Review of Industrial Organization **23:** 217–232, 2003.

Common Properties and Sectoral Specificities in the Dynamics of U.S. Manufacturing Companies

GIULIO BOTTAZZI and ANGELO SECCHI
S.Anna School of Advanced Studies, P.zza Martiri della Libertà, 33, 56100, Pisa, Italy

Abstract. The size distribution and growth rate dynamics of U.S. companies have been extensively studied by many authors. In this paper, using the COMPUSTAT database, we extend the analysis to disaggregated data, studying 15 sectors of the U.S. manufacturing industry. The sectoral investigation reveals the presence of general statistical properties that can be considered valid across all the studied sectors. In particular, the probability density of firms growth rates invariably displays a characteristic tent shape and the relation between the size of a firm and the variance of its rates of growth is characterized, in different sectors, by very similar scaling relations. The presence of characteristics that are robust and sectoral invariant hints at the existence of generic statistical properties shaping the dynamic of firms across the whole industry.

Keywords: Firm growth, industrial sectors, Laplace distribution, power law.

JEL Classifications: L1,C1,D2

I. Introduction

The statistical analysis of firm size and its dynamics constitutes one of the traditional problems in the Industrial Organization literature. Early investigations were conducted over datasets at a high level of aggregation, typically including large firms operating in very different sectors. Hart and Prais (1956), for instance, study the distribution of the whole U.K. manufacturing industry while Simon and Bonini (1958) explore the size distribution of the top U.S. companies across the whole manufacturing industry. These investigations were primarily focused on the statistical characterization of the size distribution of firms and on the analysis of firm growth dynamics in terms of autoregressive stochastic processes (in a large body of contributions see for instance Dunne et al. (1988), Evans (1987), Hall (1987)).

More recent contributions (Stanley et al., 1996b; Amaral et al., 1997) extend these studies to the analysis of the growth rates distribution and to the relation between the size of a company and its rates of growth. In particular they show that the probability density of growth rates of U.S. firms aggregated over all the manufacturing sectors displays, on a log–log scale, a tent-like shape that can be represented by a Laplace (symmetric exponential) distribution. They also identify a robust scaling relation between the size of a firm and the variance of its growth rates.

A common shortcoming in these studies resides in their aggregate nature. From an economic point of view one could argue that, due to the sectoral specificities in the nature of production and in the demand structure, the "pooling" of firms operating in different industrial sectors may lead to ambiguous conclusions. In fact, considering a large collection of heterogenous firms may introduce statistical regularities that are only the result of the aggregation procedure (e.g., via Central Limit Theorem), and, at the same time, conceal the specific characteristics of the dynamics of firms operating in different sectors. The existence of such a strong sectoral specificity in the growth dynamics of manufacturing firms was recognized quite early. For instance, Hymer and Pashigian (1962), analyzing disaggregated data, found a high heterogeneity in firms size distribution across different sectors. They conclude that it is quite unclear whether any "stylized fact" concerning the size distribution actually exists.

The present work follows in part the spirit of this last contribution and compares aggregated and disaggregated (sectoral) data. We perform a set of parametric and nonparametric statistical analyses of firms growth dynamics using the COMPUSTAT database and considering the whole U.S. manufacturing industry. We study the stationarity and shape of firms size distribution, the autoregressive structure of the growth dynamics and the statistical properties of the probability density of growth rates. Our analyses on aggregated data largely confirm the findings present in the literature. We then repeat the same analyses at a more disaggregated level, considering one industrial sector at a time. Comparing the results obtained for the different sectors with the results at the aggregate level we are able to address two relevant issues. First, we can identify the statistical features that survive the disaggregation process, i.e., that are valid both in the aggregate and at sectoral level. Second, we can account for the degree of sectoral heterogeneity, studying to what extent the features that do not appear in the aggregate differ among different sectors.

This paper is organized as follows. In Section II we provide a short description of the data source. The analysis at the aggregate level is presented in Section III while the results of the disaggregated analysis conducted on each sector separately are discussed in Section IV. In Section V we tentatively put forward some possible interpretations of the findings, before concluding in Section VI.

II. Data Description

In this paper we consider US publicly traded firms in the manufacturing sector (SIC codes ranging from 2000 to 3999) during the time window 1982–2001 as reported in the COMPUSTAT database. We proxy firm size using total sales in thousands of US dollars at current prices. To elude the problem of low numbers of firms in the first years and at the same time to fully exploit the database we build two different panels: one balanced and one unbalanced. In the first we shorten the relevant time range to 1993–2001 ending up with a sample of 1025 firms (cfr. Table I for details). The unbalanced panel is built in order to maximize the number of firms considered over the whole period and consists of more than three thousands firms (cfr. again Table I). The balanced panel will be used for the parametric analysis of the autoregressive structures while we use the unbalanced version for the nonparametric analyses to satisfy the higher requirements of these techniques in terms of number of observations.

III. Aggregate Analyses

In this section we perform some simple statistical analyses on firms size distribution and firms dynamics pooling data of all the industrial sectors together. We structure our analysis in three successive steps, analyzing in turn three aspects of the firm growth process that have found large coverage in the literature cited above. First, we start with a descriptive characterization of the size distribution of firms, in particular focusing on its stationarity in time and its shape. Second, we take a parametric approach and analyze the autoregressive structure of the firm growth process. Third, we explore the probability distribution of the firm growth rates and the scale relation existing between the growth rates variance and the size of the firm.

1. SIZE DISTRIBUTION

Our first analysis concerns the basic properties of the firms size distribution. In the left-side of Figure 1 we plot, on a log-scale, the probability density of the log of firm size, $\log(S)$, in four different years 1993, 1995, 1996 and 2001. Visual inspection of this plot reveals two distinct features. First, the densities display a remarkable degree of stationarity. This finding is confirmed by the right-hand side of Figure 1 where the time evolution of the first four moments of the distribution is plotted over the whole range 1993–2001. Apart from an upward trend in the value of the mean,[1] the other three moments appears very stable over time. Second, the log normal fit reported in Figure 1 seems to approximate the empirical densities very well.

2. AUTOREGRESSIVE STRUCTURE

Moving to a parametric approach, we now focus on the autoregressive structure of the time series of firm size. We estimate an AR(1) process on $\log(S_i)$. For this

[1] This is essentially due to the fact that firms size is measured in nominal terms.

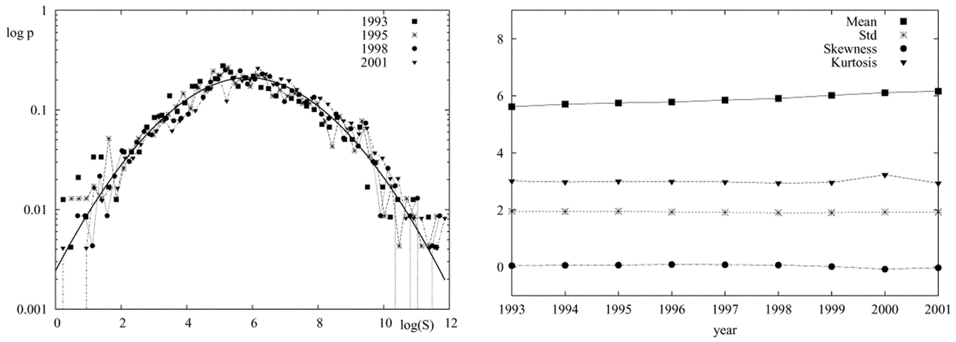

Figure 1. Left: Log empirical probability densities of (log) sales $p(\log S)$ for 4 different years together with a log normal density fitted over all the 9 years. Right: The time evolution over the range 1993–2001 of the first four moments of the log-size distribution.

purpose we use the balanced panel defined in Section II. In order to eliminate the aforementioned trend in the average size we consider the normalized (log) size

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

obtained by 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 the balanced panel.

On these rescaled observations we estimate the AR(1) model

$$s_i(t) = \phi\, s_i(t-1) + \epsilon_i(t) \tag{2}$$

where we do not introduce any firm specific term for the (log) size average. We interpret different firms as different realizations of the same stochastic process. We use a Four-Stage Instrumental Variables Estimator (Ljung, 1987, p. 403) in order to get around possible problems due to heteroskedasticity and/or autocorrelation in the error terms.[2] This procedure returns an approximately optimal set of instruments. We obtain an estimate of $\phi = 0.9576$ with a standard error of 0.005147. Our analysis confirms the existence of a unit root in the growth process. Notice that this result is well in accordance with many previous studies (cfr. among many others Hart and Prais (1956), Hymer and Pashigian (1962), Mansfield (1962), Simon and Bonini (1958)) on similar databases.

The unit-root nature of the size time series implies that the firm growth process can be well described by a geometric Brownian motion. It is interesting then to investigate the possible autoregressive structure of the first differenced process. We consider the (log) growth rates defined as the first difference of (log) size according to

$$g_i(t) = s_i(t+1) - s_i(t)\ . \tag{3}$$

[2] Indeed, as will be clear below, the process under analysis does show both these effects.

Notice that from (1) the distribution of the g's is by construction centered around 0 for any t.

We estimate

$$g_i(t) = \phi^g \, g_i(t-1) + \epsilon_i(t) \tag{4}$$

using the same multi-step procedure used for equation (2). We obtain an estimated autoregressive coefficient $\phi^g = 0.0621$ with a standard error of 0.0140, significantly different from 0 even if very small (the variance of g being approximately 0.24). We also estimate an AR(2) process on the same observations

$$g_i(t) = \phi_1^g \, g_i(t-1) + \phi_2^g \, g_i(t-2) + \epsilon_i(t) \tag{5}$$

obtaining a two-lag coefficient not significantly different from zero, $\phi_2^g \sim 0$, and a one lag coefficient not significantly different from the previous one, $\phi_1^g \sim \phi_g$. We conclude that the AR(1) model completely accounts for the autoregressive structure in the data.

3. GROWTH DENSITY AND VARIANCE-SIZE RELATION

From the early investigations of Hymer and Pashigian (1962) it has been suggested that the variance of business firms growth rates does decrease when their size increases; in particular Amaral et al. (1997) show that the standard deviation of firms growth rates $g_i(t)$ scales with size according to a Power Law $\sigma(g|S) \sim S^{\beta}$ with $\beta = -0.20$ with a standard error of 0.03. We consider the time frame 1993-2001 and we split the $s_i(t)$'s in 20 bins (quantiles). We then compute the standard deviation of the associated growth rates $g_i(t)$ in each bin. We fit the linear relation

$$\log(\sigma(g|s)) = \alpha + \beta \, s \tag{6}$$

of the standard deviation on the average bin size and we find an exponent $\beta = -0.19$ with a standard error of 0.01 that is strikingly similar to the one found by Amaral et al. (1997).[3] In Figure 2 we report on a log-log scale the sample standard deviation for each bin versus the average bin size. The linear fit (6) provides a good description of the variance-size relation. Notice that we also performed the same exercise using the mean value of the growth rates in each bin, instead of the standard deviation, but we did not find any relationship with the average size. Moreover, no robust relationship was observed between average size and higher moments (skewness and kurtosis) of the distribution of growth rates for firms inside the bin.

From the previous analysis we conclude that the dependence of the growth rates on the size of the firm is completely described by the variance-size relation in (6).

[3] In Bottazzi et al. (2001) an analogous investigation has been conducted over the worldwide pharmaceutical industry and a value of $\beta = 0.20$ with a standard error of 0.02 was found.

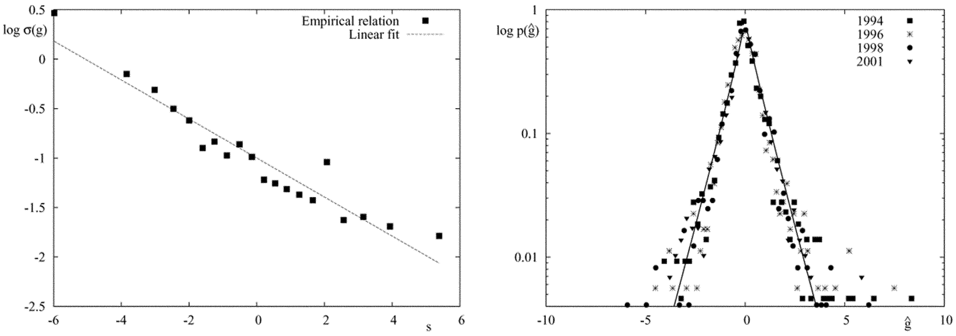

Figure 2. Left: Log of the standard deviation of the one year growth rates as a function of s. The linear fit $\log(\sigma(g|S)) \sim \beta \log(S)$ with β equal to -0.19 is also shown. Right: Log of the empirical densities of rescaled growth rates for 4 different years together with a Laplacian fit.

We use this relation to define a new variable, the “rescaled” growth rate, according to

$$\hat{g}_i(t) = g_i(t) / \left(\alpha\, e^{\beta\, s_i(t)}\right) \tag{7}$$

whose distribution has, by construction, unit variance. These new rescalled growth shocks possess statistical properties that are independent from the size of the firm, and it is then possible to pool together growth rates coming from firms in different size bins.

We plot the empirical probability densities of the rescaled growth rates $\hat{g}$ for four different years in Figure 2. These densities display a characteristic tent shape that, according to Stanley et al. (1996b), can be described using a Laplace (symmetric exponential) density

$$f_{\mathrm{L}}(x; \mu, a) = \frac{1}{2a} e^{-\frac{|x-\mu|}{a}}. \tag{8}$$

In Figure 2 we report the Laplace fit obtained using all the nine years of data pooled together. As can be seen, the Stanley et al. (1996b) result seems largely confirmed as the Laplace density well describes the observations. In order to quantify the agreement between the empirical density and the Laplace fit we follow a parametric approach and we consider the Subbotin family of densities (Subbotin, 1923), already used in Bottazzi et al. (2002), that contains the Laplace as a special case. This family is defined by three parameters: a positioning parameter μ, a scale parameter a and a shape parameter b. Its functional form reads:

$$f_{\mathrm{S}}(x) = \frac{1}{2ab^{1/b}\Gamma(1/b+1)}\, e^{-\frac{1}{b}\left|\frac{x-\mu}{a}\right|^{b}}, \tag{9}$$

where $\Gamma(x)$ is the Gamma function. The smaller the shape parameter b, the fatter the tails of the density. For $b < 2$ the density is leptokurtic and is platikurtic for

$b > 2$. It is straightforward to check that for $b = 2$ this density reduces to a Gaussian and for $b = 1$ to a Laplace.

We estimate the b parameter by maximizing the likelihood of the observed data. Notice that even if (9) is a three parameter family of densities its estimation on the empirical density of the $\hat{g}$'s is a one parameter problem since μ is set to 0 by the normalization condition in (1) and the relationship between a and b is fixed by the rescaling procedure in (7), which imposes a unit variance. We find an estimated value of $b = 1.06$ which is very close to the theoretical Laplace value of 1.

To summarize, we find that the size distribution of firms is stationary and characterized by a log-normal shape, confirming the results in Stanley et al. (1995a). We show that the growth process possesses a unit root and that the standard deviation of growth rates declines with size according to a Power Law. The $g_i(t)$ dynamics is characterized by a very small autoregressive coefficient and the growth rates, once rescaled according to the proper variance/size relation, display a characteristic Laplace shape.

IV. Sectoral Analyses

The analyses run in the previous Section considered data aggregated across the whole manufacturing industry. In the present Section we perform exactly the same statistical investigations but at a sectoral level. In this way we can check to what extent the findings of the previous Section survive to more disaggregated analysis.

Mimicking the structure above, we split this Section in three parts to study, in turn, the size distribution of firms, the autoregressive nature of the growth process and the properties of the growth rates distribution. Retaining the same structure will help the following comparative discussion.

Analogously to what done in the aggregate case, we use total sales as a definition of firms size. We consider total sales of U.S. manufacturing firms disaggregated up to a two-digit level (SIC codes range from 20 to 39) in the time frame 1982–2001. For statistical reliability we restrict our analysis to the sectors with more than 21 firms: under this constraint the number of two-digit sectors is reduced from 19 to 15. For each sector we build both a balanced and an unbalanced panel. The third column of Table I reports for each sector the number of firms in the balanced panel and the maximum number of firms present at the same time in the unbalanced panel. In the second column a brief description of the sector activity is provided.

1. SIZE DISTRIBUTION

We begin our disaggregate analysis with the inspection of the firm size distribution in the different sectors.

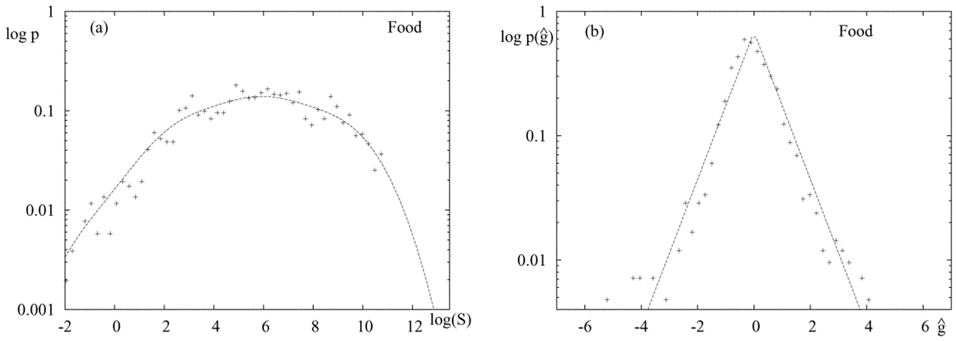

Figure 3. Binned probability density and kernel density estimation of (a) log firm size and (b) rescaled growth rates for the Food sector. The scale on y axis is logarithmic.

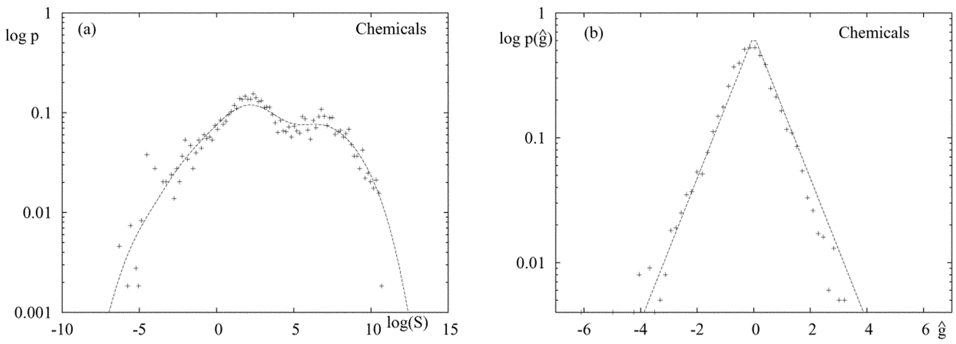

Figure 4. Binned probability density and kernel density estimation of (a) log firm size and (b) rescaled growth rates for the Chemicals sector. The scale on y axis is logarithmic.

Table I reports the first four moments of the size distributions for all the fifteen sectors analyzed. From this table it is clear that the degree of heterogeneity in the size distribution for the different sectors is rather high. For instance, the mean value and the standard deviation of firms (log) sizes range respectively from 3.09 to 7.81 and from 1.88 to 3.33. The fourth and the fifth column of Table I also show that remarkable sectoral specificities, concerning the degree of symmetry and the relative weight of the density tails, are present in the different size distributions.

To provide a visual hint of this sectoral diversity we present in the left panels of Figure 3, Figure 4 and Figure 5 the size distributions of three different sectors (Food and Kindred Products, Chemicals and Allied Products and Industrial Machinery and Equipment). Aside from the Food sector, which displays approximately a log-normal size distribution, we observe in the Machinery and Chemicals sectors respectively left and right skewed distributions with a high probability of presence of bimodality.

Table I. Summary table of the 15 sectors under analysis. Mean, standard deviation, skewness and kurtosis of the size distribution are reported.

SIC code	Sector	No. of firms (bal-unbal.)	Average size	Std of size	Skewness	Kurtosis
20	Food and kindred products	61–169	5.74	2.54	−0.28	−0.18
23	Apparel and other textile products	23–72	5.34	1.88	−0.96	2.33
26	Paper and allied products	37–80	6.16	2.21	−1.46	4.39
27	Printing and publishing	47–106	5.27	2.24	−0.89	0.75
28	Chemicals and allied products	128–646	3.35	3.33	−0.015	−0.56
29	Petroleum and coal products	30–54	7.81	3.07	−1.35	2.17
30	Rubber and miscellaneous plastics products	34–96	4.71	2.11	−0.52	1.26
32	Stone, clay, glass, and concrete products	21–46	5.21	2.17	−0.56	−0.09
33	Primary metals	51–122	5.99	2.05	−1.30	3.29
34	Fabricated metal products	45–98	4.78	1.99	−0.37	0.48
35	Industrial machinery and equipment	148–494	4.25	2.57	−0.19	0.65
36	Electrical and electronic equipment	144–611	4.07	2.31	0.20	0.89
37	Transportation equipment	57–157	6.14	2.87	−0.14	−0.16
38	Instruments and related products	95–484	3.09	2.41	−0.05	0.56
39	Miscellaneous manufacturing industries	24–89	4.09	1.92	−0.66	1.75
–	Whole industry (balanced panel)	1025	5.87	1.93	0.03	0.00

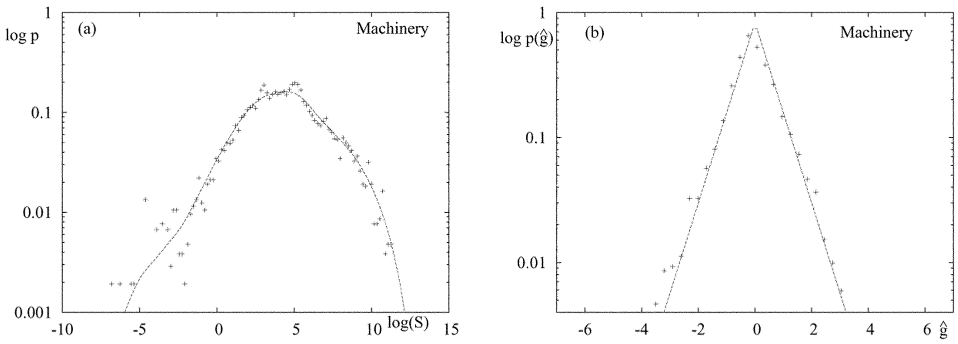

Figure 5. Binned probability density and kernel density estimation of (a) log firm size and (b) rescaled growth rates for the Machinery sector. The scale on y axis is logarithmic.

2. AUTOREGRESSIVE STRUCTURE

Let $S_{ij}(t)$ represent the sales of the i-th firm, belonging to the j-th sector, at time t. Here $j \in \{1, \ldots, 15\}$ and if N_j is the number of firms in the j-th sector then $i \in \{1, \ldots, N_j\}$. We define the normalized sectoral (log) sales as:

$$s_{ij}(t) = \log(S_{ij}(t)) - \frac{1}{N_j} \sum_{i=1}^{N_j} \log(S_{ij}(t)) \tag{10}$$

subtracting from the (log) size of each firm the average (log) size of all the firms operating in the same sector. Applying the same methodology used in Section III we estimate the model

$$s_{ij}(t) = \phi_j \, s_{ij}(t-1) + \epsilon_{ij}(t) \tag{11}$$

and we report the results in the first column of Table II. We observe a substantial homogeneity in the estimated AR coefficients in different sectors: they are all significant and very close to 1. We conclude that also at sectoral level the firm growth process is well described by a geometric Brownian motion.

We then consider the (log) growth rates defined according to

$$g_{ij}(t) = s_{ij}(t+1) - s_{ij}(t) \tag{12}$$

and we estimate the AR(1) model

$$g_{ij}(t) = \phi^g \, g_{ij}(t-1) + \epsilon_{ij}(t). \tag{13}$$

The results for the different sectors are reported in the second column of Table II and display a moderate degree of sectoral heterogeneity. Most of the sectors do not show any AR structure in growth rates (coefficients in sectors like Food and

Table II. The estimated AR coefficients, the estimated b parameters and the exponents β of the scaling relation $\sigma(g) \sim S^{\beta}$ are reported together with their standard errors for the 15 sectors under analysis.

hline SIC code	Sector	AR(1) levels		AR(1) differences		Estimated b	Scale Exponent	
		Coeff.	Std Err.	Coeff.	Std Err.		Coeff.	Std Err.
20	Food and kindred products	0.97	0.02	−0.08	0.05	1.08	−0.159	0.018
23	Apparel and other textile products	0.97	0.03	0.18	0.08	1.05	−0.201	0.035
26	Paper and allied products	0.99	0.02	0.03	0.06	1.06	−0.171	0.024
27	Printing and publishing	0.96	0.02	−0.12	0.07	0.95	−0.188	0.026
28	Chemicals and allied products	0.95	0.01	0.07	0.04	1.02	−0.203	0.012
29	Petroleum and coal products	0.96	0.03	−0.20	0.07	1.21	−0.138	0.021
30	Rubber and miscellaneous plastics products	0.97	0.03	−0.16	0.07	0.87	−0.214	0.033
32	Stone, clay, glass, and concrete products	0.93	0.04	0.16	0.12	1.13	−0.148	0.21
33	Primary metals	0.88	0.03	−0.04	0.05	1.09	−0.217	0.028
34	Fabricated metal products	0.96	0.02	−0.06	0.06	0.85	−0.184	0.030
35	Industrial machinery and equipment	0.96	0.01	0.17	0.03	1.00	−0.196	0.019
36	Electrical and electronic equipment	0.98	0.01	0.05	0.04	0.80	−0.146	0.026
37	Transportation equipment	0.96	0.02	0.06	0.05	0.93	−0.149	0.019
38	Instruments and related products	0.92	0.02	0.02	0.04	1.06	−0.193	0.020
39	Miscellaneous Manufacturing Industries	0.90	0.04	−0.20	0.09	1.02	−0.193	0.029

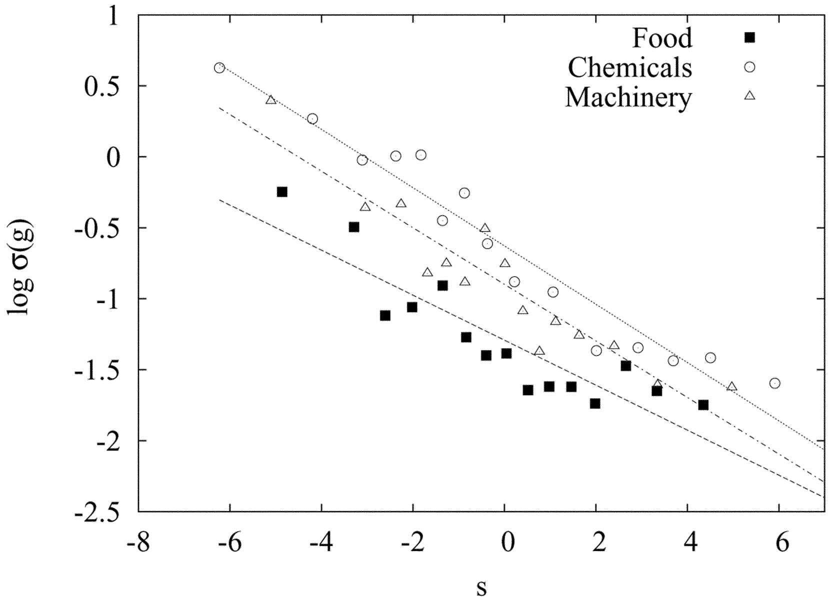

Figure 6. Log of standard deviation of the one year growth rates as a function of (log) size for the Food sector (SIC 20), the Chemicals sector (SIC 28) and the Industrial Machinery sector (SIC 35). Linear fits are also shown (see Table II for the values of the estimated coefficients).

Kindred Products, Fabricated Metals Products and some others are not significantly different from 0), other present mild positive autoregressive coefficients (for instance Industrial Machinery and Equipment) while in some other cases it is possible to observe mild negative autoregressive coefficients (for instance Petroleum and Coal Products and Rubber and Miscellaneous Plastic Products).

3. GROWTH RATES DISTRIBUTION AND VARIANCE-SIZE RELATION

We conclude our investigation of the firms growth process inside the different sectors with the analysis of the growth rates distribution. We repeat exactly the same analysis performed at the aggregate level, but now we consider a single sector at a time.

We start with the study of the first central moments of the growth rates conditional on the size of the firm. To this purpose, the (log) firm sizes $s_{ij}(t)$ are split in equipopulated bins (quantiles) and, inside each bin, the statistics of the associated growth rates $g_{ij}(t)$ are computed. For all sectors we obtain a result analogous to the aggregate case: the variance of the growth rates robustly depends on size while a relationship between size and average growth rate seems absent.

In Figure 6 we report on a log-log scale the growth rates variance in each bin $\sigma_j(g)$ versus the average size of the bin for three different sectors. The observed relationships are very similar.

Table III. Comparison between the results of the aggregate and disaggregate analyses. The second column reports which results remain valid at different levels of aggregation. The third column shows the degree of heterogeneity found at sectoral level.

Property	Robust under disaggregation	Sectoral heterogeneity
Stationarity of size distribution	yes	no
Shape of size density	no	high
AR structure in size levels	yes	no
AR structure in size differences	no	moderate
Scaling of growth rates std	yes	low
Tent shape of growth rates density	yes	low

We fit (6) using 20 bins in each sector separately to obtain sector-specific coefficients β_j. The results are reported in the last column of Table II. The obtained scaling coefficients range from the value -0.138 for Petroleum and Coal Products (SIC 29) to the value -0.217 for Primary Metals (SIC 33). In the majority of cases, (see Figure 6) the distance from the aggregate value of -0.2 is not large, even if there are noticeable deviations.

Following (7) and using the sectoral coefficients β_j it is then possible to define, for each sector, the rescaled growth rates $\hat{g}_{ij}(t)$. As a last exercise, we analyze the shape of the empirical density of this variable in each sector. In the right-hand panels of Figure 3, Figure 4 and Figure 5 the $\hat{g}_{ij}(t)$ density is reported for three sectors. The symmetric exponential fit seems to describe the observations with rather high accuracy in all three cases.

As in the aggregate case, we use the Subbotin family of densities defined in (9) to obtain a quantitative estimation of this agreement. Via maximum likelihood estimation we obtain a value of the parameter b for each sector. The results are reported in Table II. They are all very close to the value 1 that characterizes the Laplace density.

V. The Presence of Common Properties: Tentative Interpretations

The most interesting finding among the results presented in Section III and Section IV is the existence of characteristic properties of the dynamics of firms that are present both at aggregate and at disaggregate level and are essentially the same across all the sectors analyzed.

Table III provides a succinct account of the robustness of the findings at aggregate and disaggregate level, and of the degree of heterogeneity found at sectoral level for the various analyses.

In the group of results surviving disaggregation we find the stationarity of the size and growth rates distributions, the unit-root nature of the growth process, the dependence between variance of growth and firm size and the shape of the growth rates density. Notice that the similarity of these results at aggregate and disaggregate level also implies their similarity across all the sectors under study. This generality suggests that their nature has to do with some fundamental property of the economic dynamics and of the firms behavior.

While the general nature of some of these findings has long been recognized, as in the case of the unit root nature of firms growth, typically referred to as the Gibrat hypothesis (Gibrat, 1931), for other findings only recent investigations have been conducted (Stanley et al., 1996b; Amaral et al., 1997) and some tentative explanations proposed. In Bottazzi and Secchi (2003a, b) the tent-shape of firms growth rates density is explained as an emerging feature due to the existence of an underlying positive-feedback effect in the growth of firms, while in Bottazzi (2001) the relationship between growth rates variance and firm size is explained as a diversification effect, i.e., as a relation between firm size and the number of sub-markets in which the firm operates. Both these explanations ignore the details of the different economic frameworks and rely on quite general assumptions that can be plausibly considered valid across all the U.S. manufacturing sectors. The existence of robust and invariant properties in the dynamics of firms in all the sectors constitutes an argument in favor of these simple models.

Together with sector-invariant properties, we have also found statistical properties that display, on the contrary, a clear sectoral nature. The sectoral specificities are particularly strong in the firms' size distribution and in the autoregressive nature of the growth process. In this respect, it is interesting to notice the existence of a high degree of heterogeneity in the distributions of firms sizes as opposed to the essential homogeneity in the distributions of firms growth rates. This is particularly striking as the (log) size can be roughly considered as generated by a successive cumulation of growth rates so that one would expect the distributions of size and of growth rates to have the same degree of sectoral specificity. A possible explanation of this counterintuitive evidence can be found in the same nature of the "stochastic description" of firms dynamics. Suppose, indeed, that the growth dynamics of business firms can be described, in its essential nature, by a stochastic model, except from events that have nothing to do with the "generic" economic behavior, including, for instance, earthquakes or the discovery of balance sheet accounting frauds. If these events are, as it is the case, rare and of large magnitude, they can in fact permanently modify the shape of the size distribution. On the other hand, their impact on the growth rates distribution remains small and proportional to their sheer number. These arguments suggest that the growth rates structure, being less affected by "historical" events, can carry more relevant information concerning the nature of the underlying economic process.

VI. Conclusions

The parallel statistical investigations of the firm growth dynamics in the aggregate and at sectoral level allowed us to identify which properties are present at both levels of analysis and which properties, on the other hand, disappear under the aggregation or disaggregation procedure. At the sectoral level, our analysis reveals the presence of extremely robust properties, such as the tent-like shape of the growth rates density, together with properties of a more heterogeneous nature, like the firms size probability density.

We provide tentative explanations for the existence of robust characteristics across sectors. Essentially, we suggest that these characteristics can be explained with simple models, whose parsimonious requirements are so generically satisfied that they can be considered a good first approximation to many different industries. This, of course, is not intended as the end of the story, but simply as a suggestion of which facts and which directions can be considered more relevant in, first, designing and, second, comparing micro-founded models of industrial evolution.

On the other hand, we are at present unable to provide reasonable explanations for the observed heterogeneity with respect to many analytical dimensions. Concerning the autoregressive nature of growth process, for instance, its sectoral specificity may come from many distinct plausible sources, like the effect of economic cycle, the sector life-cycle, the dynamics of relative prices or shocks in the demand. The AR structure with a time horizon of several years can be indeed more affected by macroeconomic shocks or prices movements, effects that play a minor role in the intra-year dynamics shaping the growth rates density. Our time series are not long enough to allow for a macroeconomic analysis of the different sectors, but this would surely constitute a relevant addition to the present analysis.

Another interesting extension of the present work is constituted by the analysis of the degree of heterogeneity of single firms inside a given sector. In this paper, consistent with the Gibrat and Simon tradition, we consider firms as different realizations of the same stochastic process. Maybe, this assumption can conceal interesting differences in firms behavior. Again, the major obstacle in pursuing this line of research rests in the need of having quite long time series for a quite large number of firms, requirements that are not satisfied by our present data.

Finally, the work presented provides supporting evidence for the limitations of investigating aggregated data only. The degree of heterogeneity of the results for different sectors hints at a rich economic structure that is hidden by aggregated analysis, but that can be recovered when sectoral data are investigated.

Acknowledgements

The authors thank A. Beber, C. Castaldi and G. Dosi for many insightful discussions and C. Snyder and an anonymous referee for helpful comments. The usual disclaimers apply. Support from Italian Ministry of University and Research (grant

A.AMCE.E4002GD) and from Sant'Anna School of Advanced Studies (grant E6003GB) are gratefully acknowledged.

References

Amaral, L. A. N., S. V. Buldyrev, S. Havlin, M. A. Salinger, H. E. Stanley, and R. Stanley (1997) 'Scaling Behavior in Economics: The Problem of Quantifying Company Growth', *Physica A*, **244**, 1–24.

Bottazzi, G. (2001) 'Firm Diversification and the Law of Proportionate Effect', *LEM Working Paper* S. Anna School of Advanced Studies, Pisa, 2001-1

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

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 (2003a) 'A Stochastic Model of Firm Growth', *Physica A*, **324**, 213–219.

Bottazzi G., and A. Secchi (2003b) 'Why Are Distributions of Firm Growth Rates Tent-Shaped?', *Economics Letters*, **80**, 415–420.

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 Recueil Sirey.

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

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

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

Ljung, L. (1987) *System Identification. Theory for the User*. New Jersey: Prentice Hall.

Mansfield, E. (1962) 'Entry, Gibrat's Law, Innovation and the Growth of Firms', *American Economic Review*, **52**, 1023–1051.

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 (1996) 'Zipf Plots and the Size Distribution of Firms', *Economics Letters*, **49**, 453–457.

Stanley, M. H. R., L. A. N. Amaral, S. V. Buldyrev, S. Havlin, H. Leschhorn, P. Maass, M. A. Salinger, and H. E. Stanley (1996) 'Scaling Behavior in the Growth of Companies', *Nature*, **379**, 804–806.

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