The distribution of wealth
and fiscal policy in economies with finitely lived agents[*]

Jess Benhabib
Department of Economics
New York University

Alberto Bisin
Department of Economics
New York University

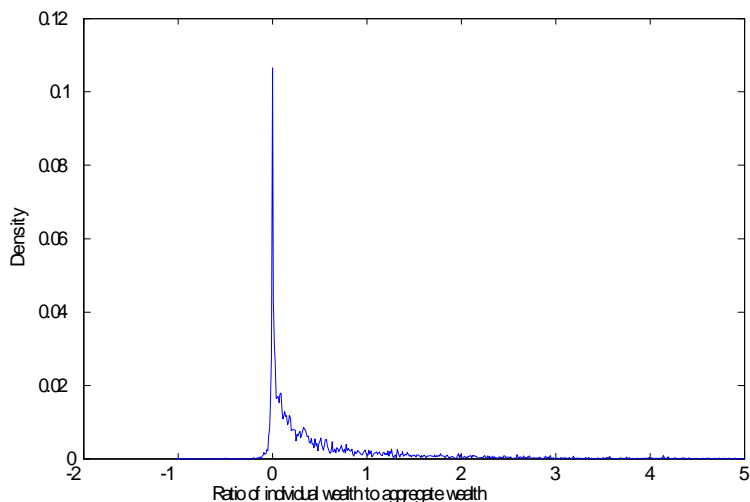This draft: December 2008

Abstract

We study the dynamics of the distribution of wealth in an overlapping generation economy with finitely lived agents and inter-generational transmission of wealth. Financial markets are incomplete, exposing agents to both labor income and capital income risk. We show that the stationary wealth distribution is a Pareto distribution in the right tail and that it is capital income risk, rather than labor income, that drives the properties of the right tail of the wealth distribution. We also study analytically the dependence of the distribution of wealth, of wealth inequality in particular, on various fiscal policy instruments like capital income taxes and estate taxes. We show that capital income and estate taxes can significantly reduce wealth inequality. Finally, we study the effects of different degrees of social mobility on the wealth distribution.

1   Introduction

Rather invariably across a large cross-section of countries and time periods income and wealth distributions are skewed to the right[1] and display heavy upper tails,[2] that is, slowly declining top wealth shares. The top 1% of the richest households in the U.S. hold over 33% of wealth[3] and the top end of the wealth distribution obeys a Pareto law, the standard statistical model for heavy upper tails.[4] The Figure below shows the distribution of wealth in the U.S. based on data from the *Survey of Consumer Finances* in 2004:

Which characteristics of the wealth accumulation process are responsible for these stylized facts? In a dynamic overlapping-generation economy with finitely lived agents we study the relationship between wealth inequality and the deep structural parameters of the economy, including fiscal policy parameters. We aim at understanding first of

---

[1] Atkinson (2002), Moriguchi-Saez (2005), Piketty (2001), Piketty-Saez (2003), and Saez-Veall (2003) document skewed distributions of income with relatively large top shares consistently over the last century, respectively, in the U.K., Japan, France, the U.S., and Canada. Large top wealth shares in the U.S. since the 60's are also documented e.g., by Wolff (1987, 2004).

[2] Heavy upper tails (power law behavior) for the distributions of income and wealth are also well documented, for example by Nirei-Souma (2004) for income in 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.

[3] See Wolff (2004). While income and wealth are correlated and have qualitatively similar distributions, wealth tends to be more concentrated than income. For instance the Gini coefficient of the distribution of wealth in the U.S. in 1992 is .78, while it is only .57 for the distribution of income (Diaz Gimenez-Quadrini-Rios Rull, 1997); see also Feenberg-Poterba (2000).

[4] 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 with an average exponent of 1.49.

all heavy upper tails, as they represent one of the main empirical features of wealth inequality.[5]

Stochastic labor incomes can in principle generate some skewness in the distribution of wealth, especially if the labor income is itself skewed and persistent. A lucky streak of high labor incomes throughout life or across generations can allow the wealth of some families to grow large. A large literature on incomplete markets studies in fact models with agents that face uninsurable idiosyncratic labor income (typically referred to as *Bewley models*). Yet the standard Bewley models of Aiyagari (1994) and Huggett (1993) produce low Gini coefficients and cannot generate heavy tails in wealth. The reason is that precautionary savings tapers off too quickly with high wealth. In order to generate skewness with heavy tails in wealth distribution, a number of authors have therefore introduced new features like e.g., heterogeneity of entrepreneurial talent (Quadrini (1999, 2000), Cagetti and De Nardi, 2006).[6]

In fact, capital income risk is a significant component of the lifetime income uncertainty of individuals and households, in addition to labor income risk. Owner-occupied housing prices have a large idiosyncratic component (see Flavin and Yamashita, 2002) and so do private equity holdings of entrepreneurs (Moskowitz and Vissing-Jorgensen, 2002).[7] In this paper we introduce uninsurable idiosyncratic shocks to capital income in addition to labor income in a overlapping generation economy where agents are finitely lived and have a "joy of giving" bequest motive.[8] Capital income risk, by inducing idiosyncratic returns to wealth accumulation, also generates lucky streaks. But the multiplicative nature of rates of return on wealth generates skewness and thick tails without relying e.g., on non-homogeneous bequest functions or other forms of heterogeneity across agents.

---

[5]A related question in the mathematics of stochastic processes and in statistical physics asks which stochastic difference equations produce stationary distributions which are Pareto; see e.g., Sornette (2000) for a survey. For early applications to the distribution of wealth see e.g., Champernowne (1953), Rutherford (1955) and Wold-Whittle (1957). For the recent econo-physics literature on the subject, see e.g., Mantegna-Stanley (2000). The stochastic processes which generate Pareto distributions in this whole literature are exogenous, that is, they are not the result of agents' optimal consumption-savings decisions. This is problematic, as e.g., the dependence of the distribution of wealth on fiscal policy in the context of these models would necessarily disregard the effects of policy on the agents' consumption-saving decisions.

[6]See also Krusell and Smith (1998) for heterogenous discount rates, Castaneda, Gimenez and Rios-Rull (2003) for life cycle features with social security and progressive taxes, De Nardi (2004) for non-homogeneous bequest functions, Becker and Tomes (1979) for heterogeneous propensities to save correlated with propensities to bequeath.

[7]The sum of primary residence (home), private business equity, and investment real estate account for more than half the total asset of U.S. households in 1998 (Bertaut and Starr-McCluer (2002) from the *Survey of Consumer Finances).* See Angeletos (2007) and Benhabib and Zhu (2008) for more evidence on the macroeconomic relevance of idiosyncratic capital income risk.

[8]Angeletos (2007) also studies an economy with idiosyncratic capital income risk, though with a focus on aggregate savings. In McKay (2008) idiosyncratic capital income risk is derived endogenously from agents' optimal search for asset in financial markets.

More specifically, in our economy the stationary wealth distribution is a Pareto distribution in the right tail.[9] Furthermore, consistently with the difficulties of Bewley models in capturing heavy tails, we analytically show that it is capital income risk, rather than labor income, that drives the properties of the right tail of the wealth distribution. By means of comparative statics exercises we study the dependence of the distribution of wealth, of wealth inequality in particular, on various fiscal policy instruments like capital income taxes and estate taxes. We show that capital income and estate taxes reduce wealth inequality. This is in contrast with the conclusions of Becker and Tomes (1979).[10]In Becker and Tomes (1979), in fact, because parents can adjust their bequests, capital and estate taxes have little, if any, impact of wealth inequality. In our model capital and estate taxes have instead an effect on the wealth distribution because they dampen the lucky streaks of persistent high realizations in the rates of return on wealth. We show by means of simulation that this effect is potentially very strong. Finally, we study the effects of different degrees of social mobility on wealth distribution.

Section 2 introduces our model and derives the solution to the individual agent's optimization problem. Section 3 gives the characterization of the stationary wealth distribution as a Pareto law and a discussion of the assumptions underlying the result. In Section 4 our results for the effects of capital income and estate taxes on wealth inequality are stated. Section 4 also reports on comparative statics for the bequest motive, the volatility of returns, and the degree of social mobility as measured by the correlation of rates of returns on capital across generations. Most proofs and several technical details are buried in Appendices A-C.

2   An OLG economy with bequests

Consider the overlapping generation (OLG) economy in Yaari (1965) and Blanchard (1985).[11] But consider the case in which each agent lives T periods.[12] We assume that both the rate of return on wealth as well earnings are stochastic across agents

---

[9]In Benhabib and Bisin (2006) we instead obtain a Pareto distribution of wealth in an economy with no income shocks and a constant probability of death.

[10]Similarly to Becker and Tomes (1979), Castaneda, Diaz Jimenez, and Rios Rull (2003) and Cagetti-De Nardi (2007) find, in calibrated models, small effects of estate taxes on the wealth distribution.

[11]More specifically, we consider the formulation with endogenous bequests in Yaari (1965). Bequests however do not fully reflect the intergenerational transfers, in particular inter-vivos transfers, that play an important role in the aggregate capital accumulation. Kotlikoff and Summers (1981) find that intergenerational transfers account for sustaining the vast majority, up to 70%, of the aggregate U.S. capital formation. See also Gale and Scholtz (1994) for more moderate findings on this topic, estimating intergenerational transfers at 50%. For an account of the the role of inhertance for the Forbes 400 see Elwood et al. (1997) and Burris (2000).

[12]In Blanchard (1985) and Yaari (1965) each agent at time t has instead a constant probability of death, p. This assumption about the demographic structure is often referred to as *perpetual youth*. We study the dynamics of the wealth distribution in this case in a companion paper, Benhabib-Bisin (2008).

and generations in the economy. We shall impose however the rate of return that an agent earns during his lifetime is constant, and that his income at birth, drawn from a distribution, grows deterministically until his death at age T. We first describe then the deterministic problem of a generic agent born at time s.

An agent born at time s, besides inheriting initial wealth $w(s,s)$, also receives labor income $y(t)$ during $t \in [s, T]$. We assume that there is no stochastic component affecting wealth accumulation once the rate of return on wealth r is drawn at birth, and that each agent's initial income endowment, also drawn from a distribution at birth, grows at a deterministic rate $g: \ y(t) = ye^{g(t-s)}$. The rate of return on wealth r, is net of the capital income tax ζ, imposed for simplicity on wealth holdings. Let $c(s,t)$ and $w(s,t)$ denote, respectively, consumption and wealth at t of an agent born at s.

An agent born at time s dies, at $s + T$, with wealth $w(s, s + T)$. Each agent has a single child, born in the economy at the agent's death. Let b denote the estate tax. At birth, each child inherits $w(s,s) = (1 - b)w(s - T, s)$ from his parent. Each agent's momentary utility function, $u(c(s,t))$, satisfies the standard monotonicity and concavity assumptions. Agents also have a preference for leaving bequests to their children. In particular, we assume "joy of giving" preferences for bequests: the parent's utility from bequests is $\chi\phi((1 - b)w(s, s + T))$, where ϕ denotes an increasing bequest function.[13]

Assumption 1 *Preferences satisfy:*

$$u(c) = \frac{c^{1-\sigma}}{1 - \sigma}, \quad \phi(w) = \frac{w^{1-\sigma}}{1 - \sigma},$$

with elasticity $\sigma \geq 1$. Furthermore, we require $r \geq \rho + \sigma g$.

The condition $\sigma \geq 1$ is required to produce a stationary non-degenerate wealth distribution. It guarantees in particular that the interest rate r on wealth is larger than the endogenous rate of growth of consumption, $\frac{r-\rho}{\sigma}$.[14] The condition $r \geq \rho + \sigma g$ guarantees that agents will not borrow during their lifetime.

The maximization problem of an agent born at time s involves choosing a consumption path, $c(s,t)$, to maximize

$$U = \int_{t=s}^{s+T} e^{-\rho t} \frac{c(s,t)^{1-\sigma}}{1 - \sigma} dt + \chi e^{-\rho T} \frac{[(1 - b) w(s, s + T)]^{1-\sigma}}{1 - \sigma} \tag{1}$$

subject to

$$\dot{w}(s,t) = rw(s,t) + y(t) - c(s,t) \tag{2}$$

---

[13] Note that we assume that the argument of the parents' preferences for bequests is after-tax bequests. We also assume that parents correctly anticipate that bequests are taxed and that this accordingly reduces their "joy of giving."

[14] The assumption could be relaxed if we allowed the elasticity of substitution for consumption and bequest to differ, at a notational cost.

2.1 The optimal consumption path

Let an individual's age be denoted $\tau = t - s$. Let human capital of an agent born at s at time t, $h(s,t)$, be defined as:

$$h(s,t) = \int_t^{s+T} y(\tau)e^{-(r)\tau}d\tau$$

We aim at solving for consumption $c(s,t)$ in terms of financial and human capital $w(s,t) + h(s,t)$: We have the following characterization of consumption.

Proposition 1 *The optimal consumption path satisfies*

$$c(s,t) = m(\tau)(w(s,t) + h(s,t)),$$

where the propensity to consume out of financial and human wealth, $m(\tau)$, is independent of $w(s,t)$ and $h(s,t)$. Furthermore, $m(\tau)$ is i) decreasing in age τ; ii) decreasing in the estate tax b and in capital income tax ζ, iii) independent of b for $\sigma = 1$.

See Appendix A for the proof. The closed form solution for $m(\tau)$ is equation (12).

2.2 The dynamics of individual wealth

Let $w(\tau)$ be the wealth of an agent of age τ born with wealth $w(0)$. Substituting the optimal consumption path into (2), we can write the dynamics of individual wealth as a function of age τ, to obtain the following linear differential equation with variable coefficients:

$$\dot{w}(\tau) = \tilde{r}(\tau)w(\tau) + q(\tau)y \tag{3}$$

It has a solution of the form

$$w(\tau) = \sigma_w(r,\tau)w(0) + \sigma_y(r)y$$

The closed form solution is given in Appendix A, Proposition A.**1**.

3 The distribution of wealth

We now characterize the properties of the wealth distribution in our OLG economy with finitely lived agents, inter-generational transmission of wealth, and redistributive fiscal policy. We show that the stationary wealth distribution obeys a Pareto law in the right tail. As we already noted, the Pareto distribution is the standard statistical model for random variables displaying a thick upper tail, whose density declines as a power law.

3.1 The initial wealth of dynasties

Exploiting the solution for $w(T)$ in terms of $w(0)$, we can construct a discrete time map for each dynasty equating post-tax bequests from parents with initial wealth of children. Let $w_n = w((n-1)T, nT)$ be the initial wealth of the n'th dynasty. As noted before, a generic individual of the n'th dynasty faces constant rate of return of wealth and initial earnings over his/her lifetime. On the other hand, the rate of return of wealth and earnings are stochastic across individuals and generations; we let $(r_n)_n$ and $\left(ye^{g'n}\right)_n$ denote, respectively the stochastic process for the rate of return of wealth and initial earnings, over dynasties n.

We can then construct a stochastic difference equation for the initial wealth of dynasties, mapping w_{n-1} into w_n.[15] It is in fact convenient to work with discounted variables:

$$z_n = \left(e^{-g'}\right)^n w_n, \quad z_{n-1} = \left(e^{-g'}\right)^{n-1} w_{n-1}$$

We thus obtain the following a stochastic difference equation of the form:

$$z_{n+1} = \alpha_n z_n + \beta_n \tag{4}$$

where $(\alpha_n, \beta_n)_n = (\alpha(r_n), \beta(r_n, y_n))_n$ are stochastic processes induced by $(r_n, y_n)_n$; see Appendix B, equations (15-17) for closed form solutions of $\alpha(r_n)$ and $\beta(r_n, y_n)$. A simple form for $\alpha(r_n)$ and $\beta(r_n, y_n)$ is obtained for logarithmic preferences, $\sigma = 1$, if we also require that $\rho = \chi$. In this case

$$\alpha_n = (1-b)e^{(r_n-\rho)T-g'}$$

$$\beta_n = (1-b)\,e^{r_n T-g'}\frac{1-e^{-(r_n-g)T}\left(1+e^{\rho\tau}\left(e^{(r_n-g)(T-\tau)}-1\right)\right)}{r_n-g}y_n$$

3.1.1 The stationary distribution of initial wealth

In this section we study conditions on the stochastic process $(r_n, y_n)_n$ which guarantee that the initial wealth process defined by (4) is ergodic.[16] We then apply a theorem from Saporta (2004, 2005) to characterize the tail of the stationary distribution of initial wealth. While the tail of the stationary distribution of initial wealth is easily characterized in the special case in which $(r_n)_n$ and $(y_n)_n$ are $i.i.d.$,[17] more general stochastic processes are required for a theory of distribution of wealth. A positive auto-correlation

---

[15] See Appendix A for the derivation.

[16] We avoid as much as possible the notation required for formal definitions on probability spaces and stochastic processes. The costs in terms of precision seems overwhelmed by the gain of simplicity. Given a random variable x_n, for instance, we simply denote the associated stochastic process as $(x_n)_n$.

[17] The characterization is an application of the well-known Kesten-Goldie Theorem in this case, as α_n and β_n are i.i.d. if r_n and y_n are. See Appendix C.

of r_n and y_n is required to capture variations in social mobility in the economy, e.g., economies in which returns on wealth and labor earning abilities are in part transmitted across generations. Similarly, it is important to allow for the correlation between r_n and y_n, that is, to allow e.g., for agents with high labor income to have better opportunities for higher returns on wealth in financial markets.[18]

We proceed therefore with appropriately weaker assumptions on $(r_n, y_n)_n$.

Assumption 2 *The stochastic process $(r_n, y_n)_n$ is a real, irreducible, aperiodic, stationary Markov chain with finite state space $\bar{\mathbf{r}} \times \bar{\mathbf{y}} := \{\bar{r}_1, ..., \bar{r}_m\} \times \{\bar{y}_1, ..., \bar{y}_l\}$. Furthermore satisfies:*

$$\Pr\left(r_n, y_n \mid r_{n-1}, y_{n-1}\right) = \Pr\left(r_n, y_n \mid r_{n-1}\right),$$

where $\Pr\left(r_n, y_n \mid r_{n-1}, y_{n-1}\right)$ denotes the conditional probability of (r_n, y_n) given (r_{n-1}, y_{n-1}).

While Assumption 2 requires r_n to be independent of $(y_{n-1}, y_{n-2}...)$, it leaves the auto-correlation of $(r_n)_n$ unrestricted, in the space of Markov chains.[19] Also, Assumption 2 allows for (a restricted form of) auto-correlation of $(y_n)_n$ and for the correlation of y_n and r_n. This assumption would be satisfied, for instance, if a single Markov process, corresponding e.g., to productivity shocks, drove returns on capital $(r_n)_n$, as well as labor income $(y_n)_n$. Following Roitershtein (2007), we say that a stochastic process $(r_n, y_n)_n$ which satisfies Assumption 2 is a *Markov Modulated chain*.[20]

Recall the stochastic difference equation $z_{n+1} = \alpha_n z_n + \beta_n$, where z_n is the discounted initial wealth of generation n. The multiplicative component α_n can be interpreted as the effective lifetime rate of return on initial wealth from one generation to the next, after subtracting the fraction of lifetime wealth consumed, and before adding effective lifetime earnings, netted for the affine component of lifetime consumption.[21] The additive component β_n can in turn be interpreted as a measure of effective lifetime labor income (again after subtracting the affine part of consumption). Intuitively then, to induce a limit stationary distribution of $(z_n)_n$ it is required that the contractive and expansive components of the effective rate of return tend to balance, i.e., that the distribution of α_n display enough mass on $\alpha_n < 1$ as well some as on $\alpha_n > 1$; and that effective earnings β_n be positive and bounded, hence acting as a reflecting barrier. We say that a stochastic process $(\alpha_n, \beta_n)_n$ which statisfies this properties is *reflective*. We hence

---

[18]See Arrow (1987) and McKay (2008) for models in which such correlations arise endogenously from non-homogeneous portolio choices in financial markets.

[19]In fact the restriction to Markov chains is just for convenience. See Roitersthein (2007) for a related analysis applicable to general Markov processes.

[20]For the use of Markov Modulated chains, see Saporta (2005) in her remarks following Theorem 2, or Saporta (2004), section 2.9, p.80. See instead Roitersthein (2007) for general Markov Modulated processes.

[21]A realization of $\alpha_n = \alpha(r_n) < 1$ should not, however, be interpreted as a negative return in the conventional sense. At any instant the rate of return on wealth for an agent is a realization of $r_n > 0$, that is, positive.

impose assumptions on $(r_n, y_n)_n$ which guarantee that the induced process $(\alpha_n, \beta_n)_n = (\alpha(r_n), \beta(r_n, y_n))_n$ is reflective.

Let P denote the transition matrix of $(r_n)_n$. Let $\boldsymbol{\alpha}(\bar{\mathbf{r}})$ denote the state space of $(\alpha_n)_n$ as induced by the map $\alpha(r_n)$.[22]

Assumption 3 $\bar{\mathbf{r}} = \{\bar{r}_1, ..., \bar{r}_m\}$, $\bar{\mathbf{y}} = \{\bar{y}_1, ..., \bar{y}_l\}$ *and P are such that: (i) $\bar{r}_i, \bar{y}_j > 0$, for $i = 1, ...m$ and $j = 1, ...l$, (ii) $P\boldsymbol{\alpha}(\bar{\mathbf{r}}) < \mathbf{1}$, (iii) $\exists \bar{r}_i$ such that $\alpha(\bar{r}_i) > 1$, (iv) the elements of the trace of the transition matrix P are positive, that is $P_{ii} > 0$, for any i.*

Let the shorthand $\bar{\boldsymbol{\alpha}} = \{\alpha(\bar{r}_1), ..., \alpha(\bar{r}_m)\} = \{\bar{\alpha}_1, ...\bar{\alpha}_m\}$ denote the induced state space of $(\alpha_n)_n$ and $\bar{\boldsymbol{\beta}} = \{\bar{\beta}_1, ...\bar{\beta}_{lm}\}$ the state space of $(\beta_n)_n$.

Proposition 2 *Assumptions 2 on $(r_n, y_n)_n$ imply that $(\alpha_n, \beta_n)_n$ is a Markov Modulated chain. Furthermore, Assumption 3 implies that $(\alpha_n, \beta_n)_n$ is reflective, that is, it satisfies: (i) $(\alpha_n, \beta_n)_n$ is > 0, (ii) $P\bar{\boldsymbol{\alpha}} < \mathbf{1}$, (iii) $\bar{\alpha}_i > 1$ for some $i = 1, ...m$, (iv) the diagonal elements of the transition matrix P of α_n are positive.*

Let A be the diagonal matrix with elements $A_{ii} = \bar{\alpha}_i$, and $A_{ij} = 0, j \neq i$. Having established that $(\alpha_n, \beta_n)_n$ is a reflective process, we can prove the following Proposition, based on a Theorem by Saporta (2005).

Theorem 1 *(Saporta (2005),Thm 1).Consider*

$$z_{n+1} = \alpha_n z_n + \beta_n, \qquad z_0 > 0$$

Let $(\alpha_n, \beta_n)_n$ be a reflective and regular (as in Assumption A.1, Appendix B in Section 7.1) Markov Modulated process.[23] Then the tail of the stationary distribution of z_n, $P_{.>}(z_n > z)$, is asymptotic to a Pareto law

$$\Pr{}_{>}(z_n > z) \sim cz^{-\mu}$$

where $\mu > 1$ satisfies
$$\lambda(A^\mu P') = 1$$

and where $\lambda(A^\mu P')$ is the dominant root of $A^\mu P'$.

---

[22] In the proof of Proposition 2, Appendix B, we show that the state space of $(\alpha_n, \beta_n)_n$ is well defined. Note also that, by Assumption 2, $(r_n)_n$ converges to a stationary distribution and hence $(\alpha(r_n))_n$ also converges to a stationary distribution.

[23] The conditions required on the state-space of the process $(r_n, y_n)_n$ to guarantee that $(\alpha_n, \beta_n)_n$ is regular are innocuous and hold generically. We impose them throughout. See Appendix B for technical details.

Proof. The Proposition follows from Saporta (2005), Theorem 1, if we show i) that there exists a μ that solves $\lambda\left(A^{\mu}P'\right)=1$, and that ii) such μ is >1. Saporta shows that $\mu=0$ is a solution to $\lambda\left(A^{\mu}P'\right)=1$, or equivalently to $\ln\left(\lambda\left(A^{\mu}P'\right)\right)=0$. This follows from $A^0=I$ and P being a stochastic matrix. Let $E\alpha\left(r\right)$ denote the expected value of α_n at its stationary distribution (which exists as it is implied by the ergodicity of $\left(r_n\right)_n$, in turn a consequence of Assumption 2). Saporta, under the assumption $E\alpha\left(r\right)<1$, shows that $\frac{d\ln\lambda\left(A^{\mu}P'\right)}{\delta\mu}<0$ at $\mu=0$, and that $ln\left(\lambda\left(A^{\mu}P'\right)\right)$ is a convex function of μ.[24] Therefore, if there exists another solution $\mu>0$ for $\ln\left(\lambda\left(A^{\mu}P'\right)\right)=0$, it is positive and unique.

To assure that $\mu>1$ we replace the condition $E\alpha\left(r\right)<1$ with (ii) of Proposition 3, $P\bar{\boldsymbol{\alpha}}<\mathbf{1}$. This implies that the column sums of AP' are <1. Since AP' is positive and irreducible, its dominant root is smaller than the maximum column sum. Therefore for $\mu=1$, $\lambda\left(A^{\mu}P'\right)=\lambda\left(AP'\right)<1$. Now note that if $\left(\alpha_n,\beta_n\right)_n$ is reflective, by Proposition 2, $P_{ii}>0$ and $\overline{\alpha}_i>1$, for some i. This implies that the trace of $A^{\mu}P'$ goes to infinity if μ does (see also Saporta (2004) Proposition 2.7). But the trace is the sum of the roots so the dominant root of $A^{\mu}P'$, $\lambda\left(A^{\mu}P'\right)$, goes to infinity with μ. It follows that for the solution of $\ln\left(\lambda\left(A^{\mu}P'\right)\right)=0$, we must have $\mu>1$. This proves ii). ∎

The matrix AP' has the property that the $i'th$ column sum gives the expected value of α_n conditional on $\alpha_{n-1}=\overline{\alpha}_i$. When $\left(\alpha_n\right)_n$ is $i.i.d.$, P has identical rows, so transition probabilities do not depend on the state α_i. In this case $A^{\mu}P'$ has identical column sums given by $E\alpha^{\mu}$ and equal to $\lambda\left(A^{\mu}P'\right)$.

Remark 1 *The analysis of this section holds more generally, when $\left(\alpha_n,\beta_n\right)_n$ is not restricted to be a finite Markov chain. For general Markov processes, an appropriate definition of Markov Modulated processes, as well as of the regularity and reflectivity conditions, allows us to apply Theorem 1 in Roitershtein (2007) to show that the tail of the distribution is asymptotic to a Pareto law*

$$\Pr{}_{>}(z_n>z)\sim cz^{-\mu}$$

where $\mu>1$ satisfies

$$\lim_{N\to\infty}\left(E\prod_{n=0}^{N-1}\left(\alpha_{-n}\right)^{\mu}\right)^{\frac{1}{N}}=1.^{25} \tag{5}$$

In fact, Saporta (2005, Proposition 1, section 4.1) establishes that, for finite Markov chains, $\lim_{N\to\infty}\left(E\prod_{n=0}^{N-1}\left(\alpha_{-n}\right)^{\mu}\right)^{\frac{1}{N}}=\lambda\left(A^{\mu}P'\right)$. *Note also that, when $\left(\alpha_n\right)_n$ is i.i.d., condition*

---

[24] This follows because $\lim_{n\to\infty}\frac{1}{n}\ln E\left(\alpha_0\alpha_{-1}...\alpha_{-n-1}\right)^{\mu}=\ln\left(\lambda\left(A^{\mu}P'\right)\right)$ (see Remark 1 below) and the log-convexity of the moments of non-negative random variables (see Loeve(1977), p. 158).

[25] Of course the term $\prod_{n=0}^{N-1}\alpha_{-n}$ arises from using repeated substitions for z_n. See Brandt (1986) for general conditions to obtain an ergodic solution for stationary stochastic processes satisfying (4).

(5) reduces to $E\left(\alpha\right)^{\mu} = 1$, a result established by Kesten (1973) and Goldie (1991); see Appendix C.

We now turn to the characterization of the stationary wealth distribution.

3.2 The stationary wealth distribution

We have shown that the stationary distribution of initial wealth in our economy has a power tail. The stationary wealth distribution can be constructed as follows.

By Proposition A. 1 in Appendix A, the wealth at age τ of an agent born with wealth $z\left(0\right)$, return r, and income y is

$$z\left(\tau\right) = \sigma_w(r,\tau)z\left(0\right) + \sigma_y(r)y \tag{6}$$

Note that (6) is a deterministic map, as we assumed that r and y are fixed for any agent during his/her lifetime. If we denote with $f_0(z(0))$ the density of the stationary distribution of initial wealth, therefore, the density of the stationary distribution of wealth of agent of age τ, $f_\tau(z(\tau))$ is obtained simply by a change of variable through the map (6).[26]

The density of the distribution of wealth z in the population is then

$$f(z) = \int_0^T f_\tau(z)d\tau$$

But the asymptotic power law property with the same power μ for each age is preserved under integration. We can then conclude the following:

Proposition 3 *Suppose the tail of the limiting distribution of initial wealth z_n is asymptotic to a Pareto law, $P_{.>}(z_n > z) \sim cz^{-\mu}$, then the tail of the distribution of wealth in the population is also asymptotic to a Pareto law, with the same exponent μ.*

---

[26]The change of variable implies:

$$f_\tau(z(\tau)) = f_0(\frac{z(\tau) - \sigma_y(r)y}{\sigma_w(r,\tau)})\frac{1}{\sigma_w(r,\tau)}$$

Recall that $(r_n, y_n)_n$ is bounded and therefore so are $\sigma_w(r,\tau)$ and $\sigma_y(r)y$. Therefore, $f_\tau(z(\tau))$ is asymptotic to a Pareto law with tail μ if $f_0(z(0))$ is.

4   Wealth inequality: some comparative statics

We study in this section the tail of the stationary wealth distribution as a function of preference parameters and fiscal policies. In particular, we study stationary wealth inequality as measured by the Gini coefficient of the tail of the distribution of wealth. Theorem 1 allows us to solve for the exponent of the Pareto distribution, μ, which characterizes the tail of the distribution of normalized wealth. But, for a Pareto distribution, the exponent μ is inversely linked to the Gini coefficient G of the distribution, a measure of its inequality:[27]

$$G = \frac{1}{2\mu - 1}. \tag{7}$$

However, the effects of the structural parameters and of fiscal policy on wealth inequality depend in turn on the auto-correlation of returns and earnings, a measure of social mobility. We distinguish two different cases in our analysis. First consider the case in which (α_n, β_n) are independent over generations (not auto-correlated). Independence implies that the children of generations which experienced high (resp. low) effective returns on wealth and/or high (resp. low) effective earnings will not, on average, experience high (resp. low) returns on wealth and/or high (resp. low) earnings. We refer to this case, abusing words somewhat, as the case of *perfect social mobility.*[28] The second case we study is naturally one in which (α_n, β_n) are positively auto-correlated, so that social mobility is reduced. We refer to this case as the case of *moderate social mobility.* We may in fact use the auto-correlation in the stochastic process of (α_n, β_n) as an inverse measure of social mobility.

We have therefore the tools to study how wealth inequality, as measured by the Gini coefficient of the tail of the stationary distribution of wealth G, depends on the structural parameters of our economy and on fiscal policies. First, we shall study how different compositions of capital and labor income risk affect wealth inequality. Second, we will study the effects of preferences, in particular the intensity of the bequest motive. Third, we will characterize the effects of both capital income and estate taxes on wealth inequality. Finally, we will address the relationship between social mobility and wealth inequality.

4.1   Capital and labor income risk

Theorem 1 characterizes the tail of the wealth distribution when the process $(\alpha_n, \beta_n)_n$ is *reflective.* It follows from the Theorem that, as long as $(\alpha_n, \beta_n)_n$ is *reflective,* the stochastic properties of labor income risk, $(\beta_n)_n$, have no effect on the tail stationary

---

[27]See e.g., Chipman (1976).

[28]Words are abused because in priciple an economic environment in which (α_n, β_n) are negatively auto-correlated could represent *more* social mobility . We do not consider this case of any relevance in practice.

distribution of wealth. In fact heavy tails in the stationary distribution require that the economy has sufficient capital income risk, with some $\bar{\alpha}_i > 1$. Consider instead an economy with limited capital income risk, where $\alpha_i \leq \bar{\alpha} < 1$ for all i and where $\bar{\beta}$ is the upper bound of β_n. In this case it is straightforward to show that the stationary distribution of wealth would be bounded above by $\frac{\bar{\beta}}{1-\bar{\alpha}}$.[29]

More generally, we can also show that wealth inequality increases with the capital income risk agents face in the economy, as measured by a "mean preserving spread" on the distribution of the stochastic process of "effective return on wealth" $(\alpha_n)_n$.

Definition 1 *Define a mean preserving spread on the distribution of the stochastic process of $(\alpha_n)_n$ as any change of the state space $\bar{\boldsymbol{\alpha}}$ and/or of the transition matrix P which increases the variance of $\lim_{N \to \infty} \left(\prod_{n=0}^{N-1} \alpha_{-n} \right)^{\frac{1}{N}}$ while keeping the mean constant.*

Note that, for a finite Markov chain, the mean of the random variable $\lim_{N \to \infty} \left(\prod_{n=0}^{N-1} \alpha_{-n} \right)^{\frac{1}{N}}$ is equal to the dominant root of AP', $\lambda(AP')$. In the *i.i.d.* case, $E(\alpha) = \lambda(AP')$.

Proposition 4 *Wealth inequality, as measured by the Gini coefficient of the tail G, increases with a mean preserving spread on the distribution of the stochastic process of $(\alpha_n)_n$.*

Proof. Define the random variable $\gamma = \lim_{N \to \infty} \left(\prod_{n=0}^{N-1} \alpha_{-n} \right)^{\frac{1}{N}}$ (the limit is well-defined under Assumption 2). Since $\mu > 1$, γ^μ is a convex function, and hence $-\gamma^\mu$ a concave function in γ. Suppose we perform a mean preserving spread of the random variable γ, given by γ' (let a prime denote a variable after the spread). By second order stochastic dominance we have $E(-\gamma^\mu) > E(-(\gamma')^\mu)$ so $E(\gamma^\mu) < E((\gamma')^\mu)$ and $1 = E\gamma^\mu \leq E(\gamma')^\mu$. It follows that if μ' solves $E\left((\gamma')^{\mu'}\right) = 1$ we must have $\mu' \leq \mu$, and a higher associated Gini coefficient $G' > G$. ∎

We conclude that it is capital income risk (idiosyncratic risk on return on capital), and not labor income risk, that determines the wealth inequality of the tail of the stationary distribution given by G: the higher capital income risk, the more unequal is wealth.

4.2 The bequest motive

Wealth inequality depends on the bequest motive, as measured by the preference parameter χ. An agent with a higher preference for bequests will save more and accumulate wealth faster with a higher α_n, which in turn will lead to higher wealth inequality.

---

[29]Of course this is true a fortiori in the case where there is no capital risk and $\alpha_n = \bar{\alpha} < 1$.

Proposition 5 *Wealth inequality, as measured by the Gini coefficient of the tail G, increases with the bequest motive χ.*

Proof. From the definition of α_n,[30] we obtain $\frac{\partial \alpha_n}{\partial \chi} > 0$. Thus an infinitesimal increase in χ shifts the state space a to the right. Therefore elements of $[A^\mu P']$ increase, which implies that the dominant root $\lambda\left(A^\mu P'\right)$ increases. However we know from Saporta (2005) that $\ln\left(\lambda\left(A^\mu P'\right)\right)$ is a convex function of μ, and is increasing at the positive value of μ which solves $\ln\lambda\left(A^\mu P'\right) = 0$. Therefore to preserve $\ln\left(\lambda\left(A^\mu P'\right)\right) = 0$, μ must decline and G must increase. ∎

4.3 Fiscal policy

Fiscal policies in our economy are captured by the parameters b and ζ, representing, respectively, the estate tax and the capital income tax.[31]

Proposition 6 *Wealth inequality, as measured by the Gini coefficient of the tail G, is decreasing in the estate tax b and in the capital tax ζ.*

Proof. From (16), we have

$$\alpha_n = (1-b)e^{-g'}\frac{A(r_n)B(b)e^{r_nT}}{A(r_n)B(b) - 1 + e^{(A(r_n)T}}$$

where $A(r_n) = r_n - \frac{r_n - \rho}{\sigma}$ and $B(b) = \chi^{\frac{-1}{\sigma}}(1-b)^{\frac{1-\sigma}{\sigma}}$. Dropping the dependence on r_n and b for notational simplicity, denoting $\frac{dB}{db}$ with B', and computing $\frac{\partial \alpha_n}{\partial b}$, after tedious simplifications, we obtain:

$$sign\left(\frac{\partial \alpha_n}{\partial b}\right) < 0 \text{ if } (1-b)B' - B < 0$$

But $(1-b)B' = \frac{\sigma - 1}{\sigma}B(b)$ and under Assumption 1, $\frac{\sigma-1}{\sigma} < 1$. We conclude then that $\frac{\partial \alpha_n}{\partial b} < 0$. Now the proof is identical to the proof of Proposition 5 in the reverse direction since $\frac{\partial \alpha_n}{\partial b} < 0$ whereas $\frac{\partial \alpha_n}{\partial \chi} > 0$. ∎

Furthermore, let $\zeta(r_n)$ denote a non-linear tax on capital, such that the net rate of return of monetary wealth for generation n becomes $r_n\left(1 - \zeta(r_n)\right)$. Since $\frac{\partial \alpha_n}{\partial r} > 0$, the Corollary below follows immediately from the argument used in the proof of Proposition 6.

Corollary 1 *Wealth inequality, as measured by the Gini coefficient of the tail G, is decreasing if a non-linear tax on capital $\zeta(r_n)$ is applied which induces a relative shift of the state space.to the left.*

---

[30]See Appendix B, equation (16).

[31]Recall that the random rate of return r_n in our economy is defined net of the capital income ζ.

The results above indicate that taxes have a dampening effect on the tail of wealth distribution. This is the case despite the presence of bequest motives. Becker and Tomes (1979), on the contrary, find that tax increases ambiguous effects on wealth inequality at the stationary distribution. In their model the utility of parents depends on the expected income of children, and parents can anticipate and essentially offset any fiscal policy, dampening any wealth equalizing aspects of these policies. In our model both capital income and estate taxes dampen the effect of luck acting through the stochastic returns on capital. The effect of a streak of luck acting multiplicatively on wealth can be powerful and in fact generates the skewness and fat tails of the wealth distribution. Any dampening either of returns to wealth through capital taxes or of the transmission of wealth through estate taxes will therefore tend to flatten the heavy tails of the wealth distribution. We conclude that capital income risk, inducing a stochastic return on capital, is the main reason why our results on fiscal policies differ substantially from those of Becker and Tomes (1979). In Section 5, we will also show by means of a simple calibration that the tail of the stationary wealth distribution, hence its inequality, is in fact quite sensitive to variations in fiscal policy, both capital income taxes as well as estate taxes.

4.4   Social mobility

We consider here an example for the case in which *social mobility* is *moderate* to study how social mobility affects the wealth inequality index G. Consider the simple case in which α_n is an two-state irreducible Markov Chain on $[\alpha_l, \alpha_h]$, with $\alpha_l < 1 < \alpha_h$, and[32]

$$P = \begin{pmatrix} \rho_l & 1 - \rho_l \\ 1 - \rho_h & \rho_h \end{pmatrix}$$

In this case the ergodic distribution of α_n is $\left(\frac{1-\rho_h}{2-\rho_l-\rho_h}, \frac{1-\rho_l}{2-\rho_l-\rho_h} \right)$ .To assure that that there exists a positive μ that solves $\lambda \left(A^\mu P' \right) = 1$,we require that $E\alpha \left(r \right) < 1$, or that the mean of α_n, taken with respect to its stationary distribution is less than 1. This implies

$$(1 - \rho_h) \ln \alpha_l + (1 - \rho_l) \ln \alpha_h < 0 \tag{8}$$

The solution of μ is given by:

$$
\begin{aligned}
\lambda \left(A^\mu P' \right) &= \frac{(\alpha_l)^\mu \rho_l + (\alpha_h)^\mu \rho_h + D^{0.5}}{2} = 1 \\
D &= (\alpha_l)^{2\mu} (\rho_l)^2 + (\alpha_h)^{2\mu} (\rho_h)^2 + 2 (\alpha_l \alpha_h)^\mu (\rho_l \rho_h - 2\rho_l - 2\rho_h + 2)
\end{aligned}
$$

where D is the discriminant of the characteristic equation of $A^\mu P'$ which must be positive since $\lambda \left(A^\mu P' \right)$ is real.

---

[32]See Saporta (2004), sections 2.10.1 and 2.10.2 for an example with a two state Markov chain.

We can now do comparative statistics on μ. If for simplicity we assume $\rho_l = \rho_h = \rho$, then μ solves

$$\frac{(\alpha_l)^\mu \rho + (\alpha_h)^\mu \rho + \left((\rho)^2 \left((\alpha_l)^{2\mu} + (\alpha_h)^{2\mu}\right) + 2 (\alpha_l \alpha_h)^\mu \left(\rho^2 - 4\rho + 2\right)\right)^{0.5}}{2} = 1.$$

and can then compute:

$$\frac{d\mu}{d\rho} = \frac{-\left((\alpha_h)^u + (\alpha_l)^\mu + (\alpha_h \alpha_l)^\mu (4\rho - 8) + \rho \left(2 (\alpha_h)^{2\mu} + 2 (\alpha_l)^{2\mu}\right)\right)}{\rho \left((\ln \alpha_h) (\alpha_h)^u + (\ln \alpha_l) (\alpha_l)^\mu\right) + 2 \left[\begin{array}{c} (\ln \alpha_h \alpha_l) (\alpha_h \alpha_l)^\mu (\rho^2 - 4\rho + 2) \\ +\rho^2 \left((\ln \alpha_h) (e_h)^{2\mu} + (\ln \alpha_l) (\alpha_l)^{2\mu}\right) \end{array}\right]}$$

While $\frac{d\mu}{d\rho}$ is in general hard to sign analytically, all the simulations performed produce $\frac{d\mu}{d\rho} < 0$; see next section.
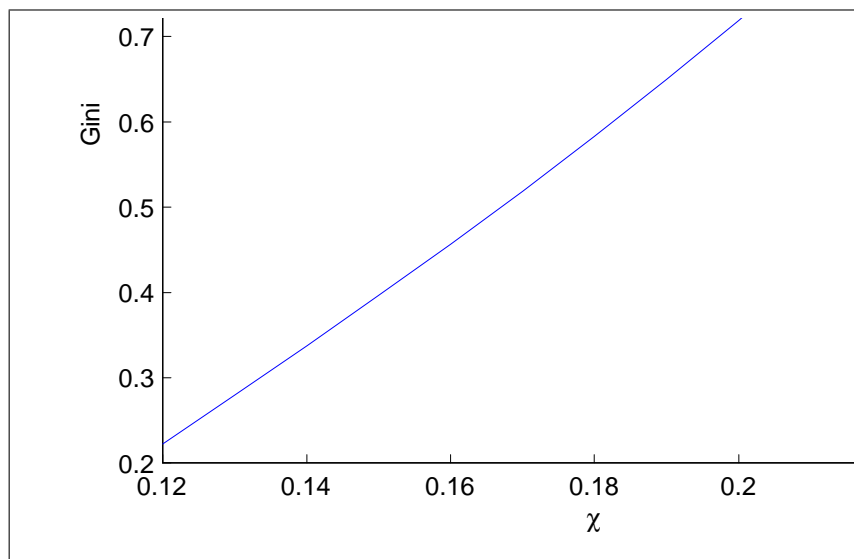
4.5   A simple calibration exercise

As we have already discussed in the Introduction, it has proven hard for standard macro-economic models, when calibrated to the U.S. economy, to produce wealth distributions with tails as thick as in the data. While we do not attempt here a full calibration of our model, in this section we aim at showing that a large range of Gini coefficients of the tail of the wealth distribution can be obtained by varying preferences for bequests and tax parameters. In fact, we easily (e.g. under perfect social mobility) obtain Gini coefficients of the order of $.8 - .9$, larger than observed in the data.

Suppose social mobility is perfect. Calibrate the stationary return r_n with a discrete uniform distribution, taking values $0.01, 0.11$ and 0.21, with a mean return $\bar{r} = 0.11$. Set $\sigma = 1$, $g' = 0.01$, $\rho = 0.04$, and the estate tax rate $b = 0.20$. Finally, set working life span $T = 45$. Under this calibration (with perfect social mobility) we obtain for $\chi = .2$ a Gini coefficient of 0.72 (see the Figure).As χ is increased to $.25$, the Gini coefficient rapidly grows over $.95$.

The Table below illustrates instead the effects of taxes on the Gini coefficient G of the tail of the distribution of wealth. We calibrate the economy as before, under perfect social mobility. We also fix $\chi = 0.2$. The table shows G as we vary b and ζ.

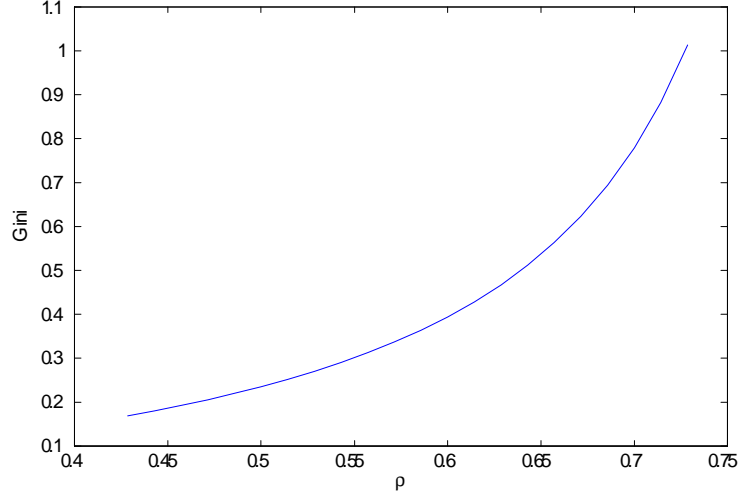| $b \backslash \zeta$	0.15	0.175	0.20
0.10	$G = 0.91$	$G = 0.57$	$G = 0.34$
0.15	$G = 0.81$	$G = 0.50$	$G = 0.29$
0.20	$G = 0.72$	$G = 0.44$	$G = 0.24$

Once again, even under perfect social mobility, a 10% estate tax b and a 15% capital income tax ζ induce $G = .91$. Furthermore, wealth inequality is quite sensitive to fiscal policy in our economy: for the fiscal policies considered we obtain Gini coefficients ranging from .24 to .91. Such high sensitivity is consistent with several studies of the effects of the tax regimes introduced in Europe after World War II. Higher and more progressive taxes did in fact significantly reduce income and wealth inequality in this historical context; notably, e.g., Lampman (1962) and Kuznets (1955). Most recently, Piketty (2001) has argued that redistributive taxation may have prevented large income shares from recovering after the shocks that they experienced during World War II in France.[33]

In our calibration exercise G appears more sensitive to capital income taxes than estate taxes. Nonetheless, an increase of the estate tax rate from 10 to 20 percent decreases the Gini index between 20 and 30 percent, depending on the level of the capital income tax.[34]

This is in contrast with the results of Becker and Tomes (1979), as we have discussed. It is in constrast also with e.g., Castaneda, Diaz Jimenez, and Rios Rull (2003) and Cagetti and De Nardi (2007) , who find very small (or even perverse) effects of bequest taxes in their calibrations, in models with a skewed distribution of earnings but no capital income risk. This suggests a word of caution in evaluating the effects on wealth inequality of proposed fiscal policies like the abolition of estate tax: in economies in which the capital income risk component is substantial, such policies can have a sizeable

---

[33]This line of argument has been extended to the U.S., Japan, and Canada, respectively, by Piketty-Saez (2003) and Moriguchi-Saez (2005), Saez-Veall (2003).

[34]The effect is larger the larger is the capital income tax rate, in our calibration.

effect in increasing wealth inequality in the tail of the distribution.

Finally, to study the effects of social mobility on wealth inequality, consider the special economy introduced in the previous section, where α_n is a two-state irreducible Markov Chain on $[\alpha_l, \alpha_h]$, with $\alpha_l < 1 < \alpha_h$, and

$$P = \begin{pmatrix} \rho_l & 1 - \rho_l \\ 1 - \rho_h & \rho_h \end{pmatrix}$$

Let $\rho_l = \rho_h = \rho$, $\alpha_h = 1.15$ and $\alpha_l = 0.8$, and note that Condition 8 is always satisfied for $\rho \in (0,1)$ since $\ln 1.15 + \ln 0.8 < 0$. The Figure below illustrates $\frac{dG}{d\rho} > 0$.[35]

In the case illustrated in the Figure where $\rho_l = \rho_h = \rho$, *perfect social mobility* corresponds to $\rho = 0.5$, *moderate mobility* to $.5 < \rho < 1$, and finally *no mobility* corresponds to $\rho = 1$. From the Figure, in the range of *perfect mobility* to *no mobility*, $\rho \in [0.5,\ 1]$, inequality (which is inversely related to μ) increases as mobility declines (mobility is inversely related to ρ).[36]

---

[35] Note that for $\rho > 0.73$ we can have $G > 1$ ($\mu < 1$): the stationary distribution of wealth does not have a mean, and the Gini coefficient is not well defined. However. replacing Condition (8) above with (ii) in Proposition 2 gives $P\bar{\alpha} < 1$ which implies:

$$\rho_l(\alpha_l) + (1 - \rho_l)(\alpha_h) \quad < \quad 1, \text{ which implies } \rho_l > \frac{\alpha_h - 1}{\alpha_h - e_l}$$

$$(1 - \rho_h)(\alpha_l) + (\rho_h)(\alpha_h) \quad < \quad 1, \text{ which implies } \rho_h < \frac{1 - \alpha_l}{\alpha_h - \alpha_l}$$

Therefore if $\rho_l = \rho_h = \rho$, $\alpha_h = 1.15$ and $\alpha_l = 0.8$ we obtain $\rho \in \left(\frac{\alpha_h - 1}{\alpha_h - \alpha_l}, \frac{1 - \alpha_l}{\alpha_h - \alpha_l} \right) = (0.428, 0.571)$ and $1 < \mu \in (1.98, 3.46)$ or $G \in (0.17,\ 0.34)$.

[36] The same relation would hold in the implausible case of *cyclic mobility* where $\rho \in [0, 0.5)$, provided

18

5 Conclusion

The main conclusion of this paper is that capital income risk, that is, idiosyncratic returns on wealth, has a fundamental role in affecting the distribution of wealth. Capital income risk appears crucial in generating the thick tails observed in wealth distributions across a large cross-section of countries and time periods. Furthermore, when the wealth distribution is shaped by capital income risk, wealth inequality is very sensitive to fiscal policies; a result which is often documented empirically but hard to generate in many classes of models without capital income risk. Higher taxes in effect dampen the multiplicative stochastic return on wealth, which is critical to generate the thick tails through lucky runs.

Interestingly, this role of capital income risk as a determinant of the distribution of wealth seems to have been lost also by Vilfredo Pareto. He explicitly noted that an identical stochastic process for wealth across agents will not induce the skewed wealth distribution that we observe in the data (See Pareto (1897), Note 1 to #962, p. 315-316). He therefore introduced skewness into the distribution of talents or the endowments of agents (1897, Notes to #962, p. 416). Left with the distribution of talents and endowments as the main determinant of the wealth distribution, he was perhaps lead to his "Pareto's Law," enunciated e.g., by Samuelson (1965) as follows:

> *In all places and all times, the distribution of income remains the same. Neither institutional change nor egalitarian taxation can alter this fundamental constant of social sciences.*[37]

---

we interpret cyclic mobility as social mobility that is higher than *perfect mobility*. More generally, there are various measures of social mobility for Markov transitions in the literature; see for example, Shorrocks (1978), Sommers-Conlisk (1979) or Dardanoni (1993), Section 5. They all satisfy the following: i) perfect mobility corresponds to a matrix with identical rows, ii) no mobility corresponds to the identity matrix. An example, suggested by Sommers-Conlisk (1979), is the modulus of the second eigenvalue of the transition matrix: for our 2×2 transition matrix, it corresponds to $\lambda_2 = |\rho_h + \rho_l - 1|$.

[37] See Chipman (1976) for a discussion on the controversy between Pareto and Pigou regarding the interpretation of the Law. To be fair to Pareto, he also had a "political economy" theory of fiscal policy (determined by the controlling elites) which could also explain the "Pareto Law;" see Pareto (1901, 1909).

References

Aiyagari, S.R. (1994): "Uninsured Idiosyncratic Risk and Aggregate Savings," *Quarterly Journal of Economics*,109 (3), 659-684.

Angeletos, G. (2007), "Uninsured Idiosyncratic Investment Risk and Aggregate Saving", *Review of Economic Dynamics*, 10, 1-30.

Arrow, K. (1987): "The Demand for Information and the Distribution of Income," *Probability in the Engineering and Informational Sciences,* 1, 3-13.

Atkinson, A.B. (2002): "Top Incomes in the United Kingdom over the Twentieth Century," mimeo, Nuffield College, Oxford.

Becker, G.S. and N. Tomes (1979): "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility," *Journal of Political Economy*, 87, 6, 1153-1189.

Benhabib, J. and A. Bisin (2006), "The Distribution of Wealth and Redistributive Policies", Manuscript, New York University.

Benhabib, J. and S. Zhu (2008): "Age, Luck and Inheritance," NBER Working Paper No. 14128.

Bertaut, C. and M. Starr-McCluer (2002): "Household Portfolios in the United States", in L. Guiso, M. Haliassos, and T. Jappelli, Editor, *Household Portfolios*, MIT Press, Cambridge, MA.

Blanchard, O.J. (1985): "Debt, Deficits, and Finite Horizons," *Journal of Political Economy*, 93, 223-47.

Brandt, A. (1986): "The Stochastic Equation $Y_{n+1} = A_n Y_n + B_n$ with Stationary Coefficients," *Advances in Applied Probability*, 18, 211–220.

Burris, V. (2000): "The Myth of Old Money Liberalism: The Politics of the "Forbes" 400 Richest Americans", *Social Problems*, 47, 360-378.

Cagetti, C. and M. De Nardi (2005): "Wealth Inequality: Data and Models," Federal Reserve Bank of Chicago, W. P. 2005-10.

Cagetti, M. and M. De Nardi (2006), "Entrepreneurship, Frictions, and Wealth", *Journal of Political Economy*, 114, 835-870.

Cagetti, C. and M. De Nardi (2007): "Estate Taxation, Entrepreneurship, and Wealth," NBER Working Paper 13160.

Castaneda, A., J. Diaz-Gimenez, and J. V. Rios-Rull (2003): "Accounting for the U.S. Earnings and Wealth Inequality," *Journal of Political Economy*, 111, 4, 818-57.

Champernowne, D.G. (1953): "A Model of Income Distribution," *Economic Journal*, 63, 318-51.

Chipman, J.S. (1976): "The Paretian Heritage," *Revue Europeenne des Sciences Sociales et Cahiers Vilfredo Pareto*, 14, 37, 65-171.

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.

Dagsvik, J.K. and B.H. Vatne (1999): "Is the Distribution of Income Compatible with a Stable Distribution?," D.P. No. 246, Research Department, Statistics Norway.

Dardanoni, V. (1993): "Measures of Social Mobility," *Journal of Economic Theory*, 61, 372-394.

De Nardi, M. (2004): "Wealth Inequality and Intergenerational Links," *Review of Economic Studies*, 71, 743-768.

Diaz-Gimenez, J., V. Quadrini, and J. V. Rios-Rull(1997): "Dimensions of Inequality: Facts on the U.S. Distributions of Earnings, Income, and Wealth," *Federal Reserve Bank of Minneapolis Quarterly Review*, 21(2), 3-21.

Elwood, P., S.M. Miller, M. Bayard, T. Watson, C. Collins, and C. Hartman (1997): "Born on Third Base: The Sources of Wealth of the 1996 Forbes 400," Boston: Unified for a Fair Economy. See also

http://www.faireconomy.org/press/archive/Pre_1999/forbes_400_study_1997.html

Feenberg, D. and J. Poterba (2000): "The Income and Tax Share of Very High Income Household: 1960-1995," *American Economic Review*, 90, 264-70.

Feller, W. (1966): *An Introduction to Probability Theory and its Applications*, 2, Wiley, New York

Flavin, M. and T. Yamashita (2002): "Owner-Occupied Housing and the Composition of the Household Portfolio", *American Economic Review*, 92, 345-362.

Flodén, M. (2008): "A note on the accuracy of Markov-chain approximations to highly persistent AR(1) processes" *Economics Letters*, 99 (3), 2008, 516-520.

Gale, W.G. and J. K. Scholz (1994): "Intergenerational Transfers and the Accumulation of Wealth", *Journal of Economic Perspectives*, 8, 145-160.

Goldie, C. M. (1991): "Implicit Renewal Theory and Tails of Solutions of Random Equations," *Annals of Applied Probability*, 1, 126–166.

Huggett, M. (1993): "The Risk-Free Rate in Heterogeneous-Agent Incomplete-Insurance Economies, *Journal of Economic Dynamics and Control*, 17, 953-69.

Huggett, M. (1996):, "Wealth Distribution in Life-Cycle Economies," *Journal of Monetary Economics,* 38, 469-494.

Kesten, H. (1973): "Random Difference Equations and Renewal Theory for Products of Random Matrices," *Acta Mathematica.* 131 207–248.

Klass, O.S., Biham, O., Levy, M., Malcai O., 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(2), 143-7.

Kotlikoff, L.J. and L. H. Summers (1981): "The Role of Intergenerational Transfers in Aggregate Capital Accumulation," *Journal of Political Economy*, 89, 706-732.

Krusell, P. and A. A. Smith (1998): "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 106, 867-896.

Kuznets, S. (1955): 'Economic Growth and Economic Inequality,' *American Economic Review*, 45, 1-28.

Lampman, R.J. (1962): *The Share of Top Wealth-Holders in National Wealth, 1922-1956*, Princeton, NJ, NBER and Princeton University Press.

Loève, M. (1977), *Probability Theory*, 4th ed., Springer, New York.

McKay, A. (2008): "Household Saving Behavior, Wealth Accumulation and Social Security Privatization," mimeo, Princeton University.

Moriguchi, C. and E. Saez (2005): "The Evolution of Income Concentration in Japan, 1885-2002: Evidence from Income Tax Statistics," mimeo, University of California, Berkeley.

Moskowitz, T. and A. Vissing-Jorgensen (2002): "The Returns to Entrepreneurial Investment: A Private Equity Premium Puzzle?", *American Economic Review*, 92, 745-778.

Nirei, M. and W. Souma (2004): "Two Factor Model of Income Distribution Dynamics," mimeo, Utah State University.

Pareto, V. (1897): *Cours d'Economie Politique*, II, F. Rouge, Lausanne.

V. Pareto (1901): "Un' Applicazione di teorie sociologiche", *Rivista Italiana di Sociologia*. 5. 402-456, translated as *The Rise and Fall of Elites: An Application of Theoretical Sociology* , Transaction Publishers, New Brunswick, New Jersey, (1991).

V. Pareto (1909): *Manuel d'Economie Politique*, V. Girard et E. Brière, Paris.

Piketty, T. (2001): "Income Inequality in France, 1901-1998," *Journal of Political Economy.*

Piketty, T. and E. Saez (2003): "Income Inequality in the United States, 1913-1998," *Quarterly Journal of Economics*, CXVIII, 1, 1-39.

Quadrini, V. (1999): "The importance of entrepreneurship for wealth concentration and mobility," *Review of Income and Wealth*, 45, 1-19.

Quadrini, V. (2000): "Entrepreneurship, Savings and Social Mobility," *Review of Economic Dynamics*, 3, 1-40.

Roitershtein, A. (2007): "One-Dimensional Linear Recursions with Markov-Dependent Coefficients," *The Annals of Applied Probability,* 17(2), 572-608.

Rutherford, R.S.G. (1955): "Income Distribution: A New Model," *Econometrica*, 23, 277-94.

Saez, E. and M. Veall (2003): "The Evolution of Top Incomes in Canada," NBER Working Paper 9607.

P.A. Samuelson (1965): 'A Fallacy in the Interpretation of the Pareto's Law of Alleged Constancy of Income Distribution,' *Rivista Internazionale di Scienze Economiche e Commerciali*, 12, 246-50.

Saporta, B. (2004): "Etude de la Solution Stationnaire de l'´Equation $Y_{n+1} = a_n Y_n + b_n$, a Coefficients Aleatoires," (Thesis), http://tel.archives-ouvertes.fr/docs/00/04/74/12/PDF/t 00007666.pdf

Saporta, B. (2005): "Tail of the stationary solution of the stochastic equation $Y_{n+1} = a_n Y_n + b_n$ with Markovian coefficients," *Stochastic Processes and Application* 115(12), 1954-1978.

Shorrocks, A. (1978): "The Measurement of Mobility," *Econometrica*, 46, 1013-1024.

Sommers, P.S. and J. Conlisk (1979): "Eigenvalue Immobility Measures for Markov Chains", *Journal of Mathematical Sociology*, 6, 253-276.

Sornette, D. (2000): *Critical Phenomena in Natural Sciences,* Berlin, Springer Verlag.

Tauchen, G. (1986): "Finite State Markov-Chain Approximations to Univariate and Vector Autoregressions," *Economics Letters*, 20(2), 177-81.

Tauchen, G. and R. Hussey (1991): "Quadrature-Based Methods for Obtaining Approximate Solutions to Nonlinear Asset Pricing Models," *Econometrica*, 59(2), 371-96.

Yaari, M. (1965): "Uncertain Lifetime, Life Insurance, and the Theory of the Consumer," *Review of Economic Studies*, 32, No. 2, 137-150.

Wold, H.O.A. and P. Whittle (1957): "A Model Explaining the Pareto Distribution of Wealth," *Econometrica*, 25, 4, 591-5.

Wolff, E. (1987): "Estimates of Household Wealth Inequality in the U.S., 1962-1983," *The Review of Income and Wealth*, 33, 231-56.

Wolff, E. (2004): "Changes in Household Wealth in the 1980s and 1990s in the U.S.," mimeo, NYU.

6   Appendix A: The agent's problem

Proof of Prop. 1: The Euler equation for consumption is

$$\dot{c}(s,t) = \frac{r-\rho}{\sigma} c(s,t) \tag{9}$$

and the (interior) terminal transversality condition:

$$c(s,s+T) = \chi^{\frac{-1}{\sigma}} (1-b)^{\frac{\sigma-1}{\sigma}} w(s,s+T) \tag{10}$$

Recall $\tau = t - s$, and

$$h(s,t) = \int_t^{s+T} y(\tau) e^{-(r)\tau} d\tau$$

Since labor income grows at a constant rate, $y(t) = e^{gt} y$, we have

$$h(s,t) = \frac{y(t)}{r-g} \left(1 - e^{-(r-g)(s+T-t)}\right) \tag{11}$$

Integrating 9, solving and then substituting for terminal wealth from 10 into 2, we can solve for consumption in terms of financial and human capital:

$$c(s,t) = m(\tau)(w(s,t) + h(s,t)),$$

where the propensity to consume out of financial and human wealth, $m(\tau)$, satisfies:

$$m(\tau) = \frac{1}{\left(r - \frac{r-\rho}{\sigma}\right)^{-1} \left(1 - e^{\left(\frac{r-\rho}{\sigma} - r\right)(T-\tau)}\right) + \chi^{\frac{1}{\sigma}} (1-b)^{\frac{1-\sigma}{\sigma}} e^{\left(\frac{r-\rho}{\sigma} - r\right)(T-\tau)}} \tag{12}$$

The comparative statics with respect to τ and b in the statement are now straightforward. As for the comparative statics with respect to ζ, recall that $\frac{\partial m}{\partial r} = -\frac{\partial m}{\partial \zeta}$. ∎

6.1   The dynamics of individual wealth in closed form

Substituting the optimal consumption path into (2), we can write the dynamics of individual wealth as a function of age τ, to obtain the following linear differential equation with variable coefficients:

$$\dot{w}(\tau) = \tilde{r}(\tau) w(\tau) + yq(\tau) \tag{13}$$

where

$$\tilde{r}(\tau) = \left(r - \frac{1}{A^{-1}(1 - e^{-A(T-\tau)}) + e^{-A(T-\tau)}B}\right)$$

$$q(\tau) = \left(1 - \frac{\frac{1}{r-g}(1 - e^{-(r-g)(T-\tau)})}{A^{-1}(1 - e^{-A(T-\tau)}) + e^{-A(T-\tau)}B}\right) e^{g\tau}$$

and

$$A(r) = r - \frac{r - \rho}{\sigma}, \quad B(b) = \chi^{\frac{-1}{\sigma}} (1 - b)^{\frac{1-\sigma}{\sigma}}. \tag{14}$$

The dynamics of individual wealth therefore satisfies the indefinite integral:

$$w(\tau) = e^{\int \tilde{r}(\tau) d\tau} \left(Z + y \int q(\tau) e^{-\int \tilde{r}(\tau) d\tau} d\tau \right)$$

where Z is a constant to be determined by initial conditions. We argued in the text it has a solution of the form

$$w(\tau) = \sigma_w(r, \tau) w(0) + \sigma_y(r) y$$

In fact, we can solve for the dynamics of individual wealth in exact closed form:

Proposition A. 1 *The wealth of an individual of age τ satisfies:*

$$w(\tau) = \sigma_w(r, \tau) w(0) + \sigma_y(r) y$$

with

$$\sigma_w(r, \tau) = \frac{e^{A(r)T} + (A(r)B(b) - 1) e^{A(r)\tau}}{e^{A(r)T} + A(r)B(b) - 1} e^{(r - A(r))\tau}$$

$$\sigma_y(r) = \frac{e^{A(r)T} + (A(r)B(b) - 1) e^{A(r)\tau}}{(A(r)B(b) - 1) e^{A(r)T}} e^{A(r)(T-\tau)+r\tau} y(Q^T(r) - Q^0(r))$$

where

$$Q^T(r) = \{Q(r, \tau)\}_{\tau=T}, \quad Q^0(r) = \{Q(r, \tau)\}_{\tau=0}, \quad and$$

$$Q(r, \tau) = \int q(\tau) \frac{(A(r)B(b) - 1) e^{A(r)T}}{e^{A(r)T} + (A(r)B(b) - 1) e^{A(r)\tau}} e^{-A(r)(T-\tau)-r\tau} d\tau$$

Furthermore, under Assumption 1, $w(\tau) > 0$, for any $\tau > 0$.

Proof of Proposition A.1. Let $A(r) = \left(r - \frac{r-\rho}{\sigma}\right)$ and $B(b) = \chi^{\frac{-1}{\sigma}} ((1-b))^{\frac{1-\sigma}{\sigma}}$. We drop the argument of $A(r)$ and of $B(b)$ for simplicity in the following.
We first compute

$$\int \tilde{r}(\tau) d\tau = \int \left(r - \frac{1}{(A)^{-1}(1 - e^{-A(T-\tau)}) + e^{-A(T-\tau)}B} \right) d\tau$$

$$= A(T - \tau) + r\tau + \ln \frac{1}{(AB - 1) e^{AT}} \left(e^{AT} + (AB - 1) e^{A\tau} \right) + C_1$$

for C_1 to be determined by initial conditions. Therefore,

$$e^{-\int \tilde{r}(\tau)d\tau} = \frac{(AB-1)e^{AT}}{e^{AT}+(AB-1)e^{A\tau}}e^{-C_1}e^{-A(T-\tau)-r\tau}$$

We then compute

$$q(\tau) = \left(1 - \frac{\frac{1-e^{-(r-g)(T-\tau)}}{r-g}}{\left(r-\frac{r-\rho}{\sigma}\right)^{-1}\left(1-e^{\left(\frac{r-\rho}{\sigma}-r\right)(T-\tau)}\right)+e^{\left(\frac{r-\rho}{\sigma}-r\right)(T-\tau)}\chi^{\frac{-1}{\sigma}}((1-b))^{\frac{1-\sigma}{\sigma}}}\right)e^{g\tau}$$

$$q(\tau) = \left(1 - \frac{\frac{1}{r-g}\left(1-e^{-(r-g)(T-\tau)}\right)}{A^{-1}\left(1-e^{-A(T-\tau)}\right)+e^{-A(T-\tau)}B}\right)e^{g\tau}$$

We have, therefore,

$$\tilde{Q}(r,\tau) = \int q(\tau)e^{-\int \tilde{r}(\tau)d\tau}d\tau$$

$$= \int q(\tau)\frac{(AB-1)e^{AT}}{e^{AT}+(AB-1)e^{A\tau}}e^{-C_1}e^{-A(T-\tau)-r\tau}d\tau$$

and

$$\tilde{Q}^T(t) = \left\{\int q(\tau)\frac{(AB-1)e^{AT}}{e^{AT}+(AB-1)e^{A\tau}}e^{-C_1}e^{-A(T-\tau)-r\tau}d\right\}|_{\tau=T}$$

$$\tilde{Q}^0(r) = \left\{\int q(\tau)\frac{(AB-1)e^{AT}}{e^{AT}+(AB-1)e^{A\tau}}e^{-C_1}e^{-A(T-\tau)-r\tau}d\right\}|_{\tau=0}$$

so that

$$w(\tau) = \frac{\left(e^{AT}+(AB-1)e^{A\tau}\right)}{(AB-1)e^{AT}}e^{C_1}e^{A(T-\tau)+r\tau}\left(Z+y\tilde{Q}(r,\tau)\right)$$

$$w(T) = \frac{\left(e^{AT}+(AB-1)e^{AT}\right)}{(AB-1)e^{AT}}e^{C_1}e^{rT}\left(Z+y\tilde{Q}^T(r)\right)$$

$$= e^{rT}e^{C_1}\left(Z+y\tilde{Q}^T(r)\right)$$

$$w(0) = \frac{\left(e^{AT}+(AB-1)\right)}{(AB-1)}e^{C_1}\left(Z+y\tilde{Q}^0(r)\right)$$

We can now solve for Z :

$$e^{-C_1}\frac{(AB-1)}{(e^{AT}+AB-1)}w(0)-y\tilde{Q}^0(r)=Z;$$

and hence,

$$
\begin{aligned}
w\left(T\right) &= e^{C_1}e^{rT}\left(e^{-C_1}\frac{\left(AB-1\right)}{\left(e^{AT}+AB-1\right)}w\left(0\right)-y\tilde{Q}^0\left(r\right)+y\tilde{Q}^T\left(r\right)\right)\\
&= \frac{e^{rT}\left(AB-1\right)}{\left(e^{AT}+\left(AB-1\right)\right)}w\left(0\right)+e^{C_1}e^{rT}y\left(\tilde{Q}^T\left(r\right)-\tilde{Q}^0\left(r\right)\right)
\end{aligned}
$$

We can factor out e^{-C_1} from $\tilde{Q}\left(r,\tau\right)$, so that $\tilde{Q}\left(r,\tau\right)=e^{-C_1}Q\left(r,\tau\right)$. Then

$$
w\left(T\right)=\left[\frac{e^{rT}\left(AB-1\right)}{\left(e^{AT}+\left(AB-1\right)\right)}\right]w\left(0\right)+\left[e^{rT}y\left(Q^T\left(r\right)-Q^0\left(r\right)\right)\right]
$$

and

$$
w\left(\tau\right)=\frac{\left(e^{AT}+\left(AB-1\right)e^{A\tau}\right)}{\left(AB-1\right)e^{AT}}e^{A\left(T-\tau\right)+r\tau}\left(\frac{\left(AB-1\right)}{\left(e^{AT}+AB-1\right)}w\left(0\right)+y\left(Q^T\left(r\right)-Q^0\left(r\right)\right)\right)
$$

We now show that if $w\left(\tau\right)>0$. We have

$$
\begin{aligned}
w\left(\tau\right) &= \frac{\left(e^{AT}+\left(AB-1\right)e^{A\tau}\right)}{e^{AT}}e^{A\left(T-\tau\right)+r\tau}\times\\
&\quad\times\left(\frac{1}{\left(e^{AT}+AB-1\right)}w\left(0\right)+\frac{y(Q^T\left(r\right)-Q^0\left(r\right))}{\left(AB-1\right)}\right)
\end{aligned}
$$

We compute $\left(Q^T\left(r_n\right)-Q^0\left(r_n\right)\right)$. We have

$$
Q(r_n,\tau)=\int q\left(\tau\right)\frac{\left(AB-1\right)e^{AT}}{e^{AT}+\left(AB-1\right)e^{A\tau}}e^{-A\left(T-\tau\right)-r\tau}d\tau
$$

$$
\begin{aligned}
&\left(Q^T\left(r_n\right)-Q^0\left(r_n\right)\right)\\
&= \int\frac{q\left(\tau\right)\left(AB-1\right)e^{AT}}{e^{AT}+\left(AB-1\right)e^{A\tau}}e^{-A\left(T-\tau\right)-r\tau}d\tau|_{\tau=T}-\int\frac{q\left(\tau\right)\left(AB-1\right)e^{AT}}{e^{A)T}+\left(AB-1\right)e^{A\tau}}e^{-A\left(T-\tau\right)-r\tau}d\tau|_{\tau=0}
\end{aligned}
$$

Then

$$
w\left(\tau\right)
$$

$$
= \frac{\left(e^{AT}+\left(AB-1\right)e^{A\tau}\right)}{e^{AT}}e^{A\left(T-\tau\right)+r\tau}\times
$$

$$
\left(\frac{w\left(0\right)}{\left(e^{AT}+AB-1\right)}+\frac{\left[y\int\frac{q(\tau)(AB-1)e^{AT}e^{-A\left(T-\tau\right)-r\tau}}{e^{AT}+(AB-1)e^{A\tau}}d\tau|_{\tau=T}-\int\frac{q(\tau)(AB-1)e^{AT}e^{-A\left(T-\tau\right)-r\tau}}{e^{AT}+(AB-1)e^{A\tau}}d\tau|_{\tau=0}\right]}{\left(AB-1\right)}\right)
$$

$$w\left(\tau\right) = \frac{\left(e^{AT} + \left(AB - 1\right)e^{A\tau}\right)}{e^{AT}}e^{A(T-\tau)+r\tau} \times$$

$$\times \left(\frac{1}{\left(e^{AT} + AB - 1\right)}w\left(0\right) + y\int_0^T \frac{q\left(\tau\right)}{e^{AT} + \left(AB - 1\right)e^{A\tau}}e^{(A-r)\tau}d\tau\right)$$

Note that $\left(e^{AT} + \left(AB - 1\right)e^{A\tau}\right) \geq 0$ since $T \geq \tau$ and $A = r - \left(\frac{r-\rho}{\sigma}\right) = r\left(1 - \sigma^{-1}\right) + \rho\sigma > 0$. It follows that $w\left(\tau\right) > 0$ if $q\left(\tau\right) > 0$. We have

$$q\left(\tau\right) = \left(1 - \frac{\frac{1}{r-g}\left(1 - e^{-(r-g)(T-\tau)}\right)}{A^{-1}\left(1 - e^{-A(T-\tau)}\right) + e^{-A(T-\tau)}B}\right)e^{g\tau}$$

So a sufficient condition for $q\left(\tau\right) > 0$ is $\frac{1}{r-g}\left(1 - e^{-(r-g)(T-\tau)}\right) < A^{-1}\left(1 - e^{-A(T-\tau)}\right)$. Since $x^{-1}\left(1 - e^{-x(T-\tau)}\right)$ is declining in x, $q\left(\tau\right) > 0$ if $A > 0$, that is if $\frac{r-\rho}{\sigma} > g$ or $r > \rho + \sigma g$. $\blacksquare$

7 Appendix B: Dynamics of Wealth

We can construct a stochastic difference equation for the initial wealth of dynasties, mapping w_{n-1} into w_n. From Proposition 1, in fact, dynasty n's initial wealth w_n satisfies:

$$w_{n+1} = (1 - b)\frac{A(r_n)B(b)e^{r_n T}}{(A(r_n)B(b) - 1) + e^{A(r_n)T}}w_n + (1 - b)e^{r_n T}y_n e^{g'n}\left(Q^T\left(r_n\right) - Q^0\left(r_n\right)\right).$$

Working with discounted variables,

$$z_n = \left(e^{-g'}\right)^n w_n, \quad z_{n-1} = \left(e^{-g'}\right)^{n-1} w_{n-1}$$

we obtain the following stochastic difference equation:

$$z_{n+1} = \alpha_n z_n + \beta_n \tag{15}$$

where

$$\alpha_n\left(r_n\right) = (1 - b)\frac{A(r_n)B(b)e^{r_n T}}{A(r_n)B(b) - 1 + e^{A(r_n)T}}e^{-g'} \tag{16}$$

$$\beta_n\left(r_n, y_n\right) = (1 - b)e^{r_n T}y_n e^{-g'}\left(Q^T\left(r_n\right) - Q^0\left(r_n\right)\right) \tag{17}$$

Thus $\left(\alpha_n\right)_n$ is driven by $\left(r_n\right)_n$ and is independent of $\left(y_n\right)_n$.

Proof of Proposition 2: Let $\mathbf{R} = \{\bar{r}_1, ..., \bar{r}_m\}$ denote the state space of r_n. Similarly, let $\mathbf{Y} = \{\bar{y}_1, ..., \bar{y}_l\}$ denote the state space of y_n. Let $\mathbf{A} = \{\bar{\alpha}_1 =, ...\bar{\alpha}_m\}$ and

$\mathbf{B} = \{\overline{\beta}_1, ...\overline{\beta}_{lm}\}$ denote the state spaces of, respectively, α_n and β_n, as they are induced through the maps (15-17). We shall show that the maps (15-17) are bounded in r_n and y_n. Therefore the state spaces of α_n and β_n are well defined. It immediately follows then that, if $(r_n, y_n)_n$ is a Markov Modulated chain (Assumption 2), so is $(\alpha_n, \beta_n)_n$.

We now show that under Assumption 3 (i), $(\alpha_n, \beta_n)_n$ is > 0; we also show that $(\alpha_n, \beta_n)_n$ are bounded with probability 1 in r_n and y_n. Recall that $A(r_n) = r_n - \frac{r_n - \rho}{\sigma} > 0$ with probability 1 (w.p. 1); since $\sigma \geq 1$ by Assumption 1. Also, $B(b) = \chi^{\frac{-1}{\sigma}} (1 - b)^{\frac{1-\sigma}{\sigma}} > 0$. The denominator of 16 is positive since $A(r_n)B(b) > 0$ and $e^{A(r_n)T} > 1$ and therefore $\alpha_n > 0$ and bounded. Therefore (α_n, β_n) is a Markov Modulated Process provided $(\beta_n)_n$ is positive and bounded.

We now show that $(\beta_n)_n \geq 0$ and is bounded. Recall that

$$Q^T(r_n) = \{Q(r_n, \tau)\}|_{\tau=T}, \ Q^0(r_n) = \{Q(r_n, \tau)\}|_{\tau=0},$$

$$Q(r_n, \tau) = \int q(\tau) \frac{(A(r_n)B(b) - 1) e^{A(r_n)T}}{e^{A(r_n)T} + (A(r_n)B(b) - 1) e^{A(r_n)\tau}} e^{-A(r)(T-\tau)-r\tau} d\tau,$$

$$q(\tau) = \left(1 - \frac{\frac{A(r)}{r-g}\left(1 - e^{-(r-g)(T-\tau)}\right)}{\left(1 + e^{-A(r)(T-\tau)}(A(r)B(b) - 1)\right)}\right) e^{g\tau}$$

It is straightforward to see that $q(\tau)$ is bounded for $\tau \in [0, T]$. Therefore $q(\tau) \frac{(A(r_n)B(b)-1)e^{A(r_n)T}}{e^{A(r_n)T}+(A(r_n)B(b)-1)e^{A(r_n)}}$ is bounded for $\tau \in [0, T]$. It follows that $Q(r_n, \tau) = \int q(\tau) \frac{(A(r_n)B(b)-1)e^{A(r_n)T}}{e^{A(r_n)T}+(A(r_n)B(b)-1)e^{A(r_n)\tau}} e^{-A(r)(T-\tau)-r\tau} d\tau$ is bounded for $\tau \in [0, T]$. We conclude then that $Q^T(r_n) - Q^0(r_n)$ is bounded. But since the support of y_n is bounded by Assumption 2, $\beta_n = (1 - b) e^{r_n T} e^{-g'} y_n \left(Q^T(r_n) - Q^0(r_n)\right)$ is also bounded.

We now show $\beta_n = \left((1 - b) e^{r_n T} e^{-g'} y_n\right) \left(Q^T(r_n) - Q^0(r_n)\right) > 0$. To see this, first note that $\left(Q^T(r_n) - Q^0(r_n)\right)$ is independent of $w(s, t)$, and in particular, of $w(s, s) = w(0)$. So set $w(0) = 0$ If $w(0) = 0$, then

$$w(T) - \left(\frac{e^{rT} A(r_n) B(b)}{(e^{A(r_n)T} + (A(r_n) B(b) - 1))}\right) w(0) = w(T) = w(s, s + T)$$

$$= e^{rT} \left(Q^T(r_n) - Q^0(r_n)\right)$$

So we have to show that $w(T) = w(s, s + T) > 0$ if $w(0) = 0$. Using (9) and (2), integrating and using (10) to eliminate $w(s, s + T)$, we get

$$c(s, t) \int_t^{s+T} e^{-\left(r\left(1-\sigma^{-1}\right)+\sigma^{-1}\rho\right)(T-\tau)} d\tau \tag{18}$$

$$= w(s, t) + h(s, t) - w(s, s + T)e^{-r(s+T-t)}$$

$$= w(s, t) + h(s, t) - c(s, t) e^{-\left(\left(1-\sigma^{-1}\right)r+\sigma^{-1}\rho\right)(s+T-t)} \chi^{\frac{1}{\sigma}} ((1 - b))^{\frac{1-\sigma}{\sigma}} \tag{19}$$

30

$$c\left(s,t\right)=\frac{w\left(s,t\right)+h\left(s,t\right)}{\left(r\left(1-\sigma^{-1}\right)+\sigma^{-1}\rho\right)^{-1}\left(1-e^{-\left(r\left(1-\sigma^{-1}\right)+\sigma^{-1}\rho\right)\left(s+T-t\right)}\right)}$$
$$+e^{-\left(\left(1-\sigma^{-1}\right)r+\sigma^{-1}\rho\right)\left(s+T-t\right)}\chi^{\frac{1}{\sigma}}\left(\left(1-b\right)\right)^{\frac{1-\sigma}{\sigma}} \tag{20}$$

From (10) we also have

$$c\left(s,s+T\right)=\frac{w\left(s,s+T\right)}{\chi^{\frac{1}{\sigma}}\left(\left(1-b\right)\right)^{\frac{1-\sigma}{\sigma}}} \tag{21}$$

$$c\left(s,s+T\right)=c\left(s,s\right)e^{\left(\sigma^{-1}(r-\rho)\right)T} \tag{22}$$

$$c\left(s,s\right)e^{\left(\sigma^{-1}(r-\rho)\right)T}=\frac{w\left(s,s+T\right)}{\chi^{\frac{1}{\sigma}}\left(\left(1-b\right)\right)^{\frac{1-\sigma}{\sigma}}} \tag{23}$$

Using definition of h, and the assumption that $w\left(s,s\right)=0$,

$$c\left(s,s\right)\begin{array}{c}\left(r\left(1-\sigma^{-1}\right)+\sigma^{-1}\rho\right)^{-1}\left(1-e^{-\left(r\left(1-\sigma^{-1}\right)+\sigma^{-1}\rho\right)T}\right)\\+e^{-\left(\left(1-\sigma^{-1}\right)r+\sigma^{-1}\rho\right)T}\chi^{\frac{1}{\sigma}}\left(\left(1-b\right)\right)^{\frac{1-\sigma}{\sigma}}\end{array}=y\left(t\right)\left(r-g\right)^{-1}\left(1-e^{-(r-g)T}\right)$$

Substituting into 10,

$$\frac{y\left(r-g\right)^{-1}\left(1-e^{-(r-g)T}\right)e^{\left(\sigma^{-1}(r-\rho)\right)T}}{\left(r\left(1-\sigma^{-1}\right)+\sigma^{-1}\rho\right)^{-1}\left(1-e^{-(r(1-\sigma^{-1})+\sigma^{-1}\rho)T}\right)+e^{-((1-\sigma^{-1})r+\sigma^{-1}\rho)T}\chi^{\frac{1}{\sigma}}\left(\left(1-b\right)\right)^{\frac{1-\sigma}{\sigma}}}$$
$$=\left(\chi^{\frac{1}{\sigma}}\left(\left(1-b\right)\right)^{\frac{1-\sigma}{\sigma}}\right)w\left(s,s+T\right)$$

$$\frac{y\left\{\left(r-g\right)^{-1}\left(1-e^{-(r-g)T}\right)\right\}e^{\left(\sigma^{-1}(r-\rho)\right)T}\chi^{\frac{1}{\sigma}}\left(\left(1-b\right)\right)^{\frac{1-\sigma}{\sigma}}}{\left\{\left(r\left(1-\sigma^{-1}\right)+\sigma^{-1}\rho\right)^{-1}\left(1-e^{-(r(1-\sigma^{-1})+\sigma^{-1}\rho)T}\right)\right\}+e^{-((1-\sigma^{-1})r+\sigma^{-1}\rho)T}\chi^{\frac{1}{\sigma}}\left(\left(1-b\right)\right)^{\frac{1-\sigma}{\sigma}}}$$
$$=w\left(s,s+T\right)$$

But the left side is positive since the brackets $\left\{\left(r-g\right)^{-1}\left(1-e^{-(r-g)T}\right)\right\}$ and $\left\{\left(r\left(1-\sigma^{-1}\right)+\sigma^{-1}\rho\right)^{-1}\left(1-e^{-\left(r\left(1-\sigma^{-1}\right)+\sigma^{-1}\rho\right)T}\right)\right\}$ are always positive.[38] So we conclude that $\left(\alpha_{n},\beta_{n}\right)_{n}$ is >0 and bounded with probability 1.

Furthermore, Assumption 3 (ii) implies directly that (ii) $P\bar{\alpha}<1$. Assumption 3 (iii) also directly implies $\bar{\alpha}_{i}>1$ for some $i=1,...m$. Finally P is the transition matrix of both r_{n} as well as of α_{n}. Therefore Assumption 3 (iv) implies that the elements of the trace of the transition matrix of α_{n} are positive. ∎

---

[38]Note that $\left\{\left(r-g\right)^{-1}\left(1-e^{-(r-g)T}\right)\right\}$ =1 if $r-g\rightarrow0$, and $\left\{\left(r\left(1-\sigma^{-1}\right)+\sigma^{-1}\rho\right)^{-1}\left(1-e^{-\left(r\left(1-\sigma^{-1}\right)+\sigma^{-1}\rho\right)T}\right)\right\}=1$ if $\left(r\left(1-\sigma^{-1}\right)+\sigma^{-1}\rho\right)^{-1}\rightarrow0$

7.1 Regularity conditions on the Markov Modulated process $\left(\alpha_n\, ,\beta_n\right)_n$

In singular cases, particular correlations between α_n and β_n can create degenerate distributions that eliminate the randomness wealth. We rule this out by means of technical regularity conditions.[39]

Assumption A 1 *The Markov Modulated process $(\alpha_n, \beta_n)_n$ is regular, that is*

$$\Pr\left(\alpha_0 x + \beta_0 = x\right) \neq 1 \ for \ any \ x \in \mathbb{R}_+$$

and the elements of the vector $\bar{\boldsymbol{\alpha}} = \{\ln \overline{\alpha}_1 ... \ln \overline{\alpha}_m\} \subset \mathbb{R}_+^m$ are not integral multiples of the same number.

Theorems which characterize the tails of distributions generated by equations with random multiplicative coefficients of the type (3) rely on this type of "non-lattice" assumptions from Renewal Theory; see for example Saporta (2005). Versions of these assumption are standard in this literature; see Feller (1966).)

8 Appendix C: The Kesten-Goldie Theorem

Consider the special case in which $(\alpha_n)_n$, and $(\beta_n)_n$ are > 0, bounded with probability 1, *i.i.d.*, and α_n satisfies: $E\alpha_n < 1$ and $\alpha_i > 1$ for some i. In this case, Kesten (1973) and Goldie (1991) show that the tail of the distribution is asymptotic to a Pareto law

$$\Pr{}_>(z_n > z) \sim cz^{-\mu}$$

where $\mu > 1$ satisfies

$$E\left(\alpha_n\right)^\mu = 1.$$

A simple very heuristic proof of the Theorem follows. Consider the stochastic difference equation

$$z_{n+1} = \alpha_n z_n + \beta_n$$

It has solution:

$$z_{n+N} = \left(\prod_{l=0}^{N-1} \alpha_{n+l}\right) z_n + \sum_{l=0}^{N-1} \beta_{n+l} \prod_{m=l+1}^{N-1} \alpha_{n+m}$$

where $\alpha_{n+N} = 1$ by definition for the special value $l = N - 1$ in the last term.

---

[39]We formulate these regularity conditions on $(\alpha_n, \beta_n)_n$, but they can be immediately mapped back into conditions on the stochastic process $(r_n, y_n)_n$.

Assume β_n has bounded support for any n. Given realizations of α_n, β_n, to attain $z_{n+1} = z$ the prior period value must be $z_n = \frac{z - \beta_n}{\alpha_n}$. For simplicity take $\beta_n = \beta$ constant (but constant support will do as well). Then, letting $P(z)$ denote the stationary distribution of $(z_n)_n$,

$$P(z_n) = \int P\left(\frac{z - \beta}{\alpha_n}\right) \pi(\alpha_n) \, d\alpha_n$$

where $\pi(\alpha_n)$ is the density of α_n. For large z, in the tail, we can approximate the solution by the solution to

$$P(z_n) = \int P\left(\frac{z}{\alpha_n}\right) \pi(\alpha_n) \, d\alpha_n$$

by $P(z_n) = Cz^{-\mu}$:

$$Cz^{-\mu} = \int Cz^{-\mu} (\alpha_n)^{\mu} \pi(\alpha_n) \, d\alpha_n = Cz^{-\mu} \int (\alpha_n)^{\mu} \pi(\alpha_n) \, d\alpha_n$$

Therefore, μ solves

$$\int (\alpha_n)^{\mu} \pi(\alpha_n) \, d\alpha_n = 1. \ \blacksquare$$