Pareto Distribution of Income in Neoclassical Growth Models

Makoto Nirei

Institute of Innovation Research, Hitotsubashi University, 2-1 Naka, Kunitachi, Tokyo 186-8603, Japan.

Shuhei Aoki

Faculty of Economics, Hitotsubashi University, 2-1 Naka, Kunitachi, Tokyo 186-8603, Japan.

Abstract

This paper constructs a Bewley model, a dynamic general equilibrium model of heterogeneous households with production, which accounts for the Pareto distributions of income and wealth. We emphasize the role played by concavity of the consumption function in generating the Pareto distribution. We show that the Pareto distribution is obtained when households face idiosyncratic investment shocks on household assets and are subject to the borrowing constraint, which leads to concavity of the consumption function. The model can quantitatively account for the observed income distribution in the U.S. under reasonable calibration. In this model, labor income shocks account for the low and middle parts of the distribution, while investment shocks mainly affect the upper tail.

Keywords: income distribution; wealth distribution; Pareto exponent; idiosyncratic investment risk; borrowing constraint

JEL codes: D31, O40

Email address: nirei@iir.hit-u.ac.jp (Makoto Nirei)

1. Introduction

The issue of national income and wealth distribution has become an increasingly prominent subject of both scholarly and public attention. Scholars investigating this topic, such as Piketty and Saez [43], have been particularly concerned with understanding these distributions for the wealthiest individuals in the economy. It has been commonly observed that the income and wealth of this segment follow Pareto distributions. An important property of Pareto distributions is that they have very thick tails. In the real world, this means that the one percent of population accounted for by very rich persons possesses a substantially larger portion of the national income and wealth than would be predicted by extrapolating the distribution of middle income earners. Accordingly, greater understanding of the overall concentration of income and wealth requires increased attention to why the distributions of top earners universally follow the Pareto distribution.

The importance of addressing this issue is further highlighted by the fact that the Pareto distribution of top earners has not been explained in the standard workhorse model in macroeconomics. Researchers typically employ dynamic general equilibrium (DGE) models with heterogeneous households and production, the so-called Bewley models, to account for observed income distributions. While these models are relatively successful in accounting for the distribution of low and middle incomes, most of them do not effectively explain the distribution in the upper tail (Aiyagari [1]; Huggett [28]; Castañeda, Díaz-Giménez, and Ríos-Rull [14]; and Quadrini and Ríos-Rull [46]). One exception is Castañeda, Díaz-Giménez, and Ríos-Rull [15], who construct a DGE model that is consistent with the observed income distribution including the upper tail. However, they do not address why the top segments of income and wealth follow Pareto distributions. Moreover, their model relies on income shocks that do not derive from micro-level evidence. Panousi [41] provides another exception. Extending Angeletos' [3] model, she builds a DGE model incorporating idiosyncratic investment shocks, whose income distribution percentile predictions comport with data. However, she does not attempt to explain whether the model can account for observed Pareto distributions of income and wealth.

Some researchers have accounted for Pareto distributions of income and wealth by using multiplicative idiosyncratic shocks in partial equilibrium models that abstract from production. Since the classic work of Champernowne [16], it has been recognized that multiplicative idiosyncratic shocks on income or wealth can generate the Pareto distribution when combined with some mechanism that prevents the distribution from diverging. One such mechanism is the overlapping generations (OLG) setup. Wold and Whittle [51] and Dutta and Michel [19] show that the discontinuities of households stemming from death, combined with shocks to wealth or income, create the Pareto distribution. Recently, Benhabib, Bisin, and Zhu [8, 10] embed this mechanism into standard models wherein households solve intertemporal decision problems. Another proposed mechanism is concavity of the household consumption function. Nirei and Souma [40] employ this mechanism to construct a model of households that accounts for Pareto distributions of income and wealth. However, they rely on an ad hoc consumption function and pay little attention to the role played by concavity of consumption function.

The purpose of this paper is to construct a Bewley model that accounts for the observed Pareto distribution.[1] We derive our results by combining the literature on

[1]It came to our attention that Benhabib et al. [9] derive similar results. This paper differs from theirs in clarifying the role of the concave consumption function, which generates the Pareto distribution in Nirei and Souma [40]. Moreover, we analyze how varying borrowing limits affect the Pareto

the Bewley models with insights from research on multiplicative idiosyncratic shocks and Pareto distributions. Idiosyncratic investment shocks and a concave consumption function, the two elements that generate the Pareto distribution as discussed above, fit naturally into the standard Bewley model. Following Quadrini [44] and Cagetti and De Nardi [11] in spirit, and adopting the modeling strategy of Covas [17], Angeletos [3], and Panousi [41], we construct an entrepreneurial economy, wherein households engage in "backyard" production. In each period, each household bears income risk by investing physical capital in its own firm. In addition, as in the standard Bewley models, all households earn labor income subject to idiosyncratic earning shocks. The investment activity of households and the risks they bear are the key factors behind accumulation and concentration of wealth and income.

To develop our model, we first clarify the mechanism in Nirei and Souma [40] that generates the Pareto distribution. In Section 2, we show how a concave consumption function with investment shocks generates the Pareto distribution by assuming an analytically tractable Solow-type consumption function. The slope of the Pareto distribution, which is called Pareto exponent and characterizes the concentration of top income and wealth, is determined by two forces in the model: an inequalization effect within the upper tail due to risky investments, and an equalization effect due to the savings at the lower bound of household wealth accumulation.

The results obtained in Section 2 continue to hold in the model where households optimally solve intertemporal consumption problem. Carroll and Kimball [13], and the papers cited therein, show that a household's consumption function is generically concave if the household faces a borrowing constraint, as is usually assumed in the

distribution, and show that our model accounts for the observed income distribution in the U.S. The basic results of the present paper are derived in the working paper version (Nirei [39]).

Bewley models. Using this property, we show in Section 3 that the Bewley model with the borrowing constraint and idiosyncratic investment shocks generates Pareto distributions of wealth and income in the upper tail. The tightness of the borrowing constraints affects the concentration of wealth and income by changing the lower bound of household wealth levels.

We further examine quantitatively whether our model can account for the observed income distribution in the U.S. when the model incorporates other features such as idiosyncratic labor income shocks and progressive taxation. We assume the perpetual youth setting, which is another source of the Pareto distribution as shown in previous studies (Wold and Whittle [51]; Benhabib et al. [10]). Under reasonably calibrated parameter values, we show that the model can account for detailed distribution characteristics such as the Pareto exponent, the quintiles of income distribution, and the Gini coefficient. In our model, investment shocks mainly affect the top part of the distribution, while the low and middle parts of the distribution are shaped mostly by labor income shocks, as in the previous Bewley models of income distribution.

The rest of the paper is organized as follows. To develop the intuition underlying why a concave consumption function is important, Section 2 introduces a basic version of the model wherein households choose consumption and investment following a Solow-type consumption function. We analytically show that the combination of idiosyncratic investment shocks and the concave consumption function generates the Pareto distribution in the upper tail of the wealth and income distributions. Section 3 provides a more elaborate Bewley model wherein households optimally choose consumption and investment. We show that our model, with the borrowing constraint for households and idiosyncratic investment and labor income shocks, can account for the observed properties in the top as well as the remaining parts of the income distribution.

Finally, Section 4 concludes.

2. Analytical results in a simple model

2.1. Solow model with idiosyncratic investment risk

In this section, we present a Solow growth model with heterogeneous households who face uninsurable idiosyncratic investment risk. Here, we assume a fixed savings rate and i.i.d. productivity and labor shocks. At the expense of these assumptions, the Solow model is analytically tractable for deriving the Pareto exponent. These assumptions are relaxed in Section 3 where we study the Bewley model, wherein the savings rate is optimally chosen by households.

In the Bewley model in Section 3, we will argue that the borrowing constraint and the concavity of consumption function play an important role in determining the tail distribution. The concave consumption function can be featured in a tractable manner in the Solow model, since its consumption function has a kinked linear form as depicted in Figure 1. Thus, the Solow model is useful in interpreting the mechanism for generating the Pareto distribution when the households face a binding borrowing limit.

Consider a continuum of infinitely-living households $i \in [0,1]$. Household i is endowed with 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 detrended capital owned by i. The labor-augmenting productivity of the production function $\tilde{a}_{i,t}$ has a common trend $\gamma > 1$:

$$\tilde{a}_{i,t} = \gamma^t a_{i,t}, \tag{2}$$

where $a_{i,t}$ is an i.i.d. productivity shock. Because of the common productivity growth γ, other variables such as output, consumption, capital, bond holding, and real wage will grow, on an average, at γ along the balanced growth path. Thus, we employ the notation wherein these variables are detrended by γ^t.

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). Labor can be hired at wage w_t, and the labor contract is struck after the realization of $a_{i,t}$. By profit maximization conditions, we obtain the goods supply function:

$$y_{i,t} = ((1-\alpha)a_{i,t}/w_t)^{(1-\alpha)/\alpha} k_{i,t}. \tag{3}$$

Then, we obtain $\pi_{i,t} = \alpha y_{i,t}$ and $w_t l_{i,t} = (1-\alpha)y_{i,t}$. Detrended aggregate output and capital are denoted as $Y_t \equiv \int_0^1 y_{i,t} di$ and $K_t \equiv \int_0^1 k_{i,t} di$, respectively. The labor share of income is constant:

$$w_t/Y_t = 1 - \alpha. \tag{4}$$

Substituting into (3) and integrating, we obtain an aggregate relation:

$$Y_t = AK_t^{\alpha}, \tag{5}$$

where

$$A \equiv \left(\mathrm{E}\left(a_{i,t}^{(1-\alpha)/\alpha}\right)\right)^{\alpha}. \tag{6}$$

Households inelastically supply labor $h_{i,t}$, which is an i.i.d. random variable over i and t. The savings rate is exogenously fixed at s. There is no capital market in this model. The capital of household i, detrended by γ^t, accumulates as follows:

$$\gamma k_{i,t+1} = (1-\delta)k_{i,t} + s(\pi_{i,t} + w_t h_{i,t}) \tag{7}$$

where $\pi_{i,t}$ is the stochastic profit from production and $\pi_{i,t} + w_t h_{i,t}$ is the income of household i.

The mean labor endowment $\mathrm{E}(h_{i,t})$ is normalized to 1. Thus, aggregate labor supply is $\int_0^1 h_{i,t} di = 1$. By aggregating the capital accumulation equation (7) across households, and by using (5), we reproduce the equation of motion for aggregate capital in the Solow model,

$$\gamma K_{t+1} = (1-\delta)K_t + sAK_t^{\alpha}, \tag{8}$$

where K_t is detrended by γ^t. Equation (8) shows that K_t follows deterministic dynamics with steady state $\bar{K}$, which is stable and uniquely solved in the domain $K > 0$ as

$$\bar{K} = \left(\frac{sA}{\gamma - 1 + \delta} \right)^{1/(1-\alpha)}. \tag{9}$$

Thus, the model preserves the standard implications of the Solow model on the aggregate characteristics of the balanced growth path. The long-run output-capital ratio Y/K is equal to $(\gamma - 1 + \delta)/s$. The golden-rule savings rate is equal to α.

2.2. Deriving the Pareto distribution

The dynamics of individual capital is derived by using (2,3,4,5,7) and $\pi_{i,t} = \alpha y_{i,t}$ as follows:

$$\gamma k_{i,t+1} = \left(1 - \delta + s\alpha K_t^{\alpha-1}(a_{i,t}/A)^{(1-\alpha)/\alpha}\right) k_{i,t} + s(1-\alpha)AK_t^{\alpha} h_{i,t}. \tag{10}$$

The system of equations (8,10) defines the dynamics of $(k_{i,t}, K_t)$. As is shown above, K_t deterministically converges to $\bar{K}$. At $\bar{K}$, the dynamics of $k_{i,t}$ (10) follows

$$k_{i,t+1} = g_{i,t} k_{i,t} + z h_{i,t}, \tag{11}$$

where

$$g_{i,t} \equiv \frac{1-\delta}{\gamma} + \frac{\alpha(\gamma - 1 + \delta)}{\gamma} \frac{a_{i,t}^{(1-\alpha)/\alpha}}{\mathrm{E}(a_{i,t}^{(1-\alpha)/\alpha})}, \tag{12}$$

$$z \equiv \frac{(1-\alpha)sA\bar{K}^{\alpha}}{\gamma} = \frac{(1-\alpha)sA}{\gamma} \left(\frac{sA}{\gamma - 1 + \delta} \right)^{\alpha/(1-\alpha)}. \tag{13}$$

$g_{i,t}$ is the return to detrended capital $(1-\delta+s\pi_{i,t}/k_{i,t})/\gamma$ and $zh_{i,t}$ is the savings from detrended labor income $sw_t h_{i,t}/\gamma$. We note that z is determined by the intercept of the Solow-type consumption function in Figure 1. For a fixed s, larger wage w induces larger z and higher intercept $(1-s)w$. Thus, given s, larger z corresponds to larger concavity of the overall consumption function.

Equation (11) is called a Kesten process, which is a stochastic process with a multiplicative shock and an additive positive shock. At the stationary distribution of $k_{i,t}$,

$$\mathrm{E}(g_{i,t}) = 1 - z/\bar{k} \tag{14}$$

must hold, where the mean capital $\bar{k}$ is equal to the aggregate steady state $\bar{K}$. $\mathrm{E}(g_{i,t}) = \alpha + (1-\alpha)(1-\delta)/\gamma < 1$ holds from the definition of $g_{i,t}$ (12), and hence, the Kesten process is stationary. The following proposition is obtained by applying the theorem shown by Kesten [30] (see also Levy and Solomon [33] and Gabaix [23]):

Proposition 1. *The household's detrended capital $k_{i,t}$ has a stationary distribution whose tail follows a Pareto distribution:*

$$\Pr(k_{i,t} > k) \propto k^{-\lambda}, \tag{15}$$

where the Pareto exponent λ is determined by the condition

$$\mathrm{E}\left(g_{i,t}^{\lambda}\right) = 1. \tag{16}$$

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

Condition (16) is understood as follows (see Gabaix [23]). When $k_{i,t}$ has a power-law tail $\Pr(k_{i,t} > k) = c_0 k^{-\lambda}$ for a large k, the cumulative probability of $k_{i,t+1}$ satisfies $\Pr(k_{i,t+1} > k) = \Pr(k_{i,t} > (k-z)/g_{i,t}) = c_0(k-z)^{-\lambda}\int g_t^{\lambda} F(dg_t)$ for a large k and a fixed z, where F denotes the distribution function of $g_{i,t}$. Thus, $k_{i,t+1}$ has the same distribution as $k_{i,t}$ in the tail only if $\mathrm{E}(g_{i,t}^{\lambda}) = 1$. The household's income also follows the same tail 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.

2.3. Determination of the Pareto exponent and comparative statics

We further characterize λ by assuming that the productivity shock $a_{i,t}$ follows a log-normal distribution with mean 1. Let σ^2 denote the variance of $\log a_{i,t}$. Thus, $\mathrm{E}(a_{i,t}) = 1$ implies $\mathrm{E}(\log a_{i,t}) = -\sigma^2/2$. We first show that λ is decreasing in σ and bounded below by 1.

Proposition 2. *The Pareto exponent λ is uniquely determined by Equation (16) for any σ. The Pareto exponent always satisfies $\lambda > 1$ and the stationary distribution has a finite mean. Moreover, λ is decreasing in σ.*

The proof is deferred to Appendix A.

Proposition 2 provides a comparative static of λ with respect to σ. In the proof, we show that $\mathrm{E}(g_{i,t}^{\lambda})$ is strictly increasing in λ. Establishing this is easy when $\delta = 1$, since $g_{i,t}$ then follows a two-parameter log-normal distribution. Under 100% depreciation, we obtain a closed-form solution for λ as follows.

Proposition 3. *If $\delta = 1$, the Pareto exponent is explicitly determined as*

$$\lambda = 1 + \left(\frac{\alpha}{1-\alpha}\right)^2 \frac{\log(1/\alpha)}{\sigma^2/2}. \tag{17}$$

The proof is in Appendix B.

This expression captures the essential result that λ is greater than 1 and decreasing in σ.[2] Proposition 2 establishes this property in a more realistic case of partial depreciation under which $g_{i,t}$ follows a shifted log-normal distribution.

An analytical solution is obtained for an important special case $\lambda = 2$ as follows.

[2]Moreover, it can be shown by (17) that λ is decreasing in α for $\alpha < 0.5$.

Proposition 4. *The Pareto exponent λ is greater than (less than) 2 when $\sigma < \hat{\sigma}$ ($> \hat{\sigma}$) where*

$$\hat{\sigma}^2 = \left(\frac{\alpha}{1-\alpha}\right)^2 \log\left(\frac{1}{\alpha^2}\left(1 + \frac{2(1-\alpha)}{\gamma/(1-\delta) - 1}\right)\right). \tag{18}$$

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

The proof is deferred to Appendix C.

Proposition 4 relates the Pareto exponent λ with the productivity shock variance σ^2, growth rate γ, and depreciation rate δ. The Pareto exponent is smaller when the variance is larger. Both γ and δ negatively affect λ around $\lambda = 2$. That is, faster growth or faster wealth depreciation helps inequalization in the tail if λ is around 2.

Proposition 4 determines the magnitude of risk that generates the Pareto exponent $\lambda = 2$. The risk magnitude is intuitively derived as follows. At $\lambda = 2$, $\mathrm{E}(g_{i,t}^2) = 1$ must hold given (16). Using $\mathrm{E}(g_{i,t}) = 1 - z/\bar{k}$, this leads to the condition $\mathrm{Var}(g_{i,t})/2 = z/\bar{k} - (z/\bar{k})^2/2$. The key variable $z/\bar{k}$ is equivalently expressed as

$$\frac{z}{\bar{k}} = \frac{(1-\alpha)s}{\gamma}\left(\frac{\bar{Y}}{\bar{K}}\right) = \frac{(1-\alpha)(\gamma - 1 + \delta)}{\gamma}. \tag{19}$$

Under the benchmark parameters $\alpha = 0.36$, $\delta = 0.1$, and $\gamma = 1.02$, we obtain $z/\bar{k}$ to be around 0.08. We can thus neglect the second-order term $(z/\bar{k})^2$ and obtain $z \approx \bar{k}\mathrm{Var}(g_{i,t})/2$ as the condition for $\lambda = 2$. Under the calibration above, the condition implies that the standard deviation of g is 0.4. This value is not unreasonable. Moskowitz and Vissing-Jørgensen [36] estimate the annual standard deviation of returns for the smallest decile of public firm in the period 1953–1999 to be 41.4%, and Davis, Haltiwanger, Jarmin, and Miranda [18] estimate the dispersion of employment growth rates across firms to be 39% for 1984–1986, while Pareto exponent, estimated by top 1 percentile and 0.1 percentile income, is 1.98 in 1985.

The condition $z = \bar{k}\mathrm{Var}(g_{i,t})/2$ is further interpreted as follows. The right-hand side expresses the growth of capital due to the diffusion effect. We interpret this term as capital income due to the risk-taking behavior. The left-hand side z represents savings from the labor income. Then, the Pareto exponent is determined as 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 taking. The ratio of the two contributions, $(z/\bar{k})/(\mathrm{Var}(g_{i,t})/2)$, is inversely related to $1-(1-\delta)/\gamma$, as can be derived from (12) and (19). Thus, both growth γ and depreciation δ enhance wealth accumulation more by the risk-taking income than by wage income. This provides the mechanism for comparative statics in Proposition 4.

When $a_{i,t}$ follows a log-normal distribution, g is approximated in the first order by a log-normal distribution around the mean of $a_{i,t}$. We explore the formula for λ under the first-order approximation. From condition (16), we obtain

$$\lambda \approx -\frac{\mathrm{E}(\log g)}{\mathrm{Var}(\log g)/2}. \tag{20}$$

Note that for a log-normal g, we have $\log \mathrm{E}(g) = \mathrm{E}(\log g) + \mathrm{Var}(\log g)/2$. Thus, (20) indicates that λ is determined by the relative importance of the drift and diffusion of capital growth rates, both of which contribute to the overall growth rate. Using the condition $\mathrm{E}(g) = 1 - z/\bar{k}$, we obtain an alternative expression $\lambda \approx 1 + \frac{-\log(1-z/\bar{k})}{\mathrm{Var}(\log g)/2}$ as in Gabaix [23]. We observe that the Pareto exponent λ is always greater than 1, and it declines to 1 as savings z decreases to 0 or the diffusion effect $\mathrm{Var}(\log g)$ increases to infinity. For a small $z/\bar{k}$, the expression is further approximated as

$$\lambda \approx 1 + \frac{z}{\bar{k}\mathrm{Var}(\log g)/2}. \tag{21}$$

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

2.4. Implications of analytical results on Pareto exponent

The intuition for the mechanism to generate a Pareto distribution is as follows. As indicated by the extensive literature on Pareto distribution, the most natural mechanisms for the right-skewed, heavy-tailed distribution of income and wealth is the multiplicative process. However, without some modification, the multiplicative process leads to a log-normal process and neither generates the Pareto distribution nor the stationary variance of relative income. Incorporating a concave consumption function results in this modification. In the present Solow model, savings from wage income z serve as a reflective lower bound of the multiplicative wealth accumulation.

The close connection between the multiplicative process and the Pareto distribution may be illustrated as follows. The Pareto distribution implies a self-similar structure of distribution in terms of change of units. If we consider a "millionaire club" where all the members earn more than a million, under Pareto exponent λ, $10^{-\lambda}$ of the club members earn 10 times more than a million. If $\lambda = 2$, this is one percent of all members. Now consider a ten-million earners club, and we find again that one percent of the club members earn 10 times more than ten million. This observation is in contrast with the "memoryless" property of an exponential distribution that characterizes the middle-class distribution well. For the population who earn more than x in the exponential region, the fraction of population who earn more than $x + y$ is constant regardless of x. The contrast between the Pareto distribution and the exponential distribution corresponds to the fact 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 (see Levy and Solomon [33]).

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 $k_{i,t}$ has a finite mean in this 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. The Pareto distribution is "top heavy" in that a sizable fraction of the total wealth is possessed by the richest few. The richest P fraction of population owns $P^{1-1/\lambda}$ fraction of the total wealth when $\lambda > 1$ (Newman [37]). For $\lambda = 2$, this implies that the top 1 percent owns 10% of the 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. Further, when $\lambda < 1$, the expected ratio of the single richest person's wealth to the economy's total wealth converges to $1 - \lambda$ (Feller [21, p.172]). Such an economy almost resembles an aristocracy where a single person owns a big fraction of the 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 indicate that the Pareto exponent transits below and above 2, in the range between 1.5 and 3.[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

[3] See, for example, Alvaredo et al. [2], Fujiwara et al. [22], and Souma [49].

qualitatively, since for $\lambda < 2$, almost the entire sum of the variances of idiosyncratic risks is borne by the wealthiest few whereas the risks are more evenly distributed for $\lambda > 2$. This can be seen as follows. In this economy, the households do not diversify investment risks. Thus, their income variance increases as the square of their wealth $k_{i,t}^2$, which follows a Paretian tail with exponent $\lambda/2$. Thus, given $\lambda < 2$, 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 the entire sum of the variances is borne by the richest few percentiles. Thus, in this model, the concentration of wealth can be interpreted as the result of the concentration of risk bearings in terms of the variance of income.

Equation (21) demonstrates that the Pareto exponent is determined by the balance between two forces: the contributions of an additive term (z) and a diffusion term ($k\mathrm{Var}(g)/2$). Influx of wealth from labor income constitutes the additive term, which increases mobility between the tail wealth group and the rest and thus, has an equalization effect in the tail. An inequalizing diffusion effect results from capital income due to risk-taking. These two forces are depicted in Figure 2. In Section 3, we use this mechanism for interpreting the comparative statics obtained in the numerical simulations of the Bewley model. Moreover, we will compare the simulated Pareto exponents with the estimate given by (21).

3. Quantitative investigation of the Pareto distribution

3.1. Bewley model with idiosyncratic investment shocks and borrowing constraints

In reality, household saving behavior depends on wealth level, tax rate, and risk environment, and it has important implications on the Pareto exponent. In order to

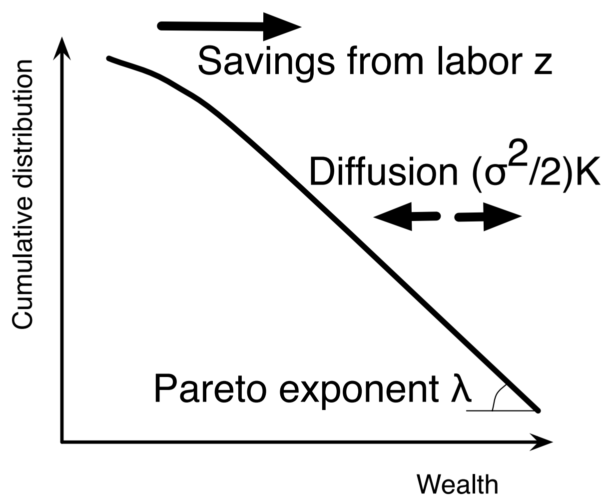

Figure 2: Determination of the Pareto exponent λ. Influx of wealth by savings raises λ, while diffusion effect lowers λ

incorporate the households' optimal savings choice, we depart from the Solow model and develop a Bewley model with idiosyncratic investment shocks and borrowing constraints. The model specification is largely unchanged from Section 2, except for the formulation of the household's dynamic optimization and serially correlated exogenous shocks on productivity and employment hours.

Household i inelastically supplies $e_{i,t}$ unit of labor, which follows an exogenous autoregressive process: $e_{i,t} = 1-\zeta+\zeta e_{i,t-1}+\epsilon_{i,t}$. The unconditional mean of individual labor supply, and thus the aggregate labor supply at the steady state, is normalized to 1. Households' production function bears idiosyncratic productivity shock, $a_{i,t}$, which follows a two-state Markov process. The households have no means to insure against idiosyncratic shocks $a_{i,t}$ and $\epsilon_{i,t}$ except for their own savings.

Household i can hold assets in the form of physical capital $k_{i,t}$ and bonds $b_{i,t}$. At the optimal labor hiring $l_{i,t}$, the return to physical capital is defined as

$$r_{i,t} \equiv \pi_{i,t}/k_{i,t} + 1 - \delta = \alpha(1-\alpha)^{(1-\alpha)/\alpha}(a_{i,t}/w_t)^{(1-\alpha)/\alpha} + 1 - \delta. \tag{22}$$

The bond bears a risk-free interest R_t. The households can engage in lending and borrowing through bonds, but the borrowing amount (detrended) must not exceed a borrowing limit ϕ, that is, $b_{i,t+1} > -\phi$.

Each household lineage is discontinued with a small probability μ in each period. At this event, a new household is formed at the same index i with no wealth. Following the perpetual youth model, we assume that the households participate in a pension program. The households contract all the non-human wealth to be confiscated by the pension program at the discontinuation of the lineage, and they receive in return a premium at rate p per unit of wealth they own in each period of continued lineage. The pension program is a pure redistribution system, and must satisfy the zero-profit condition $(1-\mu)p = \mu$. Thus, the pension premium rate is determined as

$$p = \mu/(1-\mu). \tag{23}$$

We incorporate progressive income taxation using a variation of Bénabou's [7] specification. The net tax payment is a function of household income $I_{i,t} = (r_{i,t}-1)k_{i,t} + (R-1)b_{i,t} + w_t e_{i,t}$ as follows:

$$T_{i,t} = \begin{cases} I_{i,t} - \tau_0 I_{i,t}^{1-\tau_1} & \text{if } I_{i,t} < I^* \\ I^* - \tau_0 I^{*1-\tau_1} + \tau_2(I_{i,t} - I^*) & \text{if } I_{i,t} \geq I^*. \end{cases} \tag{24}$$

The first convex part and the second linear part smoothly join at $I^* \equiv (\tau_0(1-\tau_1)/(1-\tau_2))^{1/\tau_1}$ with derivative τ_2, which denotes the highest marginal tax rate applied for the highest income bracket. We assume that the tax proceeds are spent on unproductive government purchase of goods.

Given the optimal operation of physical capital in each period, the households solve the following dynamic programming problem:

$$V(W, a, \epsilon) = \max_{c,k',b',W'} \frac{c^{1-\sigma}}{1-\sigma} + \tilde{\beta} E\left(V(W', a', \epsilon') \mid a, \epsilon\right) \tag{25}$$

subject to

$$c + \gamma(k' + b' + \phi) = W, \tag{26}$$

$$W = (1+p)(rk + Rb + we - T) + \gamma\phi, \tag{27}$$

$$b' + \phi > 0, \tag{28}$$

where $\tilde{\beta}$ is a modified discount factor $\tilde{\beta} \equiv \beta\gamma^{1-\sigma}(1-\mu)$. $W_{i,t}$ denotes the total resources available to i at t (the cash-at-hand). The control variables k_i and b_i can be equivalently expressed by i's total financial assets $x_i \equiv k_i + b_i + \phi$ and portfolio $\theta_i \equiv k_i/x_i$. Thus, the dynamic programming solves the optimal savings problem for x_i and the portfolio choice for θ_i.

An equilibrium is defined as a value function V, policy functions (x, θ), price functions (w, R), a joint distribution function Λ, and the law of motion Γ for Λ such that $V(W_i, a_i, \epsilon_i; \Lambda)$, $x(W_i, a_i, \epsilon_i; \Lambda)$, and $\theta(W_i, a_i, \epsilon_i; \Lambda)$ solve the household's dynamic programming, such that prices $w(\Lambda)$ and $R(\Lambda)$ clear the markets for goods, labor $\int_0^1 l_{i,t} di = \int_0^1 e_{i,t} di = 1$, and bonds $\int_0^1 b_{i,t} di = 0$, and such that the policy functions and the exogenous Markov processes of a_i and ϵ_i constitute Γ, which maps the joint distribution of $\Lambda(W_i, a_i, \epsilon_i)$ to that in the next period. A stationary equilibrium is defined as a particular equilibrium, wherein Λ is a fixed point of Γ.

3.2. Bewley model without borrowing constraints

The Bewley model is analytically tractable when there is no borrowing constraint. We will show that wealth follows a log-normal process if there is no limit on borrowing and if $\mu = 0$. This log-normal process implies that no stationary distribution of relative wealth exists. When $\mu > 0$, the stationary distribution of wealth is shown to have a Pareto tail, and the Pareto exponent is analytically derived.

We concentrate on a special case with no tax (i.e., $T_{i,t} = 0$), constant labor supply $e_{i,t} = 1$, and i.i.d. productivity $a_{i,t}$ over i and t. Because of the i.i.d. shocks, we have the aggregate production relation (5) as in the Solow model. Since this model features a utility exhibiting constant relative risk aversion, the savings rate and portfolio decisions are independent of wealth levels if there is no limit on borrowing (Samuelson [48]; Merton [35]). Here, we draw on Angeletos' [3] analysis. Let H_t denote human wealth, defined as the present value of future wage income stream:

$$\tilde{H}_t \equiv \sum_{\tau=t}^{\infty} \gamma^{\tau} w_{\tau} (1-\mu)^{\tau-t} \prod_{s=t+1}^{\tau} R_s^{-1}. \tag{29}$$

where wage w_τ is detrended by the growth factor γ. Define the detrended human wealth $H_t = \tilde{H}_t/\gamma^t$. Then, the evolution of human wealth satisfies $H_t = w_t + (1-\mu)\gamma R_{t+1}^{-1} H_{t+1}$. We define a household's total wealth (detrended) as

$$W_{i,t} = (1+p)(r_{i,t}k_{i,t} + R_t b_{i,t}) + H_t. \tag{30}$$

Consider a balanced growth path at which R_t, w_t, and H_t are constant over time. In this case, the dynamic programming problem allows the following linear solution with constants s and ϕ:

$$\begin{aligned} c &= (1-s)W, && (31)\\ k' &= \frac{\phi s}{\gamma} W, && (32)\\ b' &= \frac{(1-\phi)s}{\gamma} W - (1-\mu)R^{-1}H. && (33) \end{aligned}$$

By substituting the policy functions in the definition of wealth (30), and by noting that $(1-\mu)(1+p) = 1$ holds from the zero-profit condition for the pension program (23), I obtain the equation of motion for the detrended individual total wealth:

$$W_{i,t+1} = \begin{cases} \tilde{g}_{i,t+1} W_{i,t} & \text{with prob. } 1-\mu \\ H & \text{with prob. } \mu, \end{cases} \tag{34}$$

where the growth rate is defined as

$$\hat{g}_i' \equiv \frac{(\phi r_i' + (1-\phi)R)s}{(1-\mu)\gamma}. \tag{35}$$

Thus, at the balanced growth path, household wealth evolves multiplicatively according to (34) 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 [34], the Pareto exponent of the wealth distribution is determined as follows.[4]

Proposition 5. *A household's detrended total wealth $W_{i,t}$ has a stationary distribution with Paretian tail exponent λ, which is determined by*

$$(1-\mu)\mathrm{E}(\hat{g}_{i,t}^{\lambda}) = 1 \tag{36}$$

if $\mu > 0$. If $\mu = 0$, $W_{i,t}$ has no stationary distribution and asymptotically follows a log-normal distribution with diverging variance.

Proof: See Appendix E.

We note that if there is no discontinuation event (i.e., $\mu = 0$) as in Angeletos' [3] benchmark model, individual wealth follows a log-normal process with log-mean and log-variance increasing linearly in t. Therefore, the relative wealth $W_{i,t}/\int W_{j,t}dj$ does not have a stationary distribution. In this case, a vanishingly small fraction of individuals eventually possesses almost all the wealth. This is not consistent with the empirical observations that the variance of log-income is stationary over time, as Kalecki [29] pointed out. One way to avoid the diverging variance is to introduce $\mu > 0$ as seen above. Another way is to introduce borrowing constraints, as we show in the next section.

3.3. Borrowing constraints and Pareto distribution

In this section, we show that the Bewley model generates the Pareto distribution even when $\mu = 0$, if households face borrowing constraints. The key element for generating the Pareto distribution is a concave consumption function, as implied in the Solow

[4] We thank Wataru Souma for pointing to this reference. This result can be seen as a discrete-time analogue of the stationary Pareto distribution of a geometric Brownian motion with random life-time, as explained in Reed [47] and applied to the overlapping generations model by Benhabib et al. [10], although the geometric Brownian model differs in that it generates a double Pareto distribution.

model. When there is a borrowing constraint, the consumption function is concave in wealth whereas it is linear without borrowing constraints. As Carroll and Kimball [13] argue, the linear consumption function arises in a quite narrow specification of the Bewley model. For example, a concave consumption function arises when the labor income is uncertain or when the household's borrowing is constrained. This implies that the log-normal process of wealth is a special case whereas the Pareto distribution characterizes a wide class of model specifications.

Since the Bewley model with borrowing constraint is difficult to solve analytically, we numerically solve for a stationary equilibrium. This model features a multiplicative investment shock, in addition to an endowment shock that enters the wealth accumulation process additively as in Aiyagari [1]. Thus, stationary wealth distribution has a fat tail unlike the Aiyagari economy. This means that the simulation of wealth accumulation process suffers a slow convergence of aggregate wealth, since the aggregated noise in a fat tail does not decrease as quickly as the simulated population increases. However, if the wealth state is discretized in logarithmic space, the stationary distribution can be computed well simply by iterating the multiplication of the Markov transition matrix. Intuitively, this is because the logarithm of a multiplicative process falls back to an additive process.

To manage the computation of portfolio choice, we follow a two-step approach similar to Barillas and Fernández-Villaverde [6], who solve the neoclassical growth model with labor choice using the endogenous gridpoints method used by Carroll [12] for the savings problem and the standard value function iteration for the labor choice. The autoregressive process of labor supply $e_{i,t}$ is approximated by a five-state Markov process following the Rouwenhorst method (Kopecky and Suen [32]).

With autocorrelation in productivity $a_{i,t}$, households with high productivity will invest in capital at a high rate of borrowing, while the households with low productivity will shift their assets to risk-free bonds. Thus, this model captures an economy wherein a fraction of the households choose to become entrepreneurs while the other households rely on wage and returns from safe assets as their main income source. Since the entrepreneurs bear the investment shocks that generate the fat tail of wealth distribution in this model, we observe that the tail population largely consists of current and past entrepreneurs. As a model of entrepreneurship, the model presented here is not as rich as the one with occupational choice (see Quadrini [45] for a survey). Nonetheless, in this model, the entrepreneurs (households with high productivity) do not diversify much of their investment risks while workers choose to bear substantially smaller risks.

We compute the stationary equilibrium distributions of wealth $W_{i,t}$ and income $I_{i,t}$. To calibrate the taxation function, we use the estimate by Heathcote, Storesletten, and Violante [26], $\tau_1 = 0.151$. We set $\tau_0 = 0.9$ so the government expenditure is about 10% of GDP. The highest marginal tax rate is specified as $\tau_2 = 0.5$ to emulate the rate before the tax cut in 1986 in the U.S. The labor endowment process is calibrated as $\zeta = 0.82$ and $\mathrm{Std}(\epsilon_{i,t}) = 0.29$, following Guvenen [25]. The transition matrix Π for the productivity shock a_i is set by $\pi_{11} = 0.9727$ and $\pi_{22} = 0.8$, for which the stationary fraction of households with high productivity is 12% and the average exit rate from the high productivity group is 20%. These numbers correspond to the fraction and exit rate of entrepreneurs in the U.S. data (Kitao [31]). The states of $a_{i,t}$ are set at $\{0.75, 1.25\}$, which corresponds to 10% standard deviation in risky asset returns. At this volatility of productivity shocks, the stationary wealth distribution in the model with tax rate $\tau_2 = 0.5$ generates a Pareto exponent 2, which roughly matches with the U.S. level right before the tax cut in 1986. The lineage discontinuation rate μ is set at 2%. The borrowing constraint is set at $\phi = 0.19$, which is worth three months wage income. At this value, the difference in consumption growth rates between low-asset and high-asset groups matches with Zeldes' estimate [38, 52]. The parameters on technology and preferences are set at standard values as $\alpha = 0.36$, $\delta = 0.1$, $\sigma = 3$, $\beta = 0.96$, and $\gamma = 1.02$.

The wealth distribution at stationary equilibrium for this benchmark calibration is plotted in the top left panel of Figure 3. Pareto distributions are clearly observed in the right tail of both income and wealth for top 1% of the distribution (i.e., beneath 10^{-2} in the inverse cumulative distribution). The Pareto exponent $\lambda = 2$ is seen in the plot, since the logarithm of inverse cumulative probability decreases by 2 decades as the log-wealth increases by a decade.

In the same panel, the wealth distribution for the case $\mu = 0$ is also shown. This plot demonstrates that the Pareto distribution can be generated even when the households are infinitely living. As proven in the previous section, the wealth distribution of infinitely living households follows a log-normal distribution with diverging log-variance if the consumption function is linear. Thus, the borrowing constraint and the concave consumption function play a key role in generating Pareto distribution when $\mu = 0$. The borrowing constraint also has a quantitatively considerable impact on the Pareto exponent. In the same panel in Figure 3, wealth distribution for the case $\phi = 0.75$ is plotted. This value of the borrowing limit corresponds to an annual wage income. The plot indicates that relaxing the borrowing constraint from three months wage to a year's wage has roughly the same impact on the Pareto exponent as reducing the death rate from $\mu = 0.02$ to 0.

Table 1 compares income distribution for the benchmark case ($\tau_2 = 0.5$) with income distribution in the U.S. in 1985. As can be seen, the simulated and empirical

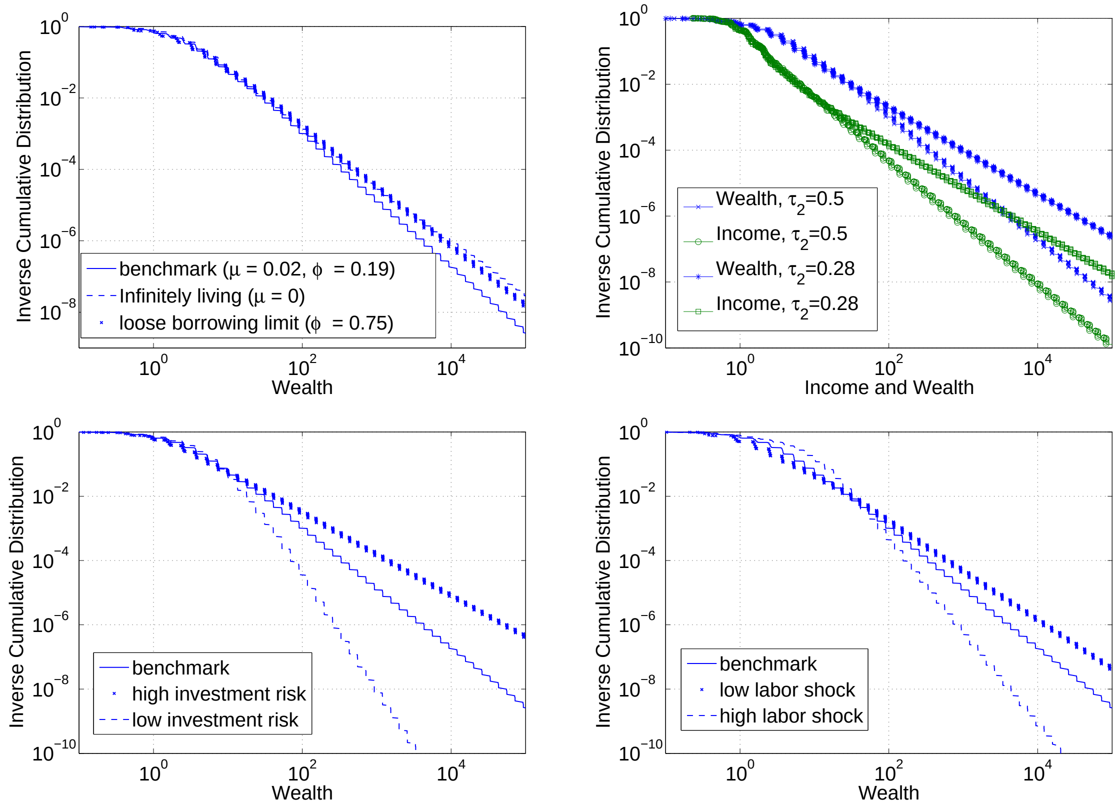

Figure 3: Simulated stationary distributions of wealth and income. *Top left*: The benchmark case ($\mu = 0.02$ and $\phi = 0.19$), the infinitely living case ($\mu = 0$ and $\phi = 0.19$), and the case of relaxed borrowing limit ($\mu = 0.02$ and $\phi = 0.75$). *Top right*: The benchmark case and the case of lower tax rate ($\tau_2 = 0.28$). *Bottom left*: Wealth distributions when standard deviation of productivity $\mathrm{Std}(a_{i,t})$ is set at 0.15 (low), 0.25 (benchmark), and 0.35 (high). *Bottom right*: Wealth distributions when variance of labor income shock $\mathrm{Var}(\epsilon_{i,t})$ is set at 0.145 (low), 0.29 (benchmark), and 0.58 (high).

	p20	p40	p60	p80	p95	Gini	Top I share	Top W share
Benchmark	0.660	0.928	1.334	1.691	3.051	0.402	0.100	0.176
Low tax	0.622	0.890	1.316	1.606	2.906	0.424	0.143	0.305
US 1985	0.421	0.792	1.227	1.845	3.049	0.419	0.127	$(0.301)_{1980}$
US 2010	0.406	0.771	1.248	2.030	3.663	0.469	0.198	0.338

Table 1: Characteristics of simulated and U.S. distributions. The table lists quintiles of income $I_{i,t}$, 95 percentile income, Gini index of income, and top 1% shares of income and wealth $W_{i,t}$. Percentile income is measured relative to the median income. Sources of the U.S. estimates are Census for the percentiles and Gini index, and Piketty [42] for the top shares. The estimate for 1980 is shown in parentheses; the top W share estimate for the U.S. for 1985 is missing.

distributions reasonably resemble each other at the quintiles and at the 95 percentile, as well as in the Gini index and the top one percent shares of income and wealth. Table 1 also reports the simulated distribution for the low tax case ($\tau_2 = 0.28$) and the empirical distribution in 2010. We observe that the wealth share increases significantly in our simulation with low tax rate.

The top-right panel of Figure 3 plots the distributions of income and wealth at stationary equilibrium for top marginal tax rate $\tau_2 = 0.5$ (benchmark) and 0.28. We observe that the Pareto exponents for income and wealth coincide. That is because in this model high income earners earn most of the income from capital. The Pareto exponent is significantly smaller in the low tax regime than in the high tax regime: 1.8 for $\tau_2 = 0.28$ and 2 for $\tau_2 = 0.5$.

We conduct further sensitivity analyses on the stationary wealth distributions. The bottom-left panel of Figure 3 shows that the increased variance of productivity shock ($\mathrm{Var}(a_{i,t})$) leads to less equal tail distributions, indicated by the flatter tail.

The bottom-right panel shows that the increased variance of labor endowment shock ($\mathrm{Var}(\epsilon_{i,t})$) results in more equal tail distributions.

3.4. Interpretation of comparative statics

Summarizing the observations in Figure 3, we find that the Pareto exponent decreases (i.e., the tail is inequalized) under low capital tax rate, high investment risk, low labor risk, or loose borrowing constraint. We interpret these comparative statics by using the scheme depicted by Figure 2: tax and investment risk are categorized as the diffusion effect, while labor risk and borrowing constraint as the influx effect. We explain them in turns.

When the investment risk is high, the volatility of capital return increases, because the mitigating effect of reduced capital portfolio is weak under our calibrations. Thus, the volatility of growth rate of wealth increases, which results in the lower Pareto exponent. The low tax rate also strengthens the diffusion effect, because it increases the volatility of after-tax returns of capital.

Two tax rates used for the top-right panel of Figure 3 emulate the U.S. Tax Reform Act in 1986. As studied by Feenberg and Poterba [20], an unprecedented decline in the Pareto exponent is observed right after the tax reform. Although the stable Pareto exponent right after the downward leap may suggest 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 Reform Act. Taxation has a direct effect on wealth accumulation by lowering the after-tax increment of wealth and the effect through the altered incentives that households face. Piketty and Saez [43] suggested that the imposition of progressive tax around the Second World War was the possible cause of decline in the top income share during this period, which continued to remain at a low level for a long time until the 1980s. Our simulations and the above analytical results are consistent with the view that the tax cut substantially reduces the stationary Pareto exponent. However, our analysis is limited to the comparison of stationary distributions, and the transition dynamics is out of the scope of this paper.[5]

The low labor risk and loose borrowing constraint affect the influx effect through the precautionary motive of savings. Household has less incentive for precautionary savings when the labor risk is low or the borrowing constraint is loose. Hence, the saving rate among the low and middle asset groups falls, which reduces the influx of labor income into wealth and decreases the Pareto exponent at the tail. This result contrasts with Benhabib et al. [8] who claimed that the labor income risk does not affect tail. The irrelevance of labor risk holds only in an environment where the consumption function is linear, just as in our Solow model. When there is a borrowing limit, the labor risk affects the tail, and our numerical result shows that its impact is quantitatively considerable.

3.5. Discussions

The above interpretations of the sensitivity analysis assume that the analysis in the Solow can be extended to the simulated Bewley economy, at least qualitatively. Since Bewley economy is complex enough, we cannot justify this extension rigorously. However, the derived formula for the Pareto exponent in the Solow model shows some qualitative agreement with the simulated results. Table 2 shows the Pareto exponents obtained in simulations and those obtained by calculating $(wK/(K+Y))/(Kr_{high}(1-\tau_2))$. The numerator expresses the savings from labor income, by approximating the saving rate by $K/(K+Y)$. The denominator proxies for the after-tax capital income due to risk premium. While admittedly these approximations are rough, especially in

[5]We tackle this issue in a different paper Aoki and Nirei [4].

	benchmark	low τ	low ϕ	high a volatility	high e volatility
Simulation	2.13	1.82	2.17	1.71	2.41
Solow	1.94	1.36	1.95	1.36	2.90

Table 2: Pareto exponents in simulations of Bewley models and predictions by the Solow model

not taking account of non-linear saving and taxation functions, Table 2 suggests that the formula predicts the direction of change in Pareto exponent reasonably well.

A note is in order for the effect of savings s on the Pareto exponent in the Solow model. Proposition 6 showed that the savings rate per se does not affect λ at the stationary distribution. This is because the savings rate in the Solow model affects both returns to wealth, through reduced reinvestment, and savings from labor income. These two effects cancel out in the determination of λ. In the Bewley model, we argue that precautionary savings serve as the reflective lower bound for wealth accumulation. When precautionary savings are present, an exogenous change in, say, investment risks, induces more savings in the low-wealth group than the high-wealth group, which affects the balance between savings from labor income and asset income and thus, changes the Pareto exponent. What actually matters for the comparative statics of λ is the differential response in saving rate between the high- and low-wealth groups.

Finally, we discuss an implication of Proposition 5 for unit-root process of income. In the benchmark model, we employ heterogeneous income profiles specification for the exogenous labor endowment process. An alternative is the restricted income profiles (RIP) specification, in which the logarithm of labor income process exhibits unit root. If the log labor income follows a unit root process with stochastic death, Proposition 5 implies that the stationary labor income distribution follows a Pareto distribution. Therefore, it is possible that RIP specification generates Pareto exponent quantitatively comparable to empirical income distribution. A back-of-the-envelope calculation of the Pareto exponent of labor endowment by using (36) generates $\lambda = 1.76$ for a RIP process $\log e' = \log e + \epsilon$, when the variance of ϵ is set at 0.03, following Hryshko [27] and $\mu = 0.02$. However, the implication on wealth distribution is not immediately clear unless we incorporate this process in the Bewley model to determine the savings and portfolio policies. This would require an extension of the state space for labor endowment as wide as wealth state. Therefore, fully implementing RIP in the current model is computationally too demanding. We leave it for future research to further investigate the alternative specification for income process.

4. Conclusion

This paper demonstrates that the neoclassical growth model with idiosyncratic investment risks is able to generate the Pareto distribution as the stationary distributions of income and wealth 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 two factors: savings from labor income, which determines the influx of population from the middle class to the tail part, and 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 this paper features risky investments in physical capital, the Paretian tail is similarly obtained when the risky asset takes the form of human capital. The essential feature of the model is that the households own a stock factor with risky returns and a flow factor for production. The risky returns generate the diffusion effects, while the flow factor provides the influx effect. The redistribution policy financed by income or bequest tax raises the Pareto exponent, because the tax reduces the diffusion effect. Similarly, increased risk sharing raises the Pareto exponent.

The analytical results shown in the Solow model hold in a Bewley model, wherein the savings rate is optimally determined by the households. In a benchmark case without borrowing constraints, the Bewley model generates a log-normal process for individual wealth that implies a counterfactual "escaping" inequalization. By incorporating a random event by which each household lineage is discontinued, we analytically reestablish the Pareto distribution of wealth and income. When borrowing constraints are introduced, the model generates the Pareto distribution due to two forces: households' discontinuation and precautionary savings. We conduct sensitivity analyses of the Pareto exponent for death rate, tax rate, return volatility, and labor endowment volatility by simulations. The simulated results agree with the mechanism of the determination of the Pareto exponent analyzed by the Solow model.

The agreement between the Solow model and the Bewley model with borrowing constraints points to the key role played by the concavity of consumption function in generating the Pareto distribution. The tighter borrowing limit leads to greater concavity of consumption function and larger precautionary savings. Savings by the low wealth group correspond to the savings of households with no wealth in the Solow model, which serve as a reflective lower bound of wealth accumulation and exert the influx effect on the Pareto exponent. Our simulations with varying borrowing constraints show that this effect can be quantitatively considerable.

Appendix A. Proof of Proposition 2

We first show the unique existence of solution λ for (16). Note that $(d/d\lambda)\mathrm{E}(g^\lambda) = \mathrm{E}(g^\lambda \log g)$ and $(d^2/d\lambda^2)\mathrm{E}(g^\lambda) = \mathrm{E}(g^\lambda(\log g)^2)$. Since $g > 0$, the second derivative is positive, and thus, $\mathrm{E}(g^\lambda)$ is convex in λ. 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, $\mathrm{E}(g^\lambda)$ eventually becomes greater than 1 as λ increases to infinity. Further, recall that $\mathrm{E}(g) < 1$. Thus, for the range $\lambda > 1$, $\mathrm{E}(g^\lambda)$ is a continuous convex function that travels from below 1 to above 1. This establishes that the solution for $\mathrm{E}(g^\lambda) = 1$ exists uniquely in the range $\lambda > 1$, and that the solution $\bar{\lambda}$ satisfies $(d/d\lambda)\mathrm{E}(g^\lambda)|_{\lambda=\bar{\lambda}} > 0$.

Next, 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,

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

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

$$\begin{aligned} \int^{\bar{g}} \frac{\partial F}{\partial \sigma_u} dg &= \int^{\bar{x}} \phi(x)(-x/\sigma_u + 1) dx dg/dx & \text{(A.2)} \\ &= \sigma_u e^{u_0} \int^{\bar{x}} \frac{-x/\sigma_u + 1}{\sqrt{2\pi}} e^{-(x-\sigma_u)^2/2} dx, & \text{(A.3)} \end{aligned}$$

The last line reads as a partial moment of $-x/\sigma_u + 1$, wherein x follows a normal distribution with mean σ_u and variance 1. The integral tends to 0 as $\bar{x} \to \infty$, and the integrand is positive for x below σ_u and negative above 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 the assertion 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 was observed, $\mathrm{E}(g^\lambda)$ is also strictly increasing in λ locally at $\lambda = \bar{\lambda}$. Thus, an increase in σ_u, and thus, an increase in σ while α is fixed, results in a decrease in $\bar{\lambda}$ that satisfies $\mathrm{E}(g^{\bar{\lambda}}) = 1$.

Appendix B. Proof of Proposition 3

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

$$
\begin{aligned}
1 = \mathrm{E}(g^\lambda) &= e^{\lambda(\log \alpha - \log \mathrm{E}(a_{i,t}^{(1-\alpha)/\alpha}) - (\sigma^2/2)((1-\alpha)/\alpha)) + \lambda^2((1-\alpha)/\alpha)^2\sigma^2/2} && \text{(B.1)} \\
&= e^{\lambda(\log \alpha - (\sigma^2/2)((1-\alpha)/\alpha)^2) + \lambda^2((1-\alpha)/\alpha)^2\sigma^2/2}. && \text{(B.2)}
\end{aligned}
$$

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

$$
\lambda = 1 - \left(\frac{\alpha}{1-\alpha}\right)^2 \frac{\log \alpha}{\sigma^2/2}. \tag{B.3}
$$

Appendix C. Proof of Proposition 4

From the definition of $g_{i,t}$ (12), we obtain

$$\mathrm{E}(g_{i,t}^2) = \left(\frac{1-\delta}{\gamma}\right)^2 + 2\frac{1-\delta}{\gamma}\frac{\alpha}{\gamma}(\gamma - 1 + \delta) + \left(\frac{\alpha}{\gamma}(\gamma - 1 + \delta)\right)^2 \frac{\mathrm{E}(a_{i,t}^{2(1-\alpha)/\alpha})}{\left(\mathrm{E}(a_{i,t}^{(1-\alpha)/\alpha})\right)^2}. \quad \text{(C.1)}$$

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

$$\frac{\mathrm{E}(a_{i,t}^{2(1-\alpha)/\alpha})}{\left(\mathrm{E}(a_{i,t}^{(1-\alpha)/\alpha})\right)^2} = \frac{e^{-\sigma^2(1-\alpha)/\alpha + 2\sigma^2(1-\alpha)^2/\alpha^2}}{e^{-\sigma^2(1-\alpha)/\alpha + \sigma^2(1-\alpha)^2/\alpha^2}} = e^{\left(\frac{\sigma(1-\alpha)}{\alpha}\right)^2}. \quad \text{(C.2)}$$

Combining these results with the condition $\mathrm{E}(g^2) = 1$, after some manipulation we obtain

$$\hat{\sigma}^2 = \left(\frac{\alpha}{1-\alpha}\right)^2 \log\left(\frac{1}{\alpha^2}\left(1 + \frac{2(1-\alpha)}{\gamma/(1-\delta) - 1}\right)\right). \quad \text{(C.3)}$$

We observe that an increase in γ or δ decreases $\hat{\sigma}$. By Proposition 2, λ is decreasing in σ. Thus, in the neighborhood of $\lambda = 2$, the stationary λ is decreasing in γ or δ, because an increase in either γ or δ decreases $\hat{\sigma}$ relative to the current level of σ.

Appendix D. Redistribution policies in the Solow model

In this section, we extend the Solow framework to redistribution policies financed by taxes on income and bequest. This extension provides an analytical framework to interpret the numerical results on income tax in the Bewley model in Section 3. Moreover, by this extension we incorporate the classic argument by Stiglitz [50] on complete equalization in the Solow model as a limiting case.

Let τ_y denote the flat-rate tax on income, and let τ_b denote the flat-rate bequest tax on inherited wealth. The tax proceeds are equally redistributed among the households. We assume that a household changes generations with probability μ in each period.

Thus, the household wealth is taxed at flat rate τ_b with probability μ and remains intact with probability $1-\mu$. We 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 (7) is modified as follows:

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

By aggregating, we recover the law of motion for K_t as in (8). Therefore, the redistribution policy does not affect $\bar{K}$ or aggregate output at the steady state. Combining with (D.1), the accumulation equation for individual wealth is rewritten at $\bar{K}$ as follows:

$$k_{i,t+1} = \tilde{g}_{i,t}k_{i,t} + \tilde{z}e_{i,t}, \tag{D.2}$$

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

$$\tilde{g}_{i,t} \equiv \frac{1-\delta-\mathbf{1}_b\tau_b}{\gamma} + \frac{(1-\tau_y)\alpha(\gamma-1+\delta)}{\gamma}\frac{a_{i,t}^{(1-\alpha)/\alpha}}{\mathrm{E}(a_{i,t}^{(1-\alpha)/\alpha})}, \tag{D.3}$$

$$\tilde{z} \equiv \frac{(1-\alpha+\alpha\tau_y)sA}{\gamma}\left(\frac{sA}{\gamma-1+\delta}\right)^{\alpha/(1-\alpha)} + \tau_b\mu\left(\frac{sA}{\gamma-1+\delta}\right)^{1/(1-\alpha)}. \tag{D.4}$$

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

Proposition 6. *Under the redistribution policy, a household's wealth $k_{i,t}$ 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 the 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 equalize the tail distribution.

The redistribution financed by bequest tax τ_b has an effect similar to that of a random discontinuation of household lineage. By setting τ_b accordingly, we can incorporate the situation where a household may have no heir, and 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, larger population longevity has an inequalizing effect on the tail wealth.

The redistribution financed by income tax τ_y essentially collects a fraction of profits and equally transfers the proceeds to the households. Thus, 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. The importance of capital market imperfections in determining income distributions is emphasized by Banerjee and Newman [5] and Galor and Zeira [24]. In this context, we obtain the following result.

Proposition 7. *Consider a risk-sharing mechanism that collects τ_s fraction of profits $\pi_{i,t}$ and refunds its ex-ante mean $\mathrm{E}(\tau_s \pi_{i,t})$ as rebate. Then, an increase in τ_s raises λ.*

The proof is as follows. Partial risk sharing ($\tau_s > 0$) reduces the weight on $\epsilon_{i,t}$ in (D.3) 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, 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, the model converges to the classic case of Stiglitz [50], wherein a complete equalization of wealth distribution takes place.

Appendix E. Proof of Proposition 5

In this section, we solve the Bewley model and show the existence of the balanced growth path. Then the proposition obtains directly by applying Manrubia and Zanette [34].

Household problem with "natural" borrowing constraint and pension program on the balanced growth path is formulated in a recursive form:

$$V(W,a) = \max_{c,k',b',W'} \frac{c^{1-\sigma}}{1-\sigma} + \tilde{\beta} E\left(V(W',a')\right) \tag{E.1}$$

subject to

$$\begin{aligned} c + \gamma(k'+b') + (1-\mu)\gamma R^{-1} H' &= W, & \text{(E.2)} \\ W &= (1+p)(rk + Rb) + H. & \text{(E.3)} \end{aligned}$$

At the steady state of detrended aggregate capital $\bar{K}$, the return to physical capital (22) is written as:

$$r_{i,t} = \alpha (a_{i,t}/A)^{(1-\alpha)/\alpha} \bar{K}^{\alpha-1} + 1 - \delta, \tag{E.4}$$

which is a stationary process. The average return is:

$$\bar{r} \equiv \mathrm{E}(r) = \alpha A \bar{K}^{\alpha-1} + 1 - \delta. \tag{E.5}$$

The lending market must clear in each period, which requires $\int b_{i,t} di = 0$ for any t. We also note that r_i is independent of k_i. 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:

$$\bar{W} = (1-\mu)^{-1} \bar{r} \bar{K} + \bar{H}. \tag{E.6}$$

Combining the market clearing condition for lending with the policy function for lending (33), we obtain the equilibrium risk-free rate:

$$R = \frac{\gamma(1-\mu)}{s(1-\phi)}\frac{\bar{H}}{\bar{W}}. \tag{E.7}$$

By using the conditions above and substituting the policy function (31), the budget constraint (E.2) becomes in aggregation:

$$(\gamma - s(1-\mu)^{-1}\bar{r})\bar{K} = (s - (1-\mu)R^{-1}\gamma)\bar{H}. \tag{E.8}$$

Plugging into (E.7), we obtain the relation:

$$R = \frac{\gamma(1-\mu)}{s(1-\phi)} - \frac{\phi}{1-\phi}\bar{r}. \tag{E.9}$$

Thus, the mean return to the risky asset and the risk-free rate are determined by $\bar{K}$ from (E.5,E.9). The expected excess return is solved as:

$$\bar{r} - R = \frac{1}{1-\phi}\left(\alpha A\bar{K}^{\alpha-1} + 1 - \delta - (1-\mu)\gamma/s\right). \tag{E.10}$$

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

By using (4,5), the human wealth is written as:

$$\bar{H} = \gamma^{-t}\left(\sum_{\tau=t}^{\infty}\bar{w}\gamma^{\tau}(1-\mu)^{\tau-t}\prod_{s=t+1}^{\tau}R_s^{-1}\right) = \frac{\bar{w}}{1-(1-\mu)\gamma R^{-1}} = \frac{(1-\alpha)A\bar{K}^{\alpha}}{1-(1-\mu)\gamma R^{-1}}. \tag{E.11}$$

Equations (E.5,E.8,E.9,E.11) determine $\bar{K}, \bar{H}, R, \bar{r}$. 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 (E.5,E.9,E.11), we have:

$$\frac{\bar{K}}{\bar{H}} = \frac{1 - \frac{(1-\mu)\gamma s(1-\phi)}{\gamma(1-\mu) - s\phi(1-\delta) - s\phi\alpha A\bar{K}^{\alpha-1}}}{(1-\alpha)A\bar{K}^{\alpha-1}}. \tag{E.12}$$

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

Now, the right hand side of (E.8) is transformed as follows:

$$\begin{aligned}\bar{H}(s-(1-\mu)\gamma R^{-1}) &= \bar{H}\left(s-s(1-\phi)\frac{\bar{W}}{\bar{H}}\right)=\bar{H}s\left(1-(1-\phi)\left((1-\mu)^{-1}\frac{\bar{r}\bar{K}}{\bar{H}}+1\right)\right)\\ &= \bar{H}s\left(\phi-(1-\phi)(1-\mu)^{-1}\frac{\bar{r}\bar{K}}{\bar{H}}\right). \qquad (E.13)\end{aligned}$$

Then we rearrange (E.8) as:

$$\frac{\gamma}{s\phi}\frac{\bar{K}}{\bar{H}}=1+(1-\mu)^{-1}\frac{\bar{r}\bar{K}}{\bar{H}}. \qquad (E.14)$$

By (E.5), $\bar{r}$ is strictly decreasing in $\bar{K}$, and R is strictly increasing by (E.9). Thus, $\bar{W}/\bar{H}$ is strictly decreasing by (E.7), and so is $\bar{r}\bar{K}/\bar{H}$ by (E.6). Thus, the right hand side of (E.14) is positive and strictly decreasing in $\bar{K}$. The left hand side is monotonically increasing from 0 to $+\infty$. Hence, there exists the steady-state solution $\bar{K}$ uniquely. This verifies the unique existence of the balanced growth path.

The law of motion (34) for the detrended individual total wealth $k_{i,t}$ is now completely specified at the balanced growth path:

$$k_{i,t+1}=\begin{cases}\tilde{g}_{i,t+1}k_{i,t} & \text{with prob. } 1-\mu\\ \bar{H} & \text{with prob. } \mu,\end{cases} \qquad (E.15)$$

where,

$$\tilde{g}_{i,t+1}\equiv(\phi r_{i,t+1}+(1-\phi)R)s/(1-\mu). \qquad (E.16)$$

This is the stochastic multiplicative process with reset events studied by Manrubia and Zanette [34]. By applying their result, we obtain the proposition.

Appendix F. Details of numerical computation

This section explains the computation procedure used for Section 3. Wealth W_i is discretized by 100 grid points separated equally in log-scale in the range between 10^{-2}

and 10^{10}. The autoregressive process of $e_{i,t}$ is discretized by Rouwenhorst's quadrature method (Kopecky and Suen [32]). To compute the stationary equilibrium of the Bewley model with portfolio choice, we use a two-step algorithm similar to Barillas and Fernández-Villaverde [6]. In the first step, we solve the savings choice given a portfolio policy, and we solve the portfolio policy given the savings choice in the second step.

1. Initialize $\theta(W, a, \epsilon)$
   (a) Initialize K
      i. Compute w and $r(a)$
      ii. Solve for household's savings policy $x(W, a, \epsilon)$
      iii. Compute the stationary distribution of (W_i, a_i, ϵ_i)
      iv. Compute stationary K. Repeat (a) until K converges
   (b) Initialize R
      i. Solve for household's portfolio policy $\theta(W, a, \epsilon)$
      ii. Compute aggregate bond demand $\int_0^1 b_i di$. Adjust R and repeat (b) until the aggregate bond demand converges to 0
2. Repeat 1 until $\theta(W, a, \epsilon)$ converges

In the algorithm above, (a-i) uses (E.4) and the profit maximization condition. In (a-ii), households' dynamic programming problem is solved by endogenous grid-point method with linear interpolation. In (a-iii), we first compute $W_{i,t+1}$ for given $W_{i,t}, a_{i,t}, a_{i,t+1}, e_{i,t}, e_{i,t+1}$, where $W_{i,t}$ is chosen at a grid point of wealth. The probability for this transition is computed by combining the transition matrices for a_i and the discretized e_i. Thus, we obtain a transition matrix for W_i. The stationary distribution of W_i is obtained by forward simulation, that is, by iterating the multiplication of the transition matrix. The resulting stationary distribution of wealth has a fat tail, but the average wealth always exists since the Pareto exponent is greater than 1. We take the maximum grid for wealth quite large (10^{10}) so that the aggregate impact of the computation error due to the finiteness is negligible. The measure of households occupying the largest grid is roughly $10^{-10\lambda}$. The wealth share held by those households, about $10^{-10(\lambda-1)}$, becomes negligible when λ is around 2.

The convergence of aggregate capital in (a) and aggregate bond in (b) are obtained by the bisection method applied to K and R, respectively. The criterion for convergence is set as 10^{-9} for savings and portfolio policy functions and 10^{-4} for aggregate capital and bond.

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

[2] Alvaredo, F., Atkinson, A.B., Piketty, T., Saez, E., 2014. The world top incomes database. Http://topincomes.g-mond.parisschoolofeconomics.eu/.

[3] Angeletos, G.M., 2007. Uninsured idiosyncratic investment risk and aggregate saving. Review of Economic Dynamics 10, 1–30.

[4] Aoki, S., Nirei, M., 2014. Zipf's law, Pareto's law, and the evolution of top incomes in the U.S. TCER Working Paper Series E-74.

[5] Banerjee, A.V., Newman, A.F., 1993. Occupational choice and the process of development. Journal of Political Economy 101, 274–298.

[6] Barillas, F., Fernández-Villaverde, J., 2007. A generalization of the endogenous grid method. Journal of Economic Dynamics & Control 31, 2698–2712.

[7] Bénabou, R., 1996. Inequality and growth. NBER Macroeconomics Annual 11.

[8] Benhabib, J., Bisin, A., Zhu, S., 2011. The distribution of wealth and fiscal policy in economies with finitely lived agents. Econometrica 79, 123–157.

[9] Benhabib, J., Bisin, A., Zhu, S., 2013. The wealth distribution in Bewley models with investment risk. Mimeo.

[10] Benhabib, J., Bisin, A., Zhu, S., 2014. The distribution of wealth in the Blanchard-Yaari model. Macroeconomic Dynamics Doi: 10.1017/S1365100514000066.

[11] Cagetti, M., Nardi, M.D., 2009. Estate taxation, entrepreneurship, and wealth. American Economic Review 99, 85–111.

[12] Carroll, C.D., 2006. The method of endogenous gridpoints for solving dynamic stochastic optimization problems. Economics Letters 91, 312–320.

[13] Carroll, C.D., Kimball, M.S., 1996. On the concavity of the consumption function. Econometrica 64, 981–992.

[14] Castañeda, A., Díaz-Giménez, J., Ríos-Rull, J.V., 1998. Exploring the income distribution business cycle dynamics. Journal of Monetary Economics 42, 93–130.

[15] Castañeda, A., Díaz-Giménez, J., Ríos-Rull, J.V., 2003. Accounting for the U.S. earnings and wealth inequality. Journal of Political Economy 111, 818–857.

[16] Champernowne, D., 1953. A model of income distributon. Economic Journal 63, 318–351.

[17] Covas, F., 2006. Uninsured idiosyncratic production risk with borrowing constraints. Journal of Economic Dynamics & Control 30, 2167–2190.

[18] Davis, S.J., Haltiwanger, J., Jarmin, R., Miranda, J., 2006. Volatility and dispersion in business growth rates: Publicly traded versus privately held firms. NBER Macroeconomics Annual 21.

[19] Dutta, J., Michel, P., 1998. The distribution of wealth with imperfect altruism. Journal of Economic Theory 82, 379–404.

[20] Feenberg, D.R., Poterba, J.M., 1993. Income inequality and the incomes of very high income taxpayers: Evidence from tax returns, in: Poterba, J.M. (Ed.), Tax Policy and the Economy. MIT Press, pp. 145–177.

[21] Feller, W., 1966. An Introduction to Probability Theory and its Applications. volume II. Second ed., Wiley, NY.

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

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

[24] Galor, O., Zeira, J., 1993. Income distribution and macroeconomics. Review of Economic Studies 60, 35–52.

[25] Guvenen, F., 2007. Learning your earning: Are labor income shocks really very persistent? American Economic Review 97, 687–712.

[26] Heathcote, J., Storesletten, K., Violante, G.L., 2014. Optimal tax progressivity: An analytical framework. NBER Working Paper 19899.

[27] Hryshko, D., 2012. Labor income profiles are not heterogeneous: Evidence from income growth rates. Quantitative Economics 3, 177–209.

[28] Huggett, M., 1996. Wealth distribution in life-cycle economies. Journal of Monetary Economics 38, 469–494.

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

[30] Kesten, H., 1973. Random difference equations and renewal theory for products of random matrices. Acta Mathematica 131, 207–248.

[31] Kitao, S., 2008. Entrepreneurship, taxation and capital investment. Review of Economic Dynamics 11, 44–69.

[32] Kopecky, K.A., Suen, R.M., 2010. Finite state Markov-chain approximations to highly persistent processes. Review of Economic Dynamics 13, 701–714.

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

[34] Manrubia, S.C., Zanette, D.H., 1999. Stochastic multiplicative processes with reset events. Physical Review E 59, 4945–4948.

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

[36] Moskowitz, T.J., Vissing-Jørgensen, A., 2002. The returns to entrepreneurial investment: A private equity premium puzzle? American Economic Review 92, 745–778.

[37] Newman, M., 2005. Power laws, Pareto distributions and Zipf's law. Contemporary Physics 46, 323–351.

[38] Nirei, M., 2006. Quantifying borrowing constraints and precautionary savings. Review of Economic Dynamics 9, 353–363.

[39] Nirei, M., 2011. Investment risk, Pareto distribution, and the effects of tax. RIETI Discussion Paper Series 11-E-015.

[40] Nirei, M., Souma, W., 2007. A two factor model of income distribution dynamics. Review of Income and Wealth 53, 440–459.

[41] Panousi, V., 2012. Capital taxation with entrepreneurial risk. Mimeo.

[42] Piketty, T., 2014. Capital in the twenty-first century. Belknap Press.

[43] Piketty, T., Saez, E., 2003. Income inequality in the United States, 1913-1998. Quarterly Journal of Economics CXVIII, 1–39.

[44] Quadrini, V., 1999. The importance of entrepreneurship for wealth concentration and mobility. Review of Income and Wealth 45, 1–19.

[45] Quadrini, V., 2009. Entrepreneurship in macroeconomics. Annals of Finance 5, 295–311.

[46] Quadrini, V., Ríos-Rull, J.V., 1997. Understanding the U.S. distribution of wealth. Federal Reserve Bank for Minneapolis Quarterly Review 21, 22–36.

[47] Reed, W.J., 2001. The Pareto, Zipf and other power laws. Economics Letters 74, 15–19.

[48] Samuelson, P.A., 1969. Lifetime portfolio selection by dynamic stochastic programming. Review of Economics and Statistics 51, 239–246.

[49] Souma, W., 2002. Physics of personal income, in: Takayasu, H. (Ed.), Empirical Science of Financial Fluctuations. Springer-Verlag.

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

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

[52] Zeldes, S.P., 1989. Consumption and liquidity constraints: An empirical investigation. Journal of Political Economy 97, 305–346.