Econometrica, Vol. 79, No. 1 (January, 2011), 123–157

THE DISTRIBUTION OF WEALTH AND FISCAL POLICY IN ECONOMIES WITH FINITELY LIVED AGENTS

BY JESS BENHABIB, ALBERTO BISIN, AND SHENGHAO ZHU[1]

We study the dynamics of the distribution of wealth in an overlapping generation economy with finitely lived agents and intergenerational 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.

KEYWORDS: Wealth distribution, Pareto, fat tails, capital income risk.

1. INTRODUCTION

RATHER INVARIABLY ACROSS A LARGE CROSS SECTION of countries and time periods income and wealth distributions are skewed to the right[2] and display heavy upper tails,[3] that is, slowly declining top wealth shares. The top 1% of the richest households in the United States hold over 33% of wealth[4] and the

[1]We gratefully acknowledge Daron Acemoglu's extensive comments on an earlier paper on the same subject, which led us to the formulation in this paper. We also acknowledge the ideas and suggestions of Xavier Gabaix and five referees that we incorporated into the paper, as well as conversations with Marco Bassetto, Gerard Ben Arous, Alberto Bressan, Bei Cao, In-Koo Cho, Gianluca Clementi, Isabel Correia, Mariacristina De Nardi, Raquel Fernandez, 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, Benoite 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.

[2]Atkinson (2002), Moriguchi and Saez (2005), Piketty (2003), Piketty and Saez (2003), and Saez and Veall (2003) documented skewed distributions of income with relatively large top shares consistently over the last century, respectively, in the United Kingdom, Japan, France, the United States, and Canada. Large top wealth shares in the United States since the 1960s were also documented, for example, by Wolff (1987, 2004).

[3]Heavy upper tails (power law behavior) for the distributions of income and wealth are also well documented, for example, by Nirei and Souma (2004) for income in the United States and Japan from 1960 to 1999, by Clementi and Gallegati (2005) for Italy from 1977 to 2002, and by Dagsvik and Vatne (1999) for Norway in 1998.

[4]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

 DOI: 10.3982/ECTA8416

top end of the wealth distribution obeys a Pareto law, the standard statistical model for heavy upper tails.[5]

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 savings paths. We aim at understanding first of all heavy upper tails, as they represent one of the main empirical features of wealth inequality.[6]

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 indeed studies 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 by Carroll (1997) and by Quadrini (1999), is that at higher wealth levels, the incentives for further precautionary savings taper off and the tails of wealth distribution remain thin. To generate skewness with heavy tails in wealth distribution, a number of authors have, therefore, successfully introduced new features, like, for example, preferences for bequests, entrepreneurial talent that generates stochastic returns (Quadrini (1999, 2000), De Nardi (2004), Cagetti and De Nardi (2006)),[7] and heterogeneous discount rates that follow an exogenous stochastic process (Krusell and Smith (1998)).

Our model is related to these papers. We study an overlapping generations economy where households are finitely lived and have a "joy of giving" bequest motive. Furthermore, to capture entrepreneurial risk, we assume households

of the distribution of wealth in the United States in 1992 is .78, while it is only .57 for the distribution of income (Diaz-Gimenez, Quadrini, and Rios-Rull (1997)); see also Feenberg and Poterba (2000).

[5]Using the richest sample of the United States, the Forbes 400—during 1988–2003, Klass, Biham, Levy, Malcai, and Solomon (2007) found, for example, that the top end of the wealth distribution obeys a Pareto law with an average exponent of 1.49.

[6]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, for example, Sornette (2000) for a survey. For early applications to the distribution of wealth, see, for example, Champernowne (1953), Rutherford (1955), and Wold and Whittle (1957). For the recent econophysics literature on the subject, see, for example, Mantegna and 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, for example, 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–savings decisions.

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

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,[8] and (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, although persistent income shocks are often allowed to explain the cross-sectional distribution of consumption; see, for example, Storesletten, Telmer, and Yaron (2004). While some authors (e.g., Primiceri and van Rens (2006)), adopted a nonstationary specification for individual income, it seems hardly the case that such a specification is suggested by income and consumption data; see for example, the discussion of Primiceri and van Rens (2006) by Heathcote (2008).[9]

The assumption that capital income contains a relevant idiosyncratic component is not standard in macroeconomics, although Angeletos and Calvet (2006) and Angeletos (2007) introduced it to study aggregate savings and growth.[10] 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 United States according to the 2001 *Survey of Consumer Finances* (SCF; Wolff (2004) and Bertaut and Starr-McCluer (2002)).[11] Case and Shiller (1989) documented a large standard deviation, on the order of 15%, of yearly capital gains or losses on owner-occupied housing. Similarly, Flavin and Yamashita (2002) measured the standard deviation of the return on housing, at the level of individual houses, from the 1968–1992 waves of the *Panel Study of Income Dynamics* (see http://psidonline.isr.umich.edu/), obtaining a similar number, 14%. Returns on private equity have an even higher idiosyncratic dispersion across households, 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), both the capital gains and earnings on private equity exhibit very substantial variation, as does excess returns to private over public equity investment, even conditional

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

[9]See Heathcote, Storesletten, and Violante (2008) for an extensive survey.

[10]See also Angeletos and Calvet (2005) and Panousi (2008).

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

on survival.[12] Evidently, the presence of moral hazard and other frictions renders complete risk diversification or concentration of each household's wealth under the best investment technology hardly feasible.[13]

Under these assumptions on labor and capital income risk,[14] 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 then tends to dominate the distribution of wealth in the tail (for high wealth). This is why Bewley models, calibrated to earnings 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.[15]

An overview of our analysis is useful to navigate over technical details. If w_{n+1} is the initial wealth of an nth 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 representing, respectively, the effective rate of return on wealth across generations and the permanent income of a generation. If α_{n+1} and β_{n+1} are independent and identically distributed (i.i.d.) processes, this dynamics of wealth converges to a stationary distribution with a Pareto law

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

with an explicit expression for μ in terms of the process for α_{n+1} (μ turns out to be independent of β_{n+1}).[16]

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

[13]See Bitler, Moskowitz, and Vissing-Jorgensen (2005).

[14]Although we emphasize the interpretation with stochastic returns, our model also accommodates a reduced form interpretation of stochastic discounting as in Krusell and Smith (1998).

[15]An alternative approach to generate fat tails without stochastic returns or discounting is to introduce a "perpetual youth" model with bequests, where the probability of death (and/or retirement) is independent of age. In these models, the stochastic component is not stochastic returns or discount rates but the length of life. For models that embody such features, see Wold and Whittle (1957), Castaneda, Diaz-Gimenez, and Rios-Rull (2003), and Benhabib and Bisin (2006).

[16]See Kesten (1973) and Goldie (1991).

But α_{n+1} and β_{n+1} are endogenously determined by the life-cycle savings and bequest behavior of households. Only by studying the life-cycle choices of households can we characterize the dependence of the distribution of wealth—and of wealth inequality in particular—on the various structural parameters of the economy, for example, 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' savings 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, for example, 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 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, 2005) and to Roitershtein (2007)), we are able to show that even in this case the stationary wealth distribution has a Pareto tail, and we can compute the effects of social mobility on the tail analytically.[17]

Finally, we calibrate and simulate our model to obtain the full wealth distribution, rather than just the tail. The model performs well in matching the (Lorenz curve of the) empirical distribution of wealth in the United States.[18]

Section 2 introduces the household's life-cycle consumption and savings 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 the 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 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 United States, and to study the effects of capital

[17]Champernowne (1953) authored the first paper to explore 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, studied a stochastic multiplicative process for returns and characterized the resulting stationary distribution; see also Levy and Solomon (1996) for more formal arguments and Fiaschi and 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.

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

income tax and estate tax on wealth inequality. Most proofs and several technical details are buried in Appendices A and B. Replication files are posted as Supplemental Material (Benhabib, Bisin, and Zhu (2011)).

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. Any household born at time s has a single child entering the economy at time $s+T$, that is, at his parents' deaths. Generations of households are overlapping but are linked to form dynasties. A household born at time s belongs to the $n = \frac{s}{T}$th generation of its 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 rates of return on wealth and earnings at birth, which remain, however, constant in its lifetime. Generation n is, therefore, associated to a rate of return on wealth r_n and to earnings y_n.[19]

Consumption and wealth at t of a household born at s depend on the generation of the household n through r_n and y_n, and on its age $\tau = t - s$. We adopt the notation $c(s,t) = c_n(t-s)$ and $w(s,t) = w_n(t-s)$, respectively, for consumption and wealth for a household of generation $n = \frac{s}{T}$ at time t. Such household inherits wealth $w(s,s) = w_n(0)$ at s from its previous generation. If $b < 1$ denotes the estate tax, then $w_n(0) = (1-b)w(s-T,s) = (1-b)w_{n-1}(T)$. Each household's momentary utility function is denoted $u(c_n(\tau))$. Households also have a preference for leaving bequests to their children. In particular, we assume "joy of giving" preferences for bequests: generation n's parents' utility from bequests is $\phi(w_{n+1}(0))$, where ϕ denotes an increasing bequest function.[20]

A household of generation n born at time s chooses a lifetime consumption path $c_n(t-s)$ to maximize

$$\int_0^T e^{-\rho\tau} u(c_n(\tau))\, d\tau + e^{-\rho T}\phi(w_{n+1}(0))$$

[19]Without loss of generality, we can add a deterministic growth component $g > 0$ to lifetime earnings: $y(s,t) = y(s,s)e^{g(t-s)}$, where $y(s,t)$ denotes the earnings at time t of an agent born at time s (in generation n) with $y_n = y(s,s)$. In fact, this is the notation we use in Appendix A. Importantly, the aggregate growth rate of the economy is independent of g. We can also easily allow for general trend-stationary earning processes across generations (with trend g' not necessarily equal to g^T). In this case, our results hold for the appropriately discounted measure of wealth (or, equivalently, for the ratio of individual and aggregate wealth); see the NBER version of this paper (Benhabib and Bisin (2009)). Finally, Zhu (2010) allowed for stochastic returns of wealth inside each generation.

[20]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.

subject to

$$\dot{w}_n(\tau) = r_n w_n(\tau) + y_n - c_n(\tau),$$
$$w_{n+1}(0) = (1-b)w_n(T),$$

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 $\sigma \geq 1$. *Furthermore, we require* $r_n \geq \rho$ *and* $\chi > 0$.[21]

The dynamics of individual wealth is easily solved for; 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 $w_n = w_n(0)$ denote the initial wealth of the n'th dynasty. Since w_n is inherited from generation $n-1$,

$$w_n = (1-b)w_{n-1}(T).$$

The rates of return of wealth and earnings are stochastic across generations. We assume they are also idiosyncratic across individuals. Let $(r_n)_n$ and $(y_n)_n$ denote, respectively, the stochastic processes for the rates of return of wealth and earnings over generations n.[22] 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]The condition $r_n \geq \rho$ (on the whole support of the random variable r_n) is sufficient to guarantee that agents will not want to 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 nondegenerate wealth distribution and could be relaxed if we allowed the elasticities of substitution for consumption and bequest to differ, at a notational cost. Finally, $\chi > 0$ guarantees positive bequests.

[22]We avoid as much as possible the notation required for formal definitions on probability spaces and stochastic processes. The costs in terms of precision seem 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 on the deep parameters of our economy; see Appendix A, equations (5) and (6), 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.[23] 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 labor income, again after subtracting the affine part of consumption.

3.1. *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.,[24] we study more general stochastic processes which naturally arise when studying the distribution of wealth. A positive autocorrelation in r_n and y_n, in particular, can capture variations in social mobility in the economy, for example, 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, for example, for households with high labor income to have better opportunities for higher returns on wealth in financial markets.[25]

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, that is, that the distribution of α_n display enough mass on $\alpha_n < 1$ as well as some on $\alpha_n > 1$, and that effective earnings β_n be positive, hence acting as a reflecting barrier.

We impose assumptions on $(r_n, y_n)_n$ which are sufficient to guarantee the existence and uniqueness of a limit stationary distribution of $(w_n)_n$; see Assumptions 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 $E(\alpha_n|\alpha_{n-1}) < \mathbf{1}$ for any α_{n-1}, and finally that

[23]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, is positive. Also, note that because bequests are positive under our assumptions, α_n is also positive; see the Proof of Proposition 1.

[24]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.

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

$\alpha_n > 1$ with positive probability; see Lemma A.1 in Appendix B.[26] In terms of fundamentals, these assumptions require an upper bound on the (log of the) mean of r_n as well as that r_n be large enough with positive probability.[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*

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

where $\mu > 1$ *satisfies*

$$\lim_{N\to\infty}\left(E\prod_{n=0}^{N-1}(\alpha_{-n})^{\mu}\right)^{1/N} = 1. \tag{2}$$

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.

[26]We also assume that β_n is bounded, although the assumption is stronger than necessary. In Proposition 1, 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.

[27]Suppose preferences are logarithmic. Then it is required that

$$E(e^{r_n T}) < \frac{e^{\rho T} + \rho\chi - 1}{(1-b)\rho\chi},$$

$$r_n > \frac{1}{T}\log\left(\frac{e^{\rho T} + \rho\chi - 1}{(1-b)\rho\chi}\right) \quad \text{with positive probability.}$$

We thank an anonymous referee for pointing this out. As an example of parameters that satisfy these conditions for the log utility case, suppose that $\rho = .04$, $\chi = .25$, $T = 45$, $b = .2$, $\zeta = .15$, and that the rate of return on wealth is i.i.d. with four states (see Section 5 for details regarding the model's calibration along these lines). The probabilities with these four states are .8, .12, .07, and .01. The first three states of before-tax rate of return are .08, .12, and .15. The above two inequalities imply that the fourth state of before-tax rate of return could belong to the open interval (.169, .286).

[28]See Appendix B, Proof of Theorem 1, for details.

[29]The 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).

3.2. *The Stationary Distribution of Wealth in the Population*

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 by aggregating over the wealth of households of all ages τ from 0 to T. The wealth of a household of generation n and age τ, born with wealth $w_n = w_n(0)$, 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 its lifetime. In Appendix B, we show that, under our assumptions, the process $(w_n, r_n)_n$ is ergodic and has a unique stationary distribution. Let ν denote the product measure of the stationary distribution of $(w_n, r_n)_n$. In Appendix A, we derive the closed form for $w_n(\tau)$, the wealth of household of generation n and age τ (equation (4)):

$$w_n(\tau) = \sigma_w(r_n, \tau)w_n + \sigma_y(r_n, \tau)y_n.$$

We can then define $F(w; \tau) = 1 - \Pr(w_n(\tau) > w)$, the cumulative distribution function of the stationary distribution of $w_n(\tau)$, as

$$F(w; \tau) = \sum_{j=1}^{l} \left(\Pr(y_j) \int I_{\{\sigma_w(r_n,\tau)w_n + \sigma_y(r_n,\tau)y_j \le w\}}\, d\nu \right),$$

where I is an indicator function. The cumulative distribution function of wealth w in the population is then defined as

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

We can now show that the power tail of the initial wealth distribution implies that the distribution of wealth w in the population displays a tail with exponent μ in the following sense:

THEOREM 2: *Suppose the tail of the stationary distribution of initial wealth* $w_n = w_n(0)$ *is asymptotic to a Pareto law*, $\Pr(w_n > w) \sim kw^{-\mu}$. *Then the stationary distribution of wealth in the population has a power tail with the same exponent* μ.

Note that this result is independent of the demographic characteristics of the economy, that is, of the stationary distribution of the households by age. The intuition is that the power tail of the stationary distribution of wealth in the population is as thick as the thickest tail across wealth distributions by age. Since under our assumptions each wealth distribution by age has a power tail with the same exponent μ, this exponent is inherited by the distribution of wealth in the population as well.[30]

[30] The tail of the stationary wealth distribution of the population is independent of any deterministic growth component $g > 0$ to lifetime earning as introduced in Appendix A.

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 $G = \frac{1}{2\mu-1}$, the classic statistical measure of inequality.[31]

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

4.1. *Capital and Labor Income Risk*

If 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 $\frac{\overline{\beta}}{1-\overline{\alpha}}$, where $\overline{\alpha}$ is the upper bound of α_n.[32]

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.*[33] *If r_n second order stochastically 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 is capital income risk, the more unequal is wealth.

[31]See, for example, Chipman (1976). Since the distribution of wealth in our economy is typically Pareto only in the tail, we refer $G = \frac{1}{2\mu-1}$ as the Gini of the tail.

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

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

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*

To study the effects of fiscal policy, first we redefine the random rate of return r_n as the pre-tax rate and introduce a capital income tax, ζ, so that the post-tax return on capital is $(1-\zeta)r_n$. Fiscal policies in our economy are then captured by the parameters b and ζ, representing, respectively, the estate tax and the capital income tax.

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 nonlinear 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 nonlinear 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, found 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. This compensating effect of bequests is present in our economy as well, although 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) 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. Social mobility is higher when $(r_n)_n$ and $(y_n)_n$ (and hence when $(\alpha_n)_n$ and $(\beta_n)_n$) are less autocorrelated over time.

We provide here expressions for the tail index of the wealth distribution as a function of the autocorrelation of $(\alpha_n)_n$ in the following two distinct cases,[34] where $0 < \theta < 1$ and $(\eta_n)_n$ is an i.i.d. process with bounded support.[35]

$$\text{MA}(1) \quad \ln \alpha_n = \eta_n + \theta \eta_{n-1},$$

$$\text{AR}(1) \quad \ln \alpha_n = \theta \ln \alpha_{n-1} + \eta_n.$$

PROPOSITION 4[36]: *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*

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

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

$$Ee^{(\mu_{\text{AR}}/(1-\theta))\eta_n} = 1.$$

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 fatter is the tail of the wealth distribution.[37]

5. A SIMPLE CALIBRATION EXERCISE

As we have already discussed in the Introduction, it has proven hard for standard macroeconomic 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. Our theoretical results are, however, limited to a characterization of the tail of the wealth distribution, and questions remain about the ability of our model to match the entire wealth distribution. To this end, we report on a simulation exercise which illustrates

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

[35] We thank Zheng Yang for pointing out that boundedness of η_n guarantees boundedness of α_n under our assumptions.

[36] We thank Xavier Gabaix for suggesting the statement of this proposition and outlining an argument for its proof.

[37] The results easily extend to MA(k) and AR(k) processes for $\ln \alpha_n$.

the ability of the model to match the Lorenz curve of the wealth distribution in the United States.[38]

We calibrate the parameters of the models as follows. First of all, we set the fundamental preference parameters in line with the macroeconomic literature: $\sigma = 2$, $\rho = 0.04$. We also set the preference for bequest parameter $\chi = 0.25$ and working life span $T = 45$.

The labor earnings process, y_n, is set to match mean earnings in \$10,000 units, 4.2.[39] We pick a standard deviation of y_n equal to 9.5 and we also assume that earnings grow at a yearly rate g equal to 1% over each household lifetime.[40]

The calibration of the cross-sectional distribution of the rate of return on wealth, r_n, is rather delicate, as capital income risk typically does not appear in calibrated macroeconomic models. We proceed as follows. First of all, we map the model to the data by distinguishing two components of r_n: a common economy-wide rate of return r^E and an idiosyncratic component r_n^I. The common component of returns, r^E, represents the value-weighted returns on the market portfolio, including, for example, cash, bonds, and public equity. The idiosyncratic component of returns, r_n^I, is composed for the most part of returns on the ownership of a principal residence and on private business equity. According to the *Survey of Consumer Finances*, ownership of a principal residence and private business equity account for about 50% of household wealth portfolios in the United States. We then map r_n into data according to

$$r_n = \frac{1}{2}r^E + \frac{1}{2}r_n^I.$$

For the common economy-wide rate of return r^E, which is assumed to be constant over time in the model, we choose a range of values between 7 and 9 percent before taxes, about 1–3 percentage points below the rate of return on public equity. Unfortunately, no precise estimate exists for the distribution of the idiosyncratic component of capital income risk to calibrate the distribution of r_n^I. Flavin and Yamashita (2002) studied the after-tax return on housing at the level of individual houses from the 1968–1992 waves of the *Panel Study of Income Dynamics*. They obtained a mean after-tax return of 6.6% with a standard deviation of 14%. Returns on private equity were estimated by Moskowitz and Vissing-Jorgensen (2002) from the 1989-1998 *Survey of Consumer Finances* data. They found mean returns comparable to those on public equity, but they lacked enough time series variation to estimate their standard deviation, which

[38]For the data on the U.S. economy, the tail index is from Klass et al. (2007), who used the Forbes 400 data. The rest of data for the United States economy are from Diaz-Gimenez, Quadrini, Ríos-Rull, and Rodríguez (2002), who used the 1998 *Survey of Consumer Finances*.

[39]More specifically, we choose a discrete distribution for y_n, taking values .75, 2.51, 5.01, 12.54, 25.07, and 75.22 with probability $\frac{14}{64}$, $\frac{36}{64}$, $\frac{11}{64}$, $\frac{1}{64}$, $\frac{1}{64}$, and $\frac{1}{64}$, respectively.

[40]This requires straightforwardly extending the model along the lines delineated in footnote 19.

they end up proxying with the standard deviation of an individual publicly traded stock. Based on these data, Angeletos (2007) adopted a baseline calibration for capital income risk with an implied mean return around 7% and a standard deviation of 20%. Allowing for a private equity risk premium, we choose mean values for r_n^I between 7 and 9 percent. With regard to the standard deviation, in our model r_n^I is constant over an agent's lifetime. Interpreting r_n^I as a mean over the yearly rates of return estimated in the data and assuming independence, a 3% standard deviation of r_n^I corresponds to a standard deviation of yearly returns on the order of 20% as in Angeletos (2007). We then choose a range of standard deviations of r_n^I between 2 and 3 percent.

With regard to social mobility, we present results for the case in which r_n is i.i.d. across generations (perfect social mobility), as well as for different degrees of autocorrelation of r_n (imperfect social mobility). The capital income risk process r_n is formally modelled as a discrete Markov chain. In the case in which r_n is i.i.d., the Markov transition matrix for r_n has identical rows.[41] We then introduce frictions to social mobility by moving a mass ε_{low} of probability from the off-diagonal terms to the diagonal term in the first row of the Markov transition matrix for r_n, that is, the row corresponding to the probability distribution of r_{n+1} conditional on r_n being lowest. We do the same shift of a mass $\varepsilon_{\text{high}}$ of probability in the last row of the Markov transition matrix for r_n, that is, the row corresponding to the probability distribution of r_{n+1} conditional on r_n being highest. This introduces persistence of low and high rates of return of wealth across generations.

For our baseline simulation, in Table I we report the relevant statistics of the r_n process at the stationary distribution for $\varepsilon_{\text{low}} = 0, .01$, and $\varepsilon_{\text{high}} = 0, .01, .02, .05$, respectively.

Finally, we set the estate tax rate $b = .2$ (which is the average tax rate on bequests), and the capital income tax $\zeta = .15$ in the baseline, but in Section 5.2 we study various combinations of fiscal policy.

With this calibration we simulate the stationary distribution of the economy.[42] 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. While we are mostly concerned with the wealth distribution, we also report the capital income to labor income ratio implied in the simulation as an extra check. We aim at a ratio not too distant from .5, the value implied by the standard calibration of macroeconomic production models (with a constant return to scale Cobb–Douglas production

[41] We choose two discrete Markov processes for r_n, the first with mean (at the stationary distribution) on the order of 9 percent and the second on the order of 7 percent. More specifically, the first process takes values [.08, .12, .15, .32] with probability rows (in the i.i.d. case) of the transition equal to [.8, .12, .07, .01]; the second process has support [.065, .12, .15, .27] with probability rows (in the i.i.d. case) equal to [.93, .01, .01, .05].

[42] We note that under these calibrations for r_n and other parameters, we check that the conditions of Assumptions 2 and 3 are satisfied and, therefore, that the restrictions on α hold.

TABLE I

BASELINE CALIBRATION OF r_n[a]

Economy	$E(r_n)$	$\sigma(r_n)$	$\text{corr}(r_n, r_{n-1})$
$\varepsilon_{\text{low}} = 0$, $\varepsilon_{\text{high}} = 0$	.0921	.0311	0
$\varepsilon_{\text{high}} = .01$	.0922	.0313	.0148
$\varepsilon_{\text{high}} = .02$	.0922	.0316	.0342
$\varepsilon_{\text{high}} = .05$	.0925	.0325	.0812
$\varepsilon_{\text{low}} = .01$, $\varepsilon_{\text{high}} = 0$	.0892	.0223	.0571
$\varepsilon_{\text{high}} = .01$	.0892	.0224	.0613
$\varepsilon_{\text{high}} = .02$	.0892	.0224	.0619
$\varepsilon_{\text{high}} = .05$	.0893	.0227	.0952

[a]All the statistics are obtained from the simulated stationary distribution of r_n except the auto-correlation $\text{corr}(r_n, r_{n-1})$ when $\varepsilon_{\text{low}} = \varepsilon_{\text{high}} = 0$, which is 0 analytically.

function with capital share equal to $\frac{1}{3}$). We report first, as a baseline, the case with $\varepsilon_{\text{low}} = .01$ and various values for $\varepsilon_{\text{high}}$.

First of all, note that the wealth distributions which we obtain in the various simulations in Table II match quite successfully the top percentiles of the United States. Furthermore, note that the tail of the simulated wealth distribution economy gets thicker by increasing $\varepsilon_{\text{high}}$, that is, by increasing $\text{corr}(r_n, r_{n-1})$. In particular, the better fit is obtained with substantial imperfections in social mobility ($\varepsilon_{\text{high}} = .02$), in which case the 99th–100th percentile of wealth in the U.S. economy is matched almost exactly.

More surprisingly, perhaps, the Lorenz curve (in quintiles) of the simulated wealth distributions, Table III, matches reasonably well that of the United States; and so does the Gini coefficient. Once again, $\varepsilon_{\text{high}} = .02$ appears to represent the better fit in terms of the Lorenz curve and the Gini coefficient

TABLE II

PERCENTILES OF THE TOP TAIL; $\varepsilon_{\text{low}} = .01$

	Percentiles		
Economy	90th–95th	95th–99th	99th–100th
United States	.113	.231	.347
$\varepsilon_{\text{high}} = 0$	.118	.204	.261
$\varepsilon_{\text{high}} = .01$	.116	.202	.275
$\varepsilon_{\text{high}} = .02$	.105	.182	.341
$\varepsilon_{\text{high}} = .05$	.087	.151	.457

TABLE III

TAIL INDEX, GINI, AND QUINTILES; $\varepsilon_{\text{low}} = .01$

			Quintiles				
Economy	Tail Index μ	Gini	First	Second	Third	Fourth	Fifth
United States	1.49	.803	−.003	.013	.05	.122	.817
$\varepsilon_{\text{high}} = 0$	1.796	.646	.033	.058	.08	.123	.707
$\varepsilon_{\text{high}} = .01$	1.256	.655	.032	.056	.078	.12	.714
$\varepsilon_{\text{high}} = .02$	1.038	.685	.029	.051	.071	.11	.739
$\varepsilon_{\text{high}} = .05$	.716	.742	.024	.042	.058	.09	.786

(even though the tail index of this calibration is lower than the U.S. economy's, but the tail index is imprecisely estimated with wealth data).[43]

Furthermore, the capital income to labor income ratio implied by the simulations takes on reasonable values: it goes from .3 for $\varepsilon_{\text{high}} = 0$ to .6 for $\varepsilon_{\text{high}} = .05$. In the $\varepsilon_{\text{high}} = .02$ calibration, the capital–labor ratio is almost exactly .5.

5.1. *Robustness*

As a robustness check, we report the calibration with $\varepsilon_{\text{low}} = 0$. In this case, the simulated wealth distributions also have Gini coefficients close to that of the U.S. economy and Lorenz curves which also match that of the United States rather well. Table IV reports the top percentiles of the U.S. economy and of the simulated wealth distribution. Table V reports instead the tail index, the Gini coefficient, and the Lorenz curve of the U.S. economy and of the simulated wealth distribution.[44] Note that the calibration with i.i.d. capital income risk r_n ($\varepsilon_{\text{low}} = \varepsilon_{\text{high}} = 0$) does particularly well.

TABLE IV

PERCENTILES OF THE TOP TAIL; $\varepsilon_{\text{low}} = 0$

	Percentiles		
Economy	90th–95th	95th–99th	99th–100th
United States	.113	.231	.347
$\varepsilon_{\text{high}} = 0$	.1	.207	.38
$\varepsilon_{\text{high}} = .01$	.082	.173	.49
$\varepsilon_{\text{high}} = .02$	.073	.154	.544
$\varepsilon_{\text{high}} = .05$	.026	.06	.836

[43]The calibration with $\varepsilon_{\text{high}} = .05$, with even more frictions to social mobility, also fares well, although in this case the tail index is < 1, which implies that the tails are so thick that the theoretical distribution has no mean. In this case, Assumption 3(ii) in Appendix B is violated.

[44]Again, for $\varepsilon_{\text{high}} = .05$, we have $\mu < 1$. See footnote 43.

TABLE V

TAIL INDEX, GINI, AND QUINTILES; $\varepsilon_{\text{low}} = 0$

Economy	Tail Index μ	Gini	Quintiles				
			First	Second	Third	Fourth	Fifth
United States	1.49	.803	−.003	.013	.05	.122	.817
$\varepsilon_{\text{high}} = 0$	1.795	.738	.023	.041	.057	.092	.788
$\varepsilon_{\text{high}} = .01$	1.254	.786	.018	.033	.046	.074	.827
$\varepsilon_{\text{high}} = .02$	1.036	.808	.017	.003	.042	.067	.844
$\varepsilon_{\text{high}} = .05$	.713	.933	.006	.01	.014	.023	.947

We also report the simulation for the economy with a different Markov process for r_n, with pre-tax mean of 7%. Table VI reports the relevant statistics of the r_n process at the stationary distribution, in this case, for $\varepsilon_{\text{low}} = 0, .1$ and $\varepsilon_{\text{high}} = .2$, respectively.[45] Tables VII and VIII collect the results regarding the simulated wealth distribution for this process of capital income risk.

While still in the ballpark of the U.S. economy, these calibrations match it much more poorly than the previous ones with a higher mean of r_n. Interestingly, they induce a higher Gini coefficient than in the U.S. distribution, suggesting that our model, in general, does not share the difficulties experienced by standard calibrated macroeconomic models to produce wealth distributions with tails as heavy as those observed in the data.

5.2. *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, as before, with r_n as in Table I, $\varepsilon_{\text{high}} = .02$, and $\varepsilon_{\text{low}} = .01$, and we vary b and ζ. Table IX reports the effects of capital income taxes and estate taxes on the 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

TABLE VI

CALIBRATION OF r_n WITH MEAN 7%

Economy	$E(r_n)$	$\sigma(r_n)$	$\text{corr}(r_n, r_{n-1})$
$\varepsilon_{\text{low}} = 0$, $\varepsilon_{\text{high}} = .02$	.772	.467	.0356
$\varepsilon_{\text{low}} = .01$, $\varepsilon_{\text{high}} = .02$	.0738	.0415	.0542

[45]A more extensive set of results is available from the authors upon request.

TABLE VII

PERCENTILES OF THE TOP TAIL

Economy	Percentiles 90th–95th	95th–99th	99th–100th
United States	.113	.231	.347
$\varepsilon_{\text{low}} = .01$, $\varepsilon_{\text{high}} = .02$	.066	.232	.675
$\varepsilon_{\text{low}} = 0$, $\varepsilon_{\text{high}} = .02$	.076	.236	.646

tax, which directly affects the stochastic returns on wealth. The implied Gini of the tail[46] is very high with no (or low) taxes,[47] while it is reduced to .66 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 X. 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. A tax increase has the effect of reducing the concentration of wealth in the tail of the distribution. This effect is, however, partly offset by greater inequality at lower wealth levels. In general, a decrease in the rate of return on wealth (e.g., due to a tax increase) has the effect of increasing the permanent labor income of households, because future labor earnings are discounted at a lower rate. For rich households, whose wealth consists mainly of physical wealth rather than labor earnings, a lower capital income tax rate generates an approximately proportional wealth effect on consumption and savings. On the other hand, the positive wealth effect of a tax reduction has a relatively large effect 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 reduction by decumulating

TABLE VIII

TAIL INDEX, GINI, AND QUINTILES

Economy	Tail Index μ	Gini	Quintiles First	Second	Third	Fourth	Fifth
United States	1.49	.803	−.003	.013	.05	.122	.817
$\varepsilon_{\text{low}} = .01$, $\varepsilon_{\text{high}} = .02$	1.514	.993	−.022	.003	.009	.016	.994
$\varepsilon_{\text{low}} = 0$, $\varepsilon_{\text{high}} = .02$	1.514	.978	−.016	.003	.008	.015	.991

[46]As before, the tail Gini is $G = \frac{1}{2\mu-1}$.

[47]When the tail index μ is less than 1, the wealth distribution has no mean, so that again, Assumption 3(ii) in Appendix B is violated. In this case, theoretically the Gini coefficient is not defined. In Table X, however, we report the simulated value, computed from the simulated wealth distribution.

TABLE IX

TAX EXPERIMENTS—TAIL INDEX μ

$b\backslash\zeta$	0	.05	.15	.2
0	.68	.76	.994	1.177
.1	.689	.772	1.014	1.205
.2	.7	.785	1.038	1.238
.25	.706	.793	1.051	1.257

physical wealth proportionately faster than households 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. Of course, the effects of a tax increase on relatively poor households would be moderated (perhaps eliminated) if tax revenues were to be redistributed toward the less wealthy.

Nonetheless the results of Table X suggests a word of caution in evaluating the effects on wealth inequality of proposed fiscal policies like the abolition of estate taxes or the reduction of capital taxes. For instance, Castaneda, Diaz-Gimenez, and Rios-Rull (2003) and Cagetti and De Nardi (2007) found 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.[48] If the capital income risk component is a substantial fraction of idiosyncratic risk, such fiscal policies could have sizeable effects in increasing wealth inequality in the top tail of the distribution of wealth which may not show up in measurements of the Gini coefficient.[49]

TABLE X

TAX EXPERIMENTS—GINI

$b\backslash\zeta$	0	.05	.15	.2
0	.779	.769	.695	.674
.1	.768	.730	.693	.677
.2	.778	.724	.679	.674
.3	.754	.726	.680	.677

[48]See also our discussion of the results of Becker and Tomes (1979) previously in this section.

[49]Empirical studies also indicate that higher and more progressive taxes did, in fact, significantly reduce income and wealth inequality in the historical context; notably, for example, Lampman (1962) and Kuznets (1955). Most recently, Piketty (2003) and Piketty and Saez (2003) 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 United States and

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 to be 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 is 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 No. 962, p. 315–316). He therefore introduced skewness into the distribution of talents or labor earnings of households (1897, notes to No. 962, p. 416). Left with the distribution of talents and earnings as the main determinant of the wealth distribution, he was perhaps lead to his Pareto law, enunciated 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.[50]

APPENDIX A: CLOSED FORM SOLUTIONS

We report here only the closed form solutions for the dynamics of wealth in the paper.

Let the age at time t of a household born at time $s \leq t$ be denoted $\tau = t - s$. An agent born at time s belongs to generation $n = \frac{s}{T}$. Let the human capital at time t of a household born at s, $h(s, t) = h_n(t - s) = h_n(\tau)$, be defined as $h_n(\tau) = \int_0^T y_n e^{-(r_n - g)\tau}\, d\tau$.[51] We adopt the notation $w_n(0) = w_n$. The optimal consumption path satisfies

$$c_n(\tau) = m(\tau)(w_n(\tau) + h_n(\tau)).$$

Japan, and the United States and Canada, respectively, by Moriguchi and Saez (2005) and Saez and Veall (2003).

[50] 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).

[51] To save on notation in the text, we restrict to the case in which $g = 0$.

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

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

The dynamics of individual wealth as a function of age τ satisfies

$$w_n(\tau) = \sigma_w(r_n, \tau) w_n + \sigma_y(r_n, \tau) y_n \tag{4}$$

with

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

$$\sigma_y(r_n, \tau) = e^{r_n \tau} \frac{e^{(g-r_n)T} - 1}{g - r_n} \times \left(\frac{e^{A(r_n)(T-\tau)} + A(r_n)B(b) - 1}{e^{A(r_n)T} + A(r_n)B(b) - 1} - \frac{e^{(r_n-g)(T-\tau)} - 1}{e^{(r_n-g)T} - 1} \right),$$

and

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

The dynamics of wealth across generation is then

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

with

$$\alpha(r_n) = (1-b) e^{r_n T} \frac{A(r_n)B(b)}{e^{A(r_n)T} + A(r_n)B(b) - 1} \tag{5}$$

and

$$\beta(r_n, y_n) = (1-b) y_n \frac{e^{(g-r_n)T} - 1}{g - r_n} e^{r_n T} \frac{A(r_n)B(b)}{e^{A(r_n)T} + A(r_n)B(b) - 1}. \tag{6}$$

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}).[52]

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, for example, to productivity shocks, drove returns on capital $(r_n)_n$ as well as labor income $(y_n)_n$.[53]

ASSUMPTION 3: *Let* P *denote the transition matrix of* $(r_n)_n$: $P_{ii'} = \Pr(r_{i'}|r_i)$. *Let* $\boldsymbol{\alpha}(\bar{\mathbf{r}})$ *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) $P\boldsymbol{\alpha}(\bar{\mathbf{r}}) < \mathbf{1}$, (iii) $\exists \bar{r}_i$ *such that* $\alpha(\bar{r}_i) > 1$, *and* (iv) $P_{ii} > 0$ *for any* i.

We are now ready to show the following lemma.

LEMMA A.1: *Assumption* 2 *on* $(r_n, y_n)_n$ *implies 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) $(\alpha_n, \beta_n)_n > 0$, (ii) $E(\alpha_n|\alpha_{n-1}) < 1$ *for any* α_{n-1},[54] (iii) $\bar{\alpha}_i > 1$ *for some* $i = 1, \dots, m$, *and* (iv) *the diagonal elements of the transition matrix* P *of* α_n *are positive*.

PROOF: Let A be the diagonal matrix with elements $A_{ii} = \bar{\alpha}_i$ and $A_{ij} = 0$, $j \neq i$. Note that $E(\alpha_n|\alpha_{n-1})$ for any α_{n-1} can be written as $P\boldsymbol{\alpha}(\bar{\mathbf{r}}) < \mathbf{1}$. 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 $\overline{\boldsymbol{\alpha}} = \{\bar{\alpha}_1, \dots, \bar{\alpha}_m\}$ and $\overline{\boldsymbol{\beta}} = \{\bar{\beta}_1, \dots, \bar{\beta}_l\}$ denote the state spaces of, respectively, α_n and β_n as they are induced through the

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

[53] 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 Roitershtein (2007) for general Markov modulated processes.

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

maps (5) and (6). We shall show that the maps (5) and (6) are bounded in r_n and y_n. Therefore, the state spaces of α_n and β_n are well defined. It immediately follows 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 greater than 0 and bounded with probability 1 in r_n and y_n. Recall that $B(b) = \chi^{1/\sigma}(1 - b)^{(1-\sigma)/\sigma} > 0$. Note that

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

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) $P\bar{\boldsymbol{\alpha}} < \mathbf{1}$. Assumption 3(iii) also directly implies $\overline{\alpha}_i > 1$ for some $i = 1, \dots, m$. Finally P is the transition matrix of r_n as well as α_n. Therefore, Assumption 3(iv) implies that the elements of the trace of the transition matrix of α_n are positive. *Q.E.D.*

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 of wealth. We rule this out by means of the following technical regularity conditions[55]:

CONDITIONS: The Markov modulated process $(\alpha_n, \beta_n)_n$ is regular, that is,

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

and the elements of the vector $\bar{\alpha} = \{\ln\overline{\alpha}_1 \cdots \ln\overline{\alpha}_m\} \subset \mathbb{R}_+^m$ are not integral multiples of the same number.[56]

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

[56]Theorems which characterize the tails of distributions generated by equations with random multiplicative coefficients rely on this type of nonlattice assumption from renewal theory; see for example Saporta (2005). Versions of these assumption are standard in this literature; see Feller (1966).

Saporta (2005, Proposition 1, Section 4.1) established that, for finite Markov chains, $\lim_{N\to\infty}(E\prod_{n=0}^{N-1}(\alpha_{-n})^{\mu})^{1/N} = \lambda(A^{\mu}P')$, where $\lambda(A^{\mu}P')$ is the dominant root of $A^{\mu}P'$.[57] 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 (ii) that such $\mu > 1$. Saporta showed 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 $E\alpha(r) < 1$, showed that $\frac{d\ln\lambda(A^{\mu}P')}{\delta\mu} < 0$ at $\mu = 0$ and that $\ln(\lambda(A^{\mu}P'))$ is a convex function of μ.[58] Therefore, if there exists another solution $\mu > 0$ for $\ln(\lambda(A^{\mu}P')) = 0$, it is positive and unique.

To assure that $\mu > 1$, we replace the condition $E\alpha(r) < 1$ with Proposition 3(ii), $P\bar{\boldsymbol{\alpha}} < \mathbf{1}$. 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 $\mu = 1$, $\lambda(A^{\mu}P') = \lambda(AP') < 1$. Now note that if $(\alpha_n, \beta_n)_n$ is reflective, by Proposition 1, $P_{ii} > 0$ 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 $A^{\mu}P'$, $\lambda(A^{\mu}P')$, 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). *Q.E.D.*

PROOF OF THEOREM 2: We first show by Lemma A.2 that the process $(w_n, r_{n-1})_n$ is ergodic[59] and thus has a unique stationary distribution. If we denote with ϕ the product measure of the stationary distribution of $(w_n, r_{n-1})_n$, and we denote with ν the product measure of the stationary distribution of $(w_n, r_n)_n$, the relationship between ϕ and v is

$$v(dw, r_n) = \sum_{r_{n-1}}(\Pr(r_n|r_{n-1})\phi(dw, r_{n-1})).$$

Ergodicity of $(w_n, r_{n-1})_n$ then implies ergodicity of $(w_n, r_n)_n$, which then also has a unique stationary distribution. Actually, Lemma A.1 shows that $(w_n, r_{n-1})_n$ is V-uniformly ergodic, which is stronger than ergodicity. For the

[57]Recall that the matrix AP' has the property that the ith column sum equals the expected value of α_n conditional on $\alpha_{n-1} = \overline{\alpha}_i$. 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 $\lambda(A^{\mu}P')$.

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

[59]Actually Lemma A.2 shows that $(w_n, r_{n-1})_n$ is V-uniformly ergodic, which is stronger than ergodicity. For the mathematical concepts such as V-uniform ergodicity, ψ-irreducibility, and petite sets, see Meyn and Tweedie (2009).

mathematical concepts such as V-uniform ergodicity, ψ-irreducibility, and petite sets which we use in the proof, see Meyn and Tweedie (2009).

LEMMA A.2: *The process* $(w_n, r_{n-1})_n$ *is* V*-uniformly ergodic.*

PROOF: As in Theorem 1,

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

As assumed in Theorem 1, the process $(r_n, y_n)_n$ satisfies Assumptions 2 and 3.

Let $\alpha^L = \min_{i=1,2,\dots,m}\{\alpha(r_i)\}$ and $\beta^L = \min_{i=1,2,\dots,m;j=1,2,\dots,l}\{\beta(r_n, y_n)\}$. Thus $\frac{\beta^L}{1-\alpha^L}$ is the lower bound of the state space of w_n. Let $X = [\frac{\beta^L}{1-\alpha^L}, +\infty) \times \{\bar{r}_1, \dots, \bar{r}_m\}$. Assumptions 2 and 3, and the *regularity* assumption of $(\alpha_n, \beta_n)_n$ guarantee that the process visits, with positive probability in finite time, a dense subset of its support; see Brandt (1986) and Saporta (2005, Theorem 2, p. 1956). The stochastic process $(w_n, r_{n-1})_n$ is then ψ-irreducible and aperiodic.[60]

Let $\tilde{\alpha} = \max_{i=1,2,\dots,m}\{E(\alpha(r_n)|\alpha(r_i))\}$. From Lemma 3(ii), we know $E(\alpha_n|\alpha_{n-1}) < 1$ for any α_{n-1}. Thus $\tilde{\alpha} < 1$. Let $\hat{w} = \frac{\beta^U+1}{1-\tilde{\alpha}}$, where $\beta^U = \max_{i=1,2,\dots,m;j=1,2,\dots,l}\{\beta(r_n, y_n)\}$. Let $C = [\frac{\beta^L}{1-\alpha^L}, \hat{w}] \times \{\bar{r}_1, \dots, \bar{r}_m\}$. Pick a function $V(w_n, r_{n-1}) = w_n$ so that

$$\begin{aligned}
E(V(w_{n+1}, r_n)|(w_n, r_{n-1})) &= E(w_{n+1}|(w_n, r_{n-1})) \\
&= E(\alpha(r_n)|r_{n-1})w_n + E(\beta(r_n, y_n)|r_{n-1}) \\
&\le w_n - 1 + (\beta^U + 1)I_C(w_n, r_{n-1}) \\
&= V(w_n, r_{n-1}) - 1 + (\beta^U + 1)I_C(w_n, r_{n-1}).
\end{aligned}$$

Thus $(w_n, r_{n-1})_n$ satisfies the *drift condition* of Tweedie (2001).

For a sequence of measurable set B_n with $B_n \downarrow \emptyset$, there are two cases: (i) B_n is contained in a compact set in X and (ii) B_n has forms of $(x_n, +\infty) \times \bar{r}_i$ or of the union of such sets. In both cases it is easy to show that

$$\lim_{n\to\infty} \sup_{(w,r)\in C} P((w, r), B_n) = 0,$$

where $P(\cdot, \cdot)$ is the one-step transition probability of the stochastic process $(w_n, r_{n-1})_n$. Thus $(w_n, r_{n-1})_n$ satisfies the *uniform countable additivity condition* of Tweedie (2001).

[60]Alternatively to the *regularity* assumption, we could assume a continuous distribution for y_n (and hence for β_n). Irreducibility would then easily follow; see Meyn and Tweedie (2009, p. 76).

As a consequence, $(w_n, r_{n-1})_n$ satisfies *condition A* of Tweedie (2001): $V(w_n, r_{n-1}) = w_n$ is everywhere finite and $(w_n, r_{n-1})_n$ is ψ-irreducible. By Theorem 3 of Tweedie (2001), we know that the set C is petite.[61]

Also we have

$$\begin{aligned} &E(V(w_{n+1}, r_n)|(w_n, r_{n-1})) - V(w_n, r_{n-1}) \\ &\quad = E(w_{n+1}|(w_n, r_{n-1})) - w_n \\ &\quad = E(\alpha(r_n)|r_{n-1})w_n + E(\beta(r_n, y_n)|r_{n-1}) - w_n \\ &\quad \leq -(1-\hat{\alpha})w_n + \beta^U I_C(w_n, r_{n-1}) \\ &\quad = -(1-\hat{\alpha})V(w_n, r_{n-1}) + \beta^U I_C(w_n, r_{n-1}). \end{aligned}$$

We then have that $(w_n, r_{n-1})_n$ is ψ-irreducible and aperiodic, $V(w_n, r_{n-1}) = w_n$ is everywhere finite, and the set C is petite. By Theorem 16.1.2 of Meyn and Tweedie (2009), we then obtain that $(w_n, r_{n-1})_n$ is V-uniformly ergodic. *Q.E.D.*

The wealth of a household of age τ, $w_n(\tau)$, is given by (4). Recall that we use the notational shorthand $w_n = w_n(0)$. The cumulative distribution function of the stationary distribution of wealth of a household of age τ, $F(w;\tau)$, is then given by

$$F(w;\tau) = \sum_{j=1}^{l} \left(\Pr(y_j) \int I_{\{\sigma_w(r_n,\tau)w_n + \sigma_y(r_n,\tau)y_j \leq w\}} \, d\nu \right),$$

where I is an indicator function and ν is the product measure of the stationary distribution of (w_n, r_n), which exists and is unique as a direct consequence of Lemma A.2. 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

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

[61]Note that (i) every subset of a petite set is petite and (ii) when we pick any w, such that $w > \frac{\beta^U + 1}{1-\hat{\alpha}}$, to replace $\hat{w}$, the proof goes through. By these two facts we could show that every compact set of X is petite. Thus by Theorem 6.2.5 of Meyn and Tweedie (2009) we know that $(w_n, r_{n-1})_n$ is a T-chain. For another example of stochastic process in economics with the property that every compact set is petite, see Nishimura and Stachurski (2005).

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 $0 < \sigma_w^L \le \sigma_w(r_n, \tau) \le \sigma_w^U$ and $\sigma_y(r_n, \tau) \le \sigma_y^U$. Let $y^U = \max\{\bar{y}_1, \dots, \bar{y}_l\}$. We have

$$I_{\{\sigma_w(r_n,\tau)w_n+\sigma_y(r_n,\tau)y_j>w\}} \ge I_{\{\sigma_w^L w_n>w\}}$$

and

$$I_{\{\sigma_w(r_n,\tau)w_n+\sigma_y(r_n,\tau)y_j>w\}} \le I_{\{\sigma_w^U w_n+\sigma_y^U y^U>w\}}.$$

Hence

$$P\left(w_n > \frac{w}{\sigma_w^L}\right) \le P(w_n(\tau) > w) \le P\left(w_n > \frac{w - \sigma_y^U y^U}{\sigma_w^U}\right).$$

We then have

$$1 - F(w) = \int_0^T P(w_n(\tau) > w)\frac{1}{T}\, d\tau.$$

Thus

$$P\left(w_n > \frac{w}{\sigma_w^L}\right) \le 1 - F(w) \le P\left(w_n > \frac{w - \sigma_y^U y^U}{\sigma_w^U}\right)$$

and

$$(\sigma_w^L)^\mu k \le \lim \inf_{w\to+\infty} \frac{1 - F(w)}{w^{-\mu}} \le \lim \sup_{w\to+\infty} \frac{1 - F(w)}{w^{-\mu}} \le (\sigma_w^U)^\mu k$$

since $\lim_{w\to+\infty} \frac{P(w_n>w)}{w^{-\mu}} = k$. We conclude that the wealth distribution in the population has a power tail with the same exponent μ, that is,

$$0 < k_1 \le \lim \inf_{w\to+\infty} \frac{1 - F(w)}{w^{-\mu}} \le \lim \sup_{w\to+\infty} \frac{1 - F(w)}{w^{-\mu}} \le k_2. \qquad \textit{Q.E.D.}$$

We can also show the following claim:

CLAIM 1: *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*: $\exists \tilde{k} > 0$ *such that*

$$\lim_{w\to+\infty} \frac{1 - F(w)}{w^{-\mu}} = \tilde{k}.$$

PROOF: When $(r_n)_n$ is i.i.d.,

$$
\begin{aligned}
&1 - F(w)\\
&\quad = \int_0^T P(w_n(\tau) > w)\frac{1}{T}\,d\tau\\
&\quad = \sum_{i=1}^{m}\sum_{j=1}^{l}\left(\Pr(r_i)\Pr(y_j)\int_0^T P\left(w_n > \frac{w - \sigma_y(r_i,\tau)y_j}{\sigma_w(r_i,\tau)}\right)\frac{1}{T}\,d\tau\right).
\end{aligned}
$$

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

$$
P\left(w_n > \frac{w - \sigma_y(r_i,\tau)y_j}{\sigma_w(r_i,\tau)}\right) \le P\left(w_n > \frac{w - \sigma_y(r_i,\hat{\tau}_i)y_j}{\sigma_w(r_i,\tilde{\tau}_i)}\right).
$$

When w is sufficiently large,

$$
\frac{P\left(w_n > \dfrac{w - \sigma_y(r_i,\hat{\tau}_i)y_j}{\sigma_w(r_i,\tilde{\tau}_i)}\right)}{w^{-\mu}} \quad \text{is bounded}
$$

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

$$
\begin{aligned}
&\lim_{w\to+\infty}\int_0^T \frac{P\left(w_n > \dfrac{w - \sigma_y(r_i,\tau)y_j}{\sigma_w(r_i,\tau)}\right)}{w^{-\mu}}\frac{1}{T}\,d\tau\\
&\quad = \int_0^T \lim_{w\to+\infty}\frac{P\left(w_n > \dfrac{w - \sigma_y(r_i,\tau)y_j}{\sigma_w(r_i,\tau)}\right)}{w^{-\mu}}\frac{1}{T}\,d\tau.
\end{aligned}
$$

Thus

$$
\begin{aligned}
&\lim_{w\to+\infty}\frac{1 - F(w)}{w^{-\mu}}\\
&\quad = \sum_{i=1}^{m}\sum_{j=1}^{l}\left(\Pr(r_i)\Pr(y_j)\int_0^T \lim_{w\to+\infty}\frac{P\left(w_n > \dfrac{w - \sigma_y(r_i,\tau)y_j}{\sigma_w(r_i,\tau)}\right)}{w^{-\mu}}\frac{1}{T}\,d\tau\right)\\
&\quad = k\sum_{i=1}^{m}\left(\Pr(r_i)\int_0^T (\sigma_w(r_i,\tau))^{\mu}\,d\tau\right).
\end{aligned}
$$

Q.E.D.

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; hence, $-\alpha(r_n)^\mu$ is a concave function of r_n. By 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 $E\alpha(r'_n)^\mu < (E\alpha(r'_n)^{\tilde{\mu}})^{\mu/\mu'} = 1$. This is a contradiction. Thus we have $\mu' \leq \mu$. *Q.E.D.*

PROOF OF PROPOSITION 2: From the definition of α_n, we have

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

and 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 nonnegative 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.[62] 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. *Q.E.D.*

PROOF OF PROPOSITION 3: From (5), we have

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

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

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

Note that $A(r_n) = r_n - \frac{r_n-\rho}{\sigma}$ and $B(b) = \chi^{1/\sigma}(1-b)^{(1-\sigma)/\sigma}$. Then $\frac{\partial A(r_n)}{\partial r_n} = \frac{\sigma-1}{\sigma} \geq 0$, 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 $\frac{\partial \alpha_n}{\partial \zeta} < 0$. Now the proof is identical to the proof

[62] See Loeve (1977, p. 158).

of Proposition 2 in the reverse direction since $\frac{\partial \alpha_n}{\partial b} < 0$ and $\frac{\partial \alpha_n}{\partial \zeta} < 0$, whereas $\frac{\partial \alpha_n}{\partial \chi} > 0$. *Q.E.D.*

PROOF OF PROPOSITION 4: We apply the results of Roitershtein (2007) on exponents of the tails of the limiting distribution. In the MA(1) case where

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

we have

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

Thus

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

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

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

Consider, in turn, the AR(1) case

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

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

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

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

$$Ee^{(\mu/(1-\theta))\eta_t} = 1. \qquad \textit{Q.E.D.}$$

REFERENCES

AIYAGARI, S. R. (1994): "Uninsured Idiosyncratic Risk and Aggregate Savings," *Quarterly Journal of Economics*, 109, 659–684. [124]

ANGELETOS, G. (2007): "Uninsured Idiosyncratic Investment Risk and Aggregate Saving," *Review of Economic Dynamics*, 10, 1–30. [125,126,137]

ANGELETOS, G., AND L. E. CALVET (2005): "Incomplete-Market Dynamics in a Neoclassical Production Economy," *Journal of Mathematical Economics*, 41, 407–438. [125]

——— (2006): "Idiosyncratic Production Risk, Growth and the Business Cycle," *Journal of Monetary Economics*, 53, 1095–1115. [125]

ARROW, K. (1987): "The Demand for Information and the Distribution of Income," *Probability in the Engineering and Informational Sciences*, 1, 3–13. [130]

ATKINSON, A. B. (2002): "Top Incomes in the United Kingdom Over the Twentieth Century," Mimeo, Nuffield College, Oxford. [123]

BECKER, G. S., AND N. TOMES (1979): "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility," *Journal of Political Economy*, 87, 1153–1189. [134,142]

BENHABIB, J., AND A. BISIN (2006): "The Distribution of Wealth and Redistributive Policies," Unpublished Manuscript, New York University. [126]

——— (2009): "The Distribution of Wealth and Fiscal Policy in Economies With Finitely Lived Agents," Working Paper 14730, NBER. [128]

BENHABIB, J., AND S. ZHU (2008): "Age, Luck and Inheritance," Working Paper 14128, NBER. [126]

BENHABIB, J., A. BISIN, AND S. ZHU (2011): "Supplement to 'The Distribution of Wealth and Fiscal Policy in Economies With Finitely Lived Agents'," *Econometrica Supplemental Material*, 79, http://www.econometricsociety.org/ecta/Supmat/8416_data and programs.zip. [128]

BERTAUT, C., AND M. STARR-MCCLUER (2002): "Household Portfolios in the United States," in *Household Portfolios*, ed. by L. Guiso, M. Haliassos, and T. Jappelli. Cambridge, MA: MIT Press. [125]

BITLER, M. P., T. J. MOSKOWITZ, AND A. VISSING-JØRGENSEN (2005): "Testing Agency Theory With Entrepreneur Effort and Wealth," *Journal of Finance*, 60, 539–576. [126]

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. [131,148]

CAGETTI, M., AND M. DE NARDI (2006): "Entrepreneurship, Frictions, and Wealth," *Journal of Political Economy*, 114, 835–870. [124]

——— (2007): "Estate Taxation, Entrepreneurship, and Wealth," Working Paper 13160, NBER. [142]

CARROLL, C. D. (1997): "Buffer-Stock Saving and the Life Cycle/Permanent Income Hypothesis," *Quarterly Journal of Economics*, 112, 1–56. [124]

CASE, K., AND R. SHILLER (1989): "The Efficiency of the Market for Single-Family Homes," *American Economic Review*, 79, 125–137. [125]

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, 818–857. [126,142]

CHAMPERNOWNE, D. G. (1953): "A Model of Income Distribution," *Economic Journal*, 63, 318–351. [124,127]

CHIPMAN, J. S. (1976): "The Paretian Heritage," *Revue Europeenne des Sciences Sociales et Cahiers Vilfredo Pareto*, 14, 65–171. [133,143]

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. [123]

DAGSVIK, J. K., AND B. H. VATNE (1999): "Is the Distribution of Income Compatible With a Stable Distribution?" Discussion Paper 246, Research Department, Statistics Norway. [123]

DE NARDI, M. (2004): "Wealth Inequality and Intergenerational Links," *Review of Economic Studies*, 71, 743–768. [124]

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, 3–21. [124]

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, 2–35. [136]

FEENBERG, D., AND J. POTERBA (2000): "The Income and Tax Share of Very High Income Household: 1960–1995," *American Economic Review*, 90, 264–270. [124]

FELLER, W. (1966): *An Introduction to Probability Theory and Its Applications*, Vol. 2. New York: Wiley. [146]

FIASCHI, D., AND M. MARSILI (2009): "Distribution of Wealth and Incomplete Markets: Theory and Empirical Evidence," Working Paper 83, University of Pisa. [127]

FLAVIN, M., AND T. YAMASHITA (2002): "Owner-Occupied Housing and the Composition of the Household Portfolio," *American Economic Review*, 92, 345–362. [125,136]

GOLDIE, C. M. (1991): "Implicit Renewal Theory and Tails of Solutions of Random Equations," *The Annals of Applied Probability*, 1, 126–166. [126,131]

GUVENEN, F. (2007): "Learning Your Earning: Are Labor Income Shocks Really Very Persistent?" *American Economic Review*, 97, 687–712. [125]

HEATHCOTE, J. (2008): "Discussion Heterogeneous Life-Cycle Profiles, Income Risk, and Consumption Inequality, by G. Primiceri and T. van Rens," Mimeo, Federal Bank of Minneapolis. [125]

HUGGETT, M. (1993): "The Risk-Free Rate in Heterogeneous-Household Incomplete-Insurance Economies," *Journal of Economic Dynamics and Control*, 17, 953–969. [124]

KESTEN, H. (1973): "Random Difference Equations and Renewal Theory for Products of Random Matrices," *Acta Mathematica*, 131, 207–248. [126,131]

KLASS, O. S., O. BIHAM, M. LEVY, O. MALCAI, AND S. SOLOMON (2007): "The Forbes 400, the Pareto Power-Law and Efficient Markets," *The European Physical Journal B—Condensed Matter and Complex Systems*, 55, 143–147. [124,136]

KRUSELL, P., AND A. A. SMITH (1998): "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 106, 867–896. [124,126]

KUZNETS, S. (1955): "Economic Growth and Economic Inequality," *American Economic Review*, 45, 1–28. [142]

LAMPMAN, R. J. (1962): *The Share of Top Wealth-Holders in National Wealth, 1922–1956*. Princeton, NJ: NBER and Princeton University Press. [142]

LEVY, M. (2005): "Market Efficiency, The Pareto Wealth Distribution, and the Levy Distribution of Stock Returns," in *The Economy as an Evolving Complex System*, ed. by S. Durlauf and L. Blume. Oxford, U.K.: Oxford University Press. [127]

LEVY, M., AND S. SOLOMON (1996): "Power Laws Are Logarithmic Boltzmann Laws," *International Journal of Modern Physics, C*, 7, 65–72. [127]

LOÈVE, M. (1977): *Probability Theory* (Fourth Ed.). New York: Springer. [147,152]

MANTEGNA, R. N., AND H. E. STANLEY (2000): *An Introduction to Econophysics*. Cambridge, U.K.: Cambridge University Press. [124]

MCKAY, A. (2008): "Household Saving Behavior, Wealth Accumulation and Social Security Privatization," Mimeo, Princeton University. [130]

MEYN, S., AND R. L. TWEEDIE (2009): *Markov Chains and Stochastic Stability* (Second Ed.). Cambridge: Cambridge University Press. [147-149]

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. [123, 143]

MOSKOWITZ, T., AND A. VISSING-JORGENSEN (2002): "The Returns to Entrepreneurial Investment: A Private Equity Premium Puzzle?" *American Economic Review*, 92, 745–778. [125,136]
NIREI, M., AND W. SOUMA (2004): "Two Factor Model of Income Distribution Dynamics," Mimeo, Utah State University. [123]
NISHIMURA, K., AND J. STACHURSKI (2005): "Stability of Stochastic Optimal Growth Models: A New Approach," *Journal of Economic Theory*, 122, 100–118. [149]
PANOUSI, V. (2008): "Capital Taxation With Entrepreneurial Risk," Mimeo, MIT. [125]
PARETO, V. (1897): *Cours d'Economie Politique*, Vol. II. Lausanne: F. Rouge. [143]
——— (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*. New Brunswick, NJ: Transaction Publishers (1991). [143]
——— (1909): *Manuel d'Economie Politique*. Paris: V. Girard et E. Brière. [143]
PIKETTY, T. (2003): "Income Inequality in France, 1901–1998," *Journal of Political Economy*, 111, 1004–1042. [123,142]
PIKETTY, T., AND E. SAEZ (2003): "Income Inequality in the United States, 1913–1998," *Quarterly Journal of Economics*, 118, 1–39. [123,142]
PRIMICERI, G., AND T. VAN RENS (2006): "Heterogeneous Life-Cycle Profiles, Income Risk, and Consumption Inequality," Discussion Paper 5881, CEPR. [125]
QUADRINI, V. (1999): "The Importance of Entrepreneurship for Wealth Concentration and Mobility," *Review of Income and Wealth*, 45, 1–19. [124]
——— (2000): "Entrepreneurship, Savings and Social Mobility," *Review of Economic Dynamics*, 3, 1–40. [124,126]
ROITERSHTEIN, A. (2007): "One-Dimensional Linear Recursions With Markov-Dependent Coefficients," *The Annals of Applied Probability*, 17, 572–608. [127,145,153]
RUTHERFORD, R. S. G. (1955): "Income Distribution: A New Model," *Econometrica*, 23, 277–294. [124]
SAEZ, E., AND M. VEALL (2003): "The Evolution of Top Incomes in Canada," Working Paper 9607, NBER. [123,143]
SAMUELSON, P. A. (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, 246–250. [143]
SAPORTA, B. (2004): "Etude de la Solution Stationnaire de l'Equation $Y_{n+1} = a_n Y_n + b_n$, a Coefficients Aleatoires," Thesis, Université de Rennes I. Available at http://tel.archives-ouvertes.fr/docs/00/04/74/12/PDF/tel-00007666.pdf. [127,130,145,147]
——— (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 Their Applications*, 115, 1954–1978. [127,130,131,145-148,152]
SORNETTE, D. (2000): *Critical Phenomena in Natural Sciences*. Springer Verlag: Berlin. [124]
STORESLETTEN, K., C. I. TELMER, AND A. YARON (2004): "Consumption and Risk Sharing Over the Life Cycle," *Journal of Monetary Economics*, 51, 609–633. [125]
TWEEDIE, R. L. (2001): "Drift Conditions and Invariant Measures for Markov Chains," *Stochastic Processes and Their Applications*, 92, 345–354. [148,149]
WOLD, H. O. A., AND P. WHITTLE (1957): "A Model Explaining the Pareto Distribution of Wealth," *Econometrica*, 25, 591–595. [124,126]
WOLFF, E. (1987): "Estimates of Household Wealth Inequality in the U.S., 1962–1983," *The Review of Income and Wealth*, 33, 231–256. [123]
——— (2004): "Changes in Household Wealth in the 1980s and 1990s in the U.S.," Mimeo, NYU. [123,125]
ZHU, S. (2010): "Wealth Distribution Under Idiosyncratic Investment Risk," Mimeo, NYU. [128]

Dept. of Economics, New York University, 19 West 4th Street, 6th Floor, New York, NY 10012, U.S.A. and NBER; jb2@nyu.edu, *Dept. of Economics, New York University, 19 West Street, 6th Floor, New York, NY 10012, U.S.A. and NBER; alberto.bisin@nyu.edu,*

and

Dept. of Economics, National University of Singapore, Faculty of Arts & Social Sciences, AS2 Level 6, 1 Arts Link, Singapore 117570; ecszhus@nus.edu.sg.

Manuscript received February, 2009; final revision received June, 2010.