JSPS Grants-in-Aid for Scientific Research (S)

Understanding Persistent Deflation in Japan

Working Paper Series

No. 036

January 2014

Analytical Derivation of Power Laws in Firm Size Variables from Gibrat's Law and Quasi-inversion Symmetry: A Geomorphological Approach

Atushi Ishikawa

Shouji Fujimoto

Takayuki Mizuno

Tsutomu Watanabe

UTokyo Price Project
702 Faculty of Economics, The University of Tokyo,
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
Tel: +81-3-5841-5595
E-mail: watlab@e.u-tokyo.ac.jp
http://www.price.e.u-tokyo.ac.jp/english/

Analytical Derivation of Power Laws in Firm Size Variables from Gibrat's Law and Quasi-inversion Symmetry: A Geomorphological Approach

Atushi Ishikawa[1 *], Shouji Fujimoto[1], Takayuki Mizuno[2], and Tsutomu Watanabe[3]

[1]*Department of Informatics and Business, Faculty of Business Administration and Information Science, Kanazawa Gakuin University, 10 Sue, Kanazawa, Ishikawa 920-1392, Japan*

[2]*National Institute of Informatics, Tokyo 101-8430, Japan*

Department of Informatics, Graduate University for Advanced Studies, Kanagawa 240-0193, Japan

[3]*Graduate School of Economics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan*

We start from Gibrat's law and quasi-inversion symmetry for three firm size variables (i.e., tangible fixed assets K, number of employees L, and sales Y) and derive a partial differential equation to be satisfied by the joint probability density function of K and L. We then transform K and L, which are correlated, into two independent variables by applying surface openness used in geomorphology and provide an analytical solution to the partial differential equation. Using worldwide data on the firm size variables for companies, we confirm that the estimates on the power-law exponents of K, L, and Y satisfy a relationship implied by the theory.

KEYWORDS: econophysics, power law, Gibrat's law

1. Introduction

In econophysics, it is well-known that the cumulative distribution functions (CDFs) of capital K, labor L, and production Y of firms obey power laws in large scales that exceed certain size thresholds, which are given by K_0, L_0, and Y_0:

$$P_>(K) \propto K^{-\mu_K} \quad \text{for} \quad K > K_0 \,, \tag{1}$$

$$P_>(L) \propto L^{-\mu_L} \quad \text{for} \quad L > L_0 \,, \tag{2}$$

$$P_>(Y) \propto Y^{-\mu_Y} \quad \text{for} \quad Y > Y_0 \,. \tag{3}$$

*E-mail: ishikawa@kanazawa-gu.ac.jp

These are called Pareto's law,[1] and positive power-law indices μ_K, μ_L, and μ_Y are called Pareto's indices. In condensed matter physics, various power laws are observed in physical quantities near the critical points of phase transitions. It is fascinating that power laws are also observed in social quantities.[2–25] On the one hand, the renormalization group approach has been employed to confirm power laws observed in physical quantities as critical phenomena of phase transitions.[26] Various models have also been proposed to explain power laws observed in social phenomena (see, e.g., Refs.[22–25]).

Under this circumstance, without any model, Fujiwara et al. showed that power laws arise as a result of Gibrat's law[27,28] and inversion symmetry, both of which are observed between variables in some year and the same variables the next year: (K_T, K_{T+1}), (L_T, L_{T+1}), (Y_T, Y_{T+1}).[29,30] By extending this discussion, we reported that power laws with different indices are similarly obtained from Gibrat's law and quasi-inversion symmetry observed between variables in some year and different variables the same year: (K, Y), (L, Y), (K, L).[31–33]

By extending the discussion on two variables to three variables, we derived the following Cobb-Douglas production function[34] from Gibrat's law and quasi-inversion symmetry for three variables:

$$Y = AK^{\alpha}L^{\beta} \ . \tag{4}$$

In economics, the product of firm Y is determined by its capital K and its labor L. This relation is called the production function, and the Cobb-Douglas form (4) is one of the most useful functions. Here, α and β are positive parameters that are constant in the analyzed firms. They are called the output elasticities of capital and labor, respectively. A in Eq. (4) is called the total factor productivity and represents a firm's technology that cannot be estimated by capital K or labor L. The Cobb-Douglas production function has been extensively used in various areas of economics.

Here, quasi-inverse symmetry and Gibrat's law are represented using joint probability density function (PDF) $P_{KLY}(K, L, Y)$ and conditional PDF $Q(R|K, L)$ as follows:[32,33]

$$\begin{aligned} &P_{KLY}(K, L, Y)dKdLdY \\ &= P_{KLY}\left(\left(\frac{Y}{aL^{\beta}}\right)^{1/\alpha}, \left(\frac{Y}{aK^{\alpha}}\right)^{1/\beta}, aK^{\alpha}L^{\beta}\right) \, d\left(\left(\frac{Y}{aL^{\beta}}\right)^{1/\alpha}\right) d\left(\left(\frac{Y}{aK^{\alpha}}\right)^{1/\beta}\right) d\left(aK^{\alpha}L^{\beta}\right) \ , \end{aligned} \tag{5}$$

$$Q(R|K, L) = Q(R) \ , \tag{6}$$

where R is the rate of three variables $R = Y/aK^{\alpha}L^{\beta}$ and a is a constant parameter. Quasi-inverse symmetry is a invariance of the system under the exchange of variables $aK^{\alpha}L^{\beta} \leftrightarrow Y$.

The point of Eq. (5) is that the functional form of right-hand side P_{KLY} is not different from the functional form of left-hand side. At the same time, infinitesimal volume elements, which were not taken into account in Ref.,[33] must be associated with the PDFs.

In previous work, we clarified that Gibrat's law and quasi-inversion symmetry in three variables (K, L, Y) is the theoretical explanation why (K, L, Y) data are well approximated by the Cobb-Douglas production function (4). In this discussion, one issue is the identification of the functional form of joint PDF $P_{KL}(K, L)$, which leads to power laws (1)–(3). To confirm the analytical result using empirical data, a transformation must be found that absolutely eliminates the correlation between K and L. In previous works, the transformation was approximately estimated using the least square method applied to the mean values of the data that were divided into logarithmically equal sized bins.[35] However, the resulting power-law indices did not follow the analytical results.

In this paper, using surface openness,[36–38] which is used in landform classification, we determined the accurate transformation from a geomorphologic point of view and found that the estimated power-law indices numerically do follow the theoretical relations among power-law indices that were analytically obtained from $P_{KL}(K, L)$.

The rest of this paper is organized as follows. In Section 2, we determine the functional form of joint PDF $P_{KL}(K, L)$ from quasi-inversion symmetry and Gibrat's law in three variables (K, L, Y), both of which will be confirmed in Section 3. At the same time, power laws (1)–(3) are directly derived from $P_{KL}(K, L)$ and relations among the power-law indices are provided. In Section 3, we observe power laws, quasi-inversion symmetry, and Gibrat's law of (K, L, Y) using ORBIS, a commercially available firm-level worldwide data set compiled by Bureau van Dijk.[39] Next, we confirm our analytical results in Section 2. What is important is the accurate determination of the transformation that completely eliminates the correlation between K and L with surface openness in landform classification. We numerically show that the new transformed variables are not only uncorrelated but also independent in the power-law region. Finally, in Section 4, we summarize our findings and discuss some additional issues.

2. Derivation of Power Laws from Quasi-inverse Symmetry and Gibrat's Law in Three Variables

In this section, we review previous works[31–33] and analytically discuss the identification of the functional form of $P_{KL}(K, L)$. As a result, the power laws of three valuables (1)-(3) are directly derived from quasi-inverse symmetry and Gibrat's law in three variables through

$P_{KL}(K, L)$. At the same time, the three indices, μ_K, μ_L, and μ_Y, are related to each other.

Quasi-inverse symmetry is also represented by using joint PDF $P_{KLR}(K, L, R)$ as follows:

$$P_{KLR}(K, L, R) dK dL dR = P_{KLR}\left(R^{1/\alpha}K, R^{1/\beta}L, R^{-1}\right)\ d\left(R^{1/\alpha}K\right) d\left(R^{1/\beta}L\right) d\left(R^{-1}\right)\ . \tag{7}$$

We can rewrite Eq. (7) as

$$P_{KLR}(K, L, R) = R^{1/\alpha+1/\beta-2} P_{KLR}\left(R^{1/\alpha}K, R^{1/\beta}L,\ R^{-1}\right)\ . \tag{8}$$

From the definition of conditional PDF $Q(R|K, L) = P_{KLR}(K, L, R)/P_{KL}(K, L)$, Eq. (8) is rewritten:

$$\frac{P_{KL}(K, L)}{P_{KL}\left(R^{1/\alpha}K, R^{1/\beta}L\right)} = R^{1/\alpha+1/\beta-2} \frac{Q\left(R^{-1} \mid R^{1/\alpha}K,\ R^{1/\beta}L\right)}{Q(R \mid K, L)} = R^{1/\alpha+1/\beta-2} \frac{Q(R^{-1})}{Q(R)}\ . \tag{9}$$

Here, Gibrat's law (6) is used. The last term of Eq. (9) is only a function of R and is denoted by $G(R)$. Therefore, we obtain

$$P_{KL}(K, L) = G(R) P_{KL}(R^{1/\alpha}K,\ R^{1/\beta}L)\ . \tag{10}$$

The definition of $G(R)$ is different from the definition in Ref.[33] Because infinitesimal volume elements in Eq. (7) or (5) were not taken into account in Ref.[33]

By expanding the right-hand side of Eq. (10) for R around 1 (namely, $R = 1 + \epsilon$ where $\epsilon \ll 1$) we obtain a differential equation:

$$G'(1) P_{KL}(K, L) + \frac{K}{\alpha}\frac{\partial}{\partial K} P_{KL}(K, L) + \frac{L}{\beta}\frac{\partial}{\partial L} P_{KL}(K, L) = 0\ . \tag{11}$$

Here, $G'(\cdot)$ is a derivative of $G(\cdot)$ with respect to R . The incorrect description of $G(R)$ in Ref.[33] has been absorbed in the difference of the difinition of $G(R)$, then the resulting differential equation is the same form in Ref.[33]

We show in the next section that variables K strongly correlate with L, and therefore, they are not independent. To eliminate the correlation, we transform variables (K, L) to (Z_1, Z_2):

$$\log_{10} Z_1 = \log_{10} k + \log_{10} l\ , \quad \log_{10} Z_2 = -\log_{10} k + \log_{10} l\ , \tag{12}$$

with

$$\log_{10} k = \frac{\log_{10} K - m_K}{\sigma_K}\ , \quad \log_{10} l = \frac{\log_{10} L - m_L}{\sigma_L}\ . \tag{13}$$

Next we introduce parameters m_K, m_L, σ_K, and σ_L. We divide the power-law region of K and L into logarithmically equal sized cells and consider the topographic map by regarding the logarithm of the number of firms in a cell as its altitude. Parameters m_K and m_L are the mean values of the central coordinates of the cells that constitute the landform's ridge. Parameters σ_K and σ_L are $\sqrt{2}$ times of the standard deviations. In the next section, we describe them in detail using empirical data.

No correlation between Z_1 and Z_2 does not necessarily mean that they are independent variables. However, in the power-law region of K and L, Z_1 and Z_2 are not only uncorrelated but also independent. This is numerically confirmed using empirical data. We will also describe them in detail in the next section.

The independence of the variables Z_1 and Z_2, which will be numerically confirmed in the next section, can be analytically explained as follows. Two variables (K, L), the coordinate system of which has quasi-inverse symmetry $a_{KL}K^{\theta_{KL}} \leftrightarrow L$, are transformed into two variables (k, l), the coordinate system of which has inverse symmetry $k \leftrightarrow l$. Note that the definitions of k and l are given in Eq. (13), and that a_{KL} and θ_{KL} are constant parameters. Inverse symmetry is a invariance under the exchange of variables with respect to the line with its slop being $\pi/4$. Since Eq. (12) rotates the coordinate system by $-\pi/4$, the new variables (Z_1, Z_2) must be independent. Also, the system (Z_1, Z_2) has $Z_2 \leftrightarrow 1/Z_2$ symmetry.

By setting

$$\theta_1 = \frac{1}{2}(\sigma_K\alpha + \sigma_L\beta) \ , \tag{14}$$

$$\theta_2 = \frac{1}{2}(-\sigma_K\alpha + \sigma_L\beta) \ , \tag{15}$$

$K^\alpha L^\beta$ is reduced to ${Z_1}^{\theta_1}{Z_2}^{\theta_2}$ and quasi-inverse symmetry is rewritten as $Y \leftrightarrow a'{Z_1}^{\theta_1}{Z_2}^{\theta_2}$. Here, $a'(= 10^{\alpha m_K+\beta m_L}a)$ is a constant parameter. As a result, quasi-inverse symmetry and Gibrat's law are similarly observed in the (Z_1, Z_2, Y) coordinates. Therefore, using a similar discussion, the following partial differential equations are obtained:

$$G_+'(1)\, P_{Z_1Z_2}(Z_1, Z_2) + \frac{Z_1}{\theta_1}\frac{\partial}{\partial Z_1}P_{Z_1Z_2}(Z_1, Z_2) + \frac{Z_2}{\theta_2}\frac{\partial}{\partial Z_2}P_{Z_1Z_2}(Z_1, Z_2) = 0 \ , \tag{16}$$

$$G_-'(1)\, P_{Z_1Z_2}(Z_1, {Z_2}^{-1}) + \frac{Z_1}{\theta_1}\frac{\partial}{\partial Z_1}P_{Z_1Z_2}(Z_1, {Z_2}^{-1}) - \frac{{Z_2}^{-1}}{\theta_2}\frac{\partial}{\partial {Z_2}^{-1}}P_{Z_1Z_2}(Z_1, {Z_2}^{-1}) = 0 \ , \tag{17}$$

where $G_\pm(R) = R^{1/\theta_1 \pm 1/\theta_2 - 2}Q(R^{-1})/Q(R)$ and $R = Y/a'{Z_1}^{\theta_1}{Z_2}^{\theta_2}$. Equation (17) is obtained from $Z_2 \leftrightarrow 1/Z_2$ symmetry.

As mentioned above, Z_1 and Z_2 are independent variables. Therefore, with the variable separation method, solution $P_{Z_1Z_2}(Z_1, Z_2)$ of Eqs. (16) and (17) is uniquely determined to be the product of the power-law functions of Z_1 and Z_2 . This will be numerically verified in the next section. What is important here is that the power-law index of Z_2 for $Z_2 > 1$ is different from the index of Z_2 for $Z_2 < 1$. Therefore, the analytical solution is expressed as follows:

$$P_{Z_1Z_2}(Z_1, Z_2) = C\, {Z_1}^{-\mu_1-1}{Z_2}^{-\mu_2-1} \ \ \text{for} \ \ \log_{10} Z_2 > 0 \ , \tag{18}$$

$$P_{Z_1Z_2}(Z_1, Z_2) = C\ {Z_1}^{-\mu_1-1}{Z_2}^{+\mu_2-1}\ \text{ for }\ \log_{10} Z_2 < 0\ , \tag{19}$$

with the conditions of variable separation $\frac{\mu_1+1}{\theta_1} + \frac{\mu_2\pm1}{\theta_2} = G_{\pm}'(1)$. This solution satisfies $Z_2 \leftrightarrow 1/Z_2$ symmetry, namely, $P_{Z_1Z_2}(Z_1, Z_2)dZ_1dZ_2 = P_{Z_1Z_2}(Z_1, {Z_2}^{-1})dZ_1d({Z_2}^{-1})$. In Ref.,[33) Eq. (18), which is valid for the case of $\log_{10} Z_2 > 0$, was only presented, because $Z_2 \leftrightarrow 1/Z_2$ symmetry was not discussed. In order to consider the case of $\log_{10} Z_2 < 0$, $P_{Z_1Z_2}$ must be written by Eq. (19). This was suggested in Figs. 7 and 8 in Ref.[33)

Note that Eqs. (18) and (19) continue to be a general solution to $P_{Z_1Z_2}(Z_1, Z_2) = G(R)P_{Z_1Z_2}(R^{1/\theta_1}Z_1, R^{1/\theta_2}Z_2)$, even if R deviates substantially from $R = 1$, as long as $Q(R)$ satisfies $Q(R) = R^{-G_{\pm}'(1)+1/\theta_1\pm1/\theta_2-2}Q(R^{-1}) = R^{-\mu_1/\theta_1-\mu_2/\theta_2-2}Q(R^{-1})$. Using the transformation of integration measure $dZ_1dZ_2 =\mid \frac{\partial(Z_1,Z_2)}{\partial(K,L)} \mid dKdL = \frac{2}{\sigma_K\sigma_L}K^{-1}L^{\frac{2}{\sigma_L}-1}dKdL$, $P_{KL}(K, L)$ is expressed as follows:

$$P_{KL}(K, L) = C_+\ K^{\theta_K^+-1}L^{-\theta_L^+-1}\quad \text{for}\quad l > k\ , \tag{20}$$

$$P_{KL}(K, L) = C_-\ K^{-\theta_K^--1}L^{\theta_L^--1}\quad \text{for}\quad l < k\ . \tag{21}$$

Here,

$$\theta_K^+ = \frac{\mu_2 - \mu_1}{\sigma_K}\ ,\quad \theta_L^+ = \frac{\mu_2 + \mu_1}{\sigma_L}\ , \tag{22}$$

$$\theta_K^- = \frac{\mu_2 + \mu_1}{\sigma_K}\ ,\quad \theta_L^- = \frac{\mu_2 - \mu_1}{\sigma_L}\ . \tag{23}$$

Note that $P_{KL}(K, L)$ is also the product of the power-law functions of K and L, but variables K and L are not independent because the power-law indices for $l > k$ are different from the indices for the case that $l < k$.

By integrating $P_{KL}(K, L)$ by L or K , from the leading order terms, power-law functions K and L are obtained:

$$P(K) = \int_0^\infty dLP_{KL}(K, L) \sim \left(\frac{C_+{a_{KL}}^{-\theta_L^+}}{\theta_L^+} + \frac{C_-{a_{KL}}^{\theta_L^-}}{\theta_L^-}\right)K^{-2\frac{\mu_1}{\sigma_K}-1}\ , \tag{24}$$

$$P(L) = \int_0^\infty dKP_{KL}(K, L) \sim \left(\frac{C_+{a_{LK}}^{\theta_K^+}}{\theta_K^+} + \frac{C_-{a_{LK}}^{-\theta_K^-}}{\theta_K^-}\right)L^{-2\frac{\mu_1}{\sigma_L}-1}\ , \tag{25}$$

where $\log_{10} a_{KL} = m_L - m_K\sigma_L/\sigma_K$, $\log_{10} a_{LK} = m_K - m_L\sigma_K/\sigma_L$. By comparing Eq. (1) with Eq. (24) and Eq. (2) with Eq. (25), the relations among power-law indices are found:

$$\mu_K = 2\ \frac{\mu_1}{\sigma_K}\ ,\quad \mu_L = 2\ \frac{\mu_1}{\sigma_L}\ . \tag{26}$$

At the same time, from the result in Refs.[40,41)] , we conclude the following. Under quasi-inverse symmetry $Y \leftrightarrow a'{Z_1}^{\theta_1}{Z_2}^{\theta_2}$, Eqs. (18) and (19) imply that Y obeys power law (3) and index μ_Y is identified as[31,33]

$$\mu_Y = \min\left\{\frac{\mu_1}{\theta_1}, \frac{\mu_2}{\theta_2}\right\} \ . \tag{27}$$

From the data analyses of various countries from 2004 to 2009, which are discussed in the next section, the relations among power-law indices are observed as $\mu_1/\theta_1 < \mu_2/\theta_2$, Eq. (27) is reduced to

$$\mu_Y = \frac{\mu_1}{\theta_1} \ . \tag{28}$$

Furthermore, by writing $A = R\, a$, we can regard the definition of R as the Cobb-Douglas production function (4). In this case, Gibrat's law (6) guarantees that the distribution of total factor productivity A does not depend on K and L . Although we have made several corrections on the formulation in Ref.,[33] the results there remain intact.

3. Data Analyses

In this section, an empirical analysis is carried out using the ORBIS database from 2004 to 2009 provided by Bureau van Dijk.[39] First, we observe power laws, quasi-inversion symmetry, and Gibrat's law in three variables (K, L, Y). After that, we confirm the analytical discussion of the previous section using tangible fixed assets K, number of employees L, and firm sales Y.

3.1 Power Laws

Figures 1-3 show that the CDFs of K, L, and Y of Japanese firms in 2004 to 2009 obey power laws (1)-(3) in large scales that exceed certain size thresholds K_0, L_0, and Y_0. On the one hand, we determine the lower bounds of power-law ranges K_0, L_0, and Y_0 using the method in Ref.[42] , which is a modified version of the method in Ref.[43] . In this algorithm, the boundary between the power-law range and the log-normal range is detected with a statistical test.

The upper bounds of the power-law ranges are set at the top 0.1% of the data. The data in the power-law range between the upper and lower bounds are divided into logarithmically equal sized bins, and the power-law indices are estimated using the least square method, which is applied to the collective data in the bins.

3.2 Quasi-inversion Symmetry

Let us observe quasi-inversion symmetry $Y \leftrightarrow aK^{\alpha}L^{\beta}$ among three variables (K, L, Y) in the power-law region determined in 3.1. In this paper, this is called the quasi-inversion

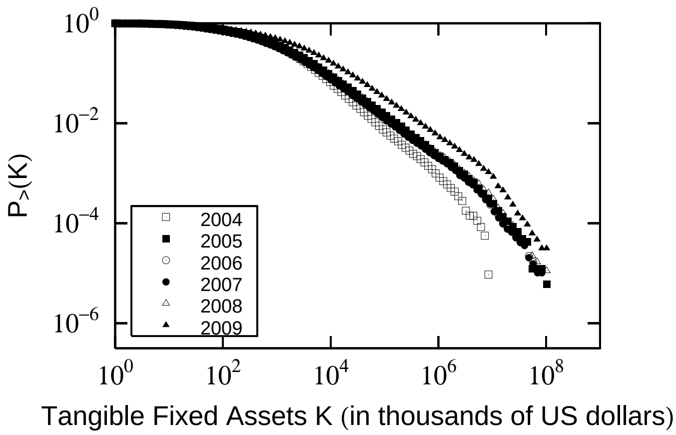

Fig. 1. Distribution of tangible fixed assets K for Japanese firms from 2004 to 2009. Number of firms changes across years but averages 148,186.

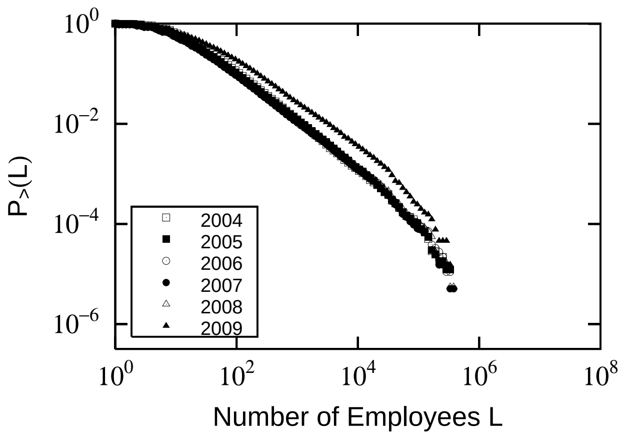

Fig. 2. Distribution of number of employees L for Japanese firms from 2004 to 2009. Number of firms changes across years but averages 155,465.

symmetry. Here, a, α, and β are the constant parameters that are identified as follows.

Figure 4 depicts a scatter plot between K and L of Japanese firms in 2008. As shown in Fig. 4, since we divide the power-law ranges of K and L into logarithmically equal sized bins, the power-law region is divided into cells enclosed by dashed lines. In each cell, we calculated the logarithmic mean of Y. By applying a least square method to the mean values of cells, we obtained a regression plane surface. By identifying the plane as the plane of the quasi-inverse symmetry, namely, $\log_{10} Y = \alpha \log_{10} K + \beta \log_{10} L + \log_{10} a$, values a, α, and β are estimated from the intercept and two components of the normal vector. For Japanese firms

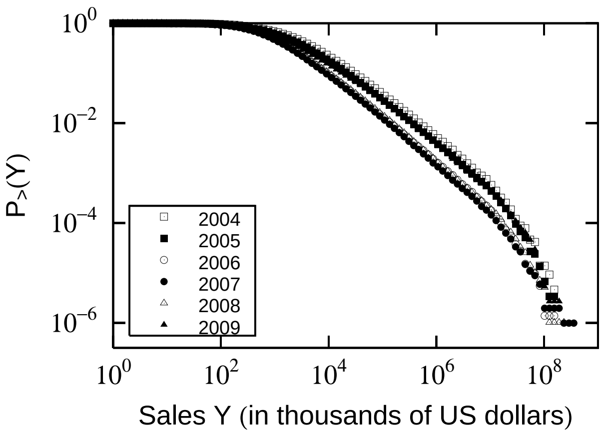

Fig. 3. Distribution of sales Y for Japanese firms from 2004 to 2009. Number of firms changes across years but averages 601,211.

in 2008, the values were calculated as $a = 68.3 \pm 1.1$, $\alpha = 0.36 \pm 0.02$, and $\beta = 0.70 \pm 0.02$.

With a, α, and β determined above, consider a scatter plot between Y and $Y'(\equiv aK^{\alpha}L^{\beta})$ (see Fig. 5). Quasi-inverse symmetry $Y \leftrightarrow Y'$ is observed by conducting a Kolmogorov-Smirnov test as follows. We first divide the range of Y and Y' into 18 bins of logarithmically equal size, given by $Y, Y' \in \left[10^{2.4+0.2(n-1)}, 10^{4+0.2n}\right), n = 1, 2, \cdots, 18$, as illustrated in Fig. 5. The null hypothesis, that two distributions $P(\left[10^{2.4+0.2(n-1)}, 10^{4+0.2n}\right), Y')$ and $P(Y, \left[10^{2.4+0.2(n-1)}, 10^{4+0.2n}\right))$ are identical, cannot be rejected at the 5% significance level in the power-law range ($n \geq 10$ in Fig. 6).

Consequently, quasi-inverse symmetry $Y \leftrightarrow aK^{\alpha}L^{\beta}$ $\left(K \leftrightarrow (Y/aL^{\beta})^{1/\alpha}, L \leftrightarrow (Y/aK^{\alpha})^{1/\beta}\right)$ is numerically confirmed in three valuables (K, L, Y) and is expressed using joint PDF $P_{KLY}(K, L, Y)$ as Eq. (5).

3.3 Gibrat's Law

We define the rate of three valuables $R = Y/aK^{\alpha}L^{\beta}$. The property, under which conditional PDF $Q(R|K, L)$ does not depend on K and L, is called Gibrat's law: Eq. (6). Figure 7 shows conditional PDFs $Q(R|K, L)$ of six blocks, which contain more than 50 data points (K, L), out of 25 blocks in Fig. 4. The difference of the blocks corresponds to the difference of (K, L) . From Fig. 7, we observe that PDFs $Q(R|K, L)$ are identical. Therefore, in the power-law region, we confirmed Gibrat's law.

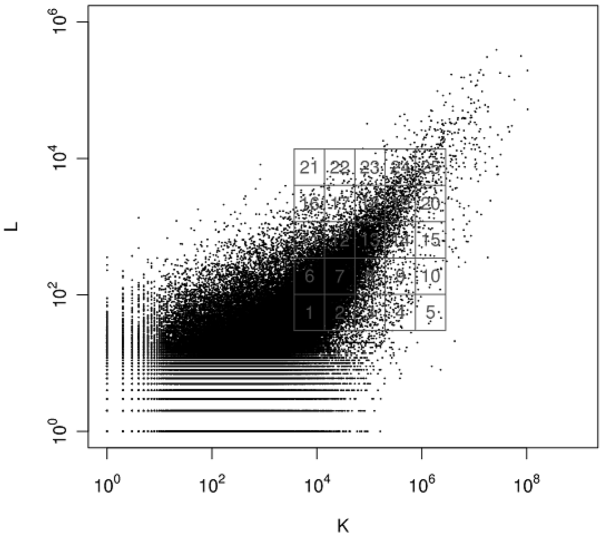

Fig. 4. Scatter plot between K (in thousands of US dollars) and L of Japanese firms in 2008. Amount of data is 176,980. As an example, power-law ranges of K and L are divided into 5 logarithmically equal sized bins, and the power-law region is divided into 25 cells. Amount of data in the power-law region is 26,286.

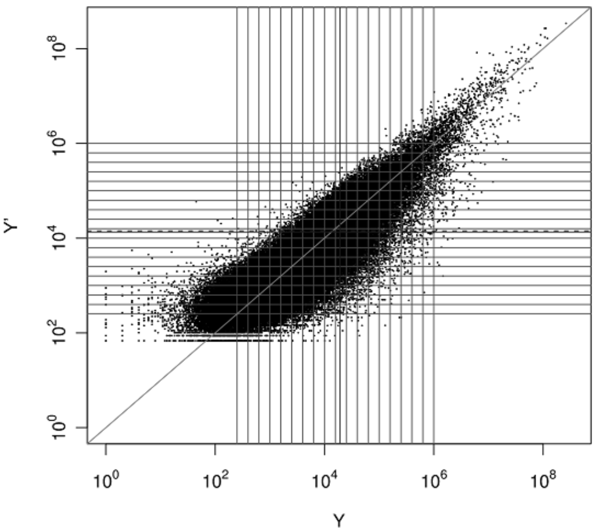

Fig. 5. Scatter plot between Y and $Y'(\equiv aK^{\alpha}L^{\beta})$ of Japanese firms in 2008. Dashed lines indicate bins of Y and Y', given by $Y, Y' \in \left[10^{2.4+0.2(n-1)}, 10^{4+0.2n}\right), n = 1, 2, \cdots, 18$, respectively.

3.4 Ridge of KL Plane

K and L data points are scattered in the KL plane (see Fig. 4). To clearly comprehend the density, we divided it into logarithmically equal sized cells and expressed the amount of data points in the cells in the light and the shade (see Fig. 8). Consider the logarithm of a cell's density as its height. Then the ridge is observed from the upper right to the lower left in the

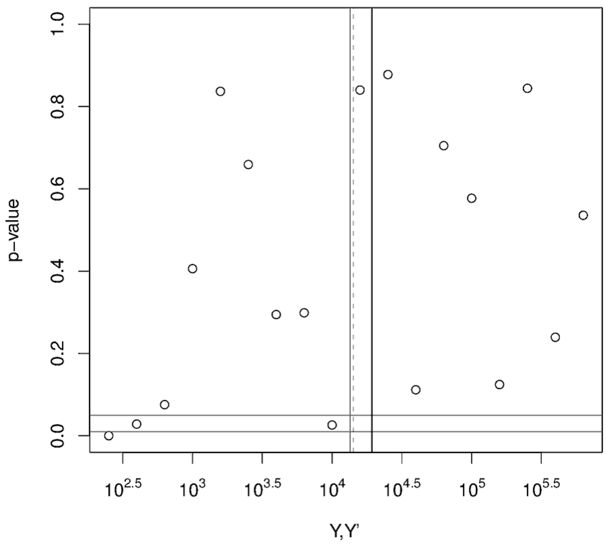

Fig. 6. Null hypothesis that two distributions $P(\left[10^{2.4+0.2(n-1)}, 10^{4+0.2n}\right), Y')$ and $P(Y, \left[10^{2.4+0.2(n-1)}, 10^{4+0.2n}\right))$ ($n = 1, 2, \cdots, 18$) are identical, cannot be rejected at the 5% significance level in the power-law range ($n \geq 10$). Vertical lines indicate lower bounds of Y and Y' .

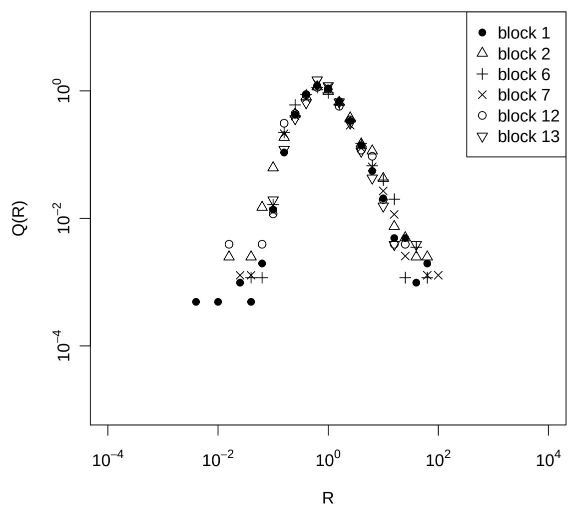

Fig. 7. Conditional PDFs $Q(R|K, L)$ of six blocks that contain more than 50 data points (K, L) out of 25 blocks in Fig. 4.

KL plane. As the steepest-ascent line in the profit space, a ridge is discussed in Ref.[44] . In this paper, we determine the cells that constitute the ridge using the surface openness defined as follows.[36–38]

Figure 9 depicts the grid linked by the center points of the cells. From grid point A, we counterclockwisely represent each azimuth as $D = 1, 2, \cdots, 8$. As shown in Fig. 10, the minimum zenith and nadir angles at grid point A within distance L in azimuth D are represented by ${}_D\phi_L$ and ${}_D\psi_L$. Positive openness Φ_L is defined by the mean value of ${}_D\phi_L$ along the eight azimuth, and negative openness Ψ_L is the corresponding mean of ${}_D\psi_L$. The surface openness is defined by the difference:

$$\Phi_L - \Psi_L = \frac{1}{8}\sum_{D=1}^{8} {}_D\phi_L - \frac{1}{8}\sum_{D=1}^{8} {}_D\psi_L \ . \tag{29}$$

The surface openness takes a negative value at the depressions and the valleys, zero at the level surface, the saddle point, and the uniform slope, and positive values at the ridge and the summit. In this analysis, by setting $L = 10$, we estimate the surface openness for each cell and extract the cells of the opennesses that exceed 0.8 . In Fig. 8, the cells in the power-law region are expressed by black dots.

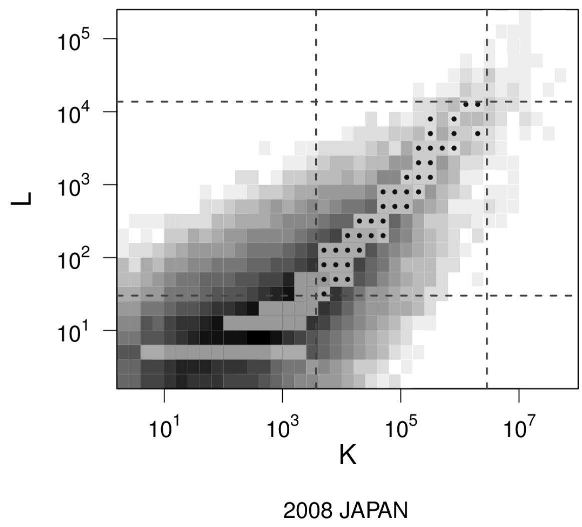

Fig. 8. To clearly comprehend the density of Fig. 4, we divide the KL plane into logarithmically equal sized cells and express the number of data points in them in light and shade. Broken lines represent upper and lower bounds of power-law ranges of K and L .

3.5 Independence of Z_1 and Z_2 and the Distributions

The black dots in Fig. 8 are the centers of cells $(\log_{10} K_i, \log_{10} L_i)$ that constitute the ridge in the power-law region of the KL plane $(i = 1, 2, \cdots, N)$. Here, N is number of cells which constitute the ridge. We signify the means and the $\sqrt{2}$ times of the standard deviations as

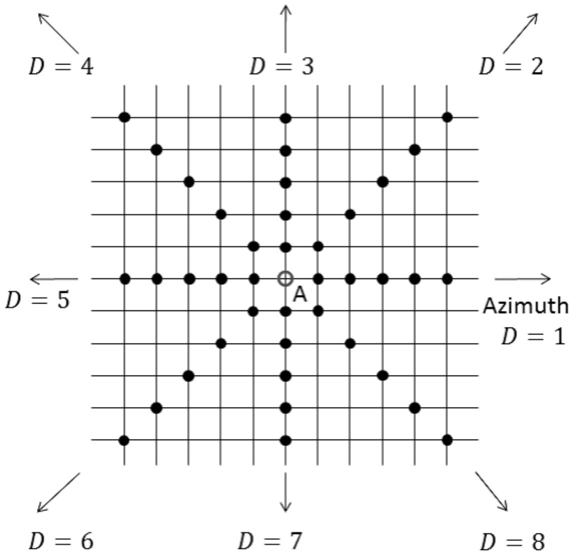

Fig. 9. Grid linked by center points of cells. From a grid point, we counterclockwisely represent each azimuth as $D = 1, 2, \cdots, 8$.

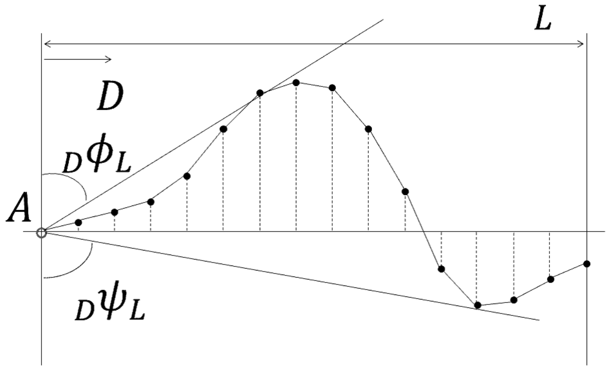

Fig. 10. Dots represent height of cells within distance L in azimuth D . From grid point A within distance L, we estimate zenith angles and denote the minimum as ${}_D\phi_L$. Similarly, the minimum nadir angle is expressed by ${}_D\psi_L$.

(m_K, m_L) and (σ_K, σ_L), respectively as follows:

$$m_K = \frac{1}{N}\sum_{i=1}^{N}\log_{10} K_i\ , \quad m_L = \frac{1}{N}\sum_{i=1}^{N}\log_{10} L_i\ , \tag{30}$$

$$\sigma_K = \sqrt{\frac{2}{N}\sum_{i=1}^{N}(\log_{10} K_i - m_K)^2}\ , \quad \sigma_L = \sqrt{\frac{2}{N}\sum_{i=1}^{N}(\log_{10} L_i - m_L)^2}\ . \tag{31}$$

In Fig. 8, $N = 39$, $m_K = 4.87$, $m_L = 2.78$, $\sigma_K = 1.10$, and $\sigma_L = 1.05$. Using the parameters, Eq. (12) transforms all data $(\log_{10} K, \log_{10} L)$ into $(\log_{10} Z_1, \log_{10} Z_2)$. Figure 11 shows the scatter plots between Z_1 and Z_2 .

Variables Z_1 and Z_2, which were obtained by the transformation (12), are not only uncorrelated but also independent. To confirm the independence between Z_1 and Z_2, we divide the range of Z_1 into 15 logarithmically equal sized bins: $Z_1 \in [10^{1+0.2(n-1)}, 10^{1+0.2n})$ ($n = 1, \ldots, 15$) as shown in Fig. 11. Conditional PDFs $P_{Z_2}(Z_2 \mid Z_1 \in [10^{-1+0.2(n-1)}, 10^{-1+0.2n}))$ are depicted in Fig. 12. The distributions for the different values of n are almost identical. In fact, using the Kolmogorov-Smirnov tests, with the null hypothesis that, for any pair of 15 distributions, two distributions are identical cannot be rejected at the 5% significance level. $P_{Z_2}(Z_2|Z_1) = P_{Z_2}(Z_2)$ is equivalent with $P_{Z_1Z_2}(Z_1, Z_2) = P_{Z_1}(Z_1)P_{Z_2}(Z_2)$, and then the independence between Z_1 and Z_2 is confirmed numerically. At the same time, from Fig. 12, we also confirm $Z_2 \leftrightarrow 1/Z_2$ symmetry.

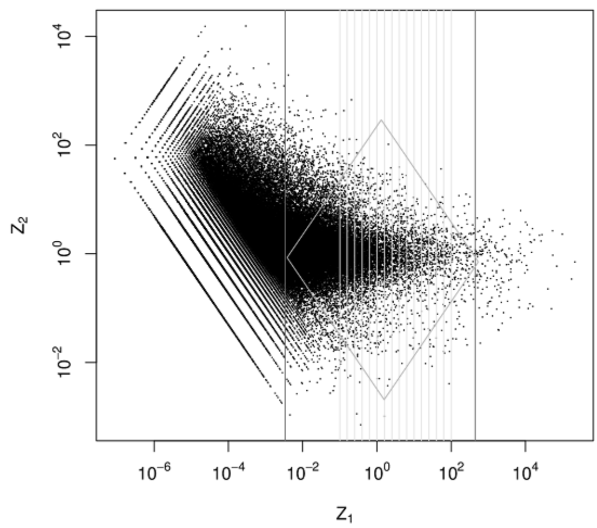

Fig. 11. Scatter plots between Z_1 and Z_2 are transformed from K and L of Japanese firms in 2008. Thick vertical lines indicate upper and lower bounds of power-law range of Z_1 . Thin vertical lines indicate bins of Z_1, given by $Z_1 \in [10^{-1+0.2(n-1)}, 10^{-1+0.2n})$ ($n = 1, \ldots, 15$) . Rhombus indicates corresponding power-law region of K and L .

3.6 Consistency of Power-law Indices

When variables Z_1 and Z_2 are independent, the solution of partial differential Eqs. (16) and (17) is uniquely determined to be the product of the power-law functions of Z_1 and Z_2 . Figure 13 shows the CDF of Z_1, which is the numerical Z_2 integration of the scatter plots between Z_1 and Z_2 . In the figure, the power law of Z_1 is observed in a range that corresponds to the power-law region of K and L .

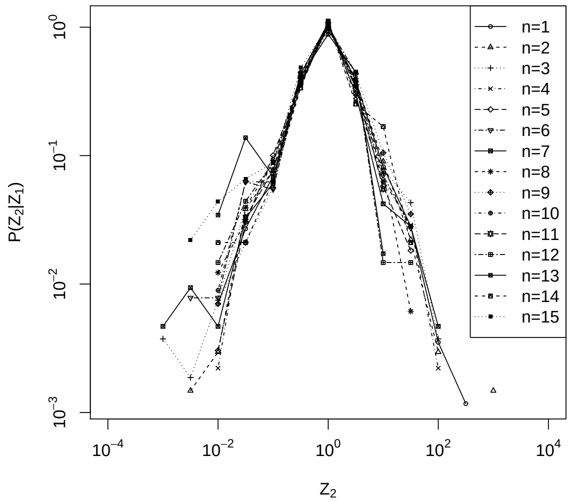

Fig. 12. Conditional PDFs $P_{Z_2}(Z_2 \mid Z_1 \in [10^{1+0.2(n-1)}, 10^{1+0.2n}))$ $(n = 1, \ldots, 15)$

Figures 14 and 15 depict the CDFs of Z_2 and $1/Z_2$, which are the numerical Z_1 integrations of the scatter plots between Z_1 and Z_2 . In the figures, the power laws of Z_2 and $1/Z_2$ are observed in the $Z_2 > 1$ and $Z_2 < 1$ ranges that correspond to the power-law region of K and L, respectively. These features did not change in different years or in other countries.

From these observations, the power laws, which are the solution of Eqs. (16) and (17), are represented as Eqs. (18) and (19). For Japanese firms in 2008, the parameters in Eq. (27) are estimated as $\mu_1 = 0.463 \pm 0.001$, $\mu_2 = 1.22 \pm 0.01$, $\mu_2{}' = 1.19 \pm 0.02$, $\theta_1 = 0.569 \pm 0.018$, and $\theta_2 = 0.171 \pm 0.018$. Here, $\mu_2{}'$ is an estimate from $1/Z_2$ distribution. In this case, $G_+{}'(1) = 15.6 \pm 1.68$, $G_-{}'(1) = 3.68 \pm 0.35$, and $\mu_1/\theta_1 + \mu_2/\theta_2 = 7.95 \pm 0.93$. Using these parameters Eq. (27) is reduced to Eq. (28), and this feature also did not change in different years or in other countries.

Finally, let us numerically confirm the validity of Eqs. (26) and (28). Figures 16–18 represent the relations between μ_K and $2\mu_1/\sigma_K$, μ_L and $2\mu_1/\sigma_K$, μ_Y and μ_1/θ_1 for ten countries from 2004 to 2009, except for years when the amount of data was not sufficient. Figures 16 and 17 show the validity of Eq. (26), and we also confirmed the validity of Eq. (28) from Fig. 18.

4. Conclusion

We directly observed quasi-inverse symmetry and Gibrat's law of three variables: tangible fixed assets K, number of employees L, and sales Y of firms from all over the world. These

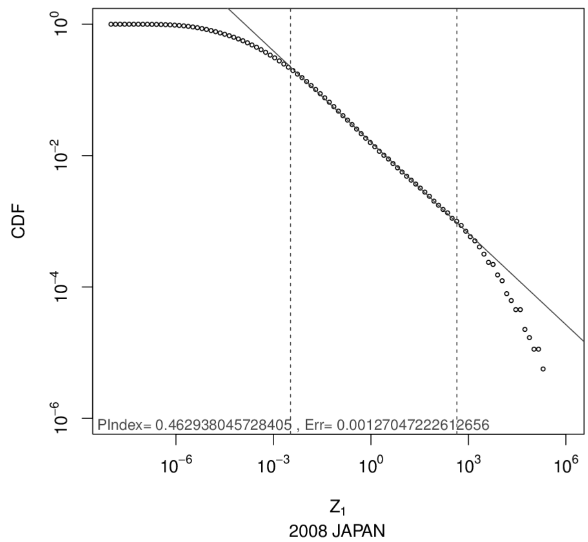

Fig. 13. CDF of Z_1 for Japanese firms in 2008. Number of firms is 176,980. Vertical broken lines indicate upper and lower bounds of power-law range of Z_1 . Number of firms in range is 36,961.

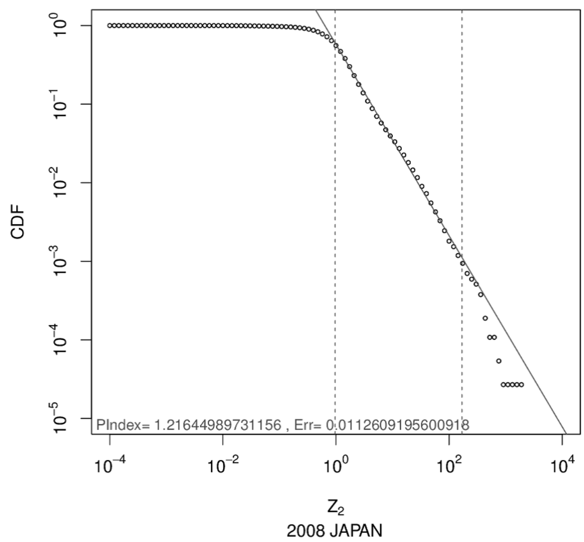

Fig. 14. CDF of Z_2 for Japanese firms in 2008. Vertical broken lines indicate upper and lower bounds of power-law range of Z_2 . Number of firms in range is 21,239.

two laws in two variables have already been confirmed; however, the laws in three variables were directly observed for the first time.

From the laws, the partial differential equation of joint PDF $P_{KL}(K, L)$ is derived. To solve it, variables K and L must be transformed into independent variables. In a previous study, using the regression line derived from the least square method that was applied to the mean values in bins, we transformed the variables. However, this procedure was not sufficiently

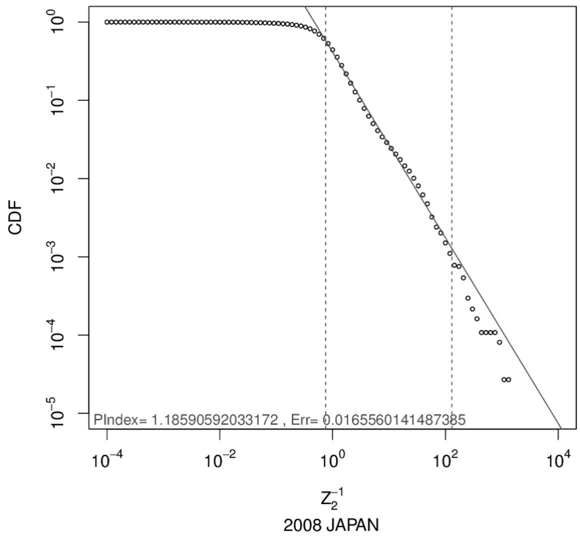

Fig. 15. CDF of $1/Z_2$ for Japanese firms in 2008. Vertical broken lines indicate upper and lower bounds of power-law range of $1/Z_2$. Number of firms in range is 21,498.

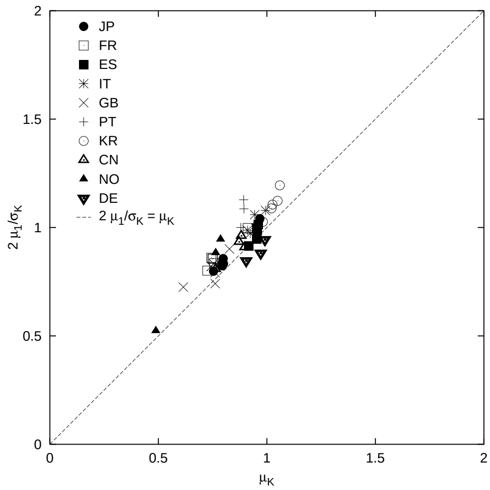

Fig. 16. Relationship between $2\mu_1/\sigma_K$ and μ_K for ten countries from 2004 to 2009: Japan (JP, average number of observations, 146,658), France (FR, 352,755), Spain (ES, 443,179), Italy (IT, 204,139), United Kingdom (GB, 49,288), Portugal (PT, 222,924), Korea (KR, 67,967), China (CN, 188,161), Norway (NO, 54,887) and Germany (DE, 21,994). Dashed line represents $2\mu_1/\sigma_K = \mu_K$. Error bars are extremely small, so they are omitted.

accurate to confirm that the numerically estimated power-law indices follow the analytical results.

In this study, we applied the surface openness used in geomorphology and accurately identified the transformation from a geomorphologic point of view. As a result, the relations

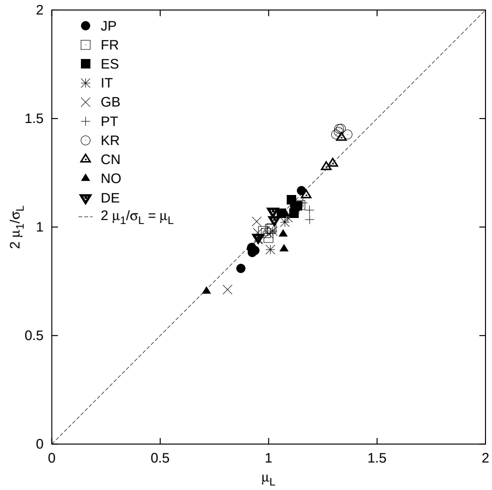

Fig. 17. Relationship between $2\mu_1/\sigma_L$ and μ_L for ten countries from 2004 to 2009. Dashed line represents $2\mu_1/\sigma_L = \mu_L$. Error bars are extremely small, so they are omitted.

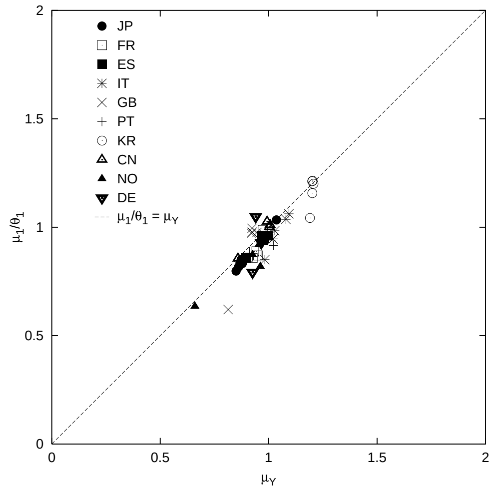

Fig. 18. Relationship between μ_1/θ_1 and μ_Y for ten countries from 2004 to 2009. Dashed line represents $\mu_1/\theta_1 = \mu_Y$. Error bars are extremely small, so they are omitted.

among power-law indices, whose observation was difficult, can be confirmed in empirical data. Consequently, we verified our analytical discussion and concluded that the functional form of $P_{KL}(K, L)$ is valid.

In our analyses, cells with surface openness over threshold 0.8 were extracted because they constitute the ridge of the KL plane. To extract the cells systematically for each KL plane, we should find the threshold value at which the cells, which constitute the ridge, decompose small clusters.[45] We hope to address this task in the future.

Acknowledgment

This study was supported by JSPS KAKENHI Grant Number 24510212, 24710156.

References

1) V. Pareto: *Cours d'Économie Politique* (Macmillan, London, 1897).

2) M. E. J. Newman: Contemporary Physics **46** (2005) 323.

3) A. Clauset, C. R. Shalizi, and M. E. J. Newman: SIAM Review **51** (2009) 661.

4) E. Bonabeau and L. Dagorn: Phys. Rev. E **51** (1995) R5220.

5) S. Render: Eur. Phys. J. B **4** (1998) 131.

6) M. Takayasu, H. Takayasu, and T. Sato: Physica A **233** (1996) 824.

7) A. Saichev, Y. Malevergne, and D. Sornette: *Theory of Zipf's Law and Beyond* (Springer, Berlin, 2009).

8) T. Kaizoji: Physica A **326** (2003) 256.

9) T. Yamano: Eur. Phys. J. B **38** (2004) 665.

10) A. Ishikawa: Physica A **371** (2006) 525.

11) A. Ishikawa: Prog. Theor. Phys. Supple. **179** (2009) 103.

12) R. N. Mantegna and H. E. Stanley: Nature **376** (1995) 46.

13) R. L. Axtell: Science **293** (2001) 1818.

14) B. Podobnik, D. Horvatic, A. M. Petersen, B. Urošević, and H. E. Stanley: Proc. Natl. Acad. Sci. **107** (2010) 18325.

15) D. Fu, F. Pammolli, S. V. Buldyrev, M. Riccaboni, K. Matia, K. Yamasaki, and H. E. Stanley: Proc. Natl. Acad. Sci. **102** (2005) 18801.

16) B. Podobnik, D. Horvatic, F. Pammolli, F. Wang, H. E. Stanley, and I. Grosse: Phys. Rev. E **77** (2008) 056102.

17) K. Okuyama, M. Takayasu, and H. Takayasu: Physica A **269** (1999) 125.

18) J. J. Ramsden and G. Kiss-Haypál: Physica A **277** (2000) 220.

19) T. Mizuno, M. Katori, H. Takayasu, and M. Takayasu: *Empirical Science of Financial Fluctuations* (Springer, Tokyo, 2002) p.321.

20) E. Gaffeo, M. Gallegati, and A. Palestrinib: Physica A **324** (2003) 117.

21) J. Zhang, Q. Chen, and Y. Wang: Physica A **388** (2009) 2020.

22) H. Kesten: Acta Math. **131** (1973) 207.

23) M. Levy and S. Solomon: Int. J. Mod. Phys. C **7** (1996) 595.

24) D. Sornette and R. Cont: J. Phys. I **7** (1997) 431.

25) H. Takayasu, A. Sato, and M. Takayasu: Phys. Rev. Lett. **79** (1997) 966.

26) H. E. Stanley: *Introduction to Phase Transitions and Critical Phenomena* (Clarendon Press, Oxford, 1971).

27) R. Gibrat: *Les Inégalités Économiques* (Sirey, Paris, 1932).

28) J. Sutton: J. Econ. Lit. **35** (1997) 40.

29) Y. Fujiwara, W. Souma, H. Aoyama, T. Kaizoji, and M. Aoki: Physica A **321** (2003) 598.

30) Y. Fujiwara, C. D. Guilmi, H. Aoyama, M. Gallegati, and W. Souma: Physica A **335** (2004) 197.

31) T. Watanabe, T. Mizuno, A. Ishikawa, and S. Fujimoto: Economic Review (Keizai Kenkyuu) **62** (2011) 193 [in Japanese].

32) T. Mizuno, T. Watanabe, A. Ishikawa, and S. Fujimoto: Prog. Theor. Phys. Supple. **194** (2012) 122.

33) A. Ishikawa, S. Fujimoto, T. Watanabe, and T. Mizuno: Physica A **392** (2013) 2104.

34) C. W. Cobb and P. H. Douglas: Amer. Econ. Rev. **18** (1928) 139.

35) H. Watanabe, H. Takayasu, and M. Takayasu: Statistical Studies on Interrelationships of Financial Indicators of Japanese Firms, The Physical Society of Japan 2009 Autumn Meeting [in Japanese].

36) R. Yokoyama, M. Sirasawa, and Y. Kikuchi: Journal of the Japan Society of Photogrammetry and Remote Sensing **38** (4) (1999) 26 [in Japanese with English abstract].

37) R. Yokoyama, M. Sirasawa, and R. Pike: Photogrammetric Engineering and Remote Sensing **68** (3) (2002) 257.

38) O. D. A. Prima, A. Echigo, R. Yokoyama, and T. Yoshida: Geomorphology **78** (2006) 373.

39) Bureau van Dijk, http://www.bvdinfo.com/Home.aspx.

40) A. H. Jessen and T. Mikosch: Inst. Math. **94** (2006) 171.

41) X. Gabaix: Annu. Rev. Econ. **1** (2009) 255.

42) S. Fujimoto, A. Ishikawa, T. Mizuno, and T. Watanabe: Economics E-Journal **5** (2011) 2011-20.

43) Y. Malevergne, V. Pisarenko, and D. Sornette: Phys. Rev. E **83** (2011) 036111.

44) W. Souma: *Econophysics of Markets and Business Networks* (Springer, Italia, 2007) p.149.

45) D. H. Kim and H. Jeong: Phys. Rev. E **72** (2005) 046133.