Available online at www.sciencedirect.com

SCIENCE DIRECT®

Physica A 321 (2003) 598 – 604

PHYSICA A

www.elsevier.com/locate/physa

Growth and fluctuations of personal income

Yoshi Fujiwara[a,*], Wataru Souma[b], Hideaki Aoyama[c], Taisei Kaizoji[d], Masanao Aoki[e]

[a] *Communications Research Laboratory, Keihanna Center, Kyoto 619-0289, Japan*
[b] *ATR Human Information Science Laboratories, Kyoto 619-0288, Japan*
[c] *Faculty of Integrated Human Studies, Kyoto University, Kyoto 606-8501, Japan*
[d] *Division of Social Sciences, International Christian University, Tokyo 181-8585, Japan*
[e] *Department of Economics, University of California, Los Angeles, CA 90095-1477, USA*

Received 15 October 2002

Abstract

Pareto's law states that the distribution of personal income obeys a power-law in the high income range. Its dynamical nature has been little studied hitherto, mostly due to the lack of empirical work. Using an exhaustive list of taxpayers in Japan for two consecutive years, when the economy was relatively stable, we report here that the law is a consequence from universal distribution of the growth rate of income and approximate time-reversal symmetry of incomes in the successive years. We also find a relation between positive and negative growth rates that shows good agreement with the data.

PACS: 89.90.+n; 02.50.−r; 05.40.+j; 47.53.+n

Keywords: Pareto's law; Personal income; Growth-rate; Multiplicative process; Econophysics

Flow and stock are the fundamental concepts in economics. They refer to a certain economic quantity over a given period of time and its accumulation at a point in time. Personal income and wealth can be regarded as the flow and stock of each individual in a giant dynamical network of people, which is open to various economic activities. The Italian social economist Vilfredo Pareto [1], more than a century ago, studied the distribution of personal income and wealth in society as a characterization of a country's

* Corresponding author. Present address: ATR, Human Information Science Laboratories, Kyoto 619-0288, Japan. Fax: +81-774-95-2647.

E-mail address: yfujiwar@atr.co.jp (Y. Fujiwara).

PII: S0378-4371(02)01663-1

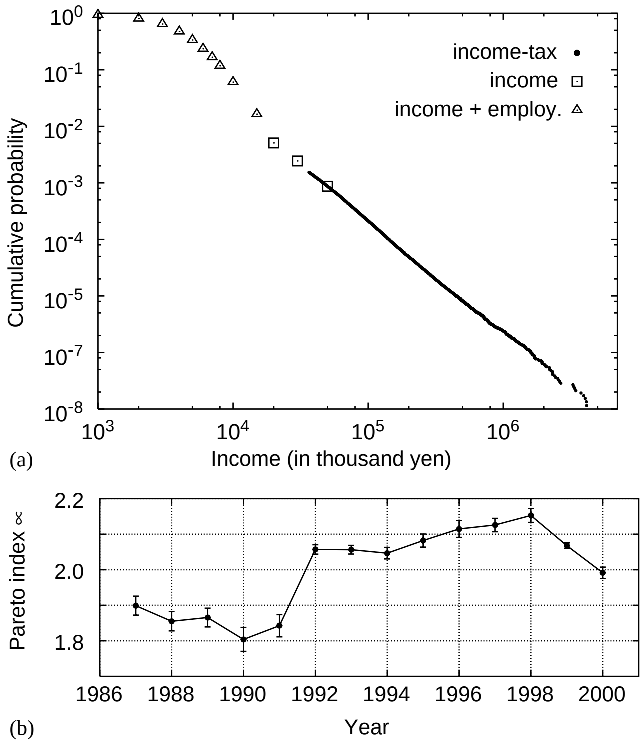

Fig. 1. Personal income in Japan. (a) Cumulative probability distribution of personal income from low to high income range in the year 2000. A data-point represents the probability (vertical axis) that a person has income equal to or more than the income of the horizontal value. Three datasets available from the Japanese National Tax Administration (NTA) were used (see footnote 1 for details). (b) Annual change of Pareto index μ from the year 1987 to 2000. The complete list of income tax data in each year was used. Excluding the top 0.1% and bottom 10%, samples equally spaced in logarithm of rank were plotted, from which slopes were estimated by least-square-fit. Error bars shown are standard error (90% level) of the estimate μ (dots).

economic status. He found that the high-income distribution follows a power-law: the probability that a given individual has income equal to or greater than x, denoted by $P_>(x)$, which obeys

$$P_>(x) \propto x^{-\mu} , \tag{1}$$

with a constant μ called the Pareto index. This phenomenon, now known as a classic example of fractals, has been observed [2–7] in many different countries, where μ varies typically around 2 reflecting economic conditions.

Recent high-quality digitized data prove that the law holds for the high income range, often with remarkable accuracy, and allows a precise estimate of Pareto index over years. Fig. 1a shows the distribution of Japanese personal income in the year 2000, derived from available data of the Japanese National Tax Administration

(NTA)[1] (corresponding to UK Board of Inland Revenue). Power-law behavior is a salient feature characterizing the high income range at nearly three orders of magnitude.

Understanding the origin of the law has importance in economics because of its linkage with consumption, business cycles, and other macro-economic activities. It is also important for assessments of economic inequality [4]. Many researchers, recently including those in non-equilibrium statistical physics, have proposed models for power-law [8–14]. Some theories were based on multiplicative stochastic processes [15,16]. A classic theory by Gibrat [8] assumed that personal income depends on a number of causes, each of which has a proportional effect that is, independent of the proportional effects of the others and also of initial income (law of proportionate effect). This theory, basically a random walk in logarithmic scale of income, predicts log-normal distribution of income with a Gaussian growth rate, both in disagreement with actual data for high income (see also Ref. [17]). One could introduce to the process a boundary effect that income should not be less than a value and derive a power-law distribution [9,10]. Another approach is to construct a simple but minimal economic model in a network of wealth [13]. Actually many kinds of proposed scenarios [16] have predicted a power-law distribution as a static snapshot. However, in order to test models, it has been highly desirable to have direct observation of the dynamical process of growth and fluctuations of personal income.

For that purpose, we employ Japanese income tax data covering most of the power-law region in Fig. 1a. It is an exhaustive list of all taxpayers—full names, addresses and tax amounts—who paid 10 million yen or more a year through tax offices of the Japanese National Tax Administration (NTA). The data were gathered from every NTA office. Fig. 1b plots Pareto indices, estimated from this income tax data, from 1987 to 2000. μ changes annually around 2 with an abrupt jump between 1991 and 1992. Before these years, the Japanese economy experienced an abnormal rise in prices in the risky assets of lands and shares due to speculative investment (“bubble”), after which those prices fell rapidly (“burst”). We examined a relatively stable period in the economy, namely, 1997 and 1998. The complete datasets of 93,394 persons in 1997 and 84,571 in 1998 were used. Identification of individuals listed in both years was done if and only if his/her full name uniquely and exactly matched in both years with the same address (zip-code). Duplicate matches were only a few cases that were discarded. We assumed that the change in address and name was negligible as a fraction of the

[1] The cumulative probability distribution of personal income is based on three datasets in the year 2000 available from the Japanese National Tax Administration (NTA). (i) Income tax data (dots) is from an exhaustive list of all taxpayers, about 80,000, who paid income tax of 10 million yen or more. Tax value was converted to income uniformly by the same proportionality following the previous work [5]. (ii) Income data (squares) is a coarsely tabulated data for all the persons, about 7,273,000, who filed tax returns. (iii) Employment income data are from a sample survey for the salaried workers in private enterprises, about 44,940,000. Under the Japanese taxation, all persons with income exceeding 20 million yen have an obligation to a file final declaration to the NTA each year. Thus dataset (ii) includes all persons listed in (i), so we have a reliable profile in the high income range ($>$ 20 million yen). For lower income, an upper-bound estimate (triangles) was given by overlapping datasets (ii) and (iii), and this was found to be relatively good [6].

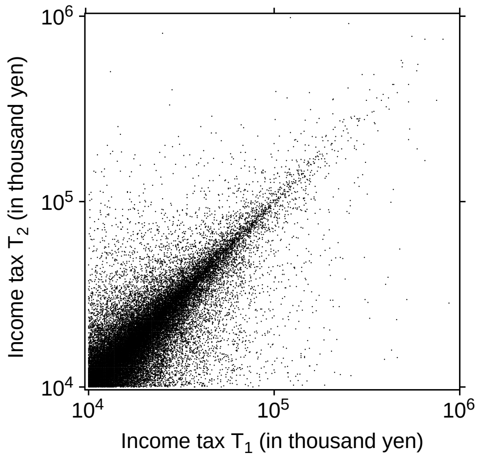

Fig. 2. Scatter-plot of all individuals whose income tax exceeds 10 million yen in both 1997 and 1998. These points (52,902) were identified from the complete list of high-income taxpayers in 1997 (93,394) and in 1998 (84,571) (numbers in parentheses), with income taxes T_1 and T_2 in each year. A few points with T_1 and/or T_2 exceeding 10^6 exist but are not shown here.

whole. The number of the common set of those appearing in the two consecutive years was 52,902. The rest of the persons in 1997 and 1998 can thus be regarded as those disappearing from or novel in the list.

The common set is shown by the scatter plot in Fig. 2, where each point represents a person who paid income tax of T_1 in 1997 and T_2 in 1998 (both in units of thousand yen). This represents the joint distribution $P_{12}(T_1, T_2)$. The plot is consistent with approximate time-reversal symmetry in the sense that the joint distribution is invariant under the exchange of values T_1 and T_2.

The quantity of our concern is the annual change of individual income-tax, or growth. Growth rate is defined as $R = T_2/T_1$. It is customary to use the logarithm of R, $r \equiv \log_{10} R$. We examine the probability density for the growth rate $P(r|T_1)$ on the condition that the income T_1 in the initial year is fixed. The result is shown in Fig. 3. Here we divide the range of T_1 into logarithmically equal bins as $T_1 \in [10^{4+0.2(n-1)}, 10^{4+0.2n}]$ with $n = 1, \ldots, 5$. For each bin, the probability density for r was calculated. As shown in the figure, different plots for n collapse onto each other. This means that the distribution for growth rate r is statistically independent of the initial value of T_1. In a mathematical notation, we found that

$$P_{1R}(T_1, R) = P_1(T_1) P_R(R) , \tag{2}$$

where P_{1R} is the joint distribution for T_1 and R, P_1 and P_R are the distributions for T_1 and R, respectively.

This “universal” distribution for the growth rate has a skewed and heavy-tailed shape with a peak at $R = 1$. How is such a functional form consistent with the approximate time-reversal symmetry shown in Fig. 2? The answer to this question leads us to an important bridge from the fluctuations of growth rates to the Pareto law as follows.

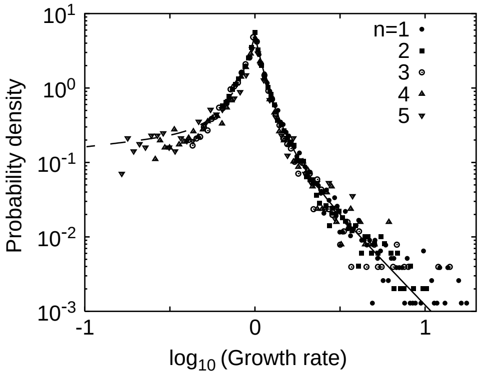

Fig. 3. Probability density $P(r|T_1)$ of growth rate $r \equiv \log_{10}(T_2/T_1)$ from 1997 to 1998. Note that due to the limit $T_1 > 10^4$ (in thousand yen), the data for large negative growth, $r < 4 - \log_{10} T_1$, are not available. Different bins of initial income-tax with equal size in logarithmic scale were taken as $T_1 \in [10^{4+0.2(n-1)}, 10^{4+0.2n}]$ $(n = 1, \ldots, 5)$ to plot probability densities separately for each such bin. The solid line in the portion of positive growth $(r > 0)$ is an analytic fit. The dashed line $(r < 0)$ on the other side is calculated from the fit by the relation given in (4).

The time-reversal symmetry (Fig. 2) claims that $P_{12}(T_1, T_2) = P_{12}(T_2, T_1)$. One can easily see that under the variable transformation from (T_1, T_2) to (T_1, R), the equality $P_{1R}(T_1, R) = T_1 P(T_1, T_2)$ holds. This equality, together with the time-reversal symmetry and the statistical independence of Eq. (2), leads us to the relation:

$$P_1(T_2)/P_1(T_1) = R P_R(R)/P_R(1/R) . \tag{3}$$

The left-hand side is a function of T_1 and T_2, while the right-hand side is a function of ratio R only. We can then conclude that distribution P_1 obeys a power-law: $P_1(x) \propto x^{-(\mu+1)}$, whose integral form gives the expression of Eq. (1). Thus the independence in the growth rate of the past value and the time-reversal symmetry requires the Pareto law.

In addition, we have a scaling relation following immediately from the above relation Eq. (3) and Eq. (1):

$$P_R(R) = R^{-(\mu+2)} P_R(1/R) . \tag{4}$$

This equation relates the positive and negative growth rates through the Pareto index μ. In Fig. 3, we fitted $P_R(R)$ for the region of positive growth $r > 0$ with an analytic function and then plotted its counter part for negative growth rate $r < 0$ derived from the scaling relation, Eq. (4). The result fits the data in the region quite satisfactorily.

In summary, the statistical independence of growth rate, the approximate time-reversal symmetry and the power-law are consistent with one another. According to a sample survey by NTA on income earners with total income exceeding 50 million yen and on sources of earning, their sources are employment income, income from real estate, and capital gains from lands and stock shares. In fraction of income amount, capital gains from risky assets considerably exceeded non-risky income sources. It would be expected that asymmetric behavior of price fluctuations in those risky assets and the

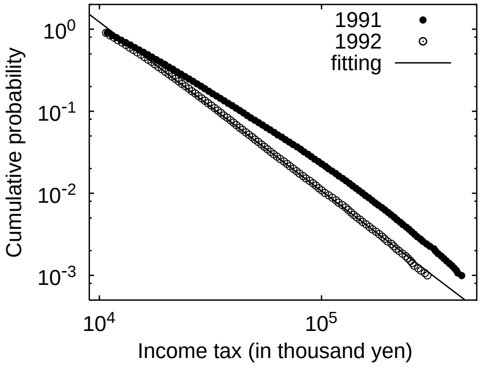

Fig. 4. Cumulative probability distributions of income tax in 1991 and 1992. The Pareto index for 1992 data was estimated by excluding the top 0.1% and bottom 10%, sampling equally in logarithmic scale, and estimating by least-square-fit, which is the fitted line ($\mu = 2.057 \pm 0.005$).

accompanying increase in high-income persons cause the breakdown of time-reversal symmetry, which necessarily invalidates Pareto's law. This was actually the case in the "bubble" phase of the Japanese economy, during which the prices of risky assets, especially of lands, rose abnormally relative to their fundamental values. Fig. 4 shows the cumulative distributions of income tax in 1991 (peak of speculative bubble) and 1992. One can observe that the 1991 data cannot be fitted by Pareto's law in the entire range of high income, compared to the 1992 data.

Our findings in this work shall serve as an empirical test for models of personal income and wealth, where people make choices among assets with different risks and returns, with changing degrees of freedom. Personal income is only a single example of such systems, but other systems composed of economic agents [18], including companies, institutions and nations, might be worth examining from a new look. Indeed, a comparison with and similar analysis in company growth, which has been studied extensively [19–24], would be an interesting subject, where the Zipf law ($\mu = 1$) is observed.

Acknowledgements

Supported in part by grants from the Telecommunications Advancement Organization in Japan. We thank M. Ojima and Teikoku Databank, Ltd. for kindly providing high-quality data, H. Fujisaka, H. Takayasu and H. Tomita for reviewing the paper. Conversations with J.-P. Bouchaud, M. Gallegati, J. Sutton were helpful.

References

[1] V. Pareto, Le Cours d'Économie Politique, Macmillan, London, 1897.
[2] C. Gini, Biblioteca delli'ecoomista 20 (1922).

[3] W.W. Badger, in: B.J. West (Ed.), Mathematical Models as a Tool for the Social Science, Gordon and Breach, New York, 1980, pp. 87–120.
[4] D.G. Champernowne, F.A. Cowell, Economic Inequality and Income Distribution, Cambridge University Press, Cambridge, 1998.
[5] H. Aoyama, W. Souma, Y. Nagahara, M.P. Okazaki, H. Takayasu, M. Takayasu, Fractals 8 (2000) 293.
[6] W. Souma, Fractals 9 (2001) 463.
[7] A. Drăgulescu, V.M. Yakovenko, Physica A 299 (2001) 213.
[8] R. Gibrat, Les Inégalités Économiques, Sirey, Paris, 1932.
[9] D.G. Champernowne, Econometric J. 63 (1953) 318.
[10] B.B. Mandelbrot, Econometrica 29 (1961) 517.
[11] E.W. Montroll, M.F. Shlesinger, J. Stat. Phys. 32 (1983) 209.
[12] M. Levy, S. Solomon, Int. J. Mod. Phys. C 7 (1996) 595.
[13] J.-P. Bouchaud, M. Mézard, Physica A 282 (2000) 536.
[14] S. Solomon, P. Richmond, in: A. Kirman, J.-B. Zimmermann (Eds.), Economics with Heterogeneous Interacting Agents, Lecture notes in Economics and Mathematical Systems, Vol. 503, Springer, Berlin, 2001, pp. 141–145.
[15] H. Kesten, Acta Math. 131 (1973) 207.
[16] D. Sornette, Critical Phenomena in Natural Sciences, Springer, Heidelberg, 2000.
[17] J. Sutton, J. Econ. Lit. 35 (1997) 40.
[18] M. Aoki, Modeling Aggregate Behaviour and Fluctuations in Economics, Cambridge University Press, New York, 2002.
[19] Y. Ijiri, H.A. Simon, Skew Distributions and the Sizes of Business Firms, North-Holland, New York, 1977.
[20] M.H.R. Stanley, L.A.N. Amaral, S.V. Buldyrev, S. Havlin, H. Leschhorn, P. Maass, M.A. Salinger, H.E. Stanley, Nature 379 (1996) 804.
[21] L.A.N. Amaral, S.V. Buldyrev, S. Havlin, M.A. Salinger, H.E. Stanley, Phys. Rev. Lett. 80 (1998) 1385.
[22] K. Okuyama, M. Takayasu, H. Takayasu, Physica A 269 (1999) 125.
[23] M. Mizuno, M. Katori, H. Takayasu, M. Takayasu, in: H. Takayasu (Ed.), Empirical Science of Financial Fluctuations: The Advent of Econophysics, Springer, Tokyo, 2002, pp. 321–330.
[24] R.L. Axtell, Science 293 (2001) 1818.