Total Factor Productivity in Firms around the Globe

Kanazawa Gakuin University Shouji Fujimoto, Atushi Ishikawa^{*)}, Hitotsubashi University Takayuki Mizuno and Tsutomu Watanabe

Abstract

In this study, by employing the database which contains comprehensive information on 63,000,000 firms worldwide in the longitudinal period 1999–2009, it is observed that sales, plant assets and the number of employee obey power-law. At the same time, by using Cobb-Douglas production function, the authors estimate total factor productivity in firms around the globe and find that it also obeys power-law. In the analysis, it is pointed out that Cobb-Douglas production function is 2-dim symmetrical surface, which relates power-laws of sales, plant assets and the number of employee to each other, in 3-dim (Plant Assets, The Number of Employee, Sales) space.

The authors employ the database, Bureau van Dijk's ORBIS which contains comprehensive information on 63,000,000 firms worldwide (Western Europe, Eastern Europe, North America, South and Central America, Japan, China, and so on) for the years 1991–2009.¹⁾ By using this database, it is found that not only sales, plant assets and the number of employee but also total factor productivity in firms around the globe obey power-law in the large-scale range. At the same time, it is pointed out that Cobb-Douglas production function,²⁾ which defines total factor productivity, can be interpreted as 2-dimensional symmetrical surface under detailed quasi-balance in 3-dimensional (Sales, Plant Assets, The Number of Employee) space.

It is well known that, in the large-scale range of sales (denoted by Y), plant assets (denoted by K) and the number of employee (denoted by L), the probability density functions (pdfs) obey power-laws

$$P(Y) \propto Y^{-(\mu_Y+1)} \quad \text{for} \quad Y > Y_0 , \qquad (1)$$

$$P(K) \propto K^{-(\mu_K+1)} \quad \text{for} \quad K > K_0 , \qquad (2)$$

$$P(L) \propto L^{-(\mu_L+1)} \qquad \text{for} \quad L > L_0$$
, (3)

respectively. Here, Y_0 , K_0 and L_0 are some size thresholds. The power-law and the exponent μ are called Pareto's Law and Pareto index, respectively.^{3),4)}

For instance, these power-laws in 2000–2009 Japan are depicted in Figs 1–3 ^{**)}. In many countries, it is observed that $\mu_L > \mu_Y \simeq \mu_K$ (Fig. 4). In Fig. 4, JP means JAPAN, FR FRANCE, ES SPAIN, CA CANADA, IT ITALY, RU RUSSIAN FEDERATION, GB UNITED KINGDOM, PT PORTUGAL, KR KOREA, REPUBLIC OF, CN CHINA, UA UKRAINE,

^{*)} E-mail: ishikawa@kanazawa-gu.ac.jp

^{**)} The database does not contain plant assets in 1999–2003. At the same time, the database contains small number of sales and the number of employee data in 1999.

NO NORWAY, DE GERMANY, SE SWEDEN, BE BELGIUM, FI FINLAND, CZ CZECH REPUBLIC, PL POLAND, BG BULGARIA, EE ESTONIA, AT AUSTRIA, LV LATVIA, DK DENMARK, HU HUNGARY, HR CROATIA, GR GREECE, NL NETHERLANDS, IE IRELAND, SK SLOVAKIA and SI SLOVENIA. Pareto indices in Fig. 4 are estimated by using uniformly most powerful unbiased test.⁵⁾ Figures 5–7 show scatter plots of Pareto indices μ_L , μ_K and μ_Y in 2008. There are weak correlations between μ_L and μ_Y (Fig. 5) and between μ_K and μ_Y (Fig. 6). However, there is little correlation between μ_L and μ_K (Fig. 7).

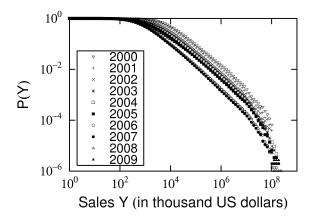


Fig. 1. Pdfs of sales Y in 2000–2009 Japan.

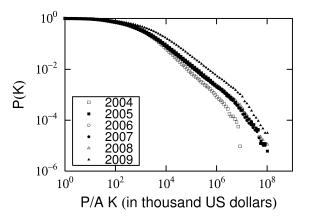


Fig. 2. Pdfs of plant assets K in 2004–2009 Japan.

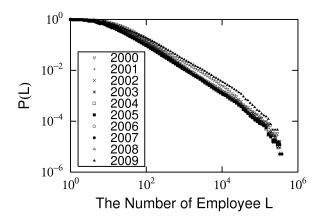


Fig. 3. Pdfs of the number of employee L in 2000–2009 Japan.

In this database, detailed quasi-balance⁶⁾ and Gibrat's Law^{7),8)} are also confirmed. In this paper, detailed quasi-balance is designated as symmetry under exchange $Y \leftrightarrow A_{LY}L^{\nu_{LY}}$, $K \leftrightarrow A_{YK}Y^{\nu_{YK}}$ and $K \leftrightarrow A_{LK}L^{\nu_{LK}}$ observed in the joint pdfs $P_J(L,Y)$, $P_J(Y,K)$ and $P_J(L,K)$, respectively, as follows:

$$P_J(L,Y) = P_J\left(\left(\frac{Y}{A_{LY}}\right)^{1/\nu_{LY}}, A_{LY}L^{\nu_{LY}}\right) , \qquad (4)$$

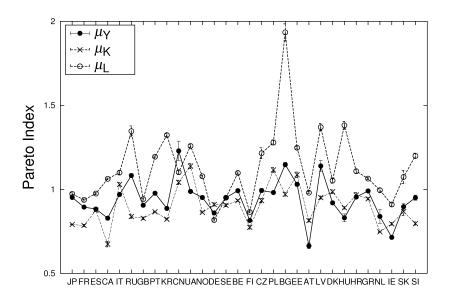


Fig. 4. Pareto indices of sales μ_Y , plant assets μ_K and the number of employee μ_L of 30 countries in 2008.

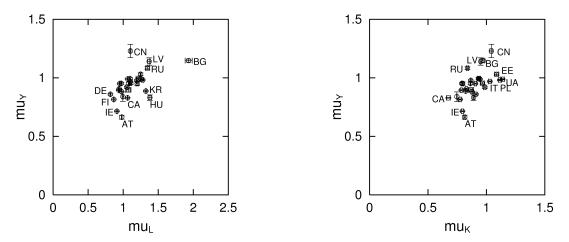


Fig. 5. Scatter plot of Pareto indices μ_L and μ_Y in 2008. Fig. 6. Scatter plot of Pareto indices μ_K and μ_Y in 2008.

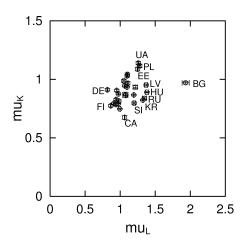


Fig. 7. Scatter plot of Pareto indices μ_L and μ_K in 2008.

4

$$P_J(Y,K) = P_J\left(\left(\frac{K}{A_{YK}}\right)^{1/\nu_{YK}}, A_{YK}Y^{\nu_{YK}}\right) , \qquad (5)$$

$$P_J(L,K) = P_J\left(\left(\frac{K}{A_{LK}}\right)^{1/\nu_{LK}}, A_{LK}L^{\nu_{LK}}\right)$$
(6)

Here, A_{LY} , A_{YK} , A_{LK} , ν_{LY} , ν_{YK} and ν_{LK} are parameters. For instance, scatter plots of sales, plant assets and the number of employee and the symmetric lines

$$\log Y = \nu_{LY} \log L + \log A_{LY} , \qquad (7)$$

$$\log K = \nu_{YK} \log Y + \log A_{YK} , \qquad (8)$$

$$\log K = \nu_{LK} \log L + \log A_{LK} \tag{9}$$

in 2008 Japan are depicted in Figs 8–10. Here, $\log x$ is common logarithm $\log_{10} x$. The symmetric lines (7)–(9) in Figs. 8–10 are obtained as follows. The horizontal axis in the power-law range is divided into logarithmically equal bins with width $10^{0.4}$. In each bin, the average of vertical scale is calculated. The least-square method is applied to the averages in bins, and the symmetric line is obtained. In this method, the large- and small-scale points in the power-law range are estimated equally. This equally weighted estimation is absolutely imperative to obtain the symmetric line which leads to detailed quasi-balance and Gibrat's Law. The symmetric line is not obtained by the least-square method simply applied to points of scatter plot. In Figs 8–10, detailed quasi-balance (4)–(6) are approximately confirmed by using Kolmogorov-Smirnov test.

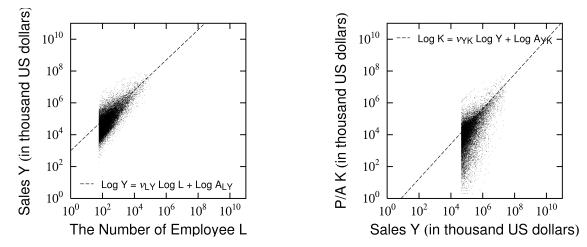


Fig. 8. Scatter plot of the number of employee L in the power-law range and sales Y in 2008 Japan and the symmetric line (7).

Fig. 9. Scatter plot of sales Y in the power-law range and plant assets K in 2008 Japan and the symmetric line (8).

Gibrat's Law means that the conditional pdfs $Q(R_{LY}|L) \equiv Q(R_{LY}|A_{LY}L^{\nu_{LY}}), Q(R_{YK}|Y) \equiv Q(R_{YK}|A_{YK}Y^{\nu_{YK}})$ and $Q(R_{LK}|L) \equiv Q(R_{LK}|A_{LK}L^{\nu_{LK}})$ of the rates $R_{LY} = Y/(A_{LY}L^{\nu_{LY}}),$ $R_{YK} = K/(A_{YK}Y^{\nu_{YK}})$ and $R_{LK} = K/(A_{LK}L^{\nu_{LK}})$ do not depend on the initial values L, Y

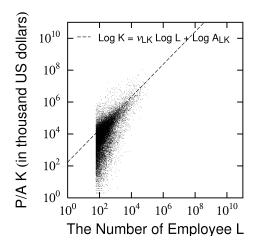


Fig. 10. Scatter plot of the number of employee L in the power-law range and plant assets K in 2008 Japan and the symmetric line (9).

and L, respectively, as follows;

$$Q(R_{LY}|L) = Q(R_{LY}) , \qquad (10)$$

$$Q(R_{YK}|Y) = Q(R_{YK}) , \qquad (11)$$

$$Q(R_{LK}|L) = Q(R_{LK}) . (12)$$

In order to verify Gibrat's Law (10), we divide the range of the initial value $A_{LY}L^{\nu_{LY}}$ into logarithmically equal bins as $A_{LY}L^{\nu_{LY}} \in [10^{4+0.4(n-1)}, 10^{4+0.4n})$ $(n = 1, 2, \dots, 5)$. The conditional pdfs $q(r_{LY}|L)$ of the logarithmic rate $r_{LY} = \log R_{LY}$ are shown in Fig. 11. In Fig. 11, the rate distribution barely changes as n increases. This means that Gibrat's Law (10) is valid. Gibrat's Laws (11), (12) are also confirmed as in the same manner (Figs. 12, 13).

By using detailed quasi-balance and Gibrat's Law, Pareto indices μ_Y , μ_K and μ_L in Pareto's Laws (1)–(3) are related to each other as follows;

$$\mu_L + 1 = \nu_{LY}(\mu_Y + 1) , \qquad (13)$$

$$\mu_Y + 1 = \nu_{YK}(\mu_K + 1) , \qquad (14)$$

$$\mu_L + 1 = \nu_{LK}(\mu_K + 1) . \tag{15}$$

These relations are observed in empirical data (Figs. 14–16 for instance).

Let us derive the relation (13) between μ_L and μ_Y by using detailed quasi-balance (4) and Gibrat's Law (10) *). From the relation $P_J(L, Y)dLdY = P_J(L, R_{LY})dLdR_{LY}$ under the change of variables $(L, Y) \longleftrightarrow (L, R_{KY})$, these two joint pdfs are related to each other

$$P_J(L, R_{LY}) = A_{LY} L^{\nu_{LY}} P_J(L, Y) .$$
(16)

^{*)} This method is firstly used to derive Pareto's Law by using detailed balance and Gibrat's Law.⁹⁾

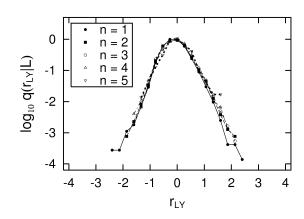


Fig. 11. Conditional pdfs of the logarithm of the rate $r_{LY} = \log R_{LY} = \log (Y/(A_{LY}L^{\nu_{LY}}))$ in 2008 Japan. Here, L is the number of employee and Y is sales in thousand US dollars. The range of the initial value $A_{LY}L^{\nu_{LY}}$ is divided into logarithmically equal bins as $A_{LY}L^{\nu_{LY}} \in [10^{4+0.4(n-1)}, 10^{4+0.4n})$ $(n = 1, 2, \cdots, 5)$.

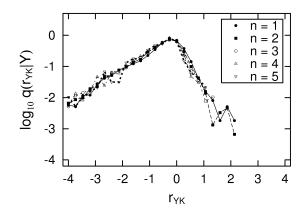


Fig. 12. Conditional pdfs of the logarithm of the rate $r_{YK} = \log R_{YK} = \log (K/(A_{YK}Y^{\nu_{YK}}))$ in 2008 Japan. Here, Y is sales and K is plant assets in thousand US dollars. The range of the initial value $A_{YK}Y^{\nu_{YK}}$ is divided into logarithmically equal bins as $A_{YK}Y^{\nu_{YK}} \in [10^{4+0.4(n-1)}, 10^{4+0.4n})$ $(n = 1, 2, \dots, 5)$.

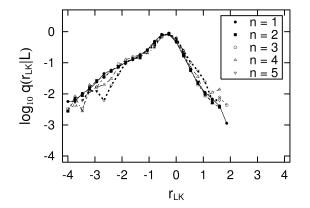


Fig. 13. Conditional pdfs of the logarithm of the rate $r_{LK} = \log R_{LK} = \log (K/(A_{LK}L^{\nu_{LK}}))$ in 2008 Japan. Here, L is the number of employee and K is plant assets in thousand US dollars. The range of the initial value $A_{LK}L^{\nu_{LK}}$ is divided into logarithmically equal bins as $A_{LK}L^{\nu_{LK}} \in [10^{4+0.4(n-1)}, 10^{4+0.4n})$ $(n = 1, 2, \cdots, 5)$.

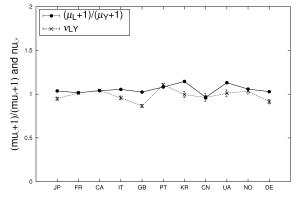


Fig. 14. The comparison between $(\mu_L + 1)/(\mu_Y + 1)$ and ν_{LY} of 11 countries in 2008.

By using this relation, detailed quasi-balance (4) is rewritten as

$$P_J(L, R_{LY}) = R_{LY}^{-1} P_J\left(\left(\frac{Y}{A_{LY}}\right)^{1/\nu_{LY}}, R_{LY}^{-1}\right) .$$
(17)

Substituting $P_J(L, R_{LY})$ for $Q(R_{LY}|L)$ by using the definition of the conditional pdf $Q(R_{LY}|L) \equiv$

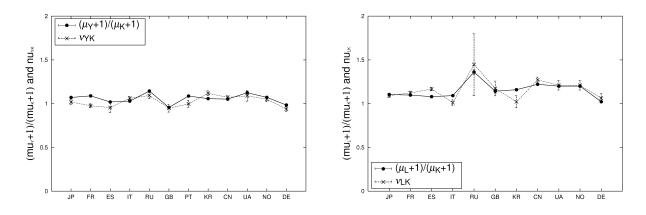


Fig. 15. The comparison between $(\mu_Y + 1)/(\mu_K + 1)$ and ν_{YK} of 12 countries in 2005.

Fig. 16. The comparison between $(\mu_L + 1)/(\mu_K + 1)$ and ν_{LK} of 11 countries in 2005.

 $P_J(L, R_{LY})/P(L)$, detailed quasi-balance is reduced to be

$$\frac{P(L)}{P\left((Y/A_{LY})^{1/\nu_{LY}}\right)} = \frac{1}{R_{LY}} \frac{Q\left(R_{LY}^{-1}|\left(Y/A_{LY}\right)^{1/\nu_{LY}}\right)}{Q(R_{LY}|L)}$$
(18)

$$= \frac{1}{R_{LY}} \frac{Q(R_{LY}^{-1})}{Q(R_{LY})} \equiv G(R_{LY}) .$$
(19)

Here, Gibrat's Law (10) is used in Eq. (18). By expanding Eq. (19) around $R_{LY} = 1$, the following differential equation is obtained

$$G'(1)\nu_{LY}P(L) + L P'(L) = 0.$$
(20)

The solution is given by

$$P(L) = C_1 L^{-G'(1)\nu_{LY}} {.} {(21)}$$

From Eqs. (3) and (21), we identify that $G'(1)\nu_{LY} = \mu_L + 1$. On the other hand, by taking $L \to (Y/A_{LY})^{1/\nu_{LY}}$ in Eq. (21), we obtain

$$P\left((Y/A_{LY})^{1/\nu_{LY}}\right) = \frac{C_1}{A_{LY}^{-G'(1)}}Y^{-G'(1)} .$$
(22)

Identifying that $C_2 = C_1 / A_{LY}^{-G'(1)}$ and $P((Y/A_{LY})^{1/\nu_{LY}}) = P(Y)$, we obtain $G'(1) = \mu_Y + 1$. As a result, the relation (13) is derived. The relations (14) and (15) are also derived in the same manner.

Finally, by using Cobb-Douglas production function:

$$Y = AK^{\alpha}L^{\beta} , \qquad (23)$$

let us estimate total factor productivity A in firms. Total factor productivity is a residual which cannot be caused by K or L, and is considered to be related to technology growth and efficiency

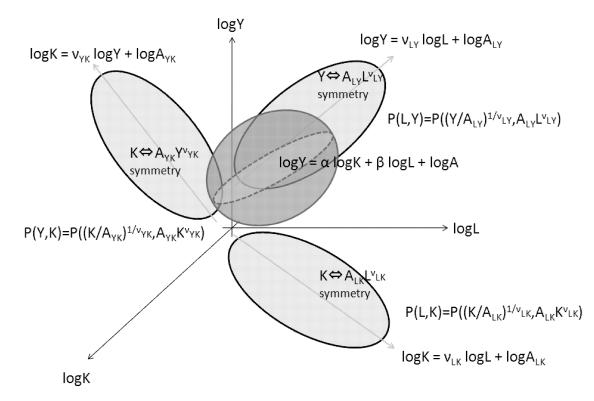


Fig. 17. Detailed quasi-balances (4)–(6) are maps of symmetry in 3-dim (K, L, Y) space to (L, Y), (Y, K) and (L, K) planes, respectively.

(see Refs. 10), 11) as recent studies for instance). In this investigation, the authors point out that Cobb-Douglas production function (23) can be interpreted as 2-dimensional symmetrical surface under detailed quasi-balance in 3-dimensional space (K, L, Y). Detailed quasi-balances (4)-(6) are symmetry of the joint pdfs $P_J(L, Y)$, $P_J(Y, K)$ and $P_J(L, K)$ under $Y \leftrightarrow A_{LY}L^{\nu_{LY}}$, $K \leftrightarrow A_{YK}Y^{\nu_{YK}}$ and $K \leftrightarrow A_{LK}L^{\nu_{LK}}$ exchanges in (L, Y), (Y, K) and (L, K) planes, respectively (Fig. 17). These three symmetry are maps of symmetry in 3-dim (K, L, Y) space to (L, Y), (Y, K)and (L, K) planes. The symmetric lines (7)-(9) relate power-laws of Y, K and L (1)-(3) to each other. Therefore, the symmetric plane with respect to Cobb-Douglas production function in 3-dim space (K, L, Y):

$$\log Y = \alpha \log K + \beta \log L + \log A \tag{24}$$

is also related to the power-laws.

In order to calculate total factor productivity A, we need to know a production elasticity of capital α and a production elasticity of labor β in Eq. (24). In the case the number of data points is not large, α and β can be estimated by multiple regression analysis. In our database in which the number of data points is enormous large, however, we concentrate ourselves on estimating large- and small-scale data points equally. This cannot be achieved by simple multiple regression analysis. At the same time, we must pay attention the correlation between plant assets K and

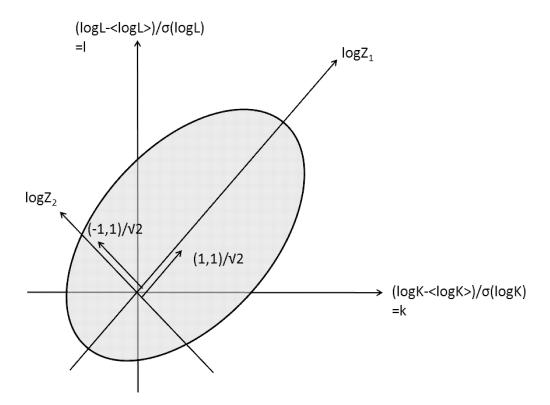


Fig. 18. By principle component analysis, we take first and second principle components $\log Z_1$ and $\log Z_2$, respectively.

the number of employee L (Fig. 10). This correlation generally causes multicollinearity. In order to avoid these difficulties, by using principal component analysis, we take first and second principle components as follows:

$$\log Z_1 = \frac{1}{\sqrt{2}}k + \frac{1}{\sqrt{2}}l$$
$$= \frac{\log K}{\sqrt{2}\sigma_K} + \frac{\log L}{\sqrt{2}\sigma_L} - \frac{\langle \log K \rangle}{\sqrt{2}\sigma_K} - \frac{\langle \log L \rangle}{\sqrt{2}\sigma_L} , \qquad (25)$$

$$\log Z_2 = -\frac{1}{\sqrt{2}}k + \frac{1}{\sqrt{2}}l$$
$$= -\frac{\log K}{\sqrt{2}\sigma_K} + \frac{\log L}{\sqrt{2}\sigma_L} + \frac{\langle \log K \rangle}{\sqrt{2}\sigma_K} - \frac{\langle \log L \rangle}{\sqrt{2}\sigma_L} , \qquad (26)$$

where $k = (\log_{10} K - \langle \log_{10} K \rangle) / \sigma_K$, $l = (\log_{10} L - \langle \log_{10} L \rangle) / \sigma_L$, $\sigma_K = \sigma (\log_{10} K)$ and $\sigma_L = \sigma (\log_{10} L)$ (Fig. 18).

On one hand, the first principle component Z_1 correlates strongly with sales Y. Figure 19 shows the scatter plot of Z_1 and Y in 2008 JP, for example. On the other hand, the second principle component Z_2 correlates weakly with sales Y. Figure 20 shows the scatter plot of Z_2 and Y in 2008 JP, for example. Figures 19 and 20 are maps of 3-dim data points (K, L, Y) to (Z_1, Y) and (Z_2, Y) planes, respectively. From Figs. 19 and 20, we observe that the correlation between Z_2 and Y is negligible compared with the correlation between Z_1 and Y. If we take the correlation between Z_2 and Y forcibly, an inappropriately correlation may be incidentally estimated. In this analysis, therefore, we disregard the second principle component Z_2 . In this case, form the symmetric line:

$$\log Y = \gamma_1 \log Z_1 + \log A_1 \tag{27}$$

in scatter plot of Z_1 in the power-law range and Y (Fig. 19), α and β are identified as follows:

$$\alpha = \frac{\gamma_1}{\sqrt{2}\sigma_K}, \quad \beta = \frac{\gamma_1}{\sqrt{2}\sigma_L} \ . \tag{28}$$

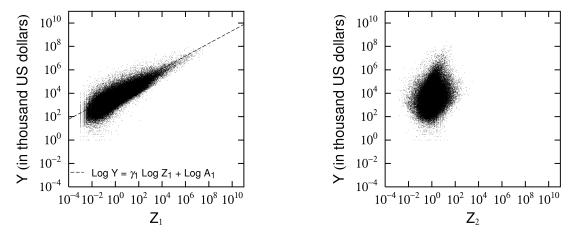


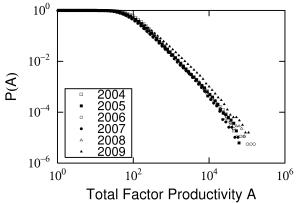
Fig. 19. Scatter plot of the first principle component Z_1 and sales Y and the symmetric line (27) in 2008 Z_2 and sales Y in 2008 Japan. Japan.

In the analysis, first we estimate γ_1 in the symmetric line (27) of scatter plot of Z_1 and Y, and calculate α and β by Eqs. (28). Then, total factor productivity A is obtained by Cobb-Douglas production function (23) or (24). It is observed that the pdf also obeys power-law:

$$P(A) \propto A^{-(\mu_A+1)}$$
 for $A > A_0$.

For instance, power-laws in 2004–2009 Japan are depicted in Fig. 21. Annual changes of Pareto indices μ_A , μ_Y , μ_K , μ_L and elasticities α , β in 2004–2009 Japan are also depicted in Fig. 22. It is found that Pareto indices hardly change annually. In many countries, it is observed that $\mu_A > \mu_L$ (Fig. 23). Figure 24 shows scatter plots of Pareto indices μ_L and μ_A in 2008, and Fig. 25 shows scatter plot of elasticities α and β in Cobb-Douglas production function (23) in 2008. In these figures, there are little correlations between μ_L and μ_A and between α and β .

In this study, by employing the database which contains comprehensive information on 63,000,000 firms worldwide in the longitudinal period 1999–2009, first it is observed that sales Y, plant assets K and the number of employee L obey power-law in the large-scale range. The Pareto indices of sales μ_Y , plant assets μ_K and the number of employee μ_L are estimated around the globe in 1999–2009. It is found that Pareto indices hardly change annually in most countries. This phenomenon has been reported in Ref. 12). The authors have found that $\mu_L > \mu_K \simeq \mu_Y$



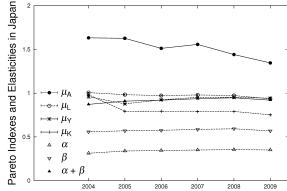


Fig. 21. Pdfs of total factor productivity A in 2004–2009 Japan.

Fig. 22. Annual changes of Pareto indices μ and elasticities α , β in 2004–2009 Japan.

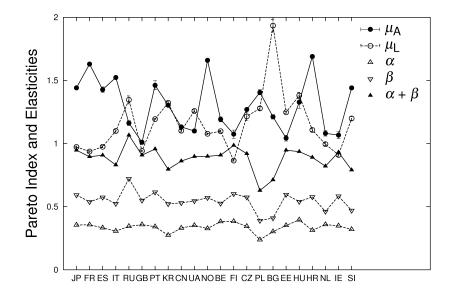


Fig. 23. Pareto indices of total factor productivity μ_A and the number of employee μ_L and production elasticities of capital α and production elasticities of labor β in 2008.

in many countries. There are weak positive correlations between μ_L and μ_Y and between μ_K and μ_Y . On the other hand, there is little correlation between μ_L and μ_K .

Second, detailed quasi-balances are found in joint pdfs $P_J(L, Y)$, $P_J(Y, K)$ and $P_J(L, K)$. Gibrat's Laws are also found in rate distributions of Y, K and L. By using detailed quasibalance and Gibrat's Law, Pareto indices μ_Y , μ_K and μ_L are related to each other analytically. These relations are also confirmed by empirical data. In the analysis, the symmetrical line under detailed quasi-balance in the joint pdf plays a central role. The authors have pointed out that the symmetrical lines in (L, Y), (Y, K) and (L, K) planes are maps of symmetrical surface in (K, L, Y) space, and that the symmetric surface is precisely Cobb-Douglas production function.

Third, by using principle component analysis, a production elasticity of capital α , a pro-

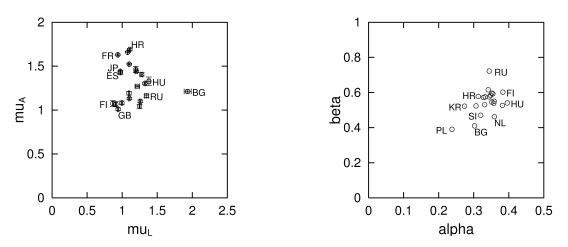


Fig. 24. Scatter plot of Pareto indices μ_L and μ_A in Fig. 25. Scatter plot of elasticities α and β in Cobb-2008. Douglas production function in 2008.

duction of elasticity of labor β and total factor product A in Cobb-Douglas production function are estimated around the globe in the period 2004–2009. It is found that the pdf of total factor product A also obeys power-law in the large-scale range and that the Pareto index μ_A hardly changes annually in most countries. The authors have found that $\mu_A > \mu_L > \mu_K \simeq \mu_Y$ in many countries. There is little correlation between μ_L and μ_A . It is also observed that $\beta > \alpha$ in all courtiers. This means that the number of employee L influences sales Y more than plant assets K in the globe. There is also little correlation between α and β .

These results are firstly obtained by using large-scale business data worldwide. The powerlaw distribution of total factor productivity A is firstly observed around the glove. At the same time, the authors have clarified the origin of Cobb-Douglas production function, which relates to power-laws of sales, plant assets and the number of employee.

Acknowledgment

The authors thank Research Institute of Economy, Trade and Industry, IAA (RIETI). The data set used in this work was supplied by RIETI. 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.

References

- 1) Bureau van Dijk, http://www.bvdinfo.com/Home.aspx.
- 2) C. W. Cobb and P. H. Douglas, American Economics Review, 18 (1928) 139.
- 3) V. Pareto, Cours d'Economique Politique, Macmillan, London, 1897.
- 4) M. E. J. Newman, Contemporary Physics 46 (2005) 323;
 - A. Clauset, C. R. Shalizi and M. E. J. Newman, SIAM Review 51 (2009) 661.
- 5) Y. Malevergne, V. Pisarenko and D. Sornette, arXiv:0909.1281v1.
- 6) A. Ishikawa, Physica A 371 (2006) 525;
- A. Ishikawa, Prog. Theor. Phys. Supple. 179 (2009) 103.
- 7) R. Gibrat, Les inegalites economiques (Sirey, Paris, 1932).
- 8) J. Sutton, J. Econo. Lit. 35 (1997) 40.

- 9) Y. Fujiwara, W. Souma, H. Aoyama, T. Kaizoji and M. Aoki, Physica A321 (2003) 598;
- Y. Fujiwara, CD. Guilmi, H. Aoyama, M. Gallegati and W. Souma, Physica A335 (2004) 197.
- 10) E. J. Bartelsman and P. J. Dhrmes, Journal of Productivity Analysis 9(1) (1998) 5.
- 11) B. Y. Aw, X. Chen and M. J. Roberts, Journal of Development Economics 66(1) (2001) 51.
- 12) T. Mizuno, M. Katori, H. Takayasu and M. Takayasu, "Statistical and Stochastic Laws in the Income of Japanese Companies" in Empirical Science of Financial Fluctuations – The Advent of Econophysics, (Springer Verlag, Tokyo, 2002) 321.