

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

Jess Benhabib
NYU and NBER

Alberto Bisin
NYU and NBER

Shenghao Zhu
NYU

This draft: October 2009

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 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, and on different degrees of social mobility. We show that capital income and estate taxes can significantly reduce wealth inequality, as do institutions favoring social mobility. Finally, we calibrate the economy to match the Lorenz curve of the wealth distribution of the U.S. economy.

*We gratefully acknowledge Daron Acemoglu's extensive comments on an earlier paper on the same subject, which have lead us to the formulation in this paper. We also acknowledge the ideas and suggestions of five referees that we incorporated into the paper, as well as the conversations with Marco Bassetto, Alberto Bressan, Bei Cao, In-Koo Cho, Gianluca Clementi, Isabel Correia, Mariacristina De Nardi, Raquel Fernandez, Xavier Gabaix, Leslie Greengard, Frank Hoppensteadt, Boyan Jovanovic, Stefan Krasa, Nobu Kiyotaki, Guy Laroque, John Leahy, Omar Licandro, Andrea Moro, Jun Nie, Chris Phelan, Alexander Roitershtein, Hamid Sabourian, Benoit de Saporta, Tom Sargent, Ennio Stacchetti, Pedro Teles, Viktor Tsyrennikov, Gianluca Violante, Ivan Werning, Ed Wolff, and Zheng Yang. Thanks to Nicola Scalzo and Eleonora Patacchini for help with "impossible" Pareto references in dusty libraries. We also gratefully acknowledge Viktor Tsyrennikov's expert research assistance. This paper is part of the Polarization and Conflict Project CIT-2-CT-2004-506084 funded by the European Commission-DG Research Sixth Framework Programme.

1 Introduction

Rather invariably across a large cross-section of countries and time periods income and wealth distributions are skewed to the right¹ 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 wealth³ and the top end of the wealth distribution obeys a Pareto law, the standard statistical model for heavy upper tails.⁴

Which characteristics of the wealth accumulation process are responsible for these stylized facts? To answer this question, we study the relationship between wealth inequality and the structural parameters in an economy in which households choose optimally their life cycle consumption and saving paths. We aim at understanding first of all heavy upper tails, as they represent one of the main empirical features of wealth inequality.⁵

Stochastic labor endowments can in principle generate some skewness in the distribution of wealth, especially if the labor endowment process is itself skewed and persistent. A large literature studies indeed models in which households face uninsurable idiosyncratic labor income risk (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, as discussed in Carroll (1997) and in Quadrini (1999), is that at higher wealth levels, the incentives for further precautionary savings tapers off and the tails of wealth distribution remain thin. In order to generate skewness with heavy tails in wealth distribution, a number of authors have therefore successfully

¹Atkinson (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).

²Heavy upper tails (power law behavior) for the distributions of income and wealth are also well documented, for example by Nirei-Souma (2007) 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.

³See 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).

⁴Using 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.

⁵A 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.

introduced new features, like for example preferences for bequests, entrepreneurial talent that generates stochastic returns (Quadrini (1999, 2000), Cagetti and De Nardi, 2006),⁶ or heterogenous discount rates that follow an exogenous stochastic process (Krusell and Smith (1998)).

Our model is related to these papers. We study an overlapping generation economy where households are finitely lived and have a "joy of giving" bequest motive. Furthermore, to capture entrepreneurial risk, we assume households face stochastic stationary processes for both labor and capital income. In particular, we assume *i)* (the log of) labor income has an uninsurable idiosyncratic component and a trend-stationary component across generations,⁷ *ii)* capital income also is governed by stationary idiosyncratic shocks, possibly persistent across generations. This specification of labor and capital income requires justification.

The combination of idiosyncratic and trend-stationary components of labor income finds some support in the data; see Guvenen (2007). Most studies of labor income require some form of stationarity of the income process, though persistent income shocks are often allowed to explain the cross-sectional distribution of consumption; see e.g., Storesletten, Telmer, Yaron (2004).⁸

The assumption that capital income contains a relevant idiosyncratic component is not standard in macroeconomics, though Angeletos and Calvet (2006) and Angeletos (2007) introduce it to study aggregate savings and growth.⁹ Idiosyncratic capital income risk appears however to be a significant element of the lifetime income uncertainty of individuals and households. Two components of capital income are particularly subject to idiosyncratic risk: ownership of principal residence and private business equity, which account for, respectively, 28.2% and 27% of household wealth in the U.S., according to the 2001 *Survey of Consumer Finances* (Wolff, 2004 and Bertaut-Starr-McCluer, 2002).¹⁰ Case and Shiller (1989) document a large standard deviation, of the order of 15%, of yearly capital gains or losses on owner-occupied housing. Similarly, Flavin and Yamashita (2002) measure the standard deviation of the return on housing, at the level of individual houses, from the 1968-92 waves of the Panel Study of Income Dynamics, obtaining a similar number, 14%. Returns on private equity have an even

⁶In Quadrini (2000) the entrepreneurs receive stochastic idiosyncratic returns from projects that become available through an exogenous Markov process in the "non-corporate" sector, while there is also a corporate sector that offers non-stochastic returns.

⁷In fact, trend-stationarity of income is assumed mostly for simplicity. More general stationary processes can be accounted for.

⁸While some authors, e.g., Primiceri and van Rens (2006), adopt a non-stationary specification for individual income, it seems hardly the case that such a specification is suggested by income and consumption data; see e.g., the discussion of Primiceri and van Rens (2006) by Heathcote (2008). See also Heathcote, Storesletten, and Violante (2009) for an extensive survey.

⁹See also Angeletos and Calvet (2005) and Panousi (2008).

¹⁰From a different angle, 67.7% of households own principal residence (16.8% own other real estate) and 11.9% of household own unincorporated business equity.

higher idiosyncratic dispersion across household, a consequence of the fact that private equity is highly concentrated: 75% of all private equity is owned by households for which it constitutes at least 50% of their total net worth (Moskowitz and Vissing-Jorgensen, 2002). In the 1989 *SCF* studied by Moskowitz and Vissing-Jorgensen (2002), the median of the distribution of returns on private equity is 6.9%, while the first quartile is 0 and the third quartile is 18.6%.¹¹ Evidently, the presence of moral hazard and other frictions render risk diversification, as well as concentrating each household's wealth in the hands of the household with the best investment technology, hardly feasible.

Under these assumptions on labor and capital income risk,¹² the stationary wealth distribution is a Pareto distribution in the right tail. The economics of this result is straightforward. When labor income is stationary, it accumulates additively into wealth. The multiplicative process of wealth accumulation will then tend to dominate the distribution of wealth in the tail (for high wealth). This is why Bewley models, calibrated to earning shocks with no capital income shocks, have difficulties producing the observed skewness of the wealth distribution. The heavy tails in the wealth distribution, in our model, are populated by the dynasties of households which have realized a long streak of high rates of return on capital income. We analytically show that it is capital income risk rather than stochastic labor income that drives the properties of the right tail of the wealth distribution.¹³

An overview of our analysis is useful to navigate over technical details. If w_{n+1} is the initial wealth of an n -th generation household, we show that the dynamics of wealth follows

$$w_{n+1} = \alpha_{n+1}w_n + \beta_{n+1}$$

where α_{n+1} and β_{n+1} are stochastic processes which can be interpreted, respectively, as the effective rate of return on wealth across generations (the *life-cycle return on wealth*) and the permanent income of a generation, after subtracting lifetime consumption. If α_{n+1} and β_{n+1} are *i.i.d.* processes, under appropriate assumptions¹⁴ this dynamic of wealth converges to a stationary distribution with a Pareto law

$$\Pr(w_n > w) \sim cw^{-\mu}$$

¹¹See Angeletos (2007) and Benhabib and Zhu (2008) for more evidence on the macroeconomic relevance of idiosyncratic capital income risk. Quadrini (2000) also extensively documents the role of idiosyncratic returns and entrepreneurial talent for explaining the heavy tails of wealth distribution.

¹²Although we emphasize the interpretation with stochastic returns, our model also accomodates a reduced form interpretation of stochastic discounting, as in Krusell-Smith (1998).

¹³An alternative approach to generate fat tails without stochastic returns or discounting is to introduce a "perpetual youth" model with bequests, where the probability death (and or retirement) is independent of age. In these models, the stochastic component is not stochastic returns or discount rates 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).

¹⁴In particular $E(\alpha_{n+1}) < 1$, $\Pr(\alpha_{n+1} > 1) > 0$, and $\beta_{n+1} > 0$.

with an explicit expression for μ in terms of the processes for α_{n+1} and β_{n+1} .¹⁵

But α_{n+1} and β_{n+1} are endogenously determined by the life-cycle saving and bequest behavior of households. Only by studying the life-cycle choices of households we can characterize the dependence of the distribution of wealth, and of wealth inequality in particular, on the various structural parameters of the economy, e.g., technology, preferences, and fiscal policy instruments like capital income taxes and estate taxes. We show that capital income and estate taxes reduce the concentration of wealth in the top tail of the distribution. Capital and estate taxes have an effect on the top tail of wealth distribution because they dampen the accumulation choices of households experiencing lucky streaks of persistent high realizations in the stochastic rates of return. We show by means of simulations that this effect is potentially very strong.

Furthermore, once α_{n+1} and β_{n+1} are obtained from households' saving and bequest decisions, it becomes apparent that the *i.i.d.* assumption is very restrictive. Positive autocorrelations in α_{n+1} and β_{n+1} 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 possibility of a correlation between α_{n+1} and β_{n+1} , to capture institutional environments where households with high labor income to have better opportunities for higher returns on wealth in financial markets. By using some new results in the mathematics of stochastic processes (due to Saporta, 2004 and 2005, and to Roitershtein, 2007) we are able to show that even in this case the stationary wealth distribution has a Pareto tail, and to compute the effects of social mobility on the tail analytically.¹⁶

Finally, we illustratively calibrate and simulate our model to obtain the full wealth distribution, rather than just the tail. The model performs quite well in matching the (Lorenz curve of the) empirical distribution of wealth in the U.S.¹⁷

Section 2 introduces the household's life-cycle consumption and saving decisions. Section 3 gives the characterization of the stationary wealth distribution with power tails, and a discussion of the assumptions underlying the result. In Section 4 our results for the effects of capital income and estate taxes on tail index are stated. Section 4 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 return on capital

¹⁵See Kesten (1973) and Goldie (1991). Gabaix (1999, Appendix 1) and Nirei-Souma (2007) have used applications of Kesten processes in economics.

¹⁶Champernowne (1953) is the first paper exploring the role of stochastic returns on wealth that follow a Markov chain to generate an asymptotic Pareto distribution of wealth. Recently Levy (2005), in the same tradition, studies a stochastic multiplicative process for returns and characterizes the resulting stationary distribution; see also Levy and Solomon (1996) for more formal arguments and Fiaschi-Marsili (2009). These papers however do not provide the microfoundations necessary for consistent comparative static exercises. Furthermore, they all assume *i.i.d.* processes for α_{n+1} and β_{n+1} and an exogenous lower barrier on wealth.

¹⁷We also explore the differential effects of capital and estate taxes and of social mobility on the tail index for top wealth shares and the Gini coefficient for the whole wealth distribution.

across generations. In Section 5 we do a simple calibration exercise to match the Lorenz curve and the fat tail of the wealth distribution in the U.S., and to study the effects of capital income tax and estate tax on wealth inequality. Most proofs and several technical details are buried in Appendices A-B.

2 Saving and bequests

Consider an economy populated by households who live for T periods. At each time t households of any age, from 0 to T are alive. We assume the age distribution of the population is uniform to avoid demographic cycles. Any household born at time s has a single child entering the economy at time $s+T$, that is, at his parents' death. Generations of households are overlapping but are linked to form dynasties. Consider in this section an household born at time s , say in the n -th generation of his dynasty. It solves a savings problem which determines its wealth at any time t in its lifetime, leaving its wealth at death to its child. The household faces idiosyncratic rate of return on wealth and earnings at birth, which remain however constant in his lifetime. Generation n is therefore associated to a rate of return on wealth r_n ¹⁸ and to earnings y_n at birth, which grow during the agent lifetime at a deterministic rate $g \geq 0$: $y(s, t) = y(s, s)e^{g(t-s)}$, where $y(s, t)$ denotes the earnings at time t of an agent born at time s (in generation n) with $y_n = y(s, s)$.¹⁹

Let $c(s, t)$ and $w(s, t)$ denote, respectively, consumption and wealth at t of the household born at s . It inherits wealth $w(s, s)$ at s from its previous generation. If $b < 1$ denotes the estate tax, $w(s, s) = (1 - b)w(s - T, s)$. Each household's momentary utility function is denoted $u(c(s, t))$. Households also have a preference for leaving bequests to their children. In particular, we assume "joy of giving" preferences for bequests: the parents' utility from bequests is $\phi(w(s, s))$, where ϕ denotes an increasing bequest function.²⁰

An household of generation n born at time s chooses a lifetime consumption path $c(s, t)$ to maximize

$$\int_t^{s+T} e^{-\rho(v-t)} u(c(s, v)) dv + \phi(w(s + T, s + T))$$

¹⁸Let r_n be net of a capital income tax ζ .

¹⁹We can also easily allow for general trend stationary earning processes across generations (with trend $g' \neq g$). Zhu (2009) allows for stochastic returns of wealth inside each generation.

²⁰Note 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."

subject to

$$\begin{aligned}\dot{w}(s, v) &= r_n w(s, v) + y_n - c(s, v) \\ w(s + T, s + T) &= (1 - b)w(s, s + T)\end{aligned}$$

where $\rho > 0$ is the discount rate and r_n and y_n are constant from the point of view of the household. In the interest of closed form solutions we make the following assumption.

Assumption 1 *Preferences satisfy:*

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

with elasticity $\sigma \geq 1$.²¹ Furthermore, we require $r_n \geq \rho$ and $\chi > 0$.²²

The dynamics of individual wealth is solved for, as a function of age $t - s$, $w(t - s)$; see Appendix A.

3 The distribution of wealth

In our economy, after-tax bequests from parents are initial wealth of children. We can construct then a discrete time map for each dynasty's wealth accumulation process. Let $w_n = w(nT, nT)$ denote the initial wealth of the n 'th dynasty. Since w_n is inherited from generation $n - 1$,

$$w_n = (1 - b)w((n - 1)T, nT) = (1 - b)w(T).$$

The rate of return of wealth and earnings are stochastic across generations. We assume they are also idiosyncratic across individual. Let $(r_n)_n$ and $(y_n)_n$ denote, respectively the stochastic process for the rate of return of wealth and earnings, over generations n .²³ We obtain a difference equation for the initial wealth of dynasties, mapping w_n into w_{n+1} :

$$w_{n+1} = \alpha_n w_n + \beta_n \tag{1}$$

²¹ $u(c) = \ln c$ and $\phi(w) = \chi \ln w$ for $\sigma = 1$.

²²The condition $r_n \geq \rho$ (on the whole support of the random variable r_n) is sufficient to guarantee that agents will not want borrow during their lifetime. The condition $\sigma \geq 1$ guarantees that r_n is larger than the endogenous rate of growth of consumption, $\frac{r_n - \rho}{\sigma}$. It is required to produce a stationary non-degenerate wealth distribution and could be relaxed if we allowed the elasticity of substitution for consumption and bequest to differ, at a notational cost. Finally, $\chi > 0$ guarantees positive bequests.

²³We 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 $(x_n)_n$.

where $(\alpha_n, \beta_n)_n = (\alpha(r_n), \beta(r_n, y_n))_n$ are stochastic processes induced by $(r_n, y_n)_n$. They are obtained as solutions of the households' savings problem and hence they endogenously depend from the deep parameters of our economy; see Appendix A, equations (4-5), for closed form solutions of $\alpha(r_n)$ and $\beta(r_n, y_n)$.

The multiplicative term 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.²⁴ We refer to α_n as the *life-cycle return to wealth*. It can be shown that $\alpha(r_n)$ is increasing in r_n . The additive component β_n can in turn be interpreted as a measure of effective lifetime (permanent) labor income, again after subtracting the affine part of consumption.

3.0.1 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 (1) is ergodic. 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.*,²⁵ we study more general stochastic processes which naturally arise when studying the distribution of wealth. A positive auto-correlation in r_n and y_n , in particular, can 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, correlation between r_n and y_n , allows e.g., for households with high labor income to have better opportunities for higher returns on wealth in financial markets.²⁶

To induce a limit stationary distribution of $(w_n)_n$ 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 impose assumptions on $(r_n, y_n)_n$ which are sufficient to guarantee the existence of a limit stationary distribution of $(w_n)_n$; see Assumption 2 and 3 in Appendix B. In terms of $(\alpha_n, \beta_n)_n$ these assumptions guarantee that $(\alpha_n, \beta_n)_n > 0$, that $E(\alpha_n | \alpha_{n-1}) < 1$ for any α_{n-1} , and finally that $\alpha_n > 1$ with positive probability; see Lemma 3 in Appendix

²⁴A 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. Also, note that, because bequests are positive under our assumptions, α_n is also positive; see the Proof of Proposition 3.

²⁵The 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.

²⁶See Arrow (1987) and McKay (2008) for models in which such correlations arise endogenously from non-homogeneous portfolio choices in financial markets.

B.²⁷ Under these assumptions we can prove the following theorem, based on a theorem in Saporta (2005).

Theorem 1 *Consider*

$$w_{n+1} = \alpha(r_n) w_n + \beta(r_n, y_n), \quad w_0 > 0.$$

Let $(r_n, y_n)_n$ satisfy Assumption 2 and 3 as well as a regularity assumption.²⁸ Then the tail of the stationary distribution of w_n , $\Pr(w_n > w)$, is asymptotic to a Pareto law

$$\Pr(w_n > w) \sim cw^{-\mu},$$

where $\mu > 1$ satisfies

$$\lim_{N \rightarrow \infty} \left(E \prod_{n=0}^{N-1} (\alpha_{-n})^\mu \right)^{\frac{1}{N}} = 1. \quad (2)$$

When $(\alpha_n)_n$ is *i.i.d.*, condition (2) reduces to $E(\alpha)^\mu = 1$, a result established by Kesten (1973) and Goldie (1991).²⁹

We now turn to the characterization of the stationary wealth distribution of the economy, aggregating over households of different ages.

3.1 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 of the economy can be constructed aggregating over the wealth of households of all ages τ from 0 to T (recall we assumed the age distribution in the population is uniform). The wealth of an household of age τ , born with wealth $w(0)$, return r_n , and income y_n , is a deterministic map, as the realizations of r_n and y_n are fixed for any household during his lifetime; see Appendix A for the closed form of $w(\tau)$.

The power tail of the initial wealth distribution implies that the wealth distribution of the whole economy displays a tail with exponent μ in the following sense:

Theorem 2 *Suppose the tail of the stationary distribution of initial wealth w_n is asymptotic to a Pareto law, $\Pr(w_n > w) \sim cw^{-\mu}$, then the stationary wealth distribution of the population has a power tail with the same exponent μ .*

²⁷In Proposition 3 we also show that the state space of $(\alpha_n, \beta_n)_n$ is well defined. Furthermore, by Assumption 2, $(r_n)_n$ converges to a stationary distribution and hence $(\alpha(r_n))_n$ also converges to a stationary distribution.

²⁸See Appendix B, proof of Theorem 1, for details.

²⁹The term $\prod_{n=0}^{N-1} \alpha_{-n}$ in 2 arises from using repeated substitutions for w_n . See Brandt (1986) for general conditions to obtain an ergodic solution for stationary stochastic processes satisfying (1).

Note that this result is independent of the demographic characteristics of the economy, that is, of the stationary distribution of the households by age.

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 tail index of the distribution of wealth, μ , which is analytically characterized in Theorem 1.

The tail index μ is inversely related to wealth inequality, as a small index μ implies a heavier top tail of the wealth distribution (the distribution declines more slowly with wealth in the tail). In fact, the exponent μ is inversely linked to the Gini coefficient G : $G = \frac{1}{2\mu-1}$, the classic statistical measure of inequality.³⁰

First, we shall study how different compositions of capital and labor income risk affect the tail index μ . Second, we shall study the effects of preferences, in particular the intensity of the bequest motive. Third, we shall characterize the effects of both capital income and estate taxes on μ . Finally, we shall address the relationship between social mobility and μ .

4.1 Capital and labor income risk

It follows from Theorem 1 that the stochastic properties of labor income risk, $(\beta_n)_n$, have no effect on the tail of the stationary wealth distribution. In fact heavy tails in the stationary distribution require that the economy has sufficient capital income risk, with $\alpha_n > 1$ with positive probability. Consider instead an economy with limited capital income risk, in which $\alpha_n < 1$ with probability 1 and $\bar{\beta}$ 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 $\frac{\bar{\beta}}{1-\bar{\alpha}}$, where $\bar{\alpha}$ is the upper bound of α_n .³¹

More generally, we can also show that wealth inequality increases with the capital income risk households face in the economy.

Proposition 1 *Consider two distinct i.i.d. processes for the rate of return on wealth, $(r_n)_n$ and $(r'_n)_n$. Suppose $\alpha(r_n)$ is a convex function of r_n .³² If r_n second order stochas-*

³⁰See e.g., Chipman (1976). A distribution with a lower μ has also a higher Lorenz curve at the top end of wealth levels; see Zhu (2009) for a proof. Since the distribution of wealth in our economy is typically Pareto only in the tail, we refer to $G = \frac{1}{2\mu-1}$ as to the "Gini of the tail."

³¹Of course this is true a fortiori in the case where there is no capital risk and $\alpha_n = \bar{\alpha} < 1$.

³²This is typically the case in our economy if constant relative risk aversion parameter σ is not too high. A sufficient condition is $2(\sqrt{2}-1)T \int_0^T t e^{A(r_n)t} dt - \frac{\sigma-1}{\sigma} \int_0^T t^2 e^{A(r_n)t} dt > 0$, which holds, e.g., if $\sigma < (1 - 2(\sqrt{2}-1))^{-1} = 4.8284$.

atically dominates r'_n , the tail index μ of the wealth distribution under $(r_n)_n$ is smaller than under $(r'_n)_n$.

We conclude that it is capital income risk (idiosyncratic risk on return on capital), and not labor income risk, that determines the heaviness of the tail of the stationary distribution given by the tail index: 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 .

Proposition 2 *The tail index μ decreases with the bequest motive χ .*

A household with a higher preference for bequests will save more and accumulate wealth faster. This saving behavior induces an higher effective rate of return of wealth across generations α_n , on average, which in turn leads to higher wealth inequality.

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.³³

Proposition 3 *The tail index μ increases with the estate tax b and with the capital income tax ζ .*

Furthermore, let $\zeta(r_n)$ denote a non-linear tax on capital, such that the net rate of return of wealth for generation n becomes $r_n(1 - \zeta(r_n))$. Since $\frac{\partial \alpha_n}{\partial r_n} > 0$, the Corollary below follows immediately from Proposition 3.

Corollary 1 *The tail index μ increases with the imposition of a non-linear tax on capital $\zeta(r_n)$.*

Taxes have therefore a dampening effect on the tail of the wealth distribution in our economy: the higher are taxes, the lower is wealth inequality. The calibration exercise in Section 2 documents that in fact the tail of the stationary wealth distribution is quite sensitive to variations in both capital income taxes and estate taxes. Becker and Tomes (1979), on the contrary, find that taxes have ambiguous effects on wealth inequality at the stationary distribution. In their model, bequests are chosen by parents to essentially offset the effects of fiscal policy, limiting any wealth equalizing aspects of these policies.

³³Recall that the random rate of return r_n in our economy is defined net of the capital income tax ζ .

This compensating effect of bequests is present in our economy as well, though it is not sufficient to offset the effects of estate and capital income taxes on the stochastic returns on capital. In other words, the power of Becker and Tomes (1979)'s compensating effect is due to the fact that their economy has no capital income risk. The main mechanism through which estate taxes and capital income taxes have an equalizing effect on the wealth distribution in our economy is by reducing the capital income risk, along the lines of Proposition 1, not its average return.

4.4 Social mobility

We turn now to the study of the effects of different degrees of social mobility on the tail of the wealth distribution. We say social mobility is higher when $(r_n)_n$ and $(y_n)_n$ (and hence when $(\alpha_n)_n$ and $(\beta_n)_n$) are less auto-correlated over time.

We provide here expressions for the tail index of the wealth distribution as a function of the auto-correlation of $(\alpha_n)_n$ in two distinct cases:³⁴

MA(1)

$$\ln \alpha_n = \eta_n + \theta \eta_{n-1}$$

AR(1)

$$\ln \alpha_n = \theta \ln \alpha_{n-1} + \eta_n$$

where $\theta < 1$ and $(\eta_n)_n$ is an *i.i.d.* process with bounded support.

Proposition 4 *Suppose that $\ln \alpha_n$ satisfies MA(1). The tail of the limiting distribution of initial wealth w_n is then asymptotic to a Pareto law with tail exponent μ_{MA} which satisfies*

$$Ee^{\mu_{MA}(1+\theta)\eta_n} = 1.$$

If instead $\ln \alpha_n$ satisfies AR(1), the tail exponent μ_{AR} satisfies

$$Ee^{\frac{\mu_{AR}}{1-\theta}\eta_n} = 1.³⁵$$

In either the MA(1) or the AR(1) case, the higher is θ , the lower is the tail exponent. That is, the more persistent is the process for the rate of return on wealth (the higher are frictions to social mobility), the thicker is the tail of the wealth distribution.³⁶

³⁴The stochastic properties of $(y_n)_n$, and hence of $(\beta_n)_n$, as we have seen, do not affect the tail index.

³⁵We thank an anonymous referee for an outline of this Proposition.

³⁶The results will easily extend to $MA(k)$ and $AR(k)$ processes for $\ln \alpha_n$.

5 An illustrative 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 heavy as those observed in the data.

The analytical results in the previous sections suggest that capital income risk should prove very helpful in matching the heavy tails. More generally, however, our results are limited to a characterization of the tail of the wealth distribution and questions might remain about the ability of our model to match the entire wealth distribution. To this end we report on an illustrative calibration exercise to match the Lorenz curve of the wealth distribution in the U.S. We shall then study the effects of capital income taxes, estate taxes, and social mobility on wealth inequality.

Our benchmark calibration postulates a rate of return r_n which is *i.i.d.* across generations (perfect social mobility). We set the before-tax rate of return to a discrete distribution, with mean 1.0921.³⁷ The distribution of labor earnings is also discrete, with mean 1.675.³⁸ These values are chosen so that the labor income share be about twice the capital income share: 1.94 in fact at the stationary distribution. We set the life span to the average working span, $T = 45$, and we set $\sigma = 2$, $\rho = 0.04$, $g = .01$, and $\chi = 0.25$. We also set the estate tax rate $b = 0.20$ (which is the average tax rate on bequests), and the capital income tax $\zeta = 0.15$. Finally, we introduce the possibility of frictions to social mobility: starting from our benchmark *i.i.d.* case for r_n , expressed as a Markov transition matrix with identical rows, we move a mass ε of probability from the off-diagonal terms in each row to the diagonal term of that row in order to introduce persistence across generations.

With this calibration we simulate the stationary distribution of the economy. We then calculate the top percentiles of the simulated wealth distribution, the Gini coefficient of the whole distribution (not just the "Gini of the tail"), the quintiles, and the tail index μ . We compare these statistics of the simulated distribution with those of the U.S. economy.³⁹

First of all, note that the simulated wealth distribution matches successfully the top percentiles of the U.S. economy.

³⁷The values of r_n are 1.08, 1.12, 1.15, and 1.32 with probability 0.8, 0.12, 0.07, and 0.01 respectively.

³⁸The values of y_n are 0.3, 1, 2, 5, 10, and 30 with probability $\frac{14}{64}$, $\frac{36}{64}$, $\frac{11}{64}$, $\frac{1}{64}$, $\frac{1}{64}$, and $\frac{1}{64}$ respectively.

³⁹For the data of U.S. economy, the tail index is from Klass et al. (2007) who use the Forbes 400 data. The rest of data for the U.S. economy are from Diaz-Gimenez et al. (2002) who use the 1998 Survey of Consumer Finances (SCF).

<i>Economy</i>	<i>Percentiles</i>		
	<i>90th – 95th</i>	<i>95th – 99th</i>	<i>99th – 100th</i>
<i>U.S.</i>	0.113	0.231	0.347
<i>Model with $\varepsilon = 0$</i>	0.099	0.207	0.385
$\varepsilon = 0.001$	0.093	0.212	0.418
$\varepsilon = 0.004$	0.088	0.230	0.461

Table 5.1: Percentiles of the top tail

More surprisingly, perhaps, the simulated wealth distribution has a Gini coefficient close to that of the U.S. economy and its Lorenz curve matches that of the U.S. economy rather well in terms of quintiles.⁴⁰

<i>Economy</i>	<i>Tail index μ</i>	<i>Gini</i>	<i>Quintiles</i>				
			<i>First</i>	<i>Second</i>	<i>Third</i>	<i>Fourth</i>	<i>Fifth</i>
<i>U.S.</i>	1.49	0.803	−0.003	0.013	0.05	0.122	0.817
<i>Model with $\varepsilon = 0$</i>	1.795	0.740	0.022	0.040	0.057	0.091	0.789
$\varepsilon = 0.001$	1.693	0.767	0.020	0.035	0.050	0.082	0.813
$\varepsilon = 0.004$	1.489	0.814	0.015	0.027	0.038	0.064	0.856

Table 5.2: Tail Index, Gini, and Quintiles

Note in particular that, by reducing social mobility (increasing ε), we can rather easily generate economies with even a thicker tail than the U.S. economy (at the benchmark parameters values). The literature on the distribution of wealth has on the contrary struggled to find significant effects of social mobility frictions and other intergenerational links; see e.g., the discussion in De Nardi (2004) on the issue. This inconsistency is in fact only apparent, once it is realized that the intergenerational links studied in the previous literature operate on labour earnings rather than on the rate of return on wealth as in our model.

5.1 Tax experiments

The Tables below illustrate the effects of taxes on the tail index and the Gini coefficient. We calibrate the parameters of the economy other than b and ζ , at the benchmark, and we vary tax rates b and ζ . Table 5.3 reports on the effects of capital income taxes and

⁴⁰The tail index μ reported in Table 5.2 is computed analytically using the results of Saporta (2005, Proposition 1, section 4.1) discussed in the proof of Theorem (1) in Appendix B.

estate taxes on the tail index μ .

$b \setminus \zeta$	0	0.05	0.1	0.15
0	1.155	1.301	1.483	1.718
0.1	1.174	1.323	1.510	1.753
0.2	1.195	1.348	1.542	1.795
0.3	1.219	1.378	1.579	1.844

Table 5.3: Tax experiments - Tail index μ

Taxes have a significant effect on the inequality of the wealth distribution as measured by the tail index. This is especially the case for the capital income tax, which directly affects the stochastic returns on wealth. The implied "Gini of the tail" changes from .763 with no taxes to .372 with a 30% estate tax and a 15% capital income tax.

We now turn to the Gini coefficient of the whole distribution. The results are in Table 5.4.

$b \setminus \zeta$	0	0.05	0.1	0.15
0	.838	.790	.772	.734
0.1	.842	.787	.760	.741
0.2	.821	.785	.757	.740
0.3	.832	.783	.761	.735

Table 5.4: Tax experiments - Gini Coefficient

We see that the Gini coefficient consistently declines as the capital income tax increases, but the decline is quite moderate, and the estate taxes can even have ambiguous effects.⁴¹ The mechanism which might induce such counterintuitive effects of a tax increase is intuitively understood as follows. A tax increase has the effect of substantially reducing the concentration of wealth in the tail of the distribution (this is clear from Table 5.3). This effect is however partly compensated for by greater inequality at lower wealth levels. The decrease in the rate of return on wealth due to the increase in tax rates has in fact a positive wealth effect, as it increases the permanent labor income of households (future labor earnings are discounted at a lower rate). This wealth effect is relatively larger for households whose physical wealth is relatively low. These households will smooth their consumption based on their lifetime labor earnings, and will hence react to a tax increase by decumulating physical wealth proportionately faster than households

⁴¹In 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. The effect is larger the larger is the capital income tax rate.

that are relatively rich in physical wealth. As a result of this effect, wealth inequality between rich and poor households as measured by physical wealth tends to increase.

This result has to be interpreted with caution, however. The counterintuitive effects of a tax increase on wealth inequality in Table 5.4 would be moderated (perhaps eliminated) if tax revenues were to be redistributed in the economy. Nonetheless Table 5.4 suggests that it might be worth re-evaluating the effects on wealth inequality of various proposed fiscal policies available in the literature. For instance, Castaneda, Diaz Jimenez, and Rios Rull (2003) and Cagetti and De Nardi (2007) find very small (or even perverse) effects of eliminating bequest taxes in their calibrations in models with a skewed distribution of earnings but no capital income risk.⁴² If the capital income risk component is a substantial fraction of idiosyncratic risk, such fiscal policies could have a sizeable effect in increasing the wealth concentrated in the top tail of the distribution which may not show up in measurements of the Gini coefficient.⁴³

6 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 heavy 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, the top tail of wealth distribution 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 heavy tails.

Interestingly, this role of capital income risk as a determinant of the distribution of wealth seems to have been lost by Vilfredo Pareto. He explicitly noted that an identical stochastic process for wealth across households 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 labor earnings of households (1897, Notes to #962, p. 416). Left with the distribution of talents and earnings as the main determinant of the wealth distribution, he was perhaps

⁴²See also our discussion of the results of Becker and Tomes (1979), previously in this section.

⁴³Empirical studies also indicate that higher and more progressive taxes did in fact significantly reduce income and wealth inequality in the historical context; notably, e.g., Lampman (1962) and Kuznets (1955). Most recently, Piketty (2001) and Piketty and Saez (2003) have argued that redistributive capital and estate taxation may have prevented holders of very large fortunes from recovering from the shocks that they experienced during the Great Depression and World War II because of the dynamic effects of progressive taxation on capital accumulation and pre-tax income inequality. This line of argument has been extended to the U.S., Japan, and Canada, respectively, by Moriguchi-Saez (2005), Saez-Veall (2003).

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.*⁴⁴

⁴⁴See 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.
- Angeletos, G. and L.E. Calvet (2005), "Incomplete-market dynamics in a neoclassical production economy," *Journal of Mathematical Economics*, 41(4-5), 407-438.
- Angeletos, G. and L.E. Calvet (2006), "Idiosyncratic Production Risk, Growth and the Business Cycle", *Journal of Monetary Economics*, 53, 1095-1115.
- 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.
- 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.
- Carroll, C. D. (1997). "Buffer-Stock Saving and the Life Cycle/Permanent Income Hypothesis," *Quarterly Journal of Economics*, 112, 1–56.
- 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.
- 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.
- Díaz-Giménez, J., V. Quadrini, J. V. Ríos-Rull, and S. B. Rodríguez (2002): "Updated Facts on the U.S. Distributions of Earnings, Income, and Wealth", *Federal Reserve Bank of Minneapolis Quarterly Review*, 26(3), 2-35.
- 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.
- 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.
- Fiaschi, D. and M. Marsili (2009): "Distribution of Wealth and Incomplete Markets: Theory and Empirical Evidence," University of Pisa Working paper 83.

- 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.
- Gabaix, X. (1999): "Zipf's law for cities: An Explanation," *Quarterly Journal of Economics*, 114(3), 739-67.
- Goldie, C. M. (1991): "Implicit Renewal Theory and Tails of Solutions of Random Equations," *Annals of Applied Probability*, 1, 126-166.
- Guvenen, F. (2007): "Learning Your Earning: Are Labor Income Shocks Really Very Persistent?," *American Economic Review*, 97(3), 687-712.
- Heathcote, J. (2008): "Discussion *Heterogeneous Life-Cycle Profiles, Income Risk, and Consumption Inequality*, by G. Primiceri and T. van Rens," Federal Bank of Minneapolis, mimeo.
- Huggett, M. (1993): "The Risk-Free Rate in Heterogeneous-household 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.
- 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.
- Levy, M. (2005), " Market Efficiency, The Pareto Wealth Distribution, and the Levy Distribution of Stock Returns" in *The Economy as an Evolving Complex System* eds. S. Durlauf and L. Blume, Oxford University Press, USA.

- Levy, M. and S. Solomon (1996): "Power Laws are Logarithmic Boltzmann Laws," *International Journal of Modern Physics*, C,7, 65-72.
- 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 (2007): "Two Review of Income and Wealth, 53(3), 440-59.
- 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.
- Panousi, V. (2008): "Capital Taxation with Entrepreneurial Risk," mimeo, MIT.
- Primiceri, G. and T. van Rens (2006): "Heterogeneous Life-Cycle Profiles, Income Risk, and Consumption Inequality," CEPR Discussion Paper 5881.
- 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/t00007666.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.
- Sornette, D. (2000): *Critical Phenomena in Natural Sciences*, Berlin, Springer Verlag.
- Storesletten, K., C.I. Telmer, and A. Yaron (2004): "Consumption and Risk Sharing Over the Life Cycle," *Journal of Monetary Economics*, 51(3), 609-33.
- 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.
- 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.
- Zhu, Shenghao (2009): "Comparative Statics of the Stationary Distribution of a Linear Model," mimeo, NYU.

7 Appendix A: Closed form solutions

We report here only the closed form solutions for the dynamics of wealth in the paper. The details of the computations are available online at <http://www.nyu.edu/econ/user/bisina/> and at <http://www.econ.nyu.edu/user/benhabib/>.

Let an individual's age be denoted $\tau = t - s$. Let human capital of an household born at s at time t , $h(s, t)$, be defined as $h(s, t) = \int_t^{s+T} y_n e^{-r\tau} d\tau$. The optimal consumption path satisfies

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

The propensity to consume out of financial and human wealth, $m(\tau)$, is independent of $w(s, t)$ and $h(s, t)$, is decreasing in age τ , in the estate tax b , and in capital income tax ζ :

$$m(\tau) = \frac{1}{\frac{1}{r_n - \frac{r_n - \rho}{\sigma}} \left(1 - e^{-(r_n - \frac{r_n - \rho}{\sigma})(T - \tau)} \right) + \chi^{\frac{1}{\sigma}} (1 - b)^{\frac{1 - \sigma}{\sigma}} e^{-(r_n - \frac{r_n - \rho}{\sigma})(T - \tau)}} \quad (3)$$

Let $w(t - s)$ be the wealth of an household of age $t - s$ born with wealth $w(0)$. The dynamics of individual wealth as a function of age τ satisfies

$$w(t - s) = \sigma_w(r_n, t - s)w(0) + \sigma_y(r_n, t - s)y(0)$$

with

$$\begin{aligned} \sigma_w(r_n, \tau) &= e^{r_n \tau} \frac{e^{A(T - \tau)} + AB - 1}{e^{AT} + AB - 1} \\ \sigma_y(r_n, \tau) &= e^{r_n \tau} \frac{e^{(g - r)T} - 1}{g - r} \left(\frac{e^{A(T - \tau)} + AB - 1}{e^{AT} + AB - 1} - \frac{e^{r_n(T - \tau)} - 1}{e^{r_n T} - 1} \right) \end{aligned}$$

The dynamics of wealth across generation is then:

$$w_{n+1} = \alpha_n w_n + \beta_n$$

with

$$\alpha(r_n) = (1 - b)e^{r_n T} \frac{A(r_n)B(b)}{e^{A(r_n)T} + A(r_n)B(b) - 1} \quad (4)$$

and

$$\beta(r_n, y_n) = (1 - b)y_n \frac{e^{(g - r)T} - 1}{g - r} e^{r_n T} \frac{A(r_n)B(b)}{e^{A(r_n)T} + A(r_n)B(b) - 1} \quad (5)$$

8 Appendix B: Proofs

The stochastic processes for (r_n, y_n) and the induced processes for $(\alpha_n, \beta_n)_n = (\alpha(r_n), \beta(r_n, y_n))_n$ are required to satisfy the following assumptions.

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

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

where $\Pr(r_n, y_n \mid r_{n-1}, y_{n-1})$ denotes the conditional probability of (r_n, y_n) given (r_{n-1}, y_{n-1}) .⁴⁵

A stochastic process $(r_n, y_n)_n$ which satisfies Assumption 2 is a *Markov Modulated chain*. This assumption would be satisfied, for instance, if a single Markov chain, corresponding e.g., to productivity shocks, drove returns on capital $(r_n)_n$, as well as labor income $(y_n)_n$.⁴⁶

Assumption 3 *Let P denote the transition matrix of $(r_n)_n$: $P_{i\bar{y}} = \Pr(r_i \mid r_{\bar{y}})$. Let $\alpha(\bar{\mathbf{r}})$ denote the state space of $(\alpha_n)_n$ as induced by the map $\alpha(r_n)$. Then $\bar{\mathbf{r}}, \bar{\mathbf{y}}$ and P are such that: (i) $\bar{\mathbf{r}} \times \bar{\mathbf{y}} \gg \mathbf{0}$, (ii) $P\alpha(\bar{\mathbf{r}}) < \mathbf{1}$, (iii) $\exists \bar{r}_i$ such that $\alpha(\bar{r}_i) > 1$, (iv) $P_{ii} > 0$, for any i .*

We are now ready to show:

Lemma 3 *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) $E(\alpha_n \mid \alpha_{n-1}) < 1$, for any α_{n-1} ,⁴⁷ (iii) $\bar{\alpha}_i > 1$ for some $i = 1, \dots, m$, (iv) the diagonal elements of the transition matrix P of α_n are positive.*

Proof of Lemma 3. Let A be the diagonal matrix with elements $A_{ii} = \bar{\alpha}_i$, and $A_{ij} = 0, j \neq i$. Note that $E(\alpha_n \mid \alpha_{n-1})$, for any α_{n-1} can be written as $P\alpha(\bar{\mathbf{r}}) < \mathbf{1}$. Let $\bar{\mathbf{r}} = \{\bar{r}_1, \dots, \bar{r}_m\}$ denote the state space of r_n . Similarly, let $\bar{\mathbf{y}} = \{\bar{y}_1, \dots, \bar{y}_l\}$ denote the state space of y_n . Let $\bar{\alpha} = \{\bar{\alpha}_1, \dots, \bar{\alpha}_m\}$ and $\bar{\beta} = \{\bar{\beta}_1, \dots, \bar{\beta}_l\}$ denote the state spaces of, respectively, α_n and β_n , as they are induced through the maps 4 and 5. We shall show that the maps 4 and 5 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 and bounded with probability 1 in r_n and y_n . Recall that $B(b) = \chi^{\frac{1}{\sigma}} (1 - b)^{\frac{1-\sigma}{\sigma}} > 0$. Note that

$$\alpha(r_n) = (1 - b) \frac{B(b)}{e^{-\frac{r_n - \rho}{\sigma} T} \int_0^T e^{-A(r_n)(T-t)} dt + e^{-r_n T} B(b)}$$

⁴⁵While Assumption 2 requires r_n to be independent of $(y_{n-1}, y_{n-2}, \dots)$, it leaves the auto-correlation of $(r_n)_n$ unrestricted, in the space of Markov chains. Also, Assumption 2 allows for (a restricted form of) auto-correlation of $(y_n)_n$ as well as the correlation of y_n and r_n .

⁴⁶For 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 Roiterstein (2007) for general Markov Modulated processes.

⁴⁷We could only require that the mean of the unconditional distribution of α is less than 1, that is if $E(\alpha) < 1$. But in this case the stationary distribution of wealth may not have a mean.

Therefore $\alpha_n > 0$ and bounded. Furthermore, note that

$$\beta(r_n, y_n) = \alpha(r_n) y_n \int_0^T e^{(g-r_n)t} dt$$

and the support of y_n is bounded by Assumption 2. Thus $(\beta_n)_n \geq 0$ and is bounded. Therefore (α_n, β_n) is a Markov Modulated Process provided $(\beta_n)_n$ is positive and bounded.

Furthermore, Assumption 3 (ii) implies directly that (ii) $P\bar{\alpha} < \mathbf{1}$. Assumption 3 (iii) also directly implies $\bar{\alpha}_i > 1$ for some $i = 1, \dots, m$. 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. ■

Proof of Theorem 1. We first define rigorously the Regularity of the Markov Modulated process $(\alpha_n, \beta_n)_n$. 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:⁴⁸

The Markov Modulated process $(\alpha_n, \beta_n)_n$ is regular, that is

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

*and the elements of the vector $\bar{\alpha} = \{\ln \bar{\alpha}_1 \dots \ln \bar{\alpha}_m\} \subset \mathbb{R}_+^m$ are not integral multiples of the same number.*⁴⁹

Saporta (2005, Proposition 1, section 4.1) establishes that, for finite Markov chains,

$$\lim_{N \rightarrow \infty} \left(E \prod_{n=0}^{N-1} (\alpha_{-n})^\mu \right)^{\frac{1}{N}} = \lambda(A^\mu P'), \text{ where } \lambda(A^\mu P') \text{ is the dominant root of } A^\mu P'.^{50}$$

Condition 2 can then be expressed as $\lambda(A^\mu P') = 1$. The Theorem then follows directly from Saporta (2005), Theorem 1, if we show i) that there exists a μ that solves $\lambda(A^\mu P') = 1$, and that ii) such μ is > 1 . Saporta shows that $\mu = 0$ is a solution to $\lambda(A^\mu P') = 1$, or equivalently to $\ln(\lambda(A^\mu P')) = 0$. This follows from $A^0 = I$ and P being a stochastic matrix. Let $E\alpha(r)$ denote the expected value of α_n at its stationary distribution (which exists as it is implied by the ergodicity of $(r_n)_n$, in turn a consequence of Assumption 2). Saporta, under the assumption $E\alpha(r) < 1$, shows that $\frac{d \ln \lambda(A^\mu P')}{d\mu} < 0$ at $\mu = 0$, and that $\ln(\lambda(A^\mu P'))$ is a convex function of μ .⁵¹ Therefore, if there exists another solution $\mu > 0$ for $\ln(\lambda(A^\mu P')) = 0$, it is positive and unique.

⁴⁸We formulate these regularity conditions on $(\alpha_n, \beta_n)_n$, but they can be immediately mapped back into conditions on the stochastic process $(r_n, y_n)_n$.

⁴⁹Theorems which characterize the tails of distributions generated by equations with random multiplicative coefficients 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).

⁵⁰Recall that the matrix AP' has the property that the i 'th column sum equals the expected value of α_n conditional on $\alpha_{n-1} = \bar{\alpha}_i$. When $(\alpha_n)_n$ 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 $\lambda(A^\mu P')$.

⁵¹This follows because $\lim_{n \rightarrow \infty} \frac{1}{n} \ln E(\alpha_0 \alpha_{-1} \dots \alpha_{-n})^\mu = \ln(\lambda(A^\mu P'))$ and because the moments of non-negative random variables are log-convex (in μ); see Loeve(1977), p. 158.

To assure that $\mu > 1$ we replace the condition $E\alpha(r) < 1$ with (ii) of Proposition 3, $P\bar{\alpha} < \mathbf{1}$. 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 $\mu = 1$, $\lambda(A^\mu P') = \lambda(AP') < 1$. Now note that if $(\alpha_n, \beta_n)_n$ is reflective, by Proposition 3, $P_{ii} > 0$ and $\bar{\alpha}_i > 1$, 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 $A^\mu P'$, $\lambda(A^\mu P')$, goes to infinity with μ . It follows that for the solution of $\ln(\lambda(A^\mu P')) = 0$, we must have $\mu > 1$. This proves ii). ■

Proof of Theorem 2. If we denote with ν the product measure of the stationary distribution of $(w(0), r_n)$, the cumulative distribution function of the stationary distribution of wealth of household of age τ , $F_\tau(w)$ is given by the map (7).

$$F_\tau(w) = \sum_{i=1}^l \left(\Pr(y_i) \int I_{\{\sigma_w(r_n, \tau)w(0) + \sigma_y(r_n, \tau)y_i \leq w\}} d\nu \right)$$

where I is an indicator function.

The cumulative distribution function of wealth w in the population is then

$$F(w) = \int_0^T F_\tau(w) \frac{1}{T} d\tau$$

Note that

$$P(w(\tau) > w) = \sum_{i=1}^l \left(\Pr(y_i) \int I_{\{\sigma_w(r_n, \tau)w(0) + \sigma_y(r_n, \tau)y_i > w\}} d\nu \right)$$

and $\sigma_w(r_n, \tau)$ and $\sigma_y(r_n, \tau)$ are continuous functions of r_n and τ . Since the number of states of r_n is finite and $\tau \in [0, T]$, there exist σ_w^L , σ_w^U , and σ_y^U such that $0 < \sigma_w^L \leq \sigma_w(r_n, \tau) \leq \sigma_w^U$ and $\sigma_y(r_n, \tau) \leq \sigma_y^U$. We have

$$I_{\{\sigma_w(r_n, \tau)w(0) + \sigma_y(r_n, \tau)y_i > w\}} \geq I_{\{\sigma_w^L w(0) > w\}}$$

and

$$I_{\{\sigma_w(r_n, \tau)w(0) + \sigma_y(r_n, \tau)y_i > w\}} \leq I_{\{\sigma_w^U w(0) + \sigma_y^U \bar{y}_l > w\}}$$

Hence

$$P\left(w(0) > \frac{w}{\sigma_w^L}\right) \leq P(w(\tau) > w) \leq P\left(w(0) > \frac{w - \sigma_y^U \bar{y}_l}{\sigma_w^U}\right)$$

We know

$$P(W > w) = \int_0^T P(w(\tau) > w) \frac{1}{T} d\tau$$

Thus

$$P\left(w(0) > \frac{w}{\sigma_w^L}\right) \leq P(W > w) \leq P\left(w(0) > \frac{w - \sigma_y^U \bar{y}_l}{\sigma_w^U}\right)$$

Thus

$$(\sigma_w^L)^\mu c \leq \liminf_{w \rightarrow +\infty} \frac{P(W > w)}{w^{-\mu}} \leq \limsup_{w \rightarrow +\infty} \frac{P(W > w)}{w^{-\mu}} \leq (\sigma_w^U)^\mu c$$

since $\lim_{w \rightarrow +\infty} \frac{P(w(0) > w)}{w^{-\mu}} = c$. We conclude that the wealth distribution in the population, W , has a power tail with the same exponent μ , i.e.,

$$0 < c_1 \leq \liminf_{w \rightarrow +\infty} \frac{\Pr(W > w)}{w^{-\mu}} \leq \limsup_{w \rightarrow +\infty} \frac{\Pr(W > w)}{w^{-\mu}} \leq c_2. \quad \blacksquare$$

We can also show that,

When $(r_n)_n$ is *i.i.d.*, the asymptotic power law property with the same power μ is preserved for each age cohort and the whole economy: $\exists \tilde{c} > 0$ such that

$$\lim_{w \rightarrow +\infty} \frac{\Pr(W > w)}{w^{-\mu}} = \tilde{c}.$$

Proof. When $(r_n)_n$ is *i.i.d.*

$$\begin{aligned} P(W > w) &= \int_0^T P(w(\tau) > w) \frac{1}{T} d\tau \\ &= \sum_{i=1}^l \sum_{j=1}^m \left(\Pr(y_i) \Pr(r_j) \int_0^T P\left(w(0) > \frac{w - \sigma_y(r_j, \tau) y_i}{\sigma_w(r_j, \tau)}\right) \frac{1}{T} d\tau \right) \end{aligned}$$

Since $\sigma_w(r_j, \tau)$ and $\sigma_y(r_j, \tau)$ are continuous functions of τ on $[0, T]$, there exist $\tilde{\tau}_j, \hat{\tau}_j \in [0, T]$ such that for $\forall t \in [0, T]$, $\sigma_w(r_j, \tau) \leq \sigma_w(r_j, \tilde{\tau}_j)$ and $\sigma_y(r_j, \tau) \leq \sigma_y(r_j, \hat{\tau}_j)$. Thus

$$P\left(w(0) > \frac{w - \sigma_y(r_j, \tau) y_i}{\sigma_w(r_j, \tau)}\right) \leq P\left(w(0) > \frac{w - \sigma_y(r_j, \hat{\tau}_j) y_i}{\sigma_w(r_j, \tilde{\tau}_j)}\right)$$

When w is sufficiently large,

$$\frac{P\left(w(0) > \frac{w - \sigma_y(r_j, \hat{\tau}_j) y_i}{\sigma_w(r_j, \tilde{\tau}_j)}\right)}{w^{-\mu}} \text{ is bounded,}$$

since $\lim_{w \rightarrow +\infty} \frac{P(w(0) > w)}{w^{-\mu}} = c$. Thus by bounded convergence theorem, we have

$$\lim_{w \rightarrow +\infty} \int_0^T \frac{P\left(w(0) > \frac{w - \sigma_y(r_j, \tau) y_i}{\sigma_w(r_j, \tau)}\right)}{w^{-\mu}} \frac{1}{T} d\tau = \int_0^T \lim_{w \rightarrow +\infty} \frac{P\left(w(0) > \frac{w - \sigma_y(r_j, \tau) y_i}{\sigma_w(r_j, \tau)}\right)}{w^{-\mu}} \frac{1}{T} d\tau$$

Thus

$$\begin{aligned} \lim_{w \rightarrow +\infty} \frac{P(W > w)}{w^{-\mu}} &= \sum_{i=1}^l \sum_{j=1}^m \left(\Pr(y_i) \Pr(r_j) \int_0^T \lim_{w \rightarrow +\infty} \frac{P\left(w(0) > \frac{w - \sigma_y(r_j, \tau) y_i}{\sigma_w(r_j, \tau)}\right) \frac{1}{T} d\tau \right) \\ &= c \sum_{j=1}^m \left(\Pr(r_j) \int_0^T (\sigma_w(r_j, \tau))^\mu d\tau \right). \blacksquare \end{aligned}$$

Proof of Proposition 1. Since $\mu > 1$, $(\alpha_n)^\mu$ is an increasing convex function in α_n . If $\alpha(r_n)$ is a convex function of r_n , then $\alpha(r_n)^\mu$ is also a convex function of r_n . And hence $-\alpha(r_n)^\mu$ is a concave function of r_n . By the second order stochastic dominance we have $E(-\alpha(r_n)^\mu) \geq E(-\alpha(r'_n)^\mu)$ so $E\alpha(r_n)^\mu \leq E\alpha(r'_n)^\mu$ and $1 = E\alpha(r_n)^\mu \leq E\alpha(r'_n)^\mu$. Let μ' solve $E\alpha(r'_n)^{\mu'} = 1$. Suppose $\mu' > \mu$. By Holder's inequality we have $E\alpha(r'_n)^\mu < (E\alpha(r'_n)^{\mu'})^{\frac{\mu}{\mu'}} = 1$. This is a contradiction. Thus we have $\mu' \leq \mu$. \blacksquare

Proof of Proposition 2. From the definition of α_n , we have

$$\alpha(r_n) = \frac{(1-b)e^{r_n T}}{\chi^{-\frac{1}{\sigma}} (1-b)^{\frac{\sigma-1}{\sigma}} \int_0^T e^{A(r_n)t} dt + 1}$$

it is easy to show that $\frac{\partial \alpha_n}{\partial \chi} > 0$. Thus an infinitesimal increase in χ shifts the state space a to the right. Therefore elements of the non-negative matrix $[A^\mu P']$ increase, which implies that the dominant root $\lambda(A^\mu P')$ increases. However we know from Saporta (2005) that $\ln(\lambda(A^\mu P'))$ is a convex function of μ .⁵² At $\mu = 0$ it is equal to zero, since A^0 is the identity matrix and P is a stochastic matrix with dominant root equal to unity. At $\mu = 0$ the function $\ln(\lambda(A^\mu P'))$ is also decreasing. (See Saporta (2005), Proposition 2, p.1962.) Then $\ln(\lambda(A^\mu P'))$ must be increasing at the positive value of μ which solves $\ln \lambda(A^\mu P') = 0$. Therefore to preserve $\ln(\lambda(A^\mu P')) = 0$, μ must decline. \blacksquare

Proof of Proposition 3. From (4), we have

$$\alpha(r_n) = \frac{e^{r_n T}}{\chi^{-\frac{1}{\sigma}} (1-b)^{-\frac{1}{\sigma}} \int_0^T e^{A(r_n)t} dt + (1-b)^{-1}}$$

Thus $\frac{\partial \alpha_n}{\partial b} < 0$. To see $\frac{\partial \alpha_n}{\partial \zeta} < 0$, we rewrite the expression of $\alpha(r_n)$ as

$$\alpha(r_n) = (1-b) \frac{B(b)}{e^{-\frac{r_n - \rho}{\sigma} T} \int_0^T e^{-A(r_n)(T-t)} dt + e^{-r_n T} B(b)}$$

Note that $A(r_n) = r_n - \frac{r_n - \rho}{\sigma}$ and $B(b) = \chi^{\frac{1}{\sigma}} (1-b)^{\frac{1-\sigma}{\sigma}}$. $\frac{\partial A(r_n)}{\partial r_n} = \frac{\sigma-1}{\sigma} \geq 0$, since $\sigma \geq 1$ by Assumption 1. And also $B(b) > 0$. Thus $\frac{\partial \alpha_n}{\partial r_n} > 0$. Higher ζ means lower r_n . We have

⁵²See Loeve (1977), page 158.

$\frac{\partial \alpha_n}{\partial \zeta} < 0$. Now the proof is identical to the proof of Proposition 2 in the reverse direction since $\frac{\partial \alpha_n}{\partial b} < 0$ and $\frac{\partial \alpha_n}{\partial \zeta} < 0$ whereas $\frac{\partial \alpha_n}{\partial \chi} > 0$. ■

Proof of Proposition 4. From

$$\ln \alpha_n = \eta_n + \theta \eta_{n-1}$$

we have

$$\sum_{t=1}^n \ln \alpha_t = \theta \eta_0 + \eta_n + \sum_{t=1}^{n-1} (1 + \theta) \eta_t$$

Thus

$$\begin{aligned} \lim_{n \rightarrow +\infty} \frac{1}{n} \ln \left(E \left(\prod_{t=1}^n \alpha_t \right)^\mu \right) &= \lim_{n \rightarrow +\infty} \frac{1}{n} \ln \left(E e^{\mu \sum_{t=1}^n \ln \alpha_t} \right) = \\ \lim_{n \rightarrow +\infty} \frac{1}{n} \ln E e^{\mu \sum_{t=1}^{n-1} (1+\theta) \eta_t} &= \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{t=1}^{n-1} \ln E e^{\mu(1+\theta) \eta_t} = \\ \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{t=1}^{n-1} \ln E e^{\mu(1+\theta) \eta_t} &= \ln E e^{\mu(1+\theta) \eta_t} \end{aligned}$$

Thus $\lim_{n \rightarrow +\infty} \frac{1}{n} \ln \left(E \left(\prod_{t=1}^n \alpha_t \right)^\mu \right) = 0$ implies

$$E e^{\mu(1+\theta) \eta_t} = 1.$$

Consider in turn the case in which

$$\ln \alpha_n = \theta \ln \alpha_{n-1} + \eta_n.$$

We have

$$\sum_{t=1}^n \ln \alpha_t = \frac{\theta(1 - \theta^n)}{1 - \theta} \ln \alpha_0 + \sum_{t=1}^n \frac{1 - \theta^{n-t+1}}{1 - \theta} \eta_t$$

Thus

$$\begin{aligned} \lim_{n \rightarrow +\infty} \frac{1}{n} \ln \left(E \left(\prod_{t=1}^n \alpha_t \right)^\mu \right) &= \lim_{n \rightarrow +\infty} \frac{1}{n} \ln \left(E e^{\mu \sum_{t=1}^n \ln \alpha_t} \right) = \\ \lim_{n \rightarrow +\infty} \frac{1}{n} \ln \left(E e^{\mu \sum_{t=1}^n \frac{1 - \theta^{n-t+1}}{1 - \theta} \eta_t} \right) &= \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{t=1}^n \ln \left(E e^{\frac{1 - \theta^{n-t+1}}{1 - \theta} \mu \eta_t} \right) = \ln \left(E e^{\frac{1}{1 - \theta} \mu \eta_t} \right) \end{aligned}$$

Thus $\lim_{n \rightarrow +\infty} \frac{1}{n} \ln \left(E \left(\prod_{t=1}^n \alpha_t \right)^\mu \right) = 0$ implies

$$E e^{\frac{\mu}{1 - \theta} \eta_t} = 1. \quad \blacksquare$$