Laboratory of Economics and Management
Sant'Anna School of Advanced Studies
Piazza Martiri della Libertà, 33 - 56127 PISA (Italy)
Tel. +39-050-883-343 Fax +39-050-883-344
Email: lem@sssup.it Web Page: http://www.lem.sssup.it/

LEM

Working Paper Series

Corporate Growth and Industrial Dynamics: Evidence from French Manufacturing

Giulio Bottazzi*
Alex Coad°
Nadia Jacoby°
Angelo Secchi*

*Scuola Superiore Sant'Anna, Pisa
°Universitè Paris 1 Pantheòn-Sorbonne, Paris, France

2005/21 March 2007

ISSN (online) 2284-0400

Corporate growth and industrial dynamics: evidence from French manufacturing*

Giulio Bottazzi [a,†] Alex Coad[a,b] Nadia Jacoby[b] Angelo Secchi[a]

[a] S.Anna School of Advanced Studies, Pisa, Italy
[b] MATISSE, Univ. Paris 1 Panthéon-Sorbonne, Paris, France

March 16, 2007

Abstract

This work explores basic properties of the size and growth rates distributions of firms at the aggregate and disaggregate levels. Using an extensive dataset on French manufacturing firms, we investigate which properties of firm size distributions and growth dynamics characterize the aggregate dynamics and are, at the same time, robust under disaggregation. Our analysis is based on non-linear robust regression methods which have never been applied before to this kind of data. The growth rates distributions we observe are well described by a Subbotin distribution with a shape parameter significantly lower than 1, suggesting a noticeable departure from the Laplace behavior reported in previous works on Italian and US data. At the same time, the variance of growth rates depends negatively on size and the relationship does not seem to be linear, with larger firms possibly displaying lower variability in their growth dynamics. At the disaggregate level, we observe significant heterogeneity in the firm size distributions across sectors while the shape of the sectoral growth rates density displays a surprising degree of homogeneity.

1 Introduction

A number of early studies into industrial structure focused on the adequation of theoretical distribution functions (in particular the Pareto, the Yule, and the log-normal) to aggregate distributions of firm size. This pioneering line of research began with Gibrat (1931) investigation of the French manufacturing sector, and was later applied to the UK manufacturing (Hart and Prais, 1956) and also to US (Simon and Bonini, 1958; Quandt, 1966) and Austrian data (Steindl, 1965).

Another strand of literature has focused on the well-known 'Law of Proportionate Effect', a statistical process formulated by the engineer Gibrat as he attempted to explain the emergence of the aggregate size distribution. This 'law' states that, in a context of constant returns to scale, firm growth follows a purely stochastic process, with growth rates being independent of

*Part of the research that has lead to this work has been undertaken as part of the activities of the DIME Network of Excellence, sponsored by the European Union.

†*Corresponding Author*: Giulio Bottazzi, Scuola Superiore S.Anna, P.za Martiri della Liberta' 33, 56127 Pisa, Italy. *E-mail*: bottazzi@sssup.it. *Phone*: +39-050-883343. *Fax*: +39-050-883344.

firm size. Despite its blatant lack of economic content, Gibrat's Law is nonetheless very useful, as it provides a sort of 'null hypothesis' against which corporate growth can be compared. A large body of research (see for example Mansfield (1962); Evans (1987); Hall (1987) and Dunne *et al.* (1988); see also Sutton (1997) for a review) generally seems to suggest that the 'Gibrat Law' benchmark can be taken as a rough first approximation of firm growth. However, a closer inspection reveals that firm size usually experiences a slight reversion to the mean (i.e. small firms having higher average growth rates than larger ones), and that several other econometric issues require special attention (such as heteroskedasticity, autocorrelation, and a sampling bias due to higher exit rates of small firms). As further investigation of this topic, several recent contributions have explored the distribution of growth rates. Using data on US manufacturing firms, Amaral *et al.* (1997) observe that the distribution is 'tent-shaped' on log-log plots and closely resembles the Laplace. This line of research has been extended to consider the Subbotin family of distributions, of which the Laplace is a special case. Growth rate distributions close to the Laplace have been observed using US data (Bottazzi and Secchi, 2003), Italian manufacturing data (Bottazzi *et al.*, 2003), and also data from the worldwide pharmaceutical industry (Bottazzi *et al.*, 2001)).

Concerning the comparison between aggregate and disaggregate properties, recent theorizing Dosi *et al.* (1995) and evidence from disaggregated analysis Bottazzi *et al.* (2003) suggests that the characteristics of the size distribution are not a robust feature of the different industries but appear, instead, as a mere statistical effect of aggregation. As a result, the distribution of firm size seems to be of limited interest to economists. On the other hand, the Laplace distribution of growth rates appears to be an extremely robust characteristic of industrial dynamics, with a high homogeneity of the distribution which holds at various levels of aggregation. Speculation emerging from the findings on US, Italian and pharmaceutical databases suggests that the Laplace distribution of corporate growth rates seems to be something of a 'stylized fact'. In this vein, Bottazzi and Secchi (2006) construct a theoretical model capable of reproducing a Laplace distribution of growth rates.

More than 70 years after Gibrat's seminal book, we return to the study of the French manufacturing sector. The timing of our work is important because it helps in understanding the degree of generality and the robustness of previous results. For instance, contrary to prior results, the present analysis provides evidence that the Laplace distribution of growth rates cannot be considered as a universal property of industrial dynamics. Looking at French manufacturing, we observe growth rates distributions with tails that are consistently fatter than those of the Laplace.

In many respects, the statistical characteristics which emerge from the present analysis seem to arbitrate between previous findings. For example, whilst variance of growth rates decreased with size in the American case (Bottazzi and Secchi, 2003), it did not for Italian firms (Bottazzi *et al.*, 2003). Here we find that a negative, though weak, relationship does exist. Also, whilst previous research had found growth rate autocorrelation that was either positive (for US data, Ijiri and Simon (1967)) or negative (for Italian data, Bottazzi *et al.* (2003)), the evidence presented here suggests that French firms experience a slight negative autocorrelation in their growth patterns.

After a brief description of the data in Section 2, Section 3 provides the results at the aggregate level on firm size distribution and growth rates distributions while Section 4 focuses on a sectoral analysis. Finally, Section 5 summarizes our findings and sketches several future directions of research.

Year	Std. Dev.	Skewness	Kurtosis
1996	1.11	1.04	1.47
1997	1.12	1.03	1.42
1998	1.12	1.01	1.37
1999	1.12	0.99	1.35
2000	1.14	0.96	1.32
2001	1.14	0.94	1.31
2002	1.16	0.91	1.25

Table 1: Descriptive statistics of $s_i(t)$ in different years. Size measured in terms of Total Sales.

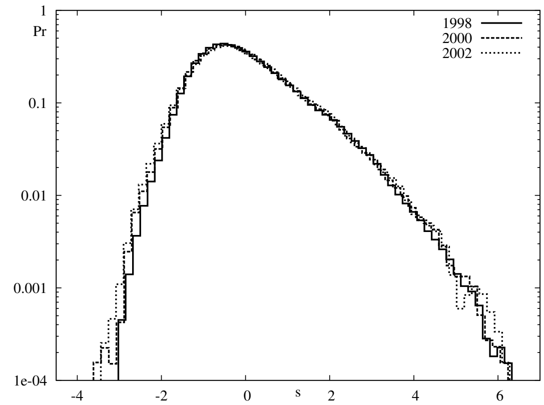

Figure 1: Kernel estimates of the density of firm size in 1998, 2000 and 2002. Densities are computed in 64 equispaced points using an Epanenchnikov kernel. Note the logarithmic scale on the y-axis.

2 Data description

This research draws upon the EAE databank collected by SESSI and provided by the French Statistical Office (INSEE).[1] This database contains longitudinal data on a virtually exhaustive panel of French firms with 20 employees or more over the period 1989-2002. We restrict our analysis to the manufacturing sectors. For statistical consistency, we only utilize the period 1996-2002 and we consider only continuing firms over this period. Firms that entered midway through 1996 or exited midway through 2002 have been removed. Since we want to focus on internal, 'organic' growth rates, we exclude firms that have undergone any kind of modification of structure, such as merger or acquisition. Because of limited information on restructuring activities and in contrast to some previous studies (e.g. Bottazzi *et al.* (2001)), we do not attempt to construct 'super-firms' by treating firms that merge at some stage during the period under study as if they had been merged from the start of the period. Firms are classified according to their sector of principal activity.[2] To start with we had observations for around 22000 firms per year for each year of the period.[3] In the final balanced panel constructed for the period 1996-2002, we have exactly 10000 firms for each year.

3 Aggregate properties

This section is devoted to the statistical analyses of the firm size distribution and of firm dynamics considering data aggregated over all the industrial sectors. We use the firms' total sales as a measure of size and we define $S_i(t)$ the size of firm i at time t.

[1]The EAE databank has been made available to Nadia Jacoby and Alex Coad under the mandatory condition of censorship of any individual information.

[2]The French NAF classification matches with the international NACE and ISIC classifications.

[3]22319, 22231, 22305, 22085, 21966, 22053, and 21855 firms respectively

3.1 Size distribution

We develop our analysis of firm size along different but complementary directions. To begin with, we explore the firm size distribution, studying its stationarity and its shape, paying particular attention to the shape of the upper tail for which we have more reliable data. We then focus on the autoregressive structure of firm size by investigating how the French data measures up to Gibrat's Law. Finally we explore the existence of relations between size and growth.

In order to eliminate the common trend in the average size we define the normalized (log) size $s(t)$ as

$$s_i(t) = \log(S_i(t)) - \frac{1}{N}\sum_{i=1}^{N}\log(S_i(t)) \qquad (1)$$

where N stands for the total number of firms.

Table 1 presents some summary statistics for the rescaled sizes $s_i(t)$ over the period 1996-2002 clearly suggesting that their distribution is remarkably stationary. There are at least two other properties of firms size deserving to be highlighted. First, we confirm once again (cfr. among many others Hart and Prais (1956); Ijiri and Simon (1977), and Bottazzi *et al.* (2003)) that the distribution of firm sizes is right-skewed as indicated by the positive values for the skewness. Second, the high values for the excess kurtosis statistics provide evidence of distribution tails fatter than in the Gaussian case.

Figure 1 presents the kernel density estimate[4] of firm size in three different years, at the beginning, middle and end of the period. Intuitively, a kernel density estimate can be considered to be a smoothed version of the histogram, obtained by counting the observations in the different bins as the width of the bins varies. This estimate requires the provision of two objects: the kernel function K and the bandwidth h of the bin. Formally, we have

$$\hat{f}(x,t;h) = \frac{1}{n\,h}\sum_{i=1}^{n} K\left(\frac{x - s_i(t)}{h}\right) \qquad (2)$$

where $s_1(t),\ldots,s_N(t)$ are the number of observations n in each sector, h is a bandwidth parameter controlling the degree of smoothness of the density estimate, and where K is a kernel density, i.e. $K(x) \geq 0, \forall x \in (-\infty, +\infty)$ and $\int dx\, K(x) = 1$.[5]

Figure 1 confirms again that the size distribution presents a strong right-skewed shape that does not seem to change over time. Bearing in mind that the data is truncated and excludes firms having less than 20 employees, we then focus our attention on the upper tail only, beyond the mode. We observe the existence of a power-like tail which can be seen as a roughly straight line of negative slope linking the density (on a log scale) with size. This simple visual inspection reveals the coexistence of many relatively small firms coexisting with a few very large ones. As noticed by many authors (see for example Dosi (2005)), the naïve notion of an 'optimal size' around which firms will fluctuate does not sit comfortably with empirical results.

	OLS	Chesher Regression			
		OLS		LAD	
Year	α_{OLS}	β	ρ	β	ρ
1998	0.9807 0.0019	0.9907 0.0013	-0.251 0.025	0.9941 0.0011	-0.0710 0.005
1999	0.9845 0.0019	0.9893 0.0014	-0.166 0.024	0.9967 0.0011	-0.0082 0.006
2000	0.9965 0.0019	1.0005 0.0015	-0.194 0.024	1.0062 0.0011	-0.0535 0.006
2001	0.9869 0.0018	0.9945 0.0014	-0.202 0.023	0.9976 0.0011	-0.0572 0.006
2002	0.9926 0.0020	0.9978 0.0015	-0.222 0.029	1.0036 0.0011	-0.0532 0.006

Table 2: Gibrat law regression coefficients using OLS (see equation (3)) and also Chesher (1979) method (equation (5) estimated using OLS and LAD).

Investigating Gibrat's Law - The autoregressive structure of firm size

How does the French dataset compare to the Gibrat Law benchmark? We investigate this by regression analysis. To begin with, we use normalized (log) sales to estimate an AR(1) model

$$s(t) = \beta\, s(t-1) + \epsilon(t) \tag{3}$$

where ϵ is an error term. Note that we have no need for a constant term, because we have already normalized the observations, removing their mean. Gibrat's Law is usually said to hold if β has a value not different from 1. Values smaller than 1 imply that small firms grow faster, on average, than large firms, whilst values larger than 1 imply the opposite.

The results of the OLS estimation of equation (3) are reported in Table 2 (errors are corrected for heteroskedasticity using the jackknife method described in MacKinnon and White (1985)). It is apparent that even if the coefficient β is very close to 1 it is always statistically different from it. However, Chesher (1979) shows that OLS estimation of the Gibrat Law coefficient may imply an estimation bias, if autocorrelation is present in the error term. He also advances that the Gibrat Law cannot be said to hold if this autocorrelation exists, because size and growth are no longer independent. In order to correct for such autocorrelation, he proposes to fit the following system

$$\begin{cases} s(t) = \beta s(t-1) + \epsilon(t) \\ \epsilon(t) = \rho\epsilon(t-1) + u(t) \end{cases} \tag{4}$$

where $\epsilon(t)$ is an autocorrelated error term and $u(t)$ is an *i.i.d.* error term. Noting that $\epsilon(t)$ may be expressed in terms of $s(t-1)$ and $s(t-2)$, we can rewrite the above system as the equivalent equation:

$$s(t) = \gamma_1 s(t-1) + \gamma_2 s(t-2) + u(t) \tag{5}$$

[4] These estimates are built following Silverman (1986).

[5] Throughout this paper the kernel function will always be the Gaussian density. The use of different kernels, such as the Epanechnikov or the Triangular, does not change noticeably our results. Where not specified otherwise, the bandwidth h has been chosen according to Silverman (1986), Section 3.4.

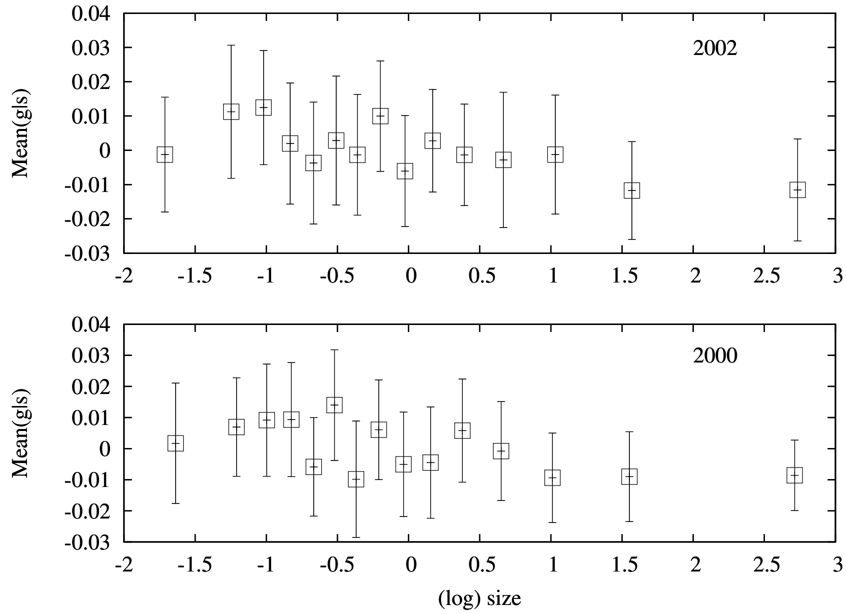

Figure 2: Scaling relation of the conditional mean growth rate with respect to firms' (log) size computed using 15 equipopulated bins in 2000 and 2002. Confidence intervals are reported as two standard errors.

where $\gamma_1 = \beta + \rho$ and $\gamma_2 = -\beta\rho$. We estimate β using OLS estimation of the parameters γ_1 and γ_2 in equation (5), and obtain the results reported in Table 2.[6] There is, however, a further problem affecting our estimates. The procedure just applied assumes $u(t)$ to be an i.i.d. Gaussian error term which, as we will show later, is not the case here. Indeed the analyses of the next sections will suggest that the Laplace distribution would be a far better assumption. In order to take the non-normality of the error term into account we also estimate equation (5) using the Least Absolute Deviation (LAD) approach assuming that the error term is distributed according to the Laplace. These results are reported again in Table 2.

The main finding of our analyses of the Gibrat's regression on French data is that the coefficient β, even if still statistically different from 1, becomes closer to one when one uses a regression technique that takes explicitly into account the possible existence of autocorrelation in the $\epsilon(t)$ error term. Indeed, this autocorrelation is actually present in our data. The ρ statistics presented in Table 2 are all significantly negative (though not very large), and roughly speaking they suggest that French firms experience a negative growth rate autocorrelation of magnitude no larger than around 7%. (Strictly speaking, the ρ coefficients correspond to the magnitude of growth rate autocorrelation once the dependence of growth on size has been controlled for.) These preliminary results suggests the need for further investigations of this issue allowing in equation (4) for a more general AR structure of the error term ϵ possibly including also Moving Average (MA) components.

Exploring non-linearities and the Scaling Effect

In this section we continue our analysis investigating the existence of non-linear relations between firm size and characteristics of growth rates. Accordingly with what done in the

[6] In estimating equation (5) we checked for possible autocorrelation in the error term $u(t)$, but we did not find any. Had it been present, such autocorrelation would have given us unreliable results.

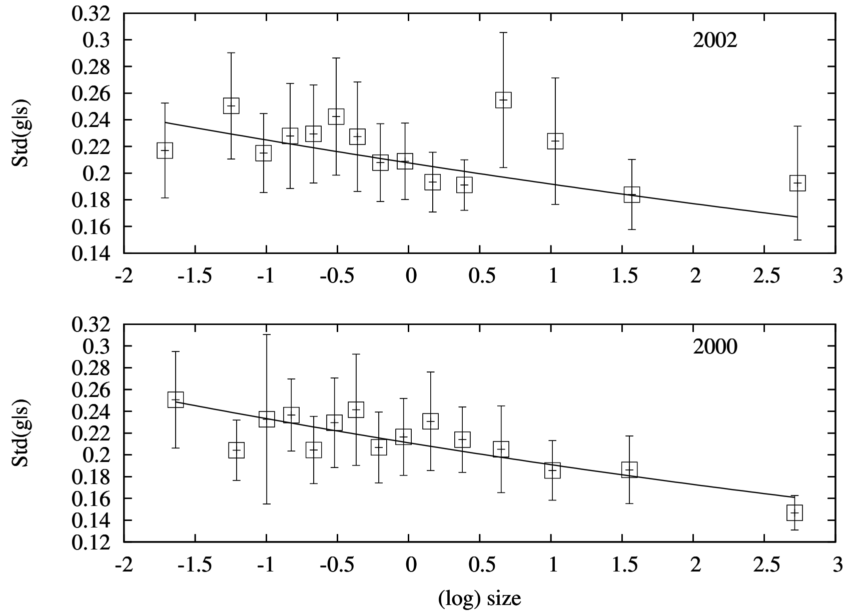

Figure 3: Scaling relation of the conditional standard deviation of growth rate with respect to firms' (log) size computed using 15 equipopulated bins in 2000 and 2002. Confidence intervals are reported as two standard errors.

previous section we define firms growth shocks as the residuals of regression equation (4), $\hat{u}(t)$. Our search for relations between $s(t)$ and $\hat{u}(t)$ is organized in two steps. First, we use a graphical analysis to obtain some hints on the existence and on the shape of such relations. Second we assess the robustness of any observed relationship applying regression techniques.

Since any linear relationship between firm size and growth rates has been captured by (4), it only remains to assess if any residual non-linear effect is present. To explore this issue we group our observations into 15 bins according to firm size and we plot in Figure 2, for two different years choosen as examples, the average growth rate in each bin against the (log) size. As expected we do not observe evidence of any linear relation between size and average growth. Moreover the visual inspection of Figure 2 rules out also the possibility that such a relation presents a nonlinear nature.

Next we consider the question of whether or not the variance of growth rates is related to firm size. Some previous studies (e.g. Amaral *et al.* (1997)), although not all (e.g. Bottazzi *et al.* (2003)), have observed a significant negative exponential relationship between $s(t)$ and the standard deviation of $\hat{u}(t)$. To investigate this for the French data we group again our observations into 15 bins according to size and we plot the conditional standard deviation of growth rates in each bin against the (log) size. Figure 3 shows that also for French firms a clear negative relationship emerges: the standard deviation of growth rates decreases with size suggesting that bigger firms present lower variability in their growth rates compared with the smaller ones.

In order to assess the statistical significance of the apparent nonlinear relation between $s(t)$ and the standard deviation of $\hat{u}(t)$ we opt for a nonlinear regression. In order to provide comparability with previous works (cfr. Amaral *et al.* (1997); Bottazzi *et al.* (2003) and Bottazzi and Secchi (2003)), we estimate the model:

$$\hat{u}(t) = e^{-\alpha\, s(t-1)}\, g(t) \tag{6}$$

Year	Scaling Relation		Subbotin fit	
	Type of regression	α	b coefficient	a coefficient
1998	non-linear OLS	-0.077 0.019		
	non-linear LAD	-0.075 0.004	0.774 0.176‰	0.110 0.137‰
1999	non-linear OLS	-0.060 0.020		
	non-linear LAD	-0.068 0.004	0.763 0.176‰	0.111 0.138‰
2000	non-linear OLS	-0.098 0.020		
	non-linear LAD	-0.074 0.004	0.800 0.177‰	0.115 0.136‰
2001	non-linear OLS	-0.072 0.022		
	non-linear LAD	-0.062 0.004	0.790 0.177‰	0.111 0.137‰
2002	non-linear OLS	-0.038 0.024		
	non-linear LAD	-0.055 0.004	0.807 0.178‰	0.119 0.136‰

Table 3: Estimated coefficient α in (6) obtained with non linear regressions under the assumption of Gaussian(OLS) and Laplacian(LAD) error term. Standard errors are also reported. We also report the maximum likelihood estimate (and coefficient of variation) of the Subbotin density (see equation (7)) on firms growth rates rescaled as in (6).

where $\hat{u}(t)$ is the residual of the regression in (4) and $g(t)$ is an error term. Equation (6) describes a regression model with a heteroskedastic error term $e^{-\alpha\, s(t-1)}\ g(t)$ which, in line with our visual inspection of Figure 3, assumes that the variance of growth rates is greater among smaller firms. We fit the data to this econometric specification to estimate the value of α. First we estimate the model in (6) assuming the normality of the error term $g(t)$, using a standard OLS approach. Furthermore we perform a LAD regression under the assumption that error terms are distributed according to the Laplace distribution.[7] Again the results are reported in Table 3. In all cases we observe a small though statistically significant negative relationship between size and growth rate variance, independently from the estimation method adopted.

3.2 Growth rates distribution

In this section, we analyze the shape and the evolution in time of the growth rates density, adopting a non-parametric approach. In the previous section we showed that the variance of growth rates decreases with firms size following an exponential decay. We use this finding to define a rescaled version of the growth rate $\hat{g}(t)$ obtained as the residual in the estimation of equation (6). Notice that the statistical properties of $\hat{g}(t)$ are by construction independent of firm size. One important implication of this rescaling is the possibility of pooling together growth rates of firms belonging to different size bins.

Figure 4 reports, on a log scale, the kernel estimates of the empirical density of $\hat{g}(t)$ in three different years. We observe a characteristic tent-shape, although the fat tails make the

[7] We will argue in the next section why this second assumption is much more appropriate in this case.

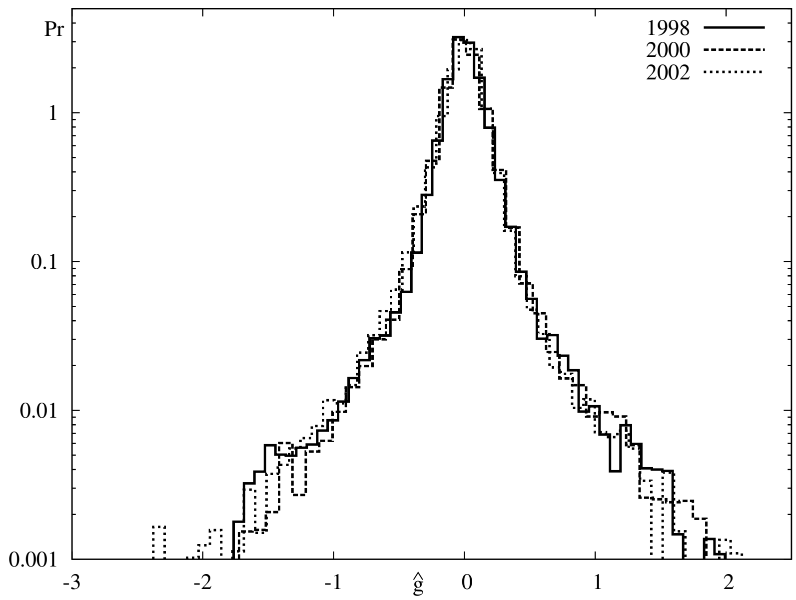

Figure 4: Kernel estimates of the growth rates density in 1998, 2000 and 2002. Densities are computed for 64 equispaced points using an Epanenchnikov kernel. Note the logarithmic scale on the y-axis.

tent-shape appear rather 'droopy'. This fat-tailed distribution of growth rates corresponds to a high frequency of extreme growth events for French manufacturing firms.

Previous studies have considered growth rates as being distributed according to the Laplace, which can be considered as a special case of the Subbotin family of distributions (Bottazzi *et al.*, 2003)). Having observed the growth rate distribution in Figure 4 we now turn to parametric methods of quantifying the distribution. To do this, we estimate the Subbotin parameters of the growth rates distribution.

The Subbotin distribution can be formally presented by the following equation:

$$f_s(x) = \frac{1}{2ab^{1/b}\Gamma(1/b+1)} e^{-\frac{1}{b}\left|\frac{x-\mu}{a}\right|^b} \tag{7}$$

where $\Gamma(x)$ is the Gamma function. The distribution has three parameters - the mean μ, the dispersion parameter a and the shape parameter b. As the shape parameter b decreases, the tails of the density become fatter. The density is leptokurtic for $b < 2$, and platykurtic for $b > 2$. Two noteworthy special cases of the Subbotin distribution are the Gaussian distribution (for $b = 2$) and the Laplace distribution (with $b = 1$).

We estimate the values of the parameters using the maximum likelihood procedure discussed in Bottazzi and Secchi (2006). Results are reported in Table 3. The robust conclusion is that the distribution of growth rates appears to be even more fat-tailed than the Laplace distribution (for which b would equal 1). This surprising result distinguishes growth patterns of French firms from those observed elsewhere, where distributions close to the Laplace are observed (Amaral *et al.*, 1997; Bottazzi *et al.*, 2003; Bottazzi and Secchi, 2003). This fat-tailed distribution of growth rates corresponds to a higher frequency of extreme growth events. Compared to results reported for Italian or US manufacturing firms, French firms are much more likely to undergo significant positive or negative changes in size.

4 Sectoral properties

The preceding analysis can be repeated at a disaggregated level. We consider this to be a worthwhile enterprise because there may well be a tension between regularities observed in aggregated data and much 'messier' results at a disaggregated level (see Dosi *et al.* (1995) for a discussion). Our results show that some properties of industrial dynamics, such as the growth rates distribution, survive disaggregation i.e. are present also at a sectoral level. However, for the firm size distribution, the smooth shape that emerges from aggregated data disappears, and we observe that significant multimodality is rife at the sectoral level. Looking at the 2-digit level of ISIC industry classification, we retain sectors 17-36 which correspond to manufacturing activities.[8] Table 4 gives a description of these sectors. Note that sectors 23 and 30 have only a small number of observations, which disqualifies them from detailed quantitative analysis.

4.1 Size distribution

We start by looking at the size distribution, using a non-parametric method to explore the shape of the firm size distribution at the disaggregate level. We test the size distribution for multimodality, and then present concentration statistics based on the properties of the distribution's upper tails. We present some kernel density plots of exemplary sectors, that have been chosen to highlight inter-sectoral diversity. We also look at Gibrat's Law statistics for each sector.

4.1.1 Firm size distribution

Similar to the previous methodology, we take the log of sales and then normalize the observations by deducting the sectoral mean. The normalized sectoral (log) sales of firm i in sector j can thus be defined as

$$s_{ij}(t) = log(S_{ij}(t)) - \frac{1}{N_j}\sum_{i=1}^{N_j} log(S_{ij}(t)) \tag{8}$$

where N_j is the number of firms in the j-th sector.

We use these normalized observations to examine the size distribution of firms in the same sectors. Although at the aggregate level we observe a rather regular unimodal distribution, previous studies suggest that this unimodality may not hold at a finer level of analysis (see for example Bottazzi and Secchi (2003)). To begin with, we build a kernel estimate (Silverman, 1986) of the probability density of firm size, in order to visualize the shape of the sectoral-level size distribution.

Although we observe stationarity of the sectoral size distribution over the 7-year time period, the shape of the distribution varies greatly across sectors. In particular, we may observe multimodality and/or different shapes and gradients for the upper tails. The presence of multimodality is not unusual. In their study of the worldwide pharmaceutical industry, for example, Bottazzi and Secchi (2005) observe significant bimodality in the size distribution and relate this to a cleavage between the industry leaders and fringe competitors. Figures 5 and

[8]Strictly speaking, the 2-digit sector '37', which corresponds to the recycling industry, is also included in the manufacturing sector. However, only a small number of firms are reported in this sector. As a consequence, it was dropped from the analysis.

Table 4: Description of the manufacturing sectors studied

ISIC class	Description	No. obs.	Mean size €'000 in 2002	Bimodality test (p-values)	D^4_{20}
17	Manuf. of textiles	730	9703	0.594	0.3892
18	Manuf. of wearing apparel, dressing and dyeing of fur	498	9623	0.000	0.5461
19	Tanning and dressing of leather, manuf. of luggage, handbags, ...	205	14629	0.045	0.5995
20	Manuf. of wood and products of wood and cork, except furniture; ...	314	9083	0.002	0.3269
21	Manuf. of paper and paper products	364	22428	0.031	0.3938
22	Publishing, printing and reproduction of recorded media	820	13745	0.022	0.4173
23	Manuf. of coke, refined petroleum products and nuclear fuel	19	73547	-	-
24	Manuf. of chemicals and chemical products	496	52378	0.003	0.3819
25	Manuf. of rubber and plastics products	685	17964	0.000	0.4676
26	Manuf. of other non-metallic mineral products	426	21624	0.052	0.5093
27	Manuf. of basic metals	265	34411	0.006	0.3475
28	Manuf. of fabricated metal products, except machinery and equipment	2276	8041	0.001	0.4174
29	Manuf. of machinery and equipment n.e.c.	987	19343	0.040	0.4374
30	Manuf. of office, accounting and computing machinery	23	39850	-	-
31	Manuf. of electrical machinery and apparatus n.e.c.	357	26740	0.000	0.4216
32	Manuf. of radio, television and communication equipment and apparatus	218	25159	0.000	0.7194
33	Manuf. of medical, precision and optical instruments, watches and clocks	354	12452	0.008	0.3988
34	Manuf. of motor vehicles, trailers and semi-trailers	280	49195	0.022	0.5796
35	Manuf. of other transport equipment	137	68192	0.000	0.7149
36	Manuf. of furniture; manufacturing n.e.c.	546	14411	0.000	0.3870

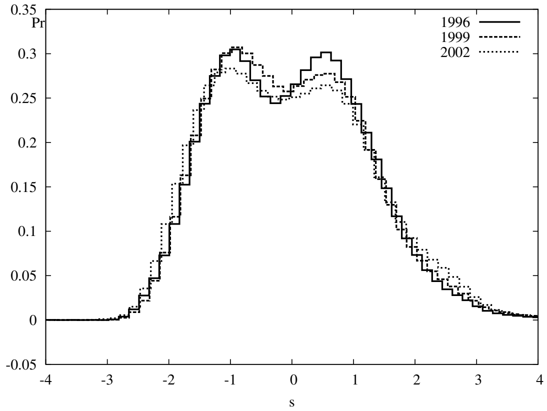

Figure 5: Size distribution of ISIC sector 18 (Manuf. of wearing apparel, dressing and dyeing of fur)

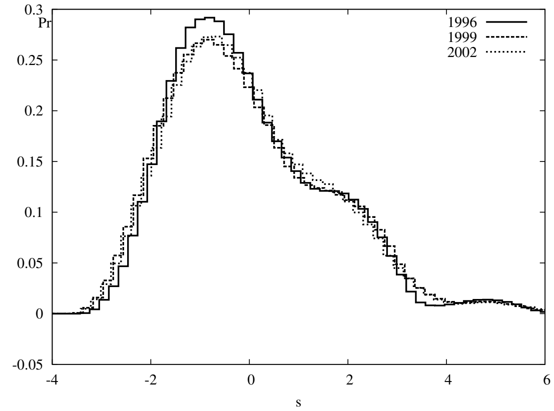

Figure 6: Size distribution of ISIC sector 35 (Manuf. of other transport equipment)

6 present some kernel density plots of exemplary sectors, that have been chosen to highlight inter-sectoral diversity (subsequent tests reveal these sectors to be significantly multimodal).

In an attempt to quantify this inter-sectoral heterogeneity, we will use the non-parametric multimodality test presented in Silverman (1981), which is constructed as follows. Consider a dataset made of n observations independently drawn from a common density f. Suppose that we wish to test the null hypothesis that the density f possesses at most k modes against the alternative that the same f possesses more than k modes. First, we need to compute the 'critical value' h^* for the bandwidth parameter, defined as the largest value of the parameter h that guarantees a kernel density estimate $\hat{f}(h^*)$ as defined in equation (2) with at least k modes. This definition is meaningful since the number of modes is a decreasing function of the bandwidth parameter: for $h > h^*$ the formula in equation (2) would give an estimated density with less than k modes while for $h \leq h^*$ the estimated density would have at least k modes.[9] Note also that, as the sample size n tends to infinity, h^* will tend to zero if the distribution is unimodal, but will be bounded away from zero otherwise.

Second, once the value h^* has been found, we need to assess its significance. Assuming known the true density f, one can repeatedly draw n observations from the true density f and count the modes of the kernel density estimate $\hat{f}(h_0)$ obtained from these observations. The fraction of times in which these modes are greater than k is an estimate of the p-value associated with h^*. The problem of this method is that, in general, the underlying true density is not known. Silverman (1981) suggests the natural candidate density function to use in the simulations is a rescaled version of $\hat{f}(s; h_0)$, derived from data equating the variance of $\hat{f}$ with the sample variance. Hall and York (2001) show that this choice is biased towards conservatism and propose an improved procedure to achieve asymptotic accuracy. Following their suggestion, we compute in each year the critical bandwidth h^* and the p-values of the test where the null is 'the (log) size distribution is unimodal' and the alternative is 'the (log) size distribution presents more than one mode'.

Column 5 in Table 4 reports the results of the bimodality tests (at the 5% significance level for the Hall-York procedure). Unimodality can be rejected in an overwhelming 18 out of 20 sectors, if we look at the 5% significance level. We conclude that the rather 'regular' shape of the aggregate size distribution does not hold at a finer, sectoral level of analysis, and is

[9]The result has been proved for a small family of kernels of which the Gaussian kernel is a member. See Silverman (1981).

ISIC class	1998				2000				2002			
	β	Std. Error	ρ	Std. Error	β	Std. Error	ρ	Std. Error	β	Std. Error	ρ	Std. Error
17	0.9962	0.0042	0.0529	0.0190	0.9980	0.0047	0.0148	0.0232	1.0036	0.0042	-0.0902	0.0202
18	0.9878	0.0042	-0.0736	0.0212	**1.0129**	0.0050	-0.1005	0.0296	1.0073	0.0049	0.0063	0.0254
19	0.9956	0.0055	-0.5304	0.0298	1.0070	0.0074	0.0001	0.0348	0.9931	0.0078	0.0085	0.0397
20	1.0090	0.0066	0.0570	0.0274	1.0028	0.0068	0.0299	0.0380	1.0065	0.0071	-0.0317	0.0423
21	0.9951	0.0039	-0.0439	0.0264	1.0140	0.0046	0.0273	0.0287	0.9980	0.0041	-0.1433	0.0246
22	0.9933	0.0041	-0.1215	0.0169	0.9987	0.0039	-0.0342	0.0178	0.9971	0.0037	-0.0088	0.0178
24	**0.9919**	0.0040	-0.0124	0.0244	1.0076	0.0034	-0.1081	0.0216	0.9939	0.0048	0.0402	0.0284
25	**0.9847**	0.0046	-0.0391	0.0221	1.0002	0.0043	-0.0286	0.0207	1.0018	0.0047	-0.0082	0.0241
26	0.9954	0.0042	-0.0174	0.0247	1.0011	0.0048	0.0137	0.0294	0.9969	0.0044	0.0136	0.0244
27	1.0002	0.0043	-0.0999	0.0273	**1.0150**	0.0044	-0.1318	0.0490	1.0029	0.0060	0.0990	0.0454
28	**0.9925**	0.0032	-0.1006	0.0120	1.0010	0.0030	-0.0919	0.0123	1.0055	0.0029	-0.1161	0.0136
29	0.9985	0.0036	-0.1325	0.0178	1.0009	0.0044	-0.0021	0.0190	**1.0092**	0.0039	-0.1220	0.0203
31	**0.9870**	0.0047	-0.0073	0.0340	0.9997	0.0050	-0.1395	0.0367	1.0017	0.0053	-0.0433	0.0296
32	0.9917	0.0079	-0.0279	0.0324	**1.0207**	0.0097	-0.0564	0.0504	0.9897	0.0125	0.0366	0.0560
33	0.9946	0.0075	-0.1082	0.0349	**1.0199**	0.0079	-0.1176	0.0338	1.0044	0.0081	-0.1275	0.0495
34	0.9987	0.0052	-0.1931	0.0359	**0.9988**	0.0058	0.0182	0.0398	1.0053	0.0064	-0.0932	0.0426
35	1.0141	0.0096	-0.0822	0.0650	0.9944	0.0080	-0.1358	0.0565	**0.9804**	0.0069	-0.0572	0.0563
36	1.0030	0.0047	-0.1426	0.0231	1.0086	0.0046	-0.0552	0.0290	1.0010	0.0046	-0.0378	0.0298

Table 5: Sectoral analysis: Estimation of the Gibrat Law coefficients, using Chesher's (1979) procedure (estimation of equation (5) using LAD). Estimates of β which significantly differ from 1 (5% significance level) are reported in bold.

primarily a result of statistical aggregation. This finding is in line with Hymer and Pashigian (1962) on UK data, and more recently with the results of Bottazzi *et al.* (2003) on Italian data and Bottazzi and Secchi (2003) on US data.

4.1.2 Sectoral concentration

Another way of comparing the size distributions of the different sectors is by looking at the upper tail of the distribution. We do this by calculating the concentration statistics. This can be done by using data on the upper tail of the distribution. Although we do not have reliable information on the market share of the largest firms (because the dataset is incomplete in the sense that it excludes firms with less than 20 employees, we can nonetheless investigate sectoral concentration using the following concentration index:

$$d^4_{20}(t) = \frac{C_4}{C_{20}} \qquad t = 1996, \ldots, 2002 \tag{9}$$

where C_4 and C_{20} are the sums of the market shares of the top 4 and top 20 firms in a sector, respectively. It is trivial to see that this simplifies to the ratio between the combined sales of the largest 4 and largest 20 firms in a sector. Notice that the possible values range from 0.2 (i.e. many firms of equal size) to 1.0 (i.e 4 firms totally dominate the sector), with higher values of d^4_{20} for more concentrated sectors. In order to obtain a more robust indicator of sectoral concentration, we take the average value of d^4_{20} over the 7 years from 1996-2002:

$$D^4_{20} = \frac{1}{7} \sum_{t=1996}^{2002} d^4_{20}(t) \ . \tag{10}$$

The values of D^4_{20} have been calculated and reported in column 6 of Table 4. Whilst the support of possible values ranges from 0.2 to 1.0, we observe that the sectoral concentration indices vary greatly from 0.33 to 0.72. This provides further evidence of heterogeneity of the firm-size distribution across sectors. However, we do not observe any close relationship between the average firm size of a sector and the concentration index in (9).

4.1.3 Gibrat's Law - autoregression of size

Using a similar methodology to that described above, we extend our investigation of Gibrat's Law to the sectoral level. We perform Chesher's (1979) calculations, equations (4) and (5), and report the results in Table 5. Again, we observe heterogeneity as the sectoral results fluctuate around the values obtained in the aggregate analysis. Generally speaking, the β values are close to the Gibrat value of 1, whilst the ρ values (which carry information on growth rate autocorrelation) are mostly negative, though often not statistically significant. Indeed, in the three different years reported in Table 5 only in few cases (4 in 1998 and 2000 and only 2 in 2002) the estimated β coefficient is significantly different from 1. Moreover, it worths to note that in none of the sectors we observe deviations from the unit root hypothesis in more than one year probably suggesting that these deviations do not entangle any economic interpretation.

4.2 Distribution of growth rates

The methodology presented in section 3.2 is now extended to the disaggregated level. All sectors have growth rates distributions that are particularly fat-tailed, although we observe heterogeneity between sectors.

	1998				**2000**				**2002**			
ISIC class	**Scaling**	Std. Error	b	Std. Error	**Scaling**	Std. Error	b	Std. Error	**Scaling**	Std. Error	b	Std. Error
17	-0.0330	0.0180	0.931	2.506‰	-0.0854	0.0161	0.937	2.509‰	-0.0642	0.0149	0.791	2.426‰
18	0.0315	0.0174	0.849	3.605‰	0.0441	0.0169	0.717	3.493‰	0.0540	0.0176	0.697	3.476‰
19	-0.0429	0.0296	0.757	8.568‰	-0.1018	0.0286	0.892	8.845‰	-0.1216	0.0295	0.860	8.780‰
20	-0.0547	0.0359	0.990	5.903‰	-0.1624	0.0346	0.786	5.633‰	-0.1700	0.0302	0.720	5.544‰
21	-0.1845	0.0226	0.705	4.764‰	0.0062	0.0233	1.078	5.189‰	-0.2158	0.0198	0.685	4.741‰
22	-0.1385	0.0156	0.578	2.048‰	-0.1116	0.0154	0.633	2.077‰	-0.1708	0.0157	0.696	2.110‰
24	-0.1607	0.0138	0.606	3.411‰	-0.1005	0.0146	0.813	3.589‰	-0.1556	0.0128	0.673	3.469‰
25	-0.1728	0.0183	0.719	2.541‰	-0.1196	0.0186	0.805	2.594‰	-0.1786	0.0174	0.811	2.597‰
26	-0.1043	0.0195	0.837	4.203‰	-0.1082	0.0216	0.718	4.085‰	-0.1627	0.0181	0.801	4.167‰
27	-0.0413	0.0197	0.876	6.817‰	0.0247	0.0207	1.154	7.242‰	-0.1106	0.0217	0.949	6.930‰
28	-0.0190	0.0125	0.813	0.782‰	-0.0921	0.0122	0.860	0.791‰	-0.0602	0.0113	0.921	0.802‰
29	-0.0515	0.0140	0.818	1.806‰	-0.0492	0.0127	0.776	1.788‰	0.0292	0.0133	0.880	1.832‰
31	-0.0832	0.0169	0.862	5.044‰	-0.2109	0.0192	1.003	5.207‰	-0.0736	0.0204	1.067	5.278‰
32	0.0077	0.0296	1.141	8.779‰	0.0831	0.0296	1.163	8.818‰	0.1090	0.0256	1.000	8.520‰
33	-0.1029	0.0276	0.738	4.940‰	-0.0446	0.0215	0.683	4.873‰	0.0251	0.0203	0.851	5.074‰
34	-0.0471	0.0216	1.011	6.649‰	-0.1462	0.0219	0.805	6.346‰	-0.0699	0.0229	0.797	6.334‰
35	-0.1455	0.0315	0.710	12.675‰	0.0473	0.0308	0.859	13.136‰	0.0006	0.0275	0.791	12.927‰
36	-0.0722	0.0210	0.769	3.227‰	-0.1312	0.0194	0.660	3.141‰	-0.0617	0.0196	0.919	3.341‰

Table 6: Sectoral analysis: Scaling coefficients (relation between size and growth rate variance) and estimated Subbotin b parameters (with coefficients of variation).

We estimate the parameters of the sectoral growth rates distribution as follows. To begin with, we investigate the possibility of a relationship between growth rate variance and size, and correct for such 'scaling effects'. The results are reported in Table 6. We observe that scaling effects are not significant in each sector. We then estimate the subbotin distribution b parameters on the basis of these rescaled error terms. These values are also shown in Table 6. Again, sectoral-level heterogeneity is observed, with most of the values being smaller than the Laplace value of 1.00.

How can we account for the differences in growth rate profiles for different sectors? There appears to be no relation between the growth rate distribution coefficients and average firm size. Also, distinguishing between upstream and downstream sectors does not help us to better understand differences in growth rate distributions (results not shown). Furthermore, grouping the sectors according to a Pavitt-type taxonomy of industries (Pavitt, 1984; see also Marsili, 2001) does not help to explain the differences in the estimated coefficients. A deeper understanding of the economic significance of growth rate distribution coefficients is clearly warranted.

5 Conclusion

In this study we have investigated some of the key quantities of the structure and dynamics of the French manufacturing industry, using an extensive longitudinal database for the period 1996-2002. We examined the size distribution, Gibrat's Law, the growth rates distribution, and growth rate autocorrelation at both an aggregate and disaggregate level. Our findings corroborate well-known stylized facts already observed with Italian and US data, but they also highlight some particularities of the French manufacturing industry.

Gibrat's Law appears to be a useful summary metric , although technically speaking it does not appear to hold for our database. Growth rate autocorrelation is observed to be negative and statistically significant (although rather small in practical terms), and this leads us to reject the proposition that growth is independent of size (Chesher, 1979). Another main finding is the peculiar shape of the growth rate distribution of French manufacturing firms. Whilst the Laplace distribution of growth rates was repeatedly found in previous studies and appeared to be emerging as something of a 'stylized fact', we observe here that the growth rates of French firms are even fatter tailed than expected, a property which holds with disaggregation. The variance of these growth rates decreases with size, which corroborates many, though not all, previous findings.

It is of interest to contrast the growth rate distribution with the size distribution. Whilst we observe that the former is a very robust property of industrial dynamics, the same cannot be said of the size distribution, which is fairly ordered at the aggregate level but quite disorganized as we move down to analyze individual sectors. Of course, there is a strong link between growth rates and the resulting size. This tension between the two serves to emphasize that firm size is not only due to growth rates but to the *initial size distribution* - size and date of entry, and also due to mergers and acquisitions. These factors are outside the scope of this study, although their effect on economic *dynamics* does not appear to be so important.

Although we de-emphasize the need to explain the aggregate firm size distribution, it seems that the distribution of growth rates is a subject ripe for future investigation. The analysis of growth rates presented here gives us important insights into the competitive process, emphasizing the importance of extreme growth events in the French manufacturing industry. However, we had difficulty in finding a connection between the growth rate distribution coefficients and other economic characteristics. For example, at a sectoral level, there appears to be no relation between the distribution parameters and average firm size. Also, our dataset suggests that there is no relationship between the distribution parameters and the distinction between upstream and downstream sectors. In addition, variation in the growth rate distribution coefficients does not seem to correspond to a Pavitt-type taxonomy (Pavitt, 1984) of industrial sectors. Mapping the growth rate distribution coefficients to economic concepts would merit further work. Furthermore, this paper provides results that would be useful in the context of a more detailed international comparison.

6 Acknowledgments

We thank Giovanni Dosi for helpful comments at various stages of this research. This work would not have been possible without data provided by the INSEE. Precious help from Danièle Bastide at the SESSI is gratefully acknowledged. The usual caveat applies.

References

Amaral, L.A.N., S.V. Buldyrev, S. Havlin, M.A. Salinger, H.E. Stanley, M.H.R. Stanley, (1997) *Scaling behavior in economics: the problem of quantifying company growth*, Physica A, 244, 1-24

Bottazzi, G., E. Cefis, G. Dosi, A. Secchi, (2003) *Invariances and Diversities in the Patterns of Industrial Evolution: Some Evidence from Italian Manufacturing Industries*, Pisa, Sant'Anna School of Advanced Studies, LEM Working Paper Series 2003-21

Bottazzi, G., G. Dosi, M. Lippi, F. Pammolli, M. Riccaboni, (2001) *Innovation and Corporate Growth in the Evolution of the Drug Industry*, International Journal of Industrial Organization, 19, 1161-1187

Bottazzi, G., A. Secchi, (2003) *Common Properties and Sectoral Specificities in the Dynamics of U. S. Manufacturing Companies*, Review of Industrial Organization, 23, pp. 217-232

Bottazzi, G., A. Secchi, (2005) *Growth and Diversification Patterns of the Worldwide Pharmaceutical Industry*, Review of Industrial Organization, 26, 195-216

Bottazzi, G., A. Secchi, (2006) *Explaining the Distribution of Firms Growth Rates*, The RAND Journal of Economics, forthcoming.

Chesher, A., (1979), *Testing the Law of Proportionate Effect*, The Journal of Industrial Economics, Vol. 27 (4), June, pp. 403-411

Dosi, G., (2005), *Statistical Regularities in the Evolution of Industries: A Guide through some Evidence and Challenges for the Theory*, Pisa, Sant'Anna School of Advanced Studies, LEM Working Paper Series, 2005/17

Dosi, G., O. Marsili, L. Orsenigo, R. Salvatore, (1995), *Learning, Market Selection and the Evolution of Industrial Structures*, Small Business Economics, 7 (6), 411-436

Dunne, T., M. J. Roberts, and L. Samuelson, (1988) *The Growth and Failure of U.S. Manufacturing Plants*, Quarterly Journal of Economics, 104, 671-698

Evans, D. S., (1987) *The Relationship between Firm Growth, Size and Age: Estimates for 100 Manufacturing Industries*, Journal of Industrial Economics, 35, 567-581

Gibrat, R., (1931) *Les inégalités économiques*, Paris: Librairie du Receuil Sirey

Hall, B., (1987) *The Relationship between Firm Size and Firm Growth in the U.S. Manufacturing Sector*, Journal of Industrial Economics, 35 (4), 583-600

Hall, P., and M. York, (2001) *On the Calibration of Silverman's Test for Multimodality*, Statistica Sinica, 11, 515-536

Hart, P. E., (1962) *The Size and Growth of Firms*, Economica, 29, 29-39

Hart, P. E., (1965) *Studies in Profit, Business Saving and Investment in the United Kingdom, 1920-1962*, Volume 1, London: Allen and Unwin

Hart, P. E., S. J. Prais, (1956) *The Analysis of Business Concentration: A Statistical Approach*, Journal of the Royal Statistical Society, 119, 150-191

Hymer, S., P. Pashigian, (1962) *Firm Size and Rate of Growth*, Journal of Political Economy, 70 (6), 556-569

Ijiri, Y., H. A. Simon, (1967) *A Model of Business Firm Growth*, Econometrica, 35 (2), 348-355

Ijiri, Y., H. A. Simon, (1977) *Skew Distributions and the Sizes of Business Firms*, Amsterdam: North Holland

MacKinnon J. G.,H. White, (1985), *Some Heteroskedasticity-consistent Covariance Matrix Estimators with Improved Finite Sample Properties*, Journal of Econometrics, 29, 305-325

Mansfield, E., (1962) *Entry, Gibrat's law, innovation and the growth of firms*, American Economic Review, 52, 1023-1051

Marsili, O., (2001) *The Anatomy and Evolution of Industries*, Edward Elgar: Cheltenham

Pavitt, K., (1984) *Sectoral patterns of technical changes: towards a taxonomy and a theory*, Research Policy, 13, 343-375

Prais, S. J., (1974) *A New Look at the Growth of Industrial Concentration*, Oxford Economic Papers, 26, 273-288

Quandt, R. E., (1966) *On the Size Distribution of Firms*, American Economic Review, 56, 416-432

Silverman, B. W., (1981) *Using Kernel Density Estimates to Investigate Multimodality*, Journal of the Royal Statistical Society, Series B, 43, 97-99

Silverman, B. W., (1986) *Density Estimation for Statistics and Data Analysis*, London: Chapman and Hall.

Simon, H. A., C. P. Bonini, (1958) *The Size Distribution of Business Firms*, American Economic Review, 48, 607-617

Singh, A., G. Whittington, (1975) *The Size and Growth of Firms*, Review of Economic Studies, 42 (1), 15-26

Steindl, J., (1965) *Random Processes and the Growth of Firms*, London: Griffin

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