Hitotsubashi University
Institute of Innovation Research

Pareto Distributions in Economic Growth Models

Makoto Nirei

IIR Working Paper WP#09-05

July 2009

Institute of Innovation Research
Hitotsubashi University
Tokyo, Japan
http://www.iir.hit-u.ac.jp

Pareto Distributions in Economic Growth Models

Makoto Nirei*

Institute of Innovation Research, Hitotsubashi University

July 22, 2009

Abstract

This paper analytically demonstrates that the tails of income and wealth distributions converge to a Pareto distribution in a variation of the Solow or Ramsey growth model where households bear idiosyncratic investment shocks. The Pareto exponent is shown to be decreasing in the shock variance, increasing in the growth rate, and increased by redistribution policies by income or bequest tax. Simulations show that even in the short run the exponent is affected by those fundamentals. We argue that the Pareto exponent is determined by the balance between the savings from labor income and the asset income contributed by risk-taking behavior.

1 Introduction

It is well known that the tail distributions of income and wealth follow a Pareto distribution in which the frequency decays in power as $\Pr(x) \propto x^{-\lambda-1}$. Researchers widely hold the view that the Pareto distribution is generated by multiplicative shocks in the

*Address: 2-1 Naka, Kunitachi, Tokyo 186-8603, Japan. Phone: +81 (42) 580-8417. Fax: +81 (42) 580-8410. E-mail: nirei@iir.hit-u.ac.jp. I am grateful to Yoshi Fujiwara and Hideaki Aoyama for generously sharing data. I have benefited by comments from Xavier Gabaix and Wataru Souma. I wish to thank Professor Tsuneo Ishikawa for introducing me to this research field.

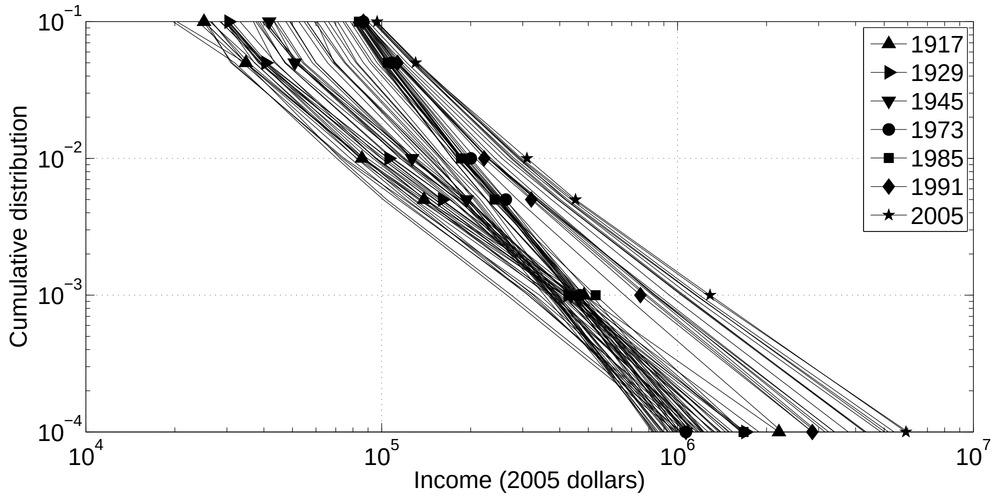

Figure 1: Income distributions above the top ten percentile

wealth accumulation process. In this paper, we formalize this idea in the framework of the standard models of economic growth.

There are renewed interests in the tail distributions of income and wealth. Recent studies by Feenberg and Poterba [9] and Piketty and Saez [24] investigated the time series of top income share in the U.S. by using tax returns data. Figure 1 shows the income distribution tails from 1917 to 2006 in the data compiled by Piketty and Saez [24]. The distribution is plotted in log-log scale and cumulated from above to the 90th percentile. We note that a linear line fits the tail distributions well in the log-log scale. This implies that the tail part follows $\log \Pr(X > x) = a - \lambda \log x$, which is called Pareto distribution and its slope λ is called Pareto exponent. The tail distribution is more egalitarian when λ is larger and thus when the slope is steeper. The good fit of the Pareto distribution in the tail is more evident in Figure 2, which plots the income distributions in Japan based on the tax returns data as in Souma [30], Fujiwara et al.[11], and Moriguchi and Saez [20]. The long tail covers more than a factor of twenty in relative income and displays a clear Paretian tail regardless of the economic situations Japan underwent, the bubble era or the lost decade.

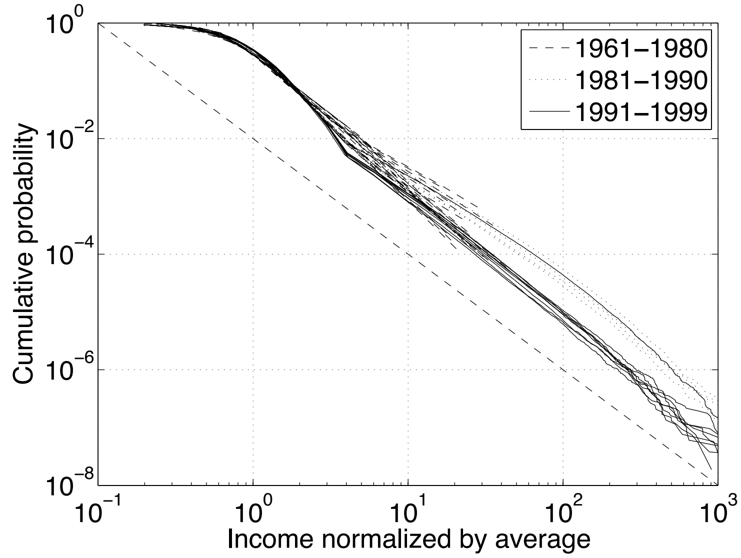

Figure 2: Income distributions in Japan (Source: Fujiwara et al. [11], Nirei and Souma [22].)

Figure 3 plots the time series of the U.S. Pareto exponent λ estimated for the various ranges of the tail.[1] We observe a slow movement of λ decade by decade: the U.S. Pareto exponent has increased steadily until 1960s, and since then declined fairly monotonically. The Pareto exponent stays in the range 1.5 to 3. But there is no clear time-trend for equalization or inequalization. In fact, we observe that the distributions in 1917 and 2006 are nearly parallel in Figure 1. This means that the income at each percentile grew at the same average rate in the long run. The Pareto exponent in Japan also stays around 2 as shown in Figure 2.

The stationarity of the Pareto exponent implies that the income distribution is characterized as a stationary distribution when normalized by the average income each year, since the Pareto distribution is "scale-free" (i.e., invariant in the change of units). This

[1] We estimate λ by a linear fit on the log-log plot. The linear fit is often problematic for estimating the tail exponent, because there are few observations on the tail for sampled data. This sampling error is not a big problem for our binned data, however, because each bin includes all the population within the bracket and thus contains a large number of observations.

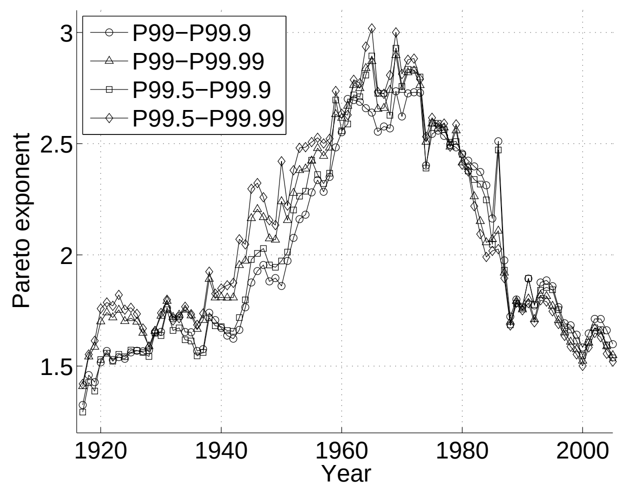

Figure 3: Pareto exponents for the U.S. income distributions

contrasts with the log-normal hypothesis of income distribution. If income incurs a multiplicative shock that is independent of income level ("Gibrat's law"), then the income follows a log-normal distribution. On the one hand, a log-normal distribution approximates a power-law distribution in an upper-middle range well (Montroll and Shlesinger [19]). On the other hand, the cross-sectional variance of log-income must increase linearly over time in the log-normal development of the distribution, and hence the dispersion of relative income diverges over time. Also, the log-normal development implies that the Pareto exponent becomes smaller and smaller as time passes. To the contrary, our plot exhibits no such trend in the Pareto exponent, and the variance of log-income is stationary over time as noted by Kalecki [14]. Thus, Gibrat's hypothesis of log-normal development need be modified.

We can obtain a stationary Pareto exponent by slightly modifying the multiplicative process in such a way that there is a small dragging force that prevents the very top portion of the distribution diverging. This was the idea of Champernowne [7], and subsequently a number of stochastic process models are proposed such as the preferential attachment and birth-and-death process (Simon [29], Rutherford [25], Sargan [27], Shorrocks [28]), and stochastic savings (Vaughan [32]). Models of randomly divided inheritance, along the line of Wold and Whittle [33], make a natural economic sense of the dragging force for the stationarity, as later pursued by Blinder [4] and Dutta and Michel [8]. Pestieau and Possen [23] presented a model of rational bequests with random returns to asset, in which the wealth distribution follows log-normal development. They showed that the variance of log-wealth can be bounded if their model incorporates a convex estate tax.

Another mechanism for the bounded log-variance is a reflexive lower bound, which pushes up the bottom of the distribution faster than the drift growth (total growth less diffusion effects) of the rest and thus achieves the stationary relative distribution. The analysis is based on the work of Kesten [15] and used by Levy and Solomon [16] and Gabaix [12]. The present paper extends this line of research by articulating the economic meanings of the multiplicative shocks and the reflexive lower bound in the context of the standard growth model. While we mainly focus on the reflexive lower bound, we also incorporate the effect of random inheritances in the Ramsey model by utilizing the multiplicative process with reset events studied by Manrubia and Zanette [17].

In the growth literature, there are relatively few studies on the shape of cross-section distribution of income and wealth, compared to the literature on the relation between inequality and growth. Perhaps the most celebrated is Stiglitz [31] which highlights the equalizing force of the Solow model. Numerical studies by Huggett [13] and Castañeda, Díaz-Giménez, and Ríos-Rull [6] successfully capture the overall shape of earnings and wealth distributions in a modern dynamic general equilibrium model. The present paper complements theirs by concentrating on the analytical characterization of the tail distribution, which covers a small fraction of population but carries a big impact on the inequality measures because of its large shares in income and wealth.

In this paper, we develop a simple theory of income distribution in the Solow and Ramsey growth models. We incorporate an idiosyncratic asset return shock, and show that the Solow model generates a stationary Pareto distribution for the detrended household income at the balanced growth path. Pareto exponent λ is analytically determined by fundamental parameters. The determinant of λ is summarized by the balance between the inflow from the middle class into the tail part and the inequalization force within the tail part. The inflow is represented by the savings from wage income that pushes up the bottom of the wealth distribution. The inequalization force is captured by the capital income that is attributed to risk taking behaviors of the top income earners. When the two variables are equal, Pareto exponent is determined at the historically stationary level, $\lambda = 2$.

We establish some comparative statics for the Pareto exponent analytically. The Pareto exponent is decreasing in the returns shock variance, and increasing in the technological growth rate. This is interpreted by the balance between the middle-class savings and the diffusion: the shock variance increases the diffusion, while the growth contributes to the savings size. The savings rate per se turns out to be neutral on the determination of the Pareto exponent, because it affects the savings amount and the variance of the asset growth equally. However, a heterogeneous change in savings behavior does affect the Pareto exponent. For example, a lowered savings rate in the low asset group reduces the savings amount and thus decreases the Pareto exponent. Redistribution policies financed by flat-rate taxes on income and bequest are shown to raise the Pareto exponent. Finally, a decline in mortality rate lowers the Pareto exponent.

The rest of the paper is organized as follows. In Section 2, the Solow model with uninsurable investment risk is presented and the main result is shown. In Section 3, the Solow model is extended to incorporate redistribution policies by income or bequest tax, and the sensitivity of the exponent is investigated. In Section 4, the stationary Pareto distribution is derived in the Ramsey model with random discontinuation of household lineage. Section 5 concludes.

2 Solow model

In this section, we present a Solow growth model with an uninsurable and undiversifiable investment risk. Consider a continuum of infinitely-living consumers $i \in [0, 1]$. Each consumer is endowed with one unit of labor, an initial capital $k_{i,0}$, and a "backyard" production technology that is specified by a Cobb-Douglas production function:

$$y_{i,t} = k_{i,t}^{\alpha}(a_{i,t}l_{i,t})^{1-\alpha} \tag{1}$$

where $l_{i,t}$ is the labor employed by i and $k_{i,t}$ is the capital owned by i. The labor-augmenting productivity $a_{i,t}$ is an i.i.d. random variable across households and across periods with a common trend $\gamma > 1$:

$$a_{i,t} = \gamma^t \epsilon_{i,t} \tag{2}$$

where $\epsilon_{i,t}$ is a temporary productivity shock with mean 1. Households do not have a means to insure against the productivity shock $\epsilon_{i,t}$ except for its own savings. Note that the effect of the shock is temporary. Even if i drew a bad shock in $t-1$, its average productivity is flushed to the trend level γ^t in the next period.

Each household supplies one unit of labor inelastically. The capital accumulation follows:

$$k_{i,t+1} = (1-\delta)k_{i,t} + s(\pi_{i,t} + w_t) \tag{3}$$

where s is a constant savings rate and $\pi_{i,t} + w_t$ is the income of household i. $\pi_{i,t}$ denotes the profit from the production:

$$\pi_{i,t} = \max_{l_{i,t}, y_{i,t}} y_{i,t} - w_t l_{i,t} \tag{4}$$

subject to the production function (1). The first order condition of the profit maximization yields a labor demand function:

$$l_{i,t} = (1-\alpha)^{1/\alpha} a_{i,t}^{(1-\alpha)/\alpha} w_t^{-1/\alpha} k_{i,t}. \tag{5}$$

Plugging into the production function, we obtain a supply function:

$$y_{i,t} = (1-\alpha)^{1/\alpha-1} a_{i,t}^{(1-\alpha)/\alpha} w_t^{1-1/\alpha} k_{i,t}. \tag{6}$$

By combining (4,5,6), we obtain that $\pi_{i,t} = \alpha y_{i,t}$ and $w_t l_{i,t} = (1-\alpha)y_{i,t}$.

The aggregate labor supply is 1, and thus the labor market clearing implies $\int_0^1 l_{i,t} di = 1$. Define the aggregate output and capital as $Y_t \equiv \int_0^1 y_{i,t} di$ and $K_t \equiv \int_0^1 k_{i,t} di$. Then we obtain the labor share of income to be a constant $1-\alpha$:

$$w_t = (1-\alpha)Y_t. \tag{7}$$

Plugging into (6) and integrating, we obtain an aggregate relation:

$$Y_t = E(a_{i,t}^{(1-\alpha)/\alpha})^\alpha K_t^\alpha = E(\epsilon_{i,t}^{(1-\alpha)/\alpha})^\alpha \gamma^{t(1-\alpha)} K_t^\alpha. \tag{8}$$

By aggregating the capital accumulation equations (3) across households, we reproduce the equation of motion for the aggregate capital in the Solow model:

$$K_{t+1} = (1-\delta)K_t + sY_t. \tag{9}$$

Define a detrended aggregate capital as:

$$X_t \equiv \frac{K_t}{\gamma^t}. \tag{10}$$

Then the equation of motion is written as:

$$\gamma X_{t+1} = (1-\delta)X_t + s\eta X_t^\alpha \tag{11}$$

where

$$\eta \equiv \mathrm{E}\left(\epsilon_{i,t}^{(1-\alpha)/\alpha}\right)^\alpha. \tag{12}$$

Equation (11) shows the deterministic dynamics of X_t, and its steady state is solved as:

$$\bar{X} = \left(\frac{s\eta}{\gamma - 1 + \delta}\right)^{1/(1-\alpha)}. \tag{13}$$

Since $\alpha < 1$, the steady state is stable and unique in the domain $X > 0$. The model preserves the standard implications of the Solow model on the aggregate characteristics of the balanced growth path. The actual investment equals the break-even investment on the balanced growth path as $s(Y_t/\gamma^t) = (\gamma - 1 + \delta)\bar{X}$. The long-run output-capital ratio $(\bar{Y/K})$ is equal to $(\gamma - 1 + \delta)/s$. The golden-rule savings rate is equal to α under the Cobb-Douglas technology.

We now turn to the dynamics of individual capital. Define a detrended individual capital:

$$x_{i,t} \equiv \frac{k_{i,t}}{\gamma^t}. \tag{14}$$

Substituting in (3), we obtain:

$$\gamma x_{i,t+1} = \left(1 - \delta + s\alpha X_t^{\alpha-1}(\epsilon_t/\eta)^{(1-\alpha)/\alpha}\right) x_{i,t} + s\eta(1-\alpha)X_t^\alpha. \tag{15}$$

The system of equations (11,15) defines the dynamics of $(x_{i,t}, X_t)$. We saw that X_t deterministically converges to the steady state $\bar{X}$. At the steady state, the dynamics of an individual capital $x_{i,t}$ (15) becomes:

$$x_{i,t+1} = g_{i,t} x_{i,t} + z \tag{16}$$

where

$$g_{i,t} \equiv \frac{1-\delta}{\gamma} + \frac{\alpha(\gamma - 1 + \delta)}{\gamma} \frac{\epsilon_{i,t}^{(1-\alpha)/\alpha}}{\mathrm{E}(\epsilon_{i,t}^{(1-\alpha)/\alpha})} \tag{17}$$

$$z \equiv \frac{s\eta(1-\alpha)}{\gamma} \bar{X}^{\alpha} = \frac{s\eta(1-\alpha)}{\gamma} \left(\frac{s\eta}{\gamma - 1 + \delta} \right)^{\alpha/(1-\alpha)}. \tag{18}$$

$g_{i,t}$ is the return to detrended capital $(1 - \delta + s\pi_{i,t}/x_{i,t})/\gamma$, and z is the savings from detrended labor income sw_t/γ^{t+1}. Equation (16) is called a Kesten process, which is a stochastic process with a multiplicative shock and an additive shock (a constant in our case). At the steady state,

$$\mathrm{E}(g_{i,t}) = 1 - z/\bar{x} \tag{19}$$

must hold, where the steady state $\bar{x}$ is equal to the aggregate steady state $\bar{X}$. $\mathrm{E}(g_{i,t}) = \alpha + (1-\alpha)(1-\delta)/\gamma < 1$ holds by the definition of $g_{i,t}$ (17), and hence the Kesten process is stationary. The following proposition is obtained by applying the theorem shown by Kesten [15] (see also [16] and [12]):

Proposition 1 *Household's detrended capital $x_{i,t}$ has a stationary distribution whose tail follows a Pareto distribution:*

$$\Pr(x_{i,t} > x) \propto x^{-\lambda} \tag{20}$$

where the Pareto exponent λ is determined by the condition:

$$\mathrm{E}(g_{i,t}^{\lambda}) = 1. \tag{21}$$

Household's income $\pi_{i,t} + w_t$ also follows the same tail distribution.

The condition (21) is understood as follows (see Gabaix [12]). When x_t has a power-law tail $\Pr(x_t > x) = c_0 x^{-\lambda}$ for a large x, the cumulative probability of x_{t+1} satisfies

$\Pr(x_{t+1} > x) = \Pr(x_t > (x-z)/g_t) = c_0(x-z)^{-\lambda} \int g_t^\lambda F(dg_t)$ for a large x and a fixed z, where F denotes the distribution function of g_t. Thus, x_{t+1} has the same distribution as x_t in the tail only if $E(g_t^\lambda) = 1$. Household's income also follows the same distribution, because the capital income $\pi_{i,t}$ is proportional to $k_{i,t}$ and the labor income w_t is constant across households and much smaller than the capital income in the tail part.

In what follows we assume that the returns shock $\log \epsilon_{i,t}$ follows a normal distribution with mean $-\sigma^2/2$ and variance σ^2. We first establish that $\lambda > 1$ and that λ is decreasing in the shock variance.

Proposition 2 *There exists $\bar{\sigma}$ such that for $\sigma > \bar{\sigma}$ Equation (21) determines the Pareto exponent λ uniquely. The Pareto exponent always satisfies $\lambda > 1$ and the stationary distribution has a finite mean. Moreover, λ is decreasing in σ.*

Proof is deferred to Appendix A.

Proposition 2 conducts a comparative static of λ with respect to σ. To do so, we need to show that $\mathrm{E}(g_{i,t}^\lambda)$ is strictly increasing in λ. Showing this is easy if $\delta = 1$, since then $g_{i,t}$ follows a two-parameter log-normal distribution. Under the 100% depreciation, we obtain the analytic solution for λ as follows (see Appendix B for the derivation):

$$\lambda|_{\delta=1} = 1 + \left(\frac{\alpha}{1-\alpha}\right)^2 \frac{\log(1/\alpha)}{\sigma^2/2}. \tag{22}$$

This expression captures our result that λ is greater than 1 and decreasing in σ. Proposition 2 establishes this property in a more realistic case of partial depreciation under which $g_{i,t}$ follows a shifted log-normal distribution. The property is shown for a sufficiently large variance σ^2. This is required so that $g > 1$ occurs with a sufficiently large probability. The lower bound $\bar{\sigma}$ takes the value about 1.8 under the benchmark parameter set ($\alpha = 1/3$, $\delta = 0.1$, $\gamma = 1.03$). This is a sufficient condition, however, and the condition can be relaxed. In simulations we use the value $\sigma^2 = 1.1$ and above, and we observe that $\mathrm{E}(g_{i,t}^\lambda)$ is still increasing in λ.

The Pareto distribution has a finite mean only if $\lambda > 1$ and a finite variance only if $\lambda > 2$. Since $\mathrm{E}(g_{i,t}) < 1$, it immediately follows that $\lambda > 1$ and that the stationary distribution of $x_{i,t}$ has a finite mean in our model. When λ is found in the range between 1 and 2, the capital distribution has a finite mean but an infinite variance. The infinite variance implies that, in an economy with finite households, the population variance grows unboundedly as the population size increases.

Proposition 2 shows that the idiosyncratic investment shocks generate a "top heavy" distribution, and at the same time it shows that there is a certain limit in the wealth inequality generated by the Solow economy, since the stationary Pareto exponent cannot be smaller than 1. Pareto distribution is "top heavy" in the sense that a sizable fraction of total wealth is possessed by the richest few. The richest P fraction of population owns $P^{1-1/\lambda}$ fraction of total wealth when $\lambda > 1$ (Newman [21]). For $\lambda = 2$, this implies that the top 1 percent owns 10% of total wealth. If $\lambda < 1$, the wealth share possessed by the rich converges to 1 as the population grows to infinity. Namely, virtually all of the wealth belongs to the richest few. Also when $\lambda < 1$, the expected ratio of the single richest person's wealth to the economy's total wealth converges to $1 - \lambda$ (Feller [10]). Such an economy almost resembles an aristocracy where a single person owns a fraction of total wealth. Proposition 2 shows that the Solow economy does not allow such an extreme concentration of wealth, because λ cannot be smaller than 1 at the stationary state.

Empirical income distributions show that the Pareto exponent transits below and above 2, in the range between 1.5 and 3. This implies that the economy goes back and forth between the two regimes, one with finite variance of income ($\lambda > 2$) and one with infinite variance ($\lambda \leq 2$). The two regimes differ not only quantitatively but also qualitatively, since for $\lambda < 2$ almost all of the sum of the variances of idiosyncratic risks is born by the wealthiest few whereas the risks are more evenly distributed for $\lambda > 2$. This is seen as follows. In our economy, the households do not diversify investment risks. Thus the variance of their income increases proportionally to their wealth $x_{i,t}^2$, which follows a Paretian tail with exponent $\lambda/2$. Thus, the income variance is distributed as a Pareto distribution with exponent less than 1, which is so unequal that the single wealthiest household bears a fraction $1 - \lambda/2$ of the sum of the variances of the idiosyncratic risks across households, and virtually all of the sum of the variances are born by the richest few percentile. Thus, in the Solow model, the concentration of the wealth can be interpreted as the result of the extraordinary concentration of risk bearings.

We can obtain an intuitive characterization for the threshold point $\lambda = 2$ for the two regimes under the assumption that $\log \epsilon_{i,t}$ follows a normal distribution with mean $-\sigma^2/2$ and variance σ^2 as follows.

Proposition 3 *The Pareto exponent λ is greater than (less than) 2 when $\sigma < \hat{\sigma}$ ($> \hat{\sigma}$) where:*

$$\hat{\sigma}^2 = \left(\frac{\alpha}{1-\alpha}\right)^2 \log\left(\frac{1}{\alpha^2}\left(1 - \frac{2\alpha}{\gamma/(1-\delta)-1}\right)\right). \tag{23}$$

Moreover, λ is increasing in γ and in δ in the neighborhood of $\lambda = 2$.

Proof is deferred to Appendix C.

Proposition 3 relates the Pareto exponent with the variance of the productivity shock, the growth rate, and the depreciation rate. The Pareto exponent is smaller when the variance is larger. At the variance $\hat{\sigma}^2$, the income distribution lays at the boarder between the two regimes where the variance of cross-sectional income is well defined and where the variance diverges. We also note that γ and δ both affect positively to λ around $\lambda = 2$. That is, a faster growth or a faster wealth depreciation helps equalization in the tail. A faster economic growth helps equalization, because the importance of the risk-taking behaviors (σ^2) is reduced relative to the drift growth. In other words, a faster trend growth increases the inflow from the middle class to the rich while the diffusion among the rich is unchanged. We explore this intuition further in the following.

The threshold variance is alternatively shown as follows. At $\lambda = 2$, $\mathrm{E}(g_{i,t}^2) = 1$ must hold by (21). Using $\mathrm{E}(g_{i,t}) = 1 - z/\bar{x}$, this leads to the condition $\mathrm{Var}(g_{i,t})/2 = z/\bar{x} - (z/\bar{x})^2/2$. The key variable $z/\bar{x}$ is equivalently expressed as:

$$\frac{z}{\bar{x}} = \frac{s(1-\alpha)}{\gamma}\left(\frac{\bar{Y}}{K}\right) = \frac{(1-\alpha)(\gamma-1+\delta)}{\gamma}. \tag{24}$$

Under the benchmark parameters $\alpha = 1/3$, $\delta = 0.1$, and $\gamma = 1.03$, we obtain $z/\bar{x}$ to be around 0.08. We can thus neglect the second order term $(z/\bar{x})^2$ and obtain $z \approx \bar{x}\mathrm{Var}(g_{i,t})/2$ as the condition for $\lambda = 2$. The right hand side expresses the growth of capital due to the diffusion effect. As we argue later, we may interpret this term as the capital income due to the risk taking behavior. The left hand side z represents the savings from the labor income. Then, the Pareto exponent is determined at 2 when the contribution of labor to capital accumulation balances with the contribution of risk taking. In other words, the stationary distribution of income exhibits a finite or infinite variance depending on whether the wage contribution to capital accumulation exceeds or falls short of the contribution from risk takings.

When ϵ follows a log-normal distribution, g is approximated in the first order by a log-normal distribution around the mean of ϵ. We explore the formula for λ under the first-order approximation. By the condition (21), we obtain:

$$\lambda \approx -\frac{\mathrm{E}(\log g)}{\mathrm{Var}(\log g)/2}. \tag{25}$$

Note that for a log-normal g we have:

$$\log \mathrm{E}(g) = \mathrm{E}(\log g) + \mathrm{Var}(\log g)/2. \tag{26}$$

Thus, (25) indicates that λ is determined by the relative importance of the drift and the diffusion of the capital growth rates both of which contribute to the overall growth rate. Using the condition $\mathrm{E}(g) = 1 - z/\bar{x}$, we obtain an alternative expression:

$$\lambda \approx 1 + \frac{-\log(1 - z/\bar{x})}{\mathrm{Var}(\log g)/2} \tag{27}$$

as in Gabaix [12]. We observe that the Pareto exponent λ is always greater than 1, and it declines to 1 as the savings z decreases to 0 or the diffusion effect $\mathrm{Var}(\log g)$ increases to infinity. For a small $z/\bar{x}$, the expression is further approximated as:

$$\lambda \approx 1 + \frac{z}{\bar{x}\mathrm{Var}(\log g)/2}. \tag{28}$$

$\mathrm{Var}(\log g)/2$ is the contribution of the diffusion to the total return of asset. Thus, the Pareto exponent is equal to 2 when the savings z is equal to the part of capital income contributed by the risk-taking behavior.

The intuition of our mechanism to generate Pareto distribution is following. The multiplicative process is one of the most natural mechanisms for the right-skewed, heavy tailed distribution of income and wealth as the extensive literature on the Pareto distribution indicates. Without some modification, however, the multiplicative process leads to a log-normal development and does not generate the Pareto distribution nor the stationary variance of log income. Incorporating the wage income in the accumulation process just does this modification. In our model, the savings from wage income (z) serves as a reflexive lower bound of the multiplicative wealth accumulation. Moreover, we find that the Pareto exponent is determined by the balance between the contributions of this additive term (savings) and the diffusion term (capital income). The savings from the wage income determines the mobility between the tail wealth group and the rest. Thus, we can interpret our result as the mobility between the top and the middle sections of income determines the Pareto exponent.

The close connection between the multiplicative process and the Pareto distribution may be illustrated as follows. Pareto distribution implies a self-similar structure of the distribution in terms of the change of units. Suppose that you belong to a "millionaire club" where all the members earn more than a million. In the club, you find that $10^{-\lambda}$ of the club members earn 10 times more than you. If $\lambda = 2$, this is one percent of the all members. Suppose that, after some hard work, you now belong to a ten-million earners' club. But then, you find again that one percent of the club members earn 10 times more than you do. If your preference for wealth is ordered by the relative position of your wealth among your social peers, you will never be satisfied by climbing up the social ladder of millionaire clubs. This observation is in good contrast with the "memoryless" property of an exponential distribution in addition. An exponential distribution characterizes the middle-class distribution well. In any of the social clubs within the region of the exponential distribution, you will find the same fraction of club members who earn \$10000 more than you. The contrast between the Pareto distribution and the exponential distribution corresponds to that the Pareto distribution is generated by a multiplicative process with lower bound while the exponential distribution is generated by an additive process with lower bound [16]. Under this perspective, Nirei and Souma [22] showed that the middle-class distribution as well as the tail distribution can be reproduced in a simple growth model with multiplicative returns shocks and additive wage shocks.

3 Sensitivity Analysis

3.1 Redistribution

In this section, we extend the Solow framework to redistribution policies financed by taxes on income and bequest. Tax proceeds are redistributed to households equally. Let τ_y denote the flat-rate tax on income, and let τ_b denote the flat-rate bequest tax on inherited wealth. We assume that a household changes generations with probability μ in each period. Thus, the household's wealth is taxed at flat rate τ_b with probability μ and remains intact with probability $1-\mu$. Denote the bequest event by a random variable $\mathbf{1}_b$ that takes 1 with probability μ and 0 with probability $1-\mu$. Then, the capital accumulation equation (3) is modified as follows:

$$k_{i,t+1} = (1-\delta-\mathbf{1}_b\tau_b)k_{i,t} + s((1-\tau_y)(\pi_{i,t}+w_t)+\tau_y Y_t + \tau_b\mu K_t). \tag{29}$$

By aggregating, we recover the law of motion for X_t as in (11). Therefore, the redistribution policy does not affect $\bar{X}$ or aggregate output at the steady state. Combining with (29), the accumulation equation for an individual wealth is rewritten at $\bar{X}$ as follows:

$$x_{i,t+1} = \tilde{g}_{i,t}x_{i,t} + \tilde{z} \tag{30}$$

where the newly defined growth rate $\tilde{g}_{i,t}$ and the savings term $\tilde{z}$ are:

$$\tilde{g}_{i,t} \equiv \frac{1-\delta-\mathbf{1}_b\tau_b}{\gamma} + \frac{(1-\tau_y)\alpha(\gamma-1+\delta)}{\gamma}\frac{\epsilon_{i,t}^{(1-\alpha)/\alpha}}{\mathrm{E}(\epsilon_{i,t}^{(1-\alpha)/\alpha})} \tag{31}$$

$$\tilde{z} \equiv \frac{s\eta(1-\alpha+\alpha\tau_y)}{\gamma}\left(\frac{s\eta}{\gamma-1+\delta}\right)^{\alpha/(1-\alpha)} + \tau_b\mu\left(\frac{s\eta}{\gamma-1+\delta}\right)^{1/(1-\alpha)}. \tag{32}$$

This is a Kesten process, and the Pareto distribution immediately obtains.

Proposition 4 *Under the redistribution policy, a household's wealth $x_{i,t}$ has a stationary distribution whose tail follows a Pareto distribution with exponent λ that satisfies $\mathrm{E}(\tilde{g}_{i,t}^{\lambda}) = 1$. An increase in income tax τ_y or bequest tax τ_b raises λ, while λ is not affected by a change in savings rate s.*

Since the taxes τ_y and τ_b both shift the density distribution of $\tilde{g}_{i,t}$ downward, they raise the Pareto exponent λ and help equalizing the tail distribution. The effect of savings s is discussed in a separate section.

The redistribution financed by bequest tax τ_b has a similar effect to a random discontinuation of household lineage. By setting τ_b accordingly, we can incorporate the situation where a household may have no heir, all its wealth is confiscated and redistributed by the government, and a new household replaces it with no initial wealth. A decrease in mortality (μ) in such an economy will reduce the stationary Pareto exponent λ. Thus, the greater longevity of the population has an inequalizing effect on the tail wealth.

The redistribution financed by income tax τ_y essentially collects a fraction of profits and transfers the proceeds to households equally. Thus, the income tax works as a means to share idiosyncratic investment risks across households. How to allocate the transfer does not matter in determining λ, as long as the transfer is uncorrelated with the capital holding.

If the transfer is proportional to the capital holding, the redistribution scheme by the income tax is equivalent to an institutional change that allows households to better insure against the investment risks. In that case, we obtain the following result.

Proposition 5 *Consider a risk-sharing mechanism which collects τ_s fraction of profits $\pi_{i,t}$ and rebates back its ex-ante mean $\mathrm{E}(\tau_s \pi_{i,t})$. Then, an increase in τ_s raises λ.*

Proof is following. A partial risk sharing ($\tau_s > 0$) reduces the weight on $\epsilon_{i,t}$ in (31) while keeping the mean of $\tilde{g}_{i,t}$. Then $\tilde{g}_{i,t}$ before the risk sharing is a mean-preserving spread of the new $\tilde{g}_{i,t}$. Since $\lambda > 1$, a mean-preserving spread of $\tilde{g}_{i,t}$ increases the expected value of its convex function $\tilde{g}_{i,t}^{\lambda}$. Thus the risk sharing must raise λ in order to satisfy $\mathrm{E}(\tilde{g}_{i,t}^{\lambda}) = 1$. When the households completely share the idiosyncratic risks away, the model converges to the classic case of Stiglitz [31] in which a complete equalization of wealth distribution takes place.

3.2 Speed of Convergence

We now examine the response of the income distribution upon a shift in fundamental parameters such as the variance of return shocks and the tax rate. By the analytical result, we know that λ will increase by an increase in the shock variance or by a decrease in tax at the stationary distribution. By numerically simulating the response, we can check the speed of convergence toward a new stationary distribution. It is important to study the speed of convergence, because some parameters may take a long time for the distribution to converge and thus may not explain the decade-by-decade movement of Pareto exponents.

We first compute the economy with one million households with parameter values $\gamma = 1.03$, $\delta = 0.1$, $\alpha = 1/3$, and $s = 0.3$. Set $\sigma^2 \equiv \mathrm{Var}(\log \epsilon)$ at 1.1 so that the stationary Pareto exponent is equal to $\lambda = 2$. Then, we increase the shock variance σ^2 by 10% to 1.2. Figure 4 plots the transition of the income distribution after this jump in σ^2 for 30 years. The plot at $t = 1$ shows the initial distribution with $\lambda = 2$. We observe that the distribution converges to a flatter one (lower λ) in less than 10 years. Therefore, the variation in σ^2 affects the tail of income distribution fast enough to cause the historical movement of Pareto exponent.

It is natural that σ^2 affects λ quickly, because σ^2 represents the diffusion effect that directly affects the tail as an inequalization factor. The diffusion effect shifts the wealth from the high density region to the low density region in the wealth distribution. Thus, the diffusion effect transports wealth toward the further tail. The wealth shifts faster under the greater gradient of the density which is determined by λ. At the stationary λ, the diffusion effect is balanced with the negative drift effect $E(g) - 1 = -z/\bar{x} < 0$. When σ^2 increases, the diffusion effect exceeds the drift effect, and thus the wealth starts to move toward the tail immediately. The shift of the wealth continues until the gradient is reduced and the wealth shifted by diffusion regains balance with the negative drift.

We now turn to income tax. Taxation has a direct effect on wealth accumulation by lowering the annual increment of wealth as well as the effect through the altered incentives that households face. The impact of income tax legislation in the 1980s has been extensively discussed in the context of the recent U.S. inequalization. In Figure 1, we observe an unprecedented magnitude of decline in Pareto exponent right after the Tax Reform Act in 1986, as studied by Feenberg and Poterba [9]. Although the stable exponent after the downward leap suggests that the sudden decline was partly due to the tax-saving behavior, the steady decline of the Pareto exponent in the 1990s may suggest more persistent effects of the tax act. Piketty and Saez [24] suggests that the imposition

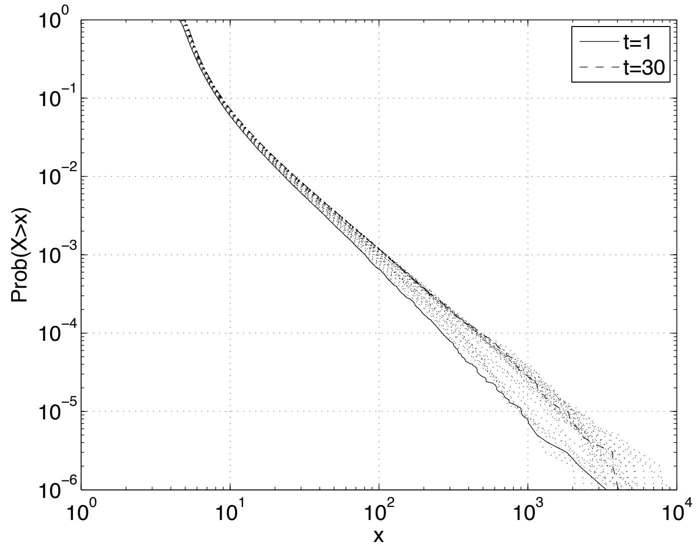

Figure 4: Transition of income distribution after an increase in σ

of progressive tax around the second world war was the possible cause for the top income share to decline during this period and stay at the low level for a long time until 1980s.

We incorporate a simple progressive tax scheme in our simulated model. We set the marginal tax rate at 50% which is applied for the income that exceeds three times worth of z. The tax proceed is consumed for non-productive activities. Starting from the stationary distribution under the 50% marginal tax, we reduce the tax rate to 20%. The result is shown in Figure 5. We observe the decline of Pareto exponent and its relatively fast convergence. This confirms that the change in progressive tax has an impact on Pareto exponent as well as on the income share of top earners.

The effect of the marginal tax on the highest bracket is similarly understood as the effect of the diffusion. The highest bracket tax affects $g_{i,t}$ for the high income households while it leaves the saving amount of most of the people unchanged. Thus a decrease in the highest marginal tax raises the volatility of $g_{i,t}$, while affecting other factors little for the capital accumulation.

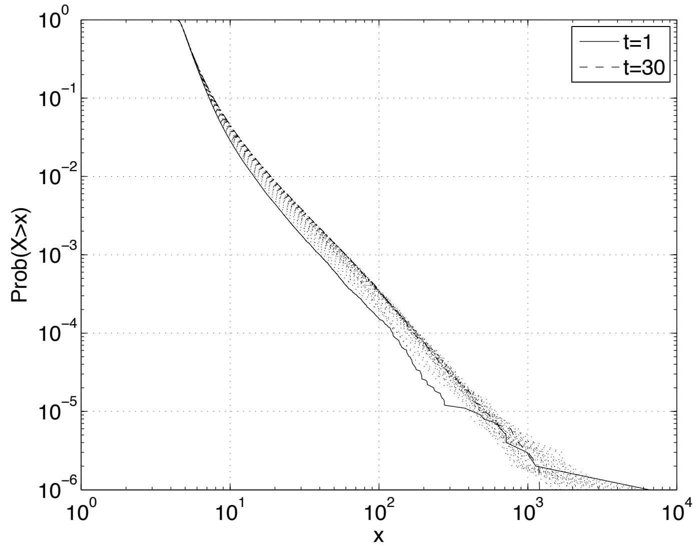

Figure 5: Transition of income distribution after a reduction in the highest marginal tax rate

3.3 Savings

Finally, we investigate the impact of the savings rate on Pareto exponent. As Proposition 4 has shown, the savings rate per se does not affect λ at the stationary distribution. It turns out that the savings rate does not affect λ in the transition either. This is because the savings rate affects the returns to wealth through the reduced reinvestment as well as the savings from the labor income. The left panel of Figure 6 shows the case in which the savings rate starts at 0.3 and is reduced by 1% every year. We observe that the Pareto exponent is preserved during the transition.

Previously we explained the mechanism that determines λ by a balance between the income from diffusion effect of assets and the savings from labor income. What matters for the determination of λ is the behavioral difference between the high and low wealth groups. We can see this point by examining the case where the change in the savings rate are different between the labor income and the capital income as in the "Cambridge" growth model. Consider that the rate of savings from labor income s_L is lower than the savings rate of capital income s_K. Suppose that $s_K = 1$ and $s_L = 0.45$ initially, and

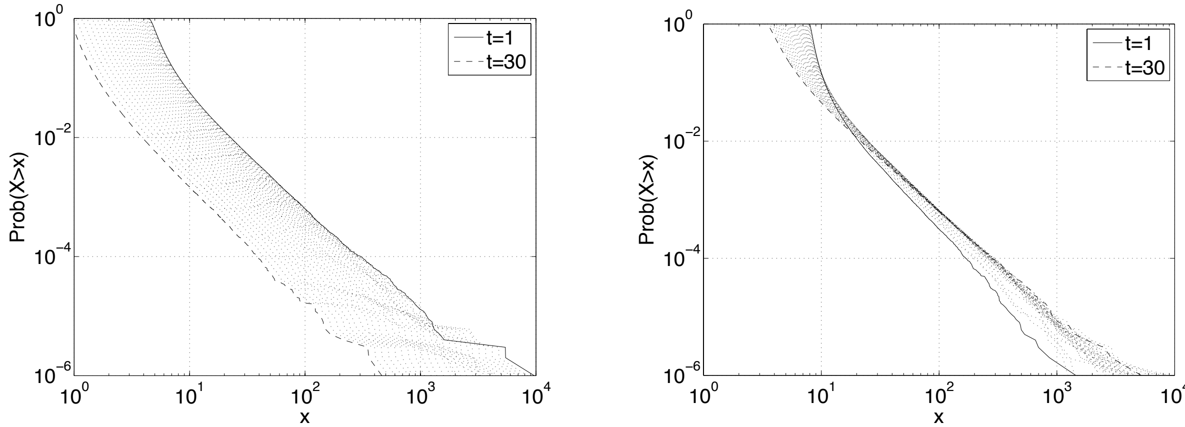

Figure 6: Impact of reduced savings rate on income distribution. Solow model (left) and Cambridge growth model (right)

suppose that s_L declines by 1 percent every year. The right panel in Figure 6 shows the transition of income distribution. We observe a steady decline in the Pareto exponent, unlike the previous case.

It is interesting that the factor such as s_L that solely affects the additive term z generates an immediate impact on the tail exponent. In a multiplicative process with reflexive lower bound, it is sometimes argued that the movement in the lower bound can take a long time to have an impact on the tail. This is because the tail is usually "too far" from the lower bound to "feel" the change in the lower bound. It seems not the case for the wealth distribution, whose exponent λ is considerably greater than 1 and the additive term z is also sizable.

All the variations investigated above show that the Pareto exponent in the Solow model has a lower limit: $\lambda > 1$. It suggests that there is a certain limit to an inequalization process. Is there a case when this limit is not warranted? Namely, is there a possibility in our framework for an "escaping" inequalization in which the log-variance of income grows without bound? Such a case requires $E(g) \geq 1$. In this case, the lower bound is never felt by the tail, the distribution of $x_{i,t}$ fails to converge, and it follows a log-normal development with log-variance linearly increasing in time.

Sources for such an escaping case are limited. It requires $E(g) \geq 1$, namely, the mean returns to wealth is greater than the overall growth rate of the economy. This could happen in an extreme case like, say, when the investment is heavily subsidized and financed by labor income tax. The Cambridge growth model studied above provides an interesting alternative case. Consider a simple case $s_K = 1$ and $s_L = 0$ in which the capital income is all reinvested and the labor income is all consumed. Then, the wealth accumulation process (16) becomes a pure multiplicative process. The income distribution follows a log-normal development, the log-variance of income grows linearly in time, and the Pareto exponent falls toward zero. Whereas the assumption $s_K = 1$ and $s_L = 0$ is extreme, this seems worth noting once we recall that the personal savings rate in the U.S. experiences values below zero recently.

4 Ramsey model

In the previous section, we note the possibility that the different savings behavior across the wealth group may have important implications on the Pareto exponent. In order to pursue this issue, we need to depart from the Solow model and incorporate optimization behavior of households. We work in a Ramsey model in this section. The model is analytically tractable when the savings rate and portfolio decisions are independent of wealth levels. Since Samuelson [26] and Merton [18], we know that this is the case when the utility exhibits a constant relative risk aversion. We draw on Angeletos [2] in which Aiyagari model [1] is modified to incorporate undiversifiable investment risks. Each consumer is endowed with an initial capital $k_{i,0}$, and a "backyard" production technology that is the same as before (1). The labor endowment is constant at 1 and supplied inelastically.

Households can engage in lending and borrowing $b_{i,t}$ at the risk-free interest R_t. Labor can be hired at wage w_t. The labor contract is contingent on the realization of the technology shock $a_{i,t}$. We assume that there is a small chance μ by which household lineage is discontinued at each i. At this event, a new household is formed at the same index i with no non-human wealth. Following the perpetual youth model [3], we assume that households participate in a pension program. Households contract all of the non-human wealth to be taken by the pension program at the event of discontinued lineage, and they receive in return the premium in each period of continued lineage at rate p per each unit of non-human wealth they own.

In each period, a household maximizes its profit from physical capital $\pi_{i,t} = y_{i,t} - w_t l_{i,t}$ subject to the production function (1). At the optimal labor hiring $l_{i,t}$, the return to capital is:

$$r_{i,t} \equiv \alpha(1-\alpha)^{(1-\alpha)/\alpha}(a_{i,t}/w_t)^{(1-\alpha)/\alpha} + 1 - \delta. \tag{33}$$

Household's non-human wealth is defined as $F_{i,t} \equiv r_{i,t}k_{i,t} + R_t b_{i,t}$. Household's total wealth $W_{i,t}$ is defined as:

$$\begin{aligned} W_{i,t} &\equiv (1+p)F_{i,t} + H_t && (34) \\ &= (1+p)(r_{i,t}k_{i,t} + R_t b_{i,t}) + H_t && (35) \end{aligned}$$

where p is the pension premium and H_t is the human wealth defined as the expected discounted present value of future wage income stream:

$$H_t \equiv \sum_{\tau=t}^{\infty} w_\tau (1-\mu)^{\tau-t} \prod_{s=t+1}^{\tau} R_s^{-1}. \tag{36}$$

The evolution of the human wealth satisfies $H_t = w_t + (1-\mu)R_{t+1}^{-1}H_{t+1}$. The pension program is a pure redistribution system, and must satisfy the zero-profit condition: $(1-\mu)p\int F_{i,t}di = \mu \int F_{i,t}di$. Thus, the pension premium rate is determined at:

$$p = \mu/(1-\mu). \tag{37}$$

Given an optimal operation of physical capital in each period, a household solves the following dynamic maximization problem:

$$\max_{c_{i,t}, b_{i,t}, k_{i,t+1}} \mathrm{E}_0 \left(\sum_{t=0}^{\infty} \beta^t \frac{c_{i,t}^{1-\theta}}{1-\theta} \right) \tag{38}$$

subject to:

$$c_{i,t} + k_{i,t+1} + b_{i,t+1} = r_{i,t}k_{i,t} + R_t b_{i,t} + w_t \tag{39}$$

and the no-Ponzi-game condition. For this maximization problem, consumption $c_{i,t}$ is set at zero once the household lineage is discontinued. Using H_t and W_t, the budget constraint is rewritten as:

$$c_{i,t} + k_{i,t+1} + b_{i,t+1} + (1-\mu)R_{t+1}^{-1}H_{t+1} = W_{i,t}. \tag{40}$$

Consider a balanced growth path at which R_t is constant over time. The household's problem has a recursive form:

$$V(W_i) = \max_{c_i, k_i', b_i', W_i'} \frac{c_i^{1-\theta}}{1-\theta} + (1-\mu)\beta E\left(V(W_i')\right) \tag{41}$$

subject to

$$W_i = c_i + k_i' + b_i' + (1-\mu)R^{-1}H' \tag{42}$$

$$W_i = (1+p)(r_i k_i + R b_i) + H. \tag{43}$$

This dynamic programming allows the following solution as shown in Appendix D:

$$c = (1-s)W \tag{44}$$

$$k' = s\phi W \tag{45}$$

$$b' = s(1-\phi)W - (1-\mu)R^{-1}H'. \tag{46}$$

Define $x_{i,t} \equiv W_{i,t}/\gamma^t$ as household i's detrended total wealth. By substituting the policy functions in the definition of wealth (35), and by noting that $(1-\mu)(1+p) = 1$ holds from the zero-profit condition for the pension program (37), we obtain the equation of motion for the individual total wealth:

$$x_i' = \hat{g}_i' x_i \tag{47}$$

where the growth rate is defined as:

$$\hat{g}_i' \equiv \frac{(\phi r_i' + (1-\phi)R)s}{(1-\mu)\gamma}. \tag{48}$$

Thus, at the balanced growth path, the household wealth evolves multiplicatively according to (47) as long as the household lineage is continued. When the lineage is discontinued, a new household with initial wealth $W_i = H$ replaces the old one. Therefore, the individual wealth W_i follows a log-normal process with random reset events where H is the resetting point. Using the result of Manrubia and Zanette [17], we can establish the Pareto exponent of the wealth distribution.

Proposition 6 *A household's detrended total wealth $x_{i,t}$ has a stationary Pareto distribution with exponent λ which is determined by:*

$$(1-\mu)\mathrm{E}(\hat{g}_{i,t}^{\lambda}) = 1. \tag{49}$$

Proof: See Appendix E.

The result is quite comparable to the previous results in the Solow model. The difference is that we do not have the additive term in the accumulation of wealth and instead we have a random reset event. As seen in (49), the Pareto exponent is large when μ is large. If there is no discontinuation event (i.e., $\mu = 0$), then the individual wealth follows a log-normal process as in [2]. In that case, the relative wealth $W_{i,t}/\int W_{j,t}dj$ does not have a stationary distribution. Eventually, a vanishingly small fraction of individuals possesses almost all the wealth. This is not consistent with the empirical evidence.

The difference from the Solow model occurs from the fact that the consumption function is linear in income in the Solow model whereas it is linear in wealth in the Ramsey model with CRRA preference. The linear consumption function arises in a quite narrow specification of the Ramsey model, however, as Carroll and Kimball [5] argues. For example, a concave consumption function with respect to wealth arises when labor income is uncertain or when the household's borrowing is constrained.

The concavity of the consumption means that the savings rate is high for the households with low wealth. This precautionary savings by the low wealth group serves as the reflexive lower bound of the wealth accumulation process. Then, the Pareto exponent of the tail distribution of wealth is determined by the balance of the precautionary savings and the diffusion of the wealth as in the Solow model. Further analysis requires numerical simulation, however, which is out of scope of the present paper.

5 Conclusion

This paper demonstrates that the Solow model with idiosyncratic investment risks is able to generate the Pareto distribution as the stationary distribution of income at the balanced growth path. We explicitly determine the Pareto exponent by the fundamental parameters, and provide an economic interpretation for its determinants.

The Pareto exponent is determined by the balance between the two factors: the savings from labor income, which determines the influx of population from the middle class to the tail part, and the asset income contributed by risk-taking behavior, which corresponds to the inequalizing diffusion effects taking place within the tail part. We show that an increase in the variance of the idiosyncratic investment shock lowers the Pareto exponent, while an increase in the technological growth rate raises the Pareto exponent. Redistribution policy financed by income or bequest tax raises the Pareto exponent, because the tax reduces the diffusion effect. Similarly, an increased level of risk-sharing raises the Pareto exponent. Savings rate per se does not affect the Pareto exponent, but a reduced savings rate from labor income lowers the Pareto exponent in a model of differential savings rate.

The effect of different savings behavior across wealth groups can be investigated in a Ramsey model. As a framework for such future studies, we present a Ramsey model with CRRA utility which is analytically tractable but in which the savings rate is constant across wealth groups. By incorporating a random event by which each household lineage is discontinued, we reestablish the Pareto distribution of wealth and income.

Appendix

A Proof of Proposition 2

We first show that $\mathrm{E}(g_{i,t}^{\lambda})$ is strictly increasing in λ for $\sigma > \bar{\sigma}$. This property is needed to show the unique existence of λ as well as to be the basis of all the comparative statics we conduct in this paper.

As $\lambda \to \infty$, g^{λ} is unbounded for the region $g > 1$ and converges to zero for the region

$g < 1$, while the probability of $g > 1$ is unchanged. Thus $\mathrm{E}(g^\lambda)$ grows greater than 1 eventually as λ increases to infinity. Also recall that $\mathrm{E}(g) < 1$. Thus, $\mathrm{E}(g^\lambda)$ travels from below to above 1. Hence, there exists unique λ that satisfies $\mathrm{E}(g) = 1$ if $\mathrm{E}(g^\lambda)$ is monotone in λ, and the solution satisfies $\lambda > 1$.

To establish the monotonicity, we show that the first derivative of the moment $(d/d\lambda)\mathrm{E}(g^\lambda) = \mathrm{E}(g^\lambda \log g)$ is strictly positive. Since $(d/d\lambda)(g^\lambda \log g) = g^\lambda(\log g)^2 > 0$, we have $\mathrm{E}(g^\lambda \log g) \geq \mathrm{E}(g \log g)$. Denote $g = a + u$ where $a = (1-\delta)/\gamma$ and $\log u$ follows a normal distribution with mean $\mu_u = u_0 - \sigma_u^2/2$ and variance $\sigma_u^2 = ((1-\alpha)/\alpha)^2\sigma^2$ where $u_0 \equiv \log(\alpha(1-(1-\delta)/\gamma))$. Then, noting that $a > 0$ and $u > 0$, we obtain:

$$\mathrm{E}(g \log g) = a\mathrm{E}(\log(a+u)) + \mathrm{E}(u \log(a+u)) > a \log a + \mathrm{E}(u \log u). \tag{50}$$

The last term is calculated as follows:

$$\begin{aligned}
\int_0^\infty \frac{u \log u}{u\sigma_u\sqrt{2\pi}} e^{-\frac{(\log u - \mu_u)^2}{2\sigma_u^2}} du &= \int_0^\infty \frac{\log u}{u\sigma_u\sqrt{2\pi}} e^{-\frac{(\log u - \mu_u)^2 - 2\sigma_u^2 \log u}{2\sigma_u^2}} du && (51)\\
&= e^{\mu_u + \sigma_u^2/2} \int_0^\infty \frac{\log u}{u\sigma_u\sqrt{2\pi}} e^{-\frac{(\log u - (\mu_u + \sigma_u^2))^2}{2\sigma_u^2}} du && (52)\\
&= e^{\mu_u + \sigma_u^2/2}(\mu_u + \sigma_u^2). && (53)
\end{aligned}$$

Collecting terms, we obtain:

$$\mathrm{E}(g \log g) > \frac{1-\delta}{\gamma} \log \frac{1-\delta}{\gamma} + \alpha\left(1 - \frac{1-\delta}{\gamma}\right)\left(\log\left(\alpha\left(1 - \frac{1-\delta}{\gamma}\right)\right) + \frac{\sigma^2}{2}\left(\frac{1-\alpha}{\alpha}\right)^2\right). \tag{54}$$

Since $1 - (1-\delta)/\gamma > 0$, the right hand side is positive for a sufficiently large σ. Thus, by defining:

$$\bar{\sigma}^2 \equiv \frac{\log\left[\left(\frac{1-\delta}{\gamma}\right)^{-\frac{1-\delta}{\gamma}} \left(\alpha\left(1 - \frac{1-\delta}{\gamma}\right)\right)^{-\alpha\left(1-\frac{1-\delta}{\gamma}\right)}\right]}{\frac{\alpha}{2}\left(1 - \frac{1-\delta}{\gamma}\right)\left(\frac{1-\alpha}{\alpha}\right)^2}, \tag{55}$$

we obtain that $\mathrm{E}(g \log g)$ is positive for $\sigma \geq \bar{\sigma}$. Thus, $\mathrm{E}(g^\lambda)$ is strictly increasing in λ under the sufficient condition $\sigma \geq \bar{\sigma}$.

Finally, we show that λ is decreasing in σ by showing that an increase in σ is a mean-preserving spread in g. Recall that g follows a shifted log-normal distribution where $\log u = \log(g - a)$ follows a normal distribution with mean $u_0 - \sigma_u^2/2$ and variance σ_u^2.

Note that the distribution of u is normalized so that a change in σ_u is mean-preserving for g. The cumulative distribution function of g is $F(g) = \Phi((\log(g-a) - u_0 + \sigma_u^2/2)/\sigma_u)$ where Φ denotes the cumulative distribution function of the standard normal. Then:

$$\frac{\partial F}{\partial \sigma_u} = \phi\left(\frac{\log(g-a) - u_0 + \sigma_u^2/2}{\sigma_u}\right)\left(-\frac{\log(g-a) - u_0}{\sigma_u^2} + \frac{1}{2}\right) \tag{56}$$

where ϕ is the derivative of Φ. By using the change of variable $x = (\log(g-a) - u_0 + \sigma_u^2/2)/\sigma_u$, we obtain:

$$\begin{aligned} \int^{\bar{g}} \frac{\partial F}{\partial \sigma_u} dg &= \int^{\bar{x}} \phi(x)(-x/\sigma_u + 1) dx dg/dx & (57) \\ &= \sigma_u e^{u_0} \int^{\bar{x}} \frac{-x/\sigma_u + 1}{\sqrt{2\pi}} e^{-(x-\sigma_u)^2/2} dx & (58) \end{aligned}$$

The last line reads as a partial moment of $-x/\sigma_u + 1$ in which x follows a normal distribution with mean σ_u and variance 1. The integral tends to 0 as $\bar{x} \to \infty$, and it is positive for any $\bar{x}$ below σ_u. Thus, the partial integral achieves the maximum at $\bar{x} = \sigma_u$ and then monotonically decreases toward 0. Hence, the partial integral is positive for any $\bar{x}$, and so is $\int^{\bar{g}} \partial F/\partial \sigma_u dg$. This completes the proof for that an increase in σ_u is a mean-preserving spread in g. Since g^λ is strictly convex in g for $\lambda > 1$, a mean-preserving spread in g strictly increases $\mathrm{E}(g^\lambda)$. As we saw, $\mathrm{E}(g^\lambda)$ is also strictly increasing in λ. Thus, an increase in σ_u, and thus an increase in σ while α is fixed, results in a decrease in λ that satisfies $\mathrm{E}(g^\lambda) = 1$ in the region $\lambda > 1$.

B Derivation of Equation (22)

We repeatedly use the fact that, when $\log \epsilon$ follows a normal distribution with mean $-\sigma^2/2$ and variance σ^2, $a_0 \log \epsilon$ also follows a normal with mean $-a_0\sigma^2/2$ and variance $a_0^2\sigma^2$. When $\delta = 1$, the growth rate of $x_{i,t}$ becomes a log-normally distributed variable $g = \alpha\epsilon^{(1-\alpha)/\alpha}/\mathrm{E}(\epsilon^{(1-\alpha)/\alpha})$. Then, g^λ also follows a log-normal with log-mean $\lambda(\log \alpha - \log \mathrm{E}(\epsilon^{(1-\alpha)/\alpha}) - (\sigma^2/2)((1-\alpha)/\alpha))$ and log-variance $\lambda^2((1-\alpha)/\alpha)^2\sigma^2$. Then we obtain:

$$\begin{aligned} 1 = \mathrm{E}(g^\lambda) &= e^{\lambda(\log \alpha - \log \mathrm{E}(\epsilon^{(1-\alpha)/\alpha}) - (\sigma^2/2)((1-\alpha)/\alpha)) + \lambda^2((1-\alpha)/\alpha)^2\sigma^2/2} & (59) \\ &= e^{\lambda(\log \alpha - (\sigma^2/2)((1-\alpha)/\alpha)^2) + \lambda^2((1-\alpha)/\alpha)^2\sigma^2/2}. & (60) \end{aligned}$$

Taking logarithm of both sides, we solve for λ as:

$$\lambda = 1 - \left(\frac{\alpha}{1-\alpha}\right)^2 \frac{\log \alpha}{\sigma^2/2}. \tag{61}$$

C Proof of Proposition 3

From the definition of $g_{i,t}$ (17), we obtain:

$$\mathrm{E}(g_{i,t}^2) = \left(\frac{1-\delta}{\gamma}\right)^2 + 2\frac{1-\delta}{\gamma}\frac{\alpha}{\gamma}(\gamma - 1 + \delta) + \left(\frac{\alpha}{\gamma}(\gamma - 1 + \delta)\right)^2 \frac{\mathrm{E}(\epsilon_{i,t}^{2(1-\alpha)/\alpha})}{\left(\mathrm{E}(\epsilon_{i,t}^{(1-\alpha)/\alpha})\right)^2}. \tag{62}$$

By applying the formula for the log-normal, we obtain:

$$\frac{\mathrm{E}(\epsilon_{i,t}^{2(1-\alpha)/\alpha})}{\left(\mathrm{E}(\epsilon_{i,t}^{(1-\alpha)/\alpha})\right)^2} = \frac{e^{-\sigma^2(1-\alpha)/\alpha + 2\sigma^2(1-\alpha)^2/\alpha^2}}{e^{-\sigma^2(1-\alpha)/\alpha + \sigma^2(1-\alpha)^2/\alpha^2}} = e^{\left(\frac{\sigma(1-\alpha)}{\alpha}\right)^2} \tag{63}$$

Combining these results with the condition $E(g^2) = 1$, we obtain:

$$\hat{\sigma}^2 = \left(\frac{\alpha}{1-\alpha}\right)^2 \log\left(\frac{1}{\alpha^2}\left(1 - \frac{2\alpha}{\gamma/(1-\delta) - 1}\right)\right). \tag{64}$$

We observe that an increase in γ or δ raises $\hat{\sigma}$. By Proposition 2, $E(g^\lambda)$ is monotonically increasing in λ when $\sigma > \bar{\sigma}$, and λ is decreasing in σ. Thus, in the neighborhood of $\lambda = 2$, the stationary λ is increasing in γ or δ, because an increase in either of them raises $\hat{\sigma}$ relative to the current level of σ.

D Derivation of the policy function in the Ramsey model

First, we guess and verify the policy functions (44,45,46) at the balanced growth path along with a guess on the value function $V(W) = BW^{1-\theta}/(1-\theta)$. The guessed policy functions for c, k', b' are consistent with the budget constraint (42).

The first order conditions and the envelope condition for the Bellman equation (41) are:

$$c^{-\theta} \quad = \quad \beta \mathrm{E}[r' V'(W')] \tag{65}$$

$$
\begin{aligned}
c^{-\theta} &= \beta R \mathrm{E}[V'(W')] && (66)\\
V'(W) &= c^{-\theta} && (67)
\end{aligned}
$$

Note that we used the condition $(1-\mu)(1+p) = 1$ from (37). By imposing the guess on these conditions, and by using $W' = (\phi r' + (1-\phi)R)(1+p)sW$ from (47), we obtain the equations that determine the constants:

$$
\begin{aligned}
0 &= E[(r' - R)(\phi r' + (1-\phi)R)^{-\theta}] && (68)\\
s &= (1-\mu)\left(\beta E[r'(\phi r' + (1-\phi)R)^{-\theta}]\right)^{1/\theta} && (69)\\
B &= (1-s)^{-\theta}. && (70)
\end{aligned}
$$

Thus we verify the guess.

E Proof of Proposition 6

In this section, we solve the Ramsey model and show the existence of the balanced growth path. Then the proposition obtains directly by applying Manrubia and Zanette [17].

At the balanced growth path, Y, K, and wealth grow at the same rate γ. The wage and output are determined similarly to the Solow model (7,8). As in the Solow model, define $X_t = K_t/\gamma^t$ as detrended aggregate physical capital. At the steady state $\bar{X}$, the return to physical capital (33) is written as:

$$
r_{i,t} = \alpha \epsilon_{i,t}^{(1-\alpha)/\alpha} \mathrm{E}(\epsilon_{i,t}^{(1-\alpha)/\alpha})^{\alpha-1} \bar{X}^{\alpha-1} + 1 - \delta, \tag{71}
$$

which is a stationary process. The average return is:

$$
\bar{r} \equiv \mathrm{E}(r) = \alpha\eta \bar{X}^{\alpha-1} + 1 - \delta. \tag{72}
$$

The lending market must clear in each period, which requires $\int b_{i,t} di = 0$ for any t. By aggregating the non-human wealth and using the market clearing condition for lending, we obtain: $\int F_{i,t} di = \bar{r} K_t$. Thus the aggregate total wealth satisfies $\int W_{i,t} di = (1-\mu)^{-1}\bar{r}K_t + H_t$. At the balanced growth path, aggregate total wealth, non-human wealth and human wealth grow at rate γ. Let $\bar{W}$, $\bar{H}$, and $\bar{w}$ denote the aggregate total wealth, the human capital and the wage rate detrended by γ^t at the balanced growth path, respectively. Then we have:

$$\bar{W} = (1-\mu)^{-1}\bar{r}\bar{X} + \bar{H}. \tag{73}$$

Combining the market clearing condition for lending with the policy function for lending (46), we obtain the equilibrium risk-free rate:

$$R = \frac{\gamma(1-\mu)}{s(1-\phi)}\frac{\bar{H}}{\bar{W}}. \tag{74}$$

By using the conditions above and substituting the policy function (44), the budget constraint (40) becomes in aggregation:

$$(\gamma - s(1-\mu)^{-1}\bar{r})\bar{X} = (s - (1-\mu)R^{-1}\gamma)\bar{H}. \tag{75}$$

Plugging into (74), we obtain the relation:

$$R = \frac{\gamma(1-\mu)}{s(1-\phi)} - \frac{\phi}{1-\phi}\bar{r}. \tag{76}$$

Thus, the mean return to the risky asset and the risk-free rate are determined by $\bar{X}$ from (72,76). The expected excess return is solved as:

$$\bar{r} - R = \frac{1}{1-\phi}\left(\alpha\eta\bar{X}^{\alpha-1} + 1 - \delta - (1-\mu)\gamma/s\right). \tag{77}$$

If $\log\epsilon \sim N(-\sigma^2/2, \sigma^2)$, then we have $\eta = e^{\frac{\sigma^2}{2}(1-\alpha)(1/\alpha-2)}$. This shows a relation between the expected excess return and the shock variance σ^2.

By using (7,8), the human wealth is written as:

$$\bar{H} = \gamma^{-t}\left(\sum_{\tau=t}^{\infty}\bar{w}\gamma^{\tau}(1-\mu)^{\tau-t}\prod_{s=t+1}^{\tau}R_s^{-1}\right) = \frac{\bar{w}}{1-(1-\mu)\gamma R^{-1}} = \frac{(1-\alpha)\eta\bar{X}^{\alpha}}{1-(1-\mu)\gamma R^{-1}}. \tag{78}$$

Equations (72,75,76,78) determine $\bar{X}, \bar{H}, R, \bar{r}$. In what follows, we show the existence of the balanced growth path in the situation when the parameters of the optimal policy s, ϕ reside in the interior of $(0,1)$. By using (72,76,78), we have:

$$\frac{\bar{X}}{\bar{H}} = \frac{1 - \frac{(1-\mu)\gamma s(1-\phi)}{\gamma(1-\mu) - s\phi(1-\delta) - s\phi\alpha\eta\bar{X}^{\alpha-1}}}{(1-\alpha)\eta\bar{X}^{\alpha-1}}. \tag{79}$$

The right hand side function is continuous and strictly increasing in $\bar{X}$, and travels from 0 to $+\infty$ as $\bar{X}$ increases from 0 to $+\infty$.

Now, the right hand side of (75) is transformed as follows:

$$\begin{aligned} \bar{H}(s-(1-\mu)\gamma R^{-1}) &= \bar{H}\left(s - s(1-\phi)\frac{\bar{W}}{\bar{H}}\right) = \bar{H}s\left(1-(1-\phi)\left((1-\mu)^{-1}\frac{\bar{r}\bar{X}}{\bar{H}}+1\right)\right) \\ &= \bar{H}s\left(\phi-(1-\phi)(1-\mu)^{-1}\frac{\bar{r}\bar{X}}{\bar{H}}\right). \end{aligned} \tag{80}$$

Then we rearrange (75) as:

$$\frac{\gamma}{s\phi}\frac{\bar{X}}{\bar{H}} = 1+(1-\mu)^{-1}\frac{\bar{r}\bar{X}}{\bar{H}}. \tag{81}$$

By (72), $\bar{r}$ is strictly decreasing in $\bar{X}$, and R is strictly increasing by (76). Thus $\bar{W}/\bar{H}$ is strictly decreasing by (74), and so is $\bar{r}\bar{X}/\bar{H}$ by (73). Thus, the right hand side of (81) is positive and strictly decreasing in $\bar{X}$. The left hand side is monotonically increasing from 0 to $+\infty$. Hence there exists the steady-state solution $\bar{X}$ uniquely. This verifies the unique existence of the balanced growth path.

The law of motion (47) for the detrended individual total wealth $x_{i,t}$ is now completely specified at the balanced growth path:

$$x_{i,t+1} = \begin{cases} \tilde{g}_{i,t+1}x_{i,t} & \text{with prob. } 1-\mu \\ \bar{H} & \text{with prob. } \mu, \end{cases} \tag{82}$$

where,

$$\tilde{g}_{i,t+1} \equiv (\phi r_{i,t+1} + (1-\phi)R)s/((1-\mu)\gamma). \tag{83}$$

This is the stochastic multiplicative process with reset events studied by Manrubia and Zanette [17]. By applying their result, we obtain our proposition.

References

[1] S. Rao Aiyagari. Uninsured idiosyncratic risk and aggregate saving. *Quarterly Journal of Economics*, 109:659–684, 1994.

[2] George-Marios Angeletos. Uninsured idiosyncratic investment risk and aggregate saving. *Review of Economic Dynamics*, 10:1–30, 2007.

[3] Olivier J. Blanchard. Debt, deficits, and finite horizons. *Journal of Political Economy*, 93:223–247, 1985.

[4] Alan S. Blinder. A model of inherited wealth. *Quarterly Journal of Economics*, 87:608–626, 1973.

[5] Christopher D. Carroll and Miles S. Kimball. On the concavity of the consumption function. *Econometrica*, 64:981–992, 1996.

[6] Ana Castañeda, Javier Díaz-Giménez, and José-Victor Ríos-Rull. Accounting for the U.S. earnings and wealth inequality. *Journal of Political Economy*, 111:818–857, 2003.

[7] D.G. Champernowne. A model of income distributon. *Economic Journal*, 63:318–351, 1953.

[8] Jayasri Dutta and Philippe Michel. The distribution of wealth with imperfect altruism. *Journal of Economic Theory*, 82:379–404, 1998.

[9] Daniel R. Feenberg and James M. Poterba. Income inequality and the incomes of very high income taxpayers: Evidence from tax returns. In James M. Poterba, editor, *Tax Policy and the Economy*, pages 145–177. MIT Press, 1993.

[10] William Feller. *An Introduction to Probability Theory and Its Applications*, volume II. Wiley, NY, second edition, 1966.

[11] Y. Fujiwara, W. Souma, H. Aoyama, T. Kaizoji, and M. Aoki. Growth and fluctuations of personal income. *Physica A*, 321:598–604, 2003.

[12] Xavier Gabaix. Zipf's law for cities: An explanation. *Quarterly Journal of Economics*, 114:739–767, 1999.

[13] Mark Huggett. Wealth distribution in life-cycle economies. *Journal of Monetary Economics*, 38:469–494, 1996.

[14] M. Kalecki. On the Gibrat distribution. *Econometrica*, 13:161–170, 1945.

[15] Harry Kesten. Random difference equations and renewal theory for products of random matrices. *Acta Mathematica*, 131:207–248, 1973.

[16] M. Levy and S. Solomon. Power laws are logarithmic boltzmann laws. *International Journal of Modern Physics C*, 7:595, 1996.

[17] Susanna C. Manrubia and Demián H. Zanette. Stochastic multiplicative processes with reset events. *Physical Review E*, 59:4945–4948, 1999.

[18] Robert C. Merton. Lifetime portfolio selection under uncertainty: The continuous-time case. *Review of Economics and Statistics*, 51:247–257, 1969.

[19] Elliott W. Montroll and Michael F. Shlesinger. Maximum entropy formalism, fractals, scaling phenomena, and 1/f noise: A tale of tails. *Journal of Statistical Physics*, 32:209–230, 1983.

[20] Chiaki Moriguchi and Emmanuel Saez. The evolution of income concentration in japan, 1886-2005: Evidence from income tax statistics. *Review of Economics and Statistics*, 90:713–734, 2008.

[21] M.E.J. Newman. Power laws, pareto distributions and zipf's law. *Contemporary Physics*, 46:323–351, 2005.

[22] Makoto Nirei and Wataru Souma. A two factor model of income distribution dynamics. *Review of Income and Wealth*, 53:440–459, 2007.

[23] Pierre Pestieau and Uri M. Possen. A model of wealth distribution. *Econometrica*, 47:761–772, 1979.

[24] Thomas Piketty and Emmanuel Saez. Income inequality in the United States, 1913-1998. *Quarterly Journal of Economics*, CXVIII:1–39, 2003.

[25] R.S.G. Rutherford. Income distribution: A new model. *Econometrica*, 23:277–294, 1955.

[26] Paul A. Samuelson. Lifetime portfolio selection by dynamic stochastic programming. *Review of Economics and Statistics*, 51:239–246, 1969.

[27] J. D. Sargan. The distribution of wealth. *Econometrica*, 25:568–590, 1957.

[28] A. F. Shorrocks. On stochastic models of size distributions. *Review of Economic Studies*, 42:631–641, 1975.

[29] H.A. Simon. On a class of skew distribution functions. *Biometrika*, 52:425–440, 1955.

[30] W. Souma. Physics of personal income. In H. Takayasu, editor, *Empirical Science of Financial Fluctuations*. Springer-Verlag, 2002.

[31] J. E. Stiglitz. Distribution of income and wealth among individuals. *Econometrica*, 37:382–397, 1969.

[32] R. N. Vaughan. Class behaviour and the distribution of wealth. *Review of Economic Studies*, 46:447–465, 1979.

[33] H. O. A. Wold and P. Whittle. A model explaining the Pareto distribution of wealth. *Econometrica*, 25:591–595, 1957.