Small Business Economics (2007) 29:137–159
DOI 10.1007/s11187-006-0014-y

Invariances and Diversities in the Patterns of Industrial Evolution: Some Evidence from Italian Manufacturing Industries

G. Bottazzi
E. Cefis
G. Dosi
A. Secchi

ABSTRACT. This work explores and compares some basic properties of corporate growth process at both aggregate industry level and disaggregated sectoral levels. Using an extensive dataset on Italian manufacturing firms, we investigate which properties of firm growth dynamics are robust under disaggregation. We compare the results obtained with three different definitions of firm size, namely total sales, number of employees and value added. Our analysis suggests that while different sectors are characterized by significant differences in firm size distributions, in the degrees of concentration and in the autoregressive structure of the growth processes, there are also regularities which hold across all of them, such as the approximate unit root nature of the growth process and the power exponential shape of the growth rates density. Together, these "stylized facts" suggest challenging puzzles on the drivers of corporate growth and the resulting industrial structures.

KEY WORDS: aggregation, concentration, firm growth, size distribution

JEL CODES: C1, D2, L1

1. Introduction

In this work we explore some basic properties of the size distribution of Italian manufacturing firms and of their growth processes both at the level of the whole industry and, separately, in each three digit sector of the ISIC classification.

Our analysis of firm size distributions links with that stream of investigations highlighting the widespread occurrence of skewed distributions involving the coexistence of firms of widely different sizes. Within a vast literature, see Hart and Prais (1956), Simon and Bonini (1958), Steindl (1965), Quandt (1966), Ijiri and Simon (1977) and more recently Stanley et al. (1995), Axtell (2001), Cabral and Mata (2003) and for an extensive critical review Kleiber and Kotz (2003). A general finding in the literature is that, at least at the level of aggregation of the whole manufacturing industry, distributions appear to persistently display Zipf (or more generally Pareto) power-law profiles. However, a common limitation of most of these studies is precisely their high level of aggregation. Furthermore no attention is generally devoted to the influence upon the results of the proxy chosen for size. A first question that we shall address is indeed the sensitivity of firm size distributions to different measures of size and different levels of aggregation. In this paper, using an extensive database on medium/large firms, we refine on previous studies on firm size distribution considering explicitly three different size proxies: Total Sales, Number of Employees and Value Added. Together we investigate which properties of firm size distributions and growth dynamics are robust under disaggregation, i.e. do show the same character when analyzed at the coarse level of the whole manufacturing industry and at the finer sectoral level. Finally, at the disaggregated level, we try to identify those features which are generic and hold across all or most of the considered sectors distinguishing them from sector-specific ones.

Final version accepted on February 27, 2006.

G. Bottazzi
Scuola Superiore Sant' Anna
Piazza Martiri della Liberta 33, 56127, Pisa, Italy
E-mail: bottazzi@sssup.it

E. Cefis
University of Bergamo, Italy and USE, Utrecht University
The Netherlands

G. Dosi · A. Secchi
Scuola Superiore Sant' Anna
Pisa, Italy

We find that size distributions are characterized by different shapes and possess different tail behaviors in the different sectors under investigation. Such differences forbid any parametric fitting based on densities families depending on a reasonably small number of parameters. Hence, in order to give a quantitative account of such inter-sectoral differences, we revert to two statistics derived from the empirical distributions, namely the number of modes and a concentration index based on the largest quantiles. These two statistics, even if not fully characterizing the densities, provide information on their behaviors in their central part and their upper tail, respectively. We show that several sectors possess multimodal densities, while others are characterized by unimodal, bell-shaped ones. Moreover, the analysis of market concentration reveals that the differences in the upper-tails of sectoral size distributions are persistent and do not change significantly in the time window under analysis. While the limited time span of our database does not allow the investigation of the dynamic origin of the different observed shapes, our analysis clearly supports a notion of structural diversity across industrial sectors, in tune with the old intuition of ''structuralist'' approaches to industrial analysis from the '50 and '60 (Bain, 1956).

A second question concerns the underlying firm growth processes including, in particular, the relationship between firm size and growth dynamics and the distribution of growth rates themselves.

The literature on the former subject, since the earliest studies, has mostly tackled the validity of the so called Law of Proportionate Effects (Gibrat, 1931) postulating the independence between the size of the firm and its logarithmic growth rate: see, for instance, Hymer and Pashigian (1962), Mansfield (1962), Kumar (1985), Evans (1987), Hall (1987), Dunne et al. (1988), Wagner (1992), Dunne and Hughes (1994), Audretsch et al. (1999) and the reviews in Sutton (1997) and Lotti et al. (2003).

In this work, following Chesher (1979), we analyze the temporal structure of size dynamics through the estimation of autoregressive models on the size time series. Our results suggest a general validity of the Gibrat's hypothesis at both aggregate and disaggregated levels, for the three size proxies considered and across the different sectors. That is, we do not find any substantial, negative or positive, correlation between growth rates and size. Conversely, the degree of persistency of firm's growth seems to depend strongly on both the sector considered and the definition used for size.

Next, using a parametric method based on the estimation of a flexible class of probability densities, the Subbotin family (originally presented in Bottazzi et al. 2002), we characterize and compare the empirical densities of growth rates at sectoral level and for the whole manufacturing industry. Our analysis reveals a high homogeneity in the growth rates densities, which commonly display power exponential shapes, often well approximated by a Laplace distribution.

The paper is organized as follows. After a brief description of the data (Section 2), we study via non-parametric methods the empirical size distributions, in the aggregate and by sectors (Section 3). Section 4 introduces the analysis of growth processes, testing the Gibrat's hypothesis and investigating the degrees of persistency in growth rates. In Section 5 we analyse the growth rates distributions, applying both parametric and non-parametric methods. Finally, Section 6 summarizes our findings, compare them with previous contributions highlights some interpretative puzzles and sketches a few future directions of research.

2. Data description

This research draws upon the MICRO.1 databank developed by the Italian Statistical Office (ISTAT).[1] This databank contains the result of an annual Industrial Census conducted in all industrial sectors during the period 1989–1997 on firms with 20 employees or more. Data on the Profit and Loss Account and the Balance Sheet are collected and organized along the normative guidelines of the *IV* Directive of the European Community Council.

The response rate of the census is, during the entire period, between 50% and 60% on a total population of firms with at least 20 employees approximately equal to 70,000. In this paper we restrict our analysis to the period 1989–1996 and only to the manufacturing industry (with ISIC code between 151 and 366) for which, according to Bartelsman et al. (2004), firms with at least 20 employees represent, in Italy, more than 68% of the total employment. Moreover, for statistical reliability, we further restrict our analysis to sectors with more than 44 firms[2] reducing the total number of sectors under study from 97 to 55. Table I. reports the list of the sectors together with the number of considered firms. Since the panel is open, due to entry, exit and fluctuations around the 20 employees threshold, we consider only those firms that are present both at the beginning and at the end of our window of observation.[3]

In this work we are exclusively interested in the process of *internal* growth, as opposed to the growth due to mergers, acquisitions and divestments. In order to control for these phenomena we build ''super-firms'' which account throughout the period for the union of the entities which undertake such changes. So, for example, if two firms merged at some time, we consider them merged throughout the whole period. Conversely, if a firm is spun off from another one, we ''re-merge'' them following the separation event. After the application of this procedure, we end up with a sample of 8091 ''super-firms'' observed for 7 years.

3. Size distributions

The analysis of firm size distributions is organized in two steps. First, using kernel estimates, we investigate the shape of firm (log) size distributions in the manufacturing aggregate and at sectoral level, discuss their stationarity and compare the results obtained considering different measures of size, namely Total Sales, Number of Employees and Value Added.

Second, in order to highlight the differences in shapes among different sectors, we use two derived statistics, the number of modes and a sectoral concentration index. Notice that both these measures are unaffected by a mere shift of the distribution: hence they capture those aspects of the density behavior which are not due to a simple change of scale but rather to a different relative distribution of ''shares'' of the overall production, however measured, among the different firms in each sector. Notice that these statistics are built using different regions of the empirical distribution: the central part for the number of modes and the upper quantiles for the concentration index. We have on purpose ignored the lower tail of the distributions since this is the region which is more likely affected by the 20 employees threshold in the database.

3.1. *Firms size distributions*

We start with a non-parametric exploration of the shape of the aggregate size distribution over the whole manufacturing industry. To this purpose, we build a kernel estimate (Silverman, 1986) of the probability density of (log) firm size for any period t. Let $S^{\mathrm{TS}}_{i,j}(t), S^{\mathrm{L}}_{i,j}(t)$ and $S^{\mathrm{VA}}_{i,j}(t)$ stand for size of firm i in sector j at time t, respectively in terms of Total Sales (TS), Number of Employees (L) and Value Added (VA). The index $j \in \{1, \ldots, 55\}$ runs over the sector under analysis.

The kernel density estimate is defined according to

$$\hat{f}(y,t;h)=\frac{1}{nh}\sum_{i=1}^{n}K\left(\frac{y-\log S_i^X(t)}{h}\right)\; X\in\{\mathrm{TS},L,\mathrm{VA}\}, \tag{1}$$

where h is a bandwidth parameter controlling the degree of smoothness of the estimate and K is a kernel density, i.e. $K(y) \geq 0, \forall y \in (-\infty, +\infty)$ and $\int \mathrm{d}y\, K(y) = 1$. The kernel density estimate can be considered a smoothed version of the histogram obtained counting the observations in different bins. In the same way as the histogram depends on the number of considered bins, the kernel estimate depends on the specification of two objects, namely the kernel K and the bandwidth h. Throughout this paper the kernel function will always be the Gaussian density; however the use of different kernels, such as the Epanenchnikov 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 1, Figure 3 and Figure 5 report estimated densities in 4 different years of $\log(S^{\mathrm{TS}}_{\mathrm{i}}(t))$, $\log(S^{\mathrm{L}}_{i}(t))$ and $\log(\mathrm{S}^{\mathrm{VA}}_{\mathrm{i}}(\mathrm{t}))$ of firms aggregated over all the 55 sectors. Visual

TABLE I
Definition and number of firms of the 55 sectors under analysis

ISIC code	Sector	Number of firms
151	Production, processing and preserving of meat	114
155	Dairy products	85
158	Production of other foodstuffs (brad, sugar, etc...)	157
159	Production of beverages (alcoholic and not)	94
171	Preparation and spinning of textiles	154
172	Textiles weaving	171
173	Finishing of textiles	181
175	Carpets, rugs and other textiles	90
177	Knitted and crocheted articles	162
182	Wearing apparel	379
191	Tanning and dressing of leather	87
193	Footwear	245
202	Production of plywood and panels	52
203	Wood products for construction	59
205	Production of other wood products (cork, straw, etc...)	56
211	Pulp, paper and paperboard	46
212	Articles of paper and paperboard	180
221	Publishing	72
222	Printing	199
241	Production of basic chemicals	80
243	Paints, varnishes, printing inks and mastics	58
244	Pharmaceuticals, medicinal chemicals and botanical products	97
245	Soap and detergents, cleaning and toilet preparations	46
246	Other chemical products	51
251	Rubber products	87
252	Plastic products	352
261	Glass and glass products	87
262	Ceramic goods not for construction	59
263	Ceramic goods for construction	91
264	Bricks, tiles and construction products in baked clay	84
266	Articles in concrete, plaster and cement	141
267	Cutting, shaping and finishing of stone	69
273	First processing of iron and steel	82
275	Casting of metals	125
281	Structural metal products	156
284	Forging, pressing, stamping and roll forming of metal	132
285	Treatment and coating of metals	182
286	Cutlery, tools and general hardware	149
287	Other fabricated metal products	265
291	Machinery for the production and the use of mechanical power	224
292	Other general purpose machinery	199
293	Agricultural and forestry machinery	54
294	Machine tools	114
295	Other special purpose machinery	424
297	Domestic appliances not elsewhere classified	59
311	Electric motors, generators and transformers	71
312	Manufacture of electricity distribution and control equipment	70
316	Electrical equipment not elsewhere classified	91
322	TV and radio transmitters and lines for telephony and telegraphy	44
332	Measure, control and navigation instruments	51
342	Production of bodies for cars, trailers and semi-trailers	50
343	Production of spare parts and accessories for cars	125
361	Furniture	444
362	Jewelry and related articles	84

TABLE I
Continued

ISIC code	Sector	Number of Firms
366	Miscellaneous manufacturing not elsewhere classified	68
Aggregate		8091

inspection reveals that all these densities display considerable stationarity over our window of observation for all the size proxies considered. This property is confirmed when we consider the time evolution of their first 4 moments, over the whole time range 1989–1996, reported in Figures 2, 4 and 6. Apart from an upward trend in the mean log size (in fact displayed only by Total Sales and Value Added, as one should expect due to the presence of both real growth and nominal trend in prices), all the moments display very stable time profiles.

Concerning the shape in the central part, Figures 1, 3 and 5 show that all the aggregate size densities are characterized by a right-skewed bell-shape.[4] The choice of the size variable, however, influences the shape of the upper tail. When Total Sales are considered, one observes a monotonic decreasing behavior in the largest quantiles. This monotonic behavior is less clear in the case of Value Added and, when the Number of Employees is chosen as the size proxy, a clear and statistically significant bimodal structure appears (see below and Table II).

The picture looks quite different when size distributions are analyzed at more disaggregated level and each of the 55 3-digit sectors are considered separately. Such an analysis reveals that sectoral size distributions possess clearly different profiles not only in the upper tails but also in their central parts: while some sectors present distributions quite similar to the aggregate ones, others are unimodal, symmetric and almost LogNormal, and yet others are bi-modal or even multi-modal. Figure 7 show size densities for 3 different sectors chosen to highlight the high degree of observed heterogeneity.[5]

Obviously, we cannot report here the estimated density for all the 55 sectors under investigation. Hence, in order to provide a succinct quantitative account of the differences emerging across the different sectors, we have to revert to some derived statistics

3.2. *Multimodality analysis*

Let us start with what is typically considered one of the prominent aspect of a density: the number of modes. Using a non-parametric statistical test based on kernel density estimation, we test, in each sector, the hypothesis that the empirical size distribution possesses a single mode.

Following the methodology described in Silverman (1981), 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 it 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 (1) with at least k modes. This definition is meaningful since the number of modes is a decreasing function of the bandwidth parameter:[6] for $h > h^*$ the formula in (1) would give an estimated density with less than k modes while for $h \leq h^*$ the estimated density would have at least k modes. Second, once the value h^* has been found, we need to establish its significance. Assuming the true density f to be known, one could repeatedly draw n observations from it 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 (1981a, b) suggests as a natural candidate density, from which to simulate, 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 conservatively biased and propose an improved procedure to achieve asymptotic

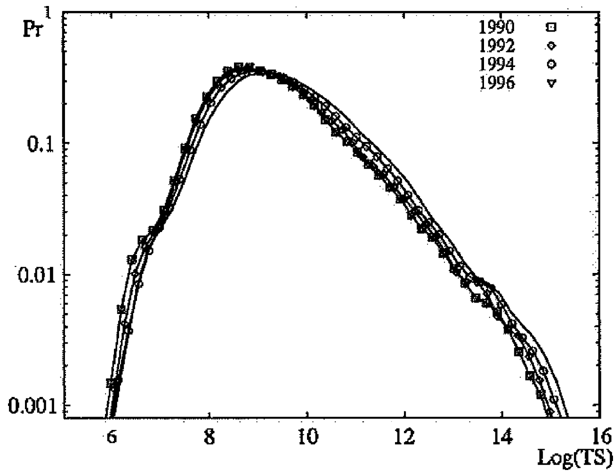

Figure 1. Kernel density estimations of log (S_i^{TS}) in different years. Size measured in terms of Sales (log scal on the y-axis).

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".

Table II reports P-values for this test and shows that in about one fourth of the sectors analyzed we observe size distributions characterized by more than one mode. The size proxy used, however, seems to play an important role in deciding the shape of the central part of the density. Indeed, in 22 sectors out of 55 the unimodality hypothesis can either be rejected or not depending on the proxy used.

3.3. *Sectoral concentration*

As a second step let us consider the behavior of the upper tail of the size distributions in the different sectors. For this purpose we introduce a concentration index based on the size of largest firms in the sector, i.e. the firms contributing to the upper tail of the size distribution itself.

The choice to focus on the upper tail is dictated both by the nature of our database, which does not contain a reliable sample of small firms, and by reasons of comparability with somewhat similar proxies historically used in industrial economics. Indeed, the distribution of market shares among the largest firms

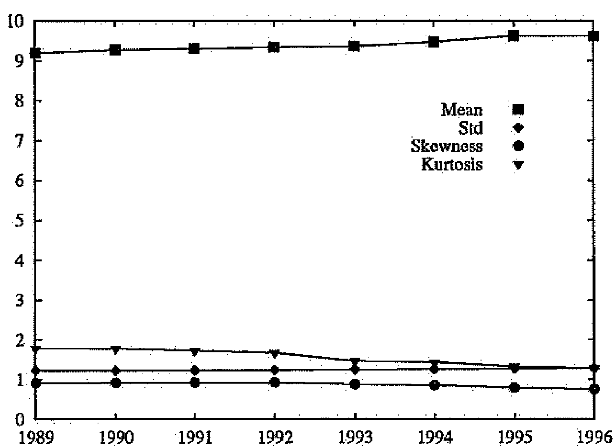

Figure 2. Time evolution (1989–1996) of the first 4 moments of the size (Sales) distribution.

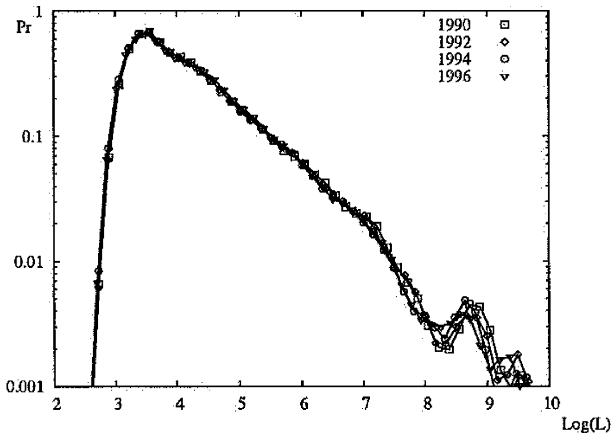

Figure 3. Kernal density estimation of log (S_i^L) in different years. Size measured in terms of Employees (log scale on the y-axis).

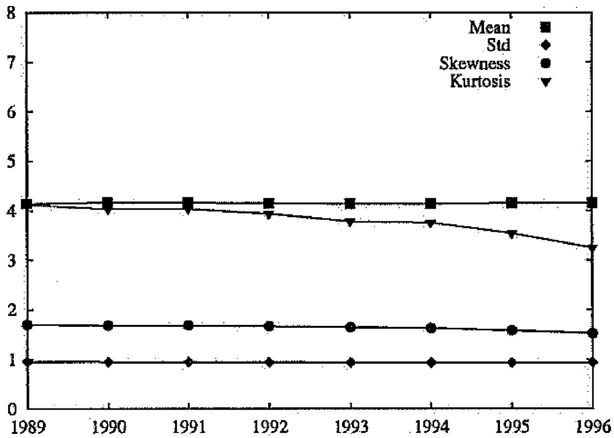

Figure 4. Time evolution (1989–1996) of the first 4 moments of the size (Number of employees) distribution.

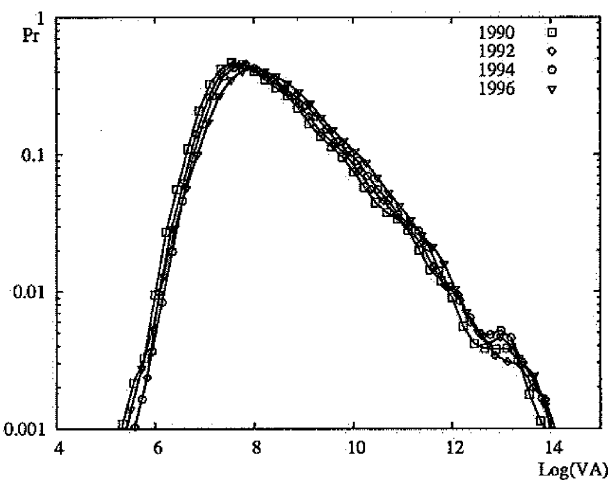

Figure 5. Kernel density estimations of log (S_i^{VA}) in different years. Size measured in terms of Value Added (log scale on the y-axis).

operating in a sector is a natural measure of market concentration and, it is often suggested, an indirect proxy for the strength of market competition.

Let us consider

$$d^4_{20}(t) = \frac{C_4}{C_{20}} \quad t = 1989, \ldots, 1996, \tag{2}$$

where C_4 and C_{20} are the sums of market shares of the top 4 and 20 firms in a sector, respectively. This particular measure of concentration also overcomes the drawbacks stemming from the lack of information about the lower tail of the size in the MICRO.1 database. Indeed if a sector is highly concentrated, d^4_{20} has to be near to 1: in fact, it is easy to verify that if the entire market is shared by less than 5 firms, d^4_{20} is exactly 1. On the contrary, for a sector where at least the first 20 firms have exactly the same size, one gets $d^4_{20} = 0.2$. In order to obtain a more robust characterization of the upper tail of the size distribution we consider the average value of d^4_{20} over all the 8 years (1989–1996) under analysis

$$D^4_{20} = \frac{1}{8}\sum_{t=1}^{8} d^4_{20}(t). \tag{3}$$

The values of D^4_{20} for all the sectors under the three size proxies are reported, in Table II. It is apparent that, in line with the patterns of the sectoral size distributions, this concentration index displays a high degree of heterogeneity across sectors. In order to give a synthetic account of this property consider that, in the case of Total Sales, D^4_{20} ranges from 0.28 in the rubber products sector to 0.86 in the fabricated metal product sector. Furthermore even if one cautiously neglects the highest and lowest 5% observed values, the relevant range is still 0.35–0.79, covering almost completely the notional support of our concentration measure. Similar results emerge if we consider Employees or Value Added as size proxies.

Notice, however, that the estimation of a linear regression between the concentration index D^4_{20} and the sectoral average firm size produce a slope which is *not* significantly different from zero. Putting it in another way, sectoral average firm size, contrary to a good deal of conventional wisdom, does not bear any implication in terms of market concentration and indirectly of the "intensity of competition" of a sector.

To sum up, the evidence on aggregate manufacturing displays a strong right-skewed size distribution, in agreement with findings of many previous contributions (among others Hart and Prais (1956), Ijiri and Simon (1977), Quandt (1966), Simon and Bonini (1958)). However, our results suggest that the upper tail does not robustly decrease linearly with the logarithm of size. Note also that our analysis offers a piece of evidence which goes against Stanley et al. (1995), who suggest that the upper tail of the size distribution should be thinner than a LogNormal. Rather, we observe a profile which is clearly fatter.

The most important result of this section comes, however, from the analysis on sectoral

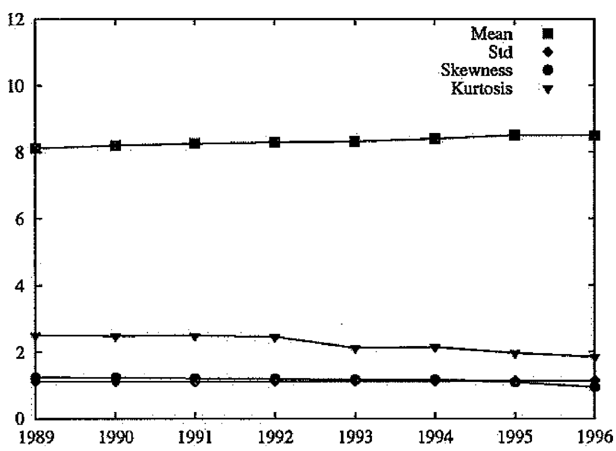

Figure 6. Time evolution (1989–1996) of the first 4 moments of the size (Value Added) distribution.

TABLE II
Average (log) size, Concentration measures D^4_{20} and P-values for the multimodality test at 5% significance level

ISIC Code	Average (log) size			D^4_{20}			Bimodality test P-values ($\alpha = 0.05$)		
	TS	L	VA	TS	L	VA	TS	L	VA
151	10.21	4.02	8.43	0.42	0.45	0.45	0.720	0.062	0.460
155	10.58	4.23	8.66	0.68	0.75	0.77	**0.024**	0.230	0.149
158	10.09	4.32	8.79	0.62	0.54	0.61	0.503	0.290	0.061
159	10.40	4.12	8.76	0.39	0.43	0.40	0.486	0.068	0.077
171	9.44	4.14	8.35	0.38	0.49	0.41	0.082	0.053	**0.001**
172	9.85	4.35	8.67	0.56	0.62	0.56	0.050	**0.016**	**0.018**
173	9.02	4.01	8.28	0.49	0.43	0.52	**0.028**	0.336	0.461
175	9.33	4.05	8.26	0.47	0.43	0.44	**0.031**	0.654	0.475
177	9.09	4.07	8.02	0.50	0.41	0.51	0.172	0.241	0.101
182	8.69	3.99	7.79	0.53	0.55	0.55	0.563	0.164	0.060
191	9.92	3.88	8.20	0.37	0.31	0.38	0.341	0.152	0.410
193	9.12	3.97	7.88	0.50	0.43	0.50	0.569	0.061	0.326
202	9.88	4.16	8.62	0.47	0.52	0.55	0.425	0.180	0.214
203	8.93	3.82	7.94	0.48	0.44	0.47	0.190	0.050	0.687
205	8.88	3.82	7.95	0.39	0.30	0.31	0.619	**0.015**	0.387
211	10.22	4.41	9.14	0.77	0.76	0.76	0.223	0.231	0.242
212	9.81	4.17	8.66	0.55	0.47	0.54	0.578	0.302	0.533
221	10.66	4.74	9.41	0.59	0.56	0.58	0.504	0.315	**0.007**
222	9.04	3.90	8.25	0.62	0.67	0.65	0.257	**0.016**	0.110
241	10.97	4.71	9.67	0.45	0.42	0.44	0.789	0.452	0.762
243	9.98	4.15	8.74	0.50	0.44	0.44	0.367	0.355	0.371
244	11.11	5.14	9.97	0.40	0.35	0.41	**0.035**	**0.020**	0.065
245	10.62	4.59	9.21	0.79	0.75	0.77	0.359	0.199	0.497
246	10.51	4.27	9.19	0.51	0.51	0.48	0.369	0.399	0.837
251	9.67	4.34	8.64	0.77	0.73	0.77	0.503	0.299	0.198
252	9.63	4.08	8.52	0.28	0.30	0.29	0.267	0.065	0.078
261	9.50	4.36	8.67	0.59	0.58	0.56	0.869	0.700	0.072
262	9.07	4.32	8.44	0.65	0.54	0.60	0.384	0.342	0.234
263	10.04	4.62	9.03	0.41	0.44	0.44	0.796	0.438	0.053
264	8.96	3.75	8.02	0.64	0.60	0.57	0.216	**0.036**	0.193
266	9.34	3.89	8.19	0.35	0.38	0.38	0.696	0.088	**0.004**
267	9.01	3.57	7.87	0.56	0.48	0.56	0.085	0.435	0.446
273	10.08	4.05	8.69	0.45	0.33	0.31	**0.034**	0.154	0.204
275	9.57	4.11	8.60	0.74	0.73	0.72	0.213	0.890	0.252
281	9.06	3.81	8.07	0.40	0.40	0.33	0.059	0.134	0.691
284	9.50	4.01	8.55	0.37	0.39	0.40	0.248	**0.042**	0.269
285	8.73	3.66	8.03	0.50	0.30	0.32	0.478	**0.047**	0.339
286	9.23	4.00	8.40	0.38	0.43	0.44	0.498	0.772	0.969
287	9.40	4.06	8.45	0.86	0.82	0.84	0.052	0.107	**0.001**
291	9.91	4.31	8.88	0.70	0.68	0.71	**0.022**	**0.018**	**0.028**
292	9.84	4.24	8.77	0.41	0.46	0.42	0.951	0.313	0.666
293	9.34	3.89	8.24	0.48	0.45	0.46	0.839	0.555	0.226
294	9.52	4.06	8.52	0.69	0.65	0.67	0.208	0.132	**0.017**
295	9.66	4.15	8.59	0.55	0.39	0.51	**0.038**	0.361	0.230
297	10.50	4.99	9.34	0.64	0.62	0.63	0.781	0.538	0.620
311	9.45	4.12	8.42	0.59	0.59	0.60	**0.028**	0.179	0.105
312	10.00	4.61	9.05	0.56	0.52	0.55	0.371	0.150	0.142
316	9.62	4.32	8.68	0.79	0.74	0.81	**0.008**	**0.060**	**0.030**
322	10.44	5.01	9.46	0.80	0.84	0.80	0.250	0.330	0.156
332	9.76	4.41	8.87	0.53	0.56	0.66	0.519	0.069	0.766
342	9.62	4.03	8.40	0.50	0.44	0.43	0.059	0.190	0.684
343	10.10	4.66	9.07	0.62	0.61	0.58	**0.028**	0.038	0.050

TABLE II
Continued

ISIC Code	Average (log) size			D^4_{20}			Bimodality test P-values ($\alpha = 0.05$)		
	TS	L	VA	TS	L	VA	TS	L	VA
361	9.17	3.88	7.99	0.44	0.23	0.34	**0.030**	0.368	0.054
362	9.50	3.85	7.98	0.51	0.44	0.47	0.124	**0.021**	0.179
366	9.28	3.99	8.20	0.35	0.42	0.43	0.750	0.567	0.330
Aggregate	9.61	4.15	8.50	0.59	0.53	0.47	0.471	**0.041**	0.093

The Concentration measure is an average over the period 1989–1997 while the average size and the bimodality test concern the year 1997.

data. Indeed, we are able to provide statistically robust evidence that the characteristic shape observed in the aggregate for the size distribution is mainly an outcome of aggregation, as it appears for manufacturing as a whole but not in most sector-specific distributions (a similar conclusion, based on U.K. data, was reached in Hymer and Pashigian (1962) while recently Bottazzi and Secchi (2003) shows that it applies also to U.S. manufacturing sectors).

At the same time, the profiles of the upper tails of the size distributions, described with the help of a concentration index, turn out to be extremely heterogeneous across the different sectors. In turn such an heterogeneity cannot be simply explained as a sheer "size" effect due to the different dimensions of the "modal" firm in different sectors.

4. Autoregressive profiles in firm sizes

A critical aspect of the investigation of firm sizes concerns their inter-temporal dynamics. We have already shown that the unconditional distribution of sizes is stationary. In this Section we analyze the profile of sizes level of each firm over time considering the autoregressive structure of the size time series at both aggregate and sectoral levels.

In the aggregate, we estimate an AR(1) process on log (S_i^X) where $x \in \{S, L, \mathrm{VA}\}$. For this purpose, in order to eliminate possible trends in the average size, we use the normalized (log) size

$$s_i^X(t) = \log\left(S_i^X(t)\right) - \frac{1}{N}\sum_{i=1}^{N} \log\left(S_i^X(t)\right) \; x \in \{\mathrm{TS}, L, \mathrm{VA}\}, \tag{4}$$

subtracting from the (log) size of each firm the average (log) size of all the firms. (Here N stands for the total number of firms in our panel). On these observations we consider the AR(1) model

$$s_i^X(t) = \beta\, s_i^X(t-1) + \epsilon_i(t). \tag{5}$$

We do not introduce any firm-specific term for the (log) size average, thus implicitly assuming different firms as different realizations of the same stochastic process.[7] Traditionally, in order to obtain reliable estimates for the parameter of model (5) two main econometric issues have been addressed, namely serial correlation and heteroskedasticity of the error terms $\epsilon_i(t)$. To cope with the first source of problems we adopt here the approach described in Chesher (1979). We assume that the stochastic variable $\epsilon_i(t)$ is characterized by an AR(1) structure

$$\epsilon_i(t) = \rho \epsilon_i(t-1) + u_i(t), \tag{6}$$

where $u_i(t)$ are *i.i.d.* disturbances and we rewrite (5) as

$$s_i^X(t) = \gamma_1\, s_i^X(t-1) + \gamma_2\, s_i^X(t-2) + u_i(t), \tag{7}$$

where $\gamma_1 = \beta + \rho$ and $\gamma_2 = -\rho\beta$. We estimate the model in (7) minimizing the Least Absolute Deviation of the errors terms u_i, which is equivalent to a Maximum Likelihood estimation under the hypothesis that the residuals are Laplace (or Double Exponential) distributed. As we will see later, this represent a much better approximation than the Gaussian distribution implied by an Ordinary Least Square estimate.[8] Furthermore we estimate the elements of the covariance matrix using a robust "sandwich"

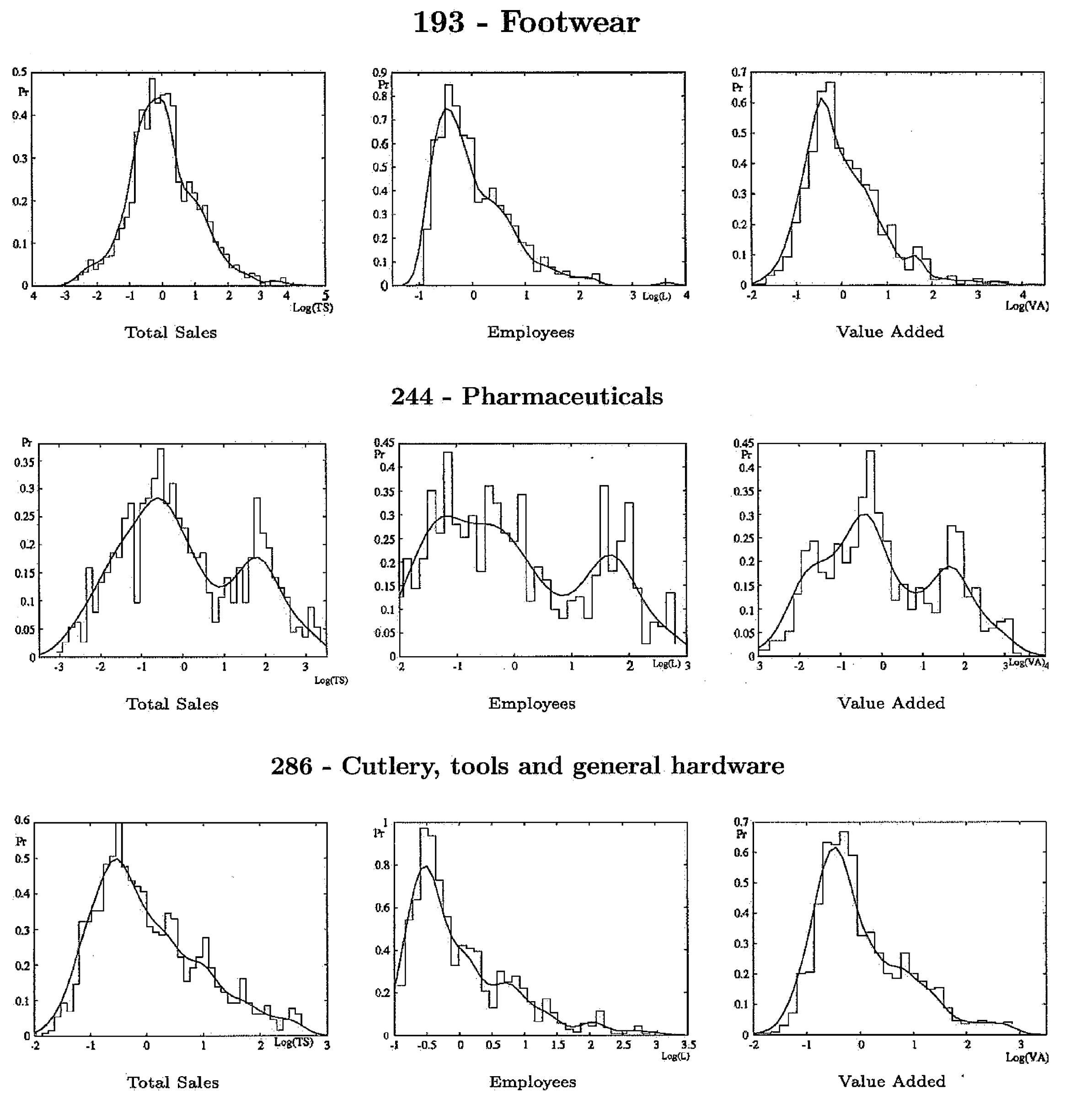

Figure 7. Binned probability density and kernel estimation of log (S_i^{TS}), log (S_i^{L}), log (S_i^{VA}) in three different manufacturing sectors. The kernel function used is the Gaussian density.

estimator (cfr. Huber (1967)). Finally the parameters β and ρ are identified, according to Chesher (1979) through

$$\beta = \frac{1}{2}\left[\gamma_1 + \sqrt{\gamma_1^2 + 4\gamma_2}\right] \quad \rho = \frac{1}{2}\left[\gamma_1 - \sqrt{\gamma_1^2 + 4\gamma_2}\right], \tag{8}$$

with corresponding errors easily obtained propagaging $(\sigma_{\gamma 1}^2, \sigma_{\gamma 2}^2, \sigma_{\gamma 1 \gamma 2})$ to β and ρ via the Taylor's expansion of (8). The estimated autoregressive coefficient is found to be $\beta = 0.9992$ with a standard error of 0.0005 for Total Sales, $\beta = 0.9950$ with a standard error of 0.0004 for Employees and $\beta = 0.9918$ with a standard error of 0.0006 for Value Added. The value of β for Total Sales is not significantly different from 1. Conversely, the estimates for Employees and Value Added, even if very near, significantly differ from 1. Due to its very small value, however, this difference can be considered of little economic relevance. Moreover our results seem to suggest that growth rates are mildly first-order autocorrelated for Total Sales ($\rho = 0.038$ with a standard error of 0.003) and for Employees ($\rho = 0.042$ with a standard error of 0.003), while mild anti-correlation is found for Value Added ($\rho = -0.025$ with a standard error of 0.003).

Let us now consider the sectoral dynamics and let $S_{ij}^X(t)$ represent the size of the i-th firm belonging to the j-th sector at time t, where again X represents in turn Total Sales (TS), Number of Employees (L) and Value Added (VA). Define the normalized (log) size as

$$s_{ij}^X(t) = \log\left(S_{ij}^X(t)\right) - \frac{1}{N_j}\sum_{i=1}^{N_j}\log\left(S_{ij}^X(t)\right), \quad (9)$$

where $j \in \{1, \ldots, 55\}$ and N_j is the number of firms in the j-th sector.

Applying the same methodology used for the aggregate case we estimate the model

$$s_{ij}^X(t) = \gamma_{1j} s_{ij}^X(t-1) + \gamma_{2j} s_{ij}^X(t-2) + u_{ij}(t) \quad (10)$$

and derive the estimates of β and ρ according to (8).

The results are reported in Table III. One observes a good degree of homogeneity in the estimated β_j coefficients in different sectors, irrespectively of the size proxies: even when they are statistically different from 1, they remains close to this value, the extremal values being 1.017 for Total Sales in sector 175 (“Carpets, rugs and other textiles”) and .968 for Value Added in sector 285 (“Treatment and coating for metals”). In general, one can conclude that, also at sectoral level, firm growth dynamics are quite well described by a unit root process.

Concerning the ρ_j estimates, it is more difficult to propose a clear-cut interpretation of the results. First, the estimated coefficients depend on the size proxy used. Indeed, while some sectors are quite similar across proxies in their ρ coefficients, there are others for which the values of ρ strongly differ. An example of the latter case is the “Textiles weaving” sector which presents a mild positive autocorrelation coefficient in terms of Total Sales, a positive autocorrelation coefficient in terms of Number of Employees and a negative coefficient in terms of Value Added.

Second, irrespectively of the variable considered, the estimated ρ coefficient displays a significant degree of sectoral heterogeneity: in the case of Total Sales the range of variation goes from −0.185 to 0.280; in the case of Number of Employees from –0.113 to 0.192; and from −0.229 to 0.131 when Value Added is considered. For the latter variable, however, when a deviation from zero is observed, it is negative in all but three cases. The opposite is true for Number of Employees, for which there are only four cases of statistically significant negative autocorrelation.

Broadly speaking, it is safe to conclude that the autocorrelation in growth rates, when present, is in general small. It is larger than 15% in less then the 10% of the cases.

As a final step of our analysis of the intertemporal dynamics of firm sizes we consider the relationships between the size of firms on the one hand, and the standard deviation of their growth rates on the other. To this purpose we divided all the firms in our database in 25 equipopulated size classes and we ran the following regressions

$$\begin{aligned} &\log\ \sigma\left(g^X(t)|s^X(t)\right) = \eta_c + \eta \bar{s}^X(t) + u(t), \\ &X \in \{\mathrm{TS}, L, \mathrm{VA}\}, \end{aligned} \quad (11)$$

where η_c s a constant term, σ is the standard deviation of growth rates $g^X(t) = s^X(t) - s^X(t-1)$ in each size class, $\bar{s}^x$ is the average size class and $u(t)$ is an i.i.d. error term. When Total Sales and Value Added are used the η's are not significantly different from zero, being −0.004 with an error of 0.028 and −0.010 with an error of 0.045, respectively. Conversely, when the size proxy is the Number of Employees a significant positive relation emerges, $\eta = 0.121$with an error of 0.017.

In conclusion, the results obtained reveal that if a relationship between the size of firms and their average growth rate exists at all, it is very weak indeed and also sector-specific. Moreover, contrary to most previous evidence (cfr. Bottazzi and Secchi (2006a) and Bottazzi and Secchi (2005) on the international pharmaceutical industry and

TABLE III
Estimates of the AR(1) size coefficient β in (5) and of AR(1) the growth rates coefficient ρ in (6) simultaneously obtained from (7) together with standard errors

ISIC Code	β coeff.			ρ coeff.		
	TS	*L*	VA	TS	*L*	VA
151	1.006 0.004	0.998 0.003	0.995 0.005	**0.054** 0.023	−0.045 0.024	−0.083 0.026
155	0.995 0.004	**0.992** 0.003	**0.987** 0.005	**0.087** 0.025	−0.019 0.029	−0.054 0.028
158	1.000 0.002	0.966 0.002	**0.990** 0.003	0.025 0.020	0.025 0.018.	**−0.134** 0.021
159	0.995 0.005	**0.987** 0.004	**0.981** 0.006	**0.205** 0.029	**−0.054** 0.025	−0.023 0.029
171	0.995 0.004	**0.990** 0.003	**0.981** 0.005	**0.074** 0.025	**0.062** 0.019	**−0.087** 0.022
172	0.995 0.004	**0.995** 0.002	**0.991** 0.004	**0.055** 0.022	**0.120** 0.018	**−0.102** 0.024
173	**0.987** 0.005	0.994 0.003	**0.975** 0.006	**0.156** 0.022	0.004 0.017	0.039 0.022
175	**1.017** 0.005	0.994 0.004	1.004 0.006	**0.070** 0.028	**0.055** 0.026	**−0.109** 0.032
177	1.005 0.003	**0.992** 0.003	**0.987** 0.005	0.027 0.021	0.015 0.017	**−0.110** 0.022
182	0.998 0.002	**0.995** 0.002	**0.989** 0.002	**0.054** 0.014	**0.043** 0.013	**−0.040** 0.013
191	0.995 0.008	0.995 0.006	1.003 0.009	0.058 0.033	0.039 0.031	**−0.172** 0.033
193	**1.016** 0.004	1.001 0.003	1.003 0.004	−0.011 0.018	−0.028 0.016	−0.045 0.017
202	0.994 0.006	0.991 0.005	0.982 0.009	−0.019 0.040	−0.056 0.036	−0.037 0.037
203	1.010 0.006	1.011 0.007	1.016 0.008	**−0.129** 0.037	0.007 0.036	**−0.150** 0.032
205	1.004 0.007	1.003 0.009	0.990 0.010	**0.088** 0.042	**0.080** 0.035	−0.082 0.044
211	1.006 0.005	1.000 0.002	1.013 0.009	−0.072 0.043	**0.080** 0.040	−0.089 0.052
212	0.999.0.003	0.997 0.003	**0.994** 0.004	0.000 0.021	**0.048** 0.020	**−0.057** 0.023
221	0.997 0.004	**0.984** 0.003	0.992 0.005	**−0.046** 0.020	**0.046** 0.020	**−0.185** 0.023
222	1.004 0.003	0.997 0.002	**0.990** 0.004	0.041 0.022	0.025 0.016	−0.016 0.022
241	0.996 0.004	**0.990** 0.003	**0.985** 0.006	0.038 0.029	**−0.072** 0.019	−0.039 0.032
243	1.007 0.004	0.995 0.004	0.998 0.006	0.005.0.037	**0.112** 0.035	**−0.118** 0.041
244	**0.991** 0.004	**0.992** 0.003	**0.987** 0.004	**0.084** 0.027	0.016 0.024	**−0.158** 0.029
245	0.999 0.005	0.995 0.004	1.003 0.006	**0.280** 0.042	0.035 0.039	0.016 0.044
246	1.000 0.006	0.992 0.005	0.993 0.009	0.039 0.036	0.042 0.037	−0.032 0.031
251	1.007 0.004	**0.991** 0.003	0.995 0.004	**0.176** 0.023	**0.079** 0.019	−0.028 0.032
252	1.002 0.003	0.997 0.002	0.996 0.004	0.008 0.014	0.023 0.013	**−0.058** 0.014
261	0.999 0.003	0.994 0.003	0.997 0.004	−0.015 0.031	0.007 0.028	−0.0.07 0.032
262	0.997 0.005	0.993 0.004	0.998 0.005	**0.209** 0.039	**0.192** 0.035	0.075 0.041
263	1.010 0.005	0.998 0.004	0.998 0.006	**0.211** 0.029	**0.096** 0.027	**0.079** 0.030
264	1.005 0.006	0.996 0.003	**0.978** 0.009	−0.001 0.033	0.003 0.030	**−0.098** 0.041
266	0.989 0.006	0.993 0.004	**0.984** 0.006	**−0.116** 0.025	−0.010 0.020	**−0.098 0.026**
267	1.004 0.006	0.989 0.006	0.998 0.009	−0.023 0.036	**−0.113 0.032**	**−0.079 0.033**
273	**0.973** 0.007	**0.988** 0.005	**0.974** 0.010	**0.094** 0.026	0.030 0.020	−0.008 0.035
275	0.990 0.005	**0.989** 0.003	**0.983** 0.005	0.014 0.023	−0.000 0.024	**−0.080** 0.026
281	1.000.0.006	**0.981** 0.005	**0.982** 0.007	**−0.185** 0.024	0.002 0.023	**−0.118** 0.023
284	1.003 0.005	1.000 0.004	0.996 0.006	−0.009 0.028	−0.059 0.024	**−0.064** 0.026
285	**0.982** 0.006	**0.985** 0.005	**0.968** 0.007	−0.022 0.021	0.004 0.021	−0.011 0.022
286	1.002 0.004	1.001 0.003	0.991 0.005	0.006 0.023	**0.045** 0.022	−0.014 0.022
287	**0.993** 0.003	**0.992** 0.002	**0.992** 0.003	0.003 0.017	−0.016 0.013	−0.000 0.017
291	0.996 0.003	**0.991** 0.002	**0.980** 0.004	**0.079** 0.018	**0.119** 0.017	**0.055** 0.018
292	1.000 0.004	0.997 0.003	**0.990** 0.004	**−0.161** 0.020	**0.117** 0.020	**−0.061** 0.022
293	1.002 0.008	1.001 0.007	1.014 0.009	**−0.164** 0.040	−0.042 0.035	**−0.229** 0.039
294	0.998 0.006	0.993 0.004	**0.985** 0.007	**−0.059** 0.027	0.024 0.025	0.014 0.027
295	0.996 0.003	0.996 0.002	**0.986** 0.003	**−0.149** 0.014	**0.082** 0.012	**−0.085** 0.014
297	**1.013** 0.004	1.001 0.004	1.004 0.004	0.018 0.035	**0.179** 0.035	−0.047 0.040
311	1.005 0.006	1.006 0.004	1.003 0.006	0.066 0.037	**0.067** 0.032	0.040 0.037
312	1.005 0.007	0.994 0.005	**1.014** 0.006	−0.013 0.034	−0.001 0.030	**−0.164** 0.035
316	0.999 0.005	0.994 0.003	**0.972**.0.005	0.041 0.033	0.045 0.027	−0.012 0.031
322	1.000 0.007	**0.988** 0.003	**0.983** 0.007	**0.085** 0.041	**0.078** 0.037	0.033 0.037
332	1.001 0.006	**0.989** 0.004	0.996 0.006	**0.114** 0.041	**0.072** 0.032	**0.131** 0.044
342	0.995 0.012	0.986 0.007	0.983 0.012	0.017 0.045	**0.113** 0.042	−0.014 0.048
343	**0.984** 0.004	**0.989** 0.003	**0.982** 0.004	−0.004 0.026	**0.052** 0.024	−0.026 0.026

TABLE III
Continued

ISIC Code	β coeff.			ρ coeff.		
	TS	L	VA	TS	L	VA
361	1.005 0.003	**0.994** 0.002	**0.990** 0.003	**0.050** 0.014	**0.048**.0.013	**−0.063** 0.014
362	1.001 0.006	0.997 0.004	0.996 0.006	**0.078** 0.033	−0.006 0.028	−0.130 0.030
366	0.993 0.007	**0.986** 0.006	0.987 0.009	**0.094** 0.037	**0.115** 0.034	−0.024 0.039
Aggregate	0.9992 0.0005	**0.9950** 0.0004	**0.9918** 0.0006	**0.038** 0.003	**0.042** 0.003	**−0.025** 0.003

Estimate of β which significantly differ from 1 (5% significance level) and of ρ which significantly differ from 0 (5% significance level) are reported in bold.

Bottazzi and Secchi (2003) on the U.S. manufacturing industry) we are not able to identify any systematic negative relation between the size of firms and the (log) standard deviation of their growth rates. In fact, rather counter-intuitively, when Number of employees is used, a clear *positive* dependence emerges: that is the growth of bigger firms fluctuates more than that of smaller ones.

5. Growth rates distribution

Let us turn to the analysis of firm growth process and growth rates distributions.

First, with non-parametric methods, we analyze whether any characteristic shape for the growth rates densities emerge, whether it is persistent over time and whether it is robust under disaggregation. Second, we focus on disaggregated data and, using a parametric approach, we compare the growth rates distributions in different sectors, based on different size proxies.

In what follows we define growth rates as the first difference of normalized (log) size according to

$$g^X_{i,t} = s^X_i(t) - s^X_i(t-1) \quad X \in \{\mathrm{TS}, L, \mathrm{VA}\} \qquad (12)$$

at the aggregate level and

$$g^X_{ij,t} = s^X_{ij}(t) - s^X_{ij}(t-1) \quad X \in \{\mathrm{TS}, L, \mathrm{VA}\} \qquad (13)$$

at the sectoral level. Notice that from (4) the distribution of the g's is by construction centered around 0 for any t.

5.1. *Density shapes*

Figures 8, 10 and 12 report the empirical densities of firm growth rates, for 4 different years, aggregated over all the sectors (as from equation (12)), in terms of Sales, Number of Employees and Value Added, respectively. These plots offer strong support to the stationarity of the growth rates distribution. This is broadly confirmed by the time profile of their standard deviation and skewness, reported in Figrues 9, 11 and 13. Conversely, the dynamics of the kurtosis displays more variability (see top panels in Figures 9, 11 and 13). Note, however, that such a volatility in the kurtosis is mainly due to a small number of extreme growth events. Similar regularities emerge when we analyze the growth rates distributions at the sectoral level: they are stationary in all the sectors and for all the considered size proxies.

A first interesting property regards the distribution of growth rates. Bottazzi and Secchi (2003), studying the aggregate U.S. manufacturing industry, found a tent-shape that is well approximated by a Laplace distribution. Note that such a density implies tails that are fatter than Gaussian ones, meaning that the underlying drivers of corporate growth ought to involve relatively frequent and relatively "big" events, unaccountable by a Gaussian process. Figures 8, 10 and 12 show that this is the case also for Italian manufacturing firms irrespectively of the proxy used.

What happens with firm growth rates distributions at sectoral level? Is this picture different at finer grained disaggregation? The answer is negative. We find that the distributions display a tent-shape which is remarkably similar across different sectors. This is in line with similar findings on U.S. manufacturing sectors.(cfr. Bottazzi and Secchi, 2003).

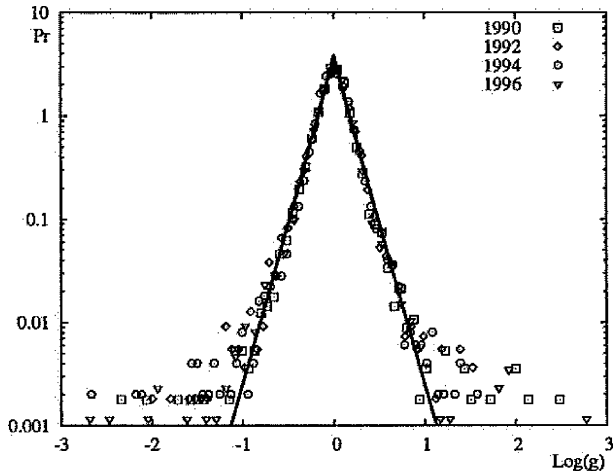

Figure 8. Growth rates densities together with the maximum likelihood Subbotin fit in different years. Size measured in terms of Sales. Notice the log scale on the y-axes.

Figure 14 reports the empirical growth rates densities[9] for the same three illustrative sectoral examples whose size distributions were displayed in Figure 7. Notice again that, in line with the results on size distribution, the shape of the growth rates densities does not seem to depend very much on the size proxy used.

Moreover, contrary to the evidence discussed above on the distribution of sizes, the evidence on growth rates densities suggests that they are rather *homogeneous* across sectors.

In order to obtain a better statical characterization of this regularity, it is useful to revert to different, parametric, methods.

5.2. *Subbotin estimations*

We adopt the parametric approach suggested in Bottazzi anal Secchi (2006) and consider a flexible family of probability densities, known as the Subbotin family (Subbotin, 1923), which includes as particular cases the Laplace and the Gaussian densities. This family is defined by 3 parameters, namely a *positioning* parameter μ a *scale* parameter a and a *shape* parameter b. Its functional form reads

$$f_s(x) = \frac{1}{2ab^{1/b}\Gamma(1/b+1)} e^{-\frac{1}{b}\left|\frac{x-\mu}{a}\right|^b} \tag{14}$$

where $\Gamma(x)$ is the Gamma function. The lower is the shape parameter b, the fatter are the density

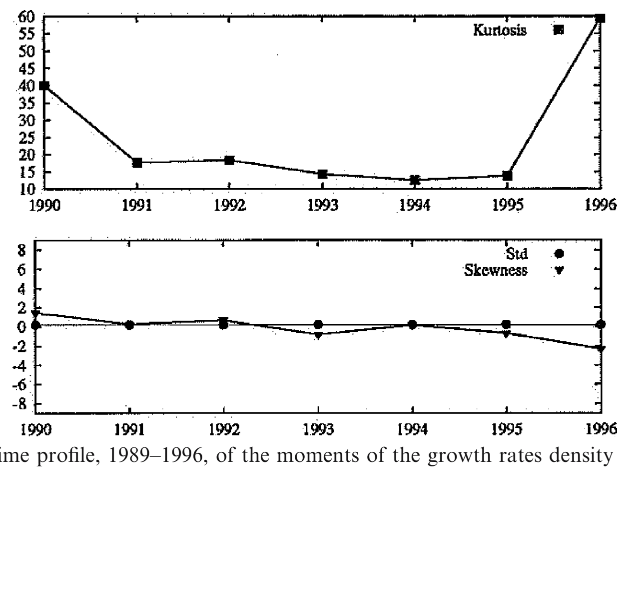

Figure 9. Time profile, 1989–1996, of the moments of the growth rates density (Total Sales).

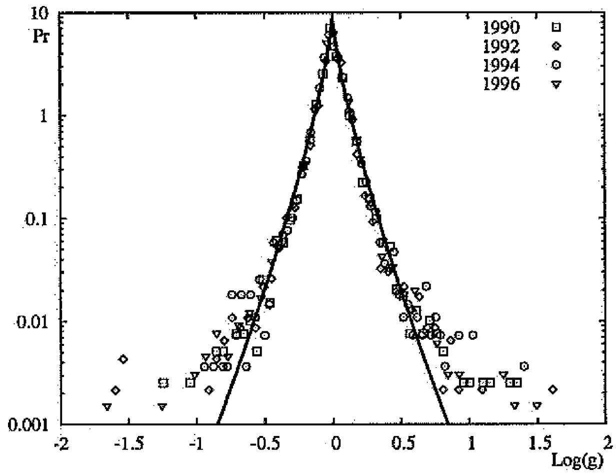

Figure 10. Growth rates densities together with the maximum likelihood Subbotin fit in different years. Size measured in terms of numbers of employees. Notice the log scale on the *y*-axes.

tails. For $b < 2$ the density is leptokurtic, while it is platykurtic for $b > 2$. For $b = 2$ this density reduces to a Gaussian and for $b = 1$ to a Laplace distribution. We undertake maximum likelihood estimations of the a and b parameters for each sector (the parameter μ is set to 0 by the normalization in (4)) following the procedure detailed in Bottazzi and Secchi (2006). The results are reported in Table IV.

Regarding the shape parameter b, it turns out to be fairly concentrated around the value of 1, suggesting a remarkable *inter-sectoral similarity* in the shapes of growth rates distributions. Notice that the value $b = 2$, associated to a Gaussian density, is ruled out with high significance (more than .999) for all the 55 sectors analyzed. Very similar results have been reported for U.S. companies in Bottazzi and Secchi (2003).

Note, at the same time, that the shape of the distribution depends on the proxy used. So, at the aggregate level, if Total Sales are used as size proxy, the Subbotin estimate of the growth rate density gives $b = 0.96 \pm 0.007$ and $a = 0.13 \pm 0.0007$. If Employees are used instead, one obtains $b = 0.76 \pm 0.006$ and $a = 0.07 \pm 0.0004$. Finally, with Value Added as measure of size the estimation gives

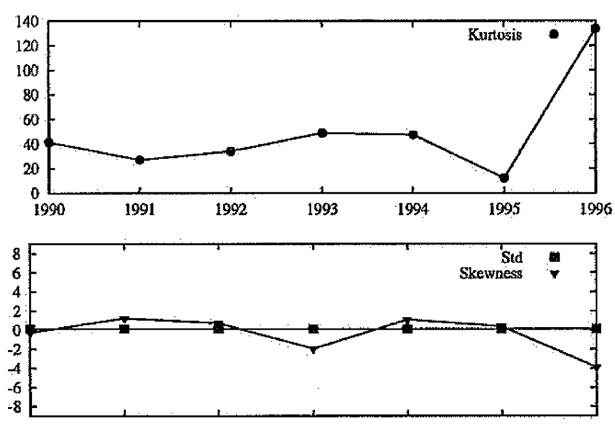

Figure 11. Time profile, 1989–1996, of the first moments of the growth rates density (Number of Employees).

$b = 0.87 \pm 0.007$ and $a = 0.14 \pm 0.0008$. The resulting densities are reported in Figures 8, 10 and 12 respectively. While the results based on Sales and Value Added are rather similar with each other, the density of growth rates measured by Number of Employees appears noticeably more fat-tailed. The same behavior is also apparent if one looks at the estimated values of the parameter b obtained at sectoral level with different measures of size (see the last three columns in Table IV). The fatter tails (i.e. the lower value of b) and the smaller width (i.e. lower a) characterizing the growth rates densities in terms of employees are plausibly the effect of higher degrees of "lumpiness" associated to employment growth: after all, indivisible plants come together with relatively lumpy changes in employment. Hence (positive and negative) growth events, when described in terms of number of employees, are more likely to appear in relatively large "chunks" (also in absolute terms).

Is there any relationship between the parameters which characterize the sectoral growth rates density, a and b, and the sectoral average firm size? For the parameter a, running a linear regression, we find a slope not significantly different from zero, irrespectively of the size proxy used. This is also the case for the parameter b when Value Added is used, while using Sales one obtain a slope equal to -1.91 ± 0.47 and a slope of -0.95 ± 0.38 using the number of employees. These negative correlations imply that sectors with higher average size are also the ones characterized by lower b's, i.e. where the tails of the growth rates distribution are fatter.

6. Conclusion

In this work, using a large longitudinal database on Italian manufacturing firms, we have tried to highlight some fundamental properties of both industrial structures and corporate growth disentangling those features which are invariant across sectors and independent of the measure of size used from others which, on the contrary, are measure- and/or sector-specific.

At the aggregate level, in contrast with Stanley et al. (1995), we found that the variable used to proxy the size of the firm has a non-negligible effect on the resulting size distributions. In particular, when the Number of Employees is used, a significant bimodal shape appears. Interestingly, this means that, at least for the database considered here, neither the log-normal nor the Yule or Pareto densities represent a robust, variable-invariant, description of firm sizes.

Moreover, corroborating a conjecture from Dosi et al. (1995), we find that the upper tail behavior in the aggregate distribution is mainly

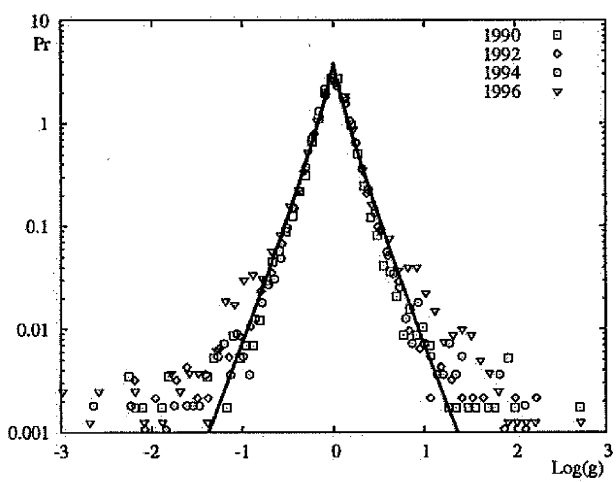

Figure 12. Growth rates densities together with the maximum likelihood Subbotin fit in different years. Size measured in terms of Value Added. Notice the log scale on the y-axes.

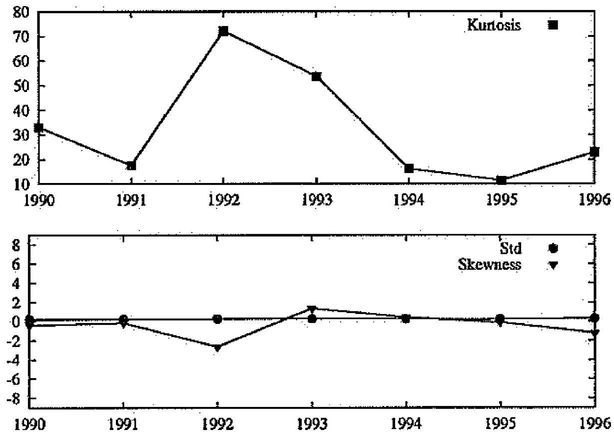

Figure 13. Time profile, 1989–1996, of the first moments of the growth rates density (Value Added).

the outcome of aggregation over different industrial sectors. In fact, at sectoral level, we unveil a high degree of heterogeneity in the shapes of different distributions. Using a non-parametric method based on kernel density estimation, we show that several sectors possess multimodal densities, while others are characterized by unimodal, bell-shaped ones. Together, the large support of all these distributions highlights the persistent coexistence of firms of very different sizes – also within similar lines of activity. At the same time, well in line with previous evidence, the analysis of market concentration confirms that the upper-tails of size distributions display significant and persistent inter-sectoral differences. They in turn hints at the existence of powerful and persistent underlying structural differences among different industrial activities, possibly due to technological and organizational factors.

Concerning the process of corporate growth we found that, first, irrespectively of the size proxy used, the size dynamic process is *approximately* characterized by a unit root. This result is well in accordance with many previous studies (cfr. among many others Hart and Prais (1956), Simon and Bonini (1958), Hymer and Pashigian (1962), Mansfield (1962), Bottazzi and Secchi (2003)) on different databases. Note, however, that our analyses only consider relatively large firms (with more than 20 employees): hence our evidence, as such, can easily come together with violations of Gibrat-type processes characterizing the growth patterns of younger, and typically smaller, business firms (cfr. Audretsch et al., 1999; Heshmati, 2001; Lotti et al., 2003).

Second, at the aggregate level, again in tune with previous studies (cfr. Dunne and Hughes, 1994; Kumar, 1985; Waguer, 1992), we find a mild autocorrelation in the growth process, whose sign seems to depend on the size proxy used.

However, third, at a disaggregated level, we identify substantial differences among different sectors. When Number of Employees or Value Added is used as size proxy, the autocorrelation coefficients are either not significantly different from zero or significantly negative. In few sectors this anti-correlation effect is rather strong. When Total Sales is used as size proxy, we also observe few cases of strong positive correlations. However, we do not have to offer any straightforward interpretation of this piece of evidence.

Fourth, contrary to the findings in Hymer and Pashigian (1962) on UK data and Bottazzi and Secchi (2003) on U.S. data, we do not identify any relation between the variance of firm growth rates and its size. There are different possible interpretations of this phenomenon. If one accepts that the negative scaling of growth variance and size is mainly due to the diversification process among weakly correlated lines of business (cfr. Bottazzi and Secchi 2006a), one may interpret the lack of this relation in our data as due to a weaker role of the diversification in the process of growth of Italian firms or to the possibility that even when diversification occurs, it happens in sectors whose demand profiles are highly correlated.

Fifth, concerning the growth rate distributions, confirming previous investigation on U.S. data (Bottazzi and Secchi, 2003), we found that they are well described by "tent-shape" (that is exponential) distributions. It must be emphasized that such regularity, contrary to what we find for size distributions, does *not* depend on the level of disaggregation at which data are observed. The non-normality of growth shocks and the fatness of its tails hint at the presence of some powerful and general correlating mechanisms. In principle, they are likely to be of two types. A first source of correlation is likely to be the very competition mechanism: if the market share of one firm grows, those of its competitors are bound to shrink (for a formulation based on a self-reinforcing mechanism explicitly yielding tent-shape growth rates cfr. Bottazzi and Secchi (2006)). A second source is likely to stem from some intrinsic "lumpiness" of the growth processes associated with the discreteness of events like entering/leaving a new market, building a

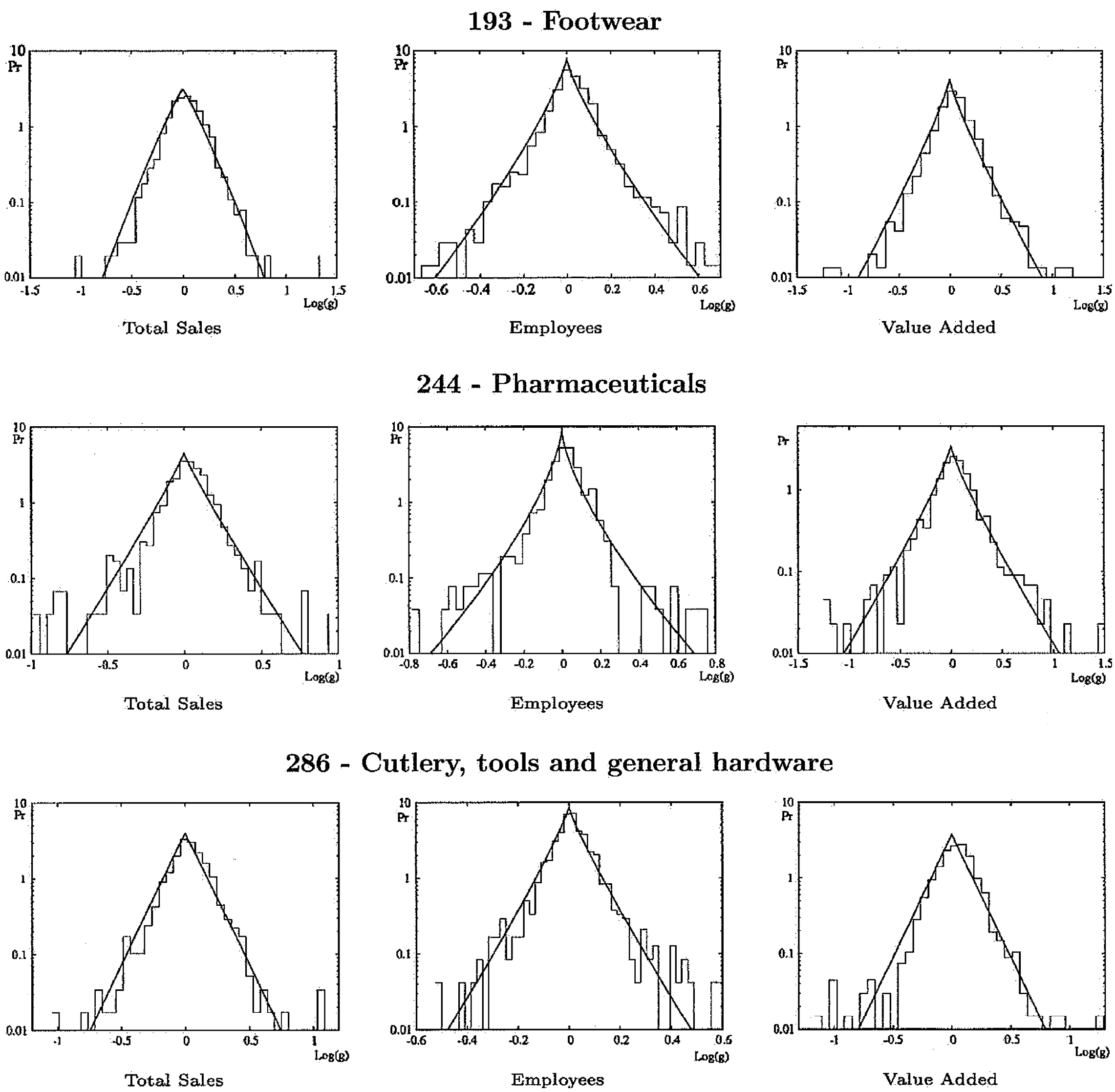

Figure 14. Binned probability density and subbotin maximum likelihood estimation of firm growth rates g in three different manufacturing sectors. Size proxies are Total Sales, Employees and Value Added. Notice the log scale on the y-axes.

TABLE IV
Estimated a and b Subbotin coefficients and standard errors

ISIC Code	a coeff.			b coeff.		
	TS	L	VA	TS	L	VA
151	0.089 0.004	0.073 0.003	0.127 0.006	0.83 0.05	0.84 0.05	0.87 0.06
155	0.080 0.004	0.065 0.004	0.143 0.008	0.91 0.07	0.80 0.06	0.87 0.06
158	0.097 0.004	0.067 0.003	0.133 0.005	0.89 0.05	0.70 0.04	0.90 0.05
159	0.109 0.006	0.077 0.004	0.158 0.008	0.88 0.06	0.70 0.05	0.86 0.06
171	0.142 0.005	0.066 0.003	0.156 0.006	1.19 0.07	0.85 0.05	0.86 0.05
172	0.122 0.004	0.051 0.002	0.147 0.005	1.12 0.06	0.85 0.04	1.11 0.06
173	0.107 0.004	0.056 0.002	0.131 0.005	1.11.0.06	0.71 0.04	1.02 0.05
175	0.118 0.006	0.064 0.003	0.138 0.007	1.02 0.07	0.83 0.06	0.90 0.07
177	0.124 0.005	0.062 0.003	0.141 0.006	0.97 0.05	0.70 0.04	0.81 0.04
182	0.120 0.003	0.073 0.002	0.121 0.003	0.92 0.03	0.84 0.03	0.76.0.03
191	0.140 0.007	0.070 0.003	0.161 0.008	1.12 0.09	1.12 0.09	1.03 0.08
193	0.150 0.004	0.074 0.002	0.133 0.004	1.13 0.05	0.80 0.03	0.86 0.04
202	0.104 0.007	0.056 0.004	0.141 0.010	0.98 0.09	0.70 0.06	0.79 0.07
203	0.105 0.007	0.073 0.005	0.121 0.008	0.94 0.08	1.02 0.09	0.74 0.06
205	0.106 0.006	0.076 0.005	0.117 0.008	1.31 0.13	0.82 0.07	0.79 0.07
211	0.120 0.009	0.037 0.003	0.194 0.014	0.89 0.09	0.83 0.08	0.92 0.10
212	0.103 0.004	0.062 0.002	0.130 0.005	0.93 0.05	0.82 0.04	1.00 0.05
221	0.079 0.005	0.064 0.004	0.126 0.009	0.62 0.05	0.66 0.05	0.60 0.05
222	0.107 0.004	0.056 0.002	0.120 0.004	1.25 0.07	0.72 0.03	0.94 0.05
241	0.114 0.006	0.062 0.004	0.161 0.010	0.88 0.07	0.63 0.05	0.75 0.06
243	0.080 0.005	0.053 0.003	0.105 0.007	1.05 0.10	0.88 0.08	0.95 0.09
244	0.117 0.006	0.074 0.004	0.161 0.008	0.91 0.06	0.70 0.05	0.86 0.06
245	0.098 0.007	0.060 0.005	0.142 0.010	0.99 0.10	0.65 0.06	1.06 0.11
246	0.108 0.008	0.066 0.005	0.137 0.010	0.80 0.07	0.85 0.08	0.81 0.08
251	0.097 0.005	0.066 0.004	0.125 0.007	0.93 0.07	0.65 0.05	0.97 0.07
252	0.113 0.003	0.069 0.002	0.132 0.004	0.95 0.04	0.81 0.03	0.81 0.03
261	0.099 0.005	0.063 0.003	0.122 0.006	1.08 0.08	0.86 0.06	1.03.0.08
262	0.107 0.006	0.065 0.004	0.147 0.007	1.20 0.11	0.99 0.09	1.06 0.10
263	0.109 0.005	0.063 0.003	0.129 0.007	1.04 0.08	0.89 0.06	0.79 0.06
264	0.103 0.005	0.050 0.003	0.159 0.008	1.17 0.09	0.83.0.06	1.00 0.08
266	0.160 0.007	0.072 0.003	0.157 0.007	0.86 0.05	0.74 0.04	0.89 0.05
267	0.116 0.006	0.071 0.004	0.128 0.007	1.19 0.10	0.83 0.07	0.96 0.08
273	0.126 0.007	0.064 0.004	0.167 0.009	0.84 0.06	0.67 0.05	0.94 0.07
275	0.164 0.005	0.069 0.003	0.133 0.006	0.99 0.06	0.95 0.06	1.07 0.07
281	0.185 0.007	0.086 0.003	0.160 0.006	1.26 0.07	0.92 0.05	1.01 0.06
284	0.126 0.005	0.071 0.003	0.133 0.006	1.13 0.07	0.94 0.06	0.96 0.06
285	0.135 0.005	0.078 0.003	0.147 0.005	1.01 0.05	0.93 0.05	1.13 0.06
286	0.125 0.005	0.063 0.003	0.134 0.005	1.01 0.06	0.89·0.05	1.00 0.06
287	0.107 0.003	0.066 0.002	0.126 0.004	0.86 0.04	0.76 0.03	0.910 0.04
291	0.121 0.004	0.066 0.002	0.136 0.005	0.93 0.04	0.85 0.04	0.88 0.04
292	0.158 0.005	0.067 0.002	0.142 0.005	1.04 0.05	0.83 0.04	0.90 0.04
293	0.139 0.009	0.080 0.006	0.143 0.010	0.96 0.09	0.83 0.08	0.90 0.08
294	0.170 0.008	0.068 0.003	0.148 0.007	1.06 0.07	0.89 0.06	0.99 0.07
295	0.185 0.004	0.066 0.002	0.159 0.004	1.06 0.04	0.83 0.03	0.98 0.03
297	0.108 0.006	0.085 0.005	0.123 0.008	1.19 0.11	1.09 0.10	0.99 0.09
311	0.131 0.007	0.071 0.004	0.140 0.007	1.01 0.08	0.91 0.07	1.33 0.12
312	0.149 0.009	0.079 0.005	0.155 0.009	0.84 0.07	0.66 0.05	0.96 0.08
316	0.137 0.007	0.077 0.004	0.137 0.007	0.97 0.07	0.83 0.06	0.90 0.06
322	0.175 0.013	0.070 0.005	0.168 0.013	0.92 0.10	0.76 0.08	0.75 0.08
332	0.119 0.008	0.064 0.005	0.128 0.008	1.20 0.12	0.77 0.07	1.39 0.15
342	0.153 0.010	0.074 0.005	0.150 0.010	1.09 0.11	1.06 0.11	1.26 0.13
343	0.137 0.006	0.084 0.004	0.153 0.007	1.09 0.07	0.90 0.06	1.02 0.06
361	0.121 0.003	0.068 0.002	0.128 0.003	1.04 0.03	0.89 0.03	0.82 0.03

TABLE IV
Continued

ISIC Code	*a* coeff.			*b* coeff.		
	TS	*L*	VA	TS	*L*	VA
362	0.163 0.008	0.060 0.003	0.130 0.007	1.41 0.12	0.93 0.06	0.93 0.07
366	0.117 0.007	0.077 0.005	0.131 0.008	1.14 0.10	1.00 0.08	0.92 0.08
Aggregate	0.13 0.0007	0.07 0.0004	0.14 0.0008	0.96 0.07	0.76 0.006	0.87 0.007

new plant, introducing a new product and so on. In this respect it is interesting to note that the tails of the growth rates density result fatter when the proxy for size is the Number of Employees (which in turn are likely to be quasi-fixed given a plant technique).

Finally again, on the side of puzzles, as already noted, rather homogeneous processes of growth appear to yield significantly diverse sectoral size distributions. How can it happen? Which more subtle differences in the dynamics are responsible for such patterns? One could reasonably hope to trace back the differences in industry structures (and the smaller ones in growth processes) into some underlying intersectoral differences in the process of technological and organizational innovation (e.g. the levels of technological opportunities and the degree of cumulativeness of innovation profiles), the degree of "lumpiness" of investments, and the specificities in the competition process. As a first attempt to identify some possible technological determinants of the structure and processes described above, we have checked whether the parameters D^4_{20}, a and b clustered in any way resembling the taxonomy of industrial sectors – based on distinct patterns of innovation – developed by Pavitt (1984) and refined in Marsili (2001). However, the evidence we have to report so far is largely negative or at least inconclusive. This negative result hints also at the urgency of finding finer proxies for the sector-specific characteristics of both technologies and competitive mechanisms. The task of building sound bridges between theoretical interpretations of the processes of innovation, competition and corporate growth, on one hand, and empirical regularities, on the other, is indeed still an open challenge. In this respect, also novel sources of evidence regarding e.g. innovative activities, investment rules, pricing behaviors and forms of market competition are probably required in order to achieve a finer description of the dynamics of different industrial activities.

Acknowledgements

Support by the Italian Ministry of Education, University and Research MIUR (grant no. A.AM-CE.E4002GD), the Sant' Anna School of Advanced Studies, Pisa (grant n. E6003GB) and the University of Bergamo (grant, no. 60CEFI06, Dept. of Economics) is gratefully acknowledged. The research would not have been possible without the precious help of the Italian Statistical Office (ISTAT – Industrial Statistics Division), and in particular A. Mancini and R. Monducci. Comments by Mariana Mazzuccato, Sid Winter and an anonymous referee have helped along the revision of this work. The usual disclaimers apply.

Notes

[1] The database has been made available to our team under the mandatory condition of censorship of any individual information.

[2] This further restriction is implemented only for sectoral analyses.

[3] Fluctuations around the 20 employees threshold and variability in response rates are the main sources of the observed entry/exit dynamics in our database. Consequently, we do not consider sample selection bias a severe problem in our analyses, given also the ways we handle apparent "exit" due to merger and acquisition: see next paragraph.

[4] This profile is plausibly also influenced by the 20 employees threshold in the database.

[5] For these three sectors we performed also a standard Kolmogorov–Smirnov test: the null hypothesis that the densities were equal is always rejected at 5% significance level.

[6] The result has been proved for a small family of kernels of which Gaussian kernel, which we shall use in the following, is a member. See Silverman (1986).

[7] For a different approach taking into account firm-specific components see Cefis et al. (2002).

[8] Estimates were obtained using *gbutils*, a set of programs for parametric and non-parametric statistical analysis. They are distributed under the General Public License and are freely available at www.sssup.it/~bottazzi/software.

[9] In the figures all the 7 years of data are pooled together under the assumption of stationarity of the growth process.

References

Audretsch, D., E. Santarelli and M. Vivarelli, 1999, 'Start-up Size and Industrial Dynamics: Some Evidence from Italian Manufacturing', *International Journal of Industrial Organization* **17**, 965–983.

Axtell, R. L., 2001, 'Zipf Distribution of U.S. Firm Sizes', *Science* **293**, 1818–1820.

Bain, J. S., 1956, *Barriers to New Competition*, Cambridge, Mass: Harvard University Press.

Bartelsman, E. J., J. Haltiwanger and S. Scarpetta, 2004, 'Microeconomic Evidence of Creative Destruction in Industrial and Developing Countries', *Tinbergen Institute, W.P.* **04–114/3**, 1–49.

Bottazzi, G., E. Cefis and G. Dosi, 2002, 'Corporate Growth and Industrial Structure Some Evidence from the Italian Manufacturing Industry', *Industrial and Corporate Change* **11**, 705–723.

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

Bottazzi, G. and A. Secchi, 2005, 'Growth and Diversification Patterns of the worldwide Pharmaceutical Industry', *Review of industrial Organization* **26**, 195–216.

Bottazzi, G. and A. Secchi, 2006, Explaining the Distribution of Firms Growth Rates, *The RAND Journal of Economics*, 37, 234–263.

Bottazzi, G. and A. Secchi, 2006a, 'Gibrat's Law and Diversification', *Industrial and Corporate Change* **15**, 847–875.

Cabral, L. M. B. and J. Mata, 2003, 'On the Evolution of the Firm Size Distribution: Facts and Theory', *American Economic Review* **93**, 1075–1090.

Cefis, E., M. Ciccarelli and L. Orsenigo, 2002, *From Gibrat's Legacy to Gibrat's Fallacy A Bayesian approach to study the growth of firms WP-AD* 2002–19, IVIE, Alicante: Universitat de Alicante.

Chesher, A., 1979, 'Testing the Law of Proportionate Effects', *Journal of Industrial Economics* **27**, 403–411.

Dosi, G., O. Marsili, L. Orsenigo and R. Salvatore, 1995, 'Learning, Market Selection and the Evolution of industrial Structures', *Small Business Economics* **7**, 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.

Dunne, T. and A. Hughes, 1994, 'Age, Size, Growth and Survival: UK Companies in the 1980s', *Journal of Industrial Economics* **42**, 115–139.

Evans, DS., 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 Recueil Sirey.

Hall, B. H., 1987, 'The Relationship Between Firm Size and Firm Growth in the U.S. Manufacturing Sector', *Journal of Industrial Economics* **35**, 583–606.

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

Hart, P. E. and S. J. Prais, 1956, 'The Analysis of Business Concentration', *Journal of the Royal Statistical Society* **119**, 150–191.

Heshmati, A., 2001, 'On the Growth of Micro and Small Firms: Evidence from Sweden', *Small Business Economics* **17**, 213–228.

Hymer, S. and P. Pashigian, 1962, 'Firm Size and Rate of Growth', *Journal of Political Economy* **70**, 556–569.

Huber, P. J., 1967, 'The behavior of maximum likelihood estimates under nonstandard conditions', *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability* **I**, 221–233.

Ijiri, Y. and H. A. Simon, 1977, *Skew Distributions and the Sizes of Business Firms*, New York: North Holland Publishing Company.

Kumar, M. S., 1985, 'Growth, Acquisition Activity and Firm Size: Evidence from the United Kingdom', *Journal of Industrial Economics* **33**, 327–338.

Kleiber, C. and S. Kotz, 2003, *Statistical Size Distributions in Economics and Actuarial Sciences*, New Jersey: John Wiley & Sons.

Lotti, F., E. Santarelli and M. Vivarelli, 2003, 'Does Gibrat's Law hold among young, small firms?', *Journal of Evolutionary Economics* **13**, 213–235.

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*, Cheltenhem: Edward Elgar.

Pavitt, K., 1984, 'Sectoral Patterns of technical Change: towards a Taxonomy and a Theory', *Research Policy* **13**, 343–373.

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 B* **43**, 97–99.

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

Simon, H. A. and C. P. Bonini, 1958, 'The Size Distribution of Business Firms', *American Economic Review* **48**, 607–617.

Stanley, M. H. R., S. V. Buldyrev, S. Havlin, R. Mantegna, M. A. Salinger and H. E. Stanley, 1995, 'Zipf plots and the size distribution of firms', *Economic Letters* **49**, 453–457.

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

Subbotin, M. T., 1923, 'On the Law of Frequency of Errors', *Matematicheskii Sbornik*. **31**, 296–301.

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

Wagner, J., 1992, 'Firm Size, Firm Growth and Persistence of chance: Testing Gibrat's Law with establishment Data from Lower Saxony', *Small Business Economics* **4**, 125–131.