Economics Bulletin

On the relationship between firms' size and growth rate

Giulio Bottazzi
Scuola Superiore Sant'Anna

Abstract

Since the seminal work of Pareto, many empirical analyses suggestedthat the distribution of firms size is characterized by an asymptoticpower like behavior. At the same time, several investigations showthat the distribution of annual growth rates of firms displays aremarkable double-exponential shape. Recently it has been suggestedthat both these statistical properties can be explained by assuming abivariate Marshall-Olkin power-like distribution for the size of firmsin subsequent time steps. Through analytical investigation, I showthat the marginal distribution of growth rates implied by thisassumption does not possess, in general, a Laplace shape and becomesdegenerate when subsequent size levels are perfectly correlated. Assuch, the bivariate Marshall-Olkin distribution is unable to properlyaccount for the observed regularities. The original suggestion isfaulty as it treats firm size levels as stationary stochasticvariables and neglects their integrated nature.

I would like to thank Angelo Secchi for a number of helpful comments. Research support by Scuola Superiore Sant'Anna (grant n.E6006GB) and by European Commission (CO3 NEST Pathfinder project, 6-th Framework program) is thankfully acknowledged.
Citation: Bottazzi, Giulio, (2008) "On the relationship between firms' size and growth rate." *Economics Bulletin,* Vol. 3, No. 8 pp. 1-7
Submitted: June 7, 2007. **Accepted:** February 1, 2008.
URL: http://economicsbulletin.vanderbilt.edu/2008/volume3/EB-07C00002A.pdf

1 Introduction

Since the seminal work of Pareto on wealth distribution of households, many empirical studies suggested that, at least above some minimal size threshold, the distribution of a number of economic relevant variables (like wealth, income, revenues or capital assets) follow a power like behavior (see Kleiber and Kotz (2003) for an extended review). This statistical fact is commonly named, after his early discoverer, "Pareto Law". At the same time, recent investigations show that the distribution of annual growth rates of business firms displays a remarkable double-exponential shape, called "Laplace" distribution (see Bottazzi and Secchi (2006) and reference therein). This result is robust to different levels of aggregation and to the use of different variables (revenues, number of employees, value added, etc.) to proxy firm size.

The question naturally arises if these two distinct "stylized facts" can be somehow reconciled so that a more "unifying" theoretical view of the dynamics of firms is recovered. This is certainly a relevant issue, since, as vigorously pointed out in Brock (1999), moving the analysis from "unconditional" quantities, like distributions of random variables, to more "conditional" (if not causal) relationships between the different regularities will obviously improve our understanding of the underlying drivers of the observed dynamics.

A recent letter (Palestrini, 2007) tries to directly address the "unifying" question by proposing a possible relationship between the power-like nature of the distribution of firm sizes and the Laplace character of the firm growth rates distribution. The author assumes a joint Pareto distribution for the size of firms in two different time steps, shaped according to the bivariate distribution proposed in Marshall and Olkin (1967). Then, he derives the tail behavior of the implied distribution of growth rates. From his asymptotic analysis, he wrongly concludes that the implied distribution does, in general, "belong to the Laplace family" (Palestrini, 2007, p.370). After briefly reviewing, in Section 2, the property of the Marshall-Olkin bivariate distribution, in Section 3 I will show how to derive, under the same hypothesis of Palestrini (2007), the *exact* distribution of growth rates. As my computation reveals, the implied distribution is not, in general, Laplacian nor everywhere continuous, having a finite atomic component for null growth rates. In Section 4 I will briefly discuss the implication of my findings on the possibility to reconcile the Pareto distribution of size and the observed Laplace density of growth rates.

2 The Marshall-Olkin bivariate distribution

Abstracting from precise economic definitions, let $\mathbf{S}_1$ and $\mathbf{S}_2$ be the size of a firm in two successive time steps. If the distribution of firm size is Paretian, the distribution of its logarithm follows an exponential distribution. Then, taking $\mathbf{s}_1 = \log(\mathbf{S}_1)$ and $\mathbf{s}_2 = \log(\mathbf{S}_2)$, the validity of the Pareto laws implies that $\mathbf{s}_1$ and $\mathbf{s}_2$ are exponentially distributed

$$\log\left(\text{Prob}\left\{\mathbf{s}_i > x\right\}\right) \sim x \quad i = 1, 2 \, .$$

Consider the joint distribution of the couple of random variables $(\mathbf{s}_1, \mathbf{s}_2)$. For the Pareto Law to be valid, this distribution should possesses exponential marginals. As suggested in Palestrini (2007), a natural candidate for a distribution of this kind is constituted by the multivariate exponential distribution proposed in Marshall and Olkin (1967), which in the bivariate case reads

$$\begin{aligned} \text{Prob}\left\{\mathbf{s}_1 > s_1, \mathbf{s}_2 > s_2\right\} &= 1 - F(s_1, s_2) \\ &= \exp\left\{-\lambda_1 s_1 - \lambda_2 s_2 - \lambda_{12} \max\{s_1, s_2\}\right\} \, . \end{aligned} \tag{1}$$

It is immediate to see that, according to (1), the marginal distribution for the random variables $\mathbf{s}_1$ and $\mathbf{s}_2$ are exponential with parameter $\bar{\lambda}_1 = \lambda_1 + \lambda_{12}$ and $\bar{\lambda}_2 = \lambda_2 + \lambda_{12}$, respectively. Hence, the expression for expected value and the variance of the two random variables immediately follow. The covariance between $\mathbf{s}_1$ and $\mathbf{s}_2$ is given by

$$\text{Cov}(\mathbf{s}_1, \mathbf{s}_2) = \frac{\lambda_{12}}{\lambda \, \bar{\lambda}_1 \, \bar{\lambda}_2} \, , \tag{2}$$

where $\lambda = \lambda_1 + \lambda_2 + \lambda_{12}$. Thus, the associated correlation coefficient becomes equal to λ_{12}/λ and is different from zero as long as $\lambda_{12} \neq 0$. As discussed in Marshall and Olkin (1967), the distribution F defined in (1) is not everywhere continuous. The atomic component comes from the presence of the discontinuous "max" term in the exponential argument. The presence of this term will play a major role in the subsequent analysis. For easy of reference I report here the expression for the generating function (Laplace transform) of the distribution F derived in Marshall and Olkin (1967). I will use it in the first Theorem of the next Section. It reads

$$\psi(\omega_1,\omega_2) = \int_0^{+\infty}\int_0^{+\infty} e^{-s_1\omega_1 - s_2\omega_2} dF(s_1,s_2) \qquad (3)$$

$$= \frac{\bar\lambda_1\bar\lambda_2(\lambda+\omega_1+\omega_2)+\omega_1\omega_2\lambda_{12}}{(\lambda+\omega_1+\omega_2)(\bar\lambda_1+\omega_1)(\bar\lambda_2+\omega_2)} \; .$$

3 The distribution of growth rates

In this Section the analytical expression of the distribution of the (logarithmic) growth rate is derived under the assumption that the firm log size is distributed, in two successive time steps, according to the bivariate Marshall-Olkin distribution defined in (1). The firm growth rate r over a given period of time is, by definition, the difference of the logarithm of the size at the end and at the beginning of said period. Then, using the notation introduced above, one has $\mathbf{r} = \mathbf{s}_2 - \mathbf{s}_1$. Let $G(r) = \text{Prob}\,\{\mathbf{r} \le r\}$ be the probability distribution of growth rates and $\tilde g(k) = \text{E}[\exp i\,k\,\mathbf{r}]$ the associated characteristic function. Using the generating function in (3) one can easily derive the expression for $\tilde g(k)$. One has the following

Theorem 1 *If $\mathbf{s}_1$ and $\mathbf{s}_2$ follow a bivariate Marshall-Olkin distribution the characteristic function $\tilde g(k)$ of their difference $\mathbf{r} = \mathbf{s}_2 - \mathbf{s}_1$ is given by*

$$\tilde g(k) = \frac{\lambda\,\bar\lambda_1\,\bar\lambda_2 + \lambda_{12}\,k^2}{\lambda\left(\bar\lambda_1 - ik\right)\left(\bar\lambda_2 + ik\right)} \; . \qquad (4)$$

PROOF. Considering the distribution function $G(r)$ one has

$$G(r) = \text{Prob}\,\{\mathbf{r} \le r\} = \text{Prob}\,\{\mathbf{s}_2 - \mathbf{s}_1 \le r\}$$

$$= \int_0^{+\infty}\int_0^{+\infty} \theta(r - s_1 + s_2) dF(s_1,s_2)$$

where θ is the right continuous Heaviside theta function, which is equal to 1 if its argument is positive or null and zero otherwise. From the definition of characteristic function it is

$$\tilde g(k) = \int_{-\infty}^{+\infty} e^{ikr}\, dG(r) \; ,$$

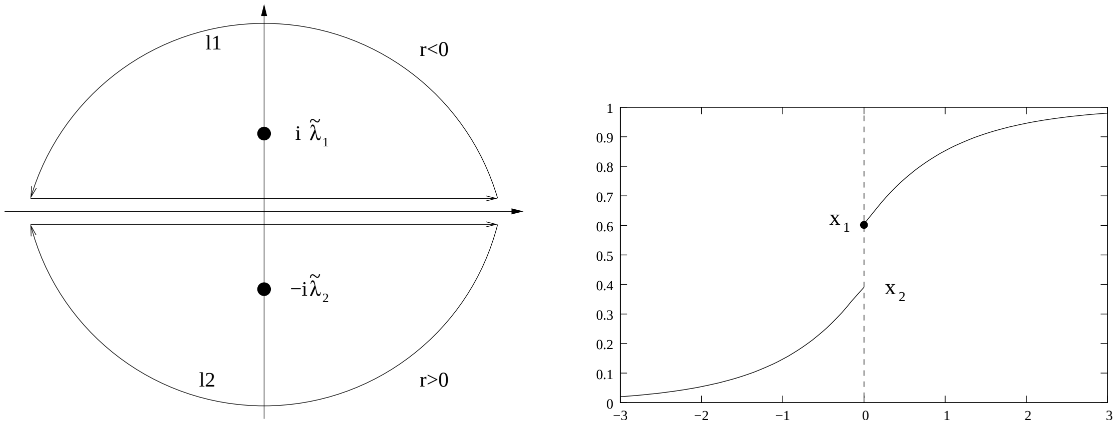

Fig. 1. **Left:** Complex plane for the anti-transform of the characteristic function. The poles are shown together with the paths of the Cauchy integral l1 and l2 used when $r < 0$ and $r > 0$, respectively. **Right:** Growth rates distribution function G for $\lambda_1 = 1$, $\lambda_2 = 1$ and $\lambda_{12} = .5$.

substituting the previous expression and inverting the order of integrations one has

$$\tilde{g}(k) = \int_0^{+\infty} \int_0^{+\infty} e^{ik(s_1 - s_2)} \, dF(s_1, s_2)$$

which, remembering (3), gives $\tilde{g}(k) = \psi(-ik, ik)$. Finally, direct substitution of (3) proves the assertion. □

The term proportional to k^2 in the numerator of $\tilde{g}(k)$ in (4) is what makes the implied growth rate distribution not Laplacian. Only in the case in which $\lambda_{12} = 0$ the expression in (4) reduces to the characteristic function of an asymmetric (or symmetric, if λ_1=λ_2) Laplace distribution (Kotz et al., 2001, p.141). As shown in the next Theorem, the effect of this term is to introduce a finite probability for the occurrence of zero growth rates, that is an atomic component in the point $r = 0$. In general, one has the following

Theorem 2 *If* $\mathbf{s}_1$ *and* $\mathbf{s}_2$ *follow a bivariate Marshall-Olkin distribution the distribution function* $G(r)$ *of their difference* $\mathbf{r} = \mathbf{s}_2 - \mathbf{s}_1$ *is given by*

$$G(r) = \begin{cases} e^{\bar{\lambda}_1 r} \frac{\lambda_2}{\lambda} & \text{if} \quad r < 0 \,, \\ 1 - e^{-\bar{\lambda}_2 r} \frac{\lambda_1}{\lambda} & \text{if} \quad r \geq 0 \,. \end{cases} \tag{5}$$

PROOF. Consider the formal definition of the density $g(r) = G'(r)$ as the anti-Fourier transform of the characteristic function

$$g(r) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dk \, e^{-ikr} \, \tilde{g}(k) \ .$$

Since $\lim_{k\to\infty} |\tilde{g}(k)| = 0$ on the whole complex plane, for Jordan's lemma the previous integral, and consequently the density function, exists for any $r \neq 0$. The characteristic function $\tilde{g}(k)$ possess two simple poles on the imaginary axis, in $i\bar{\lambda}_1$ and $-i\bar{\lambda}_2$. Then, the previous expression can be written as a Cauchy integral on the upper or lower half plane when the value of r is respectively lower or greater then zero. The two closed curves are depicted in Fig. 1 (left panel). According to Cauchy integral theorem, in each case the value of the integral is proportional to the residue of the function computed in the internal pole. After a little algebra one has

$$g(r) = \begin{cases} \frac{\bar{\lambda}_1 \lambda_2}{\lambda} \, e^{\bar{\lambda}_1 r} & \text{if} \quad r < 0 \ , \\ \frac{\bar{\lambda}_2 \lambda_1}{\lambda} \, e^{-\bar{\lambda}_2 r} & \text{if} \quad r \geq 0 \ . \end{cases}$$

Using the previous expression, the distribution function G can be computed as

$$G(r) = \int_{-\infty}^{r} dr' \, g(r') \quad \text{if} \quad r < 0$$

or

$$G(r) = 1 - \int_{r}^{+\infty} dr' \, g(r') \quad \text{if} \quad r > 0 \ .$$

Since the distribution function is by definition right continuous, the assertion follows. □

An example of the shape of the distribution G is reported in Fig. 1 (right panel). Notice that the finite weight at $r = 0$ can be easily computed using (5). Indeed one has

$$\text{Prob}\,\{\mathbf{r} = 0\} = \lim_{\delta\to 0^+} G(\delta) - G(-\delta) = \frac{\lambda_{12}}{\lambda} \ , \qquad (6)$$

that is the discontinuity in zero of the distribution function is proportional to the correlation coefficient of the two variables $\mathbf{s}_1$ and $\mathbf{s}_2$.

4 Conclusions

The exact distribution function $G(r)$ and probability density $g(r)$ of firm growth rates have been obtained under the assumption that firm size follows, in two subsequent time steps, a bivariate Marshall-Olkin distribution. Contrariwise to what reported in Palestrini (2007), this assumption implies, in general, a growth rates distribution which is not compatible with a Laplace, symmetric or asymmetric, shape. The source of this incompatibility is the discontinuous nature of the distribution function $G(r)$, which possesses an atomic probability weight in zero. The discontinuity at the origin is proportional to the correlation coefficient between the size levels $\mathbf{s}_1$ and $\mathbf{s}_2$. This implies that a bivariate Marshall-Olkin distribution for the logarithm of firm sizes could generate a Laplace distribution of growth rates only in the very particular case in which firm sizes at subsequent time steps were uncorrelated. However, as many empirical studies have shown, the correlation coefficient of subsequent size levels is in general very near (and often statistically equal) to one. Consequently, in real cases, the assumption that logarithmic sizes follow a Marshall-Olkin distribution would give a sort of degenerate growth rates density, having almost the entire probability weight in the origin. This fact greatly reduces the possibility of this distribution to ever provide an effective statistical description of empirical data.

Summarizing, the unification of the two "stylized facts" concerning the Pareto distribution of firms size and the Laplace density of growth rates proposed in Palestrini (2007) would be possible only disregarding the integrated (or unit-root) nature of firm's growth process, which constitutes an extremely robust regularity. In general, indeed, one cannot consider the logarithm of the size of the firm at subsequent time steps as generated by a stationary random process. It is the difference [1] of these logarithms, rather than their levels, which are likely to display, at least some degree of, stationarity.

The lesson to be learned is that in order to obtain a reliable phenomenological description of the growth dynamics of firms, or, to that extent, of any economic

[1] Or some other more complicate manipulation of the original variables in the case of fractional integration. The discussion of this point is outside the scope of the present note.

process, one has to start from data and carefully investigate their regularities while, at the same time, abstaining himself, as far as possible, from introducing untested theoretical hypothesis.

References

Bottazzi, G., A. Secchi, (2006) Explaining the Distribution of Firms Growth Rates, *Rand Journal of Economics*, 37, pp. 234-263.

Brock W. (1999) Scaling in Economics: a Reader's Guide, *Industrial and Corporate Change* 8, 409-446.

Marshall, A. and Olkin, I. (1967) A Multivariate Exponential Distribution *Journal of the American Statistical Association* , 62, 30-44.

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

Kotz, S., Kuzubowski T.J. and Podgórski K. (2001) *The Laplace Distribution and Generalizations*, Birkhauser, Boston.

Palestrini, A. (2007) Analysis of industrial dynamics: A note on the relationship between firms' size and growth rate, *Economic letters*, 94, pp. 367-371.