

The Wealth Distribution in Bewley Models with Investment Risk

Jess Benhabib
New York University

Alberto Bisin
New York University and NBER

Shenghao Zhu
National University of Singapore

First draft: October 2013 - This draft: July 2014*

Abstract

We study the wealth distribution in Bewley economies with idiosyncratic capital income risk. We show analytically that under rather general conditions on the stochastic structure of the economy, a unique ergodic distribution of wealth displays a fat tail; more precisely, a Pareto distribution in the right tail.

Key Words: Wealth distribution; Aiyagari, Bewley, Pareto; fat tails; capital income risk

JEL Numbers: E13, E21, E24

*Thanks to Ben Moll for kindly providing us with very useful detailed comments and to Yi Lu for precious suggestions regarding various computational aspects. We also gratefully acknowledge the exceptionally expert research assistantship of Simon Mongey on the simulations. Shenghao Zhu acknowledges the financial support from National University of Singapore (R-122-000-175-112).

1 Introduction

Bewley economies, as e.g., in Bewley (1977, 1983) and Aiyagari (1994),¹ represent one of the fundamental workhorses of modern macroeconomics, its main tool when moving away from the study of efficient economies with a representative agent to allow e.g., for incomplete markets.² In these economies the evolution of aggregate variables does not generally constitute a sufficient representation of equilibrium, which instead requires the characterization of the dynamics of the distributions across heterogeneous agents.

In Bewley economies each agent faces a stochastic process for labor earnings and solves an infinite horizon consumption-saving problem with incomplete markets. Typically, agents are restricted to save by investing in a risk-free bond and are not allowed to borrow. The postulated process for labor earnings determines the dynamics of the equilibrium distributions for consumption, savings, and wealth.³

Bewley models have been successful in the study of several macroeconomic phenomena of interest. Calibrated versions of this class of models have been used to study welfare costs of inflation (Imrohoroglu, 1992), asset pricing (Mankiw, 1986 and Huggett, 1993), unemployment benefits (Hansen and Imrohoroglu, 1992), fiscal policy (Aiyagari, 1995 and Heathcote, 2005), and labor productivity (Heathcote, Storesletten, and Violante, 2008a, 2008b; Storesletten, Telmer, and Yaron, 2001; and Krueger and Perri, 2008).⁴

On the other hand, Bewley models are hardly able to reproduce the observed distribution of wealth in many countries; see e.g., Aiyagari (1994) and Huggett (1993). More specifically, they cannot reproduce the high inequality and the fat right tail that empirical distributions of wealth tend to display.⁵ This is because at high wealth levels, the incentives for precautionary savings taper off and the tails of the wealth distribution remain thin; see Carroll (1997) and Quadrini (1999) for a discussion of these issues.⁶

¹The *Bewley economy* terminology is rather generally adopted and has been introduced by Ljungqvist and Sargent (2004).

²The assumption of complete markets is generally rejected in the data; see e.g., Attanasio and Davis (1996), Fisher and Johnson (2006) and Jappelli and Pistaferri (2006).

³More recent specifications of the model allow for aggregate risks and an equilibrium determination of labor earnings and interest rates; see Huggett (1993), Aiyagari (1994), Rios-Rull (1995), Krusell and Smith (2006, 2008). and Ljungqvist and Sargent (2004), Ch. 17, for a review of results.

⁴See Heathcote-Storesletten-Violante (2010) for a recent survey of the quantitative implications of Bewley models.

⁵Large top wealth shares in the U.S. since the 60's are documented e.g., by Wolff (1987, 2004) and, more recently, by Kopczuk and Saez (2014) using estate tax return data; Piketty and Zucman (2014) find large and increasing wealth-to-income ratios in the U.S. and Europe in 1970-2010 national balance sheets data. Fat tails for the distributions of wealth are also well documented, for example by Nirei-Souma (2004) for the U.S. and Japan from 1960 to 1999, by Clementi-Gallegati (2004) for Italy from 1977 to 2002, and by Dagsvik-Vatne (1999) for Norway in 1998. Using the richest sample of the U.S., the Forbes 400, during 1988-2003 Klass et al. (2007) find e.g., that the top end of the wealth distribution obeys a Pareto law.

⁶Stochastic labor earnings can in principle generate some skewness in the distribution of wealth, es-

In the present paper we analytically study the wealth distribution in the context of Bewley economies extended to allow for idiosyncratic capital income risk, which is naturally interpreted as entrepreneurial risk.⁷ To this end we provide first an analysis of the standard *Income Fluctuation problem*, as e.g., in Chamberlain-Wilson (2000), extended to account for capital income risk.⁸ We restrict ourselves to idiosyncratic labor earnings and capital income for simplicity. As in Aiyagari (1994), the borrowing constraint together with stochastic incomes assures a lower bound to wealth acting as a reflecting barrier.⁹ Furthermore, we show that the consumption function under borrowing constraints is strictly concave at lower wealth levels, consistent with, e.g. Saez and Zuchman (2014)'s evidence of substantial saving rate differentials across wealth levels. In this environment, therefore, the rich can get richer through savings, while the poor may not save enough to escape a poverty trap. Such non-ergodicity however would imply no social mobility between rich and poor, which seems incompatible with observed levels of social mobility in income over time and across generations; see for example Chetty, Hendren, Kline, and Saez (2014). We analytically show however that enough idiosyncratic capital income risk induces an ergodic stationary wealth distribution which is fat tailed, more precisely, a Pareto distribution in the right tail. We also show that it is capital income risk, rather than labor earnings, that drives the properties of the right tail of the

pecially if the earnings process is itself skewed and persistent. Extensive evidence for the skewedness of the income distribution has been put forth in a series of papers by Emmanuel Saez and Thomas Piketty (some with co-authors), starting with Saez and Piketty (2003) on the U.S. We refer to Atkinson, Saez, and Piketty (2011) for a survey and to the excellent website of the database they have collected (with Facundo Alvarez), The World Top Incomes Database. However, most empirical studies of labor earnings find some form of stationarity of the earnings process; see Guvenen (2007) and e.g., the discussion of Primiceri and van Rens (2006) by Heathcote (2008). Persistent income shocks are often postulated to explain the cross-sectional distribution of consumption but seem hardly enough to produce fat tailed distributions of wealth; see e.g., Storesletten, Telmer, Yaron (2004). See also Cagetti and De Nardi (2008) for a survey.

⁷Capital income risk has been introduced by Angeletos and Calvet (2005) and Angeletos (2007) and further studied by Panousi (2008) and by ourselves (Benhabib, Bisin, and Zhu, 2011 and 2013). Quadrini (1999, 2000) and Cagetti and De Nardi (2006) study entrepreneurial risk explicitly. Jones and Kim (2014) study entrepreneurs in a growth context under risk introduced by creative destruction. Relatedly, Krusell and Smith (1998) introduce heterogeneous discount rates to numerically produce some skewness in the distribution of wealth. We refer to these papers and our previous papers, as well as to Benhabib and Bisin (2006) and Benhabib and Zhu (2008), for more general evidence on the macroeconomic relevance of capital income risk.

⁸The work by Levhari and Srinivasan (1969), Schectman (1976), Schectman and Escudero (1977), Chamberlain-Wilson (2000), Huggett (1993), Rabault (2002), Carroll and Kimball (2005) has been instrumental to provide several incremental pieces to our characterization of the solution of (various specifications of) the Income Fluctuation problem; see Ljungqvist and Sargent (2004), Ch. 16, as well as Rios-Rull (1995) and Krusell-Smith (2006), for a review of results regarding the standard Income Fluctuation problem.

⁹See also Achdou, Lasry, Lions and Moll (2014) for numerical solutions of a model with stochastic returns and borrowing constraints, exploring the interaction of aggregate shocks and inequality on the transition dynamics of the macroeconomy.

wealth distribution.¹⁰

The rest of the paper is organized as follows. We present the basic setup of our economy in Section 2. Section 3 contains the characterization of the income fluctuation problem with idiosyncratic capital income risk. In Section 4 we show that the wealth accumulation process has a unique stationary distribution and the stationary distribution displays a fat right tail. In Section 5 we show that our analysis of the wealth distribution induced by the income fluctuation problem can be embedded in general equilibrium. Section 6 contains some simulation results regarding the stationary wealth distribution and the social mobility of the wealth accumulation process.

2 The economy

Consider an infinite horizon economy with a continuum of agents uniformly distributed with measure 1.¹¹ Let $\{c_t\}_{t=0}^\infty$ denote an agent consumption process and $\{a_{t+1}\}_{t=0}^\infty$ his/her wealth process. In the economy, each agent faces a no-borrowing constraint at each time t :

$$c_t \leq a_t.$$

Let $\{y_t\}_{t=0}^\infty$ represent an agent's earnings process and $\{R_{t+1}\}_{t=0}^\infty$ his/her idiosyncratic rate of return on wealth process.

Each agent in the economy solves the *Income Fluctuation (IF) problem* which is obtained under Constant Relative Risk Aversion (CRRA) preferences,

$$u(c_t) = \frac{c_t^{1-\gamma}}{1-\gamma},$$

constant discounting $\beta < 1$, and *capital income risk and earnings* processes, $\{R_{t+1}\}_{t=0}^\infty$ and $\{y_t\}_{t=0}^\infty$:

$$\max_{\{c_t\}_{t=0}^\infty, \{a_{t+1}\}_{t=0}^\infty} E \sum_{t=0}^{\infty} \beta^t \frac{c_t^{1-\gamma}}{1-\gamma} \tag{IF}$$

$$\text{s.t. } a_{t+1} = R_{t+1}(a_t - c_t) + y_{t+1}$$

¹⁰This complements the results in our previous papers (Benhabib, Bisin, and Zhu, 2011 and 2013), which focus on overlapping generation economies. An alternative approach to generate fat tails without stochastic returns is to introduce a model with bequests, where the probability of death (and/or retirement) is independent of age. In these models, the stochastic component is not stochastic returns but the length of life. For models that embody such features, see Wold and Whittle (1957), Castaneda, Gimenez and Rios-Rull (2003), and Benhabib and Bisin (2006). On the other hand, sidestepping the income fluctuation problem by assuming a constant savings rate, Nirei (2014) shows that thick tails are a direct consequence of the linearity of the wealth equation.

¹¹We avoid introducing notation to index agents in the paper.

$$c_t \leq a_t$$

a_0 given.

The following assumptions characterize formally the stochastic properties of the economic environment:

Assumption 1 R_t and y_t are stochastic processes, identically and independently distributed (i.i.d.) over time and across agents: y_t has probability density function $f(y)$ on bounded support $[\underline{y}, \bar{y}]$, with $\underline{y} > 0$ and R_t has probability density function $g(R)$ on support $[\underline{R}, \bar{R}]$. Furthermore, y_t satisfies **i)** $(\bar{y})^{-\gamma} < \beta E [R_t (y_t)^{-\gamma}]$, while R_t satisfies: **ii)** $\bar{R} > \underline{R} > 0$ large enough,¹² **iii)** $\beta E R_t^{1-\gamma} < 1$; **iv)** $(\beta E R_t^{1-\gamma})^{\frac{1}{\gamma}} E R_t < 1$; and **v)** $\Pr(\beta R_t > 1) > 0$ and any finite moment of R_t exists.

2.1 Outline

It is useful to briefly outline the role of our assumptions and our strategy to obtain the main results in the paper. Assumptions 1.i) and 1.ii) guarantee that an agent with zero wealth at some time t will not consume all his/her income at time $t + 1$ for high enough realizations of earnings and rates of return; as a consequence, the lower bound of the wealth space is a reflecting barrier, i.e. the wealth accumulation process is not stuck at the lower end of the wealth space, $a = 0$ (see Lemma 7 in the Appendix). The stochastic process for wealth is then ergodic.

Assumptions 1.iii) and 1.iv) guarantee that the wealth accumulation process is stationary. In particular, Assumption 1.iii) guarantees that the aggregate economy displays no unbounded growth in consumption and wealth.¹³ Assumption 1.iv) implies that

$$\beta E R_t < 1.$$

This is enough to guarantee that the economy contracts, giving rise to a stationary distribution of wealth. However, since we cannot obtain explicit solutions for consumption or savings policies, we have to explicitly show that under suitable assumptions there are no disjoint invariant sets or cyclic sets in wealth, so that agents do not get trapped in subsets of the support of the wealth distribution. In other words we have to show that the stochastic process for wealth is ergodic, and that a unique stationary distribution exists. We show this in Lemmata 6 and 8 in the Appendix. We then have to show that

¹²While $\bar{R} = \infty$ is allowed for, a finite \bar{R} , as derived in the proof of Theorem 4, is sufficient for all our results.

¹³We can allow for exogenous growth $g > 1$ in earnings, as in Aiyagari and McGrattan (1998). To this end, we need to deflate the variables by the growth rate and let the borrowing constraint grow at growth rate. In our context, since we allow for no borrowing, no modification of the constraint is needed. However, Assumption 1.2.iii) would have to be modified so that $\Pr(\frac{\beta R_t}{g^\gamma} > 1) > 0$.

idiosyncratic capital income risk can give rise to a fat-tailed wealth distribution. Since in our economic environment policy functions are not linear and explicit solutions are not available even under CRRA preferences, we cannot use the results of Kesten (1973), for example as in Benhabib, Bisin and Zhu (2011). We are nonetheless able to show that consumption and savings policies are asymptotically linear; a result which, under appropriate assumptions, allow us to apply Mirek (2011)'s generalization of Kesten (1973). We do this in Propositions 3, 4 and 5. The fat right tail of the stationary distribution of wealth, obtained in Theorem 3, exploits crucially that $\Pr(\beta R_t > 1) > 0$, that is, Assumption 1.v).

3 The income fluctuation problem with idiosyncratic capital income risk

In this section we show several technical results about the consumption function $c(a)$ which solves problem IF, as a build-up for its characterization of the wealth distribution in the next section. All proofs are in the Appendix.

Theorem 1 *A consumption function $c(a)$ which satisfies the constraints of the IF problem and furthermore satisfies*

i) *the Euler equation*

$$u'(c(a)) \geq \beta ER_{t+1}u'(c[R(a - c(a)) + y]) \text{ with equality if } c(a) < a, \quad (1)$$

and

ii) *the transversality condition*

$$\lim_{t \rightarrow \infty} E\beta^t u'(c_t)a_t = 0, \quad (2)$$

represents a solution of the IF problem.

By strict concavity of $u(c)$, there exists a unique $c(a)$ which solves the IF problem.

The study of $c(a)$ requires studying two auxiliary problems. The first is a version the IF problem in which the stochastic process for earnings $\{y_t\}_0^\infty$ is turned off, that is, $y_t = 0$, for any $t \geq 0$. The second is a finite horizon version of the IF problem. In both cases we naturally maintain the relevant specification and assumptions imposed on our main IF problem.

3.1 The IF problem with no earnings

The formal IF problem with no earnings is:

$$\begin{aligned} \max_{\{c_t\}_{t=0}^{\infty}, \{a_{t+1}\}_{t=0}^{\infty}} E \sum_{t=0}^{\infty} \beta^t \frac{c_t^{1-\gamma}}{1-\gamma} & \quad (\text{IF with no earnings}) \\ \text{s.t. } a_{t+1} = R_{t+1}(a_t - c_t) & \\ c_t \leq a_t & \\ a_0 \text{ given.} & \end{aligned}$$

This problem can indeed be solved in closed form, following Levhari and Srinivasan (1969). Note that for this problem the borrowing constraint is never binding because Inada conditions are satisfied for CRRA utility.

Proposition 1 *The unique solution to the IF problem with no earnings is*

$$c^{no}(a) = \phi a, \text{ for some } 0 < \phi < 1.$$

3.2 The finite IF problem

For any $T > 0$, let the finite IF problem be:

$$\begin{aligned} \max_{\{c_t\}_{t=0}^T, \{a_{t+1}\}_{t=0}^{T-1}} E \sum_{t=0}^T \beta^t \frac{c_t^{1-\gamma}}{1-\gamma} & \quad (\text{finite IF}) \\ \text{s.t. } a_{t+1} = R_{t+1}(a_t - c_t) + y_{t+1}, \quad \text{for } 0 \leq t \leq T-1 & \\ c_t \leq a_t, \quad \text{for } 0 \leq t \leq T & \\ a_0 \text{ given.} & \end{aligned}$$

With some notational abuse, let c^t denote consumption t periods from the end-period T , that is, consumption at time $T - t$.

Proposition 2 *The unique solution to the finite IF problem is a consumption function $c^t(a)$ which is continuous and increasing in a . Furthermore, let s^t denote the induced savings function,*

$$s^t(a) = a - c^t(a).$$

Then $s^t(a)$ is also continuous and increasing in a .

3.3 Characterization of $c(a)$

We can now derive a relation between $c^t(a)$, $c^{no}(a)$ and $c(a)$. The following Lemma is a straightforward extension of Proposition 2.3 and Proposition 2.4 in Rabault (2002).

Lemma 1 $\lim_{t \rightarrow \infty} c^t(a)$ exists, it is continuous and satisfies the Euler equation. Furthermore,

$$\lim_{t \rightarrow \infty} c^t(a) \geq c^{no}(a).$$

The main result of this section follows:

Theorem 2 *The unique solution to the IF problem is the consumption function $c(a)$ which satisfies:*

$$c(a) = \lim_{t \rightarrow \infty} c^t(a).$$

Let the induced savings function $s(a)$ be

$$s(a) = a - c(a).$$

Proposition 3 *The consumption and savings functions $c(a)$ and $s(a)$ are continuous and increasing in a .*

Carroll and Kimball (2005) show that $c^t(a)$ is concave.¹⁴ But Lemma 2 guarantees that $c(a) = \lim_{t \rightarrow \infty} c^t(a)$ and thus $c(a)$ is also a concave function of a .

Proposition 4 *The consumption function $c(a)$ is a concave function of a .*

The most important result of this section is that the optimal consumption function $c(a)$, in the limit for $a \rightarrow \infty$, is linear and has the same slope as the optimal consumption function of the income fluctuation problem with no earnings, ϕ .

Proposition 5 *The consumption function $c(a)$ satisfies $\lim_{a \rightarrow \infty} \frac{c(a)}{a} = \phi$.*

The proof, in the Appendix, is non-trivial.

¹⁴See also Carroll, Slacalek, and Tokuoka (2014).

4 The stationary distribution

In this section we study the distribution of wealth in the economy. The wealth accumulation equation of the IF problem is

$$a_{t+1} = R_{t+1}(a_t - c(a_t)) + y_{t+1}. \quad (3)$$

It is useful to compare it with the IF with no earnings. Using Lemma 1 we have:

$$\begin{aligned} a_{t+1} &= R_{t+1}(a_t - c(a_t)) + y_{t+1} \\ &\leq R_{t+1}(a_t - c^{no}(a_t)) + y_{t+1} \\ &= R_{t+1}(1 - \phi)a_t + y_{t+1}. \end{aligned}$$

Let

$$\mu = 1 - \phi = (\beta ER^{1-\gamma})^{\frac{1}{\gamma}}.$$

Thus $\mu < 1$ by Assumption 1.iii). We have

$$a_{t+1} \leq \mu R_{t+1}a_t + y_{t+1}.$$

The main results in this section are the following two theorems.

Theorem 3 *There exists a unique stationary distribution for a_{t+1} which satisfies the stochastic wealth accumulation equation (3).*

The proof, in the Appendix, requires several steps. First we show that the wealth accumulation process $\{a_{t+1}\}_{t=0}^{\infty}$ induced by equation (3) above is φ -irreducible, *i.e.*, there exists a non trivial measure φ on $[y, \infty)$ such that if $\varphi(A) > 0$, the probability that the process enters the set A in finite time is strictly positive for any initial condition (see Chapter 4 of Meyn and Tweedie (2009)). We also show that $a = y$ represents a reflecting barrier for the process. To show that there exists a unique stationary wealth distribution we exploit the results in Meyn and Tweedie (2009) and show that the process $\{a_{t+1}\}_{t=0}^{\infty}$ is ergodic.

Finally, in the next theorem we show that the wealth accumulation process $\{a_{t+1}\}_{t=0}^{\infty}$ has a fat tail. We use the characterization of $c(a)$ and $s(a)$ in Section 3.3, and in particular the fact that $\frac{s(a)}{a}$ is increasing in a and $\frac{s(a)}{a}$ approaches μ as a goes to infinity; this allows us to apply some results by Mirek (2011) regarding conditions for asymptotically Pareto stationary distributions for processes induced by non-linear stochastic difference equations.

Theorem 4 *The unique stationary distribution for a_{t+1} which satisfies the stochastic wealth accumulation equation (3) has a fat tail; *i.e.*, there exist $1 < \alpha < \infty$ and an $\epsilon > 0$ arbitrarily small such that*

$$E(M^\epsilon)^\alpha = 1, \quad M^\epsilon = \mu^\epsilon R_t \quad \text{and} \quad \mu - \mu^\epsilon < \epsilon.$$

and

$$\liminf_{a \rightarrow \infty} \frac{\Pr(a_{t+1} > a)}{a^{-\alpha}} \geq C,$$

where C is a positive constant.

Proof. Since $\frac{s(a)}{a}$ is increasing in a and $\frac{s(a)}{a}$ approaches μ as a goes to infinity, we can pick a large a^ϵ such that

$$\mu - \frac{s(a^\epsilon)}{a^\epsilon} < \epsilon.$$

Let

$$\mu^\epsilon = \frac{s(a^\epsilon)}{a^\epsilon}.$$

Thus $\mu - \epsilon < \mu^\epsilon \leq \mu$.

Let

$$l(a) = \begin{cases} s(a), & a \leq a^\epsilon \\ \mu^\epsilon a, & a \geq a^\epsilon \end{cases}. \quad (4)$$

Note that $l(a) \leq s(a)$ for $\forall a \in [\underline{y}, \infty)$, since $\frac{s(a)}{a}$ is increasing in a ; furthermore, the function $l(a)$ in (4) is Lipschitz continuous, since $s(a)$ is Lipschitz continuous.

Let

$$\psi(a) = R_t l(a) + y.$$

Now we apply Theorem 1.8 of Mirek (2011), to show that the stochastic process $\{\tilde{a}_{t+1}\}_{t=0}^\infty$, induced by $\tilde{a}_{t+1} = \psi(\tilde{a}_t)$, has a unique stationary distribution and that the tail of the stationary distribution for \tilde{a}_{t+1} is asymptotic to a Pareto law, *i.e.*

$$\lim_{a \rightarrow \infty} \frac{\Pr(\tilde{a}_{t+1} > a)}{a^{-\alpha}} = C,$$

where C is a positive constant.

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

By the definition of $\psi(\cdot)$ we have

$$\lim_{\tau \rightarrow 0} \left[\tau \psi \left(\frac{1}{\tau} a \right) \right] = M^\epsilon a \quad \text{for } \forall a \in [\underline{y}, \infty).$$

Let

$$N_t = \Omega R_t + y_t$$

where

$$\Omega = \max_{a \in [\underline{y}, a^\epsilon]} |s(a) - \mu^\epsilon a|.$$

It is easy to verify that

$$|\psi(a) - M^\epsilon a| < N_t \quad \text{for } \forall a \in [\underline{y}, \infty).$$

Thus $\psi(\cdot)$ satisfies Assumption 1.6 (Shape of the mappings) of Mirek (2011).

Obviously, the conditional law of $\log M^\epsilon$ is non arithmetic. Let $h(d) = \log E(M^\epsilon)^d$. By Assumption 1.iv) we have $E(\mu R_t) < 1$. Thus $h(1) = \log E(M^\epsilon) \leq \log E(\mu R) < 0$. We now show that Assumption 1.iii) and Assumption 1.iv) imply that there exists $\kappa > 1$ such that $\mu^\kappa E(R_t)^\kappa > 1$. By Jensen's inequality we have $E(R_t)^{1-\gamma} \geq (ER_t)^{1-\gamma}$. Also, Assumption 1.2.iv) implies that $\beta ER_t < 1$. Thus

$$\mu = (\beta E(R_t)^{1-\gamma})^{\frac{1}{\gamma}} \geq \beta.$$

Thus

$$E(\mu R_t)^\kappa \geq E(\beta R_t)^\kappa \geq \int_{\{\beta R_t > 1\}} (\beta R_t)^\kappa.$$

By Assumption 1.v), $\Pr(\beta R_t > 1) > 0$. Thus there exists $\kappa > 1$ such that $\mu^\kappa E(R_t)^\kappa > 1$. We could pick μ^ϵ such that $(\mu^\epsilon)^\kappa E(R_t)^\kappa > 1$. Thus $h(\kappa) = \log E(M^\epsilon)^\kappa > 0$. By Assumption 1.v), any finite moment of R_t exists. Thus $h(d)$ is a continuous function of d . Thus there exists $\alpha > 1$ such that $h(\alpha) = 0$, *i.e.* $E(M^\epsilon)^\alpha = 1$. Also we know that $h(d)$ is a convex function of d . Thus there is a unique $\alpha > 0$, such that $E(M^\epsilon)^\alpha = 1$.

Moreover, $E[(M^\epsilon)^\alpha | \log M^\epsilon] < \infty$, since M^ϵ has a lower bound, and, by Assumption 1.v), any finite moment of R_t exists.

We also know that $E(N_t)^\alpha < \infty$ since y_t has bounded support and, by Assumption 1.v), any finite moment of R_t exists.

Thus M^ϵ and N_t satisfy Assumption 1.7 (Moments condition for the heavy tail) of Mirek (2011).

By Lemma 7 in the Appendix, $a = \underline{y}$ is a reflecting barrier of the process $\{a_{t+1}\}_{t=0}^\infty$. Assumption 1.ii) requires that the upper bound of the support of R_t , \bar{R} is large enough. Consider first the case in which R_t is unbounded, $\bar{R} = \infty$. It follows that in this case the support of the stationary distribution for \tilde{a}_{t+1} is unbounded.

Applying Theorem 1.8 of Mirek (2011), we find that the stationary distribution \tilde{a}_{t+1} has a Pareto tail. Finally, we show that the stationary wealth distribution a_{t+1} , has a fat tail.

Pick $a_0 = \tilde{a}_0$. The stochastic process $\{a_{t+1}\}_{t=0}^\infty$ is induced by

$$a_{t+1} = R_{t+1}s(a_t) + y_{t+1}.$$

And the stochastic process $\{\tilde{a}_{t+1}\}_{t=0}^\infty$ is induced by

$$\tilde{a}_{t+1} = R_{t+1}l(\tilde{a}_t) + y_{t+1}.$$

For a path of (R_t, y_t) , we have $a_t \geq \tilde{a}_t$. Thus for $\forall a > \underline{y}$, we have

$$\Pr(a_t > a) \geq \Pr(\tilde{a}_t > a).$$

This implies that

$$\Pr(a_{t+1} > a) \geq \Pr(\tilde{a}_{t+1} > a),$$

since the stochastic processes $\{a_{t+1}\}_{t=0}^{\infty}$ and $\{\tilde{a}_{t+1}\}_{t=0}^{\infty}$ are ergodic. Thus

$$\liminf_{a \rightarrow \infty} \frac{\Pr(a_{t+1} > a)}{a^{-\alpha}} \geq \liminf_{a \rightarrow \infty} \frac{\Pr(\tilde{a}_{t+1} > a)}{a^{-\alpha}} = \lim_{a \rightarrow \infty} \frac{\Pr(\tilde{a}_{t+1} > a)}{a^{-\alpha}} = C.$$

Consider now the case in which the distribution of R is on a bounded support $[\underline{R}, \bar{R}]$. Along with Assumption 1.ii), we can now construct a sufficient condition on \bar{R} . Evaluate $\frac{c(a)}{a}$ at $a = \bar{y}$ and compute $\bar{R} \left(1 - \frac{c(\bar{y})}{\bar{y}}\right)$. If $\bar{R} \left(1 - \frac{c(\bar{y})}{\bar{y}}\right) > 1$, then the economy's stationary distribution of wealth has fat tails. But suppose instead that $\bar{R} \left(1 - \frac{c(\bar{y})}{\bar{y}}\right) < 1$. In this case pick an $\hat{R} > \bar{R}$ and such that $\hat{R} \left(1 - \frac{c(\bar{y})}{\bar{y}}\right) > 1$ and perturb g , the distribution of R , as follows: $g(R; \varepsilon) = (1 - \varepsilon)g(R)$ for any $R \in [\underline{R}, \bar{R}]$ and $g(\hat{R}; \varepsilon)$ has mass ε . Note that $g(\cdot, 0) = g$, so that we effectively produced a continuous parametrization of the distribution g . The parametrization is continuous in the sense that $\int h(r)g(R; \varepsilon)dR$ is continuous in ε for any continuous function h . Now this construction guarantees that wealth a can escape to the expanding region with positive probability ε . Indeed by Berge's maximum theorem $\frac{c(\bar{y})}{\bar{y}}$ is continuous in ε and \hat{R} can be chosen large enough to compensate any local variation in $\frac{c(\bar{y})}{\bar{y}}$. As a consequence, this construction produces an economy whose stationary distribution of wealth has fat tails even with a distribution of R which is bounded above. What is really needed is that the distribution of R has any positive density above the R^* such that $R^* \left(1 - \frac{c(\bar{y})}{\bar{y}}\right) = 1$, even if the support is not connected. ■

5 General equilibrium

In this section we embed the analysis of the distribution of wealth induced by the IF problem in general equilibrium. Following Angeletos (2007) we assume that each agent acts as entrepreneur of his own individual firm. Each firm has a constant returns to scale neo-classical production function

$$F(k, n, A)$$

where k, n are, respectively, capital and labor, and A is an idiosyncratic productivity shock. Agents can only use their own savings as capital in their own firm. In each period $t + 1$, each agent observes his/her firm's productivity shock A_{t+1} and decides how much labor to hire in a competitive labor market, n_{t+1} . Therefore, each firm faces the same market wage rate w_{t+1} . The capital he/she invests is instead predetermined, but the agent can decide not to engage in production, in which case $n_{t+1} = 0$ and the capital

invested is carried over (with no return nor depreciation) to the next period. The firm's profits in period $t + 1$ are denoted π_{t+1} :

$$\pi_{t+1} = \max \{F(k_{t+1}, n_{t+1}, A_{t+1}) - wn_{t+1} + (1 - \delta)k_{t+1}, k_{t+1}\}. \quad (5)$$

Letting each agent's earnings in period $t + 1$ are denoted $w_{t+1}e_{t+1}$, where e_{t+1} is his/her idiosyncratic (exogenous) labor supply, we have

$$a_{t+1} = \pi_{t+1} + w_{t+1}e_{t+1}.$$

Furthermore,

$$k_{t+1} = a_t - c_t.$$

Given a sequence $\{w_t\}_{t=0}^\infty$, each agent solves the following modified IF problem:

$$\max_{\{c_t, n_t\}_{t=0}^\infty, \{k_{t+1}, a_{t+1}\}_{t=0}^\infty} E \sum_{t=0}^{\infty} \beta^t \frac{c_t^{1-\gamma}}{1-\gamma} \quad (\text{Gen Eq IF})$$

s.t. $a_{t+1} = \pi_{t+1} + w_{t+1}e_{t+1}$ where π_{t+1} is defined in (5)

$$k_{t+1} = a_t - c_t$$

$$c_t \leq a_t$$

k_0 given.

Definition 2 *A stationary general equilibrium in our economy consists of a constant wage rate w , sequences $\{c_t, n_t\}_{t=0}^\infty, \{k_{t+1}, a_{t+1}\}_{t=0}^\infty$ which constitute a solution to the Gen Eq IF problem under $w_t = w$ for any $t \geq 0$, and a distribution $v(a_{t+1})$, such that the following conditions hold:*

- (i) *labour markets clear: $En_t = Ee_t$;*¹⁵
- (ii) *v is a stationary distribution of a_{t+1} .*

We can now illustrate how such an equilibrium can be constructed, inducing a stationary distribution of wealth $v(a_{t+1})$ with the same properties, notably the fat tail, as the one characterized in the previous section under appropriate assumptions for the stochastic processes $\{A_{t+1}\}_{t=0}^\infty$ and $\{e_t\}_{t=0}^\infty$. The first order conditions of each agent firm's labor choice requires

$$\frac{\partial F}{\partial n}(k_{t+1}, n_{t+1}, A_{t+1}) = w_{t+1};$$

which, under constant returns to scale implies,

$$\frac{\partial F}{\partial n} \left(1, \frac{n_{t+1}}{k_{t+1}}, A_{t+1} \right) = w_{t+1}. \quad (6)$$

¹⁵The usual abuse of the Law of Large Number guarantees that the market clearing condition as stated holds in the cross-section of agents.

Equation (6) can be solved to give

$$\frac{n_{t+1}}{k_{t+1}} = n(w_{t+1}, A_{t+1}); \text{ or } n_{t+1} = g(w_{t+1}, A_{t+1}) k_{t+1}.$$

The market clearing condition i) in Definition 2 is then satisfied by a constant wage rate w such that

$$En_{t+1} = E(g(w, A_{t+1})) Ek_{t+1},$$

as long as the process $\{A_{t+1}\}_{t=0}^{\infty}$ is i.i.d. over time and in the cross-section and Ek_{t+1} is constant over time.

From the constant returns to scale assumption, once again, we can write profits π_{t+1} as:

$$\pi_{t+1} = R_{t+1}k_{t+1}$$

where $\{R_{t+1}\}_{t=0}^{\infty}$ is induced by the process $\{A_{t+1}\}_{t=0}^{\infty}$ as follows:

$$R_{t+1} = \max \left\{ \frac{\partial F}{\partial k} \left(1, \frac{n_{t+1}}{k_{t+1}}, A_{t+1} \right) + 1 - \delta, 1 \right\}.$$

Letting then $y_{t+1} = we_{t+1}$, the dynamic equation for wealth can be written as

$$a_{t+1} = R_{t+1}(a_t - c_t) + y_{t+1}.$$

We conclude that the solution to Gen Eq IF induces a stochastic process $\{a_{t+1}\}_{t=0}^{\infty}$ which has the same properties as the one induced by the IF problem as long as i) Ek_{t+1} is constant and ii) the process $\{R_{t+1}\}_{t=0}^{\infty}$ induced by $\{A_{t+1}\}_{t=0}^{\infty}$ and the process $\{y_t\}_{t=0}^{\infty}$ induced by $\{e_t\}_{t=0}^{\infty}$ satisfy Assumption 1. In particular, in this case, $\{a_{t+1}\}_{t=0}^{\infty}$ has a unique stationary distribution. The stationary distribution of $\{a_{t+1}\}_{t=0}^{\infty}$ induces in turn a stationary distribution of k_{t+1} . The aggregate capital Ek_{t+1} is the first moment of the stationary distribution of k_{t+1} and is therefore constant. As a consequence, the labor market indeed clear with a constant wage w as postulated. It is verified then that at a stationary general equilibrium, as long as ii) above is satisfied, the stochastic process $\{a_{t+1}\}_{t=0}^{\infty}$ has the same properties as the one induced by the IF problem; it displays, in particular, a fat tail.

6 Simulation

In this section we carry out a simulation of the Bewley economy we studied in the paper. The objective is simply to illustrate that a reasonable (though not calibrated) parameter set produces a stationary wealth distribution which is Pareto in the tail with an exponent

in the same order of magnitude as the one estimated for various developed economies, that is, about 2.¹⁶ We shall defer a careful calibration exercise to future work.¹⁷

Consider a Bewley economy as introduced in Section 2, with the following specifications: log preferences ($\gamma = 1$); a uniformly distributed earnings process over support $[1, 44]$; a rate of return process R_t defined on $[0.8, 1, 1.08, 9]$ with associated probabilities $[0.2, 0.4, 0.39, 0.01]$; $\beta = 0.92$.¹⁸

Under this specification, we solve for the consumption function which, consistently with Propositions 4 and 5, is a concave function linear in the tail, as shown in Figure 1.

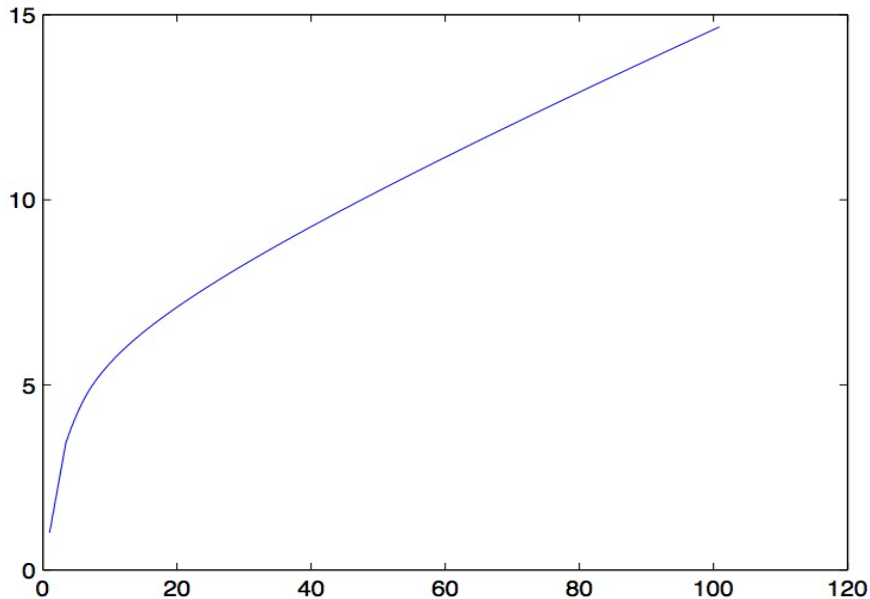


Figure 1: Consumption function

The stationary distribution of wealth, obtained for a million households after iterating for 400 periods, is shown in Figure 2.¹⁹

¹⁶See the references in footnote 5.

¹⁷Calibrating the (idiosyncratic component of) the rate of return on wealth, in particular, appears delicate. However, Saez and Zucman (2014)'s method produces a large idiosyncratic variance consistent with the large documented variability of two major components of capital income, ownership of principal residence (see Case and Shiller, 1989; and Flavin and Yamashita, 2002) and private business equity (Moskowitz and Vissing-Jorgensen, 2002; and Bitler, Moskowitz and Vissing-Jorgensen, 2005). See Angeletos (2007), Quadrini (2000), and Benhabib and Zhu (2008) for more evidence on the macroeconomic relevance of idiosyncratic capital income risk.

¹⁸The process R_t implies an expected return of 7% and hence $\beta ER < 1$ is satisfied.

¹⁹We determine that a stationary distribution once we find convergence of cut-offs for wealth quintiles as well as top decile and percentile. Given the large returns available, convergence requires a large number of simulated households.

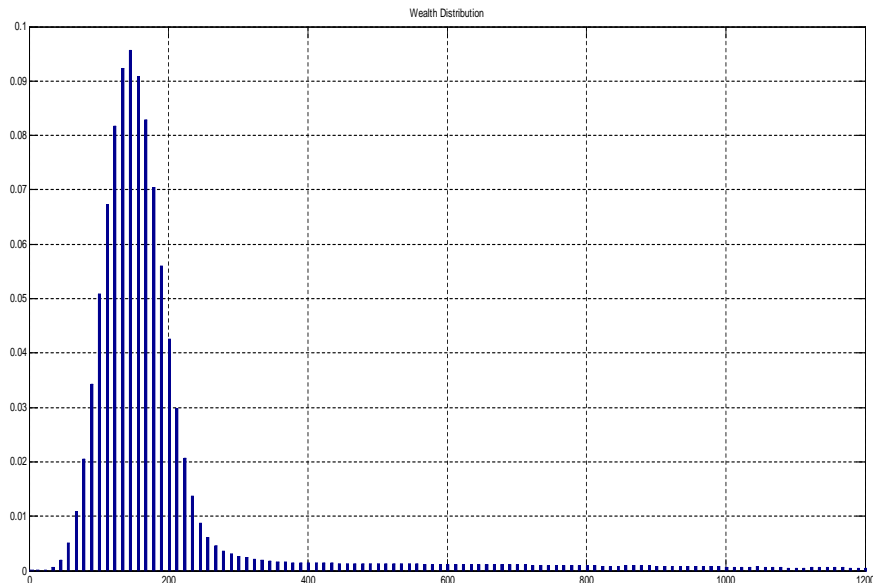


Figure 2: Stationary distribution of wealth

The distribution of wealth by quintiles and for top shares for the U.S. economy and the model is given below:

<i>Economy</i>	<i>Quintiles</i>				
	<i>First</i>	<i>Second</i>	<i>Third</i>	<i>Fourth</i>	<i>Fifth</i>
<i>U.S.</i>	-.003	.013	.05	.122	.817
<i>Model</i>	.058	.077	.091	.107	.668

<i>Economy</i>	<i>Percentiles</i>		
	<i>90th – 95th</i>	<i>95th – 99th</i>	<i>99th – 100th</i>
<i>U.S.</i>	.113	.231	.347
<i>Model</i>	.077	.135	.387

While the match for quintiles and top shares seems quite reasonable,²⁰ the tail index or Pareto tail for the U.S. economy is smaller than in the simulation (between 1.5 – 2.0 rather than 2.1).²¹

²⁰This comparison implicitly assumes that the wealth distribution for the U.S. is close to stationary. This might in general not be the case if the wealth distribution is hit frequently enough by aggregate shocks like wars, major business cycle events (e.g., a depression), changes in tax schemes, social insurance institutions, and so on; see Saez and Piketty (2003). We leave the study of the transition of the distribution of wealth for future work.

²¹The tail exponent is computed using a Matlab package, *plfit*, based on Clauset, Shalizi, and Newman (2009).

Our analysis of the stochastic process of wealth induced by stochastic earnings and capital income risk has direct implications with regards to the social mobility at the stationary distribution. From our simulations we compute the 5-year quintile transition matrix which we compare to the corresponding transition matrix for the U.S., estimated by Klevmarken, Lupton and Stafford (2003) for 1994-1999 (see Table 9, p. 342).²² There are several methods to more formally measure social mobility for transition matrices.²³ We adopt the Shorrocks index, which gives the average probability over all wealth classes of leaving the initial wealth class over the observation period.²⁴ This index takes value 0.8375 for the simulated 5-year quintile transition matrix, while only 0.6610 for the corresponding matrix for the U.S., indicating that our specification overestimates social mobility. A similar result is obtained when we compare the Shorrocks index of social mobility for the transition matrix from our simulation with the corresponding one for the U.S. estimated by Hurst and Kerwin (2003), Tables 2 and 5.²⁵ In this case, however, the overestimation of social mobility in the simulation is more moderate: the Shorrocks index is .9516 and .885 respectively for the simulated matrix and the U.S. matrix.²⁶ We conclude that, while i.i.d. shocks can generate a fat right tail not too different from the one observed in the data, further and careful calibrations would be required to more closely match other aspects of the data, in particular the tail index and social mobility jointly. To this end it seems of first order importance to allow for persistence in both the earnings and in the capital income risk processes. We in fact experimented along these lines, introducing a small degree of persistence in the returns with a transition matrix that slightly deviates from the i.i.d. case. Persistence indeed shifts the distribution to the right and reduces social mobility, thereby helping the model to better fit both the

²²The (one-year) quintile transition matrix we obtain from the simulation is

$$T = \begin{bmatrix} .7244 & .2440 & .0214 & 0 & .0103 \\ .2137 & .4511 & .3020 & .0222 & .0109 \\ .0588 & .2286 & .4250 & .2770 & .0106 \\ .0031 & .0753 & .2268 & .5771 & .1178 \\ 0 & .0009 & .0249 & .1237 & .8505 \end{bmatrix}.$$

Since the quintile transition matrix is obtained from the stationary distribution it is time homogeneous and the t -years transition matrix, for any $t > 1$, is then simply computed as T^t .

²³See, e.g., Section 5 in Dardanoni (1993).

²⁴Formally, for an square matrix A with m rows the Shorrocks index given by $s(A) = \frac{m - \sum_j a_{jj}}{m-1}$; see Shorrocks(1978).

²⁵Hurst and Kerwin (2003)'s matrix is a sample cross-section average over 10–15 years, from 1985-89 to 1999. The simulated matrix we construct for comparison is a 12-year quintile transmission matrix.

²⁶The same qualitative result is obtained comparing our simulated quintile transition process with the 6-year quintile matrix obtained by Kennickell and Starr-McCluer (1997), Table 6. Also, measures of social mobility alternative to Shorrocks index, like the second largest eigenvalue of the transition matrix (see Dardanoni (1993), yield no qualitative differences in all these comparisons .

stationary distribution of wealth as well as relevant aspects of social mobility. Also, some non-homogeneity in savings, beyond that implied by the concave consumption function of the Bewley model, may be introduced to account for fatter tails and less social mobility in wealth distribution.

7 Conclusions

In this paper we construct a general equilibrium model with idiosyncratic capital income risks in a Bewley economy and analytically demonstrate that the resulting wealth distribution has a fat right tail under well defined and natural conditions on the parameters and stochastic structure of the economy. Simulations confirm that, once idiosyncratic capital income risk is taken into account, Bewley models can reproduce fundamental stylized properties of the wealth distribution observed in the data for the U.S. and other developed economies.

References

- [1] Achdou, Y., Lasry, J-M, Lions, P-L, and B. Moll, 2014. "Wealth Distribution and the Business Cycle: The Role of Private Firms", Princeton University, mimeo, available at <http://www.princeton.edu/~moll/WDBC.pdf>
- [2] Aiyagari, S.R. (1994): "Uninsured idiosyncratic risk and aggregate saving," *Quarterly Journal of Economics*, 109(3), 659-84.
- [3] Aiyagari, S.R. (1995): "Optimal capital income taxation with incomplete markets and borrowing constraints," *Journal of Political Economy*, 103(6), 1158-75.
- [4] Atkinson, T., E. Saez, and T. Piketty (2011): "Top Incomes in the Long Run of History," *Journal of Economic Literature*, 49(1), 2011, 3-71.
- [5] Angeletos, G. (2007): "Uninsured Idiosyncratic Investment Risk and Aggregate Saving," *Review of Economic Dynamics*, 10, 1-30.
- [6] Angeletos, G. and L.E. Calvet (2005): "Incomplete-Market Dynamics in a Neoclassical Production Economy," *Journal of Mathematical Economics*, 41, 407-438.
- [7] Attanasio, O.P. and S.J. Davis(1996): "Relative wage movements and the distribution of consumption," *Journal of Political Economy*, 104(6), 1227-62.
- [8] Benhabib, J. and A. Bisin (2006): "The Distribution of Wealth and Redistributive Policies," New York University, mimeo.
- [9] Benhabib, J. and S. Zhu (2008): "Age, Luck and Inheritance, NBER W.P. 14128.
- [10] Benhabib, J., A. Bisin, and S. Zhu (2011): "The Distribution of Wealth and Fiscal Policy in Economies with Finitely Lived Agents," *Econometrica*, 79(1), 122-57.
- [11] Benhabib, J., A. Bisin, and S. Zhu (2013): "The distribution of wealth in the Blanchard-Yaari model," forthcoming in *Macroeconomic Dynamics*, special issue on Complexity in Economic Systems.
- [12] Bewley, T. (1977): "The permanent income hypothesis: A theoretical formulation," *Journal of Economic Theory*, 16(2), 252-92.
- [13] Bewley, T. (1983): "A difficulty with the optimum quantity of money," *Econometrica*, 51(5), 1485-504.
- [14] Bitler, Marianne P., Tobias J. Moskowitz, and Annette Vissing-Jørgensen (2005): "Testing Agency Theory With Entrepreneur Effort and Wealth", *Journal of Finance*, 60, 539-576.

- [15] Cagetti, M. and M. De Nardi (2006): "Entrepreneurship, frictions, and wealth," *Journal of Political Economy*, 114(5), 835-70.
- [16] Cagetti, M. and M. De Nardi (2008): "Wealth inequality: Data and models," *Macroeconomic Dynamics*, 12(52), 285-313.
- [17] Carroll, C. D. (1997): "Buffer-Stock Saving and the Life Cycle/Permanent Income Hypothesis," *Quarterly Journal of Economics*, 112, 1-56.
- [18] Carroll, C. D. and M.S. Kimball (2005): "Liquidity constraints and precautionary saving," Johns Hopkins University, mimeo.
- [19] Carroll, C. D., J. Slacalek, and K. Tokuoka (2014). "The Distribution of Wealth and the Marginal Propensity to Consume", Johns Hopkins University, mimeo.
- [20] Case, K., and Shiller, R.: (1989): "The Efficiency of the Market for Single-Family Homes," *American Economic Review*, 79, 125–137.
- [21] Castaneda, A., J. Diaz-Gimenez, and J.V. Rios-Rull (2003): "Accounting for the US earnings and wealth inequality," *Journal of Political Economy*, 111(4), 818-57.
- [22] Chamberlain G. and C.A. Wilson (2000): "Optimal intertemporal consumption under uncertainty," *Review of Economic Dynamics*, 3(3), 365-95.
- [23] Chetty, R., N. Hendren, P. Kline, and E. Saez, (2014): "Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States", University of California, Berkeley, mimeo, available at http://obs.rc.fas.harvard.edu/chetty/mobility_geo.pdf.
- [24] Clauset, A., Shalizi, C.R., and M. E. J. Newman (2009): "Power-law Distributions in Empirical Data," *SIAM Review*, 51(4), 661-703.
- [25] Clementi, F. and M. Gallegati (2005): "Power Law Tails in the Italian Personal Income Distribution," *Physica A: Statistical Mechanics and Theoretical Physics*, 350, 427-438.
- [26] Dagsvik, J.K., and B.H. Vatne (1999): "Is the Distribution of Income Compatible With a Stable Distribution?" Discussion Paper 246, Research Department, Statistics Norway.
- [27] Dardanoni, V. (1993): "Measuring Social Mobility," *Journal of Economic Theory*, 61, 372-394.
- [28] Fisher, J. and D. Johnson (2006): "Consumption mobility in the United States: Evidence from two panel data sets," *B.E. Journal of Economic Analysis and Policy*, 6(1), art. 16.

- [29] Flavin, M. and T. Yamashita (2002): "Owner-Occupied Housing and the Composition of the Household Portfolio", *American Economic Review*, 92, 345-362.
- [30] Guvenen, F. (2007): "Learning your earning: Are labor income shocks really very persistent?" *American Economic Review* 97(3), 687-712.
- [31] Hansen, G.D. and A. Imrohoroglu (1992): "The role of unemployment insurance in an economy with liquidity constraints and moral hazard," *Journal of Political Economy*, 100(1), 118-42.
- [32] Heathcote, J. (2005): "Fiscal policy with heterogeneous agents and incomplete markets," *Review of Economic Studies*, 72(1), 161-88.
- [33] Heathcote, J. (2008): "Discussion of Heterogeneous Life-Cycle Profiles, Income Risk, and Consumption Inequality, by G. Primiceri and T. van Rens," Federal Bank of Minneapolis, mimeo.
- [34] Heathcote, J., K. Storesletten, and G.L. Violante (2005): "Two views of inequality over the life cycle," *Journal of the European Economic Association*. 32(3), 765-75.
- [35] Heathcote, J., K. Storesletten, and G.L. Violante (2007): "Consumption and labor supply with partial insurance: An analytical framework," New York University, mimeo.
- [36] Heathcote, J., K. Storesletten, and G.L. Violante (2008a): "Insurance and opportunities: a welfare analysis of labor market risk," *Journal of Monetary Economics*, 55(3), 501-25.
- [37] Heathcote, J., K. Storesletten, and G.L. Violante (2008b): "The macroeconomic implications of rising wage inequality in the United States," New York University, mimeo.
- [38] Huggett, M. (1993): "The risk-free rate in heterogeneous-agent incomplete-insurance economies," *Journal of Economic Dynamics and Control*, 175(6), 953-69.
- [39] Huggett, M. (1996): "Wealth distribution in life-cycle economies," *Journal of Monetary Economics*, 38(3), 469-94.
- [40] Hurst, E. and K.F. Kerwin (2003): "The Correlation of Wealth across Generations," *Journal of Political Economy*, 111, 1155-1183.
- [41] Imrohoroglu, A. (1992): "The welfare cost of inflation under imperfect insurance," *Journal of Economic Dynamics and Control*, 16(1), 79-91.

- [42] Jones, C. I. and Kim, J., (2014). "A Schumpeterian Model of Top Income Inequality", Stanford University, mimeo, available at <http://web.stanford.edu/~chadj/inequality.pdf>
- [43] Kennickell, A.B. and M. Starr-McCluer (1997): "Household Saving and Portfolio Change: Evidence from the 1983-89 SCF Panel," *Review of Income and Wealth*, 43, 1-19.
- [44] Klevmarck, N.A., J.P. Lupton, and F.P. Stafford (2003): "Wealth Dynamics in the 1980s and 1990s: Sweden and the United States", *Journal of Human Resources*, XXXVIII, 322-353.
- [45] Krueger, D. and F. Perri (2003): "On the welfare consequences of the increase in inequality in the United States," *NBER Macroeconomics Annual 2003*, edited by M. Gertler and K. Rogoff, 83-121, Cambridge, MA: MIT Press.
- [46] Jappelli, T. and L. Pistaferri (2006): "Intertemporal choice and consumption mobility." *Journal of the European Economic Association*, 4(1), 75-115.
- [47] Kesten, H. (1973): "Random Difference Equations and Renewal Theory for Products of Random Matrices," *Acta Mathematica*, 131, 207-248.
- [48] Klass, O.S., O. Biham, M. Levy, O. Malcai, and S. Solomon (2007): "The Forbes 400, the Pareto Power-Law and Efficient Markets," *The European Physical Journal B - Condensed Matter and Complex Systems*, 55, 143-147.
- [49] Kopczuk, W. and E. Saez (2014): "Top Wealth Shares in the United States, 1916-2000: Evidence from Estate Tax Returns," *National Tax Journal*, 57(2), Part 2, 445-487 (long NBER Working Paper No. 10399, March 2004).
- [50] Krusell, P. and A.A. Smith (1998): "Income and wealth heterogeneity in the macroeconomy." *Journal of Political Economy*, 106(5), 867-96.
- [51] Krusell, P. and A.A. Smith (2006): "Quantitative macroeconomic models with heterogeneous agents," in *Advances in Economics and Econometrics: Theory and Applications*, edited by R. Blundell, W. Newey, and T. Persson, Cambridge: Cambridge University Press.
- [52] Levhari, D. and T. N. Srinivasan (1969): "Optimal savings under uncertainty," *Review of Economic Studies*, 36, 153-163.
- [53] Ljungqvist, L. and T.J. Sargent (2004): *Recursive Macroeconomic Theory*, Cambridge, MA: MIT Press.

- [54] Mankiw, N.G. (1986): "The equity premium and the concentration of aggregate shocks," *Journal of Financial Economics*, 17(1), 211-19.
- [55] Meyn, S.P. and Tweedie, R.L. (2009): *Markov chains and stochastic stability*, Cambridge University Press.
- [56] Mirek, M. (2011): "Heavy tail phenomenon and convergence to stable laws for iterated Lipschitz maps," *Probability Theory and Related Fields*, 151, 705-34.
- [57] Moskowitz, T. and A. Vissing-Jorgensen (2002): "The Returns to Entrepreneurial Investment: A Private Equity Premium Puzzle?", *American Economic Review*, 92, 745-778.
- [58] Nirei, M. (2014): "Pareto Distribution of Income in Neoclassical Growth Models," Hitotsubashi University, mimeo, available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2436858
- [59] Nirei, M. and W. Souma (2004): "Two Factor Model of Income Distribution Dynamics," Utah State University, mimeo.
- [60] Piketty, T. and G. Zucman (2014): "Capital is Back: Wealth-Income Ratios in Rich Countries 1700-2010," *Quarterly Journal of Economics*, forthcoming.
- [61] Pistaferri, L. (2003): "Anticipated and unanticipated wage changes, wage risk, and intertemporal labor supply," *Journal of Labor Economics*, 21(3), 729-82.
- [62] Primiceri, G. and T. van Rens (2013): "Heterogeneous life-cycle profiles," forthcoming in *Journal of Monetary Economics*.
- [63] Quadrini, V. (1999): "The Importance of Entrepreneurship for Wealth Concentration and Mobility," *Review of Income and Wealth*, 45, 1-19.
- [64] Quadrini, V. (2000): "Entrepreneurship, saving and social mobility," *Review of Economic Dynamics*, 3(1), 1-40.
- [65] Rabault, G. (2002): "When do borrowing constraints bind? Some new results on the income fluctuation problem," *Journal of Economic Dynamics and Control*, 26, 217-45.
- [66] Rios-Rull, J.V. (1995): "Models with heterogeneous agents," In *Frontiers of Business Cycle Research*, edited by T.F. Cooley, Princeton, NJ: Princeton University Press.
- [67] Royden, H. L.(1988): *Real Analysis*, (third ed.), Upper Saddle River, New Jersey: Prentice-Hall, Inc.

- [68] Saez, E. and T. Piketty (2003): "Income Inequality in the United States, 1913-1998," *Quarterly Journal of Economics*, 118(1), 2003, 1-39.
- [69] Saez, E. and G. Zucman (2014): "The Distribution of US Wealth, Capital Income and Returns since 1913," University of California, Berkeley, unpublished slides available at <http://gabriel-zucman.eu/files/SaezZucman2014Slides.pdf>.
- [70] Schechtman, J. (1976): "An income fluctuation problem," *Journal of Economic Theory*, 12(2), 218-41.
- [71] Schechtman, J. and V.L.S. Escudero (1977): "Some results on an income fluctuation problem." *Journal of Economic Theory*, 16(2), 151-66.
- [72] Shorrocks, A.F. (1978): "Income inequality and income mobility," *Journal of Economic Theory*, 19, 376-393
- [73] Storesletten, K., C.L. Telmer, and A. Yaron (2001): "The welfare cost of business cycles revisited: Finite lives and cyclical variation in idiosyncratic risk," *European Economic Review*, 45(7), 1311-39.
- [74] Storesletten, K., C.L. Telmer, and A. Yaron (2004): "Consumption and risk sharing over the life cycle," *Journal of Monetary Economics*, 51(3), 609-33.
- [75] Wold, H.O.A. and P. Whittle (1957): "A Model Explaining the Pareto Distribution of Wealth," *Econometrica*, 25, 591-595.
- [76] Wolff, E. (1987): "Estimates of Household Wealth Inequality in the U.S., 1962-1983," *Review of Income and Wealth*, 33, 231-256.
- [77] Wolff, E. (2004): "Changes in Household Wealth in the 1980s and 1990s in the U.S.," New York University, mimeo.

Appendix

Proof of Theorem 1. A feasible policy $c(a)$ is said to overtake another feasible policy $\hat{c}(a)$ if starting from the same initial wealth a_0 , the policies $c(a)$ and $\hat{c}(a)$ yield stochastic consumption processes (c_t) and (\hat{c}_t) that satisfy

$$E \left[\sum_{t=0}^T \beta^t (u(c_t) - u(\hat{c}_t)) \right] > 0 \quad \text{for all } T > \text{some } T_0.$$

Also, a feasible policy is said to be optimal if it overtakes all other feasible policies.

Proof: For an a_0 , the stochastic consumption process (c_t) is induced by the policy $c(a)$. Let (\hat{c}_t) be an alternative stochastic consumption process, starting from the same initial wealth a_0 . By the strict concavity of $u(\cdot)$, we have

$$E \left[\sum_{t=0}^T \beta^t (u(c_t) - u(\hat{c}_t)) \right] \geq E \left[\sum_{t=0}^T \beta^t u'(c_t)(c_t - \hat{c}_t) \right].$$

From the budget constraint we have

$$a_{t+1} = R_{t+1}(a_t - c_t) + y_{t+1}$$

and

$$\hat{a}_{t+1} = R_{t+1}(\hat{a}_t - \hat{c}_t) + y_{t+1}.$$

For a path of (R_t, y_t) , we have

$$\frac{a_{t+1} - \hat{a}_{t+1}}{R_{t+1}} = a_t - c_t - (\hat{a}_t - \hat{c}_t) \tag{7}$$

and

$$c_t - \hat{c}_t = a_t - \hat{a}_t - \frac{a_{t+1} - \hat{a}_{t+1}}{R_{t+1}}.$$

Therefore we have

$$\sum_{t=0}^T \beta^t u'(c_t)(c_t - \hat{c}_t) = \sum_{t=0}^T \beta^t u'(c_t) \left(a_t - \hat{a}_t - \frac{a_{t+1} - \hat{a}_{t+1}}{R_{t+1}} \right).$$

Using $a_0 = \hat{a}_0$ and rearranging terms, we have

$$\sum_{t=0}^T \beta^t u'(c_t)(c_t - \hat{c}_t) = - \sum_{t=0}^T \beta^t [u'(c_t) - \beta R_{t+1} u'(c_{t+1})] \frac{a_{t+1} - \hat{a}_{t+1}}{R_{t+1}} - \beta^T u'(c_T) \frac{a_{T+1} - \hat{a}_{T+1}}{R_{T+1}}.$$

Using equation (7) we have

$$\begin{aligned}
\sum_{t=0}^T \beta^t u'(c_t)(c_t - \hat{c}_t) &= - \sum_{t=0}^T \beta^t [u'(c_t) - \beta R_{t+1} u'(c_{t+1})] \{a_t - c_t - (\hat{a}_t - \hat{c}_t)\} \\
&\quad - \beta^T u'(c_T) [a_T - c_T - (\hat{a}_T - \hat{c}_T)] \\
&\geq - \sum_{t=0}^T \beta^t [u'(c_t) - \beta R_{t+1} u'(c_{t+1})] \{a_t - c_t - (\hat{a}_t - \hat{c}_t)\} - \beta^T u'(c_T) a_T.
\end{aligned}$$

Thus we have

$$E \left[\sum_{t=0}^T \beta^t u'(c_t)(c_t - \hat{c}_t) \right] \geq -E \left(\sum_{t=0}^T \beta^t [u'(c_t) - \beta E R_{t+1} u'(c_{t+1})] \{a_t - c_t - (\hat{a}_t - \hat{c}_t)\} \right) - E \beta^T u'(c_T) a_T. \quad (8)$$

By the Euler equation (1) we have $u'(c_t) - \beta E R_{t+1} u'(c_{t+1}) \geq 0$. If $c_t < a_t$, then $u'(c_t) = \beta E R_{t+1} u'(c_{t+1})$. If $c_t = a_t$, then $a_t - c_t - (\hat{a}_t - \hat{c}_t) = -(\hat{a}_t - \hat{c}_t) \leq 0$. Thus

$$-E \left(\sum_{t=0}^T \beta^t [u'(c_t) - \beta E R_{t+1} u'(c_{t+1})] \{a_t - c_t - (\hat{a}_t - \hat{c}_t)\} \right) \geq 0. \quad (9)$$

Combining equations (8) and (9) we have

$$E \left[\sum_{t=0}^T \beta^t u'(c_t)(c_t - \hat{c}_t) \right] \geq -E \beta^T u'(c_T) a_T.$$

By the transversality condition (2) we know that for large T ,

$$E \left[\sum_{t=0}^T \beta^t (u(c_t) - u(\hat{c}_t)) \right] \geq 0. \quad \blacksquare$$

Proof of Proposition 1. The Euler equation of this problem is

$$c_t^{-\gamma} = \beta E R_{t+1} c_{t+1}^{-\gamma}. \quad (10)$$

Guess $c_t = \phi a_t$. From the Euler equation (10) we have

$$\phi = 1 - (\beta E R^{1-\gamma})^{\frac{1}{\gamma}},$$

which is > 0 by Assumption 1.iii).

It is easy to verify the transversality condition,

$$\lim_{t \rightarrow \infty} E(\beta^t c_t^{-\gamma} a_t) = 0. \quad \blacksquare$$

Let $V^t(a)$ be the optimal value function of an agent who has wealth a and has t periods to the end T . Thus we have

$$V^t(a) = \max_{c \leq a} \{u(c) + \beta E V^{t-1}(R(a - c) + y)\} \quad \text{for } t > 1$$

and

$$V^1(a) = \max_{c \leq a} u(c).$$

We have the Euler equation of this problem, for $t > 1$

$$u'(c^t(a)) \geq \beta E[Ru'(c^{t-1}(R(a - c^t(a)) + y))] \text{ with equality if } c^t(a) < a.$$

Proof of Proposition 2. Continuity is a consequence of the Theorem of the Maximum and mathematical induction. The proof that $c^t(a)$ and $s^t(a)$ are increasing can be easily adapted from the proof of Theorem 1.5 of Schechtman (1976); it makes use of the fact that $c^t(a) > 0$, a consequence of Inada conditions which hold for CRRA utility functions. \blacksquare

Proof of Theorem 2. By Lemma 1 we know that $c(a)$ satisfies the Euler equation. Now we verify that $c(a)$ satisfies the transversality condition (2).

By Lemma 1 and Theorem 2 we have

$$c_t \geq \phi a_t.$$

Note that $a_t \geq \underline{y}$ for $t \geq 1$. We have

$$u'(c_t) a_t \leq \phi^{-\gamma} (\underline{y})^{1-\gamma} \quad \text{for } t \geq 1.$$

Thus

$$\lim_{t \rightarrow \infty} E \beta^t u'(c_t) a_t = 0. \quad \blacksquare$$

Proof of Proposition 3. By Lemma 1, $c(a)$ is continuous. Thus $s(a)$ is continuous since $s(a) = a - c(a)$.

Also, by Lemma 1, $\lim_{t \rightarrow \infty} s^t(a) = s(a)$, since $\lim_{t \rightarrow \infty} c^t(a) = c(a)$, $s^t(a) = a - c^t(a)$, and $s(a) = a - c(a)$. The conclusion that $c(a)$ and $s(a)$ are increasing in a follows from part (ii) of Proposition 2. \blacksquare

Note that Proposition 3 implies that $c(a)$ and $s(a)$ are Lipschitz continuous. For \tilde{a} , $\hat{a} > 0$, without loss of generality, we assume that $\tilde{a} < \hat{a}$. We have $c(\tilde{a}) \leq c(\hat{a})$ and $s(\tilde{a}) \leq s(\hat{a})$. Also $c(\tilde{a}) + s(\tilde{a}) = \tilde{a}$ and $c(\hat{a}) + s(\hat{a}) = \hat{a}$. Thus

$$c(\hat{a}) - c(\tilde{a}) + s(\hat{a}) - s(\tilde{a}) = \hat{a} - \tilde{a}.$$

Thus we have

$$0 \leq c(\hat{a}) - c(\tilde{a}) \leq \hat{a} - \tilde{a}$$

and

$$0 \leq s(\hat{a}) - s(\tilde{a}) \leq \hat{a} - \tilde{a}.$$

Thus

$$|c(\hat{a}) - c(\tilde{a})| \leq |\hat{a} - \tilde{a}|$$

and

$$|s(\hat{a}) - s(\tilde{a})| \leq |\hat{a} - \tilde{a}|.$$

Proof of Proposition 5. The proof involves several steps, producing a characterization of $\frac{c(a)}{a}$.

Lemma 2 $\exists \zeta > \underline{y}$, such that $s(a) = 0$, $\forall a \in (0, \zeta]$.

Proof. Suppose that $s(a) > 0$ for $a > \underline{y}$. Pick $a_0 > \underline{y}$. For any finite $t \geq 0$, we have $a_t > \underline{y}$ and $u'(c_t) = \beta ER_{t+1}u'(c_{t+1})$. Thus

$$u'(c_0) = \beta^t ER_1 R_2 \cdots R_{t-1} R_t u'(c_t). \quad (11)$$

By Lemma 1 and Theorem 2 we have

$$c_t \geq \phi a_t > \phi \underline{y}.$$

Thus equation (11) implies that

$$u'(c_0) \leq (\phi \underline{y})^{-\gamma} (\beta ER)^t. \quad (12)$$

Thus the right hand side of equation (12) approaches 0 as t goes to infinity. A contradiction. Thus $s(\zeta) = 0$ for some $\zeta > \underline{y}$. By the monotonicity of $s(a)$, we know that $s(a) = 0$, $\forall a \in (0, \zeta]$. ■

We can now show the following:

Lemma 3 $\frac{c(a)}{a}$ is decreasing in a .

Proof. By Lemma 2 we know that $c(\underline{y}) = \underline{y}$. For $\forall a > \underline{y}$, $\frac{c(a)}{a} \leq 1 = \frac{c(\underline{y})}{\underline{y}}$. Note that $-c(a)$ is a convex function of a , since $c(a)$ is a concave function of a . For $\hat{a} > \tilde{a} > \underline{y}$, we have²⁷

$$\frac{c(\hat{a}) - c(\underline{y})}{\hat{a} - \underline{y}} \leq \frac{c(\tilde{a}) - c(\underline{y})}{\tilde{a} - \underline{y}}.$$

This implies that

$$c(\hat{a})\tilde{a} \leq c(\tilde{a})\hat{a} - [\hat{a} - \tilde{a} - (c(\hat{a}) - c(\tilde{a}))]\underline{y}. \quad (13)$$

Since $c(a)$ is Lipschitz continuous, we have

$$c(\hat{a}) - c(\tilde{a}) \leq \hat{a} - \tilde{a}. \quad (14)$$

Combining inequalities (13) and (14) we have

$$c(\hat{a})\tilde{a} \leq c(\tilde{a})\hat{a},$$

i.e.

$$\frac{c(\hat{a})}{\hat{a}} \leq \frac{c(\tilde{a})}{\tilde{a}}.$$

■

By Theorems 1 and Proposition 1 we know that $\frac{c(a)}{a} \geq \phi$. Thus we have

$$\lim_{a \rightarrow \infty} \frac{c(a)}{a} \text{ exists.}$$

Let

$$\lambda = \lim_{a \rightarrow \infty} \frac{c(a)}{a}. \quad (15)$$

Note that $\lambda \leq 1$ since $c(a) \leq a$.

The Euler equation of this problem is

$$c_t^{-\gamma} \geq \beta ER_{t+1} c_{t+1}^{-\gamma} \text{ with equality if } c_t < a_t. \quad (16)$$

Lemma 4 $\lambda \in [\phi, 1)$.

²⁷See Lemma 16 on page 113 of Royden (1988).

Proof. Suppose that $\lambda = 1$. Thus

$$\liminf_{a_t \rightarrow \infty} \frac{c(a_t)}{a_t} = \lim_{a_t \rightarrow \infty} \frac{c(a_t)}{a_t} = 1.$$

From the Euler equation (16) we have

$$c_t^{-\gamma} \geq \beta ER_{t+1} c_{t+1}^{-\gamma} \geq \beta ER_{t+1} a_{t+1}^{-\gamma}$$

since $c_{t+1} \leq a_{t+1}$ and $\gamma \geq 1$.

Thus

$$\left(\frac{c(a_t)}{a_t} \right)^{-\gamma} \geq \beta ER_{t+1} \left(R_{t+1} \left(1 - \frac{c(a_t)}{a_t} \right) + \frac{y_{t+1}}{a_t} \right)^{-\gamma}.$$

By Fatou's lemma we have

$$\begin{aligned} & \liminf_{a_t \rightarrow \infty} ER_{t+1} \left(R_{t+1} \left(1 - \frac{c(a_t)}{a_t} \right) + \frac{y_{t+1}}{a_t} \right)^{-\gamma} \\ & \geq E \liminf_{a_t \rightarrow \infty} \left[R_{t+1} \left(R_{t+1} \left(1 - \frac{c(a_t)}{a_t} \right) + \frac{y_{t+1}}{a_t} \right)^{-\gamma} \right]. \end{aligned}$$

Thus

$$\begin{aligned} 1 & = \lim_{a_t \rightarrow \infty} \left(\frac{c(a_t)}{a_t} \right)^{-\gamma} \\ & \geq \beta \lim_{a_t \rightarrow \infty} ER_{t+1} \left(R_{t+1} \left(1 - \frac{c(a_t)}{a_t} \right) + \frac{y_{t+1}}{a_t} \right)^{-\gamma} \\ & = \beta \lim_{a_t \rightarrow \infty} \inf ER_{t+1} \left(R_{t+1} \left(1 - \frac{c(a_t)}{a_t} \right) + \frac{y_{t+1}}{a_t} \right)^{-\gamma} \\ & \geq \beta E \lim_{a_t \rightarrow \infty} \inf \left[R_{t+1} \left(R_{t+1} \left(1 - \frac{c(a_t)}{a_t} \right) + \frac{y_{t+1}}{a_t} \right)^{-\gamma} \right] \\ & = \beta E \lim_{a_t \rightarrow \infty} \left[R_{t+1} \left(R_{t+1} \left(1 - \frac{c(a_t)}{a_t} \right) + \frac{y_{t+1}}{a_t} \right)^{-\gamma} \right] \\ & = \infty. \end{aligned}$$

A contradiction. ■

From Lemma 4 we know that $c_t < a_t$ when a_t is large enough. Thus the equality of the Euler equation holds

$$c_t^{-\gamma} = \beta ER_{t+1} c_{t+1}^{-\gamma}.$$

Thus

$$\left(\frac{c_t}{a_t}\right)^{-\gamma} = \beta ER_{t+1} \left(\frac{c_{t+1}}{a_t}\right)^{-\gamma}. \quad (17)$$

Taking limits on both sides of equation (17) we have

$$\lim_{a_t \rightarrow \infty} \left(\frac{c_t}{a_t}\right)^{-\gamma} = \beta \lim_{a_t \rightarrow \infty} ER_{t+1} \left(\frac{c_{t+1}}{a_t}\right)^{-\gamma}.$$

Thus

$$\lambda^{-\gamma} = \beta \lim_{a_t \rightarrow \infty} ER_{t+1} \left(\frac{c_{t+1}}{a_t}\right)^{-\gamma}. \quad (18)$$

We turn to the computation of $\lim_{a_t \rightarrow \infty} ER_{t+1} \left(\frac{c_{t+1}}{a_t}\right)^{-\gamma}$.

In order to compute $\lim_{a_t \rightarrow \infty} ER_{t+1} \left(\frac{c_{t+1}}{a_t}\right)^{-\gamma}$, we first show a lemma.

Lemma 5 *For $\forall H > 0, \exists J > 0$, such that $a_{t+1} > H$ for $a_t > J$. Here J does not depend on realizations of R_{t+1} and y_{t+1} .*

Proof. Note that

$$\frac{a_{t+1}}{a_t} = \frac{R_{t+1}(a_t - c_t) + y_{t+1}}{a_t} \geq R_{t+1} \left(1 - \frac{c_t}{a_t}\right).$$

From equation (15) we know that for some $\varepsilon > 0, \exists J_1 > 0$, such that

$$\frac{c_t}{a_t} < \lambda + \varepsilon$$

for $a_t > J_1$. Thus

$$\frac{a_{t+1}}{a_t} \geq R_{t+1} \left(1 - \frac{c_t}{a_t}\right) \geq R_{t+1}(1 - \lambda - \varepsilon). \quad (19)$$

And

$$\frac{a_{t+1}}{a_t} \geq R_{t+1}(1 - \lambda - \varepsilon) \geq \underline{\mathbb{R}}(1 - \lambda - \varepsilon).$$

We pick $J > J_1$ such that $\underline{\mathbb{R}}(1 - \lambda - \varepsilon) \geq \frac{H}{J}$. Thus for $a_t > J$, we have

$$\frac{a_{t+1}}{a_t} \geq \frac{H}{J}.$$

This implies that

$$a_{t+1} \geq \frac{H}{J} a_t > H.$$

■

From equation (15) we know that for some $\eta > 0$, $\exists H > 0$, such that

$$\frac{c_{t+1}}{a_{t+1}} > \lambda - \eta \quad (20)$$

for $a_{t+1} > H$.

From Lemma 5 and equations (19) and (20) we have

$$R_{t+1} \left(\frac{c_{t+1}}{a_t} \right)^{-\gamma} = R_{t+1} \left(\frac{c_{t+1} a_{t+1}}{a_{t+1} a_t} \right)^{-\gamma} \leq (\lambda - \eta)^{-\gamma} (1 - \lambda - \varepsilon)^{-\gamma} R_{t+1}^{1-\gamma}$$

for $a_t > J$. And

$$(\lambda - \eta)^{-\gamma} (1 - \lambda - \varepsilon)^{-\gamma} ER_{t+1}^{1-\gamma} < \infty$$

since $\gamma \geq 1$. Thus when a_t is large enough, $(\lambda - \eta)^{-\gamma} (1 - \lambda - \varepsilon)^{-\gamma} R_{t+1}^{1-\gamma}$ is a dominant function of $R_{t+1} \left(\frac{c_{t+1}}{a_t} \right)^{-\gamma}$.

Note that

$$\lim_{a_t \rightarrow \infty} \frac{c_{t+1}}{a_{t+1}} = \lim_{a_t \rightarrow \infty} \frac{c(a_{t+1})}{a_{t+1}} = \lambda \quad a.s.$$

by Lemma 5 and equation (15). And

$$\lim_{a_t \rightarrow \infty} \frac{a_{t+1}}{a_t} = \lim_{a_t \rightarrow \infty} \left(\frac{R_{t+1}(a_t - c_t) + y_{t+1}}{a_t} \right) = R_{t+1}(1 - \lambda) \quad a.s.$$

since $y_{t+1} \in [\underline{y}, \bar{y}]$. Thus

$$\lim_{a_t \rightarrow \infty} \frac{c_{t+1}}{a_t} = \lim_{a_t \rightarrow \infty} \frac{c_{t+1} a_{t+1}}{a_{t+1} a_t} = \lambda(1 - \lambda)R_{t+1} \quad a.s.$$

Thus by the Dominated Convergence Theorem, we have

$$\lim_{a_t \rightarrow \infty} ER_{t+1} \left(\frac{c_{t+1}}{a_t} \right)^{-\gamma} = ER_{t+1} \left(\lim_{a_t \rightarrow \infty} \frac{c_{t+1}}{a_t} \right)^{-\gamma} = \lambda^{-\gamma} (1 - \lambda)^{-\gamma} ER_{t+1}^{1-\gamma}. \quad (21)$$

Combining equations (18) and (21) we have

$$\lambda^{-\gamma} = \beta \lambda^{-\gamma} (1 - \lambda)^{-\gamma} ER_{t+1}^{1-\gamma}. \quad (22)$$

By Lemma 4 we know that $\lambda \geq \phi > 0$. Thus we find λ from equation (22)

$$\lambda = 1 - (\beta ER^{1-\gamma})^{\frac{1}{\gamma}}.$$

Thus $\lambda = \phi$.

Proof of Theorem 3. The proof requires several steps.

Lemma 6 *The wealth accumulation process (a_t) is ψ -irreducible.*

Proof. First we show that the process (a_t) is φ -irreducible. We construct a measure φ on $[\underline{y}, \infty)$ such that

$$\varphi(A) = \int_A f(y)dy.$$

where $f(y)$ is the density of labor earnings $\{y\}$. Note that the borrowing constraint binds in finite time with a positive probability for $\forall a_0 \in [\underline{y}, \infty)$. Suppose not. For any finite $t \geq 0$, we have $a_t > \underline{y}$ and $u'(c_t) = \beta ER_{t+1}u'(c_{t+1})$. Following the same procedure as in the proof of Lemma 2, we obtain a contradiction. If the borrowing constraint binds at period t , then $a_{t+1} = y_{t+1}$. Thus any set A such that $\int_A f(y)dy > 0$ can be reached in finite time with a positive probability. The process (a_t) is φ -irreducible.

By Proposition 4.2.2 in Meyn and Tweedie (2009), there exists a probability measure ψ on $[\underline{y}, \infty)$ such that the process $\{a_{t+1}\}_{t=0}^{\infty}$ is ψ -irreducible, since it is φ -irreducible. ■

Lemma 7 *$a = \underline{y}$ is a reflecting barrier of the process (a_t) .*

Proof. If $a_t = \underline{y}$, then there exists \hat{y} close to \bar{y} such that $\Pr(a_{t+1} \in [\hat{y}, \bar{y}] | a_t = \underline{y}) = \Pr(y_{t+1} \in [\hat{y}, \bar{y}]) > 0$, since $s(\underline{y}) = 0$. To show that a_{t+2} can be greater than \bar{y} with a positive probability, it is sufficient to show that $s(\bar{y}) > 0$. Suppose that $s(\bar{y}) = 0$. Thus $s(a) = 0$ for $a \in [\underline{y}, \bar{y}]$. Thus by the Euler equation we have

$$(\bar{y})^{-\gamma} \geq \beta E [R_t (y_t)^{-\gamma}].$$

This is impossible under Assumption 1.i). Thus $s(\bar{y}) > 0$ and $a = \underline{y}$ is a reflecting barrier of the process $\{a_{t+1}\}_{t=0}^{\infty}$. ■

To show that there exists a unique stationary wealth distribution, we have to show that the process (a_t) is ergodic. Actually, we can show that it is geometrically ergodic.

Lemma 8 *The process $\{a_{t+1}\}_{t=0}^{\infty}$ is geometrically ergodic.*

Proof. To show that the process (a_t) is geometrically ergodic, we use part (iii) of Theorem 15.0.1 of Meyn and Tweedie (2009). We need to verify that

a the process $\{a_{t+1}\}_{t=0}^{\infty}$ is ψ -irreducible;

b the process $\{a_{t+1}\}_{t=0}^{\infty}$ is aperiodic;²⁸ and

²⁸For the definition of aperiodic, see page 114 of Meyn and Tweedie (2009).

c there exists a petite set C ,²⁹ constants $b < \infty$, $\rho > 0$ and a function $V \geq 1$ finite at some point in $[\underline{y}, \infty)$ satisfying

$$EV(a_{t+1}) - V(a_t) \leq -\rho V(a_t) + bI_C(a_t), \quad \forall a_t \in [\underline{y}, \infty).$$

By Lemma 6, the process $\{a_{t+1}\}_{t=0}^\infty$ 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; see Meyn and Tweedie (2009, p. 114). We construct a measure v_1 on $[\underline{y}, \infty)$ such that

$$v_1(A) = \int_A f(y) dy.$$

By Lemma 2, we know that $s(a) = 0$, $\forall a \in [\underline{y}, \zeta]$. Thus $[\underline{y}, \zeta]$ is v_1 -small and $v_1([\underline{y}, \zeta]) = \int_{\underline{y}}^{\zeta} f(y) dy > 0$. The process (a_t) is strongly aperiodic.

We now show that an interval $[\underline{y}, B]$ is a petite set for $\forall B > \underline{y}$. To show this, we first show that $\underline{R}s(a) + \underline{y} < a$ for $a \in (\underline{y}, \infty)$. For $s(a) = 0$, this is obviously true. For $s(a) > 0$, suppose that $\underline{R}s(a) + \underline{y} \geq a$, we have

$$u'(c(a)) = \beta ER_t u'(c(R_t s(a) + y)) \leq \beta ER_t u'(c(a)).$$

We obtain a contradiction since Assumption 1.iv) implies that $\beta ER_t < 1$. Also by Lemma 2, there exists an interval $[\underline{y}, \zeta]$, such that $s(a) = 0$, $\forall a \in [\underline{y}, \zeta]$. For an interval $[\underline{y}, B]$, $\forall a_0 \in [\underline{y}, B]$, there exists a common t such that the borrowing constraint binds at period t with a positive probability. Then for any set $A \subset [\underline{y}, \bar{y}]$, $\Pr(a_{t+1} \in A | s(a_t) = 0) = \int_A f(y) dy$. Note that a t -step probability transition kernel is the probability transition kernel of a specific sampled chain. Thus we construct a measure v_a on $[\underline{y}, \infty)$ such that v_a has a positive measure on $[\underline{y}, \bar{y}]$ and $v_a((\bar{y}, \infty)) = 0$. The t -step probability transition kernel of a process starting from $\forall a_0 \in [\underline{y}, B]$ is greater than the measure v_a . An interval $[\underline{y}, B]$ is a petite set for $\forall B > \underline{y}$.

We pick a function $V(a) = a + 1$, $\forall a \in [\underline{y}, \infty)$. Thus $V(a) > 1$ for $a \in [\underline{y}, \infty)$. Pick $0 < q < 1 - \mu ER_t$. Let $\rho = 1 - \mu ER_t - q > 0$ and $b = 1 - \mu ER_t + Ey$. Pick $B > \underline{y}$, such that $B + 1 \geq \frac{b}{q}$. Let $C = [\underline{y}, B]$. Thus C is a petite set. Therefore, for $\forall a_t \in [\underline{y}, \infty)$, we have

$$\begin{aligned} EV(a_{t+1}) - V(a_t) &= E(a_{t+1}) - a_t \\ &\leq -(1 - \mu ER_t)V(a_t) + 1 - \mu ER_t + Ey \\ &\leq -\rho V(a_t) + bI_C(a_t) \end{aligned}$$

where $I_C(\cdot)$ is an indicator function.

By Theorem 15.0.1 of Meyn and Tweedie (2009) the process (a_t) is geometrically ergodic. ■

²⁹For the definition of petite sets, see page 117 of Meyn and Tweedie (2009).

³⁰For the definition of small sets, see page 102 of Meyn and Tweedie (2009).