Inherited Access: Early Generation-Skipping Transfers and the Persistence of Top Wealth*

Donghoon Koo†

MEA-SHARE gGmbH

August, 2026

Abstract

The U.S. taxes bequests with two instruments, an estate and a generation-skipping transfer (GST) tax. I build a life-cycle model to study how these transfers sustain top wealth, and how taxes shape that persistence. The model has two distinctive features: high returns require a financial capability that a large early inheritance switches on, and dynasties may transfer directly to grandchildren, a margin the GST tax regulates. The link raises the probability a wealthy grandchild reaches the top 1 percent to 62 percent, barely moving the Gini. The exemption governs persistence, the rate governs concentration, and every relaxation lowers newborn welfare.

JEL Codes: D15, D31, D64, E21, H24, H31

Keywords: Wealth inequality, intergenerational transfers, estate and generation-skipping taxation, return heterogeneity, dynastic persistence, life-cycle model.

*An earlier version of this paper circulated under the title “Estate Taxation, Inherited Wealth and Rising Wealth Inequality.” and was part of Ph.D. Dissertation “Essays on the Macroeconomic Implications of Inheritances” at the University of Melbourne. I thank Shuyun (May) Li, Chris Edmond, Lawrence Uren, Daeha Cho, Mei Dong, Faisal Sohail and Jonathan Payne for their valuable comments. I also thank all participants at Ph.D. Student Brownbag Seminar Series, Macro Seminar Series and Department Brownbag Seminar Series at the University of Melbourne for their questions, comments, and feedback. I am grateful to three anonymous referees for their valuable comments and feedback.

†Munich Research Institute for the Economics of Aging and SHARE Analyses (MEA-SHARE gGmbH), Leopoldstraße 139, München 80804, Germany. E-mail: d.koo@mea-share.eu

Why is wealth inequality so persistent across generations? How does tax policy shape both the concentration and the persistence of wealth? Wealth in the United States is not only increasingly concentrated but also persistent: the same families tend to stay at the top over long horizons (Adermon, Lindahl and Waldenström, 2018; Clark, 2014; Pfeffer and Killewald, 2018). Beyond the long-established dynasties, each wave of technological and financial expansion has also seen new entrepreneurs emerge with an extraordinary stock of wealth. While many of these fortunes are self-made, the wealth that is not spent down will eventually be passed on, whatever its origin; today's self-made rich thus become the founders of tomorrow's dynasties. Over the same period, federal estate taxation has been steadily reduced, with only estates above $15 million now subject to estate tax under current law (Section I).

Yet the long-running debate over whether inheritances raise or lower inequality has focused almost entirely on the cross-section, treating bequests as a single parent-to-child flow taxed by the estate tax alone. This overlooks two features of the U.S. environment. Wealthy dynasties choose not only how much to bequeath but *to whom*: wealth can skip a generation to grandchildren, a margin taxed by a second instrument, the generation-skipping transfer (GST) tax, layered on top of the estate tax. Also, an early inheritance delivers not just wealth but access to the high returns that drive concentration, and its timing therefore matters as much as its size. The paper therefore asks whether generation-skipping transfers, and the GST tax that regulates them, govern the timing and dynastic *persistence* of wealth as much as its concentration, and whether the estate and GST taxes are distinct enough that the effect of relaxing estate taxation depends on which lever moves.

To answer these questions, I develop a general equilibrium life-cycle model in which wealth is generated both by the earnings process and by heterogeneity in rates of return, with return heterogeneity central to the concentration at the very top of the distribution, consistent with growing evidence that returns rise systematically with wealth.[1] The model has two distinctive features. First, high returns are not earned on wealth automatically. A household earns them only after acquiring a financial *capability*, a high-return technology it can access only by paying a fixed cost, one it can afford only once its wealth, whether from its own saving or from an inheritance, is high enough. Inheritance therefore grants not only wealth but *access* to high returns, and because that access can be inherited, *when* a transfer arrives matters as much as its size: an early inheritance lets an heir earn high returns for decades, whereas a household that must build the wealth itself gains that access only later in life, if at all. For instance, a young heir who inherits a large equity stake in an established private business starts earning high returns immediately, without first spending years building the wealth such access usually requires.

Second, a dynasty chooses to *whom* it transfers. Empirical evidence, which I document in Section I, shows that grandparents matter especially for younger households: among those under age 30 who inherit, transfers from grandparents are common and often larger than those from parents. Allowing a grandparent to transfer directly to a grandchild, the grandparent–grandchild link (G-G link, hereafter), introduces heterogeneity in bequests along the two dimensions the first feature makes relevant: their size and their timing. These

[1] See Benhabib, Bisin and Zhu (2011); Fagereng et al. (2020); Bach, Calvet and Sodini (2020); Hubmer, Krusell and Smith Jr (2021), among others. This contrasts with an alternative that generates concentration from a '*superstar*' productivity state targeting both earnings and wealth (Castaneda, Diaz-Gimenez and Rios-Rull, 2003); this approach tends to leave little role for inheritance by loading concentration onto the earnings process.

transfers are not solely a preference in the model; they arise in part as rational, dynasty-level tax planning, consistent with extensive evidence that the wealthy actively plan around estate and gift taxation (Gale, Hines and Slemrod, 2001). When a household's child already holds wealth above the estate-tax exemption, wealth passed through the child is taxed twice before reaching the grandchild, and a direct transfer skips one round of estate taxation. The model incorporates the U.S. GST tax, which neutralises this saving above the GST exemption, thereby capping tax-motivated skipping at the exemption, exactly the planning margin U.S. law is designed to regulate.

I calibrate the model to the U.S. economy over 1990–2000, a period of relatively stable estate-tax policy, matching the earnings and wealth distributions, the fraction and taxation of estates, and, importantly, the incidence and relative size of grandparent transfers to the young. I then use the calibrated benchmark to ask what the G-G link does to the distribution of wealth and to welfare, and how relaxing the estate tax reshapes both. The central finding is that the G-G link is the engine of dynastic persistence at the top. Only about 0.1 percent of newborns have a wealthy, living grandparent, but for them an early grandparent transfer is pivotal: they accumulate large fortunes by their 30s, and about 62 percent reach the top 1 percent of the wealth distribution by age 65, drawing high returns through inherited wealth rather than a rare productivity draw. Since that access is inherited, an early transfer buys decades of compounding, worth nearly 300 percent of lifetime consumption to the few newborns it reaches, but only about 1 percent to a newborn behind the veil of ignorance. Yet all of this operates through timing and persistence rather than dispersion: removing the link barely changes the wealth Gini or the top 1 percent share, though it lowers aggregate capital by about 3.7 percent. The link, therefore, governs who stays at the top and how early they arrive, far more than how unequal the distribution is at any one moment.

The same asymmetry governs policy. I consider revenue-neutral reforms that relax the estate tax, cutting the rate by about 1 percentage point and doubling the exemption in the proportions observed in the U.S. over the sample period, and their strongest effect is again on persistence, not the cross-section. Relaxing the tax lets wealth pass down the dynasty more durably, and the G-G link is what carries it: the link roughly doubles the response of the top 99–99.9 share to reform, even as the aggregate Gini and the extreme upper tail barely move. Point-in-time concentration rises only modestly (the top 1 percent share by about 1 percentage point under the main reform, from 31.7 to 32.7 percent, and by 2.5 points under full repeal), and, since the estate tax raises little revenue, the capital income tax that replaces it roughly offsets the direct effect of the cut on aggregate saving; newborn welfare, however, falls in every case. The reforms thus bite less on how concentrated wealth is than on how long a family keeps its place at the top.

Separating the two instruments shows why they are distinct, and locates the persistence margin precisely. Raising the exemption lifts the moderately wealthy, whose estates sit just above the exemption, out of the tax net and, by widening the window through which wealth passes untaxed to grandchildren, substantially increases dynastic persistence: doubling the exemption raises the probability that a wealthy grandparent's grandchild reaches the top 1 percent from 62 to 85 percent. Cutting the rate does none of this to persistence (it leaves that probability unchanged) and instead benefits the very wealthiest, who remain taxable but pay less on each dollar; this is what raises concentration. The exemption is thus a persistence instrument and the rate a concentration instrument.

The paper makes three contributions. First, on the modelling side, it distinguishes the two instruments through which the U.S. taxes inherited wealth: standard frameworks focus only on a single parent-to-child flow, taxed by the estate tax alone, whereas here a dynasty chooses to whom it transfers, and returns are generated by an acquired capability rather than by wealth mechanically. Second, the framework provides an economic mechanism for a pattern the multigenerational-mobility literature finds standard models cannot reproduce: that grandparent wealth predicts grandchild wealth beyond the parent.[2] Since a large early transfer switches on high-return access, returns become correlated within dynasties, and I use the calibrated model to show that the G-G link governs the timing, dynastic persistence, and welfare value of wealth far more than its concentration. Third, on the policy side, only a model with both instruments can separate the margins they act on, which reframes how estate-tax reform should be evaluated.[3]

Inheritance and Estate Taxes in Life-Cycle Models Whether inheritances raise or lower wealth inequality has been debated for decades. One line of work finds inheritances equalising, meaning estate taxes raise inequality unless their revenue is redistributed (Gokhale et al., 2001; Wolff and Gittleman, 2014; Elinder, Erixson and Waldenström, 2018); another finds inheritances concentrate wealth, though even here views differ on how much estate taxes can do (Nishiyama, 2002; De Nardi and Yang, 2016; Guo, 2022; Benhabib, Bisin and Zhu, 2011).[4] My paper is closest to De Nardi and Yang (2016) and Guo (2022), who study estate-tax reform in a life-cycle model and find limited effects on cross-sectional concentration. I share that conclusion for the cross-section but depart in two ways: I generate the thick wealth tail through return heterogeneity rather than earnings, giving inheritance a larger role; and neither study considers grandparent transfers or dynastic persistence, which is where I find estate taxation matters most.

Sources of Top Wealth A large literature asks what generates concentration at the very top. One influential approach relies on a fat-tailed 'superstar' earnings state (Castaneda, Diaz-Gimenez and Rios-Rull, 2003; Kindermann and Krueger, 2021), which loads concentration onto labour income and leaves little room for the inheritance and return channels I study. I instead generate concentration through return heterogeneity, in the tradition of Gokhale et al. (2001); but where they assign returns randomly, leaving parent and child returns uncorrelated, in my model returns depend on a capability an inheritance can switch on, and are correlated within a dynasty. This also bears on whether top wealth is self-made or inherited (Kaplan and Rauh, 2013; Cagetti and De Nardi, 2006; Piketty, 2014): my framework accommodates both and takes no stand on which dominates at a point in time, since the dynamic point is that a self-made household transmits access to high returns, and today's self-made rich found tomorrow's dynasties.

Multigenerational Wealth Mobility Family wealth is far more persistent than a parent–child link implies: grandparent wealth predicts grandchild wealth even after conditioning on the parent (Adermon, Lindahl and Waldenström, 2018; Pfeffer and Killewald, 2018), and wealth ranks persist across surnames over centuries (Clark, 2014; Barone and Mocetti, 2021). Benhabib, Bisin and Fernholz (2022) show this long-run persistence challenges the standard framework: i.i.d. returns reproduce the cross-section but not the grandparent

[2] See, for example, Adermon, Lindahl and Waldenström (2018); Pfeffer and Killewald (2018); Benhabib, Bisin and Fernholz (2022).

[3] On the impact of bequests, see Huggett (1996); Gokhale et al. (2001); De Nardi (2004), among others; on the distributional effects of estate-tax reform in a life-cycle framework, Cagetti and De Nardi (2009); De Nardi and Yang (2016); Guo (2022).

[4] Using Swedish administrative data, Nekoei and Seim (2023) show that the equalising effect reverses in the long run, as return heterogeneity makes inherited wealth diverge over time.

effect, which requires return heterogeneity correlated within dynasties. I provide an economic mechanism for exactly that: a large early transfer switches on a household's high-return capability, and returns are then correlated across generations, since access to them is inherited. The resulting G-G persistence is untargeted yet reproduces the observed pattern (Table 13), and it locates the source of that persistence in a transfer margin that taxation can reach.

The remainder of the paper proceeds as follows. Section I sets out the institutional background and the two motivating SCF facts, that grandparent transfers are common and often large among the young, and that returns rise with wealth, stemming from heterogeneity in portfolio composition. Section II develops the model, Section III calibrates it and compares the benchmark with the data, and Section IV presents the main results on the G-G link and estate-tax reform, including the untargeted mobility patterns the model reproduces (Appendix E.1). Section V concludes.

I. Data and Institutional Background

This section sets out the background the analysis rests on: the institutional structure of U.S. transfer taxation, and the empirical patterns in household wealth and intergenerational transfers, documented with the Survey of Consumer Finances (SCF), that motivate the model.

A. Estate and Generation-Skipping Transfer Taxes

The U.S. taxes wealth transfers at death through two federal instruments. The estate tax applies to a decedent's estate above an exemption threshold, and the GST tax applies, on top of the estate tax, to transfers that skip a generation, such as a bequest passing from a grandparent directly to a grandchild. Both have grown steadily more lenient over recent decades. The real exemption rose from about \$1.2 million in 1989 to \$11.4 million in 2019 (in 2019 dollars) while the top marginal rate fell from 55 to 40 percent; most recently, the One Big Beautiful Bill Act of 2025 made a \$15 million exemption permanent from 2026. As a result, only about 2 percent of estates are now taxable.

Two features of this system shape how I model it. First, the estate tax and the GST tax act on different transfer routes: the estate tax on wealth passing through each generation, the GST tax on wealth that skips one. Relaxing estate taxation therefore operates through two distinct margins, a distinction central to the policy analysis in Section IV. Second, the GST exemption has equalled the estate-tax exemption since 2004, and since 2011 the two have moved together dollar-for-dollar under a unified exclusion; I accordingly set the GST exemption and rate equal to their estate-tax counterparts in the model.[5]

B. The Survey of Consumer Finances

I document selected stylised facts on intergenerational transfers and household wealth from the SCF, a triennial survey well suited to studying wealth and inheritance due to its oversampling of rich households and

[5] See Appendix A for the full legislative history and the year-by-year statutory parameters of estate and GST taxes.

a dedicated inheritance module recording the status, size, and, most importantly, the source of transfers.[6] I use all twelve waves, 1989–2022.[7] Two features of U.S. household wealth motivate this study. First, wealth is highly concentrated at the top and has become more so over recent decades. Across the SCF waves, the top 1 percent share of net worth rose from about 30 percent in 1989 to a peak near 38 percent in 2016, easing to 35 percent by 2022, while the top 10 percent came to hold roughly three-quarters of all net worth (Appendix B.1 reports the series and the full distribution). What sustains concentration this high, and how estate taxation bears on it, is the question this paper takes up. Second, as the stylised facts below document, a distinctive channel runs beneath the persistence of top wealth: direct transfers from grandparents to grandchildren.

C. Stylised Facts from the SCF (1989 - 2022)

Stylised Fact 1. *Intergenerational transfers, predominantly taking the form of inheritance, apply only to a small proportion of the population.*

Table 1 shows that only 21 percent of the population has reported receiving substantial assets in the past, with more than 80 percent of the transfers taking the form of inheritances. Notably, for some households, such transfers are not one-time events and may occur multiple times. In contrast, the majority, 70 percent, neither have received nor expect to receive such transfers in the future.

Table 1: Status and Types of Intergenerational Transfers

	Future Transfers		Types of Transfers (%)		
	Expected	Not expected	Inheritance	Gift	Trust
(%) Received	4.3	16.7	80.8	13.6	4.3
(%) Not received	9.4	69.6			

Source. Federal Reserve Board, 1989 - 2022 SCF.

Stylised Fact 2. *The link between households and their grandparents is particularly pronounced for younger households.*

Table 2 highlights the importance of grandparent inheritances for the young. Among households under 30 who report having received an inheritance, 50 percent inherited from a grandparent, likely because most people in this age range still have at least one parent alive. The share of heirs rises with age, and the parent–child link comes to dominate by middle age. Since Table 2 classifies households by age at the survey, someone who inherited before 30 but is now older appears in a later row; even after reclassifying such households by age at receipt, the grandparent link remains more prevalent at early ages. Furthermore, the timing gap is stark. The median inheritance from a parent arrives when the heir is about 50, but from a grandparent at about 30 (Table 3). Grandparent inheritances are accordingly front-loaded: about 46 percent

[6] Wealth is net worth, as in Kuhn, Schularick and Steins (2020) and Bricker et al. (2020). See Appendix B for data and methodology.

[7] Unless stated otherwise, figures in this subsection average across the 12 SCF waves.

arrive before the heir turns 30 and 78 percent before 40, against only 12 and 28 percent of parent inheritances.[8] A parent–child-only model cannot capture this: since a parent is roughly 30 years older and bequests are largely accidental, that channel delivers wealth in middle age, once a household has already accumulated wealth and the return-access margin has little left to do. The transfers that arrive early enough to reshape a lifetime come from grandparents.

Table 2: Percentage of Heirs across Different Age groups

Age	% of Population Inherited	% Inherited from Grandparents	Parents	Others
Under 30	7.0	49.6	34.4	16.0
31 - 40	10.5	42.7	42.4	14.9
41 - 50	14.2	26.4	53.9	19.7
51 - 60	20.3	10.2	72.1	17.7
61 - 70	28.6	4.4	78.3	17.3
71 - 80	28.7	2.7	73.5	23.7
81 and over	25.6	4.5	65.3	30.2

Source. Federal Reserve Board, 1989 - 2022 SCF.

Table 3: Timing of Inheritance Receipt, by Source

Source	Median heir age at receipt	Received before age (%) 30	40
Grandparent	30	46	78
Parent	50	12	28

Note. "Received before age 30/40" is the count-weighted share of each source's own inheritance receipts arriving before that age, pooled across the SCF waves.
Source. Federal Reserve Board, 1989–2022 SCF.

Stylised Fact 3. *The distribution of bequest is highly skewed, predominantly concentrated within the highest wealth quintile.*

Among under-30 heirs, roughly 50 percent belong to the 5th wealth quintile, yet their inheritances constitute over 90 percent of the total received. Table 4 shows that, even within the highest wealth quintile, those who are positioned in the top 1 percent alone are allocated more than half of all inheritances by value. Grandparent and parent inheritances are of broadly comparable total value, but the two are different in kind at these ages: for an heir under 30, an inheritance from a parent requires the parent to have died prematurely, so the parent figures reflect accidental bequests from a small number of early parental deaths, occasionally large, rather than a systematic early-transfer channel. The grandparent channel is systematic by contrast, since grandparents are old enough that dying while the grandchild is young is the norm, which is the distinction

[8]Seen the other way, grandparents bequeath only about a tenth of inheritance value over the whole life cycle, yet a third to a half of the inheritance received before age 30 (34 percent by value and 41 percent by count, pooled across waves; 48 and 49 percent in the 2001 wave to which I calibrate).

the model draws in treating parent-to-child bequests as accidental and grandparent-to-grandchild transfers as the chosen margin.

Table 4: Distributions of Inheritances (Under 30 Heirs)

	Wealth Quintiles					
	1st	2nd	3rd	4th	5th	Top 1%
(%) Heirs	10.4	7.6	13.1	19.1	49.8	5.5
From grandparents	7.5	4.0	5.1	10.2	22.5	1.9
From parents	1.0	2.4	5.2	5.7	19.8	1.9
From both	0.0	0.0	0.4	0.2	1.1	0.7
From others	1.9	1.2	2.4	3.0	6.4	0.9
(%) Total Inheritance	1.5	0.9	1.9	5.1	90.6	51.9
From grandparents	1.3	0.4	0.7	1.8	37.3	15.6
From parents	0.1	0.3	1.0	2.3	48.4	33.6
From others	0.2	0.1	0.2	1.1	4.9	2.8

Note. Each figure is the share of total real inheritance value (or heir count) received by under-30 heirs, pooled across the SCF waves in 2019 dollars. "Top 1%" is a subset of the 5th quintile.
Source. Federal Reserve Board, 1989–2022 SCF.

The last motivating fact concerns return heterogeneity arising from portfolio composition. Bach, Calvet and Sodini (2020) and Fagereng et al. (2020), using Swedish and Norwegian administrative data, document a strong relation between wealth and returns, and Fagereng et al. (2016), Xavier (2021), and Kartashova and Zhou (2021) find a similar pattern in the U.S. SCF.[9]

Stylised Fact 4. *Substantial heterogeneity exists in returns on wealth, stemming from differences in the composition of wealth portfolios across different wealth groups.*

If different asset classes yield different rates, heirs from different wealth groups earn different returns on their inherited wealth depending on its composition. Over the sample period, private business (both corporate and non-corporate) yielded persistently higher returns than other assets, averaging about 21 percent, against roughly 7 percent on public equity and real estate and 2 percent on interest-earning assets. The SCF reveals that portfolios differ substantially across wealth groups (the composition by wealth group is reported in Appendix B.2): households at the top hold a disproportionately large share of public equity and private business, whereas real estate, much of it mortgaged, and interest-bearing assets dominate the portfolios of the bottom half. With the portfolio shares and average returns for each wealth component in hand, Table 5 presents the average portfolio-weighted returns to net worth across different wealth groups over the sample period. Both returns and standard deviations generally exhibit an increasing trend with overall wealth, which is consistent with findings from Bach, Calvet and Sodini (2020) and Fagereng et al. (2020).

[9] I follow their methodology; see Appendix B.2 for the definition of each wealth component and the portfolio-weighted return computation.

Table 5: Heterogeneity in Returns across Different Wealth Groups

Wealth Percentile	0 - 50	50 - 90	90 - 95	95 - 99	99 - 99.9	99.9 - 100
Average Return	(0.01)	0.06	0.07	0.08	0.10	0.12
Stand. Dev	0.01	0.01	0.02	0.03	0.05	0.07

Note. Values in parenthesis indicate negative returns.
Source. Federal Reserve Board, 1998 - 2022 SCF.

Taken together, these facts point to a channel that operates through the *timing* of transfers, not just their size. Stylised Facts 1–3 show that, while the literature has focused on the parent–child link, transfers can arrive much earlier: young heirs often inherit from grandparents while their parents are still alive, and are likely to inherit again from their parents in middle age. Stylised Fact 4 shows that the wealth these heirs receive is precisely the kind that earns high returns, namely the equity and private-business holdings concentrated in wealthy portfolios, which persistently out-yield the real estate that dominates everyone else's. An early inheritance therefore does more than raise a young household's wealth: it grants early access to high returns and lets that advantage compound for decades. This is why the G-G link bears on the persistence of top wealth (which families stay at the top, and how early they get there) rather than only on how concentrated wealth is at a point in time. To this end, the next section develops, calibrates, and simulates a general-equilibrium life-cycle model in which returns are generated by an acquired capability that a large early transfer switches on, and in which dynasties choose not only how much to bequeath but to whom. I use it to examine how the G-G link, and the estate and generation-skipping taxes that bear on it, shape the timing, persistence, and concentration of wealth.

II. Quantitative Life-Cycle Model

I develop a general-equilibrium life-cycle model with bequests to study how inheritance shapes the persistence of wealth across generations and to evaluate estate-tax policy. Return heterogeneity is the central ingredient: high returns require a financial capability that a household must pay a fixed cost to acquire, so that a large *early* inheritance, by making that cost affordable long before a household could save its way to it, buys *access* to high returns rather than just a larger balance. The economy is populated by overlapping generations of finitely-lived households with age-dependent survival. Households derive utility from consumption and from bequests net of taxes, which they begin planning at retirement and which are luxury goods not every household can afford. If they live to see their grandchildren enter the economy, they also value bequests to grandchildren; and when the wealth they pass down would push their children's estates into the taxable range, they have an additional tax motive to transfer directly to grandchildren, a margin disciplined by the GST tax. Workers supply labour inelastically, households save by lending capital to firms, and premature death leaves accidental bequests.

Model Demographics One period in the model is equivalent to 5 years. Households enter the economy at age 25 ($j = 1$) as workers, then consume and accumulate wealth. Workers supply labour inelastically until age 60 ($j = 8$) and retire from age 65 ($j = J_r = 9$). The maximum age is chosen to be 100 ($J = 16$) after which they pass away with certainty. Since a household gives birth at age 30, the grandparent generation is 60 years (12 model periods) older than the grandchild. A grandparent is therefore alive when the grandchild enters the economy at age 25 and may survive until age 100, so the two generations overlap for up to four model periods, when the grandchild is aged 25, 30, 35, or 40. Therefore, a grandchild can receive a bequest from a grandparent at any of these ages, depending on when the grandparent dies, a multi-period overlap that lets G-G transfers arrive early and at different points in the grandchild's life.

At any age $j < J$, households face a conditional probability of surviving from age j to $j + 1$ denoted as γ_j and hence $\gamma_J = 0$. The unconditional probability of reaching age j is therefore denoted by $\gamma^j = \Pi_{m=1}^{j} \gamma_m$. For simplicity, I assume households survive with certainty until their children enter the economy such that $\gamma_j = 1$ for $j \le 6$.[10] The population grows at a constant rate $n = (1 + f)^{1/6} - 1$; the stationary age distribution and the population-accounting identities are given in Appendix C.

Earnings Process Each household supplies one unit of labour inelastically in each period until it retires. The total labour productivity of worker i at age j is given by

$$\log e_j^i = z_j^i + \epsilon_j \tag{1}$$

where ϵ_j is the deterministic age-efficiency profile and z_j^i denotes the stochastic earnings process that follows AR(1):

$$z_j^i = \rho_z z_{j-1}^i + \mu_j^i, \quad \mu_j^i \sim \mathcal{N}(0, \sigma_z^2) \tag{2}$$

Moreover, a household's productivity at entry (at $j = 1$) is drawn from that of its parent at age 55 (at $j = 7$), to capture imperfect intergenerational transmission of earnings ability as in De Nardi and Yang (2016). Transmission is also stochastic: the child's entry state is drawn according to a parent-to-child transmission matrix Q_h, so that a child's earnings ability is correlated with, but not a deterministic function of, its parent's. The matrix Q_h is calibrated to the parent–child earnings correlation together with the concentration of the earnings distribution.

Rate-of-return on Capital Return heterogeneity in the model is generated by a household's *financial-capability* state $p \in \{0, 1\}$. A household with capability p and wealth a that draws the idiosyncratic shock η earns the gross return

$$1 + r(p, a, \eta) = 1 + \underline{r} + p \left(r_e + r_{e2} \cdot \mathbb{1}\{a \ge \underline{a}\} + \sigma_e \cdot \eta \right), \quad \eta \sim \mathcal{N}(0, 1) \tag{3}$$

where $\underline{r}$ is a safe return earned by every household, r_e is a first excess return earned by capable households ($p = 1$), r_{e2} is a further excess return earned on wealth above a threshold $\underline{a}$, and $\sigma_e \cdot \eta$ is idiosyncratic return risk with η an $i.i.d.$ standard-normal shock. Households without capability ($p = 0$) earn only the safe return

[10] This assumption is to ensure that all households have their parents alive when entering and no households die before their parents.

$\underline{r}$, which adjusts to clear the capital market, so the specification remains consistent with the firm's marginal product of capital equalling the aggregate return.[11]

Acquisition of capability Capability is acquired by paying a one-time fixed cost, and, once acquired, is permanent: p transitions from 0 to 1 but never reverts.[12] The one-time fixed cost κ_e of accessing high-return investment captures the combined costs of obtaining financial knowledge, social networks, relationships, and meeting participation requirements.[13] This cost may be paid out of a household's own accumulated savings. A discrete choice is available only to households wealthy enough to afford it ($a \geq \kappa_e$). Denoting $\widetilde{V}(a, \cdot; p)$ for the value of a household that enters the period with capability p and behaves optimally thereafter, a household with $p = 0$ solves the acquisition problem

$$V(a, \cdot; 0) = \max\Big\{ \widetilde{V}(a, \cdot; 0)\ ,\ \ \widetilde{V}(a - \kappa_e, \cdot; 1) \Big\}, \qquad \text{feasible only if } \ a \geq \kappa_e, \tag{4}$$

switching on the high-return technology whenever this raises its value. Crucially, a household may also pay the cost *at the moment it receives an inheritance*: an heir with $p = 0$ receiving a bequest b enters the high-return state by paying the same fixed cost κ_e out of its augmented wealth whenever it can afford to do so, that is, whenever $a + b \geq \kappa_e$. An inheritance therefore does not switch on the high-return technology for free; by augmenting the heir's wealth it brings the fixed cost within reach, letting a young heir acquire access years before it could have afforded κ_e out of its own saving. This is the paper's central mechanism: early inheritance buys the *means to access* high returns, a head start in timing rather than a free pass.

For heirs of wealthy families, however, the requirement to pay κ_e is barely a constraint: a large enough transfer covers it comfortably, so these heirs gain high-return access almost as soon as they inherit. Interpreted through its components, the cost paid out of a bequest also captures in reduced form what inherited wealth tends to carry with it: an established family business, financial relationships, and the social networks and connections that come bundled with a wealthy background. Since κ_e is pinned to a concrete institutional threshold rather than chosen freely, the rise of returns with wealth is an emergent, untargeted feature of the model: capability correlates with wealth only through selection and through the early access that large inheritances make affordable.[14]

Technology The supply side of the economy is modelled as standard. A representative firm hires capital and labour from households and produces consumption goods in this economy using the constant returns-to-scale

[11] See Gabaix et al. (2016) and Xavier (2021) for alternative mechanisms generating return heterogeneity. Unlike the scale-dependent schedule of Hubmer, Krusell and Smith Jr (2021), in which returns rise mechanically with wealth, here *expected* returns are flat within a capability type (apart from the threshold $\underline{a}$), and the rise of returns with wealth arises endogenously through selection, as capable households come to hold more wealth.

[12] This absorbing-state assumption is relaxed in the robustness analysis of Section IV, where making the capable state impermanent leaves the main results intact.

[13] This capability device is a dynastic, inheritance-triggerable analogue of the endogenous financial knowledge in Lusardi, Michaud and Mitchell (2017): rather than being built through costly effort, capability can be switched on by a sufficiently large early transfer, producing returns correlated within dynasties. See also Lusardi and Mitchell (2014) and Van Rooij, Lusardi and Alessie (2011).

[14] The threshold has a concrete institutional counterpart: U.S. securities law restricts the highest-return private vehicles (private placements, hedge funds, and private-equity and venture funds) to "accredited investors" whose net worth exceeds \$1 million, so κ_e is set to this legal cutoff (Section III). The wealth threshold to high-return access is thus a real barrier in how Americans access high-return investment rather than only a modelling device.

Cobb-Douglas production technology:

$$Y = AK^{\alpha} L^{1-\alpha} \tag{5}$$

where A and α denote the level of productivity and capital share in production respectively, and capital depreciates at rate δ_k in each period.

A. Government

The government collects taxes from labour earnings, capital income, and intergenerational transfers to finance its consumption and the interest payments on its debt. In addition to the tax on labour earnings, the government operates a balanced-budget social security program, which collects separate social security taxes from workers to provide benefits to retirees.

Labour Income Tax In the model, the progressive labour earnings tax is characterised by a two-parameter tax function of the form:

$$T_{\ell}(w \cdot e) = w \cdot e - (1 - \chi_e)(w \cdot e)^{1-\mu_e} \tag{6}$$

as in Benabou (2002) and Wu and Krueger (2021), where μ_e and χ_e are parameters governing the progressivity and the level of income tax respectively.

Estate Tax I assume the estate tax in the model is characterised by an exemption threshold χ_b and the tax rate τ_b such that only estates above the threshold are taxed at rate τ_b as in De Nardi and Yang (2016). Let $T_b(\cdot)$ be the estate tax function in the model,

$$T_b(W) = \max\{0, \tau_b(W - \chi_b)\} \tag{7}$$

where W is the total estate, the unspent wealth left by older households. Importantly, the tax is levied once on the total estate before its division among heirs, mirroring the estate-based design of the U.S. federal estate tax (as opposed to an inheritance tax levied separately on each beneficiary's receipt).[15]

Generation-skipping Transfer Tax Transfers that skip a generation (here, bequests left directly to grandchildren) are subject to an additional GST tax, levied on top of the estate tax. Let b_{gc} denote the amount received by a grandchild; the GST tax function is

$$T_{gst}(b_{gc}) = \max\{0, \tau_{gst}(b_{gc} - \chi_{gst})\} \tag{8}$$

[15] In practice, the U.S. estate tax is also progressive. In this study, I follow a commonly used specification in the existing literature by targeting relevant moments using τ_b and χ_b. The flat-rate approximation is accurate for the population to which it applies in the model: the top statutory marginal rate is reached at approximately \$1 million of taxable estate above the exemption threshold, and the model's calibrated bequester population is concentrated well above χ_b, so the flat rate reflects the effective marginal rate faced by households for which the bequest and generation-skipping decisions are operative. The estimation of changes in actual estate tax progressivity is left for future research.

where, following U.S. law, the GST tax rate equals the estate tax rate and the GST exemption mirrors the estate tax exemption: $\tau_{gst} = \tau_b$ and $\chi_{gst} = \chi_b$.[16] As the household problem below makes clear, this tax is central to the economics of GSTs in the model: direct transfers to grandchildren below χ_{gst} allow a dynasty to skip one round of estate taxation, while the GST tax neutralises this saving above the exemption, capping tax-motivated GSTs at χ_{gst}.

Social Security Tax The government also separately collects social security taxes from workers to finance social security benefits to retirees. I assume a simplified social security system that pays a common retirement benefit to all retirees in the model.[17] Following Wu and Krueger (2021), the flat payroll tax in the model τ_{ss} is set to 7.65% based on the actual Social Security and Medicare tax rates on pre-tax income of employees. The amount of benefit is then computed based on,

$$\int s \cdot \mathbb{1}_{\{j \geq J_r\}} \, d\Phi = \tau_{ss} \int w \cdot e(j, z) \, d\Phi \tag{9}$$

Capital Income Tax Finally, the government also collects taxes on capital income earned by households at a constant rate τ_a. I assume that the government in the model consumes a constant fraction of GDP and maintains a constant government debt-to-GDP ratio. Therefore, for any changes in government revenues induced by counterfactual estate tax reforms, the government will adjust its capital income tax rate to rebalance the government budget.[18]

B. Household's Problems

Preference The utility from consumption takes the standard CRRA form and the *warm-glow* bequest preference extends De Nardi and Yang (2016),

$$u(c) = \frac{c^{1-\sigma} - 1}{1-\sigma}, \qquad h_j(b_c, b_{gc}; d_c) = \frac{\phi_1}{1-\sigma}\left[\left((b_c + \phi_2^c)^{\kappa_j(d_c)}(b_{gc} + \phi_2^{gc})^{1-\kappa_j(d_c)}\right)^{1-\sigma} - 1\right] \tag{10}$$

where σ is the risk-aversion coefficient, $\phi_1 > 0$ measures the bequest intensity, and b_c and b_{gc} denote the bequest to child and grandchild respectively. The parameters ϕ_2^c and ϕ_2^{gc} are additive shifters that make the two bequest legs luxury goods. With $\phi_2^c > 0$, bequests to children are a luxury: sufficiently poor households leave nothing when they die, so the model endogenously segments households into bequesters and non-bequesters. The grandchild shifter ϕ_2^{gc} enters symmetrically but does *not* determine whether a household leaves anything to a grandchild.[19] Directing any bequest to a grandchild requires a fixed cost $\xi > 0$, the

[16] The GST tax (Internal Revenue Code, Chapter 13) was introduced precisely to prevent wealthy dynasties from avoiding one round of estate taxation by transferring wealth directly to grandchildren. Each individual has a GST exemption equal to the basic estate tax exclusion amount, and transfers above the exemption are taxed at a flat rate equal to the top estate tax rate.

[17] This assumption is to avoid an additional continuous state variable (e.g., average indexed monthly earnings) to track an agent's future retirement benefit.

[18] I take capital-income-tax closure as the conservative benchmark: a capital tax falls disproportionately on the high-return households who gain most from relaxing the estate tax, so it offsets part of their gain and yields smaller measured effects on concentration and welfare (see the robustness analysis in Section IV).

[19] This departs from a pure luxury-good formulation, in which one would impose $\phi_2^{gc} > \phi_2^c$ to make direct transfers to grandchildren a stronger luxury, and hence rare; here, instead, the extensive margin of generation-skipping is supplied by a separate fixed cost.

administrative and estate-planning cost of establishing a generation-skipping arrangement.

A household that dies once its grandchild has entered the economy therefore faces a discrete choice: leave the entire estate to the child, or pay ξ and divide the estate between child and grandchild. Only estates for which the warm-glow gain from skipping exceeds ξ leave a GST, so the *prevalence* of such transfers is an equilibrium outcome disciplined by the data, governed by ξ separately from the division *share* κ_j introduced below.[20]

The weight κ_j governs the allocation between children and grandchildren:

$$\kappa_j = 1 \quad \text{if} \quad j \leq 12, \qquad\qquad \kappa_j(d_c) = \begin{cases} \kappa_0 & \text{if} \quad d_c = 1 \\ \kappa_1 & \text{if} \quad d_c = 0 \end{cases} \quad \text{if} \quad j \geq 13, \qquad \kappa_0 < \kappa_1 \leq 1 \tag{11}$$

where $d_c = \mathbb{1}\{a_c + b_c \geq \chi_b\}$ indicates whether the household's child will itself face estate taxation (that is, whether the child's own wealth a_c *together with* the bequest b_c it receives from the grandparent exceeds the estate tax exemption), and $j \geq 13$ is the overlap window in which the grandchild is present; the grandparent has then reached age 85 or beyond. Setting $\kappa_j = 1$ for all j shuts down the G-G link.

The dependence of κ_j on d_c captures dynasty-level estate tax planning. When the child's estate will itself be subject to estate taxation ($d_c = 1$), wealth passed to the grandchild through the child is taxed twice on its way down the dynasty; transferring directly to the grandchild saves τ_b per dollar in present value. Since the trigger is evaluated on the child's wealth *inclusive* of the bequest it receives, a large bequest can by itself lift the child above the exemption, making the child's estate taxable when it otherwise would not be. The households with the most wealth to pass down are therefore precisely those whose direct transfers to grandchildren avoid an extra round of taxation, so the tax motive bites hardest exactly where transfers are largest. Households whose children remain below the exemption even after the bequest have no such tax motive, and their GSTs reflect the warm glow alone (κ_1). The GST tax introduced above completes the picture: since the GST neutralises the estate tax saving above χ_{gst}, even strongly tax-motivated households face an interior optimum, with tax-driven transfers capped at the GST exemption.

Rather than treat κ_0 and κ_1 as independent parameters, I tie the tax-motivated wedge to the estate-tax rate,

$$\kappa_1 - \kappa_0 = g\,\tau_b, \qquad g > 0, \tag{12}$$

so that the additional weight a grandparent places on the grandchild when the child is taxable ($d_c = 1$) is proportional to the estate tax that skipping avoids. This specification makes the strength of the avoidance motive respond to the estate-tax rate: an estate-tax reform alters not only how much is bequeathed but how strongly dynasties are induced to redirect wealth toward grandchildren, which is the margin the generation-skipping analysis in Section IV turns on.[21]

[20]The fixed cost is essential for matching the data. Economically, ξ captures the fixed legal and administrative cost of a generation-skipping vehicle (for example, establishing and maintaining a trust) so that generation-skipping is worthwhile only for estates above a threshold size.

[21]The warm-glow division κ_1 is calibrated to the observed value share of grandparent transfers ($\kappa_1 = 0.29$ and $\kappa_0 = 0.19$ at the benchmark rate). Since the cross-section pins only this composite share, the wedge coefficient g is not separately identified by steady-state moments; I set it *a priori* to $g = 0.92$ and report it among the externally set parameters (Section III, Table 6).

Here, I make two simplifying assumptions. First, as mentioned above, the G-G motive is active only once the grandchild has entered the economy: a household that dies before age 85 ($j \leq 12$) has no grandchild in the economy yet, so its entire estate passes to its child. Second, I assume households begin their estate planning as they transition into retirement. Given the model demographics, households live with certainty until their children enter the economy at age 55 and they retire from age 65. While households face positive mortality rates between ages 55 and 65, I consider this as part of accidental bequests.

The above specification lets a household become a bequester regardless of whether its parents were wealthy, conditional on surviving long enough: even without wealthy parents or grandparents, a household that is "self-made" wealthy at retirement becomes a bequester and leaves bequests to its offspring. At the same time, the specification captures some persistence within wealthy families, since households that receive substantial inheritances are more likely to become bequesters at retirement. Finally, it lets the composition of bequesters and non-bequesters vary as changes in estate tax policy shift the stationary distribution.

Next, I describe households' problem at each age $j = 1$ to $J = 16$. Since it is a finite-horizon life-cycle model, I start from retirees' problems then move to workers' problems. Bequests are specified for each dynasty, and agents anticipate the bequests they may receive from their parents' (and grandparents') saving behaviour. I separately describe households whose parents (or grandparents) are alive, distinguishing further between those whose parents are retired and those whose parents are working. The recursive problems below are written with the ancestors' full states as arguments for illustrative purposes.[22]

Every value function carries the household's financial-capability state $p \in \{0, 1\}$ as a state variable. At the start of each period a household without capability solves the acquisition problem above, so the value functions below are understood to be the post-acquisition values $V(\cdot\,; p)$; an heir without capability may in addition acquire capability at the moment a bequest b is received, paying the fixed cost κ_e whenever its augmented wealth covers it ($a + b \geq \kappa_e$), as described above. For notational simplicity, I suppress p from the continuation arguments where it plays no additional role, and write the net-of-tax return on capital as $r^n(p, a, \eta) = (1 - \tau_a)\big(\underline{r} + p\left(r_e + r_{e2} \cdot \mathbb{1}\{a \geq \underline{a}\} + \sigma_e \cdot \eta\right)\big)$.

Retirees The set of state variables for retirees are age j, current asset holdings a, the idiosyncratic return shock η, and the financial-capability state p; once the grandchild is present ($j \geq 13$), the state additionally includes the child's estate-tax status $d_c = \mathbb{1}\{a_c + b_c \geq \chi_b\}$, evaluated on the child's wealth inclusive of the bequest b_c it receives. Let V_r denote the value function. *Before the grandchild enters*, for $J_r \leq j \leq 12$, the recursive problem is given by:

$$V_r(j, a, \eta, p) = \max_{c, a'} \left\{u(c) + \beta \mathbb{E}\left[\gamma_j V_r(j + 1, a', \eta') + (1 - \gamma_j) \cdot h_j(b_c, 0)\right]\right\} \tag{13}$$

subject to

$$\begin{aligned} c + a' &= s + (1 + r^n) \cdot a, \\ b_c &= a' - T_b(a') \end{aligned} \tag{14}$$

[22] In the numerical solution, however, I do not track each ancestor's continuous state: I summarise ancestors by a small number of wealth bins and replace the inferred bequest with a draw from the conditional receipt distribution implied by that bin. This approximation also significantly reduces computation costs. See Appendix D for the detailed fixed-point procedure.

where $r^n(p, a, \eta)$ is the net-of-tax return on capital defined above and s is the social security benefit provided by the government. In this range the grandchild has not yet entered the economy, so the entire estate net of the estate tax (7) (levied once on the total unspent wealth a') is inherited by the child ($\kappa_j = 1$, so b_{gc} does not enter the warm glow). At $j = 12$, the period before the overlap window begins, the expectation over the next period's state includes d_c, drawn from the equilibrium conditional distribution $\pi(d_c \mid a')$.[23]

During the overlap window, for $13 \leq j \leq J$, the grandchild is present and a household that dies divides its estate between child and grandchild, choosing whether to incur the fixed cost ξ of a GST:

$$V_r(j, a, \eta, d_c, p) = \max_{c,\, a',\, b_c,\, b_{gc},\, \iota \in \{0,1\}} \Big\{ u(c) + \beta \mathbb{E} \Big[\gamma_j V_r(j+1, a', \eta', d_c') + (1 - \gamma_j)\, h_j(b_c, b_{gc}; d_c) \Big] \Big\} \tag{15}$$

subject to

$$\begin{gathered} c + a' = s + (1 + r^n) \cdot a, \qquad W = a' \\ b_c + b_{gc} + T_{gst}(b_{gc}) = W - T_b(W) - \iota \xi, \qquad b_c \geq 0,\ b_{gc} \geq 0,\ \ b_{gc} = 0 \text{ if } \iota = 0 \end{gathered} \tag{16}$$

where $W = a'$ denotes the total estate (the unspent wealth of a household that dies) and $\iota \in \{0, 1\}$ indicates whether the household sets up a GST, incurring the fixed cost ξ; with $\iota = 0$ the entire post-tax estate passes to the child ($b_{gc} = 0$). At the maximum age $j = J$, death is certain ($\gamma_J = 0$), so the continuation term vanishes and the household divides its entire estate. The estate tax (7) is applied once to the total estate W (consistent with the estate-based design of U.S. federal estate taxation), and the GST tax (8) is paid out of the estate on the amount b_{gc} received by the grandchild. Conditional on skipping ($\iota = 1$), the optimal division satisfies

$$\frac{1 - \kappa_j(d_c)}{b_{gc} + \phi_2^{gc}} = \frac{\kappa_j(d_c)\left(1 + \tau_{gst} \cdot \mathbb{1}\{b_{gc} > \chi_{gst}\}\right)}{b_c + \phi_2^c} \tag{17}$$

at interior optima. Two features follow. First, the household skips ($\iota = 1$) only when the resulting warm-glow gain exceeds the fixed cost ξ; estates below the implied threshold leave nothing to the grandchild ($b_{gc} = 0$), so the *prevalence* of generation-skipping is governed by ξ, separately from the division *share* κ_j, which is what allows the model to match both the frequency and the size of GSTs. Second, the kink in T_{gst} at χ_{gst} generates a mass of estates whose GST equals exactly the GST exemption, the model counterpart of estate planners exhausting the GST exemption.

[23] $\pi(d_c \mid a')$ is the probability that the household's child will face estate taxation, conditional on the household's own savings a'. Since the trigger is bequest-inclusive, this is the probability that the child's own wealth plus the bequest $b_c = a' - T_b(a')$ it stands to receive exceeds χ_b, i.e. $\pi(d_c = 1 \mid a') = \Pr\big(a_c \geq \chi_b - (a' - T_b(a'))\big)$, which rises toward one as the household's own estate grows. It is computed from the stationary joint distribution of linked parent–child states, and is therefore an equilibrium object determined as part of the fixed point.

Workers without parents Let V_w be the value function of a working agent without living parents,

$$V_w(j,a,z,\eta,p) = \begin{cases} \max_{c,a'} \left\{u(c) + \gamma_j \beta \mathbb{E}\left[V_r(j+1,a',\eta')\right]\right\} & \text{if} \quad j = J_r - 1 \\ \max_{c,a'} \left\{u(c) + \gamma_j \beta \mathbb{E}\left[V_w(j+1,a',z',\eta')\right]\right\} & \text{if} \quad j < J_r - 1 \end{cases} \tag{18}$$

subject to:

$$c + a' = (1-\tau_{ss}) \cdot w \cdot e(j,z) - T_\ell(w \cdot e) + (1+r^n) \cdot a \tag{19}$$

where $T_\ell(w \cdot e)$ and τ_{ss} are progressive labour income tax and social security tax as in (6) and (9) respectively and the productivity z is part of their state variables.

Workers with living parents A worker whose parent is still alive carries the parent's state S_p and anticipates the accidental bequest it receives at the parent's death. With a *retired* parent, $S_p = \{a_p, \eta_p\}$ and

$$\begin{aligned} V_w^{pr}\left(j,a,z,\eta,p,S_p\right) = \max_{c,a'} \Bigg\{ u(c) + \gamma_j \beta \mathbb{E}\Bigg[& \gamma_{j+6} \cdot V_w^{pr}\left(j+1,a',z',\eta',S'_p\right) \\ & + \left(1-\gamma_{j+6}\right) \cdot V_w\left(j+1,a' + \tfrac{b_p}{1+f},z',\eta'\right)\Bigg]\Bigg\} \end{aligned} \tag{20}$$

subject to (19). The parent dies with probability $1-\gamma_{j+6}$, leaving the accidental bequest $b_p = a'(j+6,S_p)$, after which the heir drops the parent's state and continues as V_w.[24] At a worker's earliest ages the parent may still be *working*, in which case its state also carries productivity ($S_p = \{a_p, z_p, \eta_p\}$) and the continuation follows the parent's own work-then-retire path; this variant, denoted V_w^{pw}, is otherwise identical.

During the overlap window ($1 \leq j \leq 4$, ages 25–40), a grandparent is still alive and the worker additionally carries the grandparent's state $S_g = \{a_g, \eta_g\}$. The grandparent dies with probability $1-\gamma_{j+12}$, in which case the grandchild receives its skip leg $b_g/(1+f)^2$ of the estate:

$$\begin{aligned} V_w^{pg}\left(j,a,z,\eta,p,S_p,S_g\right) = \max_{c,a'} \Bigg\{ u(c) + \gamma_j \beta \mathbb{E}\Bigg[& \gamma_{j+12} \cdot \widetilde{V}\left(j+1,a',z',\eta',S'_p,S'_g\right) \\ & + \left(1-\gamma_{j+12}\right) \cdot \widehat{V}\left(j+1,\ a' + \tfrac{b_g}{(1+f)^2},z',\eta',S'_p\right)\Bigg]\Bigg\} \end{aligned} \tag{21}$$

subject to (19), where $S_g = \{a_g, \eta_g\}$ denotes the grandparent's state. Here $\widetilde{V} = V_w^{pg}$ while the grandparent survives and the window remains open ($j+1 \leq 4$), and reverts to the corresponding with-parent value function (V_w^{pw} or V_w^{pr}, according to whether the parent is still working) once the window closes; $\widehat{V}$ is that same with-parent function, entered once the grandparent has died and b_g has been received. The parent's own survival and bequest b_p are handled within these continuation functions exactly as above. Since $\gamma_J = \gamma_{16} = 0$, a grandparent surviving to the grandchild's age $j = 4$ transfers with certainty, so a grandchild whose grandparent lives into the overlap receives b_g at some age in $\{1,2,3,4\}$, the multi-period receipt window described in the demographics above.

[24] Parents' age is redundant given the constant age gap; and, as in De Nardi and Yang (2016), children anticipate parental bequests.

C. Stationary Equilibrium

I consider a stationary equilibrium of the model economy in which factor prices $(r^*, \underline{r}, w)$, demographic processes (f, n, γ), and age-wealth distribution are constant over time.

A *stationary recursive competitive equilibrium* consists of value functions $\{V_w^{pg}, V_w^{pw}, V_w^{pr}, V_w, V_r\}$ and policy functions $\{c(x), a'(x)\}$, firm choices $\{K, L\}$, prices $(r^*, \underline{r}, w)$, and an invariant distribution Φ such that households and the representative firm optimise, the government and social-security budgets balance, and the labour, capital, and goods markets clear, with capital-market clearing pinning down the aggregate return $\underline{r}$.[25]

III. Calibration

This section describes the parameter choices for the baseline model economy. The model is calibrated to match the characteristics of the U.S. economy in 1990–2000, a period when estate-tax policy was relatively stable. I choose a subset of parameters based on model-exogenous information and calibrate remaining parameters internally so that model moments closely match their data counterparts. Table 6 summarises the parameters that are externally calibrated. I report parameters at an annual frequency (unless stated otherwise) though they are converted to 5-year frequency in the computation.

First, I set the population growth rate to $n = 1.1\%$, the long-run U.S. average, and obtain the age-dependent survival probabilities $\{\gamma_j\}$ from the U.S. life table for 2001. The capital share $\alpha = 0.36$ and the risk-aversion coefficient $\sigma = 1.5$ take standard values from the quantitative macroeconomics literature. The child luxury parameter ϕ_2^c, which segments households into bequesters and non-bequesters, is set at a value that generates a non-degenerate mass of non-bequesters.

Since labour supply is inelastic, the earnings process is calibrated externally to the earnings distribution. The persistence ρ_z and variance σ_z^2 are from De Nardi (2004), and the intergenerational transmission matrix Q_h is calibrated to a parent–child earnings correlation of 0.33 that I compute from the PSID. Following Castaneda, Diaz-Gimenez and Rios-Rull (2003) and Kindermann and Krueger (2021), the stochastic component z is a five-state Markov chain: Tauchen (1986) yields the lowest three states, to which I append two high-earnings states reproducing the concentration of the earnings distribution. Then I calibrate remaining parameters (including the earnings transmission matrix, Q_h) to match earnings concentration observed in the data. In doing so, I differ from the usual "superstar" process in that I only use information from the empirical earnings distribution, and I do not target wealth distribution. This yields five productivity states,

$$z = [0.3923,\ 1.0000,\ 2.5492,\ 15.7425,\ 110.5233]$$

and the implied initial distribution at age 1,

$$\mu_1 = [38.3\%,\ 0.1\%,\ 61.0\%,\ 0.2\%,\ 0.5\%]$$

[25] The full equilibrium conditions, including the government budget constraint (with estate and GST revenue) and the goods-market resource cost $\mathcal{K}$ of capability acquisition and generation-skipping arrangements, are stated in Appendix C.

Table 6: Externally Calibrated Parameters

Parameter		Value	Target/Data
Annual population growth rate	n	1.10%	U.S. long-run average
Age-dependent survival probability	γ_j	$*$	U.S. Life table (2001)
Capital share in production	α	0.360	Standard
Risk aversion coefficient	σ	1.500	Standard
Bequest luxury parameter (child)	ϕ_2^c	1.100	Set (bequest luxury; see text)
Tax-response coefficient (skip wedge)	g	0.920	Set; $\kappa_1 - \kappa_0 \approx 0.10$ at benchmark τ_b
Productivity state	z	$*$	See text
Age-efficiency profile	ϵ_j	$*$	SCF (2001)
5-year labour productivity persistences	ρ_z	0.850	De Nardi (2004)
5-year labour productivity variance	σ_z^2	0.300	De Nardi (2004)
Productivity transition probability	Q_z	$*$	See text
Productivity transmission probability	Q_h	$*$	See text
i.i.d. idiosyncratic shock	η	5 nodes	See text
Capability acquisition cost	κ_e	\$1.00 M	SEC accredited-investor threshold
Government consumption-to-GDP	G/Y	18.0%	FRED (A822RE1Q156NBEA)
Government debt-to-GDP	D/Y	55.0%	FRED (DEBTTLUSA188A)
Labour income tax progressivity	μ_e	0.1327	Wu and Krueger (2021)
Labour income tax level	χ_e	0.1575	Wu and Krueger (2021)
Social security payroll tax rate	τ_{ss}	0.0765	Wu and Krueger (2021)

Note. Unless stated otherwise, parameters are reported at an annual frequency.

I use the mean earnings for households aged 25–60 from the SCF 2001 as the model's deterministic age-efficiency profile of labour productivity, ϵ_j.[26] I then normalise the age-efficiency profile such that the average earnings (before tax) at age-1 is equal to unity.[27]

Third, return heterogeneity is generated by a household's *financial-capability state* $p \in \{0, 1\}$. The net return on assets held by household i with capability p_i and wealth a_i is

$$1 + r_i = 1 + \underline{r} + p_i \left(r_e + r_{e2} \cdot \mathbb{1}\{a_i \geq \underline{a}\} + \sigma_e \cdot \eta \right)$$

where $\underline{r}$ is the safe return earned by every household and $p_i = 1$ denotes a household that has acquired financial capability. A capable household earns a first excess return r_e and, on wealth above a threshold $\underline{a}$, an additional second excess return r_{e2}, both carrying idiosyncratic risk $\sigma_e \cdot \eta$. The $i.i.d.$ standard-normal shock η is discretised by 5-point Gauss–Hermite quadrature, with nodes $\{\pm 2.857, \pm 1.356, 0\}$ and weights $\{0.011, 0.222, 0.533, 0.222, 0.011\}$, matching the standard-normal moments up to fourth order ($\mathbb{E}[\eta^2] = 1$, $\mathbb{E}[\eta^4] = 3$), so that σ_e is directly interpretable as the standard deviation of the idiosyncratic return shock.

[26] I set $\epsilon_j = 0$ for $j \geq J_r$ due to mandatory retirement at age 65 (i.e. $j = 9$) in the model.

[27] The earnings profile is normalised so that the model's average earnings at age 25 equal one model unit, which corresponds to \$46,472 per year (\$232,360 per 5-year period, in 2019 US dollars); this is the value at which the capability cost matches the SEC accredited-investor net-worth threshold (κ_e = \$1.00 million). The SCF-2001 mean earnings at age 25 are about 1.12 model units. See Appendix D for more information on the calibration of the earnings process.

Capability is acquired by paying the one-time fixed cost κ_e out of own saving, or, for an heir at the moment of receipt, the same cost κ_e, exactly as described in Section II; once acquired, $p = 1$ is absorbing.[28]

Since a household's *expected* return does not rise with the *level* of wealth within a capability type (the idiosyncratic shock $\sigma_e \eta$ is mean-zero and independent of wealth), the empirical rise of returns with wealth (Table 5) is not imposed but *emerges* through selection: wealthier households, and especially those who inherit early, are more likely to have acquired capability. The wealth–return gradient is therefore an untargeted validation moment rather than a modelling input, and it sharpens the central mechanism, namely that early inheritance does not simply transfer wealth but buys *access* to high returns. The acquisition cost κ_e is set externally to the Securities and Exchange Commission (SEC) accredited-investor net-worth threshold ($1.00 million), while the remaining return parameters are calibrated internally to wealth concentration at the top; the calibrated premia are broadly consistent with the portfolio-weighted excess returns estimated from the SCF (Table 5).[29]

Next, the social security payroll tax rate, $\tau_{ss} = 7.65\%$, which implies a common retiree benefit of $27,310 (2019 dollars), and the parameters governing the progressivity of labour earnings income, $\mu_e = 0.1327$ and $\chi_e = 0.1575$, are taken from Wu and Krueger (2021).[30] Lastly, I assume government consumes a constant fraction of output ($G/Y = 18\%$) in each period and outstanding government debt D is set such that the government debt-to-GDP ratio is 55% in the initial equilibrium. Both are taken from the FRED.

The remaining parameters are chosen jointly to match the target moments, with some fixed directly by equilibrium conditions: a target real interest rate $r^* = 7\%$ requires depreciation $\delta_k = 3.85\%$, and technology $A = 1.580$ normalises the wage per efficiency unit to $w = 1$. Table 7 below summarises the parameters that are internally calibrated. First, the discount factor, β and the bequest intensity parameter, ϕ_1 are chosen to match the capital-output ratio of 3.10 and bequest-wealth ratio of 2% respectively.

The grandparent-transfer parameters are calibrated jointly to the SCF transfer moments. The extensive margin of generation-skipping is governed by the fixed cost ξ, chosen to target the fraction of young households who inherit from grandparents; the grandchild shifter ϕ_2^{gc} is calibrated to the *source* (count) share of grandparent transfers; and the division share κ_1 (applied when the child's estate is below the exemption threshold) is chosen to match the *value* (flow) share of grandparent transfers, while κ_0 (child above the threshold) is tied to κ_1 by a wedge proportional to the estate-tax rate ($\kappa_1 - \kappa_0 = g\,\tau_b$), so that the tax-motivated component of skipping scales with the tax it avoids (see Section II).[31] Since a fixed cost

[28] This type-based specification also speaks to the finding that pure scale dependence explains only a small share of return variation in the data (Bach, Calvet and Sodini, 2020): here return differences are driven by capability type and transitory risk, with scale entering only through the second-tier threshold $\underline{a}$.

[29] Calibrating the return dispersion internally is standard for this class of models (Hubmer, Krusell and Smith Jr, 2021). Since the model's return states are far coarser than the wealth bands underlying Table 5, σ_e measures dispersion within a broad capable type and need not equal any single band's standard deviation. A household-level computation on the SCF yields a within-band return standard deviation of roughly 0.04–0.05; this is a lower bound, because the SCF records the current value of private businesses but not their cost basis and so cannot measure realised business returns, the dominant source of return dispersion. The calibrated value, $\sigma_e = 0.146$, sits below the cross-sectional standard deviation of net-worth returns of about 0.22 that Fagereng et al. (2020) document with Norwegian administrative panel data, and within the idiosyncratic, largely business-driven return dispersion that they and Bach, Calvet and Sodini (2020) measure once realised business returns are observed.

[30] Wu and Krueger (2021) estimate the two parameters, μ_e and χ_e for the U.S. by running the OLS regression: $\ln(e - T_\ell(e)) = \ln(1-\chi_e) + (1-\mu_e)\ln(e)$ where the tax liabilities $T_\ell(e)$ are defined as federal income taxes minus eligible amounts of the Earned Income Tax Credit (EITC) and food stamp benefits.

[31] The grandparent-transfer moments are computed from the SCF waves spanning 1989–2019; including the 2022 wave leaves

Table 7: Internally Calibrated Parameters

Parameter		Value	Target/Data
Time discount factor (annualised)	β	0.982	$K/Y = 3.10$
Bequest intensity	ϕ_1	0.81	$B/W = 0.02$
Fixed cost of generation-skipping	ξ	\$ 0.14 M	Fraction of young inheriting from grandparents
Grandchild luxury shifter	ϕ_2^{gc}	0.265	Source (count) share of G-G transfers
Division share at $j \geq 13$ ($d_c = 1$)	κ_0	0.19	$\kappa_0 = \kappa_1 - g\,\tau_b$ (tax-responsive wedge; see text)
Division share at $j \geq 13$ ($d_c = 0$)	κ_1	0.29	Value share of G-G transfers
Technology level	A	1.580	$w = 1$
Capital depreciation rate	δ_k	3.85%	$r^* = 7\%$
Aggregate return component	$\underline{r}$	5.54%	$r^* \int a \, d\Phi = \int \big(\underline{r} + p(r_e + r_{e2}\mathbb{1}\{a \geq \underline{a}\} + \sigma_e \eta)\big) \cdot a \, d\Phi$
First excess return (annual)	r_e	3.1%	Share held by 95th–99th percentiles
Second excess return (annual)	r_{e2}	0.6%	99–99.9 share (2nd-tier premium, with $\underline{a}$)
Second-tier wealth threshold	$\underline{a}$	\$ 5.0 M	99–99.9 share (2nd-tier threshold, with r_{e2})
Std. dev. of excess return	σ_e	0.146	Top 0.1% share
Social security benefit	s	\$ 27,310	Social security budget balance
Capital income tax rate	τ_a	15.24%	Government budget balance
Estate tax rate	τ_b	10.6%	Estate tax revenue (as % of GDP)
Estate tax exemption threshold	χ_b	\$ 5.95 M	% of estates that are subject to tax

Note. Unless stated otherwise, parameters are reported at an annual frequency.

creates an extensive margin that is distinct from the intensive division share, the model can match both the prevalence and the size of GSTs, which a smooth division rule alone cannot.[32]

The excess-return premia and the capability-acquisition parameters are calibrated jointly to the concentration at the top of the wealth distribution: the first excess return r_e targets the share held by the 95th–99th percentiles, the second-tier threshold $\underline{a}$ = \$5.0 million targets the 99th–99.9th percentile share, and the return dispersion $\sigma_e = 0.146$ targets the top 0.1% share, while the acquisition cost κ_e is set externally to the SEC accredited-investor net-worth threshold (\$1.00 million), so that high-return participation among the wealthy is a resulting, untargeted moment. The aggregate safe return component $\underline{r}$ is found to be 5.54% in the baseline. The main estate parameters, τ_b and χ_b are chosen to match the fraction of estate tax revenue to output (0.33%) and the fraction of estates that are subject to tax (2.0%) respectively, as in De Nardi and Yang (2016). In the baseline, $\tau_b = 0.106$ and χ_b = \$5.95 million, which corresponds closely to the statutory highest tax bracket exemption prevailing in 1990–2000 (in 2019 dollars). Finally, the capital income tax rate of $\tau_a = 15.24\%$ balances the government budget.

A. Calibration Results

Table 8 reports shares of total earnings and wealth held by percentile group, in the data and the benchmark model, with Gini coefficients. The baseline model produces cross-sectional distributions of both earnings and

the source (count) share of grandparent transfers essentially unchanged and moves the value share by at most about two percentage points, so the calibration targets are robust to the sample period.

[32] With the calibrated luxury parameter ϕ_2^c, the intended-bequest motive is active only for households whose end-of-life wealth is sufficiently high, so that poorer households leave no intended bequest; the model reproduces the empirical fraction of estates that leave a bequest.

wealth that closely align with empirical counterparts. In particular, it generates a substantial concentration of wealth at the upper end of the wealth distribution. The benchmark matches the top 0.1% share of the earnings distribution and closely reproduces it for wealth (about 11 percent in the model against 10 in the data); the model is modestly less concentrated than the data across the 90th–99th percentiles of wealth, holding correspondingly more in the upper-middle of the distribution (the 50th–90th band), neither of which is targeted, reflecting the discrete high-return capability state. The final row makes the role of return heterogeneity explicit: an otherwise-identical economy with a single interest rate (the same earnings process, but without capability-based return heterogeneity) produces a top-1% wealth share of only 18 percent (against 32 in the data) and a wealth Gini of 0.68. Earnings heterogeneity alone, even with the fat upper tail that reproduces the earnings distribution, cannot account for the concentration of wealth; return heterogeneity, and the inherited access to it, is essential. The same decoupling appears *within* earnings groups: near retirement the model's average wealth Gini conditional on earnings is about 0.58, close to the 0.54 that Hendricks (2007) documents across lifetime-earnings deciles in the U.S., so the wide dispersion of wealth among households with similar earnings is an untargeted feature the model reproduces.

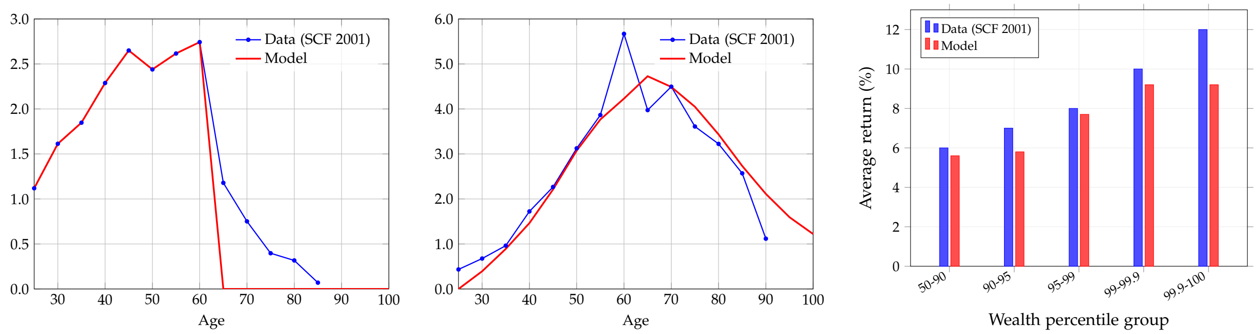
Table 8: Earnings and Wealth Distributions in Benchmark Economy

	Share of total (%)						
	0–50	50–90	90–95	95–99	99–99.9	99.9–100	Gini
Earnings							
Data (2001)	9.6	44.1	11.3	16.1	12.8	6.1	0.63
Model	9.5	45.0	10.7	16.0	12.7	6.1	0.62
Wealth							
Data (2001)	2.8	27.6	12.2	25.2	21.8	10.5	0.81
Model (Benchmark)	3.3	34.2	9.9	20.9	21.0	10.7	0.77
Model (Single rate)	4.3	47.5	12.5	17.8	12.5	5.4	0.68

Note. The *Single rate* row holds all parameters at their benchmark values except the discount factor β, which is re-cleared to the same capital–output ratio.

Figure 1 reports the model's life-cycle fit. The calibrated earnings process replicates the mean earnings profile until retirement, and average wealth by age tracks the empirical hump and gradual old-age decumulation, neither profile targeted beyond the earnings distribution. As a final untargeted check, panel (c) compares the model's average return by wealth group with the SCF profile (Table 5): the model reproduces the upward gradient through selection alone, its mean return rising from 5.6 to 9.2 percent across the five groups against 6.0 to 12.0 percent in the data, with the very top carried by return *dispersion* rather than a higher mean.

Figure 1: Model Fit: Life-cycle Earnings and Wealth Profiles, and Returns by Wealth Group



(a) Earnings (Targeted) **(b)** Wealth (Untargeted) **(c)** Returns (Untargeted)

Table 9 reports the targeted moments in the baseline model; the implied macroeconomic aggregates are reported in Table 10. The untargeted macroeconomic ratios are in line with postwar U.S. averages: a capital–output ratio of 3.1, and consumption and investment shares of 67 and 15 percent of output. The calibrated earnings process generates both parent–child earnings correlation and the earnings mean-to-median ratio that are close to the data. The calibrated model is also able to generate overall bequest characteristics. In the model, 2.00% of estates are taxable and the ratio of government estate tax revenue to output is 0.33%. Moreover, it closely reproduces both the extensive and intensive margins of inheritances from grandparents, matching each to within a few percentage points.

Table 9: Targeted Moments and Empirical Counterparts

Target	Source	Target/Data		Model	
Capital-output ratio	·	3.10		3.10	
Aggregate interest rate	·	7.00		7.00	
Wealth share, 95-99th Percentile	SCF (2001)	0.25		0.21	
Wealth share, 99-99.9th Percentile	SCF (2001)	0.22		0.21	
Wealth share, Top 0.1%	SCF (2001)	0.10		0.11	
Wealth share, Top 1%	SCF (2001)	0.32		0.32	
Earnings share, Top 0.1%	SCF (2001)	6.1		6.1	
Parent-child earnings correlation	PSID	0.33		0.33	
Earnings mean-median ratio	SCF (2001)	1.71		1.71	
Bequest-to-wealth ratio	SCF (2001)	2.00%		2.07%	
Estate tax revenues (as % of GDP)	Gale, Hines and Slemrod (2001)	0.33%		0.33%	
% of estates that are subject to tax	Gale, Hines and Slemrod (2001)	2.00%		2.00%	
		From Grandparents	Parents	From Grandparents	Parents
Flow of Transfers to under 30	SCF (1989 - 2022)	39%	61%	39%	61%
Sources of Transfers to under 30	SCF (1989 - 2022)	52%	48%	52%	48%

Note. "Flow of Transfers" is the value (flow) share received by under-30 heirs; "Sources of Transfers" is the count share.

The calibrated model reproduces the U.S. earnings and wealth distributions of the 1990s and, without further targeting, their life-cycle profiles and the rise of returns with wealth. The next section puts it to

Table 10: Macroeconomic Variables

Parameter	Value	Parameter	Value
Capital	310.1%	Tax revenues	
Government debt	55.0%	- Labour income	16.97%
		- Capital income	4.48%
Consumption	67.3%	- Estate	0.33%
Investment	14.7%	Pension System	
Government Consumption	18.0%	Contribution rate (in %)	7.65%

Note. All variables as a share of GDP unless indicated otherwise.

work, isolating what the G-G link does and how estate-tax reform reshapes the timing, persistence, and concentration of wealth.

IV. Results

This section presents the main results of the paper: the G-G link reshapes the *timing* of wealth accumulation and the *persistence* of wealth across generations, and is worth a substantial amount ex ante to a small group of heirs, all while leaving cross-sectional concentration almost unchanged. It then examines estate-tax reform, showing how the link shapes the aggregate, distributional, and welfare consequences of relaxing the tax, and how the tax's two instruments, the rate and the exemption, act on concentration and persistence through different margins.

A. The Grandparent–Grandchild Link

Early Inheritance and High-Return Access The defining feature of the G-G link is its *timing*. High returns in the model must be *bought*: a household earns the excess return only after paying the fixed capability cost κ_e = \$1.00 million, which the calibrated earnings and progressive labour income tax system puts beyond the reach of all but a small group of exceptional earners in a single model period.[33] Grandparent transfers arrive early (at ages 25 to 40); therefore, apart from those few self-made earners who clear the cost on their own, it is the link that buys most young households early high-return access, years before they could otherwise afford it. Also recall that return heterogeneity operates through the household's financial-capability state: a capable household earns a first excess return, plus a second excess return on wealth above \$5.0 million, while a household that has not acquired capability earns only the safe return.

Since high returns must be bought, the households that earn them are those that could afford the cost, from their own saving or from an inheritance. Table 11 shows the decomposition of the source of acquired high-return capability. "Self-made" households paid the fixed cost κ_e out of own saving; the remaining columns classify each capable household by the receipt that carried it over the threshold, from a grandparent,

[33] In the absence of inheritance, self-funding the capability cost in a single period is out of reach for all but the very top earners: the median entrant (in the third productivity state, earning about \$53,000 a year) optimally saves only tens of thousands of dollars, and even a household in the fourth state (about \$330,000 a year) saves under \$40,000, far short of the \$1.00 million cost. Only the top-earnings state (z_5, about \$2.3 million a year) saves enough to clear the threshold immediately (see Appendix D).

a parent, or both. Inheritance is the gateway: of the top 1 percent, only 22 percent acquired access on their own and 78 percent through an inheritance (about 10 percent via a grandparent and 67 via a parent), and measured at the moment of acquisition 90 percent of all capable households were carried over the threshold by a receipt rather than by own purchase.

Table 11: Source of High-Return Access at the Top

Group	Capable (%)	Source of capability (% of capable)			
		Self-made	Grandparent	Parent	Both
All capable households	5.4	9.8	7.8	81.0	1.4
Top 10%	36.6	12.0	8.8	78.0	1.2
Top 1%	100.0	21.8	9.9	67.2	1.1

Note. These four categories are mutually exclusive; "Both" denotes households carried across by grandparent *and* parent transfers, not by a mix of saving and inheritance.

The large parent share is dominated by accidental bequests that arrive late in life and, as the persistence results below show, do not sustain a dynasty; what matters is *when* access arrives. The grandparent channel is small in the cross-section, about a tenth of top-1-percent access, since a living and wealthy grandparent is rare, yet it is the one channel that delivers access *early*: over the overlap window (ages 30–40) capability is acquired predominantly through inheritance, the self-made labour-saving route contributing only a small component from exceptional top earners until mid-life, so that at age 30 the link raises capability access from 0.79 to 1.17 percent almost entirely through the grandparent channel. The full acquisition profile by age is shown in Table 12.[34]

Table 12: Capability Access by Age, by Source of Transfer (%)

Age	With G-G Link				Without G-G Link			
	Grandparent	Parent	Self-made	Total	Grandparent	Parent	Self-made	Total
25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
30	0.34	0.55	0.24	1.17	0.00	0.50	0.22	0.79
35	0.48	1.45	0.35	2.36	0.00	1.34	0.35	1.78
40	0.50	2.60	0.35	3.54	0.00	2.44	0.35	2.89
45	0.50	4.15	0.36	5.10	0.00	3.96	0.36	4.42
55	0.50	7.57	0.89	9.06	0.00	7.27	0.74	8.11
65	0.50	8.13	0.98	9.71	0.00	8.56	0.88	9.54

Note. A small dual-receipt category (households carried over the threshold by both a grandparent and a parent in the same period, at most 0.1 percent) is omitted, so the source columns need not sum to Total. Without the G-G link, the grandparent channel is absent.

Dynastic Persistence and the Value of the Link The link matters most not for cross-sectional inequality but for intergenerational *persistence*: the two are distinct objects, and the link moves persistence while leaving the cross-section essentially unchanged (as I show below). Table 13, panel A, traces where a household ends up as a function of its ancestors' wealth, with and without the link. With the link active, a grandchild

[34] "Self-made" here is the labour-saving route; the model abstracts from entrepreneurship, through which the empirically salient young self-made wealthy create high returns directly rather than saving into them (see the concluding remarks).

of a grandparent in the top wealth bins reaches the top 1 percent of the wealth distribution with probability 62.2 percent, roughly 60 times the unconditional rate.[35] Removing the link collapses that probability to 1.6 percent, no higher than the 2.6 percent rate for a grandchild who receives no grandparent transfer at all: without the transfer, a wealthy grandparent confers no advantage, and essentially the entire G-G wealth advantage runs through this single channel.[36] By contrast, the link barely touches the parent-to-child chain: a child of a top-quintile parent reaches the top 1 percent with probability 4.8 percent with the link and 4.6 percent without it. The link therefore does not re-route persistence that would otherwise flow through the parent; it creates a generation-skipping channel that does not exist in its absence. In short, the G-G link creates dynastic persistence directly from grandparent to grandchild, and it does so through the transfer itself rather than through inherited earning ability: removing the transfer channel collapses G-G persistence to the no-transfer baseline, since without a transfer a wealthy grandparent leaves no imprint on where the grandchild ends up. What sustains a dynasty is thus the inheritance of *access* to high returns, not of earning *ability*; the self-made rich are real, but it is their descendants, launched by an early transfer over the capability threshold, who persist at the top.

Table 13: Dynastic Persistence and the Heirs of Wealthy Grandparents

Panel A. Probability of reaching the top 1% by age 65

Transmission channel	With link	Without link
Grandparent → grandchild (grandparent in top wealth bins)	62.2%	1.6%
Parent → child (parent in top wealth quintile)	4.8%	4.6%

Panel B. Heirs by grandparent wealth

Grandparent's wealth (economy-wide)	Share of newborns (%)	W_{30} (\$M)	W_{65} (\$M)	Pr(1%) by 65	Newborn CEV (%)
With G-G link					
Top bins ($> 2\chi_b \approx$\$12M)	0.1	4.34	12.82	62.2	+291.3
Below top bins	21.0	0.13	0.86	1.3	+ 4.0
No living grandparent	78.9	0.07	1.14	2.5	—
Without G-G link					
Top bins ($> 2\chi_b \approx$\$12M)	0.1	0.06	0.88	1.6	− 6.2
Below top bins	21.0	0.05	0.79	1.1	− 7.6
No living grandparent	78.9	0.08	1.19	2.6	—

Note. Panel A conditions on an ancestor's wealth, with the link on and switched off ($\kappa_j = 1$); the unconditional probability of reaching the top 1 percent is 1 percent, and the grandparent figure is untargeted. Panel B: W_{30} and W_{65} are mean wealth (2019 US\$ millions) at ages 30 and 65, Pr(1%) is each group's share in the top 1 percent at age 65, and the newborn CEV is relative to the no-transfer group; the share of newborns is identical across the two blocks.

[35] Throughout, *persistence* conditions on grandparents in the top two wealth bins ($> 2\chi_b \approx$ \$12 million) and scores the grandchild against the top 1 percent of all households (the same measure as the cohort accounting in Table 13 panel B). This conditioning is deliberately narrower than the top-quintile cut used for the parent-to-child chain below: the direct GST operates only among wealthy grandparents, for whom skipping clears its fixed cost, so a broader conditioning would fold in grandparents who never skip and understate the channel where it is active. Appendix E.1 reports both channels under a common conditioning.

[36] Without a transfer the wealthy-grandparent lineage is, if anything, slightly less likely to reach the top than a household with no living grandparent (1.6 versus 2.6 percent), since the dynasty's wealth is still held by the grandparent and reaches the grandchild only later, through the parent.

The link thus governs which lineages occupy the top of the distribution across generations. It also gives a structural account of a pattern documented in the data, that grandparent wealth predicts grandchild wealth beyond what parental wealth alone would imply (Adermon, Lindahl and Waldenström, 2018; Pfeffer and Killewald, 2018): in the model that excess grandparent effect *is* the direct GST. Since nothing in the calibration disciplines the G-G relationship, the magnitude of this persistence is a validation of the mechanism rather than an artefact of it.[37]

The persistence documented above also survives an external check. Table 14 reports the same-age wealth rank correlation (a Spearman correlation, with each generation ranked within its own age cohort) between generations by the age at which wealth is measured, in the benchmark (heirs anticipate transfers) and under a no-anticipation counterfactual. Although the model targets no mobility moment, a dynasty simulation that tracks continuous wealth reproduces the empirical G-G wealth rank correlation in the Panel Study of Income Dynamics, about 0.28, against 0.39 for the parent–child link,[38] and matches conventional mobility estimates at the ages the literature measures. It concentrates the grandparent effect among the heirs of wealthy grandparents rather than spreading it as a broad gradient. This match holds even though the model's grandparent transfers are more front-loaded over the life cycle than in the data: the fixed intergenerational age gap and finite lifespan confine transfers to ages 25 to 40, whereas empirically a declining minority continue into their 40s. That timing simplification is immaterial to persistence, however, which is carried by the early, capability-clearing transfer rather than by the late grandparent bequests the model omits. The gap between the two rows is itself an anticipation channel: a young heir who receives a grandparent transfer while its parents are still alive rationally anticipates a later parental bequest, and so consumes more and saves less than a household with no living ancestors.

Table 14: Wealth rank correlation by measurement age

Measurement age	35	40	45	50	55	65
Parent–child						
Benchmark (bequests anticipated)	0.05	0.15	0.28	0.40	0.45	0.49
No anticipation	0.14	0.24	0.37	0.50	0.54	0.56
Grandparent–grandchild						
Benchmark (bequests anticipated)	0.06	0.10	0.15	0.20	0.22	0.23
No anticipation	0.10	0.14	0.20	0.26	0.28	0.29

Note. Entries are Spearman wealth rank correlations between two generations, each generation ranked within its own age cohort, at the measurement age in the column heading. "Benchmark" is the calibrated model, in which heirs anticipate future bequests; "No anticipation" shuts that channel off. Empirically, parent–child wealth rank correlations are about 0.3–0.4, typically measured in the 30s and 40s (Charles and Hurst, 2003; Adermon, Lindahl and Waldenström, 2018); the corresponding grandparent–grandchild correlation in the PSID is about 0.28.

The analysis so far has concerned the wealth distribution, and I now turn to welfare implications, measured as the consumption-equivalent variation (CEV) for a household entering the economy behind the

[37] The model's grandparent state is coarsely discretised, so the relevant object is the conditional attainment reported here rather than a smooth population-wide G-G wealth correlation; consistent with the welfare results, the persistence is concentrated in the small group that receives a large, capability-clearing transfer.

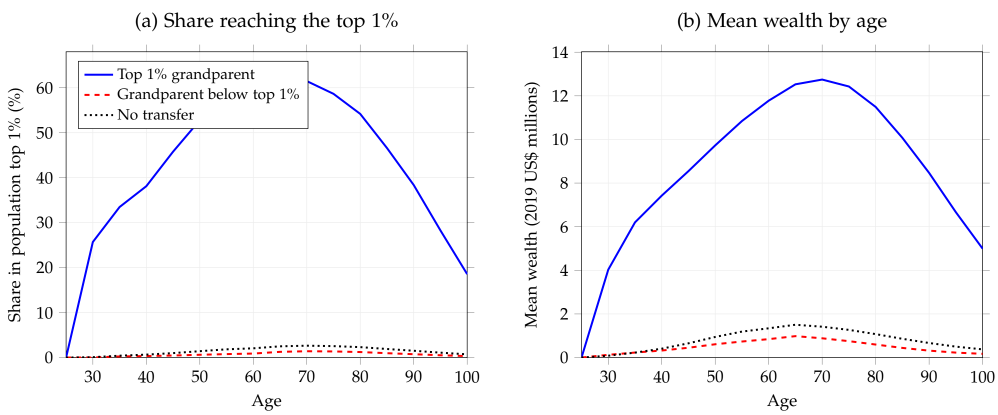
[38] With earnings transmission alone (bequests switched off), the parent–child correlation is 0.17 (benchmark) and 0.27 (no anticipation). As an empirical benchmark, the parent–child wealth rank correlation is about 0.3–0.4 (Charles and Hurst, 2003; Adermon, Lindahl and Waldenström, 2018), typically measured for children in their 30s and 40s.

veil of ignorance, before its family background, productivity, and inheritances are realised.[39]

Consider first the value of the G-G link. In the aggregate it is small: a newborn behind the veil would give up only about 1 percent of lifetime consumption to retain access to it. But this average is the net of two starkly unequal effects. Being born to a wealthy, living grandparent is the equivalent of winning a life-changing lottery: the roughly 0.1 percent of newborns who hold one end up worth nearly 300 percent of lifetime consumption more than a newborn with no grandparent transfer. The 79 percent with no living grandparent, by contrast, are if anything slightly worse off with the link than without it, since its extra capital accumulation lowers the return that ordinary savers earn. The link is thus almost invisible in the aggregate, yet decisive for the handful of lineages it launches.

The welfare value of the G-G link, and the timing effects documented above, are concentrated in a very small group. To make this concrete, I partition newborns by the wealth of their grandparent, measured against the economy-wide wealth distribution, and follow each group over the life cycle; Figure 2 plots each group's trajectory and Table 13, panel B, reports the corresponding levels.

Figure 2: Grandchild Outcomes over the Life Cycle, by Grandparent Wealth



Note. Newborns are grouped by grandparent wealth, as in the text. Panel (a) plots each group's share in the population top 1% of wealth by age; panel (b) plots mean wealth in levels. The decline after age 65 reflects ordinary retirement decumulation, common to all groups.

About 0.1 percent of newborns have a living grandparent in the top wealth bins of the model, grandparents holding more than roughly twice the estate-tax exemption ($2\chi_b \approx$ \$11.9 million).[40] For these children the link is decisive: they hold \$4.3 million by age 30, and 62.2 percent reach the top 1 percent of the overall wealth distribution by age 65 (against only 1.6 percent of the same group without the link). Behind the veil, having such a grandparent is worth 291 percent of lifetime consumption relative to a household with no grandparent transfer, the "lottery" component of the link's ex-ante value, enormous but realised by very few.[41]

[39] The newborn CEV is the uniform percentage change in period consumption, in every future state, that makes a household indifferent between being born into the two economies. It is the natural ex-ante measure here since the objects of interest (access to the G-G link, and the incidence of estate taxation) are realised only after birth. The closed form is $\lambda = (\mathcal{W}_a/\mathcal{W}_b)^{1/(1-\sigma)} - 1$, where $\mathcal{W}$ is the ex-ante value of being born into an economy; Appendix D.4 gives the formal definition.

[40] Only a small fraction of newborns have such a grandparent since a GST requires the grandparent to be both wealthy and still alive when the grandchild enters the economy; most grandparents have died by then, so 79 percent of newborns receive no grandparent transfer at all.

[41] This group figure is an upper bound on the welfare gain: since heirs anticipate transfers perfectly and face no liquidity constraints,

It is worth noting that the effect is specific to wealthy grandparents, not to grandparent transfers in general. Newborns who receive a grandparent transfer from outside the top wealth bins (about 21 percent of the cohort) are barely distinguishable from those who receive none, and in fact end up slightly poorer in wealth by age 65 (\$0.86 million versus \$1.14 million) and no more likely to reach the top 1 percent. A transfer of this size is too small to afford the high-return access cost, and less wealthy families tend to transmit lower earnings ability, so it raises early consumption (a positive but small welfare gain of 4.0 percent), without launching the heir onto a high-accumulation path.

The model's mechanics produce this threshold directly. A grandchild acquires high-return capability only if its transfer is large enough to afford the cost $\kappa_e = \$1.00$ million early in life; for the amount actually reaching the grandchild to cover this (after the estate tax and the child–grandchild division), the grandparent's estate must exceed roughly twice the exemption ($2\chi_b \approx \$11.9$ million). This is why the group the link launches to the top coincides with the wealthiest grandparents and why grandparent wealth just below does little. It is therefore not inheritance in itself but a *sufficiently large* early inheritance that reshapes an heir's lifetime wealth, which is why the G-G link is so valuable ex ante yet nearly invisible in the aggregate: its force is concentrated in the small group whose inheritance clears the capability threshold.

The Link and Cross-Sectional Concentration What does the G-G link do to the cross-sectional wealth distribution? Is it needed to reproduce the upper tail at all? Table 15 answers both by reporting three economies: the benchmark (link on); a no-link economy that sets $\kappa_j = 1$ while holding every parameter at its benchmark value, isolating the link's effect; and a no-link economy re-calibrated to the same external targets.

Holding parameters fixed, removing the link lowers aggregate capital: the capital-output ratio falls from 3.10 to 3.03 and the pre-tax return rises from 7.0 to 7.2 percent as capital becomes scarcer. Its effect on the cross-sectional distribution, by contrast, is negligible and, if anything, mildly equalising: the aggregate Gini rises only slightly (from 0.766 to 0.774) and the top 1 percent share barely moves (from 31.7 to 32.3 percent). The link matters for how much capital the economy accumulates, not for how unequal it is. Its force falls on the persistence of wealth across generations, which collapses from 62.2 to 1.6 percent, not on inequality at a point in time.

Table 15: The G-G Link: A Persistence Device, Not Needed to Fit the Cross-Section

	Aggregate		Wealth shares (%)					Persistence
Economy	K/Y	r^*	Gini	Top 1	95–99	99–99.9	Top 0.1	GP→GC
Benchmark (link on)	3.10	7.00	0.766	31.7	20.9	21.0	10.7	62.2
No-link, parameters fixed	3.03	7.20	0.774	32.3	20.9	21.5	10.8	1.6
No-link, re-calibrated	3.10	7.00	0.776	32.2	21.1	21.5	10.7	1.6

Note. r^* and the wealth shares are in percent. GP→GC is the probability that the grandchild of a top-bins grandparent reaches the top 1 percent of wealth by age 65, an untargeted moment. The re-calibrated no-link economy hits the same external targets with nearly unchanged parameters (see text).

they front-load consumption more than they likely would under illiquid inherited assets or uncertainty about the size and timing of bequests. Either friction would temper the magnitude while leaving the pattern intact.

The distributional effect of the link is concentrated early in life: at age 30 the wealth Gini is 0.811 with the link versus 0.784 without it, a gap that closes, and indeed reverses, by the mid-30s as the rest of the cohort receives parental bequests, leaving the no-link economy marginally more unequal in the aggregate. The link reshapes the *timing* of inequality far more than its long-run level.

Furthermore, the link is not needed to *fit* the cross-section. Re-calibrating a no-link economy to the same external targets reproduces the capital-output ratio, the Gini, and the entire top-share profile, and it does so with almost unchanged parameters: the discount factor rises by about 0.3 percent, the estate-tax exemption falls by 4 percent, and the return dispersion by 2 percent, with the excess-return schedule left untouched. What supplies the concentration is return heterogeneity, not the link: as Table 8 earlier shows, an otherwise-identical economy with a single interest rate collapses the top 1 percent share to 18 percent. The division of labour is clean: return heterogeneity concentrates wealth at a point in time, while the link governs *whose* descendants hold it across generations. This is the paper's central finding: the link matters since it operates on the timing of wealth and the fortunes of a very small group, not on the breadth of the distribution.

B. The Generation-Skipping Transfer Tax

Subsection A established that the G-G link is what generates dynastic persistence at the top: shutting the channel collapses G-G transmission almost entirely. But what happens when the *tax* that governs the channel, rather than the channel itself, is changed? The link operates through two forces. The first is a warm-glow motive to leave wealth directly to grandchildren. The second is a tax-planning *tilt*: when a grandparent's child is already wealthy enough to owe estate tax, wealth routed through the child is taxed twice before reaching the grandchild, so a direct transfer skips a round of estate taxation (the dependence of κ_j on the child's taxable status d_c). The GST tax caps this saving above its exemption. Table 16 isolates each force by switching it off in turn, re-solving the benchmark in general equilibrium.

Table 16: Decomposing the Link: Tax Planning and the Generation-Skipping Tax

Economy	GG flow	GG source	Top 1%	Gini	Persistence
Benchmark	0.394	0.523	31.7	0.766	62.2
Tax-planning tilt off	0.382	0.523	31.5	0.764	53.6
GST tax off	0.398	0.522	31.8	0.766	64.7
Tilt off & GST off	0.386	0.522	31.7	0.765	59.0
No G-G link ($\kappa = 1$)	0.000	0.000	32.3	0.774	1.6

Note. "GG flow" is the grandparent share of transfer *value* received by the under-30, and "GG source" the grandparent share of inheritance *count*; persistence is defined in the text. Each row switches off one force at the benchmark and re-solves in general equilibrium: tilt off sets $\kappa_0 = \kappa_1$, GST off sets $\tau_{gst} = 0$, and no link sets $\kappa_0 = \kappa_1 = 1$.

Three results follow. First, the tax-planning tilt is a quantitatively meaningful driver of skipping: removing it (setting $\kappa_0 = \kappa_1$, so the division of the estate no longer responds to the child's tax status, while warm-glow transfers remain) lowers the generation-skipping flow and cuts G-G persistence from 62.2 to 53.6 percent. Second, the GST tax is what *regulates* this margin: removing the GST cap alone, leaving the estate tax and the planning motive in place, lets tax-motivated skipping expand and *raises* persistence to 64.7 percent. The GST does not create the generation-skipping channel; it restrains how far dynasties exploit it, exactly as U.S. law intends. Third, by contrast, shutting the link entirely collapses persistence to 1.6 percent: the mechanism, not the tax, is what generates dynastic persistence, and the transfer taxes only modulate it.

Across the tax variations, the top 1 percent wealth share stays within a narrow 31.5 to 31.8 percent, while persistence swings from 53.6 to 64.7 percent. The generation-skipping margin (and the GST tax that governs it) operates on *persistence*, on how long dynasties hold top wealth across generations, not on the cross-sectional concentration of that wealth. This is the counterpart to the estate-tax reform studied next, which moves concentration: the two instruments act on genuinely different margins, rather than being two labels for the same lever.

C. Estate-Tax Reform

I now turn to estate-tax reform. Whether the estate tax effectively restrains the concentration of wealth has long been debated, among policymakers and in the academic literature. The U.S. experience since 2000 offers a natural reference point: a sequence of reforms, chiefly the 2001 and 2017 federal tax acts, greatly increased the estate-tax exemption, from about \$1 million in the 1990s to \$11.4 million by 2019 (\$12.06 million by 2022), while the top statutory rate fell from 55 to 40 percent and the average effective rate from about 18 to 16 percent.

The estate tax has two instruments in the model, the rate τ_b and the exemption χ_b, with an estate b owing $T(b) = \max\{0, \tau_b(b - \chi_b)\}$. The two act on households differently: raising the exemption both lifts estates out of the tax net and lightens the burden on those still in it, whereas cutting the rate only lowers the bill on estates already taxed. That difference maps onto the two margins of the previous subsection, concentration and persistence, and it is the organising theme of what follows. Table 17 shows the policy changes considered in this section, alongside the U.S. data that motivate them.

Table 17: Estate Tax Policy: U.S. Data and Model Counterfactuals

	Data			Model	
	1990-2000	2019		Baseline	Counterfactuals
Effective Estate Tax Rate	18%	16%	τ_b	0.106	0.095
Estate Exemption Threshold	\$1.04	\$11.40	χ_b	\$5.95	\$11.89
Highest Tax Bracket Threshold	\$6.16	\$12.40			

Note. All dollar figures in 2019 US millions.
Source. U.S. Internal Revenue Service.

Since I abstract from the progressivity of estate taxes, the model's τ_b and χ_b are best understood as the flat-tax analog of this schedule: they are calibrated jointly to reproduce estate-tax revenue as a share of GDP and the fraction of estates taxed (about 2 percent), following De Nardi and Yang (2016), rather than as the statutory rate and exemption. My main counterfactual mirrors the observed 1990s-to-2019 reform, doubling the exemption threshold and lowering τ_b in the same proportion as the decline in the average effective rate ($18 \rightarrow 16$ percent). Although the model abstracts from progressivity, its calibrated exemption ($\chi_b = \$5.95$

million) sits close to the highest statutory bracket threshold of the 1990s ($6.16 million), the level above which the top rate applied; doubling it to $11.89 million is therefore a reasonable counterfactual, landing near the 2019 exemption ($11.40 million). Since the model's GST mirrors the estate tax by construction ($\tau_{gst} = \tau_b$ and $\chi_{gst} = \chi_b$, as under U.S. law), this experiment relaxes the estate tax and its generation-skipping counterpart together: it is a reform of the entire transfer-tax system, reaching the skipping margin at the center of this paper, rather than of the estate tax alone. The resulting flat rate, $\tau_b = 10.6$ percent, lies below the statutory schedule because the model's tax base is deduction-free, not because the tax is lenient; measured as the IRS reports effective rates, the model's burden is about 5 percent, within the range observed in the data.[42]

I consider revenue-neutral reforms, adjusting the capital income tax to keep the government budget balanced and solving for the new stationary distribution. Table 18 reports the steady-state effects of the reform and of its two components separately, all under the capital-income-tax closure. The table is organised around the paper's central policy finding: the rate τ_b governs cross-sectional concentration, while the exemption χ_b governs dynastic persistence. The last two columns make this visible by pairing the familiar concentration measure (the top 1 percent wealth share) with a mobility measure, the probability that a grandchild of a top-bin grandparent reaches the top 1 percent by age 65, which is 62.2 percent at the benchmark.

Table 18: Estate-Tax Reforms: Concentration versus Persistence

Reform	τ_b	χ_b ($M)	τ_a (%)	K/Y	Top 1% (%)	Gini	Persistence (%)	Newborn CEV (%)
Benchmark*	0.106	5.95	15.2	3.10	31.7	0.766	62.2	—
Cut rate only	0.095	5.95	15.4	3.10	31.9	0.767	62.3	−0.20
Double exemption only	0.106	11.89	15.7	3.08	32.6	0.770	85.4	−0.57
Both (main reform)	0.095	11.89	15.7	3.09	32.7	0.771	85.3	−0.61
Full repeal	0.000	5.95	16.4	3.10	34.2	0.775	61.0	−0.97

Note. Steady-state effects of revenue-neutral estate-tax reforms under the capital-income-tax closure; * marks the benchmark. *Persistence* (untargeted) and *newborn CEV* are defined in the text; χ_b is in 2019 US$ millions.

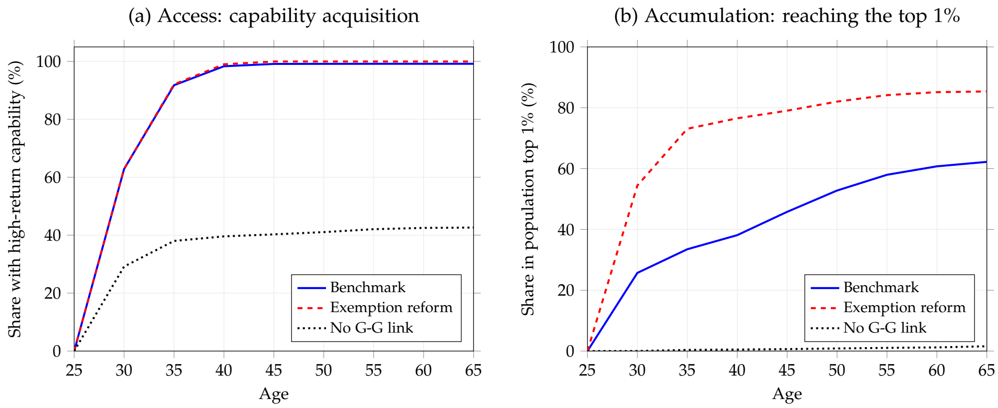
The aggregate effects are small throughout. Since the estate tax raises little revenue (0.33 percent of GDP), the revenue-neutral increase in the capital income tax needed to offset each reform (τ_a rises by at most about 1.2 percentage points) roughly cancels the direct effect of the cut on saving: the capital-output ratio is essentially flat, edging slightly down in every case, and the aggregate safe return $\underline{r}$ declines modestly. Relaxing the estate tax thus buys no aggregate efficiency dividend; its consequences are entirely distributional and dynastic.

Since a reform does not change who has a wealthy ancestor, it leaves the early acquisition of high-return

[42] Matching estate-tax revenue with a deduction-free base (wealth at death, with no marital, charitable, or other deductions) necessarily calls for a flat rate below the statutory schedule, since the model's tax base is broader than the taxable estate reported to the IRS. In IRS Statistics of Income data the effective rate on the gross estate is hump-shaped, peaking near 17 percent for mid-sized estates and falling toward 13 percent for the largest, as deductions and sheltering claim up to two-thirds of the gross estate at the top. This is the standard mapping from a graduated statutory schedule to a tractable flat tax: De Nardi and Yang (2016), for example, calibrate an effective rate of 21 percent against a 55 percent top statutory rate.

capability essentially unchanged; its consequences operate later in the life cycle, through accumulation rather than access, as Figure 3 makes explicit.

Figure 3: How the Exemption Reform Raises Persistence: Access versus Accumulation



Note. Grandchildren of top-bin grandparents (grandparents above roughly $2\chi_b$), followed over the life cycle. Panel (a) plots the share holding high-return capability (access) by age; panel (b) plots the share reaching the population top 1 percent (the accumulation outcome that defines persistence).

Concentration. Every relaxation raises wealth concentration, and the ranking across reforms identifies the rate as the operative instrument. Cutting the rate alone barely moves the top 1 percent share (from 31.7 to 31.9 percent), since the proportional cut considered here is modest; but repealing the tax, setting the rate to zero, raises it by 2.5 percentage points (to 34.2), the largest concentration effect in the table. Raising the exemption occupies an intermediate position (a 0.9-point rise), as estates between the old and new thresholds drop out of the tax net. The extreme upper tail responds to the rate: the very largest estates remain taxable until the rate itself falls, so it is rate relief, and ultimately repeal, that most enlarges the top of the distribution.[43]

Persistence. The mobility column, on the other hand, tells a different story: cutting the rate leaves dynastic persistence essentially unchanged (62.2 to 62.3 percent), and even full repeal leaves it essentially flat (61.0 percent). Doubling the exemption, by contrast, sends persistence from 62.2 to 85.4 percent. The mechanism is generation-skipping: a higher exemption widens the window through which wealth passes untaxed to grandchildren, so far more grandchildren of the wealthy clear the top-1-percent threshold, whereas cutting the rate lowers the bill on estates already in the net without changing who transfers what to whom. *The rate is a concentration instrument; the exemption is a persistence instrument.* Repeal produces the highest top-1-percent share yet delivers far less persistence than a mere doubling of the exemption: concentration at a moment in time and the durability of dynastic advantage are distinct objects, governed by distinct margins, and only the exemption, through the GST tax it anchors, governs the second.[44]

Welfare. Every reform lowers newborn welfare (last column): behind the veil of ignorance a household

[43]This repeal effect is larger than in De Nardi and Yang (2016), who find the top 1 percent share rises by 1.7 percentage points following complete abolition of the estate tax; here it rises by 2.5 points. The difference reflects the different engines of concentration: De Nardi and Yang (2016) target both the earnings and wealth distributions, whereas here earnings heterogeneity alone cannot generate the observed concentration, so return heterogeneity, and the inheritance that buys access to it, carries more of the load. It is also consistent with Nekoei and Seim (2023), where heterogeneous returns lead inheritances to widen long-run wealth inequality.

[44]Repealing the tax removes the estate-tax *motive* to skip generations, so dynastic wealth passes down the ordinary parent–child line and the top accumulates as a stock; raising the exemption instead keeps the skip motive alive while widening the untaxed transfer window, routing wealth directly and early to grandchildren.

would sacrifice between 0.2 and 1.0 percent of lifetime consumption to keep the benchmark rather than move to a relaxed regime, the loss rising with the size of the cut to 0.61 percent for the main reform and 0.97 percent for full repeal. This is a matter of incidence. Since the reforms are revenue-neutral, the forgone estate revenue is recovered through a higher capital income tax that falls on all wealthy savers in every period, whereas only the small group whose estates actually reach the estate tax benefits from the cut. The welfare loss therefore rises with wealth and peaks in middle age, when households hold the most taxable capital: relaxing the estate tax redistributes from the moderately wealthy, who bear the replacement tax throughout their lives, to the tiny ultra-wealthy who escape estate taxation, a trade that a newborn, not knowing which group it will join, evaluates as a loss.[45]

Finally, the G-G link shapes *where* within the top the reform's gains land. Measuring each economy's response against its own benchmark, the link amplifies the effect in the upper-middle of the top percentile: following the main reform the 99–99.9 percentile share rises by +0.8 percentage points with the link but only +0.4 without it, twice as large, and the same amplification carries over to full repeal (+1.3 with the link versus +0.9 without). The aggregate Gini and the top 0.1 percent share, by contrast, move almost identically in the two economies (the main reform raises the Gini by +0.005 with the link and +0.003 without). The link thus concentrates the reform's effect where its generation-skipping channel operates, while leaving the broad Gini and the top 0.1 percent share little changed.

The Estate Tax and Capital A revenue-neutral estate-tax cut slightly lowers aggregate capital, which may seem to sit uneasily with the usual argument that estate taxation discourages saving. Both hold at once. In partial equilibrium the cut does raise capital (the estate tax lowers the after-tax value of bequests), but the higher capital income tax that restores revenue discourages saving by about as much, so in general equilibrium the capital-output ratio ends at or just below its benchmark. Within the model the estate tax is thus a *more efficient* way to raise revenue from the top than the capital income tax that would replace it. The three-step partial-equilibrium, revenue-neutral, and full-GE decomposition is reported in Table 19.[46]

Table 19: Decomposition of the Aggregate Effect of Estate Tax Relaxation

	Partial equilibrium		Revenue-neutral		Full GE	
Reform	K/Y	τ_a	K/Y	τ_a	K/Y	τ_a
B1 (main relaxation)	3.128	15.2	3.094	15.8	3.086	15.7
Full repeal	3.177	15.2	3.107	16.4	3.101	16.4

Note. Benchmark $K/Y = 3.10$, $\tau_a = 15.2$ percent; the partial-equilibrium, revenue-neutral, and full-GE columns are defined in the text above.

Transition Dynamics The steady-state comparisons above describe where the economy eventually settles;

[45] These welfare signs do not depend on how the lost revenue is recovered: replacing it with a labour income tax, in the robustness analysis below, leaves every sign unchanged and, if anything, deepens the losses. Tightening the GST tax is close to welfare-neutral (a CEV of about −0.01 percent), consistent with its negligible effect on concentration and reinforcing that the GST operates on a different margin from the estate tax.

[46] The reforms use the capital-income-tax closure; a labour-income-tax closure is examined for robustness below. Consistent with this, the partial-equilibrium concentration effect for the main reform (a 1.2 percentage-point rise in the top 1 percent share) is close to the labour-closure effect (1.2 points) and larger than the capital-closure effect (1.0 point), since the labour closure does not offset concentration through the capital tax.

the transition to that new steady state reveals a second dimension of the concentration-versus-persistence distinction, since the two margins move on very different time scales. Figure 4 traces the perfect-foresight transition after the exemption reform, the component that drives persistence.[47] The response is almost immediate: persistence rises to essentially its reform steady state within a single generation, reaching that level by year ten and holding there. Concentration at the extreme upper tail, by contrast, adjusts far more slowly: the top 0.1 percent wealth share, whose steady-state change is itself small (from 10.7 to 10.8 percent), drifts up only gradually and does not fully settle for roughly two centuries.

The asymmetry is mechanical, and it reinforces the steady-state finding. Persistence is a *flow* phenomenon: the moment the exemption is raised, the very next cohort of wealthy grandparents makes larger untaxed transfers, and their grandchildren clear the top-1-percent threshold within their own lifetimes: one transfer event suffices, so the persistence response is essentially complete within a generation. Concentration at the extreme upper tail is a *stock* phenomenon: enlarging the top 0.1 percent requires dynasties to compound their advantage across many generations, a process that plays out over a century or more. Estate-tax relaxation thus reshapes *who stays rich* almost at once, while reshaping *how concentrated* wealth is only slowly, a timing distinction that steady-state comparisons alone cannot reveal.[48]

The transition also sharpens the welfare accounting, but in the opposite direction to the concentration story. A newborn entering at the reform date, who lives through the entire adjustment, suffers a consumption-equivalent loss of only 0.17 percent, smaller than the 0.57 percent computed from the steady-state comparison alone. The cohorts alive during the transition inherit the pre-reform capital stock and the wages it supports, while the replacement capital income tax rises only gradually as the base rebuilds; the early cohorts therefore bear less of the reform's cost than a household born into the fully-adjusted reform steady state. Steady-state welfare comparisons thus overstate the near-term cost of relaxing the estate tax for the generations who actually live through the change, even as they understate its long-run concentration effect.

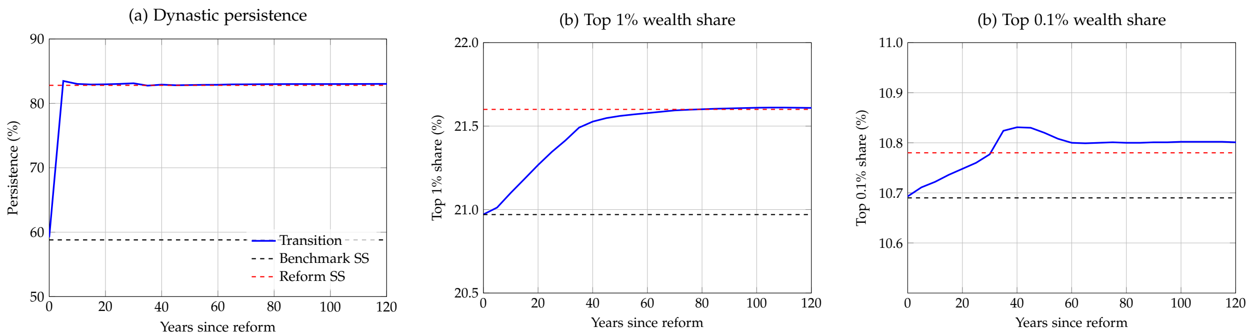
D. Robustness

I evaluate the sensitivity of the main findings across three dimensions: the method of government revenue replacement, the estate-tax schedule, and the capability process, and the main results remain robust in all cases (Table 20 and Table 21). First, re-balancing the government budget through the labour income tax rather than the capital income tax leaves all welfare signs identical, with losses slightly larger (-0.70 versus -0.61%) and concentration rising further (+1.2 versus +1.0 points), since the capital tax falls on the high-return households who gain most from the cut; the main-text estimates are thus conservative. Second, a revenue-neutral two-bracket progressive estate (and GST) tax schedule yields nearly identical results, with a slightly larger concentration effect. Finally, perturbing the capability process — making financial capability

[47] The transition is measured by perfect-foresight simulation, so the persistence levels it reports (about 59 percent at the benchmark rising to 83 percent) sit a few points below the stationary values in Table 18 (62.2 and 85.4 percent); the figure is read for the speed of adjustment rather than the precise level.

[48] The transition is solved on a fixed asset grid over a four-hundred-year horizon ($T = 80$ five-year periods). The aggregate factor prices and the persistence response converge over this window; the top 0.1 percent share, a slow-moving stock, does not fully reach its new steady state, consistent with the slow-tail property of this class of models. Dynastic persistence is a top-1-percent snapshot measured with simulation noise and is lightly median-filtered in the figure to remove two isolated spikes; beyond the first generation it carries residual noise from the still-adjusting cross-section.

Figure 4: Transition after the Estate-Tax Reform



Note. Perfect-foresight transition after the exemption reform (doubling χ_b), from the benchmark steady state toward the reform steady state. Panel (a) plots dynastic persistence; panels (b)–(c) the top-1% and top-0.1% wealth shares.

an impermanent rather than an absorbing state, or varying the return risk (σ_e) and acquisition cost (κ_e) — leaves both main outcomes unaffected: persistence stays between the high 50s and mid 60s, and concentration moves monotonically with σ_e. Together with the ability-independence check in Subsection A, these analyses confirm that persistence reflects the inheritance of high-return access, not a mechanical artefact of the setup.

Table 20: Robustness of the Estate-Tax Reform

	Capital-tax closure			Labour-tax closure			Progressive schedule	
	ΔTop 1%	ΔGini	CEV	ΔTop 1%	ΔGini	CEV	ΔTop 1%	ΔGini
Benchmark	31.7	0.766	—	31.7	0.766	—	31.1	0.764
Main reform	+1.0	+0.005	−0.61	+1.2	+0.006	−0.70	+1.9	+0.007
Full repeal	+2.5	+0.009	−0.97	+3.2	+0.012	−1.45	—	—

Note. The *Benchmark* row reports levels, the reform rows changes; CEV is the newborn consumption-equivalent variation (percent). The budget closures and the progressive schedule are defined in the text above; full repeal under the progressive schedule is not computed.

Table 21: Robustness of the Capability Mechanism

	Top 1%	Gini	$P(p{=}1)$	GP→GC persist.
Benchmark (δ=0, σ_e=0.146, κ_e=\$1.00M)	31.7	0.766	5.4	62.2
Capability loss $\delta = 0.10$	32.3	0.760	3.9	57.6
Capability loss $\delta = 0.25$	33.1	0.754	2.6	58.1
Return risk $\sigma_e = 0.11$	29.5	0.760	5.7	62.8
Return risk $\sigma_e = 0.18$	33.6	0.770	5.0	59.1
Acquisition cost κ_e = \$0.77M	32.6	0.789	6.4	62.2
Acquisition cost κ_e = \$1.23M	31.3	0.752	4.8	66.5

Note. Each row perturbs one feature of the benchmark general equilibrium. $P(p{=}1)$ is the stationary share of capable households (percent); δ is the per-period probability that capability is lost ($\delta = 0$ in the benchmark).

V. Concluding Remarks

In this paper I study why wealth persists at the very top of the distribution across generations, and how transfer taxation shapes that persistence. Motivated by evidence from the SCF that G-G transfers are large and arrive early, and that wealthy households hold high-return portfolios, I develop a general equilibrium life-cycle model with two novel features. Returns are generated by an acquired financial-capability state, so that a large early inheritance can switch on high returns years before a household could accumulate on its own. Dynasties also choose not only how much to bequeath but to whom: they may transfer directly to grandchildren to skip a round of estate taxation, a margin disciplined by the GST tax. The model is calibrated to match the size and timing of grandparent transfers in the data.

Three findings stand out. First, the G-G link is primarily a mechanism of dynastic *persistence* rather than of cross-sectional inequality: it governs which lineages remain at the top across generations while leaving the Gini coefficient almost unchanged. Second, its welfare value is highly concentrated, large for the small group of newborns with a wealthy, living grandparent and negligible for everyone else, since a sufficiently large early transfer buys *access* to high returns decades ahead of ordinary saving. Third, relaxing the estate tax raises wealth concentration with no aggregate efficiency dividend, and its two instruments act on different margins, the rate on concentration and the exemption, through the generation-skipping tax it anchors, on persistence; evaluated ex ante, every such relaxation lowers welfare, a conclusion robust to how the lost revenue is recovered.

The model can be extended in several directions. It represents the estate tax as a flat rate above an exemption; a fuller treatment of the deductions and sheltering that separate wealth at death from the taxable estate, which the flat rate absorbs in reduced form, is an important avenue. Returns are governed by a binary capability state acquired through a fixed cost or a large inheritance. Since that cost is met out of *labour* income, self-made high-return access emerges only in mid-life: the model reproduces the self-made *share* of top wealth but not the young, entrepreneurial self-made rich, for whom founder equity is itself the high-return asset. A natural extension adds an entrepreneurial channel, tying capability to entrepreneurial ability so that a large early transfer relaxes would-be entrepreneurs' borrowing constraints (Cagetti and De Nardi, 2006, 2009), letting self-made access arise young and sharpening the self-made-versus-inherited comparison. The model also abstracts from inter vivos transfers, though these are roughly 14 percent of intergenerational giving; incorporating the timing of giving is a natural companion to the generation-skipping margin.

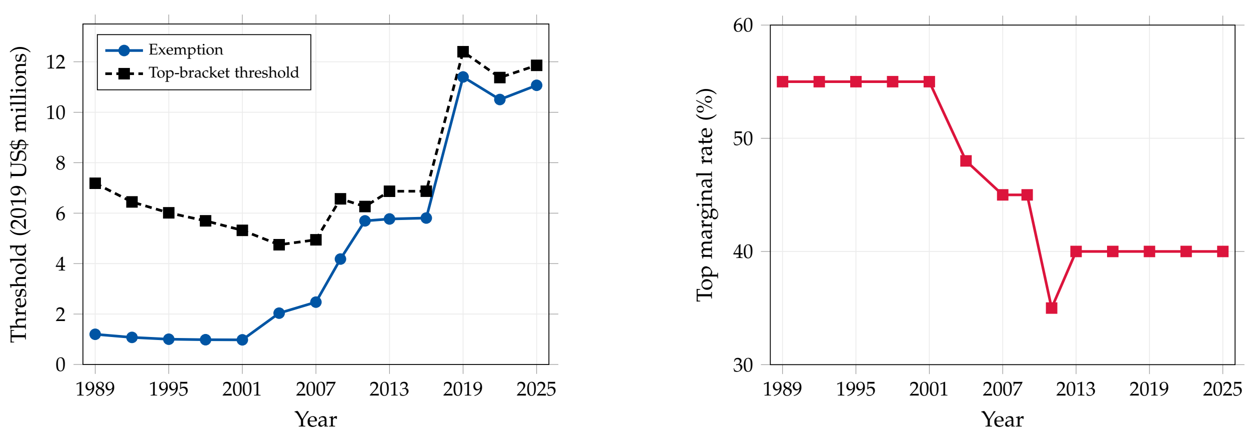
On the policy side, this model provides a laboratory for the joint design of the estate and GST taxes, which I leave for future research: the two instruments operate on different margins, the level and the destination of transfers within a dynasty; therefore, their optimal combination is an open question. More broadly, the generation-skipping margin studied here is the simplest form of a wider phenomenon, whereby the wealthy arrange transfers through legal vehicles in practice, most notably trusts, designed to move wealth across generations at a lower tax cost. Extending the framework to incorporate these instruments offers a promising avenue for understanding both the persistence of wealth at the top and the effectiveness of transfer taxation.

Supplemental Appendix

A. Estate Tax Policy in the United States

Estate tax policy in the U.S. has undergone various reforms and changes over the years depending on the political party in power, with reforms often amended by succeeding administrations.[49] The estate tax policy was steadily relaxed until 2009, accompanied by plans to repeal in 2010 then revert back to 2002 provisions in 2011. The plans were then reversed under the Obama Administration: the 2010 Act maintained the exemption level at $5 million with a top tax rate of 35 percent, which then increased to 40 percent under the 2012 Act, with a scheduled return to $3.5 million in 2018 that never took effect. The Tax Cuts and Jobs Act of 2017 temporarily doubled the exemption amount to $11 million from 2018, and most recently the One Big Beautiful Bill Act of 2025 instead made a higher exemption permanent, setting it at $15 million per individual from 2026 and indexing it to inflation thereafter.

Figure A1: Federal Estate Tax Policy, 1989–2025



(a) Exemption and top-bracket thresholds **(b)** Top marginal rate

Note. Panel (a): the federal estate-tax exemption threshold (solid) and the top-bracket threshold, the estate value at which the top marginal rate applies (dashed), both in constant 2019 dollars. Panel (b): the top statutory marginal rate. Both panels span 1989–2025;

A.1. Inheritance/Estate Tax at the State Level

While they are often used interchangeably in the literature, inheritance tax and estate tax are different in practice. The estate tax applies to the decedent's estate, while the inheritance tax applies to the beneficiaries who inherit the assets. As of 2022, there are 12 states[50] and the District of Columbia that impose estate taxes at the state level and 6 states[51] impose inheritance taxes at the state level. Maryland is unique in having

[49]Major legislation includes the Economic Growth and Tax Relief Reconciliation Act of 2001 and Jobs and Growth Tax Relief Reconciliation Act of 2003 during the Bush Administration, Tax Relief, Unemployment Insurance Reauthorization, and Job Creation Act of 2010 and American Taxpayer Relief Act of 2012 during the Obama Administration and Tax Cuts and Jobs Act 2017 during the Trump Administration.

[50]These include Connecticut, Hawaii, Illinois, Maine, Maryland, Massachusetts, Minnesota, New York, Oregon, Rhode Island, Vermont and Washington.

[51]These include Iowa, Kentucky, Maryland, Nebraska, New Jersey and Pennsylvania.

Table A1: Evolution of Federal Estate Tax Policy in the U.S.

Year	Tax Rate Range	Exemption Threshold	Exemption Threshold (2019 USD)	Highest Bracket Threshold	Highest Bracket Threshold (2019 USD)	Effective Estate Tax Rate
1989	(18 - 55)	0.600	1.197	3.600	7.185	0.180
1992	(18 - 55)	0.600	1.074	3.600	6.445	0.177
1995	(18 - 55)	0.600	1.002	3.600	6.013	0.179
1998	(18 - 55)	0.625	0.982	3.625	5.695	0.190
2001	(18 - 55)	0.675	0.977	3.675	5.318	0.181
2004	(18 - 48)	1.500	2.035	3.500	4.748	0.205
2007	(18 - 45)	2.000	2.472	4.000	4.943	0.199
2009	(18 - 45)	3.500	4.180	5.500	6.568	0.186
2011	(18 - 35)	5.000	5.694	5.500	6.263	0.142
2013	(18 - 40)	5.250	5.769	6.250	6.868	0.171
2016	(18 - 40)	5.450	5.805	6.450	6.870	0.173
2019	(18 - 40)	11.400	11.400	12.400	12.400	0.158
2022	(18 - 40)	12.060	10.505	13.060	11.376	—
2025	(18 - 40)	13.990	11.069	14.990	11.860	—

Note. The years shown are the SCF survey waves. Values in parentheses indicate the statutory marginal-rate range. Exemption and highest-bracket thresholds are reported in nominal dollars and in 2019 dollars. The effective estate-tax rate is from the IRS Statistics of Income estate-tax study, whose most recent year-of-death estimates are for 2019; it is therefore unavailable (—) for 2022 and 2025.
Source. Internal Revenue Service (statutory parameters; effective rates from IRS Statistics of Income).

both of these taxes, and the specific rules and exemptions can vary depending on the relationship between the decedent and the beneficiary.

A.2. The generation-skipping transfer (GST) tax

The generation-skipping transfer (GST) tax evolved in parallel with the estate tax. Table A2 lists the top statutory rate, the estate and GST exemptions, and the introduction of portability. Two institutional features justify the model's treatment of the GST as mirroring the estate tax ($\chi_{gst} = \chi_b$, $\tau_{gst} = \tau_b$): since 2004 the GST exemption has equalled the estate-tax exemption, and since 2011 the two have moved together dollar-for-dollar under a unified exclusion.

Table A2: Federal Estate and Generation-Skipping Transfer (GST) Tax Parameters

Year	Top Rate	Estate Exemption (nominal $M)	GST Exemption (nominal $M)	DSUE Portability
1998	55%	0.625	1.000	No
2001	55%	0.675	1.060	No
2004	48%	1.500	1.500	No
2006	46%	2.000	2.000	No
2009	45%	3.500	3.500	No
2011	35%	5.000	5.000	Yes
2013	40%	5.250	5.250	Yes
2019	40%	11.400	11.400	Yes
2022	40%	12.060	12.060	Yes
2025	40%	13.990	13.990	Yes

Note. Nominal exemption amounts. The GST exemption was set separately (and lower) before 2004, and set equal to it, with portability (DSUE), from 2011.
Source. Internal Revenue Service;

B. Survey of Consumer Finances

I utilise a total of 12 SCF survey datasets spanning from 1989 to 2022. Wealth is defined as the net worth of households, as in Kuhn, Schularick and Steins (2020) and Bricker et al. (2020). Regarding demographics, I categorise households into seven age groups. The youngest group comprises households 30 years or younger, while the oldest group includes households 81 years or older. The five intermediate groups are divided according to 10-year age intervals. In terms of intergenerational transfers, I analyse responses to the following questions in Section X: 'Inheritances and Charity'. These questions are included in each SCF dataset, as shown in Table B1 below.

Table B1: List of Questions in Section X, SCF

SCF Code	Questions
x5801	Have you ever received an inheritance, or been given substantial assets?
x5803	Was that an inheritance, a trust or something else?
x5804	What was its approximate value at the time it was received?
x5805	In what year was it received?
x5806	From whom was it received?
x5818	How much altogether were any others you have received?
x5819	Do you expect to receive a substantial inheritance in the future?

Source. Federal Reserve Board, 1989 - 2022 SCF.

The SCF thoroughly records intergenerational transfer status for up to three past transfers including the sources and the timing of transfers, while any additional transfers are aggregated into a single entry (i.e., see SCF code x5818 above).[52] For each reported past transfer, I utilise the reported approximate values at the time of receipt, along with the year of receipt and the year of the survey, to adjust all past transfers to 2019 U.S. dollars using the Bureau of Labour Statistics (BLS) Consumer Price Index for Urban Consumers (CPI-U-RS).

B.1. Additional Stylised Facts

This section presents more stylised facts based on the empirical analysis of the U.S. household wealth using the SCF data. First, Table B2 presents the evolution of U.S. wealth inequality according to the SCF data from 1989 to 2022. Wealth became increasingly concentrated over the three decades to 2019: the wealthiest 5 percent saw their share of wealth rise from 54 percent to a peak of roughly 65 percent, while all other groups saw their shares decline, lifting the aggregate Gini coefficient from 0.79 in 1989 to a peak of 0.86 in the mid-2010s. This trend partially reversed in the most recent wave, with the top-1 percent share easing from 37.2 percent in 2019 to 35.1 percent in 2022 and the Gini declining to 0.83.

Table B2: Evolution of Wealth Inequality in the U.S.

Year	Gini	HH in Wealth Quintiles					The Wealth-Rich		
		1st	2nd	3rd	4th	5th	90-95%	95-99%	Top 1%
1989	0.790	-0.2	1.2	5.2	13.0	80.7	12.9	24.3	29.9
1992	0.786	-0.2	1.5	5.4	13.2	80.1	12.6	24.3	30.1
1995	0.791	-0.2	1.6	5.5	12.5	80.6	11.8	21.2	34.8
1998	0.800	-0.3	1.4	5.1	12.4	81.4	11.4	23.3	33.8
2001	0.805	-0.1	1.2	4.5	11.8	82.5	12.2	25.1	32.2
2004	0.809	-0.2	1.1	4.4	11.8	82.9	12.0	24.2	33.2
2007	0.816	-0.2	1.1	4.5	11.2	83.4	11.1	26.7	33.6
2010	0.846	-0.7	0.7	3.3	9.9	86.7	13.5	26.8	34.1
2013	0.850	-0.7	0.6	3.2	9.8	87.0	12.1	27.3	35.5
2016	0.860	-0.5	0.6	2.9	8.6	88.3	11.9	26.5	38.5
2019	0.852	-0.5	0.8	3.4	9.0	87.4	11.5	27.7	37.2
2022	0.830	-0.2	1.0	3.8	10.1	85.4	12.4	25.9	35.1

Source. Federal Reserve Board, 1989 - 2022 SCF.

Furthermore, I observe that inequality has become even more severe post-Great Recession. This can partially be explained by wealth portfolio heterogeneity: For the majority of U.S. households, real estate is the single most important asset in their portfolios, whereas wealthier households hold portfolios that are mostly comprised of public and private equities. Therefore, the burst of the housing market during the Great Recession negatively affected bottom and middle-class households more severely, while wealthy households were less impacted and recovered quickly as the financial sector rebounded after the Great Recession.

[52] For robustness, I compared two different calculations: (i) excluding this entry from the calculation, and (ii) including this entry, assuming the same year of receipt and sources of transfers as their third transfer. The results were quite similar.

Table B3 below presents the evolution of wealth inequality—as measured by the Gini coefficient—across different age groups over the sample period.[53] The increasing trends in inequality apply to all age groups. The results also suggest that inequality tends to decrease with age. The youngest group exhibits the highest inequality in all SCF datasets, while the oldest group, on average, has the lowest Gini coefficient.

Table B3: Gini Coefficients by Age

Year	Under 30	31 - 40	41 - 50	51 - 60	61 - 70	71 - 80	81 +
1989	0.882	0.758	0.741	0.737	0.769	0.753	0.748
1992	0.853	0.761	0.757	0.749	0.739	0.712	0.742
1995	0.835	0.740	0.753	0.764	0.761	0.713	0.785
1998	0.961	0.773	0.762	0.772	0.765	0.739	0.678
2001	0.900	0.785	0.761	0.780	0.786	0.737	0.699
2004	0.917	0.809	0.773	0.764	0.774	0.770	0.691
2007	0.971	0.774	0.779	0.781	0.784	0.786	0.736
2010	1.012	0.896	0.846	0.809	0.781	0.766	0.744
2013	1.118	0.875	0.846	0.815	0.801	0.800	0.758
2016	1.144	0.880	0.820	0.847	0.818	0.797	0.763
2019	1.061	0.877	0.839	0.828	0.815	0.787	0.778
2022	0.885	0.811	0.785	0.789	0.810	0.800	0.803
Total	0.962	0.812	0.789	0.786	0.784	0.763	0.744

Source. Federal Reserve Board, 1989 - 2022 SCF.

A simple comparison between the Gini coefficients in 1989 and 2019 shows that the inequality in the under-30 age group increased the most, by 0.179 over the past three decades. Moreover, the increase in inequality tends to decrease with age, with the oldest group experiencing the smallest increase, by 0.03.

Table B4 presents the sources and sizes of early inheritances (under age 30). This includes households who are over 30 at the time of the survey but had received inheritances before turning 30, restricted to inheritances received within the preceding 30 years. The size columns report the household-level average share of inheritance value by source, each heir's inheritance value is split across sources and then averaged across heirs, so that a small number of unusually large transfers does not dominate, with all amounts deflated to 2019 dollars using the CPI-U-RS.

[53] Note that the Gini coefficient for the under-30 age group goes above 1, as some households hold negative wealth.

Table B4: Sources and Size of Early Inheritances

	Sources of Inheritances			Size of Inheritances		
Year	Grandparents	Parents	Others	Grandparents	Parents	Others
1989	0.42	0.37	0.21	0.50	0.36	0.14
1992	0.43	0.37	0.20	0.29	0.54	0.17
1995	0.31	0.50	0.19	0.14	0.59	0.27
1998	0.34	0.46	0.20	0.48	0.41	0.11
2001	0.49	0.43	0.08	0.48	0.46	0.06
2004	0.44	0.43	0.13	0.54	0.40	0.06
2007	0.43	0.45	0.12	0.22	0.73	0.05
2010	0.47	0.40	0.13	0.18	0.69	0.13
2013	0.36	0.47	0.17	0.30	0.57	0.13
2016	0.36	0.54	0.10	0.32	0.60	0.08
2019	0.41	0.47	0.12	0.39	0.49	0.12
2022	0.44	0.48	0.08	0.19	0.77	0.04
Average	0.41	0.45	0.14	0.34	0.55	0.11
	0.48	0.52		0.38	0.62	

Note. Early inheritances are those received before age 30. "Sources" are shares of receipt counts and "Size" shares of real (2019 US$) inheritance value, by giver; the last row renormalises grandparents versus parents.

Source. Federal Reserve Board, 1989 - 2022 SCF.

Table B5 reports the composition of wealth portfolios by wealth group. Note that The weights of each component are computed as a share of net worth, so that total assets and total debt sum to one. While the share of real estate for the bottom 50% appears very high, their net worth is negative, as the value of debt exceeds the value of total assets.

Table B5: Heterogeneity in Wealth Portfolio Composition

	Financial assets			Non-financial assets			Total Debt	
	Interest-earning asset	Public equity	Other	Real estate	Private business	Other		
Top 0.1%	0.11	0.23	0.04	0.14	0.49	0.01	(0.01)	1.000
Next 0.9%	0.13	0.26	0.06	0.26	0.32	0.01	(0.04)	1.000
Next 4%	0.13	0.28	0.09	0.36	0.20	0.02	(0.07)	1.000
Next 5%	0.13	0.26	0.12	0.47	0.10	0.03	(0.12)	1.000
Next 40%	0.13	0.18	0.11	0.74	0.06	0.08	(0.30)	1.000
Bottom 50%	0.26	0.18	0.16	2.55	0.04	0.63	(2.83)	1.000
Overall	0.13	0.24	0.09	0.45	0.21	0.04	(0.16)	1.000

Note. Portfolio shares of net worth by wealth group; assets and debt sum to one. Shares are aggregate ratios (component value over net worth) pooled across waves; for the bottom 50%, real estate exceeds one and debt is strongly negative because aggregate net worth is small relative to gross holdings.
Source. Federal Reserve Board, 1998 - 2022 SCF.

B.2. Return on Wealth: Data and Methodology

Another important exercise I consider in this appendix is documenting evidence of return heterogeneity stemming from portfolio heterogeneity. To address this, I use the SCF datasets to compute portfolio-weighted average rates of return on wealth for different wealth groups. In doing so, I generally follow the methodologies employed by existing studies.[54]

Let $a_{i,j,t}$ and $d_{i,k,t}$ denote the value of asset j and the value of debt k held by each wealth group i at the beginning of period t. Given J assets and K debt categories, each wealth group i's net worth is given by

$$w_{i,t} = \sum_{j=1}^{J} a_{i,j,t} - \sum_{k=1}^{K} d_{i,k,t} \tag{B1}$$

First, for each SCF dataset, households are grouped into 6 wealth bins: 0-50, 50-90, 90-95, 95-99, 99-99.9 and 99.9-100 percentiles. Next, I define households' assets into 6 categories as shown in Table B6 such that (B1) holds: (i) interest-earnings asset, (ii) public equity (iii) other financial assets (iv) real estate, (v) private businesses and (vi) other non-financial assets. For debts, I follow pre-defined categories by the SCF: (i) debt secured by primary residences, (ii) debt secured by other real estate (iii) other lines of credit, (iv) credit card debt (v) instalment loans (including education loans and vehicle loans) and (vi) other debt.

Table B6: Classification of Assets and Debt Components

Total Assets	Total Debt
Financial Assets	Debt secured by primary residences
Interest-earning assets	Debt secured by other real estate
Public equity	Other lines of credit
Other financial assets	Credit card debt
Non-financial assets	Installment loans
Real estates	Education loans
Private Businesses	Vehicle loans
Other non-financial assets	Other debt

Note. Asset and debt components follow the Survey of Consumer Finances classification; indented rows are subcomponents of the category listed above them. This mapping underlies the portfolio shares in Table B5 and the component returns in Table B8.

If we let $r^A_{i,j,t}$ and $r^D_{i,j,t}$ denote the average return on asset j and debt k earned (or paid) by the wealth group i respectively. Then the return on net worth $r_{i,t}$ is then computed as the weighted average of the returns

[54] It is worth noting that this methodology requires two consecutive SCF datasets for each episode due to its non-panel structure. For example, the SCF 2016 and 2019 are utilised to compute the average return over the 2017-2019 episode. Consequently, I use eight sets of SCF data to calculate average returns spanning from 1999-2001 to 2017-2019.

on individual wealth components,

$$r_{i,t} = \sum_{j=1}^{J} \omega^{A}_{i,j,t} \times r^{A}_{i,j,t} - \sum_{k=1}^{K} \omega^{D}_{i,k,t} \times r^{D}_{i,k,t} \tag{B2}$$

where the weights $\omega^{A}_{i,j,t}$ and $\omega^{D}_{i,k,t}$ of any wealth group i in (B2), are given by the share of net worth invested in an asset or debt category. Therefore, for each SCF dataset and for each wealth group, I aggregate the values of each asset and each debt and divide by the aggregated value of net worth.[55]

The return on any asset j (or any debt k) is defined as the sum of yield component and capital gain component.

$$r_{i,j,t} = y^{*}_{i,j,t} + g_{i,j,t} \tag{B3}$$

where the yield component $y^{*}_{i,j,t}$ in (B3) is computed as,

$$y^{*}_{i,j,t} = \left(1 + \frac{y_{i,j,t}}{a_{i,j,t} + 0.5\left(a_{i,j,t+3} - a_{i,j,t} - \tilde{y}_{i,j,t}\right)}\right)^{\frac{1}{3}} - 1$$

following Fagereng et al. (2016) and Kartashova and Zhou (2021). Here, $a_{i,j,t}$ and $a_{i,j,t+3}$ are the reported market values of an asset (or debt) in two consecutive surveys and $y_{i,j,t}$ in the numerator is the total flow income (payment) generated over 3 years. The denominator takes into account potential net flows invested between t and $t+3$. Note that $y_{i,j,t}$ and $\tilde{y}_{i,j,t}$ are specific to the asset type.

Interest-earning assets. This includes all types of transaction account (liquid assets), certificate of deposit, directly and indirectly held bonds, cash value of whole life insurance and other interest-earning assets invested through annuities and trusts. The income flow generated from these assets is the total interest income reported by households. In the SCF 2019, households report annual interest income earned for 2018. Following Kartashova and Zhou (2021), I use growth rates of interest income from the Bureau of Economic Analysis (BEA) to obtain annual interest income earned in 2017 and 2019 and sum all to obtain $y_{i,j,t}$ over 2017-2019. As in Fagereng et al. (2020) and Kartashova and Zhou (2021), all interest incomes are capitalised into $a_{i,j,t+3}$ and no capital gains for interest-earning assets hence, $\tilde{y} = y_{i,j,t}$.

Public-equities. Public equities are directly and indirectly held stocks including stocks invested through annuities and trusts. The income flow generated from public equities is the total dividend income reported by households. Similar to the interest incomes, I use the growth rates of dividend income from the BEA to obtain total flow dividend income over 2017-2019. I use Wilshire 5000 Total Market Index to obtain the capital gains component of public equities. Following Kartashova and Zhou (2021), I assume dividend income and capital gains are capitalised into $a_{i,j,t+3}$, hence $\tilde{y}_{i,j,t} = y_{i,j,t} + g_{i,j,t}a_{i,j,t}$.

[55] Then the portfolio shares of each wealth group for a given time period, is simply the average of the values obtained from two consecutive SCF surveys.

Other financial assets. Other financial assets is computed as the total financial assets minus the interest-earning assets and public equity. I assume no income is generated from these assets.

Real estate. This category includes primary residence, other residential properties, and net equity in non-residential real estate. Since the yield on primary residence is unobserved, I use aggregate counterparts to compute the yield on primary residences. I divide imputed rental of owner-occupied housing by the market value of owner-occupied real estate to compute the annual yield on primary residences. Following Kartashova and Zhou (2021), I subtract the annual depreciation rate of 2.3% (estimated by the BEA) and the average effective property tax rate 1.03%. For non-primary residential properties, I use the annual rent income reported in SCF by households. Since the rent income collected by the SCF include rent income plus other incomes from trusts or royalties, I condition on households who own non-primary residential properties (but not the commercial properties and/or any other miscellaneous assets such as royalties) to isolate the rent income. As before, I use growth rates of rent income to obtain $y_{i,j,t}$. For this asset, $\tilde{y}_{i,j,t} = (g_{i,j,t} - 3\delta^R)a_{i,j,t}$ where $\delta^R = 3.64\%$ is the annual depreciation of residential rental properties. As before, I further subtract the average effective property tax rate of 1.03%. Finally, I use the commercial property return index (CPRI) from the Green Street as the return on non-residential real estate as in Kartashova and Zhou (2021). For the 2020–2022 episode, over which the Green Street series is unavailable, I substitute the NCREIF Property Index, a comparable transaction-based commercial-property total-return benchmark.

Private Business. This includes the share of net equity in the non-publicly traded businesses owned reported by households. This can be sub-divided into two: corporate equities (S and C corporations) and non-corporate equities (proprietorship and partnerships). I follow Moskowitz and Vissing-Jørgensen (2002) and Kartashova (2014) to compute the yield component. From the reported net business income, I adjust for corporate taxes (30% for C corporations and 0% for S-corporations and non-corporate businesses) and retained earnings (40% for C-corporations and 20% for S-Corporations and non-corporate businesses). In order to account for labour incomes of entrepreneurs who actively manage the business but report no salary, I regress the wage rate of households (who actively manage and report salary) on households' age, gender, education, working hours and size of business to impute the labour income of these households. As before, I use growth rates of corporate profit and non-corporate profit to obtain $y_{i,j,t}$ over 2017-2019 and I assume $\tilde{y}_{i,j,t} = g_{i,j,t}$ to compute the return on each type of business.

Other non-financial assets. Other non-financial assets is computed as the total non-financial assets minus the private businesses and real estate. I assume constant return at 2% as in Kartashova and Zhou (2021) over the sample period.

Debt. The income flow (i.e. payments) generated from each debt category is computed as the reported annual interest rate multiplied by the total amount owed. If households report multiple loans within each debt component, I used geometric average of the interest rates reported. Once the interest payments are obtained, I use growth rates of mortgages (if secured by residential properties) and consumer credits (for other types of debt) to obtain $y_{i,j,t}$ and set $\tilde{y}_{i,j,t} = y_{i,j,t}$ to compute the return on each debt category.

For the capital gains component $g_{i,j,t}$ in (B3), I infer their changes based on the growth rates of their aggregate counterparts. Following external data series are used to compute capital gains component when calibrating the excess return schedules.

Table B7: List of External Data Series

Data Series (FRED code)	Source
Financial assets	
Personal interest income (PII)	BEA
Personal dividend income (PDI)	BEA
Wilshire 5000 Total Market Index (WILL5000IND)	Wilshire Associates
Non-financial assets	
Imputed Rental of Owner-occupied housing (A2013C1A027NBEA)	BEA
Personal rental income (RENTIN)	BEA
Corporate Profits After Tax (without IVA and CCAdj) (CP)	BEA
Net profit of nonfarm proprietorships and partnerships (B1207C1A027NBEA)	BEA
Commercial Property Return Index (CPRI)	Green Street
Households and Nonprofit Organisations;	
Owner-Occupied Real Estate (inc. vacant land and mobile homes) (HOOREVLMHMV)	Financial Accounts
Real Estate at Market Value (HNOREMV)	Financial Accounts
Corporate Equities; Asset, Market value Levels (HNOCEA)	Financial Accounts
Proprietors' Equity in Noncorporate Business (ENBABSHNO)	Financial Accounts
Debts	
Households and Nonprofit Organisations;	
One-to-Four-Family Residential Mortgages (HMLBSHNO)	Financial Accounts
Consumer Credit (CCLBSHNO)	Financial Accounts

Note. External aggregate series used to construct the component returns in Table B8; FRED mnemonics are shown in parentheses. Income flows are from the BEA, asset stocks from the Financial Accounts of the United States, the public-equity index from Wilshire Associates, and commercial-property returns from Green Street.

Table B8 reports the resulting average return on each wealth component by episode; private business has yielded persistently higher returns than other assets throughout the sample period.[56]

Table B8: Average Returns on Each Wealth Component

	Financial assets			Non-financial assets			Total Debt
	Interest-earning asset	Public equity	Other	Real estate	Private business	Other	
1999 - 2001	0.04	0.02	-	0.11	0.29	0.02	0.07
2002 - 2004	0.02	0.00	-	0.12	0.27	0.02	0.06
2005 - 2007	0.03	0.10	-	0.09	0.40	0.02	0.06
2008 - 2010	0.02	(0.06)	-	(0.07)	0.06	0.02	0.06
2011 - 2013	0.02	0.14	-	0.05	0.28	0.02	0.04
2014 - 2016	0.02	0.10	-	0.10	0.26	0.02	0.04
2017 - 2019	0.02	0.12	-	0.06	0.36	0.02	0.04
2020 - 2022	0.02	0.15	-	0.13	0.23	0.02	0.04
Average	0.02	0.07	-	0.07	0.21	0.02	0.06

Note. Values in parenthesis indicate negative returns.
Source. Federal Reserve Board, 1998 - 2022 SCF.

Since the 2022 wave of the Survey of Consumer Finances became available, I extend the return series through the 2020–2022 episode, computed as above with two data substitutions. Since the Green Street commercial-property series ends in 2019, I proxy the non-residential real-estate return with the NCREIF Property Index; and, absent a separate corporate-business return for the episode, I apply the non-corporate business return to private business as a whole, while the interest-earning return is held at its long-run level. These substitutions affect only a small share of the wealthiest households' portfolios: adding the episode raises the top-group returns by about one percentage point (Table B8) and leaves the wealth–return gradient, the standard deviations, and every conclusion drawn from them unchanged.[57]

[56]The lower average returns on public equity and real estate over the full sample reflect substantial declines during the Great Recession.

[57]Households with greater wealth are likely to earn higher returns even within the same wealth component, possibly reflecting advantages in education, skills, or information access. I abstract from such within-class heterogeneity, as it is beyond the scope of this exercise.

Table B9: Heterogeneity in Returns across Different Wealth Groups and Episodes

	Percentile (%)					
	0-50	50-90	90-95	95-99	99-99.9	99.9-100
1999 - 2001	0.07	0.08	0.10	0.09	0.09	0.08
	(0.01)	(0.01)	(0.02)	(0.03)	(0.04)	(0.06)
2002 - 2004	0.08	0.08	0.07	0.08	0.10	0.11
	(0.00)	(0.01)	(0.01)	(0.02)	(0.03)	(0.04)
2005 - 2007	0.00	0.07	0.09	0.13	0.16	0.18
	(0.01)	(0.01)	(0.02)	(0.04)	(0.07)	(0.10)
2008 - 2010	-0.48	-0.08	-0.05	-0.04	-0.04	-0.03
	(0.01)	(0.01)	(0.02)	(0.03)	(0.05)	(0.08)
2011 - 2013	-0.10	0.05	0.07	0.10	0.13	0.15
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.03)
2014 - 2016	0.12	0.08	0.09	0.11	0.11	0.14
	(0.01)	(0.01)	(0.01)	(0.03)	(0.04)	(0.06)
2017 - 2019	0.04	0.07	0.09	0.11	0.12	0.13
	(0.01)	(0.02)	(0.03)	(0.06)	(0.09)	(0.14)
2020 - 2022	0.22	0.11	0.10	0.11	0.13	0.17
	(0.02)	(0.06)	(0.05)	(0.05)	(0.05)	(0.05)
Average	-0.01	0.06	0.07	0.08	0.10	0.12
	(0.01)	(0.01)	(0.02)	(0.03)	(0.05)	(0.07)

Note. Values in parentheses indicate standard deviations. Group returns are portfolio-weighted and pooled by episode; the bottom-50% level (leveraged real-estate gains net of debt) is the least reliable and least relevant to the top-concentrated mechanism.

C. Model

C.1. Stationary Population

The total population alive at any given time t is given by:

$$
\begin{aligned}
N_t &= \sum_{j=1}^{J} N_{j,t} \\
N_{j+1,t+1} &= \gamma_{j,t} N_{j,t}, \quad \forall j \in \{1, \ldots, J\} \\
N_{1,t} &= (1+f) N_{1,t-6}
\end{aligned}
$$

where the last equation describes the total population of new households entering at $t = 1$ that is given by the population of their parents' generation ($N_{1,t-6}$) multiplied by the fertility rate f. I assume that the population grows at a constant rate n such that:

$$
N_{1,t+1} = (1+n) N_{1,t}
$$

Hence, in a steady-state where age-dependent survival probabilities and fertility rate are constant, the rate of total population growth is determined as,

$$
n = (1+f)^{\frac{1}{6}} - 1
$$

and the population composition becomes stationary. Normalising the total population to 1, we have:

$$
\begin{aligned}
N &= \sum_{j=1}^{J} N_j = 1 \\
N_{j+1} &= \frac{\gamma_j}{1+n} N_j, \quad \forall j \\
N_1 &= \bar{N}_1 = \left(\sum_{j=1}^{J} \frac{\gamma^{j-1}}{(1+n)^{j-1}} \right)^{-1}
\end{aligned}
$$

where $\bar{N}_1$ is the share of age-1 households in the stationary equilibrium.

C.2. Definition of the Stationary Equilibrium

Definition: Given government consumption G, government debt D, a tax system characterised by (τ_a, $T_\ell(e)$, τ_b, χ_b, τ_{gst}, χ_{gst}) and a social security system characterised by (τ_{ss}, s), a stationary recursive competitive equilibrium with population growth is a collection of value functions $\{V_w^{pg}, V_w^{pw}, V_w^{pr}, V_w, V_r\}$ and relevant policy functions $\{c(x), a'(x)\}$, optimal input choices $\{K, L\}$ of firms, and equilibrium prices $(r^*, \underline{r}, w)$ with the following properties:

i. Given prices $(r^*, \underline{r}, w)$ and government policies $(\tau_a, T_\ell(e), \tau_b, \chi_b, \tau_{gst}, \chi_{gst}, \tau_{ss}, s)$, the functions $V(x)$, $c(x)$ and $a'(x)$ solve the household's maximisation problem in state x.

ii. Given prices $(r^*, \underline{r}, w)$, the optimal choices of the representative firm satisfy:

$$r^* = A\alpha \left(\frac{K}{L}\right)^{\alpha-1} - \delta_k, \qquad w = A(1-\alpha)\left(\frac{K}{L}\right)^{1-\alpha}$$

iii. Φ is the invariant distribution of households over the state variables, $\Phi(X) = 1$

iv. Government policies satisfy the government budget constraints:

$$\begin{aligned} G + ((1-\tau_a)r^* - n)D = \tau_a r^* K &+ \int T_\ell(w \cdot e(j,z))\ d\Phi \\ &+ \int (1-\gamma_j)\, T_b(a'(x))\ d\Phi + \int (1-\gamma_j)\, \mathbb{1}_{\{j \geq 13\}}\, T_{gst}\big(b_{gc}(x)\big)\ d\Phi \end{aligned}$$

$$\int s \cdot \mathbb{1}_{\{j \geq J_r\}}\ d\Phi = \tau_{ss} \int (w \cdot e(j,z))\ d\Phi$$

v. All markets clear:

- The labour market clears,

$$L = \int e(j,z)\ d\Phi$$

- The capital market clears,[58]

$$(1+n)(K+D) = \int a'(x)\ d\Phi$$

$$r^* \int a(x)\ d\Phi = \int \left(\underline{r} + p\left(r_e + r_{e2}\mathbb{1}\{a \geq \underline{a}\} + \sigma_e \eta\right)\right) \cdot a\ d\Phi$$

- The goods market clears,[59]

$$\int c(x)\ d\Phi + (n+\delta)K + G + \mathcal{K} = Y$$

[58] The condition pins down the aggregate return component, $\underline{r}$.

[59] Here, $\mathcal{K} = \int \left[\kappa_e\, \mathbb{1}_{\text{acq}}(x) + (1-\gamma_j)\, \mathbb{1}_{\{j \geq 13\}}\, \iota(x)\, \xi\right] d\Phi$ is the aggregate resource cost of acquiring financial capability and of establishing generation-skipping arrangements, with $\mathbb{1}_{\text{acq}}(x)$ indicating a household that acquires capability in state x. These include payments for real financial, legal, and advisory services and could equivalently be counted within aggregate consumption; they amount to less than 1 percent of output in the calibration.

D. Calibration

D.1. Earnings Process

In the model, households in working stage supply one unit of labour inelastically and earn labour income which is given by:

$$\log e(j, z) = z + \epsilon_j$$

where z is the idiosyncratic stochastic component that follows an AR(1) process and ϵ_j is the deterministic age-efficiency profile.

I calibrate the model earnings process following Kindermann and Krueger (2021). The stochastic component z is approximated by a **five-state** Markov chain: the lowest three states and their 3×3 transition probabilities are obtained from Tauchen (1986), and to these I append two high-productivity "superstar" states, z_4 and z_5, whose levels and transition probabilities are calibrated to reproduce the concentration of the earnings distribution. As in the treatment of top earnings there, I use only information from the empirical earnings distribution and do not target the wealth distribution.

Since the model features imperfect transmission of productivity from parents to children as in De Nardi and Yang (2016), the stationary age-1 distribution μ_1 solves $\mu_1 = Q'_h\,(Q'_z)^6\mu_1$, where Q_z is the within-life productivity transition matrix and Q_h the parent-to-child transmission matrix (from an age-55 parent to an age-25 child at entry). I impose the following restrictions to reduce the parameter space:

- Households in the lowest three states cannot reach the top state z_5 directly; z_5 is reached only through z_4 (a ladder restriction).
- Top-state earnings are largely transitory: from z_5 a household reverts to z_4 with high probability, and the children of top-state parents (z_4, z_5) regress toward the middle of the distribution rather than inheriting the top.
- Parents outside the top states (states z_1–z_3) cannot have a top-state child; a child reaches z_5 only from a top-state parent.

I calibrate the free parameters to minimise the squared distance to a set of earnings moments: points on the earnings Lorenz curve, the earnings Gini coefficient, the mean-to-median ratio, the share held by the top 0.1 percent (6.1% in the SCF), and the parent–child earnings correlation, which I set to 0.33 from my own PSID estimate (Section III). The calibrated grid is

$$z = \begin{bmatrix} 0.3923 & 1.0000 & 2.5492 & 15.7425 & 110.5233 \end{bmatrix}$$

with implied age-1 stationary distribution

$$\mu_1 = [\,38.3\%,\ 0.1\%,\ 61.0\%,\ 0.2\%,\ 0.5\%\,].$$

The resulting transition and transmission matrices are

$$Q_z = \begin{bmatrix} 0.7238 & 0.2639 & 0.0105 & 0.0017 & 0 \\ 0.1962 & 0.6060 & 0.1962 & 0.0017 & 0 \\ 0.0104 & 0.2617 & 0.7177 & 0.0102 & 0 \\ 0 & 0.0003 & 0.0008 & 0.9941 & 0.0048 \\ 0 & 0 & 0 & 0.4500 & 0.5500 \end{bmatrix}, \quad Q_h = \begin{bmatrix} 0.9858 & 0.0013 & 0.0129 & 0 & 0 \\ 0.2839 & 0.0009 & 0.7152 & 0 & 0 \\ 0.0009 & 0.0022 & 0.9965 & 0.0004 & 0 \\ 0 & 0 & 0.8318 & 0.0430 & 0.1253 \\ 0 & 0 & 0.8318 & 0.0430 & 0.1253 \end{bmatrix}.$$

The calibrated process reproduces the earnings distribution (including the top-0.1 percent share of 6.1%) and the mean-to-median ratio of 1.71, at a parent–child earnings correlation of 0.33. Two features of the transition dynamics are worth noting. The within-life matrix Q_z makes middle productivity persistent over five years while the top state z_5 is markedly *transitory* (a five-year persistence of 0.55, against 0.99 for z_4), reflecting the temporary character of very high earnings. The transmission matrix Q_h shows that the children of top-state parents regress toward the middle with high probability, so that top-tail earnings persistence is low even though the average parent–child correlation matches the data. This is what allows the fat earnings tail to coexist with a disciplined intergenerational transmission: the model contains realistic self-made high earners, but their earning ability does not by itself propagate down the dynasty.

I use mean earnings for households age between 25 to 60 from the SCF (2001) as the deterministic age-efficiency profile. Since the mandatory retirement in the model starts is 65, I set $\epsilon_j = 0$ for $j \geq J_r$. I then normalise the age-efficiency profile such that average earnings (before tax) at age 1 equal one unit. In dollars, one unit corresponds to $\$46,472 \times 5 = \$232,360$ per 5-year period (in 2019 US dollars), the value at which the capability cost κ_e equals the SEC accredited-investor net-worth threshold ($1.00 million).

D.2. Return Heterogeneity

The return process is described in Section 4: a household earns only the safe return $\underline{r}$ unless it has acquired the financial-capability state $p = 1$, in which case it earns a first excess return r_e, a second excess return r_{e2} on wealth above the threshold $\underline{a}$, and bears idiosyncratic return risk $\sigma_e \eta$:

$$1 + r_i = 1 + \underline{r} + p_i \left(r_e + r_{e2} \cdot \mathbb{1}\{a_i \geq \underline{a}\} + \sigma_e \cdot \eta \right).$$

The idiosyncratic shock η is discretised by 5-point Gauss–Hermite quadrature, with nodes and weights

$$\{\pm 2.857, \pm 1.356, 0\} \quad \text{and} \quad \{0.011, 0.222, 0.533, 0.222, 0.011\},$$

matching the standard-normal moments up to fourth order ($\mathbb{E}[\eta^2] = 1$, $\mathbb{E}[\eta^4] = 3$).

The magnitudes of the excess returns are disciplined by the portfolio-weighted returns to net worth computed from the SCF (Table B8; see Appendix B.2 for the methodology). In the data, returns rise with wealth, from roughly 6 percent for households in the 50th–90th percentiles to about 8, 10, and 12 percent for the 95th–99th, 99th–99.9th, and top 0.1 percent groups, and their dispersion widens in parallel. The model's calibrated excess returns are comparable to the excess of these group returns over the safe rate: a capable household earns $r_e = 3.1$ percent above the safe return, a further $r_{e2} = 0.6$ percent on wealth above $\underline{a}$, and bears idiosyncratic dispersion σ_e around these; the calibrated premia are broadly consistent with SCF excess returns of roughly 2 to 6 percent for the wealthy groups. Within the model, the first excess return r_e and the threshold $\underline{a}$ are calibrated to the shares held by the 95th–99th and 99th–99.9th percentiles, the return dispersion σ_e to the top 0.1 percent share, while the capability-acquisition cost κ_e is set externally to the SEC accredited-investor threshold, as detailed in Section 4. Since a household's *expected* return does not rise with the *level* of wealth within a capability type (the idiosyncratic shock $\sigma_e \eta$ is mean-zero and independent of wealth), the empirical rise of average returns with wealth is not imposed but reproduced through selection, as an untargeted feature of the model.

D.3. Computational Details

The stationary equilibrium for the baseline economy is computed using constant real interest rate and capital income tax rate. I start first by guessing the capital income tax rate and compute the optimal policy and value functions for the last period $j = J$, then solve the household problems at earlier ages using the method of endogenous grid and backward induction.

Two features of the revised model augment the state space described below. First, every individual state vector additionally carries the financial-capability indicator $p \in \{0, 1\}$, so the dimension of each state block is doubled. Since capability is absorbing, its law of motion is deterministic given the household's policy and requires no additional stochastic kernel: at the start of a period an uncapable household ($p = 0$) either pays the fixed cost κ_e and moves to $p = 1$, whenever the acquisition problem in Section II makes this optimal and $a \geq \kappa_e$, or remains uncapable, and any uncapable heir whose augmented wealth covers the acquisition cost ($a + b \geq \kappa_e$) pays κ_e and has its capability switched on at receipt, $p' = \max\{p, \mathbb{1}\{a + b \geq \kappa_e\}\}$, entering the next period with assets $a + b - \kappa_e$. Both transitions are encoded as $0/1$ indicators in the relevant transition matrices. Second, in the overlap window the retiree state carries the child's estate-tax status $d_c = \mathbb{1}\{a_c + b_c \geq \chi_b\}$ (evaluated on the child's wealth inclusive of the bequest it receives, as in Section II), discussed further below.

In order to simulate the economy and find its stationary distribution, I avoid carrying ancestors' full states. A living household's individual state is (a, z, η, p): assets, earnings productivity, the return shock, and the capability indicator. Family history enters only through two *discrete* tags: the parent's wealth bin $\sigma_p \in \{1, \ldots, K_p\}$ (the parent's wealth quintile measured at the parent's age 7, the child's entry) and the grandparent's wealth bin $\sigma_g \in \{1, \ldots, K_g\}$ (the grandparent's wealth at age 13, with edges $\chi_b \cdot [1, 2, 6]$), with $K_p = 5$ and $K_g = 4$. The population is carried as four age-indexed mass blocks according to which ancestors are still linked: μ^1 (no living-ancestor link), $\mu^2(\sigma_p)$ (parent alive), $\mu^g(\sigma_g)$ (grandparent alive), and $\mu^3(\sigma_g, \sigma_p)$ (both), together with a retiree block μ^r over (a, η, p).

Rather than simulate each ancestor's wealth path, I summarise the bequests a household of a given tag will receive by a set of *receipt distributions*. For each parent bin σ_p and child age j, $F_p(b \mid \sigma_p, j)$ gives the distribution of parental receipts, the whole net estate for accidental parental deaths at ages 7–12, and the child's leg of the split for deaths at ages 13–14. For each grandparent bin σ_g and age $j \leq n_{GG}$, $F_g(b \mid \sigma_g, j)$ gives the distribution of the grandchild's leg of the split, net of estate and GST tax. These distributions are built by pushing each ancestor wealth-bin cohort forward with its own policy functions and recording, at each age, the death-event mass and the associated post-tax estate on the asset grid.

The economy is then advanced one age at a time. Each block's mass is moved by the household's own saving policy $a' = g(a, z, \eta, j)$ and the earnings transition Q_z (with the η shock drawn *i.i.d.*); when a linked ancestor of tag σ dies at age j (an event of probability $1 - \gamma_j$), the household receives a bequest drawn from $F(b \mid \sigma, j)$, moves to the corresponding no-longer-linked block, and has its capability switched on whenever the receipt lifts augmented wealth above κ_e. Receipts and savings are assigned to the nearest two grid points with distance weights, so that mass is never quantised to a single node, and cohort masses obey $\sum \mu_{j+1} = \frac{\gamma_j}{1+n} \sum \mu_j$. Each newborn draws a productivity level z_1 and a pair of ancestor tags (σ_p, σ_g) from a joint entry distribution $\lambda(z_1, \sigma_p, \sigma_g)$.

The receipt distributions F_p and F_g and the entry distribution λ are recomputed from the stationary distribution after each pass over the household problem, and iterated to convergence jointly with the conditional probability $\pi(d_c \mid a)$ and the budget-balancing capital income tax described below.

It is worth noting that, I discretise the asset grid into 70 points ranging from 0 to 1600.[60] I discretise all 70 points geometrically based on

$$a_i = \bar{a} \cdot \frac{(1 + g_{na})^{i-1} - 1}{(1 + g_{na})^{na-1} - 1}$$

where I set $\bar{a} = a_{max} = 1600$, $na = 70$ and $g_{na} = 0.15$. Hence the distance between two successive grid points becomes significantly larger in the upper region. although I described using an indicator function when assigning probabilities in transition matrices, in fact, I assign positive probabilities to the nearest two grid points, using the distance as weights to ensure a smooth transition and prevent the wealthiest from being stuck at a certain state.

The estate tax, the GST tax, and the skipping decision enter the computation through the retiree's terminal problem rather than through the transition kernels. For a household that dies in the overlap window ($j \geq 13$), I solve the discrete generation-skipping choice $\iota \in \{0, 1\}$ by comparing the warm-glow value under two options: leaving the entire post-estate-tax estate to the child, and paying the fixed cost ξ and dividing the estate between child and grandchild according to the interior first-order condition in Section II, with the GST tax T_{gst} levied on the grandchild's share. The household skips only when the second option dominates, which makes the prevalence of generation-skipping an equilibrium object governed by ξ. The one piece that must be solved jointly with the rest of the equilibrium is the conditional distribution $\pi(d_c \mid a')$, the probability that a dying grandparent's child holds wealth above the estate-tax exemption, since it determines the division weight $\kappa_j(d_c)$. I compute $\pi(d_c \mid a')$ from the stationary joint distribution of linked parent–child states, update it after each pass over the household problem, and iterate until $\pi(d_c \mid a')$, the capital income tax rate, and the aggregate distribution have jointly converged.

Once all the transition matrices are obtained, I then simulate the economy until a stationary distribution of households over the state space is achieved. The stationary distribution is then obtained when the age distribution, productivity distribution and average wealth across the households are all stabilised. Using the stationary distribution, I check whether the government budget is balanced and update the capital income tax accordingly. The procedure described above is repeated until the government budget is balanced.

[60] The normalisation is such that 1 unit in the model corresponds to $\$46,472 \times 5 = \$232,360$ per 5-year period (in 2019 US dollars). Hence $a_{max} = 1600$ corresponds to about \$372 million, ensuring agents in the model can become as wealthy as in the data.

D.4. Welfare Computation

Welfare is measured by the consumption-equivalent variation (CEV) of a household evaluated behind the veil of ignorance, before its productivity, family background, and inheritances are realised. Let Ψ_0 denote the entry (age-1) distribution over newborn states $x_0 = (z, S_p, S_g)$ (the household's initial productivity draw together with the states of its parent and, where alive, its grandparent) and let $V_1(x_0)$ be the corresponding age-1 value function (V_w^{pg}, V_w^{pw}, or V_w depending on which ascendants are alive at entry). The ex-ante value of being born into an economy is

$$\mathcal{W} = \int V_1(x_0)\, d\Psi_0(x_0).$$

Consider two economies, a benchmark b and an alternative a, a counterfactual reform, or the model with the G-G link switched off. Since preferences are CRRA and the value function is homogeneous of degree $1 - \sigma$ in the consumption–bequest scale, the uniform proportional adjustment λ to consumption in every date and state that equates ex-ante values has the closed form

$$\lambda = \left(\frac{\mathcal{W}_a}{\mathcal{W}_b}\right)^{\frac{1}{1-\sigma}} - 1,$$

reported in percent throughout.[61] A positive λ means a newborn prefers economy a; a negative λ means it prefers the benchmark. For the reform experiments b is the benchmark and a the reformed economy, so the negative CEV values in Table 18 are welfare losses. For the value of the G-G link I set a to the economy with the link active and b to the economy with it switched off, so the positive CEV is the amount a newborn would sacrifice to retain access to the link.

The same formula applied to a *conditional* entry distribution yields the group CEVs of the cohort analysis (Table 13). Restricting Ψ_0 to newborns whose grandparent falls in a given wealth bin and renormalising gives $\mathcal{W}^g$ for that group; comparing $\mathcal{W}^g$ across grandparent-wealth bins isolates the "lottery" component of the link's value, for instance the CEV of being born to a top-1% grandparent relative to receiving no grandparent transfer. Since these conditional groups differ enormously in inherited wealth, the group CEVs are large, whereas the unconditional CEV above averages the same effect over the roughly four-fifths of newborns who receive no grandparent transfer at all.

[61] The additive constants in the CRRA and warm-glow functions are immaterial normalisations of the utility index and are dropped when forming the ratio, so that $\mathcal{W}$ is homogeneous of degree $1 - \sigma$ in the consumption–bequest scale.

E. Additional Results

This appendix collects supporting results referenced in the main text.

E.1. Intergenerational Wealth Mobility

The benchmark is calibrated to cross-sectional concentration and to the size and timing of transfers; it uses no intergenerational *mobility* moment as a target. This appendix asks whether the model nonetheless reproduces the mobility patterns documented in the data, clarifies how its grandparent effect is distributed, and isolates the role of bequest anticipation. Since the numerical stationary distribution summarises ancestors by a small number of wealth bins, which mechanically attenuates any rank statistic, I recover the mobility moments from a Monte Carlo simulation of dynasties that tracks each generation's continuous wealth.

Simulation. I simulate a large panel of lineages. Each household enters at age 25 with a productivity level inherited from its parent through Q_h, accumulates wealth over the life cycle under the benchmark saving and capability-acquisition policies (with the return shock η drawn i.i.d. each period), and dies according to the survival schedule γ_j. Saving policies are selected by family state, whether the parent and grandparent are alive and their wealth bins, and at death the net estate is divided by the fertility rate, the child receiving $b/(1+f)$ and, when the household dies inside its grandchild's overlap window, the grandchild receiving the skip leg $b_{gc}/(1+f)^2$, exactly as in the equilibrium law of motion. Wealth is recorded at every age and ranked *within each generation's own age cohort*; the parent–child (G-G) rank correlation at age a is the Spearman correlation of the two generations' ranks, both measured at age a but one generation apart in calendar time. The simulation reproduces the conditional statistics of Table 13 to within sampling error: $P(\text{top } 1\% \mid \text{parent in top quintile}) = 5.0\%$ against 4.8%, and among newborns with a living grandparent above $2\chi_b$, 55% reach the top 1 percent by age 65 against 62% in the benchmark, which validates it as a faithful representation of the model's transmission.

Anticipation and the life-cycle profile of mobility. Table 14 in the main text reports the same-age rank correlation from ages 35 to 65 under two saving regimes. In the benchmark, households *anticipate* bequests: knowing a transfer is coming, a prospective heir saves less early in life, so young-age wealth reflects mainly its own earnings (the parent–child correlation sits near its earnings-transmission floor of 0.17) and rises only as inheritances arrive in the late 40s and 50s. The correlation therefore reaches the conventional empirical range of 0.3–0.4 (Charles and Hurst, 2003; Adermon, Lindahl and Waldenström, 2018) only at midlife ages ($\approx$ 48–53). The literature, however, measures children in their 30s and 40s, where full anticipation is implausible: inheritance timing and size are uncertain, many heirs do not know their parents' finances, and inheritances generate responses *at receipt*. The second regime switches anticipation off, raising the young-age correlations into the conventional range precisely at the ages the literature measures (0.24 at 40, 0.37 at 45). Reality lies between the two, closer to the no-anticipation line for the young; on either accounting the model brackets the conventional estimates and is not over-persistent.

Transition probabilities. Table E1 reports transition cuts measured at age 65 under both regimes. Persistence is strong where it should be: over half of the children of top-1% parents are themselves in the top 1 percent, and roughly nine in ten remain in the top quintile.

Table E1: Intergenerational transition probabilities (measured at age 65)

Transition	Benchmark	No anticipation
P(child top 1% | parent top 1%)	52%	52%
P(child top 1% | parent top quintile)	4.9%	5.0%
P(child top quintile | parent top quintile)	52%	55%
P(child top quintile | parent top 1%)	90%	91%
P(grandchild top 1% | grandparent top 1%)	34%	34%

Note. Snapshot at age 65; each generation ranked within its own cohort. Unconditional probabilities are 1% (top 1 percent) and 20% (top quintile). The G-G row is unconditional on the parent and is therefore largely parent-mediated; the direct skip channel is isolated by the conditional statistics of Table 13.

The grandparent effect is concentrated, not broad. Holding the parent fixed, the model generates no positive population-wide excess-grandparent effect: the partial rank correlation between grandchild and grandparent wealth, controlling for the parent, is -0.08. Grandparent wealth predicts grandchild wealth almost entirely *through* the parent. The distinctive direct channel operates only for the small group with a wealthy *living* grandparent: among the newborns whose living grandparent holds more than $2\chi_b$ (about 0.1% of the cohort, as in the benchmark), 55% reach the top 1 percent by age 65, against a 1% baseline, reproducing the conditional attainment of Table 13. This is precisely the generation-skipping mechanism of the paper: a concentrated effect among the heirs of wealthy grandparents rather than a smooth grandparent gradient across the distribution.

Comparison with the data. Since the model's grandparent effect is untargeted, it can be taken directly to the data. Using the PSID biological genealogy (FIMS) and wealth panel, I compute the Spearman rank correlation of grandparent and grandchild net worth, ranking each generation over its adult years and measuring a grandchild only when it heads its own household, so co-residence does not transmit wealth mechanically.[62] The empirical G-G rank correlation is 0.28 (about 7,100 lineages) and the parent–child correlation 0.39 (about 12,600), the latter squarely within the range Charles and Hurst (2003) report. Both lie inside the model's simulated range in Table 14 (G-G up to 0.29, parent–child up to 0.56), and the ordering, parent–child stronger than G-G, holds in model and data alike. In transition terms, a grandchild whose grandparent is in the top wealth quintile reaches the top quintile with probability 36 percent, against a 20 percent baseline. The comparison is coarse: the PSID does not oversample the wealthy, so the top-1-percent attainment that defines persistence (Table 13) cannot be recovered from survey data. Even so, the model reproduces grandparent wealth persistence in the data without having been asked to.

[62]Without this restriction the correlation is spuriously inflated (to 0.65) since co-resident grandchildren are recorded with their ancestor's household net worth; the restriction is the wealth analogue of the independent-household requirement standard in intergenerational income studies.

References

Adermon, Adrian, Mikael Lindahl, and Daniel Waldenström. 2018. "Intergenerational wealth mobility and the role of inheritance: Evidence from multiple generations." *The Economic Journal*, 128(612): F482–F513.

Bach, Laurent, Laurent E Calvet, and Paolo Sodini. 2020. "Rich pickings? Risk, return, and skill in household wealth." *American Economic Review*, 110(9): 2703–2747.

Barone, Guglielmo, and Sauro Mocetti. 2021. "Intergenerational mobility in the very long run: Florence 1427–2011." *The Review of Economic Studies*, 88(4): 1863–1891.

Benabou, Roland. 2002. "Tax and education policy in a heterogeneous-agent economy: What levels of redistribution maximize growth and efficiency?" *Econometrica*, 70(2): 481–517.

Benhabib, Jess, Alberto Bisin, and Ricardo T Fernholz. 2022. "Heterogeneous dynasties and long-run mobility." *The Economic Journal*, 132(643): 906–925.

Benhabib, Jess, Alberto Bisin, and Shenghao Zhu. 2011. "The distribution of wealth and fiscal policy in economies with finitely lived agents." *Econometrica*, 79(1): 123–157.

Board of Governors of the Federal Reserve System. 1989–2022. "Survey of Consumer Finances." Washington, DC: Board of Governors of the Federal Reserve System. https://www.federalreserve.gov/econres/scfindex.htm (accessed [MONTH YEAR]).

Bricker, Jesse, Sarena Goodman, Alice Henriques Volz, and Kevin B Moore. 2020. "Wealth and Income Concentration in the SCF: 1989–2019." Board of Governors of the Federal Reserve System FEDS Notes.

Cagetti, Marco, and Mariacristina De Nardi. 2006. "Entrepreneurship, frictions, and wealth." *Journal of political Economy*, 114(5): 835–870.

Cagetti, Marco, and Mariacristina De Nardi. 2009. "Estate taxation, entrepreneurship, and wealth." *American Economic Review*, 99(1): 85–111.

Castaneda, Ana, Javier Diaz-Gimenez, and Jose-Victor Rios-Rull. 2003. "Accounting for the US earnings and wealth inequality." *Journal of political economy*, 111(4): 818–857.

Charles, Kerwin Kofi, and Erik Hurst. 2003. "The correlation of wealth across generations." *Journal of Political Economy*, 111(6): 1155–1182.

Clark, Gregory. 2014. *The Son Also Rises: Surnames and the History of Social Mobility.* Princeton University Press.

De Nardi, Mariacristina. 2004. "Wealth inequality and intergenerational links." *The Review of Economic Studies*, 71(3): 743–768.

De Nardi, Mariacristina, and Fang Yang. 2016. "Wealth inequality, family background, and estate taxation." *Journal of Monetary Economics*, 77: 130–145.

Elinder, Mikael, Oscar Erixson, and Daniel Waldenström. 2018. "Inheritance and wealth inequality: Evidence from population registers." *Journal of Public Economics*, 165: 17–30.

Fagereng, Andreas, Luigi Guiso, Davide Malacrino, and Luigi Pistaferri. 2016. "Heterogeneity in returns to wealth and the measurement of wealth inequality." *American Economic Review*, 106(5): 651–55.

Fagereng, Andreas, Luigi Guiso, Davide Malacrino, and Luigi Pistaferri. 2020. "Heterogeneity and persistence in returns to wealth." *Econometrica*, 88(1): 115–170.

Gabaix, Xavier, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll. 2016. "The dynamics of inequality." *Econometrica*, 84(6): 2071–2111.

Gale, William G, Jr. James R. Hines, and Joel Slemrod. 2001. *Rethinking the Estate and Gift Tax.* Brookings Institution Press.

Gokhale, Jagadeesh, Laurence J Kotlikoff, James Sefton, and Martin Weale. 2001. "Simulating the transmission of wealth inequality via bequests." *Journal of Public economics*, 79(1): 93–128.

Guo, Yanran. 2022. "Inheritance, wealth distribution, and estate taxation." Working Paper.

Hendricks, Lutz. 2007. "Retirement wealth and lifetime earnings." *International Economic Review*, 48(2): 421–456.

Hubmer, Joachim, Per Krusell, and Anthony A Smith Jr. 2021. "Sources of US wealth inequality: Past, present, and future." *NBER Macroeconomics Annual*, 35(1): 391–455.

Huggett, Mark. 1996. "Wealth distribution in life-cycle economies." *Journal of Monetary Economics*, 38(3): 469–494.

Internal Revenue Service. 2001. "Statistics of Income: Estate Tax Statistics." Washington, DC: US Department of the Treasury. https://www.irs.gov/statistics/soi-tax-stats-estate-tax-statistics (accessed [MONTH YEAR]).

Kaplan, Steven N, and Joshua D Rauh. 2013. "Family, education, and sources of wealth among the richest Americans, 1982–2012." *American Economic Review*, 103(3): 158–162.

Kartashova, Katya. 2014. "Private equity premium puzzle revisited." *American Economic Review*, 104(10): 3297–3334.

Kartashova, Katya, and Xiaoqing Zhou. 2021. "Wealth inequality and return heterogeneity during the COVID-19 pandemic." Federal Reserve Bank of Dallas Working Paper.

Kindermann, Fabian, and Dirk Krueger. 2021. "High Marginal Tax Rates on the Top 1? Lessons from a Life Cycle Model with Idiosyncratic Income Risk." *American Economic Journal: Macroeconomics*.

Kuhn, Moritz, Moritz Schularick, and Ulrike I Steins. 2020. "Income and wealth inequality in America, 1949–2016." *Journal of Political Economy*, 128(9): 3469–3519.

Lusardi, Annamaria, and Olivia S. Mitchell. 2014. "The economic importance of financial literacy: Theory and evidence." *Journal of Economic Literature*, 52(1): 5–44.

Lusardi, Annamaria, Pierre-Carl Michaud, and Olivia S. Mitchell. 2017. "Optimal financial knowledge and wealth inequality." *Journal of Political Economy*, 125(2): 431–477.

Moskowitz, Tobias J, and Annette Vissing-Jørgensen. 2002. "The returns to entrepreneurial investment: A private equity premium puzzle?" *American Economic Review*, 92(4): 745–778.

National Center for Health Statistics. 2004. "United States Life Tables, 2001." National Vital Statistics Reports, vol. 52, no. 14. Hyattsville, MD: National Center for Health Statistics. https://www.cdc.gov/nchs/data/nvsr/nvsr52/nvsr52_14.pdf (accessed [MONTH YEAR]).

Nekoei, Arash, and David Seim. 2023. "How do inheritances shape wealth inequality? Theory and evidence from Sweden." *The Review of Economic Studies*, 90(1): 463–498.

Nishiyama, Shinichi. 2002. "Bequests, inter vivos transfers, and wealth distribution." *Review of Economic Dynamics*, 5(4): 892–931.

Panel Study of Income Dynamics. 2019. "Public Use Dataset." Produced and distributed by the Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI. https://psidonline.isr.umich.edu (accessed [MONTH YEAR]).

Pfeffer, Fabian T, and Alexandra Killewald. 2018. "Generations of advantage. Multigenerational correlations in family wealth." *Social Forces*, 96(4): 1411–1442.

Piketty, Thomas. 2014. *Capital in the Twenty-First Century.* Cambridge, MA:Harvard University Press.

Tauchen, George. 1986. "Finite state markov-chain approximations to univariate and vector autoregressions." *Economics letters*, 20(2): 177–181.

Van Rooij, Maarten, Annamaria Lusardi, and Rob Alessie. 2011. "Financial literacy and stock market participation." *Journal of Financial Economics*, 101(2): 449–472.

Wolff, Edward N, and Maury Gittleman. 2014. "Inheritances and the distribution of wealth or whatever happened to the great inheritance boom?" *The Journal of economic inequality*, 12(4): 439–468.

Wu, Chunzan, and Dirk Krueger. 2021. "Consumption insurance against wage risk: Family labor supply and optimal progressive income taxation." *American Economic Journal: Macroeconomics*, 13(1): 79–113.

Xavier, Inês. 2021. "Wealth inequality in the US: the role of heterogeneous returns." *Available at SSRN 3915439*.