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

*Atushi Ishikawa
Kanazawa Gakuin University, Kanazawa 920-1392, Japan

Abstract

We have shown that signed temporal changes of firm size variables follow not only power-law in the large scale region but also the log-normal distribution in the middle scale one. In the analyses, we employ 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 which is observed in assets and sales of firms, the number of employees and personal income data. An extended-Gibrat's law is also found in the growth rate distributions of temporal changes of firm size variables. This leads the power-law and the log-normal distributions in the temporal changes of firm size under the detailed balance.

JEL classification numbers : D30, D31, D39
Keywords : Econophysics; temporal change of firm size; Pareto's law; log-normal distribution; (non-)Gibrat's law; detailed balance

*e-mail address: ishikawa@kanazawa-gu.ac.jp

1 Introduction

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

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

where C is a normalization and the power μ is called Pareto index. In general, the power-law is valid only in the large scale region (Badger 1980; Montrll and Shlesinger 1983). The threshold of the large scale region is denoted by $x_{\rm th}$. In the middle scale region below the threshold $x_{\rm th}$, the pdf allegedly follows the log-normal distribution:

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

Here, $\bar{x}$ is a mean value and σ^2 is a variance. The study for these two distributions is highly required. Because a large amount of total economic quantities are occupied by a few percent of firms or persons included in the large scale region. At the same time, a large number of firms or persons exist within the middle scale region.

Recently it is found that these distributions can be explained by laws observed in massive amount of digitized economic data. Fujiwara et al. (2003, 2004) point out that the Pareto's law can be derived from the law of detailed balance and Gibrat's law (Gibrat 1932). Along this line, Ishikawa (2006a, 2007a) shows that the log-normal distribution is also deduced from the detailed balance and Non-Gibrat's law. The detailed balance is time-reversal symmetry observed in the equilibrium system. The Gibrat's law means that the conditional pdf of the growth rate is independent of the initial value. On the other hand, the Non-Gibrat's law describes the dependence of the initial value. The Gibrat's law is observed only in the large scale region, and the Non-Gibrat's law in the middle scale one.

It is interesting to note that there are two types in growth rate distributions. The figure of the growth rate distribution of profits or income of firms (Fig. 1) 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 in the large scale region but also in the middle scale one. The point is that the difference might be related to the difference between Non-Gibrat's laws in the middle scale region. In Fig. 1, the probability of the positive growth decreases and the probability of the negative one increases as the classification of x increases in the middle scale region (Ishikawa 2006a, 2007a). On the other hand in Fig. 2, the probability of the positive and negative growths decreases simultaneously as the classification of x increases (Aoyama 2004a, 2004b). This size dependence in the middle scale region is significant because a large number of firms or persons are included in this region.

In this study, we propose that the figure of the growth rate distribution is determined by the character of economic variables. In concrete terms, the variables are calculated by any subtraction or not. By employing sales, profits and income data of firms, we confirm this proposition.

2 Firm size distributions

In this section, we 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 were announced publicly as "high-income firms" every year, the number of which is about 70 thousand. 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. In the database, positive and negative profits data are also included. The number of positive data is about 300 thousand and that of negative data is about 40 thousand. We exclude the negative data, the number of which is much less than that of the positive data. Because the negative data gathered from high-sales data are exclusive as profits data. The positive data in the middle scale region are not thought to be completely exhaustive. In order to investigate the consistency between laws in the data, however, we employ the positive profits data. In this study, we investigate these 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 the Gibrat's law (Gibrat 1932) are represented as follow:

- **Detailed balance**

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

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

- **Gibrat's law**

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

$$Q(R|x_1) = Q(R) \ , \tag{4}$$

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

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

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

These laws are confirmed in the databases I – III. In order to compare analyses in the next section, we investigate firms data which exist in successive three years 2003 (x_0), 2004 (x_1) and 2005 (x_2). In the scatter plot in each database, the detailed balance (3) is approximately confirmed by taking the one-dimensional Kolmogorov–Smirnov (K–S) test. The details are given in the Appendix. Figures 3 – 5 show the time-reversal symmetry under the exchange $x_1 \leftrightarrow x_2$.[1] The Gibrat's law (4) is also confirmed in each database. Figures 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 $x_{\rm th}$. Here the pdf for $r = \log_{10} R$ defined by $q(r|x_1)$ is related that for R by

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

Note that the 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 data I. This is also observed in Figs. 4 and 7 for high-sales data II; however, the lower bound is probably obscure.[2] The detailed balance and the Gibrat's law have been 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 the 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

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

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

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

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

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

[1] At the same time, the symmetry under the exchange $x_0 \leftrightarrow x_1$ is also confirmed in each database.

[2] These analyses with respect to the Gibrat's law are also valid in the analyses from 2003 to 2004.

By the use of the Gibrat's law (4), the detailed balance is reduced to

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

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

$$G^{'}(1)P(x) = -xP^{'}(x) \ , \tag{11}$$

where x denotes x_1. The solution is given by

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

This is identical to the Pareto's law (1) with $G'(1) = \mu + 1$. Note that the Gibrat's law is valid only in the case that the initial value is larger than some threshold $x_{\rm th}$.[3] This threshold is coincident with the threshold in the Pareto's law, because there is no threshold in the detailed balance (Fig. 5 and Appendix).

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

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

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

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

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

the middle scale region, we approximate $\log_{10} q(r|x_1)$ in Fig. 8 by linear functions of r as follows:

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

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

These approximations are not appropriate for $n = 1, \cdots, 5$; therefore, we consider the case for $n = 6, \cdots, 20$. Equations (14) and (15) are expressed as so-called exponential functions:

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

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

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

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

for $R > 1$ case. 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

$$\Big[1 + t_+(x) - t_-(x)\Big] P(x) + x\ P^{'}(x) = 0\ , \tag{19}$$

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

$$t_+{}^{'}(x) + t_-{}^{'}(x) = 0\ , \quad t_+{}^{'}(x) + x\ t_+{}^{''}(x) = 0\ . \tag{20}$$

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

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

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

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

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

These analytic results are confirmed in the 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). Figure 12 shows that $t_\pm(x)$ hardly responds to x for $n = 15, \cdots, 20$. This means that Gibrat's law is valid in the large scale region. On the other hand, $t_+(x)$ linearly increases and $t_-(x)$ linearly decreases symmetrically with $\log_{10} x$ for $n = 6, \cdots, 10$. This is the Non-Gibrat's law (21) derived analytically by the linear approximations (14) and (15).

The Non-Gibrat's law (21) and the resultant pdf (22) are considered as Gibrat's law and Pareto's law, respectively, for the case $\alpha = 0$. We take Eqs. (21) and (22) not only in the middle scale region but also in the large scale one. In this sense, we call Eq. (21) extended-Gibrat's law. The parameters are estimated as follows: $\alpha \sim 0$ for $x > x_{\rm th}$, $\alpha \sim 0.14$ for $x_{\rm min} < x < x_{\rm th}$, $t_+(x_{\rm th}) \sim 2$, $t_-(x_{\rm th}) \sim 1$, $x_{\rm th} \sim 10^{2+0.2(16-1)} = 10^5$ thousand (= 100 million) yen and $x_{\rm min} \sim 10^{2+0.2(6-1)} = 10^3$ 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 in two successive years. This parameterization is approximately valid for describing the pdf. This is confirmed in Fig. 13. In this figure "14,800" firms (about 8.3% of the data), the profits of which are about 91.6% of the total profits in the data, are included in the large scale region ($x \geq x_{\rm th}$). In the middle scale region ($x_{\rm min} \leq x_1 < x_{\rm th}$), there are "130,018" firms (about 73.3% of the data), the profits of which are about 8.3% of the total profits in the database. Similar analysis is 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, hardly approximated by the linear functions because the distributions with 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 been also 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? 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 form a subtraction. In order to verify this assumption, we investigate the temporal change of firm size data. If the assumption is appropriate, 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 I – III. In the analysis, we take sales data more than 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 the 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

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

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

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

$$t_\pm(|v_{01}|) = t_\pm(|v_{\rm th}|) \pm \alpha \, \ln \frac{|v_{01}|}{|v_{\rm th}|} \;. \tag{25}$$

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 represents that Pareto indices for $|v_{01}|$ and $|v_{12}|$ are approximately the same value in each figure. This fact and the extended-Gibrat's law (25) suggest that there is a detailed balance under exchange $|v_{01}| \leftrightarrow |v_{12}|$ in each case.[4] 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:

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

In the temporal sales change data, the detailed balance (26) and the 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 one:

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

As the same manner in profits data, we confirm this in Fig. 18. The parameters are estimated as follows: $\alpha \sim 0$ for $|v| > |v_{\rm th}|$, $\alpha \neq 0$ for $|v_{\rm min}| < |v| < |v_{\rm th}|$, $t_+(|v_{\rm th}|) - t_-(|v_{\rm th}|) \sim 1$, $|v_{\rm th}| \sim 10^{4+0.5(5-1)} = 10^6$ thousand (=1 billion) yen and $x_{\rm min} \sim 10^{4+0.5(2-1)} = 10^{4.5}$ thousand

[4] If Pareto indices vary, there is thought to be a detailed quasi-balance (Ishikawa 2006b, 2007b).

(=10 million) yen. In each case, about 5∼10% data are included in the large scale region and about 65∼70% data exist within the middle scale one.

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

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 one. In the analyses, we employ 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 which is observed in 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 property is mutually 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 is also confirmed. This leads the extended-Gibrat's law. At the same time, Pareto indices are almost the same value in the large scale regions of two successive temporal change data. The detailed balance and the extended-Gibrat's law lead the 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 conceivable to derive the followings analytically or numerically (Tomoyose et al. 2008). (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). (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 be clear. In the firm size growth rate distributions with no wide tail (Fig. 1), the probability of the positive growth decreases and the probability of the negative growth increases symmetrically as 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 the positive and negative growth decreases simultaneously as the classification of x increases.

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

Acknowledgments

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.

Appendix

We approximately confirm 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 compare 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 $P(x_1, x_2 \in 4 \times [10^{4+0.4(n-1)}, 10^{4+0.4n}))$ $(n = 1, 2, \cdots, 10)$ 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 is not rejected in 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 compare 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 $P(x_1, x_2 \in 2 \times [10^{5+0.4(n-1)}, 10^{5+0.4n}))$ $(n = 1, 2, \cdots, 10)$. Each p value is shown in Fig 20. For the scatter plot in positive-profits database III (Fig. 5), we compare the distribution sample for $P(x_1 \in [10^{3+0.25(n-1)}, 10^{3+0.25n}), x_2)$ with another sample for $P(x_1, x_2 \in [10^{3+0.25(n-1)}, 10^{3+0.25n}))$ $(n = 1, 2, \cdots, 20)$. 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 take into account the trends in the average growth to test the detailed balance truthfully. However, there are many same data at round figures in the database originally. If we subtract the trends from data, the values are displaced. As a result, p values are underestimated by the displacements. In the databases, the effect of the trends in the average growth is not too large, so we can confirm the detailed balance approximately without subtracting the trends.

From this reason, we do not subtract the trends from the scatter plots in the temporal sales changes (Fig. 17). We compare the distribution sample for $P(|v_{01}| \in [10^{4+0.5(n-1)}, 10^{4+0.5n}), |v_{12}|)$ with another sample for $P(|v_{01}|, |v_{12}| \in [10^{4+0.5(n-1)}, 10^{4+0.5n}))$ $(n = 1, 2, \cdots, 10)$ 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.

References

[1] 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.

[2] Aoyama, H. (2004). Ninth Annual Workshop on Economic Heterogeneous Interacting Agents (WEHIA 2004).

[3] Aoyama, H., Fujiwara, Y., and Souma, W. (2004). The Physical Society of Japan 2004 Autumn Meeting.

[4] 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.

[5] 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.

[6] Diamond Inc., http://www.diamond.co.jp/index.shtml.

[7] Fujiwara, Y., Souma, W., Aoyama, H., Kaizoji, T., and Aoki, M. (2003). Growth and Fluctuations of Personal Income. *Physica A321*: 598–604.

[8] 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 A335*: 197–216.

[9] Gabaix, X. (2005). The Granular Origins of Aggregate Fluctuations. MIT and NBER Working Paper.

[10] Gibrat, R. (1932). *Les inegalites economiques.* Paris: Sirey.

[11] Ishikawa, A. (2006). Derivation of the distribution from extended Gibrat's law. *Physica A367*: 425–434.

[12] Ishikawa, A. (2006). Annual change of Pareto index dynamically deduced from the law of detailed quasi-balance. *Physica A371*: 525–535.

[13] Ishikawa, A. (2007). The uniqueness of firm size distribution function from tent-shaped growth rate distribution. *Physica A383/1*: 79–84.

[14] Ishikawa, A. (2007). Quasistatically varying log-normal distribution in the middle scale region of Japanese land prices. arXiv:0710.1893. To appear in *Prog. Theor. Phys. Suppl.*

[15] 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.

[16] Montrll, 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.

[17] Okuyama, K., Takayasu,M., and Takayasu, H. (1999). Zip's law in income distribution of companies. *Physica A269*: 125–131.

[18] Pareto, V. (1897). *Cours d'Economique Politique*. London: Macmillan.

[19] TOKYO SHOKO RESEARCH, LTD., http://www.tsr-net.co.jp/.p

[20] 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.*

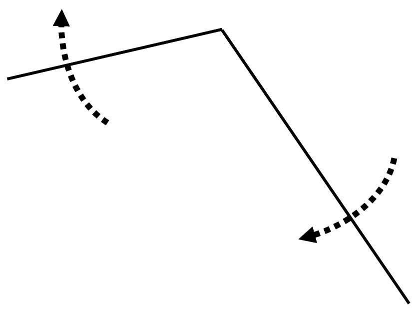

Figure 1: The growth rate distribution of profits or income of firms. The horizontal axis is the logarithm of the growth rate and the vertical axis is the logarithm of its pdf.

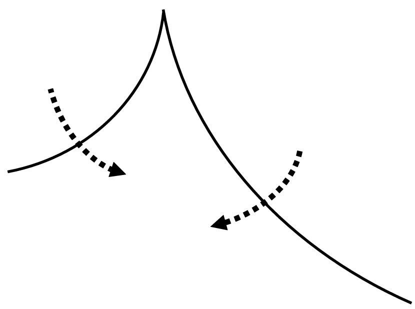

Figure 2: The growth rate distribution of assets and sales of firms, the number of employees or personal income.

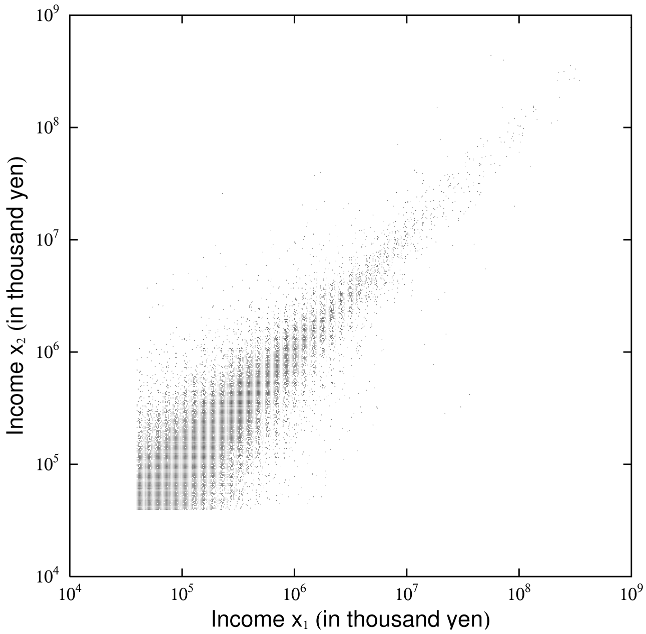

Figure 3: The scatter plot of firms in the database I, the income of which in 2003 (x_0), 2004 (x_1) and 2005 (x_2) exceeded 4×10^4 thousand yen: $x_0 > 4 \times 10^4$ and $x_1 > 4 \times 10^4$ and $x_2 > 4 \times 10^4$. The number of firms is "40,829".

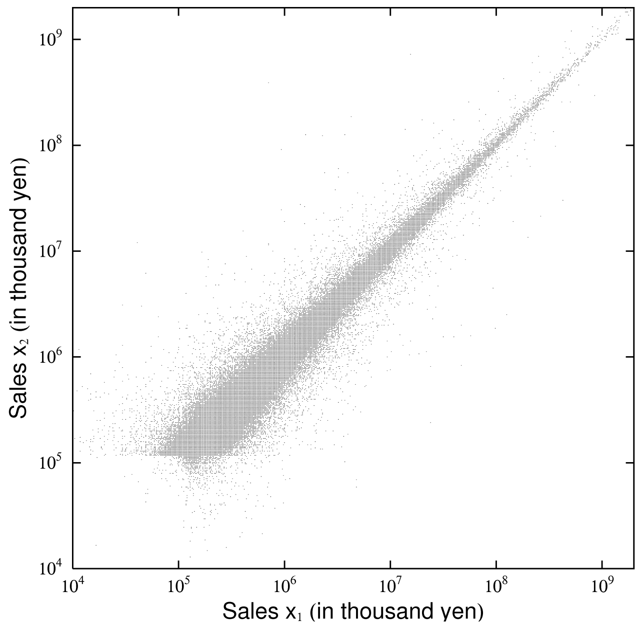

Figure 4: The scatter plot of firms in the database II, the sales of which in 2003 (x_0), 2004 (x_1) and 2005 (x_2) exceeded 0 yen: $x_0 > 0$ and $x_1 > 0$ and $x_2 > 0$. The number of firms is "406,385".

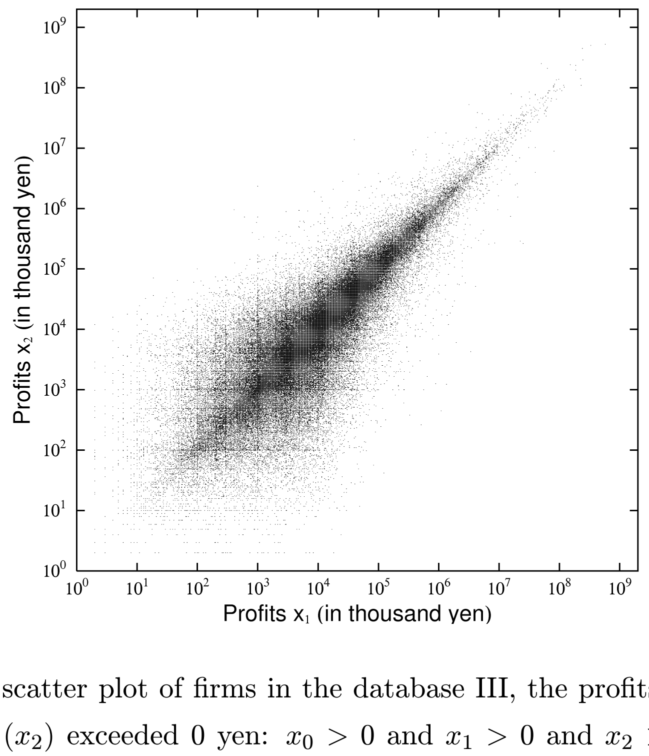

Figure 5: The scatter plot of firms in the database III, the profits of which in 2003 (x_0), 2004 (x_1) and 2005 (x_2) exceeded 0 yen: $x_0 > 0$ and $x_1 > 0$ and $x_2 > 0$. The number of firms is "177,492".

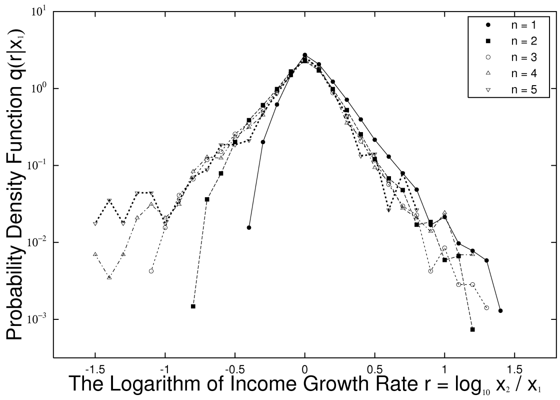

Figure 6: Conditional pdfs $q(r|x_1)$ of the log income growth rate $r = \log_{10} x_2/x_1$ from 2004 to 2005. The data points are classified into five bins of the initial income with equal magnitude in logarithmic scale, $x_1 \in 4 \times [10^{4+0.4(n-1)}, 10^{4+0.4n}]$ $(n = 1, 2, \cdots, 5)$ thousand 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.

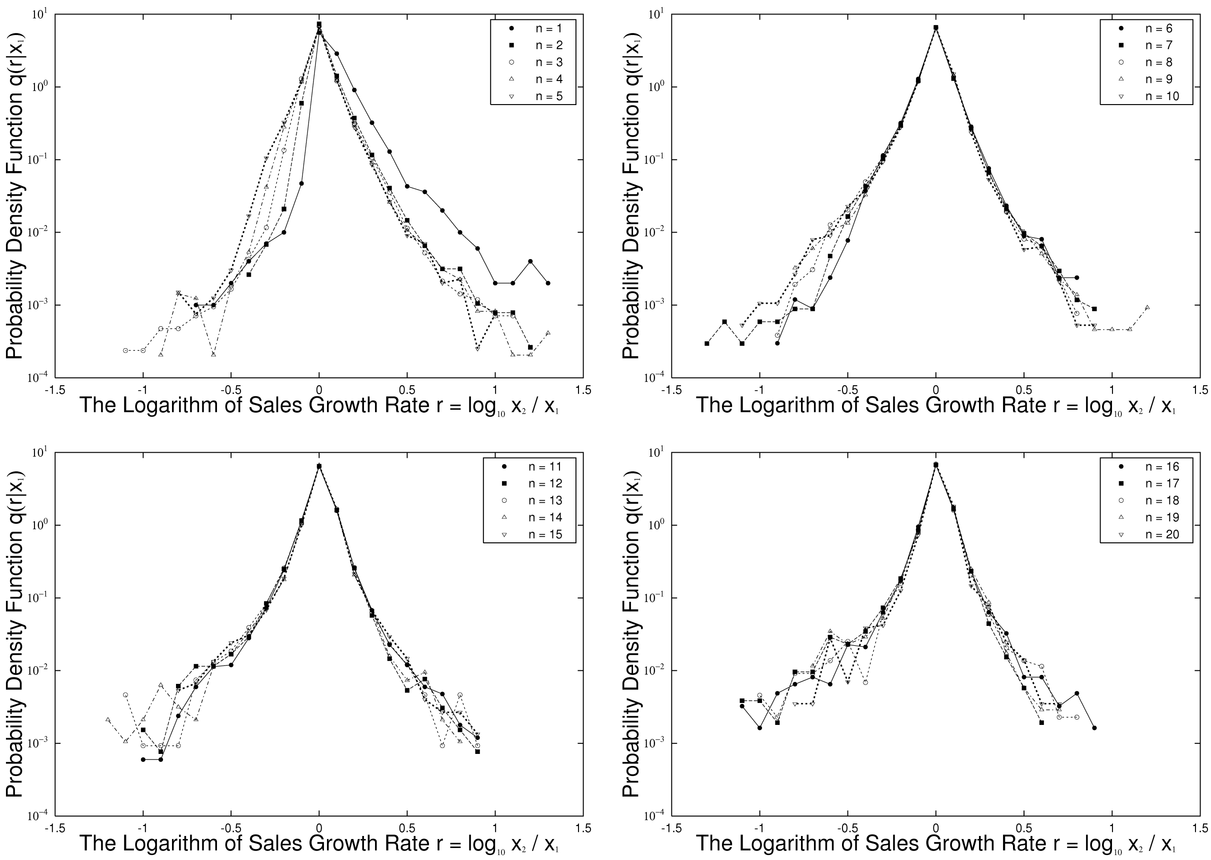

Figure 7: Conditional pdfs $q(r|x_1)$ of the log sales growth rate $r = \log_{10} x_2/x_1$ from 2004 to 2005. The data points are classified into twenty bins of the initial sales with equal magnitude in logarithmic scale, $x_1 \in [10^{5+0.2(n-1)}, 10^{5+0.2n}]$ $(n = 1, 2, \cdots, 20)$ thousand yen.

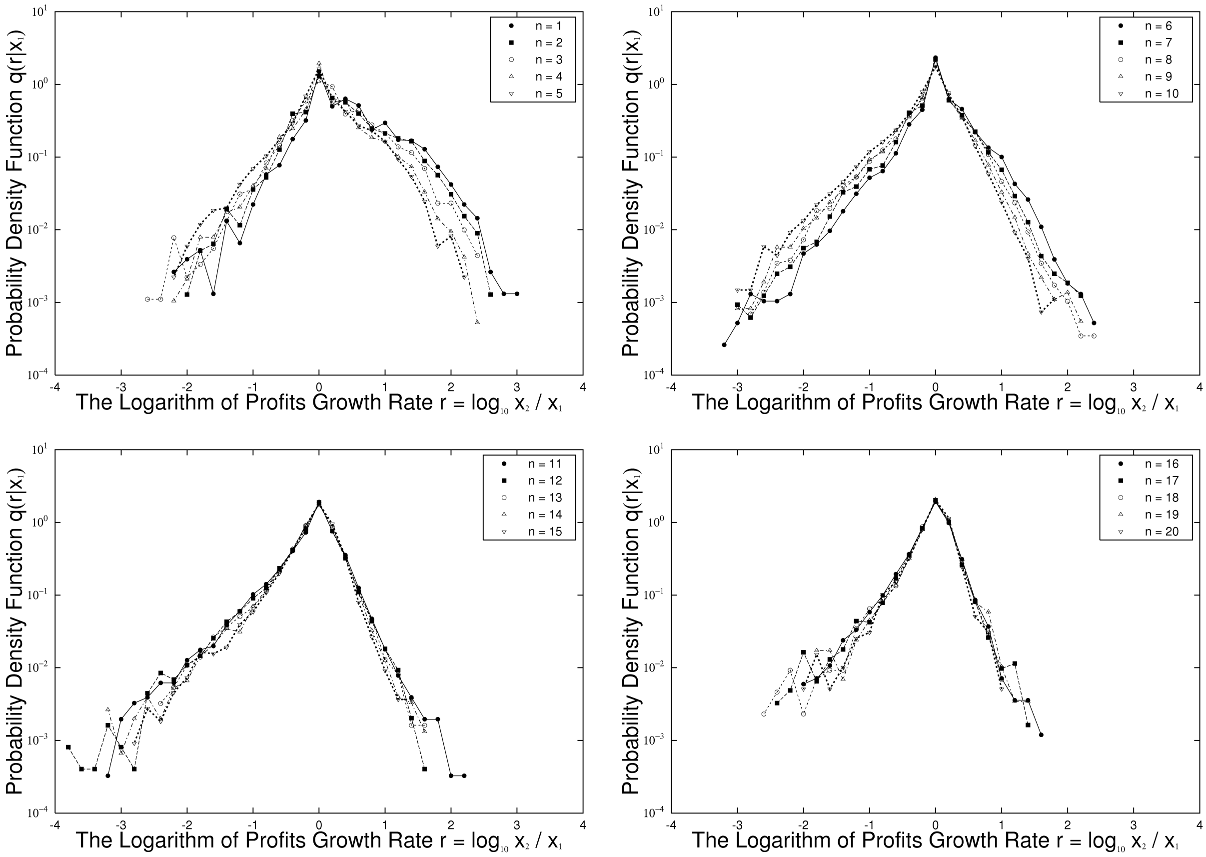

Figure 8: Conditional pdfs $q(r|x_1)$ of the log profits growth rate $r = \log_{10} x_2/x_1$ from 2004 to 2005. The data points are classified into twenty bins of the initial profits with equal magnitude in logarithmic scale, $x_1 \in [10^{2+0.2(n-1)}, 10^{2+0.2n}]$ $(n = 1, 2, \cdots, 20)$ thousand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

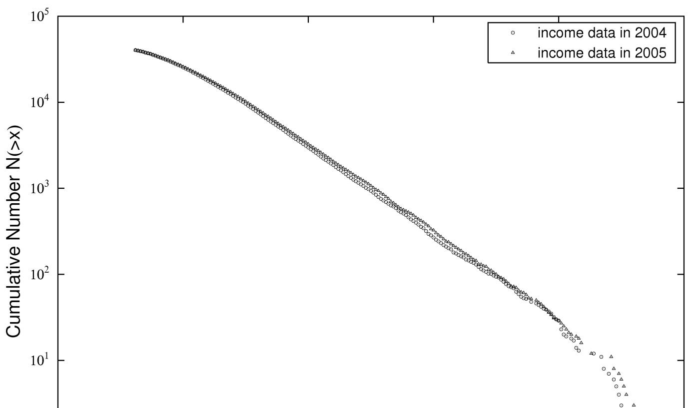

Figure 9: Cumulative number plots of income in the database I, the income of which in 2003 (x_0), 2004 (x_1) and 2005 (x_2) exceeded 4×10^4 thousand yen: $x_0 > 4 \times 10^4$ and $x_1 > 4 \times 10^4$ and $x_2 > 4 \times 10^4$. In the large scale region over about 10^5 thousand (=100 million) yen, Pareto's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

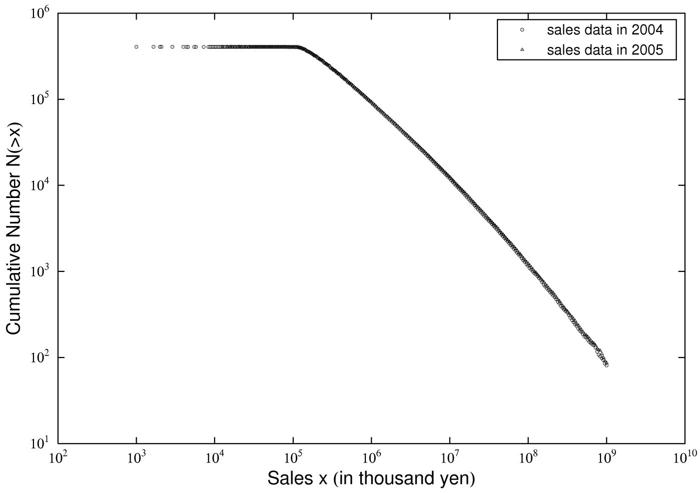

Figure 10: Cumulative number plots of sales in the database II, the sales of which in 2003 (x_0), 2004 (x_1) and 2005 (x_2) exceeded 0 yen: $x_0 > 0$ and $x_1 > 0$ and $x_2 > 0$. In the large scale region over about 2×10^5 thousand (=200 million) yen, Pareto's law is observed. Each Pareto index is estimated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

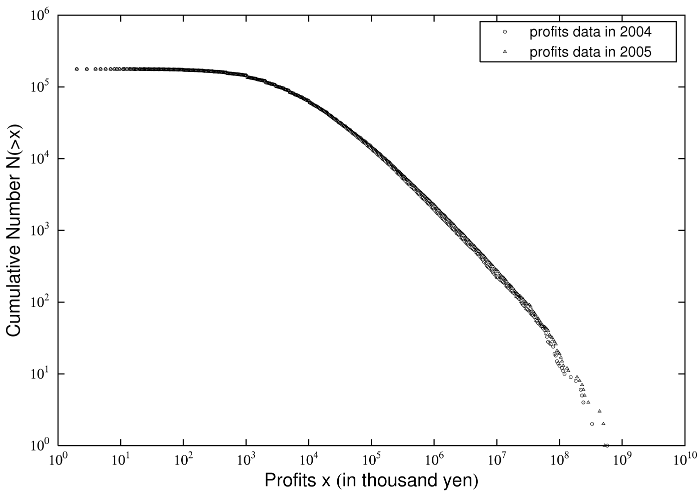

Figure 11: Cumulative number plots of positive-profits in the database III, the profits of which in 2003 (x_0), 2004 (x_1) and 2005 (x_2) exceeded 0 yen: $x_0 > 0$ and $x_1 > 0$ and $x_2 > 0$. In the large scale region over about 10^5 thousand (=100 million) yen, Pareto's law is observed. Each Pareto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

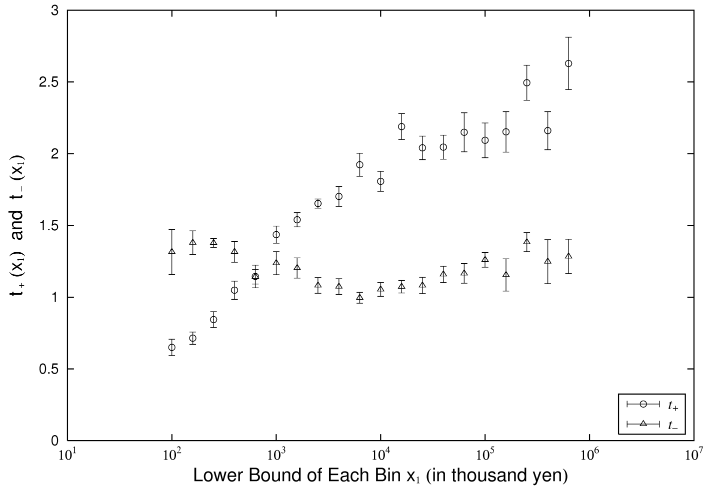

Figure 12: The relation between the lower bound of each bin $x_1 \in [10^{2+0.2(n-1)}, 10^{2+0.2n}]$ and $t_{\pm}(x_1)$. From the left, each data point represents $n = 1, 2, \cdots, 20$. The values are measured by the least square method in the region $0 \leq |r| \leq 2$ in Fig. 8.

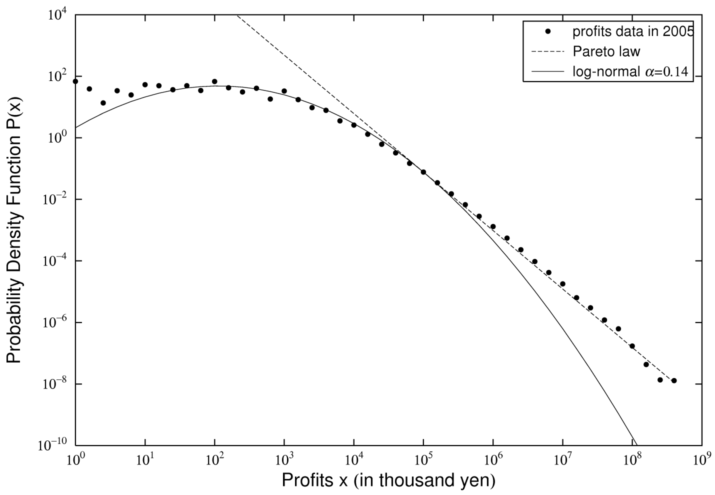

Figure 13: The pdf of positive-profits in the database III.

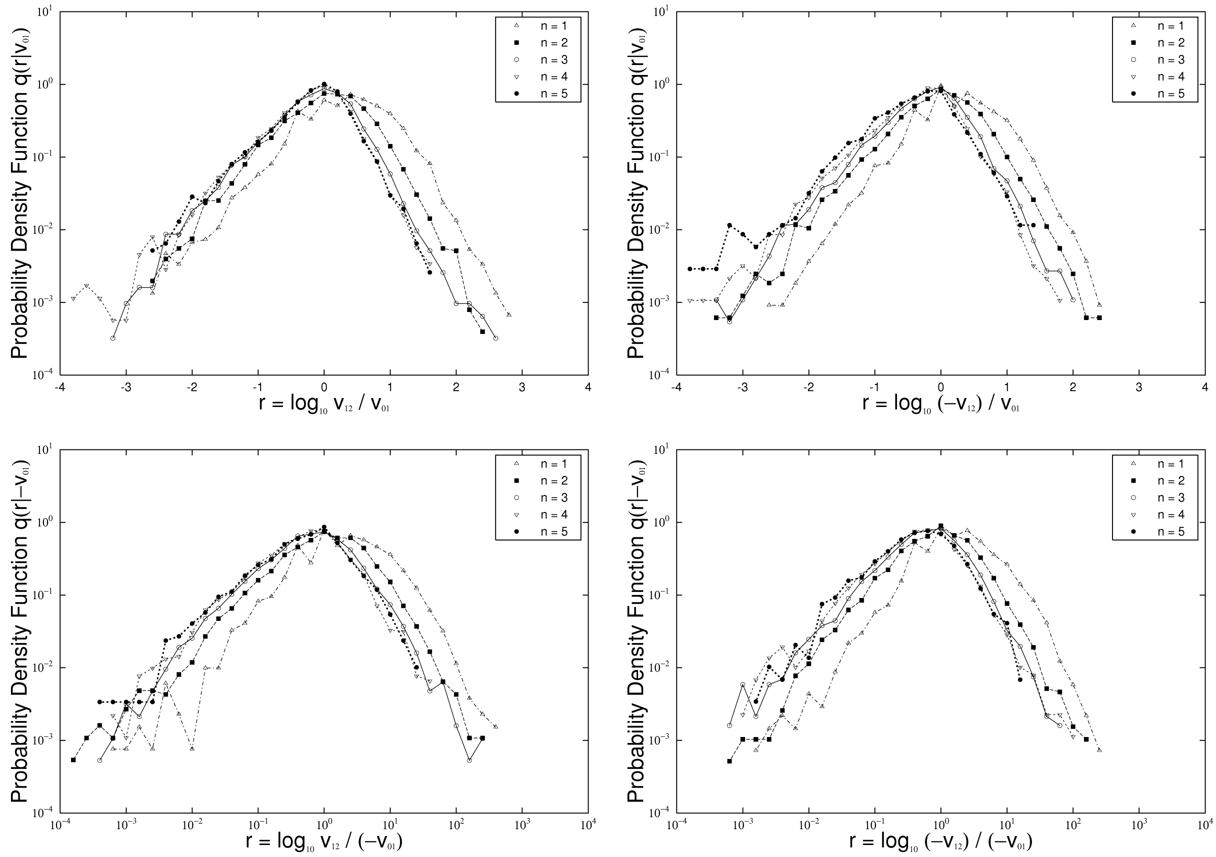

Figure 14: Conditional pdfs $q(r||v_{01}|)$ in the log growth rate of the temporal sales change $r = \log_{10}|v_{12}/v_{01}|$ for cases $(v_{01} > 0, v_{12} > 0)$, $(v_{01} > 0, v_{12} < 0)$, $(v_{01} < 0, v_{12} > 0)$ and $(v_{01} < 0, v_{12} < 0)$. The number of data is "54,181", "32,959", "35,218" 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, $|v_{01}| \in [10^{4+0.5(n-1)}, 10^{4+0.5n}]$ $(n = 1, 2, \cdots, 5)$ 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 $x_2 > 4 \times 10^5$.

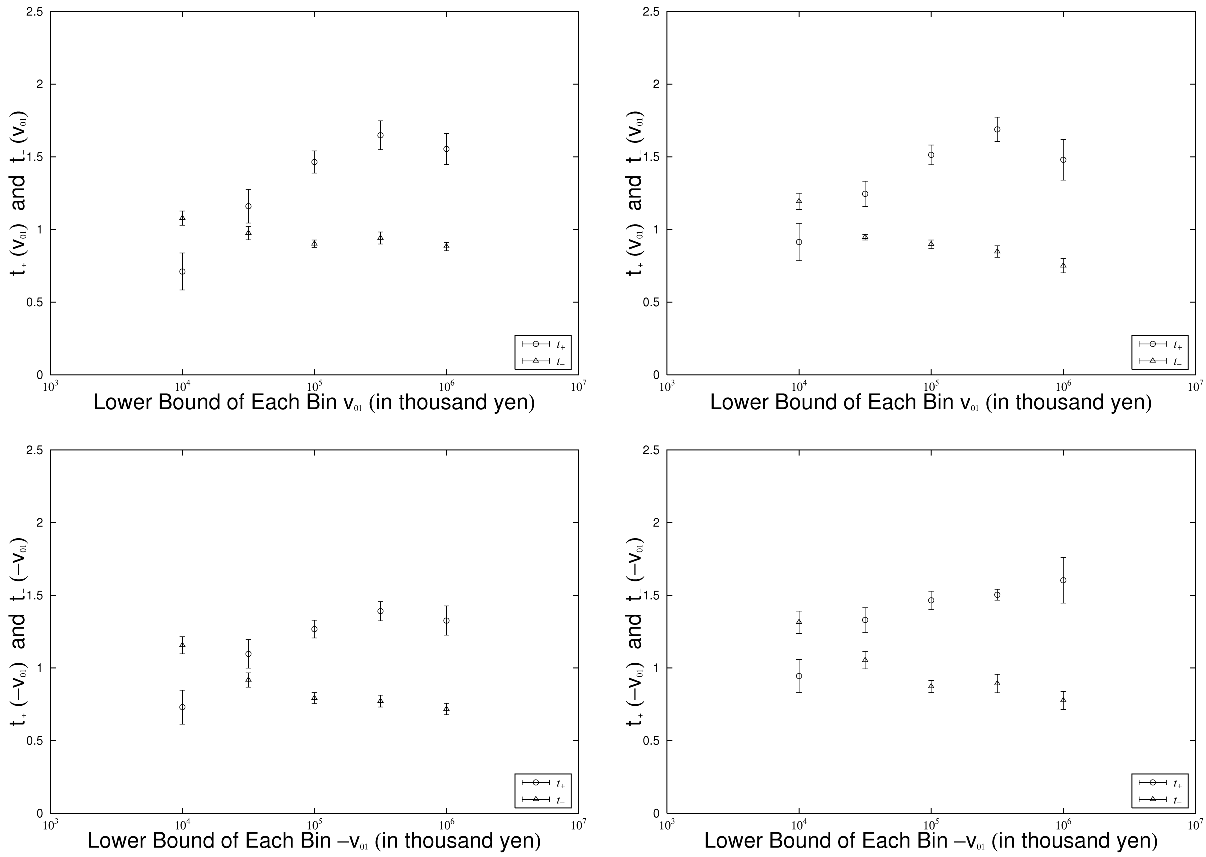

Figure 15: The relation between the lower bound of each bin $|v_{01}| \in [10^{4+0.5(n-1)}, 10^{4+0.5n}]$ and $t_{\pm}(|v_{01}|)$. In each figure, from the left each data point represents $n = 1, 2, \cdots, 5$. The values are measured by the least square method in the region $0 \le |r| \le 2$ in Fig.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

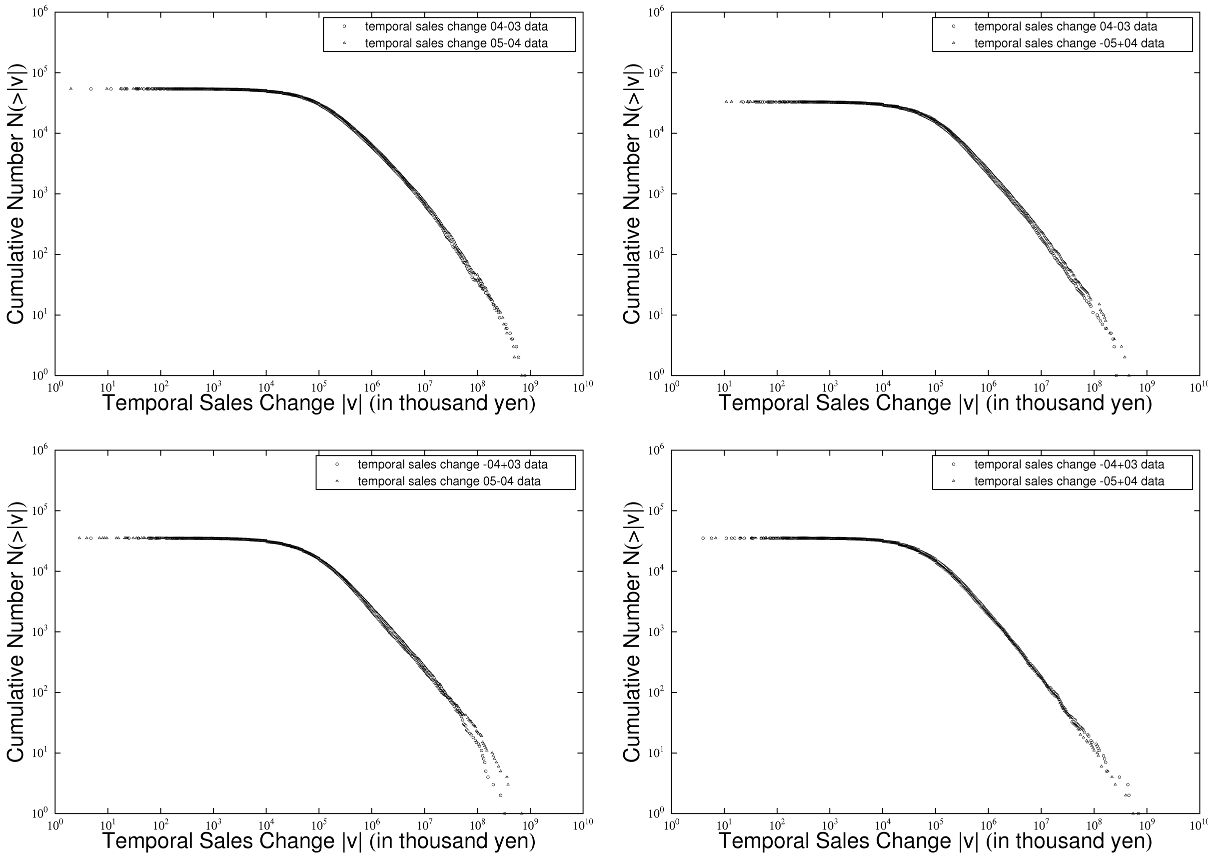

Figure 16: Cumulative number plots in the temporal sales changes for cases $(v_{01} > 0, v_{12} > 0)$, $(v_{01} > 0, v_{12} < 0)$, $(v_{01} < 0, v_{12} > 0)$ and $(v_{01} < 0, v_{12} < 0)$.

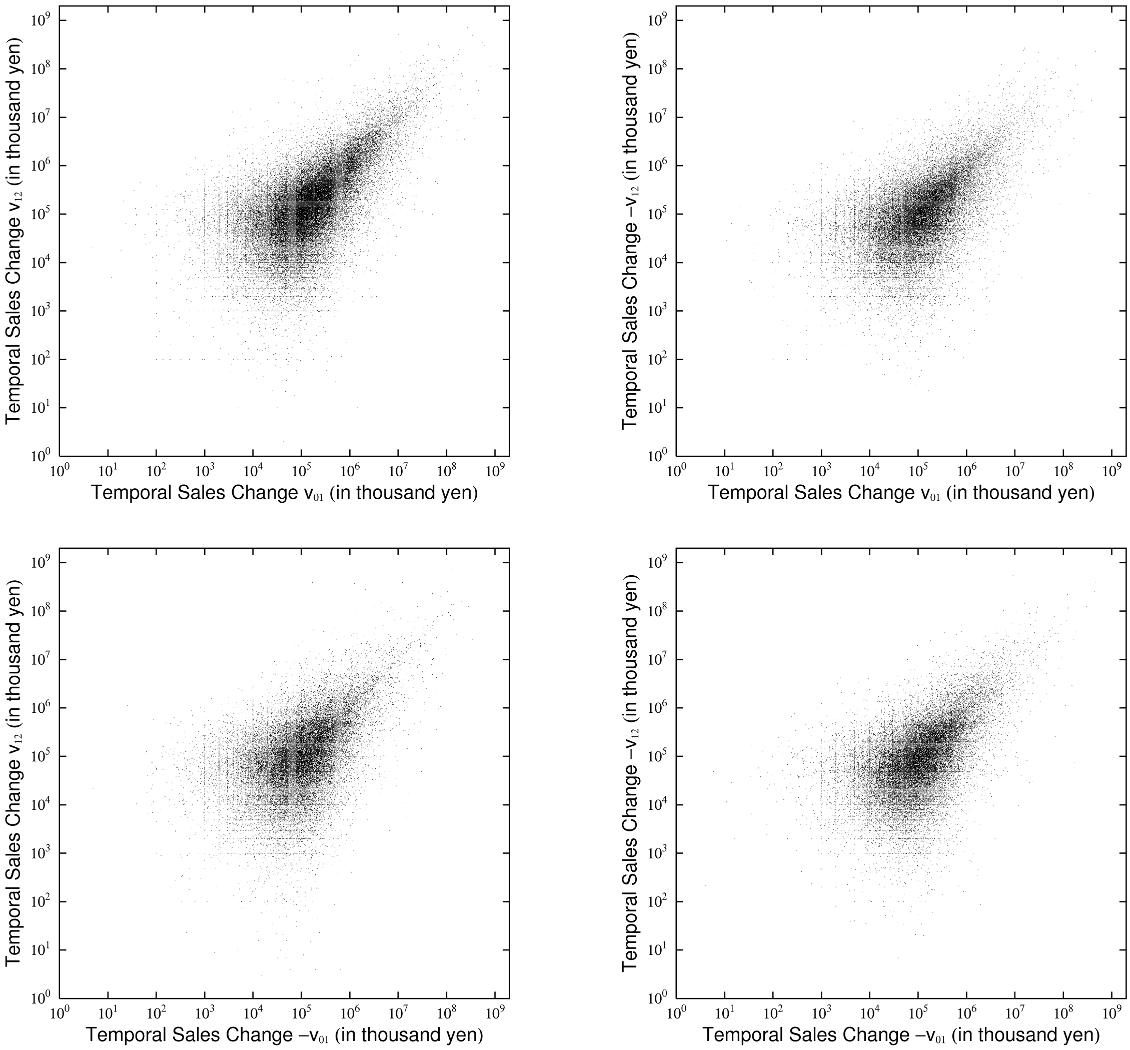

Figure 17: Scatter plots in the temporal sales changes for cases $(v_{01} > 0, v_{12} > 0)$, $(v_{01} > 0, v_{12} < 0)$, $(v_{01} < 0, v_{12} > 0)$ and $(v_{01} < 0, v_{12} < 0)$.

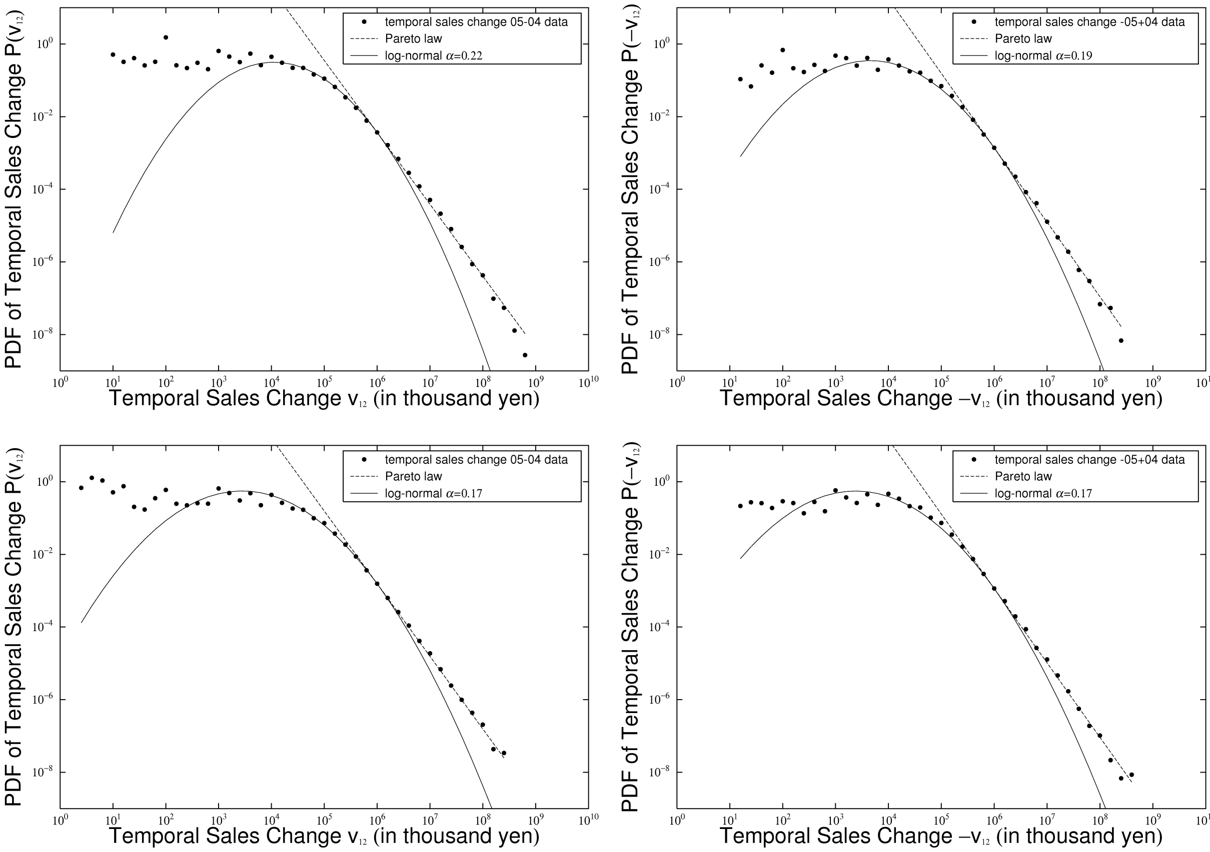

Figure 18: The pdf of the temporal sales change data.

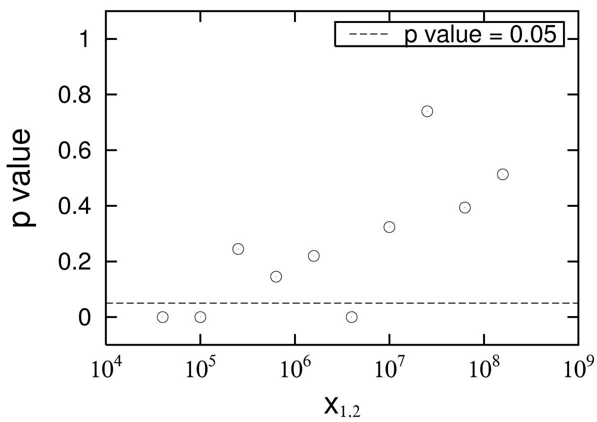

Figure 19: Each p value of the one-dimensional K–S test for the scatter plot of high-income data points (Fig. 3).

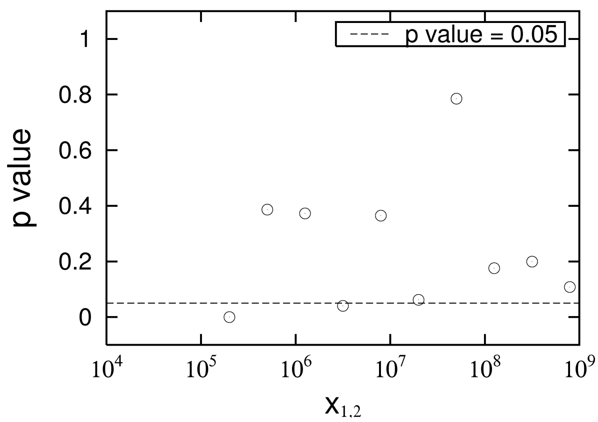

Figure 20: Each p value of the one-dimensional K–S test for the scatter plot of high-sales data points (Fig. 4).

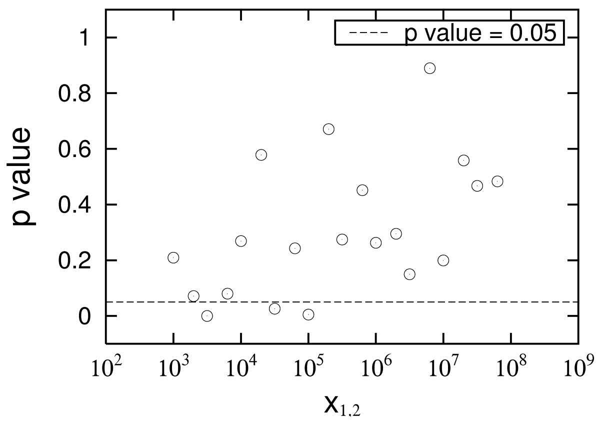

Figure 21: Each p value of the one-dimensional K–S test for the scatter plot of positive-profits data points (Fig. 5).

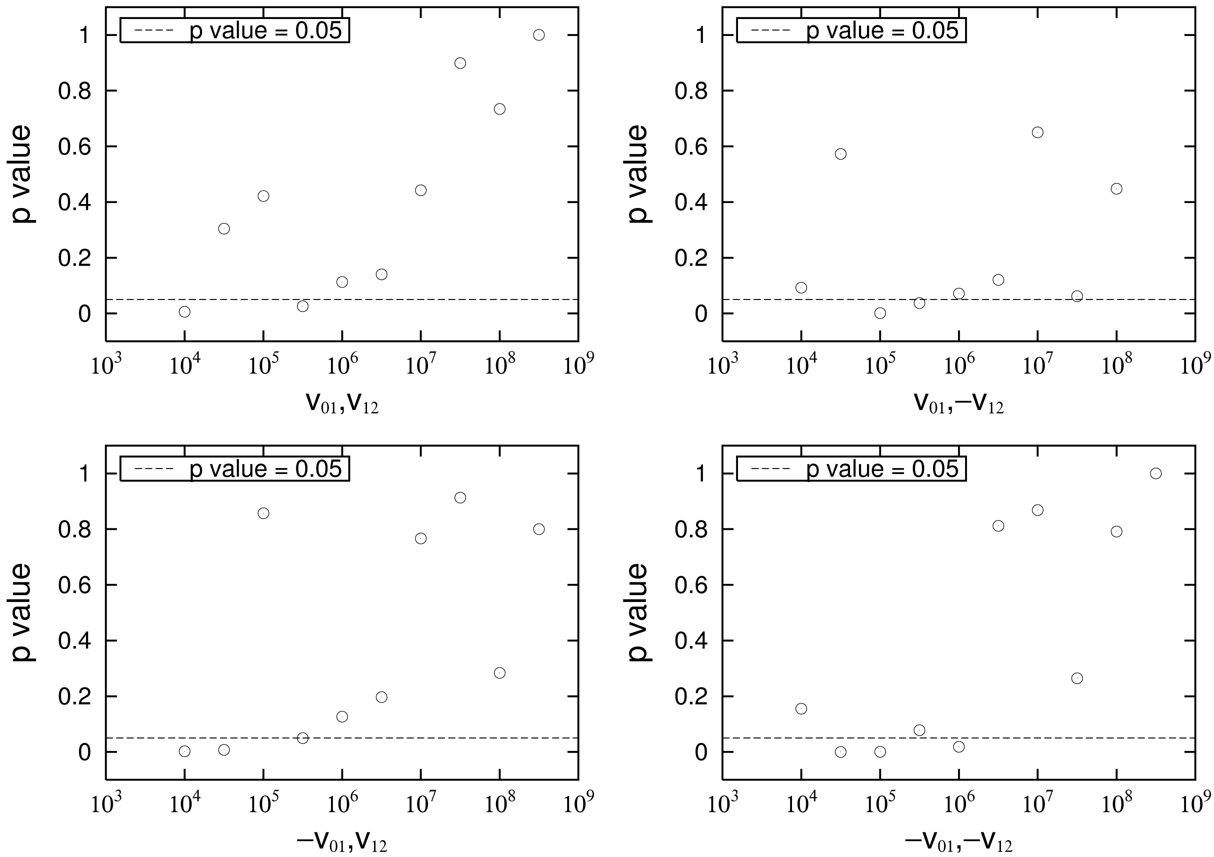

Figure 22: Each p value of the one-dimensional K–S test for the scatter plots in the temporal sales changes (Fig. 17).