Economic Theory
https://doi.org/10.1007/s00199-024-01581-w

RESEARCH ARTICLE

Left and right: a tale of two tails of the wealth distribution

Marcello D'Amato[1,2] · Christian Di Pietro[3,4] · Marco M. Sorge[2,5]

Received: 29 September 2023 / Accepted: 13 May 2024

Abstract
We study a model of wealth accumulation in altruistic lineages, in which households face uninsurable risk, investment indivisibilities and borrowing constraints. A thick upper tail of the stationary distribution of wealth is shown to emerge as a robust prediction, irrespective of (1) the presence of multidimensional (wealth and ability) heterogeneity and non-convexities in human capital formation, and (2) the nature of parental bequest motives (joy-of-giving vs. paternalism). Additionally, (3) we identify conditions under which the unique, ergodic wealth distribution exhibits a mass point at the bottom of its support, where credit market imperfections continue to affect, along the convergence process, the structure of wealth transitions at the lineage level. Motivated by these results, we then analyze the sensitivity of the left tail to various frictions and fiscal instruments that affect bequest strategies and the ensuing transmission of wealth across generations. In particular, capital income or bequest taxes with no redistribution may reinforce economic mechanisms underpinning mobility traps in the left tail, thereby increasing the persistence of households in the lowest tiers of the wealth distribution.

Keywords Wealth distribution · Wealth inequality · Capital income risk · Credit market imperfections · Educational investment

✉ Marco M. Sorge
msorge@unisa.it

Marcello D'Amato
marcello.damato@unisob.na.it

Christian Di Pietro
christian.dipietro@uniparthenope.it

1 Department of Law, University of Napoli Suor Orsola Benincasa, C.so Vittorio Emanuele, 334, 80135 Naples, Italy

2 CSEF, Via Cintia, Monte S. Angelo, 80126 Naples, Italy

3 Department of Business and Economic Studies, University of Napoli Parthenope, Via Generale Parisi, 13, 80132 Naples, Italy

4 CRISEI, Via Generale Parisi, 13, 80132 Naples, Italy

5 Department of Economics and Statistics, University of Salerno, Via Giovanni Paolo II, 132, 84084 Fisciano, SA, Italy

Published online: 14 June 2024

JEL Classification D31 · H20 · I24

"Money, says the proverb, makes money. When you have a little, it is often easier to get more. The great difficulty is to make that little."
A. Smith, The Wealth of Nations, Chapter XI, p. 111.

1 Introduction

Household wealth data from a large panel of countries reveal two fairly general and robust patterns: first, a tilt to the top in the concentration of wealth over the last decades, see e.g. Wolff (1987), Davies and Shorrocks (2000), Klass et al. (2007), Piketty (2014), Vermeulen (2018); second, a sharp and growing divide among the richest and poorest households, with large shares of the population owning zero or even negative non-housing wealth, see e.g. Balestra and Tonkin (2018).[1]

Common to many advanced economies, this evidence appears to suggest a major role for general economic mechanisms, rather than specific political or institutional factors, in forging long-term features of wealth inequality. In this paper, we develop a unified theoretical framework for assessing the relative importance of two well-established market imperfections (non-insurability of income shocks and credit constraints) in producing patterns of wealth evolution in line with the observed structure of the data. Focusing on educational investment and financial (voluntary) bequests as means of intergenerational wealth transmission, we aim at studying the determinants of unequal opportunities in the lower rungs of the wealth ladder, where the fortunes of children might be shaped by family characteristics and economic factors other than those governing the process of wealth accumulation among the wealthiest.[2]

Our analysis explicitly addresses the issue, largely under-explored in previous work, of identifying conditions for the *co-existence* of a fat tail at the top *in conjunction with* a mass point at the bottom of the stationary distribution, in a model which preserves social mobility across time and generations (Chetty et al. 2014). To this end, we characterize the bequest mechanisms governing the transmission of human and non-human wealth along the cross-sectional distribution in the presence of multidimensional heterogeneity (in abilities and wealth) and idiosyncratic income risk. A main contribution of our analysis is to clarify that models that predict a fat upper tail due to exposure to non-insurable investment shocks are fully consistent with the occurrence of mobility

[1] Survey data from the Survey of Consumer Finances 2017 (Federal Reserve Board) reveals that the wealth gap between US richest and poorer families more than doubled from 1989 to 2016, with a record high 30% of zero-wealth households in 2016. Exploiting the second wave of the OECD Wealth Distribution Database, Balestra and Tonkin (2018) document that up to a quarter of the total population households in a number of OECD countries report negative net worth (i.e. liabilities exceeding the value of their assets).

[2] Other empirically relevant mechanisms that contribute to shaping wealth accumulation patterns and wealth inequality are due to life-cycle saving motives, individual lifetime uncertainty and accidental bequests, heterogeneous lifespans and medical expense risk at retirement—see De Nardi and Fella (2017) and Jappelli and Pistaferri (2017) for excellent surveys on the topic. While such forces certainly help to explain the evolution of consumption inequality over the life cycle and to rationalize observed wealth disparities between rich and poor households, they appear not to be driving the kind of saving behavior that is necessary to generate extreme concentration at the top and mobility traps at the bottom of the long-run wealth distribution, which is the focus of the present work.

traps at the bottom based on ex ante heterogeneity, so that credit market imperfections continue to affect, in the economic growth process, the structure of wealth transitions at the lineage level in the lower tiers of the wealth distribution.

To study intergenerational mobility and its relation with the evolution of wealth inequality, we incorporate ability heterogeneity and indivisible human capital investment into an otherwise standard framework of wealth accumulation. Specifically, we consider a simple economy populated by a large number of family lineages who differ with respect to innate learning ability ('ex ante heterogeneity') and wealth, due to the presence of uninsured income shocks ('ex post heterogeneity'). Parents decide both on financial bequests and investments in their children's education, based on their own preferences for altruism and investment costs. Given the presence of multiple channels of intergenerational wealth transmission, we explicitly study the implications of a *paternalistic bequest motive*, which engenders the enjoyment of a child's economic status through the lens of her parents' preferences, e.g. Becker and Tomes (1979). Unlike the *joy-of-giving* hypothesis (Andreoni 1990), by which parents benefit from the pure act of giving, paternalistic altruism entails portfolio choice considerations on the part of risk-averse parents, that render the educational investment problem more complicated to solve, and the implications of capital income risk and ex ante heterogeneity on wealth dynamics more interesting to evaluate.[3]

In our model, the combined effects of borrowing constraints, investment indivisibilities and idiosyncratic returns to financial and human wealth prompt heterogeneous bequest strategies in the cross section of lineages. We first show that the force of capital income risk is strong enough to deliver a *fat right tail* of the limit distribution of wealth, notwithstanding the presence of multidimensional heterogeneity (wealth and ability) and local non-convexities in human capital formation. However, these features crucially interact with the nature and extent of parental altruism to determine the transmission of economic status across generations *at the bottom* of the wealth distribution. Specifically, we establish rather general conditions on the model's fundamentals—conditions that admit natural economic interpretations—under which a unique, ergodic wealth distribution will emerge (Proposition 1) that exhibits *an atom*—i.e. a mass point—at (almost) zero wealth. In that wealth state upward mobility can only occur through occupational upgrading within lineages (Proposition 3). The intuition for this result is as follows. Under paternalistic altruism, the risky nature of investment opportunities and the presence of indivisibilities in human capital formation fully deter any kind of bequest choice at low wealth states, producing a mass of unskilled workers with no wealth inheritances at any time period—a mobility trap. Absent any direct wealth transfer in the group of the least wealthy families, the support of the stationary distribution of wealth will therefore inherit the structural properties of the left end of the support of the endogenous distribution of income, as determined by the educational choices made by parents. As a result, the unique steady state of the wealth accumulation process will display a mass of households bunched in the lowest tier of the cross-sectional distribution.

[3] The role of alternative bequest formulations in shaping the intergenerational transmission of wealth has only recently been addressed from an analytical standpoint, e.g. Pestieau and Thibault (2012). What contributes to shaping bequest motives of altruistic agents mostly is an empirical question, which appears to still lack a conclusive answer (Kopczuk 2013).

Based on credit market imperfections and the indivisibility of educational investment, this mechanism ceases to operate over larger wealth holdings, where financial wealth transfers become an overwhelmingly important source of asset accumulation within lineages; the cumulative effects of high returns on financial bequests (market luck) will then produce large and slowly declining wealth shares at the upper end of the stationary distribution (Proposition 2).

We show that modeling family altruism in paternalistic form is crucial for the existence of a mass point in the left tail of the stationary distribution. Under joy-of-giving, in fact, financial bequests have no compensatory goals in terms of consumption opportunities of later generations—see e.g. De Nardi (2004), Benhabib et al. (2011), De Nardi and Fella (2017). Any degree of altruism will then trigger positive financial bequests all along the equilibrium path, and the mechanics of compounded multiplicative shocks will sustain upward mobility from all states of the wealth space, implying that the constraining effects of credit market imperfections and ability heterogeneity are bound to vanish along the convergence process (Proposition 4).

Under paternalism, by contrast, the presence of multiple channels to invest in the future generation affects both the level of bequests and their composition. We identify conditions under which bequest incentives are not operative for a positive measure subset of lineages at each point in time, implying a positive probability for the wealth transitions to reach the lowest state of the wealth space. The distribution of abilities will then dictate the size of the atom in the left tail of the cross-sectional distribution of wealth: human capital formation among the least wealthy families, rather than exposure to uninsured capital income risk, will then prevail as a mobility engine that warrants ergodicity of the wealth dynamics. These results suggest that distinct accumulation factors and mechanisms, typically studied in isolation from each other, play distinct roles on within-lineage wealth transitions and the ensuing long-run traits of inequality that are consistent with the tail characteristics of empirical wealth distributions and with the degree of intergenerational mobility documented for advanced economies, e.g. Charles and Hurst (2003), Benhabib et al. (2019).

In cases where credit market imperfections persistently affect household bequest strategies and the ensuing mobility patterns, a natural question revolves around the kinds of fiscal policies needed to enhance economic trajectories of the group of the poorest vis-à-vis instruments for taxing the rich. In this respect, of particular interest are the implications of multidimensional heterogeneity for the design of public interventions aimed at reducing wealth disparities along the cross-sectional distribution. In such a setting, in fact, policy design should reckon with changing incentives to saving by taxed households in the lower and middle class, in particular those populating wealth states where no intergenerational transmission of financial wealth occurs.

Our second contribution is to link the properties of the left tail of the wealth distribution to various structural parameters of the model, such as the intensity of parental altruism, the heterogeneity in educational investment opportunities and fiscal policies that affect the saving and bequest behaviour of altruistic individuals.

According to a recent narrative that emphasizes market incompleteness as a key driver of top wealth concentration, proportional taxation of the realized rates of return on wealth unambiguously reduces inequality in the upper tail by curbing exposure to uninsured investment risk; numerical evidence from the simulation of calibrated models suggests, by contrast, that the sign of the response of the Gini coefficient for the whole distribution to a tax increase depends on the kind of fiscal instrument employed (estate versus capital income tax), see Benhabib et al. (2011). We complement these results by showing that estate and capital income taxes both produce ambiguous effects on the size of the atom in the left tail of the stationary wealth distribution (Proposition 7). The tax intervention influences the overall level of total bequests and their composition, encouraging wealth-constrained individuals to substitute financial for human capital investment in response to changes in their risk-return structure. In particular, households in the vicinity of the borrowing constraint, for whom labour earnings are the main component of permanent income, will decumulate non-human wealth to a greater extent than relatively richer households; this wealth effect on consumption can prove sufficiently strong to exacerbate downward mobility flows towards the bottom end of the wealth space, thereby increasing the measure of the poorest in the long run. Even when abstracting from general equilibrium effects on the risk-return structure of investment opportunities, the framework does not deliver a theoretically unambiguous characterization of the sensitivity of the left tail of the wealth distribution to the same kind of fiscal policies that have been recently called for to dampen the concentration of wealth among the wealthiest. This indeterminacy suggests that the tension between markets and governments and the relevance of policy prescriptions to narrow wealth inequality rest on foundational principles that still deserve to be explicitly discussed and more clearly understood.

The remainder of the paper is organized as follows. Section 2 provides the basic setup and assumptions of the framework of analysis. Sections 3 and 4 characterize optimal bequest strategies under alternative bequest motives (paternalism vs. joy-of giving), and work out the main theoretical results about the tail properties of the stationary distribution of wealth. Section 5 studies the effects of various frictions and fiscal policies on intergenerational wealth transmission and the ensuing patterns of social mobility. Section 6 connects our work to the extant literature and discusses some potentially interesting extensions. Section 7 offers concluding remarks. For ease of exposition, all the proofs and technical details are confined to the Appendix.

2 Model environment

We consider a simple heterogeneous-agent model with a measure-one continuum of individual lineages i in each period $t \geq 0$. In each lineage, parental wealth is allocated to current consumption needs as well as financial and educational bequests. Upon entering adulthood, agents inelastically supply labor in one of two different occupations, requiring different skill profiles (high and low). As in Loury (1981), the labor earnings technology depends on human capital; for simplicity, it is assumed that agents accessing (respectively, not accessing) education always become skilled (resp. unskilled) workers. Parents' own wealth thus is made of wealth inheritance, capital income (i.e. the return on financial bequests) and labor income (i.e. the occupation-specific wage).

Agent heterogeneity in the model comes in two concurrent forms: ex ante heterogeneity in abilities, which is relevant for educational investment decisions; and ex post return (hence wealth) heterogeneity, due to the occurrence of idiosyncratic income shocks against which agents cannot fully insure. Moreover, parents are unable to borrow against their children's future earnings or human capital (credit market imperfection), since inherited debts are not enforceable.

For reasons discussed above, we let the economic environment be described by the following

Assumption 1 (Capital income risk) The rate of return on wealth R is a stochastic process, independent and identically distributed (i.i.d.) over time and across lineages; it has a cumulative distribution function H and strictly positive density h on the bounded support $\Delta r \equiv [-1, \overline{r}]$, with $\overline{r} < \infty$; any finite moment of R is assumed to exist (with $\mathbb{E}[R] > 0$).

Assumption 2 (Labour income risk) The wage in the high-skill occupation Y is a stochastic process, i.i.d. over time and across lineages; it has a cumulative distribution function F with strictly positive density f on the bounded support $\Delta y = [\underline{y}, \overline{y}]$ ($\underline{y} > 0$), with Y and R being mutually independent. The wage in the low-skill occupation is equal to $\underline{y}$.

Assumption 3 (Heterogeneity in abilities) The cost of educational investment X is a stochastic process, i.i.d. over time and across lineages; it has a cumulative distribution function G with strictly positive density g on the bounded support $\Delta x = [0, \overline{x}]$, with $\overline{x} > \underline{y}$ and X, Y, and R mutually independent.

The first and second assumptions are rather standard, and have found empirical support (see discussion in Benhabib et al. 2011); a main implication of idiosyncratic shocks to capital income and labour earnings is the emergence of a positive ex post correlation between realized returns/wages and wealth. The third assumption generalizes Galor and Zeira (1993) by introducing heterogeneous abilities in population, as, for example, in Mookherjee and Napel (2007) and D'Amato and Di Pietro (2014). The empirical literature in the economics of education provides extensive evidence of heterogeneity in educational costs across children - see e.g. Henderson et al. (2011). Notice that we assume abilities not to be correlated over time, and posit that the most able individual is assigned a zero educational cost; these features simplify the analytical characterization of upward mobility via occupational upgrading but can be easily relaxed, and relatively sharper conditions devised for all of our results to hold true.[4]

2.1 Preferences

Let $\omega_{i,t}$ denote parental wealth at time t within lineage i, to be allocated among current consumption $c_{i,t}$, financial bequests $b_{i,t}$ and educational investment $e_{i,t} \cdot x_{i,t+1}$, where $e_{i,t} = 0$ (resp. $e_{i,t} = 1$) when parents in generation t decide not to invest (resp. invest)

[4] While linking heterogeneity of educational costs to cross-sectional variation in cognitive learning abilities, we do acknowledge the existence of other determinants such as fundamental personality traits and/or household characteristics (home-school proximity, parental monitoring capacity, etc), see e.g. Heckman et al. (2006).

in human capital of their children in generation $t+1$, who face known educational costs $x_{i,t+1}$.

Agents have separable preferences of the form

$$U(c_{i,t},\ b_{i,t},\ e_{i,t}) := u(c_{i,t}) + \chi \mathbb{E}[v(\omega_{i,t+1})], \tag{1}$$

under *paternalistic altruism*, where

$$\omega_{i,t+1} = (1 + r_{i,t+1})b_{i,t} + \underline{y} + e_{i,t} \cdot (y_{i,t+1} - \underline{y}). \tag{2}$$

where the parameter $\chi > 0$ captures the intensity of the bequest motive. The expectation operator $\mathbb{E}$ is conditional on the educational choice given the probability distribution of their offspring's wealth components (financial and human capital): bequest strategies must be formulated before uncertainty is resolved.

Different from Loury (1981), we assume that the child's ability is known to parents at the time of the educational investment decision, and that the latter produces a random rate of return which does not correlate with ability. Also, Loury (1981) allows the educational choice variable to be continuous rather than discrete, as we do in order to capture empirically plausible indivisibilities in the process of human capital formation.

An alternative formulation also considered in the ensuing analysis features a *joy-of-giving* bequest motive, whereby both financial and educational bequests enter the utility function directly (*bequest-as-last-consumption*), i.e.

$$U(c_{i,t},\ b_{i,t},\ e_{i,t}) := u(c_{i,t}) + \chi v\left(b_{i,t} + e_{i,t} \cdot x_{i,t+1}\right) \tag{3}$$

We further adopt the following

Assumption 4 Preferences satisfy

$$u = \frac{c_{i,t}^{1-\gamma}}{1-\gamma}, \qquad v = \begin{cases} \frac{\omega_{i,t+1}^{1-\gamma}}{1-\gamma}, & \text{under paternalism} \\ \frac{(b_{i,t}+e_{i,t}\cdot x_{i,t+1})^{1-\gamma}}{1-\gamma}, & \text{under joy-of-giving} \end{cases} \tag{4}$$

where $\gamma \geq 1$ is the coefficient of relative risk aversion.[5]

The class of CRRA preferences described in Assumption 4 represents a cornerstone of theoretical and applied models in finance and macroeconomics, and have been extensively used for the analysis of wealth inequality, transitional dynamics and long-run distributions, see, among many others, Benhabib et al. (2011), Benhabib et al. (2015), Zhu (2019), Wan and Zhu (2019), Birkner et al. (2023). While allowing for explicit characterizations of some of the objects of interest in our analysis (such as e.g. the local curvature of bequest policy functions), Assumption 4 can be relaxed, at

[5] When $\gamma = 1$, the limiting logarithmic forms for both u and v are adopted.

the cost of some complications of the analysis, in favour of a general state separable class of increasing, smooth and strictly concave functions u and v.[6]

2.2 Intergenerational transfers and wealth accumulation

In each lineage i, parents face the same choice about whether to bequeath educational investment or financial bequests or both. Optimal bequests emerge from a portfolio choice among alternative investments where one form of investment (educational) has an indivisible component, the other (financial) has not.

For each proposed formulation of the bequest motive, the parents' utility maximization problem is

$$\max_{c_{i,t},\, b_{i,t},\, e_{i,t}} \quad U\left(c_{i,t},\ b_{i,t},\ e_{i,t}\right) \tag{5}$$

$$s.to \quad c_{i,t} + b_{i,t} + e_{i,t} \cdot x_{i,t+1} \leq \omega_{i,t} \tag{6}$$

$$b_{i,t} \geq 0 \tag{7}$$

$$e_{i,t} \in \{0,\ 1\} \tag{8}$$

where (6) is the resource constraint defining feasible choices, and (7) underscores the inability of parents to borrow from their offspring.

Since the distributional features of exogenous variables (R, Y, X) are stationary over time, the expected utility maximization problem (5)-(8) is static at the family level, implying that optimal consumption and bequest policies are time invariant functions of the state variable ω_t. Keeping this in mind, we shall omit from now on the subscripts i and t on policy functions, in the interest of better readability.

In general, constrained optimization problems involving both discrete and continuous variable sets and nonlinear objectives such as (5)-(8) raise non-trivial analytic challenges. In order to allow for an intuitive understanding of the properties of financial and educational bequest policies, we adopt a straightforward branch-and-bound approach by which (i) the feasible set is partitioned by fixing the binary variable e (to either 0 or 1), involving a corresponding partition of the constrained optimization problem into two distinct sub-problems; (ii) each of these sub-problems is solved for continuous *pseudo policy functions* (c^*_e, b^*_e) via Karush-Kuhn-Tucker (KKT) optimality conditions that handle non-negativity constraints, and finally (iii) the *optimal bequest plan* (b^*, e^*)—where $b^* = b^*_e$ for $e = e^*$—is selected as the one producing the maximal objective value between the two partitions. For fixed e, the budget constraint (6) necessarily bites at the optimum and the constrained set of feasible points is convex; hence the pseudo consumption and pseudo bequest policies c^*_e and b^*_e solving the associated problems are fully characterized by the first-order KKT conditions. Direct comparison of indirect utilities induced by distinct bequest choices will then define the optimal bequest strategies as a function of parental wealth ω_t.

6 Loury (1981) works with a completely general specification of parental utility, as do, among others, Ray (2006), Mookherjee and Ray (2003), Mookherjee and Napel (2007) for the study of inequality dynamics and mobility issues in different contexts.

Once properly characterized, the set of optimal bequest policy functions will entail wealth transition patterns at the lineage level, whose limiting properties (ergodicity and structure of the tails) can be studied by standard techniques from the theory of Markov chains, e.g. Meyn and Tweedie (2009).

3 Paternalistic altruism

3.1 Pseudo bequest policies

Consider the following sub-program obtained from (5)-(8) by fixing $e \in \{0, 1\}$:

$$\max_{c_e,\, b_e} \quad U\left(c_e,\ b_e;\ e\right) \tag{9}$$

$$s.to \quad c_e + b_e \leq \omega_{i,t} - e \cdot x \tag{10}$$

$$b_e \geq 0 \tag{11}$$

Using (10) to substitute consumption out, the KKT first-order conditions for each partition $e \in \{0,\ 1\}$ are as follows

$$\left(\omega_t - b_e^* - e \cdot x\right)^{-\gamma} = \chi\left(1 + \mathbb{E}[R]\right)\left(\omega_{t+1}(b_e^*)\right)^{-\gamma} \quad \text{and} \quad b_e^* \geq 0 \tag{12}$$

or

$$\left(\omega_t - b_e^* - e \cdot x\right)^{-\gamma} \geq \chi\left(1 + \mathbb{E}[R]\right)\left(\omega_{t+1}(b_e^*)\right)^{-\gamma} \quad \text{and} \quad b_e^* = 0 \tag{13}$$

with condition (12) pinning down all candidate interior solutions, and (13) implied by the corner solution $b_e^* = 0$ and $c_e^* = \omega_t - e \cdot x$.[7]

We next characterize the properties of the pseudo financial bequest policies:

Lemma 1 *For fixed $e \in \{0, 1\}$, there exist a unique pseudo bequest policy b_e^* solving (9)-(11); moreover b_e^* (i) is continuous, non-decreasing and convex in ω_t, and (ii) satisfies $\lim_{\omega_t \to \infty} b_e^*/\omega_t = \phi$, with $\phi \in (0, 1)$.*

Proof See the Appendix. □

A key insight from Lemma 1 is that pseudo bequest policies b_e^* are convex functions of parental wealth—due to the precautionary effects triggered by income risk, as discussed in the next section—and become asymptotically linear over infinitely large wealth holdings, implying heterogeneous bequest behavior in the cross-sectional distribution of household wealth and an asymptotically constant marginal propensity to save for the wealthiest households.

To figure out conditions under which parents optimally choose to undertake educational investment, let us label *educational threshold cost* any $\tilde{x}(\omega_t) > 0$ at which

[7] At the initial date $t = 0$, we assume $\omega_{i,0} > 0$ in each lineage i. Observe that, from $t = 1$ onward, parental wealth will be larger than or equal to $\underline{y}$, i.e. the wage in the low-skilled occupation. That is, $\omega_t > 0$ for all $t \geq 0$ and all $x \in \Delta x$.

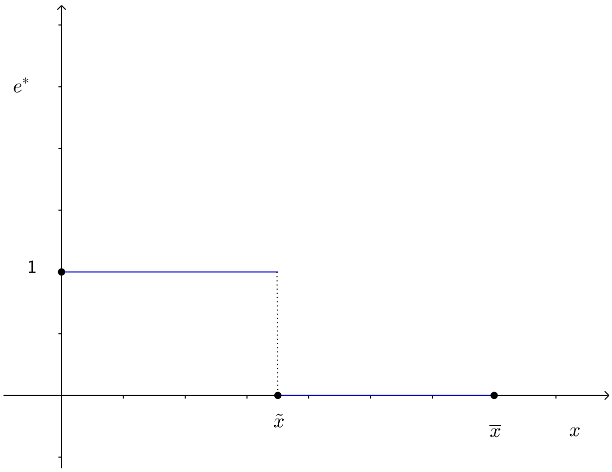

Fig. 1 The function e^*. In the figure, $\omega_t > 0$ is fixed and such that $\tilde{x}(\omega_t) < \bar{x}$

parents with wealth ω_t would enjoy the same utility by undertaking educational investment as by not doing so, i.e. let $\tilde{x}(\omega_t)$ solve

$$U\left(b_0^*;\ e=0\right) = U\left(b_1^*;\ e=1;\tilde{x}\right) \tag{14}$$

Under our stated assumptions about the distribution of abilities, a unique threshold $\tilde{x}$ exists as a function of parental wealth, and it entails the following: parents with equally talented kids but different wealth holdings can undertake different educational investment choices, so that heterogeneity in wealth background influences the distribution of earning capacities across workers; all else equal, the higher parental wealth, the larger the set of innate abilities conducive to human capital formation. The educational investment policy, part of the optimal bequest plan, is formally characterized as follows:

Lemma 2 $e^* = 1$ *if and only if* $x < \tilde{x}(\omega_t)$, *where* $\tilde{x}(\omega_t) : \mathbb{R}_+ \mapsto \mathbb{R}_+$ *is differentiable, and for all* $\omega_t > 0$ *it satisfies (i)* $0 < \tilde{x}(\omega_t) < \omega_t$, *and (ii)* $d\,\tilde{x}(\omega_t)/d\,\omega_t \in (0, 1)$.

Proof See the Appendix. □

Observe that, if $x = 0$, $e^* = 1$ prevails at any wealth level $\omega_t > 0$. By monotonicity of the threshold $\tilde{x}(\omega_t)$, there exist finite wealth levels ω_t such that $\tilde{x}(\omega_t) > \overline{x}$, i.e. parents who are wealthy enough will always find it profitable to engage in educational investment. Figure 1 below offers a depiction of the educational investment policy as a function of educational costs, for given parental wealth ω_t.

The optimal bequest plan (b^*, e^*)—where $b^* = b_e^*$ when $e = e^*$—exhibits the following property: at any given level of wealth, financial bequests of parents who also engage in human capital formation are never larger than those chosen by parents who rather abstain from educational investment. Intuitively, in the face of indivisibilities, parents use financial bequests to compensate kids with high educational costs, taking

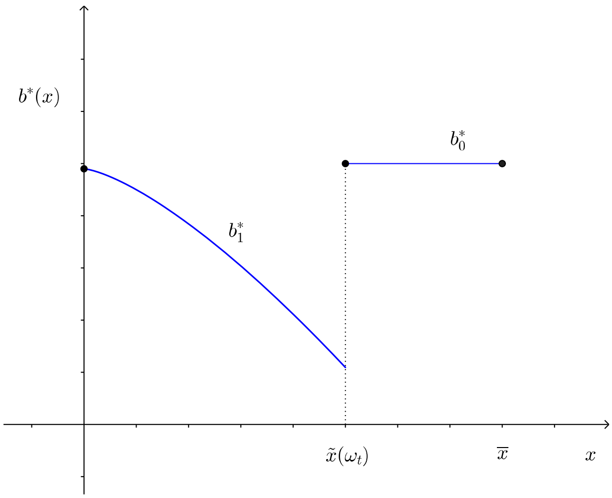

Fig. 2 The function $b^*(x)$. In the figure, $\omega_t > 0$ is fixed and such that $b_e^* > 0$ for each $e \in \{0, 1\}$

into account the risks and expected returns in either kind of investment. This results in a discontinuity in the optimal level of financial bequests given wealth, which in turn influences the transition of wealth within heterogeneous lineages—see Fig. 2 below. More formally, we state the following

Lemma 3 *Consider pseudo bequest policies b_e^* and, for fixed $\omega_t > 0$ at which $b_e^* > 0$ for $e \in \{0, 1\}$, define the mapping $b^*(x) : \Delta x \mapsto \Re_+$ as*

$$b^*(x) = \begin{cases} b_1^*, \textit{for} \quad x < \tilde{x}(\omega_t) \\ \\ b_0^*, \textit{for} \quad x \geq \tilde{x}(\omega_t) \end{cases} \tag{15}$$

Then (i) $b_1^* \leq b_0^*$ for all $x \in \Delta x$; and (ii) $b^*(x)$ strictly decreases over $[0, \tilde{x})$ and exhibits an upward jump discontinuity at $\tilde{x}(\omega_t)$.

Proof See the Appendix. □

We now turn to investigating how the economy's fundamentals (preferences and market characteristics) shape financial bequest strategies along the wealth distribution. A preliminary step of our analysis entails the identification of the wealth states for which optimal financial bequests are zero irrespective of the distribution of abilities and thus of the expected returns to human capital. In those states, borrowing constraints and indivisibilities will affect the intergenerational wealth transmission at the lineage level, and thus influence the evolution of the wealth distribution over time.

3.2 Bequest strategies and individual wealth transitions

In order to characterize wealth transitions at the lineage level we evaluate the KKT conditions in (13) at the point in which the pseudo financial bequest is optimally chosen to be zero conditional on the indivisible education choice. For $x \in \Delta x$, let us denote as $\tilde{\omega}_e(x)$, for each $e = \{0, 1\}$, the supremum of the set of wealth states ω for which b_e^* is zero. From (13) and the mutual independence between R and Y, any such $\tilde{\omega}_e(x)$ (for $\gamma \neq 1$) must solve

$$\tilde{\omega}_e(x) = \chi^{-\frac{1}{\gamma}} (1 + \mathbb{E}[R])^{-\frac{1}{\gamma}} \left(\mathbb{E}\left[\underline{y} + e\left(y - \underline{y}\right)\right]^{-\gamma} \right)^{-\frac{1}{\gamma}} + e \cdot x \tag{16}$$

where (as before) the expectation is taken with respect to the distributions of R and Y. Notice that, by virtue of Lemma 1 and Lemma 3, it holds $b_e^* = 0$ for all $\omega_t \leq \tilde{\omega}_e(x)$, $e \in \{0, 1\}$. We can accordingly interpret $\tilde{\omega}_0$ as the threshold below which financial transfers are not part of the optimal bequest plan, regardless of the educational choice undertaken by the parents; and $\tilde{\omega}_1(x)$ as the cost-specific threshold above which both financial and educational bequests occur whenever convenient, with $\tilde{\omega}_0 < \tilde{\omega}_1(x)$ for any educational investment cost $x \in \Delta x$.[8]

Recall that, in each lineage i, $\omega_t \geq \underline{y}$ for all $t > 0$ and all $x \in \Delta x$, i.e. the wage in the unskilled occupation defines the lowest level of parental wealth. The following Lemma characterizes the existence of wealth states entailing zero financial bequests irrespective of the educational investment choice of parents:

Lemma 4 *There exists $\tilde{\omega}_0 > \underline{y}$ such that $b_e^* = 0$ for all $\omega_t \leq \tilde{\omega}_0$ and each $e \in \{0, 1\}$ if and only if $\chi(1 + \mathbb{E}[R]) < 1$; $\tilde{\omega}_0$ decreases with the intensity of the bequest motive (χ) and with the coefficient of relative risk aversion (γ).*

Proof See the Appendix. □

The economic interpretation of Lemma 4 is straightforward. In the model, two key forces shape individual bequest decisions. The first is the effective preference for altruism, summarized by the term $\chi(1 + \mathbb{E}[R])$, see the optimality conditions (12) and (13): all else equal, the larger χ, the stronger the incentive for parents to sacrifice their own consumption in order to leave positive financial bequests. The second force is the desire to self-insure against income risk in a world of incomplete markets: intergenerational wealth transmission (and thus the accumulation of wealth at the lineage level) reflects the precautionary saving behavior of parents who exhibit aversion to downside risk—or *prudence*, after Kimball (1990)—and thus wish to shield their offspring's wealth from bad outcomes (low returns on financial and non-financial wealth). All else equal, the larger γ, the stronger the precautionary motive that stimulates positive financial transfers to children.[9]

Three possible cases could in principle be obtained:

[8] From now on, we shall not consider the log-log preference specification characterized by $\gamma = 1$. All of our results can nonetheless be explicitly restated to encompass it.

[9] It must be stressed that, in our setting, individual attitudes toward risk and toward intergenerational variations are automatically connected: the degree of relative risk aversion, the intensity of the precautionary saving motive and the elasticity of substitution between consumption and bequests are all governed by the single parameter γ. Under these circumstances, if parents exhibit a strong tendency to avoid risk, they will necessarily feature a relatively stronger preference for saving against income risk (higher prudence) and will also be less willing to replace transfers to kids with their own consumption. In addition, even for homothetic, additively (in arguments and states) separable preferences, a low degree of prudence may fail to

(i) $\tilde{\omega}_0 \leq \underline{y}$—At any level of wealth $\omega_t \geq \underline{y}$ the bequest motive is operative and $b_e^* > 0$ for each educational choice $e \in \{0, 1\}$. This occurs if paternalistic altruism is sufficiently strong (χ sufficiently high) and/or if the average rate of return on bequests is high enough ($\mathbb{E}[R]$ sufficiently large), irrespective of the degree of relative risk aversion. In this case, the structure of individual wealth transitions becomes independent of credit market imperfections over time, so that, provided it entails convergence to a non-degenerate limit distribution, this latter will exhibit no mass point on the left, mirroring the characterization in Benhabib et al. (2011)—a case we thus choose not to investigate further.

(ii) $\tilde{\omega}_0 \in \left(\underline{y}, \overline{y}\right)$—At wealth levels $\omega_t \leq \tilde{\omega}_0$ the bequest motive is inactive for households facing large enough educational costs (i.e. those with $x \geq \tilde{x}(\omega_t)$): households below the wealth threshold will hit the borrowing constraint and abstain from human capital investment. Positive financial bequests require parental preferences to be sufficiently altruistic towards children and prudent vis-à-vis the market risk and the rate of return attached to the financial bequest. When the choice also entails educational bequests, such returns and risks are relative to those in the skilled labor market. For those families who invest in education, financial bequests are part of the optimal bequest plan only at sufficiently large levels of wealth. When this is the case, wealthier households will leave higher bequests, producing persistence in wealth.

(iii) $\tilde{\omega}_0 \geq \overline{y}$—At wealth levels $\omega_t \leq \tilde{\omega}_0$, it holds $b_e^* = 0$ for each educational choice $e = \{0, 1\}$ and $e^* = 1$ only for lineages with $x < \tilde{x}(\omega_t)$. In this case, however, even the largest return on human capital (i.e. the highest wage in the high-skill occupation) will prevent households from escaping the lowest part of the wealth space where direct wealth transfers do not take place. Credit market imperfections and ability heterogeneity will then fully govern individual wealth transitions, while capital income risk ceases to operate as an engine of mobility.

Information from Lemma 1 and Lemma 4 is summarized in Fig. 3, representing an example of the weakly increasing, convex and asymptotically linear financial bequest policy for given educational cost $x \in \Delta x$.

Given optimal bequest strategies (b^*, e^*), individual wealth transitions are in the form

$$\omega_{t+1} = \begin{cases} (1+r)\, b_0^*(\omega_t) + \underline{y}, & \omega_t \leq \hat{\omega} \\ (1+r)\, b_1^*(\omega_t) + y, & \omega_t > \hat{\omega} \end{cases} \tag{17}$$

where $\hat{\omega} = \left\{ \omega_t > \underline{y} \,\middle|\, \tilde{x}(\omega_t) = x \right\}$ is the threshold below which agents facing educational costs x optimally decide not to engage in educational investment ($e^* = 0$). Observe that this threshold is unique by virtue of Lemma 2.

Footnote9 continued

induce precautionary behavior in the presence of multiple endogenous risks (Gollier 2001). Intuitively, since investing in assets with uncertain returns increases children's exposure to overall risk, risk-averse parents might want to compress savings so as to scale down the variance of their offspring's wealth—a substitution effect due to risk aversion, that counteracts the income effect related to prudence, e.g. Rothschild and Stiglitz (1971).

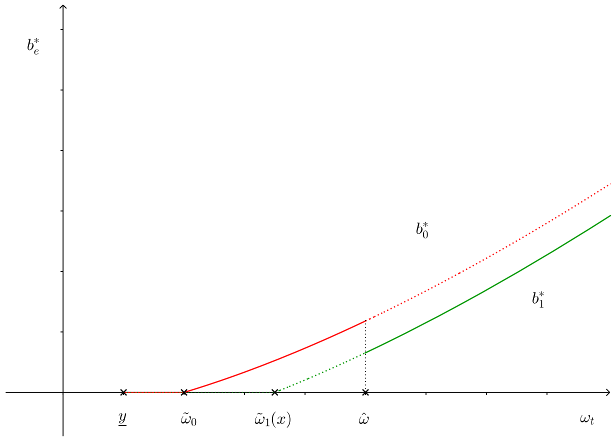

Fig. 3 The function b_e^*. In the figure, $x \in \Delta x$ is fixed, $\hat{\omega} = \left\{\omega_t > \underline{y} \,\middle|\, \tilde{x}(\omega_t) = x\right\}$, and structural parameters are such that $\underline{y} < \tilde{\omega}_0$ and $\tilde{\omega}_1(x) < \hat{\omega}$

Provided lineage i's initial wealth (at $t = 0$) and the sequences of random returns/wages (R, Y, X) are independent, the wealth transition (17) defines a Markov chain $\{\omega_t\}$ that evolves on a general state space $W \subseteq [\underline{y}, \infty)$ according to some probability law μ, whereby the equilibrium wealth level ω_{t+1} achievable by agents receiving financial and educational bequests (b^*, e^*), as a function of current wealth ω_t, depends on i.i.d. shocks (R, Y) hitting at time $t+1$, for different values of $x \in \Delta x$.

The wealth accumulation process (17) involves two kinds of non-linearity: one pertains to the occurrence of zero bequests in the lower states of the wealth space (Lemma 4), and to the convexity of the pseudo bequest policies (Lemma 1); the other, more troublesome from the analytic point of view, hinges on the indivisible nature of educational investment and the ensuing threshold rule (Lemma 2 and Lemma 3), which induce a discontinuity in the financial bequest policy. These peculiar features will require some nontrivial adaptations of the arguments put forward in Benhabib et al. (2015) for the derivation of the limit properties of the wealth distribution.

As will be made clear in the following, the occurrence of a mass point at the bottom of the stationary distribution demands that the forces mentioned above (altruism, risk aversion, prudence) interact so as to generate *non-saving behavior* in the lower tiers of wealth: parents for whom it is never optimal to engage in any kind of risky bequest choice against the expected benefits to their heirs. Existence of a unique invariant distribution with a fat upper tail, by contrast, requires that individual transitions be not trapped in the lowest region of the wealth space: the presence of indivisibilities in educational investment will thus call for a sufficiently large wage premium between occupations to allow labor income risk to stimulate precautionary savings and positive bequests on the part of less wealthy households.

3.3 The stationary wealth distribution

Ergodic convergence of the Markov chain (17) requires that every state in the wealth space W is accessible from any other (*irreducibility*), and that states are visited at irregular times (*aperiodicity*); a suitable non-explosion (*geometric drift*) condition is also to be satisfied to warrant convergence of (17) to a unique invariant distribution. The following result holds:

Proposition 1 *Let $\chi(1+\mathbb{E}[R]) < 1$. The stochastic process $\{\omega_t\}$ generated by (17) is ergodic, and thus converges to a unique stationary distribution ω_∞ with limit support S_∞, where*

$$S_\infty = \begin{cases} [\underline{y}, \infty) \text{ if and only if } \chi(1+\mathbb{E}[R]) > \left(\underline{y}/\overline{y}\right)^{\gamma} \\ \left[\underline{y}, \overline{y}\right] \text{ otherwise} \end{cases}$$

Proof See the Appendix. □

Proposition 1 identifies two possible cases, depending on the location of the threshold $\tilde{\omega}_0$ in the wealth space W:

Case 1. $\chi(1+\mathbb{E}[R]) \in \left(\left(\underline{y}/\overline{y}\right)^{\gamma}, 1\right)$ or equivalently $\tilde{\omega}_0 \in \left(\underline{y}, \overline{y}\right)$—In this case, the lower bound $\underline{y}$ of the wealth space W is a *partially reflecting barrier*: for any given lineage i, if, at some time $t \geq 1$, it holds $\omega_{i,t} = \underline{y}$, then $Pr\left(\omega_{i,t+1} = \underline{y} \mid \omega_{i,t}\right) \in (0, 1)$. At each point in time, at the wealth level $\underline{y}$ educational investment will be undertaken by a subgroup of low-wealth families families who hit the borrowing constraint. As established below, the left tail of the stationary distribution will accordingly be populated by all those lineages who are temporarily stuck in the unskilled occupation and enjoy no wealth inheritances from their parents. At higher wealth states, capital income risk will fuel accumulation patterns that can produce power-law decay of wealth shares in the right tail, as in Benhabib et al. (2011).

Case 2. $\chi(1+\mathbb{E}[R]) \leq \left(\underline{y}/\overline{y}\right)^{\gamma}$ or, equivalently, $\tilde{\omega}_0 \geq \overline{y}$. In this case, the lower bound $\underline{y}$ of the wealth space W is a partially reflecting barrier, but the set Δy is *absorbing*: for any given lineage i, if, at some time $t \geq 1$, $\omega_{i,t} \in \Delta y$, then $Pr\left(\omega_{i,t+1} \in \Delta y \mid \omega_{i,t}\right) = 1$. The economic interpretation of this result is straightforward: once visiting the lower region of the wealth space (i.e. below the threshold $\tilde{\omega}_0$), even skilled, top income earners would not find it optimal to save out of their resources and transfer wealth directly to their children; this in turn prevents lineages from experiencing accumulation patterns beyond a given finite level. As a result, the stationary distribution of wealth will necessarily overlap with the distribution of labor earnings across households, as dictated by the return to educational investment (the latter being driven by the heterogeneity in the educational costs), e.g. Loury (1981).

Building on these insights, we next characterize the tail behavior of the stationary distribution of wealth. To tackle the non-linearity issues pointed out above, we formalize a simple logic based on Jensen's inequality that allows us to rely on standard regularity conditions on the stochastic properties of the return on wealth and labour income. Specifically, we take advantage of the curvature and asymptotic properties of pseudo policy functions b_e^*—established in Lemma 1—and prove tail fatness of the stationary distribution ω_∞ in three steps: (i) first, we construct a piece-wise linear (in wealth) function b^a which acts, at any wealth level, as a lower bound for the pseudo bequests policies b_e^*, irrespective of the optimal educational choice e^* (and, thus, of the underlying heterogeneity in abilities across lineages); (ii) second, we set up an auxiliary process $\left\{\omega_t^a\right\}$ following the accumulation law

$$\omega_{t+1}^a = (1+r)b^a(\omega_t^a) + y \tag{18}$$

that, under mild restrictions on the income shocks (R, Y), converges to a unique stationary distribution w_∞^a with an asymptotic Pareto tail; and (iii) third, we establish that, for any admissible path (R, Y), the distribution of $\left\{\omega_t^a\right\}$, at each time period, is first-order stochastically dominated by the distribution of the actual wealth accumulation process $\{\omega_t\}$. By ergodicity, the unique stationary distribution to which the wealth accumulation process (17) converges will necessarily exhibit power-law decay in the upper tail.

We can compress this logic into the following

Proposition 2 *Let* $\chi(1+\mathbb{E}[R]) < 1$. *If* $\overline{y}$ *and* $\mathbb{E}[R^2]$ *are sufficiently large, the unique stationary distribution* ω_∞ *has a fat right tail, i.e. there exists* $\nu \in (1, 2)$ *such that*

$$\liminf_{\omega \to +\infty} \frac{Pr\left(\omega_\infty > \omega\right)}{\omega^{-\nu}} \geq C \tag{19}$$

for some constant $C > 0$.

Proof See the Appendix. □

The intuition for why the stationary distribution ω_∞ exhibits fat-tailed behavior is the same as in Benhabib et al. (2015): if upward mobility from the lowest wealth states is ensured by sufficiently large fluctuations in labour income ($\overline{y}$ high enough) relative to educational costs, favorable capital income shocks (warranted by sufficiently volatile financial returns, i.e. $\mathbb{E}[R^2]$ large enough) can compound one another to push lucky lineages towards arbitrarily high wealth levels so that, provided it contracts on average, the dynamics of wealth (17) will converge, in a strong probabilistic sense, to a limit distribution with a Pareto-like right tail.

While Proposition 2 does not pin down the tail index of the stationary wealth distribution, it entails precise bounds on the actual rate of decay of wealth shares in top percentiles. Indeed, let ξ denote the power tail index of the theoretical wealth distribution that our model predicts; by ergodicity of individual wealth transitions, it must hold $\xi > 1$, and by virtue of (19) it must hold $\xi \leq \nu < 2$ (or the right tail would decay faster than that of a Pareto distribution with exponent ν). The tail index ν of the limit distribution generated by the auxiliary linear recursion (18) is exclusively expressed in terms of preference parameters and moments of the distribution of the rates of return on non-human wealth—see equations (20) and (23) in the Appendix; upon calibrating the structural features of the model, a value for ν can be simulated and thus used to discipline the ability of our framework to replicate the documented power-law behavior of the upper tail of empirical wealth distributions.[10]

If $\tilde{\omega}_0 > \underline{y}$, financial bequests are not part of the optimal bequest plan of low-wealth households. In this case, no intergenerational wealth transmission occurs in the lowest states of the wealth space, and the stationary wealth distribution will necessarily exhibit a mass point at the lower end of its support, whose size is determined by the distribution of abilities across lineages. The following result formalizes this observation:

Proposition 3 *Let $\chi(1 + \mathbb{E}[R]) < 1$. Then the stationary wealth distribution ω_∞ exhibits an atom in $\underline{y}$, i.e.*

$$\mu^*\left(\underline{y}\right) = \frac{\int_{\omega \in (\underline{y},\, \tilde{\omega}_0]} [1 - G(\tilde{x}(\omega))] d\mu^*}{G(\tilde{x}(\underline{y}))} > 0 \tag{20}$$

where μ^* is the invariant measure of ω_∞.

Proof See the Appendix. □

From a technical standpoint, an atom in the support of the distribution of the additive shock to wealth dynamics (here, occupation-specific wages) becomes an atom in the support of the stationary distribution of wealth since the multiplicative component embodying the multiplicative shock (here, the gross return on financial bequests) can be null with positive probability (when parents' optimizing behavior entails zero financial transfers for a positive measure of lineages).

Propositions 2 and 3 jointly clarify that right-skewed wealth distributions with a fat right tail, as ultimately shaped by exposure to capital income risk, can emerge in environments where wealth backgrounds, heterogeneous abilities and credit market imperfections all interact in producing a temporary mobility trap at the bottom of the support. At steady state, even extremely poor lineages will experience upward mobility via occupational upgrading, and yet other lineages will replace them in the poorest group due to unlucky income draws and the ensuing optimal investment plan entailing zero bequests.

We can invoke Kac's theorem (Meyn and Tweedie 2009) to connect the size of the atom in the stationary distribution—i.e. $\mu^*(\underline{y})$—with the expected residence time of lineages in the lowest state of the wealth space: the larger the invariant measure of the atom, the lower the mean return time of households to the bottom end of the wealth distribution, meaning that, at steady state, the persistence of lineages in the left tail is relatively higher. In principle, this result might be useful for (i) simulation-based validation exercises, that rely on matching e.g. the left tail of the theoretical wealth

10 Klass et al. (2007) offer evidence that the top end of the wealth distribution in the US is well approximated by a Pareto law with an average exponent of 1.49.

distribution (percentage of households at zero non-housing wealth) along with standard measures of inequality along the entire distribution (e.g. the Gini coefficient) with their data counterparts; and possibly also for (ii) empirical work aimed at constructing measures of long-run poverty persistence and mobility from the bottom wealth percentiles, along the lines of e.g. Ray and Genicot (2023), something we leave to future research.[11]

4 Altruism in joy-of-giving form

For the sake of comparisons with other contributions where intergenerational altruism is assumed in the form of joy-of-giving, e.g. Benhabib et al. (2011), see equation (3), we now analyze the implications for wealth accumulation of this bequest motive.

Solving the parents' utility maximization problem for the system of the KKT conditions associated with each education choice $e \in \{0, 1\}$ delivers the following interior pseudo bequest functions

$$b_0^* = \left(1 + \chi^{-\frac{1}{\gamma}}\right)^{-1} \cdot \omega_t \tag{21}$$

$$b_1^* = b_0^* - x \tag{22}$$

Notice that in the presence of a joy-of-giving bequest motive, the utility flow obtained from leaving a bequest (of whatever nature) depends *only* on the size of the bequest; in the presence of a borrowing constraint, educational investment $e^* = 1$ in lineage i can thus be taken to occur if and only if $x \leq b_0^*$, where $b_0^* = 0$ never occurs at positive wealth levels.

In sharp contrast to the case of paternalistic altruism, educational and financial bequest choices are tightly intertwined when altruism is shaped by a joy-of-giving motivation. In fact, with an infinite marginal utility of zero bequests (either financial or educational or both), optimal bequests are always positive. Since the transition of wealth within each lineage is given by

$$\omega_{t+1} = \begin{cases} y - (1+r)\,x + (1+r)\,b_0^* & x \leq b_0^*(\omega_t) \\ \underline{y} + (1+r)\,b_0^* & x > b_0^*(\omega_t) \end{cases} \tag{23}$$

the following result is easily established

Proposition 4 *In the presence of a joy-of-giving bequest motive, if the wealth accumulation process (23) is ergodic, then the unique stationary distribution ω_∞ exhibits no mass point in $\underline{y}$.*

Proof See the Appendix. □

In the presence of a joy-of-giving bequest motive, capital income risk will provide the necessary mobility across wealth levels that prevents lineages from getting

11 We thank an anonymous Reviewer for soliciting us to ponder over the implications of Kac's theorem for our analysis.

persistently trapped in the lowest state of the wealth space, making credit market imperfections and investment indivisibilities immaterial over the long run.[12]

5 The size of the atom: some comparative statics

So far, our analysis has confirmed that capital income risk plays a fundamental role in generating fat-tailed behavior of the right tail of the stationary distribution of wealth, even when households exhibit a paternalistic bequest motive and face non-convexities in educational investment. It has further identified conditions under which occupational upgrading via human capital formation, rather than uninsured financial shocks, is the key driver of upward mobility for low-wealth lineages experiencing the vicious confluence of borrowing constraints and investment indivisibilities.

We now study the comparative statics of the size of the atom in the left tail of the stationary wealth distribution, i.e. $\mu^*(\underline{y})$, with respect to some of the structural parameters and fiscal policies that shape social mobility patterns in our economy. To this end, we will restrict our attention to cases where the support of the stationary distribution contains a mass point at its lower end, i.e. altruism is taken to be paternalistic and preference parameters/exogenous returns on investment are such that $\tilde{\omega}_0 > \underline{y}$—see Lemma 4.

5.1 Intensity of the bequest motive

As shown in Benhabib et al. (2011) and Zhu (2019), a stronger preference for altruism in economies with uninsured investment risk reinforces wealth concentration in the right tail of the stationary distribution. By the same logic, a stronger concern for the wealth status of children entails, at any wealth level, stronger educational efforts, thereby reducing the probability of transitioning toward the lower states of the wealth space. As a result, the measure of the least wealthy households in the stationary distribution decreases, as stated next:

Proposition 5 *Let $\chi(1+\mathbb{E}[R]) < 1$. The measure $\mu^*(\underline{y})$ of the atom in the stationary wealth distribution decreases with the intensity of altruism χ, all else equal.*

Proof See the Appendix. □

Notice that this result pertains to both cased identified in Proposition 1, i.e. irrespective of whether the stationary wealth distribution exhibits a fat upper tail or not.

5.2 Educational costs

Ex ante heterogeneity in abilities naturally influences upward mobility flows out of the lowest wealth states and the composition of optimal investment portfolios in the cross-section of lineages. We next study how the stochastic properties of the distribution of

[12] A different preference specification allowing for imperfect substitution between financial and educational bequests which maintains the joy-of-giving paradigm does not alter any of the qualitative results established in this section (details are available upon request).

educational costs in the population affects the size of the mass point at the bottom of the support of the stationary distribution of wealth. Intuitively, the size of the atom (stochastically) increases with the likelihood of relatively larger educational costs, as less families will afford, in each point in time and for any level of wealth, investing in human capital formation. Formally:

Proposition 6 *Let* $\chi(1+\mathbb{E}[R]) < 1$. *All else equal, consider two distinct distributions* G *and* G' *for the educational investment costs* X *such that* $G'(X \leq x) \leq G(X \leq x)$ *for all* $x \in \Delta x$*; and let* $\mu^*_{G'}(\underline{y})$ *and* $\mu^*_G(\underline{y})$ *denote the measure of the atom in* $\underline{y}$ *in the stationary wealth distribution under* G' *and* G*, respectively. Then it holds* $\mu^*_{G'}(\underline{y}) > \mu^*_G(\underline{y})$.

Proof See the Appendix. □

Again, this result applies to both the cases identified in Proposition 1.

5.3 Fiscal policies

The presence of multiple dimensions of heterogeneity across individuals, such as in underlying labor earning abilities and investment returns, has been advocated as one of the main arguments in favour of positive capital income taxation, for both equity and efficiency reasons.[13] As studied in Benhabib et al. (2011), in economies with uninsured financial shocks, top wealth concentration proves highly sensitive to fiscal policies that curb the boosting effect of capital income risk on intergenerational wealth transfers: all else equal, the higher the capital income and/or bequest taxes, the thinner the upper tail of the stationary distribution. It is straightforward to see that the same result holds in our setting too.

A focus of interest in our current setting is the impact of similar tax instruments on the left tail. To study whether and how these fiscal policies (with no redistribution) shape also the properties of the bottom end of the stationary distribution of wealth, we consider the model including a proportional tax $\tau_b \in (0, 1)$ on financial bequests (an *estate tax*, following the terminology in Benhabib et al. 2011) or, equivalently, a proportional tax $\tau_r \in (0, 1)$ on realized financial returns (*capital income tax*). In the former case, $(1-\tau_b)b^*$ will define post-tax optimal bequests enjoyed by children; in the latter, we simply re-define the random rate of return r_{t+1} as the pre-tax rate and let $(1-\tau)(1+R)$ denote the post-tax rate of return on financial bequests. The dynamics of wealth at the lineage level under either type of tax is

$$\omega_{t+1} = \begin{cases} (1-\tau_j)\,(1+r)\,b_0^{*'}(\omega_t) + \underline{y}, \; \omega_t \leq \hat{\omega}' \\ (1-\tau_j)\,(1+r)\,b_1^{*'}(\omega_t) + y, \quad \omega_t > \hat{\omega}' \end{cases} \tag{24}$$

[13] For a survey on the welfare arguments about capital taxation, see Bastani and Waldenström (2020). For the welfare analysis of public programs in the education sector in the presence of credit and insurance market imperfections, see the classic treatment in Loury (1981) and a more recent discussion in Mookherjee and Napel (2021).

where the subscript j denotes the operative tax rate (i.e. $j = r$ or $j = b$) and the superscript $'$ labels endogenous objects (bequest policies and thresholds) resulting in the economy.

A positive tax on bequests (or on realized returns on non-human wealth) has three main effects on the individual wealth transition (24) in partial equilibrium: first, it mechanically changes the average return on direct wealth transfers, as well as its variability (a first-order stochastic dominance effect); second, due to households' prudence, it might trigger a *compensating effect* by inducing parents to bequeath, for given their educational choice, a larger share of their wealth to their offspring, to counter the depressive effects of taxes on post return inheritances, e.g. Becker and Tomes (1987); and third, it affects parents' optimal portfolio choices by modifying the expected risk-return profiles of investment opportunities—a *saving composition effect*.

The first (exogenous) effect works through the accumulation law (24) by compressing the structure of returns on non-human wealth. The second and third (endogenous) effects both materialize through households' optimal bequest plan as follows: first, conditional on the educational choice, post-tax pseudo financial bequests $(1 - \tau_j)b_e^{*'}$ necessarily decrease, regardless of the strength of the compensating effect; second, the educational investment threshold $\tilde{x}'$ can increase with the tax rate and be strictly larger than that governing human capital formation in the no-tax scenario, for a positive measure subset of wealth-constrained households. Thus, the introduction of an estate or capital income tax fosters the formation of human capital across lineages at middle wealth levels, and this capital can serve as a safety net, preventing them from moving down to the bottom of the distribution. Formally it holds

Lemma 5 *Let* $\chi(1 + \mathbb{E}[R]) < 1$. *Then*

(i) at any $x \in \Delta x$, *for each* $e \in \{0, 1\}$ *the pseudo financial bequest policies* $b_e^{*'}$ *under either tax* $\tau_j \in (0, 1)$, $j = b, r$ *and their analogues* b_e^* *in the no-tax benchmark* $(\tau_j = 0)$ *satisfy*

$$(1 - \tau_j)b_e^{*'} \leq b_e^*, \quad \forall \omega_t \geq \underline{y};$$

(ii) if the maximal educational cost $\bar{x}$ *is sufficiently large, then there exist tax rates* $\tau_j \in (0, 1)$, $j = b, r$ *and a set* $B \subset W$ *of positive measure such that the educational threshold cost* $\tilde{x}'(\omega_t)$ *and its analogue* $\tilde{x}(\omega_t)$ *in the no-tax benchmark* $(\tau_j = 0)$ *satisfy*

$$\tilde{x}'(\omega_t) > \tilde{x}(\omega_t), \quad \forall \omega_t \in B$$

Proof See the Appendix. □

The impact of fiscal policies on the transmission of wealth at the lineage level naturally depends on the interplay across the three effects mentioned above. Capital income or estate taxes expand the set of wealth states where financial bequests are not part of the optimal bequest plan of households; at higher wealth levels, where the borrowing constraint is not binding, they also induce a portfolio re-allocation towards human capital investment that becomes relatively more attractive as a mean of intergenerational wealth transmission. If sufficiently strong, this effect can reduce the size of the atom in the lowest state of the wealth space: if the enhanced process of occupational upgrading and the ensuing persistence of human capital overcompensate the contraction in the intergenerational transmission of non-human wealth, this composition effect dominates the other two and ends up improving upward movements from the lower rungs of the wealth ladder. However, the wealth effect imparted by taxation on savings is relatively stronger for households who are relatively poorer in financial wealth, and might exacerbate downward mobility of lineages in the vicinity of the borrowing constraint towards the bottom end of the wealth space.

The next Proposition formalizes the foregoing arguments about the ambiguous effects of estate/capital income taxation on the left tail of the stationary distribution of wealth:

Proposition 7 *Let $\chi(1+\mathbb{E}[R]) < 1$. Depending on the risk-return structure of investment opportunities, the measure of the atom in the stationary wealth distribution can increase or decrease in response to the introduction of an estate tax τ_b and/or of a capital income tax τ_r.*

Proof See the Appendix. □

From a policy perspective, this latter result suggests a word of caution in evaluating the effects on wealth inequality of fiscal intervention that abstracts from redistributive considerations. When wealth is observable, taxing wealth at the top and redistributing government revenue lump-sum at the bottom would in fact help to mitigate the effects of the distortions on the saving behavior of the poorest, for any given distribution of (unobservable) abilities among the tax-payers. Receipts from taxation of top wealth owners could also be used to design transfers to low-wealth families that are conditioned on educational investment by parents, or to lower their educational costs by e.g. promoting public schooling, with similar effects on the left tail of the wealth distribution. In light of the previous results, the analysis is straightforward and we do not pursue it here.

6 Related literature and possible extensions

Our theoretical exploration of the tail properties of the stationary wealth distribution is inspired by a growing literature studying saving mechanisms and the transmission of bequests and human capital in the presence of idiosyncratic risk.

While Marshall (1890) explicitly acknowledged the concern for children as a key reason for saving, Pareto (1903) was the first to suggest that preferences for altruism and bequest strategies were the main driving force of the observed structure of wealth distributions. This view was later on taken by Gary Becker in a few contributions, formalized in Becker and Tomes (1979) and Becker and Tomes (1986). Though lacking a formal derivation of the equilibrium distribution of wealth, these studies trace the origins of the inequality among families back to more fundamental heterogeneity in socio-economic inheritable characteristics ("endowments", "ability", "social connections") so that the economic status tends to persist at the lineage level. On top of the transmission of endowments and the inheritability of socio-economic connections, the main channel of the persistence of inequality was the acknowledged credit market imperfection in human capital formation. In this view, little or no role is played by heterogeneity in the returns to financial wealth.[14]

Since Bewley (1977), Bewley (1983) and Aiyagari (1994), incomplete markets models of optimal life-cycle consumption-saving behavior (also known as *Bewley economies*) have been successfully employed to shed light on the determinants of the evolution of the equilibrium distributions for consumption, savings, and wealth. While able to match empirical measures of inequality (e.g. the Gini coefficient), these models predict wealth distributions that fail to display a fat enough upper tail compared to the data, since saving incentives quickly dissipate over sufficiently high wealth levels.[15]

Based on substantial evidence on the impact of capital income on wealth concentration in recent decades—e.g. Flavin and Yamashita (2002), Moskowitz and Vissing-Jørgensen (2002), Wolff (2006)—Benhabib et al. (2011) formally prove that the fundamental force driving wealth accumulation at the top end of the wealth distribution is idiosyncratic investment risk: propagating along intergenerational links via the life-cycle and bequest behaviour of forward-looking agents, uninsured shocks to financial returns accumulate multiplicatively into wealth, boosting its concentration in the upper tail of the limit distribution.[16]

Our work expands this research horizon by focusing on the interplay of non-insurable income shocks, credit market imperfections and investment indivisibilities in shaping bequest incentives of individuals at the bottom of the distribution. A main contribution of our analysis is to show that a thick right tail generated by some specific features of the wealth accumulation process is not theoretically—and hence empirically—inconsistent with the presence of extreme forms of compression of economic traits (wealth and income) of lineages at the bottom of the wealth distribution. This we deem important: in our model poor households are more similar in terms of wealth and income than other groups at different ladders in terms of socio-economic conditions and also expectations about the future. If this compression of the variability of economic traits is cast in a richer model than the one analyzed in the present paper, it may easily come to be associated to a compression in other, more general, socio-economic aspects (such as beliefs, norms of behavior, aspirations). In this case the

[14] In Loury (1981), who provided the first elegant and rigorous formalization for the emergence of the limit distribution of earnings, financial wealth plays no role at all in the determination of the limit distribution of the socio-economic status of families.

[15] See Krusell and Smith (2006) for a review of canonical income fluctuation problems. Recent studies have offered novel insights into the economic predictions of Bewley economies by generalizing the basic model setup along important dimensions. To cite a few, Stachurski and Toda (2019) formally show that Bewley economies that abstract from capital income risk necessarily generate the same tail behavior for income and wealth distributions; Ma et al. (2020) explore the key properties of a generalized income fluctuation problem where returns on assets, labor earnings and impatience exhibit state-dependence, serial as well as mutual correlation properties.

[16] This is an instance of Champernowne (1953)'s seminal result, that random growth processes with multiplicative shocks endowed with a reflecting lower barrier produce a power law at steady state. Other micro-foundations able to generate Pareto-tailed wealth and/or income distributions are studied by e.g. Nirei and Aoki (2016), where it is business productivity shocks in a world with safe and risky investment technologies that push concentration of income at the top; and Toda (2019), who exploit random discount factors to generate heterogeneity in saving rates.

structural elements that determine the compression of economic traits at the bottom of the wealth distribution can more easily lead to the poverty traps due to neighborhood effects as in Benabou (1996) or Durlauf (1996), reinforcing the persistence of lineages in the left tail.[17]

Relying on a non-interactive framework of wealth distribution, our analysis disregards the fact that heterogeneity in rates of return to wealth and skill premia are endogenous to portfolio and occupational choices of individuals. In principle, abstracting from general equilibrium considerations may undermine the robustness of the economic implications of our results in terms of e.g. the relative strength of the forces that ultimately govern the concentration of wealth in the right tail and the mobility frictions in the left tail of the stationary distribution; or of the importance of feedback effects from policies (such as income taxation) to relative wages of skilled and unskilled workers. For instance, by lowering the volatility of returns on financial wealth, capital income taxation provides partial insurance against idiosyncratic risk and thus curbs the demand for (precautionary) saving, in turn affecting asset returns via general equilibrium effects.[18]

We conjecture that the bulk of our characterizations can be extended to model environments incorporating general equilibrium features. Let us consider, for instance, the results in Proposition 1, where two possible regimes for the long-run wealth distribution are characterized according to the value taken up by the wealth threshold $\tilde{\omega}_0$. In our partial equilibrium setting that term is a fixed parameter of the model, depending on the exogenous skill premium, on the expected rate of return granted by the financial assets and on other preference parameters. Suppose, instead, that the skill premium is endogenous as where the wage in the unskilled sector is determined by educational choices by households (whereas the skilled sector is modeled, for simplicity, as a standard AK technology). In that case both $\tilde{\omega}_0$ and $\underline{y}$ (the wage in the unskilled sector) will change over time. Then, to the extent that such a model preserves ergodicity, as per a convex production function in the unskilled sector, the sequences $\{\tilde{\omega}_0\}_t$ and $\left\{\underline{y}\right\}_t$ must both settle down to their steady state level. Associated to this level, there will then exist a value of $\tilde{\omega}_0$ and Proposition 1 applies. In other words, and in a precise sense, if the general equilibrium model converges to a unique steady state level of the skill premium, then the limit distribution must exhibit the features we characterize in the partial equilibrium setting. Our results can therefore be exploited as a useful and intuitive guide for investigating the properties of the admissible regimes in the limit distribution of wealth in general equilibrium models.

[17] Exploiting dynamic models in which altruistic parents may transfer resources to their offspring by providing education and by leaving bequests, other studies have addressed different issues than the ones we are interested in, such as e.g. the implications of differential tax treatments for the optimal mix of financial bequests and human capital investment (Blinder 1976), the efficiency enhancing effects of bequest taxes in the presence of wage taxation (Grossmann and Poutvaara 2009), and the characterization of market outcomes under laissez-faire vs. the social planner's allocation (Dávila 2023).

[18] General equilibrium models of the evolution of wealth distributions have been put forth, among others, by Dutta and Michel (1998), Castaneda et al. (2003), Nirei and Aoki (2016), Cao and Luo (2017). In particular, Nirei and Aoki (2016) show that labor earnings risk plays a non-negligible role in determining the Pareto exponent of the equilibrium distribution of wealth, whereas Cao and Luo (2017) establish that top wealth inequality responds (among other things) to salient aggregate statistics like the economy-wide growth rate and the labor share.

More generally, we believe that the analysis of the implications of endogenous factor prices would be a fruitful avenue for future research, for it may clarify whether labor market dynamics has the potential to establish a connection between the right and the left tail of the wealth distribution, and thus shed further light on the relative contribution of the many mechanisms driving the dynamics of inequality.

7 Concluding remarks

When credit market imperfections constrain educational investment opportunities, individual abilities and wealth background may exert a significant influence on the formation of human capital and the ensuing patterns of social mobility within and across family lineages, e.g. Becker and Tomes (1986). The present paper has identified simple conditions on the actual nature and intensity of intergenerational altruism under which households in the lowest rungs of the wealth ladder are more likely to arrange their bequest strategies according to the mechanisms emphasized in Loury (1981), along with those highlighted in Benhabib et al. (2011).

Optimal bequest strategies in our model are indeed shown to entail two key features: first, they embody heterogeneous saving rates across the wealth distribution, a prediction that has been empirically documented in recent work, see e.g. Dynan et al. (2004) and Fagereng et al. (2019); and second, due to the interplay of wealth constraints and investment indivisibilities, bequests incentives may quickly dissipate moving down to the bottom of the cross-sectional distribution, breaking the chain of positive private transfers within family lineages. This implies that the dynamics of the aggregate saving rate is directly influenced, among other things, by the transitions in and out of the atom, motivating us to explore the deep causes that reproduce, in a given economy, the emergence of a group of non-savers in the population of households.

On the policy front, linking the shape of the left tail to the structure of social mobility via educational investment allows us to start thinking, in the simplest possible setting, about the ability of public policies to create substitutes for missing insurance and credit markets. Our analysis has in fact established that ergodicity (mobility across and within lineages) of the wealth dynamics does not necessarily entail the vanishing of credit market imperfections in the growth process of wealth accumulation due to parental altruism, implying that heterogeneity in individuals' wealth backgrounds and educational opportunities contribute to shaping the long-term structure of inequality. The characterization of the size of the mass of non-savers at the bottom and the initial exploration of its determinants is, in principle, relevant to inform government intervention in the education sector, aimed at curtailing poverty persistence (Mookherjee and Napel 2021), and also for the design of social insurance and social security institutions (Diamond and Geanakoplos 2003; Diamond 2004).

Appendix

Proof of Lemma 1

For fixed $e = \{0, 1\}$, the KKT first-order conditions (12) and (13) are sufficient for existence of a global maximum, being the objective function (twice-differentiable and) strictly concave, and the inequality constraints continuously differentiable convex functions. Existence, uniqueness and differentiability (hence continuity) of pseudo policy functions b_e^* and c_e^* are therefore obtained.

Since the cross partial derivative of $u(c_e)$ with respect to b_e and ω_t is equal to minus the second-order derivative of $u(c_e)$ with respect to b_e, applying the implicit function theorem to (12) delivers

$$\frac{d\,b_e^*}{d\,\omega_t} = -\frac{\frac{d^2\,u(c_e^*)}{d\,b_e d\,\omega_t}}{\frac{d^2\,u(c_e^*)}{d\,b_e^2} + \chi\mathbb{E}\left[(1+r)^2 \frac{d^2\,v(\omega_{t+1}(b_e^*))}{d\,b_e^2}\right]} \in (0, 1) \tag{A1}$$

implying c_e^* is non-decreasing in wealth as well.
(i) In order to establish convexity, differentiating (12) twice with respect to ω_t shows that $d^2 b_e^*/d\omega_t^2 \geq 0$ if and only if

$$\frac{u'''(c_e^*)}{u''(c_e^*)} \cdot \left(1 - \frac{d\,b_e^*}{d\,\omega_t}\right) \geq \frac{\mathbb{E}\left[(1+r)^3 v'''(\omega_{t+1}(b_e^*)) \cdot \frac{d\,b_e^*}{d\,\omega_t}\right]}{\mathbb{E}\left[(1+r)^2 v''(\omega_{t+1}(b_e^*))\right]} \tag{A2}$$

which, given (A1) and (12), is equivalent to

$$\frac{u'''(c_e^*) \cdot u'(c_e^*)}{\left[u''(c_e^*)\right]^2} \leq \frac{\mathbb{E}\left[(1+r)^3 v'''(\omega_{t+1}(b_e^*))\right] \cdot \mathbb{E}\left[(1+r)v'(\omega_{t+1}(b_e^*))\right]}{\left[\mathbb{E}\left[(1+r)^2 v''(\omega_{t+1}(b_e^*))\right]\right]^2} \tag{A3}$$

showing that, locally, the convexity of the bequest policy function depends on the comparative local curvature properties of the functions u and v. By Assumption 4 (CRRA utility), the left-hand side of the weak inequality (A3) is constant and equal to $(1+\gamma)\gamma^{-1}$; moreover, the CRRA specification implies that v' and v''' are strictly convex in their argument, while v'' is strictly concave. Define the mapping

$$k'(\omega_{t+1}) : \omega_{t+1} \to \chi(1+r)v'(\omega_{t+1})$$

and apply Jensen's inequality to obtain

$$\begin{aligned}
\mathbb{E}\left[k'(\omega_{t+1})\right] &\geq k'\left(\mathbb{E}[\omega_{t+1}]\right)\\
\mathbb{E}\left[k''(\omega_{t+1})\right] &\leq k''\left(\mathbb{E}[\omega_{t+1}]\right)\\
\mathbb{E}\left[k'''(\omega_{t+1})\right] &\geq k'''\left(\mathbb{E}[\omega_{t+1}]\right)
\end{aligned}$$

and thus

$$\frac{\mathbb{E}\left[(1+r)^3 v'''(\omega_{t+1}(b_e^*))\right]\cdot\mathbb{E}\left[(1+r)v'(\omega_{t+1}(b_e^*))\right]}{\left[\mathbb{E}\left[(1+r)^2 v''(\omega_{t+1}(b_e^*))\right]\right]^2} \geq \frac{k'''\left(\mathbb{E}[\omega_{t+1}]\right)\cdot k'\left(\mathbb{E}[\omega_{t+1}]\right)}{\left[k''\left(\mathbb{E}[\omega_{t+1}]\right)\right]^2} = \frac{1+\gamma}{\gamma}$$

showing that (A3) is fulfilled.
(ii) The proof that, for each $e \in \{0,1\}$, one has $\lim_{\omega_t\to\infty} b_e^*/\omega_t = \phi$, with $\phi \in (0,1)$, exploits an argument in Zhu (2019), that we reproduce here for the sake of completeness. We first show that the ratio b_e^*/ω_t is non-decreasing in ω_t, and bounded from above by ϕ. Since the following results hold for the CRRA specification, in this proof we exploit its explicit functional form.

Consider first the case $e = 0$, and note that, by monotonicity of b_0^*, $b_0^*/\omega_t = 0$ if and only if $\omega_t \leq \underline{y}\,[\chi(1+\mathbb{E}[R])]^{-\frac{1}{\gamma}}$—see (12) and (13); otherwise, we have $b_0^*/\omega_t > 0$ and thus $c_0^*/\omega_t = 1 - \frac{b_0^*}{\omega_t} < 1$. Consider now arbitrary wealth levels $\hat{\omega}_t$ and $\hat{\hat{\omega}}_t$ satisfying

$$\hat{\hat{\omega}}_t > \hat{\omega}_t > \underline{y}\,[\chi(1+\mathbb{E}[R])]^{-\frac{1}{\gamma}}$$

and let $\hat{b}_0^*$ and $\hat{\hat{b}}_0^*$ be the respective images under the pseudo policy b_0^*. The first-order condition (12) can be equivalently written as

$$1 = \chi\mathbb{E}\left[(1+r)\left((1+r)\frac{b_0^*}{\omega_t - b_0^*} + \frac{\underline{y}}{\omega_t - b_0^*}\right)^{-\gamma}\right] \tag{A4}$$

from which we obtain

$$1 = \chi\mathbb{E}\left[(1+r)\left((1+r)\left(\frac{\hat{\omega}_t}{\hat{c}_0^*} - 1\right) + \frac{\underline{y}}{\hat{c}_0^*}\right)^{-\gamma}\right] \tag{A5}$$

and

$$1 = \chi\mathbb{E}\left[(1+r)\left((1+r)\left(\frac{\hat{\hat{\omega}}_t}{\hat{\hat{c}}_0^*} - 1\right) + \frac{\underline{y}}{\hat{\hat{c}}_0^*}\right)^{-\gamma}\right] \tag{A6}$$

where $\hat{c}_0^* = \hat{\omega}_t - \hat{b}_0^*$ and $\hat{\hat{c}}_0^* = \hat{\hat{\omega}}_t - \hat{\hat{b}}_0^*$, with $\hat{\hat{c}}_0^* \geq \hat{c}_0^*$—see (A1).

Suppose $\frac{\hat{\hat{\omega}}_t}{\hat{\hat{c}}_t^*} < \frac{\hat{\omega}_t}{\hat{c}_0^*}$. Then from (A5) it must hold

$$\begin{aligned}1 &= \chi \mathbb{E}\left[(1+r)\left((1+r)\left(\frac{\hat{\omega}_t}{\hat{c}_0^*} - 1\right) + \frac{\underline{y}}{\hat{c}_0^*}\right)^{-\gamma}\right] \\ &< \chi \mathbb{E}\left[(1+r)\left((1+r)\left(\frac{\hat{\hat{\omega}}_t}{\hat{\hat{c}}_0^*} - 1\right) + \frac{\underline{y}}{\hat{\hat{c}}_0^*}\right)^{-\gamma}\right]\end{aligned}$$

i.e. a contradiction, see equation (A6). Thus is must be

$$\frac{\hat{\hat{\omega}}_t}{\hat{\hat{c}}_0^*} \geq \frac{\hat{\omega}_t}{\hat{c}_0^*} \quad \Leftrightarrow \quad \frac{\hat{\hat{c}}_0^*}{\hat{\hat{\omega}}_t} \leq \frac{\hat{c}_0^*}{\hat{\omega}_t} \tag{A7}$$

which in turn implies that b_0^*/ω_t is non-decreasing in ω_t. To prove that b_0^*/ω_t is also bounded from above, we first notice that $b_0^*/\omega_t = 0$ when $\omega_t \leq \underline{y}\,[\chi(1+\mathbb{E}[R])]^{-\frac{1}{\gamma}}$, whereas for $\omega_t > \underline{y}\,[\chi(1+\mathbb{E}[R])]^{-\frac{1}{\gamma}}$ the first-order condition (A4) implies

$$\begin{aligned}1 &= \chi \mathbb{E}\left[(1+r)\left((1+r)\frac{b_0^*}{\omega_t - b_0^*} + \frac{\underline{y}}{\omega_t - b_0^*}\right)^{-\gamma}\right] \\ &\leq \chi \mathbb{E}\left[(1+r)\left((1+r)\frac{b_0^*}{\omega_t - b_0^*}\right)^{-\gamma}\right]\end{aligned}$$

from which we have

$$\frac{b_0^*}{\omega_t} \leq \frac{\left(\chi \mathbb{E}\left[(1+r)^{1-\gamma}\right]\right)^{\frac{1}{\gamma}}}{1 + \left(\chi \mathbb{E}\left[(1+r)^{1-\gamma}\right]\right)^{\frac{1}{\gamma}}} = \phi \in (0,1) \tag{A8}$$

We finally establish that $\lim_{\omega_t \to \infty} b_0^*/\omega_t = \phi$. From (A4) we have

$$1 = \chi \mathbb{E}\left[(1+r)\left((1+r)\left(\frac{1}{\frac{c_0^*}{\omega_t}} - 1\right) + \frac{\underline{y}}{c_0^*}\right)^{-\gamma}\right]$$

with $\lim_{\omega_t \to \infty} c_0^* = \infty$ by virtue of (A1) and $\lim_{\omega_t \to \infty} c_0^*/\omega_t = \lambda \in [1-\phi, 1)$ from (A7) and (A8). Define $\iota = \left(\frac{1}{\lambda+\epsilon} - 1\right)^{-\gamma} (1+r)^{1-\gamma}$ for some arbitrarily small $\epsilon > 0$ such that $\lambda + \epsilon < 1$, and note that $\mathbb{E}[\iota] < \infty$ and that

$$0 < (1+r)\left(\frac{\omega_{t+1}}{c_0^*}\right)^{-\gamma} = \left[(1+r)\left((1+r)\left(\frac{1}{\frac{c_0^*}{\omega_t}} - 1\right) + \frac{\underline{y}}{c_0^*}\right)^{-\gamma}\right] \leq \iota$$

for ω_t large enough. From (A4), exploiting Lebesgue's Dominated Convergence Theorem we obtain

$$\begin{aligned} 1 &= \lim_{\omega_t \to \infty} \left\{ \chi \mathbb{E} \left[(1+r) \left((1+r) \left(\frac{1}{\frac{c_0^*}{\omega_t}} - 1 \right) + \frac{y}{c_0^*} \right)^{-\gamma} \right] \right\} \\ &= \chi \mathbb{E} \lim_{\omega_t \to \infty} \left\{ \left[(1+r) \left((1+r) \left(\frac{1}{\frac{c_0^*}{\omega_t}} - 1 \right) + \frac{y}{c_0^*} \right)^{-\gamma} \right] \right\} \\ &= \chi \left(\frac{1}{\lambda} - 1 \right)^{-\gamma} \mathbb{E} \left[(1+r)^{1-\gamma} \right] \end{aligned} \tag{A9}$$

which delivers

$$\lambda = \frac{1}{1 + \left(\chi \mathbb{E} \left[(1+r)^{1-\gamma} \right] \right)^{\frac{1}{\gamma}}}$$

and hence

$$\lim_{\omega_t \to \infty} \frac{b_0^*}{\omega_t} = 1 - \lim_{\omega_t \to \infty} \frac{c_0^*}{\omega_t} = 1 - \lambda = \phi$$

The proof for the case $e = 1$—with $b_1^* = 0$ for all $\omega_t \leq x + [\chi\,(1 + \mathbb{E}[R])]^{-\frac{1}{\gamma}}$ $\left(\mathbb{E}\left[y^{-\gamma}\right]\right)^{-\frac{1}{\gamma}}$ by monotonicity and (12), $b_1^* > 0$ satisfying (12) and $c_1^* = \omega_t - b_1^* - x$, $x \in \Delta x$ given—follows exactly the same route and is thus omitted.

Proof of Lemma 2

Consider the indifference condition (14). Fix $\omega_t > 0$. The left-hand side does not depend on x (since $e^* = 0$), whereas the right-hand side continuously decreases in $x \in \Delta x$ by the envelope theorem. At $x = 0$, $e^* = 1$ and the right-hand side is strictly larger than the left-hand side. Then, for given $\omega_t > 0$, there exists a unique $\tilde{x}(\omega_t) > 0$ such that $e^* = 1$ if and only if $x < \tilde{x}(\omega_t) < \omega_t$ (recall that $x > \omega_t$ implies $e^* = 0$ under the no-borrowing constraint).

The indifference condition (14) defines a continuously differentiable function $F(\omega_t, \tilde{x}) = 0$. By virtue of the implicit function theorem, $\tilde{x}(\omega_t)$ is a well-defined continuous and (continuously) differentiable function satisfying

$$\frac{d\,\tilde{x}}{d\,\omega_t} = -\frac{F_{\omega_t}}{F_{\tilde{x}}} = \frac{\frac{d\,u(c_1^*)}{d\,b_e} - \frac{d\,u(c_0^*)}{d\,b_e}}{\frac{d\,u(c_1^*)}{d\,b_e}} < 1 \tag{A10}$$

and

$$\text{sign}\ \frac{d\,\tilde{x}}{d\,\omega_t} = -\text{sign}\ F_{\omega_t} \tag{A11}$$

Using the first-order condition (12) we have

$$F_{\omega_t} = \left(\omega_t - b_0^*\right)^{-\gamma} - \left(\omega_t - b_1^*(\tilde{x}) - \tilde{x}\right)^{-\gamma} \tag{A12}$$

and thus $\tilde{x}(\omega_t)$ is increasing in its domain if and only if $b_1^*(\tilde{x}) + \tilde{x} > b_0^*$ for all ω_t. Suppose not, i.e. assume $b_1^*(\tilde{x}) + \tilde{x} \leq b_0^*$ for some ω_t. Then from the indifference condition (14) it must hold at $\tilde{x}$

$$\mathbb{E}\left[v\left((1+r)b_1^*(\tilde{x}) + y\right)\right] \leq \mathbb{E}\left[v\left((1+r)b_0^* + \underline{y}\right)\right] \tag{A13}$$

while the first-order condition (12) and the concavity of v jointly imply

$$\mathbb{E}\left[v'\left((1+r)b_1^*(\tilde{x}) + y\right)\cdot(1+r)\right] \leq \mathbb{E}\left[v'\left((1+r)b_0^* + \underline{y}\right)\cdot(1+r)\right] \tag{A14}$$

i.e. a contradiction (which is obtained, *a fortiori*, when the borrowing constraint bites and (13) is the relevant optimality condition). Thus, the threshold $\tilde{x}(\omega_t)$ must increase monotonically in wealth.

Proof of Lemma 3

(i) Assume there exists some $\omega_t > 0$ such that, at $x \in \Delta x$, it holds $b_1^*(x) > b_0^* > 0$. Then by the first-order condition (12)

$$\chi\mathbb{E}\left[v'(\omega_{t+1}(b_0^*))\right] = u'(\omega_t - b_0^*) < u'(\omega_t - b_1^* - x) = \chi\mathbb{E}\left[v'(\omega_{t+1}(b_1^*))\right]$$

Since by strict concavity of v one has

$$\mathbb{E}\left[v'\left((1+r)b_1^* + \underline{y}\right)(1+r)\right] \geq \mathbb{E}\left[v'\left((1+r)b_1^* + y\right)(1+r)\right]$$

we have a contradiction. Hence, at any given $\omega_t > 0$ at which $b_e^* > 0$ for each $e \in \{0, 1\}$, it must be the case that $b_1^*(x) \leq b_0^*$ for all $x \in \Delta x$.

(ii) Fix ω_t. For $x \geq \tilde{x}(\omega_t)$, $e^* = 0$ and thus $b^*(x) = b_0^*$ does not vary with the educational cost. For $x < \tilde{x}(\omega_t)$, the pseudo financial bequest b_1^* (hence $b^*(x)$) decreases monotonically in the educational cost $x \in [0, \tilde{x})$, for we have

$$\frac{\partial b_1^*}{\partial x} = -\frac{\frac{d^2\,u(c_1^*)}{d\,b_1^2}}{\frac{d^2\,u(c_1^*)}{d\,b_1^2} + \chi\,\frac{d^2\,\mathbb{E}v(\omega_{t+1}(b_1^*))}{d\,b_1^2}} \in (-1, 0)$$

From this and the fact that $b^*(0) \leq b_0^*$, it follows that $b^*(x)$ exhibits an upward jump discontinuity at $\tilde{x}(\omega_t)$.

Proof of Lemma 4

Equation (16) and Lemmata 2 to 3 imply that b_0^* and b_1^* are both zero if and only if $\omega_t \in [\underline{y}, \tilde{\omega}_0]$. From (16) one has

$$\tilde{\omega}_0 = \underline{y}\left[\chi\left(1 + \mathbb{E}[R]\right)\right]^{-\frac{1}{\gamma}}$$

and thus

$$\tilde{\omega}_0 > \underline{y} \quad \Longleftrightarrow \quad \chi(1 + \mathbb{E}[R]) < 1$$

We have

$$\tilde{\omega}_0 = \underline{y}\left[\chi\left(1 + \mathbb{E}[R]\right)\right]^{-\frac{1}{\gamma}} < \left[\chi\left(1 + \mathbb{E}[R]\right)\right]^{-\frac{1}{\gamma}}\left(\mathbb{E}\left[y^{-\gamma}\right]\right)^{-\frac{1}{\gamma}} + x = \tilde{\omega}_1(x)$$

for all $x \in \Delta x$. Clearly, $\tilde{\omega}_0$ decreases with χ and γ. This completes the proof.

Proof of Proposition 1

We adapt a proof in Zhu (2019) for a similar—albeit not identical—setup, which in turn exploits Theorem 15.0.1, part (iii) in Meyn and Tweedie (2009). Formally, Meyn and Tweedie (2009)'s argument relies on establishing three key properties for the Markov chain $\{\omega_t\}$ generated by (17): (1) irreducibility; (2) aperiodicity, and (3) geometric drift.

To facilitate sailing the technical details, in the following we consider two distinct cases.

a. $\chi\left(1 + \mathbb{E}[R]\right) > \left(\underline{y}/\overline{y}\right)^{\gamma}$ (or equivalently $\tilde{\omega}_0 < \overline{y}$).
1. Irreducibility. We first recall that (c_e^*, b_e^*) are continuous functions in ω_t, and that $\omega_{t+1} \geq \underline{y}$ for all $\omega_t > 0$ and $x \in \Delta x$. Notice also that $Pr\left(\omega_t < \tilde{\omega}_0 \,\middle|\, \omega_1 > \underline{y}\right) > 0$ at some time $t \geq 1$ since

$$(1 + \min(\Delta r))b_e^*(\omega_t) + \underline{y} + e^*\left(\min(\Delta y) - \underline{y}\right) = \underline{y} < \tilde{\omega}_0$$

for all $\omega_t > \underline{y}$ and $x \in \Delta x$. Since $b_e^* = 0$ for $\omega_t \leq \tilde{\omega}_0$ for each $e \in \{0, 1\}$, and yet $Pr\left(X < \tilde{x}(\omega_t) \mid \omega_t < \tilde{\omega}_0\right) = G(\tilde{x}(\omega_t)) > 0$ by virtue of Assumption 3 and Lemma 2, implying $e^* = 1$ and $\omega_s = y_s > \underline{y}$ at some $s > t$, any set A such that

$$\int_A f(z)dz > 0 \tag{A15}$$

can be reached in finite time with a positive probability. Letting $\varphi(A) = \int_A f(z)dz$ define a measure on $[\underline{y}, \infty)$, the process $\{\omega_t\}$ is therefore φ-irreducible, and thus ψ-irreducible for some other measure ψ on $[\underline{y}, \infty)$, which necessarily exists—see Proposition 4.2.2 of Meyn and Tweedie (2009).

2. Aperiodicity. Consider the set $C = [\underline{y}, \tilde{\omega}_0]$. For $\omega_t \in C$, $b_e^*(\omega_t) = 0$ and $e^* = 1$ for all $x < \tilde{x}(\omega_t)$, which occurs with probability $G(\tilde{x}(\omega_t)) > 0$. As a result, one has $\int_C f(z)dz > 0$; this allows constructing a non-trivial measure $v_1(C) := \int_C f(z)dz$ on the Borel σ-field of W, denoted with $\mathcal{B}(W)$, satisfying

$$v_1(B) \leq P(\omega, B), \quad \forall \omega \in C, \;\; B \in \mathcal{B}(W) \tag{A16}$$

where $P(\omega, B)$ is the one-step transition probability kernel. Hence, C is a so-called v_1-small set. Since $\{\omega_t\}$ is φ-irreducible, existence of a small set C and of a positive measure $v_1(C) > 0$ jointly imply that $\{\omega_t\}$ is strongly aperiodic—see Meyn and Tweedie (2009), p. 114.

Geometric drift. This condition ensures that the Markov process is stable in the sense of exhibiting inward drift to some small (typically compact) subset of the wealth space W, i.e. for some measurable function $V \geq 1$, finite at some point $\omega^\circ \in W$, the drift of $V(\omega_t)$ defined as

$$\Delta V(\omega) = \int P(\omega, dz)V(z) - V(\omega)$$

satisfies

$$\Delta V(\omega) \leq -\beta V(\omega) + b\mathbb{1}_C(\omega), \quad \omega \in W$$

where C is a *petite set*, $\beta > 0$ and $b < \infty$ are constants, and $\mathbb{1}_C$ is the characteristic function associated to C (i.e. $\mathbb{1}_C = 1$ if and only if $\omega \in C$). By Proposition 5.5.3 in Meyn and Tweedie (2009), a v_m-small set for some $m \geq 1$ is also petite (for some well-defined sampling distribution).[19]

Fix an arbitrary $\underline{\omega} > \underline{y}$, and consider the compact set $C = [\underline{y}, \underline{\omega}]$. Since

$$(1 + \min(\Delta r))b_e^*(\omega_t) + \underline{y} + e^* \left(\min(\Delta y) - \underline{y}\right) = \underline{y} < \tilde{\omega}_0$$

for all $\omega_t > \underline{y}$, there exists a common (across lineages) m, however large, such that $Pr(\omega_m < \tilde{\omega}_0 \mid \omega_1) \geq \epsilon > 0$. Since $b_e^* = 0$ for each $e = \{0, 1\}$ if and only if $\omega_t \in [\underline{y}, \tilde{\omega}_0]$, we have

$$\begin{aligned} Pr\,(\omega_{m+1} \in C \mid \omega_1) &\geq Pr(\omega_{m+1} \in C \mid \omega_m < \underline{y}) \times Pr(\omega_m < \underline{y} \mid \omega_1) \\ &\geq \epsilon \int_C f(z)dz > 0 \end{aligned}$$

[19] From (A16), by analogy, a v_m-small set C satisfies $v_m(B) \leq P^m(\omega, B), \quad \forall \omega \in C, \;\; B \in \mathcal{B}(W)$, where $P^m(\omega, B)$ is the m-step transition kernel.

which implies that C is υ_{m+1}-small and therefore petite.
Let us now observe that, for all ω_t, the pseudo financial bequest policy b_e^* for $e = \{0, 1\}$ satisfies the first-order conditions (12)-(13). Thus either $b_e^* = 0$ or $b_e^* > 0$ solve

$$\left(\omega_t - b_e^* - e \cdot x\right)^{-\gamma} = \chi \mathbb{E}\left[(1+r)\left((1+r)b_e^* + \underline{y} + e \cdot (y - \underline{y})\right)^{-\gamma}\right]$$

or equivalently

$$\left(c_e^*\right)^{-\gamma} = \chi \mathbb{E}\left[(1+r)\left((1+r)(\omega_t - c_e^* - e \cdot x) + \underline{y} + e \cdot (y - \underline{y})\right)^{-\gamma}\right]$$

For $e^* = 0$, by Lemma 3 part (ii), we have $c_0^* \geq (1-\phi) \cdot \omega_t$—and thus $b_0^* \leq \phi \cdot \omega_t$—where

$$1 - \phi := \frac{1}{1 + \left(\chi \mathbb{E}\left[(1+r)^{1-\gamma}\right]\right)^{\frac{1}{\gamma}}} \in (0, 1)$$

Similarly, when $e^* = 1$, one obtains $c_1^* \geq (1-\phi) \cdot (\omega_t - x)$, and $b_1^* \leq \phi \cdot (\omega_t - x)$. Notice that $\chi(1 + \mathbb{E}[R]) < 1$ (as we assume) implies $\chi < 1$. Hence, from the wealth accumulation equation (17) we can write

$$\omega_{t+1} = \begin{cases} (1+r)\, b_0^* + \underline{y} \leq \chi\phi(1+r)\omega_t + \underline{y}, & \omega \leq \hat{\omega} \\ (1+r)\, b_1^* + y \leq \chi\phi(1+r)\omega_t + y, & \omega > \hat{\omega} \end{cases} \tag{A17}$$

where apparently $\mathbb{E}\left[\chi\phi(1+r)\right] < 1$.
Define

$$\begin{aligned} &V(\omega_t) = \omega_t + 1, \quad \omega_t \in W; \\ &\beta = 1 - \chi\phi(1 + \mathbb{E}[R]) - q, \quad q \in (0, 1 - \chi\phi(1 + \mathbb{E}[R])); \\ &b = \beta + q + \mathbb{E}[Y], \end{aligned}$$

where $E[Y]$ is the unconditional average of the random process for labour income $\{y_t\}$. Pick any finite $\underline{\omega}$ such that $b \leq q \cdot (\underline{\omega} + 1)$; then $C = [\underline{y}, \underline{\omega}]$ is a petite set, and (A17) implies

$$\begin{aligned} &\mathbb{E}_t\left[V(\omega_{t+1}) - V(\omega_t)\right] \\ &= \mathbb{E}_t[\omega_{t+1}] - \omega_t \\ &\leq \mathbb{E}_t[\chi(1-\phi)(1+r)]\omega_t + \mathbb{E}[Y] - \omega_t \\ &= -\omega_t + (1-1) + \mathbb{E}[\chi(1-\phi)(1+r)](\omega_t + 1) - \mathbb{E}[\chi\phi(1+r)] + \mathbb{E}[Y] \\ &= -V(\omega_t) + \mathbb{E}[\chi\phi(1+r)]V(\omega_t) + 1 - \mathbb{E}[\chi\phi(1+r)] + \mathbb{E}[Y] \\ &= -\left[1 - \mathbb{E}[\chi\phi(1+r)]\right] V(\omega_t) + (1 - \mathbb{E}[\chi\phi(1+r) + \mathbb{E}[Y]]) \\ &\leq -\beta V(\omega_t) + b\mathbb{1}_C(\omega_t), \quad \omega_t \in W \end{aligned} \tag{A18}$$

showing that the drift condition is satisfied.
The support of the stationary distribution is $S_\infty = [\underline{y}, \infty)$ since

$$Pr\left(\omega_{t+1} > \omega_t \mid \omega_t\right) > 0, \quad \forall \omega_t > \tilde{\omega}_0$$

by Assumption 1, and

$$Pr\left(\omega_{t+1} > \omega_t \mid \omega_t\right) > 0, \quad \forall \omega_t \in \left[\underline{y}, \tilde{\omega}_0\right]$$

since $\tilde{\omega}_0 < \overline{y}$.

b. $\chi\left(1 + \mathbb{E}[R]\right) \leq \left(\underline{y}/\overline{y}\right)^\gamma$ (or equivalently $\tilde{\omega}_0 \geq \overline{y}$).
The assertion follows immediately from noticing that, for each lineage i, it holds

$$Pr(\omega_{i,t+n} < \tilde{\omega}_0 \mid \omega_{i,t}) > 0, \quad \forall \omega_{i,t} \leq \underline{y}, \;\; 1 \leq n < \infty$$

and

$$Pr\left(\omega_{i,t+1} > \overline{y} \mid \omega_{i,t}\right) = 0, \quad \forall \omega_{i,t} \in \left[\underline{y}, \tilde{\omega}_0\right]$$

Hence, in finite time, the individual wealth transitions (17) for each lineage i will reduce to $\omega_t = \bar{y} + e^*(y_t - \bar{y})$. Since the educational investment costs and labour income follow mutually independent, ergodic processes (Assumptions 2 and 3), the wealth accumulation process has a unique limiting (and thus stationary) distribution with bounded support $S_\infty = \Delta y$.

Proof of Proposition 2

Recall that $\chi(1 + \mathbb{E}[R]) \leq \left(\underline{y}/\overline{y}\right)^\gamma$ is equivalent to $\tilde{\omega}_0 \geq \bar{y}$; by Proposition 1 the stationary distribution of wealth would have bounded support and all of its moments would be finite, implying a thin right tail.

We thus focus on the case $\chi(1 + \mathbb{E}[R]) < 1$ (i.e. $\tilde{\omega}_0 > \underline{y}$), and assume that the wage in the skilled occupation $\overline{y}$ is large enough to satisfy

$$\overline{y} > \left[\chi\left(1 + \mathbb{E}[R]\right)\right]^{-\frac{1}{\gamma}} \left(\mathbb{E}\left[y^{-\gamma}\right]\right)^{-\frac{1}{\gamma}} + \bar{x} \tag{A19}$$

so that $\overline{y} > \tilde{\omega}_1(x) > \tilde{\omega}_0$ for all $x \in \Delta x$—see Lemma 4. Notice that, since $\chi(1 + \mathbb{E}[R]) < 1$, this assumption requires that the probability of very large earnings in the skilled occupation be sufficiently small.

The proof consists of three steps: (i) Construct a piece-wise linear function $b^a : \omega_t \in W \mapsto [0, \infty)$ that satisfies $b^a \leq b_e^*$ for each ω_t and $\lim_{\omega_t \to \infty} b^a/\omega_t \leq \lim_{\omega_t \to \infty} b_e^*/\omega_t$. (ii) Construct an auxiliary wealth accumulation process $\left\{\omega_t^a\right\}$ that can be shown to converge to a unique stationary distribution ω_∞^a with a Pareto right tail. (iii) Exploit an argument grounded in the theory of stochastic dominance to show that the right tail of the stationary distribution ω_∞ cannot decay at a lower rate than that of ω^a_∞.

(i) Consider the following function

$$b^a = \begin{cases} 0, & \text{for } \omega_t \le \tilde{\omega}_1(\bar{x}) \\ \underbrace{\left(\frac{1}{1+\chi^{-\frac{1}{\gamma}}(1+\mathbb{E}[R])^{1-\frac{1}{\gamma}}}\right)}_{\eta} \cdot (\omega_t - \tilde{\omega}_1(\bar{x})), & \text{for } \omega_t > \tilde{\omega}_1(\bar{x}) \end{cases} \tag{A20}$$

where $\eta \in (0, 1)$. Notice that, for $\omega_t > \tilde{\omega}_1(\bar{x})$, b^a is the unique solution to the equation

$$\left(\omega_t - b^a - \bar{x}\right)^{-\gamma} = \chi(1+\mathbb{E}[R])\left[(1+\mathbb{E}[R])\,b^a + y\right]^{-\gamma} \tag{A21}$$

which can be confronted with (12) to show that, at any wealth level $\omega_t \ge \underline{y}$, Jensen's inequality imparts $b^a \le b^*_1$ at $\bar{x}$, and thus $b^a \le b^*_e$ for each $e \in \{0, 1\}$ and for all $x \in \Delta x$—see Lemma 3. Furthermore one has

$$\lim_{\omega_t \to \infty} \frac{b^a}{\omega_t} = \eta \le \phi = \lim_{\omega_t \to \infty} \frac{b^*_e}{\omega_t} \tag{A22}$$

if and only if $\mathbb{E}\left[(1+R)^{1-\gamma}\right] \ge (1+\mathbb{E}[R])^{1-\gamma}$ which is warranted by Assumption 4. It is straightforward to see that the piece-wise linear function b^a is Lipschitz continuous.

(ii) Let the auxiliary wealth accumulation process $\left\{\omega^a_t\right\}$ evolve according to

$$\begin{aligned} \omega^a_{t+1} &= (1+r)b^a + y \\ &= \begin{cases} y, & \text{for } \omega^a_t \le \tilde{\omega}_1(\bar{x}) \\ (1+r)\eta(\omega^a_t - \tilde{\omega}_1(\bar{x})) + y, & \text{for } \omega^a_t > \tilde{\omega}_1(\bar{x}) \end{cases} \end{aligned} \tag{A23}$$

Notice that, by virtue of (A19), one has $F(y \ge \tilde{\omega}_1(\bar{x})) > 0$, which is required for the auxiliary process (A23) not to be permanently trapped in wealth states where $b^a = 0$.

We follow Mirek (2011) to establish that the auxiliary process $\left\{\omega^a_t\right\}$ admits a unique stationary distribution with an asymptotic Pareto (right) tail. To this end, we next verify Assumption 1.6 (*Shape of the mappings*) and Assumption 1.7 (*Moment conditions*) under which Theorem 1.8 in Mirek (2011) holds.

Assumption 1.6 Recall that, by virtue of (A19) and the ensuing features of the function b^a, the history of random variables X is irrelevant for the auxiliary wealth accumulation process (A23). Let $\theta = (R, Y)$ and $\psi_\theta(\omega^a) = (1+r)b^a + y$, so that the process $\{\omega_t^a\}$ is produced according to the stochastic recursion $\omega_{t+1}^a = \psi_\theta(\omega_t^a)$, which is Lipschitz continuous. For each $z > 0$ define

$$\psi_{\theta,z}(\omega^a) = z \cdot \psi_\theta(z^{-1}\omega^a), \qquad \bar{\psi}_\theta(\omega^a) = \lim_{z\to 0} \psi_{\theta,z}(\omega^a)$$

and notice that

$$\bar{\psi}_\theta(\omega^a) = \lim_{z\to 0}\left[z \cdot \psi_\theta(z^{-1}\omega^a)\right] = (1+r)\eta\omega_t^a, \quad \forall\omega_t^a \geq \underline{y} \tag{A24}$$

which in turn implies

$$|\,\psi_\theta(\omega^a) - \bar{\psi}_\theta(\omega^a)\,| = \begin{cases} |\,y - (1+r)\eta\omega_t^a\,|, & \text{for} \quad \omega_t^a \leq \tilde{\omega}_1(\bar{x}) \\ \\ |\,y - (1+r)\eta\tilde{\omega}_1(\bar{x})\,|, & \text{for} \quad \omega_t^a > \tilde{\omega}_1(\bar{x}) \end{cases}$$

Upon defining the positive random variable $N_\theta := \kappa + y$ for any fixed constant κ satisfying $\kappa \geq (1+\bar{r})\eta\tilde{\omega}_1(\bar{x})$, we notice that N_θ has bounded support and that

$$\left|\psi_\theta(\omega^a) - \bar{\psi}_\theta(\omega^a)\right| < N_\theta, \quad \forall\omega_t^a \geq \underline{y} \tag{A25}$$

Hence, conditions (H1) and (H2) of Assumption 1.6 in Mirek (2011) are both satisfied.

Assumption 1.7 We check conditions (H3) to (H7) of Assumption 1.7 in Mirek (2011):

(H3) This condition is satisfied since R is an i.i.d. (thus serially uncorrelated) random variable with bounded support.

(H4) This condition is satisfied since R has a strictly positive density h on $[-1, \bar{r}]$, implying that the law of $log(1+R)$ conditioned on $\{(1+R) > 0\}$ is non-arithmetic.

(H5) It is easy to see that $\chi(1+\mathbb{E}[R]) < 1$ implies $\mathbb{E}[(1+R)\eta] < 1$. Assume now that, for given average return $\mathbb{E}[R]$, the second moment of the return distribution $\mathbb{E}\left[R^2\right]$ is large enough to satisfy $\mathbb{E}[(1+R)^2] > 1/\eta^2$. Notice this latter restriction implies $H((1+R)\eta) > 1) > 0$, i.e. positive expansion for the auxiliary process $\{\omega_t^a\}$ (Kesten 1973). We thus have

$$\log\mathbb{E}[(1+R)\eta] < 0, \quad \log\mathbb{E}[(1+R)^2\eta^2] > 0$$

Since $\log\mathbb{E}[((1+R)\eta)^\nu]$ is a convex and continuous function of $\nu > 0$, there exists a unique $\nu \in (1,2)$ such that $\log\mathbb{E}[((1+R)\eta)^\nu] = 0$ and thus $\mathbb{E}[((1+R)\eta)^\nu] = 1$, and the condition is satisfied.

(H6) This condition is satisfied since R has bounded support and thus

$$\eta^\nu \cdot \mathbb{E}\left[(1+R)^\nu \cdot |\log((1+R)\eta\,|\right] < \infty$$

(H7) This condition is satisfied since N_θ is supported in a compact set and thus $\mathbb{E}[(N_\theta)^\nu] < \infty$.

According to Theorem 1.8 in Mirek (2011), a unique stationary solution ω_∞^a to (A23) will exist and it will feature an asymptotic Pareto tail, i.e.

$$Pr(\omega_\infty^a > \omega) \sim C\omega^{-\nu}, \quad \omega \to \infty$$

for some constant $C > 0$.

(iii) To conclude the proof, consider now the wealth accumulation process $\{\omega_t\}$ generated by (17) and the auxiliary process $\left\{\omega_t^a\right\}$ evolving according to (A23). Pick $\omega_{t=0} = \omega_{t=0}^a$ (by ergodicity, initial distributions can be arbitrarily fixed). For any admissible path $\theta = (R, Y)$—i.e. for any history of shocks to human and non-human wealth– that is identical across the two processes, we have $\omega_t \geq \omega_t^a$ for all $t \geq 1$ since $b^a \leq b_e^*$ for all $\omega_t \geq \underline{y}$. It therefore holds

$$Pr\left(w_t \geq \omega\right) \geq Pr\left(w_t^a \geq \omega\right), \quad \forall \omega > \underline{y}, \quad \forall t \geq 1$$

which in turn, by ergodicity, implies $Pr\left(w_\infty \geq \omega\right) \geq Pr\left(w_\infty^a \geq \omega\right)$, and hence

$$\liminf_{\omega \to \infty} \ \omega^\nu \cdot Pr\left(\omega_\infty > \omega\right) \geq \liminf_{\omega \to \infty} \ \omega^\nu \cdot Pr\left(\omega_\infty^a > \omega\right) = C$$

whence the assertion.

Proof of Proposition 3

Since $\tilde{x}(\underline{y})$ is larger then zero, it holds $0 < G(\tilde{x}(\underline{y}))$. We also know that

$$G(\tilde{x}(\underline{y}))\mu^*(\underline{y}) = \int_{\omega_t \in (\underline{y}, \tilde{\omega}_0]} (1 - G(\tilde{x}(\omega_t)))d\mu^*(\omega) \tag{A26}$$

since $Pr\left(\omega_{t+1} = \underline{y} \mid \omega_t > \tilde{\omega}_0\right) = 0$.

The invariant measure μ^* is such that $0 < \mu^*([\underline{y}, \tilde{\omega}_0])$, given that $\tilde{\omega}_0 > \underline{y}$. Since $1 - G(\tilde{x}(\omega_t)) > 0$, the right-hand side of (A26) is larger than zero: there exists a strictly positive probability for a generic lineage i with wealth in $(\underline{y}, \tilde{\omega}_0]$ at time t to display a wealth level equal to $\underline{y}$ in the next period $t + 1$, i.e. $\mu^*(\underline{y}) > 0$.

To characterize the size of the atom, consider the following sets of events that collect lineages i with the lowest level of wealth:

$$N_1 = \{i \mid \omega_i = \underline{y}, \ e^* = 0, \ b_0^* = 0\}, \tag{A27}$$

$$N_2 = \{i \mid \omega_i = \underline{y}, \ e^* = 0, \ 1 + r = 0\}, \tag{A28}$$

$$N_3 = \{i \mid \omega_i = \underline{y}, \ e^* = 1, \ b_1^* = 0, \ y = \underline{y}\}, \tag{A29}$$

$$N_4 = \{i \mid \omega_i = \underline{y}, \ e^* = 1, \ 1 + r = 0, \ y = \underline{y}\}. \tag{A30}$$

Since the sets N_j, $j = 2, 3, 4$ have zero measure, it follows that the mass point at the bottom of the stationary wealth distribution has measure $\mu^* (\{N_1\})$.

Proof of Proposition 4

Since $\underline{y} \leq \omega_t$ for all t, one has $b_0^* > 0$ for all levels of parental wealth. Using the same argument as in the proof of Proposition 3 shows that N_1 is a zero measure set, whence the assertion.

Proof of Proposition 5

From the optimality condition (12) it is easily seen that an infinitesimal rise in the intensity of the bequest motive, from χ to $\chi' > \chi$, increases, all else equal, the marginal utility from financial bequests relative to consumption, at any cost $x \in \Delta x$. Therefore, $\partial b_e^*/\partial \chi > 0$, $e \in \{0, 1\}$ whenever $b_e^* > 0$. Also, from (14), one has $\partial \tilde{x}/\partial \chi > 0$ for all $\omega_t \geq \underline{y}$ if and only if $b_0^* < b_1^*(\tilde{x}) + \tilde{x}$, which is always the case—see Lemma 2. Finally, at any cost $x \in \Delta x$, the wealth thresholds $\tilde{\omega}_e(x)$ decrease with χ; hence, the support $[\underline{y}, \tilde{\omega}_0]$ of wealth states at which $b_e^* = 0$ shrinks.
Consider now the individual wealth transition (17) for lineage i, letting the superscript $'$ label optimal bequest choices made by parents in lineage i when the intensity of altruism is χ':

- for lineage i with educational cost $x_i \in \Delta x$ and wealth $\omega_{i,t} \leq \tilde{\omega}_0$, pseudo financial bequests are zero for any educational choice and

$$G\left(x_i \leq \tilde{x}(\omega_{i,t}) \,|\, \omega_{i,t} \leq \tilde{\omega}_0\right) < G\left(x_i \leq \tilde{x}'(\omega_{i,t}) \,|\, \omega_{i,t} \leq \tilde{\omega}_0\right);$$

- for lineage i with educational cost $x_i \in \Delta x$ and wealth $\omega_{i,t} > \tilde{\omega}_1(x_i)$, for whom pseudo financial bequests are strictly positive for any educational choice, relatively larger altruism produces larger financial transfers, hence

$$Pr\left(\omega_{i,t+1}(b_e^{*'} > 0) \geq \underline{y} \,|\, \omega_{i,t} > \tilde{\omega}_1(x_i)\right) > Pr\left(\omega_{i,t+1}(b_e^* > 0) \geq \underline{y} \,|\, \omega_{i,t} > \tilde{\omega}_1(x_i)\right),$$

  for $e \in \{0, 1\}$;
- for lineage i with educational cost $x_i \in (\tilde{x}(\omega_{i,t}), \tilde{x}'(\omega_{i,t}]$ and wealth $\omega_{i.t} \in (\tilde{\omega}_0, \tilde{\omega}_1(x_i)]$, relatively larger altruism entails a switch from the optimal bequest plan $(b_{i.0}^* > 0, e_i^* = 0)$ to the optimal bequest plan $(b_{i,1}^{*'} = 0, e_i^{*'} = 1)$, with the property that $\tilde{x}'(\omega_{i,t}) > b_{i,0}^*$—see Lemma 1. From the indifference condition (14), for any such i at time t, it must hold

$$\mathbb{E}\left[v(y)\right] > \mathbb{E}\left[v((1+r)b_{i,0}^*)\right] \tag{A31}$$

where, as usual, the expectation is taken with respect to the distributions of R and Y. Since the latter inequality holds for any increasing and concave function v, it implies

$$\mathbb{E}[\omega_{i,t+1}(b^{*'}_{i,1} = 0, e^{*'}_{i} = 1)] \geq \mathbb{E}[\omega_{i,t+1}(b^{*}_{i,0} > 0, e^{*}_{i} = 0)]$$

by the properties of the increasing concave order (Shaked and Shanthikumar 2007). That is, the average wealth (averaging over all possible realizations of random returns) of the infra-marginal households i who substitute financial bequests with human capital investment under stronger intensity χ' of altruism cannot be lower than its counterpart under χ.

Integrating over i, it follows that, at any time period $t > 1$, the cross-sectional distribution of wealth under stronger intensity χ' (i) second-order stochastically dominates its counterpart under χ, and (ii) if the two ever cross over the support $[0, \infty)$, then at the first crossing point the former must cross the latter from below to the right of $\tilde{\omega}_0$. By ergodicity and Theorem 4.A.8(c) of Shaked and Shanthikumar (2007), these properties also characterize the invariant measures $\mu^*_{\chi'}$ and μ^*_{χ}, implying that

$$\int_{\omega \in (\underline{y}, \tilde{\omega}_0]} [1 - G(\tilde{x}(\omega))] d\mu^*_{\chi} > \int_{\omega \in (\underline{y}, \tilde{\omega}'_0]} [1 - G(\tilde{x}'(\omega))] d\mu^*_{\chi'} \qquad \text{(A32)}$$

which, coupled with $G(X \leq \tilde{x}(\underline{y})) < G(X \leq \tilde{x}'(\underline{y}))$, delivers the assertion.

Proof of Proposition 6

Notice first that, at any cost $x \in \Delta x$, the optimality conditions (12)-(13) as well as the wealth thresholds $\tilde{\omega}_e(x)$ are not conditioned on G, meaning that the pseudo financial bequest policies b^*_e (for $e \in \{0, 1\}$) do no vary with the distribution of educational costs in population. Also, the investment threshold cost $\tilde{x}(\omega_t)$ solving (14) is invariant with respect to G, implying that the threshold rule for optimal educational investment (as expressed in Lemma 2) does not vary either. As a consequence, at each time period $t > 1$ the cross-sectional distribution of wealth entails a smaller measure of parents investing in human capital formation under G' than under G, all else equal, since $G'(X \leq \tilde{x}(\omega_t)) \leq G(X \leq \tilde{x}(\omega_t))$ for all $\omega_t \geq \underline{y}$. By ergodicity, it follows that the invariant measure $\mu^*_{G'}$ of the stationary wealth distribution under G' first-order stochastically dominates its counterpart μ^*_G under G; coupled with $G'(X \leq \tilde{x}(\underline{y})) \leq G(X \leq \tilde{x}(\underline{y}))$, and the fact that the support $[\underline{y}, \tilde{\omega}_0]$ entailing zero financial bequests for each education choice is unchanged, the assertion follows immediately from (20).

Proof of Lemma 5

We first notice that, all else equal, for all $\tau_j \in (0, 1)$ $j = b, r$ we have

$$\tilde{\omega}_0' = \chi^{-\frac{1}{\gamma}} \left[\left(1 - \tau_j\right)(1 + \mathbb{E}[R])\right]^{-\frac{1}{\gamma}} \underline{y} > \tilde{\omega}_0 \tag{A33}$$

and

$$\tilde{\omega}_1'(x) = \chi^{-\frac{1}{\gamma}} \left[\left(1 - \tau_j\right)(1 + \mathbb{E}[R])\right]^{-\frac{1}{\gamma}} \left(\mathbb{E}\left[y^{-\gamma}\right]\right)^{-\frac{1}{\gamma}} + x > \tilde{\omega}_1(x), \quad x \in \Delta x \tag{A34}$$

i.e. the introduction of estate/capital income taxes increases the wealth thresholds below which pseudo optimal financial bequests are zero.

(i) Fix e to either 0 or 1. From (A33) and (A34) it follows that $b_e^{*'} = 0$ whenever $b_e^* = 0$ (and the assertion follows trivially), so that $b_e^{*'}$ can only be positive if b_e^* are. We thus focus on establishing the assertion when pseudo optimal financial bequests are strictly positive with and without taxes. Suppose, to the contrary, that $(1 - \tau_r) b_e^{*'} > b_e^*$. Then, by the relevant first-order conditions

$$\left(\omega_t - b_e^{*'} - e \cdot x\right)^{-\gamma} = \chi \mathbb{E}\left[(1 - \tau_j)(1 + r)\left(\omega_{t+1}(b_e^{*'})\right)^{-\gamma}\right]$$
$$\left(\omega_t - b_e^{*} - e \cdot x\right)^{-\gamma} = \chi \mathbb{E}\left[(1 + r)\left(\omega_{t+1}(b_e^{*})\right)^{-\gamma})\right]$$

and exploiting the concavity of v, we obtain

$$\mathbb{E}\left[(1 - \tau_j)(1 + r)\left(\omega_{t+1}(b_e^{*'})\right)^{-\gamma}\right] < \mathbb{E}\left[(1 + r)\left(\omega_{t+1}(b_e^{*})\right)^{-\gamma})\right]$$

for $e \in \{0, 1\}$. Since $\tau_j < 1$, $(1 - \tau_r) b_e^{*'} > b_e^*$ implies $b_e^{*'} > b_e^*$ and thus $c_e^{*'} < c_e^*$, from which

$$\mathbb{E}\left[(1 - \tau_j)(1 + r)\left(\omega_{t+1}(b_e^{*'})\right)^{-\gamma}\right] > \mathbb{E}\left[(1 + r)\left(\omega_{t+1}(b_e^{*})\right)^{-\gamma})\right]$$

i.e. a contradiction.

(ii) Assume $\bar{x} > 0$ is sufficiently large. Then, for some tax rates $\tau_j \in (0, 1)$, there exists a positive measure set $\Delta' \subseteq \Delta x$ such that

$$\tilde{\omega}_0' < \tilde{\omega}_1(x), \quad x \in \Delta' \tag{A35}$$

Fix now any such $x \in \Delta'$ and $t_j \in (0, 1)$. For $\omega_t \in B = \left(\tilde{\omega}_0', \ \tilde{\omega}_1(x)\right]$ we have $b_0^* > 0$, $b_0^{*'} > 0$, $b_1^* = b_1^{*'} = 0$, and the educational threshold $\tilde{x}'(\omega_t)$ cost under estate/capital income taxation solves

$$u(\omega_t - b_0^{*'}) + \chi \mathbb{E}\left[v\left((1 - \tau_j)(1 + r) b_0^{*'} + \underline{y}\right)\right] = u(\omega_t - \tilde{x}') + \chi \mathbb{E}\left[v\left(y\right)\right] \tag{A36}$$

Applying the implicit function theorem in the neighborhood of $(\tilde{x}', \tau_j)$ fulfilling (A36) shows that $d\tilde{x}/d\tau_j$ has the same sign as the following expression

$$\mathbb{E}\left((1+r)\left[\left((1+r)b_0^{*'} + \underline{y}\right)^{-\gamma} \cdot b_0^{*'}\right]\right)$$

which is clearly positive. Since $\lim_{\tau_j \to 0^+} b_0^{*'} = b_0^*$ and $\lim_{\tau_j \to 0^+} d\tilde{x}/d\tau_j > 0$ it follows that $\tilde{x}'(\omega_t) > \tilde{x}(\omega_t)$ for all $\omega_t \in B$.

Proof of Proposition 7

Let $\mu*'(\underline{y})$ denote the measure of the atom in the stationary wealth distribution in the presence of estate/capital income taxation. We next identify two simple sets of conditions pertaining to ex-ante heterogeneity in agents' characteristics and the expected returns on human capital versus financial investment under which the introduction of a tax τ_b on bequests or a tax τ_r on financial returns increases the size of the atom in the left tail of the stationary distribution (i.e. $\mu*'(\underline{y}) > \mu^*(\underline{y})$, case (i)) or rather lowers it (i.e. $\mu*'(\underline{y}) < \mu^*(\underline{y})$, case (ii)).
(i) Define $\hat{x}$, if it exists, as the solution to the following equation

$$(\tilde{\omega}_0)^{1-\gamma} + \chi \underline{y}^{1-\gamma} = \left(\tilde{\omega}_0 - \hat{x}\right)^{1-\gamma} + \chi \mathbb{E}\left[y^{1-\gamma}\right] \tag{A37}$$

where, we recall, $\tilde{\omega}_0 = \underline{y}\,[\chi\,(1+\mathbb{E}[R])]^{-\frac{1}{\gamma}}$. Notice that $\hat{x} = \tilde{x}(\tilde{\omega}_0)$: if there is no $\hat{x} \in \Delta x$ solving (A37), then all the parents with wealth $\tilde{\omega}_0$ will undertake educational investment, irrespective of the actual cost they face. Notice also that $\tilde{x}(\omega_t) > \hat{x}$ for all $\omega_t > \tilde{\omega}_0$ by virtue of Lemma 2; and that $\tilde{x}'(\omega_t) = \tilde{x}(\omega_t)$ for all $\omega_t \leq \tilde{\omega}_0$, with $\tilde{x}'(\omega_t)$ increasing in ω_t. It follows that, if $\bar{x} \leq \hat{x}$, at each time period t all the parents with wealth $\omega_t > \tilde{\omega}_0$ (and *a fortiori* those with wealth $\omega_t > \tilde{\omega}_0'$) will have $e^* = 1$ whatever their educational cost $x \in \Delta x$.

Assume now $\bar{x} \leq \hat{x}$. It follows that the introduction of a tax τ_j (on bequests or realized financial returns) does not modify the educational investment choices of households (i.e. $e^{*'} = e^* = 1$) for all wealth levels, while implying a contraction in the pseudo post-tax financial bequests $(1-\tau_j)b_e^{*'}$ for all wealth levels. For any given initial distribution of wealth and for any tax rate $\tau_j \in (0, 1)$, $j = b, r$ we therefore have

$$\begin{aligned} Pr&\left((1-\tau_j)(1+r)b_e^{*'} + \underline{y} + e^{*'}(y - \underline{y}) \leq \bar{\omega}\right) \\ &\geq Pr\left((1+r)b_e^* + \underline{y} + e^*(y - \underline{y}) \leq \bar{\omega}\right) \end{aligned}$$

for all $\bar{\omega} > \underline{y}$ and time $t \geq 1$. By ergodicity we obtain $Pr(\omega_\infty' \leq \bar{\omega}) \geq Pr(\omega_\infty \leq \bar{\omega})$, where ω_∞' denotes the stationary distribution to which the wealth accumulation process converges in the presence of estate/capital income taxation. Notice that $G(\tilde{x}(\underline{y})) = G(\tilde{x}'(\underline{y}))$ and $G(\tilde{x}(\omega_t)) = G(\tilde{x}(\omega_t))$ for all $\omega_t \leq \tilde{\omega}_0'$. Then it must be the case that if

the maximal educational cost is sufficiently small (i.e. if $\bar{x} \leq \hat{x}$) then $\mu^{*'}(\underline{y}) > \mu^{*}(\underline{y})$ for all $\tau_j \in (0, 1)$, $j = b, r$.
(ii) With little loss of generality, let us assume that the wage in the high-skill occupation is deterministic and equal to $\overline{y} > \underline{y}$. Let also the economy's fundamentals be such that $\tilde{\omega}_0 < \overline{y}$, i.e. let

$$\underline{y}\chi^{-\frac{1}{\gamma}}(1 + \mathbb{E}[R])^{-\frac{1}{\gamma}} < \overline{y}$$

Define $\check{x}$ as the solution, if it exists, to the following equation

$$\overline{y}^{1-\gamma} + \chi \underline{y}^{1-\gamma} = \left(\overline{y} - \check{x}\right)^{1-\gamma} + \chi \overline{y}^{1-\gamma} \tag{A38}$$

Notice that if there is no $\check{x} \in \Delta x$ solving (A38), then all individuals with wealth $\overline{y}$ who optimally choose not to leave financial bequests will undertake educational investment, irrespective of the actual cost they face.
Notice also that $\hat{x} < \check{x}$, provided both thresholds exist. Assume then $\hat{x} < \bar{x} \leq \check{x}$, and let the actual tax rate τ_j $(j = b, r)$ implemented in the economy satisfy

$$\tau_j \geq \bar{\tau} = 1 - \left(\frac{\underline{y}}{\overline{y}}\right)^{\gamma} \cdot \frac{1}{\chi(1 + \mathbb{E}[R])} \iff \tilde{\omega}_0' \geq \overline{y}$$

where the lower bound $\bar{\tau} \in (0, 1)$ always exists for any finite $\overline{y}$ as $\lim_{\tau_j \to 0^+} \tilde{\omega}_0' = \tilde{\omega}_0 < \overline{y}$ and $\lim_{\tau_j \to 1^-} \tilde{\omega}_0' = \infty$.

Under these circumstances, with probability one, any lineage will find herself in the lower states of the wealth space ($\omega_t \in [\underline{y}, \tilde{\omega}_0']$) entailing $b_e^* = 0$ for $e = \{0, 1\}$, and will not manage to exit those states by means of human capital formation (since $\overline{y} \leq \tilde{\omega}_0'$)—see Proposition 1. Since educational investment remains affordable by all households in the high skill occupation (as $\bar{x} \leq \check{x}$), the economy will converge to a one-point stationary distribution localized at $\omega_t = \overline{y}$, meaning no mass point at $\omega_t = \underline{y}$, i.e. $\mu^{*'}(\underline{y}) = 0$.

Consider now tax rates $\tau_j < \bar{\tau}$ (including the no tax scenario). In this case we have (see (A34))

$$\underline{y} < \tilde{\omega}_0' < \overline{y} < \tilde{\omega}_1'(\bar{x})$$

and hence $b_0^{*'}(\overline{y}) > 0$ and $b_1^{*'}(\overline{y}) = 0$. As a result we obtain

$$\begin{aligned}
(\overline{y} - \tilde{x}'(\overline{y}))^{1-\gamma} &+ \chi \overline{y}^{1-\gamma} \\
&= (\overline{y} - b_0^{*'})^{1-\gamma} + \chi \mathbb{E}\left[\left((1 + r)b_0^{*'} + \underline{y}\right)^{1-\gamma}\right] \\
&> \overline{y}^{1-\gamma} + \chi \underline{y}^{1-\gamma} \\
&= (\overline{y} - \check{x})^{1-\gamma} + \chi \overline{y}^{1-\gamma}
\end{aligned} \tag{A39}$$

where the first equality follows from the definition of the investment cost threshold $\tilde{x}'(\omega_t)$, the inequality by the optimality conditions, and the last equality from the definition of $\check{x}$.

From (A39) we immediately obtain $\tilde{x}'(\overline{y}) < \check{x}$. As a main consequence, when $\tau_j < \bar{\tau}$ and provided $\bar{x} > \tilde{x}'(\overline{y})$, there exists a positive measure subset of households—those with wealth $\omega_t = \overline{y}$ and facing investment costs $x \in [\tilde{x}'(\overline{y}), \bar{x})$—who optimally choose to refrain from educational investment while substituting it with financial bequests, whose rate of return is lower than $\overline{y}$ with positive probability. By Proposition 1, the wealth dynamics will converge to a fully-fledged distribution with limit support $S_\infty = [\underline{y}, \infty)$ and a mass point at $\omega_t = \underline{y}$: a reduction of the tax rate increases wealth inequality and involves the emergence of an atom in the left tail of the stationary distribution of wealth. This is true also for $\tau_j = 0$, $j = b, r$, the benchmark economy, and thus we obtain $\mu^{*'}(\underline{y}) < \mu^*(\underline{y})$.

Acknowledgements We are grateful to an Associate Editor and two anonymous Reviewers for several comments and suggestions, which have helped us to greatly improve the manuscript. We also thank Achille Basile, Chetan Dave, Paula Gobbi (discussant), Tullio Jappelli, Dilip Mookherjee, Andy Newman, Jacques Silber, and Emma Tominey for insightful discussions on previous versions of the manuscript. We wish to thank seminar/conference participants (CFE 2018, IASI-CNR 2019, SAET Conference 2015 & 2019, University of Trento 2019, LOrde Workshop 2022, CefES Conference 2022, RCEA Conference 2022, RES Annual Conference 2022, PET Conference 2022, EWET 2022, SITES Conference 2022, CISEPS Annual Workshop 2023, SEHO 2023, ASSET 2023, University of Napoli Parthenope 2024, University of Alberta 2024, Inequality in Rome Seminar Series - Roma Tre University) for comments. Special gratitude we wish to express to Alberto Bisin for stimulating the questions addressed in this paper. Any remaining errors are our own.

Funding Open access funding provided by Università degli Studi di Salerno within the CRUI-CARE Agreement.

Declarations

Conflict of interest The authors have no Conflict of interest directly relevant to the content of this manuscript.

References

Aiyagari, S.R.: Uninsured idiosyncratic risk and aggregate saving. Q. J. Econ. **109**(3), 659–684 (1994)

Andreoni, J.: Impure altruism and donations to public goods: a theory of warm-glow giving. Econ. J. **100**(401), 464–477 (1990)

Balestra, C., Tonkin, R.: Inequalities in household wealth across OECD countries: evidence from the OECD wealth distribution database (2018)

Bastani, S., Waldenström, D.: How should capital be taxed? J. Econ. Surv. **34**(4), 812–846 (2020)

Becker, G.S., Tomes, N.: An equilibrium theory of the distribution of income and intergenerational mobility. J. Polit. Econ. **87**(6), 1153–1189 (1979)
Becker, G.S., Tomes, N.: Human capital and the rise and fall of families. J. Law Econ. **4**(3, Part 2), S1–S39 (1986)
Benabou, R.: Heterogeneity, stratification, and growth: macroeconomic implications of community structure and school finance. Am. Econ. Rev. **86**(3), 584–609 (1996)
Benhabib, J., Bisin, A., Zhu, S.: The distribution of wealth and fiscal policy in economies with finitely lived agents. Econometrica **79**(1), 123–157 (2011)
Benhabib, J., Bisin, A., Zhu, S.: The wealth distribution in Bewley economies with capital income risk. J. Econ. Theory **159**, 489–515 (2015)
Benhabib, J., Bisin, A., Luo, M.: Wealth distribution and social mobility in the US: a quantitative approach. Am. Econ. Rev. **109**(5), 1623–1647 (2019)
Bewley, T.: The permanent income hypothesis: a theoretical formulation. J. Econ. Theory **16**(2), 252–292 (1977)
Bewley, T.: A difficulty with the optimum quantity of money. Econometrica J. Econ. Soc. 1485–1504 (1983)
Birkner, M., Scheuer, N., Wälde, K.: The dynamics of Pareto distributed wealth in a small open economy. Econ. Theor. **76**(2), 607–644 (2023)
Blinder, A.S.: Intergenerational transfers and life cycle consumption. Am. Econ. Rev. **66**(2), 87–93 (1976)
Cao, D., Luo, W.: Persistent heterogeneous returns and top end wealth inequality. Rev. Econ. Dyn. **26**, 301–326 (2017)
Castaneda, A., Diaz-Gimenez, J., Rios-Rull, J.-V.: Accounting for the us earnings and wealth inequality. J. Polit. Econ. **111**(4), 818–857 (2003)
Champernowne, D.G.: A model of income distribution. Econ. J. **63**(250), 318–351 (1953)
Charles, K.K., Hurst, E.: The correlation of wealth across generations. J. Polit. Econ. **111**(6), 1155–1182 (2003)
Chetty, R., Hendren, N., Kline, P., Saez, E.: Where is the land of opportunity? The geography of intergenerational mobility in the United States. Q. J. Econ. **129**(4), 1553–1623 (2014)
D'Amato, M., Di Pietro, C.: Occupational mobility and wealth evolution in a model of educational investment with credit market imperfections. J. Econ. Inequal. **12**, 73–98 (2014)
Davies, J.B., Shorrocks, A.F.: The distribution of wealth. Handb. Income Distrib. **1**, 605–675 (2000)
Dávila, J.: Bequests or education. Econ. Theor. **75**(4), 1039–1069 (2023)
De Nardi, M.: Wealth inequality and intergenerational links. Rev. Econ. Stud. **71**(3), 743–768 (2004)
De Nardi, M., Fella, G.: Saving and wealth inequality. Rev. Econ. Dyn. **26**, 280–300 (2017)
Diamond, P.: Social security. Am. Econ. Rev. **94**(1), 1–24 (2004)
Diamond, P., Geanakoplos, J.: Social security investment in equities. Am. Econ. Rev. **93**(4), 1047–1074 (2003)
Durlauf, S.N.: A theory of persistent income inequality. J. Econ. Growth **1**, 75–93 (1996)
Dutta, J., Michel, P.: The distribution of wealth with imperfect altruism. J. Econ. Theory **82**(2), 379–404 (1998)
Dynan, K.E., Skinner, J., Zeldes, S.P.: Do the rich save more? J. Polit. Econ. **112**(2), 397–444 (2004)
Fagereng, A., Holm, M.B., Moll, B., Natvik, G.: Saving behavior across the wealth distribution: the importance of capital gains. Technical report, National Bureau of Economic Research (2019)
Flavin, M., Yamashita, T.: Owner-occupied housing and the composition of the household portfolio. Am. Econ. Rev. **92**(1), 345–362 (2002)
Galor, O., Zeira, J.: Income distribution and macroeconomics. Rev. Econ. Stud. **60**(1), 35–52 (1993)
Gollier, C.: The Economics of Risk and Time. MIT Press, Cambridge (2001)
Grossmann, V., Poutvaara, P.: Pareto-improving bequest taxation. Int. Tax Public Finance **16**, 647–669 (2009)
Heckman, J.J., Stixrud, J., Urzua, S.: The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. J. Law Econ. **24**(3), 411–482 (2006)
Henderson, D.J., Polachek, S.W., Wang, L.: Heterogeneity in schooling rates of return. Econ. Educ. Rev. **30**(6), 1202–1214 (2011)
Jappelli, T., Pistaferri, L.: The Economics of Consumption: Theory and Evidence. Oxford University Press, Oxford (2017)
Kesten, H.: Random difference equations and renewal theory for products of random matrices. Acta Math. **131**(1), 207–248 (1973)
Kimball, M.: Precautionary saving in the small and in the large. Econometrica **58**(1), 53–73 (1990)

Klass, O.S., Biham, O., Levy, M., Malcai, O., Solomon, S.: The Forbes 400, the pareto power-law and efficient markets. Eur. Phys. J. B **55**(147), 143 (2007)
Kopczuk, W.: Incentive effects of inheritances and optimal estate taxation. Am. Econ. Rev. **103**(3), 472–477 (2013)
Krusell, P., Smith, A.A.: Quantitative macroeconomic models with heterogeneous agents. Econom. Soc. Monogr. **41**, 298 (2006)
Loury, G.C.: Intergenerational transfers and the distribution of earnings. Econometrica (1981)
Ma, Q., Stachurski, J., Toda, A.A.: The income fluctuation problem and the evolution of wealth. J. Econ. Theory **187**, 105003 (2020)
Marshall, A.: The principles of economics. Technical Report, McMaster University Archive for the History of Economic Thought (1890)
Meyn, S., Tweedie, R.L.: Markov Chains and Stochastic Stability. Cambridge University Press, Cambridge (2009)
Mirek, M.: Heavy tail phenomenon and convergence to stable laws for iterated Lipschitz maps. Probab. Theory Relat. Fields **151**(3), 705–734 (2011)
Mookherjee, D., Napel, S.: Intergenerational mobility and macroeconomic history dependence. J. Econ. Theory **137**(1), 49–78 (2007)
Mookherjee, D., Napel, S.: Welfare rationales for conditionality of cash transfers. J. Dev. Econ. **151**, 102657 (2021)
Mookherjee, D., Ray, D.: Persistent inequality. Rev. Econ. Stud. **70**(2), 369–393 (2003)
Moskowitz, T.J., Vissing-Jørgensen, A.: The returns to entrepreneurial investment: A private equity premium puzzle? Am. Econ. Rev. **92**(4), 745–778 (2002)
Nirei, M., Aoki, S.: Pareto distribution of income in neoclassical growth models. Rev. Econ. Dyn. **20**, 25–42 (2016)
Pareto, V.: Les systèmes socialistes, volume 2. V. Giard & e. Briere (1903)
Pestieau, P., Thibault, E.: Love thy children or money: reflections on debt neutrality and estate taxation. Econ. Theor. **50**, 31–57 (2012)
Piketty, T.: Capital in the Twenty-First Century. Harvard University Press, Harvard (2014)
Ray, D.: On the dynamics of inequality. Econ. Theor. **29**(2), 291–306 (2006)
Ray, D., Genicot, G.: Measuring upward mobility. Am. Econ. Rev. **113**(11), 3044–3089 (2023)
Rothschild, M., Stiglitz, J.E.: Increasing risk II: its economic consequences. J. Econ. Theory **3**(1), 66–84 (1971)
Shaked, M., Shanthikumar, J.G.: Stochastic Orders. Springer, Berlin (2007)
Stachurski, J., Toda, A.A.: An impossibility theorem for wealth in heterogeneous-agent models with limited heterogeneity. J. Econ. Theory **182**, 1–24 (2019)
Toda, A.A.: Wealth distribution with random discount factors. J. Monet. Econ. **104**, 101–113 (2019)
Vermeulen, P.: How fat is the top tail of the wealth distribution? Rev. Income Wealth **64**(2), 357–387 (2018)
Wan, J., Zhu, S.: Bequests, estate taxes, and wealth distributions. Econ. Theor. **67**, 179–210 (2019)
Wolff, E.N.: Estimates of household wealth inequality in the US, 1962–1983. Rev. Income Wealth **33**(3), 231–256 (1987)
Wolff, E.N.: 4. Changes in household wealth in the 1980s and 1990s in the united states. International Perspectives on Household Wealth, p. 107 (2006)
Zhu, S.: A Becker–Tomes model with investment risk. Econ. Theor. **67**(4), 951–981 (2019)