**QUADERNI DEL DIPARTIMENTO DI SCIENZE ECONOMICHE E SOCIALI**

# Firm size distributions and stochastic growth models: a comparison between ICT and Mechanical Italian Companies

*Lisa Crosato – Piero Ganugi – Luigi Grossi*

*Serie Blu: Metodi quantitativi e Informatica – Quaderno N.4 Marzo 2003*

**UNIVERSITÀ CATTOLICA DEL SACRO CUORE**

**PIACENZA**

I QUADERNI
Possono essere richiesti a:

Dipartimento di Scienze Economiche e Sociali, Università Cattolica, Via Emilia Parmense 84, 29100 Piacenza tel. 0523 599342. Oppure si può ottenere una copia dall'area di download del Dipartimento al seguente indirizzo: **http://www.unicatt.it/dipartimenti/ScEcoSoc/default.stm**

# Firm size distributions and stochastic growth models: a comparison between ICT and Mechanical Italian Companies.

Crosato L.*, Ganugi P.*, Grossi L.**
*Università Cattolica del Sacro Cuore
Dip. di Scienze Economiche e Sociali
Piacenza, via Emilia Parmense 84
lisa.crosato@unicatt.it, piero.ganugi@unicatt.it
**Università di Parma.
Dipartimento di Economia
Parma,Via Kennedy, 6
luigi.grossi@unipr.it

**Abstract:** In this paper we analyze the relationship between firm size distribution and stochastic processes. Three main models have been suggested by Gibrat, Kalecki and Champernowne. The first two lead to lognormal distribution and the last to Pareto distribution. We fitted lognormal and Pareto distribution to two Italian sectors: ICT and mechanical. For ICT we found that lognormal distribution must be rejected and Pareto fit reasonably well to the last 30% of largest companies. For mechanical we can not reject lognormal distribution. Furthermore, we perform some experiments to corroborate the theoretical models. By means of similarity analysis and transition matrices we found that Kalecki's model strongly fit to mechanical, while ICT shows features very close to the remaining two models.

**Key words:** Gibrat law, Pareto distribution, lognormal distribution, concentration, firm size.

**Jel:** L00, L25,D21.

## 1 Introduction.

The analysis of firm size distribution in an industrial sector answers a double need.

The first more immediate one, is to supply an analytical description of a structural feature of the sector. Scale and therefore technology, market structure, perspective of growth and decline are only a few of the aspects strictly related to the problem of firm size. On the other hand means, variability and concentration indexes are no doubt important but far from exhaustive tools for the study of size.

The second purpose for which the study of distribution is particularly suitable, concerns the necessity to fit the process of growth of the sector.

There is, in fact, a specific relation between peculiar stochastic models of firm growth and statistical models of firm size.

The problems focused by this family of stochastic models are essentially two:

- the existence and persistence of stochastic factors affecting the operating of firms;
- the possibility of a process of inequality and eventually of concentration among firms.

Pitfalls of concentration indexes and correspondence between true distribution of the sector and related growth model explain the considerable work devoted to the topic of true distribution in the second post-war period starting from the contributions of Hart and Prais (1956), to the now classic book by Ijiri and Simon (1977), the paper by Barca (1985) and the recent contribution by Hart and Oulton (1999 and 2001).

As for the possibility of moulding firm size with known statistical distributions the mainstream of applied economics can be summarized by three different phases:

'50- '70: a strong conviction (Hart and Prais, 1956; Steindl, 1965; Quandt, 1966; Ijiri and Simon, 1977);

'80-'90: growing skepticism (Schmalensee, 1992; Sutton,1998);

nowadays: renewed optimism in the light of the improvement in computational statistics (Dosi, Riccaboni, Varaldo 2001,Cipollini and Ganugi 2000; Axtell, 2001; Marsili, 2001; Marsili and Salter, 2002).

The present paper is divided into three parts:

- in the first we present three classical stochastic models: Gibrat, Kalecki and Champernowne;
- in the second part of the paper: a) we study the fitting of Lognormal and Pareto distributions in ICT and mechanical industries; b) through an analysis of residuals we test the validity of Gibrat and Kalecki's models for the two industries;
- in the third we develop two experiments: a) a similarity analysis of an index of growth in the partitioned - by means of percentiles - data sets; b) the computation of Champernowne's transition matrices. Through these experiments we want both to corroborate the previous results and to evaluate the holding of Champernowne's model for the two industries.

## 2 The data

Our paper is based on two important fix panels represented by the universe of Mechanical and ICT companies in Italy for the years 1997-1998-1999. Data are constituted by Companies Accounts of the same firms, respectively of 12296 for ICT

(Table 1 reports the branches we have selected to identify the Italian ICT) and 11250 for Mechanics (branch Dk29, construction of machines and mechanical device).
To exclude self-employment companies and not proper firms, we have removed from our panel companies with personnel expenditure equal to zero in the considered interval. After this trimming our data set is represented by 9822 units for Mechanical and 7887 for ICT. It is worth noting that reducing of dataset is far stronger in ICT than in Mechanical.

**Table 1 Branches which, in this paper, identify New Economy, making resort to the ATECO91[1] classification**

| ATECO 91 Code | Branches |
|---|---|
| I064200 | Telecommunication; |
| K072000 | Informatics and related business; |
| K072100 | Computer installation and consulting; |
| K072200 | Software supplying and computer consulting; |
| K072300 | Electronic data processing; |
| K072400 | Database activities; |
| K072500 | Computer maintenance and repair; |
| K072600 | Other business related to informatics; |
| K072601 | Telematics and robotics services, computer graphics; |
| K072602 | Other services related to informatics. |

The choice of working on fix panels must not be considered reductive for at least two reasons:

- focusing the analysis on fix panel makes it possible to analyse the core of the sector represented by companies active at least for a short time interval: "despite the large number of births and deaths, a preponderating influence is had by the changes in the sizes of those firms that are alive throughout the period" (Hart and Prais 1956 p.168);
- the three stochastic models presented in this paper are based on a constant number of firms. The introduction of demography makes it necessary to refer to the model of Simon (1955,1960) which deserves an autonomous paper we are going to develop.

As highlighted in an old contribution by Boeri (1989) the variables most often used to mould firm size distribution are assets, sales, equity, employees.
A comparative work on the distribution of the same group of firms according to different variables would be of great interest. In particular it would be really useful to compare the firm size distribution according to budget variables - after all present in the SCI as well - together with the number of employees or personnel.
The variables we have used are sales and assets. Nevertheless, since sales present remarkable problems of time-stability we have devoted more attention to assets.

[1] The ATECO91 classification is used by Istat, the Italian Statistical Institute and is closely derived from NACE Rev.1 Classification.

## PART I: THREE STOCHASTIC MODELS

### 3 Gibrat: lognormality, increasing variance and increasing concentration

The first process we will describe is the one by Gibrat, who started the study of the link between stochastic processes, statistical distributions and laws of economics.

Let $X_j$, the firm size, be a random variable, the result of the combined action of different causes, independent one each other, which continue to act in time.

Let the initial $j$ firm size be equal to $X_{j0}$ and after the $t$-th step of the process let it be equal to $X_{jt}$, reaching its final value $X_{jn}$ after $n$ steps.

We shall say a variable subject to a variation process satisfies the Gibrat law if at step t of the process the variation in the variable is a random proportion of the value attained by the variable at step t-1.

In this case we have that

$$X_{jt} - X_{jt-1} = \varepsilon_{jt} X_{jt-1} \quad (1)$$

where the $\varepsilon_{jt}$ are identically and independently distributed (i.i.d.) with mean μ and variance $\sigma^2$; furthermore $\varepsilon_{jt}$ are independent of $X_{jt}$ as well.

Now we can write (1) as

$$\frac{X_{jt} - X_{jt-1}}{X_{jt-1}} = \varepsilon_{jt}, \quad (2)$$

so that, adding through all steps, we obtain

$$\sum_{t=1}^{n} \frac{X_{jt} - X_{jt-1}}{X_{jt-1}} = \sum_{t=1}^{n} \varepsilon_{jt}. \quad (3)$$

Now, supposing that the effects at each step are small, the sum becomes an integral:

$$\sum_{t=1}^{n} \frac{X_{jt} - X_{jt-1}}{X_{jt-1}} \sim \int_{X_{j0}}^{X_{jn}} \frac{dX}{X} = \log X_{jn} - \log X_{j0}, \quad (4)$$

it follows that:

$$\log X_{jn} = \log X_{j0} + \sum_{t=1}^{n} \varepsilon_{jt}. \quad (5)$$

Since log $X_{jn}$ is the sum of i.i.d. variables with constant mean and variance, it follows, from the central limit theorem, that it is asymptotically distributed as a normal and therefore $X_{jn}$ is asymptotically lognormal.

It is also clear, referring to (5) that, increasing the number of the process steps, the logarithmic variance of the variable increases with a consequent enlargement of concentration.

In their classical work on Gibrat's law Hart and Prais (1956) analyze concentration of a growth process by means of the logarithmic variance of the size. They stress that variance-based concentration measures are very effective and particularly suitable to highlight, in the distribution structure itself, the very causes of (possible) changes in concentration.
It can besides be shown (Aitchison and Brown, 1957) that the Lorenz concentration measure is monotonically dependent on the logarithmic variance value of the considered variable, and consequently the mean-difference coefficient of Gini behaves likewise.

## 4 Gibrat's law and Italian economy

It is now necessary to remark two important aspects of Gibrat's model:

1) the growth of firms is a stochastic process i.e. it is not fostered by a particular scale of the same productive units because:
   a. in the different markets of money, commodities and labour, the firms of a given industrial sector have the same chances independently of their scale;
   b. given the nature of technology faced by the firms, the sector has no optimal scale by which unit costs are minimized and profit maximized;
2) given the same opportunities of growth in firms of different sizes, growth of firms is proportionate to their size and a concentration process is unavoidable.

The relation between the family of stochastic models and the neoclassical theory of production is hinted by Quandt (1963) and Simon and Bonini (1958). While Quandt advocates the necessity to reconcile the two approaches, Simon and Bonini are much more cautious. An empirical research comparing production theory and stochastic process of growth might contribute to explore the distance between two theoretical frameworks on the problem of firm growth.
The Italian economic history of '70s,'80s, and '90s does not apparently confirm Gibrat law given the structural impediments to growth of large firms, hence the much faster growth rates of small and medium size productive units. Econometric and applied work produced on this topic for Italy has corroborated this result (Brusco et al., 1979, Solinas, 1996, Audretsch, Santarelli e Vivarelli, 1999). Recently however, Piergiovanni et al., 2002, showed that Gibrat law holds for some business groups in Italian Hospitality services.
In spite of this, unanimously refusing Gibrat law, also with reference to Italy, some authors have been careful to separate the first aspect of the law - the stochastic nature of growth process - from the second - the tendency to concentration or at least to higher inequality (Barca, 1985). On the basis of a good fitting of a Pareto distribution on Italian firms by two censuses, Barca (1985) concludes that "la forma della distribuzione dimensionale degli stabilimenti è influenzata molto fortemente dall'operare nel tempo, in modo cumulato, di fattori stocastici".

## 5 Kalecki: lognormality, constant variance and non increasing concentration

We have seen that the Gibrat law involves lognormal size and increasing logarithmic variance.
Kalecki (1945) formulates a stochastic model of growth in which it is assumed that the logarithmic variance of size is constant: this implies a negative correlation between the

logarithm of size and the logarithm of random variables describing the variation, that is the $\varepsilon_{jt}$.
Let $\mu_{X,t-1}$ be the arithmetic mean of the distribution of the log-size at step *t-1*, and $\mu_{\varepsilon,t}$ the arithmetic mean of the distribution of the log-increments occurred from step *t-1* to step *t*.
Suppose that the variance is constant before and after the occurred increment.
The arithmetic mean of $\log X_{jt-1}+\log(\varepsilon_{jt}+1)$ is then given by $\mu_{X,t-1} + \mu_{\varepsilon,t}$ and the hypothesis of constant variance is formally given by the following equation:

$$\frac{1}{n}\sum_j[\log X_{jt-1} + \log(\varepsilon_{jt} + 1) - (\mu_{X,t-1} + \mu_{\varepsilon,t})]^2 = \frac{1}{n}\sum_j(\log X_{jt-1} - \mu_{X,t-1})^2 \ , \qquad (6)$$

Putting

$Y_{jt-1} = \log X_{jt-1} - \mu_{X,t-1},$
$y_{jt} = \log(\varepsilon_{jt} + 1) - \mu_{\varepsilon,t},$

and substituting in (6), we get:

$$\frac{1}{n}\sum_j(Y_{jt-1} + y_{jt})^2 = \frac{1}{n}\sum_j Y_{jt-1}{}^2, \qquad (7)$$

whence

$$2\sum_j Y_{jt-1}y_{jt} = -\sum_j y_{jt}{}^2 \ . \qquad (8)$$

Therefore a negative correlation between the random increment of size and the size itself is proven.
The simplest hypothesis concerning the shape of this relation is its linearity, that is:

$$y_t=-\alpha_t Y_{t-1}+z_t \qquad (9)$$

where $\alpha_t$ is a constant parameter and $z_t$ a random variable independent of $Y_{t-1}$. (We have cancelled the *j* index to simplify the notation.)
Substituting (9) in the first member of (8), we obtain

$$\alpha_t = \frac{\sum y_t{}^2}{2\sum Y_{t-1}{}^2} \qquad (10)$$

As the second moment of $y_t$ is smaller than that of $Y_{t-1}$, we have
$0 < 1-\alpha_t < 1$

and rewriting (9)

$$Y_{t-1} + y_t = (1-\alpha_t)Y_{t-1} + z_t \qquad (11)$$

Therefore, at each step, we have

$Y_1 = Y_0 + y_1 = Y_0(1-\alpha_1) + z_1$
$Y_2 = Y_0 + y_1 + y_2 = Y_0(1-\alpha_1)(1-\alpha_2) + z_1(1-\alpha_2) + z_2$
....................
$$Y_n = Y_0 + y_1 + y_2 + \ldots + y_n = Y_0(1-\alpha_1)(1-\alpha_2)\ldots(1-\alpha_n) + z_1(1-\alpha_2)\ldots(1-\alpha_n) + \ldots + z_{n-1}(1-\alpha_n) + z_n. \quad (12)$$

Since $0 < 1-\alpha_k < 1$ for each $k$ between 1 and $n$, if $n \rightarrow \infty$ the absolute value of all components in (12) is small if compared to the variance of $Y_n$ (unless $z_k$ tend to zero at the same time).
We have therefore obtained $Y_n$ as a sum of small random increments $z_k \Pi(1-\alpha_k)$, independent of each other and of the initial value $Y_0$. For the central limit theorem the distribution of the $Y_t$ will therefore be normal.

It is now important to underline two aspects of Kalecki's model:

1) the stochastic nature of the process of growth of firms with the same meaning above specified for Gibrat;
2) the existence of impediments for large firms to grow proportionately to their size, avoiding, in such way, an increase of concentration.

## 6 Champernowne: Paretian size, decreasing concentration

We are now describing a stochastic process leading to the Pareto distribution, originally proposed by Champernowne in 1937 regarding income distribution, and referred to by Steindl (1965) and Simon (1955). The application of Champernowne's model of income distributions to the asset distribution of firms was suggested by Quandt (1966, p. 418).
Suppose that firm size develops over time according to a Markovian process, in which the state of the process is the annual firm size: each state is completely determined by its previous state and a random element.
The possibility to pass from state $r$ to state $s$ is called transition probability and is generally indicated by $p_{rs}$. All possible transitions from one year to the next are stored in a matrix, called transition matrix, whose elements $p_{rs}$ must satisfy the following properties:

i) $p_{rs} \geq 0$;
ii) $\sum_s p_{rs} = 1$.

In this case the matrix rows will be the size classes in a certain year, the columns the size classes of the next year and $p_{rs}$ gives the probability of a shift from size class $r$ to size class $s$ from year to year.
The size intervals are equally spaced on a logarithmic scale, therefore of uniform proportionate extent. It follows that the $j$ state will correspond to a size in the interval ($10^{jh}Y_{min}$, $10^{(j+1)h}Y_{min}$), where $Y_{min}$ is the minimum considered size and $h$ is a positive real number, equal to class width on a logarithmic scale.
The relevance of equally spaced classes on a logarithmic scale is that, for firms being in similar positions in different classes, proportionate variations will imply equal class jumps.

Champernowne imposes to transition probabilities the hypothesis that the probability of a transition from a size range to another does not depend on the position we start from, but only on the jump width.
In other words, the probability $p_{rs}$ is a function of $s$-$r$ only, therefore putting $s-r=u$ we can determine the whole matrix from the knowledge of $p_u$, which is independent of $r$.
As a consequence, the process will develop according to the formula

$$X_s(t+1) = \sum_{u=-\infty}^{s} X_{s-u}(t) p_{s-u,s}(t) \tag{13}$$

In the model, though, it stands out that, when $n$ is fixed, transitions are possible only between $-n$ and 1: in this hypothesis $p_u$ represents the probability a firm has to diminish its size of $u$ levels downwards (for a maximum of $n$ and given that $u>-r$) or increase of 1 level upwards or remain in the same class: $p_1$, for instance, represents the probability to increase of one class upwards, $p_{-1}$ the probability to diminish of one class downwards, $p_0$ the probability to remain in the same class.
It is therefore taken that $p_u=0$ if $u>1$ *o* $u<-n$.
Should the process go on for a sufficiently long period, the size distribution would reach an equilibrium, in which the action of the transition matrix would leave the distribution unchanged: at this point the distribution would take the name of stationary distribution.
We will therefore have

$$X_s = \sum_{u=-n}^{1} p_u X_{s-u} \qquad (s > 0) \tag{14}$$

and

$$X_0 = \sum_{u=-n}^{1} q_u X_{-u} \text{ where } q_u = \sum_{v=-n}^{u} p_v \,. \tag{15}$$

Equation (14) is a finite difference equation whose solution is obtained by putting $X_s=z^s$; thus we get the characteristic equation

$$g(z) = \sum_{u=-n}^{1} p_u z^{1-u} - z = 0 \;. \tag{16}$$

Observe that, since $g(0)=p_1>0$ and $g(1)=0$, for Descartes' rule of signes, the characteristic equation has only two real positive roots one of which is clearly unity.
If we indicate the other one with $b$, the equilibrium distribution required is

$$X_s = b^s \,. \tag{17}$$

To have $b$ change between $0$ and $1$ it is necessary to introduce the following stability condition:

$$g'(1) = -\sum_{u=-n}^{1} u p_u > 0 \tag{18}$$

By summing up both members of (17), $s$ tending to the infinite we get that the total number of firms is equal to 1/(1-b).

It follows that for any given number of firms $N$ the equilibrium distribution becomes

$$X_s = N(1-b)b^s . \quad (19)$$

We have seen that the proportionate width of size classes is $10^h$ and the minimum size considered is $Y_{min}$; $X_s$ is thus the number of firms in size class $s$ whose lower bound is given by

$$Y_s=10^{sh}Y_{min} \quad (20)$$

and passing to logarithms

$$logY_s=sh+logY_{min}. \quad (21)$$

The number of firms with size exceeding $Y_s$ (i.e. the cumulative right distribution function) is obtained by the geometrical progression of (19)

$$F(Y_s)=Nb^s. \quad (22)$$

Going back to logarithms again

$$log\ F(Y_s)=logN+slogb=\log N + (\frac{\log Y_s - \log Y_{\min}}{h})\log b\ , \quad (23)$$

whence, putting

$$\alpha = -\frac{1}{h}\log b, \gamma = \log N + \alpha \log Y_{\min} \quad (24)$$

we get at last the Pareto law

$$logF(Y_s)=\gamma - \alpha logY_s. \quad (25)$$

## PART II: THE FITTING OF THE DISTRIBUTIONS

Before presenting our empirical analysis, with reference to the relationship between process of growth and concentration, it may be worth underlining:

- in each of the three models described in part I every firm faces a proper rate of growth which is a random variate;
- in Gibrat's model the rate of growth is independent of size with a consequent *increase* of variance and concentration;
- in Kalecki's model the process of growth is non-dissipative because concentration does not increase, given negative correlation between size and growth;
- Champernowne's model has features which are common to previous models:
  a) the rate of growth is independent of size like in Gibrat's;

b) the process of "growth" remains non-dissipative as in Kalecki but in a much more restrictive sense: the expected value of variation is negative for all firms with a consequent decrease of concentration (stability condition).

## 7 ICT: fitting the Pareto

### 7.1 Size-Rank

Considering sales and total assets of ICT companies we found that hypothesis of lognormality must be rejected, both using statistical tests and graphical representations.
According to Ijiri and Simon (1964,1977), but also to Steindl (1965) an alternative method to corroborate the law of proportionate effects is to measure the fitness of a Pareto curve to the largest companies of the distribution.
In this work we estimate the Pareto cumulative distribution of firm size starting from the relation between firm size and rank in the ordered data set. The technique, known as Zipf plot, is a plot of the log of the rank vs. the log of the variable being analyzed. Let $(x_1, x_2, ..., x_N)$ be a set of $N$ companies on a random variable $X$ for which the cumulative distribution function is $F(x)$ and suppose that the observations are ordered from largest to smallest so that the index $i$ is the rank of the $i$-th company. The Zipf plot of the sample is the graph of $\ln(x_i)$ against $\ln(i)$. Because of the ranking, $i/N = 1 - F(x_i)$, so $\ln(i) = \ln[1 - F(x_i)] + \ln(N)$.
Thus, the log of the rank is simply a transformation of the cumulative distribution function. It accentuates the upper tail of the distribution and therefore makes it easier to detect deviations in the upper tail from the theoretical prediction of a particular distribution. Since there has been interest in the upper tail of the size distribution of firms, the Zipf plot is particularly useful for analyzing this question.
The Pareto's curve can be formulated as (see Johnson and Kotz, 1972, pp. 573-627):

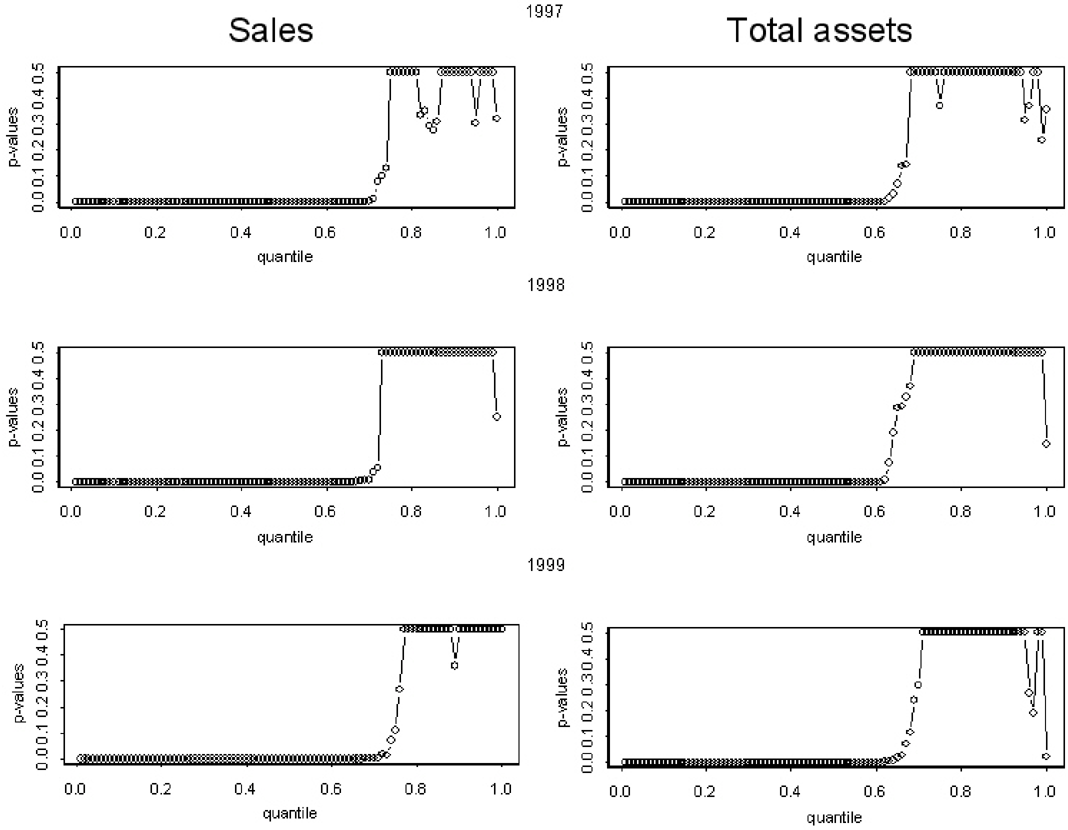
$$S_i = M\, i^{-\beta} \qquad \beta > 0 \quad i = 1,2,...,N \tag{26}$$

where $S_i$ is the size of the $i$-th firm (Sales or Total Assets), $i$ the rank of the $i$-th firm, while $\beta$ and $M$ are parameters. This relation can be estimated taking logs:

$$\log S_i = \log M - \beta \log(i) \qquad i = 1,2,...,N\ . \tag{27}$$

which is a straight line in the Zipf plot.
The parameter $\beta$, the slope of the log-line, is a measure of concentration. The larger the $\beta$, the greater the relative size of a large firm (small rank) compared with a smaller firm (large rank) (see Ijiri and Simon, 1977, p. 196).
Pareto's curve has been estimated selecting companies with size over the 70-*th* percentile. The choice of this threshold is justified by Figure 1. Plots report p-value of Kolmogorov-Smirnov test for Pareto distribution for the portion of companies with dimension over a given quantile. As it can be seen, *p*-value is very close to zero until 70-*th* quantile and the hypothesis of Pareto distribution can not be rejected considering the largest 30% of companies both for sales and total assets.

**Figure 1. P-values of Kolmogorov test for Pareto distribution with respect to different sample size. ICT, Sales 1999.**




It is useful to analyze the estimates of $\beta$ for different years to highlight possible changes of concentration during time. Estimates of $\beta$ (obtained by OLS) in (3) are reported in Table 2 for years 1997-99.

**Table 2. Concentration indexes ($\beta$) computed by log(size)-log(rank) regression on the last 30% of companies in ICT**

| | 1997 | | 1998 | | 1999 | |
|---|---|---|---|---|---|---|
| | Sales | Total assets | Sales | Total assets | Sales | Total assets |
| LogM | 21.539 | 21.835 | 21.886 | 22.100 | 22.157 | 22.314 |
| $\beta$ | 1.048 | 1.121 | 1.071 | 1.134 | 1.090 | 1.144 |

It is interesting to remark that concentration is higher for Total Asset than for Sales.

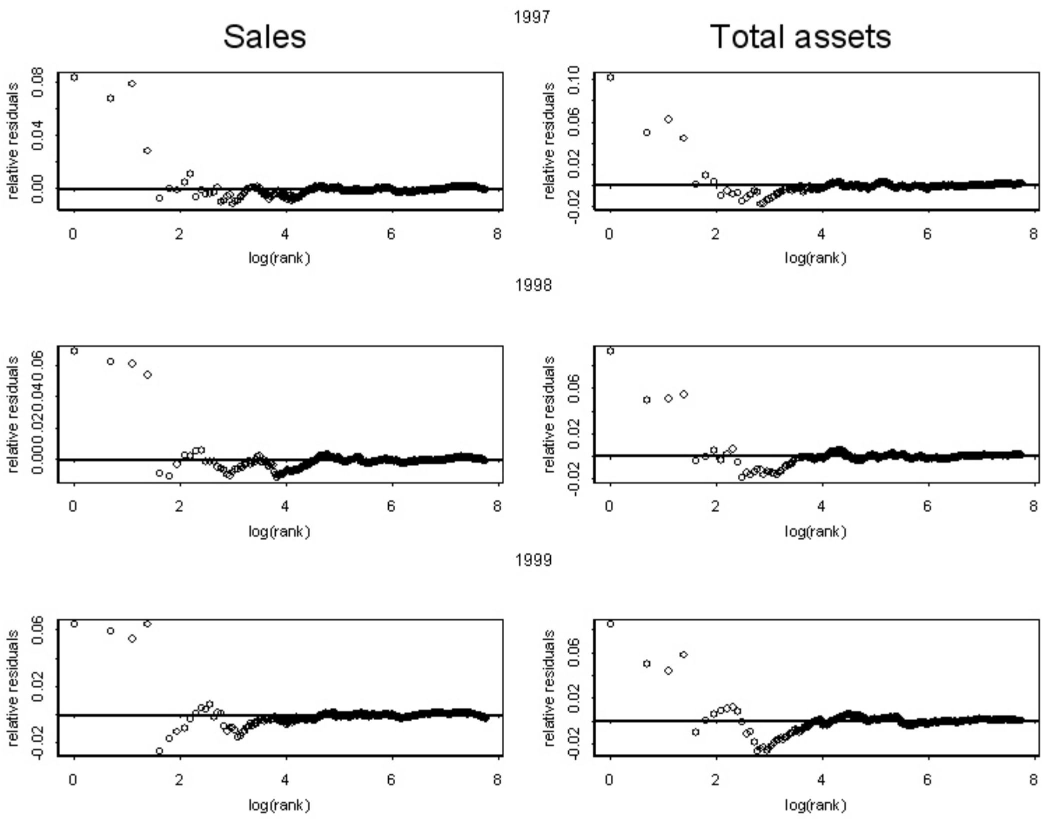
## 7.2 Residuals on Pareto: the absence of impediments for large companies

As emphasized by Ijiri and Simon (1964) the fitting of Pareto can reveal the presence of a systematic difference between effective and fitted values in the extreme tail of the distribution.

From an economic point of view the same discrepancy is extremely meaningful:

- in the case of negative difference between effective and fitted values in the tail of distribution, we have that, for a given value of the rank, the corresponding effective size rank is lower than that estimated. The economic consequence on the process of growth is important: large firms - which are in the tail of the distribution and then with very low rank - meet impediments to their growth. Thus, we deduce the stochastic nature of the growth process, but also the existence of impediments to growth for companies of this portion of the distribution. According to Barca (1985) this is the situation of the Italian Industry resulting from the Censuses of 1971 and 1981.
- In the opposite case - positive residuals in the tail of the distribution - we conclude that large size firms do not find particular impediments to their growth. Given the stochastic nature of the growth process, if repeated in time, the absence of impediments involves a process of concentration.

**Figure 2. Relative residuals of log(size)-log(rank) model for sales and total assets of ICT (1997-1999)**




From our estimates we have positive residuals for the largest four companies (having the lowest ranks) in each year (see Figure 2).

Table 3 supplies a clear picture of the importance of these 4 companies.

Our estimates provide evidence of the existence of favourable conditions to concentration for this sector. This does not imply that ICT is consistent with Gibrat's law. As we showed in the first part, Gibrat's law implies lognormality of firm size which cannot be proved for ICT.

**Table 3. Amounts of sales, total assets and employment costs of four greatest companies in ICT (1997-1999)**

| | 1997 | | 1998 | | 1999 | |
|---|---|---|---|---|---|---|
| | Percentage | Euro (mln) | Percentage | Euro (mln) | Percentage | Euro (mln) |
| Total assets | 79.39% | 45004.10 | 77.72% | 49683.88 | 79.10% | 54913.62 |
| Sales | 67.50% | 25704.89 | 64.66% | 28411.89 | 65.50% | 34316.93 |
| Personnel costs | 56.21% | 4405.24 | 52.24% | 4300.49 | 50.95% | 4530.49 |

# 8 Mechanical: Lognormality and Pareto

## 8.1 Lognormality: a bootstrap procedure

To test the lognormality hypothesis of firm size (sales and total assets) we have firstly applied two well-known tests: Kolmogorov-Smirnov and chi-squared. In all cases the tests reject the hypothesis of lognormality. Nevertheless, it is well known that statistical tests are strongly influenced by the sample size and that results of chi-squared test depend on the choice of the number of classes. For this reason, we have performed a bootstrap procedure which is described as follows.

Given a sample of $n$ firms, we randomly choose, by reintroduction, $k$ samples of size $m$, $2m$, …, $sm$, $n$, with $m<n$, where $s = \text{integer}(n/m)$. For each sample of size $m$ we compute the $p$-value of Kolmogorov-Smirnov and chi-squared test for lognormality of size variable (total assets, sales, employees). $P$-values are indexed by $pv_j$, $j$=1, 2, …, $k$. For each size $mr$, $r = 1,2,...,s$ we compute the median of $p$-values indexed by $\text{Me}(pv_r)$. We performed the experiment for total assets and sales of mechanical in 1997, 1998, 1999. As an example we report the result of the experiment for total assets in 1999 with $k$=100 (see Table 4). As can be seen, lognormality cannot be rejected at least until $m$=800 at 5%.

**Table 4. Median of p-values of Kolmogorov-Smirnov (KS) test and Chi-Squared (CS) test from bootstrap procedure on Total Assets of mechanical in 1999. *k*=100**

| | Sample size (*m*) | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 100 | 200 | 300 | 400 | 500 | 600 | 700 | 800 | 900 | 1000 |
| median p-value of KS test | 0.294 | 0.198 | 0.137 | 0.114 | 0.105 | 0.103 | 0.093 | 0.055 | 0.022 | 0.020 |
| median p-value of CS test | 0.509 | 0.286 | 0.225 | 0.181 | 0.132 | 0.114 | 0.107 | 0.073 | 0.053 | 0.039 |

Furthermore, we found that dividing the data with respect to geographical macro-regions (North-western, North-Eastern, Central and Southern Italy) the hypothesis of lognormality cannot be rejected for the macro-region with the lowest number of firms (Grossi, Ganugi, Gozzi, 2003).

Because, as it is well-known and as it has been proved by the bootstrap experiment, the greater the sample size the greater is the probability to reject the null hypothesis of lognormality, we chose to adopt alternative descriptive tools which are independent of

sample size. We then compare histograms and kernel densities with the density of a normal distribution with the same parameters. Results are reported for total assets in Figure 3.
As can be seen, the empirical distribution of log-size is very near to normal with same discrepancies in the central classes. Thus, we can say that lognormality can be accepted apart from slight differences of central values, while tails – very important for economic models of growth – are consistent with lognormality.

## 8.2 Pareto

The fitting of rank-size regression for mechanical is much less easy than in ICT. Looking at plots of *p*-values of Kolmogorov-Smirnov test for Pareto (not reported for lack of space, but similar to Figure 1) *p*-values become different from zeros only after the 90*th* percentile. For this, we fit Pareto through model (3) on the largest 10% of total companies. This narrower percentage covers however a relevant stock of 1000 companies, the largest firms of the sector.
The fitting of the Pareto confirms again the stochastic nature of the growth process.

**Figure 3. Histograms and kernel densities of log-transformed data compared with normal density. Mechanics, total assets 1997-1999**

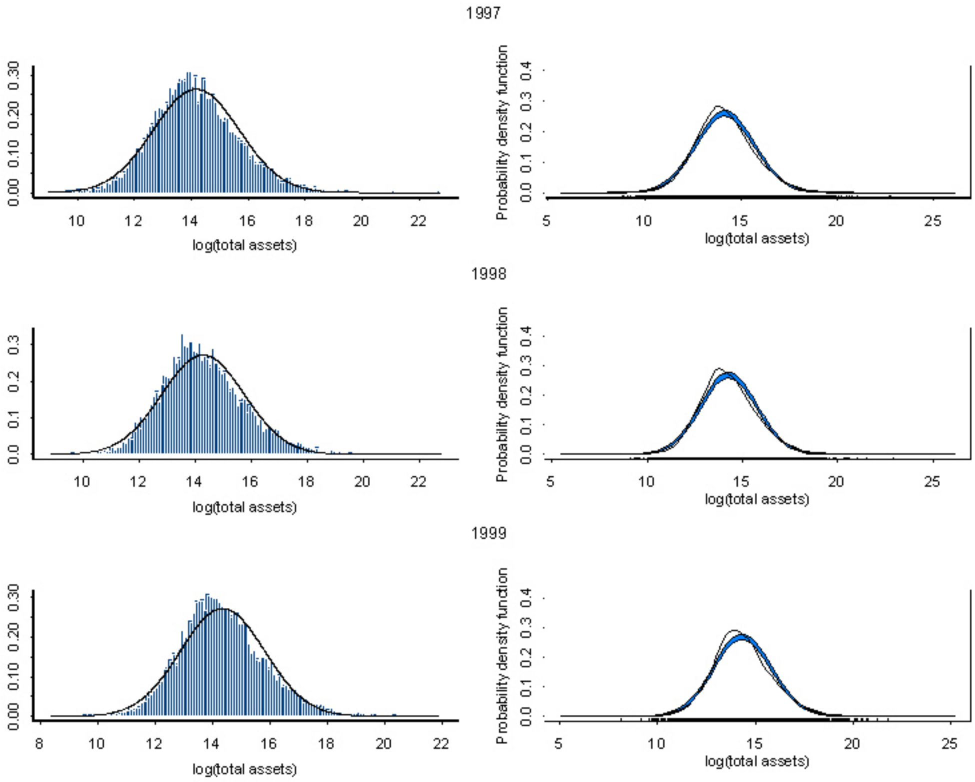
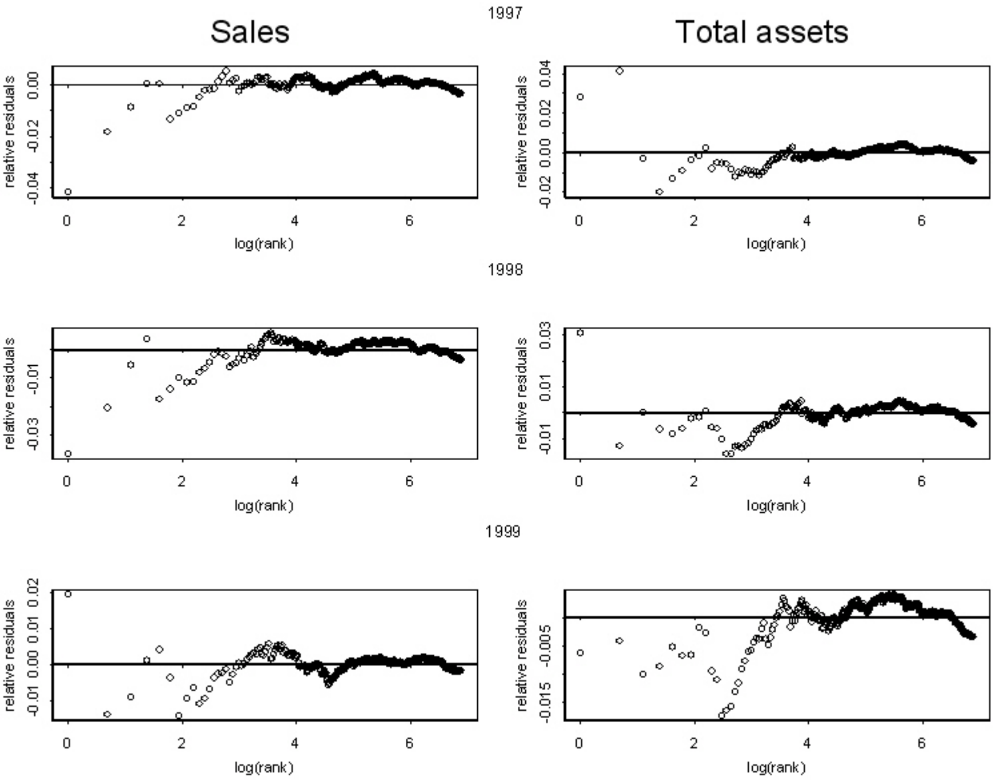


**Figure 4. Relative residuals of log(size)-log(rank) model for sales and total assets of mechanichs (1997-1999)**




## 8.3 Residuals on Pareto: the presence of impediments for large companies

The observation of residuals in the tail of the distribution reveals the existence of a structure opposite to ICT for Mechanical: for a given rank the effective size is lower than that shown by the model and so large firms meet serious impediment to their growth. It is interesting to note (see Figure 4) that, differently form ICT, in the tail of the distribution we find a lot of large relative residuals which are now almost all negative (compare Figure 4 with Figure 1).

## THIRD PART: TWO EXPERIMENTS

We want now to corroborate the previous analysis by means of two experiments:

a) the analysis of the distributions of the index of growth, *size(t)/size(t-1),* in the 8 portions of our data set (of each industry) defined by 7 equally spaced percentiles (12.5th, 25th, …, 87,5th) with respect to total assets and sales;
b) the construction of Champernowne-based transition matrices whose classes are formed, as previously explained, keeping the proportionate extent uniform.

## 9 The distributions of the index of growth

The choice of studying the distribution of the growth index arises from the necessity to supply a satisfactory measurement of the process of growth in the 8 portions of the distribution defined by the 7 percentiles.

Relevance of studying the mean of the distribution of growth index is immediate: the holding of Gibrat’s law requires the presence of, at least, equal average rates of growth in each portion of firms, regardless of their size. Conversely, the rejection of the law is corroborated by decreasing rates of growth.
The analysis of variability of the distribution of growth index is important as well.
In the works of the ‘60s and the ‘70s the necessity to study not only the mean but also the variance of growth is clearly underlined (Hart and Prais ,1956; Simon and Bonini, 1958; Mansfield, 1962; Singh and Whittinghton, 1975; Brusco, 1979).
Hart and Prais have then stressed the absence of an impact of size on the variability of growth. On the contrary, Mansfield, Singh and Whittinghton have found a decreasing path of variability with respect to size.
It is then evident that equal variability of the growth index in the different classes implies a lesser degree of stability in size: large firms have the same chances as small ones to jump to higher size classes, as well as to more modest size.
From the point of view of Gibrat’s law, it is remarkable that, with respect to their size, large firms are not more stable than small firms. According to Scherer (1980), statistical evidence points out that variability is more curbed in large firms than in small ones. Brusco reaches the same conclusion for several Emilian industries of the ‘70s. It is a credit of the above quoted works to focus attention on the necessity to study distribution of an index (rate) of growth.
In this work we aim to develop a refinement of this approach by three steps:

- we produce *robust* measures of the same distributions. For all the ‘70s and ‘80s the problem of robustness has yet not been well grounded. To this aim we substitute mean with median, standard error with MAD which is the median of the differences from median;
- we supply kernel estimates of the 8 distributions: as it is well-known, equal synthetic measures of the distributions are not sufficient to guarantee that the same distributions are similar or overlapping. The inspection of the kernel shapes produces further evidence on this aspect;
- we calculate an index of similarity between each distribution and all the other 7 combining the results in a matrix of similarity. Similarity between two distributions can be measured through different approaches. The affinity approach is well known and rich in intuitive appeal (Krzanowski 2000). In this work we prefer to use an index of Gini’s which has the great advantage of not requiring the knowledge of distributions. Given a variable $X$ observed in two populations $A$ and $B$, we can measure the dissimilarity of the corresponding cumulative distribution functions ($F_A(X)$ and $F_B(X)$) by the Gini’s index $G$, which is computed as follows (see Leti, 1983, pp. 534-550):

$$G = \sum_{i=1}^{N-1} \left(x_{i+1} - x_i\right)\left|F_{Ai} - F_{Bi}\right| \qquad (28)$$

  where $x_i$ is the $i$-th value of $X$, in a non decreasing order and N is obtained by summing the size of $A$ and $B$. Gini’s index implies that population $A$ and $B$ have the same size. This assumption is satisfied giving, in each distribution, null frequency when value $x_i$ is not present.
  It can be proved that $G$=0 when CDF in $A$ and $B$ are perfectly equal, and $G$=max($X$)-min($X$) in the case of maximum dissimilarity, that is when all units in one population

assume the minimum value and all units in the other population assume the maximum value. Thus, the relative Gini's index $G'$ is given by $G' = G / [\max(X) - \min(X)]$ which goes from 0 (maximum similarity) to 1 (maximum dissimilarity).

## 9.1 Medians of growth indexes

From Table 6, we note that 1998 medians of the first and also of the second groups (formed by smallest firms) of our empirical distribution of ICT - calculated on Assets - are higher than the other six medians. But starting from the third group the median reveals a strong stability with an increase in the last one.
In 1998-99 medians of the first and also second group of ICT continue to be higher than the other 6 portions but the differences with the same are now much more modest.
The importance of these results is much clearer if we observe the values of the percentiles which represent the limits of each group (see Table 5).

**Table 5. Percentiles of total assets of each industry in 1997**

| | I | II | III | IV | V | VI | VII |
|---|---|---|---|---|---|---|---|
| Mechanical | 2.78E+05 | 5.15E+05 | 8.35E+05 | 1.28E+06 | 2.02E+06 | 3.51E+06 | 7.95E+06 |
| ICT | 6.45E+04 | 1.10E+05 | 1.71E+05 | 2.54E+05 | 3.85E+05 | 6.18E+05 | 1.34E+06 |

**Table 6. Median and Mean Absolute Deviation for indexes of growth of Sales and Total assets of ICT divided in eight groups using percentiles of Sales and Total assets in 1997**

| | Sales | | | | | | Total Assets | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1998-97 | | | 1999-98 | | | 1998-97 | | | 1999-98 | | |
| Groups | median | MAD | MAD/ median | median | MAD | MAD/ median | median | MAD | MAD/ median | median | MAD | MAD/ median |
| I | 117.67 | 38.87 | 0.33 | 109.63 | 28.04 | 0.26 | 120.83 | 35.13 | 0.29 | 111.46 | 30.14 | 0.27 |
| II | 108.00 | 22.11 | 0.20 | 105.90 | 19.12 | 0.18 | 113.80 | 26.81 | 0.24 | 110.24 | 25.96 | 0.24 |
| III | 108.07 | 19.52 | 0.18 | 105.68 | 17.29 | 0.16 | 111.77 | 26.94 | 0.24 | 107.32 | 25.38 | 0.24 |
| IV | 107.58 | 19.65 | 0.18 | 107.01 | 19.92 | 0.19 | 110.83 | 24.54 | 0.22 | 108.58 | 23.94 | 0.22 |
| V | 107.36 | 20.64 | 0.19 | 107.54 | 21.03 | 0.20 | 109.40 | 22.11 | 0.20 | 106.41 | 22.77 | 0.21 |
| VI | 108.78 | 22.20 | 0.20 | 108.86 | 20.88 | 0.19 | 109.98 | 23.56 | 0.21 | 108.46 | 23.31 | 0.21 |
| VII | 111.14 | 23.44 | 0.21 | 108.89 | 22.46 | 0.21 | 108.58 | 23.01 | 0.21 | 106.98 | 22.46 | 0.21 |
| VIII | 113.27 | 25.52 | 0.23 | 112.22 | 24.67 | 0.22 | 110.67 | 23.70 | 0.21 | 108.33 | 21.54 | 0.20 |

**Table 7. Median and Mean Absolute Deviation for indexes of growth of Sales and Total assets of mechanical divided in eight groups using percentiles of Sales and Total assets in 1997**

| | Sales | | | | | | Total Assets | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1998-97 | | | 1999-98 | | | 1998-97 | | | 1999-98 | | |
| Groups | median | MAD | MAD/ median | median | MAD | MAD/ median | median | MAD | MAD/ median | median | MAD | MAD/ median |
| I | 130.42 | 60.1 | 0.4612 | 106.04 | 35.8 | 0.3376 | 121.51 | 39.1 | 0.322 | 111.33 | 31.3 | 0.2815 |
| II | 112.75 | 27.8 | 0.2466 | 106.18 | 25.3 | 0.2382 | 112.02 | 27.2 | 0.2431 | 107.52 | 24.9 | 0.232 |
| III | 110.79 | 24.5 | 0.2215 | 104.71 | 23.5 | 0.2245 | 107.3 | 19.3 | 0.1801 | 106.62 | 20.5 | 0.1921 |
| IV | 108.51 | 23.2 | 0.2138 | 103.31 | 20.2 | 0.1958 | 106.88 | 19.9 | 0.1862 | 104.77 | 19.9 | 0.1895 |
| V | 108.48 | 20.8 | 0.1915 | 102.89 | 19.4 | 0.1882 | 107.05 | 18.9 | 0.1769 | 104.7 | 19.3 | 0.1842 |
| VI | 107.91 | 18.5 | 0.1715 | 102.91 | 17.7 | 0.1719 | 106.99 | 16.6 | 0.155 | 103.62 | 15.7 | 0.1515 |
| VII | 107.3 | 17.1 | 0.1589 | 103.32 | 17.4 | 0.1684 | 106.93 | 15.4 | 0.1441 | 105.33 | 15.6 | 0.1484 |
| VIII | 105.38 | 15.6 | 0.1484 | 101.84 | 15.6 | 0.153 | 104.81 | 14.2 | 0.1353 | 104.25 | 14.8 | 0.1424 |

Because all the companies we have considered in this work have employees, also the units included in the first group are proper firms. However, the first group includes companies whose size is very modest. For this class of companies, it is evident that even very small investments involve a substantial increase of growth rate. Furthermore, literature has showed that many companies have the urgent need to grow in order to reach MES (Minimum Efficient Size; Audretsch *et al.*, 1999). However, from the third class onwards, the medians of the index are stable: for instance, companies whose size is included between 100.000 and 170.000 euros and companies whose size is over 1.345.000 euros present the same median (the last group has a median greater than the previous one).
In Mechanical medians are much higher in the first two classes and decrease slightly in the next ones (see Table 7).

### 9.2 MAD

It is on MAD that the two industrial sectors differ broadly (see Table 6 and Table 7).
From the first to the last class its size decreases of 1/3 in ICT for both years while it is curbed of 1/2 in 1999 and more than 1/2 in 1998 for mechanical.
If instead of an absolute measure of variability we consider a relative one - the MAD divided by median - on the same classes we get the same result.

### 9.3 Kernel distributions of the index of growth

Even if robust, median, MAD, MAD/median represent a synthesis of the distributions of our growth index.
By means of the reconstruction of the entire distributions of our index of growth in each class we are allowed to focus analytically the different patterns of growth associated with different firm sizes.
For this aim we use a kernel approach (Figure 5 and Figure 6).

**Figure 5. Distribution of growth rates (1998-99) for total assets of ICT. Groups are detected by quantiles of total assets in 1997.**

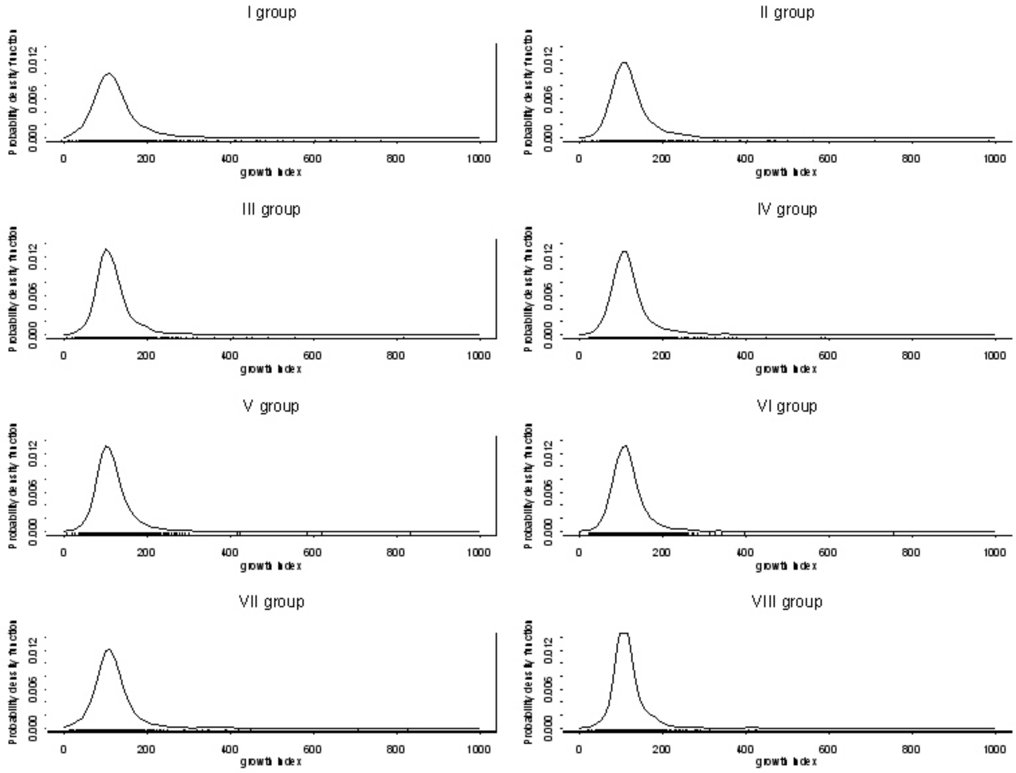
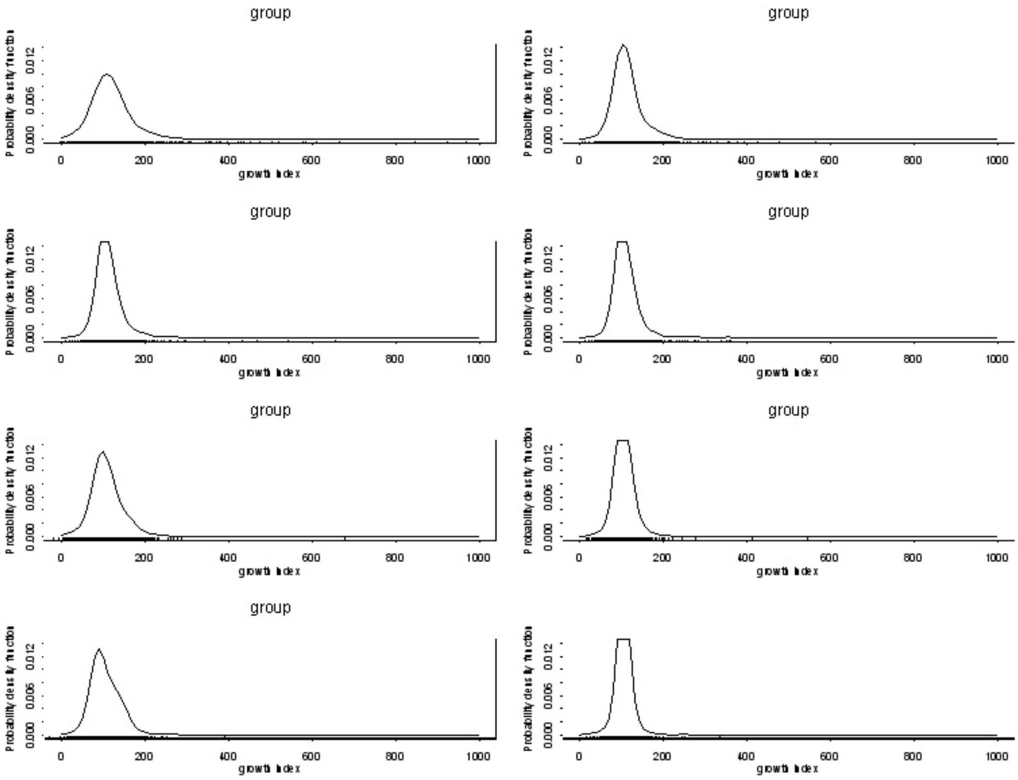



**Figure 6. Distribution of growth rates (1998-99) for total assets of mechanical. Groups are detected by quantiles of total assets in 1997.**



It is now evident that with the expected result of the first class, the distributions are almost the same in each class of ICT.
On the contrary Mechanical presents progressively more narrow distributions.
Referring to the classical quotations we have made above, we can then conclude that ICT is far closer to the results obtained by Hart and Prais than those of Mansfield, Whittinghton and Brusco and with this to the operating of Gibrat's law. The opposite holds for mechanical.
Even if the picture supplied by the kernel has a strong intuitive appeal we want now to use a synthetic measure of similarity to verify our conclusions on both sectors.

### 9.4 Similarity analysis

As we have hinted above (see equation 27), Gini's index is articulated on the difference of the cumulative distributions. It is an index of "dissimilarity". As a consequence, small values imply strong similarity. In Table 8 and Table 9 Gini's indexes on total assets have been multiplied by 1000 to make figure more legible.

**Table 8. Similarity matrices for total assets of ICT in 1999**

ICT - Total assets (1998-99)

| | I group | II group | III group | IV group | V group | VI group | VII group | VIII group |
|---|---|---|---|---|---|---|---|---|
| I group | | 7.778974 | 14.94291 | 16.29488 | 16.78602 | 19.21378 | 16.92112 | 20.99415 |
| II group | | | 6.843035 | 6.720662 | 8.518207 | 8.490543 | 8.737101 | 9.976101 |
| III group | | | | 2.6417 | 3.414634 | 4.92334 | 4.083618 | 6.659701 |
| IV group | | | | | 3.772457 | 3.870457 | 3.693389 | 6.772879 |
| V group | | | | | | 3.34305 | 2.708344 | 4.575375 |
| VI group | | | | | | | 3.200187 | 2.686766 |
| VII group | | | | | | | | 3.719558 |
| VIII group | | | | | | | | |

**Table 9. Similarity matrices for total assets of mechanical in 1999**

Mechanics - Total assets (1998-99)

| | I group | II group | III group | IV group | V group | VI group | VII group | VIII group |
|---|---|---|---|---|---|---|---|---|
| I group | | 12.69361 | 16.38432 | 19.03491 | 20.81868 | 24.06598 | 23.33458 | 25.83124 |
| II group | | | 6.451305 | 11.04344 | 12.89788 | 19.5057 | 18.65965 | 22.69289 |
| III group | | | | 5.187542 | 7.480622 | 11.69474 | 10.24287 | 13.81563 |
| IV group | | | | | 3.749566 | 9.848199 | 11.92265 | 18.50227 |
| V group | | | | | | 5.875576 | 5.50953 | 8.269595 |
| VI group | | | | | | | 4.653195 | 4.237481 |
| VII group | | | | | | | | 6.95521 |
| VIII group | | | | | | | | |

The results supplied by the matrices are particularly interesting.
First of all, we stress the strong dissimilarity of the first class with respect to the other ones for ICT (see Table 8). For mechanical we note a high dissimilarity of the first two classes with respect to the others (see Table 9). In ICT the index on assets shows a

strong similarity through the last three classes in 1998 (not reported for lack of space) and through the last four in 1999. On the contrary, in mechanical the eighth class does not show similarity with any class in 1998 and in 1999 just with the seventh.

## 10 Transition matrixes according to Champernowne.

In this section we explore transition matrices (described analytically in section 6) from 1998 to 1999 built on the basis of total assets.
To be brief, we shall report only those of the second biennium. The classes are ten in the starting year (number of rows) and 10 in the final year (number of columns), of the same proportional width, precisely equal to 4,6 in Mechanical and to 6,1 in ICT. The elements of the matrix are the relative frequencies of shift between classes from one year to the next (tables of absolute frequencies are not reported for lack of space).
The first remark to be made on the obtained classes, common to the two sectors, is that, if class width does not change in relative terms, in terms of absolute difference it greatly increases from the first classes to the last ones, creating, in the latter, a problem due to the corresponding small frequencies. This fact is to be considered while drawing conclusions: in the last classes, the probabilities depend on one or few firms, in particular in ICT; at the same time these firms represent the majority of the Assets of the whole sector. It is interesting to observe, first of all, that in the last three classes, ICT has only 4 firms in each of the three years against 13 (‘97), 12 (‘98) and 11 (‘99) of Mechanical: these results allow us to realize how different the growth processes are in the two sectors: in the very heavy size classes - the last three involving a percentage of the Total Assets between 77% and 79% - the ICT has lost no firms while the Mechanical has lost one per year.

### 10.1 Mechanical

Observing the transition matrices for the Mechanical (Table 10) results achieved through the other techniques are completely confirmed: suffice it to think that in both matrices (‘97-‘98 and ‘98-‘99) there is a firm in the largest size classes which jumps a class back, in the first two-year period from the ninth to the eighth and in the second from the tenth to the ninth. These two firms hold 18.8% of the total assets in 1997 and 7.7% in 1999.
The observations below confirm what we have previously seen: persistence probability in ‘97-‘98 increases rapidly as far as the fourth class, then remains more or less constant as far as the eighth; in the second matrix instead the growth stops at the fifth class while we notice an opposite trend from the seventh class onwards. At the same time in both matrixes the probability of going one class up decreases manifestly while the probability of going one class down increases. It can also be observed that in both matrices, from the fifth class onwards, the probabilities of a class transition are greater downwards than upwards (they are equal in one case only).

### 10.2 ICT

In ICT (see Table 11), the probability to remain in the same class from ‘97 to ‘98 stabilizes after the third class, contrary to mechanical; from ‘98 to ‘99 the trend shows a downward peak in the eighth class due, however, to the forward jump of a firm ( not a backward one as in mechanical): hence there are no impediments to growth in largest firms.

In ICT, the probability of shifting to a lower class, hence of negative percentage variations, is in practice constant with size increase in both biennia and therefore represents a full validation of the hypothesis of size shift.
As regards the probability of shifting to a higher class, its trend is similar to that of mechanical in both biennia. It must be specified, though, that in the second biennium, as many as one firm out of three passes from the eighth to the ninth class in ICT, whereas the other two remain in the same position; in mechanical, instead, only one firm out of eleven has the same jump, as another firm belonging in the same class has a jump downwards.
Besides, comparing for each class the probability of a jump upwards with that of a jump downwards, contrariwise to what we have seen in the mechanical, in ICT, in the second biennium, they remain unaltered starting from the fifth class, with the meaningful exception of the eighth class, in which the former outdoes the latter.
We can conclude, according to what we have been considering, that, in ICT, there are no impediments to growth in major size classes, quite the opposite.
Even if we have some elements supporting a good fitting of Champernowne model to ICT, the results in favour of major firms' growth and therefore of a concentration process exclude the suitability of the same model to this sector.
This fact cannot be compatible with the non-dissipative character imposed by Champernowne to his model.

**Table 10. transition matrix on Total Assets for Mechanical('98-'99)**

| Mechanics - Total assets 1998-99 | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| *T\t+1* | I | II | III | IV | V | VI | VII | VIII | IX | X |
| I | 0.5 | 0.5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| II | 0.0295 | 0.6841 | 0.2773 | 0.0068 | 0.0023 | 0 | 0 | 0 | 0 | 0 |
| III | 0.0007 | 0.0298 | 0.8271 | 0.1424 | 0 | 0 | 0 | 0 | 0 | 0 |
| IV | 0.0003 | 0.0003 | 0.0475 | 0.8972 | 0.0543 | 0.0003 | 0.0003 | 0 | 0 | 0 |
| V | 0 | 0 | 0.0005 | 0.0523 | 0.911 | 0.0361 | 0 | 0 | 0 | 0 |
| VI | 0 | 0 | 0 | 0 | 0.0474 | 0.9165 | 0.0361 | 0 | 0 | 0 |
| VII | 0 | 0 | 0 | 0 | 0 | 0.1236 | 0.8764 | 0 | 0 | 0 |
| VIII | 0 | 0 | 0 | 0 | 0 | 0 | 0.0909 | 0.8182 | 0.0909 | 0 |
| IX | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| X | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |

**Table 11. transition matrix on Total Assets for ICT ('98-'99)**

| ICT - Total assets 1998-99 | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| *t\t+1* | I | II | III | IV | V | VI | VII | VIII | IX | X |
| I | 0.5424 | 0.4576 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| II | 0.0138 | 0.7996 | 0.186 | 0.0006 | 0 | 0 | 0 | 0 | 0 | 0 |
| III | 0.0002 | 0.0348 | 0.8827 | 0.0815 | 0.0007 | 0 | 0 | 0 | 0 | 0 |
| IV | 0 | 0 | 0.0612 | 0.8866 | 0.0515 | 0.0006 | 0 | 0 | 0 | 0 |
| V | 0 | 0 | 0 | 0.0665 | 0.8703 | 0.0633 | 0 | 0 | 0 | 0 |
| VI | 0 | 0 | 0 | 0 | 0.0143 | 0.9714 | 0.0143 | 0 | 0 | 0 |
| VII | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| VIII | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.6667 | 0.3333 | 0 |
| IX | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| X | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |

## 11 Conclusions

In this paper we have analyzed and applied three classical models of stochastic growth: Gibrat, Kalecki, Champernowne to Italian economy.
All the three models have a common feature: opportunities for firm growth are largely determined by stochastic factors, with the same distributional result for Gibrat and Kalecki (lognormal) and a different one for Champernowne (Pareto).
If this is the common feature of the three models, the path of growth described by each of them is completely different: increase of concentration in Gibrat's, impediment to growth for large firms and a constant level of concentration in Kalecki's; decrease of concentration in Champernowne, the expected value of variations being negative.
On the basis of these three models we have analysed the universe of Companies for mechanical and ICT. To this purpose we have used the File of Companies accounts of Cerved-Pitagora respectively for a total of 9822 and 7887.
The approaches we have used have been the fitting of Lognormal and Pareto distributions to the data, the analysis of an index of growth, the construction of transition matrices according to Champernowne.
By means of a bootstrap procedure we have observed that sample size influences the *p*-value of classic distribution tests (chi-squared and Kolmogorov-Smirnov) leading to an increasing probability to reject the null hypothesis as the size increases. This is true for mechanical where lognormal distribution cannot be refused for sample size until 800. This is not the same for ICT. Thus, we have looked at the fitting of empirical distribution to lognormal in mechanical by means of histograms and kernel density estimation and we have observed a good fitting.
We have then been able to fit Pareto to both sectors even if on different bases: last 10% of the universe for mechanical and last 30% for ICT.
The analysis of residuals obtained in the log size/log rank regression - used to fit Pareto distribution to the data - has emphasized a tendency of the 4 largest companies of ICT - which represent roughly 60% of the whole Sales of the sector - to maintain an effective size above the estimated and an opposite behaviour of the largest companies of Mechanical.
The conclusions we draw by fitting the distributions are:

a) the prevailing stochastic nature of growth in the two sectors;
b) the existence of impediments to growth for the largest companies of Mechanical and, conversely, the tendency for the largest units of ICT to reach levels of size above the estimated;
c) the conforming of Mechanical to Kalecki's model;
d) the presence of a process of concentration in ICT which has to be interpreted neither as consistent with Gibrat law, given the impossibility to fit lognormal to the same sector, nor conforming to Champernowne. Even if ICT is well moulded by Pareto - which is the referring distribution of Champernowne's - it does show a dissipative process.

To corroborate these results we have developed two experiments.
- Through the first we have studied the distributions of the index of growth in the 8 portions of our empirical distributions obtained by means of the percentiles. On each of these eight portions we have studied the distributions of the index of growth.
This analysis has been developed using median, robust measures of variability, a kernel approach and a Gini index of similarity.

- By means of the second experiment we have constructed transition matrices à la Champernowne by which we analyze the persistence/migration of companies from one class to another.

Migration and persistence in Champernowne transition matrices confirm the above results.

Although the considered years are not many, the largeness of our data set and the methodology we have used, confer a considerable robustness to the results we have obtained.

## References

Aitchison J., Brown J.A.C. (1957) The lognormal distribution (with special references to its use in economics). Cambridge University Press, Cambridge.

Audretsch D.B., Santarelli E., Vivarelli M. (1999) Start-up size and industrial dynamics: some evidence from italian manufacturing. International Journal of Industrial Organization 17: 965-983.

Axtell R. (2001) Zipf distribution of U.S. firm sizes. Science vol. 293.

Barca F. (1985) Tendenze nella struttura dimensionale dell'industria italiana: una verifica empirica del 'Modello di specializzazione flessibile'. Politica Economica 1: 71-109.

Boeri, T. (1989) Does firm size matter?. Giornale degli Economisti e Annali di Economia 43: 477-495.

Bonini C.P., Simon H.A. (1958) The size distribution of business firms. American Economic Review 48:607-617.

Brusco S, Giovannetti E., Malagoli W. (1979) La relazione tra dimensione e saggio di sviluppo nelle imprese industriali: una ricerca empirica. Università degli Studi di Modena, Studi e Ricerche dell'Istituto Economico, No.5.

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

Cefis E., Ciccarelli M., Orsenigo L. (2001) The growth of firms: from Gibrat's legacy to Gibrat's fallacy. working paper.

Champernowne, D.G. (1973) The distribution of Income between Persons. Cambridge University Press.

Cipollini F., Ganugi P. (2001) The "true" distribution of industrial districts: a non parametric analysis. Aarhus CAED.

Cox D.R, Miller H.D. (1965) The theory of stochastic processes. Chapman and Hall.

Dosi G., Riccaboni M., Varaldo R. (2001) La dinamica dei rapporti tra grandi e piccole imprese: alcune riflessioni tra teoria e analisi storica. LEM Working Paper, Scuola Superiore Sant'Anna, Firenze.

Dunne P., Hughes A. (1994) Age, size growth and survival: UK Companies in the 1980s. The Journal of Industrial Economics, XLII: 115-140.

Faliva M. (2002) Modelling financial return distributions: a neat solution for moderately leptokurtic data. Studi in onore di Angelo Zanella a cura di Frosini, Magagnoli e Boari, editore Vita e Pensiero Università, Milano.

Feller W.A. (1968) An introduction to probability theory and its applications. Wiley and Sons, New York.

Ganugi P. Grossi L., Gozzi G., Gagliardi R. (2002) Growth with statistical regularity. The evidence of Italian ICT. IAOS, London , cd.

Ganugi P. Grossi L., Gozzi G. (2003) La distribuzione della dimensione delle società meccaniche dell'Italia: un'analisi spaziale su un panel fisso, working paper.

Geroski P. (1999) The growth of firms in theory and in practice. Industrial organization, CEPR discussion paper.

Guarini, R., Tassinari F. (2000) Statistica Economica. il Mulino, Bologna.

Hall B.H. (1987) The relationship between firm size and firm growth in the US manufacturing sector. The Journal Of industrial Economics 35: 583-606.

Hart P.E., Oulton N. (1996) Growth and size of firms. The Economic Journal 106: 1242-52.

Hart P.E.,Oulton N. (1999) Gibrat, Galton and job generation. International Journal of the Economics of Business 6, No. 2: 149-164.
Hart P.E., Prais S.J. (1956) The analyisis of business concentration: a statistical approach. Jour.Royal Stat. Soc., Ser A., 119: 150-191.
Hymer S., Pashigan P. (1962) Firm size and rate of growth. Journal of political economy 70: 556-569.
Ijiri Y., Simon H.A. (1964) Interpretations of departures from the Pareto curve firm-size distributions. Journal of Political Economy 82: 315-331.
Ijiri Y., Simon H.A. (1977) Skew Distributions and the Sizes of Business Firms. North Holland, Amsterdam
Johnson N.L., Kotz S., Balakrishnan N. (1994) Continuos Univariate distributions. Wiley, New York.
Kalecki M. (1945) On the Gibrat distribution. Econometrica 13: 161-170.
Kattuman P. (1996) On the size distribution of business of large enterprises. UK Manufacturing, Structural Change and Economic Dynamics 7: 479-494.
Krzanowski W.J., (2000) Principles of multivariate analysis: a users perspective. Rev. Ed. Oxford University Press, Oxford.
Leti G. (1983) Statistica descrittiva. Il mulino, Bologna.
Lotti, F., Santarelli, E., Vivarelli, M. (1999) Does Gibrat’s law hold in the case of young, small firms? Working Paper n° 361, Dipartimento di scienze Economiche Università degli Studi di Bologna.
Lotti F., Santarelli E. (2001) Is firm growth proportional? an appraisal of firm size distribution. Economics Bullettin 12, No. 6:1-7.
Lotti F., Santarelli E., Vivarelli M. (2001) The relationship between size and growth: the case of italian newborn firms. Applied Economics Letters, 8: 451-154.
Mansfield E. (1962) Entry, Gibrat’s law, innovation, and the growth of firms, American Economic Review. 52: 1023-1051.
Marsili O. (2001) Stability and turbolence of the size distribution of firms: evidence from dutch manufacturing. Paper prepared for the CAED’01 Conference, october 2001, Aarhus, Denmark.
Marsili O., Salter A. (2002) Is innovation democratic? Skewed distributions and the return to innovation in Dutch manufacturing. CEREM publications, Netherlands.
Piergiovanni R., Santarelli E., Klomp L., Thurik A.R., 2002, Gibrat’s law and the firm size firm growth relationship in Italian Services. Tinbergen Institute Discussion Paper.
Quandt R.E. (1966) On the size distribution of firms. American Economic Review 3, pp.416-432.
Sawyer, C. (1979) Variance of logarithms and industrial concentration. Oxford Bullettin of Economics and Statistics 41: 165-181.
Scherer F.M. (1980) Industrial market structure and economic performance. Houghton Mifflin Company.
Schmalensee, R. (1979) Using the H-index of concentration with published data. Review of Economics and Statistics 59: 186-193.
Schmalensee, R. (1992) Inter-industry studies of structure and performance. Handbook of Industrial Organization, Schmalensee eds., vol II: .951-100, North Holland, Amsterdam.
Simon H. A. (1955) On a class of skew distribution functions’. Biometrika 52:425-440.

Simon H. A. (1960) Some further notes on a class of skew distribution functions. Information and Control 3:80-88.

Singh A., Whittinghton G. (1975) The size and growth of firms. Review of Economic Studies 42: 15-26.

Solinas G. (1996) I processi di formazione, la crescita e la sopravvivenza delle piccole imprese. Francoangeli.

Stanley M.H.R. et al. (1995) Zipf Plots and the size distribution of firms. Economics Letters 49: 453-457.

Steindl J. (1965) Random processes and the growth of firms. Griffin, London.

Sutton, J. (1997) Gibrat's legacy. Journal of Economic Literature 35: 40-59.

Sutton J. (1998) Technology and market structure. MIT Press, Cambridge Massachussets.

Vennet R.V. (2001) The Law of Proportionate Effect and OECD bank sectors. Applied Economics 33: 539-546.

Voit J. (2001) The growth dynamics of German business firms. Advances in Complex Systems, Vol.4, No.1: 149-162.