A Becker-Tomes Model with Investment Risk

Shenghao Zhu*

National University of Singapore

June 20, 2017

Abstract

Recent literatures find that sufficiently volatile idiosyncratic investment risk plays an important role in generating wealth inequality. I introduce idiosyncratic investment risk into the Becker and Tomes (1979) model and find an explicit expression of the stationary wealth distribution in this simple model. This explicit expression brings us new insights of how bequest motives and estate taxes influence wealth distributions. I find that inheritance increases wealth inequality in models with idiosyncratic investment risk through exaggerating labor earnings uncertainty, while inheritance decreases wealth inequality in the Becker and Tomes (1979) model through mitigating labor earnings uncertainty. This causes estate taxes to

*I would like to thank Jess Benhabib, Alberto Bisin, Basant Kapur, Tomoo Kikuchi, Vincenzo Quadrini, Danny Quah, Thomas Sargent, Jing Wan, C.C. Yang, Hanqin Zhang, Jie Zhang, and seminar participants at ECINEQ 2013 and SAET 2013. I acknowledge the research grant support from National University of Singapore (R-122-000-175-112).

have different impacts on wealth inequality in my model and the Becker and Tomes (1979) model.

JEL classification: D31, H20

Keywords: Investment risk, estate taxes, bequest motives, wealth inequality, the Becker-Tomes model

1 Introduction

Recent literatures find that sufficiently volatile idiosyncratic investment risk plays an important role in generating wealth inequality.[1] These papers characterize the fat tail of the stationary wealth distribution. However, they do not find the explicit expression of the stationary wealth distribution.

I introduce idiosyncratic investment risk into the Becker and Tomes (1979) model and set up a heterogeneous agents model with idiosyncratic investment shocks and labor earnings shocks to investigate the impacts of bequest motives and estate taxes on wealth distributions. Thanks to linear policy functions, I obtain the explicit expression of the stationary wealth distribution.

Different from Becker and Tomes (1979) and Davies (1986), I find that in an economy with altruistic bequest motives and idiosyncratic investment risk, bequest motives increase the long-run wealth inequality. Most importantly, I find that estate taxes reduce the long-run wealth inequality.

Comparing with existing literatures of idiosyncratic investment risk,

[1] This large class of literatures includes, for example, Quadrini (2000), Cagetti and De Nardi (2006, 2009), Benhabib et al. (2011), Panousi (2012), Shourideh (2014), Achdou et al. (2015), and Benhabib, Bisin, and Zhu (2015).

this paper has two contributions. The first contribution is that the explicit expression of the stationary wealth distribution in this simple model brings us new insights of how bequest motives and estate taxes influence wealth distributions in models with idiosyncratic investment risk. This explicit expression permits me to use the decomposition technique developed by Davies (1986) to analyze the wealth accumulation process. Then I can separate the inheritance effect from the redistribution effect of estate taxation. I find that, incorporating idiosyncratic investment risk into the wealth accumulation process leads to the inheritance effect which decreases the long-run wealth inequality.

In this paper the investment risk introduces a multiplicative shock to the agent's wealth accumulation process. Even though on average the multiplicative random coefficient of the wealth accumulation equation is smaller than 1, there are some paths of high investment returns. Along these paths the multiplicative random coefficient continues to be greater than 1 for many periods. Individuals who keep drawing good luck of investment returns become the rich people in the economy. Inheritance exaggerates wealth inequality because of good luck of investment returns. However in Becker and Tomes (1979) the multiplicative coefficient of the wealth accumulation equation is deterministic and smaller than 1. Inheritance has an averaging effect on labor earnings uncertainty. Thus inheritance reduces wealth inequality.

The second contribution is that this paper uses altruistic bequest motives. Both my paper and Benhabib et al. (2011) have idiosyncratic investment risk. In this paper I extend results in the Benhabib, Bisin, and Zhu (2011) model with "joy of giving" bequest motives to models with altruistic bequest motives. I find that the extension reflects a fundamental difference between my paper and Becker and Tomes (1979) after I also compare two models without idiosyncratic investment risk, Becker and Tomes (1979) and Bossmann et al. (2007). These two papers only have idiosyncratic labor earnings shocks. Becker and Tomes (1979) use altruistic bequest motives while Bossmann et al. (2007) use "joy of giving" bequest motives. However, these two papers have opposite implications of impacts of estate taxes on the long-run wealth inequality. I find that, with only idiosyncratic labor earnings shocks, impacts of estate taxes on the long-run wealth inequality depend on formulations of bequest motives, while with idiosyncratic investment shocks, impacts of estate taxes on the long-run wealth inequality do not depend on formulations of bequest motives. This is because impacts of estate taxes on wealth inequality do not depend on the redistribution of tax revenues in models with idiosyncratic investment shocks. Empirical researches have not found evidences to distinguish these two bequest motives: altruism and "joy of giving." In a recent literature review, Kopzuk (2013) states that "Bequest motives are the key building block for theoretical analysis of taxation of transfers, but the empirical literature has not settled on a clear answer to the question about the nature of bequest motivations." (Page 331 of Kopzuk (2013))

1.1 Literatures of the impact of taxes on wealth inequality

Pestieau and Possen (1979) use a model with multiplicative investment shocks to show that the greater the degree of progressivity of the estate tax, the lower the long-run wealth inequality. I study a different tax scheme from theirs.[2] I investigate how the flat estate tax affects the long-run wealth inequality in this paper.[3]

Benhabib et al. (2011) set up an overlapping generations (OLG) model with "joy of giving" bequest motives. Each generation has investment shocks and labor earnings shocks. Their model can generate a stationary wealth distribution with a fat tail and they show that labor earnings shocks do not influence the tail.[4] They find that estate taxes reduce the long-run wealth inequality in their model.[5] In this paper I extend results in Benhabib et al. (2011) to models with altruistic bequest motives. Another important difference between Benhabib et al. (2011) and this paper is that I use the decomposition technique developed by Davies (1986) to separate the inheritance effect from the redistribution effect of estate taxation on wealth inequality in this paper.[6]

Bossmann et al. (2007), using a two-period overlapping generations (OLG) general equilibrium model with only labor earnings shocks, finds

[2] The tax scheme in Pestieau and Possen (1979) has the form

$$S_A = pS_B^c,$$

where $p \geq 1$, $0 \leq c < 1$. S_A represents the after-tax estate, and S_B the before-tax estate. The lower the value of c, the greater the degree of progressivity of the tax. The constant p is the instrument through which the government returns the tax revenues to the economy.

[3] The mechanisms generating stationary distributions in the two models are also different. Pestieau and Possen (1979) study a progressive property or estate tax. They use the concavity to generate a stationary wealth distribution. And the stationary distribution is lognormal. In this paper I study a flat estate tax and use a Kesten process to generate a stationary distribution with a Pareto tail.

[4] Benhabib, Bisin, and Zhu (2015) generate a stationary wealth distribution with a fat tail in an infinite-horizon model and their model permits agents to have the precautionary savings motive.

[5] See Benhabib, Bisin, and Luo (2015) for careful quantitative analyses of the Benhabib, Bisin, and Zhu(2011) model.

[6] To isolate the redistribution effect from their mechanism, Benhabib et al. (2011) intentionally assume that the government wastes collected revenues and does not redistribute them.

that estate taxes reduce the long-run wealth inequality. Wan and Zhu (2017) use the decomposition technique developed by Davies (1986) to analyze the impact of estate taxes on wealth inequality in Bossmann et al. (2007). In Becker and Tomes (1979), Davies (1986), and Bossmann et al. (2007) there are only labor earnings shocks (additive shocks). Thus these models share a common feature, that the inheritance effect of estate taxes increases the long-run wealth inequality.

I summarize the literature about the impact of taxes on the long-run wealth inequality in Table 1.

Table 1: The related literature

	Estate taxes decrease inequality	Estate taxes increase inequality
Multiplicative shocks	Pestieau and Possen (1979) Benhabib et al. (2011)	N.A.
Additive shocks	Bossmann et al. (2007) Wan and Zhu (2017)	Becker and Tomes (1979) Davies (1986)

Recently researchers pay attention to the impact of taxes on the transition of wealth inequality mainly after observing the striking rise of income and wealth inequalities in the United States in recent decades. Piketty and Saez (2003) conject that the decline of progressive taxation since the early 1980s in the United States could be the main reason of the increase of income inequality in recent decades, even though they cannot prove their conjecture. Aoki and Nirei (2017) find that changes in tax rates can explain the decline in the Pareto exponent of income distribution and the increasing trend of the top 1% income share in the United States in recent decades.

Saez and Zucman (2015) find that top wealth shares have followed a U-shaped evolution since the early twentieth century. Kaymak and Poschke (2016) separate the contributions of taxes, government transfers, and the wage dispersion to the increase of wealth inequality since 1970. They find that changes in taxes and transfers account for nearly half of the rise in wealth concentration between 1960 and 2010. Higher wage dispersion due to skill biased technical change is the dominant factor, explaining $50-60\%$ of the rise in wealth inequality. Cao and Luo (2016) set up a general equilibrium model with idiosyncratic investment risk to study the transition of wealth inequality.

My paper is also related to optimal capital taxation literatures. Piketty and Saez (2013) derive optimal inheritance tax formulas that capture the key equity-efficiency trade-off. They show that the optimal tax rate could be positive and quantitatively large.[7] Farhi and Werning (2010) find that optimal estate taxes should be negative. Shourideh (2014) show that the long-run wealth distribution has a fat tail under optimal bequest taxes, which should be negative. Panousi and Reis (2015) investigate optimal linear capital taxes in a model with idiosyncratic investment risk.

The rest of the paper is organized as follows. Section 2 contains the basic set-up of our model with investment risk. Section 3 presents a Becker-Tomes model with only labor earnings risk. I investigate the impacts of bequest motives on wealth distributions in Section 4. I study the impacts of estate taxes on wealth distributions in Section 5. Section 6 incorporates three extensions of the benchmark model. In Section 7 I permit the agent to have precautionary savings. Section 8 concludes the paper. All proofs

[7] In a working paper version of Piketty and Saez (2013), Piketty and Saez (2012) use a multiplicative random coefficient to generate the fat tail of the wealth distribution.

are in Appendix.

2 The benchmark model

There are a continuum of measure 1 agents in the economy. Each agent lives for one period. At the end of the period, the agent dies and gives birth to one child. The population keeps constant. At the beginning of his life the agent receives an inheritance I_t left by his father. The estate tax rate is b. His after-tax inheritance is $(1-b)I_t$. The agent's labor earnings H_t follow a stochastic process.

Assumption 1. $\{H_t\}$ is irreducible and ergodic.[8,9]

The model includes the case in which $\{H_t\}$ is independent and identically distributed (*i.i.d.*). And it also includes the interesting and realistic case of serially correlated $\{H_t\}$.[10]

Assumption 2. $H_t \in (0, \bar{H})$. In the stationary distribution of $\{H_t\}$, $E(H_t) = 1$. There exists a function $f(x,y)$ on $(0,\bar{H}) \times (0,\bar{H})$, which is uniformly bounded above, such that

$$\Pr\left(H_{t+1} \leq h \mid H_t = x\right) = \int_0^h f(x,y)dy, \;\; for\ h \in (0,\bar{H}).$$

[8] I use $\{x_t\}$ to denote a sequence.

[9] A Markov porcess $\{x_t\}$ is irreducible if there exists a measure φ such that whenever $\varphi(A) > 0$ the process $\{x_t\}$ enters the set A in finite time with a positive probability. See page 82 of Meyn and Tweedie (2009).

[10] Davies (1986) uses a mean-reverting process,

$$H_t = (1-\omega)\hat{H} + \omega H_{t-1} + \varepsilon_t,$$

with $0 < \omega < 1$. And $\hat{H}$ is the long-run mean of H_t. He assumes that ε_t is independent of H_{t-1} and has a zero mean and a constant variance. To ensure $H_t > 0$ he assumes that ε_t is strictly bounded from below by $-(1-\omega)\hat{H}$. If we furthermore assume that ε_t is bounded from above and have a continuous density function on its support, such a process $\{H_t\}$ satisfies Assumptions 1 and 2 of my model.

The agent also receives a lump-sum transfer G_t from the government. The agent's after-tax wealth is

$$L_t = H_t + (1-b)I_t + G_t.$$

The agent's consumption is C_t. He leaves bequests B_t to his child. His budget constraint is

$$C_t + B_t = L_t.$$

The agent has a gross interest rate $\tilde{R}_{t+1}$. Thus

$$I_{t+1} = \tilde{R}_{t+1} B_t.$$

Assumption 3. $\left\{\tilde{R}_t\right\}$ is *i.i.d.* along generations. $\tilde{R}_t$ and H_t are independent of each other.

Both $\tilde{R}_t$ and H_t are idiosyncratic shocks, and they are *i.i.d.* across agents. The stochastic rate of return on investments, $\tilde{R}_t$, is the novelty which I introduce into the Becker-Tomes model.[11] Assuming that $\left\{\tilde{R}_t\right\}$ is correlated across generations adds mathematical complexities. We leave discussions of serially correlated $\left\{\tilde{R}_t\right\}$ to Subsection 6.2.

Assuming that $\{H_t\}$ could be serially correlated adds much mathematical complexities.[12] However, it will be clear later that whether $\{H_t\}$ is *i.i.d.* or serially correlated does not influence underlying mechanisms of the model. The aim of incorporating the case of serially correlated $\{H_t\}$ is only to set my model as close as possible to the Becker and Tomes (1979) model.

[11] Angeletos (2007) studies the impact of idiosyncratic investment risk on the aggregate captial stock in a neoclassical growth model.

[12] The proof of Theorem 1 could be way simplified if we assume that both $\{H_t\}$ and $\left\{\tilde{R}_t\right\}$ are *i.i.d.* along generations.

Assumption 4. $\tilde{R}_t$ has a probability density function $l(\cdot)$ on $[\underline{R}, \bar{R}]$ with $\underline{R} > 0$.

Assumption 4 implies that $\tilde{R}_t$ is bounded.

The agent has an altruistic bequest motive. He cares about the total wealth of his child. The agent has a constant relative risk aversion (CRRA) utility function. The agent first draws $\tilde{R}_t$ and H_t, and then makes the consumption decision. Becker and Tomes (1979) assume that agents correctly anticipate labor earnings shocks of their children. Here I assume that parents correctly anticipate both the investment return and labor earnings of their children.[13] The agent's problem is

$$\max_{C_t, L_{t+1}} \frac{C_t^{1-\gamma}}{1-\gamma} + \chi \frac{L_{t+1}^{1-\gamma}}{1-\gamma}$$

$$s.t. \quad C_t + \frac{L_{t+1}}{\tilde{R}_{t+1}(1-b)} = Z_t, \tag{1}$$

$$Z_t = L_t + \frac{H_{t+1} + G_{t+1}}{\tilde{R}_{t+1}(1-b)}, \tag{2}$$

where χ is the bequest motive intensity and $\gamma \geq 1$ is the coefficient of relative risk aversion. $\gamma = 1$ corresponds to the logarithmic utility function. As in Davies (1986) Z_t represents "family wealth."

The optimal policy functions are

$$C_t = \frac{1}{1 + \left[\tilde{R}_{t+1}(1-b)\right]^{\frac{1-\gamma}{\gamma}} \chi^{\frac{1}{\gamma}}} Z_t, \tag{3}$$

[13] I relax this assumption in Section 7.

and

$$L_{t+1} = \frac{\tilde{R}_{t+1}(1-b)}{1 + \left[\tilde{R}_{t+1}(1-b)\right]^{\frac{\gamma-1}{\gamma}} \chi^{-\frac{1}{\gamma}}} Z_t. \tag{4}$$

Plugging equation (2) into equation (4) we have the individual wealth accumulation process

$$L_{t+1} = d_{t+1} L_t + \eta_{t+1}, \tag{5}$$

where

$$d_{t+1} = \frac{\tilde{R}_{t+1}(1-b)}{1 + \left[\tilde{R}_{t+1}(1-b)\right]^{\frac{\gamma-1}{\gamma}} \chi^{-\frac{1}{\gamma}}}, \tag{6}$$

and

$$\eta_{t+1} = \frac{1}{1 + \left[\tilde{R}_{t+1}(1-b)\right]^{\frac{\gamma-1}{\gamma}} \chi^{-\frac{1}{\gamma}}} \left(H_{t+1} + G_{t+1}\right). \tag{7}$$

The wealth accumulation equation (5) has a multiplicative random coefficient and an additive random term. The investment risk introduces a multiplicative shock to the process. In Becker and Tomes (1979) and Davies (1986), the wealth accumulation processes only have an additive shock: the labor earnings shock. This is the most important difference between the wealth accumulation process in this paper and those in Becker and Tomes (1979) and Davies (1986). Benhabib et al. (2011) also derive a wealth accumulation equation with a multiplicative random coefficient and an additive random term.

The linear policy functions (3) and (4), induced by the CRRA utility functions, have three advantages. Firstly, they lead to the linear wealth accumulation process (5). Secondly, they reduce the difficulty of aggregation. Thirdly, they keep the economy on a balanced growth path if labor earnings have a constant growth rate.[14]

The government taxes inheritances and redistributes the revenue to all agents in a lump-sum form. The government has a balanced budget in every period. Thus

$$G_t = b \int I_t dj,$$

where $\int dj$ represents the aggregation over all agents.

2.1 The stationary wealth distribution

To study the wealth inequality I concentrate on the stationary distribution of the wealth accumulation process (5). In order to guarantee that the process (5) has a stationary distribution with the observed fat tail in the wealth data, I need more assumptions.

Assumption 5. $E(d_{t+1}) < 1$.

Assumption 5 is for the stationarity of the process (5). By Theorem 1 of Brandt (1986), we know that Assumptions 1∼5 imply that the process (5) has a unique stationary distribution. The stationary wealth distribution is

$$L_\infty = \sum_{t=1}^{\infty} \left(\prod_{i=1}^{t-1} d_i \right) \eta_t, \tag{8}$$

with the assumption that $\prod_{i=1}^{0} d_i = 1$. Also, by Theorem 1 of Brandt (1986), we know that starting from any distribution, L_0, the wealth accumulation process $\{L_t\}$ converges in distribution to the stationary distribution, L_∞.

Assumption 6. $E(d_{t+1})^2 > 1$.

Assumption 6 implies that the random coefficient d_{t+1} is sufficiently

[14] I introduce an exogenous economic growth rate into the economy and extend the benchmark model in Subsection 6.1.

volatile. Note that $E(d_{t+1})^2 > 1$ implies that $\Pr(d_{t+1} > 1) > 0$. It turns out $\Pr(d_{t+1} > 1) > 0$ is crucial for main results of this paper.

Theorem 1 characterizes the stationary wealth distribution.

Theorem 1 *Under Assumptions 1∼6, the individual wealth has a unique stationary distribution with an asymptotic Pareto tail of an exponent* $1 < \mu < 2$, *i.e.*

$$\lim_{x\to\infty} \frac{\Pr(L_\infty > x)}{x^{-\mu}} = c,$$

with $c > 0$. *And* μ *solves*

$$E\left(d_{t+1}\right)^{\mu} = 1. \tag{9}$$

Theorem 1 shows that investment risk causes the fat tail of the wealth distribution. And labor earnings shocks and redistribution do not influence the tail of the wealth distribution. Individuals who keep drawing good luck of investment shocks become the rich people in the economy. As in Benhabib et al. (2011), this shows different impacts of labor earnings shocks and investment shocks on wealth distributions.

By Assumption 2, labor earnings H_t is bounded. Thus the process $\{H_t\}$ itself is not powerful enough to generate the fat tail of the wealth distribution. The multiplicative shock with large volatility generates the fat tail. The new mechanism to generating fat tails and wealth inequality of brings us different policy implications of estate taxes on wealth inequality, even though my model uses altruistic bequest motives as in Becker and Tomes (1979) and Davies (1986).

The mechanism to generating a stationary distribution with a fat tail is the combination of the stochastic growth and a lower reflecting barrier. In equation (5) the random coefficient d_{t+1} plays the role of the stochastic growth. And the additive term η_{t+1} plays the role of the lower reflecting barrier when L_t is too low. Gabaix (1999) uses this mechanism to generating Zipf's law of the city size distribution.[15] This mechanism is also used in Benhabib et al. (2011).

Klass et al. (2006), using the Forbes 400 lists during 1988-2003, find that the top wealth in the United States is distributed according to a Pareto distribution with an average exponent of 1.49. Thus I concentrate on situations in which $\mu < 2$. Assumption 6 guarantees $\mu < 2$ in the model. However, $\mu < 2$ implies that the stationary wealth distribution does not have a finite variance. Thus I could not use the coefficient of variation as the inequality measure, even though Becker and Tomes (1979) and Bossmann et al. (2007) use it. I will use the Pareto exponent as the inequality index when I investigate the comparative statics of the wealth distribution in an economy with idiosyncratic investment risk.

The investment risk in my model introduces a multiplicative shock to the agent's wealth accumulation process. The multiplicative random coefficient in the wealth accumulation equation generates the fat tail of the wealth distribution. The new mechanism to generating fat tails and inequality of wealth distributions leads to results different from those in Becker and Tomes (1979) and Davies (1986). Theorem 2 is useful for the comparative static analysis of the wealth inequality in my model.

Theorem 2 *Suppose that d_{t+1} first-order stochastically dominates d'_{t+1}.*[16]

[15] Zipf's law refers to a distribution with an asymptotic Pareto tail of an exponent close to 1.

[16] Let $F_X(x)$ and $F_Y(x)$ be the distribution functions of random variables X and Y, respectively. X first-order stochastically dominates Y, denoted as $X \succeq_{FSD} Y$, if, and only if,

$$F_X(x) \leq F_Y(x)$$

for all $x \in \mathbb{R}$. See page 2 of Müller and Stoyan (2002).

Then the Pareto exponent μ of the stationary wealth distribution under d_{t+1} is smaller than under d'_{t+1}.

A higher d_{t+1} (for almost all paths) implies that the wealth accumulation process (5) is more persistent. This leads to a fatter tail of the wealth distribution. For agents who draw $d_{t+1} > 1$, a higher d_{t+1} causes faster growth of wealth. This increases the wealth inequality.

By the linearity of the policy functions, we can calculate the aggregate wealth of the economy in the steady state,

$$E(L_t) = \int L_t dj = (1-b)\left[\frac{1-E(d_{t+1})}{E\left(\frac{1}{1+\left[\tilde{R}_{t+1}(1-b)\right]^{\frac{\gamma-1}{\gamma}}\chi^{-\frac{1}{\gamma}}}\right)} - b\right]^{-1}.$$

Also, we have

$$G = b\int I_t dj = b\left(\int L_t dj - \int H_t dj\right) = b\left[E(L_t) - 1\right].$$

3 The Becker-Tomes model

In this section I briefly review some main results of Becker-Tomes models by Becker and Tomes (1979) and Davies (1986). In the Becker-Tomes model agents only face idiosyncratic labor earnings shocks. Thus wealth accumulation equation only has an additive shock, the labor earnings shock.

Using constant R to replace $\tilde{R}_{t+1}$ in Section 2 we have the wealth accumulation process

$$L_{t+1} = \delta L_t + \theta(H_{t+1} + G_{t+1}) \tag{10}$$

where

$$\delta = \frac{R(1-b)}{1+[R(1-b)]^{\frac{\gamma-1}{\gamma}}\chi^{-\frac{1}{\gamma}}} \tag{11}$$

and

$$\theta = \frac{1}{1+[R(1-b)]^{\frac{\gamma-1}{\gamma}}\chi^{-\frac{1}{\gamma}}}. \tag{12}$$

Equation (10) is not a special case of equation (5) since we assume that $E(d_{t+1})^2 > 1$ for equation (5) (Assumption 6). In equation (10) δ is a constant. Thus it does not satisfy the volatility assumption of d_{t+1} in equation (5).

The government budget constraint is $G_t = b\int I_t dj$ where $\int dj$ represents the aggregation over all agents.

3.1 The stationary wealth distribution

To study the wealth inequality I concentrate on the stationary distribution of the wealth accumulation process (10). As in Becker and Tomes (1979) and Davies (1986), I assume that $0 < \delta < 1$. By Theorem 1 of Brandt (1986), we know that the process (10) has a unique stationary distribution. The stationary wealth distribution is

$$L_\infty = \theta \sum_{t=1}^{\infty} \delta^{t-1}\left(H_t + G_t\right). \tag{13}$$

Also, we know that starting from any distribution, L_0, the wealth accumulation process $\{L_t\}$ converges in distribution to the stationary distribution, L_∞.

By the linearity of the policy functions, we can calculate the aggregate wealth of the economy in the steady state,

$$E\left(L_{t}\right)=\int L_{t}dj=\frac{1-b}{1+\left[R(1-b)\right]^{\frac{\gamma-1}{\gamma}}\chi^{-\frac{1}{\gamma}}-\left[R(1-b)+b\right]}.$$

Also, we have

$$G=b\left[E\left(L_{t}\right)-1\right].$$

By writing the stationary wealth distribution in the form of (13), Davies (1986) invents a decomposition technique to study the impact of estate taxes on wealth inequality. He separates the inheritance effect and the redistribution effect of estate taxes on wealth inequality. The channel through which estate taxes influence δ^{t-1} in (13) is called the lag structure effect. The channel through which estate taxes influence G_t in (13) is called the transfer effect. Davies (1986) uses this decomposition technique to investigate the mechanism through which estate taxes increase the long-run wealth inequality in Becker and Tomes (1979).[17] In Subsection 5.1 I employ this decomposition technique to analyze impacts of estate taxes on wealth inequality in my model.

An important contribution of Davies (1986) is that he theoretically shows inheritance reduces wealth inequality in the Becker and Tomes (1979) model. Let $W(\cdot)$ be inequality measures defined over relative wealth that obey the Pigou-Dalton "principle of transfers." Here is Proposition 1 of Davies (1986).

Proposition 1 *(Davies, 1986)* $W(L_{\infty})<W(L'_{\infty})$ *where*

(i) $L_{\infty}=\theta\sum_{t=1}^{\infty}\delta^{t-1}\left(H_{t}+G_{t}\right)$;

[17] Wan and Zhu (2017) apply this decomposition technieque and find that different formulations of bequest motives affect the redistribution effect (the transfer effect) of estate taxes. See discussions in Subsection 5.2.

(ii) $L'_\infty = \theta' \sum_{t=1}^{\infty} (\delta')^{t-1} (H_t + G_t)$;

(iii) $0 < \delta' < \delta < 1$.

In order to see the intuition of Proposition 1, we rewrite expression (13) as

$$L_\infty = \frac{\theta}{1-\delta} \sum_{t=1}^{\infty} (1-\delta)\delta^{t-1} (H_t + G_t) . \tag{14}$$

Let $X = \sum_{t=1}^{\infty} (1-\delta)\delta^{t-1} (H_t + G_t)$. We know that L_∞ and X have the same Lorenz curve.[18] Note that X is a weighted average of $\{H_t+G_t\}$. This averaging effect helps to cancel the uncertainty of $\{H_t + G_t\}$. Proposition 1 shows that the higher δ the stronger the averaging effect. Proposition 1 does not assume that $\{H_t\}$ is *i.i.d.* along generations. It even holds for serially correlated $\{H_t\}$.

Note that $\{G_t\}$ does not change in expressions of L_∞ and L'_∞ in Proposition 1. Using the same sequence of $\{G_t\}$, Davies (1986) isolates the inheritance effect from the redistribution effect of estate taxes. The higher G_t the more equal the distribution of $\{H_t + G_t\}$. If policy reforms cause a higher G_t, the redistribution effect reduces wealth inequality. If policy reforms cause a lower G_t, the redistribution effect increases wealth inequality.[19,20]

[18] For a non-negative random variable Y with a finite positive mean and a constant $c > 0$, Y and cY have the same Lorenz curve, i.e. a Lorenz curve satisfies the scale invariance axiom.

[19] This intuition comes from the mathematical result that $X + a$ Lorenz dominates $X + b$ for any non-negative random variable X with a finite positive mean and $a > b > 0$ (See Theorem 3.A.25 of Shaked and Shanthikumar (2010)). Thus $X + a$ is more equal than $X + b$.

[20] See more discussions about the redistribution effect in Subsection 5.2.

4 Bequest motives and wealth distributions

In order to investigate the impact of bequest motives on wealth inequality, we set $b = 0$. Thus $G_t = 0$.

From expression (6) we know that d_{t+1} increases with the bequest motive χ. Note that d_{t+1} is a random variable in equation (5) while δ is deterministic in equation (10). A higher d_{t+1} causes a higher mean of wealth in the stationary distribution. A higher δ also implies that the wealth accumulation process (10) is more persistent and that the stationary wealth distribution has a higher mean. However, the increase of persistency of processes (5) and (10) has different impacts on the dispersion of the wealth distribution. Even though $\delta < 1$ has the averaging effect in equation (14), d_i in equation (8) does not. On the contrary, for those paths through which d_i continue to be greater than 1 for many periods, $\prod_{i=1}^{t-1} d_i$ is exploding in equation (8).

Applying Theorem 2 we have

Proposition 2 *In an economy with idiosyncratic investment risk, the higher the bequest motive χ the fatter the tail of the wealth distribution.*

A higher χ increases the persistency of process (5). Proposition 2 shows that the higher the bequest motive intensity, the higher the wealth inequality.

Davies (1986) finds a seemingly counterintuitive result of the Becker and Tomes (1979) model.

Proposition 3 *In an economy without idiosyncratic investment risk, the higher the bequest motive χ the more equal the wealth distribution.*

Note that δ increases with the bequest motive χ. Proposition 3 follows directly from Proposition 1. In a model with only labor earnings risk, a higher bequest motive reduces wealth inequality. Through the view of averaging, it is not difficult for us to understand this result.

Propositions 2 and 3 implies that the impact of bequest motives on wealth distributions depends on the idiosyncratic investment risk. Comparisons between expressions (8) and (14) show that inheritance plays the role of averaging labor earnings uncertainty when the investment return is deterministic. But with sufficient volatility of investment risk due to Assumption 6, inheritance exaggerates labor earnings uncertainty.

To study the impact of bequest motives on wealth inequality also helps us to understand the inheritance effect of estate taxes. When we investigate the inheritance effect of estate taxes, we have to keep redistribution constant. It is difficult to isolate the inheritance effect from the redistribution effect by adjusting the estate tax rate b. But in terms of the inheritance effect, increasing the estate tax rate b is equivalent to decreasing the bequest motive χ in expressions of (6) and (11).

5 Estate taxes and wealth distributions

Becker and Tomes (1979) and Davies (1986) show that taxing bequests increases the long-run wealth inequality in an economy with only labor earnings shocks. In this section I investigate the impact of estate taxes on the long-run wealth inequality in an economy with investment shocks and labor earnings shocks.

Proposition 4 shows a result different from those of Becker and Tomes

(1979) and Davies (1986).

Proposition 4 d_{t+1} *decreases with the estate tax rate* b. *Thus in an economy with idiosyncratic investment risk, the higher the estate tax rate* b *the thinner the tail of the wealth distribution.*

Proposition 4 implies that a higher estate tax rate reduces the long-run wealth inequality. In this paper a higher estate tax rate cuts the return to bequests. After the government increases the tax rate, wealth accumulates slower even for the agent who draws a good rate of return. Formula (9) of Theorem 1 shows that redistribution does not influence the tail of the wealth distribution. Investment risk plays a crucial role in Proposition 4.

5.1 The decomposition technique

I then use the decomposition technique developed by Davies (1986) to analyze the impact of estate taxes on wealth inequality.

By writing the stationary wealth distribution in the form of (8), I can separate the inheritance effect and the redistribution effect of estate taxes as in Davies (1986). In expression (8), the term $\prod_{i=1}^{t-1} d_i$ reflects the inheritance effect of estate taxes. Note that, d_i depends on the estate tax rate b by equation (6). In expression (8), the term η_t reflects the redistribution effect, since government redistribution G_t is embedded in η_t by equation (7). Comparing expressions of (8) and (13) I find that my model has a different inheritance effect from Becker and Tomes (1979) and Davies (1986). In these papers the inheritance effect of estate taxes increases wealth inequality because it reduces the averaging force of inheritance. With investment return luck, inheritance does not have the weighted average effect on labor income luck, since some realizations of d_i are greater than 1 in equation (8). On the contrary, inheritance exaggerates wealth inequality. Thus in my model the inheritance effect of estate taxes reduces wealth inequality because it mitigates the amplifying force of inheritance.

Davies (1986) shows that the lag structure effect (the inheritance effect) is valid in both altruistic and nonaltruistic models.[21] The finding here and that in Benhabib et al. (2011) show that the inheritance effect in economies with idiosyncratic investment risk is also valid in both altruistic and nonaltruistic models. However, the inheritance effect in my model and that in Davies (1986) are opposite.

Benhabib et al. (2011) show that a higher estate tax rate causes a thinner tail of the wealth distribution in a model with "joy of giving" bequest motives. I show here that a higher estate tax rate leads to a thinner tail of wealth distribution in an altruistic model. Discussions here also help us to understand impacts of estate taxes on wealth inequality in Benhabib et al. (2011). Adopting the decomposition technique invented by Davies (1986), I show key mechanisms which lead to different effects of estate taxes on wealth inequality in the Becker and Tomes (1979) model and in the Benhabib, Bisin, and Zhu (2011) model.

5.2 Redistribution

My model shows that estate taxes reduce wealth inequality in an economy with investment shocks. Bossmann et al. (2007) find that estate taxes reduce the long-run wealth inequality in a model with only labor earnings shocks. Both my model and Bossmann et al. (2007) find results different

[21] See comments in the last paragraph of page 547 of Davies (1986).

from Becker and Tomes (1979) and Davies (1986).

Theorem 1 shows that redistribution does not influence the tail of the wealth distribution in my model. Thus the mechanism of my paper does not depend on the redistribution effect, while results in Bossmann et al. (2007) hinge on the redistribution effect of estate taxes.

Both Becker and Tomes (1979) and Davies (1986) use altruistic bequest motives. In Becker and Tomes (1979) and Davies (1986) the redistribution effect of estate taxes increases wealth inequality. After the government increases the estate tax rate, the percentage decrease of before-tax bequests may exceed the percentage increase of the tax rate. Thus a higher tax rate does not necessarily cause higher tax revenues and redistribution.

The elasticity of before-tax bequests with respect to estate taxes are different under different formulations of bequest motives. As found in Bossmann et al. (2007), the direction of the redistribution effect of estate taxes depends on the formulations of bequest motives. Bossmann et al. (2007) uses "joy of giving" bequest motives. The redistribution effect of estate taxes reduces wealth inequality.

Wan and Zhu (2017) find that the inheritance effect of estate taxes in Bossmann et al. (2007) is in line with that in Becker and Tomes (1979) and Davies (1986). It increases wealth inequality. Wan and Zhu (2017) show that it is the different redistribution effects that cause the impact of estate taxes on wealth inequality in Bossmann et al. (2007) to be different from that in Becker and Tomes (1979) and Davies (1986).

In stead of revising the redistribution effect, my paper revises the inheritance effect in Becker and Tomes (1979). And the result in my paper does not depend on redistribution. Thus it does not depend on formulations of bequest motives.

6 Extensions

In this section I investigate three extensions of the benchmark model in Section 2. In the first extension I introduce economic growth into the benchmark model. I also discuss the effect of economic growth on wealth inequality. In the second extension I permit distributions of investment returns to be correlated across generations. In the third extension I rule out negative bequests by adding the borrowing constraint. These extensions show that the main results of the benchmark model, that bequest motives increase wealth inequality and estate taxes reduce wealth inequality, remain to be true, i.e. Propositions 2 and 4 still hold.

6.1 Economic growth

The benchmark model is a stationary economy. To permit economic growth, I assume that

$$H_t = \hat{H}_t g^t,$$

where $g > 1$ is the gross grow rate of labor earnings, and $\hat{H}_t$ is the detrended labor earnings. I assume that $\left\{\hat{H}_t\right\}$ and $\left\{\tilde{R}_t\right\}$ satisfy Assumptions 1∼6. Thus the aggregate economy also has a gross growth rate g. We divide individual variables by g^t to obtain normalized variables. Let $\hat{L}_t = (L_t) / (g^t)$ and $\hat{G}_{t+1} = (G_{t+1}) / (g^{t+1})$. Thus from equation (5) we have

$$\hat{L}_{t+1} = \frac{d_{t+1}}{g}\hat{L}_t + \hat{\eta}_{t+1}, \tag{15}$$

where d_{t+1} is the same as in equation (6) and

$$\hat{\eta}_{t+1} = \frac{1}{1 + \left[\tilde{R}_{t+1}(1-b)\right]^{\frac{\gamma-1}{\gamma}} \chi^{-\frac{1}{\gamma}}} \left(\hat{H}_{t+1} + \hat{G}_{t+1}\right).$$

After I introduce the growth rate into the economy, the wealth process $\{L_t\}$ does not have a stationary distribution. However, we can investigate the stationary distribution of the normalized wealth process $\left\{\hat{L}_t\right\}$. Thus we can obtain the counterpart of Theorem 1 for the normalized wealth process $\left\{\hat{L}_t\right\}$. The stationary distribution of the process $\left\{\hat{L}_t\right\}$ has an asymptotic Pareto tail of an exponent μ. Using $(d_{t+1})/g$ to replace d_{t+1} in equation (9) we know that μ solves

$$E\left(\frac{d_{t+1}}{g}\right)^{\mu} = 1.$$

It is easy to derive the counterparts of Propositions 2 and 4 in an economy with a positive growth rate. In an economy with idiosyncratic investment risk, the higher the bequest motive χ the fatter the tail of the wealth distribution. The higher the estate tax rate b is, the more equal the wealth distribution is.

Piketty (2014) finds that a higher economic growth rate causes the capital's share of income to be lower. This implies lower income inequality between capital owners and laborers. I can also use my model to study the impact of economic growth on wealth inequality. Applying Theorem 2 to the stochastic process $\left\{\hat{L}_t\right\}$, we have

Proposition 5 *In an economy with idiosyncratic investment risk, the higher the economic growth rate g the thinner the tail of the wealth distribution.*

Proposition 5 shows that the higher the economic growth rate g, the lower the wealth inequality. In an economy with idiosyncratic investment risk, individuals who keep drawing high rates of return for a while enter the top of the wealth distribution. A higher economic growth rate g implies higher growth of the average wealth in the economy. Thus individuals who have good luck may not differentiate them faster from the average wealth. This is the intuition behind Proposition 5.

However, I also find that the impact of economic growth on wealth inequality depends on the idiosyncratic investment risk. Using $\hat{H}_{t+1}g^{t+1}$ to replace H_{t+1} in equation (10), we obtain

$$\hat{L}_{t+1} = \frac{\delta}{g}\hat{L}_t + \theta\left(\hat{H}_{t+1} + \hat{G}_{t+1}\right), \tag{16}$$

in the Becker-Tomes model without idiosyncratic investment risk. Here δ and θ are the same as in equations (11) and (12) respectively. Using Proposition 1, Davies (1986) shows[22]

Proposition 6 *In an economy without idiosyncratic investment risk, the higher the economic growth rate g the less equal the wealth distribution.*

Proposition 6 shows that the higher the economic growth rate g, the higher the wealth inequality. The result of Proposition 6 is opposite to that of Proposition 5. Inheritance plays the role of averaging labor earnings uncertainty in the Becker-Tomes model with only labor earnings risk. From equation (16) we find that a higher growth rate g essentially reduces the averaging effect of inheritance. Thus a higher growth rate g causes a less equal wealth distribution.

[22] See footnote 15 of Davies (1986).

6.2 Markov-dependent $\tilde{R}_t$

Children could inherit parents' investment abilities through either genes or family backgrounds. Thus distributions of investment returns could be correlated across generations. I assume $\left\{\tilde{R}_t\right\}$ is *i.i.d.* along generations in the benchmark model. In this subsection I generalize this assumption and $\left\{\tilde{R}_t\right\}$ are permitted to be Markov-dependent. We need a group of assumptions, Assumptions $1' \sim 7'$, which are listed in Subsection A.6 of the Appendix.

Proposition 7 *Under Assumptions $1' \sim 7'$, the individual wealth has a unique stationary distribution with an asymptotic Pareto tail of an exponent $\mu > 1$, i.e.*

$$\lim_{x\to\infty} \frac{\Pr(L_\infty > x)}{x^{-\mu}} = c,$$

with $c > 0$. And μ solves

$$\Lambda(\mu) \equiv \lim_{n\to\infty} \frac{1}{n} \log E\left[\prod_{i=1}^{n-1} (d_i)^\mu\right] = 0.$$

Assuming Markov-dependent $\tilde{R}_t$, we can show the counterpart of Proposition 2.

Proposition 8 *In an economy with idiosyncratic investment risk, the higher the bequest motive χ the fatter the tail of the wealth distribution.*

Similarly, we can also show the counterparts of Propositions 4 and 5. In an economy with idiosyncratic investment risk, the higher the estate tax rate b the thinner the tail of the wealth distribution. The higher the economic growth rate g is, the more equal the wealth distribution is.

6.3 The borrowing constraint

The third extension is to introduce a borrowing constraint into the model. Parents cannot borrow money from their children without frictions. In this subsection I concentrate on the stationary equilibrium of the economy. Thus the lump-sum transfer from the government G is constant.

We need the following assumptions for this subsection.

Assumption 1''. $\{H_t\}$ and $\left\{\tilde{R}_t\right\}$ are $i.i.d.$ along generations. $\tilde{R}_t$ and H_t are independent of each other.

Assumption 2''. H_t has a probability density function $f(\cdot)$ on $(0, \bar{H})$.

Assumption 3''. $\tilde{R}_t$ has a probability density function $l(\cdot)$ on $[\underline{R}, \bar{R}]$ with $\underline{R} > 0$.

Assumption 4''. $E(d_{t+1}) < 1$.

Assumption 5''. $E(d_{t+1})^2 > 1$.

The agent's problem with the borrowing constraint is

$$\max_{C_t, B_t, L_{t+1}} \frac{C_t^{1-\gamma}}{1-\gamma} + \chi \frac{L_{t+1}^{1-\gamma}}{1-\gamma}$$

$$s.t. \quad C_t + B_t = L_t,$$

$$L_{t+1} = H_{t+1} + (1-b)\tilde{R}_{t+1} B_t + G,$$

$$B_t \geq -\theta L_t,$$

where $0 \leq \theta < 1$. Here $B_t \geq -\theta L_t$ is the borrowing constraint. Note that $\theta = 0$ corresponds to the non-negative constraint on bequests, $B_t \geq 0$.

Let $\phi_{t+1} = \frac{\left[\tilde{R}_{t+1}(1-b)\chi\right]^{-\frac{1}{\gamma}}(H_{t+1}+G)}{1+\theta\left(1+\left[\tilde{R}_{t+1}(1-b)\right]^{\frac{\gamma-1}{\gamma}}\chi^{-\frac{1}{\gamma}}\right)}$. The optimal policy functions are

$$C_t = \begin{cases} (1+\theta)L_t, & if\ L_t \leq \phi_{t+1} \\ \frac{1}{1+\left[\tilde{R}_{t+1}(1-b)\right]^{\frac{1-\gamma}{\gamma}}\chi^{\frac{1}{\gamma}}}\left[L_t + \frac{H_{t+1}+G}{\tilde{R}_{t+1}(1-b)}\right], & otherwise \end{cases},$$

$$B_t = \begin{cases} -\theta L_t, & if\ L_t \leq \phi_{t+1} \\ \frac{1}{1+\left[\tilde{R}_{t+1}(1-b)\right]^{\frac{\gamma-1}{\gamma}}\chi^{-\frac{1}{\gamma}}}\left(L_t - \frac{H_{t+1}+G}{\left[\tilde{R}_{t+1}(1-b)\chi\right]^{\frac{1}{\gamma}}}\right), & otherwise \end{cases},$$

and

$$L_{t+1} = \begin{cases} H_{t+1} - (1-b)\tilde{R}_{t+1}\theta L_t + G, & if\ L_t \leq \phi_{t+1} \\ d_{t+1}L_t + \eta_{t+1}, & otherwise \end{cases},$$

where d_{t+1} and η_{t+1} are the same as in equations (6) and (7) respectively.

The policy functions in this subsection are piecewise-linear, while those in the benchmark model are linear. Nevertheless, we can obtain the counterpart of Theorem 1.

Proposition 9 *Under Assumptions 1″ ∼5″, the individual wealth has a unique stationary distribution with an asymptotic Pareto tail of an exponent* $1 < \mu < 2$*, i.e.*

$$\lim_{x\to\infty} \frac{\Pr(L_\infty > x)}{x^{-\mu}} = c,$$

with $c > 0$*. And* μ *solves*

$$E\left(d_{t+1}\right)^{\mu} = 1.$$

From Proposition 9 we know that the relationship, $E\left(d_{t+1}\right)^{\mu} = 1$ still holds. Thus we can show the counterparts of Propositions 2 and 4. In an economy with idiosyncratic investment risk and borrowing constraints, the higher the bequest motive χ the fatter the tail of the wealth distribution. The higher the estate tax rate b is, the thinner the tail of the wealth distribution is.

7 Unobservable $\tilde{R}_{t+1}$

In the benchmark model I assume that parents correctly anticipate both the investment return and labor earnings of their children. Here I assume that the gross interest rate $\tilde{R}_{t+1}$ and labor earnings of children H_{t+1} are unobservable for parents. But parents know the distributions of $\tilde{R}_{t+1}$ and H_{t+1}. In this section I concentrate on the stationary equilibrium of the economy. Thus the lump-sum transfer from the government G is constant.

We need the following assumptions.

Assumption 1‴. $\{H_t\}$ and $\left\{\tilde{R}_t\right\}$ are *i.i.d.* along generations. $\tilde{R}_t$ and H_t are independent of each other.

Assumption 2‴. H_t has a probability density function $f(\cdot)$ on $(0, \bar{H})$.

Assumption 3‴. $\tilde{R}_t$ has a probability density function $l(\cdot)$ on $[\underline{R}, \bar{R}]$ with $\underline{R} > 0$.

Assumption 4‴. $E\left(\tilde{R}_{t+1}\right)\left[(1-b)\chi E\left(\tilde{R}_{t+1}\right)^{1-\gamma}\right]^{\frac{1}{\gamma}} < 1.$[23]

Assumption 5‴. $\bar{R}$ is large enough and

$$\bar{H} + G > [(1-b)\chi]^{-\frac{1}{\gamma}}\left[E\left(\tilde{R}_{t+1}\left(H_{t+1}+G\right)^{-\gamma}\right)\right]^{-\frac{1}{\gamma}}.$$

The agent's problem is

$$\max_{C_t, B_t, L_{t+1}} \frac{C_t^{1-\gamma}}{1-\gamma} + \chi E\left(\frac{L_{t+1}^{1-\gamma}}{1-\gamma}\right)$$

[23] Note that Assumption 4‴ implies that $(1-b)\chi E\left(\tilde{R}_{t+1}\right) < 1$.

$$s.t. \quad C_t + B_t = L_t,$$

$$L_{t+1} = H_{t+1} + (1-b)\tilde{R}_{t+1}B_t + G,$$

$$B_t \geq 0,$$

where $B_t \geq 0$ is the borrowing constraint. Parents have the non-negative constraint on bequests

The optimal C_t is determined by the first order condition,

$$\begin{aligned} & C_t \\ = \quad & \min\left\{L_t, [(1-b)\chi]^{-\frac{1}{\gamma}}\left\{E\left(\tilde{R}_{t+1}\left[(1-b)\tilde{R}_{t+1}\left(L_t - C_t\right) + H_{t+1} + G\right]^{-\gamma}\right)\right\}^{-\frac{1}{\gamma}}\right\} \end{aligned} \tag{17}$$

We denote

$$C_t = C\left(L_t\right),$$

and

$$B_t = B\left(L_t\right) = L_t - C\left(L_t\right).$$

Thus $C\left(L_t\right)$ is a continuous function of L_t. Also, $B\left(L_t\right)$ is a continuous function of L_t. The individual wealth process $\{L_t\}$ is generated by

$$L_{t+1} = H_{t+1} + (1-b)\tilde{R}_{t+1}B\left(L_t\right) + G.$$

Let $\underline{L} = G$. We have $L_{t+1} \geq \underline{L}$ for $\forall L_t > 0$. Thus $\underline{L} = G$ is a reflecting barrier of the process $\{L_t\}$.

In Appendix A.9 I show that $\frac{B(L_t)}{L_t} \leq 1 - \phi$, where

$$\phi = \frac{1}{1 + \left[E\left(\tilde{R}_{t+1}\right)^{1-\gamma}\right]^{\frac{1}{\gamma}}(1-b)^{\frac{1-\gamma}{\gamma}}\chi^{\frac{1}{\gamma}}}.$$

Note that $0 < \phi < 1$. Thus we have

$$\begin{aligned} L_{t+1} &= H_{t+1} + (1-b)\tilde{R}_{t+1} B\left(L_t\right) + G \\ &\leq H_{t+1} + (1-\phi)\left(1-b\right)\tilde{R}_{t+1} L_t + G. \end{aligned}$$

Let

$$\tilde{\rho}_{t+1} = (1-\phi)\left(1-b\right)\tilde{R}_{t+1}.$$

Thus

$$L_{t+1} \leq \tilde{\rho}_{t+1} L_t + H_{t+1} + G.$$

Assumption $4'''$ implies that $E\left(\tilde{\rho}_{t+1}\right) < 1$.

Theorem 3 *The individual wealth process $\{L_t\}$ is ergodic and hence has a unique stationary distribution. The support of the stationary distribution is unbounded.*

I assume that $\tilde{R}_{t+1}$ and H_{t+1} are unobservable for parents. Thus parents have precautionary savings. Thus, different from those in the benchmark model, the policy functions here are nonlinear. However, these policy functions are asymptotically linear. I have a characterization of the tail of the stationary wealth distribution L_∞.

Definition 1 *A distribution X is said to have a right fat tail if there exists $\mu > 0$ such that*

$$\lim \inf_{x \to +\infty} \frac{\Pr(X > x)}{x^{-\mu}} \geq c,$$

where c is a positive constant.

Theorem 4 *The stationary wealth distribution L_∞ has a fat tail.*

Following Benhabib, Bisin, and Zhu (2015), I use a comparison method to show the fat-tail result in Theorem 4. I also have a result of comparative statics.

Theorem 5 *Suppose that there are two estate tax rates b and b', $b < b'$. Under b', there exists $\mu' > 1$ such that*

$$\lim_{x\to+\infty}\inf \frac{\Pr(L'_\infty > x)}{x^{-\mu'}} \geq c',$$

with $c' > 0$. Then, under b we can always find $1 < \mu \leq \mu'$ such that

$$\lim_{x\to+\infty}\inf \frac{\Pr(L_\infty > x)}{x^{-\mu}} \geq c,$$

with $c > 0$.

Theorem 5 implies that, in an economy with idiosyncratic investment risk and precautionary savings, the higher the estate tax rate b the thinner the tail of the wealth distribution. Similarly, we can show that the higher the bequest motive χ is, the fatter the tail of the wealth distribution is.

7.1 A portfolio selection problem

Next I introduce a risk free asset into the model. The risk-free asset has the rate of return R^f which is constant. Now there are two assets in the economy. The risky asset has the stochastic rate of return $\tilde{R}_{t+1}$. I still assume that the stochastic rate of return $\tilde{R}_{t+1}$ and labor earnings of children H_{t+1} are unobservable for parents. But parents know the distributions of

$\tilde{R}_{t+1}$ and H_{t+1}. The agent faces a portfolio selection problem,

$$\max_{C_t,B_t,F_t,L_{t+1}} \frac{C_t^{1-\gamma}}{1-\gamma} + \chi E\left(\frac{L_{t+1}^{1-\gamma}}{1-\gamma}\right)$$

$$s.t. \quad C_t + B_t = L_t,$$

$$L_{t+1} = H_{t+1} + (1-b)\left[\tilde{R}_{t+1}\left(B_t - F_t\right) + R^f F_t\right] + G,$$

$$F_t \geq 0,$$

$$B_t \geq 0,$$

where F_t is the investment in the risk-free asset.

Let the agent's policy functions be

$$C_t = C\left(L_t\right),\ B_t = B\left(L_t\right),\ \text{and } F_t = \left[1-\omega(L_t)\right] B_t.$$

Let

$$\lim_{L_t\to\infty} \frac{C\left(L_t\right)}{L_t} = \hat{\phi},\ \text{and } \lim_{L_t\to\infty} \omega(L_t) = \hat{\omega}.$$

For large L_t, the first order conditions for the agent's problem are

$$C_t^{-\gamma} = (1-b)\chi E\left[\tilde{R}_{t+1}\left(L_{t+1}\right)^{-\gamma}\right], \tag{18}$$

and

$$E\left[\left(R^f - \tilde{R}_{t+1}\right)\left(L_{t+1}\right)^{-\gamma}\right] = 0. \tag{19}$$

From equation (19) we have

$$E\left(\left(R^f - \tilde{R}_{t+1}\right)\left[R^f\left(1-\omega\right) + \tilde{R}_{t+1}\omega\right]^{-\gamma}\right) = 0,$$

which determines ω. From equation (18) we have

$$\hat{\phi} = \frac{1}{1 + \left[E\left(R^f(1-\omega) + \tilde{R}_{t+1}\omega\right)^{1-\gamma}\right]^{\frac{1}{\gamma}} (1-b)^{\frac{1-\gamma}{\gamma}} \chi^{\frac{1}{\gamma}}}.$$

Thus policy functions are asymptotically linear.

The individual wealth process $\{L_t\}$ is generated by

$$\begin{aligned} L_{t+1} &= H_{t+1} + (1-b)\left[\tilde{R}_{t+1}(B_t - F_t) + R^f F_t\right] + G \\ &= H_{t+1} + (1-b)\left(R^f[1-\omega(L_t)] + \tilde{R}_{t+1}\omega(L_t)\right) B(L_t) + G. \end{aligned}$$

Thus the results of the stationary wealth distribution in the economy with two assets are the same as those in the economy with one risky asset except one difference. In the economy with two assets, we use the rate of return of the investment portfolio, $R^f[1-\omega(L_t)] + \tilde{R}_{t+1}\omega(L_t)$, to analyze the wealth accumulation process. A higher volatility of $\tilde{R}_{t+1}$ causes the agent to reduce the investment in the risky asset. Thus the volatility of the return of the whole portfolio decreases. This dampens the mechanisms of stochastic investment returns and of inheritance which I emphasize in the paper.[24]

8 Conclusion

There are two contributions of this paper.

Firstly, the explicit expression of the stationary wealth distribution in this simple model brings us new insights of how bequest motives and estate taxes influence wealth distributions in models with idiosyncratic invest-

[24] See also Benhabib and Zhu (2008).

ment risk. This explicit expression permits me to use the decomposition technique developed by Davies (1986) to analyze the wealth accumulation process. Then I can separate the inheritance effect from the redistribution effect of estate taxation. I find that, incorporating idiosyncratic investment risk into the wealth accumulation process leads to the inheritance effect which decreases the long-run wealth inequality.

Secondly, I extend results in the Benhabib, Bisin, and Zhu (2011) model with "joy of giving" bequest motives to models with altruistic bequest motives. In my model the impact of estate taxes on wealth inequality does not depend on the redistribution of tax revenues and thus does not depend on formulations of bequest motives.

References

[1] Achdou, Yves, Jiequn Han, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll (2015): "Heterogeneous agent models in continuous time," Mimeo, Princeton University.

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

[3] Aoki, Shuhei and Makoto Nirei (2017): "Zipf's law, Pareto law, and the evolution of top incomes in the U.S.," *American Economic Journal: Macroeconomics*, forthcoming.

[4] Becker, Gary and Nigel Tomes (1979): "An equilibrium theory of the distribution of income and intergenerational mobility," *Journal of Po-*

litical Economy, 87, 1153-1189.

[5] Benhabib, Jess, Alberto Bisin, and Mi Luo (2015): "Wealth distribution and social mobility in the US: a quantitative approach," NBER working paper 21721.

[6] Benhabib, Jess, Alberto Bisin, and Shenghao Zhu (2011): "The distribution of wealth and fiscal policy in economies with finitely lived agents," *Econometrica*, 79, 123-157.

[7] Benhabib, Jess, Alberto Bisin, and Shenghao Zhu (2015): "The wealth distribution in Bewley economies with capital income risk," *Journal of Economic Theory*, 159, 489-515.

[8] Benhabib, Jess and Shenghao Zhu (2008): "Age, luck, and inheritance," NBER working paper 14128.

[9] Bossmann, Martin, Christian Kleiber, and Klaus Wälde (2007): "Bequests, taxation and the distribution of wealth in a general equilibrium model," *Journal of Public Economics*, 91, 1247-1271.

[10] Brandt, Andreas (1986): "The stochastic equation $Y_{n+1} = A_n Y_n + B_n$ with stationary coefficients," *Advances in Applied Probability*, 18, 211-220.

[11] Cagetti, Marco and Mariacristina De Nardi (2006): "Entrepreneurship, frictions, and wealth," *Journal of Political Economy*, 114, 835-870.

[12] Cagetti, Marco and Mariacristina De Nardi (2009): "Estate taxation, entrepreneurship, and wealth," *American Economic Review*, 99, 85-111.

[13] Cao, Dan and Wenlan Luo (2016): "Sufficient statistics for top end wealth inequality," Mimeo, Georgetown University.

[14] Davies, James (1986): "Does redistribution reduce inequality?" *Journal of Labor Economics*, 4, 538-559.

[15] Farhi, Emmanuel and Iván Werning (2010): "Progressive estate taxation," *Quarterly Journal of Economics*, 125, 635-673.

[16] Gabaix, Xavier (1999): "Zipf's law for cities: an explanation," *Quarterly Journal of Economics*, 114, 739-767.

[17] Kaymak, Bariş and Markus Poschke (2016): "The evolution of wealth inequality over half a century: the role of taxes, transfers and technology," *Journal of Monetary Economics*, 77, 1-25.

[18] Klass, Oren, Ofer Biham, Moshe Levy, Ofer Malcai, and Sorin Solomon (2006): "The Forbes 400 and the Pareto wealth distribution," *Economics Letters*, 90, 290-295.

[19] Kopczuk, Wojciech (2013): "Taxation of intergenerational transfers and wealth," in *Handbook of Public Economics*, ed. by Alan Auerbach, Raj Chetty, Martin Feldstein, and Emmanuel Saez. Amsterdam: Elsevier.

[20] Loève, M. (1977): *Probability Theory I* (Fourth Ed.), Springer-Verlag Inc.

[21] Meyn, Sean and Richard Tweedie (2009): *Markov Chains and Stochastic Stability* (Second Ed.), Cambridge University Press.

[22] Mirek, Mariusz (2011): "Heavy tail phenomenon and convergence to stable laws for iterated Lipschitz maps," *Probability Theory and Related Fields*, 151, 705-734.

[23] Müller, Alfred and Dietrich Stoyan (2002): *Comparison Methods for Stochastic Models and Risks*, John Wiley & Sons, Ltd.

[24] Ok, Efe (2016): *Probability Theory with Economic Applications*, Mimeo, New York University.

[25] Panousi, Vasia (2012): "Capital taxation with entrepreneurial risk," Mimeo, Federal Reserve Board.

[26] Panousi, Vasia and Catarina Reis (2015): "Optimal capital taxation with idiosyncratic investment risk," Mimeo, Federal Reserve Board.

[27] Pestieau, Pierre and Uri Possen (1979): "A model of wealth distribution," *Econometrica*, 47, 761-772.

[28] Piketty, Thomas (2014): *Capital in the Twenty-First Century*, The Belknap press of Harvard University press.

[29] Piketty, Thomas and Emmanuel Saez (2003): "Income inequality in the United States, 1913-1998," *Quarterly Journal of Economics*, 118, 1-39.

[30] Piketty, Thomas and Emmanuel Saez (2012): "A theory of optimal capital taxation," NBER working paper 17989.

[31] Piketty, Thomas and Emmanuel Saez (2013): "A theory of optimal inheritance taxation," *Econometrica*, 81, 1851-1886.

[32] Quadrini, Vincenzo (2000): "Entrepreneurship, saving, and social mobility," *Review of Economic Dynamics*, 3, 1-40.

[33] Rointershtein, Alexander (2007): "One-dimensional linear recursions with Markov-dependent coefficients," *Annals of Applied Probability*, 17, 572-608.

[34] Royden, H. L. (1988): *Real Analysis* (Third Ed.), Prentice Hall, Inc.

[35] Saez, Emmanuel and Gabriel Zucman (2015): "Wealth inequality in the United States since 1913: evidence from capitalized income tax data," *Quarterly Journal of Economics*, forthcoming.

[36] Shaked, Moshe and George Shanthikumar (2010): *Stochastic Orders*, Springer.

[37] Shourideh, Ali (2014): "Optimal taxation of wealthy individuals," Mimeo, University of Pennsylvnia.

[38] Wan, Jing and Shenghao Zhu (2017): "Bequests, estate taxes, and wealth distributions," Mimeo, National University of Singapore.

A Appendix

A.1 The tail of the stationary wealth distribution

In the steady state of the aggregate economy, equation (5) implies

$$L_{t+1} = d_{t+1} L_t + \eta_{t+1}, \tag{A.1}$$

where

$$d_{t+1} = \frac{\tilde{R}_{t+1}(1-b)}{1 + \left[\tilde{R}_{t+1}(1-b)\right]^{\frac{\gamma-1}{\gamma}} \chi^{-\frac{1}{\gamma}}}, \tag{A.2}$$

and

$$\eta_{t+1} = \frac{1}{1 + \left[\tilde{R}_{t+1}(1-b)\right]^{\frac{\gamma-1}{\gamma}} \chi^{-\frac{1}{\gamma}}} \left(H_{t+1} + G\right). \tag{A.3}$$

In order to investigate the stationary wealth distribution with serially correlated $\{H_t\}$ and $\{\tilde{R}_t\}$, we need the following definition.

Definition 2 *Let $(\ell, \mathcal{F})$ be a measurable space and let $\{x_n\}$ be a stationary Markov process with transition kernel $Q(x, \cdot)$ defined on it. A Markov-modulated process (MMP) associated with $\{x_n\}$ is a stationary Markov process $\{(x_n, \zeta_n)\}$ defined on a product space $(\ell \times \Upsilon, \mathcal{F} \otimes \Xi)$, whose transitions depend only on the position of x_n. That is, for any $n \geq 0$, $A \in \mathcal{F}$, $B \in \Xi$,*

$$\Pr\left(x_n \in A, \zeta_n \in B \mid \sigma\left((x_i, \zeta_i) : i < n\right)\right) = \int_A Q(x, dy) \Gamma(x, y, B)|_{x = x_{n-1}},$$

where $\Gamma(x, y, \cdot) = \Pr(\zeta_1 \in \cdot \mid x_0 = 0, x_1 = y)$ is a kernel on $(\ell \times \ell \times \Xi)$.

Lemma 1 *Let*

$$m(x) = \log E(d_{t+1})^x.$$

Then $m(x)$ is a convex function of $x > 0$.

Proof: See page 158 of Loève (1977). ■

Lemma 2 *$m(x)$ is a continuous function of $x > 0$.*

Proof: By Proposition 17 of Chapter 5 in Royden (1988), Lemma 1 implies Lemma 2. ■

Proof of Theorem 1: Note that the process $\{(H_t, v_t)\}$, where $v_t = (d_t, \eta_t)$, is a Markov-modulated process associated with $\{H_t\}$.

In order to apply Theorem 1.5 of Roitershtein (2007) to the process $\{L_t\}$, we will verify (A1)-(A7) of Assumption 1.2 in Roitershtein (2007).

(A1) is obviously satisfied since the Borel sigma-algebra is countably generated.

By Assumption 1, $\{H_t\}$ is irreducible. Thus (A2) is satisfied.

By Assumption 2, we have

$$\Pr\left(H_{t+1} \le h \mid H_t = x\right) = \int_0^h f(x,y)dy = \int_0^h \bar{H} f(x,y) \frac{1}{\bar{H}} dy,$$

for $h \in (0, \bar{H})$. Let μ^{Leb} be the Lebesgue measure. We construct a probability measure λ on $(0, \bar{H})$, such that $\lambda(A) = \frac{1}{\bar{H}} \mu^{Leb}(A)$ for any Borel set A. Since $f(x,y)$ is uniformly bounded above on $(0, \bar{H}) \times (0, \bar{H})$, $\bar{H} f(x,y)$ is also uniformly bounded above on $(0, \bar{H}) \times (0, \bar{H})$. Thus the family of functions $\{\bar{H} f(x, \cdot) : (0, \bar{H}) \to [0, \infty)\}_{x \in (0, \bar{H})}$ is uniformly integrable with respect to the measure λ. Then (A3) is satisfied for $m_1 = 1$ and the measure λ we construct.

From Assumption 2 we know that $\Pr(H_{t+1} \in (0, \bar{H})) = 1$. And from Assumption 4 we know that $\Pr(\tilde{R}_{t+1} \in [\underline{R}, \bar{R}]) = 1$. Thus from equation

(A.3) we know that there exists an $\bar{\eta} > 0$ such that $\Pr(\eta_{t+1} < \bar{\eta}) = 1$. Thus (A4) is satisfied.

From Assumption 4 we know that $\Pr(\tilde{R}_{t+1} \in [\underline{\mathrm{R}}, \bar{R}]) = 1$. Thus d_{t+1} is bounded, since d_{t+1} is a continuous function of $\tilde{R}_{t+1}$ (See equation (A.2)). We also know that d_{t+1} is bounded away from zero since $\underline{\mathrm{R}} > 0$. Thus there exists an $c_\rho > 1$ such that $\Pr(\frac{1}{c_\rho} < d_{t+1} < c_\rho) = 1$. Thus (A5) is satisfied.

For $x > 0$, we have

$$\lim_{n\to\infty} \frac{1}{n} \log E\left[\prod_{i=1}^{n-1} (d_i)^x\right] = \log E(d_{t+1})^x = m(x),$$

since, by Assumption 3, $\left\{\tilde{R}_t\right\}$ is *i.i.d.* along generations. From Lemmas 1 and 2 we know that $m(x)$ is convex and continuous. From Assumption 5 we have $E\left(d_{t+1}\right) < 1$. Thus $m(1) < 0$. By Assumption 6, $E(d_{t+1})^2 > 1$. Thus $m(2) > 0$. Then we know that there exists a unique $\mu \in (1, 2)$ such that $m(\mu) = 0$, i.e.

$$E(d_{t+1})^\mu = 1.$$

Thus (A6) is satisfied.

By Assumption 4, $\tilde{R}_{t+1}$ has a probability density function $l(\cdot)$ on $[\underline{\mathrm{R}}, \bar{R}]$. Thus the distribution of $\log d_{t+1}$ is nonarithmetic. Thus (A7) is satisfied.

We have verified (A1)-(A7) of Assumption 1.2 in Roitershtein (2007). Applying Theorem 1.5 of Roitershtein (2007) to the process $\{L_t\}$, we have

$$\lim_{x\to\infty} \frac{\Pr(L_\infty > x)}{x^{-\mu}} = c,$$

with $c > 0$. ■

A.2 Proof of Theorem 2

Proof: Let $\mu' \in (1,2)$ solves

$$E(d'_{t+1})^{\mu'} = 1.$$

Since $d_{t+1} \succeq_{FSD} d'_{t+1}$ and $f(d) = (d)^{\mu'}$ is an increasing function of d, applying Theorem 1.2.8 of Müller and Stoyan (2002), we have

$$E(d_{t+1})^{\mu'} \geq E(d'_{t+1})^{\mu'} = 1.$$

Thus

$$\log E(d_{t+1})^{\mu'} \geq 0.$$

By Assumption 5 we have $\log E(d_{t+1}) < 0$. From Lemmas 1 and 2 we know that $\log E(d_{t+1})^{x}$ is convex and continuous in x. Thus there exists a $\mu > 1$ such that $\log E\left(d_{t+1}\right)^{\mu} = 0$, and $\mu \leq \mu'$. ■

A.3 Proof of Proposition 2

Proof: Suppose that $\chi > \chi'$. Thus by equation (A.2) we know that

$$d_{t+1} = \frac{\tilde{R}_{t+1}(1-b)}{1+\left[\tilde{R}_{t+1}(1-b)\right]^{\frac{\gamma-1}{\gamma}}\chi^{-\frac{1}{\gamma}}} > d'_{t+1} = \frac{\tilde{R}_{t+1}(1-b)}{1+\left[\tilde{R}_{t+1}(1-b)\right]^{\frac{\gamma-1}{\gamma}}(\chi')^{-\frac{1}{\gamma}}}.$$

For $x \in \mathbb{R}$, $d_{t+1} \leq x$ implies $d'_{t+1} \leq x$. Thus we have

$$\Pr(d_{t+1} \leq x) \leq \Pr(d'_{t+1} \leq x),$$

for $x \in \mathbb{R}$. Then we know that $d_{t+1} \succeq_{FSD} d'_{t+1}$. Applying Theorem 2 we know that the Pareto exponent μ of the wealth distribution under χ is smaller than under χ'. ■

A.4 Proof of Proposition 4

Proof: By equation (A.2) we have

$$\begin{aligned} d_{t+1} &= \frac{\tilde{R}_{t+1}(1-b)}{1+\left[\tilde{R}_{t+1}(1-b)\right]^{\frac{\gamma-1}{\gamma}}\chi^{-\frac{1}{\gamma}}} \\ &= \frac{1}{\left[\tilde{R}_{t+1}(1-b)\right]^{-1}+\left[\tilde{R}_{t+1}(1-b)\right]^{-\frac{1}{\gamma}}\chi^{-\frac{1}{\gamma}}}. \end{aligned}$$

Thus d_{t+1} decreases with b.

Suppose that $b < b'$. Thus we have $d_{t+1} > d'_{t+1}$. For $x \in \mathbb{R}$, $d_{t+1} \leq x$ implies $d'_{t+1} \leq x$. Thus we have

$$\Pr(d_{t+1} \leq x) \leq \Pr(d'_{t+1} \leq x),$$

for $x \in \mathbb{R}$. Then we know that $d_{t+1} \succeq_{FSD} d'_{t+1}$. Applying Theorem 2 we know that the Pareto exponent μ of the wealth distribution under b is smaller than under b'. ■

A.5 Proof of Proposition 5

Proof: Suppose that $g < g'$. Thus we have $\frac{d_{t+1}}{g} > \frac{d_{t+1}}{g'}$. For $x \in \mathbb{R}$, $\frac{d_{t+1}}{g} \leq x$ implies $\frac{d_{t+1}}{g'} \leq x$. Thus we have

$$\Pr\left(\frac{d_{t+1}}{g} \leq x\right) \leq \Pr\left(\frac{d_{t+1}}{g'} \leq x\right),$$

for $x \in \mathbb{R}$. Then we know that $\frac{d_{t+1}}{g} \succeq_{FSD} \frac{d_{t+1}}{g'}$. Applying Theorem 2 to the process $\left\{\hat{L}_t\right\}$ we know that the Pareto exponent μ of the wealth distribution under g is smaller than under g'. ■

A.6 Serially correlated $\left\{\tilde{R}_t\right\}$

We introduce a stationary Markov process $\{x_t\}$ into the model such that the process $\{(x_t, \psi_t)\}$, where $\psi_t = \left(\tilde{R}_t, H_t\right)$, is a Markov-modulated process associated with $\{x_t\}$. Thus the Markov process $\{x_t\}$ is the underlying process.

Assumption 1′. $\{x_t\}$ is on the measurable space $(\mathbb{R}, \mathbf{B})$, where $\mathbf{B}$ is the Borel sigma-algebra.

Assumption 2′. $\{x_t\}$ is irreducible.

Assumption 3′. Let $Q(x, \cdot)$ be the transition kernel of $\{x_t\}$. There exist a probability measure λ on $(\mathbb{R}, \mathbf{B})$, a number $m_1 \in \mathbb{N}$, and a measurable density kernel $f(x, y) : \mathbb{R}^2 \to [0, \infty)$ such that

$$Q^{m_1}(x, A) = \int_A f(x, y)\lambda(dy),$$

and the family of functions $\{f(x, \cdot) : \mathbb{R} \to [0, \infty)\}_{x \in \mathbb{R}}$ is uniformly integrable with respect to the measure λ.

Assumption 4′. $H_t \in (0, \bar{H})$.

Assumption 5′. $\tilde{R}_t \in [\underline{R}, \bar{R}]$ with $\underline{R} > 0$.

Assumption 6′. Let $\Lambda(x) = \limsup_{n\to\infty} \frac{1}{n} \log E\left[\prod_{i=1}^{n-1} (d_i)^x\right]$. There exist $\mu_1 > 1$ and $\mu_2 > 1$ such that $\Lambda(\mu_1) \geq 0$ and $\Lambda(\mu_2) < 0$.

Assumption 7′. There do not exist a constant $\alpha > 0$ and a measurable function $\beta : \mathbb{R} \to [0, \alpha)$ such that[25]

$$\Pr\left(\log(d_1) \in \beta(x_0) - \beta(x_1) + \alpha\mathbb{Z}\right) = 1.$$

Proof of Proposition 7: The process $\{(x_t, v_t)\}$, where $v_t = (d_t, \eta_t)$, is a Markov-modulated process associated with $\{x_t\}$ since the process $\{(x_t, \psi_t)\}$, where $\psi_t = \left(\tilde{R}_t, H_t\right)$, is a Markov-modulated process associated with $\{x_t\}$.

In order to apply Theorem 1.5 of Roitershtein (2007) to the process $\{L_t\}$, we will verify (A1)-(A7) of Assumption 1.2 in Roitershtein (2007).

(A1) is obviously satisfied since the Borel sigma-algebra is countably generated.

By Assumptions $2'$ and $3'$, (A2) and (A3) are satisfied.

From Assumption $4'$ we know that $\Pr(H_{t+1} \in (0, \bar{H})) = 1$. And from Assumption $5'$ we know that $\Pr(\tilde{R}_{t+1} \in [\underline{R}, \bar{R}]) = 1$. Thus from equation (A.3) we know that there exists an $\bar{\eta} > 0$ such that $\Pr(\eta_{t+1} < \bar{\eta}) = 1$. Thus (A4) is satisfied.

From Assumption $5'$ we know that $\Pr(\tilde{R}_{t+1} \in [\underline{R}, \bar{R}]) = 1$. Thus d_{t+1} is bounded, since d_{t+1} is a continuous function of $\tilde{R}_{t+1}$ (See equation (A.2)). We also know that d_{t+1} is bounded away from zero since $\underline{R} > 0$. Thus there exists an $c_\rho > 1$ such that $\Pr(\frac{1}{c_\rho} < d_{t+1} < c_\rho) = 1$. Thus (A5) is satisfied.

By Assumptions $6'$ and $7'$, (A6) and (A7) are satisfied.

We have verified (A1)-(A7) of Assumption 1.2 in Roitershtein (2007). By Lemma 2.3 of Roitershtein (2007) we know that, for $x > 0$, the following

[25] $\mathbb{Z}$ denotes the set of integers.

limit exists,

$$\Lambda(x) = \lim_{n\to\infty} \frac{1}{n} \log E\left[\prod_{i=1}^{n-1} (d_i)^x\right].$$

We then show the following lemma.

Lemma 3 $\Lambda(x)$ *is a convex function of* $x > 0$*.*

Proof: Note that $\log E\left[\prod_{i=1}^{n-1} (d_i)^x\right] = \log E\left[\left(\prod_{i=1}^{n-1} d_i\right)^x\right]$. Viewing $\left(\prod_{i=1}^{n-1} d_i\right)$ as a random variable, we know, from Lemma 1, that $\log E\left[\left(\prod_{i=1}^{n-1} d_i\right)^x\right]$ is a convex function of $x > 0$. Thus $\frac{1}{n} \log E\left[\prod_{i=1}^{n-1} (d_i)^x\right]$ is a convex function of $x > 0$. Then we know that $\Lambda(x)$ is a convex function of $x > 0$. □

Thus we have Lemma 4 as a corollary to Lemma 3.

Lemma 4 $\Lambda(x)$ *is a continuous function of* $x > 0$*.*

Proof: By Proposition 17 of Chapter 5 in Royden (1988), Lemma 3 implies Lemma 4. □

Lemmas 3 and 4 imply that $\Lambda(x)$ is convex and continuous. From Assumption 6′ we know that there exist $\mu_1 > 1$ and $\mu_2 > 1$ such that $\Lambda(\mu_1) \geq 0$ and $\Lambda(\mu_2) < 0$. Thus there exists a unique $\mu > 1$ such that

$$\Lambda(\mu) = 0.$$

Applying Theorem 1.5 of Roitershtein (2007) to the process $\{L_t\}$, we have

$$\lim_{x\to\infty} \frac{\Pr(L_\infty > x)}{x^{-\mu}} = c,$$

with $c > 0$. ■

A.7 Proof of Proposition 8

Proof: Suppose that $\chi > \chi'$. Thus by equation (A.2) we know that

$$d_{t+1} = \frac{\tilde{R}_{t+1}(1-b)}{1+\left[\tilde{R}_{t+1}(1-b)\right]^{\frac{\gamma-1}{\gamma}} \chi^{-\frac{1}{\gamma}}} > d'_{t+1} = \frac{\tilde{R}_{t+1}(1-b)}{1+\left[\tilde{R}_{t+1}(1-b)\right]^{\frac{\gamma-1}{\gamma}} (\chi')^{-\frac{1}{\gamma}}}.$$

Let

$$\Lambda_1(x) = \lim_{n\to\infty} \frac{1}{n} \log E\left[\prod_{i=1}^{n-1} (d'_i)^x\right],$$

for $x > 0$. Suppose that $\Lambda_1(\mu') = 0$. Thus we have

$$\Lambda(\mu') = \lim_{n\to\infty} \frac{1}{n} \log E\left[\prod_{i=1}^{n-1} (d_i)^{\mu'}\right] \geq \lim_{n\to\infty} \frac{1}{n} \log E\left[\prod_{i=1}^{n-1} (d'_i)^{\mu'}\right] = 0.$$

From Assumption 6′ we know that there exists $\mu_2 > 1$ such that $\Lambda(\mu_2) < 0$. Then we know that there exists a $\mu > 1$ such that $\Lambda(\mu) = 0$, and $\mu \leq \mu'$.

■

A.8 Proof of Proposition 9

Proof: I apply Theorem 1.8 of Mirek (2011) to prove Proposition 9. I need to verify Assumptions 1.6 and 1.7 of Mirek (2011).

Let $\vartheta = \left(\tilde{R}, H\right)$ and

$$\psi_\vartheta(L) = \begin{cases} H - (1-b)\tilde{R}\theta L + G, & if\ L \leq \phi \\ dL + \eta, & otherwise \end{cases}.$$

The process $\{L_t\}$ is generated by $L_{t+1} = \psi_\vartheta(L_t)$. Thus $\psi_\vartheta(L)$ is Lipschitz continuous.

Verification of Assumption 1.6 of Mirek (2011). For every $z > 0$, let

$$\psi_{\vartheta,z}(L) = z\psi_{\vartheta}\left(\frac{1}{z}L\right).$$

$\psi_{\vartheta,z}$ are called dilatations of ψ_{ϑ}. Let

$$\bar{\psi}_{\vartheta}(L) = \lim_{z\to 0}\psi_{\vartheta,z}(L).$$

Thus we have

$$\bar{\psi}_{\vartheta}(L) = \lim_{z\to 0}\psi_{\vartheta,z}(L) = \lim_{z\to 0}\left[z\psi_{\vartheta}\left(\frac{1}{z}L\right)\right] = dL, \text{ for } \forall L \geq 0.$$

Since $\psi_{\vartheta}(L)$ is piecewise-linear, It is easy to find a random variable N_{ϑ} with bounded support such that

$$|\psi_{\vartheta}(L) - dL| \leq N_{\vartheta}, \text{ for } \forall L \geq 0.$$

Assumption 1.6 of Mirek (2011) is satisfied.

Verification of Assumption 1.7 of Mirek (2011). As for Assumption 1.7 of Mirek (2011), condition (H3) is satisfied since

$$d_{t+1} = \frac{\tilde{R}_{t+1}(1-b)}{1 + \left[\tilde{R}_{t+1}(1-b)\right]^{\frac{\gamma-1}{\gamma}}\chi^{-\frac{1}{\gamma}}},$$

by equation (A.2). $\left\{\tilde{R}_t\right\}$ is $i.i.d.$ along time and the support of $\tilde{R}_t$ is closed.

The law of $\log d$ is non-arithmetic since $\tilde{R}_t$ has a probability density function $l(\cdot)$ on $[\underline{R},\bar{R}]$ by Assumption $3''$. Thus (H4) in Assumption 1.7 of Mirek (2011) is satisfied.

Let $m(x) = \log E(d)^x$. From Assumption $4''$ we have $E(d) < 1$. Thus $m(1) < 0$. By Assumption $5''$, $E(d)^2 > 1$. Thus $m(2) > 0$. From Lemmas 1 and 2 we know that $m(x)$ is convex and continuous. Thus there exists a unique $\mu \in (1, 2)$ such that $m(\mu) = 0$, i.e.

$$E(d)^\mu = 1.$$

We also know that $E(d^\mu |\log d|) < \infty$, since d is bounded.

$E[(N_\vartheta)^\mu] < \infty$, since N_ϑ is bounded.

Assumption 1.7 of Mirek (2011) is satisfied.

Applying Theorem 1.8 of Mirek (2011), we find that the stationary distribution of the process $\{L_t\}$, L_∞, has an asymptotic Pareto tail, i.e.

$$\lim_{x\to\infty} \frac{\Pr(L_\infty > x)}{x^{-\mu}} = c,$$

with $c > 0$. ■

A.9 Unobservable $\tilde{R}_{t+1}$

Here I assume that the gross interest rate, $\tilde{R}_{t+1}$, and labor earnings of children H_{t+1} are unobservable for parents. But parents know the distributions of $\tilde{R}_{t+1}$ and H_{t+1}.

Lemma 5 *$C(L_t)$ is an increasing function of L_t.*

Proof: Let

$$\begin{aligned} & f(C_t; L_t) \\ = \; & C_t - \min\left\{L_t, [(1-b)\chi]^{-\frac{1}{\gamma}}\left\{E\left(\tilde{R}_{t+1}\left[(1-b)\tilde{R}_{t+1}(L_t - C_t) + H_{t+1} + G\right]^{-\gamma}\right)\right\}^{-\frac{1}{\gamma}}\right\}. \end{aligned}$$

Thus $f(C_t; L_t)$ is increasing in C_t. The optimal C_t satisfies $f(C_t; L_t) = 0$.

Suppose that $\hat{L}_t > L_t$. Then we have

$$\begin{aligned} & f\left(C_t; \hat{L}_t\right) \\ = \;& C_t - \min\left\{L_t, [(1-b)\chi]^{-\frac{1}{\gamma}}\left\{E\left(\tilde{R}_{t+1}\left[(1-b)\tilde{R}_{t+1}\left(\hat{L}_t - C_t\right) + H_{t+1} + G\right]^{-\gamma}\right)\right\}^{-\frac{1}{\gamma}}\right\} \\ \leq \;& 0. \end{aligned}$$

Since the optimal $\hat{C}_t$ satisfies $f\left(\hat{C}_t; \hat{L}_t\right) = 0$, we know that $\hat{C}_t \geq C_t$. ■

Lemma 6 *$B(L_t)$ is an increasing function of L_t.*

Proof: Let

$$\begin{aligned} & g(B_t; L_t) \\ = \;& B_t - \max\left\{0, L_t - [(1-b)\chi]^{-\frac{1}{\gamma}}\left\{E\left(\tilde{R}_{t+1}\left[(1-b)\tilde{R}_{t+1}B_t + H_{t+1} + G\right]^{-\gamma}\right)\right\}^{-\frac{1}{\gamma}}\right\}. \end{aligned}$$

Thus $g(B_t; L_t)$ is increasing in B_t. The optimal B_t satisfies $g(B_t; L_t) = 0$.

Suppose that $\hat{L}_t > L_t$. Then we have

$$\begin{aligned} & g\left(B_t; \hat{L}_t\right) \\ = \;& B_t - \max\left\{0, \hat{L}_t - [(1-b)\chi]^{-\frac{1}{\gamma}}\left\{E\left(\tilde{R}_{t+1}\left[(1-b)\tilde{R}_{t+1}B_t + H_{t+1} + G\right]^{-\gamma}\right)\right\}^{-\frac{1}{\gamma}}\right\} \\ \leq \;& 0. \end{aligned}$$

Since the optimal $\hat{B}_t$ satisfies $g\left(\hat{B}_t; \hat{L}_t\right) = 0$, we know that $\hat{B}_t \geq B_t$. ■

Lemma 7 *Both $C(L_t)$ and $B(L_t)$ are Lipschitz continuous.*

Proof: For L_t and $\hat{L}_t$, without loss of generality, we assume that $L_t < \hat{L}_t$. By Lemmas (5) and (6) we know that $C\left(\hat{L}_t\right) \geq C(L_t)$ and $B\left(\hat{L}_t\right) \geq B(L_t)$. Also $C(L_t) + B(L_t) = L_t$ and $C\left(\hat{L}_t\right) + B\left(\hat{L}_t\right) = \hat{L}_t$. Thus

$$C\left(\hat{L}_t\right) - C(L_t) + B\left(\hat{L}_t\right) - B(L_t) = \hat{L}_t - L_t.$$

Thus we have

$$0 \leq C\left(\hat{L}_t\right) - C(L_t) \leq \hat{L}_t - L,$$

and

$$0 \leq B\left(\hat{L}_t\right) - B(L_t) \leq \hat{L}_t - L_t.$$

Thus

$$|C\left(\hat{L}_t\right) - C(L_t)| \leq |\hat{L}_t - L|,$$

and

$$|B\left(\hat{L}_t\right) - B(L_t)| \leq |\hat{L}_t - L_t|.$$

We know that both $C(L_t)$ and $B(L_t)$ are Lipschitz continuous. ■

Lemma 8 $\frac{C(L_t)}{L_t}$ *is decreasing in* L_t*.*

Proof: If $L_t \leq [(1-b)\chi]^{-\frac{1}{\gamma}}\left[E\left(\tilde{R}_{t+1}(H_{t+1}+G)^{-\gamma}\right)\right]^{-\frac{1}{\gamma}}$, from the first order condition (17) we know that $B(L_t) = 0$. Thus $\frac{C(L_t)}{L_t} = 1$. Otherwise, we have $\frac{C(L_t)}{L_t} < 1$.

For $\hat{L}_t > L_t > [(1-b)\chi]^{-\frac{1}{\gamma}}\left[E\left(\tilde{R}_{t+1}(H_{t+1}+G)^{-\gamma}\right)\right]^{-\frac{1}{\gamma}}$, let $C_t = C(L_t)$ and $\hat{C}_t = C\left(\hat{L}_t\right)$. From the first order condition (17) we have

$$1 = (1-b)\chi E\left(\tilde{R}_{t+1}\left[(1-b)\tilde{R}_{t+1}\left(\frac{L_t}{C_t} - 1\right) + \frac{H_{t+1}+G}{C_t}\right]^{-\gamma}\right), \quad (20)$$

and

$$1 = (1-b)\chi E\left(\tilde{R}_{t+1}\left[(1-b)\tilde{R}_{t+1}\left(\frac{\hat{L}_t}{\hat{C}_t}-1\right)+\frac{H_{t+1}+G}{\hat{C}_t}\right]^{-\gamma}\right). \quad (21)$$

Suppose that $\frac{\hat{C}_t}{\hat{L}_t} > \frac{C_t}{L_t}$. Thus $\frac{\hat{L}_t}{\hat{C}_t} < \frac{L_t}{C_t}$. From Lemma 5 we know that $\hat{C}_t \geq C_t$. Thus $\frac{H_{t+1}+G_{t+1}}{\hat{C}_t} \leq \frac{H_{t+1}+G_{t+1}}{C_t}$. From equation (20) we have

$$\begin{aligned} 1 &= (1-b)\chi E\left(\tilde{R}_{t+1}\left[(1-b)\tilde{R}_{t+1}\left(\frac{L_t}{C_t}-1\right)+\frac{H_{t+1}+G}{C_t}\right]^{-\gamma}\right) \\ &< (1-b)\chi E\left(\tilde{R}_{t+1}\left[(1-b)\tilde{R}_{t+1}\left(\frac{\hat{L}_t}{\hat{C}_t}-1\right)+\frac{H_{t+1}+G}{\hat{C}_t}\right]^{-\gamma}\right), \end{aligned}$$

which contradicts with equation (21). Thus we have $\frac{\hat{C}_t}{\hat{L}_t} \leq \frac{C_t}{L_t}$. ■

Thus $\frac{B(L_t)}{L_t} = 1 - \frac{C(L_t)}{L_t}$ is increasing in L_t.

Lemma 9 $\frac{C(L_t)}{L_t} \geq \phi$, *where*

$$\phi = \frac{1}{1+\left[E\left(\tilde{R}_{t+1}\right)^{1-\gamma}\right]^{\frac{1}{\gamma}}(1-b)^{\frac{1-\gamma}{\gamma}}\chi^{\frac{1}{\gamma}}}.$$

Proof: If $L_t \leq [(1-b)\chi]^{-\frac{1}{\gamma}}\left[E\left(\tilde{R}_{t+1}\left(H_{t+1}+G\right)^{-\gamma}\right)\right]^{-\frac{1}{\gamma}}$, from the first order condition (17) we know that $B(L_t)=0$. Thus $\frac{C(L_t)}{L_t}=1$.

If $L_t > [(1-b)\chi]^{-\frac{1}{\gamma}}\left[E\left(\tilde{R}_{t+1}\left(H_{t+1}+G\right)^{-\gamma}\right)\right]^{-\frac{1}{\gamma}}$, from equation (20), we have

$$\begin{aligned} 1 &= (1-b)\chi E\left(\tilde{R}_{t+1}\left[(1-b)\tilde{R}_{t+1}\left(\frac{L_t}{C_t}-1\right)+\frac{H_{t+1}+G}{C_t}\right]^{-\gamma}\right) \\ &\leq (1-b)\chi E\left(\tilde{R}_{t+1}\left[(1-b)\tilde{R}_{t+1}\left(\frac{L_t}{C_t}-1\right)\right]^{-\gamma}\right). \end{aligned}$$

Thus we have

$$\frac{C_t}{L_t} \geq \frac{1}{1 + \left[E\left(\tilde{R}_{t+1}\right)^{1-\gamma}\right]^{\frac{1}{\gamma}} (1-b)^{\frac{1-\gamma}{\gamma}} \chi^{\frac{1}{\gamma}}}.$$

■

Note that $0 < \phi < 1$. Thus $\frac{B(L_t)}{L_t} = 1 - \frac{C(L_t)}{L_t} \leq 1 - \phi < 1$.

Lemma 10 *$\lim_{L_t \to \infty} \frac{C(L_t)}{L_t} = \phi$.*

Proof: From Lemma 8 we know that $\frac{C(L_t)}{L_t}$ is decreasing in L_t. We also know that $\frac{C(L_t)}{L_t} \geq \phi$. Thus $\lim_{L_t \to \infty} \frac{C(L_t)}{L_t}$ exists. Let $\lambda = \lim_{L_t \to \infty} \frac{C(L_t)}{L_t}$. Thus $\lambda \geq \phi$. $\lim_{L_t \to \infty} C(L_t) = \infty$.

Since $\lambda < 1$, we can pick $\varepsilon > 0$ such that $\lambda + \varepsilon < 1$. Let $g = \left(\frac{1}{\lambda+\varepsilon} - 1\right)^{-\gamma} (1-b)^{-\gamma} \left(\tilde{R}_{t+1}\right)^{1-\gamma}$. Thus we have

$$0 < \tilde{R}_{t+1} \left(\frac{L_{t+1}}{C_t}\right)^{-\gamma} = \tilde{R}_{t+1} \left[(1-b)\tilde{R}_{t+1}\left(\frac{1}{\frac{C_t}{L_t}} - 1\right) + \frac{H_{t+1} + G}{C_t}\right]^{-\gamma} \leq g,$$

when L_t is large enough. Note that

$$Eg = \left(\frac{1}{\lambda + \varepsilon} - 1\right)^{-\gamma} (1-b)^{-\gamma} E\left(\tilde{R}_{t+1}\right)^{1-\gamma} < \infty.$$

Using Lebesgue's Dominated Convergence Theorem, we know that equation (20) implies

$$\begin{aligned}
1 &= (1-b)\chi E\left(\tilde{R}_{t+1}\left[(1-b)\tilde{R}_{t+1}\left(\frac{1}{\frac{C_t}{L_t}}-1\right)+\frac{H_{t+1}+G}{C_t}\right]^{-\gamma}\right) \\
&= (1-b)\chi \lim_{L_t\to\infty} E\left(\tilde{R}_{t+1}\left[(1-b)\tilde{R}_{t+1}\left(\frac{1}{\frac{C_t}{L_t}}-1\right)+\frac{H_{t+1}+G}{C_t}\right]^{-\gamma}\right) \\
&= (1-b)\chi E \lim_{L_t\to\infty}\left(\tilde{R}_{t+1}\left[(1-b)\tilde{R}_{t+1}\left(\frac{1}{\frac{C_t}{L_t}}-1\right)+\frac{H_{t+1}+G}{C_t}\right]^{-\gamma}\right) \\
&= (1-b)^{1-\gamma}\chi\left(\frac{1}{\lambda}-1\right)^{-\gamma} E\left(\tilde{R}_{t+1}\right)^{1-\gamma}.
\end{aligned}$$

Thus we have

$$\lambda = \frac{1}{1+\left[E\left(\tilde{R}_{t+1}\right)^{1-\gamma}\right]^{\frac{1}{\gamma}}(1-b)^{\frac{1-\gamma}{\gamma}}\chi^{\frac{1}{\gamma}}} = \phi.$$

■

Thus $\lim_{L_t\to\infty}\frac{B(L_t)}{L_t} = 1-\lim_{L_t\to\infty}\frac{C(L_t)}{L_t} = 1-\phi$.

Lemma 11 *For $L_t > G$, we have*

$$(1-b)\underline{R}B(L_t)+G < L_t.$$

Proof: Since $B(L_t) \geq 0$, we know that $L_{t+1} = (1-b)\underline{R}B(L_t)+G$ is the lowest policy function of wealth accumulation.

For $G < L_t \leq [(1-b)\chi]^{-\frac{1}{\gamma}} \left[E\left(\tilde{R}_{t+1}(H_{t+1}+G)^{-\gamma}\right)\right]^{-\frac{1}{\gamma}}$,[26] from the first order condition (17) we know that $B(L_t) = 0$. Thus $(1-b)\underline{R}B(L_t) + G = G < L_t$.

For $L_t > [(1-b)\chi]^{-\frac{1}{\gamma}} \left[E\left(\tilde{R}_{t+1}(H_{t+1}+G)^{-\gamma}\right)\right]^{-\frac{1}{\gamma}}$, suppose that $(1-b)\underline{R}B(L_t) + G \geq L_t$. Then $L_{t+1} = H_{t+1} + (1-b)\tilde{R}_{t+1}B(L_t) + G \geq L_t$. We also know that $C(L_t) < L_t$, since $B(L_t) > 0$. From the first order condition (17) we have

$$\begin{aligned}
& \left[(1-b)\chi E\left(\tilde{R}_{t+1}\right)\right]^{-\frac{1}{\gamma}} L_t \\
\leq\ & [(1-b)\chi]^{-\frac{1}{\gamma}} \left\{E\left[\tilde{R}_{t+1}(L_{t+1})^{-\gamma}\right]\right\}^{-\frac{1}{\gamma}} \\
=\ & C_t \\
<\ & L_t.
\end{aligned}$$

Thus we have $(1-b)\chi E\left(\tilde{R}_{t+1}\right) > 1$. Using Jensen's inequality we know that Assumption $4'''$ implies that $(1-b)\chi E\left(\tilde{R}_{t+1}\right) < 1$. We have a contradiction. Thus $(1-b)\underline{R}B(L_t) + G < L_t$. ■

Proof of Theorem 3: The proof consists of three lemmas.

Lemma 12 *The individual wealth process $\{L_t\}$ is ψ-irreducible.*

---

[26]Note that

$$\begin{aligned}
& [(1-b)\chi]^{-\frac{1}{\gamma}} \left[E\left(\tilde{R}_{t+1}(H_{t+1}+G)^{-\gamma}\right)\right]^{-\frac{1}{\gamma}} \\
\geq\ & \left[(1-b)\chi E\left(\tilde{R}_{t+1}\right)\right]^{-\frac{1}{\gamma}} G \\
>\ & G,
\end{aligned}$$

since Assumption $4'''$ implies that $(1-b)\chi E\left(\tilde{R}_{t+1}\right) < 1$.

Proof: I first show that the process $\{L_t\}$ is φ-irreducible. We construct a measure φ on $[G, \infty)$ such that

$$\varphi(A) = \int_A f(x - G)dx,$$

where $f(\cdot)$ is the probability density function of labor earnings H_t.

From Lemma 11 we know that, for $\forall L_1 > G$, there is a positive probability such that $\exists t \geq 1$ we have $L_t < [(1-b)\chi]^{-\frac{1}{\gamma}} \left[E\left(\tilde{R}_{t+1}\left(H_{t+1}+G\right)^{-\gamma}\right)\right]^{-\frac{1}{\gamma}}$. Then $B\left(L_t\right) = 0$ and $L_{t+1} = H_{t+1} + G$. Thus any set A such that $\int_A f(x-G)dx > 0$ can be reached in finite time with a positive probability. The process $\{L_t\}$ is φ-irreducible.

By Proposition 4.2.2 of Meyn and Tweedie (2009), there exists a probability measure ψ on $[G, \infty)$ such that the process $\{L_t\}$ is ψ-irreducible, since it is φ-irreducible. □

Lemma 13 *The process $\{L_t\}$ is geometrically ergodic.*

Proof: To show that the process $\{L_t\}$ is geometrically ergodic, I use part (iii) of Theorem 15.0.1 of Meyn and Tweedie (2009). We need to verify that

a the process $\{L_t\}$ is ψ-irreducible;

b the process $\{L_t\}$ is aperiodic;[27] and

c there exists a petite set C,[28] constants $e < \infty$, $\beta > 0$ and a function

[27] See page 114 of Meyn and Tweedie (2009) for the definition of aperiodic.
[28] See page 117 of Meyn and Tweedie (2009) for the definition of petite sets.

$V \geq 1$ finite at some point in $[G, \infty)$ satisfying[29]

$$E_t \left[V(L_{t+1})\right] - V(L_t) \leq -\beta V(L_t) + eI_C(L_t), \; \forall L_t \in [G, \infty).$$

By Lemma 12, the process $\{L_t\}$ is ψ-irreducible.

For a φ-irreducible Markov process, when there exists a v_1-small set A with $v_1(A) > 0$,[30] then the stochastic process is called strongly aperiodic.[31] We construct a measure v_1 on $[G, \infty)$ such that

$$v_1(A) = \int\limits_A f(x - G)dx,$$

where $f(\cdot)$ is the probability density function of labor earnings H_t. From the first order condition (17) we know that $B(L_t) = 0$ for $G < L_t \leq [(1-b)\chi]^{-\frac{1}{\gamma}} \left[E\left(\tilde{R}_{t+1}\left(H_{t+1} + G\right)^{-\gamma}\right)\right]^{-\frac{1}{\gamma}}$. Let

$$\zeta = [(1-b)\chi]^{-\frac{1}{\gamma}} \left[E\left(\tilde{R}_{t+1}\left(H_{t+1} + G\right)^{-\gamma}\right)\right]^{-\frac{1}{\gamma}}.$$

Thus $[G, \zeta]$ is v_1-small and $v_1([G, \zeta]) = \int_G^\zeta f(x - G)dx > 0$. The process $\{L_t\}$ is strongly aperiodic.

I now show that an interval $[G, y]$ is a petite set for $\forall y > G$. From Lemma 11 we know that, for $\forall L_1 \in [G, y]$, there exists a common m such that $\Pr(L_m < \zeta | L_1) \geq \epsilon > 0$. For $L_m < \zeta$, we have $B(L_m) = 0$ and $L_{m+1} = H_{m+1} + G$. Thus

$$\Pr(L_{m+1} \in A | L_1) \geq \Pr(L_{m+1} \in A | L_m < \zeta) \times \Pr(L_m < \zeta | L_1) \geq v_{m+1}(A),$$

[29] I use $I_C(\cdot)$ to denote the indicator function of the set C.

[30] See page 102 of Meyn and Tweedie (2009) for the definition of small sets.

[31] See page 114 of Meyn and Tweedie (2009).

where $v_{m+1}(A) \equiv \epsilon \int_A f(x-G)dx$. Thus $[G,y]$ is v_{m+1}-small. By Proposition 5.5.3 of Meyn and Tweedie (2009), $[G,y]$ is a petite set.

I pick a function $V(L_t) = L_t + 1$, $\forall L_t \in [G,\infty)$. Thus $V(L_t) > 1$ for $L_t \in [G,\infty)$. Pick $0 < q < 1 - E\left(\tilde{\rho}_{t+1}\right)$. Let $\beta = 1 - E\left(\tilde{\rho}_{t+1}\right) - q > 0$ and $e = 1 - E\left(\tilde{\rho}_{t+1}\right) + E\left(H_{t+1}\right) + G$. Pick $y > G$ such that $y + 1 \geq \frac{e}{q}$. Let $C = [G,y]$. Thus C is a petite set. For $\forall L_t \in [G,\infty)$ we have

$$\begin{aligned} & E_t\left[V(L_{t+1})\right] - V(L_t) \\ = \;& E(L_{t+1}) - L_t \\ \leq \;& -\left[1 - E\left(\tilde{\rho}_{t+1}\right)\right] V(L_t) + 1 - E\left(\tilde{\rho}_{t+1}\right) + E\left(H_{t+1}\right) + G \\ \leq \;& -\beta V(L_t) + eI_C(L_t). \end{aligned}$$

By Theorem 15.0.1 of Meyn and Tweedie (2009), the process $\{L_t\}$ is geometrically ergodic. □

Lemma 14 *The support of the stationary distribution L_∞ is unbounded.*

Proof: For $L_t > \zeta$, $B\left(L_t\right) > 0$. Assumption $5'''$ implies that we can pick $\bar{R}$ large enough such that

$$\Pr(L_{t+1} = H_{t+1} + (1-b)\tilde{R}_{t+1} B\left(L_t\right) + G > L_t | L_t) > 0.$$

For $L_t \in [G,\zeta]$, $B\left(L_t\right) = 0$. Assumption $5'''$ implies that

$$\Pr(L_{t+1} = H_{t+1} + G > \zeta | L_t) > 0.$$

Thus the support of the stationary distribution L_∞ is unbounded. □

This concludes the proof of Theorem 3. ■

Proof of Theorem 4: I start the proof by constructing an auxiliary process $\left\{\tilde{L}_t\right\}$.

Construction of $\left\{\tilde{L}_t\right\}$. Since $\frac{B(L_t)}{L_t}$ is increasing in L_t and $\lim_{L_t \to \infty} \frac{B(L_t)}{L_t} = 1-\phi$. Thus for $\varepsilon > 0$ we can find $\hat{L}$ such that

$$1-\phi-\varepsilon < \frac{B\left(\hat{L}\right)}{\hat{L}} \leq 1-\phi.$$

Let

$$s(L_t) = \min\left\{\frac{B\left(L_t\right)}{L_t}, \frac{B\left(\hat{L}\right)}{\hat{L}}\right\} L_t.$$

Thus $0 \leq s(L_t) \leq B\left(L_t\right)$. From Lemma 7 we know that $B\left(L_t\right)$ is Lipschitz continuous. Thus $s(L_t)$ is Lipschitz continuous.

Let $\theta = \left(\tilde{R}, H\right)$ and

$$\psi_\theta(L) = H + (1-b)\tilde{R}s\left(L\right) + G.$$

The process $\left\{\tilde{L}_t\right\}$ is generated by $\tilde{L}_{t+1} = \psi_\theta(\tilde{L}_t)$.

In order to apply Theorem 1.8 of Mirek (2011), we need to verify Assumptions 1.6 and 1.7 of Mirek (2011).

Verification of Assumption 1.6 of Mirek (2011). For every $z > 0$, let

$$\psi_{\theta,z}(L) = z\psi_\theta\left(\frac{1}{z}L\right).$$

$\psi_{\theta,z}$ are called dilatations of ψ_θ. Let

$$\bar{\psi}_\theta(L) = \lim_{z\to 0} \psi_{\theta,z}(L).$$

Thus we have

$$\bar{\psi}_\theta(L) = \lim_{z\to 0}\psi_{\theta,z}(L) = \lim_{z\to 0}\left[z\psi_\theta\left(\frac{1}{z}L\right)\right] = (1-b)\tilde{R}\frac{B\left(\hat{L}\right)}{\hat{L}}L,\ \text{for } \forall L \in [G,\infty).$$

Let $M = (1-b)\tilde{R}\frac{B(\hat{L})}{\hat{L}}$. Thus

$$\bar{\psi}_\theta(L) = ML.$$

Let

$$N_\theta = H + (1-b)\tilde{R}\Omega + G,$$

where

$$\Omega = \max_{L\in[G,\hat{L}]} |s\,(L) - \frac{B\left(\hat{L}\right)}{\hat{L}}L|.$$

It is easy to verify that

$$|\psi_\theta(L) - ML| \leq N_\theta,\ \text{for } \forall L \in [G,\infty).$$

Assumption 1.6 of Mirek (2011) is satisfied.

Verification of Assumption 1.7 of Mirek (2011). As for Assumption 1.7 of Mirek (2011), condition (H3) is satisfied since $M = (1-b)\tilde{R}\frac{B(\hat{L})}{\hat{L}}$, $\left\{\tilde{R}_t\right\}$ is *i.i.d.* along time and the support of $\tilde{R}_t$ is closed. The law of $\log M$ is non-arithmetic since $\tilde{R}_t$ has a probability density function $l(\cdot)$ on $[\underline{R},\bar{R}]$ by Assumption $3'''$. Thus (H4) in Assumption 1.7 of Mirek (2011) is satisfied.

Let $m(x) = \log E(M)^x$. Note that

$$M = (1-b)\tilde{R}\frac{B\left(\hat{L}\right)}{\hat{L}} \leq (1-\phi)\,(1-b)\tilde{R} = \tilde{\rho}.$$

Thus we have

$$m(1) = \log E(M) \leq \log E(\tilde{\rho}) < 0.$$

I now show that there exists $\kappa > 1$ such that $m(\kappa) = \log E\left(M^{\kappa}\right) > 0$. Note that

$$E\left(\left[(1-\phi)(1-b)\tilde{R}\right]^{\kappa}\right) \geq \int\limits_{\left\{(1-\phi)(1-b)\tilde{R}>1\right\}} \left[(1-\phi)(1-b)\tilde{R}\right]^{\kappa}.$$

We know that $\tilde{R}_t$ has a probability density function $l(\cdot)$ on $[\underline{R},\bar{R}]$ by Assumption 3‴. And $\bar{R}$ is large enough by Assumption 5‴. Thus

$$\Pr((1-\phi)(1-b)\tilde{R} > 1) > 0.$$

Then we know that there exists $\kappa > 1$ such that $E\left(\left[(1-\phi)(1-b)\tilde{R}\right]^{\kappa}\right) > 1$. Thus we can pick $\hat{L}$ large enough such that $E\left(M^{\kappa}\right) > 1$. Thus $m(\kappa) = \log E\left(M^{\kappa}\right) > 0$. From Lemmas 1 and 2 we know that $m(x)$ is convex and continuous. Thus there exists a unique $\mu > 1$ such that $m(\mu) = 0$, i.e. $E\left(M^{\mu}\right) = 1$.

We also know that $E\left(M^{\mu}|\log M|\right) < \infty$, since M is bounded.

$E\left[(N_{\theta})^{\mu}\right] < \infty$, since N_{θ} is bounded.

Assumption 1.7 of Mirek (2011) is satisfied.

The comparison method. Applying Theorem 1.8 of Mirek (2011), we find that the stationary distribution of the process $\left\{\tilde{L}_t\right\}$, $\tilde{L}_{\infty}$, has an asymptotic Pareto tail, i.e.

$$\lim_{x\to\infty} \frac{\Pr(\tilde{L}_{\infty} > x)}{x^{-\mu}} = c,$$

with $c > 0$.

Pick $L_1 = \tilde{L}_1$. The process $\{L_t\}$ is generated by

$$L_{t+1} = H_{t+1} + (1-b)\tilde{R}_{t+1}B\left(L_t\right) + G.$$

And the process $\left\{\tilde{L}_t\right\}$ is generated by

$$\tilde{L}_{t+1} = H_{t+1} + (1-b)\tilde{R}_{t+1}s\left(\tilde{L}_t\right) + G.$$

For a path of $\left\{\left(\tilde{R}_t, H_t\right)\right\}$, we have $L_t \geq \tilde{L}_t$, $\forall t \geq 1$. Thus for $\forall x > G$, we have

$$\Pr(L_t \geq x) \geq \Pr(\tilde{L}_t \geq x), \text{ for } \forall t \geq 1.$$

Thus we have

$$\Pr(L_\infty \geq x) \geq \Pr(\tilde{L}_\infty \geq x),$$

since processes $\{L_t\}$ and $\left\{\tilde{L}_t\right\}$ are ergodic. Thus

$$\liminf_{x\to\infty} \frac{\Pr(L_\infty > x)}{x^{-\mu}} \geq \liminf_{x\to\infty} \frac{\Pr(\tilde{L}_\infty > x)}{x^{-\mu}} = \lim_{x\to\infty} \frac{\Pr(\tilde{L}_\infty > x)}{x^{-\mu}} = c.$$

■

Proof of Theorem 5: Under b, we have

$$\phi = \frac{1}{1 + \left[E\left(\tilde{R}_{t+1}\right)^{1-\gamma}\right]^{\frac{1}{\gamma}} (1-b)^{\frac{1-\gamma}{\gamma}} \chi^{\frac{1}{\gamma}}}.$$

And under b' we have

$$\phi' = \frac{1}{1 + \left[E\left(\tilde{R}_{t+1}\right)^{1-\gamma}\right]^{\frac{1}{\gamma}} (1-b')^{\frac{1-\gamma}{\gamma}} \chi^{\frac{1}{\gamma}}}.$$

From Theorem 4 we know that there exists $\mu' > 1$ such that

$$\lim_{x\to+\infty}\inf \frac{\Pr(L'_\infty > x)}{x^{-\mu'}} \geq c',$$

with $c' > 0$. And μ' solves

$$E\left(\left[(1-b')\tilde{R}\frac{B'\left(\hat{L}'\right)}{\hat{L}'}\right]^{\mu'}\right) = 1.$$

Note that

$$(1-\phi)(1-b) = \frac{\left[E\left(\tilde{R}_{t+1}\right)^{1-\gamma}\right]^{\frac{1}{\gamma}}\chi^{\frac{1}{\gamma}}}{(1-b)^{-\frac{1}{\gamma}} + \left[E\left(\tilde{R}_{t+1}\right)^{1-\gamma}\right]^{\frac{1}{\gamma}}(1-b)^{-1}\chi^{\frac{1}{\gamma}}}.$$

Since $b < b'$, we have

$$(1-\phi')(1-b') < (1-\phi)(1-b).$$

Since $\frac{B(L_t)}{L_t}$ is increasing in L_t and $\lim_{L_t\to\infty}\frac{B(L_t)}{L_t} = 1-\phi$. Thus we can find $\hat{L}$ such that

$$(1-\phi')(1-b') < \frac{B\left(\hat{L}\right)}{\hat{L}}(1-b) \leq (1-\phi)(1-b).$$

Repeating the proof of Theorem 4 we know that there exists $\mu > 1$ such that

$$\lim_{x\to+\infty}\inf \frac{\Pr(L_\infty > x)}{x^{-\mu}} \geq c,$$

with $c > 0$. And μ solves

$$E\left(\left[(1-b)\tilde{R}\frac{B\left(\hat{L}\right)}{\hat{L}}\right]^{\mu}\right) = 1.$$

Since $\frac{B'\left(\hat{L}'\right)}{\hat{L}'}(1-b') \leq (1-\phi')\,(1-b')$, we have

$$\frac{B'\left(\hat{L}'\right)}{\hat{L}'}(1-b') \leq \frac{B\left(\hat{L}\right)}{\hat{L}}(1-b).$$

Thus we know that

$$(1-b)\tilde{R}\frac{B\left(\hat{L}\right)}{\hat{L}} \succeq_{FSD} (1-b')\tilde{R}\frac{B'\left(\hat{L}'\right)}{\hat{L}'}.$$

From the proof of Theorem 2 we have $\mu \leq \mu'$. ■