

Ishikawa, Atushi

Article

Power-Law and Log-Normal Distributions in Temporal Changes of Firm-Size Variables

Economics: The Open-Access, Open-Assessment E-Journal

Provided in Cooperation with:

Kiel Institute for the World Economy – Leibniz Center for Research on Global Economic Challenges

Suggested Citation: Ishikawa, Atushi (2009) : Power-Law and Log-Normal Distributions in Temporal Changes of Firm-Size Variables, Economics: The Open-Access, Open-Assessment E-Journal, ISSN 1864-6042, Kiel Institute for the World Economy (IfW), Kiel, Vol. 3, Iss. 2009-11, pp. 1-25, <https://doi.org/10.5018/economics-ejournal.ja.2009-11>

This Version is available at:

<https://hdl.handle.net/10419/27531>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<http://creativecommons.org/licenses/by-nc/2.0/de/deed.en>

Power-Law and Log-Normal Distributions in Temporal Changes of Firm-Size Variables

Atushi Ishikawa

Kanazawa Gakuin University, Kanazawa, Japan

Abstract

In this paper the author shows that signed temporal changes of firm size variables follow the power-law for large changes; while, for middle changes a log-normal distribution is found. In the analyses, the author employed three databases: high-income data, high-sales data and positive-profits data of Japanese firms. It is particularly worth noting that the growth rate distributions in temporal changes of the firm size data have no wide tail, unlike the distributions observed in assets and sales of firms, the number of employees and personal income data. An Extended-Gibrat's Law was also found in the growth rate distributions of temporal changes of firm size variables, which induces both the power-law and the log-normal distributions in the temporal changes of firm size under the Detailed Balance.

Special issue “[Reconstructing Macroeconomics](#)”

JEL: D30, D31, D39

Keywords: Econophysics; temporal change of firm size; Pareto's law; log-normal distribution; (non-)Gibrat's law; detailed balance

Correspondence:

Atushi Ishikawa; Kanazawa University, Kanazawa 920-1392, Japan, e-mail address: ishikawa@kanazawa-gu.ac.jp

The author is grateful to the Yukawa Institute for Theoretical Physics at Kyoto University, where this work was initiated during the YITP-W-07-16 on “Econophysics III – Physical Approach to Social and Economic Phenomena” and especially to Dr. S. Fujimoto for a lot of useful discussions and comments about K–S test. Thanks are also due to Professor H. Kasama for careful reading of the manuscript. This work was supported in part by a Grant-in-Aid for Scientific Research (C) (No. 20510147) from the Ministry of Education, Culture, Sports, Science and Technology, Japan.

1 Introduction

Power-law distributions are frequently observed in economic data such as assets, the number of employees, personal income, sales, profits and income of firms (denoted by x). The Power Law is known as Pareto's Law (Pareto 1897) and the probability density function (pdf) is represented as

$$P_{\text{PL}}(x) = Cx^{-\mu-1} \quad \text{for } x > x_{\text{th}} , \quad (1)$$

where C is a normalization and the power μ is called the Pareto index. In general, the power-law is valid only for large-sized companies (Badger 1980; Montroll and Shlesinger 1983). The size threshold of the region of large-sized firms is denoted by x_{th} . It has been hypothesized that the pdf follows the log-normal distribution for middle-sized firms below the size threshold x_{th} :

$$P_{\text{LN}}(x) = \frac{1}{x\sqrt{2\pi\sigma^2}} \exp \left[-\frac{\ln^2(x/\bar{x})}{2\sigma^2} \right] \quad x < x_{\text{th}} . \quad (2)$$

Here, $\bar{x}$ is a mean value and σ^2 is a variance. Therefore, we concluded that further research of the power-law distribution was warranted because a small number of firms dominate a large percent of the overall sales and profits. Furthermore, since the majority of firms are middle-sized, we also decided to look at the log-normal distributions.

It has been shown that these distributions can be explained by laws observed in a massive amount of digitized economic data. Fujiwara et al. (2003, 2004) point out that Pareto's Law can be derived from the Law of Detailed Balance and from Gibrat's Law (Gibrat 1931). Along these lines, Ishikawa (2006a, 2007a) shows that the log-normal distribution can also be deduced using the Law of Detailed Balance and Non-Gibrat's Law. The Detailed Balance is time-reversal symmetry as observed in the equilibrium system. On the one hand, Gibrat's Law states that the conditional pdf of the growth rate is independent of the initial value. On the other hand, Non-Gibrat's Law describes the dependence of the initial value. Gibrat's Law is observed only in companies in the large scale region, and Non-Gibrat's Law in companies in the middle scale region.

It is interesting to note that there are two types of growth rate distributions. The shape of the growth rate distribution of profits or income of firms (Fig. 1¹) is different from that of assets and sales of firms, the number of employees or personal income (Fig. 2). This difference is observed not only for companies in the large scale region but also for companies in the middle scale region. The point is that the difference might be related to the difference between two Non-Gibrat's Laws in companies in the middle scale region. On the one hand in Fig. 1, the probability of positive growth decreases and the probability of negative growth increases as the bin size of the classification of x increases in the middle scale region (Ishikawa 2006a, 2007a). On the other hand in Fig. 2, the probability of positive and negative growth decreases as the bin size of the classification of x increases (Aoyama 2004, Aoyama et al. 2004). This size dependence for companies in the middle scale region

¹All figures appear in the Appendix.

is significant due to the preponderance of companies and individuals in this size range.

In this study, we propose that the shape of the growth rate distribution is determined by the type of economic variable. In concrete terms, the variables are calculated by any subtraction or not. In this study we will prove our hypothesis by employing data from sales, profits and income.

2 Firm Size Distributions

In this section, we will review the derivation of Pareto's Law and the log-normal distribution from the Detailed Balance and (Non-)Gibrat's Law, and confirm the laws by employing data for sales, profits and income of Japanese firms.

In Japan, firms having an annual income of more than 40 million yen have been officially designated as "high-income firms"; there are approximately 70 thousand of these firms. The exhaustive database was published by Diamond Inc. Top 500 thousand sales data of Japanese firms are available on the database "CD Eyes 50" published by TOKYO SHOKO RESEARCH, LTD. This database is thought to be approximately exhaustive. The database includes both positive and negative profits data. The number of firms with positive profits is approximately 300 thousand and that with negative profits is approximately 40 thousand. In this study we will exclude firms with negative profits, the number of which is significantly less than that with positive profits. This is due to the fact that the negative profits data gathered from high-sales data are exclusive. Despite the fact that we do not have a complete picture of the positive profits data for companies in the middle scale region; nonetheless, we decided to employ this data to investigate the consistency of the laws with the data. In this study, we will examine the following three databases: high-income data (Database I), high-sales data (Database II) and positive-profits data (Database III).

2.1 Pareto's Law from the Detailed Balance and Gibrat's Law

Let firm sizes at the two successive points in time be denoted by x_1 and x_2 . The growth rate R is defined as the ratio $R = x_2/x_1$. The Detailed Balance and Gibrat's Law (Gibrat 1932) are represented as follows:

- **Detailed Balance**

The joint pdf $P_{12}(x_1, x_2)$ is symmetric under the exchange $x_1 \leftrightarrow x_2$:

$$P_{12}(x_1, x_2) = P_{12}(x_2, x_1) . \quad (3)$$

- **Gibrat's Law**

The conditional pdf of the growth rate $Q(R|x_1)$ is independent of the initial value x_1 :

$$Q(R|x_1) = Q(R) , \quad (4)$$

where the conditional pdf $Q(R|x_1)$ is defined as

$$Q(R|x_1) = \frac{P_{1R}(x_1, R)}{P(x_1)} \quad (5)$$

by using the pdf $P(x_1)$ and the joint pdf $P_{1R}(x_1, R)$.

These laws are confirmed in Databases I – III. In order to compare analyses in the next section, we will investigate data for 2003 (x_0), 2004 (x_1) and 2005 (x_2). In the scatter plot of each database, the Detailed Balance (3) is approximately confirmed by applying the one-dimensional Kolmogorov–Smirnov (K–S) test (see Appendix). Figs. 3 – 5 show time-reversal symmetry under the exchange $x_1 \leftrightarrow x_2$.² Gibrat’s Law (4) is also confirmed for each database. Figs. 6 – 8 show that the conditional pdf of the growth rate is approximately independent of the initial value, if the initial value is larger than some threshold x_{th} . Here the pdf for $r = \log_{10} R$ defined by $q(r|x_1)$ is related to that of R by

$$\log_{10} q(r|x_1) = \log_{10} Q(R|x_1) + r + \log_{10}(\ln 10) . \quad (6)$$

Note that large negative growth is not available if there is a lower bound of the data. This is notably observed in Figs. 3 and 6 for high-income Database I. This is also observed in Figs. 4 and 7 for high-sales Database II; however, the lower bound is probably obscure.³ The Detailed Balance and Gibrat’s Law are confirmed by employing personal income data in Japan (Fujiwara et al. 2003), and assets and sales data in France and the number of employees in UK (Fujiwara et al. 2004).

In the literature (Fujiwara et al. 2003, 2004), Pareto’s Law is analytically derived from the Detailed Balance and Gibrat’s Law. By using the relation $P_{12}(x_1, x_2)dx_1dx_2 = P_{1R}(x_1, R)dx_1dR$ under the exchange of variables from (x_1, x_2) to (x_1, R) , these two joint pdfs are related to each other

$$P_{1R}(x_1, R) = x_1 P_{12}(x_1, x_2) . \quad (7)$$

From this relation, the Detailed Balance (3) is rewritten in terms of $P_{1R}(x_1, R)$ as

$$P_{1R}(x_1, R) = R^{-1} P_{1R}(x_2, R^{-1}) . \quad (8)$$

Substituting the joint pdf $P_{1R}(x_1, R)$ for the conditional pdf $Q(R|x_1)$ defined by Eq. (5), the Detailed Balance is expressed as

$$\frac{P(x_1)}{P(x_2)} = \frac{1}{R} \frac{Q(R^{-1}|x_2)}{Q(R|x_1)} . \quad (9)$$

By the use of Gibrat’s Law (4), the Detailed Balance is reduced to

$$\frac{P(x_1)}{P(x_2)} = G(R) , \quad (10)$$

²At the same time, the symmetry under the exchange $x_0 \leftrightarrow x_1$ is also confirmed for each database.

³These analyses with respect to Gibrat’s Law from 2004 to 2005 are also valid in the analyses from 2003 to 2004.

where we define $G(R) \equiv Q(R^{-1})/(RQ(R))$. By setting $R = 1$ after differentiating Eq. (10) with respect to R , we obtain the following differential equation

$$G'(1)P(x) = -xP'(x) , \quad (11)$$

where x denotes x_1 . The solution is given by

$$P(x) = Cx^{-G'(1)} . \quad (12)$$

This is identical to Pareto's Law (1) with $G'(1) = \mu + 1$. Note that Gibrat's Law is valid only in cases where the initial value is larger than some threshold x_{th} .⁴ This threshold is coincident with the threshold in Pareto's Law because there is no threshold in the Detailed Balance (Fig. 5 and Appendix).

In order to make Pareto's Law clear, we consider the cumulative number:

$$\begin{aligned} N_{PL}(> x) &= N_{PL}(> x_{th})P_{PL}(> x) = N_{PL}(> x_{th}) \int_x^\infty dt P_{PL}(t) \\ &= N_{PL}(> x_{th}) \left(\frac{x}{x_{th}} \right)^{-\mu} \quad \text{for } x > x_{th} . \end{aligned} \quad (13)$$

Pareto's Law is confirmed in Databases I – III (Figs. 9 – 11). In Fig. 9 for the cumulative number plot of income, Pareto's Law holds for cases over approximately 100 million yen (The number of firms in the region is approximately 25 thousand). This is related to the fact that Gibrat's Law is observed for $n = 2, \dots, 5$ in Fig. 6. In Fig. 10 for the cumulative number plot of sales, Pareto's Law holds for cases over approximately 200 million yen (The number of firms in the region is approximately 315 thousand). This is related to the fact that Gibrat's Law is observed for $n = 3, \dots, 20$ in Fig. 7. Each threshold comes from the lower bound of the data.

In Fig. 11 for the cumulative number plot of profits, Pareto's Law holds for profits over approximately 100 million yen (The number of firms in the region is approximately 15 thousand). This corresponds to the fact that Gibrat's Law is observed for $n = 16, \dots, 20$ in Fig. 8. This threshold does not come from the lower bound of the data. For $n = 1, \dots, 15$, as n increases, the growth rate distributions change according to a law that we have designated Non-Gibrat's Law.

2.2 Log-normal Distribution from the Detailed Balance and Non-Gibrat's Law

In the literature (Ishikawa 2006a, 2007a), the log-normal distribution is analytically derived from the Detailed Balance and Non-Gibrat's Law. In order to identify Non-Gibrat's Law in the middle scale region, we approximate $\log_{10} q(r|x_1)$ in Fig. 8 by linear functions of r as follows:

$$\log_{10} q(r|x_1) = c - t_+(x_1) r \quad \text{for } r > 0 , \quad (14)$$

$$\log_{10} q(r|x_1) = c + t_-(x_1) r \quad \text{for } r < 0 . \quad (15)$$

These approximations are not appropriate for $n = 1, \dots, 5$; therefore, we consider

⁴If Gibrat's Law holds for all $x_1 \in [0, \infty]$, then $P(x_1)$ cannot be a pdf (Fujiwara et al. 2004).

the case for $n = 6, \dots, 20$. Eqs. (14) and (15) are expressed as exponential functions:

$$Q(R|x_1) = d R^{-t_+(x_1)-1} \quad \text{for } R > 1, \quad (16)$$

$$Q(R|x_1) = d R^{+t_-(x_1)-1} \quad \text{for } R < 1, \quad (17)$$

where $d = 10^c / \ln 10$. Under these approximations, the Detailed Balance (9) is reduced to

$$\frac{P(x_1)}{P(x_2)} = R^{+t_+(x_1)-t_-(x_2)+1} \quad (18)$$

for the case of $R > 1$. Interestingly, $t_{\pm}(x)$ in the approximations (14) and (15) are uniquely fixed under the Detailed Balance.

By setting $R = 1$ after differentiating Eq. (18) with respect to R , we obtain the following differential equation

$$\left[1 + t_+(x) - t_-(x)\right] P(x) + x P'(x) = 0, \quad (19)$$

where x denotes x_1 . The same differential equation is obtained for $R < 1$. Similarly, from the second and third derivatives of Eq. (18), the following differential equations are obtained:

$$t_+'(x) + t_-'(x) = 0, \quad t_+'(x) + x t_+''(x) = 0. \quad (20)$$

The solutions $t_{\pm}(x)$ are uniquely fixed as

$$t_{\pm}(x) = t_{\pm}(x_{\text{th}}) \pm \alpha \ln \frac{x}{x_{\text{th}}}. \quad (21)$$

With Eq. (19), $t_{\pm}(x)$ also uniquely fix the pdf $P(x)$ as

$$P(x) = C x^{-[t_+(x_{\text{th}})-t_-(x_{\text{th}})+1]} e^{-\alpha \ln^2 \frac{x}{x_{\text{th}}}} \quad \text{for } x > x_{\text{min}}. \quad (22)$$

These solutions satisfy Eq. (18) beyond perturbation around $R = 1$ under the restricted assumption of Eqs. (14) and (15).

These analytic results are confirmed in Database III. By applying the linear approximations (14) and (15) to the data in Fig. 8, the relation between x and $t_{\pm}(x)$ is obtained (Fig. 12). On the one hand, Fig. 12 shows that $t_{\pm}(x)$ hardly responds to x for $n = 15, \dots, 20$. This means that Gibrat's Law is valid for the region of large scale firms. On the other hand, $t_+(x)$ linearly increases and $t_-(x)$ linearly decreases symmetrically with $\log_{10} x$ for $n = 6, \dots, 10$. This is Non-Gibrat's Law (21) derived analytically by the linear approximations (14) and (15).

Non-Gibrat's Law (21) and the resultant pdf (22) are considered as Gibrat's Law and Pareto's Law, respectively, for the case of $\alpha = 0$. We apply Eqs. (21) and (22) not only in the middle scale region but also in the large scale one. In this sense, we refer to Eq. (21) as Extended-Gibrat's Law. The parameters are estimated as follows: $\alpha \sim 0$ for $x > x_{\text{th}}$, $\alpha \sim 0.14$ for $x_{\text{min}} < x < x_{\text{th}}$, $t_+(x_{\text{th}}) \sim 2$, $t_-(x_{\text{th}}) \sim 1$, $x_{\text{th}} \sim 10^{2+0.2(16-1)} = 10^5$ thousand (= 100 million) yen and $x_{\text{min}} \sim 10^{2+0.2(6-1)} = 10^3$ thousand (= 1 million) yen. Rigorously, a constant parameter α must not take different values. In the database, however, a large number of firms stay in the same

region for two successive years. This parameterization is approximately valid for describing the pdf. This is confirmed in Fig. 13. This figure shows that 14,800 firms (a number which accounts for approximately 8.3% of the total data and whose profits account for approximately 91.6% of the total profits) are included in the large scale region ($x \geq x_{th}$). In the middle scale region ($x_{min} \leq x_1 < x_{th}$), there are 130,018 firms (73.7% of data; 8.3% of total profits). A similar analysis was confirmed in the data from 2003 (x_0) to 2004 (x_1).

3 Distributions in Temporal Change of Firm Size

In analyses in the previous section, we have investigated growth rate distributions of income, sales and profits. There is a noteworthy difference between them. As depicted in Fig. 1, the growth rate distributions of profits can be approximated by linear functions (14) and (15). The validity of the approximations is confirmed by the results. In Fig. 6, these approximations are also appropriate for the growth rate distributions of income. The growth rate distributions of sales are, however, difficult to approximate by the linear functions because the distributions with a curvature have wide tails (Fig. 7) as depicted in Fig. 2. This difference has been observed in other literature by employing not only Japanese firms data but also European and North American firms data (Amaral et al. 1997, Okuyama et al. 1999, Matia et al. 2004, Gabaix 2005 for instance). This aspect has also been observed in other quantities. In the literature (Canning et al. 1998 for instance), the growth rate distributions of GDP have no wide tail. In the literature (Fujiwara et al. 2003), the growth rate distributions of personal income in Japan have wide tails. In the literature (Fujiwara et al. 2004), the growth rate distributions of assets and sales in France and the number of employees in UK have also wide tails.

Where does this difference between figures of the growth rate distributions come from? On the one hand, income and profits of firms are calculated by a subtraction of total expenditure from total sales at a rough estimate. The values can be both positive and negative. On the other hand, assets and sales of firms, the number of employees and personal income are not calculated by any subtraction. The values cannot be negative. From these facts, we make a simple assumption that the difference between figures of growth rate distributions comes from being subtracted. In order to verify this assumption, we will investigate the temporal change of firm size data. If our assumption is correct, the growth rate distributions in the temporal change of firm size data are approximated by linear functions.

Firstly, we analyze the temporal change of sales data, the number of which is the largest in three databases, Databases I – III. In the analysis, we used sales data over 400 million yen, the value of which is sufficiently larger than the obscure lower bound of the data (Figs. 4 and 7). These sales data are in Pareto's Law region (Fig. 10). Let us consider two temporal changes $v_{12} = x_2 - x_1$ and $v_{01} = x_1 - x_0$. Here, v_{12} is the change between 2004 (x_1) and 2005 (x_2), and v_{01} is between 2003 (x_0) and 2004 (x_1). The temporal changes v_{01} and v_{12} can be both negative and positive. The data are classified into the following four cases: ($v_{01} > 0, v_{12} > 0$), ($v_{01} > 0, v_{12} < 0$), ($v_{01} < 0, v_{12} > 0$) and ($v_{01} < 0, v_{12} < 0$).

In each case, distributions in the growth rate of temporal sales changes $R =$

$|v_{12}/v_{01}|$ are shown in Fig. 14. In four cases, no wide tail is observed as expected. The assumption is valid at least in this database. The distributions are approximated by linear functions as

$$\log_{10} q(r|v_{01}) = c - t_+(|v_{01}|) r \quad \text{for } r > 0, \quad (23)$$

$$\log_{10} q(r|v_{01}) = c + t_-(|v_{01}|) r \quad \text{for } r < 0. \quad (24)$$

Here, we take the absolute value of v because it can be negative. Furthermore, Extended-Gibrat's Law is approximately confirmed in each case (Fig. 15) as follows:

$$t_{\pm}(|v_{01}|) = t_{\pm}(|v_{th}|) \pm \alpha \ln \frac{|v_{01}|}{|v_{th}|}. \quad (25)$$

The distributions in the temporal sales changes $|v_{01}|$ and $|v_{12}|$ are shown in Fig. 16, in which not only Pareto's Law in the large scale region but also the log-normal distribution in the middle scale region is observed. Figure 16 shows that Pareto indices for $|v_{01}|$ and $|v_{12}|$ are approximately the same value in each figure. This fact and Extended-Gibrat's Law (25) suggest that there is a Detailed Balance under exchange $|v_{01}| \leftrightarrow |v_{12}|$ in each case.⁵ The scatter plots of the temporal sales changes are shown in Fig. 17. In each case, by using the K-S test given in the Appendix, the following Detailed Balance is approximately observed:

$$P_{12}(|v_{01}|, |v_{12}|) = P_{12}(|v_{12}|, |v_{01}|). \quad (26)$$

In the temporal sales change data, the Detailed Balance (26) and Extended-Gibrat's Law (25) are observed. The distribution of the temporal sales change data, therefore, follows the Pareto's Law in the large scale region and the log-normal distribution in the middle scale:

$$P(|v|) = C v^{-[t_+(|v_{th}|) - t_-(|v_{th}|) + 1]} e^{-\alpha \ln^2 \frac{|v|}{|v_{th}|}} \quad \text{for } |v| > |v_{min}|. \quad (27)$$

In the same manner as the analysis of the profits data, we confirm this in Fig. 18. The parameters are estimated as follows: $\alpha \sim 0$ for $|v| > |v_{th}|$, $\alpha \neq 0$ for $|v_{min}| < |v| < |v_{th}|$, $t_+(|v_{th}|) - t_-(|v_{th}|) \sim 1$, $|v_{th}| \sim 10^{4+0.5(5-1)} = 10^6$ thousand (=1 billion) yen and $x_{min} \sim 10^{4+0.5(2-1)} = 10^{4.5}$ thousand (=10 million) yen. In each case, approximately 5~10% of firms in the data are included in the large scale region and approximately 65~70% of firms in the data exist within the middle scale.

Similar phenomena are observed in Databases I and II. In our analysis of the temporal high-income change in Database I, this phenomenon is confirmed for the case in which the growth rate distribution of firm size has no wide tail and the data are completely exhaustive. In the analysis of the temporal positive-profits change in Database II, this phenomenon is also confirmed for the case in which the growth rate distribution of firm size has no wide tail and the data cover the middle scale region.

⁵If Pareto indices vary, there is thought to be a Detailed Quasi-Balance (Ishikawa 2006b, 2007b).

4 Conclusion and Future Issues

In this study, we have shown that the signed temporal change of firm size data follows not only power-law in the large scale region but also the log-normal distribution in the middle scale region. In the analyses, we employed three databases: high-income data (Database I), high-sales data (Database II) and positive-profits data (Database III) of Japanese firms. It is particularly worth noting that the growth rate distributions in the temporal change of firm size have no wide tail as is observed in the assets and sales of firms, the number of employees and personal income data. The growth rate distribution with no wide tail can be linearly approximated. This lack of a wide tail is also observed in the temporal change of the firm size, such as income and profits of firms. From these observations, we conclude that the quantity calculated by any subtraction has no wide tail in the growth rate distribution and vice versa.

In the data of temporal firm size changes, the Detailed Balance was also confirmed. This leads to Extended-Gibrat's Law. At the same time, the Pareto indices are almost the same value in the large scale regions of two successive temporal change data. The Detailed Balance and Extended-Gibrat's Law lead to Pareto's Law in the large scale region and the log-normal distribution in the middle scale one. This is consistently confirmed in the empirical data.

From the growth rate distribution in the temporal firm size changes with no wide tail, it is possible to derive the following schemes analytically or numerically (Tomoyose et al. 2008). (Scheme A) The growth rate distribution of x , which cannot be negative (assets and sales of firms, the number of employees and personal income), has wide tails (Fig. 2). (Scheme B) The growth rate of distribution x , which can be negative (profits and income of firms), has no wide tail (Fig. 1). In addition, the difference of Non-Gibrat's Laws might become more distinct. On the one hand, in the firm size growth rate distributions with no wide tail (Fig. 1), the probability of positive growth decreases and the probability of negative growth increases symmetrically as the bin size of the classification of x increases in the middle scale region. On the other hand, in the firm size distributions with wide tails (Fig. 2), the probability of positive and negative growth decreases simultaneously as the bin size of the classification of x increases.

The data analyses in this study are presumably important in the management of credit risk, and they should be considered in any system of taxation. Furthermore, the mechanism described in this paper might be useful for an understanding of aggregate phenomena in macro-economics (Gabaix 2005).

Appendix

We approximately confirmed the Detailed Balance (3) by taking the one-dimensional Kolmogorov–Smirnov (K–S) test. For the scatter plot in high-income Database I (Fig. 3), we compared the distribution sample for $P(x_1 \in 4 \times [10^{4+0.4(n-1)}, 10^{4+0.4n}), x_2)$ with another sample for $P(x_1, x_2 \in 4 \times [10^{4+0.4(n-1)}, 10^{4+0.4n}))$ ($n = 1, 2, \dots, 10$) by making the null hypothesis that these two samples are taken from the same parent distribution. Each p value is shown in Fig 19. In most cases, the null hypothesis was not rejected within the 5% significance level. We recognize that the Detailed Balance (3) in Fig. 3 is approximately observed.

For the scatter plot in high-sales Database II (Fig. 4), we compared the distribution sample for $P(x_1 \in 2 \times [10^{5+0.4(n-1)}, 10^{5+0.4n}), x_2)$ with another sample for $P(x_1, x_2 \in 2 \times [10^{5+0.4(n-1)}, 10^{5+0.4n}))$ ($n = 1, 2, \dots, 10$). Each p value is shown in Fig 20. For the scatter plot in positive-profits Database III (Fig. 5), we compared the distribution sample for $P(x_1 \in [10^{3+0.25(n-1)}, 10^{3+0.25n}), x_2)$ with another sample for $P(x_1, x_2 \in [10^{3+0.25(n-1)}, 10^{3+0.25n}))$ ($n = 1, 2, \dots, 20$). Each p value is shown in Fig 21. We also recognize that the Detailed Balances in Fig. 4 or Fig. 5 are approximately observed.

In these data analyses, we should have taken into account the trends in the average growth to test the Detailed Balance accurately. However, in the original database many of the figures in the data had been rounded off. If we had subtracted the trends from data, the values would have been displaced. As a result, p values were underestimated by the displacements. In the databases, the effect of the trends in the average growth is not too large, so we could confirm the Detailed Balance approximately without subtracting the trends.

For this reason, we did not subtract the trends from the scatter plots in the temporal sales changes (Fig. 17). We compared the distribution sample for $P(|v_{01}| \in [10^{4+0.5(n-1)}, 10^{4+0.5n}), |v_{12}|)$ with another sample for $P(|v_{01}|, |v_{12}| \in [10^{4+0.5(n-1)}, 10^{4+0.5n}))$ ($n = 1, 2, \dots, 10$) by making the null hypothesis that these two samples are taken from the same parent distribution. Each p value is shown in Fig 22. We also recognize that the Detailed Balances in Fig. 17 are approximately observed.

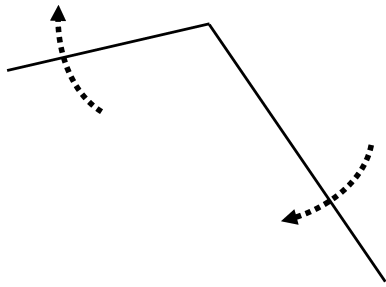


Figure 1: The growth rate distribution of profits or income of firms. The horizontal axis is the logarithm of the growth rate and the vertical axis is the logarithm of its pdf.

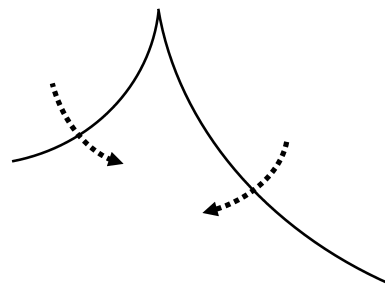


Figure 2: The growth rate distribution of assets and sales of firms, the number of employees or personal income.

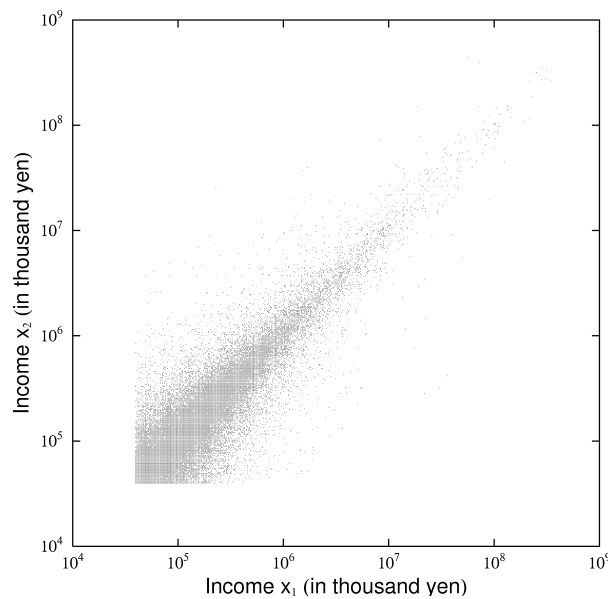


Figure 3: The scatter plot of firms in Database I, the income of which in 2003 (x_0), 2004 (x_1) and 2005 (x_2) exceeded 4×10^4 thousand yen: $x_0 > 4 \times 10^4$ and $x_1 > 4 \times 10^4$ and $x_2 > 4 \times 10^4$. The number of firms is 40,829.

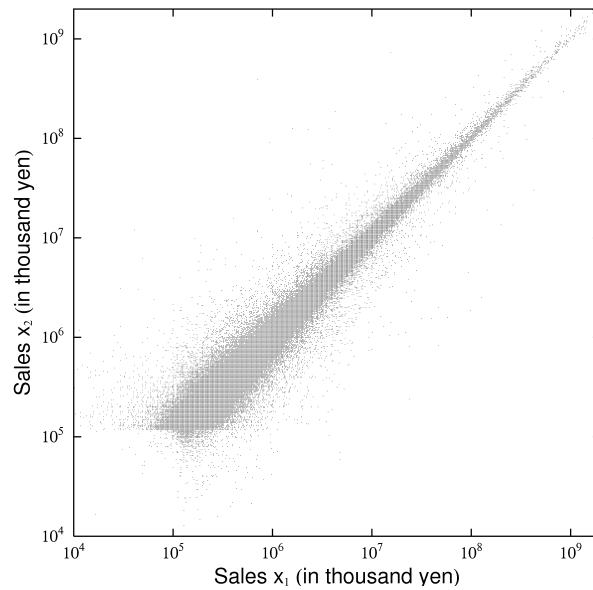


Figure 4: The scatter plot of firms in Database II, the sales of which in 2003 (x_0), 2004 (x_1) and 2005 (x_2) exceeded 0 yen: $x_0 > 0$ and $x_1 > 0$ and $x_2 > 0$. The number of firms is 406,385.

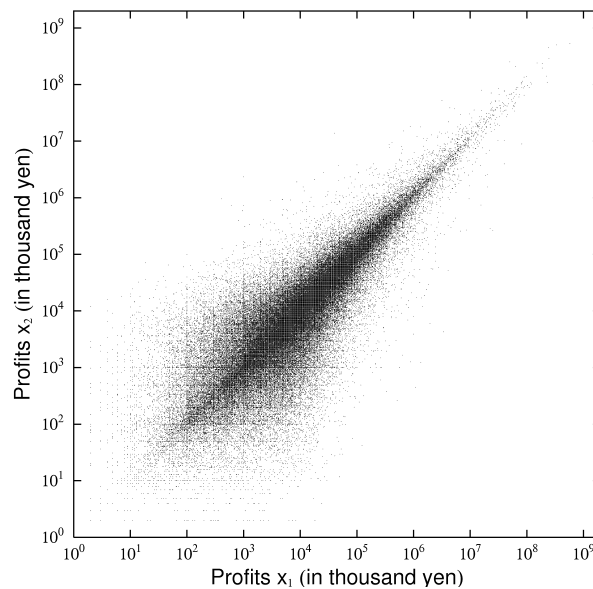


Figure 5: The scatter plot of firms in Database III, the profits of which in 2003 (x_0), 2004 (x_1) and 2005 (x_2) exceeded 0 yen: $x_0 > 0$ and $x_1 > 0$ and $x_2 > 0$. The number of firms is 177,492.

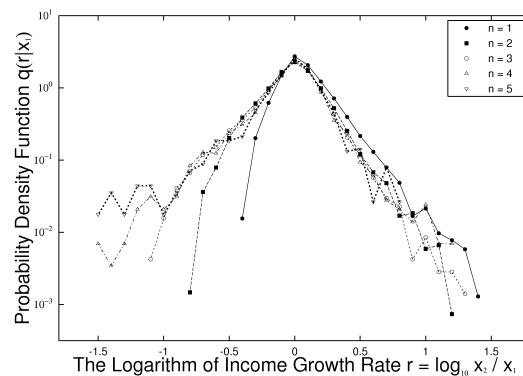


Figure 6: Conditional pdfs $q(r|x_1)$ of the log income growth rate $r = \log_{10} x_2/x_1$ from 2004 to 2005. The data points are classified into five bins of the initial income with equal magnitude in logarithmic scale, $x_1 \in 4 \times [10^{4+0.4(n-1)}, 10^{4+0.4n}]$ ($n = 1, 2, \dots, 5$) thousand yen. The data for large negative growth, $r \leq 4 + \log_{10} 4 - \log_{10} x_1$, are not available because of the lower bound of the high-income data, 4×10^4 thousand (= 40 million) yen.

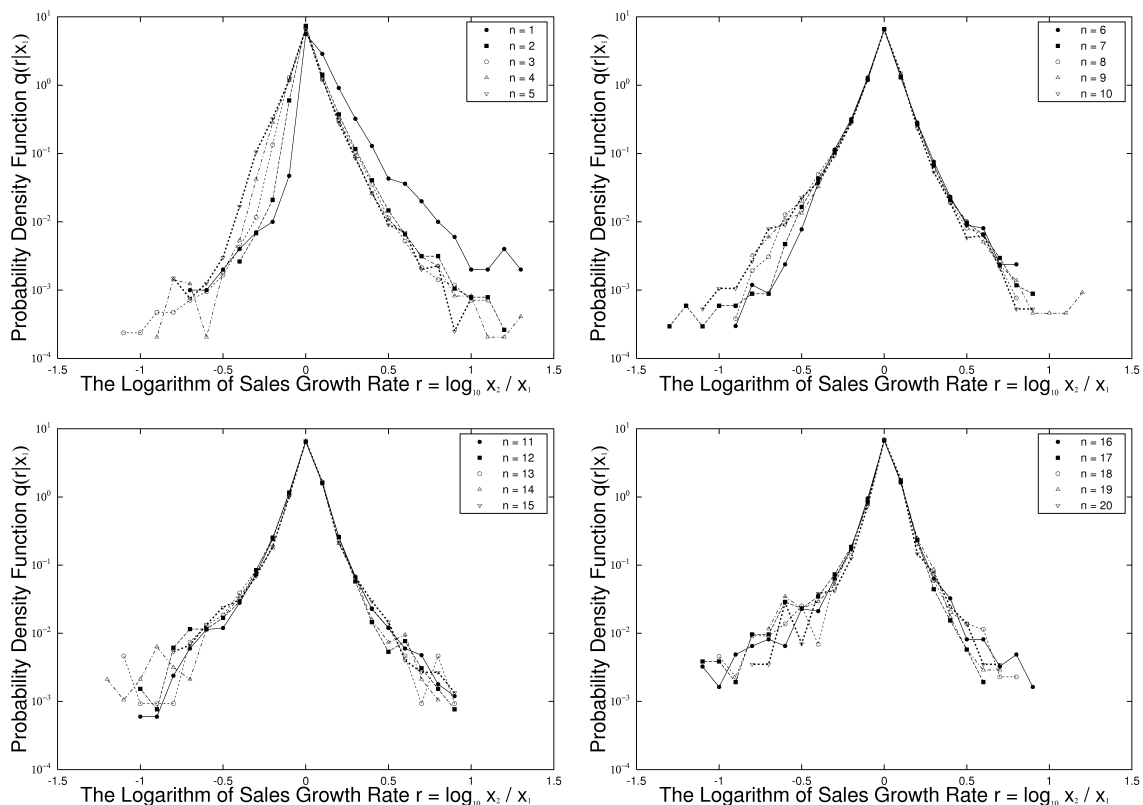


Figure 7: Conditional pdfs $q(r|x_1)$ of the log sales growth rate $r = \log_{10} x_2/x_1$ from 2004 to 2005. The data points are classified into twenty bins of the initial sales with equal magnitude in logarithmic scale, $x_1 \in [10^{5+0.2(n-1)}, 10^{5+0.2n}]$ ($n = 1, 2, \dots, 20$) thousand yen.

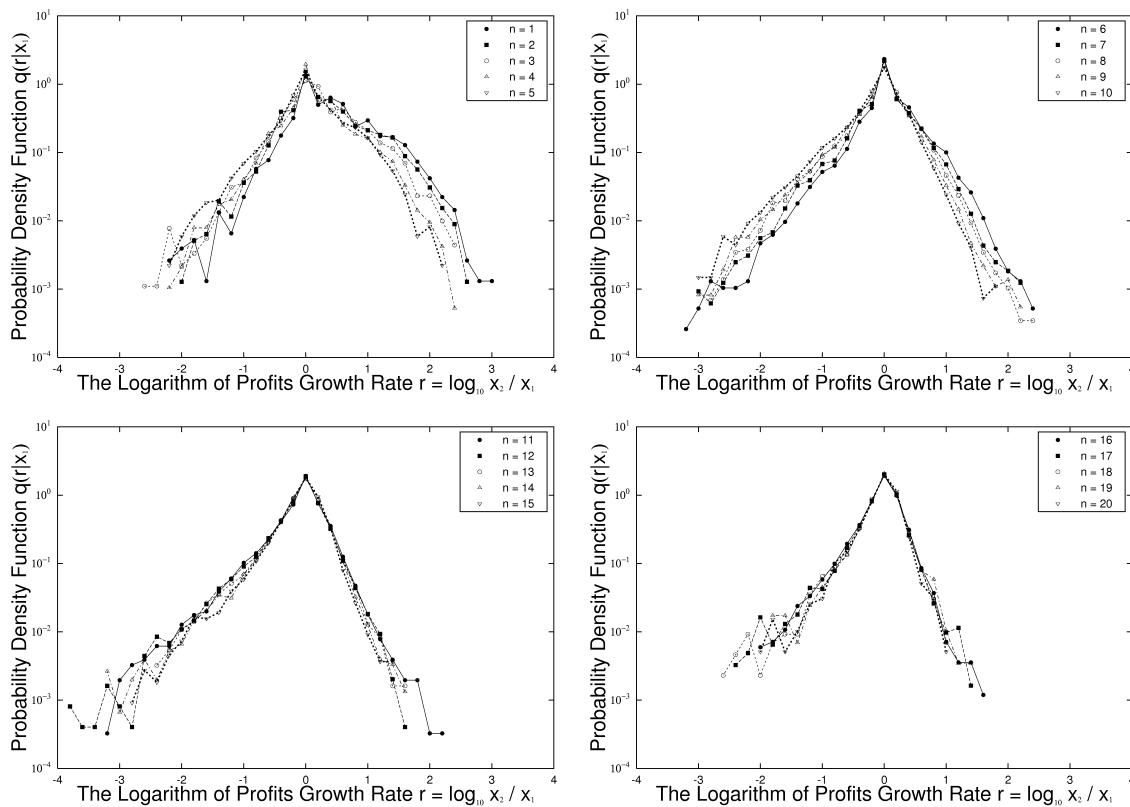


Figure 8: Conditional pdfs $q(r|x_1)$ of the log profits growth rate $r = \log_{10} x_2/x_1$ from 2004 to 2005. The data points are classified into twenty bins of the initial profits with equal magnitude in logarithmic scale, $x_1 \in [10^{2+0.2(n-1)}, 10^{2+0.2n}]$ ($n = 1, 2, \dots, 20$) thousand yen.

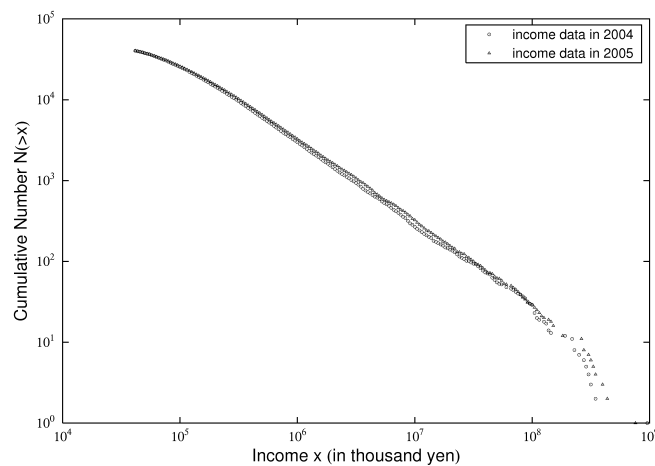


Figure 9: Cumulative number plots of income in Database I, the income of which in 2003 (x_0), 2004 (x_1) and 2005 (x_2) exceeded 4×10^4 thousand yen: $x_0 > 4 \times 10^4$ and $x_1 > 4 \times 10^4$ and $x_2 > 4 \times 10^4$. In the large scale region over approximately 10^5 thousand (=100 million) yen, Pareto's Law is observed. Each Pareto index is estimated to be nearly 1.

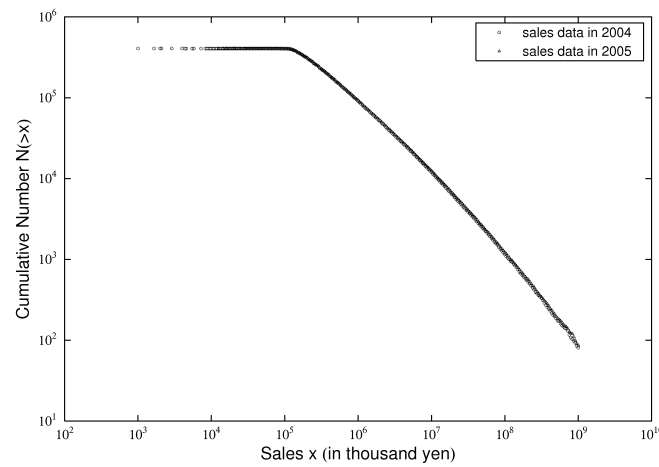


Figure 10: Cumulative number plots of sales in Database II, the sales of which in 2003 (x_0), 2004 (x_1) and 2005 (x_2) exceeded 0 yen: $x_0 > 0$ and $x_1 > 0$ and $x_2 > 0$. In the large scale region over approximately 2×10^5 thousand (=200 million) yen, Pareto's Law is observed. Each Pareto index is estimated to be nearly 1.

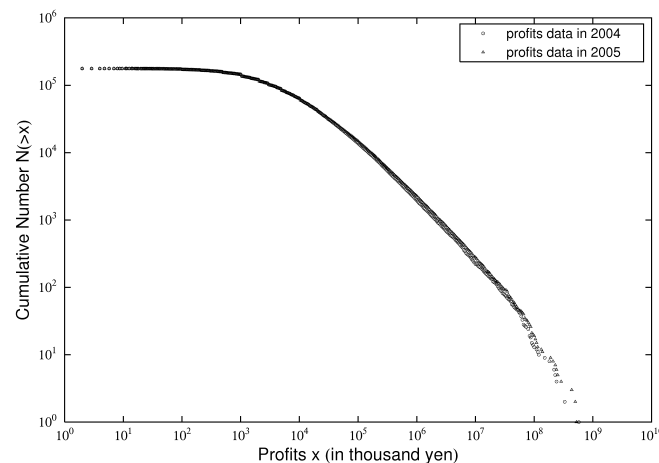


Figure 11: Cumulative number plots of positive-profits in Database III, the profits of which in 2003 (x_0), 2004 (x_1) and 2005 (x_2) exceeded 0 yen: $x_0 > 0$ and $x_1 > 0$ and $x_2 > 0$. In the large scale region over approximately 10^5 thousand (=100 million) yen, Pareto's Law is observed. Each Pareto index is estimated to be nearly 1.

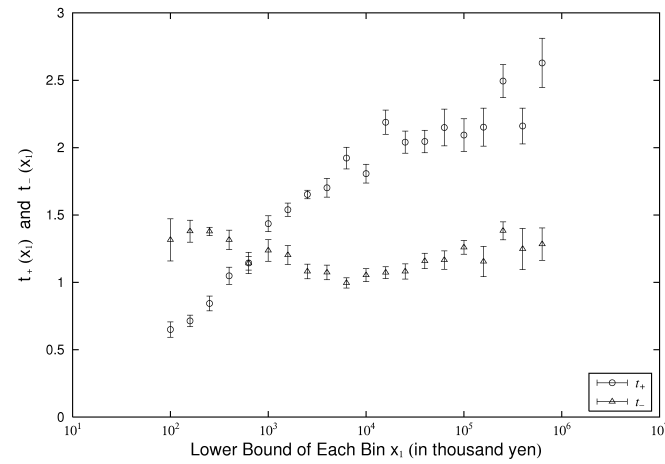


Figure 12: The relation between the lower bound of each bin $x_1 \in [10^{2+0.2(n-1)}, 10^{2+0.2n}]$ and $t_{\pm}(x_1)$. From the left, each point on the graph represents $n = 1, 2, \dots, 20$. The values are measured by the least square method in the region $0 \leq |r| \leq 2$ in Fig. 8.

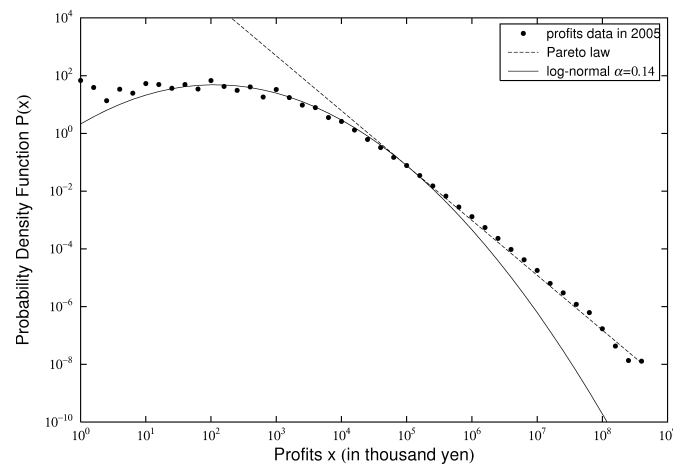


Figure 13: The pdf of positive-profits in Database III.

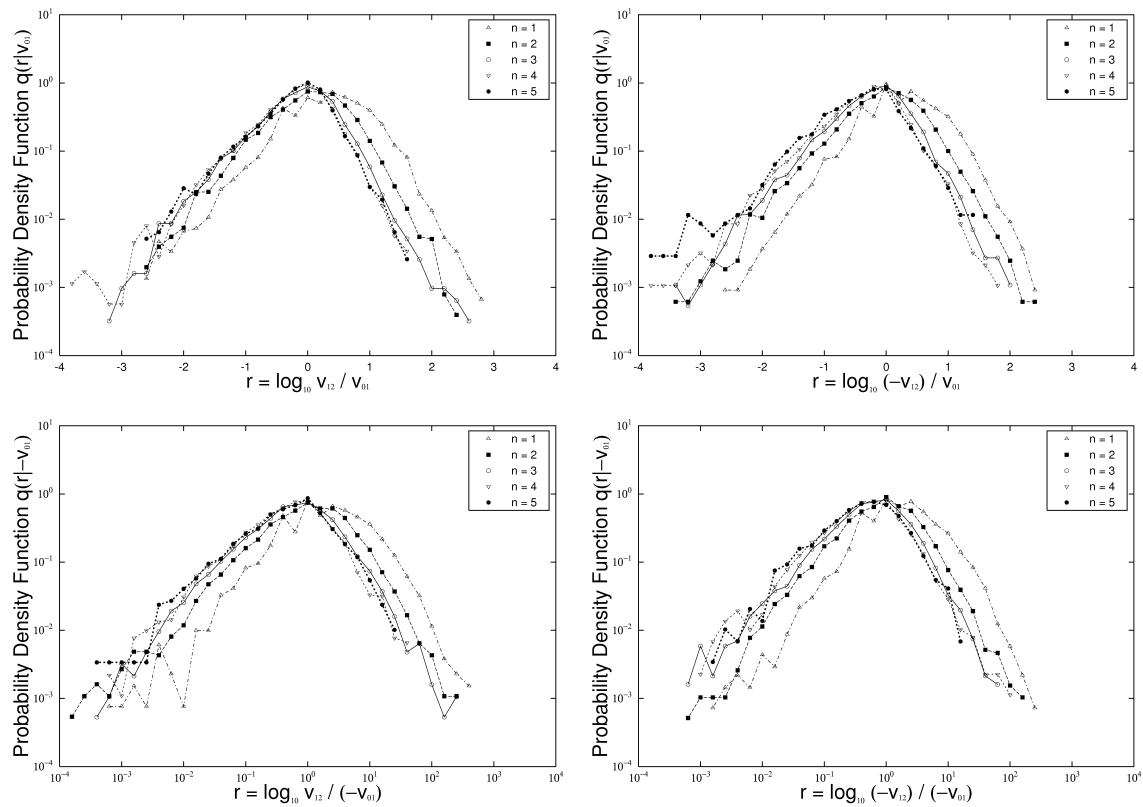


Figure 14: Conditional pdfs $q(r|v_{01})$ in the log growth rate of the temporal sales change $r = \log_{10} |v_{12}/v_{01}|$ for cases $(v_{01} > 0, v_{12} > 0)$, $(v_{01} > 0, v_{12} < 0)$, $(v_{01} < 0, v_{12} > 0)$ and $(v_{01} < 0, v_{12} < 0)$. The number of data is “54,181”, “32,959”, “35,218” and “35,272”, respectively. In each figure, data points are classified into five bins of the initial temporal sales change with equal magnitude in logarithmic scale, $|v_{01}| \in [10^{4+0.5(n-1)}, 10^{4+0.5n}]$ ($n = 1, 2, \dots, 5$) thousand yen. Here, $v_{12} = x_2 - x_1$ is the change between 2004 (x_1) and 2005 (x_2), and $v_{01} = x_1 - x_0$ is between 2003 (x_0) and 2004 (x_1). Each sales data x_0 , x_1 and x_2 exceeded 4×10^5 thousand (=400 million) yen: $x_0 > 4 \times 10^5$ and $x_1 > 4 \times 10^5$ and $x_2 > 4 \times 10^5$.

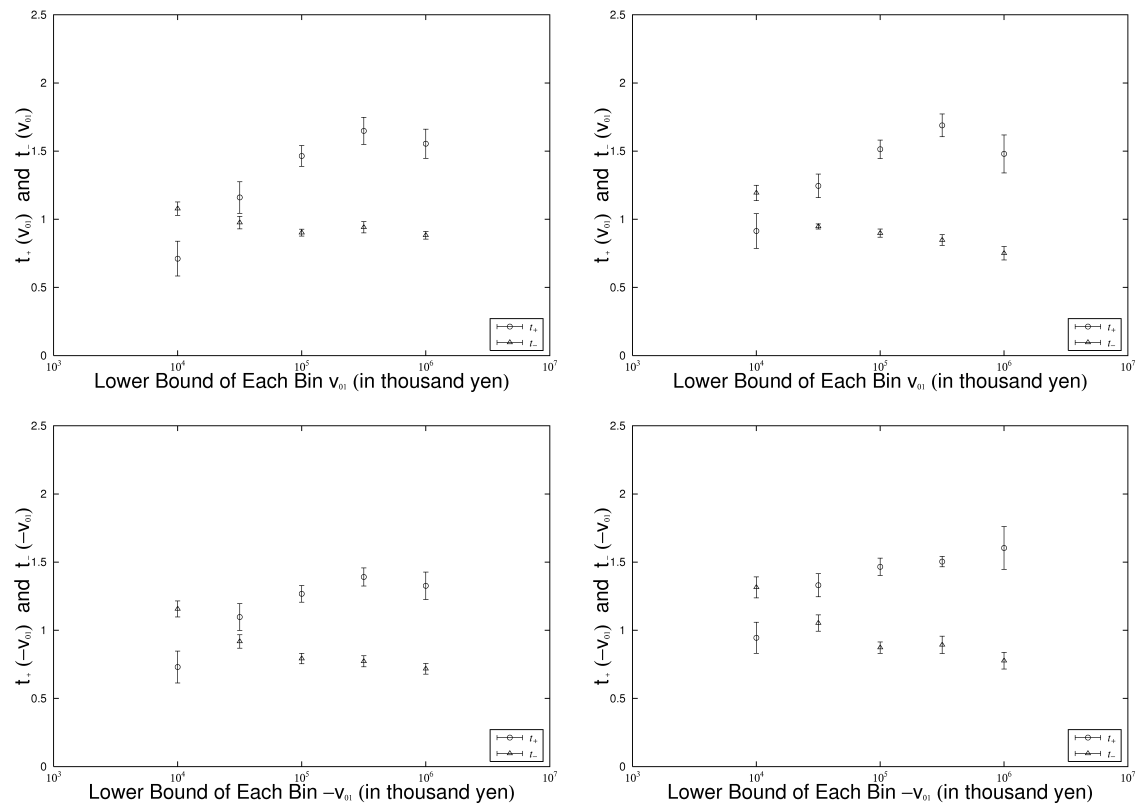


Figure 15: The relation between the lower bound of each bin $|v_{01}| \in [10^{4+0.5(n-1)}, 10^{4+0.5n}]$ and $t_{\pm}(|v_{01}|)$. In each figure, from the left each data point represents $n = 1, 2, \dots, 5$. The values are measured by the least square method in the region $0 \leq |r| \leq 2$ in Fig. 14.

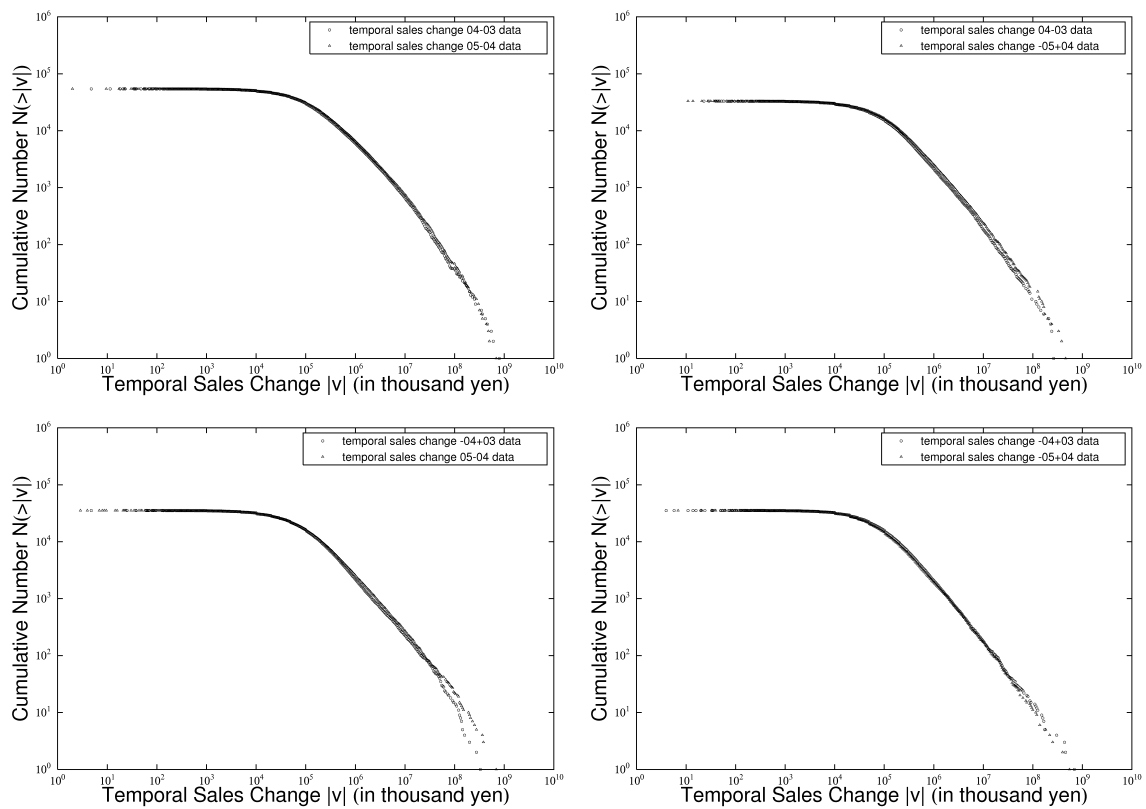


Figure 16: Cumulative number plots in the temporal sales changes for cases $(v_{01} > 0, v_{12} > 0)$, $(v_{01} > 0, v_{12} < 0)$, $(v_{01} < 0, v_{12} > 0)$ and $(v_{01} < 0, v_{12} < 0)$.

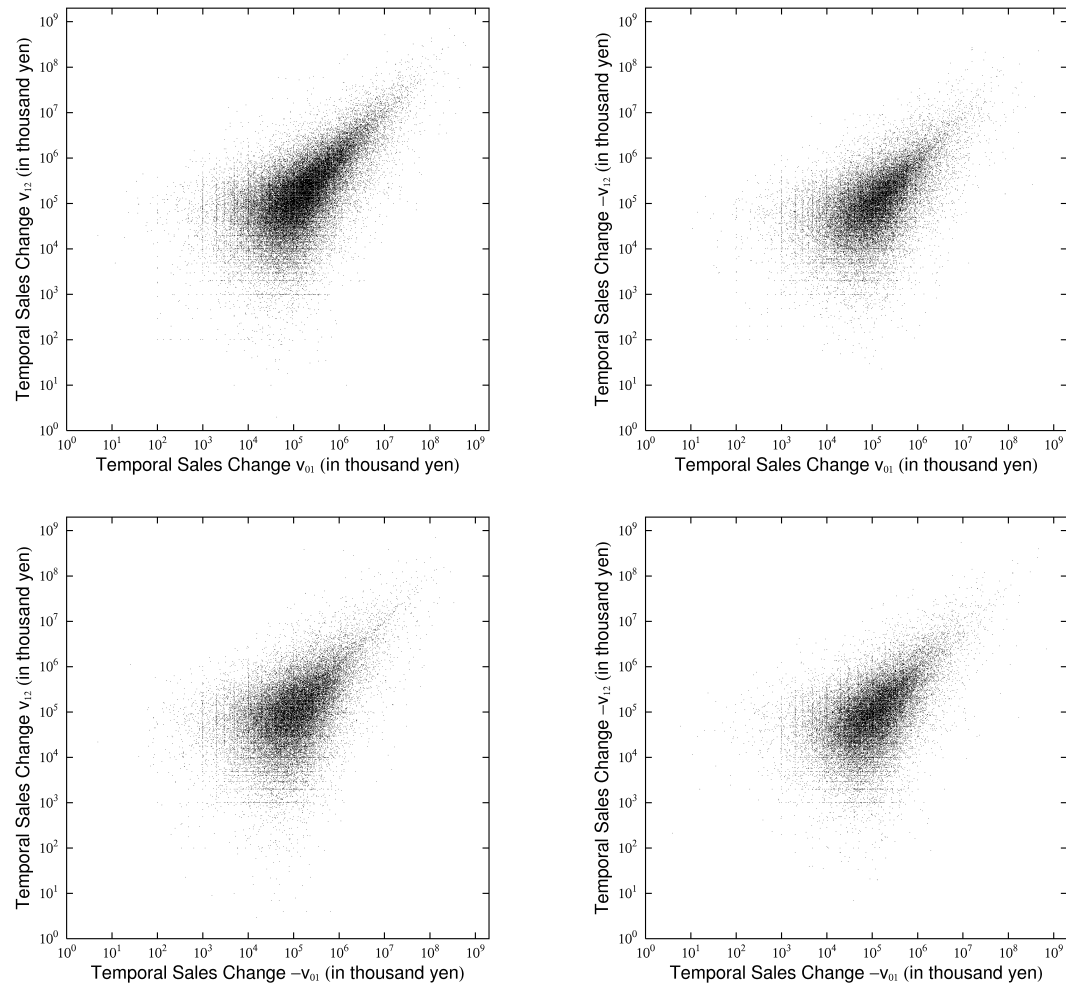


Figure 17: Scatter plots in the temporal sales changes for cases $(v_{01} > 0, v_{12} > 0)$, $(v_{01} > 0, v_{12} < 0)$, $(v_{01} < 0, v_{12} > 0)$ and $(v_{01} < 0, v_{12} < 0)$.

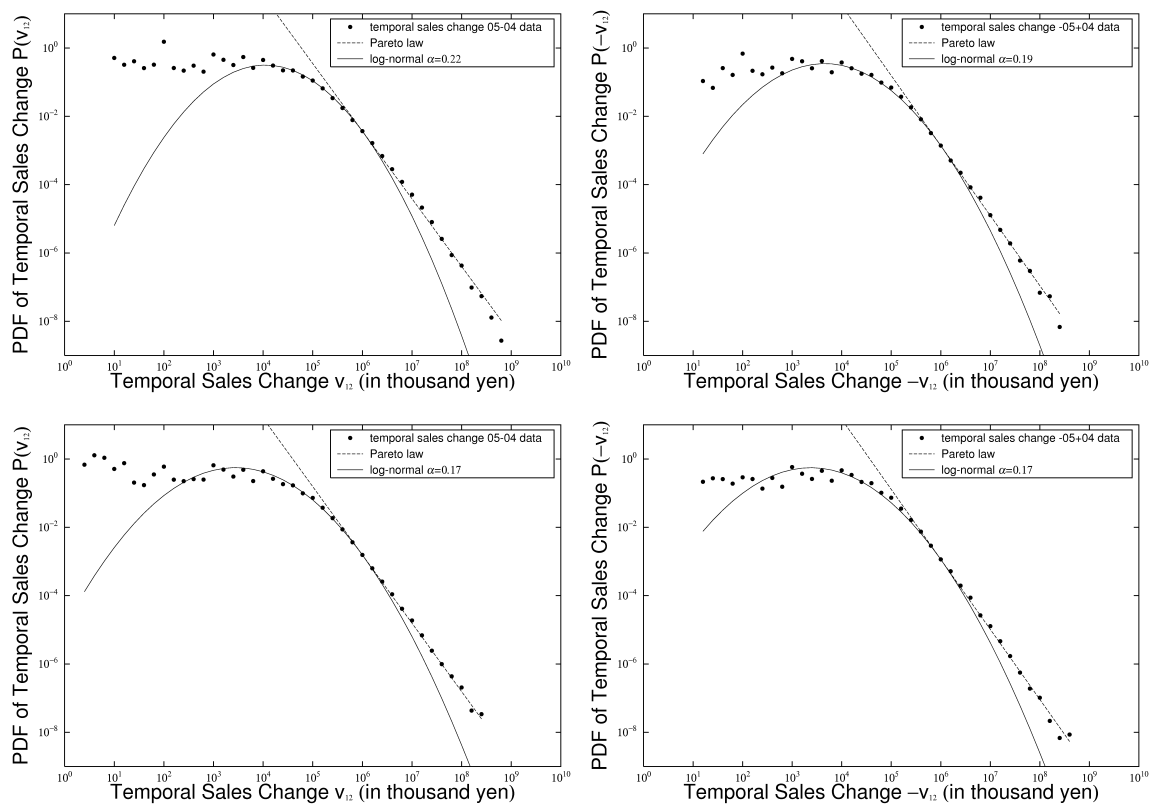


Figure 18: The pdf of the temporal sales change data.

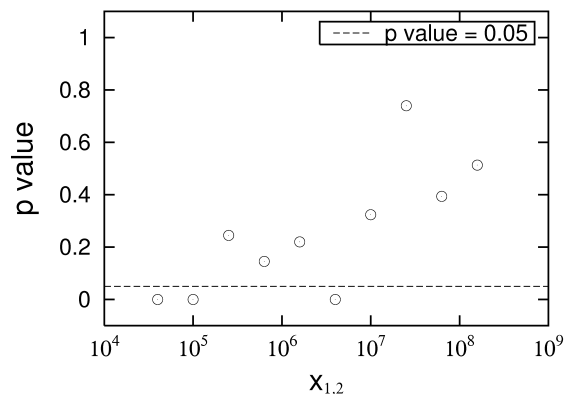


Figure 19: Each p value of the one-dimensional K-S test for the scatter plot of high-income data points (Fig. 3).

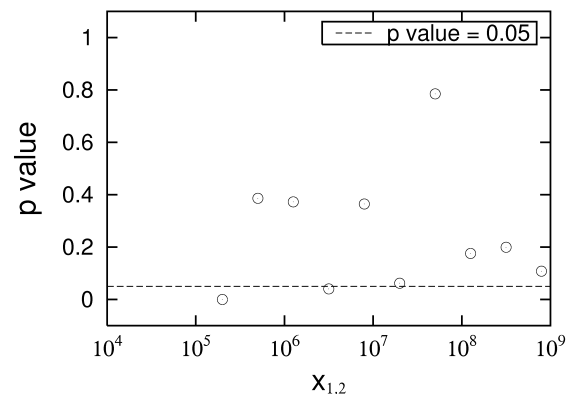


Figure 20: Each p value of the one-dimensional K-S test for the scatter plot of high-sales data points (Fig. 4).

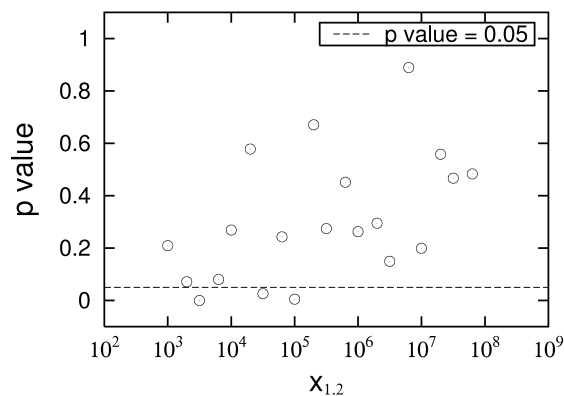


Figure 21: Each p value of the one-dimensional K-S test for the scatter plot of positive-profits data points (Fig. 5).

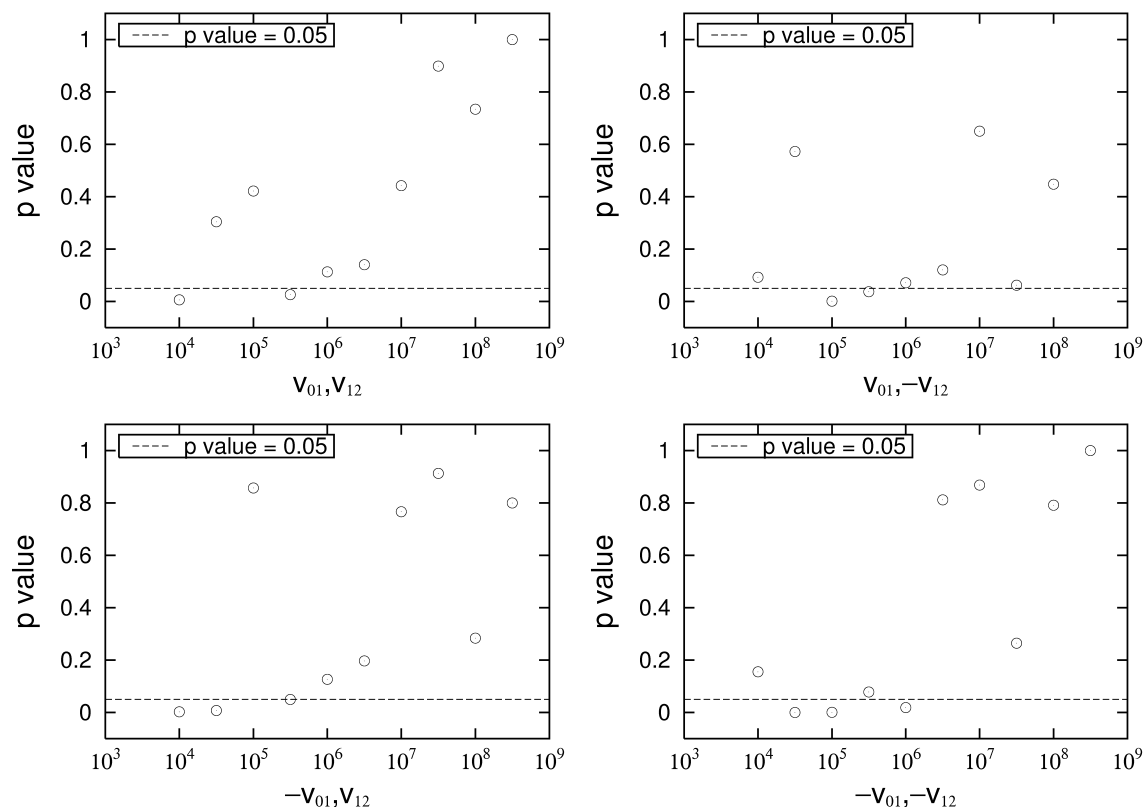


Figure 22: Each p value of the one-dimensional K-S test for the scatter plots in the temporal sales changes (Fig. 17).

References

- Amaral, L.A.N., Buldyrev, S.V., Havlin, S., Leschhorn, H., Maass, P., 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.
- Aoyama, H. (2004). Ninth Annual Workshop on Economic Heterogeneous Interacting Agents (WEHIA 2004).
- Aoyama, H., Fujiwara, Y., and Souma, W. (2004). The Physical Society of Japan 2004 Autumn Meeting.
- Badger, W.W. (1980). In B.J. West (ed.), *Mathematical Models as a Tool for the Social Science*, p. 87. New York: Gordon and Breach.
- Canning, D., Amaral, L.A.N., Lee, Y., Meyer, M., and Stanley, H.E., (1998). Scaling the Volatility of GDP Growth Rates, *Economics Lett.* 60: 335–341.
- Diamond Inc., <http://www.diamond.co.jp/index.shtml>.
- Fujiwara, Y., Souma, W., Aoyama, H., Kaizoji, T., and Aoki, M. (2003). Growth and Fluctuations of Personal Income. *Physica A*321: 598–604.
- Fujiwara, Y., Guilmi, C.D., Aoyama, H., Gallegati, M., and Souma, W. (2004). Do Pareto–Zipf and Gibrat laws hold true? An analysis with European Firms. *Physica A*335: 197–216.
- Gabaix, X. (2005). The Granular Origins of Aggregate Fluctuations. MIT and NBER Working Paper.
- Gibrat, R. (1931). *Les inegalites economiques*. Paris: Sirey.
- Ishikawa, A. (2006a). Derivation of the distribution from extended Gibrat’s law. *Physica A*367: 425–434.
- Ishikawa, A. (2006b). Annual change of Pareto index dynamically deduced from the law of detailed quasi-balance. *Physica A*371: 525–535.
- Ishikawa, A. (2007a). The uniqueness of firm size distribution function from tent-shaped growth rate distribution. *Physica A*383/1: 79–84.
- Ishikawa, A. (2007b). Quasistatically varying log-normal distribution in the middle scale region of Japanese land prices. arXiv:0710.1893. To appear in *Prog. Theor. Phys. Suppl.*
- Matia, K., Fu, D., Buldyrev, S.V., Pammolli, F., Riccaboni, M., and Stanley, H.E. (2004). Statistical properties of business firms structure and growth. *Europhys. Lett.*, 67: 498–503.
- Montroll, E.W., and Shlesinger, M.F. (1983). Maximum Entropy Formalism, Fractals, Scaling Phenomena, and $1/f$ Noise: A tale of Tails. *J. Stat. Phys.* 32: 209–230.
- Okuyama, K., Takayasu, M., and Takayasu, H. (1999). Zip’s law in income distribution of companies. *Physica A*269: 125–131.
- Pareto, V. (1897). *Cours d’Economie Politique*. London: Macmillan.

TOKYO SHOKO RESEARCH, LTD., <http://www.tsr-net.co.jp/.p>

Tomoyose, M., Fujimoto, S., and Ishikawa, A. (2008). Non-Gibrat's law in the middle scale region. arXiv:0809.3060. To appear in *Prog. Theor. Phys. Suppl.*

Please note:

You are most sincerely encouraged to participate in the open assessment of this article. You can do so by either rating the article on a scale from 1 (bad) to 5 (excellent) or by posting your comments.

Please go to:

www.economics-ejournal.org/economics/journalarticles/2009-11

The Editor