

Pareto Distribution of Income in Neoclassical Growth Models

Makoto Nirei^{a,b}, Shuhei Aoki^c

^a*Policy Research Institute, Ministry of Finance, 3-1-1 Kasumigaseki, Chiyoda-ku, Tokyo 100-8940, Japan*

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

^c*Department of Economics, Shinshu University, 3-1-1 Asahi, Matsumoto, Nagano 390-0802, Japan*

Abstract

We construct a neoclassical growth model with heterogeneous households that accounts for the Pareto distributions of income and wealth in the upper tail. In an otherwise standard Bewley model, we feature households' business productivity risks and borrowing constraints, which we find generate the Pareto distributions. Households with low productivity rely on wages and returns from safe assets, while high productivity households choose not to diversify their business risks. The model can quantitatively account for the observed income distribution in the U.S. under reasonable calibrations. Furthermore, we conduct several comparative statics to examine how changes in parameters affect the Pareto distributions. In particular, we find that the change in the top tax rates in the 1980s potentially accounts for much of the observed increase in top income dispersion in the last decades. Our analytical result provides a coherent interpretation for the numerical comparative statics.

Keywords: income distribution; Pareto exponent; idiosyncratic investment risk;

[☆]We are grateful to the editor Jonathan Heathcote, anonymous referees, and Wataru Souma for very helpful comments. We wish to thank Professor Tsuneo Ishikawa for introducing us to this research field. The views expressed herein do not reflect those of the Ministry of Finance.

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

entrepreneurs; borrowing constraint

JEL codes: D31, O40

1. Introduction

The issue of national income and wealth distribution has become a prominent subject of both scholarly inquiry and public discussion. Scholars investigating this topic, such as Piketty and Saez [45], have been particularly concerned with understanding the share of wealth held by 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 a fat tail. In the real world, this means that the wealthiest one percent of population 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 be paid to why the distributions of top earners universally follow the Pareto distribution. Moreover, it is important to analyze income distribution in a general equilibrium model, because many variables that influence the distribution, such as wage rates, capital returns, or the aggregate capital level, are endogenously determined, as emphasized by Jones [28].

The purpose of this paper is to construct a neoclassical growth model that accounts for the observed Pareto distribution and examine what variables quantitatively affect the distributions of income and wealth. For this purpose, we embed in an otherwise standard Bewley model idiosyncratic risks on households' business productivity. In the model, each household is endowed with a "backyard" production technology. Households which expect low productivity rely on wages and returns from safe assets, while households which expect high productivity choose to take business risks and invest capital into

their backyard technology. Thus, this model captures an economy wherein a fraction of households choose to become entrepreneurs, while others choose to become workers.

We show that Pareto distributions of income and wealth obtain in the Bewley model with (i) CRRA preferences and (ii) either a borrowing constraint or stochastic death, in addition to (iii) business productivity risks. CRRA preferences lead to an asymptotically constant marginal propensity to consume from wealth and an asymptotically constant level of portfolio risk as wealth becomes large. Business productivity risk combined with a linear savings function constitutes a multiplicative shock in the wealth accumulation process, leading to the diffusion effect that generates the fat-tailed distribution. It is well known that a stochastic process with a multiplicative shock generates a fat-tailed distribution, but diverges indefinitely in the absence of an influx effect. The presence of a borrowing constraint or stochastic death induces a net influx of households into the high wealth group from below, which prevents the wealth distribution from diverging. We obtain a closed-form formula for the Pareto exponent, the parameter that determines the concentration of top income and wealth in the Pareto distribution, using a reduced-form version of the model assuming a Solow-type consumption function. The formula summarizes our main argument that the Pareto exponent is determined by two opposing forces: business productivity risk that diffuses the wealth across households within the high wealth group, and the influx of households at the lower end of the high wealth group.

Given the analytical result in the reduced-form model, we examine quantitative implications of our Bewley model where households optimally solve their savings and portfolio problem. First, we check whether the model replicates the distribution characteristics of income and wealth in the U.S., such as the Pareto exponent, the quintiles of income distribution, the Gini coefficient, and the top income and wealth shares. Second, we conduct several comparative statics to examine how the income and wealth distributions are

affected by changes in parameters such as idiosyncratic business productivity and labor shock volatility, tax rates, borrowing constraints, and the exogenous growth rate. We find that the change in the top tax rates in the 1980s potentially accounts for much of the observed increase in top income dispersion in the last decades. We also find that loose borrowing constraints increase top income dispersion. Consideration of the diffusion and influx effects enables us to interpret our comparative statics results obtained through numerical analysis of the Bewley model. In particular, we obtain that a stricter borrowing constraint reduces wealth concentration through an increased influx effect, whereas an increase in the top marginal tax rate reduces wealth concentration through both a diminished diffusion effect and an augmented influx effect.

We derive our results by combining the literature on Bewley models with insights from research on multiplicative idiosyncratic shocks and Pareto distributions. Idiosyncratic business productivity shocks and an influx into the high wealth group from below, the two elements that generate the Pareto distribution as discussed above, fit naturally into the standard Bewley model. Following Quadrini [46] and Cagetti and De Nardi [11] in spirit, and adopting the modeling strategy of Covas [17], Angeletos [3], and Panousi [43], we construct an entrepreneurial economy, wherein households bear income risk by investing physical capital in its own firm. Under CRRA preferences, the fraction of the risky asset in the household portfolio is asymptotically constant among wealthy households. Thus, the distribution of the wealth growth rate becomes independent of the wealth level—i.e., Gibrat’s law applies for the wealthy group. Hence, the wealth accumulation of the wealthy households follows a multiplicative process that is driven by business productivity risks.

It is known that the multiplicative process generates the Pareto distribution, once it is combined with stochastic death or inheritance (Wold and Whittle [55]; Dutta and Michel [18]; Reed [49]; Benhabib, Bisin, and Zhu [8, 9]; Toda [54]). We note that the

savings of households in the low and middle wealth groups also generate an influx effect that prevents the multiplicative process from diverging. As shown by Carroll and Kimball [13] and the papers cited therein, a household’s consumption function is generically concave if the household with CRRA preferences faces a borrowing constraint or labor income shocks. We feature such households in a Bewley model and show that the precautionary savings induced by the borrowing constraint constitute an influx effect. The idea of the influx effect for generating a power-law distribution is formulated by Kesten [31] and used by Gabaix [22]. Nirei and Souma [42] apply this mechanism to the processes of income and wealth and argue that rate-of-return shocks generate the Pareto distribution, while additive shocks in labor income generate exponential decay in the low and middle incomes. This paper extends their result by incorporating idiosyncratic business productivity shocks and earning shocks in the Bewley model, and obtains new testable implications, such as that a strict borrowing constraint or greater risk in labor income increases the Pareto exponent.

In our model, business productivity shocks mainly affect the top part of the income 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. While the previous studies are successful in accounting for the distribution of low and middle incomes, they do not fully explain the distribution in the upper tail (Aiyagari [1]; Huggett [27]; Quadrini and Ríos-Rull [48]; Castañeda, Díaz-Giménez, and Ríos-Rull [14, 15]; Panousi [43]). Benhabib et al. [10] recently derive a Pareto distribution in a Bewley model. Their paper shows an asymptotically linear savings function under CRRA and applies a generalized version of the Kesten processes that are used in Nirei and Souma [42]. The present paper differs from theirs in exploring the general equilibrium implications of the Bewley model. In particular, we find that labor earnings play an important role in

the tail distribution, contrary to their claim. This paper examines the impacts on tail distributions of various parameters, such as borrowing limits, tax rates, the exogenous growth rate, and shock volatilities, going beyond the working paper version (Nirei [41]).

The rest of the paper is organized as follows. Section 2 lays out the general environment studied in this paper. Section 3 investigates a restricted version of the model wherein households choose consumption and investment following a Solow-type consumption function. With this model, we analytically show that the combination of idiosyncratic business productivity shocks and a lower bound for household wealth generates a Pareto distribution in the upper tail of the income and wealth distributions. Section 4 provides a more elaborate Bewley model wherein households optimally choose savings and portfolio. In Section 4.2, we study an analytically tractable case of the Bewley model. In Section 5, we show that the Bewley model with all the elements, such as the borrowing constraint for households, idiosyncratic business productivity and labor income shocks, and progressive income taxes, can account for the observed properties in the top as well as the remaining parts of the income distribution. Finally, Section 6 concludes.

2. General environment

In this section, we lay out all the model elements that we use in this paper. In Sections 3 and 4, we work on restricted versions of the model in order to derive analytical solutions. In Section 5, we investigate a full model with all of the ingredients.

Consider a continuum of households $i \in [0, 1]$. Household i is endowed with initial physical capital $k_{i,0}$, and a “backyard” production technology that is specified by a Cobb-Douglas production function:

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

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 $\gamma > 1$:

$$\tilde{a}_{i,t} = \gamma^t a_{i,t}, \quad (2)$$

where $a_{i,t}$ is an idiosyncratic productivity shock which follows a Markov process. Because of the common productivity growth γ , other variables such as output, consumption, capital, bond holding, and the 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 after the realization of $a_{i,t}$. Using the optimal level of labor input, we obtain the goods supply function:

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

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:

$$w_t/Y_t = 1 - \alpha. \quad (4)$$

In contrast to labor, capital is installed before the realization of $a_{i,t}$. Thus, there is idiosyncratic risk in capital investment, which we refer to as business productivity risk.

Household i inelastically supplies $e_{i,t}$ units 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.

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

$$r(a_{i,t}) \equiv \pi_{i,t}/k_{i,t} + 1 - \delta = \alpha(1 - \alpha)^{(1-\alpha)/\alpha} (a_{i,t}/w_t)^{(1-\alpha)/\alpha} + 1 - \delta. \quad (5)$$

There is no market for physical capital in this model. The bond bears risk-free interest R_t . The households can engage in lending and borrowing through bonds, but the borrowing amount (detrended) must not exceed the borrowing limit ϕ , that is, $b_{i,t+1} > -\phi$. Because there is no insurance market for idiosyncratic business productivity shocks $a_{i,t}$ and labor shocks $\epsilon_{i,t}$, households accommodate these risks by self-insurance through own savings.

Each household lineage is discontinued with a small probability μ in each period. In 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

$$p = \mu/(1 - \mu). \quad (6)$$

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

$$T_{i,t} = \begin{cases} I_{i,t} - \tau_0 I_{i,t}^{1-\tau_1} & \text{if } I_{i,t} < I^* \\ I^* - \tau_0 I^{*1-\tau_1} + \tau_2(I_{i,t} - I^*) & \text{if } I_{i,t} \geq I^*. \end{cases} \quad (7)$$

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 to the highest income bracket. We assume that the tax proceeds are spent on unproductive government purchase of goods.

3. Analytical results in a simple model

3.1. Solow model with idiosyncratic business productivity risk

In this section, we restrict the general environment in order to derive the Pareto distribution analytically. Specifically, we assume: (i) productivity and labor shocks are i.i.d., (ii) income tax $T_{i,t}$ and risk-free bond $b_{i,t}$ are abstracted from, (iii) households are infinitely living ($\mu = 0$), and (iv) household's savings rate is exogenously fixed. Because of the exogenous savings rate, this model may be viewed as a Solow growth model with heterogeneous households who face uninsurable idiosyncratic business productivity risk. At the expense of these restrictions, the Solow model is analytically tractable for deriving the Pareto exponent. These assumptions are relaxed in Section 4 where we study the Bewley model, wherein the savings rate is optimally chosen by households. With the Bewley model, we will argue that the Pareto exponent is affected by factors such as a borrowing constraint and a marginal tax rate as well as the volatilities of productivity and labor shocks. How those factors affect the Pareto exponent will be explained by using the analytical results obtained in the Solow model.

Households inelastically supply labor $e_{i,t}$, which is an i.i.d. random variable over i and t . The savings rate is exogenously fixed at s . The capital of household i , detrended by γ^t , accumulates as follows:

$$\gamma k_{i,t+1} = (1 - \delta)k_{i,t} + s(\pi_{i,t} + w_t e_{i,t}) \quad (8)$$

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

Substituting (4) into (3) and integrating, we obtain an aggregate relation:

$$Y_t = AK_t^\alpha, \quad (9)$$

where $A \equiv \left(\mathbb{E} \left(a_{i,t}^{(1-\alpha)/\alpha} \right) \right)^\alpha$ is constant under the i.i.d. assumption through which $a_{i,t}$ is independent of $k_{i,t}$. By aggregating the capital accumulation equation (8) across households, and by using (9), we reproduce the equation of motion for aggregate capital in the Solow model,

$$\gamma K_{t+1} = (1 - \delta)K_t + sAK_t^\alpha. \quad (10)$$

Equation (10) shows that K_t follows deterministic dynamics with steady state \bar{K} , which is stable and uniquely solved in the domain $K > 0$ as

$$\bar{K} = \left(\frac{sA}{\gamma - 1 + \delta} \right)^{1/(1-\alpha)}. \quad (11)$$

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 .

3.2. Deriving the Pareto distribution

Equations (2,3,4,8,9,10) and $\pi_{i,t} = \alpha y_{i,t}$ define 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 individual capital $k_{i,t}$ follows

$$k_{i,t+1} = g_{i,t}k_{i,t} + ze_{i,t}, \quad (12)$$

where

$$g_{i,t} \equiv \frac{1 - \delta}{\gamma} + \frac{\alpha s}{\gamma} \left(\frac{a_{i,t}}{A\bar{K}^\alpha} \right)^{(1-\alpha)/\alpha}, \quad (13)$$

$$z \equiv \frac{(1 - \alpha)sA\bar{K}^\alpha}{\gamma}. \quad (14)$$

The additive term $ze_{i,t}$ is the savings from detrended labor income, $sw_t e_{i,t}/\gamma$, and strictly positive. The multiplicative term $g_{i,t}$ is the return to detrended capital $(1 - \delta + s\pi_{i,t}/k_{i,t})/\gamma$.

The stochastic process (12) is stationary because $E(g_{i,t}) = \alpha + (1 - \alpha)(1 - \delta)/\gamma < 1$ holds from (13).

Equation (12) is called a Kesten process, which is a stochastic process with a multiplicative shock and an additive positive shock. It is known that a Kesten process has a stationary distribution with a power-law tail. Hence, we obtain the following proposition.

Proposition 1. *If the probability for $g_{i,t} > 1$ is strictly positive, the household's detrended capital $k_{i,t}$ has a stationary distribution whose tail follows a Pareto distribution:*

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

where the Pareto exponent λ is determined by the condition

$$E(g_{i,t}^\lambda) = 1. \quad (16)$$

The power-law tail distribution (15) and the condition for λ (16) are obtained by the theorem shown by Kesten [31] (see also Levy and Solomon [34] and Gabaix [22]). Because capital income $\pi_{i,t}$ is proportional to $k_{i,t}$, the household's income $\pi_{i,t} + w_t e_{i,t}$ follows the same tail distribution as $k_{i,t}$, if $e_{i,t}$ has a thin tail bounded by $e_{i,t}^{-\lambda-1}$. Note that the condition $\Pr(g_{i,t} > 1) > 0$ is required for (16) to have a solution for λ . In our Solow model, this condition is satisfied when $\alpha a_{\max}^{(1-\alpha)/\alpha} > E(a_{i,t}^{(1-\alpha)/\alpha})$.

Proposition 1 shows that idiosyncratic business productivity shocks generate a “top heavy” distribution. The Pareto distribution is top heavy in that a sizable fraction of the total income is possessed by the richest few. Under the Pareto distribution with $\lambda > 1$, the richest P fraction of the population owns fraction $P^{1-1/\lambda}$ of the total income (Newman [38]). This formula applies only when the entire distribution follows a Pareto distribution, whereas in empirical income distributions only the right tail above the 99th percentile is fitted well by a Pareto distribution. Nonetheless, estimates of historical Pareto exponents match well with historical top income shares through this formula, as

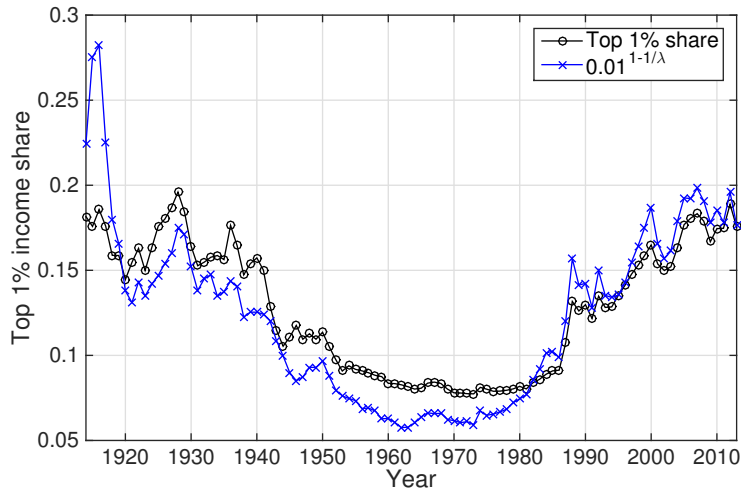


Figure 1: Top 1% income shares and $0.01^{1-1/\lambda}$ evaluated at Pareto exponent λ estimated with historical U.S. income distributions. Source for shares and λ is the World Top Incomes Database [2].

can be seen in Figure 1. This agreement suggests that factors that determine the Pareto exponent may explain the shifts in top income shares.

3.3. Determination of the Pareto exponent and comparative statics

We further characterize λ by assuming that the productivity shock follows a log-normal distribution. We assume that $a_{i,t}^{(1-\alpha)/\alpha}$ follows a log-normal distribution with log-mean $-\sigma^2/2$ and log-variance σ^2 . We choose this specification so that $A = 1$ for any σ . 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 proofs for the propositions in this section are deferred to Appendix A. Proposition 2 provides a comparative static of λ with respect to σ . In the proof, we show that $E(g_{i,t}^\lambda)$

is strictly increasing in λ . Establishing this is easy when $\delta = 1$, since $g_{i,t}$ then follows a 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 determined as $\lambda = 1 + (2/\sigma^2) \log(1/\alpha)$.*

This expression captures the essential 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.

Proposition 2 shows that the “top heavy” distribution generated by the Solow economy has a certain limit in the wealth inequality, since the stationary Pareto exponent cannot be smaller than 1. As we discussed previously, the share of the total wealth owned by the wealthiest P fraction is $P^{1-1/\lambda}$ for $\lambda > 1$. For example, if $\lambda = 2$, 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. 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 [20, p.172]). 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.

The Pareto distribution has a finite mean only if $\lambda > 1$ and a finite variance only if $\lambda > 2$. 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. Empirical income distributions indicate that the Pareto exponent historically transits below and above 2, in the range between 1.5 and 3.¹ This implies that the economy goes

¹See, for example, Alvaredo et al. [2], Fujiwara et al. [21], and Souma [52].

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 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 our model, the households do not diversify business productivity risks. Thus, their income variance increases with the square of their wealth $k_{i,t}^2$, which has a Pareto 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.

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

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

$$\hat{\sigma}^2 = \log \left(\frac{1}{\alpha^2} \left(1 + \frac{2(1-\alpha)}{\gamma/(1-\delta) - 1} \right) \right). \quad (17)$$

Moreover, λ is decreasing in γ and δ in the neighborhood of $\sigma = \hat{\sigma}$.

Proposition 4 relates the Pareto exponent λ with the productivity shock variance σ^2 , the growth rate γ , and the 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 increases inequality 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 $\lambda = 2$, the condition

for determining the Pareto exponent (16) must hold as $E(g_{i,t}^2) = 1$. Using $E(g_{i,t}) = 1 - z/\bar{K}$, this leads to $\text{Var}(g_{i,t})/2 = z/\bar{K} - (z/\bar{K})^2/2$. The key variable z/\bar{K} is equal to $(1-\alpha)(1-(1-\delta)/\gamma)$ as can be derived from (14) at the steady state. 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 $z/\bar{K} \approx \text{Var}(g_{i,t})/2$ as the condition for $\lambda = 2$. This implies that $\lambda = 2$ holds when the standard deviation of g is about 0.4 under the calibration above.² The ratio of the two contributions, $(z/\bar{K})/(\text{Var}(g_{i,t})/2)$, is inversely related to $1 - (1 - \delta)/\gamma$, as can be derived from (13,14) at the steady state. Thus, both growth γ and depreciation δ enhance wealth accumulation more through risk-taking income ($\text{Var}(g_{i,t})/2$) than through labor income (z/\bar{K}). This mechanism explains the comparative statics in Proposition 4.

We elaborate on this mechanism that determines the Pareto exponent λ . 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 $a_{i,t}$. Under the first-order approximation, we obtain $\lambda \approx -E(\log g)/(\text{Var}(\log g)/2)$ from (16). This 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 $E(g) = 1 - z/\bar{K}$, we obtain an alternative expression $\lambda \approx 1 - \frac{\log(1-z/\bar{K})}{\text{Var}(\log g)/2}$ as in Gabaix [22]. We observe that the Pareto exponent λ is always greater than 1, and it declines to 1 as savings z/\bar{K} decreases to 0 or the diffusion effect $\text{Var}(\log g)$ increases to infinity. For a small z/\bar{K} , the expression is further

²While it is difficult to obtain a reliable estimate for entrepreneurial risks, this value seems not unreasonable. For example, Moskowitz and Vissing-Jørgensen [37] cite the annual standard deviation of returns for the smallest decile of public firm in the period 1953–1999, 41.4%, in order to find the lower bound of the risk faced by entrepreneurs.

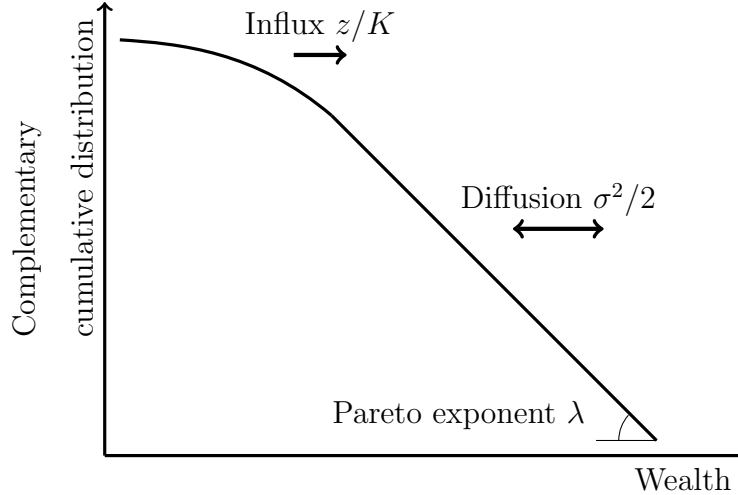


Figure 2: Determination of the Pareto exponent λ . The influx of wealth through savings from labor raises λ , while the diffusion of wealth caused by business productivity risk lowers λ .

approximated as

$$\lambda \approx 1 + \frac{z/\bar{K}}{\text{Var}(\log g)/2}. \quad (18)$$

Equation (18) demonstrates that the Pareto exponent is determined by the balance between two forces: the contributions of an additive term (z/\bar{K}) and a diffusion term ($\text{Var}(\log g)/2$) of the Kesten process (12). These two forces are depicted in Figure 2. The additive term constitutes the wealth accumulation contributed from labor income. This term determines the mass of households who enter the lower end of the high wealth group from the low and middle groups. We call this an influx effect. The diffusion term constitutes the wealth accumulation contributed from capital income due to risk taking. This is seen by noting that $\text{Var}(\log g)/2$ is the contribution of diffusion to the mean total return $\log E(g)$.

Our analysis of the Solow-version model showed that the Pareto exponent is determined as 2 when the influx effect balances with the diffusion effect. In other words, the stationary distribution of income exhibits a finite or infinite variance depending on whether the contribution of labor to capital accumulation exceeds or falls short of the contribution of risk taking. The simple Solow model allows extension in various ways. In Appendix B, we analyze the effect of redistribution and risk sharing on the Pareto exponent in an extended Solow model.

4. The Bewley model

4.1. Bewley model with idiosyncratic business productivity shocks

In reality, household saving behavior depends on the wealth level, tax rate, and risk environment, and it has important implications on the Pareto exponent. In order to incorporate the households' optimal savings and portfolio choice, we depart from the Solow-type restrictions and develop a Bewley model with idiosyncratic business productivity shocks and borrowing constraints.

Given the optimal static labor hiring and optimal production in each period as in (3), the households solve the following dynamic programming problem at the balanced growth path:

$$V(W, a, e) = \max_{c, k', b', W'} \frac{c^{1-\rho}}{1-\rho} + \tilde{\beta} \mathbb{E}(V(W', a', e') | a, e) \quad (19)$$

subject to $c, k' \geq 0$ and

$$c + \gamma(k' + b' + \phi) = W, \quad (20)$$

$$W' = (1 + p)(r(a')k' + Rb' + we' - T') + \gamma\phi, \quad (21)$$

$$b' + \phi > 0, \quad (22)$$

where $\tilde{\beta}$ is a modified discount factor $\tilde{\beta} \equiv \beta\gamma^{1-\rho}(1-\mu)$. W denotes the total resources available (the cash-at-hand). The control variables k and b can be equivalently expressed by the household's total financial assets $x \equiv k + b + \phi$ and portfolio $\theta \equiv k/x$. Thus, the household solves the optimal savings problem for x and the portfolio choice for θ .

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

We argue that the stationary wealth distribution in the Bewley economy with a sufficiently large business productivity shock and with CRRA preferences generically has a power-law tail. There is one notable exception in which the borrowing limit ϕ is not binding, the labor shock $e_{i,t}$ is degenerate, and $\mu = 0$. In this particular case, the stationary distribution does not exist, as we show in Section 4.2. Otherwise, the stationary distribution exists, and the business productivity shock leads to the Pareto distribution. This can be seen as follows. Households with CRRA preferences hold a constant portfolio of risky and risk-less assets asymptotically as their wealth increases (see Carroll and Kimball [13]; Benhabib et al. [10]). Thus, their total resources $W_{i,t}$ incur a multiplicative shock $g_{i,t}$. In this case, if the total resources $W_{i,t}$ (detrended) has a stationary distribution, the stationary distribution must satisfy $\Pr(W_{i,t+1} > w) = \Pr(W_{i,t} > w/g_{i,t})$ for large w in the case of i.i.d. shocks, as argued in Gabaix [22]. This equation is solved by a Pareto distribution with exponent λ satisfying $E(g_{i,t}^\lambda) = 1$, if the condition $\Pr(g_{i,t} > 1) > 0$ is satisfied. The existence of the stationary distribution is warranted either by $\mu > 0$ or a binding bor-

rowing constraint. The former case is discussed in Section 4.2. The latter case is related to the recent finding by Kamihigashi and Stachurski [30] on the existence of a stationary distribution for Markov processes. They establish the existence even when the domain is not compact. Their existence condition requires that the steady-state transition function of household total resources $W_{i,t}$ in our model satisfy that the stochastic kernel is increasing, order-reversing, bounded in probability and has a deficient distribution. As Aiyagari [1] showed, solvency constraints require that households' debt is limited by a natural borrowing limit $w_{\min}/(R - 1)$ if there is a non-degenerate labor shock. With this borrowing limit, the detrended total resources are bounded in probability and have a deficient distribution. The order-reversing condition is satisfied in the setup studied in Section 5, where households with high productivity always choose to invest in physical capital.

4.2. Tractable case without borrowing constraints

The Bewley model is analytically tractable when the borrowing constraint and the labor shock are absent. In this sub-section, we show that wealth follows a log-normal process if $\mu = 0$ in this case. This log-normal process implies that no stationary distribution of relative wealth exists. When $\mu > 0$, the stationary distribution of wealth exists and has a Pareto tail, and the Pareto exponent is analytically derived.

We concentrate on a special case of the general environment with: (i) constant labor supply $e_{i,t} = 1$ and i.i.d. productivity $a_{i,t}$ over i and t , (ii) no tax ($T_{i,t} = 0$), and (iii) no binding borrowing limit ϕ . However, here we have endogenous savings and portfolio choice. Because of the i.i.d. shocks, we have the aggregate production relation (9) as in the Solow model. Since this model features 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 [51]; Merton [36]). Here, we draw on Angeletos'

[3] analysis. Let H_t denote human wealth, defined as the present value of future wage income stream, $\tilde{H}_t \equiv \sum_{\tau=t}^{\infty} \gamma^\tau w_\tau (1 - \mu)^{\tau-t} \prod_{s=t+1}^{\tau} R_s^{-1}$, where wage w_τ is detrended by the growth factor γ . Define 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 $W_{i,t} = (1 + p)(r(a_{i,t})k_{i,t} + R_t b_{i,t}) + H_t$.

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 a linear solution with constants s and θ as $c = (1 - s)W$, $k' = \theta s W / \gamma$, and $b' = (1 - \theta)s W / \gamma - (1 - \mu)R^{-1}H$.

By substituting the policy functions into the definition of wealth, and by noting that $(1 - \mu)(1 + p) = 1$ holds from the zero-profit condition for the pension program (6), we obtain the equation of motion for the detrended individual total wealth:

$$W_{i,t+1} = \begin{cases} \tilde{g}'_{i,t+1} W_{i,t} & \text{with prob. } 1 - \mu \\ H & \text{with prob. } \mu, \end{cases} \quad (23)$$

where the growth rate is defined as

$$\tilde{g}'_i \equiv \frac{(\theta r'_i + (1 - \theta)R)s}{(1 - \mu)\gamma}. \quad (24)$$

Thus, at the balanced growth path, household wealth evolves multiplicatively according to (23) as long as the household lineage is continued. When the lineage is discontinued, a new household with initial wealth $W_{i,t+1} = H$ replaces the old one. Therefore, the individual wealth $W_{i,t}$ follows a log-normal process with random reset events where H is the resetting point. Using the result of Manrubia and Zanette [35], the Pareto exponent of the wealth distribution is determined as follows.³

³We 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

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

$$(1 - \mu)E(\tilde{g}_{i,t}^\lambda) = 1 \tag{25}$$

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

Proof: See Appendix C.

We note that if there is no discontinuation event (i.e., $\mu = 0$), individual wealth follows a log-normal process with 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 the case of $\mu = 0$ and no binding borrowing constraint. In this case, a vanishingly small fraction of individuals eventually possesses almost all the wealth. However, historical records in the U.S. show otherwise, where the current level of the Pareto exponent is comparable with that in 1920. Unlike the log-normal process, the empirical variance of log-income is stationary in the long run, as Kalecki [29] pointed out in a classic debate with Gibrat. 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.

5. Quantitative investigation of the Pareto exponent

5.1. Bewley model with binding borrowing constraints

In this section, we analyze how the Pareto distribution is affected by fundamental parameters in the general equilibrium model defined in Section 4.1 without restrictions imposed in Section 4.2. The key element for the Bewley model to generate the Pareto

explained in Reed [49] and applied to the overlapping generations model by Benhabib et al. [9], although the geometric Brownian model differs in that it generates a double Pareto distribution.

distribution even when $\mu = 0$ is a borrowing constraint. When households face a binding borrowing constraint, the consumption function is concave in wealth, whereas it is linear without borrowing constraints. As Carroll and Kimball [13] emphasize, the linear consumption function arises only in a 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 a borrowing constraint is difficult to solve analytically, we numerically solve for a stationary equilibrium. This model features a multiplicative shock on wealth caused by business risks, in addition to a labor hours shock that enters the wealth accumulation process additively as in Aiyagari [1]. Thus, the stationary wealth distribution has a fat tail unlike the Aiyagari economy.

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 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 [33]).

We set $a_{i,t}$ to be a two-state Markov process with persistence. 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 wages and returns from safe assets as their main income source. Since the entrepreneurs bear the business productivity

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 ones with occupational choice (see Quadrini [47] for a survey). Nonetheless, in this model, the entrepreneurs (households with high productivity) do not diversify much of their business risks while workers (households with low productivity) 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 [25], $\tau_1 = 0.151$. We set $\tau_0 = 0.9$ so 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 $\text{Std}(\epsilon_{i,t}) = 0.29$, following Guvenen [24]. The transition matrix Π for the productivity shock $a_{i,t}$ 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 [32]).⁴ The states of $a_{i,t}$ are set at $\{0.71, 1.29\}$, which corresponds to a 20% 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 of about 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 (less than two

⁴In the model, the low productivity households still hold some capital in backyard production, as a result of imposing the Inada condition. However, the level of the return to capital with low productivity is very close to the risk-free rate. Thus, we interpret those households as non-entrepreneurs.

	p20	p40	p60	p80	p95	Gini	Top I share	Top W share
Benchmark	0.611	0.931	1.223	1.666	2.919	0.369	0.083	0.152
Low tax	0.617	0.944	1.194	1.667	2.838	0.379	0.106	0.248
US 1985	0.421	0.792	1.227	1.845	3.049	0.419	0.127	0.251
US 2010	0.406	0.771	1.248	2.030	3.663	0.469	0.198	0.395

Table 1: Distribution characteristics in the model and in the U.S. 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 the Census for the percentiles and Gini index, Piketty [44] for the top income shares, and Saez and Zucman [50] for the top wealth shares.

months labor income) and high-asset groups matches with Zeldes’ estimate, 1.7% (Zeldes [56]; Nirei [39]). The parameters on technology and preferences are set at standard values as $\alpha = 0.36$, $\delta = 0.1$, $\rho = 3$, $\beta = 0.96$, and $\gamma = 1.02$.

The wealth distribution at the stationary equilibrium for this benchmark calibration is computed numerically. Appendix D details the computation method. Table 1 compares income distribution for the benchmark case with income distribution in the U.S. in 1985. As can be seen, the simulated and empirical 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. Note that the highest marginal income tax rate is set at $\tau_2 = 0.5$ in the benchmark calibration in order to be comparable to the tax rate before the tax cut in 1986. The case of a low tax ($\tau_2 = 0.28$) and the empirical distribution in 2010 are also reported in the table. We observe that the top 1% wealth share increases significantly in our model with the low tax rate.

In what follows, we conduct comparative statics of the Pareto exponent. We will show

that the Pareto exponent decreases (i.e., the tail becomes less equal) under high business productivity risk, low labor risk, low income tax rate, loose borrowing constraint, or high growth. We interpret these results by using the scheme depicted in Figure 2: business productivity risk is categorized as the diffusion effect, while labor risk and borrowing constraint as the influx effect. The tax and technological growth affect both the diffusion and influx effects. We explain them in turn.

5.2. Business productivity risk: the diffusion effect

Figure 3 shows the stationary wealth distributions under the benchmark calibration, the case with high variance of idiosyncratic productivity shocks $a_{i,t}$, the case with low variance, and the case with no productivity shocks. The plot shows a complementary cumulative distribution (i.e., a distribution cumulated from above) of the household wealth relative to the median household wealth. Pareto distributions are clearly observed in the right tail of wealth for the top 1% of the distribution (i.e., beneath 10^{-2} in the complementary cumulative distribution). The plots demonstrate that the increased business productivity risk leads to less equal tail distributions, indicated by the flatter tail. Furthermore, the Pareto tail distribution disappears when the variance of the productivity shock is set at 0. This observation confirms our hypothesis that business productivity risk is necessary for our model to generate the Pareto distribution of wealth.

We interpret the effect of productivity shocks as a diffusion effect. Some diffusion is necessary for the Pareto distribution to emerge, as shown by the case with no productivity risk. As the productivity risk becomes high, the volatility of the wealth growth rate increases, resulting in a lower Pareto exponent. Note that the higher productivity risk can reduce the capital portfolio due to risk aversion, which mitigates the increased volatility of capital return. However, this mitigating effect is relatively weak under our calibrations.

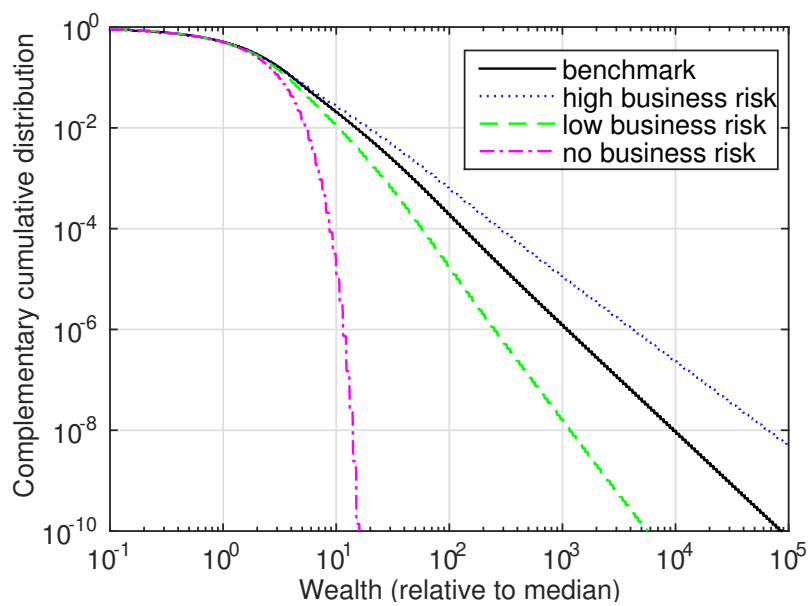


Figure 3: Stationary wealth distributions when standard deviation of productivity $\text{Std}(a_{i,t})$ is set at 0.48 (high), 0.41 (benchmark), 0.34 (low), and 0 (no productivity shock). The standard deviations of corresponding capital returns are 0.23 (high), 0.20 (benchmark), 0.16 (low), and 0, respectively.

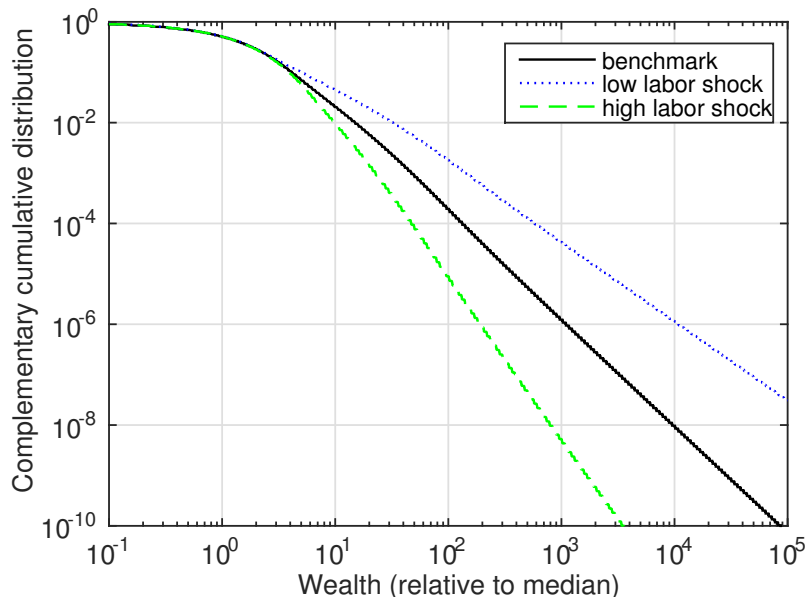


Figure 4: Stationary wealth distributions when standard deviation of labor income shock $\text{Std}(\epsilon_{i,t})$ is set at 0.145 (low), 0.29 (benchmark), and 0.435 (high).

5.3. Precautionary savings: the influx effect

Figure 4 shows the stationary wealth distributions when the standard deviation of the labor endowment shock ($\text{Std}(\epsilon_{i,t})$) varies. The plots show that the increased variance of the labor endowment shock naturally leads to a less equal distribution in the low and middle wealth region. However, in the tail region, the wealth distribution becomes more equal under the increased variance of the labor shock. We will interpret this as an increased influx effect shortly.

Our benchmark calibration for the labor income process adopts a relatively mild persistence, $\zeta = 0.82$, following Guvenen's [24] estimates. To see if our results survive when the labor income process is near unit-root, we conduct a robustness check by setting $\zeta = 0.97$ and the variance of $\epsilon_{i,t}$ at 0.03 following Hryshko [26]. Figure 5 shows the result.

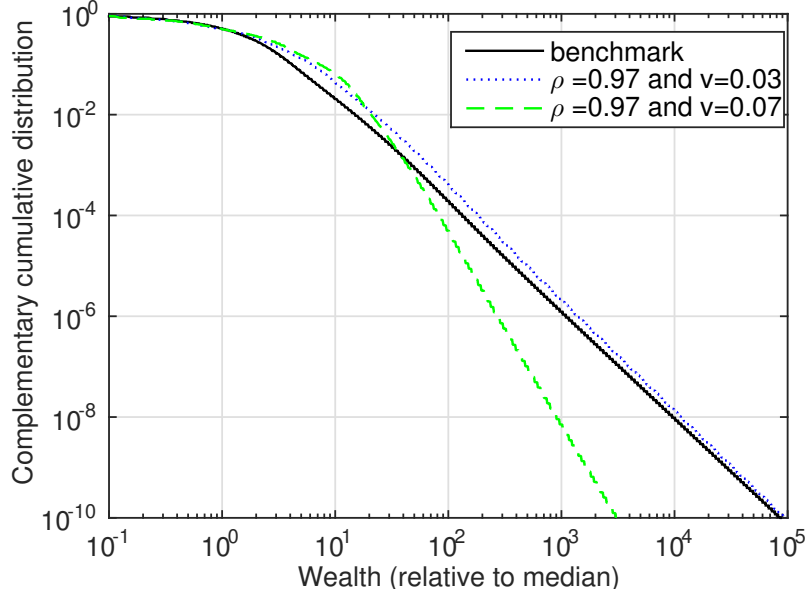


Figure 5: Stationary wealth distributions when the labor income process is persistent ($\zeta = 0.97$). The variance of shock $\text{Var}(\epsilon_{i,t})$ is set at 0.03 and 0.07.

As can be seen, the persistent labor income process does not change the Pareto exponent of the stationary wealth distribution from the benchmark case. We also use an upper-bound estimate for the variance of $\epsilon_{i,t}$, 0.07. The high variance results in a high Pareto exponent, similarly to that in Figure 4. Hence, we find that our results are robust to the persistence of labor income process.

The stationary wealth distribution is affected by other fundamental parameters, such as the death rate μ , borrowing constraint ϕ , and technological growth γ , as shown in Figure 6. The plot for $\mu = 0$ 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 savings function is linear in wealth. Thus, the bind-

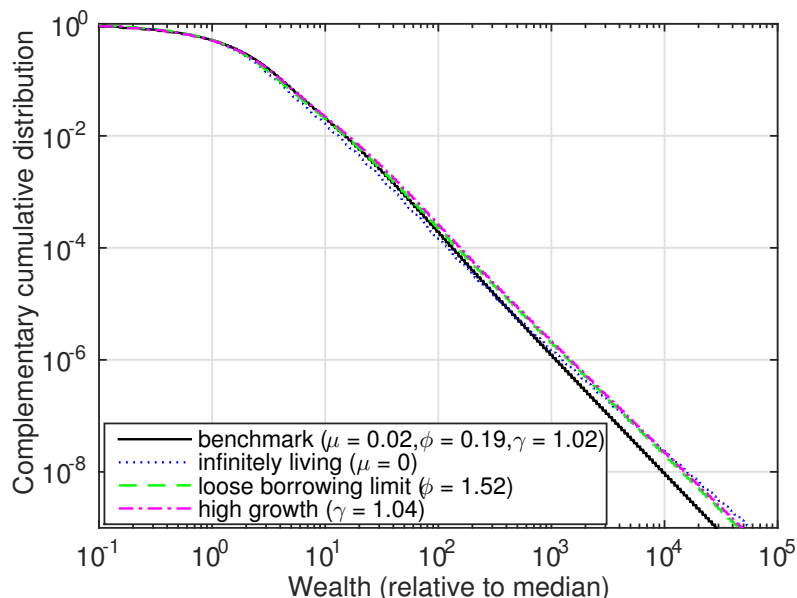


Figure 6: Stationary wealth distributions for the cases of infinitely living households ($\mu = 0$), relaxed borrowing limit ($\phi = 1.52$), and high growth ($\gamma = 1.04$), compared to the benchmark ($\mu = 0.02$, $\phi = 0.19$, and $\gamma = 1.02$).

ing borrowing constraint plays a key role in generating a stationary Pareto distribution when $\mu = 0$, by inducing precautionary savings. The borrowing constraint also has a quantitatively considerable impact on the Pareto exponent. In Figure 6, the stationary wealth distribution for the case $\phi = 1.52$ is plotted. This value of the borrowing limit corresponds to two years' worth of wage income. The plot indicates that relaxing the borrowing constraint from the benchmark three months wage to two years wage has roughly the same impact on the Pareto exponent as reducing the death rate from $\mu = 0.02$ to 0.

The low labor risk and loose borrowing constraint affect the influx effect through the precautionary motive of savings. Households have less incentive for precautionary savings when the labor risk is low or the borrowing constraint becomes lax. Hence, the saving

rate among the low and middle wealth groups falls, which reduces the influx of labor income into wealth, leading to a fall in the Pareto exponent. This result contrasts with Benhabib et al. [8] who claimed that labor income risk does not affect the tail. When there is a borrowing limit, our numerical results show that labor risk affects the tail and that the impact is quantitatively considerable.

We note that the influx effect results from the difference in the saving rates between the high wealth group and the others, rather than from the average saving rate. In the Solow model, the savings rate per se does not affect λ at the stationary distribution, as shown in Proposition 6 in Appendix B. In the Bewley model, we argue that precautionary savings serve as a reflective lower bound for wealth accumulation. When the borrowing constraint is binding, more volatile labor shocks lead to a higher saving rate among the low and middle wealth groups, while the effect for high wealth households is limited. The differential response of households in the saving rates results in an increased influx effect z/\bar{K} . Because stationarity requires $E(g) = 1 - z/\bar{K}$, the increased z/\bar{K} raises the Pareto exponent through the condition $E(g^\lambda) = 1$.⁵

Figure 6 shows that a low death rate ($\mu = 0$) leads to a low Pareto exponent. While this is straightforward given the analysis in Section 4.2, we can interpret the effect of μ as an influx effect, because the low death rate reduces the *net* inflow of households from the low and middle wealth groups to the lower end of high wealth group relative to the outflow. Figure 6 also plots the stationary wealth distribution for the case with high growth $\gamma = 1.04$. The plot shows that the higher rate of technological growth makes

⁵It may be useful to consider a “Cambridge” growth model in which the saving rate from capital income (i.e., s in (13)) differs from that from labor income (s in (14)). Nirei [40] showed that an increase in s only in (14) but not in (13) causes a relatively quick transition to a flatter Pareto tail. This occurs because the increased savings in (14) alone increases z more than aggregate capital \bar{K} .

the wealth distribution less equal in the right tail. This is consistent with our analytical result in Proposition 4. A high growth rate increases both the influx effect z/\bar{K} and the wealth growth g linearly in $1 - (1 - \delta)/\gamma$, as can be derived from (13,14) at the steady state. Since the diffusion effect depends on the variance (i.e., the square) of g , the diffusion effect overwhelms the influx effect, resulting in the lower Pareto exponent.

5.4. *The effect of tax*

Figure 7 plots the complementary cumulative distributions of wealth and income at the stationary equilibrium for the cases of a top marginal tax rate $\tau_2 = 0.5$ (benchmark) and 0.28. Both wealth and income are normalized by the median household wealth and income, respectively. We observe that the Pareto exponents for wealth and income coincide. That is because in this model high income earners earn most of their income from accumulated wealth. The Pareto exponent is significantly smaller in the low tax regime than in the high tax regime: 1.46 for $\tau_2 = 0.28$ and 1.99 for $\tau_2 = 0.5$.

The low tax rate strengthens the diffusion effect, because it increases the volatility of after-tax returns to capital. At the same time, the low tax rate weakens the influx effect, because it reduces the net inflow of households into the tail from below. Therefore, both effects induce a low Pareto exponent. The two tax rates used for Figure 7 emulate the U.S. Tax Reform Act in 1986. As studied by Feenberg and Poterba [19], 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 tax-saving behavior, the steady decline of the Pareto exponent in the 1990s may suggest more persistent effects of the tax reform. Piketty and Saez [45] suggested that the imposition of a progressive tax around the Second World War was the possible cause for the top income share remaining at a low level for a long time after the war, until the 1980s. Our simulations and the above analytical results are consistent with

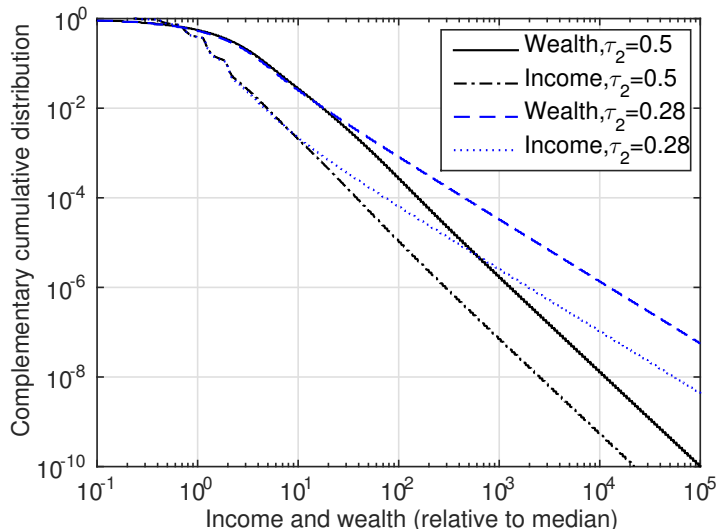


Figure 7: Stationary distributions of wealth and income for the benchmark case ($\tau_2 = 0.5$) and the case of lower tax rate ($\tau_2 = 0.28$).

the view that a tax cut substantially reduces the stationary Pareto exponent. However, our analysis is limited to the comparison of stationary distributions, and the transition dynamics are beyond the scope of this paper.⁶

The above interpretations of the sensitivity analysis assume that the analysis in the Solow model can be extended to the Bewley economy, at least qualitatively. Since our Bewley economy incorporates many features, such as non-linear savings and taxation as well as autocorrelated shocks in business returns and endowments, we cannot provide analytical justifications to this analogy. Nonetheless, we find that the comparative statics obtained by numerical analysis of the Bewley model fit well with the theoretical predictions obtained from the analytical formula shown with the Solow model.

⁶We tackle this issue in a different paper (Aoki and Nirei [4]).

6. Conclusion

This paper demonstrates that the neoclassical growth model with idiosyncratic business productivity 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 in terms of the fundamental parameters, and provide an economic interpretation for its determinants in a general equilibrium analysis. In particular, we emphasize the role played by precautionary savings in determining the Pareto exponent.

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 business income contributed by risk-taking behavior, which corresponds to the diffusion effect making the tail distribution less equal. We show that an increase in the variance of the idiosyncratic business productivity shock lowers the Pareto exponent. While this paper features households' risky business investments in physical capital, the Pareto tail should be similarly obtained when the risky asset takes the form of human capital, or executive stock option pay as studied in Aoki and Nirei [4]. The essential feature of the model is that the households own a stock factor of production with risky returns and a flow factor of production that is always positive. The risky returns generate the diffusion effects, while the flow factor provides the influx effect.

The analytical results shown in the Solow model hold in a Bewley model, wherein the savings and portfolio are optimally determined by the households. In a special case without borrowing constraints, the Bewley model generates a log-normal process for individual wealth that implies a counterfactual "escaping" path of less and less equal distributions. One way to reestablish the Pareto distribution of income and wealth is to incorporate a random reset event by which each household lineage is discontinued. We emphasize a new factor that prevents the distribution from diverging and generates the

Pareto exponent: precautionary savings generated by a binding borrowing constraint. We conduct sensitivity analyses of the Pareto exponent for the borrowing constraint, technological growth rate, death rate, tax rate, business return volatility, and labor endowment volatility. Results obtained by numerical computation agree with the mechanism of the determination of the Pareto exponent derived in the Solow model.

The agreement between the Solow model and the Bewley model with borrowing constraints points to the key roles played by the influx effect and the diffusion effect in generating the Pareto distribution. Without idiosyncratic business productivity shocks, the model lacks the diffusion effect and thus no Pareto distribution emerges. The influx effect is also necessary for generating Pareto distributions, and it is affected by a variety of economic fundamentals other than the death rate of households. For example, either a stricter borrowing limit or a riskier labor income leads to greater precautionary savings, resulting in a higher Pareto exponent. In the Solow model, savings by the low wealth group correspond to the savings from labor income, which serve as a reflective lower bound of wealth accumulation and exert the influx effect on the Pareto exponent. Our numerical analysis with varying borrowing constraints shows that this effect can be quantitatively considerable.

Appendix A. Proofs for the propositions in Section 3.3

Proof of Proposition 2. We first show the unique existence of a solution λ for (16). Note that $(d/d\lambda)E(g^\lambda) = E(g^\lambda \log g)$ and $(d^2/d\lambda^2)E(g^\lambda) = E(g^\lambda (\log g)^2)$. Since $g > 0$, the second derivative is positive, and thus, $E(g^\lambda)$ is convex in λ . As $\lambda \rightarrow \infty$, 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, $E(g^\lambda)$ eventually becomes greater than 1 as λ increases to infinity. Further, recall that $E(g) < 1$. Thus, for the range $\lambda > 1$, $E(g^\lambda)$ is a continuous

convex function that travels from below 1 to above 1. This establishes that the solution for $E(g^\lambda) = 1$ exists uniquely in the range $\lambda > 1$, and that the solution $\bar{\lambda}$ satisfies $(d/d\lambda)E(g^\lambda)|_{\lambda=\bar{\lambda}} > 0$.

Next, we show that λ is decreasing in σ by showing that an increase in σ is a mean-preserving spread in g . Define constants $g_0 \equiv (1 - \delta)/\gamma$ and $u_0 \equiv \log(\alpha(1 - (1 - \delta)/\gamma))$. Then, $\log u \equiv \log(g - g_0)$ follows a normal distribution with mean $u_0 - \sigma^2/2$ and variance σ^2 . Namely, g follows a shifted log-normal distribution. The mean of g is equal to $g_0 + e^{u_0}$ and independent of σ .

The cumulative distribution function of g is $F(g) = \Phi((\log(g - g_0) - u_0 + \sigma^2/2)/\sigma)$, where $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal. Then,

$$\frac{\partial F}{\partial \sigma} = \phi\left(\frac{\log(g - g_0) - u_0 + \sigma^2/2}{\sigma}\right) \left(-\frac{\log(g - g_0) - u_0}{\sigma^2} + \frac{1}{2}\right), \quad (\text{A.1})$$

where $\phi(\cdot)$ is the derivative of Φ . Using the change of variable $x = (\log(g - g_0) - u_0 + \sigma^2/2)/\sigma$, we obtain

$$\int^{\bar{g}} \frac{\partial F}{\partial \sigma} dg = \int^{\bar{x}} \phi(x)(-x/\sigma + 1) dx dg/dx \quad (\text{A.2})$$

$$= \sigma e^{u_0} \int^{\bar{x}} \frac{-x/\sigma + 1}{\sqrt{2\pi}} e^{-(x-\sigma)^2/2} dx, \quad (\text{A.3})$$

The last line reads as a partial moment of $-x/\sigma + 1$, wherein x follows a normal distribution with mean σ and variance 1. The integral tends to 0 as $\bar{x} \rightarrow \infty$, and the integrand is positive for x below σ and negative above σ . Thus, the partial integral achieves the maximum at $\bar{x} = \sigma$ and then monotonically decreases toward 0. Hence, the partial integral is positive for any \bar{x} , and so is $\int^{\bar{g}} \partial F/\partial \sigma dg$. This completes the proof for the assertion that an increase in σ 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 $E(g^\lambda)$. As was observed, $E(g^\lambda)$ is also strictly increasing in λ locally at $\lambda = \bar{\lambda}$. Thus, an increase in σ results in a decrease in $\bar{\lambda}$ that satisfies $E(g^{\bar{\lambda}}) = 1$.

Proof of Proposition 3. When $\delta = 1$, g is reduced to $\alpha a_{i,t}^{(1-\alpha)/\alpha}$. Then, g^λ follows a log-normal with log-mean $\lambda(\log \alpha - \sigma^2/2)$ and log-variance $\lambda^2 \sigma^2$. Then, we obtain

$$1 = \mathbb{E}(g^\lambda) = e^{\lambda(\log \alpha - \sigma^2/2) + \lambda^2 \sigma^2/2}. \quad (\text{A.4})$$

Taking the logarithm of both sides and solving for λ , we obtain $\lambda = 1 - (2/\sigma^2) \log \alpha$.

Proof of Proposition 4. From the definition of $g_{i,t}$ (13), we obtain

$$\mathbb{E}(g_{i,t}^2) = \left(\frac{1-\delta}{\gamma}\right)^2 + 2\frac{1-\delta}{\gamma} \frac{\alpha}{\gamma} (\gamma - 1 + \delta) + \left(\frac{\alpha}{\gamma} (\gamma - 1 + \delta)\right)^2 \mathbb{E}(a_{i,t}^{2(1-\alpha)/\alpha}). \quad (\text{A.5})$$

Using the log-normal distribution of $a_{i,t}^{(1-\alpha)/\alpha}$, we obtain $\mathbb{E}(a_{i,t}^{2(1-\alpha)/\alpha}) = e^{\sigma^2}$. Combining these results with the condition $\mathbb{E}(g^2) = 1$, after some manipulation we obtain

$$\hat{\sigma}^2 = \log \left(\frac{1}{\alpha^2} \left(1 + \frac{2(1-\alpha)}{\gamma/(1-\delta) - 1} \right) \right). \quad (\text{A.6})$$

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 B. Redistribution policies in the Solow model

In this section, we extend the Solow framework to redistribution policies financed by taxes on income and bequests. This extension provides an analytical framework to interpret the numerical results on income tax in the Bewley model in Section 5. Moreover, by this extension we incorporate, as a limiting case, the classic argument by Stiglitz [53] that the Solow economy leads to a completely equal distribution of wealth.

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 (8) is modified as follows:

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

By aggregating, we recover the law of motion for K_t as in (10). Therefore, the redistribution policy does not affect \bar{K} or aggregate output at the steady state. Combining with (B.1), the accumulation equation for individual wealth is rewritten at \bar{K} as follows:

$$k_{i,t+1} = \hat{g}_{i,t} k_{i,t} + \hat{z} e_{i,t}, \quad (\text{B.2})$$

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

$$\hat{g}_{i,t} \equiv \frac{1 - \delta - \mathbf{1}_b \tau_b}{\gamma} + \frac{(1 - \tau_y) \alpha (\gamma - 1 + \delta)}{\gamma} \frac{a_{i,t}^{(1-\alpha)/\alpha}}{\text{E}(a_{i,t}^{(1-\alpha)/\alpha})}, \quad (\text{B.3})$$

$$\hat{z} \equiv \frac{(1 - \alpha + \alpha \tau_y) s A}{\gamma} \left(\frac{s A}{\gamma - 1 + \delta} \right)^{\alpha/(1-\alpha)} + \tau_b \mu \left(\frac{s A}{\gamma - 1 + \delta} \right)^{1/(1-\alpha)}. \quad (\text{B.4})$$

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

Proposition 6. *Under the redistribution policy, if the probability for $\hat{g}_{i,t} > 1$ is strictly positive, a household's wealth $k_{i,t}$ has a stationary distribution whose tail follows a Pareto distribution with exponent λ that satisfies $\text{E}(\hat{g}_{i,t}^\lambda) = 1$. An increase in income tax τ_y or bequest tax τ_b raises λ , while λ is not affected by the savings rate s .*

Since the taxes τ_y and τ_b both shift the density distribution of $\hat{g}_{i,t}$ downward, they raise the Pareto exponent λ through the condition $\text{E}(\hat{g}_{i,t}^\lambda) = 1$.

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 longevity results in a less equal tail distribution.

The redistribution financed by income tax τ_y collects a fraction of profits and equally transfers the proceeds to the households. Thus, income tax serves as a means to share idiosyncratic business productivity risks across households. When the transfer is uncorrelated with the capital holding, it works as another additive term to wealth accumulation. Thus, how the transfer is allocated does not affect λ , as long as the distribution of the transferred amount has a thin tail.

If the transfer is proportional to the capital holding, this redistribution scheme by the income tax is equivalent to an insurance against the business productivity risks. The importance of capital market imperfections in determining income distributions is emphasized by Banerjee and Newman [5] and Galor and Zeira [23]. 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 $E(\tau_s \pi_{i,t})$ as rebate. Then, an increase in τ_s raises λ .*

The proof is as follows. Partial risk sharing ($\tau_s > 0$) reduces the coefficient of $a_{i,t}$ in (B.3) while keeping the mean of $\hat{g}_{i,t}$. Then, $\hat{g}_{i,t}$ before the risk sharing is a mean-preserving spread of the new $\hat{g}_{i,t}$. Since $\lambda > 1$, a mean-preserving spread of $\hat{g}_{i,t}$ increases the expected value of its convex function $\hat{g}_{i,t}^\lambda$. Thus, risk sharing must raise λ in order to satisfy $E(\hat{g}_{i,t}^\lambda) = 1$. When the households completely share the idiosyncratic risks, the model converges to the classic case of Stiglitz [53], wherein wealth is equally distributed across households.

Appendix C. Proof of Proposition 5

In this section, we solve the Bewley model with “natural” borrowing constraint and pension program, and show the existence of the balanced growth path. Then the proposition obtains directly by applying Manrubia and Zanette [35].

At the steady state of detrended aggregate capital \bar{K} , the return to physical capital (5) is written as:

$$r(a_{i,t}) = \alpha(a_{i,t}/A)^{(1-\alpha)/\alpha} \bar{K}^{\alpha-1} + 1 - \delta, \quad (\text{C.1})$$

which is an i.i.d. process. The average return is:

$$\bar{r} \equiv E(r) = \alpha A \bar{K}^{\alpha-1} + 1 - \delta. \quad (\text{C.2})$$

The lending market must clear in each period, which requires $\int b_{i,t} di = 0$ for any t . We also note that $r(a_{i,t})$ is independent of $k_{i,t}$ under the i.i.d. assumption on $a_{i,t}$. Thus, 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 γ . Let \bar{W} , \bar{H} , and \bar{w} denote aggregate total wealth, human capital, and the wage rate detrended by γ^t at the balanced growth path, respectively. Then we have:

$$\bar{W} = (1 - \mu)^{-1} \bar{r} \bar{K} + \bar{H}. \quad (\text{C.3})$$

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

$$R = \frac{\gamma(1 - \mu) \bar{H}}{s(1 - \theta) \bar{W}}. \quad (\text{C.4})$$

Aggregation of the policy function for savings across households yields:

$$\bar{K} = (s\theta/\gamma) \bar{W}. \quad (\text{C.5})$$

Combining with (C.3, C.4), we obtain the relation:

$$R = \frac{\gamma(1-\mu)}{s(1-\theta)} - \frac{\theta}{1-\theta}\bar{r}. \quad (\text{C.6})$$

Thus, the mean return to the risky asset and the risk-free rate are determined by \bar{K} from (C.2,C.6). By using (4,9), the human wealth is written as:

$$\bar{H} = \frac{\bar{w}}{1 - (1-\mu)\gamma R^{-1}} = \frac{(1-\alpha)A\bar{K}^\alpha}{1 - (1-\mu)\gamma R^{-1}}. \quad (\text{C.7})$$

In what follows, we show the existence of the balanced growth path in the situation when the parameters of the optimal policy (s, θ) reside in the interior of $(0, 1)$. By using (C.2,C.6,C.7), we have:

$$\frac{\bar{K}}{\bar{H}} = \frac{1 - \frac{(1-\mu)\gamma s(1-\theta)}{\gamma(1-\mu) - s\theta(1-\delta) - s\theta\alpha A\bar{K}^{\alpha-1}}}{(1-\alpha)A\bar{K}^{\alpha-1}}. \quad (\text{C.8})$$

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, we rearrange (C.3) using (C.5) as:

$$\frac{\gamma}{s\theta} \frac{\bar{K}}{\bar{H}} = 1 + (1-\mu)^{-1} \frac{\bar{r}\bar{K}}{\bar{H}}. \quad (\text{C.9})$$

By (C.2), \bar{r} is strictly decreasing in \bar{K} , and R is strictly increasing by (C.6). Thus, \bar{W}/\bar{H} is strictly decreasing by (C.4), and so is $\bar{r}\bar{K}/\bar{H}$ by (C.3). Thus, the right hand side of (C.9) is positive and strictly decreasing in \bar{K} . The left hand side is monotonically increasing from 0 to $+\infty$. Hence, there exists a unique steady-state solution \bar{K} . This verifies the unique existence of the balanced growth path.

The law of motion (23) for the detrended individual total wealth $k_{i,t}$ is now completely specified at the balanced growth path:

$$k_{i,t+1} = \begin{cases} \tilde{g}_{i,t+1}k_{i,t} & \text{with prob. } 1-\mu \\ \bar{H} & \text{with prob. } \mu, \end{cases} \quad (\text{C.10})$$

where,

$$\tilde{g}_{i,t+1} \equiv (\theta r(a_{i,t+1}) + (1 - \theta)R)s / (1 - \mu). \quad (\text{C.11})$$

This is the stochastic multiplicative process with reset events studied by Manrubia and Zanette [35]. By applying their result, we obtain the proposition.

Appendix D. Details of numerical computation

This section explains the computation procedure used for Section 5. Wealth W is discretized by 400 grid points separated equally in log-scale in the range between 10^{-2} and 10^{10} . The autoregressive process of e is discretized by Rouwenhorst's quadrature method (Kopecky and Suen [33]). 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 $\Lambda(W, a, \epsilon)$
 - 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 (C.1) and the profit maximization condition. In (a-ii), the households' dynamic programming problem is solved by the endogenous gridpoint method with linear interpolation. In (a-iii), we first compute W' for given W, a, a', e, e' , where W is chosen at a grid point of wealth. The probability for this transition is computed by combining the transition matrices for a and the discretized e . Thus, we obtain a transition matrix for W . The stationary distribution of W 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 make the maximum grid for wealth quite large (10^{10}) so that the marginal propensity to consume for the households with high wealth can be computed accurately.

The convergence of aggregate capital in (a) and aggregate bonds in (b) are obtained by the bisection method applied to K and R , respectively. The criterion for convergence is set as 10^{-9} for the savings policy function, 10^{-4} for aggregate capital and bonds, 10^{-3} for θ in the outer loop, and 10^{-2} for the estimated Pareto exponent of the stationary wealth distribution in the outer loop.

Figure D.8 shows a part of the computed policy functions and the Lorenz curve for the benchmark calibration. We note that the households with high productivity invest all their wealth in risky assets, and thus k' and $k' + b' + \phi$ coincide for them. We also note that the households with low productivity hold some bonds when their wealth is in the middle level or in the very high level. Bonds demanded by those households are supplied by the households with high productivity and the households with low wealth. Under the benchmark calibration, the stationary capital-output ratio is 1.746 and the consumption-output ratio is 0.725. The risk-free rate is 0.053, while the risky capital generates the net rate of return 0.048 in the low productivity state and 0.327 in the high

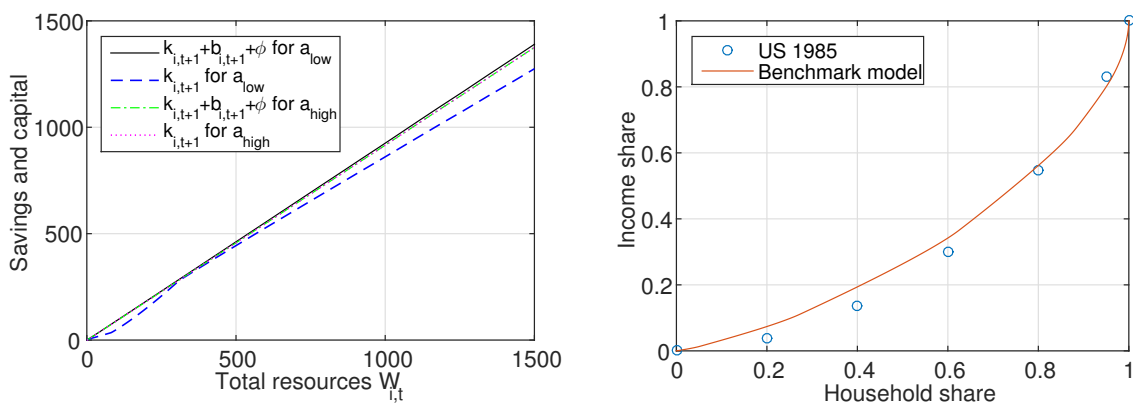


Figure D.8: Policy functions (left) and Lorenz curve for income (right) for the benchmark calibration. Policy functions show savings $k' + b' + \phi$ and capital holding k' conditional on total resources W for a household with low or high productivity a and with the median value of labor endowment e . Total resources, savings, and capital are normalized by the median total resources. Note that the savings and capital levels coincide for the households with high productivity, because they choose to hold all their wealth in capital.

productivity state.

References

- [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., 2015. The world top incomes database. [Http://topincomes.g-mond.parisschoolofeconomics.eu/](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., 2015. Zipf’s law, Pareto’s law, and the evolution of top incomes in the U.S. Available at SSRN: <http://ssrn.com/abstract=2426418>.
- [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., 2014a. The distribution of wealth in the Blanchard-Yaari model. *Macroeconomic Dynamics* Doi: 10.1017/S1365100514000066.
- [10] Benhabib, J., Bisin, A., Zhu, S., 2014b. The wealth distribution in Bewley economies with investment risk. NBER Working Paper 20157.

- [11] Cagetti, M., De Nardi, M., 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] Dutta, J., Michel, P., 1998. The distribution of wealth with imperfect altruism. *Journal of Economic Theory* 82, 379–404.
- [19] 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.
- [20] Feller, W., 1966. *An Introduction to Probability Theory and its Applications*. volume II. Second ed., Wiley, NY.

- [21] Fujiwara, Y., Souma, W., Aoyama, H., Kaizoji, T., Aoki, M., 2003. Growth and fluctuations of personal income. *Physica A* 321, 598–604.
- [22] Gabaix, X., 1999. Zipf’s law for cities: An explanation. *Quarterly Journal of Economics* 114, 739–767.
- [23] Galor, O., Zeira, J., 1993. Income distribution and macroeconomics. *Review of Economic Studies* 60, 35–52.
- [24] Guvenen, F., 2007. Learning your earning: Are labor income shocks really very persistent? *American Economic Review* 97, 687–712.
- [25] Heathcote, J., Storesletten, K., Violante, G.L., 2014. Optimal tax progressivity: An analytical framework. NBER Working Paper 19899.
- [26] Hryshko, D., 2012. Labor income profiles are not heterogeneous: Evidence from income growth rates. *Quantitative Economics* 3, 177–209.
- [27] Huggett, M., 1996. Wealth distribution in life-cycle economies. *Journal of Monetary Economics* 38, 469–494.
- [28] Jones, C.I., 2015. Pareto and Piketty: The macroeconomics of top income and wealth inequality. *Journal of Economic Perspectives* 29, 29–46.
- [29] Kalecki, M., 1945. On the Gibrat distribution. *Econometrica* 13, 161–170.
- [30] Kamihigashi, T., Stachurski, J., 2014. Stochastic stability in monotone economies. *Theoretical Economics* 9, 383–407.
- [31] Kesten, H., 1973. Random difference equations and renewal theory for products of random matrices. *Acta Mathematica* 131, 207–248.

- [32] Kitao, S., 2008. Entrepreneurship, taxation and capital investment. *Review of Economic Dynamics* 11, 44–69.
- [33] Kopecky, K.A., Suen, R.M., 2010. Finite state Markov-chain approximations to highly persistent processes. *Review of Economic Dynamics* 13, 701–714.
- [34] Levy, M., Solomon, S., 1996. Power laws are logarithmic boltzmann laws. *International Journal of Modern Physics C* 7, 595.
- [35] Manrubia, S.C., Zanette, D.H., 1999. Stochastic multiplicative processes with reset events. *Physical Review E* 59, 4945–4948.
- [36] Merton, R.C., 1969. Lifetime portfolio selection under uncertainty: The continuous-time case. *Review of Economics and Statistics* 51, 247–257.
- [37] Moskowitz, T.J., Vissing-Jørgensen, A., 2002. The returns to entrepreneurial investment: A private equity premium puzzle? *American Economic Review* 92, 745–778.
- [38] Newman, M., 2005. Power laws, Pareto distributions and Zipf’s law. *Contemporary Physics* 46, 323–351.
- [39] Nirei, M., 2006. Quantifying borrowing constraints and precautionary savings. *Review of Economic Dynamics* 9, 353–363.
- [40] Nirei, M., 2009. Pareto distributions in economic growth models. IIR Working Paper, Hitotsubashi University 09-05.
- [41] Nirei, M., 2011. Investment risk, Pareto distribution, and the effects of tax. RIETI Discussion Paper Series 11-E-015.

- [42] Nirei, M., Souma, W., 2007. A two factor model of income distribution dynamics. *Review of Income and Wealth* 53, 440–459.
- [43] Panousi, V., 2012. Capital taxation with entrepreneurial risk. Mimeo.
- [44] Piketty, T., 2014. *Capital in the twenty-first century*. Belknap Press.
- [45] Piketty, T., Saez, E., 2003. Income inequality in the United States, 1913-1998. *Quarterly Journal of Economics* CXVIII, 1–39.
- [46] Quadrini, V., 1999. The importance of entrepreneurship for wealth concentration and mobility. *Review of Income and Wealth* 45, 1–19.
- [47] Quadrini, V., 2009. Entrepreneurship in macroeconomics. *Annals of Finance* 5, 295–311.
- [48] Quadrini, V., Ríos-Rull, J.V., 1997. Understanding the U.S. distribution of wealth. *Federal Reserve Bank for Minneapolis Quarterly Review* 21, 22–36.
- [49] Reed, W.J., 2001. The Pareto, Zipf and other power laws. *Economics Letters* 74, 15–19.
- [50] Saez, E., Zucman, G., 2014. Wealth inequality in the United States since 1913: Evidence from capitalized income tax data. NBER Working Paper 20625.
- [51] Samuelson, P.A., 1969. Lifetime portfolio selection by dynamic stochastic programming. *Review of Economics and Statistics* 51, 239–246.
- [52] Souma, W., 2002. Physics of personal income, in: Takayasu, H. (Ed.), *Empirical Science of Financial Fluctuations*. Springer-Verlag.

- [53] Stiglitz, J.E., 1969. Distribution of income and wealth among individuals. *Econometrica* 37, 382–397.
- [54] Toda, A.A., 2014. Incomplete market dynamics and cross-sectional distributions. *Journal of Economic Theory* 154, 310–348.
- [55] Wold, H.O.A., Whittle, P., 1957. A model explaining the Pareto distribution of wealth. *Econometrica* 25, 591–595.
- [56] Zeldes, S.P., 1989. Consumption and liquidity constraints: An empirical investigation. *Journal of Political Economy* 97, 305–346.