ln M \tag{9}$$

with the following results from estimation

here, $\ln A = 4.841$, $\hat{\alpha} = -.0962$, $\hat{\beta} = .2613$, e $\hat{\gamma} = .5637$, meaning that the returns to scale should be negative again, very similarly than in the case of the equation estimated in level, in its non linear form; the main difference between equation (8) and (9) relies in the estimate for $\ln A$, difference attributable to the logarithmic transformation.

5 conclusions

we attempted an double empirical exercise with some data of a firm operating in the supply chain of the automobile industry, manily german, producting small metal parts out of metal pipes. we first reviewed the theory of estimation through the ordinary least squares method, building on a previous work by [2]; then specified and estimated a set of bivariate linear regressions of revenues on fixed and variable costs, both contemporaneous and lagged; we found evidence for positive impact of variable costs on revenues on impact; afterwards, we reviewed a part of the theory of production following [4], ch.5, stressing the relevance of the properties of production sets, with specific reference to the returns to scale with the cobb douglas production function; in particular, assuming the cobb douglas is a good fit in approximating the firm's production process, we tried to find out whether the returns to scale have been for the past ten years of operations, increasing, decreasing or constant, depending on the magnitude of the summed exponents of the production function, namely whether greater or smaller than one; we found a mixed evidence.

while basing the inference on the cobb douglas with only capital and labour as inputs leads us to accept the hypothesis of increasing returns to scale[12], adding as an additional input the purchase of raw materials, conducts us towards accepting the hypothesis of decreasing returns to scale, with this last variable driving most of the effect of production, both for the regression estimated in level and for the one in logarithms.

due to the relevance of the raw materials in the business operations of such a firm, we are more prone to pend towards the second set of results: that it has decreasing returns to scale.

more research is needed on this topic, perhaps on estimating the returns to scale of entire sectors of activity as well as trying to estimate some other production functions such as the constant elasticity of substitution one.

Supplementary information. The paper has an appendix with the STATA code and data set used to produce the estimation reported in the main text as well as the MATLAB code written to produce the figures contained in the paper.

Acknowledgments. I thank a collaborator in my studio who helped me making available the financial statement of the firm from which the data are taken.

Declarations

- no conflict of interests
- consent for publication of the data
- STATA data set and codes available for replications
- MATLAB code available for replications

Appendix A

An appendix contains supplementary information that is not an essential part of the text itself but which may be helpful in providing a more comprehensive understanding of the research problem or it is information that is too cumbersome to be included in the body of the paper.

[12] for both the function in levels and in natural logarithms.

Fig. A1 revenues, variable and fixed costs for $t = 2011, \ldots, 2021$

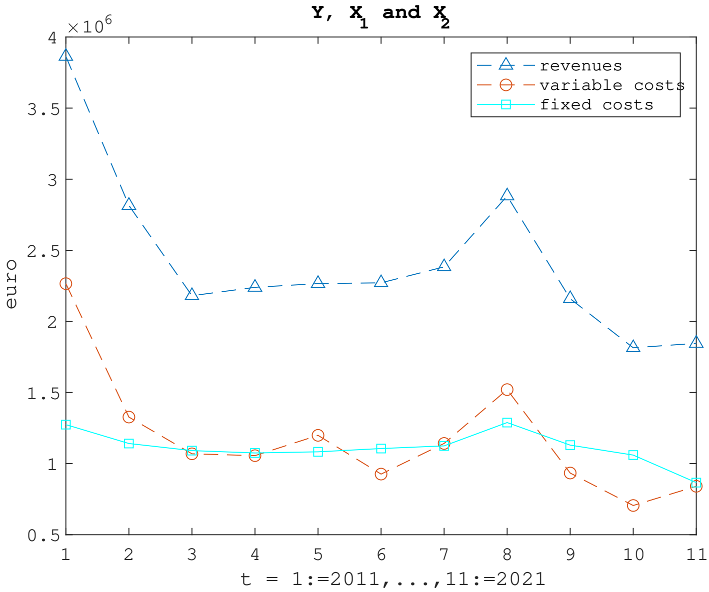


Fig. A2 variable and fixed costs for $t = 2011, \ldots, 2021$

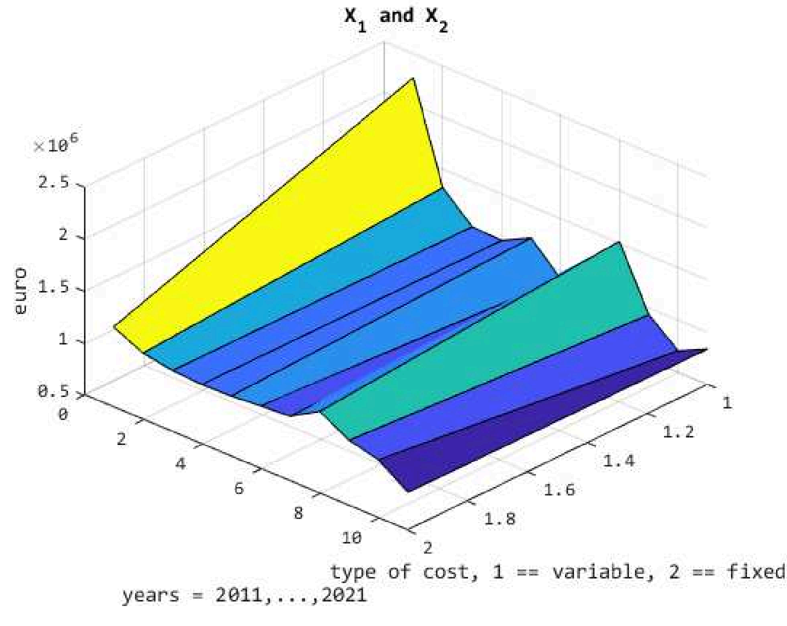


Fig. A3 revenues, variable and fixed costs for $t = 2011, \ldots, 2021$

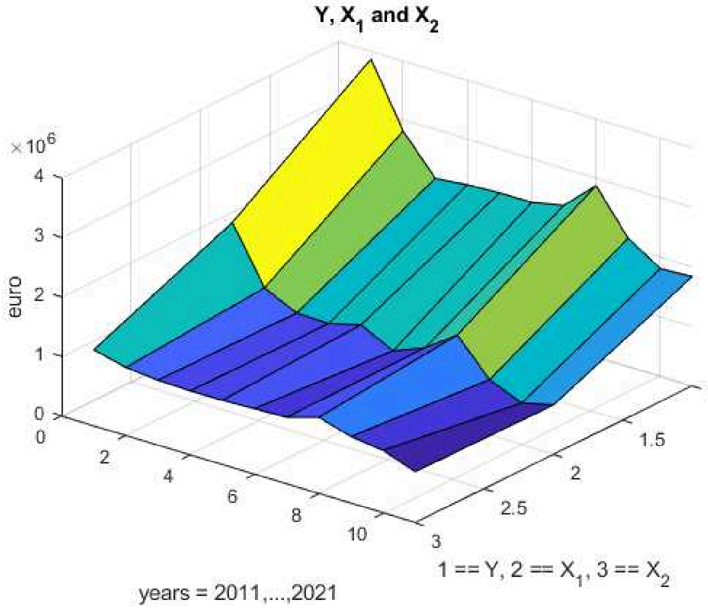


Fig. A4 revenues, capital, disaggregated labour cost

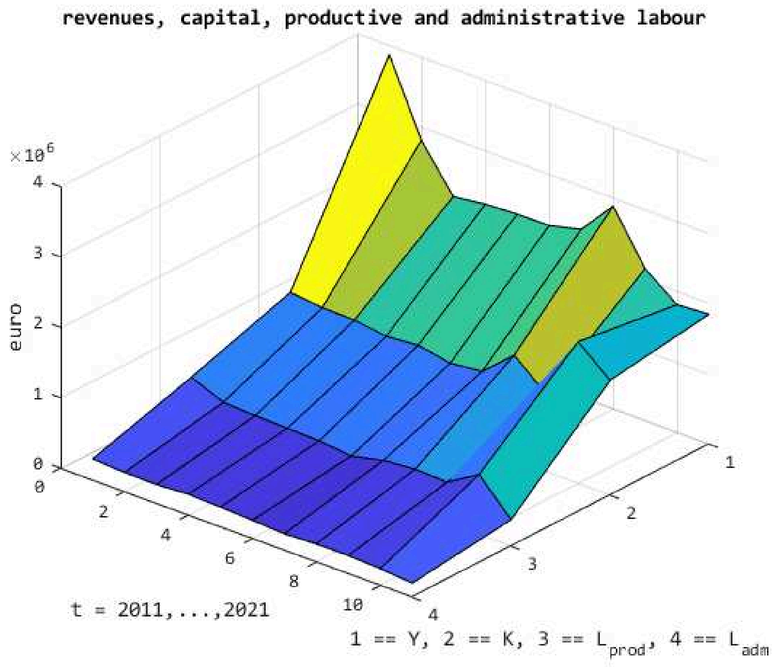


Fig. A5 revenues, capital, labour

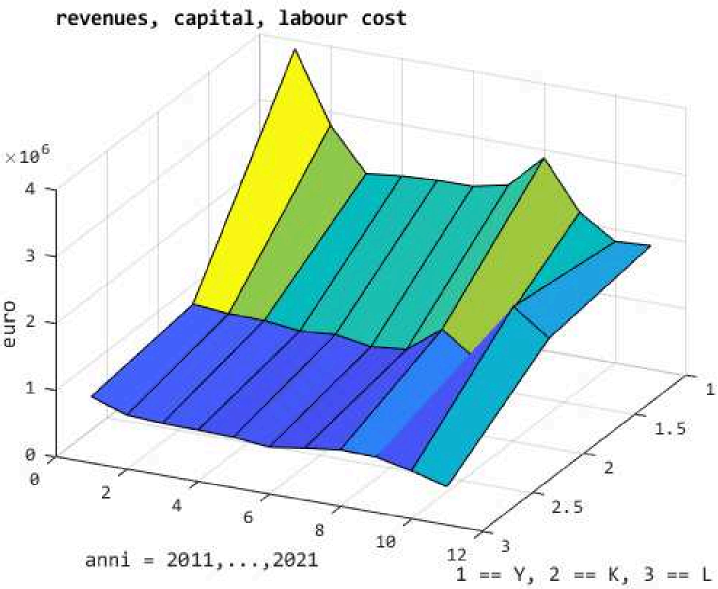


Fig. A6 revenues, capital, labour costs and purchases of raw materials

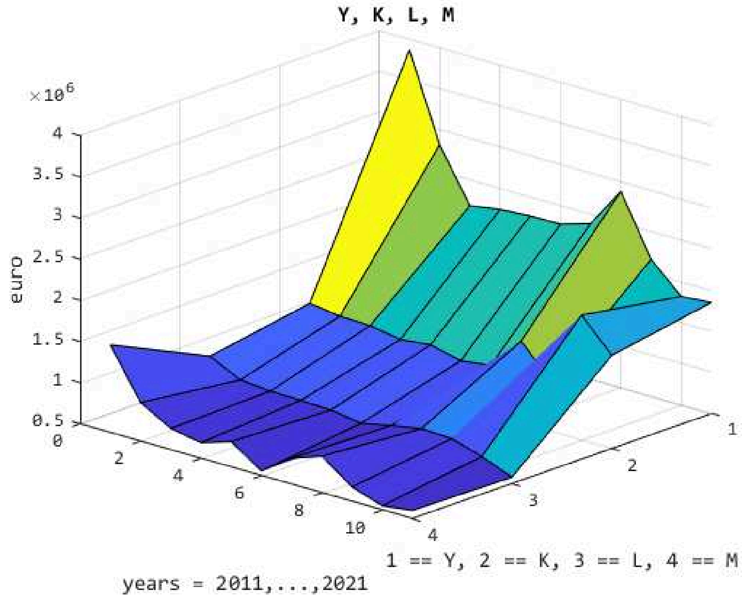
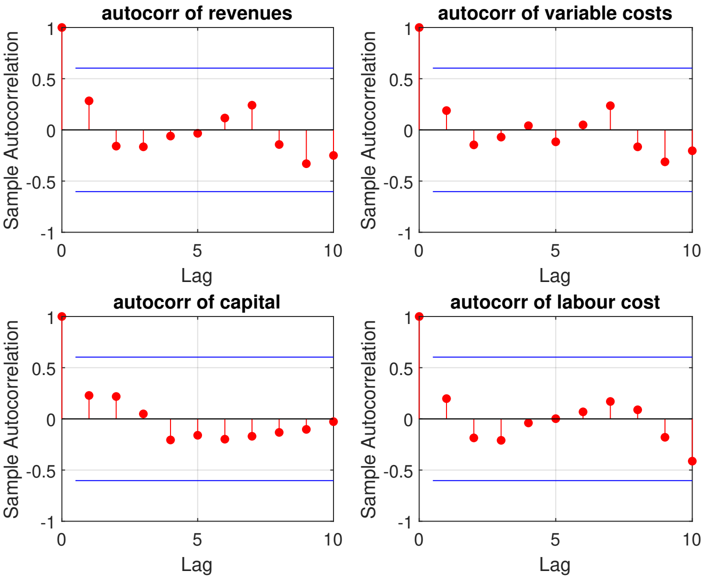


Fig. A7 sample autocorrelation functions



References

[1] Gerard Debreu, *Theory of Value*, *Wiley*, New York, 1959;

[2] Carlo A. Favero, *Econometria. Modelli e Applicazioni in Macroeconomia*, *Carocci*, Roma, 1998;

[3] Tjalling Koopmans, *Three Essays on the State of Economic Science*, Essay 1,*McGraw-Hill*, 1957;

[4] Andreu Mas Colell, Michael D. Whinston, and Jerry R. Green, *Microeconomic Theory*, *Oxford University Press*, New York, Oxford, 1995, pp. 128-136;