

Review of Income and Wealth
Series 53, Number 3, September 2007

A TWO FACTOR MODEL OF INCOME DISTRIBUTION DYNAMICS

BY MAKOTO NIREI*

Carleton University

AND

WATARU SOUMA

National Institute of Information and Communications Technology

This paper analyzes empirical income distributions and proposes a simple stochastic model to explain the stationary distribution and deviations from it. Using the individual tax returns data in the U.S. and Japan for 40 years, we first summarize the shape of the income distribution by an exponential decay up to about the 90th percentile and a power decay for the top 1 percent. We then propose a minimal stochastic process of labor and asset income to reproduce the empirical characteristics. In particular, the Pareto exponent is derived analytically and matched with empirical statistics.

1. INTRODUCTION

This paper characterizes the historical shape of income distributions in the U.S. and Japan and proposes a simple two-factor model to reproduce the stationary distribution and deviations from it. Our goal is to explain the parameters of the income distribution by a set of fundamental economic parameters.

Our research is motivated by the recent development on the analysis of power-law tail distributions. Since Pareto, it has long been known that the tail of distribution of income or wealth w universally obeys a power-law distribution $w^{-\alpha}$ for a constant α around 1.5–2.5. A multiplicative process of wealth accumulation has been a standard explanation for the heavy tail. This explanation makes good economic sense because the rate of return for asset is a stationary process. Gibrat's celebrated "law of proportionate effect" first embodied the idea that the multiplicative process generates a lognormal evolution of the distribution (see Sutton (1997) for a review). Kalecki (1945) observed, however, that the variance of empirical log income does not grow linearly in time such as in the lognormal process. A successive surge of research, notably by Champernowne (1953), Simon (1955), Rutherford (1955), and more recently Reed (2001), showed that a slight modification of the lognormal development generates a power law. The field has been stimulated recently by studies on a reflected multiplicative process (Levy and

Note: We thank Hideaki Aoyama and Yoshi Fujiwara for providing us with the data on Japanese top taxpayers. We have benefited from comments by Samuel Bowles, Dooyne Farmer, Daniel Feenberg, Duncan Foley, Xavier Gabaix, Mauro Gallegati, John Miller, Toshihiko Mukoyama, Emmanuel Saez, Spyros Skouras, D. Eric Smith, Victor Yakovenko, seminar participants at Tokyo Metropolitan University, the editor, and two anonymous referees.

*Correspondence to: Makoto Nirei, Department of Economics, Carleton University, Ottawa, Ontario K1S 5B6, Canada (makoto_nirei@carleton.ca).

© 2007 The Authors

Journal compilation © 2007 International Association for Research in Income and Wealth Published by Blackwell Publishing, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main St, Malden, MA, 02148, USA.

Solomon, 1996) or a closely related Kesten process (Sornette and Cont, 1997; Takayasu *et al.*, 1997) which have revealed the effect of a reflective lower bound on the tail of the stationary distribution.

The reflective barrier model provides economists with an interestingly sharper structure in the multiplicative processes than the previous models do. Gabaix (1999) constructed an economic model of city size distributions by utilizing this structure and suggested its application to income distribution tails. Levy (2003) also derived the power-law distribution of wealth in the same framework. We extend this literature by incorporating the labor wage process in the wealth accumulation process.

This paper builds upon the idea that the savings from labor income serve as the reflective lower bound of asset which accumulates multiplicatively. Our model of personal income consists of an asset accumulation process and a wage process. The asset accumulation process is multiplicative due to the stationary random asset returns. The wage process is assumed additive, reflecting the productivity heterogeneity. We show that this simple process can successfully reproduce the empirical distribution of income. In particular, the model can reproduce the transition of the distribution shape from the middle part which decays exponentially to the tail part which decays in power.

A novelty of the model is that it allows us to analytically derive the tail exponent of the distribution α and provides us with an economic intuition for the determinants of α . The exponent α turns out to be approximately one plus the ratio of savings from labor income to asset income. This formula intuitively captures the dynamics behind the power-law tail. The savings are the inflow of wealth into the tail part, and the asset returns are the inequalizing growth within the tail part. The tail becomes flatter when the asset returns rise, and it becomes steeper when the inflow from the middle part increases. Moreover, if asset income exceeds savings, α is less than two and the variance of the stationary distribution of income diverges to infinity. When savings are more than asset income, α is greater than two and the income distribution has a finite variance. In this sense, the stationary distribution of income qualitatively differs depending on the balance between savings and asset income. GDP statistics show that investment and asset income are about the same historically. This fact is consistent with the empirical estimate of α being about two and with our formula. The formula also shows that α is greater than one, which implies that the mean of the power-law distribution is always finite.

The multiplicative asset process generates a power-law distribution, and the additive wage process generates an exponential decay. This observation motivates our parametrization of empirical income distributions. We use the tax returns data in the U.S. and Japan for 40 years. The data sets have been utilized by Feenberg and Poterba (1993) and Souma (2002) for estimating the Pareto exponents as well as by Piketty and Saez (2003) and Moriguchi and Saez (2006) for estimating the top income shares. We also use the list of top Japanese taxpayers compiled by Fujiwara *et al.* (2003). The list covers about 80,000 individuals. It provides compelling evidence on the long held conjecture that the tail distribution follows a Pareto (or a power-law) distribution. The tax returns data are suitable for the study of the shape of income the distribution, especially in its heavy tail, due to its

complete coverage of all taxpayers, whereas the absence of demographic attributes by the nature of tax returns has kept researchers of the earning distributions from utilizing the data.

The data shows that the distribution of adjusted gross income in tax returns consists of three different parts: the top 1 percent, the middle part above about the 10th percentile, and the lower part which we exclude from our study for not being representative as an income distribution. The distinct pattern of the income distribution in the middle and in the tail has been well noted by researchers such as Singh and Maddala (1976). We parametrize the distribution by an exponential distribution for the middle part and by the Pareto distribution for the upper tail, following Dragulescu and Yakovenko (2001). Although the data does not exclude various alternative parametrizations such as discussed in Bordley *et al.* (1996), we choose this parametrization on the ground that it allows an economic interpretation with our model. We do not simply fit the income distribution by parametric functions or simulated distributions, but we do so only in the perspective of an economic model which enables us to interpret the result.

Ours is a reduced-form model in contrast to the dynamic general equilibrium models which have been carefully calibrated to match empirical income distributions (e.g. Castañeda *et al.*, 2003). At the expense of comprehensiveness, our model provides a sharper analysis of the generative mechanism of the income distribution properties, such as the Pareto exponent of the tail. Our reduced form model emphasizes two facts of an economy: that income is largely derived from labor and capital; and that the return to capital is stationary and stochastic. The two factors lead to a two-story structure of the income distribution. Accordingly, the historical fluctuation in the Gini coefficient is shown to be decomposed into two factors: the real wage growth which affects the middle part; and the real asset return which affects the tail.

The rest of the paper is organized as follows. Section 2 explains the tax returns data set we use, and characterizes the income distribution shape. Section 3 presents the model and derives the Pareto exponent analytically. Section 4 shows the simulation results. Section 5 concludes.

2. EMPIRICAL DISTRIBUTIONS OF INCOME

2.1. Data

We use the individual income tax returns data in the U.S. and Japan from 1960 to 1999. The data records the number of taxpayers for various income strata. The number of bins is 15–30 for the U.S. data and 11–14 for Japanese data. The tabulated data of the U.S. tax returns is provided by Statistics of Income database of the Internal Revenue Service. The Japanese tax returns data is provided by National Tax Agency and tabulated as in Souma (2002).

Additionally, we use Japanese tax returns data for all the individuals who pay income tax of more than 10 million yen. The data on those top taxpayers is made publicly available by the same agency. The length of the list varies around 80,000 from year to year. As far as we know, this is by far the most extensive list available

on the distribution of high income earners.¹ The method for converting the income tax to income is described in Fujiwara *et al.* (2003).

The tax returns data has been utilized in the previous studies (Feenberg and Poterba, 1993; Souma, 2002; Piketty and Saez, 2003; Moriguchi and Saez, 2006). It has a distinctive merit for the study of the shape of income distributions. First, it provides an exact rank-size relation at the threshold of the bins, unlike sampled survey data which always suffers from sampling errors. The sampling error is most severe for the high income range, because income distributions typically have a heavy tail and thus the sampled data contains few observations in this range. Secondly, the range covered by all the bins is relatively wide. The top bin corresponds to at least the 99.9 percentile and sometimes extends to the 99.999 percentile. Besides, our data on Japanese top taxpayers provides us with a decisive testing power for parametric examination of the tail distribution. The tax returns data has a disadvantage as well. Since there is no demographic data associated with the income tax data, we cannot study how the income level is attributed to demographic factors. For this reason our study concentrates on the overall shape of the distribution.

The lowest 10 percent of the tax returns data is excluded from our scope. The adjusted gross income of the taxpayers in the range is considerably different from the actual income due to exemptions. Also, that portion of tax returns data includes taxpayers who have other means to sustain their consumption, such as transfers within families. Hence the gross adjusted income is not considered representative for the distribution of income for this range.

2.2. Stationary Distribution

Figure 1 shows the distribution of adjusted gross income for the U.S. and Japan in 1999. The distribution is cumulated from the top. The top panel shows the entire distribution in log–log scale. The bottom panel shows the distribution up to 200,000 U.S. dollars in semi-log scale. The Japanese income is converted to U.S. dollars by an average exchange rate in 1999. The following two features are evident in the plot, as proposed in Dragulescu and Yakovenko (2001) for British and U.S. income distributions. The top panel shows that the distribution decays in power for the top 1 percent, which can be seen by the linearity in the log–log plot. The Japanese data with top taxpayers clearly shows the good fit with a power-law distribution:

$$(1) \quad \Pr(W > w | W \geq \theta) = (w/\theta)^{-\alpha}$$

where θ is the income level at which the power law starts. We call α the Pareto exponent. A larger α implies that the distribution among the rich is more equal. The bottom panel shows that the middle portion of the distribution is fitted well by an exponential distribution:

$$(2) \quad \Pr(W > w) = e^{-(w-w_0)/\beta}$$

¹The publication of the list started after the second world war, and ceased in 2005 due to privacy concerns.

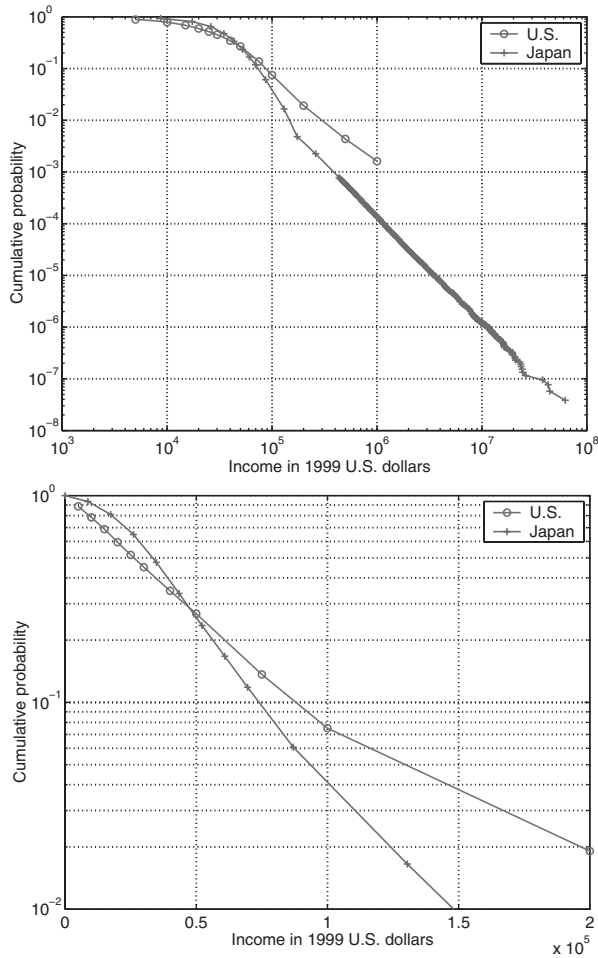


Figure 1. Income Distribution in the U.S. and Japan in 1999

The β represents the standard deviation of the exponential distribution. The w_0 is the income level at which the exponential decay starts, and we interpret it as a subsistence income in the model later.

The pattern that the distribution decays in power in the tail and decays exponentially in the middle is found for all 40 years in the U.S. and Japan. Figure 2 shows the plot for all the years. Figures 3 and 4 show distributions normalized by average income of each year in log scale and in semi-log scale, respectively. We observe that the plots collapse well by the normalization. This suggests that, in the first approximation, we can view the dynamics of the empirical income distribution as a stationary distribution with a multiplicative horizontal shift. The shape parameters α and β do not seem to have a trend over time, while the location of the distribution shifts to the right along with the lower bound income w_0 . This observation corresponds to the fact that the income distribution does not follow a

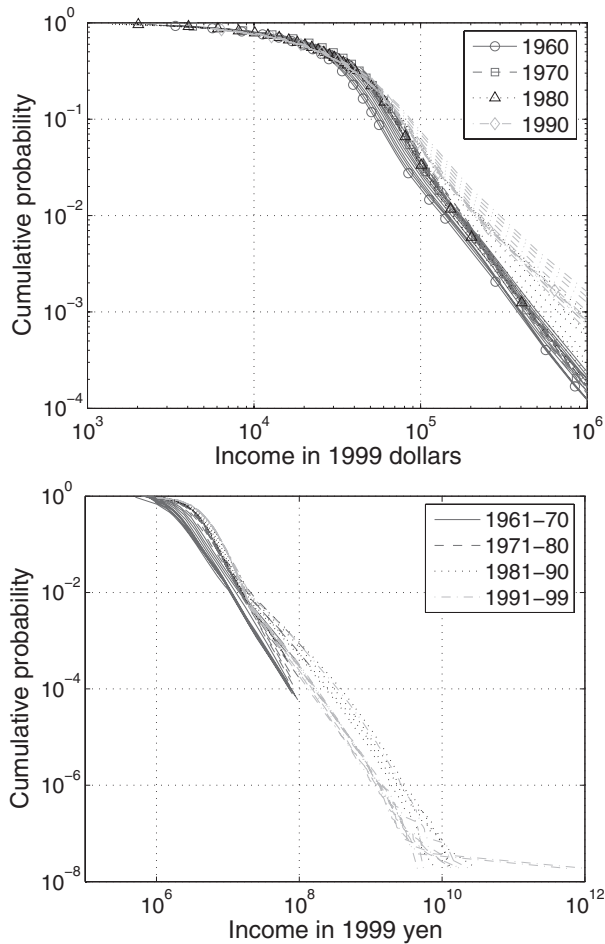


Figure 2. Development of the Income Distribution over time in the U.S. (Top) and Japan (Bottom)

lognormal process in which the log-variance grows linearly in time. The stationary levels of α and β are plotted by dashed lines in Figures 3 and 4. The stationary level of the standard deviation of the exponential distribution β is about 0.85 for the U.S. and 0.6 for Japan. The Pareto exponent α is centered around 2 in both countries. The stationary relative lower bound (w_0 divided by average income) is estimated at the intersection of the dashed line in the semi-log plot and the horizontal line at which the cumulative probability is equal to 1.

The tails of the top taxpayers for 1987–99 in the bottom panel of Figure 3 provide us with compelling evidence that the tail obeys a power-law distribution. We also notice that the tail slope around the 99th percentile extends well to the further tail during the normal years in the second half of the 1990s. However, in the years of volatile financial market such as the bubble and burst of stock and land prices around 1990 in Japan, the slope at the 99th percentile seems flatter than the further tail.

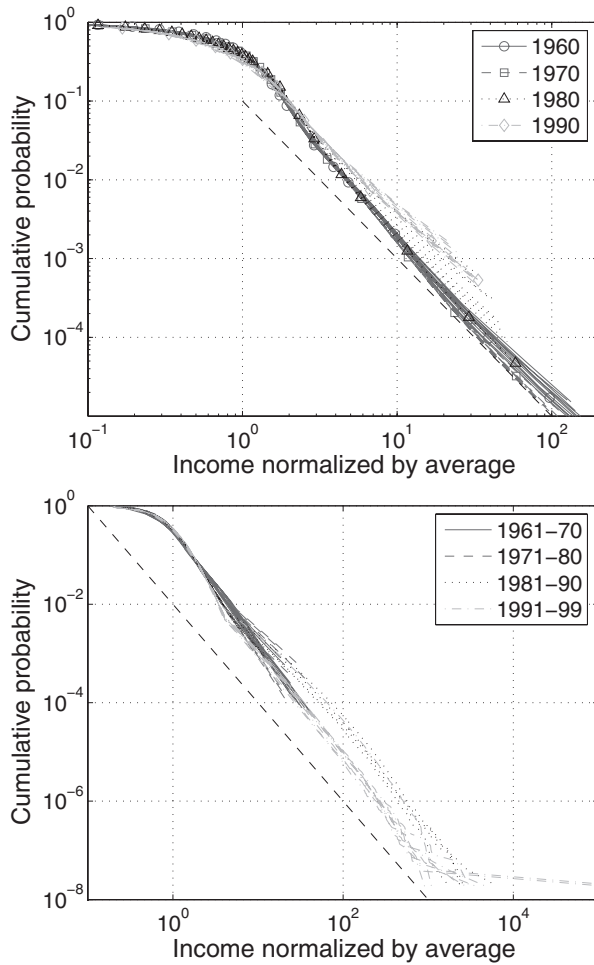


Figure 3. Normalized Income Distributions in Log Scale: U.S. (Top) and Japan (Bottom)

The empirical distributions lead us to the following hypothesis. The stationary distribution of income consists of two parts. The top 1 percent follows a power-law distribution. The middle part that covers the 10th percentile to the 90th percentile of the population is approximated by an exponential distribution. The hypothetical distribution is described by four parameters: the Pareto exponent α , the standard deviation of the exponential distribution β , the lower bound of the exponential distribution w_0 , and the cross-over income level θ at which the two distributions meet.

2.3. Fluctuation of Distribution

The fluctuation of the shape of income distributions can be quantified by estimating α and β for each year. We estimate α by a linear fit to the log-scaled cumulative distribution for the points greater than the 99th percentile, and estimate β by a linear fit to the semi-log plot of the cumulative distribution for the range

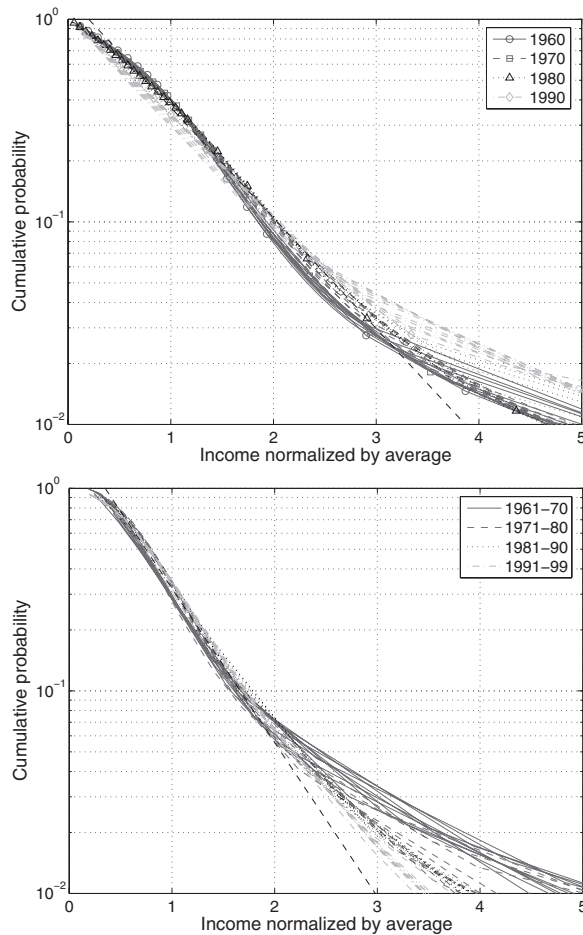


Figure 4. Normalized Income Distributions in Semi-Log Scale: U.S. (Top) and Japan (Bottom)

between the 10th and the 90th percentiles. We are restricted to the linear regression for these estimates, because our data is already binned; otherwise, other estimators such as Hill's estimator for α and the maximum likelihood estimator for β would have been available. The top row of Figure 5 shows the estimated time series of $\hat{\alpha}$ (left panel) and $\hat{\beta}$ (right). The $\hat{\alpha}$ is quite similar to the estimates by Feenberg and Poterba (1993) for the U.S. and by Souma (2002) for Japan. The bottom row shows the estimates of θ (left) and the Gini coefficients (right). The θ is defined as the intercept of the linear fit to the tail in log-scale with the horizontal line at the 99th percentile. The income level is normalized by the average. The Gini coefficient is computed from the whole distribution of the taxpayers.

The fluctuation of α appears negatively correlated with asset returns. We observe that the U.S. exponent stayed high during the stagnation in the 1970s and declined during the stock boom periods in the 80s and 90s. The U.S. also experienced a jump in θ in the late 80s. These behaviors of α and θ match with the rising

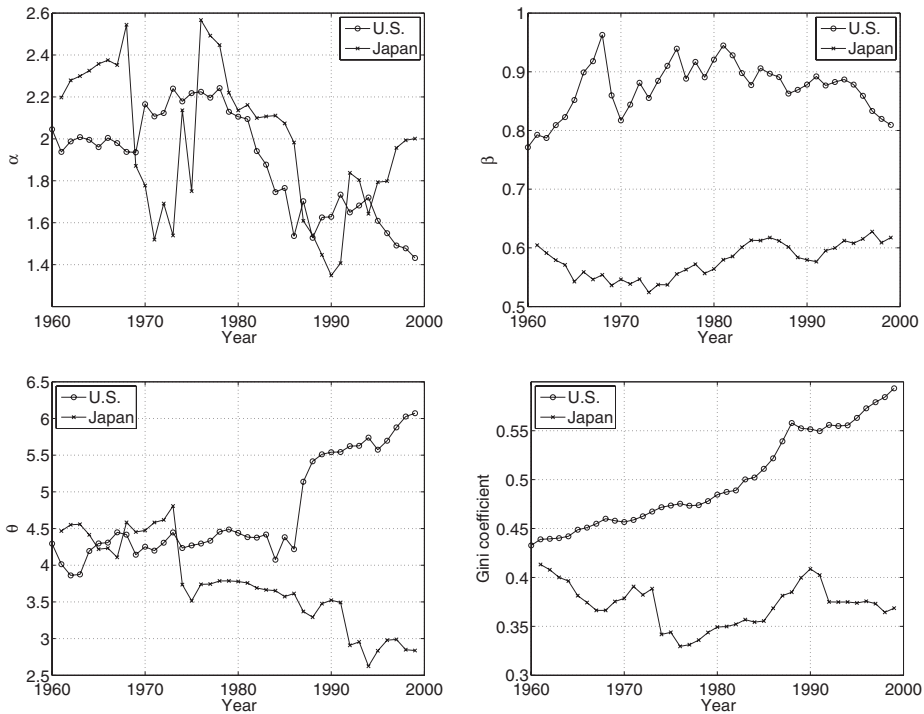


Figure 5. Fluctuations of Distribution Parameters and the Gini Coefficients

income share of top taxpayers found by Piketty and Saez (2003) and correspond to the increasing inequality of wealth during the period (Wolff, 1998). The Japanese exponent also declined during the bull years in the 80s and had a sharp rebound after the crash. The sharp decline of the exponent around 1970 in Japan seems to correspond with the real estate boom in the period. The $\hat{\beta}$ shows a persistent difference between the two countries. The U.S. distribution was less equal than the Japanese in the middle part. In the U.S., the dispersion became larger during the 60s and 70s, and a reversed trend was observed in the latter 90s. In Japan, it declined during the economic high growth years and rebounded at the end of the period in the mid 70s.

We can decompose the Gini coefficients into the effects within the tail α , within the middle β , and of the cross-over point θ . A simple regression confirms this point. Table 1 shows the ordinary least square result for the regression of the first-order difference of the Gini coefficients on the first-order difference of estimated α , β , and θ . The demeaned estimation equation is written as follows when we denote the Gini coefficient and the regression coefficient by $g_{t,i}$ and b_i respectively for year t and country i : $g_{t,i} - g_{t-1,i} = b_{\alpha,i}(\alpha_{t,i} - \alpha_{t-1,i}) + b_{\beta,i}(\beta_{t,i} - \beta_{t-1,i}) + b_{\theta,i}(\theta_{t,i} - \theta_{t-1,i}) + \varepsilon_{t,i}$ for all the years² t and country set $i \in \{\text{U.S., Japan, Pooled}\}$. In the pooled regression, the coefficients are pooled whereas dummy variables are assigned for the two

²For the U.S. data, $t = 1960, 1961, \dots, 1999$, whereas for the Japanese data, $t = 1961, 1962, \dots, 1999$.

TABLE 1
LEAST SQUARES REGRESSION OF DIFFERENCE IN GINI ON DIFFERENCES
OF α , β , AND θ

	α	β	θ	R^2
U.S. (se)	-0.0379 (0.0070)	-0.0013 (0.0188)	0.0174 (0.0030)	0.559
Japan (se)	-0.0332 (0.0034)	0.1419 (0.0654)	0.0259 (0.0031)	0.856
Pooled (se)	-0.0322 (0.0027)	0.0143 (0.0214)	0.0214 (0.0022)	0.794

countries. We confirm significant effects of α and θ on the fluctuation of the Gini coefficients, whereas the effect of β is significant only for Japan.

3. MODEL

In this section we propose a simple stochastic model of the income process. We show that the model can match the empirical distributions of income and wealth. The process is quite parsimonious as a model of personal income which in reality involves many important variables, but its simplicity gives us analytical insight and intuition into the dynamics behind the peculiar shape of the empirical income distributions.

The process consists of an asset accumulation process $a(t)$ and a labor income process $w(t)$. The asset accumulation follows a multiplicative process. Namely, an individual bears an idiosyncratic risk in asset returns. Let γ denote the asset return which is independently and identically distributed across individuals and time. Then:

$$(3) \quad a(t+1) = \gamma(t)a(t) + w(t) - c(t)$$

where $c(t)$ denotes the consumption. We assume that the log return $\log \gamma(t)$ follows a normal distribution with mean y and variance x^2 . This is consistent with the case when the asset return is stationary in continuous time. Let us recall that the mean of a lognormal distribution is $e^{y+x^2/2}$. We may thus interpret y as a risk-free common return and $x^2/2$ as a return for taking idiosyncratic risks. We do not allow the agent to choose the asset portfolio, however. The agent in our model behaves as an entrepreneur who does not diversify her risks.

We assume that the labor income evolves as an additive process with a reflective lower bound as follows.

$$(4) \quad w(t+1) = \begin{cases} uw(t) + s\varepsilon(t)\bar{w}(t) & \text{if } uw(t) + s\varepsilon(t)\bar{w}(t) > \bar{w}(t) \\ \bar{w}(t) & \text{otherwise} \end{cases}$$

$$(5) \quad \bar{w}(t) = v^t \bar{w}(0).$$

The trend growth rate of the labor income is denoted by u . The reflective lower bound $\bar{w}(t)$ grows deterministically at the rate v . The reflective lower bound is interpreted as a subsistence level of income, since the worker below the subsistence income will have to seek another job to sustain the labor force. The idiosyncratic additive shock term $s\varepsilon(t)\bar{w}(t)$ represents the heterogeneity of labor productivity across workers and time. The random shock $\varepsilon(t)$ follows a standard normal distribution, and the constant s determines the standard deviation.

In the simulation, we specify the consumption $c(t)$ as a linear function in asset and disposable income:

$$(6) \quad c(t) = \bar{w}(t) + b(a(t) + w(t) - \bar{w}(t)).$$

Note that the subsistence income $\bar{w}(t)$ determines the minimum level of consumption. The linearity is of course a crude assumption. However, our simulation results are robust to introducing concavity for the consumption function in the lower and middle levels of income. The behavior of the marginal propensity for the top income part, b , does affect the results.

Finally, we define a normalized income $I(t) = \tilde{I}(t)/E[\tilde{I}(t)]$ where:

$$(7) \quad \tilde{I}(t) = w(t) + E[\gamma(t) - 1]a(t).$$

This model offers an analytical characterization of the stationary Pareto exponent. The power-law tail distribution of income parallels the power-law distribution of asset holdings. The power law of wealth is generated by the reflected multiplicative asset accumulation process as in Levy (2003). In our model, the reflective barrier for asset accumulation is the savings from labor income. To see this point, let us normalize equation (3) by the average asset:

$$(8) \quad \hat{a}(t+1) = (\gamma(t)/g(t))\hat{a}(t) + s(t)/(g(t)\langle a(t) \rangle)$$

where $\hat{a}(t) \equiv a(t)/\langle a(t) \rangle$, $g(t)$ is the growth rate of average asset $\langle a(t) \rangle$, and $s(t)$ is savings from labor income $s(t) \equiv w(t) - c(t)$. This is called a stationary Kesten process, when $s(t)$ is positive and independent of $a(t)$ and $\gamma(t)$, and when $E[\gamma] < g$ (which is the case if $\langle s(t) \rangle > 0$). Let us define z as the steady state value of $s(t)/\langle a(t) \rangle$. Also define g as the steady state value of $g(t)$. By applying the formula in Gabaix (1999), the power-law exponent of the stationary distribution of \hat{a} is derived as $\alpha = 1 - \log[1 - z/g]/(x^2/2)$. Since the savings-to-asset ratio z is sufficiently close to zero, we can approximate the expression as $\alpha \approx 1 + z/(gx^2/2)$.

The formula for α provides rich implications on the universality of the power-law exponent of the income and wealth distributions. Recall that $x^2/2$ is the contribution of diffusion to the mean growth rate of assets: $\log E[\gamma] = y + x^2/2$. Then the formula is rewritten as $\alpha \approx 1 + s(t)/(\langle a(t+1) \rangle x^2/2)$. The term $x^2/2 = \log E[\gamma] - y$ is the average excess returns to assets. Thus, $\langle a(t+1) \rangle x^2/2$ expresses the part of asset income that accrues to the risk-taking behavior. Thus α is approximated by one plus the ratio of the savings from labor income to the asset income accrued to risk-taking.

According to this formula, the Pareto exponent becomes two roughly when savings are equal to asset income. The power-law distribution has a diverging variance when α is less than two. (The infinite variance means that a sample variance grows unboundedly as the sample size increases.) Our formula thus implies that the variance of the stationary income distribution is infinite if asset income exceeds savings, whereas it is finite otherwise. In other words, the stationary income distribution is qualitatively different depending on whether asset income exceeds savings. Our data shows that the Pareto exponent historically fluctuates around two. Thus the economy goes back and forth between the two different regimes with respect to the second moment of the income distribution.

Why does α fluctuate? The formula tells us how the asset market fluctuations affect α . The α decreases when the excess asset returns $x^2/2$ increase. Note that this takes effect quickly, because the entrepreneur's income and its distribution is affected directly by the riskier environment. The effect of the savings–asset ratio (s/a) will be much slower, since the convergence to the new stationary distribution involves a slow process of income mobility between the top and the middle parts. However, the savings–asset ratio plays an important role in determining the stationary level of α . To analyze the determinant of the stationary level of α , we need a structural model that endogenously determines the savings rate. It is possible to derive our reduced model of consumer behavior from such a structural model. For example, the linear consumption function is obtained when the preference exhibits constant relative risk aversion (Samuelson, 1969). Further, the saving rate can be endogenized in the Ramsey models. In the Brock–Mirman model as a special case, the saving rate is analytically solved as the capital share of income times the discount factor. The optimal saving rate in a more realistic environment such as a fractional depreciation of capital is also determined in a standard dynamic general equilibrium model. It is interesting to characterize the Pareto exponent in such a structural model, but it has to be deferred to a separate paper.

It is also known that the power–law distribution has a diverging mean when α is less than one. If this is the case, the highest individual income is of the same order as the average income (see Feller, 1966). In our model, the stationary α is always greater than one. Thus our model shows that the income dispersion due to stationary random asset returns does not lead to as much of inequality as $\alpha < 1$ where a single individual occupies a sizable fraction of aggregate income, as long as the savings from labor income grow fast enough to serve as the lower bound of asset accumulation.

GDP statistics support the validity of our formula for the Pareto exponent. For the U.S. data, we estimate the ratio of savings to asset income by private investment divided by the sum of asset income and proprietors' income in the next period. Our formula gives 1.8 for the average Pareto exponent whereas the estimated average Pareto exponent is 2.1. For Japan, the ratio is estimated as the private investment divided by the operating surplus. Our formula provides 2.1, which coincides with the estimated average Pareto exponent.

Let us note that the normalized asset converges to a power-law distribution only if the growth rates of savings and asset income balance. If asset income grows consistently faster than savings, then asset accumulation is a lognormal process

and the estimated α decreases over time. This fact is suggestive for the apparent constant growth in the log variance of the U.S. distribution in these two decades. A structural imbalance in the growth rates of savings and asset income can cause such a constant decrease in α rather than a temporary deviation from the stationary distribution.

It is assumed that labor income follows an additive process with a growing reflective lower bound. We obtain a normalized labor income $\hat{w}(t)$ by dividing the process (4) by average income which grows at v in the stationary growth: $\hat{w}(t+1) = (u/v)\hat{w}(t) + s\varepsilon(t)$. The normalized process is a stationary AR(1) if $u < v$, namely, if the lower bound grows faster than the trend growth. The diffusion coefficient s determines the wage differentiation effect. The diffusion effect balances out the mean-reverting force u/v at the stationary state. Thus the stationary distribution of \hat{w} has a larger variance when u/v is close to one or s is a large.

The stationary distribution of \hat{w} is not generally known. However, it is known that there exists a non-negative random variable $\varepsilon(t)$ for which the process $\hat{w}(t)$ converges to an exponential stationary distribution (Gaver and Lewis, 1980). Let us replace the reflective bound with a nonnegative shock $\varepsilon(t)$. Then $w(t)/\bar{w}(t)$ is non-negative and has a stationary mean $E[\varepsilon(t)]s/(1-u/v)$. If we fit the stationary distribution by an exponential distribution, the standard deviation is determined by the mean and it is positively affected by u/v . Namely, the standard deviation of the middle part distribution widens when the growth gap between the bottom part and the middle part decreases. Numerical simulations confirm this relation, as well as that the exponential distribution approximates the stationary distribution in the middle part.

The mean-zero shock of wage contributes to the average growth when the process is near the reflective lower bound. A job creation–destruction dynamics can be considered to be behind this process. A job disappears when it cannot afford a subsistence level of consumption of a worker, and then a new job is created at the subsistence level by a new industry. Also we can consider a minimum wage scheme, unemployment compensation, and government-funded jobs as the factors behind the reflective lower bound.

The additive process is one of the simplest representations of the wage process which generates the stationary degree of differentiation when average labor income grows. We can consider a productivity shock as the main risk of the wage process. In this sense, this is a wage process of a *job* rather than an individual. An individual labor income process is clearly affected by demographic factors such as experience and the life-cycle as well as sudden probabilistic events such as health and unemployment risks. As long as we are concerned with the overall shape of the distribution, however, the distribution of wages for jobs should coincide with the distribution of individual wages. We can consider that a job is a vehicle of an individual wage which individuals switch due to demographic or individual risk factors. This reshuffle of the job–worker match causes the discrepancy between a process of an individual's wage and of a job's wage, but the distributions of the two processes coincide. However, this interpretation does not match the correlation structure between labor income and asset accumulation specified in our model. This interpretation is useful only to interpret the exponential decline of the wage income.

Asset income is set equal to the asset multiplied by the mean growth rate of assets $E[\gamma(t) - 1]$. In reality, the realization of asset income from asset growth can happen in various ways. It is not immediately clear how the timing of the realization (and taxation) of asset income affects the shape of the distribution. To simplify, we assume a constant rate of asset income realization. The constant is chosen so that all of asset growth is eventually realized and taxed. If the rate is less than $E[\gamma(t) - 1]$, then there is a portion of asset growth which is not counted as income in the long run. If the rate is bigger, then the realized asset income eventually becomes bigger than the accumulated asset itself.

We will not try to match the factor distributions in our simulation, because the notion of labor and capital in our model does not exactly correspond to the notions in the tax account. Our tax returns data does include breakdown of the income into basic factors such as labor and assets. The shape of the distribution of labor income (salary and wage) in the U.S. is quite similar to the total income distribution. In Japan, the labor income distribution matches the lower and middle part and asset income distribution matches the power-law part. In our model, the asset income is broadly associated with any income that is derived from accumulative factor. An important example is human capital. Another example is managers' salary which should correlate strongly with capital size and the capital growth rate. Thus income classified as asset income in our model can be filed as labor income in the taxation system.

4. SIMULATION

4.1. *Stationary Distribution*

The stochastic process proposed in the previous section generates a stationary distribution of normalized income $I(t)$, and it matches the empirical distribution with a plausible set of parameter values. Figure 6 shows the simulated and empirical income distributions for Japan in 1999. The plot also shows the simulated individual wealth $a(t)/E[I(t)]$ along with the empirical household wealth distribution which is taken from the Survey of Consumers in 1999. We also plot Lorenz curves in Figure 7 for both simulated and empirical data. We see that the fit is quite good.

The parameter values for the simulation are obtained as follows. We note that average labor income, average asset income and the lower barrier must grow at the same rate at the balanced-growth path. Thus, we use the time-average growth rate of the nominal income per capita for the parameter of the bound growth rate: $\nu = 1.0673$ for the period 1961–99. The trend growth rate of labor income, u , reflects an automatic growth in nominal wage. We use an average inflation rate for u , which is 1.0422 for the same period.

The log-variance of the asset return x^2 is estimated from top taxpayers data. The data contains the growth rate of individuals who paid income tax of more than 10 million yen in 1997 and 1998. The distribution of the log of income growth rate is symmetric, and its tail follows a power law. The scatter plot of the income for the two years in Fujiwara *et al.* (2003) also exhibits symmetry in the density function ($\Pr(I_{1997}, I_{1998}) = \Pr(I_{1998}, I_{1997})$). These two observations indicate that the tail

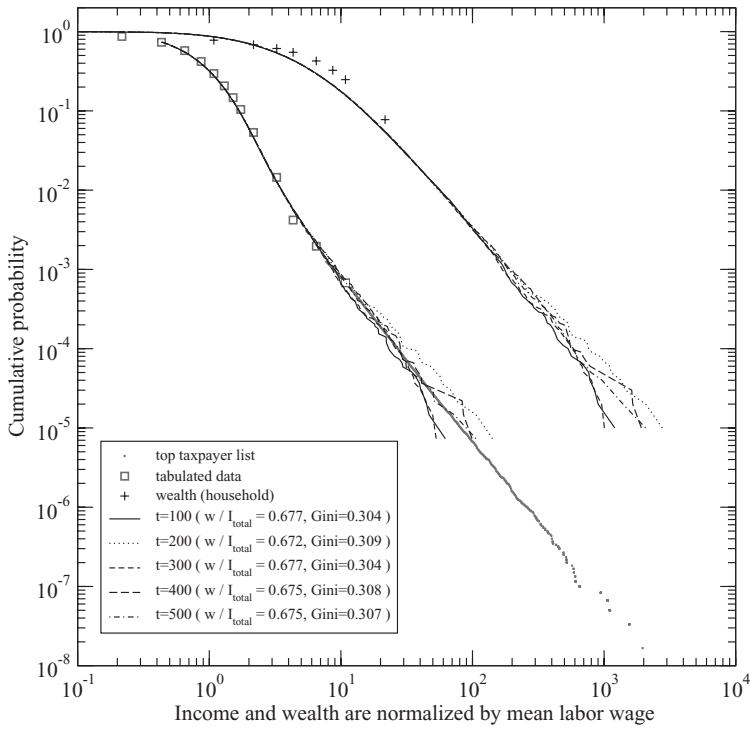


Figure 6. Simulated and Empirical Stationary Distributions of Income and Wealth

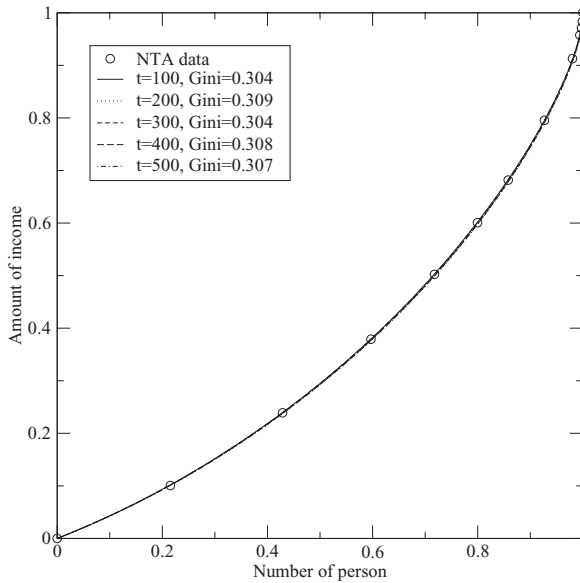


Figure 7. Simulated and Empirical Lorenz Curves

distribution of the income growth rate strongly reflects temporal income such as bequests. Hence we truncate the tail at the point where the power-law takes place, and thus only use the range $1/3 < I_{1998}/I_{1997} < 3$ to estimate the log-variance. In order to eliminate the upper bias due to the censoring of the data at 10 million yen, we use the sample only if the 1997 income tax is greater than 30 million yen. The estimated x is 0.3122.

The log-mean of the asset growth y cannot be estimated by the same growth rate data, since 1997 was not a typical year in financial markets and the average growth rate was negative. In general, asset markets suffer considerable aggregate shocks across years. Thus we estimate y by using a time-average growth rate of the Nikkei average index. The Nikkei grew by 5.95 percent on average over 1961–99. Hence the log-mean is derived from the formula of lognormal mean as $y = 0.0595 - x^2/2$.

The propensity to consume from asset, $b = 0.059$, is chosen from the empirical range (0.05–0.1) estimated from Japanese micro data in the 1990s by Hori and Shimizutani (2003). The linear specification of consumption function does not affect our simulation result much as long as the marginal propensity to consume from wealth for the high asset range is held in the empirical range, because the consumption function most crucially affects the accumulation rate of assets for the high asset group in our simulation. The standard deviation of labor income shock s determines the level of income for the middle class. We chose $s = 0.32$ to fit the middle part of the empirical distribution. We run the stochastic process for 100,000 agents.

The same model can simulate the U.S. distribution well with the different set of parameter values. Figure 8 plots the income distribution in the U.S. in 1971 and the simulated stationary distribution. We chose the year 1971 so that the bins in our data extend well in the tail part of the distribution (the tail extends to the top 0.001 percentile). The parameters are estimated by the same method as for the Japanese data. The lower bound growth rate is determined by the growth rate of average income for 1961–70: $v = 1.0525$. The trend wage growth rate is estimated by the inflation rate for the same period: $u = 1.0308$. We assume that the log-variance of the

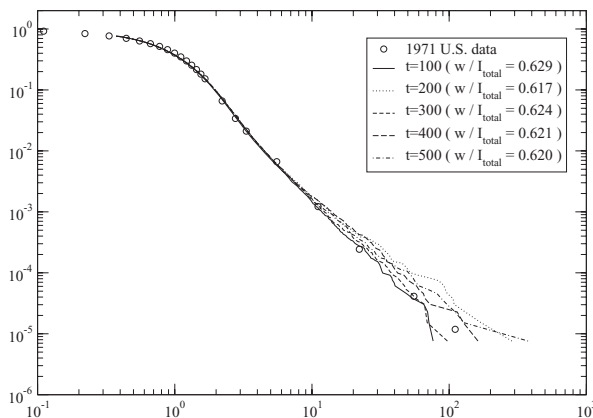


Figure 8. Simulated Stationary Distribution for the U.S.

asset returns is the same as in Japan: $x = 0.3122$. The log-mean of the asset returns is estimated by using x and the time-average growth rate of the Dow–Jones industrial index as $y = -0.029$. Parameters s and b are set free. The standard deviation of the labor shock, s , is set at 0.37 and the asset propensity to consume, b , is set at 0.018. The number of agents, 100,000, is the same as before. The good fit shown by the plot indicates that our parametric specification is versatile enough to produce realistic distributions for different economies. The model also enables us to discern what economic parameters correspond to the different shapes of income distributions. The standard deviation of the exponential distribution β is larger in the U.S. than in Japan. In the simulations, the larger β in the U.S. is caused by the larger s and the smaller $v - u$. Also, the Pareto exponent α is larger in the U.S. in 1971 than in Japan in 1999, which corresponds to the smaller b and y in the U.S.

4.2. *Fluctuations*

We conducted a sensitivity analysis by simulations to see that the changes in the parameter values can explain the fluctuations of the distribution. Figure 9 summarizes the results. In the top row, we observe that an increase in u or s increases β while the tail exponent α is unaltered. The bottom row shows that an increase in b or a decrease in x increases α while β is unaltered. Note that the log-mean of the asset returns y has the similar effect as $-b$, since it affects the asset growth rate $\gamma(t) - b$ through its mean $e^{y+x^2/2} - b$. Thus a decrease in y increases α just as an increase in b does. We also observed in simulations that an increase in v decreases β and increases α . Note that v is the growth rate of the average income which is the normalization factor. Hence the true growth parameters that determine the stationary distribution should be taken relative to v . Namely, $u - v$ instead of u determines the exponential mean β , and $y - v$ instead of y determines the Pareto exponent α . We also show the Gini coefficient for each distribution in Figure 9. The calculated coefficient varies considerably as the parameter changes. The plausible range of fluctuations of our parameters can span the range of Gini coefficients we observed in the data.

Finally, we tested whether the effects of our fundamental parameters $u - v$ and $y - v$ on the Gini coefficients are statistically significant in the tax returns data. Table 2 shows the estimates of the regression equation $g_t - g_{t-1} = b_u(u_t - v_t) + b_y(y_t - v_t) + \varepsilon_t$ where the difference in Gini coefficients is demeaned. The result is mixed. The coefficients for the Japanese data and the pooled data exhibit the predicted signs, but it is not significantly different from zero for $y - v$. For the U.S. data, the coefficient for $y - v$ is insignificant, and the coefficient for $u - v$ is significantly negative. This result shows the limitation of our data choice to represent the independent variables: asset returns volatility x^2 may be time-varying and have a significant role in determining α , and the real wage may not be a good measurement for the growth gap $u - v$ in the U.S.

5. CONCLUSION

This paper investigates the empirical shape of income distribution with a parametric specification motivated by the basic fact that income is derived from

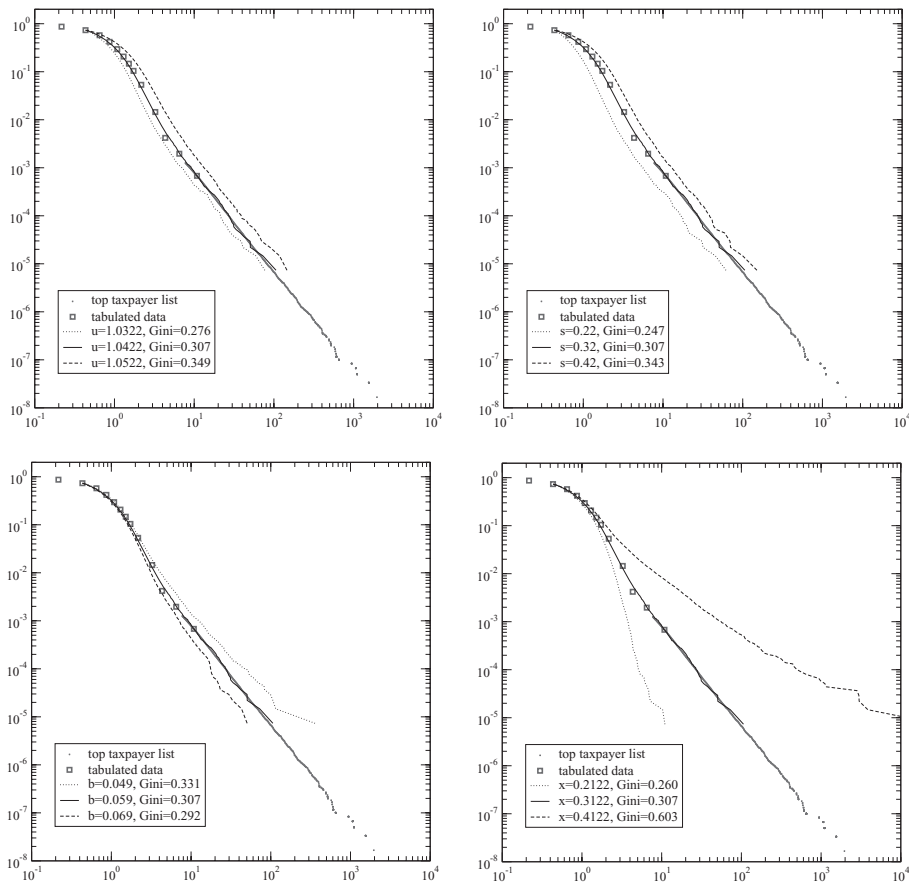


Figure 9. Parameter Sensitivity of the Stationary Distribution: Effects of u, s, x, b (Clockwise from Top Left)

TABLE 2
LEAST SQUARES REGRESSION OF THE DIFFERENCE IN GINI COEFFICIENTS

	$u - v$ (se)	$y - v$ (se)	R^2
U.S.	-0.0755 (0.0247)	0.0068 (0.0048)	0.237
Japan	0.0801 (0.0303)	0.0111 (0.0085)	0.183
Pooled	0.0452 (0.0216)	0.0082 (0.0055)	0.154

two different factors, labor and capital. Forty years of tax returns data in the U.S. and Japan reveal that the income distribution consistently obeys a particular shape of distribution if normalized by average income each year. The distribution is described by an exponential the distribution in the range from low to middle income and a power-law distribution for the top portion. In particular, the top taxpayers data clearly demonstrates the power-law in the tail. The power-law part has a stationary Pareto exponent about 2 and fluctuates in the range 1.3–2.6. The

exponential part in Japan has a stationary standard deviation about 0.6 of the average income and fluctuates in the range 0.5–0.65. The standard deviation of the exponential part in the U.S. rose from about 0.75 to 0.95 in the 1960s and 70s and then declined to 0.8 by 1999.

A simple stochastic process model of labor and asset income can explain this particular shape. We assume that the labor income follows an additive process with a growing lower bound, and the distribution of asset returns is independent of asset size. The additive process of labor income generates the stationary exponential distribution for the middle part of the distribution above a subsistence level of income. The savings from labor income behaves as the lower bound of the asset accumulation process. We characterize the tail exponent analytically. Our formula provides a rule-of-thumb relation between the Pareto exponent and macroeconomic statistics. Historically, the savings and the asset income are of the same magnitude, and it explains the stationary level of historical Pareto exponents. A simulation with calibrated parameters matches the middle and tail parts of the empirical distributions well. The model also explains the fluctuations of the shape of the distribution by the change in parameter values.

We leave it for future research to implement our mechanism of income distribution in the dynamic stochastic general equilibrium model in which the individual's policy function and the factor prices are endogenously determined. The income and wealth distribution in the DSGE models with incomplete insurance markets have been explored extensively, notably by Castañeda *et al.* (2003), Krusell and Smith (1998), and Quadrini (2000). Our conjecture is that the DSGE model with idiosyncratic returns would generate the power-law distribution in the tail even with an ex ante homogenous population, and the tail exponent would find a representation by the fundamental economic parameters.

REFERENCES

- Bordley, Robert F., James B. McDonald, and Anand Mantrala, "Something New, Something Old: Parametric Models for the Size Distribution of Income," *Journal of Income Distribution*, 6, 91–103, 1996.
- Castañeda, Ana, Javier Díaz-Giménez, and José-Victor Ríos-Rull, "Accounting for the U.S. Earnings and Wealth Inequality," *Journal of Political Economy*, 111, 818–57, 2003.
- Champernowne, D. G., "A Model of Income Distribution," *Economic Journal*, 63, 318–51, 1953.
- Dragulescu, A. A. and V. M. Yakovenko, "Exponential and Power-Law Probability Distributions of Wealth and Income in the United Kingdom and the United States," *Physica A*, 299, 213–21, 2001.
- Feenberg, Daniel R. and James M. Poterba, "Income Inequality and the Incomes of Very High Income Taxpayers: Evidence from Tax Returns," in James M. Poterba (ed.), *Tax Policy and the Economy*, MIT Press, 145–77, 1993.
- Feller, William, *An Introduction to Probability Theory and Its Applications*, Volume II, 2nd edition, Wiley, New York, 1966.
- Fujiwara, Y., W. Souma, H. Aoyama, T. Kaizoji, and M. Aoki, "Growth and Fluctuations of Personal Income," *Physica A*, 321, 598–604, 2003.
- Gabaix, Xavier, "Zipf's Law for Cities: An Explanation," *Quarterly Journal of Economics*, 114, 739–67, 1999.
- Gaver, D. P. and P. A. W. Lewis, "First-Order Autoregressive Gamma Sequences and Point Processes," *Advances in Applied Probability*, 12, 727–45, 1980.
- Hori, Masahiro and Satoshi Shimizutani, "Asset Holding and Consumption: Evidence from Japanese Panel Data in the 1990s," *ESRI Discussion Paper Series*, 55, 2003.
- Kalecki, M., "On the Gibrat Distribution," *Econometrica*, 13, 161–70, 1945.

- Krusell, Per and Anthony A. Smith, Jr., "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 106, 867–96, 1998.
- Levy, Moshe, "Are rich people smarter?" *Journal of Economic Theory*, 110, 42–64, 2003.
- Levy, M. and S. Solomon, "Power Laws are Logarithmic Boltzmann Laws," *International Journal of Modern Physics C*, 7, 595, 1996.
- Moriguchi, Chiaki and Emmanuel Saez, "The Evolution of Income Concentration in Japan, 1886–2002: Evidence from Income Tax Statistics," *NBER Working Paper*, 12558, 2006.
- Piketty, Thomas and Emmanuel Saez, "Income inequality in the United States, 1913–1998," *Quarterly Journal of Economics*, CXVIII, 1–39, 2003.
- Quadrini, Vincenzo, "Entrepreneurship, Saving, and Social Mobility," *Review of Economic Dynamics*, 3, 1–40, 2000.
- Reed, William J., "The Pareto, Zipf and Other Power Laws," *Economics Letters*, 74, 15–19, 2001.
- Rutherford, R. S. G., "Income Distribution: A New Model," *Econometrica*, 23, 277–94, 1955.
- Samuelson, Paul A., "Lifetime Portfolio Selection by Dynamic Stochastic Programming," *Review of Economics and Statistics*, 51, 239–46, 1969.
- Simon, H. A., "On a Class of Skew Distribution Functions," *Biometrika*, 52, 425–40, 1955.
- Singh, S. K. and G. S. Maddala, "A Function for Size Distribution of Incomes," *Econometrica*, 44, 963–70, 1976.
- Sornette, D. and R. Cont, "Convergent Multiplicative Processes Repelled from Zero: Power Laws and Truncated Power Laws," *Journal de Physique I France*, VII, 431–44, 1997.
- Souma, W., "Physics of Personal Income," in H. Takayasu (ed.), *Empirical Science of Financial Fluctuations*, Springer-Verlag, 2002.
- Sutton, John, "Gibrat's Legacy," *Journal of Economic Literature*, XXXV, 40–59, 1997.
- Takayasu, Hideki, Aki-Hiro Sato, and Misako Takayasu, "Stable Infinite Variance Fluctuations in Randomly Amplified Langevin Systems," *Physical Review Letters*, 79, 966–69, 1997.
- Wolff, Edward N., "Recent Trends in the Size Distribution of Household Wealth," *Journal of Economic Perspectives*, 12, 131–50, 1998.