

THE GROWTH OF BUSINESS FIRMS: FACTS AND THEORY

Sergey V. Buldyrev
Yeshiva University

Jakub Growiec
Warsaw School of Economics and
CORE

Fabio Pammolli
University of Florence and IMT,
Institute for Advanced Studies

Massimo Riccaboni
University of Florence

H. Eugene Stanley
Boston University

Abstract

We refer to the framework developed by Ijiri and Simon (1977) and to the notion of independent submarkets (Sutton 1998) to provide a simple candidate explanation for the shape of the firm growth distribution based on a model of proportional growth at the level of both the introduction of new products by firms and their size dynamics. We exploit the features of a unique longitudinal data set which covers the entire distribution of products and firms in the worldwide pharmaceutical industry to test the model at different levels of aggregation as well as at different time lags. Econometric investigations show that the model's predictions are in good agreement with empirical evidence. (JEL: L11, L65)

1. Introduction

According to the so-called Gibrat's "Law of Proportionate Effect," the expected value of the growth rate of a business firm is proportional to its current size. Models of proportional growth have been extensively applied to account for the relation between the size and expected growth rate of firms (see Sutton 1997; Rossi-Hansberg and Wright 2005). In the last decade, models in this tradition have been critically revisited and generalized by researchers in the "econophysics" community who found that the growth distribution of firms is not Gaussian as predicted

Acknowledgments: Authors are listed in alphabetical order. We thank Xavier Gabaix, Dongfeng Fu, John Sutton, and Kazuko Yamasaki for helpful discussions and suggestions. The Merck Foundation (EPRIS Program) is gratefully acknowledged for financial support.

E-mail addresses: Buldyrev: buldyrev@yu.edu; Growiec: growiec@core.ucl.ac.be; Pammolli: pammolli@gmail.com; Riccaboni: riccaboni@unifi.it; Stanley: hes@bu.edu

by Gibrat's law of proportionate effect but displays heavy tails.¹ Following the seminal paper by Stanley et al. (1996), heavy-tailed distributions have been extensively applied to describe growth rates of economic entities as diverse as GDP (Canning et al. 1998), returns in financial markets (Gabaix et al. 2006), business firm sales (Bottazzi et al. 2001), and wealth and income (Chatterjee, Yarlagadda, and Chakrabarti 2005).

In this paper, we show that the growth distribution of firms can be generated by a simple stochastic model that builds upon two assumptions in Ijiri and Simon (1977):

1. The probability that a new business opportunity is taken up by any particular existing firm is proportional to the current size of the firm.
2. The probability that a new firm enters the market by capturing a new opportunity is constant over time.

In our model, we introduce a further assumption according to which each market opportunity is assumed to grow in size by an independent random factor, that is, the market consists of a number of independent submarkets, each large enough to support a single investment opportunity (Sutton 1998, 2007).

As shown in Klette and Griliches (2000), Gibrat's law in the Ijiri and Simon (1977) framework can be related to a fully-specified model of endogenous firm growth based on optimising agents. Klette and Kortum (2004) challenge the firm-level evidence with an endogenous R&D-based growth model with heterogeneous innovating firms, Poisson processes ruling R&D output, and quality ladders for differentiated products.

Our specific contribution with this paper is to show that the Gibrat's law combined with the notion of independent submarkets (Sutton 1998, 2007) can account for the shape of the growth distribution.

Because it is difficult to discriminate among many equally plausible alternative generative mechanisms, rigorous empirical tests are needed to ascertain the stochastic properties of growth processes in specific industrial settings, controlling for different regimes of entry, exit, information diffusion, and turnover. Against this background, in Section 3 we perform a thorough econometric investigation based on a unique data set in which we have reconstructed the entire product and firm size distributions for the worldwide pharmaceutical industry. Entry and exit of both products and firms are monitored. The availability of fine-grained data sustains an accurate econometric investigation that refers to the Gibrat benchmark and, at the same time, captures some significant departures from it.

1. Econophysics is a neologism that denotes the activities of a multidisciplinary community who are working on economics problems to test a variety of new conceptual approaches deriving from the physical sciences.

2. The Model

We model business firms as portfolios of goods that they produce (see Klette and Kortum 2004). It follows that the size of each firm is a sum of the sizes of its products.

Two key assumptions in the model are the following (see also Ijiri and Simon 1977; Sutton 1997, 2007):

1. Each firm α consists of $K_\alpha(t)$ products. At time $t = 0$ there are N_0 firms consisting of a total of n_0 products. At each time step t , a new product is introduced. Thus, the number of products at time t is $n_t = n_0 + t$.
2. With probability b , a new product is assigned to a new firm, whereas with probability $1 - b$, a new product is assigned to an incumbent firm α with probability $P_\alpha = (1 - b)K_\alpha(t)/n_t$.

Moreover, we assume that each product grows in time according to an independent random process.

3. At time t , product $i \in \{1, \dots, K_\alpha(t)\}$ of firm α has size $\xi_{\alpha i}(t) > 0$, where K_α and $\xi_{\alpha i}$ are independent random variables. At time $t + 1$, the size of each product is decreased or increased by a random factor $\eta_{\alpha i}(t) > 0$, so that $\xi_{\alpha i}(t + 1) = \xi_{\alpha i}(t) \eta_{\alpha i}(t)$. The growth rate $\eta_{\alpha i}$ is independent of K_α and $\xi_{\alpha i}$ and is independent over time.

Based on the first set of assumptions, we derive $P(K)$, the probability distribution of the number of products by firm. Then, using $P(K)$ and the third assumption, we calculate the probability distribution of growth rates $P(g)$. Because the exact analytical solution of $P(K)$ is not known, we provide an approximate mean-field solution for $P(K)$ starting from two limiting cases.²

In the limit of large t when $b = 0$, the distribution of K converges to an exponential distribution

$$P_1(K) = \lambda^K \frac{1}{K(t) - 1} \approx \frac{1}{K(t)} \exp\{-K/K(t)\}, \quad (1)$$

where $\lambda = 1 - 1/K(t)$ and

$$K(t) = [(n_0 + t)/n_0]^{1-b} \cdot n_0/N_0$$

is the average number of products by firm at time t (see Yamasaki et al. 2006).

2. For simplicity, we do not explicitly consider exit here. In reality, products enter and exit; for a growing economy, on average, positive net entry rate can be assumed and proxied by a positive constant.

On the contrary, when $b > 0$ the distribution of products at large t converges to a power law with an exponential cutoff

$$P_{\Pi}(K) \sim K^{-\varphi} f(K), \quad (2)$$

where $\varphi \approx 2 + b$ for small b .

If we neglect the influx of the new products and consider $K_{\alpha} = K_{\alpha}(t + 1) = K_{\alpha}(t)$, then the resulting distribution of the growth rates of all firms is determined by

$$P_g(g) \equiv \sum_{K=1}^{\infty} P(K) P_g(g | K), \quad (3)$$

where $P(K)$ is the distribution of the number of products by firm, computed in the previous stage of the model, and $P_g(g | K)$ is the conditional distribution of growth rates of firms with a given number of products.

The conditional distribution of the logarithmic growth rates $P_g(g | K)$ for the firms consisting of a fixed number of products converges to a Gaussian distribution for $K \rightarrow \infty$:

$$P_g(g | K) \approx \frac{\sqrt{K}}{\sqrt{2\pi} V_g} \exp\{(g - \bar{g})^2 K / 2V_g\}, \quad (4)$$

where V_g is a function of parameters of the distributions $P_{\xi}(\xi)$ and $P_{\eta}(\eta)$, and \bar{g} is mean logarithmic growth rate of a unit, $\bar{g} = \langle \ln \eta_{\alpha i} \rangle$.

In Fu et al. (2005), we find analytical solutions for certain limiting cases. In particular, if $b = 0$ and $t \rightarrow \infty$ then the distribution of $P(K)$ is dominated by the exponential distribution and $K(t) \rightarrow \infty$ as defined by equation (1). In this case, $P_g(g)$ for $g \ll \sqrt{V_g}$ can be approximated by

$$P_g(g) \approx \frac{\sqrt{K(t)}}{2\sqrt{2V_g}} \left(1 + \frac{K(t)}{2V_g} g^2\right)^{-3/2}. \quad (5)$$

In the case in which $b > 0$ and $K \rightarrow \infty$, the distribution $P(K)$ behaves as a power law. The approximate solution of $P_g(g)$ is obtained by using equation (4) for $P_g(g | K)$, the mean-field solution of $P(K)$ in the limit $t \rightarrow \infty$, and replacing summation by integration in equation (3). In the $b \rightarrow 0$ case, the solution is

$$P_g(g) \approx \frac{2V_g}{\sqrt{g^2 + 2V_g} (|g| + \sqrt{g^2 + 2V_g})^2} \quad (b \rightarrow 0), \quad (6)$$

a distribution that combines a double exponential body for $g \rightarrow 0$ and a power law decay for $g \rightarrow \infty$.³

3. Note that due to replacement of summation by integration in equation (3), the approximation equation (6) holds only for $g < \sqrt{V_g}$.

TABLE 1. Exponential mixtures of Gaussian distributions.

ψ	Probability density function (PDF)
$\psi > 1$	Exponential power with shape parameter $\alpha \in (0, 1)$
$\psi = 1$	Laplace, $P(g) = \frac{1}{2} \sqrt{\frac{2\lambda}{V}} \exp \left\{ -\sqrt{\frac{2\lambda}{V}} g \right\}$
$0 < \psi < 1$	Exponential power with shape parameter $\alpha \in (1, 2)$
$\psi = 0$	Gaussian, $P(g) = \frac{1}{\sqrt{2\pi V}} \exp \left\{ -\frac{g^2}{2V} \right\}$
$\psi < 0$	Emergence of power law tails
$\psi = -1$	$\sim g^{-3}$, $P(g) = \frac{\lambda}{2\sqrt{2V}} \left(\frac{g^2}{2V} + \lambda \right)^{-\frac{3}{2}}$
$\psi = -2$	$\sim g^{-2}$, $P(g) = \frac{\lambda^2 \sqrt{V}}{2\sqrt{\pi} g^2} \left(1 - \left(1 - 2\Phi \left(\frac{\lambda \sqrt{V}}{\sqrt{2 g }} \right) \right) \sqrt{\pi V} \exp \left\{ \frac{\lambda^2 V}{2g^2} \right\} \right)$

In both cases, the growth distribution is obtained as a scale mixture of Gaussians. Gaussian scale mixtures generate a large class of distributions that includes exponential power and symmetric stable distributions; the precise distribution depends upon the mixing distribution $P(K)$. In particular, if we assume that $P(K)$ is exponentially distributed, then, by mixing it with a standard normal distribution, we obtain a Laplace distribution (Kotz, Kozubowski, and Podgorski 2001). Hence, a Laplace growth model can be interpreted as a generalization of a Gibrat growth process amplified by an event that may occur with probability K and may change the probabilistic structure of the stochastic process.

To extend these considerations, we shall now refer to the general case of exponential mixtures of Gaussians given by

$$P(g) = \int_0^\infty \lambda e^{-\lambda K} \frac{1}{\sqrt{2\pi VK}^\psi} \exp \left\{ -\frac{g^2}{2VK} \right\} dK, \tag{7}$$

where ψ captures the scaling of variance and reflects the fact that for different time frames, the strength of dependence of the firm size distribution on the number of its products varies. Indeed, for short time frames, one should expect strong dependence and thus a high ψ (due to a small number of idiosyncratic product-specific shocks), and ψ should fall accordingly with the length of the time frame. This intuition is confirmed by numerical tests and econometric results.

Table 1 summarizes some of the probability density functions that are obtained by varying the scaling parameter ψ . For $\psi \geq 0$, the growth-rate distribution is an exponential power distribution: The case $\psi < 1$ corresponds to $\alpha \in (1, 2)$, whereas the case $\psi > 1$ corresponds to $\alpha < 1$ (α is thus inversely related to ψ). Intuitively, one could expect ψ to be falling with the length of the time frame. In light of the above findings, α should then increase with the length of the time frame. This is exactly what we find empirically (see Table 2). As ψ

goes from 0 to 1, the exponential mixture of Gaussians changes its shape from purely Gaussian to Laplace. When ψ exceeds 1, the resultant probability density function (PDF) “goes past” the Laplace, exhibiting a sharp spike at zero and relatively fat tails. If $\psi < 0$, the distribution exhibits power law wings $P(g) \sim g^{-\gamma}$, with an exponent γ that increases with $|\psi|$, approaching $\gamma = +\infty$ as $\psi \rightarrow 0$. In the case $\psi = -1$, we obtain the same distribution as in equation (5).

3. The Empirical Evidence

In this section, we analyze a unique longitudinal database that records sales figures of the 340,560 products commercialized by 8,072 firms in 28 countries from 1994 to 2004, covering the whole size distribution for products and firms, and monitoring the flows of entry and exit at both levels. Data cover the worldwide pharmaceutical industry. The pharmaceutical industry offers a unique context for empirical investigation relevant to our model, because it consists of many independent submarkets corresponding to different therapeutic groups within the industry (see Sutton 1998; Bottazzi et al. 2001).

Figure 1 shows the annual growth-rate distributions for firms and products on a logarithmic scale. Both growth distributions depict marked departures from a Gaussian shape and are even more leptokurtic than Laplace.

To better characterize the properties of the growth distribution, we refer to a general family of distributions, namely, the exponential power family \mathcal{EP} , which includes the Laplace and the Gaussian distributions as particular instantiations. Following Box and Tiao (1973), the \mathcal{EP} has been applied in robustness studies and in the context of GARCH volatility modeling. Ayebo and Kozubowski (2003) generalized \mathcal{EP} to skew exponential power laws in order to estimate moments and related parameters of asymmetric empirical distributions. An asymmetric exponential power distribution has the form

$$P_g(g) = \frac{\alpha}{\sigma \Gamma(\frac{1}{\alpha})} \frac{\kappa}{1 + \kappa^2} \exp \left\{ -\frac{\kappa^\alpha}{\sigma^\alpha} (g^+)^{\alpha} - \frac{1}{\sigma^\alpha \kappa^\alpha} (g^-)^{\alpha} \right\}, \quad (8)$$

where

$$g^+ = \begin{cases} g & \text{if } g \geq 0, \\ 0 & \text{if } g < 0, \end{cases} \quad \text{and} \quad g^- = \begin{cases} -g & \text{if } g \leq 0, \\ 0 & \text{if } g > 0. \end{cases} \quad (9)$$

Parameter $\sigma > 0$ corresponds to scale while $\kappa > 0$ controls skewness and $\alpha > 0$ is the shape parameter.⁴ By varying the exponent α it is possible to describe

4. We applied a nonparametric methodology in order to identify the robust estimator of the location of the distributions and shifted the data prior to computing estimators presented in this paper.

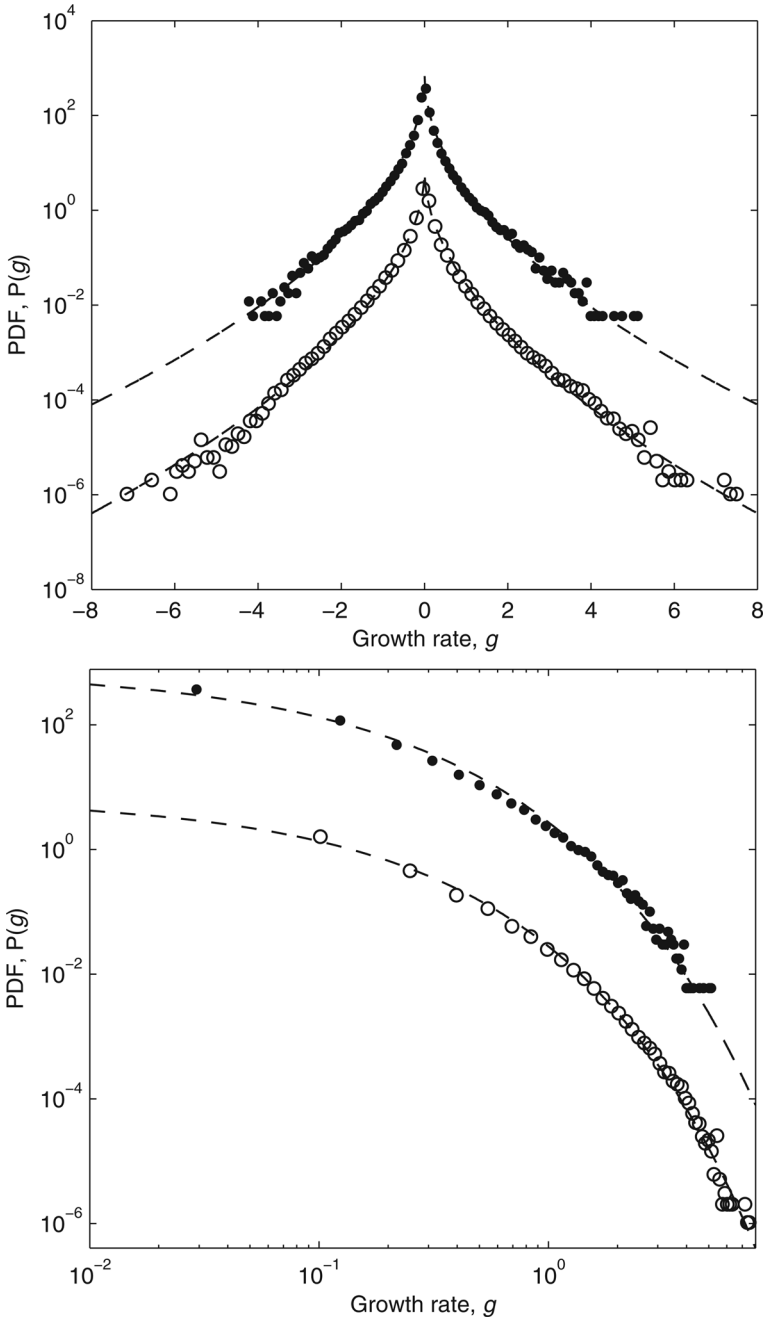


FIGURE 1. Yearly growth distributions of firms (dots) and products (circles), $P(g)$, logarithmic scale (top), and double logarithmic scale (bottom). Empirical fit of equation (8) with parameters $\sigma = 0.3$, $\alpha = 0.5$, and $\kappa = 0$ (dashed lines). For clarity, the growth distribution of firms is offset by a factor of 10^2 .

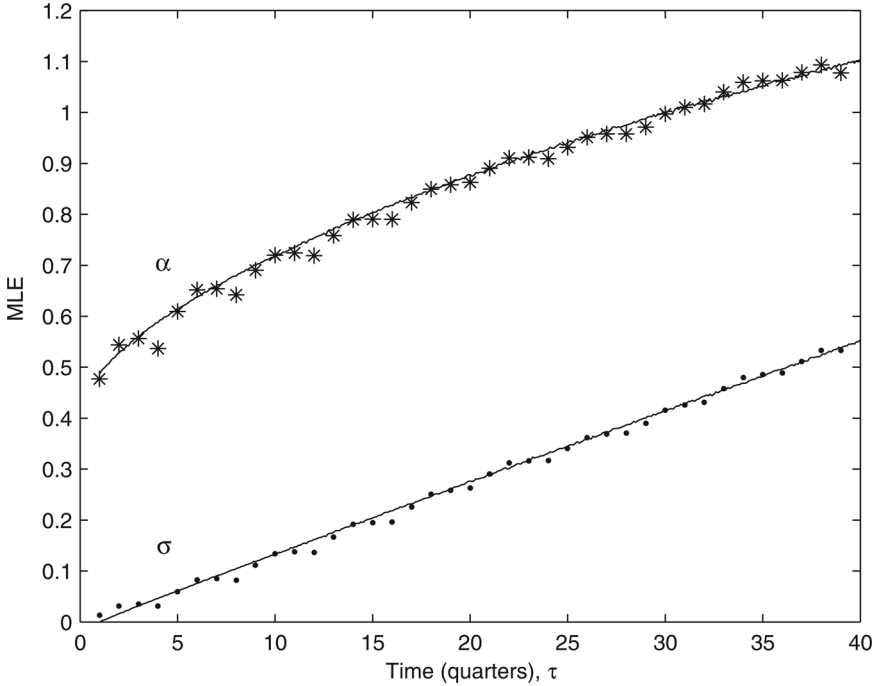


FIGURE 2. The MLEs of σ (stars) and α (dots) of the distribution of firm growth rates (equation [8]) at different time lags as compared to simulation results of an exponential scale mixture of Gaussian distributions (equation [7]) with $\psi = 2$.

Gaussian, platikurtic, and leptokurtic distributions. For $\kappa = 1$ and $\alpha = 2$, the distribution is a symmetric Gaussian. For $\kappa = 1$ and $\alpha = 1$, we obtain a Laplace distribution with standard deviation σ . If $\kappa \neq 1$, letting $\alpha = 1$ leads to a skew Laplace distribution.

We compute the maximum likelihood estimators (MLEs) of the parameters in equation (8) and control our results by carrying out a Monte Carlo experiment. In Figure 2, MLEs of parameters σ and α for real data are compared with simu-

TABLE 2. Kolmogorov–Smirnov distances for Gaussian, asymmetric Laplace (AL), exponential power (EP), and skew exponential power (SEP) distributions.

	$g_{f,1}$	$g_{f,2}$	$g_{f,3}$	$g_{f,4}$	$g_{f,5}$	$g_{f,6}$
Gaussian	.1693	.1400	.1252	.1100	.0986	.0895
AL	.0651	.0485	.0407	.0304	.0247	.0177
EP	.0272	.0269	.0296	.0299	.0306	.0305
SEP	.0126	.0146	.0163	.0200	.0157	.0163

TABLE 3. MLEs of α , κ , and σ of equation (8) at different time lags. Firm growth distributions are computed both with and without new product/firm entry.

Entry	$g_{f,1}$	$g_{f,2}$	$g_{f,3}$	$g_{f,4}$	$g_{f,5}$	$g_{f,6}$
α	.5927	.6727	.7423	.7986	.8688	.9475
κ	.9705	.9695	.9264	.9835	.9880	.9736
σ	.1077	.2198	.3424	.4662	.6173	.8298
Without Entry	$g_{f,1}$	$g_{f,2}$	$g_{f,3}$	$g_{f,4}$	$g_{f,5}$	$g_{f,6}$
α	.7092	.8098	.8812	.9487	1.0201	1.0564
κ	1.0847	1.0825	1.1184	1.1187	1.0732	1.1284
σ	.1026	.1979	.2931	.3915	.5020	.5893

relative results. The MLE of the shape parameter α increases from 0.5 to 1.1–1.2, converging to a Gaussian distribution, while the MLE of the standard deviation σ grows linearly from 0.02 to 0.45. In our simulations, we first generate a vector of exponentially distributed stochastic variables $P(K)$ whose parameter λ scales from 0 to 1. Then we generate a vector of stochastic elements drawn from a Gaussian distribution with zero mean and standard error $\sigma = K$. Finally, we estimate the MLEs of parameters for the simulated growth distribution and notice that simulation results perfectly reproduce the growth distribution of firms and its evolution over different time lags (see Figure 2).⁵

Table 2 reports the values of the Kolmogorov–Smirnov distances between the data and four distribution fits to firm growth rates over a period of one to six years. Clearly, the \mathcal{EP} fit outperforms the Gaussian and the Laplace ones for yearly growth distributions, whereas there is only a slight advantage of the skew exponential power over the Laplace fit for long-term growth rates. In order to ascertain the effect of product turnover on firm growth, in Table 3 we compare MLE of parameters α , κ and σ of firm growth distributions with and without new product entry. Notably, the removal of new product entry leads to an increase in the shape parameter of the growth distribution of about 0.15. Thus, product entry amplifies the departure from Gaussian/Laplace in our data even if it is not sufficient to explain the skewness of the firm growth distribution. Indeed, the increase in the shape parameter over time seems to be related to asynchronous product/firm arrivals.

5. For a given value of parameter $\beta_\tau \equiv (\alpha_\tau, \kappa_\tau, \sigma_\tau)$, we generate 10,000 replications of the \mathcal{EP} random variable G with density f_G . The simulations are performed using Matlab gamma random number generator with default seed values. The sample size is set equal to the number of observations in each time frame τ . For each replication, the true parameter set β_τ is estimated by $\hat{\beta}_{\tau,n} \equiv (\hat{\alpha}_{\tau,n}, \hat{\kappa}_{\tau,n}, \hat{\sigma}_{\tau,n})$.

4. Conclusions

Following Stanley et al. (1996), researchers in the “econophysics” community have found that the growth distribution of firms is not Gaussian as implied by Gibrat’s Law of Proportionate Effect but displays heavy tails. In this paper, we provide a simple candidate explanation for the shape of the firm growth distribution based on a model of proportional growth of both the introduction of new products by firms, which is supposed to be proportional to knowledge capital accumulated through past product innovations as in Klette and Kortum (2004), and product size dynamics under the assumption of independent submarkets (Sutton 1998, 2007). The pharmaceutical industry is analyzed as an ideal context for testing such a model, because (approximately) independent business units can be identified with different therapeutic groups within the industry. Our econometric investigations show that model’s predictions are in good agreement with empirical evidence.

References

- Ayebo, Abraham, and Tomasz J. Kozubowski (2003). “An Asymmetric Generalization of Gaussian and Laplace Laws.” *Journal of Probability and Statistical Science*, 1, 187–210.
- Bottazzi, Giulio, Giovanni Dosi, Marcello Lippi, Fabio Pammolli, and Massimo Riccaboni (2001). “Innovation and Corporate Growth in the Evolution of the Drug Industry.” *International Journal of Industrial Organization*, 19, 1161–1187.
- Box, George E. P., and George C. Tiao (1973). *Bayesian Inference in Statistical Analysis*. Addison-Wesley.
- Canning, David, Luis A. Nunes Amaral, Youngki Lee, Martin Meyer, and H. Eugene Stanley (1998). “A Power Law for Scaling the Volatility of GDP Growth Rates with Country Size.” *Economics Letters*, 60, 335–341.
- Chatterjee, Arnab, Sudhakar Yarlagadda, and Bikas K. Chakrabarti, eds. (2005). *Econophysics of Wealth Distributions*. Springer.
- Fu, Dongfeng, Fabio Pammolli, Sergey V. Buldyrev, Massimo Riccaboni, Kaushik Matia, Kazuko Yamasaki, H. Eugene Stanley (2005). “The Growth of Business Firms: Theoretical Framework and Empirical Evidence.” *Proceedings of the National Academy of Sciences*, 102, 18801–18806.
- Gabaix, Xavier, Parameswaran Gopikrishnan, Vasiliki Plerou, and H. Eugene Stanley (2006). “Institutional Investors and Stock Market Volatility.” *Quarterly Journal of Economics*, 121, 461–504.
- Ijiri, Yuji, and Herbert A. Simon (1977). *Skew Distributions and the Sizes of Business Firms*. North-Holland.
- Klette, Tor J., and Zvi Griliches (2000). “Empirical Patterns of Firm Growth and R&D Investment: A Quality Ladder Model Interpretation.” *Economic Journal*, 110, 363–387.
- Klette, Tor J., and Samuel Kortum (2004). “Innovating Firms and Aggregate Innovation.” *Journal of Political Economy*, 112, 986–1018.
- Kotz, Samuel, Tomasz J. Kozubowski and Krzysztof Podgorski (2001). *The Laplace Distribution and Generalizations: A Revisit with Applications to Communications, Economics, Engineering, and Finance*. Birkhauser.
- Rossi-Hansberg, Esteban, and Mark L. J. Wright (2005). “Establishment Size Dynamics in the Aggregate Economy.” NBER Working Paper No. 11261.

- Stanley, Michael H. R., Luis A. Nunes Amaral, Sergey V Buldyrev, Shlomo Havlin, Heiko Leschhorn, Philipp Maass, Michael A Salinger, and H. Eugene Stanley (1996). "Scaling Behaviour in the Growth of Companies." *Nature*, 379, 804–806.
- Sutton, John (1997). "Gibrat's Legacy." *Journal of Economic Literature*, 35, 40–59.
- Sutton, John (1998). *Technology and Market Structure*. MIT Press.
- Sutton, John (2007). "Market Structure: Theory and Evidence." In *Handbook of Industrial Organisation*, vol. 3, edited by Robert Porter and Mark Armstrong. North-Holland.
- Yamasaki, Kazuko, Kaushik Matia, Sergey V. Buldyrev, Dongfeng Fu, Fabio Pammolli, Massimo Riccaboni, and H. Eugene Stanley (2006). "Preferential Attachment and Growth Dynamics in Complex Systems." *Physical Review E*, 74, 35103–35106.