Small Bus Econ
https://doi.org/10.1007/s11187-024-00915-1

RESEARCH ARTICLE

Was Robert Gibrat right? A test based on the graphical model methodology

Marco Guerzoni · Luigi Riso · Marco Vivarelli

Accepted: 3 March 2024

Abstract Using both regression analysis and an unsupervised graphical model approach (never applied before to this issue), we confirm the rejection of Gibrat's Law (stating that a firm's growth is independent of that firm's initial size) when our firm-level data are considered over the entire investigated period, while the opposite is true when we allow for market selection; indeed, the growth behavior of the surviving most efficient firms is in line with Gibrat's Law. This evidence reconciles early and current literature and may have interesting implications in terms of both theoretical research and policy suggestions regarding subsidies to small firms, which do not necessarily grow faster than their larger counterparts.

M. Guerzoni
DEMS, University of Milano-Bicocca, Milan, Italy

M. Guerzoni
BETA, University of Strasbourg, Strasbourg, France

L. Riso · M. Vivarelli (✉)
Department of Economic Policy, Catholic University, Milan, Italy
e-mail: marco.vivarelli@unicatt.it

M. Vivarelli
UNU-MERIT, Maastricht, Netherlands

M. Vivarelli
IZA, Bonn, Germany

Plain English Summary Challenging Gibrat's Law, this study reveals that small Italian firms initially outpace larger ones in growth, but selection evens the field over time; this evidence calls for smarter and targeted policies. Indeed, our analysis challenges the widely accepted result that small firms grow faster than their larger counterparts, thus rejecting Gibrat's Law stating that a firm's growth is independent of that firm's initial size. Using a unique combination of regression analysis and innovative graphical models, we tracked newly founded Italian manufacturing firms over 11 years. We discovered that initially, smaller firms grow faster, rejecting Gibrat's Law. However, over time, as less efficient firms exit the market, the surviving, more efficient firms display growth patterns consistent with Gibrat's Law. This finding bridges the gap between previous and recent studies on firm growth. The study's key implication is for policy makers: support for young and small firms is crucial, but policy focus should shift to ensuring markets function efficiently, and aiding the most promising businesses. This approach could foster a more dynamic and robust economic environment, benefiting society by promoting sustainable business growth and stability.

Keywords Gibrat's Law · Firm survival · Market selection · Firm growth

JEL Classification L11

Published online: 04 April 2024

1 Introduction

The standard interpretation of the law put forward by Gibrat (1931) is that the growth rate of a given firm is independent of its initial size. However, while earlier studies, based on limited samples of well-established and large companies, confirmed Gibrat's Law, since the contribution by Mansfield (1962) subsequent and recent research has rejected it (see next section). Indeed, the current consensus within the extant empirical literature is that smaller firms show a higher growth rate than their larger counterparts.

One way to approach this puzzle is to take into account the fact that earlier studies focused on companies which were the outcome of a previous (not investigated) market selection and so represented the industrial "core" within which Gibrat's Law tended to be confirmed. On the other hand, the current literature, based on more comprehensive and large datasets including newly-founded and small firms, tests Gibrat's Law by investigating a given population of the same firms over time and in doing so magnifies the role of smaller and younger firms (which must grow faster in order to reach a minimum efficient size and survive), and rejects the Law.

The purpose of this study is to test whether a given population of firms tends to converge to Gibrat-like behaviour over time, allowing for market selection and for the correlated exit of the less efficient firms. In this context,—and in contrast with the studies discussed above, the final population is smaller than the initial one and is made up of only the surviving, most efficient companies. In this setting, it may well be the case that Gibrat's Law is rejected when considering the entire population of firms and the entire period examined (consistently with the current literature), while it is confirmed when considering only the tracked population of surviving firms (consistently with the earlier literature).[1]

In other words, the hypothesis tested in this work is the following: "*Gibrat's Law, although to be rejected in general, might actually be accepted when market selection generates a sort of 'steady state', where a much more homogeneous population of surviving firms may behave according to the Law*".

Another important novelty of this study is the method adopted. While authors of earlier and current empirical literature have used econometric techniques, in this work we also include an innovative unsupervised approach generating graphical models able to elicit the intrinsic structure of the data and represent them as a network. More specifically, graphical models represent the data in a fully unsupervised way without testing an a priori assumed functional form. Moreover, they offer an overview of the structural relationship between all variables and not only for each covariate with respect to a dependent variable. However, standard econometric analysis is proposed as a complementary analysis and for the sake of comparison with previous literature.

2 Past and current literature

In 1931 the French engineer and economist Robert Gibrat put forward his Law of Proportionate Effect stating that the proportional rate of growth of a given company is independent of its absolute size at the beginning of the investigated period (now well-known worldwide and since then called Gibrat's Law, or rule of proportionate growth, see Gibrat, 1931).[2]

After the Second World War, Gibrat's Law of Proportionate Effect was very popular among both economists and statisticians (Santarelli et al. 2006). The main reason was that the Law was fully consistent with a log-normal distribution of firm size (or even considered as the data generation process behind such a distribution). In turn, a log-normal distribution of firm size was (and is nowadays) actually observed in virtually all economic sectors, where a vast majority of small- and medium-sized firms coexist with few larger counterparts. Therefore, as stated by Simon and Bonini (1958), if one "…incorporates the Law of Proportionate Effect in the transition matrix of a stochastic process, […] then the resulting steady-state distribution of the process will be a highly skewed

[1] This competitive dynamic is well captured by the seminal theoretical model put forward by Jovanovic (1982), based on Bayesian passive learning, and by the models with active learning (Ericson and Pakes,1995; Pakes and Ericson, 1998).

[2] Edwin Mansfield, in his seminal paper on the *AER*, describes Gibrat's Law in the following terms: "the probability of a given proportionate change in size during a specified period is the same for all firms in a given industry—regardless of their size at the beginning of the period" (Mansfield, 1962, p. 1031).

distribution" (1958, p.609). The empirical consistency between Gibrat's Law and the observed size distribution of firms across different industries was also discussed by Steindl (1965) and treated through examples and simulations by Prais (1976, Chapter 2).

However, although in the long term Gibrat's Law certainly generates a log-normal firm size distribution, the latter does not necessarily require firms' proportional rates of growth. Indeed, if we do not limit our attention to the incumbent firms but extend it to the analysis of industrial dynamics, i.e. the entry and exit of companies within a given industry, a log-normal distribution may also emerge, as the consequence of a small group of persisting larger (core) incumbents coexisting with a large fringe of smaller firms, characterised by churning and turbulence (high entry rates, low survival rates, revolving-door firms, see Geroski, 1995). In other words, Gibrat's Law is a sufficient, but not necessary condition for generating an observable log-normal distribution of firm size. This argument permits the possibility of refuting the Law, without being in contrast with the revealed skewed distribution of firm size (see below).

Indeed, while earlier studies based on subsamples of large and mature firms had tended to confirm the Law (Hart & Prais, 1956; Hymer & Pashigian, 1962; Simon & Bonini, 1958), further research began to challenge its overall validity. It is important to note that earlier studies were based on limited databases, only comprising large incumbents, namely companies quoted on the London Stock Exchange (Hart and Prais (1956)); the largest 500 *Fortune* US corporations (Simon and Bonini (1958)); the 1,000 US largest manufacturing firms in the period from 1946–1955 (Hymer and Pashigian (1962)). In other words, earlier consensus regarding the validity of Gibrat's Law was based on empirical tests limited to the core of larger incumbent companies, so neglecting the role of both incumbent SMEs and newly-founded firms.

The turning point in the literature was the seminal contribution by Mansfield (1962), investigating the U.S. steel, petroleum and tyre sectors in different time periods and finding that Gibrat's Law was failing in the majority of cases, with smaller firms growing faster than their larger counterparts. Indeed, when studies take SMEs into account it is found that these incorporate their need to reach the minimum efficient size (MES) and so engage in accelerated growth.

Mansfield's outcome has largely been confirmed by subsequent empirical studies using more comprehensive specifications and also including firms' age and other controlling regressors. For instance, Hall (1987) studied 1,778 US manufacturing firms which had already reached a certain minimum size (measured in terms of employment) and belonged to two samples spanning the periods 1972–1979 and 1976–1983. Unlike Mansfield (1962), Hall directly regressed growth rates on the logarithm of the initial size and found that the observed negative relationship between size and growth was robust to corrections for both sample attrition and heteroskedasticity.[3]

Evans (1987a) analyzed 100 4-digit manufacturing industries using firm level data drawn from the US Small Business Data Base (42,339 firms). The novel feature of this study was the introduction of age (in addition to size measured in terms of employment) as a possible factor in explaining departure from Gibrat's Law. A negative relationship between growth and size was found in 89% of the industries examined, while a negative relationship between growth and age was verified in 76% of firms. As in the previous study, the estimation procedure controlled for sample selection bias and heteroskedasticity (see also Evans, 1987b).

The work by Dunne et al. (1989) also supported the rejection of Gibrat's Law: within each age category, growth rates turned out to decline along employment size classes. Dunne, Roberts and Samuelson obtained these results from data on 219,754 individual plants, rather than firms as in the previous studies, collected in five US censuses of manufacturing (1963-67-72-77-82).

Another important contribution to the investigation of Gibrat's Law was put forward by Dunne and Hughes (1994), who tested the Law of Proportionate Effect over the periods 1975–80 and 1980–85 using 2,149 quoted and unquoted UK companies belonging to 19 different manufacturing industries. After controlling for sample attrition and heteroskedasticity, Dunne and Hughes found further confirmation that smaller companies tend to grow faster than their larger counterparts; they also found that younger companies, for a given size, tended to grow faster than older ones.

[3] This type of econometric specification will also be adopted in the present study, see next section, Eq. (3).

By the same token, Hart and Oulton (1996) used data from 87,109 UK incumbent companies over the period 1989–93 and tested the Chesher-Mansfield specification (see next section), measuring size in terms of employment, sales and net assets. In all cases, they detected an overall estimated coefficient of less than one: on average, small firms grew more quickly than larger ones; however, they also found a not significant relationship between growth and size when considering only the larger firms. In other words, Gibrat's Law turned out to be rejected in general, but not falsified within the subsample of the core companies (see above).

Audretsch et al. (1999) used an Italian dataset comprising newly-founded manufacturing firms tracked from 1987 to 1993 and found that Gibrat's Law was indeed rejected in the vast majority of industries, both considering the entire set of firms and the limited set of surviving firms.

On the whole, by the end of the'90 s a new consensus had been reached, partially in contrast with that shared during the previous decades: this was that "*Gibrat's Legacy*" (as named by Sutton, 1997; see also Caves, 1998 and Coad, 2009) was defendable not as a general law, but only as a dynamic rule valid for large and mature firms that had already attained the MES level of output but not for their smaller counterparts operating at a sub-optimal scale (Geroski, 1995). In a nutshell, and combining the two consensuses reached by the literature, Gibrat's Law should be considered rejected when all firms are taken into account, but confirmed when only the core companies within industries are considered. Audretsch et al. (2004), for instance, analysed a sample of firms in the hospitality sector and found mixed results. However, when considering only surviving firms or large firms, Gibrat's Law was more likely to hold.

The most recent literature has generally supported this overall conclusion. For instance, Calvo (2006), analysing 1,272 Spanish manufacturing firms over the period 1990–2000, found smaller firms growing faster than larger ones. By the same token, Oliveira and Fortunato (2006), using an unbalanced panel of Portuguese manufacturing firms over the period 1990–2001, found that large and mature firms do have smaller growth rates than small and young firms. Daunfeldt and Elert (2013) studied Swedish firms within five-digit NACE-industries during the period 1998–2004 and confirmed the rejection of Gibrat's Law when considering the entire population of investigated companies; however, Gibrat's Law was more likely to be rejected for industries characterised by a higher MES, while it was more likely to hold in mature industries, in industries with a high degree of group ownership, and in industries with a high market concentration. Tang (2015) studied the Swedish energy market, using an unbalanced longitudinal dataset covering 2,185 firms during the 1997–2011 period, and found an interesting twofold result: on the one hand, Gibrat's Law was rejected, with smaller firms found to grow faster than their larger counterparts; on the other hand, when examining each firm individually, it was found that many Swedish energy firms do behave in accordance with Gibrat's Law, namely those in a *steady state,* i.e. the larger and more mature ones. Distante et al. (2018) ran quantile regression models using annual data covering US manufacturing firms over six decades (1950–2010) and found that, conditional on survival, small establishments grow faster than their larger counterparts. Arouri et al. (2020) studied the pattern of growth of Tunisian firms over the period 1996–2010: their key finding was that consistently with the extant literature, Gibrat's Law was rejected overall, with smaller firms growing faster than their larger counterparts; however, the negative impact of the initial size was found to be larger and more significant for young firms rather than for mature, larger incumbent firms. Elston and Weidinger (2023) investigated MENA companies listed on the United Arab Emirates (UAE) stock exchanges and found that in most industries smaller firms grow faster than larger firms, with three notable exceptions: energy, telecommunications and industrial manufacturing.[4]

All in all, the extant literature seems to support the general idea that Gibrat's Law should be rejected when all firms are taken into account, but can be revived when the core of larger and older incumbents is singled out. As mentioned in the previous section, the purpose (and the novelty) of this paper is to test whether a given population of firms tends to converge to Gibrat-like behaviour over time, allowing for market selection and for the correlated exit of the less

[4] Interestingly enough, these industries are characterised by larger MES and the dominant role of core companies, particularly within a sample of listed firms.

efficient firms. In particular, and differently from the previous studies discussed above, we will start from a brand-new population (1,720 newly-founded Italian manufacturing firms) and will track them over 11 years to test whether convergence to Gibrat-like behaviour emerges over time. Instead of separating *different groups of firms* (i.e. core *vs* fringe), we will deal with the *same population of companies* over time, allowing for market selection and so for the exit of the less efficient firms, the purpose being to investigate whether earlier literature can be reconciled with more recent research, i.e. to test whether the rejection of Gibrat's Law *ex ante* can be coupled with the defence of the Law *ex post* (see the hypotheses proposed in the next section).

To our knowledge, only two previous studies have attempted this kind of experiment. In a first work, Lotti et al. (2003) ran quantile regressions using data for 855 Italian manufacturing firms founded in January 1987 and tracked for six years; their main result was that in five industries out of six, Gibrat's Law fails to hold in the years immediately following start-up, whereas it holds, or fails less severely, when firms approach maturity and market selection has done its job. In a later study, the same authors (Lotti et al., 2009) focused on the Italian radio, TV, and communication equipment industry over the period 1987–1994, studying the growth patterns of all the incumbent firms which were active in the sector at the beginning of the examined period (3,285 companies). Consistently with the former study, their results were twofold: on the one hand, Gibrat's Law is rejected over the entire period, with smaller firms growing faster than larger ones; on the other hand, a convergence toward the validity of the Law occurs over time, once the annual regressions are run over the sub-population of surviving firms.

In what follows, the econometric regressions are similar in nature to the work described in Lotti et al. (2003, 2009), while the graphical model approach in testing Gibrat's Law is applied for the first time, at least to our knowledge.

Table 1 summarizes the empirical literature discussed so far and highlights samples and main results regarding the testing of Gibrat's Law from each cited paper.

While the above discussion is specifically devoted to Gibrat's Law, it is also worth briefly referring to the related vast amount of literature focusing on the so called High-Growth-Firms (HGFs, often referred to as "gazelles"). This literature is somewhat tangential but relevant to our current topic, since the role of size is often taken into account, albeit not being its main topic of analysis. As summarised by Coad et al. (2014), researchers in this field have explored whether factors such as innovation and industry affiliation affect a company's likelihood of becoming a HGF, coming to the conclusion that innovation (Audretsch et al., 2014; Cefis & Marsili, 2006; Colombelli et al., 2013; Goel & Nelson, 2022; Zhang & Mohnen, 2022) and sector-specificity (Acs, 2011) can enhance both a firm's survival prospects and its rate of growth.[5] However, even within this strand of literature, there remains a lack of consensus on how a firm's size influences growth dynamics: while most studies tend to confirm a negative relationship between size and growth, some are more nuanced. For instance, Coad et al. (2014) highlight the stylised fact that HGFs tend to be young, but not necessarily small. Additionally, Delmar et al. (2003) show that when growth is measured in absolute terms, large firms tend to grow more, and Acs (2011) demonstrates that there is a subset of high-growth firms which are large.[6]

As we will discuss later in the empirical section, we engage with this literature by incorporating controls for innovativeness, focusing on a single cohort of firms to account for the impact of firm age, and introducing sectoral dummies.

3 Data, hypotheses and preliminary econometric specification

The analysis is based on AIDA-BvD data, which contains comprehensive information on all the Italian firms required to file accounts. Specifically, we

[5] Sectoral belonging impacts firm growth by capturing the stage in the underlying technology life cycle: more specifically, after the introduction of a new technology in a sector, we expect new and small entrepreneurial firms to grow more than large incumbents, which in turn prosper in sectors characterised by mature technologies (Malerba and Orsenigo (1995) and (1996)).

[6] A noteworthy point raised by Nightingale and Coad (2014) is that entrepreneurship literature tends to overemphasise the role of HGFs, which represent a very relevant but small portion of the overall distribution of firm performance.

Table 1 Previous literature: synoptic view

Article	Sample	Gibrat law
Hart and Prais (1956)	UK quoted firms in mining, manufacturing and distribution 1885–1950	YES
Simon and Bonini (1958)	Largest Fortune firms (1954–56)	YES
Hymer and Pashigian (1962)	Largest manufacturing firms (1945–1955)	YES
Mansfield (1962)	U.S. steel, petroleum and tyre sectors	NO
Hall (1987)	US manufacturing firms in two periods, 1972–1979 and 1976–1983	NO
Evans (1987b)	US Small Business Data Base	NO
Dunne et al. (1989)	US manufacturing individual plants	NO
Dunne and Hughes (1994)	quoted and unquoted UK companies in manufacturing 1975–80 and 1980–85	NO
Hart and Oulton (1996	UK incumbent companies (1989–93)	YES, but only for the core
Calvo (2006)	Spanish manufacturing firms (1990–2000)	NO
Audretsch et al. (2004)	Dutch hospitality sector between 1987 and 1991	YES, for surviving and large firms
Oliveira and Fortunato (2006)	Portuguese manufacturing firms (1990- 2001)	NO
Daunfeldt and Elert (2013)	Swedish firms within five-digit NACE-industries (1998–2004)	YES, but only in mature industries
Tang (2015)	Swedish energy market (1997–2011)	YES, but only for large and mature firms
Distante et al. (2018)	US manufacturing firms (1950–2010)	NO
Arouri et al. (2020)	Tunisian firms (1996–2010)	NO, but stronger effect for small firms
Elston and Weidinger (2023)	MENA companies listed on the United Arab Emirates (UAE) stock exchanges	YES, but only for energy, telecommunications and industrial manufacturing
Lotti et al. (2003)	Italian manufacturing firms (1987–2003)	YES, but only in later years
Lotti et al. (2009)	Italian radio, TV and communication equipment industry (1987–1994)	YES, but only for surviving firms

acquired a dataset comprising the entire population[7] of 1,720 new (founded in 2009) Italian manufacturing firms with at least one employee, tracked for 11 years, i.e. until 2020. We selected the following variables of interest: Employees (E); Regional belonging (dummies corresponding to the NUTS-2 classification, R); Sectoral belonging (dummies corresponding to the 2 digits NACE classification, S); Innovativeness (a dummy I that indicates whether a firm is registered as "innovative" according to the Italian decree "innovative firms act 221/2012", see Guerzoni et al., 2021)[8]; Profitability, computed as the ratio between "earnings before taxes" and "revenues from sales and services" (P).

By utilising data on Italian manufacturing firms newly established in 2009, the validity of Gibrat's Law was examined over the entire period of 2010–2020, as well as year-by-year. Through this set of analyses, we aimed to test the following hypotheses jointly, with the objective of reconciling if possible the diverging evidence discussed in Section 2:

(H1) Gibrat's Law is rejected over the entire period (a priori hypothesis);
(H2) A convergence towards a Gibrat-like steady state emerges among the population of surviving firms (a posteriori hypothesis).

If both these hypotheses were to be confirmed, this would mitigate the apparently controversial debate surrounding the validity of Gibrat's Law (see

[7] Since we are dealing with an entire population, our dataset cannot be affected by sample selection.

[8] The combination of sector-specific variables, innovativeness and regional variables can at least partially (given our data limitations and the dummy nature of our controls) capture specific technological characteristics and local knowledge spillovers.

Table 2 Descriptive statistics (year by year and overall)

Datasets		2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	All
N		1720	1573	1403	1259	1126	1030	942	881	798	726	688	1720
G	*Mean*	0.20	−0.04	0.06	0.00	0.03	0.05	0.01	0.05	0.01	0.03	0.00	−0.31
	St. Dev	0.70	0.51	0.50	0.37	0.41	0.36	0.30	0.52	0.23	0.31	0.20	1.82
	Min	−0.85	−0.98	−0.99	−0.99	−0.98	−0.93	−0.97	−0.94	−0.96	−0.88	−0.75	−1.00
	Max	18.00	4.00	9.00	5.00	4.40	3.50	2.58	13.00	2.29	5.90	1.34	35.75
E	*Mean*	14.10	15.50	15.13	15.47	15.64	16.25	17.47	18.60	19.63	20.64	21.40	14.10
	St. Dev	28.31	24.35	26.57	26.81	28.14	30.08	32.23	35.24	37.69	41.82	46.12	28.31
	Min	1	1	1	1	1	1	1	1	1	1	1	1
	Max	450	436	472	480	480	539	562	587	603	636	679	450
P	*Mean*	−0.22	−0.02	−0.50	−0.11	−1.34	0.14	−0.34	0.05	−0.07	0.01	21.44	−0.22
	St. Dev	2.97	0.69	17.87	2.15	32.42	6.66	3.99	2.81	1.68	0.31	562.38	2.97
	Min	−68.34	−19.34	−669.07	−73.20	−1045.74	−11.92	−87.17	−11.78	−44.55	−6.16	−5.40	−68.34
	Max	7.62	10.51	1.00	3.68	1.71	212.52	1.69	81.10	6.09	1.11	14751.00	7.62
I	*Yes*	5	5	5	5	5	5	5	4	4	4	4	5
	No	1715	1568	1398	1254	1121	1025	937	877	794	722	688	1715

G represents the annual employment growth rate, E denotes total employment, P stands for profitability calculated as earnings before taxes, and I is a binary variable indicating whether a firm is classified as an innovative SME. The columns represent distinct years, while 'ALL' pertains to the pooled dataset

Section 2). While the Law may be rejected when examining the overall evolution of a given ex-ante population of companies (a priori hypothesis), it may still accurately describe the patterns of growth for well-established firms within the ex-post sub-population that results from market selection and learning processes (a posteriori hypothesis).

The specification used to test Gibrat's Law econometrically is the following (the same adopted by Evans, 1987a, b, Lotti et al., 2009):

$$G_{i,t} = \beta_0 + \beta_1 log\left(E_{i,t-1}\right) + \beta_1 R_{i,t} + \beta_1 S_{i,t} + \beta_1 log\left(P_{i,t-1}\right) + \beta_1 I_i + \varepsilon_{i,t}, \tag{1}$$

where $G_{i,t} = (E_{i,t} - E_{i,t-1})/E_{i,t-1}$ is the employment growth rate of firm i at time t; β_0 is the intercept of the model and serves as a reference point, indicating the expected outcome when no explanatory factors are at play; $E_{i,t-1}$ is the number of employees the firm has i at time t-1; regional dummies, sectoral dummies, profitability (P) and innovativeness (I) act as controls.[9]

[9] Controlling for age is obviously useless in our context, since we are dealing with a sole cohort, namely companies founded in 2009.

Chesher (1979) pointed to the coefficient β_1 in order to test the validity of Gibrat's Law through the significance of the relevant parameter. In particular, if β_1= 0, Gibrat's Law holds; if β_1< 0, smaller firms grow at a higher rate than their larger counterparts, while the opposite is the case if β_1> 0.

We estimated Eq. (1) for each period t on the subsamples of firms still existing at t; moreover, we estimated overall employment growth over the entire investigated period, with growth $G_{i,t} = -100\%$ for firms that have exited the market (as in Evans, 1987a, b).

Table 2 presents the summary statistics for each period, showing the progressive reduction in the population as well as the corresponding growth rate and other variables included in the model for each year. It is noteworthy that only 688 companies out of 1,720 survived until the end of the investigated period, pointing to a ten-year survival rate in Italian manufacturing, equal to 40% (this is not surprising and in line with the stylised facts pointed out by Geroski, 1995). In order to survive, newly-founded firms must grow, with an average size increasing from 14.10 employees at the beginning of the period to 21.40 at the end of the period. Profitability is hard to achieve for these young firms, although more likely in the later years, and innovativeness (as defined in this study) is a really exceptional attribute.

4 A novel graphical model approach

One of the main contributions of this study is that we made use of unsupervised graphical models (GM) to gain a different and more comprehensive perspective on whether a given population of firms tends to converge to Gibrat-like behaviour over time. In fact, GM allow us to describe jointly the overall structure of dependency between variables and in this way we can capture multiple relationships, including non-linear and conditional dependencies, and both direct and indirect influences. In particular, in our context GM can depict not only whether the proxy for size directly affects growth, but also any other mediating effects.

In more detail, graphical models are a framework combining network representation and probability theory to specify conditional independence relationships between random variables in a given dataset. These relationships are represented through a graphical representation, specifically a graph $\mathcal{G}$(V,L), where V is a finite set of nodes corresponding to the variables of interest and L is the set of links in the network, representing the conditional dependence between any pair of variables (Lauritzen, 1996).[10]

One of the central problems in GM representation is the estimation of the underlying probability distributions of the variables from a finite sample. Chow and Liu (1968) proposed an approach for discrete variables that approximates their probability functions via probability distribution of the second-order tree dependence. The connection between nodes of the tree represents the unknown joint probability of the nodes (or associated variables), providing information on their mutual dependence or mutual information. Specifically, Chow and Liu (1968) found that a probability distribution of a tree dependence approximates the true value probability of a set of discrete random variables composing the tree, if and only if the latter has maximum mutual information. Under the assumption that the cell probabilities of discrete random variables factorise according to an unknown tree τ written as $\mathcal{G}_D = (\Delta, L_\Delta)$, they can be written as:

$$p(d|\tau) = \frac{\prod_{u,v \in E_\Delta} p(d_u, d_v)}{\prod_{v \in \Delta} p(d_v)^{d_v - 1}} \quad (2)$$

where d_v is the number of links incident to node v, namely the degree of v. According to Eq. (2), the maximized log-likelihood, up to a constant, turns out to be $\sum_{(u,v)\in L_\Delta} I_{u,v}$, where $I_{u,v}$ is the mutual information between nodes u and v. It is worth noting that the mutual information between two variables is defined as a measure of their closeness (Lewis, 1959), and therefore is a dimensionless, non-negative, and symmetric quantity which measures the reduction of uncertainty about a random variable, given the knowledge of another. Put differently, Chow and Liu's (1968) algorithm employs the concept of mutual information to quantify the strength of the connections which link the variables depicted in the graphical representation (Riso et al., 2023).

Prior to the development of the Chow-Liu algorithm, various other algorithms were created to determine the probabilistic structure and corresponding maximum-likelihood estimator. Specifically, Kruskal (1956) proposed a simple and efficient solution to this problem by starting with a null graph and adding the edge with the highest weight at each step, as long as it does not form a cycle with previously chosen edges. Edwards et al. (2010) extended the Chow-Liu algorithm to be applied to mixed data sets **X**, using mutual information between discrete and continuous variables**.** This algorithm relies on the use of mutual information between a discrete variable, D_u, and a continuous variable, C_v. It is characterised by the marginal model, which turns out to be an ANOVA model (Edwards, 1995). It is worth noting that when dealing with mixed variables, the evaluation of the mutual information $I(d_u, c_v)$ between each pair of nodes requires distinguishing between the case when the variance of C_v is distributed homogeneously across the level of the discrete variable D_u from the case when it is heterogeneously distributed (Edwards, 1995). As pointed out by Edwards et al. (2010), one

[10] In order to better explain the following analysis, we introduced a mixed dataset X, composed of n observations and p variables. We split the variables into r discrete, D = (D_1,…,D_r) and q continuous C = (C_1,…,C_q). We denoted the i-observation of X = (D,C) as 〖(d〗_i,c_i) with d_i and c_i representing the i-observation of the variables D_i∈D and C_i∈C, respectively. Given the one-to-one correspondence between variables and nodes, we can write the set of nodes as V = {Δ,Γ}, where Δ and Γ are the nodes corresponding to the variables in D and C respectively.

of the disadvantages of selecting a tree on the basis of maximum likelihood is that it always includes the maximum number of edges, even if the latter are not supported by data. They therefore suggested the use of one of the following measures to avoid this drawback:

$$I^{BIC} = I(x_i, x_j) - log(n)k_{x_i,x_j}; I^{AIC} = I(x_i, x_j) - 2k_{x_i,x_j} \tag{3}$$

The degrees of freedom associated with the pair of variables x_i and x_j are represented by k_{x_i,x_j}, and are determined based on the nature of the variables involved, continuous or discrete[11]. These measures are used in an algorithm proposed by Edwards et al. (2010) to determine the best spanning tree. The algorithm stops when the graph has reached its maximum number of edges. The algorithm can generate either a tree or a forest, where a forest is a group of trees.

To test the validity of Gibrat's Law, we employed the extension of the Chow-Liu algorithm (Chow & Liu, 1968) proposed by Edwards et al. (2010) for mixed datasets. This methodology allowed us to map the conditional dependence relationships of the variables involved in this analysis onto a graph $\mathcal{G}$ (V,L), where V is a finite set of nodes with direct correspondence to the variables of interest and L is the set of links in the network (Lauritzen, 1996). The links represent the conditional dependence between any pair of variables.[12] Specifically, the GM employed in this paper belong to the class of multivariate distributions, whose conditional independence properties are encoded in a tree/forest in the following way: the absence of a link between two nodes represents conditional independence between the corresponding variables (Jordan, 2004).

In the context of this study, if there is a direct connection between node G (employment growth rate at time *t*) and node E (number of employees at time *t-1*), or the connection is mediated by another node, Gibrat's Law does not hold. Conversely, if node G is not connected with node E, Gibrat's Law holds. In other words, we are examining the dependence between variable G (employment growth rate at time t) and variable E (number of employees at time t-1) using the GM to validate Gibrat's Law over time. In more detail, the GM methodology allowed us to understand how the relationships between the variables involved in the model change over time; in particular, we built both a GM for each year and one to test the overall relationships over the entire period; therefore, we exactly mimic the econometric setting put forward in the previous section.[13]

On the whole, the unsupervised learning approach proposed in this study does not supplant the econometric approach but rather enhances it, enabling us to test Gibrat's Law without being constrained by the assumptions inherent in econometric modelling.

5 Empirical findings

Table 3 presents the output of the regressions, which can be considered as our preliminary baseline. In the last column, which displays the results over the entire investigated period, the key coefficient log (E) is negative and significant, rejecting Gibrat's Law and supporting Hypothesis 1.[14] However, the regressions on single periods tell a different story: the initial size displays a significant and negative impact only in the first seven years, whereas—allowing for market selection—this significance disappears as of 2017 (supporting Hypothesis 2). As far as the controls are

[11] For discrete random variables, the degrees of freedom are equal to $|D_u| - 1$, where D_u is the number of levels of the discrete random variable. However, for continuous random variables, there is only 1 degree of freedom. Under marginal independence, the statistic $I_{u,v}$ has an asymptotic χ^2 distribution (Edwards et al., 2010).

[12] It is important to note that we cannot take the magnitude of these links into account, but only their presence or absence, which defines the structure of the tree itself (Riso & Guerzoni, 2022).

[13] The analytical difference between GM and multivariate analysis is that GMs are based on the mutual information of random variables, while traditional regressions are based on their covariance: both mutual information and covariance are measures of distance but the correlation is based on difference in levels and mutual information in logarithms. However, the most important difference is that graphical models do not provide information on a linear relation of an a priori given dependent variable and a set of covariates, but as an exercise of unsupervised structural learning, providing information about the joint relation between all variables.

[14] In line with the recent literature (see Section 2).

Table 3 Regression analyses (year by year and overall); dependent variable: employment growth rate

	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	All
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Intercept	**0.510^{***}**	0,052	**0.269^{***}**	0,074	**0.151^{**}**	0,064	**0.139^{**}**	0,113	−0,029	−0,046	−0,045	0,013
	(0.077)	*(0.071)*	*(0.082)*	*(0.057)*	*(0.074)*	*(0.058)*	*(0.068)*	*(0.116)*	*(0.041)*	*(0.048)*	*(0.041)*	*(0.2)*
log (E)	**-0.156^{***}**	**-0.041^{**}**	**-0.071^{***}**	**-0.064^{***}**	**-0.059^{***}**	**-0.037^{***}**	**-0.023^{**}**	−0,043	−0,011	−0,016	0,004	**-0.174^{***}**
	(0.029)	*(0.016)*	*(0.017)*	*(0.015)*	*(0.014)*	*(0.012)*	*(0.01)*	*(0.027)*	*(0.008)*	*(0.011)*	*(0.008)*	*(0.063)*
I	−0,02	−0,05	−0,002	0,168	0,059	**0.095^{**}**	-0,047	0,104	-0,003	0,053	0,02	0,726
	(0.083)	*(0.112)*	*(0.086)*	*(0.107)*	*(0.082)*	*(0.037)*	*(0.065)*	*(0.102)*	*(0.059)*	*(0.072)*	*(0.035)*	*(0.642)*
log (P)	0,077	**0.095^{*}**	**0.209^{***}**	**0.297^{***}**	**0.241^{***}**	**0.133^{***}**	**0.138^{***}**	**0.138^{*}**	**0.174^{***}**	**0.162^{***}**	**0.071^{**}**	**0.335^{**}**
	(0.073)	*(0.051)*	*(0.057)*	*(0.049)*	*(0.049)*	*(0.046)*	*(0.037)*	*(0.079)*	*(0.04)*	*(0.051)*	*(0.034)*	*(0.144)*
Regional dummies	YES											
Sectoral Dummies												
Observations	1720	1573	1403	1259	1126	1030	942	881	798	726	688	1720

OLS regressions; robust standard errors reported in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%; significant coefficients in bold

The columns (1)-(11) represent distinct years, while 'All' (12) pertains to the pooled dataset

concerned, while profitability boosts growth, innovation does not seem to play a significant role.[15]

Figure 1 illustrates the GMs generated using the algorithm introduced by Edwards (1995) and depicts the investigated nodes, with the key variables growth (G) and employment (E) highlighted in orange and blue respectively.

In Fig. 1, each node represents a variable and the presence of a link indicates a relation of conditional dependences. If two nodes are connected via a third variable, the two variables are in a dependence relation, conditional to the distribution of the third one. For instance, in 2010 growth (G) depends directly on employment (E), while it is independent of the dummy variable for innovative firms (I).

Over the entire period (Fig. 1, last panel: All) growth is associated with the initial size, through the mediation of sectoral belonging. This implies that for each given sector, growth and size are not independent. Consequently, Gibrat's Law is rejected, since size does have an effect on growth, albeit varying across sectors[16] (this evidence supports Hypothesis 1).

However, shifting to the annual analyses, our results present a different picture. In the first three periods, growth is directly linked to employment: even independently from sectoral belonging, size does influence the growth rate of firms. In the subsequent four periods, this relationship persists, but is mediated by profitability (and sectoral belonging in 2015). This evidence suggests that Gibrat's Law is initially rejected but the dependency becomes weaker over time. Finally, as was the case for the regressions, as of 2017 growth and employment were no longer connected, corroborating the interpretation of Gibrat's Law as a steady state convergence (thus supporting Hypothesis 2).

It is noteworthy that the results from the graphical models not only align with the regression results, but also provide a more intuitive and interpretable representation of the data. Specifically, we were able to observe an initial direct correlation between growth

[15] This outcome may be due to the very small incidence of innovative firms within our population.

[16] The mediating role of sectoral belonging might be due to the different "minimum efficient sizes" required by the different sectors (while interesting, this research perspective is beyond the scope of the present work and cannot be tested, given the available data).

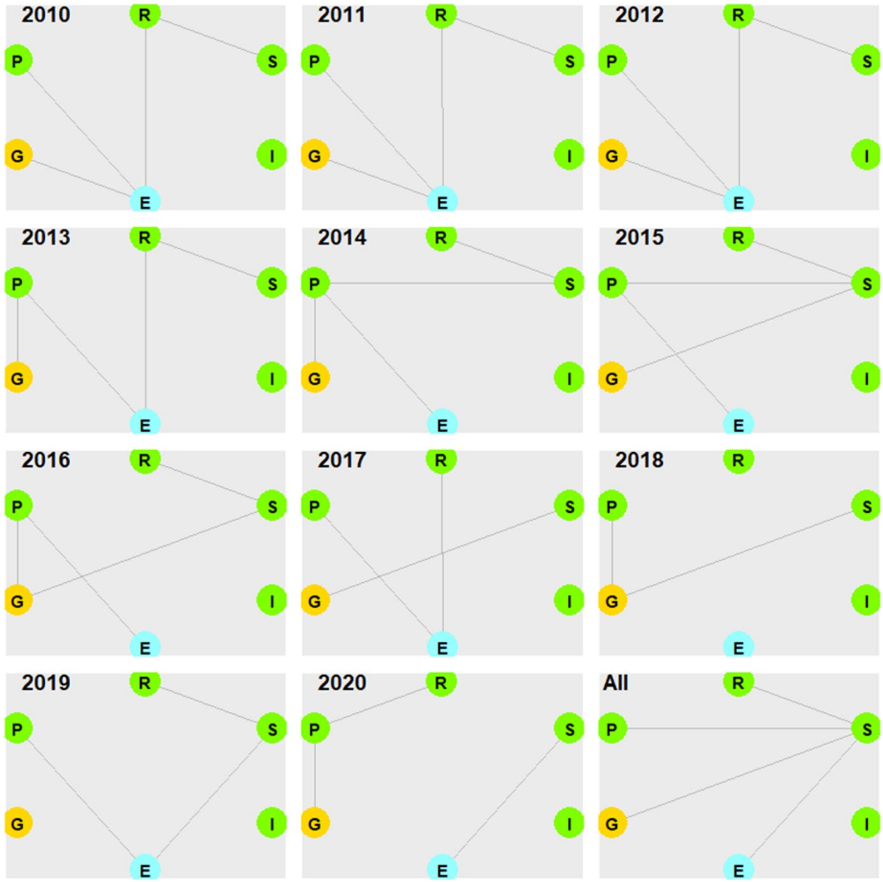

Fig. 1 Graphical Models: direct and conditional dependences between variables (year by year and overall). *Notes*: Each subfigure represents the graphical model for a distinct year, while 'ALL' pertains to the pooled dataset. Nodes represent variables, and connections indicate the absence of statistical independence. G denotes the employment growth rate, E stands for employment, I is innovativeness, S represents sectoral belonging, P denotes profitability, and R stands for the NUTS-2 region

and size, then a temporary mediation effect of profitability, while eventually any type of either direct or indirect dependency fades away as market selection proceeds.

6 Conclusions

As discussed in detail in Section 2, in the extant (and partially controversial) literature the conclusion has been reached that Gibrat's Law can be rejected when all firms are taken into account, but can be confirmed when the core of larger and older incumbents is isolated.

Differently from most previous studies, in this paper we did not single out different groups of firms (i.e. core *vs* fringe), instead tracking a brand new population of companies over time, allowing for market selection and so discovering whether a given population of firms tends to converge to Gibrat-like behaviour through time. In so doing, we tested whether Gibrat's Law, albeit being general refutable, can actually be accepted when market selection generates a sort of "steady state", where a much more homogeneous population of surviving firms may behave according to the Law (see Section 1).

Using both standard econometrics and a novel unsupervised approach generating graphical models, in this paper it has been shown that the early and current literature testing Gibrat's Law can be indeed reconciled; in particular, the rejection of Gibrat's Law *ex ante* can be coupled with a defence of the Law *ex post*. More specifically, while we have confirmed rejection of the Law when firms were considered over the entire investigated period, we have shown the opposite when we allowed for market selection and we tracked only the surviving companies. Indeed, the growth behaviour of the re-shaped (smaller) population of surviving most efficient firms was in line with Gibrat's intuition.

This twofold evidence may have interesting implications in terms of both applied and theoretical research on the one hand, and policy options on the other. In particular, policy makers should take into account the fact that employment growth crucially depends on the combination of different factors characterising industrial dynamics, such as new firm formation, firm size, survival rates and market selection. Indeed, the evidence supporting rejection of Gibrat's Law has laid the groundwork for policies aimed at bolstering young and small firms; if young and smaller firms exhibit more substantial growth than mature and larger incumbents, thereby contributing to increased employment and added value in the economy, this provides support for interventions in favour of entrepreneurial activities.

However, if this deviation from Gibrat's Law is primarily driven by the survival dynamics of underperforming firms in their initial years, the policy implications might be different. Indeed, if after this initial period surviving firms stabilise and conform to Gibrat's Law (as shown in this study), interventions should be more cautious and selective. In particular, the policy focus should not be solely on assisting small firms in general but rather on ensuring that market mechanisms operate efficiently during a company's early stages, facilitating the survival of the most promising ventures. In this context, policies guaranteeing liquidity and survival chances to struggling firms may not be the optimal choice. Conversely, greater emphasis should be placed on institutions and organisations capable of channelling funds to the most promising enterprises, such as venture capital firms or sector-specific industrial policies targeting high-potential sectors.

Acknowledgements Marco Vivarelli acknowledges support from the Italian Ministero dell'Istruzione, dell'Università e della Ricerca (PRIN-2017, project 201799ZJSN: "Technological change, industry evolution and employment dynamics"; principal investigator: Marco Vivarelli).

Funding Open access funding provided by Università Cattolica del Sacro Cuore within the CRUI-CARE Agreement.

Data availability The data that support the findings of this study are available from Bureau van Dijk (BvD), but restrictions apply to the availability of these data, which were used under license and so are not publicly available. The data are, however, available from the authors upon reasonable request and with the permission of BvD.

References

Acs, Z. (2011). High impact firms: Gazelles revisited. In Fritsch (Ed.), *Handbook of research on entrepreneurship and regional development*. Elgar. Chapter: 6. https://doi.org/10.4337/9780857936493.

Arouri, H., Youssef, A. B., Quatraro, F., & Vivarelli, M. (2020). Drivers of growth in Tunisia: Young firms vs incumbents. *Small Business Economics, 54*, 323–340. https://doi.org/10.1007/s11187-018-00133-6

Audretsch, D. B., Klomp, L., Santarelli, E., & Thurik, A. R. (2004). Gibrat's Law: Are the services different? *Review of Industrial Organization, 24*, 301–324. https://doi.org/10.1023/B:REIO.0000038273.50622.ec

Audretsch, D. B., Coad, A., & Segarra, A. (2014). Firm growth and innovation. *Small Business Economics*, *43*, 743–749. https://doi.org/10.1007/s11187-014-9560-x

Audretsch, D. B., Santarelli, E. & Vivarelli, M. (1999). Start-up size and industrial dynamics: Some evidence from Italian manufacturing. In *Entry and post-entry performance of newborn firms*. Routledge (pp. 44–61). 2007. https://doi.org/10.1016/S0167-7187(98)00002-2.

Calvo, J. L. (2006). Testing Gibrat's Law for small, young and innovating firms. *Small Business Economics, 26*, 117–123. https://doi.org/10.1007/s11187-004-2135-5

Caves, R. E. (1998). Industrial organization and new findings on the turnover and mobility of firms. *Journal of Economic Literature, 36*, 1947–1982.

Cefis, E., & Marsili, O. (2006). Survivor: The role of innovation in firms' survival. *Research Policy, 35*(5), 626–641. https://doi.org/10.1016/j.respol.2006.02.006

Chesher, A. (1979). Testing the law of proportionate effect. *Journal of Industrial Economics, 27*, 403–411. https://doi.org/10.2307/2097961

Chow, C., & Liu, C. (1968). Approximating discrete probability distributions with dependence trees. *IEEE Transactions on Information Theory, 14*, 462–467. https://doi.org/10.1109/TIT.1968.1054142

Coad, A., Daunfeldt, S. O., Hölzl, W., Johansson, D., & Nightingale, P. (2014). High-growth firms: Introduction to the special section. *Industrial and Corporate Change, 23*, 91–112. https://doi.org/10.1093/icc/dtt052

Coad, A. (2009). *The growth of firms: a survey of theories and empirical evidence*. Elgar. https://doi.org/10.4337/9781848449107.

Colombelli, A., Haned, N., & Le Bas, C. (2013). On firm growth and innovation: Some new empirical perspectives using French CIS (1992–2004). *Structural Change and Economic Dynamics, 26*, 14–26. https://doi.org/10.1016/j.strueco.2013.03.002

Daunfeldt, S. O., & Elert, N. (2013). When is Gibrat's Law a law? *Small Business Economics, 41*, 133–147. https://doi.org/10.1007/s11187-011-9404-x
Delmar, F., Davidsson, P., & Gartner, W. B. (2003). Arriving at the high-growth firm. *Journal of Business Venturing, 18*, 189–216. https://doi.org/10.1016/S0883-9026(02)00080-0
Distante, R., Petrella, I., & Santoro, E. (2018). Gibrat's Law and quantile regressions: An application to firm growth. *Economics Letters, 164*, 5–9. https://doi.org/10.1016/j.econlet.2017.12.028
Dunne, P., & Hughes, A. (1994). Age, size, growth and survival: UK companies in the 1980s. *Journal of Industrial Economics, 42*, 115–140. https://doi.org/10.2307/2950485
Dunne, T., Roberts, M. J., & Samuelson, L. (1989). The growth and failure of US manufacturing plants. *Quarterly Journal of Economics, 104*, 671–698. https://doi.org/10.2307/2937862
Edwards, D., De Abreu, G. C., & Labouriau, R. (2010). Selecting high-dimensional mixed graphical models using minimal aic or bic forests. *BMC Bioinformatics, 11*, 1–13. https://doi.org/10.1186/1471-2105-11-18
Edwards, D. (1995). *Introduction to graphical modelling*. Springer Science & Business Media. https://doi.org/10.1007/978-1-4684-0481-4.
Elston, J. A., & Weidinger, A. (2023). An empirical examination of firm growth in the MENA region through the lens of Gibrat's Law. *Small Business Economics, 60*, 121–213. https://doi.org/10.1007/s11187-022-00640-7
Ericson, R., & Pakes, A. (1995). Markov-perfect industry dynamics: A framework for empirical work. *Review of Economic Studies, 62*, 53–82. https://doi.org/10.2307/2297841
Evans, D. S. (1987a). The relationship between firm growth, size, and age: Estimates for 100 manufacturing industries. *Journal of Industrial Economics, 35*, 567–581. https://doi.org/10.2307/2098588
Evans, D. S. (1987b). Tests of alternative theories of firm growth. *Journal of Political Economy, 95*, 657–674. https://doi.org/10.1086/261480
Geroski, P. A. (1995). What do we know about entry? *International Journal of Industrial Organization, 13*, 421–440. https://doi.org/10.1016/0167-7187(95)00498-X
Gibrat, R. (1931). *Les inegalites economiques*. Librairie du Recueil Sirey.
Goel, R. K., & Nelson, M. A. (2022). Employment effects of R&D and process innovation: Evidence from small and medium-sized firms in emerging markets. *Eurasian Business Review, 12*(1), 97–123. https://doi.org/10.1007/s40821-022-00203-6
Guerzoni, M., Nava, C. R., & Nuccio, M. (2021). Start-ups survival through a crisis. Combining machine learning with econometrics to measure innovation. *Economics of Innovation and New Technology, 30*, 468–493. https://doi.org/10.1080/10438599.2020.1769810
Hall, B. (1987). The relationship between firm size and firm growth in the US manufacturing sector. *Journal of Industrial Economics, 3*, 583–606. https://doi.org/10.2307/2098589
Hart, P. E., & Oulton, N. (1996). Growth and size of firms. *Economic Journal, 106*, 1242–1252. https://doi.org/10.2307/2235518
Hart, P. E., & Prais, S. J. (1956). The analysis of business concentration: A statistical approach. *Journal of the Royal Statistical Society, 119*, 150–191. https://doi.org/10.2307/2342882
Hymer, S., & Pashigian, P. (1962). Firm size and the rate of growth. *Journal of Political Economy, 70*, 556–569. https://doi.org/10.1086/258716
Jordan, M. I. (2004). Graphical models. *Statistical Science, 19*, 140–155. https://doi.org/10.1214/088342304000000026
Jovanovic, B. (1982). Selection and evolution of Industry. *Econometrica, 50*, 649–670. https://doi.org/10.2307/1912606
Kruskal, J. B. (1956). On the shortest spanning subtree of a graph and the traveling salesman problem. *Proceedings of the American Mathematical Society, 7*(1), 48–50. https://doi.org/10.1090/S0002-9939-1956-0078686-7
Lauritzen, S. L. (1996). *Graphical models*. Oxford statistical science series. Oxford University Press. https://doi.org/10.1093/oso/9780198522195.001.0001
Lewis, P. M., II. (1959). Approximating probability distributions to reduce storage requirements. *Information and Control, 2*, 214–225. https://doi.org/10.1016/S0019-9958(59)90207-4
Lotti, F., Santarelli, E., & Vivarelli, M. (2003). Does Gibrat's Law hold among young, small firms? *Journal of Evolutionary Economics, 13*, 213–235. https://doi.org/10.1007/s00191-003-0153-0
Lotti, F., Santarelli, E., & Vivarelli, M. (2009). Defending Gibrat's Law as a long-run regularity. *Small Business Economics, 32*, 31–44. https://doi.org/10.1007/s11187-007-9071-0
Malerba, F., & Orsenigo, L. (1995). Schumpeterian patterns of innovation. *Cambridge Journal of Economics, 19*(1), 47–65. https://doi.org/10.1093/oxfordjournals.cje.a035308
Malerba, F., & Orsenigo, L. (1996). Schumpeterian patterns of innovation are technology-specific. *Research Policy, 25*(3), 451–478. https://doi.org/10.1016/0048-7333(95)00840-3
Mansfield, E. (1962). Entry, Gibrat's Law, innovation, and the growth of firms. *American Economic Review*, *52*, 1023–1051. https://www.jstor.org/stable/1812180. Accessed 22 Mar 2024.
Nightingale, P., & Coad, A. (2014). Muppets and gazelles: Political and methodological biases in entrepreneurship research. *Industrial and Corporate Change, 23*(1), 113–143. https://doi.org/10.1093/icc/dtt057
Oliveira, B., & Fortunato, A. (2006). Testing Gibrat's Law: Empirical evidence from a panel of portuguese manufacturing firms. *International Journal of the Economics of Business, 13*, 65–81. https://doi.org/10.1080/13571510500519996
Pakes, A., & Ericson, R. (1998). Empirical implications of alternative models of firm dynamics. *Journal of Economic Theory, 79*, 1–45. https://doi.org/10.1006/jeth.1997.2358
Prais, S. J. (1976). *The evolution of giant firms in Britain*. Cambridge University Press.
Riso L, Zoia M. G., & Nava C. R. (2023). Feature selection based on the best-path algorithm in high dimensional graphical models. *Information Sciences, 119601*. https://doi.org/10.1016/j.ins.2023.119601.
Riso, L., & Guerzoni, M. (2022). Concept drift estimation with graphical models. *Information Sciences, 606*, 786–804. https://doi.org/10.1016/j.ins.2022.05.056
Santarelli, E., Klomp, L., & Thurik, A. R. (2006). Gibrat's law: An overview of the empirical literature. In *Entrepreneurship, growth, and innovation: The dynamics of firms and industries* (pp. 41–73). https://doi.org/10.1007/0-387-3.
Simon, H. A., & Bonini, C. P. (1958). The size distribution of business firms. *American Economic Review*, *58*, 607–617. https://www.jstor.org/stable/1808270. Accessed 22 Mar 2024.

Steindl, J. (1965). *Random processes and the growth of firms: A study of the Pareto Law*. Griffin.
Sutton, J. (1997). Gibrat's Legacy. *Journal of Economic Literature, 35*, 40–59. https://doi.org/10.1023/A:1013309928700
Tang, A. (2015). Does Gibrat's Law hold for Swedish energy firms? *Empirical Economics, 49*, 659–674. https://doi.org/10.1007/s00181-014-0883-x
Zhang, M., & Mohnen, P. (2022). R&D, innovation and firm survival in Chinese manufacturing, 2000–2006. *Eurasian Business Review, 12*(1), 59–95. https://doi.org/10.1007/s40821-021-00200-1