Pareto-Zipf, Gibrat's Laws, Detailed-Balance and their Breakdown

Yoshi Fujiwara

ATR Network Informatics Laboratories, Kyoto 619-0288, Japan
yfujiwar@atr.jp

Summary. By employing exhaustive lists of personal income and firms, we show that the upper-tail of the distribution of income and firm size has power-law (Pareto-Zipf law), and that in this region their growth rate is independent of the initial value of income or size (Gibrat's law of proportionate effect). In addition, detailed balance holds in the power-law region; the empirical probability for an individual (a firm) to change its income (size) from a value to another is statistically the same as that for its reverse process in the ensemble. We prove that Pareto-Zipf law follows from Gibrat's law under the condition of detailed balance. We also show that the distribution of growth rate possesses a non-trivial relation between the positive and negative sides of the distribution, through the value of Pareto index, as is confirmed empirically. Furthermore, we also show that these properties break down in the non power-law region of distribution, and can possibly do so temporally according to drastic change in financial or real economy.

Key words. personal income, firm size, Pareto-Zipf distribution, Gibrat law, detailed-balance

1 Introduction

Flow and stock are 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 respectively. Personal income and wealth can be regarded as the flow and stock of each *household* in a giant dynamical network of people, which is open to various economic activities. The same is true for *firms*.

High-income distribution follows a power-law: the probability $P_>(x)$ that a given individual has income equal to, or greater than x, obeys

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

with a constant μ called the Pareto index. This phenomenon, now known as Pareto law, has been observed [5, 7, 2, 17, 8, 9] in different countries (see

Fig. 1 (a) for japanese data). On the other hand, low and middle-income distribution has been considered to obey log-normal distribution, Boltzman distribution or other functional form (see Clementi and Gallegati, Souma and Nirei, Willis, Yakovenko in this workshop).

A similar distribution has also been observed for firm size [4] (see Fig. 1 (b) for european data). μ is typically around 2 for personal income and around 1 for firm size distribution[1]. The latter is often refered to as Zipf law. We call the distribution (1) Pareto-Zipf law in this paper.

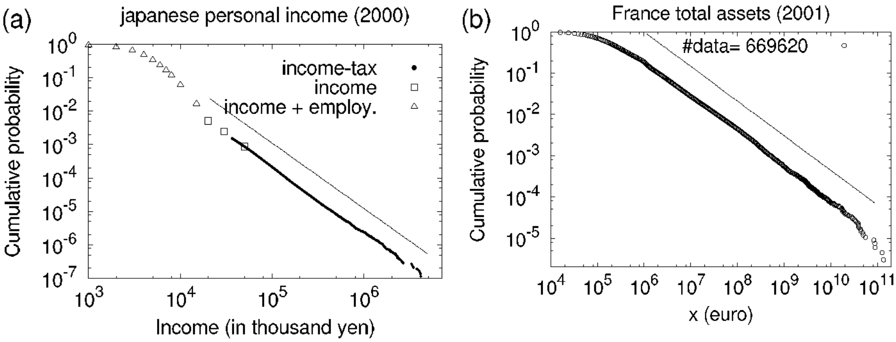

Fig. 1. (a) Cumulative probability distribution of japanese personal income in the year 2000. The line is simply a guide for eyes with $\mu = 1.96$ in (1). Note that the dots are income tax data of about 80,000 taxpayers. (See [17][9] for the details). (b) Cumulative probability distribution of firm size (total-assets) in France in the year 2001. Data consist of 669620 firms, which are exhaustive in the sense that firms exceeding a threshold are all listed. The line corresponds to $\mu = 0.84$ (See also [11]).

Understanding the origin of the law has importance in economics because of its linkage with consumption, business cycles, and other macro-economic activities. Also note that even if the range for which (1) is valid is a few percent in the upper tail of the distribution, it is often observed that such a small fraction of individuals (firms) occupies a large amount of total sum of income (size). Small idiosyncratic shock can make a considerable macro-economic impact.

Many researchers, recently including those in non-equilibrium statistical physics, have proposed models for power-law [7, 14, 15, 13, 6, 16]. Actually many kinds of proposed scenarios have predicted a power-law distribution as a static snapshot. However, in order to test models, it is highly desirable to have direct observation of the dynamical process of growth and fluctuations of personal income (or firm size). This is what we address in this paper.

[1] It would be interesting to note that debt distribution of *bankrupted* firms obeys Zipf law in accordance to firm size distribution [12].

2 Growth and Fluctuations

For the study of personal income, we employ japanese income-tax data, which is an exhaustive list of all taxpayers who paid 10 million yen (approx. 10 thousand euros/dollars) or more in a year. It is considered that the taxpayers cover most of the power-law region in Fig. 1 (a). Distribution of personal income has been studied by using the data [2, 17]. Furthermore, growth and fluctuations of *each* individual can be examined [9]. In fact, we examined a relatively stable period in japanese economy, namely 1997 and 1998.

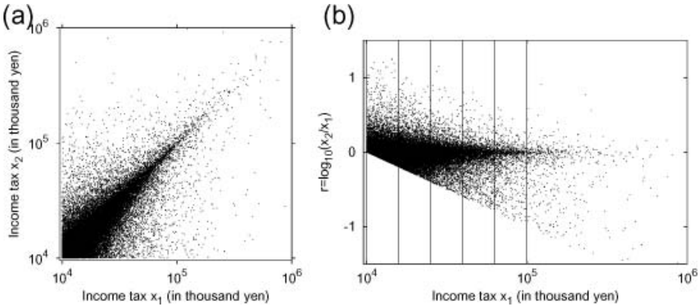

Fig. 2. (a) Scatter-plot of all individuals whose income tax exceeds 10 million yen in both 1997 and 1998. These points (52,902) were identified from the complete list of high-income taxpayers in 1997 and 1998, with income taxes x_1 and x_2 in each year. (b) The same as (a) with vertical axis for $r = \log_{10}(x_2/x_1)$. The segments are bins for $x_1 \in [10^{4+0.2(n-1)}, 10^{4+0.2n}]$ $(n = 1, \cdots, 5)$.

In the scatter plot in Fig. 2 (a) each point represents a person who is high-income taxpayers in both of the years, 1997 and 1998. The plot represents the joint distribution $P_{12}(x_1, x_2)$. The plot is consistent with *detailed-balance* in the sense that the joint distribution is invariant under the exchange of values x_1 and x_2, i.e. $P_{12}(x_1, x_2) = P_{12}(x_2, x_1)$[2]. Detailed-balance means that the empirical probability for an individual to change its income from a value to another is statistically the same as that for its reverse process in the ensemble.

Our concern is the annual change of individual income-tax. Growth rate is defined by $R = x_2/x_1$. It is customary to use its logarithm, $r \equiv \log_{10} R$. We examine the probability density for the growth rate $P(r|x_1)$ on the condition that the income x_1 in the initial year is fixed (see Fig. 2 (b)). The result shows that the distribution for growth rate r is *statistically independent* of the value of x_1, as shown in Fig. 3. This is known as *law of proportionate effect* or Gibrat's law (see [18]).

[2] Actually we can make a direct statistical test for the symmetry in the two arguments of $P_{12}(x_1, x_2)$. This can be done by two-dimensional Kolmogorov-Smirnov test, which is not widely known but has been developed by astrophysicists to test uniform distribution of galaxies appearing in the sky (see references in [11]) .

The phenomenological properties (A) detailed-balance, (B) Pareto-Zipf law, and (C) Gibrat's law are observed for firm size as well as for personal income. See [11] for such a study of large firms in European countries.

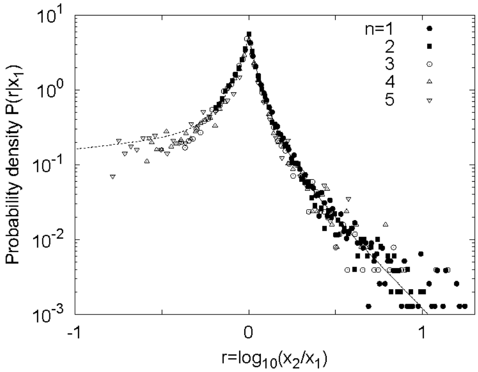

Fig. 3. Probability density $P(r|x_1)$ of growth rate $r \equiv \log_{10}(x_2/x_1)$ from 1997 to 1998. Note that due to the limit $x_1 > 10^4$ (in thousand yen), the data for large negative growth, $r < 4 - \log_{10} x_1$, are not available. Different bins of initial income-tax with equal size in logarithmic scale were taken as $x_1 \in [10^{4+0.2(n-1)}, 10^{4+0.2n}]$ ($n = 1, \cdots, 5$) 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 by the relation in (7).

The probability distribution for the growth rate, such as the one observed in Fig. 3, contains information of dynamics. One can notice that it has a skewed and heavy-tailed shape with a peak at $R = 1$. How is such a functional form consistent with the detailed-balance shown in Fig. 2? And how these phenomenological facts are consistent with Pareto's law in Fig. 1? Answers to these questions are given in the next section.

3 Pareto-Zipf and Gibrat under detailed balance

Let x be a personal income or a firm size, and let its values at two successive points in time (i.e., two consecutive years) be denoted by x_1 and x_2. We denote the joint probability distribution for the variables x_1 and x_2 by $P_{12}(x_1, x_2)$. We define conditional probability, $P_{1R}(x_1, x_2/x_1) = x_1 P_{12}(x_1, x_2)$, where $P_1(x_1)$ is marginal, i.e., $P_1(x_1) = \int_0^\infty P_{1R}(x_1, R)dR = \int_0^\infty P_{12}(x_1, x_2)dx_2$.

The phenomenological properties can be summarized as follows.

(A) *Detailed Balance*:

$$P_{12}(x_1, x_2) = P_{12}(x_2, x_1). \tag{2}$$

(B) *Pareto-Zipf's law*:

$$P_1(x) \propto x^{-\mu-1}, \tag{3}$$

for $x \to \infty$ with $\mu > 0$.

(C) *Gibrat's law*: The conditional probability $Q(R \mid x)$ is independent of x:

$$Q(R \mid x) = Q(R). \tag{4}$$

We note here that this holds only for x larger than a certain value. All the arguments below is restricted in this region.

Now we prove that the properties (A) and (C) lead to (B). Under the change of variables from (x_1, x_2) to (x_1, R), since $P_{12}(x_1, x_2) = (1/x_1)P_{1R}(x_1, R)$, one can easily see that $P_{1R}(x_1, R) = (1/R)P_{1R}(Rx_1, R^{-1})$. It immediately follows from the definition of $Q(R \mid x)$ that

$$\frac{Q(R^{-1} \mid x_2)}{Q(R \mid x_1)} = R\frac{P_1(x_1)}{P_1(x_2)}. \tag{5}$$

This equation is thus equivalent to detailed-balance condition.

If Gibrat's law holds, $Q(R \mid x) = Q(R)$, then

$$\frac{P_1(x_1)}{P_1(x_2)} = \frac{1}{R}\frac{Q(R^{-1})}{Q(R)}. \tag{6}$$

Note that while the left-hand side of (6) is a function of x_1 and $x_2 = Rx_1$, the right-hand side is a function of ratio R only. It can be easily shown that the equality is satisfied by and only by a power-law function (3)[3].

As a bonus, by inserting (3) into (6), we have a non-trivial relation:

$$Q(R) = R^{-\mu-2}Q(R^{-1}), \tag{7}$$

which relates the positive and negative growth rates, $R > 1$ and $R < 1$, through the Pareto index μ.

One can also show that $Q(R)$ has a cusp at $R = 1$; $Q'(R)$ is discontinuous at $R = 1$. Explicitly, $[Q^{+\prime}(1) + Q^{-\prime}(1)]/Q(1) = -\mu - 2$, where we denote the right and left-derivative of $Q(R)$ at $R = 1$ by the signs $+$ and $-$ in the superscript, respectively. This relation states that the shape of cusp in $Q(R)$ at $R = 1$ is determined by the Pareto index μ.

Summarizing this section, we have proved that under the condition of detailed balance (A), Gibrat's law (C) implies Pareto-Zipf law (B). The opposite (B) $\to$ (C) is not true. See [11] for several kinematic relations, and also [9, 3, 10] for the validity of our findings in personal income and firms data.

[3] Expand (6) with respect to R around $R = 1$ and to equate the first-order term to zero, which gives an ordinary differential equation for $P_1(x)$.

4 Temporal breakdown of the laws

In economically unstable period such as "bubble" and "crash", income distribution deviates from power-law as shown in Fig. 4 for japanese case. The year 1991 coincides with the peak of speculative bubble of land price. One can observe that the 1991 data cannot be fitted by Pareto's law in the entire range of high income, while one year later the distribution went back to power-law.

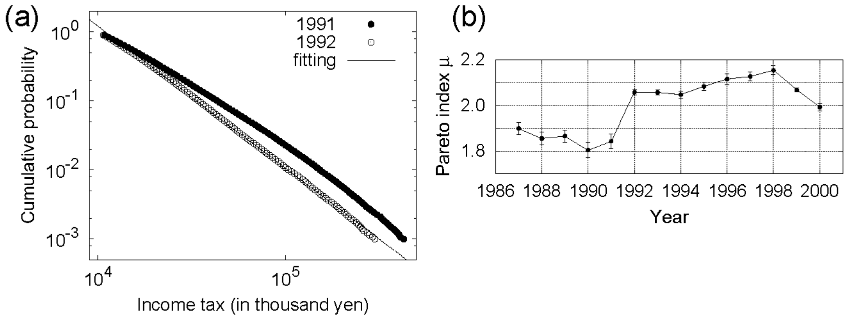

Fig. 4. (a) Cumulative probability distributions of income tax in 1991 and 1992. The fitted line is for $\mu = 2.057$. Note that the distribution does not obey power-law in 1991. (b) Annual change of Pareto index μ from the year 1987 to 2000. The abrupt change from 1991 to 1992 corresponds to abnormal rise and collapse of risky assets prices. (See [9] for estimation).

Sample survey on income earners provides information about income-sources. Picking those persons with total income exceeding 50 million yen (who are necessarily included in the exhaustive list described so far), it can be observed that in terms of the numbers chief income-sources are employment income, rental of real estate, and capital gains from lands and stock shares (Fig. 5 (a)), and that in terms of the amounts contribution comes largely from capital gains from lands and stocks (Fig. 5 (b)). It is expected that asymmetric behavior of price fluctuations in those risky assets and the accompanying increase in high-income persons cause the breakdown of detailed-balance and/or the statistical independence, which necessarily invalidates Pareto's law.

This can be verified in Fig. 6 showing the breakdown of Gibrat's law. By statistical test [11], the null hypothesis $P_{12}(x_1, x_2) = P_{12}(x_2, x_1)$ can be rejected for the pair of 1991 and 1992, but cannot be rejected for the pairs of 1992 and 1993, and of 1997 anf 1998 with significance level 95%.

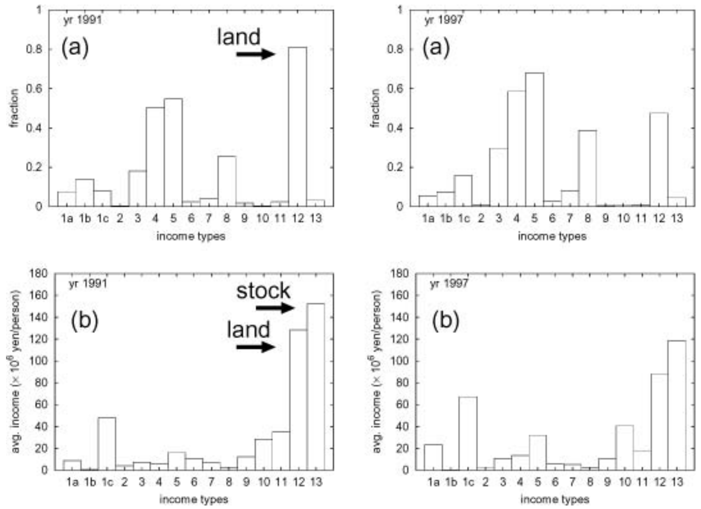

Fig. 5. See the list of income types below. *Left panels*: 1991. *Right panels*: 1992. (a) Fraction of the numbers for income earners with total income exceeding 50 million yen with a particular income type. A person can have more than a single income type. (b) Fraction of the amounts in average over all the earners. The numerals are 1a: business income, 1b: firm income (agricultural), 1c: other operating income (lawyers, doctor, entertainers, etc.), 2: interest income, 3: dividends, 4: rental income (mainly of real estate), 5: wages/salaries, 6: comprehensive capital gains, 7: sporadic income, 8: miscellaneous income (including public pension, etc.), 9: forestry income, 10: retirement income, 11: short-term separate capital gains (selling real estate possessed in 5 years), 12: long-term separate capital gains (selling real estate possessed over 5 years), 13: capital gains from stocks, etc.

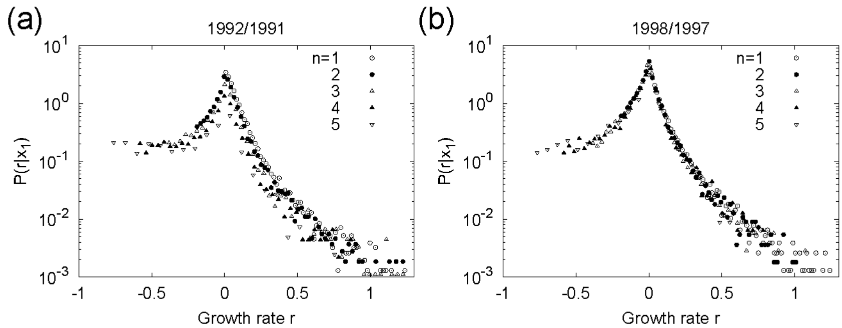

Fig. 6. (a) Probability density $P(r|x_1)$ of growth rate $r \equiv \log_{10}(x_2/x_1)$ from 1991 to 1992. It is obvious that $P(r|x_1)$ depends on x_1, thus breaking Gibrat's law. (b) The same plot for the successive years of 1997 and 1998, for which Gibrat's law holds.

5 Small and midsize firms

According to survey by statistics bureau, the number of japanese companies is approximately 1.6 million in the year 2001. Credit Risk Database (CRD) is a database of about one million japanese small-business firms. Small-business firms have qualitatively different characteristics of firm size growth from those for large firms [1]. The CRD covers the non-power-law regime and the transition region to Pareto-Zipf regime.

Fig. 7 (a)–(b) shows the breakdown of Gibrat's law by depicting the probability density function $P(r|x_1)$. The probability density has explicit dependence on x_1 showing the breakdown of Gibrat's law. In order to quantify the dependence, we examine how the standard deviation of r for each group of firms, whose size is $x_1 \sim dx_1$, scales as x_1 becomes larger. Let the standard deviation of r be denoted by σ. Fig. 7 (c)–(d) shows that σ scales as a function of x_1 ($\sigma \propto x_1^{-\beta}$), but asymptotically approaches non-scaling regime ($\sigma \sim$ const). The breakdown of Gibrat's law in the non-power-law regime and its validity in the power-law regime are consistent with what we showed in [11].

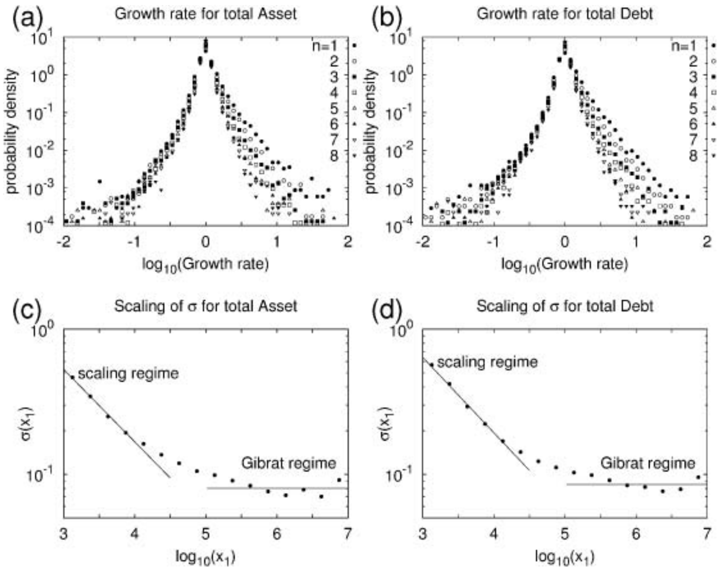

Fig. 7. *Upper panels:* Probability density function $P(r|x_1)$ for logarithmic growth-rate $r = \log_{10}(R)$. For conditioning x_1, we use different bins of initial firm size with equal interval in logarithmic scale as $x_1 \in [10^{4+0.25(n-1)}, 10^{4+0.25n}]$ ($n = 1, \cdots, 8$) for total-assets (a) and total-debts (b) (both in thousand yen). *Lower panels:* Standard deviation σ of r as a function of x_1 for total assets (c) and total debts (d).

6 Conclusion

We have shown the following stylized facts concerning distribution of personal income and firm size, their growth and fluctuations by studying exhaustive lists of high-income persons and firm sizes in Japan and in Europe.

- In power-law regime, detailed-balance and Gibrat's law hold.
- Under the condition of detailed-balance, Gibrat's law implies Pareto's law (but not *vice versa*).
- Growth-rate distribution has a non-trivial relation between its positive and negative growth sides through Pareto index. The distribution must have a cusp whose shape is related to the value of Pareto index.
- Power-law, detailed-balance and Gibrat's law break down according to abrupt change in risky asset market, such as japanese "bubble" collapse of real estate and stock.
- For firm size in non-power-law regime corresponding to small and mid-size firms, Gibrat's law does not hold. Instead, there is a scaling relation of variance in the growth-rates of those firms with respect to firm size, which asymptotically approaches to non-scaling region as firm size comes to power-law regime.

These stylized facts would serve to test models that explain personal income and firm size distributions.

acknowledgement The author would like to thank Hideaki Aoyama, Corrado Di Guilmi, Mauro Gallegati and Wataru Souma for collaboration and discussions. This research was supported in part by the National Institute of Information and Communications Technology. He is grateful to the organizers of workshop for hospitality, including Arnab Chatterjee, Bikas K. Chakrabarti and Sudhakar Yarlagadda.

References

1. Amaral LAN, Buldyrev SV, Havlin S, Leschhorn H, Maass P, Salinger MA, Stanley HE, Stanley MHR (1997) Scaling behavior in economics: I.empirical results for company growth. J Phys (France) I 7:621–633
2. Aoyama H, Souma W, Nagahara Y, Okazaki MP, Takayasu H, Takayasu M (2000) Pareto's law for income of individuals and debt of bankrupt companies. Fractals 8:293–300 (2000)
3. Aoyama H, Souma W, Fujiwara Y (2003) Growth and fluctuations of personal and company's income. Physica A 324:352–358
4. Axtell RL (2001) Zipf distribution of U.S. Firm Sizes. Science 293:1818–1820
5. Badger WW (1980) An entropy-utility model for the size distribution of income. In: West BJ (ed) Mathematical models as a tool for the social science. Gordon and Breach, New York
6. Bouchaud J-P, Mézard M (2000) Wealth condensation in a simple model of economy. Physica A282:536–545

7. Champernowne DG (1953) A model of income distribution. Econometric Journal 63:318–351
8. Drăgulescu A, Yakovenko VM (2001) Exponential and power-law probability distributions of wealth and income in the United Kingdom and the United States. Physica A299:213–221
9. Fujiwara Y, Souma W, Aoyama H, Kaizoji T, Aoki M (2003) Growth and fluctuations of personal income Physica A 321:598–604
10. Fujiwara Y, Aoyama H, Souma W (2003) In: Takayasu H (ed) Proceedings of Second Nikkei Symposium on Econophysics. Springer-Verlag, Tokyo
11. Fujiwara Y, Di Guilmi C, Aoyama H, Gallegati M, Souma W (2004) Do Pareto-Zipf and Gibrat laws hold true? An analysis with European firms. Physica A 335:197–216
12. Fujiwara Y (2004) Zipf Law in Firms Bankruptcy. Physica A 337:219–230.
13. Levy M, Solomon S (1996), Power laws are logarithmic Boltzmann laws. Int J Mod Phys C7:595–601
14. Mandelbrot BB (1961) Stable Paretian random functions and the multiplicative variation of income. Econometrica 29:517–543
15. Montroll EW, Shlesinger MF (1983) Maximum entropy formalism, fractals, scaling phenomena, and 1/f noise. J Stat Phys 32:209–230
16. Solomon S, Richmond P (2001) Stability of Pareto-Zipf Law in Non-stationary Economies. In Kirman A, Zimmermann J-B (eds) Economics with Heterogeneous Interacting Agents. Springer-Verlag, Berlin Heidelberg
17. Souma W (2001) Universal structure of the personal income distribution. Fractals 9:463–470
18. Sutton J (1997) Gibrat's legacy. J Economic Literature 35:40–59