JSPS Grants-in-Aid for Creative Scientific Research
Understanding Inflation Dynamics of the Japanese Economy
Working Paper Series No.38

The Firms as a Bundle of Barcodes

Koji Sakai
And
Tsutomu Watanabe

May 12, 2009

Research Center for Price Dynamics
Institute of Economic Research, Hitotsubashi University
Naka 2-1, Kunitachi-city, Tokyo 186-8603, JAPAN
Tel/Fax: +81-42-580-9138
E-mail: sousei-sec@ier.hit-u.ac.jp
http://www.ier.hit-u.ac.jp/~ifd/

Noname manuscript No.
(will be inserted by the editor)

The Firm as a Bundle of Barcodes *

Koji Sakai[1], Tsutomu Watanabe[2]**

[1] Institute of Economic Research, Hitotsubashi University, ksakai@ier.hit-u.ac.jp
[2] Research Center for Price Dynamics, Hitotsubashi University, tsutomu.w@srv.cc.hit-u.ac.jp

May 12, 2009

Abstract We empirically investigate the firm growth model proposed by Buldyrev et al. by using a unique dataset that contains the daily sales of more than 200 thousand products, which are collected from about 200 supermarkets in Japan over the last 20 years. We find that the empirical firm growth distribution is characterized by a Laplace distribution at the center and power-law at the tails, as predicted by the model. However, some of these characteristics disappear once we randomly reshuffle products across firms, implying that the shape of the empirical distribution is not produced as described by the model. Our simulation results suggest that the shape of the empirical distribution stems mainly from the presence of relationship between the size of a product and its growth rate.

1 Introduction

Why do firms exist? What determines a firm's boundaries? These questions have been repeatedly addressed by social scientists since Adam Smith argued more than two centuries ago that division of labor or specialization is a key to the improvement of labor productivity [1].

Specifically, Ronald Coase emphasized the importance of transactions costs, such as the cost of price discovery, arguing that the firm is a special organization that is able to save these costs by internalizing transactions that would be otherwise carried out through the market [2]. On the other hand, Edith Penrose argued that firm growth is constrained by a firm's internal resources, such as technologies, skills, and knowledge, thus defining the firm as a collection of these resources [3].

These two ideas, as well as other numerous ideas proposed by social scientists, share the following features. First, the firm is regarded as an organization with a will, making a decision by solving an optimization problem. Second, heterogeneity of firms in various respects is regarded as an important thing to be explained by the theory of the firm. For example, Penrose and her successors pay special attention to heterogeneity of firms in terms of their resources, which are acquired by individual firms through the process of learning and innovation.

Recently, however, a different view has been proposed by a group of physicists [4–9]. They regard the firm as a mere bundle of its constituent units, such as divisions, products, and individual transactions. Put differently, the firm is no longer an organization in which its constituent units are closely linked to each other. Also, they emphasize the role of stochastic elements in explaining firm dynamics,[1] thus the boundaries of a firm is determined not by a firm's optimizing behavior but by simple stochastic processes with i.i.d. property. More importantly, they pay almost no attention to individual firms: instead, their focus is on how individual firms are distributed in terms of their performance measures, such as sales growth.

Specifically, they start by assuming that firm growth is governed by two simple stochastic processes regarding the change in the number of units within a firm, as well as the change in the size of a unit, and then show that the distribution of firm growth rates is characterized by a Laplacian cusp in the central part and asymptotic power law tails [8, 9].

* We thank H. Takayasu, M. Takayasu, and H. Yoshikawa for helpful discussions and suggestions. We also thank participants at the APFA7 Conference and members of the Industrial & Financial Network Study Group at RIETI. This research is a part of the project entitled: Understanding Inflation Dynamics of the Japanese Economy, funded by JSPS Grant-in-Aid for Creative Scientific Research (18GS0101).

** *Present address:* Institute of Economic Research, Hitotsubashi University, Kunitachi, Tokyo 186-8603, Japan.

[1] The view that firm dynamics is driven by stochastic processes was first proposed by Robert Gibrat, a French economist [11], although most economists in subsequent generations have not been willing to accept it.

One of the most surprising implications obtained from their model is that the distribution of firm growth rates should be unchanged even if one randomly reshuffles products across firms. The purpose of this paper is to conduct such reshuffling in order to see whether the firm is a mere bundle of randomly chosen products or not. To do this, we employ a unique dataset that records the daily sales for each of more than 200 thousand products, which are identified by their barcodes. These are products sold at about 200 supermarkets in Japan, so they consist mainly of food, beverages, and other domestic non-durables.

2 Theoretical predictions

In ref.[8], a firm is treated as an entity consisting of a random number of units such as products. They assume that firm growth is essentially governed by two stochastic processes: (1) the number of units in a firm grows in proportion to the existing number of units; (2) the size of each unit grows in proportion to its size. They also assume no interdependence across units, as well as no autocorrelation in terms of the growth rate of a unit.

These assumptions imply various types of "independence". That is, the size of a unit of a firm in period t is independent of the number of units of the firm in period t. On the other hand, the growth rate of a unit of a firm from period t to period $t+1$ is independent of the number of units of the firm in t, the size of the unit in t, the growth rate of other units from t to $t+1$, and the growth rate of the unit from $t-1$ to t.

Their main result is that the probability distribution of the growth rate of a firm, g, is given by

$$P(g) \approx \frac{2V}{\sqrt{g^2+2V}(|g|+\sqrt{g^2+2V})^2} \qquad (1)$$

where V represents the variance of g. Eq.(1) indicates that $P(g)$ approaches $1/\sqrt{2V} - |g|/V$ as $g \to 0$, while it goes to $V/(2g^3)$ as $g \to \infty$, implying that $P(g)$ is approximated by a Laplace distribution in the body and by a power-law with an exponent of 3 in the tail.

An important testable implication of their model is that the distribution of firm growth should be unchanged even if one randomly reshuffles units across firms, which is a direct reflection of their independence assumptions. Put differently, the firm is a mere bundle of units like products, which are chosen randomly.

3 Data

The dataset we use is a store scanner data compiled jointly by Nikkei Digital Media, Co., Ltd. and Research Center for Price Dynamics. This dataset contains the daily sales for more than 200 thousand products sold at about 200 supermarkets in Japan from 1988 to 2008. The products consist mainly of food, beverages, and other domestic non-durables (like detergent, facial tissues, shampoo, soaps, toothbrushes, etc.) and their sales are recorded through the so-called point-of-sale system. Each product is identified by the JAN (Japanese Article Number) code, an equivalence of the UPS code in the United States. Each product is assigned a 6-digit product category code, and tied to a producer code indicating by whom it is produced.

We first choose samples of stores that exist during the entire period of 1998-2008. Then, we construct two kinds of aggregated sales: sales at the firm level, which is defined by the annual sales of each firm; sales at the product level, which is defined by the annual sales of products belonging to a 6-digit product category within a firm. We eventually have 4,000 observations of sales at the firm level and 14,000 observations of sales at the product level, for each year.

Eq.(1) is tested by ref.[9] using a dataset from the pharmaceutical industry, which records the sales of each product within a firm. Our dataset differs from it in some important respects. As emphasized by ref.[7], the pharmaceutical industry consists of independent submarkets corresponding to different therapeutic groups within the industry; therefore, one may safely assume that the growth processes of the constituent parts of a firm are independent from each other, which is a part of the assumptions adopted in obtaining eq.(1). There is no clear reason to beleive, at least a priori, that the assumption of independent submarkets is also satisfied in our dataset.

On the other hand, the pharmaceutical industry is *not* a typical industry in that there are various regulations including those related to the entry to the industry, the introduction of new products to the market, and pricing. This implies that firm dynamics in such an industry could be substantially influenced by the decisions made by public authorities. Our dataset is collected from more or less "standard" industries that are less regulated and more competitive compared with the pharmaceutical industry. In this sense, the industries covered by our dataset is closer to the ones frequently investigated in the previous studies about firm dynamics, such as Gibrat law, by using datasets at a high level of aggregation.

4 Empirical results

4.1 The shape of firm growth distributions

We start by looking at the shape of growth distributions. We first calculate the annual growth rates both at the firm and at the product levels, and then fit the theoretical density given by eq.(1) to them.

The results for the growth rate distribution at the firm level are presented in Fig.1, showing that (1) the entire distribution is fairly close to the theoretical one;

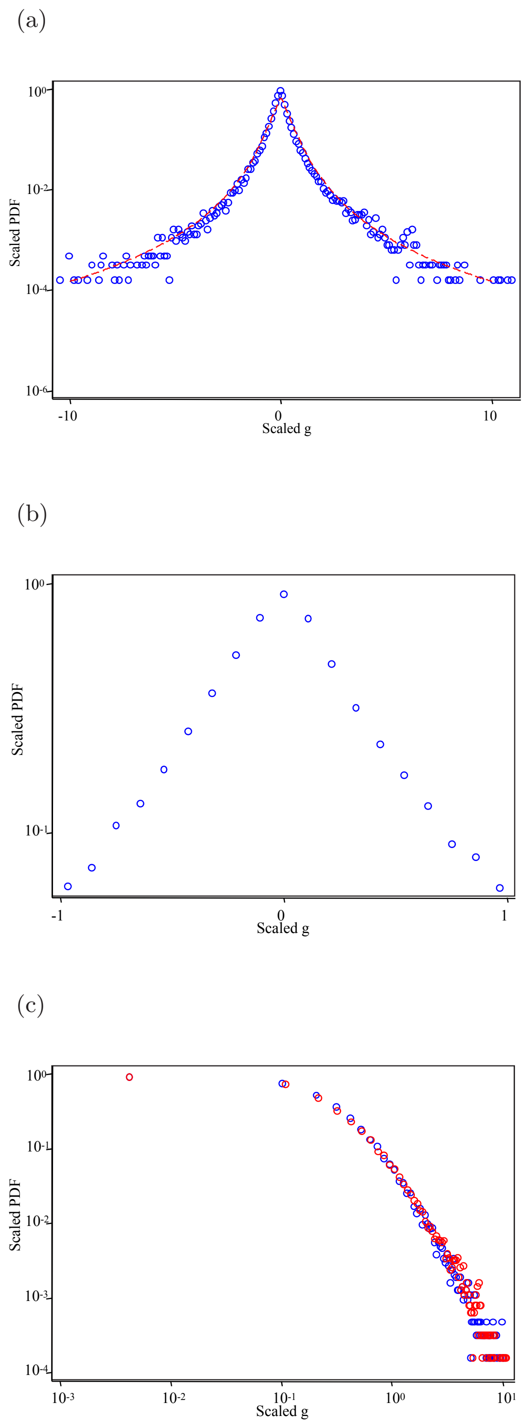

Fig. 1 The empirical probability density function $P(g)$ of firm growth rates. The PDF is rescaled by $\sqrt{V}$. The panel (a) shows the entire distribution with the markers representing the empirical PDF, and the dashed line representing a fitted line by eq.(1). The estimated value of V is 0.452. The panel (b) shows the central part of the PDF, indicating that it is well fit by a Laplace distribution. The panel (c) shows the left and right tails of the PDF (the left tail: ○, the right tail: ○), indicating that both tails are approximated by power law with an exponent slightly smaller than 3.

(2) the central part of the distribution is well approximated by a Laplace distibution; (3) the tails are symmetric and approximated by power law, although its exponent is slightly smaller than the predicted value, 3. We have obtained a similar result for the growth rate distribution at the product level, although it is not reported here due to space limitations.

4.2 Random reshuffling of products across firms

Our next task is to conduct random reshuffling of products across firms in order to see whether the shape of the firm growth distribution is unchanged or not. We run four simulations, in each of which we break relationships between the number of products within a firm, denoted by K, the size of each product, ξ, and the growth rate of each product, η, in different ways, and then recalculate firm growth distributions in each case. Similar simulations are conducted by ref.[10] in a different context, although they investigate only a part of the possible relationships between K, ξ, and η.

In the first simulation, we randomly reshuffle ξ and η *separately* by keeping the real-world distribution of number of units within each firm, $P(K)$, unchanged. As an example, consider a firm with K_0 products. We multiply a new value for the sales of a product of the firm in year t by a new value for the growth rate from t to $t+1$ to compute the sales of the product in year $t+1$. We then sum up the sales at the product level over K_0 products in t and $t+1$ to compute the corresponding growth rate at the firm level. This is a benchmark case close to the theoretical model, in the sense that the possible relationships between K, ξ, and η are all removed. In the second simulation, we randomly reshuffle η keeping the real-world $P(K)$ and ξ, thereby removing the following links: the link between K and η; the link between ξ and η; and the link between η's within a firm. Note that the link between K and ξ and the link between ξ's are not removed. In the third simulation, we randomly reshuffle ξ keeping the real-world $P(K)$ and η. The links between K and ξ, between ξ and η, and between ξ's are removed, but the link between K and η and the link between η's are not removed. Finally, in the fourth simulation, we keep the real $P(K)$, and randomly reshuffle ξ and η *jointly* (i.e. keeping the real-world links between ξ and η). The links between K and ξ, between K and η, between ξ's, and between η's are removed, but the link between ξ and η is not removed.

Fig.2 presents firm growth distributions for the four cases. The firm growth distribution obtained from the first simulation exhibits important deviations from the theoretical prediction: the growth distribution is asymmetric with positive growth being more likely than negative one, so that it can no longer be fitted well by eq.(1); the positive tail is no longer close to power law, although the negative tail still exhibits power law behavior. More

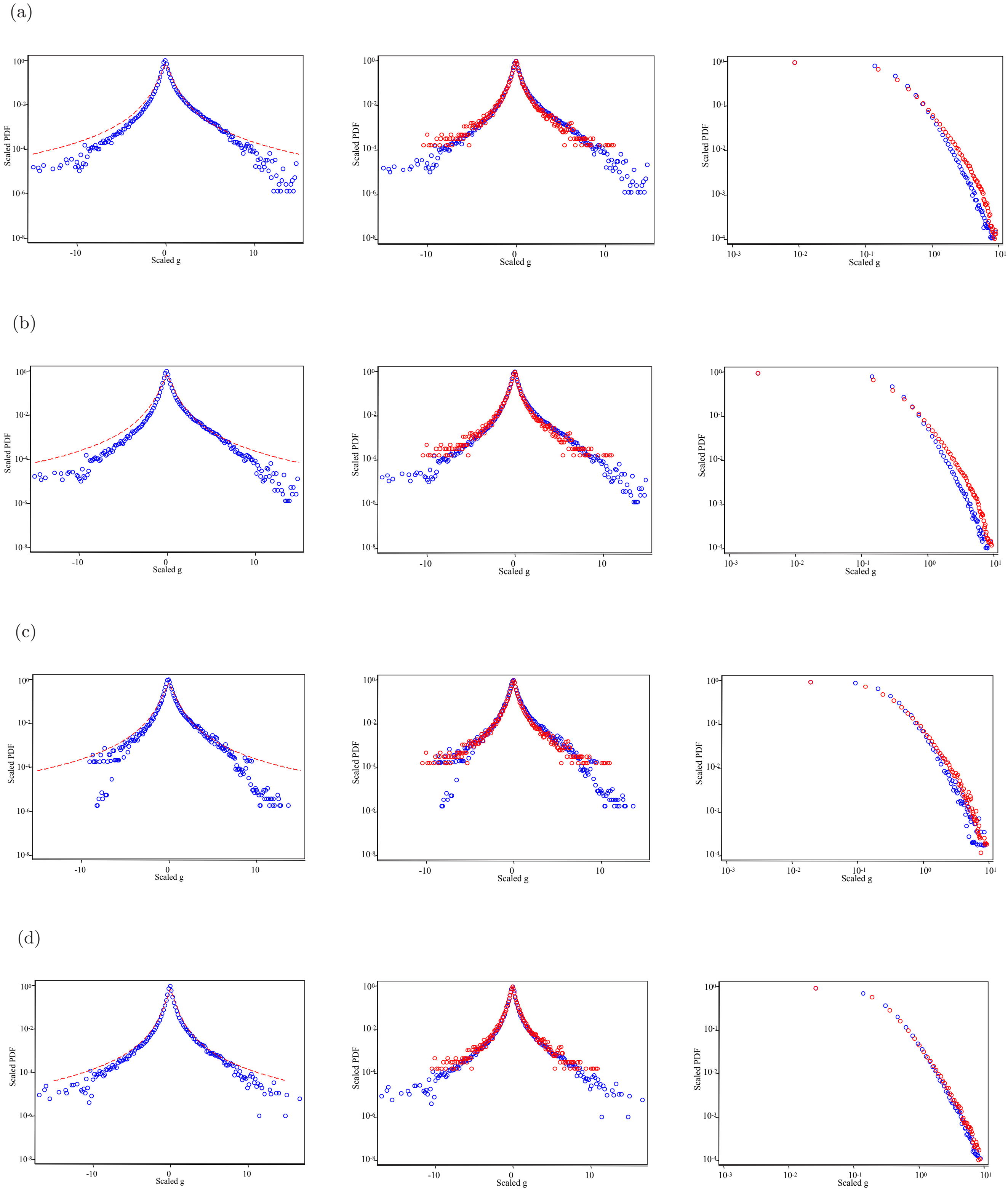

Fig. 2 The probability density functions of firm growth rates obtained from simulations. Figures on the left show the entire PDFs, with the blue markers representing the PDFs and the red dashed lines representing fitted lines by eq.(1). Figures on the middle show the simulated and empirical PDFs, with the blue markers representing the simulated PDFs and the red markers representing the empirical PDFs. Figures on the right focus on the tails of the PDFs ((the left tail: ○, the right tail: ○)). The panel (a) shows the result from the first simulation in which we reshuffle ξ and η separately keeping the real-world distribution of the number of units within a firm, $P(K)$. The panel (b) shows the result from the second simulation in which we reshuffle η keeping the real-world $P(K)$ and ξ. The panel (c) shows the result from the third simulation in which we reshuffle ξ keeping the real-world $P(K)$ and η. The panel (d) shows the result from the fourth simulation in which we reshuffle ξ and η jointly (i.e. keeping the real-world links between ξ and η) keeping the real-world $P(K)$.

specifically, as shown in the figures at the middle and the right of the first row of Fig.2, positive but small growth is more likely to occur compared with the empirical distribution, while positive and large growth is less likely. This is a surprising result, given that we have removed all possible links between K, ξ, and η in the first simulation, so that the independence assumptions in the model should be exactly satisfied, and thus we should expect a better fit to eq.(1) compared with the actual data. This result implies that the previous result presented in Fig.1 is a spurious one in the sense that the firm growth distribution observed in the actual data is *not* produced as described in the model, although its shape is very close to the theoretical prediction.

One may wonder why the distribution obtained in the first simulation deviates from the theoretical prediction. Presumably, the actual distribution of the number of units of a firm, $P(K)$, deviates from the one in the model. However, a more important question to be addressed is why the shape of the actual distribution in Fig.1 is so close to the theoretical prediction, especially in terms of the symmetry of the distribution and its power law tails. To answer to this question, we need to look for driving forces outside the model, which contribute to producing the actual distribution.

In doing this, the results from the other simulations provide useful information. The distributions obtained from the second and third simulations are again asymmetric and lack power law tails. In fact, these distributions are almost identical to the one obtained from the first simulation. Note that the second simulation differs from the first one in that the link between K and ξ as well as the link between ξ's are not removed in the second one, while both are removed in the first one. On the other hand, the third simulation differs from the first one in that the link between K and η as well as the link between η's are not removed in the third one, while both are removed in the first one. Thus, these results indicate that none of the links between K and ξ, between K and η, between ξ's, and between η's plays a critically important role in producing the shape of the actual growth distribution in Fig.1.

Put differently, the scope of a firm, measured in terms of the number of products it produces, is not significantly correlated with the growth performance of the individual products it produces, and, in this sense, the economies of scope are not so important in generating the actual firm growth distribution. Also, the interdependence of individual product growth within a firm is not strong enough to have any significant consequences on the firm growth distribution, suggesting the absence of a firm's boundaries.

Turning to the distribution obtained from the fourth simulation, we see that it is very close to the one in Fig.1: it is symmetric and fitted well by eq.(1); it exhibits power law behavior both at the positive and negative tails. Note that the fourth simulation differs from the first one only in that the link between ξ and η is not removed in the fourth one while it is removed in the first one. Therefore, this result implies that the relationship between ξ and η is a key force contributing to producing the actual distribution with symmetry and power law tails.

Where does the relationship between ξ and η come from? One possibility is the presence of size dependence at the product level. Young products might be more volatile in terms of their growth rates than old ones because of, say, learning effects, just as young firms are more volatile than old ones. Alternatively, this could come from the presence of autocorrelation in the growth rate of a product. That is, the current size of each product, ξ, is equal to the multiplication of the current and past growth rates, η's. Thus the relationship between ξ and η may stem from autocorrelation in η's. This is in line with recent theoretical developments about firm dynamics by economists, such as life-cycle, evolutionary, and innovation models, in which past idiosyncratic shocks to a firm have persistent effects upon firm growth through learning about the technology or persistent shocks to the technology [12–14].

5 Conclusion

In this paper, we have tested the theoretical predictions made by ref.[8] using a unique dataset containing the daily sales of more than 200 thousand products, which are collected from about 200 supermarkets in Japan over the last 20 years. We have found that the empirical sales growth distributions, both at the firm and product levels, are characterized by a Laplace distribution at the center and power-law at the tails, as predicted by the model. However, some of these features disappear once we randomly reshuffle products across firms; namely, the firm growth distribution becomes asymmetric and no longer exhibits power-law tails. Our simulation results suggest that the shape of the empirical sales growth distribution stems mainly from the presence of relatioship between the size of a product in period t and its growth rate from t to $t+1$.

References

1. A. Smith, *An Inquiry into the Nature and Causes of the Wealth of Nations* (Oxford Univ. Press, Oxford 1976, originally published 1776)
2. R.H. Coase, Economica **4**, (1937) 386.
3. E.T. Penrose, *The Theory of the Growth of the Firm* (Oxford Univ. Press, Oxford 1959)
4. S.V. Buldyrev, L.A.N. Amaral, S. Havlin, H. Leschhorn, P. Maass, M.A. Salinger, H.E. Stanley, M.H.R. Stanley, J. Phys. I France **7**, (1997) 635.
5. L.A.N. Amaral, S.V. Buldyrev, S. Havlin, H. Leschhorn, P. Maass, M.A. Salinger, H.E. Stanley, M.H.R. Stanley, J. Phys. I France **7**, (1997) 621.

6. D. Fu, F. Pammolli, S.V. Buldyrev, M. Riccaboni, K. Matia, K. Yamasaki, H.E. Stanley, Proc. Natl. Acad. Sci. USA **102**, (2005) 18801.
7. S.V. Buldyrev, J. Growiec, F. Pammolli, M. Riccaboni, H.E. Stanley, J. Eur. Econ. Assoc. **5**, (2008) 574.
8. S.V. Buldyrev, F. Pammolli, M. Riccaboni, K. Yamasaki, D. Fu, K. Matia, H.E. Stanley, Eur. Phys. J. B **57**, (2007) 131.
9. F. Pammolli, D. Fu, S.V. Buldyrev, M. Riccaboni, K. Matia, K. Yamasaki, H.E. Stanley, Eur. Phys. J. B **57**, (2007) 127.
10. M. Riccaboni, F. Pammolli, S.V. Buldyrev, L. Ponta, H.E. Stanley, Proc. Natl. Acad. Sci. USA **105**, (2008) 19595.
11. R. Gibrat, *Les Inégalités Économiques* (Librairie du Recueil Sirey, Paris 1931)
12. B. Jovanovic, Econometrica **50**, (1982) 642.
13. H.A. Hopenhayn, Econometrica **60**, (1992) 1127.
14. R. Ericson, A. Pakes, Rev. Econ. Stud. **62**, (1995) 53.