Wealth Distribution with Random Discount Factors

Alexis Akira Toda*

November 5, 2017

Abstract

It is well-known that empirical wealth distributions have Pareto tails. To explain this fact, the quantitative macro literature has occasionally assumed that agents have random discount factors. However, the fact that random discounting generates Pareto tails is a 'folk theorem' that has been shown only in very particular settings (*e.g.*, i.i.d. environment or affine rule-of-thumb consumption rule). Using a highly stylized but fully specified heterogeneous-agent dynamic general equilibrium model of consumption and savings with an arbitrary Markovian dynamics for the discount factor, I prove that the upper and lower tails of the wealth distribution obey power laws and characterize the Pareto exponents. I also discuss a numerical example and show that there is no clear relationship between the return on wealth and inequality and that the Pareto exponent is highly sensitive to the persistence of the discount factor process.

Keywords: consumption-savings problem, double power law, inequality, Pareto exponent.

JEL codes: C62, D31, D58, E21.

1 Introduction

One of the key features of the empirical wealth distribution in many countries is that it obeys the power law: the fraction of agents with wealth w or larger is approximately proportional to a power function $w^{-\alpha}$, where $\alpha > 0$ is called the power law (or Pareto) exponent.[1] Since a power law distribution does not have moments beyond order α, it is more skewed and has heavier tails than many commonly used distributions such as the exponential, gamma, or lognormal. A typical value for the Pareto exponent for wealth is 1.5 (Pareto, 1896; Klass et al., 2006; Vermeulen, 2017).

Economists have long been interested in explaining the wealth distribution. An early example is Champernowne (1953), who obtains a stationary wealth distribution that obeys the power law in both the upper and lower tails, although

*Department of Economics, University of California San Diego. Email: atoda@ucsd.edu.
I thank John Stachurski for suggesting some of the proof techniques and Jinhui Bai, Dan Cao, Chris Carroll, Alisdair McKay, and Tommaso Porzio for comments and suggestions that improved the paper.

[1]More precisely, a random variable X has a Paretian tail if $0 < A = \lim_{x\to\infty} x^\alpha \Pr(X > x)$ exists. See Gabaix (2009, 2016) for an introduction to power laws.

his model is not micro-founded. Since the 1990s researchers have tried to explain the wealth distribution from fully specified dynamic general equilibrium models. Early examples are Banerjee and Newman (1991), Huggett (1993), and Aiyagari (1994). In these models in which agents are subject to uninsurable income risk alone, however, it has been found that the model-generated wealth distribution does not have a fat enough upper tail compared to the data, because the precautionary savings motive is not strong at high wealth levels (Carroll, 1997). One particular mechanism introduced by Krusell and Smith (1998) in order to overcome this issue is to assume that agents have random discount factors. Since agents are more patient in some states than in others, it introduces heterogeneity in saving rates, which makes the wealth distribution more skewed. Although such heterogeneous-agent models are successful in explaining the wealth distribution, there seems to be a huge gap between the numerical results and the theoretical understanding of models.

The present paper attempts to fill this gap. I build a stylized dynamic general equilibrium model with heterogeneous agents to study the saving behavior and the wealth distribution when agents have random discount factors. In order to isolate the effect of random discounting on the wealth distribution, I abstract from all other forms of uncertainty or heterogeneity: agents have identical constant relative risk aversion (CRRA) preferences and identical, constant endowment. The only uncertainty in the economy is that agents discount future utility randomly; more precisely, the discount factor of an agent evolves according to some Markov chain.[2] We can interpret random discounting in several ways. Krusell and Smith (1998) view households as dynasties and justify random discounting as genetic differences in the population that are passed on imperfectly from parents to children. Random discounting may also be viewed as heterogeneity in the wealth portfolio returns due to differences in financial sophistication or entrepreneurship (Calvet et al., 2007; Cao and Luo, 2017). Yet another possible interpretation is the liquidity shocks during the lifetime of agents (Diamond and Dybvig, 1983; Geanakoplos and Walsh, 2017). With the last interpretation, agents trade a risk-free bond in order to self-insure against liquidity shocks. In this model I establish two main theoretical results: (i) a stationary equilibrium always exists if the "average" discount factor is less than 1 (and is unique if the relative risk aversion is bounded above by 1), and (ii) the stationary wealth distribution obeys the power law in both the upper and lower tails when agents are born and die with constant probability.[3] Furthermore, I provide an analytical characterization of the Pareto exponents.

I must acknowledge outright that these results are largely expected and not necessarily surprising. The main contribution of the paper is that I establish these results rigorously, instead of just providing the intuition or verifying through numerical simulations as is often the case in this literature. Regarding (i) existence, although in my model agents solve a standard Samuelson (1969)-type optimal consumption-savings problem, when agents have random discount factors, the Bellman operator is not a contraction depending on the level of risk

[2]In my model I also assume that agents have uncertain lifetime (Yaari (1965)-Blanchard (1985) perpetual youth model) in order to obtain a stationary wealth distribution, but since I assume perfect annuity markets, this uncertainty is perfectly insured.

[3]This "double power law" has been empirically observed in cities (Reed, 2002; Giesen et al., 2010), income (Reed and Jorgensen, 2004; Toda, 2012), and consumption (Toda and Walsh, 2015; Toda, 2017a).

aversion.[4] Consequently, it is not obvious how to show that the excess demand function is a continuous function of the interest rate and crosses zero. To attack this problem, I use a technique recently introduced by Borovička and Stachurski (2017), who give a necessary and sufficient condition for existence, uniqueness, and computability of a fixed point of a certain monotone operator. Because their condition is necessary and sufficient, we can study the behavior of the fixed point as the interest rate tends to the upper bound for the existence of a solution, which enables me to show that the excess demand function crosses zero. Regarding (ii) double power law, I use recent results in Beare and Toda (2017), who characterize the tail behavior of non-Gaussian, Markovian random growth processes. As I discuss in the related literature, theoretically rigorous derivations of power law distributions from random growth models in economics are mostly limited to i.i.d. processes, which are not applicable in models with Markovian dynamics and fully optimizing agents.

Finally, I present a numerical example and ask whether there is a positive relationship between the return on wealth and inequality (as suggested by Piketty, 2014), and which parameter is important in determining the Pareto exponent. In Proposition 4.2 I show that in a small open economy (partial equilibrium model) with relative risk aversion bounded above by 1, increasing the interest rate (return on wealth) unambiguously lowers the Pareto exponent, and hence increases inequality. However, this is not necessarily the case in a general equilibrium model. While there is a negative relationship between the interest rate and Pareto exponent when we change the average discount factor, the relation flips when we change the risk aversion, persistence of the discount factor process, or average life span of agents. Furthermore, the numerical value of the Pareto exponent is highly sensitive to the persistence: when we increase the persistence from 0.9 to 0.995, the Pareto exponent decreases from 31 to 1.5.

1.1 Related literature

In a calibrated life-cycle model with idiosyncratic income risk, Huggett (1996) shows that the model-generated wealth distribution matches the Gini coefficient in U.S. but misses the upper tail. Krusell and Smith (1998) find that a heterogeneous-agent model with idiosyncratic income risk alone cannot generate enough cross-sectional dispersion in the wealth distribution, but can do so by making the discount factors random.[5] Beare and Toda (2017) prove that in a CARA Huggett economy with birth and death, the wealth distribution has exponential tails, which are thin. Quadrini (2000) and Cagetti and De Nardi (2006) introduce idiosyncratic investment risk in quantitative models and show that they can match the upper tail of the wealth distribution. See Benhabib and Bisin (2017) for a more detailed review of this literature.

[4] To be precise, the Bellman operator is a contraction if $\gamma > 1$ and hence a standard proof applies. The log utility case ($\gamma = 1$) is exactly solvable. When $\gamma < 1$, the Bellman operator is not a contraction so proving theorems becomes tricky.

[5] Random discounting is used in Krusell and Smith (1997, 1999), Mukoyama and Şahin (2006), Erosa and Koreshkova (2007), Hendricks (2007), Krusell et al. (2009), Bachmann and Bai (2013), Carroll et al. (2017), and Hubmer et al. (2016), among others. In Piketty and Zucman (2015) agents are subject to idiosyncratic saving shocks. Another common approach in generating wealth inequality is to assume that agents have heterogeneous but nonstochastic discount factors (Krueger et al., 2016; McKay and Reis, 2016; McKay, 2017), although this type of models do not generate Pareto tails without further assumptions.

Benhabib et al. (2011) provide the first rigorous proof that idiosyncratic investment risk can generate Pareto tails. They consider an overlapping generations model in which households face idiosyncratic rates of return on wealth and earnings at birth (which remain constant in their lifetime). They show that the equation of motion for inherited wealth follows the so called Kesten (1973) process, which admits a stationary distribution with a Pareto upper tail. (An earlier paper by Nirei and Souma (2007) obtains similar results but in their model agents use a rule-of-thumb consumption policy.) Benhabib et al. (2016) consider a continuous-time perpetual youth model (agents are born and die with constant probability) in which agents are subject to idiosyncratic investment risk generated by a Brownian motion and show that the stationary wealth distribution is double Pareto (Reed, 2001). Toda (2014) considers a discrete-time heterogeneous-agent model with Markovian dynamics and shows that the stationary wealth distribution converges to double Pareto in the continuous-time limit.[6] Using a stylized model in which agents have an affine consumption rule (so the equation of motion for wealth is a Kesten process), Benhabib et al. (2017) show that the Pareto exponent is either (i) equal to the Pareto exponent for income, or (ii) entirely determined by the distribution of idiosyncratic returns on wealth (which may be due to investment risk or random discount factors). However, they do not provide an example where the affine consumption rule is optimal. Benhabib et al. (2015) show that in a Bewley (1983) model with i.i.d. idiosyncratic investment risk, the optimal consumption rule is asymptotically linear and that the stationary wealth distribution has a Pareto upper tail. In summary, there are no results in the literature that rigorously show that random discount factors with Markovian (non-i.i.d.) dynamics generate Pareto tails.

2 Optimal savings with random discounting

In this section I consider the optimal consumption-savings problem of an agent who faces a constant interest rate, which I subsequently embed into a general equilibrium model.

Time is indexed by $t = 0, 1, 2, \dots$. The agent is endowed with initial wealth $w_0 > 0$ and nothing thereafter. The agent can save at a constant gross risk-free rate $R > 0$. Letting w_t be the wealth at the beginning of time t and c_t be consumption, the budget constraint is

$$w_{t+1} = R(w_t - c_t). \tag{2.1}$$

The period utility function exhibits constant relative risk aversion (CRRA) $\gamma > 0$, so

$$u(c) = \begin{cases} \log c, & (\gamma = 1), \\ \frac{c^{1-\gamma}}{1-\gamma}. & (\gamma \neq 1) \end{cases}$$

At any point in time, the agent can be in one of the states indexed by $s \in S = \{1, \dots, S\}$. The discount factor in state s is denoted by $\beta_s > 0$.[7] The states evolve according to a time-homogeneous Markov chain with transition

[6]Krebs (2006) and Toda (2015) also use analytically tractable heterogeneous-agent models with Markovian dynamics, but they do not consider the wealth distribution.

[7]As Kocherlakota (1990) argues, there is no need to restrict attention to the case $\beta_s < 1$.

probability matrix $P = (p_{ss'})$. Throughout the paper I maintain the following assumption.

Assumption 1 (Irreducibility)**.** *The transition probability matrix $P = (p_{ss'})$ is irreducible.*

Recall that a nonnegative matrix P is irreducible if for any pair (s, s'), there exists n such that $(P^n)_{ss'} > 0$. Intuitively, irreducibility means that one can move between any two states within finite time with positive probability. The irreducibility assumption is without loss of generality since if P is reducible, there exist states that are never reached if agents start from some state. Therefore we can remove such states.

Using the budget constraint (2.1), the agent's Bellman equation becomes

$$V_s(w) = \max_c \left\{ \frac{c^{1-\gamma}}{1-\gamma} + \beta_s \operatorname{E}\left[V_{s'}(R(w-c)) \mid s\right] \right\}. \tag{2.2}$$

The following proposition characterizes the solution to the optimal consumption-savings problem when $\gamma \neq 1$. (The case with log utility ($\gamma = 1$) is treated in Appendix B, which is far simpler.) For a square matrix A, let $\rho(A)$ denote its spectral radius (the maximum modulus of all eigenvalues of A).

Proposition 2.1. *Let $R > 0$ be the gross risk-free rate. Then the Bellman equation* (2.2) *has a fixed point if and only if*

$$R^{1-\gamma}\rho(BP) < 1, \tag{2.3}$$

where $B = \operatorname{diag}(\beta_1, \dots, \beta_S)$ is the diagonal matrix of discount factors. In this case, the fixed point is unique and takes the form $V_s(w) = a_s \frac{w^{1-\gamma}}{1-\gamma}$, where $a_s > 0$ solves the nonlinear equation

$$a_s = \left(1 + (\beta_s R^{1-\gamma} \operatorname{E}\left[a_{s'} \mid s\right])^{1/\gamma}\right)^{\gamma}. \tag{2.4}$$

The optimal consumption rule is $c_s(w) = a_s^{-1/\gamma} w$.

Proof.

Step 1. If (2.2) *has a fixed point, it must be of the form $V_s(w) = a_s \frac{w^{1-\gamma}}{1-\gamma}$, where $a_s > 0$ solves* (2.4)*. The optimal consumption rule is $c_s(w) = a_s^{-1/\gamma} w$.*

By homotheticity the value function must be of the form $V_s(w) = a_s \frac{w^{1-\gamma}}{1-\gamma}$ for some $a_s > 0$. Substituting this guess into the Bellman equation (2.2), we get

$$a_s \frac{w^{1-\gamma}}{1-\gamma} = \max_c \left\{ \frac{c^{1-\gamma}}{1-\gamma} + \beta_s \frac{(R(w-c))^{1-\gamma}}{1-\gamma} \operatorname{E}\left[a_{s'} \mid s\right] \right\}. \tag{2.5}$$

For notational simplicity let $\rho_s = \beta_s R^{1-\gamma} \operatorname{E}\left[a_{s'} \mid s\right]$. Then the maximization problem in (2.5) becomes

$$\text{maximize} \quad \frac{1}{1-\gamma}(c^{1-\gamma} + \rho_s (w-c)^{1-\gamma}).$$

The objective function is clearly strictly concave. The first-order condition with respect to c is

$$c^{-\gamma} = \rho_s (w-c)^{-\gamma} \iff c = \frac{w}{1 + \rho_s^{1/\gamma}}.$$

Substituting this into (2.5), it follows that

$$c^{1-\gamma} + \rho_s (w-c)^{1-\gamma} = c^{1-\gamma} + c^{-\gamma}(w-c) = c^{-\gamma} w = (1+\rho_s^{1/\gamma})^\gamma w^{1-\gamma},$$

so cancelling $w^{1-\gamma}/(1-\gamma)$ in the Bellman equation (2.5), we obtain (2.4). The optimal consumption rule is then

$$c = \frac{w}{1+\rho_s^{1/\gamma}} = a_s^{-1/\gamma} w.$$

Step 2. The Bellman equation (2.2) *has a fixed point if and only if* (2.3) *holds.*

Define $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ by $\phi(t) = (1+t^{1/\gamma})^\gamma$ and the $S \times S$ matrix K by $K = R^{1-\gamma} BP$. Then (2.4) is equivalent to $a = \phi(Ka)$, where ϕ is applied element-wise. Since by Assumption 1 P is irreducible and B is a positive diagonal matrix, K is irreducible. By Lemma A.1 in the appendix, a necessary and sufficient condition for the existence and uniqueness of a fixed point is $1 > \rho(K) = R^{1-\gamma}\rho(BP)$.

Step 3. The transversality condition holds, and hence the above value function and consumption rule characterize the solution.

Starting from initial wealth w_0, let $\{c_t\}_{t=0}^\infty$ be the consumption plan generated by the rule $c_s(w) = a_s^{-1/\gamma} w$ and let $\{w_t\}_{t=0}^\infty$ be the associated wealth implied by the budget constraint. Since $V_s(w) = a_s \frac{w^{1-\gamma}}{1-\gamma}$ satisfies the Bellman equation (2.2), by iteration for all $n \in \mathbb{N}$ we obtain

$$V_{s_0}(w_0) = \mathrm{E}_0 \left[\sum_{t=0}^{n-1} \beta(s^t) \frac{c_t^{1-\gamma}}{1-\gamma} + \beta(s^{n-1}) V_{s_n}(w_n) \right],$$

where $\beta(s^t) = \prod_{\tau=0}^t \beta_{s_\tau}$ is the product of discount factors along the history $s^t = (s_0, \dots, s_t)$. By Toda (2014, Proposition 2), the transversality condition

$$\limsup_{n\to\infty} \mathrm{E}_0[\beta(s^{n-1}) V_{s_n}(w_n)] \le 0 \tag{2.6}$$

is sufficient for optimality. If $\gamma > 1$, this is trivial since $V_s(w) \le 0$. Suppose that $\gamma < 1$. Since $c_t \ge 0$, we have $w_{t+1} \le R w_t$, so $w_n \le R^n w_0$. Therefore letting $\bar{a} = \max_s a_s$, we obtain

$$\begin{aligned} \mathrm{E}_0[\beta(s^{n-1}) V_{s_n}(w_n)] &\le \bar{a} \frac{(R^n w_0)^{1-\gamma}}{1-\gamma} \sum_{s^n} \prod_{t=0}^{n-1} (\beta_{s_t} p_{s_t s_{t+1}}) \\ &\le \bar{a} \frac{w_0^{1-\gamma}}{1-\gamma} 1'(R^{1-\gamma} BP)^n 1 \to 0 \end{aligned}$$

as $n \to \infty$ because $R^{1-\gamma}\rho(BP) < 1$ by assumption. □

3 General equilibrium

Having solved the optimal consumption-savings problem with random discount factors, I embed the problem into general equilibrium.

I consider an infinite horizon endowment economy with heterogeneous (but ex ante identical) agents. The economy is populated by a continuum of agents with mass 1 indexed by $i \in [0,1]$ with CRRA preferences as in the previous section. Agents have constant and identical endowment $y > 0$ of a perishable good every period. In this model one can show that agents solve an optimal consumption-savings problem as in Proposition 2.1. Since the optimal consumption rule is $c_s(w) = a_s^{-1/\gamma} w$, using the budget constraint (2.1) the equation of motion for wealth is $w' = R(1 - a_s^{-1/\gamma})w$. By taking the logarithm, the log wealth evolves according to a random walk, which does not have a stationary distribution if agents are infinitely lived.

In order to obtain a stationary wealth distribution, as in Yaari (1965) and Blanchard (1985), I assume that agents are born and die with probability $p > 0$ each period (perpetual youth model). This mechanism for generating a stationary (Pareto) distribution was first discovered by Wold and Whittle (1957) in the case of deterministic growth and was generalized to the case of i.i.d. Gaussian growth and non-Gaussian, Markovian growth by Reed (2001) and Beare and Toda (2017), respectively. A big advantage of using a perpetual youth model is that heterogeneous-agent models with homogeneous problems (scale invariant problems such as optimal consumption-portfolio problems with homothetic preferences) still admit a stationary distribution, unlike models with infinitely lived agents.[8] It is possible, however, to replace the birth/death assumption by others, though at the expense of losing tractability. For example, one may assume that agents face a borrowing constraint. In that case it is often the case that the optimal consumption rule is still asymptotically linear and one obtains a Pareto upper tail as in Acemoglu and Cao (2015) and Benhabib et al. (2015). Since the goal of this paper is to provide a clean analysis with general Markovian dynamics, I maintain the perpetual youth assumption.

Because agents die with probability p, if an agent has discount factor $\beta_s > 0$ in state s, the "effective" discount factor is $\tilde{\beta}_s = \beta_s(1-p)$. Letting $R > 0$ be the equilibrium gross risk-free rate, under perfect annuity markets agents face an effective risk-free rate $\tilde{R} = \frac{R}{1-p}$. Since by Assumption 1 the transition probability matrix P is irreducible, by the Perron-Frobenius theorem P' has a unique positive eigenvector π such that $\sum_{s=1}^{S} \pi_s = 1$ corresponding to the Perron root 1, which is the stationary distribution of P. I assume that the initial states of newborn agents are drawn from π, so at any point in time the fraction π_s of agents is in state s.

Since there is no income risk (only mortality risk, which is covered by annuities), agents can just sell off their future endowments. Hence the initial wealth of a typical agent is

$$w_0 = y \sum_{t=0}^{\infty} \tilde{R}^{-t} = \frac{\tilde{R}}{\tilde{R}-1} y \tag{3.1}$$

if $\tilde{R} > 1$ (and infinite if $\tilde{R} \leq 1$). An alternative way to see this is as follows. Letting a_t be the savings (risk-free asset holdings) at the end of period t, the

[8] Recent applications of the perpetual youth model in the context of power law distributions can be found in Toda (2014), Toda and Walsh (2015, 2017a), Arkolakis (2016), Benhabib et al. (2016), Gabaix et al. (2016), Nirei and Aoki (2016), Aoki and Nirei (2017), Cao and Luo (2017), and Kasa and Lei (2017).

budget constraint becomes

$$c_t + a_t = y + \tilde{R}a_{t-1}.$$

Adding the present discounted value of future income $\sum_{t=1}^{\infty} \tilde{R}^{-t}y = \frac{1}{\tilde{R}-1}y$ to both sides, defining $w_t = c_t + a_t$, and setting $a_{-1} = 0$, we obtain $w_{t+1} = \tilde{R}(w_t - c_t)$, which is the same as the budget constraint (2.1) with initial wealth (3.1).

With this definition of wealth, each agent solves an optimal consumption-savings problem studied in Proposition 2.1. Now we can formally define the stationary equilibrium.

Definition 3.1 (Stationary equilibrium)**.** A *stationary equilibrium* consists of an effective gross risk-free rate $\tilde{R} > 1$, value functions $\{V_s(\cdot)\}_{s=1}^S$, consumption rule $\{c_s(\cdot)\}_{s=1}^S$, and the distribution of wealth and states $\Gamma(w, s)$ such that

1. (Agent optimization) Given $\tilde{R}$, $V_s(\cdot)$ satisfies the Bellman equation (2.2) and $c_s(w)$ is the argmax,

2. (Market clearing) The risk-free asset market clears, so

$$\int w \, \mathrm{d}\Gamma(w, s) = w_0, \tag{3.2}$$

   where w_0 is the initial wealth (3.1), and

3. (Stationarity) $\Gamma(w, s)$ is the stationary distribution of the law of motion

$$(w, s) \mapsto \begin{cases} (G_s w, s'), & \text{(with probability } (1-p)p_{ss'}) \\ (w_0, s'), & \text{(with probability } pp_{ss'}) \end{cases}$$

   where $G_s = \tilde{R}(1 - a_s^{-1/\gamma})$ is the gross growth rate of wealth and $a = (a_s)$ is as in Proposition 2.1 with β_s, R replaced with $\tilde{\beta}_s, \tilde{R}$.

The intuition for the market clearing condition (3.2) is as follows. In a stationary equilibrium, the aggregate wealth must be constant. Since in this model there is no production and goods are perishable, the aggregate wealth must coincide with initial wealth, which is (3.2).

In order to prove the existence of equilibrium, I maintain the following assumption.

Assumption 2 (Random discounting)**.** *Let $\beta_s > 0$ be the discount factor in state s, $\tilde{\beta}_s = \beta_s(1-p)$, and $B = \operatorname{diag}(\tilde{\beta}_1, \dots, \tilde{\beta}_S)$. Then $\rho(BP) < 1$. Furthermore, $\{\beta_s\}_{s=1}^S$ take distinct values.*

The assumption that $\{\beta_s\}_{s=1}^S$ are distinct is without loss of generality since we can merge states with identical discount factors. The assumption $\rho(BP) < 1$ ensures that the equilibrium risk-free rate $\tilde{R}$ exceeds 1, which makes the initial wealth (3.1) finite. This condition holds, for example, if $\tilde{\beta}_s < 1$ for all s. To see this, if $\tilde{\beta}_s \le \bar{\beta} < 1$ for all s, then

$$\rho(BP) \le \rho(\bar{\beta}IP) = \bar{\beta}\rho(P) = \bar{\beta} < 1.\text{[9]}$$

[9] In general, if $0 \le A \le B$, then $\rho(A) \le \rho(B)$. See Horn and Johnson (1985, Theorem 8.4.5).

However, it may be possible that agents are very patient in some state ($\tilde{\beta}_s > 1$) as long as they are impatient enough in other states so that $\rho(BP) < 1$.

Before proving the existence of equilibrium, let us rewrite the equilibrium condition (3.2) in a more convenient form. Let W_s be the aggregate wealth held by agents in state s. Since the initial state is independently drawn from the stationary distribution $\pi = (\pi_1, \dots, \pi_S)'$ of the transition probability matrix P, by accounting we must have

$$W_{s'} = p\pi_{s'} w_0 + (1-p) \sum_{s=1}^{S} p_{ss'} G_s W_s.$$

Letting $G = (G_1, \dots, G_S)'$ and $W = (W_1, \dots, W_S)'$, in matrix for this equation becomes

$$\begin{aligned} W &= pw_0 \pi + (1-p) P' \operatorname{diag}(G) W \\ \iff W &= pw_0 \left(\sum_{n=0}^{\infty} M^n \right) \pi = pw_0 (I - M)^{-1} \pi \end{aligned}$$

provided that the spectral radius of the (nonnegative and irreducible) matrix $M = (1-p)P' \operatorname{diag} G$ is less than 1. (If the spectral radius is 1 or larger, then each element of W is infinite.) Therefore the market clearing condition (3.2) becomes

$$\sum_{s=1}^{S} W_s = w_0 \iff f(\tilde{R}) = p1'(I - (1-p)P' \operatorname{diag} G)^{-1} \pi - 1 = 0, \tag{3.3}$$

where $1 = (1, \dots, 1)'$.

The following theorem establishes the existence of equilibrium (and uniqueness if $\gamma < 1$).

Theorem 3.2. *Suppose that Assumptions 1 and 2 hold. Then there exists a stationary equilibrium. If $\gamma < 1$, the equilibrium is unique.*

Proof. For notational simplicity let us write β_s, R instead of $\tilde{\beta}_s, \tilde{R}$.

Step 1. Let

$$\Omega = \left\{ R > 0 \,\middle|\, R^{1-\gamma} \rho(BP) < 1 \right\} = \begin{cases} (0, \rho(BP)^{\frac{1}{\gamma-1}}), & (\gamma < 1) \\ (\rho(BP)^{\frac{1}{\gamma-1}}, \infty) & (\gamma > 1) \end{cases}$$

be the set of gross risk-free rates for which the fixed point $a(R) = (a_s(R))_{s=1}^S$ in Proposition 2.1 exists. Then each $a_s(R)$ is continuous on Ω.

For notational simplicity let us write a instead of $a(R)$. By the proof of Proposition 2.1, $a = (a_1, \dots, a_S)$ can be obtained by iterating $a \mapsto \phi(Ka)$. Now

$$\begin{aligned} \phi(Ka)_s &= \left(1 + \left(\beta_s R^{1-\gamma} \sum_{s=1}^{S} p_{ss'} a_{s'} \right)^{\gamma} \right)^{1/\gamma} \\ &= \left(1 + \left(\beta_s \mathrm{e}^{(1-\gamma)r} \sum_{s=1}^{S} p_{ss'} a_{s'} \right)^{\gamma} \right)^{1/\gamma}, \end{aligned}$$

where $r = \log R$. By Lemma A.2, the class of all log-convex functions is closed under addition, multiplication, and raising to any positive power. Since the exponential function is log-convex, if each a_s is a log-convex function of r, so is $\phi(Ka)_s$. Since convexity is preserved by taking pointwise limits, and by the proof of Proposition 2.1 we can compute the fixed point by iterating from any point (in particular, $a = (1, \dots, 1)$, whose elements are log-convex), it follows that each element a_s of the fixed point a is also log-convex in $r = \log R$. Since a convex function is continuous in the interior of its domain and Ω is an open interval, each $a_s(R)$ is continuous on Ω.

Step 2. A stationary equilibrium exists.

By Assumption 2 we have $\rho(BP) < 1$. Since $R^{1-\gamma}\rho(BP) < 1$ at $R = 1$, we have $1 \in \Omega$. Noting that $G_s = R(1 - a_s^{-1/\gamma}) < R$ because $a_s > 0$, we have $M = (1-p)P' \operatorname{diag} G \le (1-p)RP'$ element-wise with some strict inequalities. By Footnote 9, we get $\rho(M) \le (1-p)R$ and $\rho(M) \le 1 - p < 1$ at $R = 1$. Since π is an eigenvector of P' corresponding to the eigenvalue 1, we obtain

$$\begin{aligned} f(R) &= p1'(I-M)^{-1}\pi - 1 = p1'\left(\sum_{n=0}^{\infty} M^n\right)\pi - 1 \\ &< p1'\left(\sum_{n=0}^{\infty}(1-p)^n R^n (P')^n\right)\pi - 1 = p1'\left(\sum_{n=0}^{\infty}(1-p)^n R^n\right)\pi - 1 \\ &= \frac{p}{1-(1-p)R} - 1 = \frac{(1-p)(R-1)}{1-(1-p)R}. \end{aligned} \tag{3.4}$$

Substituting $R = 1$, we get $f(1) < 0$. Since by the previous step each $a_s(R)$ is continuous, so is f on its domain. Therefore to show the existence of an equilibrium, it suffices to show that $f(R) > 0$ for some $R > 1$. Let

$$\begin{aligned} \Omega_1 &= \{R \ge 1 \mid R \in \Omega\} = \begin{cases} [1, \rho(BP)^{\frac{1}{\gamma-1}}), & (\gamma < 1) \\ [1, \infty), & (\gamma > 1) \end{cases} \\ \Omega_2 &= \{R \in \Omega_1 \mid (1-p)\rho(P' \operatorname{diag} G) < 1\} \end{aligned}$$

be the sets of effective gross risk-free rate such that the initial wealth and the optimal consumption rules are well defined, and in addition the stationary wealth distribution is finite, respectively. Clearly $\Omega_2 \subset \Omega_1$. By continuity, Ω_2 is an open subset of $[1, \infty)$ that contains 1 by the previous discussion. Let Ω_3 be the connected component of Ω_2 that contains 1. There are several cases to consider.

Case 1: $\boldsymbol{\Omega_3 \subsetneq \Omega_1}$. In this case let $\bar{R} = \sup \Omega_3$. By continuity, it must be $(1-p)\rho(P' \operatorname{diag} G) = 1$ at $R = \bar{R}$. Since $(I - M)^{-1} = \sum_{n=0}^{\infty} M^n$ for $M = (1-p)P' \operatorname{diag} G$, by (3.3) we obtain $f(R) \to \infty$ as $R \uparrow \bar{R}$.

In the remaining cases, assume that $\Omega_3 = \Omega_1$.

Case 2: $\boldsymbol{\gamma > 1}$. In this case by assumption $\Omega_3 = \Omega_1 = [1, \infty)$ so we can make R arbitrarily large. Let $\underline{a} = \min_s a_s$ and $\underline{s}$ be the corresponding state. Then $G_s \ge \underline{G} = R(1 - \underline{a}^{-1/\gamma})$ for all s. Using (2.4), for any $R > 1$ we have

$$\underline{a}^{1/\gamma} = 1 + (\beta_{\underline{s}} R^{1-\gamma} \operatorname{E}[a_{s'} \mid \underline{s}])^{1/\gamma} \ge 1 + (\beta_{\underline{s}} R^{1-\gamma} \underline{a})^{1/\gamma}.$$

Multiplying both sides by $R\underline{a}^{-1/\gamma} > 0$ and rearranging terms, we obtain

$$\underline{G} = R(1 - \underline{a}^{-1/\gamma}) \geq (\beta_{\underline{s}} R)^{1/\gamma}.$$

Hence $M = (1-p)P' \operatorname{diag} G \geq (1-p)(\beta_{\underline{s}} R)^{1/\gamma} P'$, so $\rho(M) \geq (1-p)(\beta_{\underline{s}} R)^{1/\gamma} \to \infty$ as $R \to \infty$, which implies that the assumption $\Omega_3 = \Omega_1$ is never satisfied because by definition $\rho(M) < 1$ for $R \in \Omega_3$. Hence this case never occurs.

Case 3: $\boldsymbol{\gamma < 1}$. Since $\rho(BP) < 1$, we have $\bar{R} := \rho(BP)^{\frac{1}{\gamma - 1}} > 1$. First let us show that $a_s(R) \to \infty$ for all s as $R \to \bar{R}$.

Since $\gamma < 1$, $Ta = \phi(Ka) = \phi(R^{1-\gamma} BPa)$ is increasing in R. Therefore the fixed point $a(R)$ is also increasing in R, so $\bar{a} = \lim_{R \uparrow \bar{R}} a(R)$ exists in $[0, \infty]^S$. Since ϕ is continuous and $a(R) = \phi(R^{1-\gamma} BPa(R))$ for all $R \in (0, \bar{R})$, letting $R \uparrow \bar{R}$ we obtain $\bar{a} = \phi(\bar{R}^{1-\gamma} BP\bar{a})$. Therefore $\bar{a}$ is a fixed point of $\bar{T} : a \mapsto \phi(\bar{R}^{1-\gamma} BPa)$. By the definition of $\bar{R}$, we have $\bar{R}^{1-\gamma} \rho(BP) = 1$. Since condition (a) of Lemma A.1 is violated for $K = \bar{R}^{1-\gamma} BP$, the vector $\bar{a}$ cannot be finite. Therefore $\bar{a}_{s'} = \infty$ for at least one s'. Since $\phi(t) \geq t$, we have $\bar{a} = \bar{T}^n \bar{a} \geq K^n \bar{a}$ for all n. Since P is irreducible, so is K. Therefore for any s, we can take n such that $(K^n)_{ss'} > 0$, so $\bar{a}_s \geq (K^n)_{ss'} \bar{a}_{s'} = \infty$. Since $\bar{a}_s = \infty$ for all s, it follows that $a_s(R) \to \bar{a}_s = \infty$ as $R \to \bar{R}$.

Since $G_s = R(1 - a_s^{-1/\gamma})$ and $a_s \to \infty$ as $R \to \bar{R}$, we have $G_s \to \bar{R}$. Therefore $M = (1-p)P' \operatorname{diag} G \to (1-p)RP'$, so by the same derivation as in (3.4), we obtain

$$f(\bar{R}) = \begin{cases} \frac{(1-p)(\bar{R}-1)}{1-(1-p)\bar{R}}, & ((1-p)\bar{R} < 1) \\ \infty. & ((1-p)\bar{R} \geq 1) \end{cases}$$

In the former case, $f(\bar{R}) > 0$ because $0 < p < 1$ and $\bar{R} > 1$. In the latter case, $f(\bar{R}) > 0$ trivially.

Step 3. The equilibrium is unique if $\gamma < 1$.

We have already seen that if $0 < \gamma < 1$, then each element $a_s(R)$ of the fixed point $a(R)$ is increasing in R. Since $1 - a_s^{-1/\gamma}$ is increasing, so is $G_s = R(1 - a_s^{-1/\gamma})$ for each s. Since the nonconstant terms of $f(R)$ in (3.3) are positive multiple of the elements of $(P' \operatorname{diag} G)^n$ (which are all increasing in R and some of them strictly so because P' is irreducible), it follows that $f(R)$ is strictly increasing. Therefore $f(R) = 0$ has at most one solution and the equilibrium is unique. ☐

Remark. It is well-known that $\gamma < 1$ (elasticity of intertemporal substitution larger than 1) is a sufficient condition for equilibrium uniqueness in a variety of economies.[10] With $\gamma > 1$ it is easy to construct examples of multiple equilibria.[11]

[10] Hens and Loeffler (1995) consider an Arrow-Debreu economy with additive utility; Achdou et al. (2017) consider a continuous-time Huggett economy with no borrowing.

[11] See Kubler and Schmedders (2010) and Toda and Walsh (2017b) for examples in Edgeworth box economies. Although I have not tried to construct an example of multiple equilibria with $\gamma > 1$, it seems possible to do so by using the approach in Toda (2017b).

4 Wealth distribution

4.1 Theoretical results

What does the stationary wealth distribution look like? Since agents are born and die with probability p, the stationary age distribution is geometric with mean $1/p$. While an agent is alive, his wealth follows a random growth process. In fact, substituting the optimal consumption rule $c_s(w) = a_s^{-1/\gamma} w$ into the budget constraint, individual wealth follows $w' = G_s w$, where $G_s = \tilde{R}(1-a_s^{-1/\gamma})$ is the growth rate of wealth in state s. Thus the stationary wealth distribution will be the same as that of a random growth process $w' = G_s w$ initialized at $w_0 = \frac{\tilde{R}}{\tilde{R}-1} y$ and evaluated at a geometrically distributed time with mean $1/p$.

In general, consider the random growth process for wealth

$$w_{i,t+1} = G_{i,t+1} w_{it},$$

where w_{it} is the wealth of agent i at time t and $G_{i,t+1}$ is the gross growth rate of wealth between time t and $t+1$. Assuming that agents are born and die with constant probability $p > 0$ and the distribution of the growth rate $G_{i,t+1}$ conditional on time t information depends only on the current state $s = s_{it}$, Beare and Toda (2017) show that the stationary cross-sectional wealth distribution has Pareto tails and characterize the tail exponents. Let

$$D(z) = \operatorname{diag}(\mathrm{E}\left[\mathrm{e}^{z \log G} \,\middle|\, s = 1\right], \dots, \mathrm{E}\left[\mathrm{e}^{z \log G} \,\middle|\, s = S\right])$$

be the diagonal matrix of conditional moment generating functions of log growth rates. Then the upper and lower tail Pareto exponents are determined such that

$$\rho(PD(\alpha_1)) = \rho(PD(-\alpha_2)) = \frac{1}{1-p}$$

whenever such $-\alpha_1 < 0 < \alpha_1$ exist in the interior of the domain of $\rho(PD(z))$. In our case, the growth rates are deterministic given the current state, namely $G_s = \tilde{R}(1 - a_s^{-1/\gamma})$. Therefore the upper and lower tail Pareto exponents $\alpha_1, \alpha_2 > 0$ satisfy

$$\rho(P \operatorname{diag} G^{(\alpha_1)}) = \rho(P \operatorname{diag} G^{(-\alpha_2)}) = \frac{1}{1-p} \tag{4.1}$$

whenever such $\alpha_1, \alpha_2 > 0$ exist, where $G = (G_1, \dots, G_S)'$ and $G^{(\alpha)} = (G_1^\alpha, \dots, G_S^\alpha)'$ is the vector of element-wise powers. The following theorem, which is the main result of this paper, shows that such $\alpha_1, \alpha_2 > 0$ exists under weak conditions.

Theorem 4.1. *Suppose that $S \geq 2$, Assumptions 1 and 2 hold, and the transition probability matrix $P = (p_{ss'})$ has positive diagonal elements, so $p_{ss} > 0$ for all s. Then there exist unique $\alpha_1, \alpha_2 > 0$ that satisfy (4.1). Consequently, the stationary wealth distribution has Pareto upper and lower tails with exponents α_1, α_2.*

Proof. Uniqueness is proved in Beare and Toda (2017). By the market clearing condition (3.3), we obtain

$$0 = f(\tilde{R}) = p1'(I - M)^{-1}\pi - 1 = p1' \left(\sum_{n=0}^{\infty} M^n \right) \pi - 1,$$

where $M = (1-p)P' \operatorname{diag} G$ and $G_s = \tilde{R}(1 - a_s^{-1/\gamma})$.

Suppose for the moment that $G_s > 1$ for some state s and let $M_s(z)$ be the $S \times S$ matrix whose (s,s)-th element is $p_{ss} G_s^z > 0$ and all other elements are 0. By Footnote 9, we obtain

$$\rho(P \operatorname{diag} G^{(0)}) = \rho(P) = 1 < \frac{1}{1-p},$$
$$\rho(P \operatorname{diag} G^{(z)}) \ge \rho(M_s(z)) = p_{ss} G_s^z \to \infty$$

as $z \to \infty$ because $G_s > 1$. Since the spectral radius is continuous, there exists $\alpha_1 > 0$ such that $\rho(P \operatorname{diag} G^{(\alpha_1)}) = \frac{1}{1-p}$. By the same argument, if $G_s < 1$ for some s, then there exists $\alpha_2 > 0$ such that $\rho(P \operatorname{diag} G^{(-\alpha_2)}) = \frac{1}{1-p}$. Hence there exist $\alpha_1, \alpha_2 > 0$ that satisfy (4.1).

Therefore to complete the proof it remains to show that

$$\min G_s < 1 < \max G_s.$$

First let us show $\min G_s < \max G_s$. If not, then G_s is constant, and so is a_s because $G_s = \tilde{R}(1 - a_s^{-1/\gamma})$. Letting $a_s = a$, by (2.4) we have

$$a = \left(1 + (\tilde{\beta}_s \tilde{R}^{1-\gamma} a)^{1/\gamma}\right)^{\gamma},$$

which is a contradiction because $S \ge 2$ and $\{\beta_s\}_{s=1}^S$ are distinct by Assumption 2. Therefore $\min G_s < \max G_s$.

If $G_s \ge 1$ for all s, then there exists s such that $G_s > 1$. Since $\operatorname{diag} G \ge I$ with some strict inequality for diagonal entries, it follows that

$$\begin{aligned} 0 = f(\tilde{R}) &> p1' \left(\sum_{n=0}^{\infty} (1-p)^n (P')^n \right) \pi - 1 \\ &= p1' \left(\sum_{n=0}^{\infty} (1-p)^n \right) \pi - 1 = \frac{p}{1-(1-p)} - 1 = 0, \end{aligned}$$

which is a contradiction. Therefore $\min G_s < 1$. If $G_s \le 1$ for all s, by the same argument we obtain the contradiction $0 = f(\tilde{R}) < 0$. Therefore $\max G_s > 1$. $\square$

Recently Piketty (2014) and Piketty and Zucman (2015, Section 15.5.4) have argued that the condition $r > g$ (the rate of return on wealth exceeding the growth rate of the economy) exacerbate income and wealth inequality. However, this argument is rather informal and seems to ignore general equilibrium effects. Using my model we can address this issue formally. The proposition below shows that if the relative risk aversion is bounded above by 1, then in a small open economy a higher interest rate increases inequality.

Proposition 4.2. *Let everything be as in Theorem 4.1. Suppose that $\gamma < 1$ and consider a small open economy, so the gross risk-free rate $R > 0$ is given. Then $\partial \alpha_1 / \partial R < 0$, so higher interest rate implies lower Pareto exponent (more inequality).*

Proof. Let

$$F(z, R) = \rho(P \operatorname{diag} G^{(z)}) - \frac{1}{1-p},$$

where $G = (G_1, \dots, G_S)$ and $G_s = \tilde{R}(1 - a_s^{-1/\gamma})$. By (4.1), the upper tail Pareto exponent satisfies $F(\alpha_1, R) = 0$. Beare and Toda (2017) show that $\partial F/\partial z(\alpha_1, R) > 0$. By the proof of Theorem 3.2, $G_s = \tilde{R}(1-a_s^{-1/\gamma})$ is increasing in R, so by Footnote 9 we have $\partial F/\partial R > 0$. Therefore by the implicit function theorem, we obtain $\partial\alpha/\partial R = -(\partial F/\partial R)/(\partial F/\partial z) < 0$. □

Because my model is a general equilibrium model, the risk-free rate R is endogenous. Therefore it is not clear how the interest rate R and the Pareto exponent α_1 change when we change exogenous model parameters. Below, I explore this issue using a numerical example.

4.2 Numerical example

As a numerical example, I consider parameter values and a discount factor process similar to Hubmer et al. (2016). The relative risk aversion coefficient is $\gamma = 1.5$. The discount factor follows the AR(1) process

$$\log \beta_t = (1-\rho)\mu + \rho \log \beta_{t-1} + \sigma\epsilon_t, \quad \epsilon_t \sim N(0,1), \tag{4.2}$$

where $\mu = -0.0836$, $\rho = 0.992$, and $\sigma = 0.0021$.[12] To convert this process into a finite-state Markov chain, I use the discretization method proposed by Farmer and Toda (2017)—which is more accurate and generally applicable than other discretization methods—with a 9-point even-spaced grid and treat this Markov chain as the true process. Finally, I set the death probability $p = 0.025$ so that the mean life span (which should be interpreted as the average years in the labor market) is 40. For each guess of effective gross risk-free rate $\tilde{R}$, I compute the market clearing condition (3.3) by solving for the fixed point in (2.4), and then solve the equation $f(\tilde{R}) = 0$ numerically. Finally, I compute the upper and lower tail Pareto exponents by solving (4.1) for α_1, α_2.

With the above parameter values, in equilibrium the effective gross risk-free rate is $\tilde{R} = 1.1095$, the gross risk-free rate is $R = 1.0818$, and the upper and lower tail Pareto exponents are $\alpha_1 = 2.2183$ and $\alpha_2 = 1.6058$, respectively. Figure 1 shows the log tail probabilities ($\log \Pr(w_{it} > w)$ for $w > w_0$ and $\log \Pr(w_{it} < w)$ for $w < w_0$) from a simulation with 100,000 agents and $T = 20/p = 800$ periods. Consistent with the double power law, the log tail probabilities show a tent-shaped pattern. The Pareto exponents estimated by maximum likelihood using 10% of the extreme tail observations are $\widehat{\alpha}_1 = 2.1543$ and $\widehat{\alpha}_2 = 1.6531$, which are close to the theoretical values.

Next, I conduct a comparative statics exercise by changing model parameters. The main questions here are whether there is a positive relationship between the interest rate and inequality, and which parameter is important in determining the Pareto exponent. I consider four cases: (i) change the relative risk aversion in the range $\gamma \in [0.1, 10]$, (ii) change the average discount factor in the range $\mu_\beta \in [0.86, 0.98]$, (iii) change the persistence of the AR(1) process

[12] Hubmer et al. (2016) specify the AR(1) process in levels (β_t instead of $\log \beta_t$), which has the undesirable property that β_t eventually becomes negative with probability 1. When the log discount factor follows the AR(1) process (4.2), using the property of the lognormal distribution, the unconditional mean and variance of β_t becomes $\mu_\beta = \mathrm{e}^{\mu + \frac{\sigma^2}{2(1-\rho)}}$ and $\sigma_\beta^2 = (\mathrm{e}^{\frac{\sigma^2}{1-\rho}} - 1)\mathrm{e}^{2\mu + \frac{\sigma^2}{1-\rho}}$, respectively. The parameter values μ, σ are chosen so as to match $\mu_\beta = 0.92$ and $\sigma_\beta = 0.0019$ in Hubmer et al. (2016).

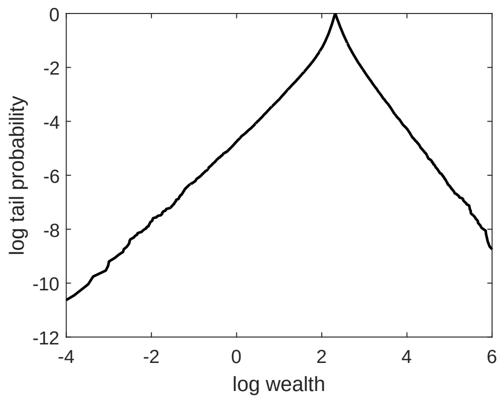


Figure 1: Log tail probabilities.

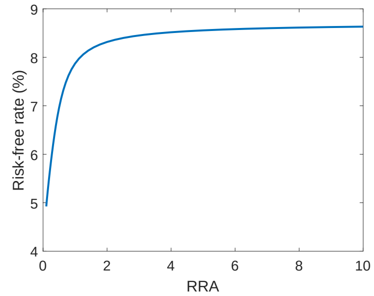
in the range $\rho \in [0.5, 0.995]$, and (iv) change the birth/death probability in the range $p \in [0.01, 0.1]$ (average life span in the range $[10, 100]$). In each case I fix all other parameters at the baseline values. Figure 2 shows the results.

The left panels show the effect of changing model parameters (relative risk aversion γ, average discount factor μ_β, persistence ρ, and average life span $1/p$) on the equilibrium gross risk-free rate. The right panels do the same for the upper tail Pareto exponent α_1. Recalling that smaller Pareto exponent implies more inequality, Piketty (2014)'s claim would be supported if the gross risk-free rate R and Pareto exponent α_1 move in opposite directions when we change parameters. However, this is the case only when we change the average discount factor μ_β. In general there is no clear relationship between the interest rate and inequality due to general equilibrium effects.

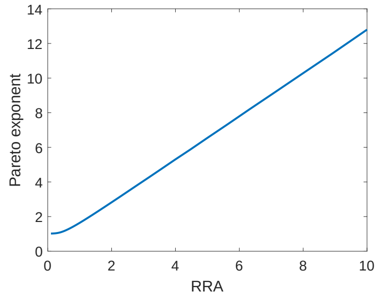
The right panels in Figure 2 also show which parameter is important in determining inequality. The average discount factor μ_β has a modest effect on α_1. Increasing the life span $1/p$ decreases the Pareto exponent but only moderately. The relative risk aversion (reciprocal of elasticity of intertemporal substitution) has a large effect on inequality: when agents are close to risk-neutral, the Pareto exponent α_1 is close to 1 (Zipf's law), and roughly linearly increases with γ. The persistence ρ of the AR(1) process has a huge impact on the Pareto exponent: even when we restrict attention to the highly persistent case, the Pareto exponent ranges from $\alpha_1 = 31$ when $\rho = 0.9$ to $\alpha_1 = 1.5$ when $\rho = 0.995$.

5 Concluding remarks

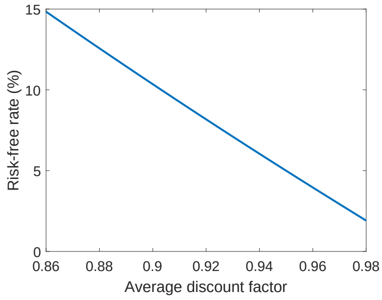
Using a simple dynamic general equilibrium model with random discount factors and lifetime alone, in this paper I proved that the stationary wealth distribution has Pareto tails and provided an analytical characterization. Since my model does not feature any idiosyncratic income risk, as the example in Benhabib et al. (2017) suggests, income risk does not likely matter for determining the top wealth inequality (although it is, of course, important for the wealth inequality in the *middle* of the distribution). Furthermore, the equilibrium Pareto exponent is highly sensitive to the persistence of the discount factor process. These



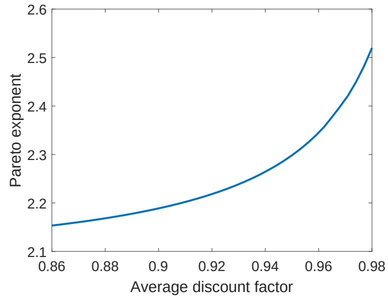
(a) Effect of γ on R.



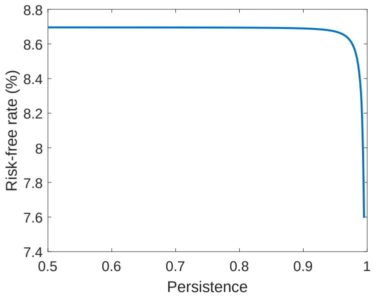
(b) Effect of γ on α_1.



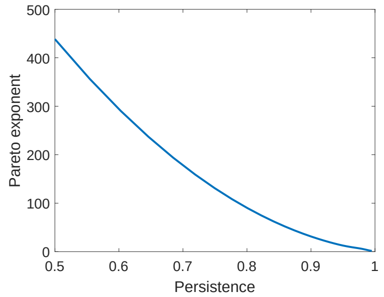
(c) Effect of μ_β on R.



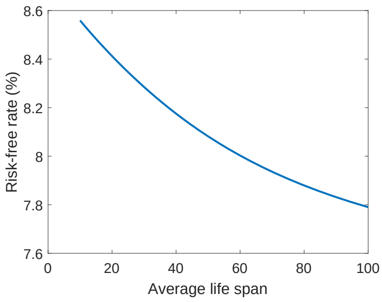
(d) Effect of μ_β on α_1.



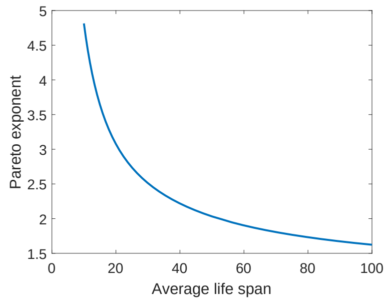
(e) Effect of ρ on R.



(f) Effect of ρ on α_1.



(g) Effect of $1/p$ on R.



(h) Effect of $1/p$ on α_1.

Figure 2: Comparative statics results.

somewhat negative results may force applied researchers to be more clear about the mechanism that generates the amount of inequality they wish to explain. In a general equilibrium model, there is no clear relationship between the return on wealth and the Pareto exponent when we change the model parameters because interest rate and inequality are endogenously determined.

A Mathematical results

Lemma A.1. *Let $\gamma > 0$, K be an $S \times S$ nonnegative irreducible matrix, and $X = \mathbb{R}_+^S$. Define $T : X \to X$ by*

$$(Tx)_s = (1 + ((Kx)_s)^{1/\gamma})^\gamma.$$

Then the following statements are equivalent:

(a) $\rho(K) < 1$.

(b) T has a fixed point in X.

(c) There exists a $g \in X$ such that $\{T^n g\}_{n=1}^\infty$ is convergent in X.

(d) The sequence $\{T^n g\}_{n=1}^\infty$ is convergent in X for all $g \in X$.

(e) T has a unique fixed point g^* in X and $\lim_{n\to\infty} T^n g = g^*$ for all $g \in X$.

Proof. ((a) $\implies$ (e)) Define $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ by $\phi(t) = (1 + t^{1/\gamma})^\gamma$. If $\gamma \le 1$, then

$$0 \le \phi'(t) = \gamma(1 + t^{1/\gamma})^{\gamma-1} \frac{1}{\gamma} t^{1/\gamma-1} = (t^{-1/\gamma} + 1)^{\gamma-1} \le 1,$$

so ϕ is Lipshitz continuous with order 1. Since the conditions of Borovička and Stachurski (2017, Proposition 4.1) are satisfied, the conclusion holds.

If $\gamma > 1$, then

$$\phi''(t) = (\gamma - 1)(t^{-1/\gamma} + 1)^{\gamma-2} \left(-\frac{1}{\gamma} t^{-1/\gamma-1}\right) < 0,$$

so ϕ is increasing and concave. Since the conditions of Borovička and Stachurski (2017, Proposition 4.2) are satisfied, the conclusion holds.

((e) $\implies$ (d) $\implies$ (c)) Trivial.

((c) $\implies$ (b)) Let $x^{(n)} = T^n g$. By definition, we have $x^{(n+1)} = \phi(Kx^{(n)})$, where ϕ is applied element-wise. Letting $n \to \infty$, since by assumption $x^{(n)} \to x^*$ for some $x^* \in X$ and ϕ is continuous, we obtain $x^* = \phi(Kx^*) = Tx^*$. Therefore x^* is a fixed point of T.

((b) $\implies$ (a)) Let $g^* \in X$ be a fixed point of T. Since $\phi(t) > 0$, clearly $g^* \gg 0$. Since $\phi(t) > t$ for $t > 0$, we have $Tx \gg Kx$ for all $x \gg 0$. In particular, $g^* = Tg^* \gg Kg^*$. By the Perron-Frobenius theorem, there exists $v \gg 0$ such that $v'K = \rho(K)v'$. Left-multiplying v' to the above inequality, we obtain $v'g^* > \rho(K)v'g^*$, so $\rho(K) < 1$ because $v'g^* > 0$. ☐

Lemma A.2 (Kingman, 1961)**.** *The class of all log-convex functions is closed under addition, multiplication, and raising to any positive power.*

Proof. Since the sum and positive multiple of convex functions are convex, it suffices to show that the sum of log-convex functions is log-convex. Let f, g be log-convex. Then for any x_1, x_2 and $t \in (0,1)$, we have

$$\log f(x) \le (1-t)\log f(x_1) + t\log f(x_2) \implies f(x) \le f(x_1)^{1-t} f(x_2)^t,$$

where $x = (1-t)x_1 + tx_2$. The same inequality holds for g. Hence by Hölder's inequality, we obtain

$$\begin{aligned} f(x) + g(x) &\le f(x_1)^{1-t} f(x_2)^t + g(x_1)^{1-t} g(x_2)^t \\ &\le (f(x_1) + g(x_1))^{1-t} (f(x_2) + g(x_2))^t, \end{aligned}$$

so $f + g$ is log-convex. □

B Log utility

In this appendix I solve the model with log utility, so $u(c) = \log c$. The following proposition shows that the optimal consumption rule does not depend on the interest rate.

Proposition B.1. *Suppose that Assumptions 1 and 2 hold. Then the optimal consumption rule is $c = w/b_s$, where $(b_1, \dots, b_S)' = b = (I - BP)^{-1}1$ and $1 = (1, \dots, 1)'$.*

Proof. For notational simplicity let us write β_s, R instead of $\tilde{\beta}_s, \tilde{R}$. By the structure of the problem, we can guess that the value function is of the form $V_s(w) = a_s + b_s \log w$. Substituting this guess into the Bellman equation (2.2), we get

$$a_s + b_s \log w = \max_c \{\log c + \beta_s \operatorname{E}[a_{s'} + b_{s'} \log R(w - c) \mid s]\}. \tag{B.1}$$

The first-order condition with respect to c is

$$\frac{1}{c} = \beta_s \frac{\operatorname{E}[b_{s'} \mid s]}{w - c} \iff c = \frac{w}{1 + \beta_s \operatorname{E}[b_{s'} \mid s]}.$$

Substituting this c into (B.1) and comparing the coefficient of $\log w$, we obtain

$$b_s = 1 + \beta_s \operatorname{E}[b_{s'} \mid s], \tag{B.2}$$

so the optimal consumption rule is $c = w/b_s$. Since $\rho(BP) < 1$, $(I - BP)^{-1}$ exists and equals $\sum_{n=0}^{\infty} (BP)^n$, which is a positive matrix. Letting $b = (b_1, \dots, b_S)'$ and expressing (B.2) in vector form, we obtain

$$b = 1 + BPb \iff b = (I - BP)^{-1}1.$$

Substituting this into (B.1) and expressing in matrix form, we obtain $a = BPa + d \iff a = (I - BP)^{-1}d$ for some vector d. To show the transversality condition, by the same argument as in the proof of Proposition 2.1, we obtain

$$\operatorname{E}_0[\beta(s^{n-1}) V_{s_n}(w_n)] \le (a_{s_{n-1}} + b_{s_{n-1}} \log R^n w_0) \sum_{s^n} \prod_{t=0}^{n-1} (\beta_{s_t} p_{s_t s_{t+1}}).$$

Since $\log R^n w_0$ grows linearly in n and

$$\sum_{s^n} \prod_{t=0}^{n-1} (\beta_{s_t} p_{s_t s_{t+1}}) \leq 1'(BP)^n 1$$

decays to 0 exponentially in n since $\rho(BP) < 1$, the transversality condition holds. □

As in the CRRA case, a stationary equilibrium exists and the stationary wealth distribution has Pareto tails if agents are born and die with probability $p > 0$.

Theorem B.2. *Suppose that Assumptions 1 and 2 hold. Then the log utility economy has a unique stationary equilibrium. If in addition $p_{ss} > 0$ for all s, then the conclusion of Theorem 4.1 holds if we define $G_s = \tilde{R}(1 - 1/b_s)$.*

Proof. As in the proof of Theorem 3.2, the equilibrium condition is (3.3), where $G = (G_1, \dots, G_S)'$ and $G_s = \tilde{R}(1 - 1/b_s)$. By the proof of Proposition B.1, clearly b_s does not depend on R and $b_s > 1$ for all s. Therefore $f(\tilde{R})$ is continuous and strictly increasing in $\tilde{R}$, $f(1) < 0$ by (3.4), and $f(\tilde{R}) > 0$ for large enough $\tilde{R}$ as in the proof of Theorem 3.2, so there exists a unique equilibrium. The proof that the stationary wealth distribution has Pareto tails is identical to Theorem 4.1. □

References

Daron Acemoglu and Dan Cao. Innovation by entrants and incumbents. *Journal of Economic Theory*, 157:255–294, May 2015. doi:10.1016/j.jet.2015.01.001.

Yves Achdou, Jiequn Han, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll. Income and wealth distribution in macroeconomics: A continuous-time approach. NBER Working Paper 23732, 2017. URL `http://www.nber.org/papers/w23732`.

S. Rao Aiyagari. Uninsured idiosyncratic risk and aggregate saving. *Quarterly Journal of Economics*, 109(3):659–684, 1994. doi:10.2307/2118417.

Shuhei Aoki and Makoto Nirei. Zipf's law, Pareto's law, and the evolution of top incomes in the U.S. *American Economic Journal: Macroeconomics*, 9(3): 36–71, July 2017. doi:10.1257/mac.20150051.

Costas Arkolakis. A unified theory of firm selection and growth. *Quarterly Journal of Economics*, 131(1):89–155, 2016. doi:10.1093/qje/qjv039.

Rüdiger Bachmann and Jinhui H. Bai. Politico-economic inequality and the comovement of government purchases. *Review of Economic Dynamics*, 16(4): 565–580, October 2013. doi:10.1016/j.red.2012.09.008.

Abhijit V. Banerjee and Andrew F. Newman. Risk-bearing and the theory of income distribution. *Review of Economic Studies*, 58(2):211–235, April 1991. doi:10.2307/2297965.

Brendan K. Beare and Alexis Akira Toda. Geometrically stopped Markovian random growth processes and Pareto tails. 2017.

Jess Benhabib and Alberto Bisin. Skewed wealth distributions: Theory and empirics. *Journal of Economic Literature*, 2017. Forthcoming.

Jess Benhabib, Alberto Bisin, and Shenghao Zhu. The distribution of wealth and fiscal policy in economies with finitely lived agents. *Econometrica*, 79(1): 123–157, January 2011. doi:10.3982/ECTA8416.

Jess Benhabib, Alberto Bisin, and Shenghao Zhu. The wealth distribution in Bewley economies with capital income risk. *Journal of Economic Theory*, 159 (A):489–515, September 2015. doi:10.1016/j.jet.2015.07.013.

Jess Benhabib, Alberto Bisin, and Shenghao Zhu. The distribution of wealth in the Blanchard-Yaari model. *Macroeconomic Dynamics*, 20:466–481, March 2016. doi:10.1017/S1365100514000066.

Jess Benhabib, Alberto Bisin, and Mi Luo. Earnings inequality and other determinants of wealth inequality. *American Economic Review Papers and Proceedings*, 107(5):593–597, May 2017. doi:10.1257/aer.p20171005.

Truman F. Bewley. A difficulty with the optimum quantity of money. *Econometrica*, 51(5):1485–1504, September 1983. doi:10.2307/1912286.

Olivier J. Blanchard. Debt, deficits, and finite horizons. *Journal of Political Economy*, 93(2):223–247, April 1985. doi:10.1086/261297.

Jaroslav Borovička and John Stachurski. Spectral conditions for existence and uniqueness of recursive utilities. 2017. URL `https://arxiv.org/abs/1710.06526`.

Marco Cagetti and Mariacristina De Nardi. Entrepreneurship, frictions, and wealth. *Journal of Political Economy*, 114(5):835–870, October 2006. doi:10.1086/508032.

Laurent-Emmanuel Calvet, John Y. Campbell, and Paolo Sodini. Down or out: Assessing the welfare costs of household investment mistakes. *Journal of Political Economy*, 115(5):707–747, October 2007. doi:10.1086/524204.

Dan Cao and Wenlan Luo. Persistent heterogeneous returns and top end wealth inequality. *Review of Economic Dynamics*, 2017. doi:10.1016/j.red.2017.10.001.

Christopher D. Carroll. Buffer-stock saving and the life cycle/permanent income hypothesis. *Quarterly Journal of Economics*, 112(1):1–55, February 1997. doi:10.1162/003355397555109.

Christopher D. Carroll, Jiri Slacalek, Kiichi Tokuoka, and Matthew N. White. The distribution of wealth and the marginal propensity to consume. *Quantitative Economics*, 2017. URL `http://qeconomics.org/ojs/forth/694/694-2.pdf`. Forthcoming.

David G. Champernowne. A model of income distribution. *Economic Journal*, 63(250):318–351, June 1953. doi:10.2307/2227127.

Douglas W. Diamond and Philip H. Dybvig. Bank runs, deposit insurance, and liquidity. *Journal of Political Economy*, 91(3):401–419, June 1983. doi:10.1086/261155.

Andres Erosa and Tatyana Koreshkova. Progressive taxation in a dynastic model of human capital. *Journal of Monetary Economics*, 54(3):667–685, April 2007. doi:10.1016/j.jmoneco.2006.01.005.

Leland E. Farmer and Alexis Akira Toda. Discretizing nonlinear, non-Gaussian Markov processes with exact conditional moments. *Quantitative Economics*, 8(2):651–683, July 2017. doi:10.3982/QE737.

Xavier Gabaix. Power laws in economics and finance. *Annual Review of Economics*, 1:255–293, 2009. doi:10.1146/annurev.economics.050708.142940.

Xavier Gabaix. Power laws in economics: An introduction. *Journal of Economic Perspectives*, 30(1):185–206, Winter 2016. doi:10.1257/jep.30.1.185.

Xavier Gabaix, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll. The dynamics of inequality. *Econometrica*, 84(6):2071–2111, November 2016. doi:10.3982/ECTA13569.

John Geanakoplos and Kieran James Walsh. Inefficient liquidity provision. *Economic Theory*, 2017. doi:10.1007/s00199-017-1059-7.

Kristian Giesen, Arndt Zimmermann, and Jens Suedekum. The size distribution across all cities—double Pareto lognormal strikes. *Journal of Urban Economics*, 68(2):129–137, September 2010. doi:10.1016/j.jue.2010.03.007.

Lutz Hendricks. Retirement wealth and lifetime earnings. *International Economic Review*, 48(2):421–456, May 2007. doi:10.1111/j.1468-2354.2007.00433.x.

Thorsten Hens and András Loeffler. Gross substitution in financial markets. *Economics Letters*, 49(1):39–43, July 1995. doi:10.1016/0165-1765(94)00647-K.

Roger A. Horn and Charles R. Johnson. *Matrix Analysis*. Cambridge University Press, New York, 1985.

Joachim Hubmer, Per Krusell, and Anthony A. Smith, Jr. The historical evolution of the wealth distribution: A quantitative-theoretic investigation. NBER Working Paper 23011, 2016. URL `http://www.nber.org/papers/w23011`.

Mark Huggett. The risk-free rate in heterogeneous-agent incomplete-insurance economies. *Journal of Economic Dynamics and Control*, 17(5-6):953–969, September-November 1993. doi:10.1016/0165-1889(93)90024-M.

Mark Huggett. Wealth distribution in life-cycle economies. *Journal of Monetary Economics*, 38(3):469–494, December 1996. doi:10.1016/S0304-3932(96)01291-3.

Kenneth Kasa and Xiaowen Lei. Risk, uncertainty, and the dynamics of inequality. *Journal of Monetary Economics*, 2017. Forthcoming.

Harry Kesten. Random difference equations and renewal theory for products of random matrices. *Acta Mathematica*, 131(1):207–248, 1973. doi:10.1007/BF02392040.

John Frank Charles Kingman. A convexity property of positive matrices. *Quarterly Journal of Mathematics*, 12(1):283–284, 1961. doi:10.1093/qmath/12.1.283.

Oren S. Klass, Ofer Biham, Moshe Levy, Ofer Malcai, and Sorin Solomon. The Forbes 400 and the Pareto wealth distribution. *Economics Letters*, 90(2): 290–295, February 2006. doi:10.1016/j.econlet.2005.08.020.

Narayana R. Kocherlakota. On the 'discount' factor in growth economies. *Journal of Monetary Economics*, 25(1):43–47, January 1990. doi:10.1016/0304-3932(90)90043-4.

Tom Krebs. Recursive equilibrium in endogenous growth models with incomplete markets. *Economic Theory*, 29(3):505–523, 2006. doi:10.1016/S0165-1889(03)00062-9.

Dirk Krueger, Kurt Mitman, and Fabrizio Perri. Macroeconomics and household heterogeneity. In John B. Taylor and Harald Uhlig, editors, *Handbook of Macroeconomics*, volume 2, chapter 11, pages 843–921. Elsevier, 2016. doi:10.1016/bs.hesmac.2016.04.003.

Per Krusell and Anthony A. Smith, Jr. Income and wealth heterogeneity, portfolio choice, and equilibrium asset returns. *Macroeconomic Dynamics*, 1(2): 387–422, June 1997. doi:10.1017/S1365100597003052.

Per Krusell and Anthony A. Smith, Jr. Income and wealth heterogeneity in the macroeconomy. *Journal of Political Economy*, 106(5):867–896, October 1998. doi:10.1086/250034.

Per Krusell and Anthony A. Smith, Jr. On the welfare effects of eliminating business cycles. *Review of Economic Dynamics*, 2(1):245–272, January 1999. doi:10.1006/redy.1998.0043.

Per Krusell, Toshihiko Mukoyama, Ayşegül Şahin, and Anthony A. Smith, Jr. Revisiting the welfare effects of eliminating business cycles. *Review of Economic Dynamics*, 12(1):393–404, July 2009. doi:10.1016/j.red.2009.01.002.

Felix Kubler and Karl Schmedders. Tackling multiplicity of equilibria with Gröbner bases. *Operations Research*, 58(4):1037–1050, 2010. doi:10.1287/opre.1100.0819.

Alisdair McKay. Time-varying idiosyncratic risk and aggregate consumption dynamics. *Journal of Monetary Economics*, 88:1–14, June 2017. doi:10.1016/j.jmoneco.2017.05.002.

Alisdair McKay and Ricardo Reis. The role of automatic stabilizers in the U.S. business cycle. *Econometrica*, 84(1):141–1194, January 2016. doi:10.3982/ECTA11574.

Toshihiko Mukoyama and Ayşegül Şahin. Costs of business cycles for unskilled workers. *Journal of Monetary Economics*, 53(8):2179–2193, November 2006. doi:10.1016/j.jmoneco.2005.08.014.

Makoto Nirei and Shuhei Aoki. Pareto distribution of income in neoclassical growth models. *Review of Economic Dynamics*, 20:25–42, April 2016. doi:10.1016/j.red.2015.11.002.

Makoto Nirei and Wataru Souma. A two factor model of income distribution dynamics. *Review of Income and Wealth*, 53(3):440–459, September 2007. doi:10.1111/j.1475-4991.2007.00242.x.

Vilfredo Pareto. *Cours d'Économie Politique*. F. Rouge, Lausanne, 1896.

Thomas Piketty. *Capital in the 21st Century*. Harvard University Press, 2014.

Thomas Piketty and Gabriel Zucman. Wealth and inheritance in the long run. In Anthony B. Atkinson and François Bouguignon, editors, *Handbook of Income Distribution*, volume 2, chapter 15, pages 1303–1368. Elsevier, 2015. doi:10.1016/B978-0-444-59429-7.00016-9.

Vincenzo Quadrini. Entrepreneurship, saving, and social mobility. *Review of Economic Dynamics*, 3(1):1–40, January 2000. doi:10.1006/redy.1999.0077.

William J. Reed. The Pareto, Zipf and other power laws. *Economics Letters*, 74(1):15–19, December 2001. doi:10.1016/S0165-1765(01)00524-9.

William J. Reed. On the rank-size distribution for human settlements. *Journal of Regional Science*, 42(1):1–17, February 2002. doi:10.1111/1467-9787.00247.

William J. Reed and Murray Jorgensen. The double Pareto-lognormal distribution—a new parametric model for size distribution. *Communications in Statistics—Theory and Methods*, 33(8):1733–1753, 2004. doi:10.1081/STA-120037438.

Paul A. Samuelson. Lifetime portfolio selection by dynamic stochastic programming. *Review of Economics and Statistics*, 51(3):239–246, August 1969. doi:10.2307/1926559.

Alexis Akira Toda. The double power law in income distribution: Explanations and evidence. *Journal of Economic Behavior and Organization*, 84(1):364–381, September 2012. doi:10.1016/j.jebo.2012.04.012.

Alexis Akira Toda. Incomplete market dynamics and cross-sectional distributions. *Journal of Economic Theory*, 154:310–348, November 2014. doi:10.1016/j.jet.2014.09.015.

Alexis Akira Toda. Asset prices and efficiency in a Krebs economy. *Review of Economic Dynamics*, 18(4):957–978, October 2015. doi:10.1016/j.red.2014.11.003.

Alexis Akira Toda. A note on the size distribution of consumption: More double Pareto than lognormal. *Macroeconomic Dynamics*, 21(6):1508–1518, September 2017a. doi:10.1017/S1365100515000942.

Alexis Akira Toda. Huggett economies with multiple stationary equilibria. *Journal of Economic Dynamics and Control*, 84:77–90, November 2017b. doi:10.1016/j.jedc.2017.09.002.

Alexis Akira Toda and Kieran Walsh. The double power law in consumption and implications for testing Euler equations. *Journal of Political Economy*, 123(5):1177–1200, October 2015. doi:10.1086/682729.

Alexis Akira Toda and Kieran James Walsh. Fat tails and spurious estimation of consumption-based asset pricing models. *Journal of Applied Econometrics*, 32(6):1156–1177, September/October 2017a. doi:10.1002/jae.2564.

Alexis Akira Toda and Kieran James Walsh. Edgeworth box economies with multiple equilibria. *Economic Theory Bulletin*, 5(1):65–80, April 2017b. doi:10.1007/s40505-016-0102-3.

Philip Vermeulen. How fat is the top tail of the wealth distribution? *Review of Income and Wealth*, 2017. doi:10.1111/roiw.12279.

Herman O. A. Wold and Peter Whittle. A model explaining the Pareto distribution of wealth. *Econometrica*, 25(4):591–595, October 1957. doi:10.2307/1905385.

Menahem E. Yaari. Uncertain lifetime, life insurance, and the theory of the consumer. *Review of Economic Studies*, 32(2):137–150, April 1965. doi:10.2307/2296058.