Laboratory of Economics and Management
Sant'Anna School of Advanced Studies

Piazza dei Martiri della Libertà, 33 - I-56127 PISA (Italy)
Tel. +39-050-883-341 Fax +39-050-883-344
Email: lem@sssup.it Web Page: http://lem.sssup.it

LEM

Working Paper Series

On the Laplace distribution of Firm Growth Rates

Giulio Bottazzi[*] *Angelo Secchi*[*]

[*] *Sant'Anna School of Advanced Studies, Pisa, Italy*

2002/20 **August 2002**

On the Laplace Distribution of Firm Growth Rates *

Giulio Bottazzi   Angelo Secchi

S.Anna School of Advanced Studies, Pisa, Italy

May 11, 2005

Abstract

Recent empirical analyses on different datasets have revealed a common exponential behaviour in the shape of the probability density of the corporate growth rates. We briefly review these analyses and present new evidence on this topic based on corporate data from Italian Manufacturing Industry. We then propose a very simple model that, under rather general assumptions, provides a robust explanation of the observed regularities. The model is based on a very simple stochastic process describing the random partition of a number of "business opportunities" among a population of identical firms. A theoretical result is presented for the limiting case in which the number of firms and opportunities go to infinity. Moreover, using simulations, we show that even in a moderately small industry the agreement with asymptotic results is almost complete.

1 Introduction

One of the most traditional problems in the Industrial Organization literature concerns the statistical properties of the size of firms and its dynamics. Early investigations were conducted over datasets at a high level of aggregation, which included large firms operating in very different sectors and the focus of the analysis was, in particular, the shape of the distribution of firms size. For instance Hart and Prais (1956) study the distribution of the whole U.K. manufacturing industry while

*Comments by Mikhail Anoufriev, Carolina Castaldi and Roberto Gabriele have helped shaping this draft. We also would like to thank Elena Cefis, Roberto Monducci(ISTAT) and the support of the Italian Ministry of University and Research.

Simon and Bonini (1958) explore the size distribution of the top manufacturing firms of the U.S. economy, across all the sectors. It is plausible that considering such aggregate data is likely (1) to introduce statistical regularities that are only the result of aggregation (e.g. via Central Limit Theorem) and (2) to conceal the true characteristics of the dynamics of business firms that are active in specific sectors. Indeed Hymer and Pashigian (1962), analyze more disaggregated data, and find 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.

More recently a new strand of analysis, pioneered by Stanley et al. (1996) (see also Bottazzi et al. (2001)), focused on the shape of the growth rates distribution and its dependence on company size. In line with these new studies the present paper analyses the growth rates distribution of business firms in the Italian manufacturing industry using data disaggregated by sectors in order to avoid the aforementioned problems. The findings are impressive. While the size distribution is strongly variable across sectors, the growth rate distribution show basically identical shape in most sectors.

In the literature one finds few standard stochastic models aimed of explaining the mentioned industry dynamics. From the seminal work of Gibrat (Gibrat, 1931) to the more recent contributions of Geroski (2000) and Amaral et al. (2001), all models share a common feature. They do not assume any interdependence between the histories of different firms. The dynamics of each firm is a stochastic process, encompassing growth, diversification, entry and exit, that, nevertheless, does not keep in consideration the dynamics of the other firms. In this sense, any firm acts as a monopolist in a sector where the growth dynamics can be represented simply as an exogenous expansion (or contraction) of the demand. A slightly different kind of models, originally proposed by Simon (see Ijiri and Simon (1977)) and later reconsidered by Sutton (1998) make the assumption that there is a finite set of pre-existing "growth" opportunities (or equivalently, a constant arrival of new opportunities) and that firms growth process is conditioned by the number of opportunities they are able to catch. Roughly speaking, one could say that these models, generically known as "islands models", try to provide a first account of the competitive behavior among firms based on the idea that firms need to seize the most out of a scarce resources.

In the present paper we build a stylized model of firm dynamics where this idea of "competition" is introduced. The model is, indeed, similar to the previously cited tradition, at least in its aspiration to meet both the requirements of simplicity and generality. A stochastic model is used and each firm is considered a different realization of the same process. Similarly to what happens in the "island" models, this symmetry is however broken at the aggregate level, in the sense that the firms population total growth is bounded by a finite set of sector-specific opportunities.

The novelty of our approach resides in the way in which we describe the random distribution of opportunities among firms. Indeed the assumption that a random "assignment" of opportunities across firms can provide a zero-level approximation of a competitive dynamics, in other terms a "competition with random outcome", solves only part of the problem, and the implied random process of "assignment" must be specified. In providing this specification, we depart from the original "splitting" procedure found in the Simon-inspired literature and we introduce a different (and, at a first glance, maybe, unusual) statistics to describe the outcome of the "random competition". The fundamental justification of choosing a different description for the "opportunities assignment" relies on the possibility of obtaining much better description of the empirical findings. Rephrasing William Feller *We have here an instructive example of the impossibility of selecting or justifying probability models by A PRIORI argument* (Feller (1968), p.41)

The extreme simplicity of the model assumptions, suggests that this model can be a good description of many economic sectors.

In Sec. 2 we examine the empirical data. In Sec. 3 we propose a new and alternative theory based on a stochastic model of the growth process whose generality and robustness is provided by the theorem proved in Sec. 4. In Sec. 5 we present numerical simulations and discuss their agreement with empirical evidence. In Sec. 6 we draw some conclusions and briefly comment on the needs for further theoretical research.

2 Empirical evidence

Some years ago in a series of papers based on the COMPUSTAT database Stanley et al. (1996) and Amaral et al. (1997) analyzed the probability distribution of the (log) growth rates of many U.S. manufacturing firms. These studies were performed using observations in the time frame $1974-93$ and on companies with primary activity [1] belonging to the SIC code range $2000-3299$. The authors found that the growth rates followed a tent-shape distribution, and in particular, they proposed to describe the empirical observations using a Laplace (symmetric exponential) functional form

$$f_L(x;\mu,a) = \frac{1}{2a} e^{-\frac{|x-\mu|}{a}} \tag{1}$$

More recently Bottazzi et al. (2001), using the PHID database, found the same exponential behavior of the growth rate distribution analyzing the largest worldwide companies in the pharmaceutical industry.

[1] The different lines of business inside the same multi-activity firm were completely aggregated.

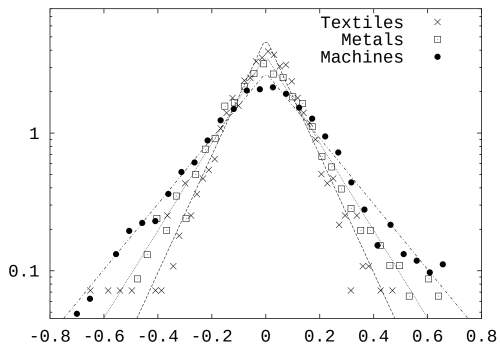

Figure 1: Binned empirical densities of the growth rates for the three sectors of textiles (181 firms), treatment of metals and metal coating (182 firms) and special purpose (metallurgy, mining, chemistry ...) machines (424 firms). The distribution pooled over all the years is reported.

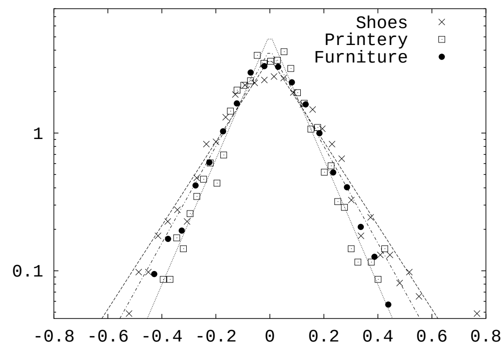

Figure 2: Binned empirical densities of the growth rates for the three sectors of shoes (245 firms), not publishing printery (199 firms) and furniture (444 firms). The distribution pooled over all the years is reported.

In the present section we extend these analysis in two directions. First, a new dataset is studied that includes many sectors of the Italian manufacturing industry. Second, by considering firms disaggregated by sectors, we are able to check to what extent the mentioned findings were a mere effect of aggregation.

The analysis presented here draws upon the MICRO.1 databank developed by the Italian Statistical Office (ISTAT)[2]. MICRO.1 contains longitudinal data on a panel of several thousands of Italian manufacturing firms, with 20 or more employees, over a decade. For statistical reliability we restrict our analysis to the period $1989 - 96$ and to the sectors with more than 44 firms. Under these constraints, the number of 3 digit sectors under study is reduced from 97 to 55.

It should be stressed that we do not present here a detailed statistical description of the business companies and their dynamics. For the purpose of the present paper we are only interested in the analysis of probability distribution of companies growth rates[3]. According to what done in Bottazzi et al. (2001) we use sales as a definition of the size of the firm.

Let $S_{i,j}(t)$ represents the sales of the i-th firm in the j-th sector at time t. In order to eliminate possible trends, both sector specific and industry-wide, we

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

[3] For an in depth analysis of some sectors and a more extensive description of the database see Bottazzi et al. (2002).

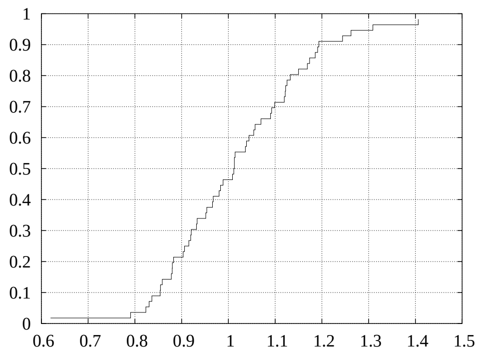

Figure 3: Distribution of the Subbotin shape parameter b estimated using maximum likelihood over the sectors population.

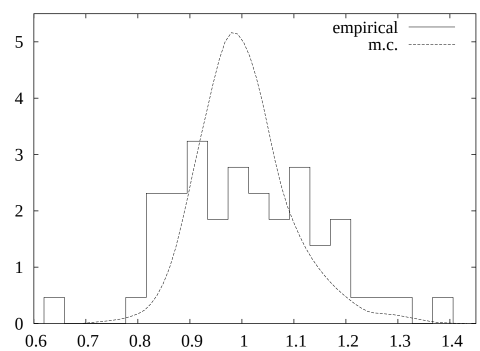

Figure 4: The binned empirical density of the b parameter values estimated using maximum likelihood over the different sectors. A Monte Carlo computation of the theoretical distribution under the hypothesis of Laplace-distributed growth rates is also shown.

consider the normalized (log) sales

$$s_{i,j}(t) = \log(S_{i,j}(t)) - < \log(S_{i,j}(t)) >_i \tag{2}$$

where $< . >_i$ stands for the average over all the firms of the j-th sector. The variable under study becomes:

$$r_{i,j}(t) = s_{i,j}(t+1) - s_{i,j}(t) \tag{3}$$

Fig. 1 and Fig. 2 show the growth rates density for 6 different 3 digit sectors chosen because both numerous and rather disparate. The activity indeed ranges from shoes making to the treatment of metals for industrial purposes. All the 7 years of data are pooled together since the differences across time in the distribution shape are negligible for any sector. As can be seen, the Laplace fit well describes the observations.

To give a synthetic account of the robustness and generality of this result (without showing 55 plots) we consider a family of distributions, introduced in Bottazzi et al. (2002), that includes the Laplace as a particular case. We fit it over the different sectors and we obtain a measure of the deviation from the exponential behavior.

The Subbotin family (Subbotin, 1923) is defined by 3 parameters: a positioning parameter μ, a scale parameter a and a shape parameter b. Its functional form reads

$$f(x) = \frac{1}{2ab^{1/b}\Gamma(1/b+1)}\, e^{-\frac{1}{b}\left|\frac{x-\mu}{a}\right|^b} \tag{4}$$

where $\Gamma(x)$ is the Gamma function. The lower is the shape parameter b, the fatter are the distribution tails. For $b < 2$ the distribution is leptokurtic and is platikurtic for $b > 2$. When $b = 2$ this distribution reduces to a Gaussian and for $b = 1$ to a Laplace.

In Fig. 3 we report the distribution function of the b's parameter estimated using maximum likelihood over the 55 sectors. Notice that almost 80% of the b values range between .85 and 1.2, suggesting that (1) provides a very good approximation of the empirical densities.

At this stage, it is important to verify if the observed departures of the b parameter from the Laplace value of 1 shown in Fig. 3 can be explained by statistical error only, i.e. if this model can be assumed not only as a proxy but as a complete statistical description of the data. The answer is partially negative. This can be straightforwardly checked using Monte Carlo techniques. Suppose that the 55 sectors are distinct and independent realizations of the same Laplace model but with different values of a and different numbers of firms. Knowing these parameters[4] it is possible to compute, using repeated simulations, the probability to obtain a given value of b via maximum likelihood procedure. In Fig. 4 the binned empirical density of the b values, together with the Monte Carlo theoretical density are shown. The agreement is not perfect, suggesting that the observed deviations from the Laplace model cannot be fully explained by statistical errors. Nevertheless this can be considered a rather robust "stylized fact" that deserves economic explanation.

3 A model of firms growth dynamics

In this Section we propose a simple model for the growth dynamics of business firms that explains the tent-shape distribution of the rates of growth discussed above. We do not claim that it can constitute a perfect description of the economic dynamics of any company, rather, it is intended to be a stylized "metaphor" of the growth dynamics observed in reality. We think that, due to the looseness and plausibility of assumptions, it can provide, at least, a "first order" approximation of the real dynamics in many different sectors.

The model is based on a simple stochastic process describing the random partition of a number of "business opportunities" among a population of identical firms. We use here the generic term "opportunity" to refer to all the variegated "accidents" that can plausibly affect the history of a business firm, encompassing, among others, the exploitation of technological novelties, the reaction to demand shocks and the effects of managerial reorganizations.

[4] We use the simple relation between the standard deviation of the distribution σ and the parameter a, namely $\sigma = \sqrt{2}\, a$, to obtain an estimation for its value.

We consider a population of N firms and suppose that at each time step, M opportunities appear that can be exploited by firms to increase their business. The specification of the model has two main elements. The way opportunities are captured by firms, and the way these affect the business of the firms.

Opportunities assignment

In order to describe how opportunities are distributed over the population of firms, we assume a random "assignment" process that can be called "random partition"[5]. Consider a starting sequence of $N+M-1$ symbols, the first $N-1$ being fences # and the other M zeros 0. Now take a random permutation of this sequence and assign to the first firm a number of opportunities that is equal to the number of 0 preceding the first #, to the second firm a number that is equal to the number of 0 between the first and the second # and so on. For example, suppose to have 3 firms and 4 opportunities, so that the starting sequence is (##0000). Now if the random permutation of its terms leads to the sequence (0#000#), the result would be an assignment of 1 opportunity to firm 1, three opportunities to firm 2 and no opportunities to the last firm.

From the previous description it is immediate to see that the total number of distinct partitions $A(N, M)$ is the number of possible permutations of $N+M-1$ elements over the number of possible permutations of the $N-1$ fences and the M zeros

$$A(N, M) = \binom{N+M-1}{N-1} \tag{5}$$

and the probability that a given firm gets exactly h opportunities becomes

$$P(h; N, M) = \frac{\binom{N+M-h-2}{N-2}}{\binom{N+M-1}{N-1}} \tag{6}$$

which corresponds to the Bose-Einstein distribution. The outcome of the previous procedure can be completely described by the occupancy N-tuple $(m_1(t), m_2(t), \ldots, m_N(t))$ where $m_i(t)$ is the number of opportunities assigned to firm i and $\sum_i m_i(t) = M$.

In order to clarify the nature of the previous process and to better understand where this description departs from the traditional approach let us briefly describe in the same framework the random "assignment" process that is typically, and mainly implicitly, assumed in the literature (Ijiri and Simon, 1977; Sutton, 1998; Amaral et al., 2001). Consider now a starting sequence of only $N-1$ fences # and think about the space between two adjacent fences as representing one single firm.

[5]The process described here is essentially the Bose-Einstein statistics used in physics to describe the behavior of a large family of subnuclear particles.

Consider the first opportunity and assign it to a randomly chosen firm, each firm having the same probability $1/N$ to be selected. This assignment can be represented placing a 0 between the fences corresponding to this firm. Repeating this procedure M times one finally obtains, analogously to the previous case, a sequence of $N+M-1$ elements which again corresponds to an occupancy N-tuple.

In this case, since the assignment of each opportunity to a given firm is an independent event with constant probability $1/N$ the resulting distribution for the number of shocks takes the well known binomial form

$$B(h;N,M) = \binom{M}{h}\left(\frac{1}{N}\right)^{h}\left(1-\frac{1}{N}\right)^{M-h} \tag{7}$$

An example can probably clarify the differences between the two assignment procedures described above better than many equations. Consider the case of 2 opportunities and 2 firms. According to the first mechanism the possible assignments are provided by the random permutations of the sequence #00, i.e. by the 3 sequences #00, 0#0 and 00#. Then, the probability of a given firm to obtain 0,1 or 2 opportunities is the same and equal to 1/3. On the contrary, according to the latter mechanism the possible assignments are obtained as follows: starting from the string # one begins by randomly placing the first opportunity and obtains, with equal probability, the sequence 0# or #0. Then one proceeds to the assignment of the second opportunity, randomly placing another 0 to the left or to the right of the #. This leads with equal probability to one of the 4 sequences 00#,0#0,#00 and 0#0. Then, the probability of a given firm to obtain 0,1 and 2 opportunities are respectively 1/4,1/2 and 1/4.

It is now clear how, starting with two equally plausible interpretation of a random " assignment", we end up with two completely different occupancy statistics as shown in Fig. 5.

Opportunities exploitation

Using the random assignment described above one obtains a partition of opportunities among firms summarized by the occupancy vector $(m_1(t), m_2(t), \ldots, m_N(t))$ where $\sum_i m_i(t) = M$ and $m_i(t)$ represents the number of opportunities assigned to firm i. These "business opportunities" can be thought of as the source of micro-shocks affecting the size of firms. This is a rather common prescription for this kind of model, and dates back to the early works of Simon (see e.g. Ijiri and Simon (1977)). We make no assumptions on the actual nature of these shocks and we want to relate "opportunities" to "growth" in the simplest way. Hence, we assume that these micro-shocks are randomly and independently drawn from a distribution with zero mean and fixed variance v_0. Then the total growth of firm i

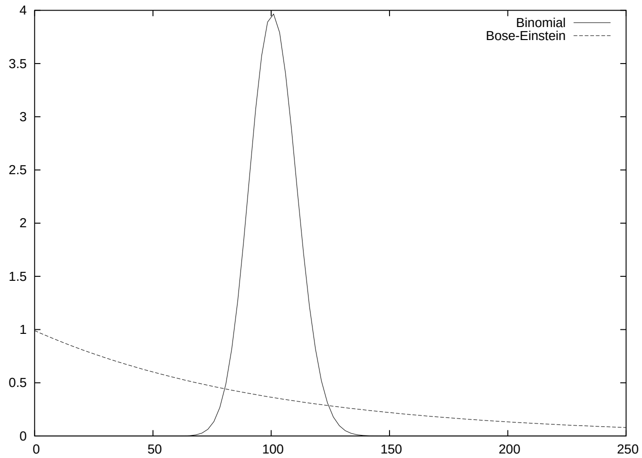

Figure 5: The comparison of a binomial (7) and a Bose-Einstein (6) distribution with $N = 100$ firms and an average number of shocks per firm $M/N = 100$.

is obtained adding $m_i(t)$ independent micro-shocks and if $s_i(t)$ stands for the (log) size of firm i at time t, its growth equation reads

$$s_i(t+1) = s_i(t) + r_i(t) \tag{8}$$

$$r_i(t) = \sum_{j=1}^{m_i(t)} \epsilon_j(t)$$

where ϵ are i.i.d. with a common distribution $f(\epsilon; v_0)$ with variance v_0. The random growth rates r_i are then identically[6] distributed across firms and their unconditional probability density reads

$$p(x; N, M, v_0) = \sum_{h=0}^{M} P(h; N, M)\ f(x; v_0)^{\star h} \tag{9}$$

where $f(x; v_0)^{\star h}$ stands for the h-time convolution of the micro-shocks density (i.e. the distribution of the sum of h micro shocks). The average number of opportunities per firm is M/N and the distribution of growth rates r has mean 0 and variance $v = v_0 M/N$.

At this point it is useful to clarify a few points about the above assumptions. First, concerning the 0 mean hypothesis, notice that the choice of a distribution with a non-zero mean m_0 would simply introduce an industry-wide growth trend

[6] But not independently, due to the global constraint $\sum_i m_i = M$

proportional to Mm_0 which is irrelevant if one is interested in the distribution of growth rates[7]. Second, both the hypotheses of no correlation among micro-shocks and of constant variance in their distribution are working hypotheses introduced to keep the discussion clearer and can be relaxed[8]. Finally, the choice of the shape for the micro-shocks distribution is, as we show in the next Section, completely irrelevant in all the case of interest.

4 The "large industry" limit

In this Section we present analytical results, concerning the growth rates distribution, in the case in which the total number of firms is large and their growth is generated by the assignment of a large number of small shocks. More precisely, we study the model in the limit $M, N \to \infty$ and $v_0 \to 0$. In order to obtain non-trivial results, however, the variance of growth rates must remain finite, and one has to impose the condition that the previous limits are performed in such a way as to keep $v = v_0 M/N$ invariant so that when M and N change the micro-shocks distribution must be rescaled proportionally to N/M. We ensure that micro-shocks can be easily rescaled by using a random variable $\mathbf{x}$ drawn from a probability density $g(x)$ with zero mean and unit variance and by taking the micro-shocks in (8) to be $\epsilon = \sqrt{v_0}\,\mathbf{x}$. The density that appears in (9) then reads

$$f(x; v_0) = \frac{1}{\sqrt{v_0}} g(\frac{x}{\sqrt{v_0}}) \quad . \tag{10}$$

This device ensures that the shape of the micro-shocks distribution is preserved when their scale changes.

Lemma

The limit of the Bose-Einstein distribution defined in (6) when $M, N \to \infty$ with constant $\lambda = M/N$ is a Geometric distribution with mean λ.

Proof

Consider the generic term in (6), expanding the binomial coefficient one obtains

$$P(h; N, M) = \frac{(N + M - h - 2)!}{(N - 2)!\,(M - h)!} \frac{(N - 1)!\,M!}{(N + M - 1)!} \tag{11}$$

that after the expansion of the factorial terms and some easy simplifications reduces to

$$P(h; N, M) = \frac{(N - 1)\,M\,(M - 1)\,\ldots\,(M - h + 1)}{(N + M - 1)\,(N + M - 2)\,\ldots\,(N + M - h - 1)} \tag{12}$$

[7]Notice that in all the empirical studies cited in Section 2 the actual variable under study is the "normalized" growth rate, i.e. the growth rate of market shares.

[8]For instance introducing a mild correlation among micro-shocks or introducing a random variance extracted from a given distribution.

Dividing both numerator and denominator by N^{h+1} the expression can be rewritten as

$$P(h;N,M) = \frac{N-1}{N} \frac{\frac{M}{N}\left(\frac{M}{N}-\frac{1}{N}\right)\ \dots\ \left(\frac{M}{N}-\frac{h-1}{N}\right)}{\left(1+\frac{M}{N}-\frac{1}{N}\right)\left(1+\frac{M}{N}-\frac{2}{N}\right)\ \dots\ \left(1+\frac{M}{N}-\frac{h+1}{N}\right)}. \qquad (13)$$

Substituting $M/N = \lambda$ and taking the limit for $N \to \infty$, the previous equation reduces to the h-th term of a geometric distribution

$$p_h(\lambda) = \frac{\lambda^h}{(1+\lambda)^{h+1}} \qquad (14)$$

Q.E.D.

From the assumptions of the lemma, the limit for $N, M \to \infty$ is obtained with constant total variance v. Recalling the previous definitions, if we now set $v_0 = v/\lambda$ where $\lambda = M/N$ and use (10), we obtain the growth rates distribution (see (9))

$$p(x;\lambda,v,N) = \sum_{h=0}^{\lambda N} P(h;N,\lambda N)\ (\sqrt{\frac{\lambda}{v}} g(\sqrt{\frac{\lambda}{v}}\, x))^{\star h} \qquad (15)$$

whose behavior in the "large industry limit" is provided by the following theorem.

Theorem

Suppose that the density g in (10) possesses all the central moments. Then the summation in (15) converges in the limit for $N, \lambda \to \infty$ to a Laplace distribution with parameter $\sqrt{v/2}$, i.e.

$$\lim_{\lambda,N\to\infty} p(x;\lambda,v,N) = f_L(x;\sqrt{v/2}) = \frac{1}{\sqrt{2v}}\, e^{-\sqrt{2/v}\,|x|} \qquad (16)$$

Proof

Let us start by considering the characteristic function of the growth rates distribution

$$\tilde{p}(k;\lambda,v,N) = \int_{-\infty}^{+\infty} dx\ e^{i\,k\,x}\ p(x;\lambda,N) \qquad (17)$$

The series in (15) is absolutely convergent and one can pass the Fourier integral inside the series to obtain

$$\tilde{p}(k;\lambda,v,N) = \sum_{h=0}^{\lambda N} P(h;N,\lambda,v,N)\ \tilde{g}(\sqrt{\frac{v}{\lambda}}\, k)^h \qquad (18)$$

where $\tilde{g}$ is the characteristic function of the micro-shocks distribution and we used the fact that the characteristic function of the h-convolution of density g is h times its characteristic function.

For any value of $\lambda > 0$ one can evaluate the limit for $N \to \infty$ and use the previous Lemma to obtain:

$$\lim_{N\to\infty} \tilde{p}(k;\lambda,v,N) = \tilde{p}(k;\lambda,v) = \frac{1}{1+\lambda}\sum_{h=0}^{\infty}\left(\frac{\lambda}{1+\lambda}\right)^h \tilde{g}(\sqrt{\frac{v}{\lambda}}\,k)^h. \qquad (19)$$

Since $\tilde{g}$ is a characteristic function it is $|g(x)| \leq 1, \forall x$ (Lemma 1, p.499, Feller (1971)), so that the geometric series in (19) can be resummed to give:

$$\tilde{p}(k;\lambda,v) = \frac{1}{1+\lambda-\lambda\,\tilde{g}(\sqrt{\frac{v}{\lambda}}\,k)}. \qquad (20)$$

If m_h stands for the h-th moment of the g distribution, using (10) and the hypothesis that m_h exist $\forall h$, the infinite series expansion for the rescaled micro-shock distribution reads

$$\tilde{g}(\sqrt{\frac{v}{\lambda}}\,k) = 1 - \frac{1}{2}\,\frac{v}{\lambda}\,k^2 + \sum_{h=3}^{+\infty}\frac{1}{h!}\,m_h\,\left(\frac{v}{\lambda}\right)^{\frac{h}{2}}\,(i\,k)^h \qquad (21)$$

under the assumption that $m_2 = 1$. Substituting this expansion in (20) and keeping only the leading term in $1/\lambda$ we obtain:

$$\tilde{p}(k;\lambda,v) = \frac{1}{1+\frac{v}{2}k^2 + o(1/\lambda)} \overset{\lambda\to+\infty}{\longrightarrow} \frac{1}{1+\frac{v}{2}k^2} \qquad (22)$$

Since the limit function in (22) is continuous at the origin we can conclude that the original series in (15) converge to a proper density function (Theorem 2, p. 508, Feller (1971))[9].

Finally, substituting the definition of the Laplace density (1) in (17) and performing the integral it is easy to check that the last expression is the characteristic function of a Laplace distribution with parameter $a = \sqrt{v/2}$.

Q.E.D.

5 Simulations

The above theorem ensures that, when the number of firms and opportunities goes to infinity, the growth rates distribution generated by the model converges to the

[9] The requirement of continuity in 0 is needed in order to assure that the limit density function is not defective.

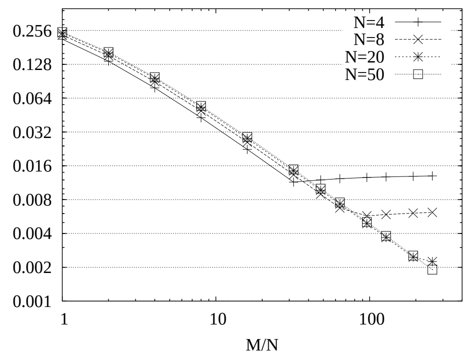

Figure 6: The absolute deviation D as a function of the average number of micro-shocks per firm M/N for different values of N. Micro-shocks are normally distributed.

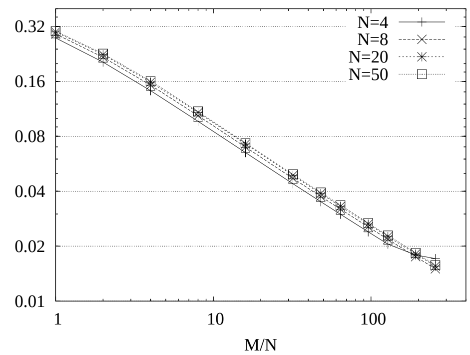

Figure 7: The absolute deviation D as a function of the average number of micro-shocks per firm M/N for different values of N. The micro-shocks distribution is a Gamma with shape parameter 1.

Laplace. If the parameters values estimated from the data were sufficiently near to the asymptotic region considered in the theorem, the model would reproduce, with a sufficiently high "degree of precision", our empirical observations.

The aim of the present Section is to provide an estimate of the fitness of our model to the empirical findings of Sec. 2. We want to verify that, setting the model parameters to the "typical" values suggested by real data, one is able to obtain, via simulations, the same kind of exponential behavior for the growth rates distribution found in empirical investigations.

First of all, we consider the probability distribution of our model $F_{model}(x; N, M)$, derived from the density defined in (9) and with micro shocks as defined in (10). We compare it with the Laplace $F_L(x)$ using different values of the parameters defining the model, namely the total number of firms composing the sector N, the total number of opportunities M, and the shape of their distribution g. Since the variance of the growth rates can be perfectly reproduced with a tuning of the micro-shocks variance, we consider unit variance distributions and set $v_0 = N/M$.

We use the maximum absolute deviation as a measure of this agreement:

$$D(N, \lambda) = max_{-\infty<x<+\infty} \left| F_{model}(x; N, N\,\lambda) - F_L(x) \right| \tag{23}$$

where $\lambda = M/N$ is the average number of opportunities per firm. The values of D for different values of N and λ are plotted in Fig. 6 and Fig. 7 for, respectively, Gaussian and Gamma distributed micro-shocks. For a fixed value of N, as λ increases, the value of D decreases and eventually reaches a constant asymptotic value. This constant regime, for the lower values of N, can be clearly seen in Fig. 6.

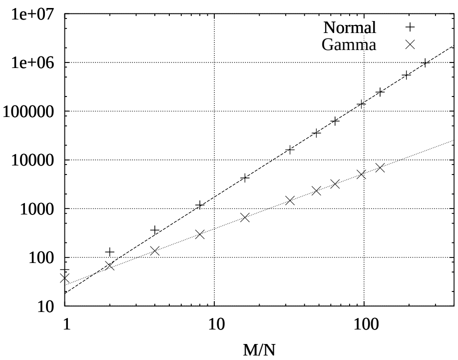

Figure 8: $N^\star$ as a function of $\lambda = M/N$ for Normal and Gamma distributed micro-shocks. The straight lines are obtained using (24) to approximate the value of D.

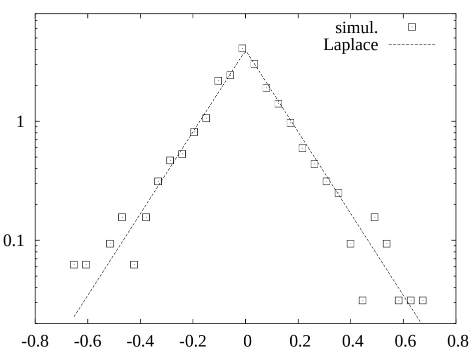

Figure 9: The model growth rates probability density with $N = 100$, $\lambda = M/N = 16$ and micro-shocks normally distributed. The statistics is collected over 7 independent realizations to provide direct comparability to the empirical plots in Sec. 2. The theoretical Laplace density with unit variance ($a = 1/\sqrt{2}$) is also shown.

When N is large, the curves for different values of N collapse and D becomes a function of λ only. This can be directly checked observing that the curves for the largest values of N are basically superimposed in both Fig. 6 and Fig. 7. The variation of D as a function of λ, until the constant regime sets in, is perfectly fitted by a power-like behavior

$$\log(D(N \gg 1, \lambda)) \sim a \log(\lambda) + b \tag{24}$$

The fitted parameters on a population of $N = 500$ firms are $a = -0.970 \pm 5.210^{-3}$, $b = -0.84 \pm 2.210^{-2}$ for the Gaussian case and $a = -0.560 \pm 2.610^{-3}$, $b = -1.04 \pm 1.110^{-2}$ for the Gamma. The remarkable increase in the value of D for the latter is due to the large skewness of the distribution.

Using the absolute deviation D one is able to measure the agreement between the model and the Laplace distribution as a function of the parameters values. More precisely, one can compare the value of D obtained in (23) with the Kolmogorov-Smirnov statistics, i.e. the maximum absolution deviation $D^\star(N)$ between the empirical distributions of two finite samples of size N composed by independent random variables extracted from a Laplacian. With this comparison we are able to assess in what region of the parameters space the two distributions, F_L and F_{model} , provide analogous results (with respect to a Kolmogorov-Smirnov test performed over a sample of a given size).

In particular, we can compute the number of observation[10] $N^\star(D,r)$ needed to reject, with a given significance r, the null hypothesis of Laplacian distribution if a maximum absolute deviation D is observed. If one chooses for r the standard value of .05 the behavior of $N^\star(D(N >> 1, \lambda), .05)$ as a function of λ is reported in Fig. 8 for the same micro-shocks distributions, Normal and Gamma, used in Fig. 6 and Fig. 7. Even for relatively small values of $\lambda \sim 20$, the discrepancies between our model and the Laplace distribution can only be statistically revealed using rather large samples, with more than 500 observations.

We now turn to study the fitness of our model with the empirical findings presented in Sec. 2. Notice that the exact tuning of all the parameters of the model is not feasible. While the number N of firms in the different sectors is known, and the final variance v can be obtained from the relevant samples, nothing can be said about the number of opportunities M and the shape of the related shocks g, whose nature was left largely unspecified in the definition of the model. Nevertheless, since the number of firms per sector is high (on average $N = 130$), one can use the N large limit in (24) for an estimate of the discrepancy to obtain a lower bound on the value of λ for different micro-shock densities g. From Fig. 7 it appears that even with skewed shapes a number of opportunities per firm ≥ 20 is enough to obtain a rather good agreement: indeed the typical number of observations for a sector ~ 1000 lies near or above the considered $N^\star$.

Finally, we check the fitness in a specific case. We run simulations of the model with $N = 100$ and $\lambda = 16$. Results reported in Fig. 9 show that the agreement is good even if the value of λ is low.

We can conclude that the model presented in Section 3 does generate, for a wide range of parameters values, growth rates distributions that are almost identical to the Laplace, hence providing a complete explanation of this "stylized fact".

6 Final Notes and Outlook

In the present paper we add crucial evidence in support of the tent-shape of the firm growth rates distribution, extending previous findings with respect to two directions. First we replicate the analysis already performed by Stanley et al. (1996) and Bottazzi et al. (2001) respectively on COMPUSTAT and PHID databases, on a new databank (MICRO.1) covering many firms in the Italian manufacturing industry. Second, using data disaggregated by sector, we prove that the shape

[10] This number can be written

$$N^\star(D,r) = \frac{1}{2}\left(Q_{ks}^{-1}(r)/d - .12 + \sqrt{(Q_{ks}^{-1}(r)/d - .12)^2 - .44}\right)^2$$

where Q_{ks} is the distribution of the Kolmogorov-Smirnov statistics (see e.g. Press et al. (1993)).

of these distributions is not a mere effect of aggregation. Although intersectoral differences clearly arise, we conclude, in line with previous studies, that the tent-shape distribution of corporate growth rates appears as an extremely robust feature of the manufacturing industry, characterized by a higher regularity than the one shown by size distributions.

A possible explanation of this impressive 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 well described, in general, by a stochastic model, apart from very rare events, ranging from earthquakes to the discovery of balance sheet accounting frauds, that have nothing to do with the "generic" economic behavior. 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 rate distribution remain small and proportional to their sheer number.

Notice however that, contrary to what reported in Stanley et al. (1996) and Amaral et al. (1997), preliminary studies on MICRO.1 seem to suggest that using different definitions of firm size, like, for instance, number of employees or value added, would introduce relevant differences in the shape of the distributions.

We also present a model of the growth dynamics of firms. The model clearly originates in the Simon inspired literature on firm dynamics with which it shares two central features. First, different firms are viewed as different realizations of the same stochastic process. Second, the model includes a very simple idea of competition among firms that try to seize the most out of a scarce resource.

The essential novelty lies in the use of a "non trivial" assignment of business opportunities to firms described by the Bose-Einstein statistics. This statistics introduces a sort of "attracting force" between the various opportunities that tends to group them in bigger chunks leading to the appearance of two noticeable properties: the presence of a fat tail, which indicates a more likely presence of extremely large number of "opportunities" assigned to a single firm and the absence of a natural "scale" in the number of opportunities, hinted by the 0 value of the distribution mode.

The ability of the model to reproduce empirical findings without requiring a fine tuning of the parameters is ensured by the Theorem in Sec. 4 and constitutes its main strength.

The analysis presented in this paper can be extended following different lines of investigation. Concerning the empirical evidence, both the intersectoral diversity and the sensitivity to the definition of size suggest the need for further investigation aimed to disentangle presumably different degrees of "lumpiness" in the dynamics of different variable . On the other hand, the proposed model can be possibly extended to include these sources of "heterogeneity" and specific tests, having more discriminatory power, can be developed to better decide its agreement with the actual shapes of the observed data.

References

Amaral L. A. N., Buldyrev S.V., Havlin S., Leschhorn H., Maass F., Salinger M.A., Stanley H.E. and Stanley M.H.R. (1997) “Scaling Behavior in Economics: I. Empirical Results for Company Growth”, J. Phys. I France, 7, 621-633

Amaral L. A. N., Gopikrishnan P., Plerou V., Stanley H.E. (2001) “A model for the growth dynamics of economic organizations”, Physica A, 299, 127-136

Bottazzi G., Cefis E., Dosi G. (2002) “Corporate Growth and Industrial Structure. Some Evidence from the Italian Manufacturing Industry” forthcoming on “Industrial and Corporate Change”.

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

Feller W. (1968) “An Introduction to Probability Theory and Its Applications” Vol.I, Third Edition, Wiley & Sons, N.Y.

Feller W. (1971) “An Introduction to Probability Theory and Its Applications” Vol.II, Second Edition, Wiley & Sons, N.Y.

Geroski P.A., (2000) “Growth of Firms in Theory and Practice” in Foss N. and Malinke V. Eds., “New Directions in Economics Strategy Research”, Oxford University Press, Oxford, U.K.

Gibrat R., (1931) “Les inégalités économiques”, Librairie du Recueil Sirey, Paris.

Hall B. H., (1987) “The Relationship Between Firm Size and Firm Growth in the US Manufacturing Sector”, Journal of Industrial Economics, 35, 4

Hart P.E. and Prais S.J. (1956) “The Analysis of Business Concentration”, Journal of the Royal Statistical Society, pp.150-191

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

Ijiri y., Simon H.A., (1977) “Skew distributions and the sizes of business firms”, North Holland Publishing Company.

Press W.H., Teukolsky S.A., Vetterling W.T. and Flannery B.P. “Numerical recipes in C” 2nd edition, Cambridge University Press

Stanley M. H. R., Nunes Amaral L. A., Buldyrev S. V., Havlin S., Leschhorn, H., Maass P., Salinger M. A., Stanley H. E., (1996) “Scaling behaviour in the growth of companies”, Nature, 379.

Simon H.A. and Bonini C.P. (1958) “The Size Distribution of Business Firms” American Economic Review XLVIII, pp. 607-617.

Subbotin M.T., (1923) “On the Law of Frequency of Errors” Matematicheskii Sbornik, 31, pp. 296-301.

Sutton J., (1997) “Gibrat’s Legacy” Journal of Economic Literature XXXV

Sutton J., (1998) “Technology and Market Structure, Theory and History”, MIT Press, Cambridge, Ma.