Consumption, Savings, and the Distribution of Permanent Income*

Ludwig Straub
Harvard University and NBER

June 5, 2019

Abstract

Rising inequality in the permanent component of labor income, henceforth *permanent income*, has been a major force behind the secular increase in U.S. labor income inequality. This paper explores the macroeconomic consequences of this rise. First, I show that in many common macroeconomic models—including precautionary savings models—consumption is a *linear* function of permanent income. This implies that macroeconomic aggregates are independent of permanent income inequality. Motivated by this neutrality result, I provide new evidence from U.S. household panel data that working-life consumption is in fact *concave* in permanent income, rather than linear, with an elasticity of around 0.7. To quantify the effects of concavity, I extend a canonical precautionary savings model to include non-homothetic preferences, capturing that permanent-income rich households save disproportionately more than their poor counterparts. The model suggests that the U.S. economy is far from neutral. In the model, the rise in U.S. permanent labor income inequality since the 1970s has caused: (a) a decline in real interest rates of 1%; (b) an increase in the wealth-to-GDP ratio of 30%; (c) a fast rise in wealth inequality. The model predicts these trends will continue.

Keywords: Inequality; permanent income; interest rates; wealth inequality; non-homothetic preferences

JEL codes: D31, D52, E21, E43

*I am indebted to Iván Werning, Jonathan Parker, Robert Townsend, and Alp Simsek for their continuous guidance and support during my PhD at MIT. I am very grateful to Daron Acemoglu, Mark Aguiar, Marios Angeletos, Adrien Auclert, Mark Bils, Mariacristina De Nardi, Sebastián Fanelli, Martin Holm, Greg Howard, Erik Hurst, Ernest Liu, Ben Moll, Matthew Rognlie, and Nathan Zorzi for many helpful conversations. I would like to thank numerous seminar participants for comments I received both during and after presentations. I thank the Macro-Financial Modeling Group for financial support. All errors are my own. *Email:* ludwigstraub@fas.harvard.edu

1 Introduction

Labor income inequality has increased substantially in the U.S. over the past few decades (Katz and Murphy 1992; Autor, Katz and Kearney 2008), with the top 10% now earning over 35% of all labor income (Piketty and Saez 2003). A significant share of this increase is the result of rising dispersion in the individual fixed component of labor income, which captures the returns to permanent skills or abilities and which I henceforth refer to as *permanent income*.[1] Indeed, Guvenen, Kaplan, Song and Weidner (2018) show that "newer cohorts enter with much higher inequality than older cohorts, which is the main force behind rising income inequality" (p. 38).[2]

According to many well-known macroeconomic models, shifts in the distribution of permanent income are exactly or approximately *neutral*: macroeconomic aggregates—such as consumption, wealth and interest rates—are independent of permanent income inequality since consumption is a *linear* function of permanent income. While this neutrality result holds almost by construction in models adhering to the permanent income hypothesis (Friedman 1957), it is much broader: even canonical precautionary-savings models (Aiyagari 1994; Carroll 1997; Gourinchas and Parker 2002), which are well known to generate concave consumption functions in current income or liquid assets (Zeldes 1989; Carroll and Kimball 1996), predict a linear consumption function in permanent income, and are therefore neutral.[3]

In this paper, I make two main contributions. First, I show that consumption is a concave function of permanent income, rather than a linear one. I do so by proposing new ways to estimate the elasticity of consumption to permanent income, finding estimates around 0.7—significantly below 1. Second, I demonstrate that the concavity in permanent income is crucial to understand the macroeconomic implications of rising income inequality. I do so by incorporating non-homothetic preferences into a canonical precautionary-savings model to match the estimated concavity. The model suggests that, among other implications, the increase in permanent income inequality since 1970 has pushed equilibrium interest rates down by around 1% through the present day, and is expected to lower interest rates by another 1% going forward (despite assuming stable inequality after 2014).

The first contribution of this paper is to propose new ways to measure the curvature of consumption in permanent income, building on previous work by Friedman (1957), Mayer (1972), and Dynan, Skinner and Zeldes (2004), among others. What distinguishes my work is the use of a large household panel data set—the Panel Study of Income Dynamics (PSID)—which since 1999 includes measures of both total consumption expenditure and income. I estimate a log-linear relationship between consumption and permanent income, which I demonstrate is a good fit to the data. This

[1] In the terminology of this paper, permanent income refers to each individual's fixed effect in log labor income and does not include returns to capital.

[2] Complementing this view, Sabelhaus and Song (2010) and Guvenen, Ozkan and Song (2014) provide evidence that both transitory and persistent shock variances have declined in recent decades; Kopczuk, Saez and Song (2010) and DeBacker, Heim, Panousi, Ramnath and Vidangos (2013) argue that either the variance of persistent shocks or the dispersion in fixed effects has increased. See also Figure 11 in Appendix A.

[3] Among the few exceptions to this are Hubbard, Skinner and Zeldes (1994, 1995) and De Nardi (2004), which are discussed below.

yields the *permanent income elasticity of consumption*, ϕ, which is equal to 1 if consumption is a linear function of permanent income.

The key challenge in estimating ϕ is that permanent income is not directly observable and must be distinguished from income shocks, especially persistent ones. This is important since consumption is naturally smoothed in response to income shocks, so that ignoring income shocks results in attenuation bias in ϕ. I propose two novel solutions to this challenge, which depend on the autocovariance structure of persistent income shocks. If persistent income shocks follow an AR(1) process, ϕ is identified and can be consistently estimated by instrumenting log current income with quasi-differenced future log incomes. If persistent income shocks follow a random walk, ϕ is partially identified, and one can estimate an upper bound using initial incomes when entering the labor market as an instrument. Both approaches suggest that ϕ is either around or below 0.7, in both cases statistically significantly below 1. I supplement these tests with a number of extensions and robustness checks that all yield similar results. Among these are specifications that include proxies for preference or rate-of-return heterogeneity, that deal with private and public transfers, and that are based on different measures of consumption expenditure.

The second contribution of this paper is to demonstrate that modeling deviations from linearity is crucial to understand the quantitative implications of rising permanent income inequality. To this end, I build a quantitative life-cycle model with idiosyncratic income risk and incomplete markets in the tradition of Deaton (1991), Huggett (1996), Carroll (1997) and Gourinchas and Parker (2002). Aside from standard sources of non-neutrality, such as nonlinear tax-and-transfer and social security systems (Hubbard et al. 1994, 1995; Scholz, Seshadri and Khitatrakun 2006), the key elements of the model are two kinds of non-homothetic preferences. The first kind follows the seminal work of De Nardi (2004) and assumes that bequests are treated as a luxury good. The second kind are "life-cycle" non-homothetic preferences, over consumption across periods, rather than over bequests. This latter kind turns out to be the most important source of non-neutrality in the model.

Non-homothetic preferences capture the idea that permanently richer agents save a larger fraction of their income, either for bequests or for other expenses later in life. While the model does not require enumeration of what those expenses are, one may think of college tuition payments for kids, expensive medical treatments later in life, or charitable giving.[4] I calibrate the strength of the life-cycle non-homotheticity to match the elasticity ϕ when estimating the same regressions on artificial panel data simulated from the model; I follow the literature for the calibration of the bequest non-homotheticity. Importantly, I find that models without life-cycle non-homotheticity, even non-neutral ones, cannot rationalize the empirical magnitude of ϕ. Although I do not target any specific moments of the wealth distribution, the model matches the (highly unequal) U.S. wealth distribution quite well.

Using the calibrated economy as a laboratory, I study the implications of rising permanent

[4] The idea of non-homothetic savings behavior goes back at least to Fisher (1930) and Keynes (1936). Fisher (1930) noted that if one person's income is simply a scaled version of another person's in all periods, then "the smaller the income, the higher the preference for present over future" consumption. Keynes (1936) argued that as long as one's "primary needs" are not satisfied, consumption is "usually a stronger motive than the motives towards [wealth] accumulation".

income inequality. I simulate the general equilibrium transitional dynamics from a steady state with the 1970 levels of income inequality to one with the 2014 levels of inequality, which I assume to remain constant after 2014.[5] This exercise allows the model to speak directly to the forces behind three important recent macroeconomic trends: (a) the decline in the real (natural) interest rate since the 1980s (Laubach and Williams 2016), (b) the rising private wealth to GDP ratio (Piketty and Zucman 2015), and (c) the large and rapid increase in U.S. wealth inequality (Saez and Zucman 2016).

Regarding the first macroeconomic trend, I find that the real interest rate declines by around 1 percentage point through 2020, explaining approximately one-third of the decline in the U.S. natural rate since the 1980s. Interestingly, despite the absence of any further increases in income inequality, the model predicts the interest rate will continue declining by another 1% over the next two decades. The reason for this result is intuitive: in the model, the generation entering the labor market today is the first to experience the highest level of permanent income inequality for their entire working lives. By the same token, from today onwards, each cohort will have amassed a larger fortune by the time they retire than previous generations of retirees. This effect causes a large and predictable decline in interest rates going forward.

Concerning the second trend, I find private wealth to GDP to rise by around 30 percentage points through 2020 (also roughly one-third of the rise in the data). Again, I find that the trend will continue, leading to an eventual total increase of 55 percentage points. Finally, the model offers an explanation for the size and speed of the increase in the top 10% wealth share and around two-thirds of the increase in the top 1% wealth share. This is an especially stark contrast to a homothetic model, according to which wealth inequality can only increase as much as income inequality.[6] In sum, the non-homothetic model suggests that rising permanent labor income inequality alone can account for a significant share of three major macroeconomic trends.

There are two limitations of my baseline model, which assumes a simple Cobb-Douglas aggregate production function with frictionless capital. First, the model cannot generate saving via capital gains; instead, it predicts rising wealth levels as the result of the accumulation of physical capital. To explore the robustness of my results, I consider a "Lucas tree" economy, with a fixed capital stock. I show that while making similar aggregate predictions, the Lucas-tree economy generates the entire increase in wealth to GDP as a result of capital gains.[7] Second, the model assumes the entire increase of income inequality to be exogenous shifts in the production function. I relax this assumption by considering an alternative production function with capital-skill complementarity (Krusell, Ohanian, Rios-Rull and Violante 2000; Autor, Levy and Murnane 2003). This is shown to lead to a novel feedback mechanism between greater income inequality, lower interest rates and

[5] The transitional dynamics are computationally non-trivial since the model has a large number of idiosyncratic states, as well as endogenous bequest distributions over which agents have rational expectations. I overcome these difficulties by improving existing algorithms along a number of margins (see Appendix H).

[6] The dynamics of the wealth distribution have recently been investigated theoretically by Gabaix, Moll, Lasry and Lions (2017), and numerically by Hubmer, Krusell and Smith (2016), Kaymak and Poschke (2016), Aoki and Nirei (2017) and Benhabib, Bisin and Luo (2017) among others.

[7] Since capital gains are not part of the personal saving rate as measured in national income accounting, this also shows that the channels in this paper are not tied to a rising personal saving rate.

rising levels of skill-complementary investments (e.g. automation).

Finally, I provide several robustness and sensitivity analyses, using a second, different, non-homothetic life-cycle economy in the spirit of Carroll (2000), as well as alternative calibrations of both models to match estimates of ϕ closer to 1. I find that both models predict similar effects of rising income inequality, and that even economies with permanent-income elasticities of 0.80 or 0.90 are still substantially non-neutral.

Literature. My paper relates to several strands of a vast literature at the intersection of inequality, consumption dynamics, and macroeconomics.[8]

First, my paper contributes to the large empirical literature that tests the permanent income hypothesis (PIH), starting with Friedman (1957). One can group the predictions of the PIH into two conceptually-distinct categories: predictions about *changes* in consumption in response to predictable or unpredictable, transitory or permanent, income *changes*; and predictions about the *level* of consumption in relation to the *level* of the permanent component of income. Throughout the 1950s and 1960s, many economists viewed the second prediction as the "most controversial aspect of the permanent income theory" (Mayer 1972, p.34) and consequently it received relatively more attention. Partly due to data quality issues, however, the evidence remained inconclusive, and the focus of empirical work on the PIH subsequently shifted almost entirely to testing the first set of predictions (Hall 1978; Flavin 1981; Johnson, Parker and Souleles 2006).

The main exception to this is Dynan et al. (2004), henceforth DSZ.[9] Their paper computes saving rates in different datasets and documents that saving rates increase across current income quintiles and when instrumented by lagged or future income, or education. My empirical exercise follows their lead, innovating along two dimensions. First, I use a more recent, higher-quality dataset (PSID, since 1999) that allows me to show that the relationship between consumption and income is roughly log-linear and thus well-described by a single elasticity parameter ϕ. Second, and more importantly, I use the panel aspect of the dataset to develop two new instruments under mild assumptions on the income process. I show that the two instruments either estimate ϕ consistently or estimate an upper bound for ϕ consistently in canonical precautionary savings models. This improves upon the simple instruments used in the literature, such as lagged or future income (which lead to downward-biased results under the assumptions of a canonical linear model), or education (which could be correlated with preferences, income profiles, etc). In fact, my approach allows for additional proxies to try to control for heterogeneity in preferences and returns.[10]

The second main contribution of this paper is the analysis of rising permanent income inequality

[8] For recent surveys and books, see among others Bertola, Foellmi and Zweimuller (2005); Krusell and Smith (2006); Heathcote, Storesletten and Violante (2009); Guvenen (2011); Quadrini and Ríos-Rull (2015); De Nardi, Fella and Yang (2015); Piketty and Zucman (2015); De Nardi and Fella (2017); Attanasio and Pistaferri (2016); Benhabib and Bisin (2017).

[9] For similar approaches see Bozio, Emmerson, O'Dea and Tetlow (2013) for the UK and Alan, Atalay and Crossley (2015) for Canada. In a recent paper, Abbott and Gallipoli (2019) estimate a different but complementary measure of "permanent income" and find an elasticity of consumption to their measure of around 0.80. Fagereng, Holm, Moll and Natvik (2019b) document rising saving rates with wealth in Norwegian registrar data (see Section 6.5).

[10] In a recent paper, Olivi (2019) estimates beliefs by income in a non-parametric consumption-savings model and shows that they can contribute to a nonlinear relationship between consumption and permanent income.

in a non-neutral, incomplete-markets economy. Here, I combine elements from two literatures, one on non-neutral economies and one on rising income inequality. Seminal work on non-neutral economies, by Hubbard et al. (1994, 1995) and De Nardi (2004), argues that a realistic social safety net and non-homothetic bequest motives, respectively, can increase wealth inequality compared to a homothetic model.[11] My paper follows their lead but argues that one needs more concavity to match the empirical estimate for ϕ. I therefore include non-homothetic preferences over consumption within the life-cycle as well, which, when calibrated to match the empirical evidence, generate a wealth distribution that fits the recent U.S. distribution quite well.

In addition, my paper builds on a recent literature that studies the quantitative effects of income inequality in (mostly) neutral economies. Auclert and Rognlie (2018) show how greater inequality has aggregate effects that crucially depend on the types of incomes (transitory, persistent or permanent) that become more unequal. Heathcote, Storesletten and Violante (2010) investigate the human capital investment and family labor supply implications of rising income inequality. Kaymak and Poschke (2016) and Hubmer et al. (2016) consider the effects of rising income inequality on wealth inequality. Favilukis (2013) studies the joint implications of greater income risk, relaxed borrowing constraints and an increase stock market participation rate. Krueger and Perri (2006) argue that insurance against shocks may improve with greater within-group income risk.[12] The key distinguishing feature of my paper is that I focus on rising *permanent* income inequality, arguably among the most important driver of rising income inequality in the US.

Finally, there is a long tradition of studying and identifying the degree to which consumption is insured from changes in incomes. Of particular relevance and inspiration to my research are Blundell, Pistaferri and Preston (2008) and Kaplan and Violante (2010). In a landmark result, Blundell et al. (2008) describe a way in which one can estimate the degrees to which consumption responds to persistent or transitory income shocks. Kaplan and Violante (2010) compare these estimates to those implied by canonical precautionary-savings models. Relative to these papers, I focus on the relationship between the level of consumption and the level of permanent income, rather than on changes. However, this paper shares the spirit of these papers by identifying important moments in similar panel data on consumption and income, and using them to inform microfounded consumption-savings models.

Layout. Section 2 presents a simple two-agent example of how a concave consumption function in permanent income can be modeled and what its qualitative effects are. Section 3 introduces a canonical precautionary-savings model and explains under what assumptions the consumption function is linear in permanent income. I empirically estimate the curvature of that function in Section 4. Extending the canonical model, Section 5 then relaxes the linearity assumptions, introducing the non-homothetic model, whose general equilibrium effects are studied in Section 6. Section 7 concludes. The appendix contains all proofs, as well as additional empirical and quantitative results.

[11] See also Huggett and Ventura (2000), Scholz et al. (2006) and De Nardi and Yang (2014). Kumhof, Rancière and Winant (2015) study income inequality in a two-agent economy with non-homothetic preferences.

[12] Holm (2018) studies the differential effects of monetary policy during times with increased persistent income risk.

2 Permanent Income Inequality in a Two-Agent Model

I begin by studying a stylized two-agent OLG framework to illustrate the main effects of rising permanent income inequality in models with linear and concave consumption functions.

2.1 Linear and concave consumption functions

Consider a dynasty of 1-period lived generations that earn a constant stream of wage incomes $w > 0$ and can save in risk-free bonds paying a constant interest rate $R > 1$. They face the following decision problem: each period $t = 0, 1, 2, \ldots$ the current generation solves

$$\max_{c_t, a_{t+1}} u(c_t) + \beta U(a_{t+1}) \tag{1}$$

$$c_t + R^{-1} a_{t+1} \leq a_t + w. \tag{2}$$

Here, a_t denotes the value of financial wealth held by the dynasty at the beginning of period t, $a_t + w$ can be regarded as the dynasty's "cash on hand", c_t denotes the consumption choice in period t, and a_{t+1} is the bequest left to the subsequent generation. Observe that in this model, w is the dynasty's permanent income level, where I use the term permanent income, as explained in the introduction, to denote the fixed-effect component of labor income. In fact, this model is so stylized that there is no other component of labor income, that is, no income shocks, no life-cycle earnings profile, and so on.

The choice of the two utility functions is critical for this model: the flow utility u and the joy-of-giving utility U. For simplicity, I assume that both have a constant elasticity,

$$u(c) = \frac{(c/o)^{1-\sigma} - 1}{1-\sigma} \qquad U(a) = \frac{(a/o)^{1-\Sigma} - 1}{1-\Sigma}, \tag{3}$$

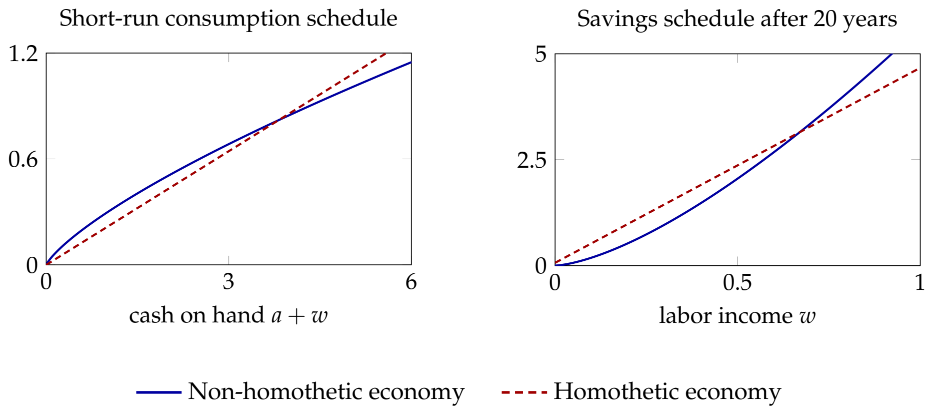
but the two (inverse) elasticities $\sigma, \Sigma > 0$ are allowed to differ. $o > 0$ is a normalization parameter, which can be used to retain aggregate scale invariance.

Differing elasticities σ, Σ allow the model to capture the fact that richer dynasties may have a greater propensity to save. To see this, observe that (1) can be reinterpreted as a simple utility maximization problem with two "goods"—consumption c_t and savings a_{t+1}. In this problem when saving is a "luxury good"—that is, its income elasticity is greater than one—a richer dynasty decides to save a larger fraction of its wealth.[13] With utilities as in (3), this is the case if $\Sigma < \sigma$, so that the utility over savings (bequests) is more linear than the utility over consumption.

Figure 1 illustrates two outcomes of the utility maximization problem (1). Panel (a) shows the optimal *short-run* consumption choice c_t as a function of current cash on hand $a_t + w$. Panel (b) shows the optimal *long-run* asset position (after 20 years) as a function of permanent income w.

[13] See also Strotz (1955), Atkinson (1971) and Blinder (1975) for early deterministic life-cycle models with non-homothetic utility over bequests.

Figure 1: Stylized model: Consumption and savings schedules.



Note. This figure shows consumption and asset choices in the stylized model. Panel (a) shows the optimal consumption policy as a function of cash on hand. Panel (b) shows the asset position after 20 years as a function of permanent income. "Non-homothetic" refers to a model where saving has an income elasticity greater than 1. "Homothetic" refers to a model where saving has an income elasticity of 1.

In both panels the agent starts with the average wealth and income position in the economy. The panels contrast two cases with each other: the *homothetic* case, where $\Sigma = \sigma$, and the *non-homothetic* case, where $\Sigma < \sigma$ and savings are treated as a luxury good. While optimal short-run consumption and long-run savings schedules are both linear in the homothetic case, consumption is concave and savings convex in the non-homothetic case.[14]

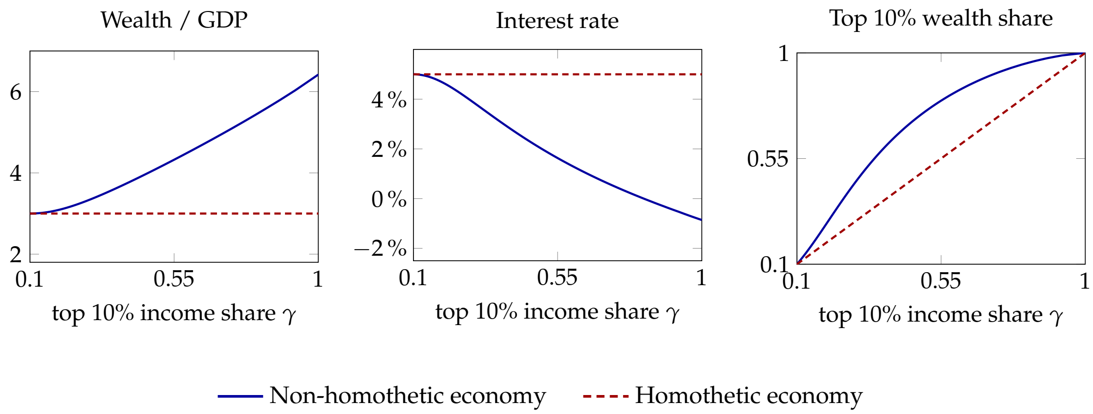
The consumption schedule turns out to be well approximated by a simple power function $c_t \approx k(a_t + w)^{\phi}$ for large values of cash on hand, where the exponent is given by the ratio of the elasticities, $\phi = \Sigma/\sigma$. The elasticity ϕ takes a central role in this paper, as it succinctly characterizes the degree of concavity in consumption as a function of permanent income.

2.2 General equilibrium effects of greater inequality

Having introduced the decision problem of a single dynasty, I now describe the effects of shifts in income inequality between two dynasties. Assume an economy is populated by two dynasties, both with the exact same preferences (1). The only difference between these dynasties is their permanent income level: one dynasty, the "rich" r, is assumed to have a strictly greater permanent (labor) income than the other, the "poor" p, that is, $w^r > w^p$. Assume that the population share of the rich dynasty is $\mu = 0.10$. The share of labor income earned by the rich dynasty, $\gamma \equiv \mu w^r/(\mu w^r + (1-\mu)w^p)$, will serve as the measure of inequality in this economy. The economy is more unequal the further away γ is from μ. To close the model, assume that income inequality γ is generated by a Cobb-Douglas aggregate production function in capital K and the dynasties' labor L^r and L^p, $Y = K^{\alpha}(L^r)^{(1-\alpha)\gamma}(L^p)^{(1-\alpha)(1-\gamma)}$. Shifts in income inequality γ are generated by shifts in a parameter of the production function.

[14]It is worth emphasizing that this model generates a consumption function that is concave in *permanent* (labor) income, which, as I argue in Section 3 is not the case in a canonical precautionary-savings model, where consumption is concave

Figure 2: Stylized model: Neutrality and non-neutrality in general equilibrium.



Imagine that the economy is initially perfectly equal, $\gamma = \mu$, and consider an increase in inequality to some $\gamma > \mu$. The three panels in Figure 2 illustrate the long-run steady-state outcomes for the homothetic and the non-homothetic economies. Due to the linear consumption and savings schedules, the homothetic economy is *neutral*, meaning all aggregate quantities—e.g. output, wealth, and interest rates—are unaffected. Wealth inequality increases, at the same rate as income inequality, reflecting the proportionality of assets and income in the model. By contrast, the non-homothetic economy is *non-neutral.* Aggregate wealth increases with greater inequality, while interest rates fall. Wealth inequality rises faster than one-for-one with income inequality.

Albeit intuitive, the model in this section is very stylized. The remainder of this paper moves step-by-step towards a quantification of this section's qualitative results. The first step is a characterization of the consumption function in permanent income in richer, less stylized models.

3 Linearity in Precautionary-Savings Models

Linearity of the consumption function in permanent income was an immediate consequence of $\Sigma = \sigma$ in the two-agent model. Yet, how general is this property? In this section, I introduce a general framework which nests a sizable set of life-cycle and infinite-horizon precautionary savings models and allows sharp characterization of when the model is linear and when it is not.

3.1 Setup

Time is discrete, $t \in \{0, 1, \ldots\}$, and there is no aggregate risk. The model is an overlapping generations (OLG) version of an Aiyagari (1994) model. It allows for an endogenous bequest distribution which agents receive at the time of their parents' death. I focus on the steady state of the economy.

Birth, death and skills. The economy is populated by a continuum of mass 1 of agents at all times, each of whom is assigned a permanent type in a finite set $\mathcal{S} \subset \mathbb{N}$. One can think of a permanent

in *current* income or assets, but not permanent income.

type $s \in \mathcal{S}$ as innate skill or ability. Agents with skill s are endowed with units of that skill and make up a constant share $\overline{\mu}_s \in [0,1]$ of the population. To allow for overlapping generations, I assume that there is a constant inflow and outflow of agents at rate $\overline{\delta} \geq 0$, where zero is included. An agent's age is indexed by $k \in \mathbb{N}$. With an OLG structure, $\overline{\delta} > 0$, each agent has a single offspring that is born at fixed parental age $k_{born} > 0$, and dies with certainty at age $K_{death} \in \mathbb{N} \cup \{\infty\}$. Henceforth I assign all agents ever to live in this economy the unique label $i \in [0, \infty)$.

Production. There is a single consumption good, which is produced using a neoclassical aggregate production function $Y = F(K, \{L_s\}_{s \in \mathcal{S}})$ from K units of capital and L_s efficiency units of skill s. I assume that F is Cobb-Douglas, that is, $F = AK^\alpha \prod_s L_s^{(1-\alpha)\gamma_s}$, where $\gamma_s > 0$ is the labor income share of skill s. I denote by w_s the price of an efficiency unit of skill s and by r the real interest rate, so that in equilibrium, $Y = (r+\delta)K + \sum_{s \in \mathcal{S}} w_s L_s$. As before, the main comparative statics exercise is a shift in the distribution of labor income shares $\{\gamma_s\}$, which induces a shift in the distribution of skill prices $\{w_s\}$, since $w_s = \gamma_s Y / L_s$.

Government. There is a government that levies a constant tax rate $\tau^b \in [0,1]$ on any bequests and applies to all agents a possibly age-dependent income tax function $T_k(y^{pre})$, where $k \geq 1$ is an agent's age, and y^{pre} an agent's pre-tax income.[15] The government holds a level of government debt B and chooses its spending G to balance its budget.

Agents. In the life-cycle case, an agent is born at some date t_0 with zero asset holdings. In the infinite-horizon case, agents are already alive at date $t = 0$. An agent faces idiosyncratic shocks captured by a Markov chain $z_t \in \mathcal{Z}$ with transition probabilities $\Pi_{zz'}$ from state z to state z', initialized at date t_0 with a fixed initial distribution $\{\pi_z\}$. The idiosyncratic shocks determine the agent's stochastic endowment of efficiency units of skill s, which is given by a function $\Theta_{t-t_0}(z_t)$ at time t. The agent's income is then $y_t^{pre} = \Theta_{t-t_0}(z_t)$ before taxes and $y_t = y_t^{pre} - T_{t-t_0}(y_t^{pre})$ after taxes. I assume the function $\Theta_k(z)$ is normalized such that it averages to 1 when averaged over the whole population of agents and over all idiosyncratic states. An agent dies after age k with probability $\delta_k \in [0,1]$. In case of death, the agent is allowed to derive utility over bequests. I denote by $u_k(c)$ the agent's possibly age-dependent, per-period utility over the consumption good, and by $U(a)$ the utility from bequeathing asset position a.

Bequests. It is assumed that each agent of skill s has an offspring with skill s', where s' is randomly drawn from a transition matrix $P_{ss'}$. The process for skills is assumed to be independent of $\{z_t\}$. Bequests are not necessarily received at the beginning of life, so it is important to specify each agent's belief about the distribution of bequests they may receive later on. I assume that $\varphi \in \{0,1\}$ is an indicator for whether an agent has already received a bequest and that $v(\cdot|s,k,\varphi)$ denotes the probability distribution over bequest sizes to be received next period conditional on age k, skill s, and indicator φ. Formally, $v(\cdot|s,k,\varphi)$ is defined over the product space of bequests and bequest indicators, $\mathbb{R}_+ \times \{0,1\}$, together with the Borel σ-algebra. Agents use $v(\cdot|s,k,\varphi)$ to predict bequests they may receive based on their own skill s and age k.[16]

[15] I allow for age-dependence to nest the case where the government provides a social security and pension system, in which case T_k would be negative for retired individuals.

[16] This assumption is not important for the linearity results in this section and is made to simplify the computation later

Agent's optimization problem. Taken together, an agent born at date t_0 with skill s solves the following optimization problem,

$$V_{k,s}(a,z,\varphi) = \max\ u_k(c) + \beta(1-\delta_k)\mathbb{E}_{z,\varphi} V_{k+1,s}(a'+b',z',\varphi') + \beta\delta_k U(a') \tag{4}$$

$$c + \frac{1}{1+r}a' \leq a + \Theta_k(z)w_s - T_k(\Theta_k(z)w_s)$$

$$(b',\varphi') \sim v(\cdot|s,k,\varphi)$$

$$a' \geq 0.$$

The definition of equilibrium is standard and relegated to Appendix C.1.

Similar to Section 2, I now show three sets of results in this economy. Under certain assumptions, consumption functions are linear in permanent income; consumption and wealth inequality move one-to-one with income inequality; and the aggregate economy in partial and general equilibrium is unaffected by changes in permanent income inequality.

3.2 Assumptions for linearity

To state the results, I formally introduce three necessary assumptions to obtain neutrality. Each of these is relaxed in Section 5 to explore their respective roles in generating a concave consumption function. The first assumption is that utility functions over consumption and bequests each have a constant elasticity; and moreover, that elasticities are the same and not age-dependent.

Assumption 1 (Homothetic utility functions)**.** *(i) The per-period utility function $u_k(c)$ is homogeneous with a constant elasticity of intertemporal substitution, that is, $u_k(c) = \frac{c^{1-\sigma}}{1-\sigma}$ for some $\sigma > 0$. (ii) The bequest utility function $U(a)$ is homogeneous with the same elasticity, that is, $U(a) = \kappa\frac{a^{1-\sigma}}{1-\sigma}$ for some parameter $\kappa \geq 0$.*

This assumption is the reason I refer to this benchmark economy as *homothetic*. The second assumption is that the income tax schedule is linear.

Assumption 2 (Linear tax schedule)**.** *The income tax function is linear in pre-tax income, that is, $T_k(y^{pre}) = \tau_k y^{pre}$ for some $\tau_k \in \mathbb{R}$, for each $k \in \{1,\ldots,K_{death}\}$.*

This assumption restricts both income taxes and any social security payments to be entirely linear. As I discuss below, however, richer, progressive tax-and-transfer schedules can still be allowed without breaking the linearity result below. The final assumption is that bequests play no redistributional role, that is, no rich person leaves any wealth to a differently-skilled offspring.

Assumption 3 (No redistribution through bequests)**.** *One of the following three assumptions is satisfied: (i) Skills are perfectly persistent, that is, the transition matrix $P_{ss'}$ is the identity, $P_{ss'} = 1$ if $s = s'$ and $P_{ss'} = 0$ otherwise; (ii) There are no bequests, that is, the model is an infinite-horizon economy or a perfectly deterministic life cycle model without bequest utility; (iii) Bequests are perfectly taxed, $\tau_b = 1$.*

on. For a similar assumption, see De Nardi (2004).

A commonly-used fourth alternative, not modeled here, is the assumption of a perfect annuities market (and no preferences for bequests). In addition to those three economic assumptions, I make a fourth, technical one to ensure that there is a unique wealth distribution for a given level of the interest rate r and wages $\{w_s\}$ (formally stated in Appendix C.2).

Having made these (arguably standard) assumptions, I can now characterize the micro implications of steady state equilibria in this economy.

3.3 Linearity and neutrality

I start by showing that the implied consumption policy function is linear in permanent income.

Proposition 1 (Linear consumption function)**.** *Under Assumptions 1–4, the consumption policy function* $\mathbf{c}_{k,s}(a, z, \varphi)$ *is homogeneous of degree 1 in current assets and permanent income,*

$$\mathbf{c}_{k,s'}\left(a\frac{w_{s'}}{w_s}, z, \varphi\right) = \frac{w_{s'}}{w_s}\mathbf{c}_{k,s}(a, z, \varphi). \tag{5}$$

Thus, the distribution of consumption choices $\{c_{ik}\}$ *across agents* i *of age* k *in skill group* $s(i)$ *scales in* $w_{s(i)}$*. In logs,*

$$\log c_{ik} = const_k + \log w_{s(i)} + \epsilon_{ik}, \tag{6}$$

where $const_k \in \mathbb{R}$ *and* $\mathbb{E}[\epsilon_{ik}|k, s] = 0$*. Moreover, the after-tax income process satisfies*

$$\log y_{ik} = \widetilde{const}_k + \log w_{s(i)} + \tilde{\epsilon}_{ik}, \tag{7}$$

where $\widetilde{const}_k \in \mathbb{R}$*, and* $\mathbb{E}[\tilde{\epsilon}_{ik}|k, s] = 0$*.*

Proposition 1 shows that if one agent has a $w_{s'}/w_s$ times greater permanent income and asset position than another, the former also consumes $w_{s'}/w_s$ as much.[17] Since asset positions scale in permanent income, this implies linearity of consumption in permanent income. In particular, motivates a simple log-linear specification that I use to measure the degree of non-linearity in Section 4. The next result directly follows from the linearity of the consumption function.

Proposition 2 (Consumption and wealth inequality under linearity)**.** *Under Assumptions 1–4, the variances of log consumption and log wealth move one-to-one with the variance of log permanent income,*

$$\begin{aligned} Var_{i,k} \log c_{ik} &= const + Var_s \log w_s \\ Var_{i,k} \log (a_{ik} + y_{ik}) &= \widetilde{const} + Var_s \log w_s, \end{aligned}$$

for all t*, where the constants are independent of the distribution of permanent incomes* $\{\log w_{s(i)}\}$*.*

[17] An interesting implication of this is that the distribution of MPCs is independent of each skill s. This follows from differentiating the equation for $\mathbf{c}_{k,s}(a, z, \varphi)$ in Proposition 1 with respect to a. I state and prove this result formally in Appendix C.4.

Despite its simplicity, this is a striking result, especially in light of the recent U.S. experience. While there is still some debate about how much consumption inequality rose compared to income inequality (Attanasio and Pistaferri 2016), there is clear evidence that wealth inequality has significantly outpaced income inequality in recent decades (Piketty and Saez 2003; Saez and Zucman 2016).

So far, I have focused on the micro predictions of the model, which are entirely independent of the supply side of the economy. I now turn to the macro predictions. To do this, I consider shifts in the distribution of labor income $\{\gamma_s\}$. This leads to the following general equilibrium result.

Proposition 3 (Neutrality)**.** *Suppose Assumptions 1–4 hold. Then, aggregate consumption and savings are linear functions of average permanent income $\mathbb{E}_s w_s$,*

$$C = \kappa_C \times \mathbb{E}_s w_s \quad \textit{and} \quad A = \kappa_A \times \mathbb{E}_s w_s$$

where $\kappa_A, \kappa_C > 0$ are two constants that do not depend on $\{\gamma_s\}$. It follows that any redistribution of permanent incomes through a change in labor income shares $\{\gamma_s\}$ leaves all aggregate quantities unchanged. This means that permanent income inequality is irrelevant for aggregate consumption, savings, investment, tax revenues, bequests, asset prices, and the interest rate.

Any individual's consumption is linear in permanent incomes w_s, so the distribution of w_s is irrelevant for aggregate consumption and savings. Therefore, all aggregate quantities are unchanged in general equilibrium.

3.4 Discussion and relation to literature

The key distinction between the aggregation result in this section and previous ones is the focus on permanent incomes—that is, individual fixed effects in log labor income. For instance, in Constantinides and Duffie (1996) there are only permanent shocks (here income shocks are very general), and there is no trade in equilibrium (whereas here there is). The approximate aggregation result in Krusell and Smith (1998) is about the asymptotic linearity of the consumption function out of assets for large levels of assets, not the linearity of consumption as a function of permanent income—in fact, in the above economy, consumption can be an arbitrarily-curved function of assets and still be linear in permanent income. Finally, the results in this section are related to the common trick in homothetic life-cycle models of dividing by the current random-walk component in income (Carroll 1997); the results here are different since they apply to arbitrary income processes, and since consumption is not necessarily linear in the random-walk component of income (Carroll 2009), while it is in permanent income.[18]

The linearity result has been stated in a fairly general way, but not as general as possible. Similar

[18] The results here also contribute benchmark results on the effects of greater permanent income inequality on consumption and wealth inequality, as well as on macroeconomic aggregates.

results hold with progressive tax systems,[19] endogenous labor supply, habit formation, aggregate risk, or non-separable preferences (e.g. Epstein-Zin preferences).[20]

One limitation of this result is that it is a long-run result. If changes in the distribution of skill prices hit currently-living generations in mid-life, rather than only affecting new cohorts, there will be a period of adjustment to the new set of skill prices. In Section 6, I explore this effect quantitatively and find it to be small. Another limitation is that the result no longer exactly applies when initial assets or borrowing constraints are nonzero *and* do not scale with an agent's permanent income level w_s. This was also explored quantitatively and found to be insignificant. Finally, I assume joy-of-giving preferences. If preferences were altruistic *and* Assumption 3 were relaxed to allow for imperfect skill persistence, the result no longer holds exactly, as parents essentially treat bequests as a luxury good due to mean reversion in skills. In reduced form, this precisely corresponds to a joy-of-giving utility U with a lower elasticity than that of u, which will be an integral part of the quantitative model in Section 5.

4 Measuring (Non-)Linearity

I introduced a general linear model in the previous section. Yet, how "linear" is actual consumption behavior? Motivated by Proposition 1, I approach this question by estimating the following system of equations,

$$\log c_{it} = \mathbf{X}_{it}'\beta + \phi \log w_{s(i)} + \epsilon_{it} \tag{8a}$$

$$\log y_{it} = \tilde{\mathbf{X}}_{it}'\tilde{\beta} + \log w_{s(i)} + \eta_{it} + \psi_{it} + \nu_{it}. \tag{8b}$$

Here, $c_{it}, y_{it}, w_{s(i)}$ denote current consumption, current income and permanent income as before, $\mathbf{X}_{it}, \tilde{\mathbf{X}}_{it}$ are sets of controls, η_{it} is a persistent income shock, ψ_{it} is a transitory income shock, and ν_{it} is measurement error. The object of interest is the slope coefficient ϕ of permanent income on consumption. Throughout this section, I will denote by $\hat{y}_{it} \equiv \log y_{it} - \tilde{\mathbf{X}}_{it}'\hat{\tilde{\beta}}$ log income residuals after partialing out observable controls $\tilde{\mathbf{X}}_{it}$; by $\hat{w}_i \equiv \log w_{s(i)}$ log permanent income residuals of agent i; and by $\hat{c}_{it} \equiv \log c_{it}$ log consumption.

4.1 Data description

Overview. I use data from the 1999 – 2013 waves of the Panel Study of Income Dynamics (PSID). The PSID began in 1968 and is currently the longest running longitudinal household survey in the world. The survey was conducted annually until 1996, and biennially since 1997. It is known for having comparatively low attrition rates and relatively high response rates (Becketti, Gould, Lillard and Welch 1988; Andreski, Li, Samancioglu and Schoeni 2014).

[19]This works as long as post-tax incomes are a power function of pre-tax incomes, which holds relatively well in U.S. data (Benabou 2000; Heathcote, Storesletten and Violante 2017). In that case, consumption is a linear function of *post-tax* incomes (not pre-tax incomes).

[20]These more general results are available upon request.

Consumption expenditure data. Until 1997, the PSID only collected information on specific consumption categories (food, housing and childcare). Since 1999, however, the PSID now collects a much wider set of consumption expenditure data that captures around 70% of the expenditures surveyed in the Consumer Expenditure Survey (CEX) and in the U.S. National Income and Product Accounts (NIPA).[21] These new categories include expenditures on food, housing, mortgages and rents, utilities, transportation and vehicles, education and health care. The largest categories missing in the revised consumption survey are home repairs and maintenance, household furnishing, and clothing. Those were added in a further update in 2005. Since 2005, the PSID consumption data captures almost all categories of the CEX (Andreski et al. 2014).

In the analysis below, I use the longer-running but slightly less comprehensive consumption expenditure data as my baseline measure. As I show in Appendix D.1, moving to the more comprehensive post-2005 consumption measure only has a minor effect on my results. In my baseline measure, I include all available expenditure categories, including durable goods (see Appendix D.1 for results without durables). Since mortgage payments are the sum of imputed rents and accumulation of housing wealth, which is a form of saving, I replace them by imputed rent, as computed by the PSID.[22]

Income data. As income in my baseline regressions, I use post-tax household labor income. This is the right income concept to use for my exercise since it is the best measure of the amount of labor income that households actually have at their disposal. In particular, it takes into account taxes, transfers and changes in labor supply. The income measure consists of labor income of all family members minus taxes (computed using NBER's TAXSIM program). I discuss robustness with respect to alternative income measures in Appendix D.1.

Sample Selection. My baseline sample includes all PSID waves from 1999 to 2013, and consists of all households whose head is between 30 and 65 years old. I exclude households without a single non-missing consumption and income observation, as well as observations with income below 5% of the yearly average income. The PSID waves prior to 1999, when no broad consumption measure is available in the dataset, are only used for their income data. My baseline sample consists of 5,881 distinct households with at least one observation. I discuss several alternative sample choices in Appendix D.1. Throughout, I use PSID's post-1999 longitudinal sample weights.

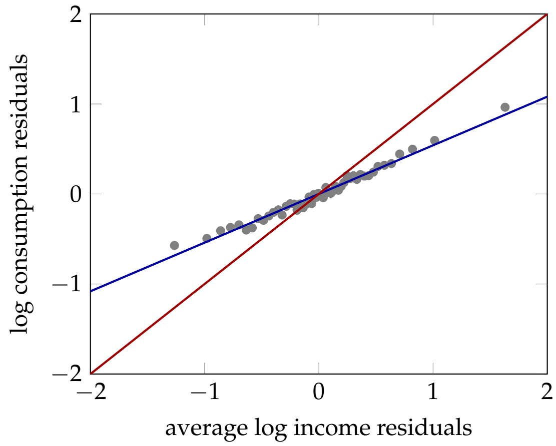
Controls. In my benchmark specifications, I use as controls $\tilde{\mathbf{X}}_{it}$ in the income equation the household head's five-year age bracket, dummies for household size, and year dummies; and as controls $\mathbf{X}_{it}$ in the consumption equation the same controls plus a location dummy to capture heterogeneity in living costs.[23] The results are robust to several other sets of controls, see Section 4.4 and Appendix D.1.

[21] Unlike the CEX, however, the post-1999 PSID consumption data does not suffer from a downward trend relative to the PCE (see, e.g. Blundell, Pistaferri and Saporta-Eksten 2016).

[22] My results are very similar if computing imputed rents as 6% of the house price, as do Blundell et al. (2016) and Poterba and Sinai (2008). When using the comprehensive post-2005 consumption measure, I exclude home repairs and maintenance costs since these are investments.

[23] The location dummy is the interaction of an urban-rural dummy and dummies for the nine Census divisions.

Figure 3: Consumption and average income.



Note. The graph shows consumption and average income in logs for the baseline sample of PSID households. To construct it, log consumption is regressed on controls (year, age, household size, location) and 50 bins for average log income residuals. Log income residuals are obtained by partialing out year, age, and household size dummies and then averaged over a symmetric 9-year interval for each household ($T = 5$). The blue line is the estimated linear relationship with slope ϕ, the red line is the 45° line.

4.2 A first glance at the data

Motivated by Proposition 1, I start by showing results for specifications in which log permanent income $\hat{w}_i$ is proxied for by a simple income average. Even if these specifications turn out to be biased under the linearity assumptions of the benchmark model (see Section 4.3), they are intuitive and set the stage for the more formal econometric investigation in the next section. Consecutive observation periods $t, t+1$ are two years apart in the PSID sample.

I use symmetrically-averaged income residuals as proxies for permanent income, constructed as

$$\overline{y}_{it}^{T} \equiv \frac{1}{T} \sum_{\tau=-(T-1)/2}^{(T-1)/2} \hat{y}_{i,t+\tau},$$

where T is the odd number of incomes being averaged.[24] When $T = 1$, $\overline{y}_{it}^{T}$ is equal to current income residuals $\hat{y}_{it}$. When $T > 1$, it is an average of T income observations over $2T - 1$ years, due to the biennial nature of the sample. The OLS specification to test (6) is then

$$\hat{c}_{it} = \mathbf{X}_{it}'\beta + \phi \overline{y}_{it}^{T} + \epsilon_{it}. \tag{9}$$

For long averages, $T \to \infty$, under a suitable law of large numbers for the income processes η_{it}, ψ_{it}, the average income residuals $\overline{y}_{it}^{T}$ are measuring $\hat{w}_i$ without noise. In that case, the linear model in Section 3.3 would correspond to $\phi = 1$, implying a linear consumption function in permanent incomes. Since T is finite, however, one can generally expect slopes ϕ below 1, even if

[24]See Kopczuk et al. (2010) for similar symmetrically-averaged income measures.

Table 1: OLS specifications with various proxies for permanent income.

	log household consumption			
	(1)	(2)	(3)	(4)
	$T = 1$	$T = 5$	$T = 9$	c and y FE
log Income	0.396	0.547	0.645	0.561
	(0.01)	(0.02)	(0.03)	(0.02)
Year FE, Age FE	Yes	Yes	Yes	Yes
Hh.size FE	Yes	Yes	Yes	Yes
Location FE	Yes	Yes	Yes	Yes
Observations	24994	7979	2050	2644
R-squared	0.47	0.56	0.60	0.66

Note. This table shows results from OLS regressions with various proxies of permanent income as regressors. In column 1, the regressor is current log income (residuals), in column 2 (3) it is log incomes averaged over 9 (17) years. Column 4 shows the results from a fixed effects regression, where log consumption fixed effects are regressed against log income fixed effects. Standard errors are corrected for heteroskedasticity and clustered by household.

the assumptions of the linear model are satisfied (see Section 4.3). An alternative way to construct a proxy for permanent income is as income fixed effects. In this case, household i's permanent income proxy is an average over all of i's observed income residuals $\hat{y}_{it}$.

Results. To avoid relying on functional form assumptions, Figure 3 shows the results of a non-parametric version of (9) spanning 9 years ($T = 5$). Specifically, it shows the results of a regression of $\hat{c}_{it}$ on controls and 50 bins of $\overline{y}_{it}^{T}$. Two observations are immediate: the relationship is almost exactly linear in logs, and its slope is significantly below 1. I further investigate this in Table 1 using the linear specification (9). Columns 1–3 show the results of specification (9) for different values of T. It is evident that longer averages push up the estimated ϕ, likely by reducing the downward bias. For the largest T shown, $T = 9$ income observations are averaged across 17 years—around half of an entire work-life. Even then, the estimated elasticity ϕ is around 0.65—far from 1. Column 4 shows results from a specification where (9) is implemented using fixed effects, regressing each household's consumption fixed effect on its income fixed effect, for the subsample of households with at least 5 non-missing income and consumption observations. The estimated ϕ is similar to the estimate for $T = 5$.

4.3 Econometric approach

I now investigate the possible biases in the OLS specifications more formally and propose solutions. The econometric model consists of a set of households $i \in \mathcal{I}$, of which each i enters the labor market at time $t_0(i)$, with initial age $k = 1$, and is observed until age $k = K_{death} > 1$. As before, their consumption and income processes are governed by,

$$\hat{c}_{it} = \phi \hat{w}_i + \mathbf{X}'_{it}\beta + \epsilon_{it} \tag{10a}$$

$$\hat{y}_{it} = \hat{w}_i + \eta_{it} + \psi_{it} + \nu_{it}, \tag{10b}$$

where $t \in t_0(i) + \{0, 1, \ldots, K_{death} - 1\}$. I make the following baseline assumptions on the model (10), all of which are satisfied in the linear model of Section (3). First, all random variables in (10) are iid across households i. Second, measurement error ν_{it} is iid over time and uncorrelated with consumption, $Cov(\epsilon_{it}, \nu_{it}) = 0$. Third, permanent income $\hat{w}_i$ and the controls $\mathbf{X}_{it}$ are uncorrelated with the income shocks η_{it}, ψ_{it}, measurement error ν_{it} and the consumption error term ϵ_{it}. Fourth, future transitory income shocks $\psi_{it+\tau}$ are uncorrelated with current consumption, that is, $Cov(\epsilon_{it}, \psi_{it+\tau}) = 0$ for $\tau > 0$; and past transitory income shocks $\psi_{it-\tau}$ are positively correlated with current consumption, that is, $Cov(\epsilon_{it}, \psi_{i,t+\tau}) \geq 0$ for $\tau \leq 0$ (positive income shocks in the past only raise consumption going forward, all else equal).

The critical assumptions are the third and fourth. The third requires permanent incomes $\hat{w}_i$ to be uncorrelated with ϵ_{it}, ruling out the presence of unobserved heterogeneity in savings preferences that could be correlated with $\hat{w}_i$. And the fourth assumption requires the unforecastability of transitory income shocks by the agent. Both assumptions are discussed in Section 4.4 and Appendix D.1.

The two biases of OLS regressions. Having introduced these assumptions, it is possible to investigate the biases of OLS regressions. Consider a simple OLS regression of $\hat{c}_{it}$ on current income $\hat{y}_{it}$ (and controls), corresponding to $T = 1$ in the previous section (column 1 of Table 1). It is straightforward to show that

$$\operatorname*{plim}_{N\to\infty} \hat{\phi}^{OLS} = \phi \underbrace{- \left\{ \phi - \frac{\text{Cov}(\epsilon_{it}, \eta_{it} + \psi_{it})}{\text{Var}(\eta_{it} + \psi_{it})} \right\} \frac{\text{Var}(\eta_{it} + \psi_{it})}{\text{Var}(\hat{y}_{it})}}_{<0 \text{ (consumption smoothing bias)}} \underbrace{-\phi \frac{\text{Var}(\nu_{it})}{\text{Var}(\hat{y}_{it})}}_{<0 \text{ (attenuation bias)}} .$$

There are two possible biases in OLS: the first is what one may call "consumption smoothing bias", since it is nonzero precisely when the agent's consumption reaction to income shocks—captured by the slope coefficient $\text{Cov}(\epsilon_{it}, \eta_{it} + \psi_{it}) / \text{Var}(\eta_{it} + \psi_{it})$—is different from the reaction to permanent income—captured by ϕ. In most reasonable models of consumption behavior, the former is less than the latter due to consumption smoothing, inducing a natural downward bias in the OLS estimate. The second bias is attenuation due to the presence of measurement error in income.

The consumption smoothing bias is difficult to overcome. Simple instruments, such as future or lagged incomes are able to eliminate attenuation by measurement error, but in the presence of persistent income shocks η_{it}, the consumption smoothing bias remains.[25] I propose new IV strategies that eliminate both biases, using additional assumptions on the autocorrelation structure of the persistent income shocks η_{it}.

[25]For instance, some of the specifications Mayer (1972) and Dynan et al. (2004) used instruments along these lines.

ARMA process for η_{it}. Assume the persistent income shocks η_{it} follow a (stationary) ARMA(p,q) process, that is, one can express the process as

$$\mathbf{a}(L)\eta_{it} = \mathbf{b}(L)\epsilon^{\eta}_{it},$$

where $\mathbf{a}$ is a polynomial of order p, $\mathbf{b}$ is a polynomial of order q, and L denotes the lag operator. One implication of stationarity is that $\mathbf{a}(1) \neq 0$. Again, assume that the agent cannot foresee future innovations ϵ^{η}_{it}, that is, $\text{Cov}(\epsilon^{\eta}_{it+\tau}, \epsilon_{it}) = 0$ for $\tau > 0$ (see Section 4.4 for a discussion). A standard example of such a process is an AR(1) process with persistence parameter $\rho < 1$, which case $\mathbf{a}(L) = 1 - \rho L$ and $\mathbf{b}(L) = 1$.

Define the process

$$z_{it} \equiv \mathbf{a}(L)\hat{y}_{it}.$$

By construction, z_{it} is independent of realizations of the persistent shock η_{it} that lie more than q periods in the past. Indeed, one can express z_{it} as

$$z_{it} = \mathbf{a}(1)\hat{w}_i + \mathbf{a}(L)\left(\psi_{it} + \nu_{it}\right) + \mathbf{b}(L)\epsilon^{\eta}_{it}.$$

This shows that since $\mathbf{a}(1) \neq 0$, z_{it} is correlated with $\hat{w}_i$, yet uncorrelated with $\epsilon_{it-\tau}$ for $\tau > \max\{p,q\}$. Thus, any future $z_{it+\tau}$ with $\tau > \max\{p,q\}$ is a valid instrument for current income $\hat{y}_{it}$ in a regression of consumption $\hat{c}_{it}$ on current income $\hat{y}_{it}$. In other words, ϕ can be recovered as $\phi = \text{Cov}(\hat{c}_{it}, \mathbf{a}(L)\hat{y}_{it+\tau})/\text{Cov}(\hat{y}_{it}, \mathbf{a}(L)\hat{y}_{it+\tau})$. When η_{it} is an AR(1) process, the instrument is simply given by quasi-differenced future incomes, $z_{it+\tau} = \hat{y}_{it+\tau} - \rho\hat{y}_{it+\tau-1}$.

Why does this strategy work? Using future incomes, rather than lagged incomes, is helpful because in a world without persistent income shocks, this would yield consistent estimates. The reason for this is that the actual realizations of future income shocks is not known to the agent in that case, so the only source of correlation between current and future incomes is the permanent component $\hat{w}_i$. When there is a persistent income shock, however, it needs to be differenced out first, which is precisely the role of quasi-differencing. Together, quasi-differenced future incomes are a valid instrument for $\hat{w}_i$ under the ARMA assumptions above.

Random walk η_{it}. A downside of this approach is that it is only consistent for non-unit-root processes. Indeed, if $\mathbf{a}(1)$ were equal to zero, the instrument z_{it} would be independent of $\hat{w}_i$ altogether. A common formulation of the persistent shock η_{it}, however, is as a random walk.[26] I consider this case now. In particular, assume that

$$\eta_{it} = \eta_{it-1} + \epsilon^{\eta}_{it}.$$

[26] See the recent survey by Meghir and Pistaferri (2011). The method proposed here can be extended to include the case where the transitory shock ψ_{it} follows an MA process, rather than being iid over time (MaCurdy 1982; Abowd and Card 1989).

Table 2: IV specifications.

	log household consumption						
	(1) $\rho = 0$	(2) $\rho = 0.88$	(3) $\rho = 0.89$	(4) $\rho = 0.9$	(5) $\rho = 0.91$	(6) $\rho = 0.92$	(7) ini.
log Income	0.599	0.675	0.686	0.699	0.717	0.741	0.732
	(0.02)	(0.04)	(0.04)	(0.05)	(0.05)	(0.06)	(0.05)
Year FE, Age FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hh.size FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7158	7158	7158	7158	7158	7158	21642
1st stage F	537.6	59.2	49.5	40.2	31.5	23.4	101.5
R-squared	0.48	0.45	0.44	0.44	0.43	0.41	0.41

Note. Columns 1-6 show IV results with ρ-differenced future incomes as instruments, for various choices of ρ. Column 7 shows results when the autocovariances of log income residuals are used without assuming a parametric form for income shocks. Standard errors are corrected for heteroskedasticity and clustered by household.

Let z_i be the agent's initial labor income when entering the labor market at time $t_0(i)$,

$$z_i \equiv \hat{y}_{i,t_0(i)} = \hat{w}_i + \eta_{i,t_0(i)} + \psi_{i,t_0(i)} + \nu_{i,t_0(i)}.$$

For an arbitrary process η_{it}, this variable is not an exogenous instrument. However, when η_{it} follows a random walk, the initial persistent draw $\eta_{i,t_0(i)}$ is indistinguishable from permanent income $\hat{w}_i$ for the agent, and thus can be set to zero without loss of generality. This then means that the only endogenous variable in z_i is $\psi_{i,t_0(i)}$, which I argued above is positively correlated with future errors in the consumption equation, ϵ_{it}. Therefore, if η_{it} follows a random walk, an IV strategy with instrument z_i provides an asymptotic upper bound of ϕ which can then also be used to test linearity.[27]

Results of the IV approaches. To operationalize the approach for a stationary η_{it}, I model η_{it} as an AR(1) process with annual persistence ρ. I use all periods for which at least three future incomes $z_{i,t+\tau}$ are observable and use all available instruments for each observation. There are many estimates of persistences ρ in the literature. In Appendix D.4, I provide estimates for ρ for precisely the income measures used here (i.e. total household labor incomes) using a similar horizon (15 years) as the maximum span of the instruments $z_{i,t+\tau}$ I use. This procedure yields an annual persistence $\rho = 0.90$, somewhat lower than standard values for the persistence, typically estimated for male wages and longer horizons. Columns 1–6 of Table 2 show the results of a two-stage-least-squares estimation for various choices of ρ around the estimated level of $\rho = 0.9$ as well as $\rho = 0$.

Overall, a consistent picture emerges. While estimates do increase with larger choices of persistences ρ (which they are expected to—see the discussion on misspecification of ρ in Appendix D.1),

[27]In my simulations in Section 5 I will show that this upper bound is generally very tight in models with random walk income processes and provides a close upper bound even in models without random walk income processes.

Table 3: Robustness checks.

	OLS with $T = 9$	IV with $\rho = 0.90$	IV with initial income
1. Baseline	0.64	0.70	0.73
	(0.03)	(0.04)	(0.05)
2. Education and preference controls	0.58	0.69	0.78
	(0.03)	(0.07)	(0.08)
3. Controlling for positive business wealth	0.55	0.68	0.72
	(0.02)	(0.05)	(0.05)
4. HIP – education specific trends	0.59	0.68	0.73
	(0.03)	(0.06)	(0.06)

Note. This table lists OLS and IV estimates for 4 different specifications. Row 1 shows the baseline specifications from Table 1 column 3 and Table 2 columns 4 and 7. Row 2 adds controls for several proxies for preferences (education, race, sex, and self-reported bequest intentions). Row 3 adds a dummy for positive business wealth. Row 4 controls for education-specific household income trends. All specifications control for year, age, household size and location dummies. All IV specifications have first stage F statistics above 10. Standard errors are corrected for heteroskedasticity and clustered by household.

they fall between 0.60 and 0.75. Importantly, for all specifications, the F-statistics are above 10, suggesting that there is no weak instruments problem for those values of ρ.[28] In column 7 of Table 2, I show the results of the second IV approach using as IV the household's labor income at the head's age of 25. The result is consistent with this estimate providing an upper bound, even if the underlying income process may not exactly be a random walk.

4.4 Discussion and robustness

There are a variety of concerns one might have about the specifications in Section 4.3. In this subsection, I address a few of the most important. Appendix D.1 executes a number of additional robustness checks, including among others: results using alternative consumption measures; group-level specifications, similar to those in the seminal work of Attanasio and Davis (1996); results with education as IV; and results using the strategy in Aguiar and Bils (2015) to strip out non-classical measurement error. The regression results for each of the following robustness exercises appear in Table 3.

Heterogeneous time or bequest preferences. One of the most immediate concerns one may have about the specifications in Section 4.3 is endogeneity due to preference heterogeneity that is correlated with permanent income levels. For instance, better-educated workers may be more financially responsible and save more conditional on a given level of permanent income. Since educated workers usually earn higher incomes, this could lead to a downward bias in $\hat{\phi}^{IV}$. To address these concerns, the second row in Table 3 provides specifications that include additional controls that proxy for factors influencing savings preferences: the household head's education, race, sex, and

[28] Values of ρ beyond around 0.94 (annual), however, do generate significant weak instrument problems.

the reported bequest intention, which was elicited in the 2007 wave of the PSID.[29] The results are broadly similar to the baseline estimates, suggesting a limited role for observed time preference heterogeneity.[30]

Heterogeneous returns on wealth. A recent literature powerfully demonstrates that returns on wealth are very heterogeneous in the population and generally increase with wealth (Fagereng, Guiso, Malacrino and Pistaferri 2019a). One relative advantage of my specification based on consumption expenditure (rather than wealth differences or wealth-to-income ratios) is that its results cannot be "mechanically" driven by heterogeneity in returns that is correlated with income. The effect of such heterogeneity on my results generally depends on the elasticity of intertemporal substitution (EIS): if the EIS is below 1, as is typically assumed in precautionary savings models, it may well be that high-return agents consume more out of their labor income, not less. Since return heterogeneity is hard to disentangle from noise in the PSID, I use as a simple proxy whether a household has positive business wealth or not. The results appear in the third row of Table 3 and fall slightly below the baseline numbers.

Heterogeneous income profiles or advance information. A recent literature argues that income age profiles may have heterogeneous slopes (Guvenen 2007, 2009; Guvenen and Smith 2014; Guvenen, Karahan, Ozkan and Song 2019), and more generally, agents may have advance information about future income realizations (Primiceri and van Rens 2009). The presence of these features do not affect my initial income IV results when η_{it} follows a random walk. When η_{it} follows an AR(1), this is likely to bias $\hat{\phi}^{IV}$ upwards: when an agent expects greater future income growth, then, all else equal, the consumption c_{it} will be higher. A straightforward way in which the robustness of my results with respect to heterogeneous income profiles can be explored is to allow in the controls $\mathbf{X}_{it}$ and $\tilde{\mathbf{X}}_{it}$ for heterogeneity in income profiles based on observables. This is exactly what I do in the fourth row of Table 3 by adding an interaction of an age trend with education to $\mathbf{X}_{it}$. The results are shifted downward—as one may expect—but not by a whole lot, at least for this rough measure of heterogeneous income profiles.

5 A Non-Homothetic Life-Cycle Economy

The evidence in the previous section suggests that consumption is better described by a concave function of permanent income than a linear one. The next natural question is whether this matters for the impact of income inequality on macroeconomic outcomes. To answer this question, I first propose a model that incorporates a concave consumption function into a quantitative life-cycle model. In order to be able to speak to macroeconomic outcomes, the model is a parsimonious

[29] The question asked was: "Some people think that leaving an estate or inheritance to their children or other relatives is very important, while others do not. Would you say this is very important, quite important, not important, or not at all important?" Answers were given on a numeric scale from 1 (very important) to 4 (not at all important).

[30] Of course, these proxies are imperfect and my estimates could still be partly due to unobserved preference heterogeneity. I explore this possibility in a quantitative model in Appendix G.3 in a setting where heterogeneity in preferences is the sole cause of $\hat{\phi} < 1$. I find that this model exhibits aggregate non-neutrality, similar to that of the non-homothetic models in Section 5.

non-homothetic extension of the standard life-cycle model introduced in Section 3.[31]

5.1 Model

Each period (or age) corresponds to one year in the model.

Agents. In the model, agents are "born" at the model age $k = 1$, corresponding to a biological age of 25 when agents enter the labor market. They have an offspring that enters the labor market at model age $k = 25$ (biological age 50). They retire at model age $K^{ret} \equiv 40$ (biological age 65), and die with certainty at model age $K_{death} \equiv 65$ (biological age 90). After retiring, all the way to the certain death, agents face a positive mortality rate δ_k from age k to $k+1$. Their pre-tax income process is stochastic subject to idiosyncratic productivity shocks and given by

$$\log y_t^{pre} = \theta_{t-t_0} + \eta_t + \psi_t$$

where θ_k captures changing productivity over the life-cycle; η_t is an AR(1) process with persistence ρ and a standard deviation of its innovation of σ_η; and ψ_t is iid over time, with standard deviation σ_ψ. The initial persistent income shock η_{t_0} is drawn from a normal distribution with variance σ_η^2. Agents inherit their parents' skill with probability $p_{inherit} \in [0, 1]$ and are assigned a random skill according to population shares with probability $1 - p_{inherit}$.

Government. Retired agents earn social security payments $T^{socsec}(\bar{y})$ which are modeled using a piecewise linear schedule according to the Old Age and Survivor Insurance component of the Social Security system (Huggett and Ventura 2000; De Nardi and Yang 2014)

$$T^{socsec}(\bar{y}, W) = 0.9 \min(\bar{y}, 0.2W) + 0.32 \left(\min(\bar{y}, 1.24W) - 0.2W\right)^+ + 0.15 \left(\min(\bar{y}, 2.47W) - 1.24W\right)^+ .$$

Here, W is the average labor income, and $\bar{y}$ is a measure of an individual agent's lifetime income. Since keeping track of an agent's actual lifetime income is computationally costly—it adds an additional continuous state variable—I predict an agent's lifetime income based on the agent's last working-age income when computing the agent's social security transfer.

All agents pay income taxes according to a progressive income tax system (Benabou 2000; Heathcote et al. 2017),

$$T^{inctax}(y^{pre}) = y^{pre} - \tau^{inctax} \left(y^{pre}\right)^{1-\lambda}, \tag{11}$$

where $\tau^{inctax} > 0$ is a constant and $\lambda \geq 0$ is a tax progressivity parameter: a larger λ corresponds to more progressive income taxation. In addition, following Hubbard et al. (1995), even households with very low incomes receive basic assistance and basic health care. I capture this by a government-provided income floor of $\underline{y}$. After-tax incomes are then

$$y_t = \max\{\underline{y}, y_t^{pre} - T^{inctax}(y_t^{pre})\}.$$

[31] Also, given the current evidence alone, it is hard to discipline richer models.

Finally, there is a proportional tax $\tau^{cap} > 0$ on capital income. Henceforth, r denotes the after-tax interest rate, $r = (1 - \tau^{cap})r^{pre}$.

Preferences. I allow households to have non-homothetic preferences, both over consumption and bequests. The model breaks Assumption 1 by assuming that preferences over consumption are given by

$$u_k(c) = \frac{(c/o)^{1-\sigma_k}}{1-\sigma_k}, \tag{12}$$

where $\sigma_k > 0$ is an age-dependent elasticity that is constant and equal to σ during retirement, and $o > 0$ is again a constant. Preferences over bequests are given by

$$U(a) = \kappa \frac{((a + \underline{a})/o)^{1-\sigma}}{1-\sigma},$$

where $\kappa > 0$ and $\underline{a} > 0$.

5.2 Discussion of preferences

These assumptions on preferences constitute the most important deviations from the linearity result of Section 3.3 and therefore merit an extensive discussion.

Age-dependent elasticities. The key idea behind the $\{\sigma_k\}$ is to change income elasticities of spending across periods. To achieve this in the most straightforward and transparent way, I use "addilog" preferences, pioneered by Houthakker (1960). In a static setup, Houthakker (1960) shows that when the utility function is iso-elastic and additively separable, with power $1 - \sigma_k$ on good k, the income elasticity of good k, ε_k, is inversely proportional to σ_k, that is, $\varepsilon_k \sim \sigma_k^{-1}$. Therefore, goods with *low* elasticity σ_k have a *high* income elasticity. Moreover, in a two-good setting, the income elasticity is equal to the ratio σ_k/σ_{k+1} for high incomes (see also Section 2).

I apply this logic to an intertemporal context. I assume that σ_k is lower for higher ages, capturing the fact that higher income households spend relatively more in the future. For simplicity, I let the ratios σ_k/σ_{k+1} be constant until retirement, with a constant that is determined by the calibration.

Examples for spending of rich households later in life include payments for college education, gifts for kids or grandkids, charitable giving, or expensive medical treatments. These expenditures are typically back-loaded and occur late in one's working life or in retirement.[32] I present direct evidence of such expenditures in Appendix D.6.

These preferences have two additional implications for household behavior, aside from the changes in income elasticities they induce. First, the preferences change attitudes towards risk both over the life cycle and as a function of permanent income, pushing down risk aversion at higher ages and at higher income or wealth levels. This is not necessarily a bug: while homothetic models are well-known to predict *increasing* curvature in the value function with age, and hence increasing risk aversion, the non-homothetic model generates a relatively flat age profile of risk aversion. In a finance context, such an age profile is more in line with the evidence, e.g. from the age structure in

[32] This is amplified by the fact that rich households are well-known to live longer than poor households.

portfolio allocation.[33] Second, the preferences imply an elasticity of intertemporal substitution (EIS) that rises with income or wealth, in line with estimates by Blundell, Browning and Meghir (1994) who find that the EIS increases in permanent income (see Attanasio and Browning 1995 for similar findings).

Non-homothetic bequest motive. The second, more standard source of non-homotheticity is in the form of bequests. Bequest utilities are allowed to have a different intercept than that of consumption. In particular, when $\underline{a}$ is relatively large, bequests are treated as a luxury good. Thus richer agents choose to save in order to leave bequests, while poorer agents do not, or less frequently so. The idea to incorporate such preferences in a canonical incomplete-markets life-cycle economy goes back to the seminal work of De Nardi (2004).

Why do I allow for both sources of non-homothetic consumption-savings behavior, and not merely focus on non-homotheticity in bequests? As I argue below using simulations, the reason is that non-homotheticity in bequests by itself cannot quantitatively account for the concavity ϕ I document in Section 4, without implying implausibly large bequests. This is because bequest flows are typically estimated to be around 5% of GDP (see Section 5.3 below), limiting their quantitative role.[34]

An alternative non-homothetic model based on status. To check my results for robustness I propose an alternative set of non-homothetic preferences in Appendix G.1, which goes back to the idea that being relatively wealthy either gives utility per se (e.g. as in Carroll 2000 and more recently Saez and Stantcheva 2017) or confers a certain status in society (e.g. as in Cole, Mailath and Postlewaite 1992, 1995, 1998). This model assumes that the working-life per-period utility is given by

$$u(c,a) = \frac{c^{1-\sigma}}{1-\sigma} + U(a/A)$$

where a/A is one's current wealth relative to average wealth A—and thus a measure of status; $U(x) = \kappa \frac{(x+\underline{x})^{1-\sigma}}{1-\sigma}$ is the utility function over status. As I argue in Section 6.7 and Appendix G.1, this model can also be calibrated to match the empirical evidence on the curvature ϕ of the consumption function, and predicts similar effects from changing income inequality.

5.3 Calibration

Wherever possible, the two models are calibrated to the U.S. economy in 2014. Even though one of my main experiments will be a comparison with 1970, it is important to calibrate the economy at

[33] Ameriks and Zeldes (2004) show that the risky share in one's portfolio does not decline with age. Such a decline would be predicted by a canonical homothetic life-cycle economy, since human capital acts like a bond position and declines over one's life. An important paper explaining this fact is Wachter and Yogo (2010), who incorporate intratemporal addilog preferences over two goods (a necessity and a luxury) into a canonical life-cycle economy.

[34] The age pattern I find in Appendix D.6 also suggests that there must be an additional force for non-homotheticity in the economy, since an economy that only has a non-homothetic bequest motive would generate a permanent income elasticity ϕ that declines with age.

roughly the same time during which the concavity parameter ϕ was estimated.[35] Table 4 summarizes the calibration.

Birth death skills. I assume there are $\mathcal{S} = 3$ skills with population shares of $\overline{\mu}_s$ of 0.90, 0.09, and 0.01, capturing the bottom 90%, the middle 9% and the top 1%. This choice is motivated by recent increases in incomes going to top income groups (Piketty and Saez 2003), but the results are very similar using different skill groups.[36] I calibrate mortality rates $\{\delta_k\}$ by age to the data from the Center for Disease Control for 2011.[37]

Production. The gross capital share α is computed for the U.S. non-financial corporate sector for 2014 (as in Rognlie 2015). This yields $\alpha = 36.7\%$. The labor income shares $\{\gamma_s\}$ are calibrated to match the bottom 90%, the middle 9% and top 1% income shares using updated data from Piketty and Saez (2003). Depreciation is set to a value of $\delta = 0.055$ implying a capital-output ratio of $K/Y = 3.5$ and thus (including government debt) a private wealth to GDP ratio of around 4.2 in line with Piketty and Zucman (2014). I use A to normalize GDP Y to 1. I adjust the normalization when changing the distribution of income $\{\gamma_s\}$ to focus on the pure redistributional effects without mechanical changes in Y. I choose a level of $r = 3\%$ for the after-tax interest rate; before taxes this corresponds to $r^{pre} = 5\%$.

Government. Government debt B is set equal to 73% of GDP to match the 2014 ratio of total federal debt held by the public to GDP, which excludes bond positions held by the social security trust fund and other government entities (not the Federal Reserve). It is difficult to calibrate estate taxes, since, as is well known, there exist numerous (legal and illegal) ways in which the estate tax burden can be reduced. I therefore choose to follow the literature and assume an intermediate value of $\tau^{beq} = 0.10$ as in De Nardi (2004). I set the income floor $\underline{y}$ to be 40% of average household labor income, corresponding to around \$13,000 per adult in 2014 dollars. This is a conservative estimate, slightly larger than that in Guvenen et al. (2019). I pick income the tax progressivity $\lambda = 0.181$ in line with estimates from Heathcote et al. (2017). The average income tax rate τ^{inctax} is set to match the sum of personal tax receipts, employers' contributions to government social insurance as a fraction of total labor income and fraction $0.5(1-\alpha)$ of tax income from production and imports.[38] Together this results in an average income tax of $\tau^{inctax} = 30\%$. Capital taxes are set to cover the remaining fraction of total government receipts, so that approximately $\tau^{cap} = 40\%$. Government spending G is set to be the residual in the government budget constraint. In the calibration, this gives a value of $G/Y = 14\%$.

[35] As it turns out, however, the model-implied concavity parameter ϕ is quite stable over time. I estimated ϕ in the 1970 steady state from which the transitional dynamics in Section 6 begin and found estimates that were around $0.65 - 0.7$.

[36] While the PSID does cover some very high incomes (the highest pre-tax income in survey year 2013 is over \$6m) the very top shares (top 1%, and especially top 0.1%) are underrepresented. The results here should therefore be understood conditional on the top 1% being still described by the same relationship between consumption and income as the bottom 99%.

[37] To avoid a somewhat larger mass of agents dying right at age 90 compared to other ages, I smooth mortality rates over the last ten years of the life (this has no effect on the results).

[38] Around 50% of the taxes levied on production and imports is property tax income and is counted towards capital taxation. The rest is split according to capital and labor income ratios.

Table 4: Calibrated parameters of the baseline non-homothetic life cycle model.

Parameter	Description	Value	Origin
Birth, death, skills			
$\mathcal{S}$	Number of permanent types	3	see text
$\{\overline{\mu}_s\}$	Population shares by type	$\{0.9, 0.09, 0.01\}$	see text
$\{\delta_k\}$	Mortality rates by age		CDC, 2011
Production			
α	Capital share	0.37	NIPA, 2014
$\{\gamma_s\}$	Labor income shares	$\{0.65, 0.24, 0.11\}$	Piketty and Saez (2003), updated to 2014
δ	Depreciation	0.055	match $K/Y = 3.05$ (NIPA, 2014)
A	Total factor productivity	0.63	normalize $Y = 1$
Government			
B/Y	Federal debt held by the public / GDP	0.73	NIPA, 2014
τ^b	Bequest tax	0.10	see text
$\underline{y}$	Income floor	$0.30W$	literature
$\overline{\lambda}$	Income tax progressivity	0.18	PSID, 2013
τ^{inctax}	Average income tax	0.30	NIPA, see text
τ^{cap}	Capital tax	0.40	NIPA, see text
G/Y	Government spending / GDP	0.13	gov. budget constraint
Productivities			
ρ	Income shock persistence	0.90	PSID
σ^2_η	Var. of innovations to persistent shock	0.028	PSID
σ^2_ψ	Var. of transitory income shocks	0.055	PSID
σ^2_v	Var. of measurement error in incomes	0.02	literature, see text
$p_{inherit}$	Prob. of intergen. skill transmission	0.35	Chetty et al (2014)
Preferences			
β	Discount factor	0.89	match interest rate $r = 0.03$
$\overline{\sigma}$	Elast. of intertemp. substitution, median age	2.5	literature
z	Scale term in utility function	0.30	30% of average income
σ^{slope}	Ratio of elasticities σ_{k+1}/σ_k	0.94	match $\phi = 0.699$
κ	Weight on bequest motive	15.84	match bequests / GDP = 0.05
$\underline{a}$	Intercept in bequest utility	1.72	30% share with beq. $\leq$ 6.25% avg. income

Table 5: Elasticity of per-period utility function u by age.

Age group	$25-44$ years	$45-64$ years	$65+$ years
Average σ	7.1	3.6	1.6

Note. This table shows the age profile of elasticities σ_k, averaged within three age groups. The declining age profile in σ_k captures the empirical fact that consumption early in one's life-cycle has a permanent income elasticity below 1, implying that permanently richer agents save relatively more.

Idiosyncratic productivity process. I determine the idiosyncratic productivity process on the pre-1997 annual sample of working-age PSID households (see Appendix D.4 for details) by estimating

$$\hat{y}_{it}^{pre} = \theta_{t-t_0(i)} + \hat{w}_i + \eta_{it} + \psi_{it} + \nu_{it}.$$

Here, $\hat{y}_{it}^{pre}$ is the log of pre-tax household labor income; I model the age efficiency profile θ_k as a cubic polynomial in age; $\hat{w}_i$ is assumed to follow a normal distribution; and ν_{it} is measurement error. Since ν_{it} and ψ_{it} are indistinguishable, I follow Heathcote et al. (2010) and assume the variance of ν_{it} is equal to $\sigma_\nu^2 = 0.02$.[39] This yields the following results. The persistence of η_{it} is found to be $\rho = 0.90$, the variance of the innovation to η_{it} is equal to 0.026, and the variance of the transitory shock ψ_{it} is given by 0.052.[40] The inheritance probability of parental skill, $p_{inherit}$, is calibrated to match the slope between between parental and child income ranks measured in Chetty, Hendren, Kline and Saez (2014). This gives $p_{inherit} = 0.35$.

Preferences. I choose a simple parametric form for the age profile in consumption elasticities, namely a simple exponential decay, $\sigma_{k+1}/\sigma_k = \sigma^{slope} > 0$, during one's working life and flat thereafter.[41] In addition to the slope, I pick the median elasticity $\overline{\sigma}$. I choose a standard parameter, $\overline{\sigma} = 2.5$. I calibrate jointly $\{\beta, \sigma^{slope}, \kappa, \underline{a}\}$ to match the following four moments: (1) an (after-tax) real interest rate of $r = 3\%$, which one should understand as the "total rate of return" in the economy; (2) a bequest flow over GDP of 5%—an intermediate value between the recent estimate in Alvaredo, Garbinti and Piketty (2017) of 8% and the estimate of 2% (see, e.g., Hendricks 2001) ; (3) a 30% share of households with bequests below 6.25% of average income (De Nardi 2004); (4) an estimate of the permanent income elasticity of consumption that matches column 5 of Table 2. This last moment is matched using Monte-Carlo simulations from the model (see Appendix E.1 for details). It is this last set of moments that ultimately determines the life-cycle non-homotheticity parameter σ^{slope}. The parameters are found to be: $\beta = 0.89$, $\sigma^{slope} = 0.94$, $\kappa = 16$, $\underline{a} = 1.7$. The parameter o, which is irrelevant in the homothetic case ($\sigma^{slope} = 1$), is set to 30% of average household income, or \$21,000 in 2014 dollars, which can be thought of as a "minimum" level of consumption below which agents barely save.[42] Table 5 shows the age-dependence of σ_k and average levels of σ_k for three stages in

[39] I also include this measurement error term in any model simulations below.

[40] Results for models with different income processes (e.g. a random walk) are very similar and available upon request.

[41] I experimented with several other parametric choices and, to the extent that σ_k is downward sloping during one's working or entire life with a flexible slope parameter, the qualitative and quantitative results are similar.

[42] It turns out that numerically, o is fairly aligned with β, so different values of o mainly shift the discount factor around

Table 6: Estimating ϕ in simulated data from the non-homothetic lifecycle model.

	OLS		IV			
	$T=1$	$T=9$	$\rho=0.89$	$\rho=0.9$	$\rho=0.91$	initial income
Data	0.40	0.64	0.69	0.70	0.72	0.73
Non-homothetic model	0.52	0.71	0.69	0.70	0.71	0.72
Homothetic model	0.67	0.89	0.98	0.99	1.01	1.00

Note. This table compares regressions testing the linearity of consumption in permanent labor income in the data and in two models: A non-homothetic and a homothetic life-cycle model. The first two columns are results from an OLS regression of log consumption on log income residuals, averaged across T observations. Columns 3–5 are IV results with quasi-differenced future incomes as instruments; Column 6 shows IV results with initial incomes as instruments.

one's life cycle.

Homothetic benchmark. In addition to the non-homothetic life-cycle model I introduced, I also calibrate a "homothetic version" of this economy. For that, I set $\sigma_k = \overline{\sigma}$ and $\underline{a} = 0$ to eliminate all forms of non-homothetic savings behavior. I then re-calibrate β and κ to still match a post-tax real interest rate of 3% and the same total bequest flow relative to GDP of 5%. This gives $\beta = 0.99$, and $\kappa = 1.5$. Moreover I assume social security payments to be a linear function of (projected) lifetime incomes, with a slope of 20% chosen to match the sum of social security expenses in the non-homothetic economy. This homothetic model is therefore very close to the linearity benchmark in Section 3.[43]

5.4 Calibration results

I explore the calibrated non-homothetic and homothetic models in a number of dimensions. Additional calibration results can be found in Appendix F.

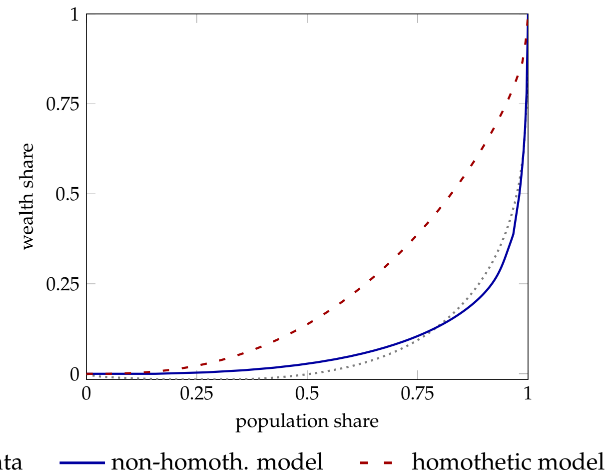
Estimating ϕ in simulated data. Table 6 shows estimates from the same regressions as those in Section 4, just run on simulated data. The table reveals that there is indeed a significant degree of bias in OLS estimates in simulated data from two two models; that the non-homothetic model matches the IV estimates well (and not just the targeted $\rho = 0.90$ estimates); and that the homothetic model is indeed very close to being linear. The table also confirms why the IV specifications in Section 4.3 are useful: even with a 9-year average income measure, the ϕ estimate according to the OLS specification is around 0.90, rather than 1.

Wealth inequality. The fact that richer people save relatively more naturally generates more wealth inequality than would occur in a linear economy. But how much more? Figure 4 shows the Lorenz curves for wealth in the models and the data. As is visible, the non-homothetic model

but do not materially affect any other results.

[43]It is not exactly linear since there is imperfect skill persistence and the income tax schedule is progressive, that is, Assumptions 2 and 3 are violated.

Figure 4: Lorenz curve for wealth.



Note. The Lorenz curve shows how much wealth the bottom x percent of the population hold, where x varies along the horizontal axis. The closer the plot is to the 45° line, the more equal the distribution is. The more this plot is pushed towards the bottom right corner, the more unequal is the distribution. The data on wealth inequality is from Saez and Zucman (2016).

generates a significant amount of wealth inequality that matches the data quite successfully overall, despite wealth inequality not having been a calibration target.[44]

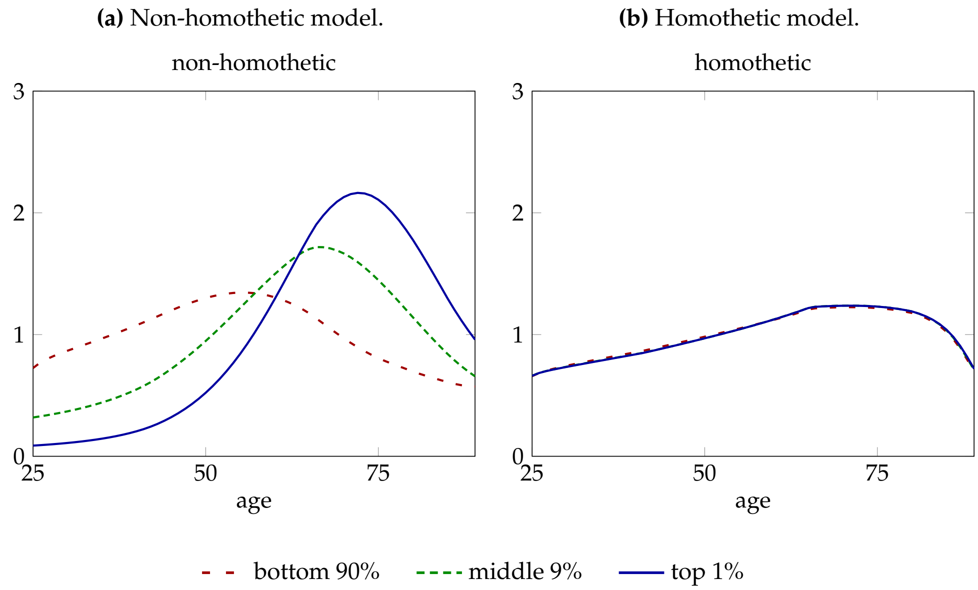
Consumption profiles. To illustrate the non-homotheticity, Figure 5 plots expenditure profiles over the life cycle for the homothetic and non-homothetic models. The profiles are normalized by their average expenditure. While in the homothetic model (Panel (b)), agents of all skills choose the same expenditure profiles (up to scale), this is not the case in the non-homothetic model (Panel (a)). There, agents with higher permanent incomes choose to shift their expenditures considerably more towards higher ages.

5.5 Comparison with alternative models

I showed in Table 6 that the non-homothetic model matches the empirical evidence in Section 4 relatively successfully, especially when compared to the homothetic benchmark economy. Yet, one may still wonder how other economies fare under the same test. For example, what happens if the persistent shocks are more persistent than an AR(1) with $\rho = 0.90$, as was assumed in the two models of the previous section? What if there is partial insurance against income shocks? What if agents are subject to shocks to their discount factors—a common model ingredient that is generally used to increase wealth concentration (Krusell and Smith 1997; Hubmer et al. 2016)? What if the income process is entirely different, e.g. one with kurtosis, inspired by those of Guvenen et al. (2019), or one with an extremely productive state, as in Castaneda, Giménez and Ríos-Rull (2003)

[44] As is intuitive, this mechanism is not able to match wealth inequality at the very top, e.g. the top 0.1% wealth share, which is more driven by entrepreneurship and return heterogeneity, as in the models of Quadrini (2000), Cagetti and De Nardi (2006), and Benhabib, Bisin and Luo (2018); Benhabib et al. (2017), among many others.

Figure 5: Expenditure profiles (normalized) by skill group.



Note. This graph shows the age profiles of expenditures, normalized by average expenditure, for each skill group. Panel (a) computes this for the non-homothetic model, Panel (b) for the homothetic model.

and Kindermann and Krueger (2017)?

This subsection answers these questions by comparing the main IV specifications—the AR(1) IV with quasi-differencing ($\rho = 0.90$) and the initial income IV—across a wide variety of other precautionary savings models. To this end, I computed a number of extensions to the homothetic benchmark economy which are designed to capture certain additional model features. All of these alternative models are calibrated to match the same post-tax steady state interest rate of $r = 3\%$, and their model parameters are set to standard values—wherever possible, to the same parameter values that are used in the non-homothetic model. I then simulated data from these models and estimated the same specifications as in the data. Details on the estimation on model-simulated data can be found in Appendix E.2, and details on the model extensions are in Appendix G.2.

Table 7 shows the results, with the columns representing the various specifications: the first column shows the true ϕ parameter in the model, which, reassuringly, exactly coincides with the AR(1) IV as long as that is the income process being used, once more confirming the usefulness of the specifications in Section 4.3. If the true ϕ is equal to 1, the model is exactly linear as in Section 3.3. The second and third columns show the two IV specifications.

The alternative models considered in Table 7 split into three blocks (labelled 2.–4.), starting with models with alternative preference or transfer assumptions. As can be the seen, a concave social security system and a non-homothetic bequest motive can generate true concavity in consumption as function of permanent income, albeit only mildly so. The next block, linear models with other income processes, shows that misspecification of the income process can bias AR(1) IV estimates of ϕ downwards, but always biases initial income IV estimates upwards, and thus is hard to reconcile

Table 7: Comparison across models.

	true ϕ	IV, $\rho = 0.9$	IV, ini
1. Data and main models			
Data		0.70	0.73
Non-homothetic	0.70	0.70	0.72
Homothetic	0.99	0.99	1.00
2. Alternative preferences or transfers			
Homothetic w/out bequests	1.00	1.00	1.03
Homothetic w/ social security	0.95	0.95	0.96
Homothetic u but luxury bequests	0.96	0.96	0.97
3. Alternative income process			
AR(1) with $\rho = 0.95$	1.00	0.93	1.02
Permanent-transitory	1.00	0.91	1.01
Heavy-tailed	1.00	0.89	1.00
Extreme productivity state	1.00	0.90	1.00
Heterogeneous income profiles	1.00	1.10	1.03
4. Other			
Partial insurance	1.00	1.00	1.06
Random discount factors	1.00	1.00	1.12

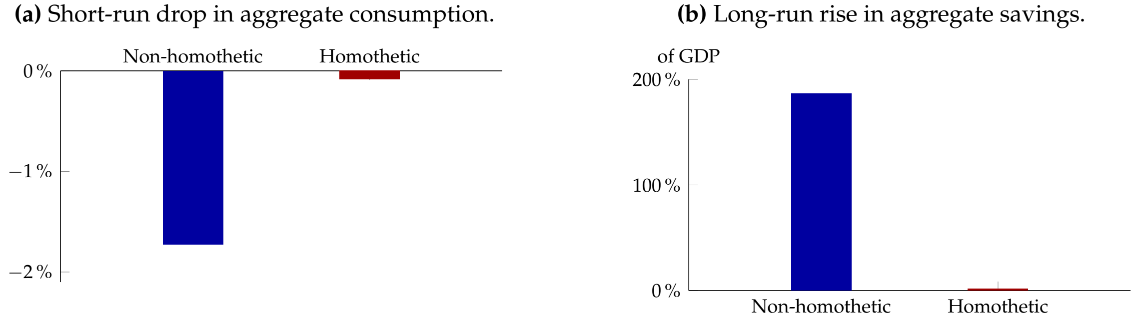
Note. This table compares the two main IV tests of the linearity of consumption in permanent income, across a wide range of models. It estimates the regressions on Monte-Carlo simulated data. The models are recalibrated to match the same equilibrium interest rate. See text and Appendix G.2 for details.

with the data. Finally, the last block of models shows that the slope of the consumption function is orthogonal to the inclusion of partial insurance (as in Guvenen and Smith 2014) or random discount factors (as in Hubmer et al. 2016).

6 The Non-Neutral Effects of Rising Permanent Income Inequality

Armed with a model that captures the relationship between consumption and permanent income in the data, it is now possible to ask one of the central questions behind this paper. What role, if any, did the recent rise in permanent income inequality play for the evolution of the macroeconomy over the past few decades? This section approaches this question from multiple angles, beginning with a simple partial equilibrium exercise, holding the interest rate fixed. Throughout the economy begins in a steady state that matches the 1970 labor income shares $\{\gamma_s\}$ from Piketty and Saez (2003) and the associated steady state interest rate.

Figure 6: The partial equilibrium effects of a rising income inequality.



Note. Panel (a) shows the simulated short-run drop in consumption in the non-homothetic and homothetic economies after income inequality permanently and unexpectedly shifts from 1970 to 2014 levels. Panel (b) shows the simulated long-run (steady-state) rise in accumulated wealth after such an experiment in both economies. Both panels are in partial equilibrium, with a fixed interest rate.

6.1 Non-neutrality in partial equilibrium

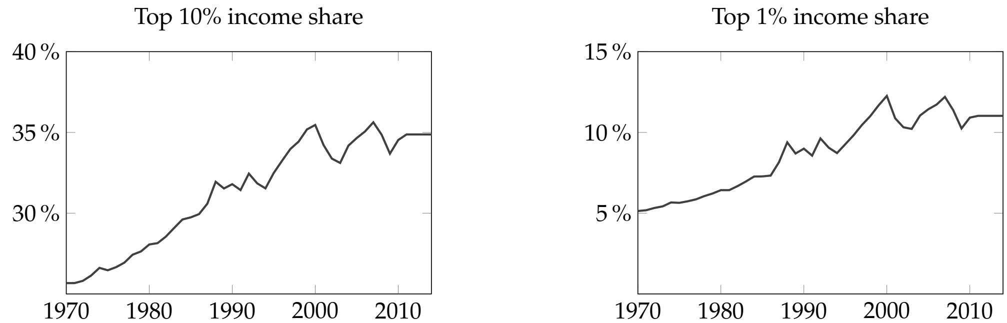
The first exercise is a partial equilibrium one, in which the interest rate is held fixed. Starting in the 1970 steady state, I let labor income shares $\{\gamma_s\}$ suddenly jump back to their 2014 levels, simulating the transition thereafter. I focus on two specific outcomes: the immediate shortfall in consumption, which should be greater in an economy with a more concave consumption function in permanent income; and the long-run accumulation of wealth in the new steady state with greater inequality. I compare both outcomes with the response in the homothetic economy (whose consumption function is almost linear in permanent income).

Short-run effect on consumption. I calculate the percentage decrease in aggregate consumption after the unanticipated shift in income inequality. Panel (a) of Figure 6 shows these results. In the non-homothetic model, initial consumption drops around 2% in total. In the homothetic model, consumption also falls, but the fall is an order of magnitude smaller, around 0.1%. This is another indicator that consumption is a concave function of permanent income in the non-homothetic model, while it is linear in the homothetic model, even in the short run and not just in the long run.

Long-run effect on savings. After the economy settled to the new steady state with higher income inequality, the non-homothetic economy has accumulated a significant amount of additional wealth, around 180% of GDP, whereas the homothetic economy has virtually not accumulated any extra wealth.

Why is the long run so much larger than the short run in the non-homothetic economy? To understand this, notice that the consumption adjustment in Panel (a) of Figure 6 is very persistent: permanently richer agents have greater saving rates, so an increase in their incomes snowballs into a large pile of additional savings. Moreover, this large pile of additional savings persists: not only do future generations also raise their total saving due to greater inequality, but they also receive greater bequests than previous generations did.

Figure 7: The evolution of labor income shares $\{\gamma_s\}$.



Note. The two figures show the two exogenous inputs into the transitional dynamics exercise: the top 10% labor income share and the top 1% labor income share, both from Piketty and Saez (2003).

6.2 General equilibrium transitional dynamics

The partial equilibrium experiment already suggests that the rise in permanent income inequality may have had a sizable effect on the U.S. aggregate economy. This raises two questions: how large are the effects in general equilibrium? And how long does it take for the effects to show?

Computing the general equilibrium (GE) transitional dynamics requires to overcome significant computational challenges. The state space of the models studied so far has 1.3 million idiosyncratic states; in addition to that, it is not only the interest rate that is endogenous along the transition path, but also the entire distribution of bequests—which matters in an economy where bequests are not received at birth, but rather in mid-life; and if this weren't enough, the transition needs to be solved over hundreds of years. In light of these challenges, I reduce its state space to 400k states and improve existing algorithms along a number of dimensions. I discuss these improvements in Appendices E.1 and H and formally define a non-stationary equilibrium in Appendix E.1.

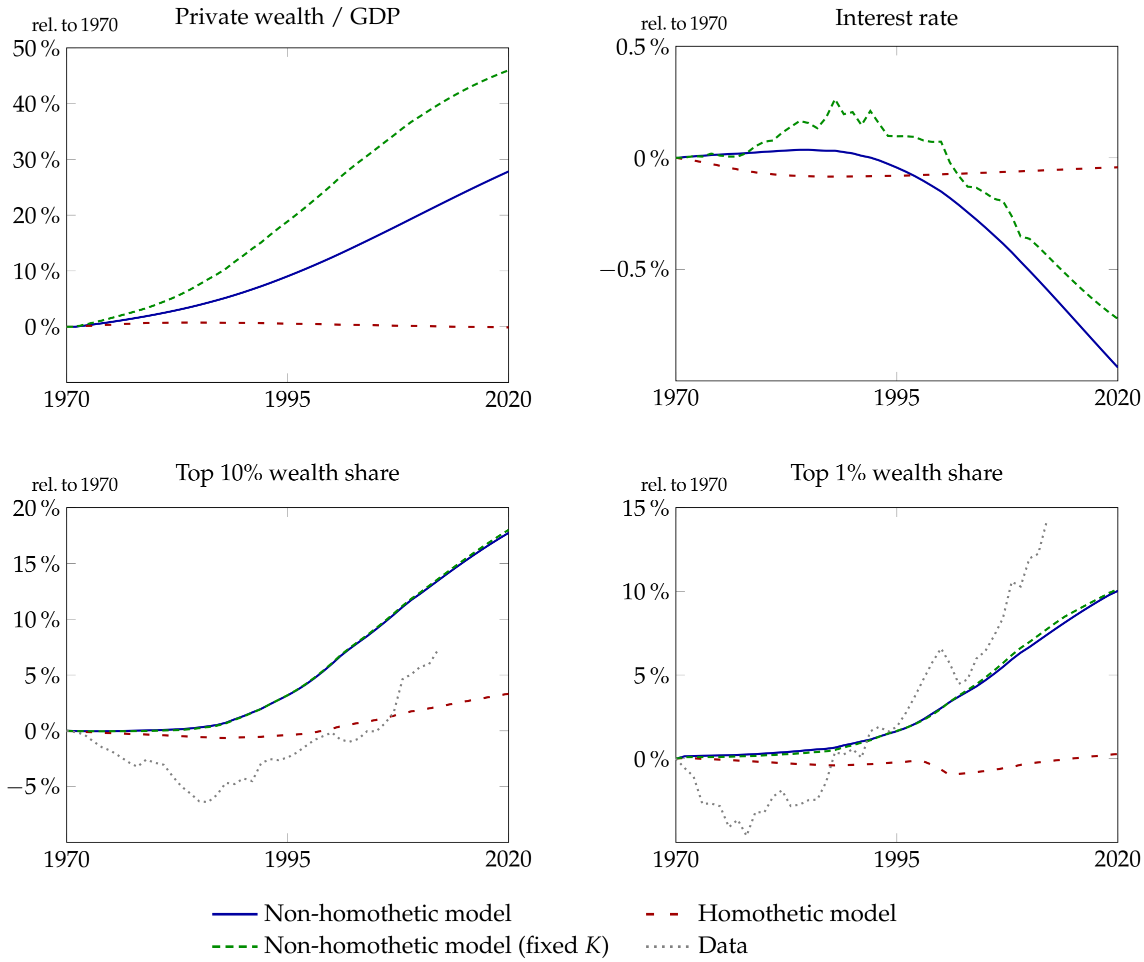
To simulate the transitional dynamics, I initiate the model in a steady state with the 1970 income distribution, and then feed in an exogenous path for labor income shares $\{\gamma_{s,t}\}$ that matches the actual one (see Figure 7). I assume that the values for $\gamma_{s,t}$ remain constant after 2014, and that agents have perfect foresight over the entire path. Finally, I assume that all agents hold the two assets—capital and bonds—in equal proportions to their wealth.[45]

The transitional dynamics exercise in this section is designed to speak directly to the origins behind three important recent macroeconomic trends: (a) the rising private wealth to GDP ratio (Piketty and Zucman 2015); (b) the decline in the real interest rate since the 1980s (Laubach and Williams 2003, 2016); (c) the large and rapid increase in U.S. wealth inequality (Saez and Zucman 2016).

Figure 8 (blue line) depicts the model-implied transitional dynamics for these three outcomes.

[45]While the two assets are perfect substitutes in this economy, they will be subject to different valuation changes in an extension below.

Figure 8: Transitional dynamics from rising permanent income inequality.

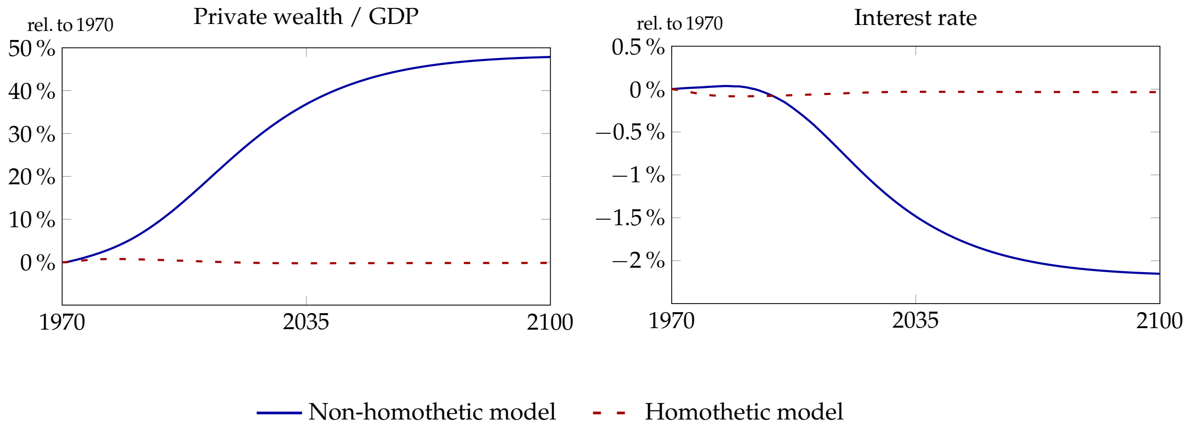


Note. The figures show the percentage-point responses of a non-homothetic and a homothetic economy to an increase in permanent income inequality. The responses are expressed relative to 1970. The bottom plots also show the evolution wealth inequality in the data from Saez and Zucman (2016) (right vertical axis, with same scale). The dashed green line is the transition in a Lucas-tree economy, with a fixed capital stock K.

The top left panel shows that there is still a significant rise in private wealth (relative to to GDP), which increases by around 30 percentage points. It is, however, not quite as pronounced as the one in partial equilibrium. This partly because in general equilibrium, interest rates are endogenous and adjust to clear asset markets. The top right panel shows that, indeed, interest rates declined by around 1 percentage point. Both numbers, the rise in wealth as well as the decrease in interest rates explain around one third of their respective recent trends in the US. Non-surprisingly, the homothetic economy does not predict much movement in wealth or interest rates in response to widening income inequality.

The bottom two panels show the evolution of wealth inequality in the models as well as in the data. As can be seen in the bottom left panel, the non-homothetic economy first exhibits a period of relative stability of the top 10% wealth share, before it starts rising strongly, with a speed after 1980

Figure 9: Long-run view on the transitional dynamics from rising inequality.



Note. This figure plots the evolution of private wealth to GDP and interest rates (in percentage point changes relative to 1970) over a longer time horizon than Figure 8. Observe the backloaded nature of the transition despite constant income inequality after 2014.

that is similar to the one in the data.[46] The model-implied top 1% wealth share follows a similar trajectory, also predicting a strong increase, although here the movements in the data are even stronger than those in the model.[47]

6.3 Long-run effects of income inequality

One of the most striking aspects of Figure 8 is how *backloaded* the responses are. None of the plots are even close to having converged by 2020—even though income inequality was assumed to stop rising in 2014—and some plots show significant delays before they even start moving. Why is this?

The key to understanding the delayed effects of income inequality lies in the life cycle. In the model, rising income inequality changes the distribution of labor incomes across skill groups. Who benefits from this the most? Young, skilled workers that just entered the labor market. They are set to benefit from rising skill premia for their entire working lives. Yet, this will not be immediately reflected in their financial wealth, as their peak earnings (and savings) occurs with a lag of 20-30 years due to the nature of the hump-shaped age-efficiency profile.

This not only explains why some of the time series in Figure 8 move with considerable delay. It also suggests that the transition of the quantities in Figure 8 is far from complete. To see that this is correct, Figure 9 plots the long-run evolution of the wealth-to-GDP ratio and interest rates. As can be seen, the former keeps rising by almost as much as it did until 2020, while the interest rate keeps falling by another full percentage point. These striking predictions demonstrate that even if income inequality were stagnant from now on, its recent surge casts a long a shadow into the future.

[46] The decline in the top 10% wealth share in during the 1970s is not modeled here and would require additional model ingredients (such as surprise inflation). This is why the focus here is on wealth inequality since the mid-1980s.

[47] There is some debate as to whether the top 1% wealth share in Saez and Zucman (2016) is measured correctly. See Kopczuk (2015).

6.4 Aggregate saving rates

The models considered so far work with a standard frictionless supply side with a flexible capital stock. This means that the price of capital is always equal to 1 so that any movements in private wealth (such as those in Figures 8 and 9) must be accommodated by a rising quantity of capital. Such an increase is counterfactual, however: it is well-known that capital and investment, if anything, have declined in recent decades, not risen. In other words, the *aggregate (personal) saving rate*—as computed in the national accounts—has fallen. This is greatly at odds with the standard frictionless supply side; various solutions are being debated in a recent literature (Philippon and Gutiérrez 2017; Karabarbounis and Neiman 2018; Farhi and Gourio 2019).

To see whether the effects of rising income inequality on wealth and interest rates rely on a quantity response, I recomputed the transitional dynamics in the non-homothetic model in a "Lucas tree" economy, where the capital stock is fixed exogenously to its 2014 level. In this economy, any movements in private wealth are due to changing prices of capital, not quantities. By design, the aggregate saving rate is constant in this economy! Capital gains due to rising prices of capital are not counted towards the aggregate saving rate in national income accounting.

The dashed green line in Figure 8 shows the evolution of wealth, interest rates and wealth inequality in that economy. Overall, the transition is similar to that in the flexible-K model, confirming that the behavior of the aggregate saving rate in the data is not contradicting a strong rise in private wealth (which has been even more pronounced in the data).

In sum, the non-homothetic model predicts a greater supply of savings in response to larger income inequality. Whether those savings manifest themselves in rising quantities—as in the baseline model—or rising prices—as in the "Lucas tree" economy—depends entirely on the demand for savings implied by the supply side.

6.5 Individual saving rates

The previous subsection has demonstrated that the non-homothetic model is consistent with the idea that capital gains have been driving aggregate saving rates in the US. But what about *individual* saving rates? Does the non-homothetic model imply that rich households save "actively"—that is, by accumulating a greater quantity of capital—or "passively"—by letting a fixed quantity of capital appreciate over time? Which of the two types of savings behavior is responsible for the recent increase in wealth inequality?

This is an important question that has prompted a number of recent interesting studies (Bach, Calvet and Sodini 2018; Fagereng et al. 2019b; Gomez 2019; Kuhn, Schularick and Steins 2019), which overall find a strong role for passive saving, i.e. capital gains. For example, Fagereng et al. (2019b) compute "gross" saving rates (inclusive of capital gains) for households in different percentiles of the wealth distribution, documenting increasing saving rates with wealth. However, when they subtract capital gains from the saving rate's numerator and denominator—a "net" saving rate—the saving rate is largely constant in wealth.

Table 8: Saving rates.

	type	Savings rate by wealth fractile				
		P0-P25	P25-P50	P50-P90	P90-P99	P99-P100
Non-homothetic model	gross	12.3%	1.8%	0.6%	21.0%	26.0%
	net	12.3%	1.1%	−2.7%	13.0%	8.5%
Homothetic model	gross	12.4%	21.0%	17.4%	10.9%	3.3%
	net	11.7%	18.6%	10.9%	0.8%	−10.6%

Note. This table shows average saving rates by wealth fractiles, computed as follows. A model household's gross saving rates is defined as the change in net worth divided by total income. A model household's net saving rate is defined as the change in net worth less capital gains divided by total income less capital gains.

As such, not even the "Lucas tree" version of my model can speak to this issue as it does not generate any capital gains in the steady state. To generate steady-state capital gains, it is necessary to introduce growth into the model.[48] I assume a growth rate of $g = 3\%$ in line with historical growth rates in the US. With an otherwise equivalent calibration, this means the pre-tax return on capital is no longer just given by a 5% dividend yield but now also includes 3% capital gains. Denote by r^{cg} the rate of capital gains after taxes.

Table 8 computes gross saving rates $\frac{\Delta a_{it+1}}{y_{it}+ra_{it}}$ and net saving rates $\frac{\Delta a_{it+1}-r^{cg}a_{it}}{y_{it}+(r-r^{cg})a_{it}}$ by wealth fractiles in the growth versions of the non-homothetic model and the homothetic model.[49] The gross saving rates in the non-homothetic economy are qualitatively similar to those documented in Fagereng et al. (2019b): they are high for poor households, decrease and then increase strongly again for wealthier households. Net saving rates also share some similarity with the evidence in Fagereng et al. (2019b) in that there is no longer a noticeable trend for saving rates to rise in wealth. Neither of the two features are matched by the homothetic economy.

In sum, this illustrates that, when equipped with a supply side that generates capital gains, the non-homothetic model attributes a sizable role of individual saving rate heterogeneity to capital gains. The exact magnitude, however, depends on the specific supply side. The only objects in the non-homothetic model that are independent of the supply side are the gross saving rates.

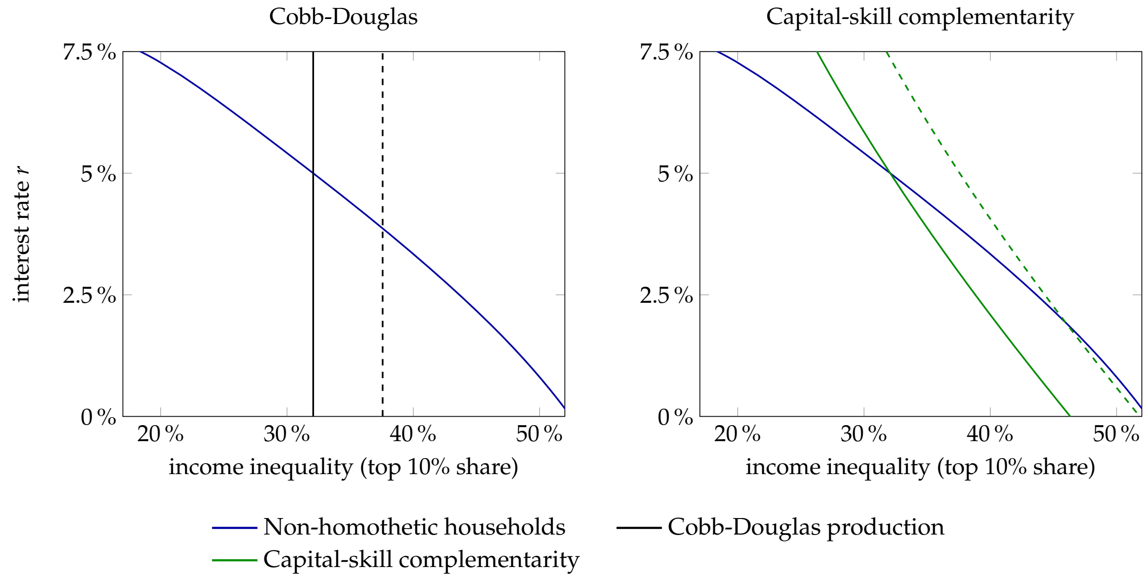
6.6 Endogenous income inequality

So far, the distribution of labor income $\{\gamma_s\}$ was assumed to be entirely exogenous. Yet, a number of recent papers have put forth the idea that capital-skill complementarity together with capital deepening may have been responsible for some of the recent changes in income inequality (see e.g. Krusell et al. 2000, Autor et al. 2003, or more recently Acemoglu and Restrepo 2017). Such production functions give rise to a role for interest rates to affect the labor income distribution, by way skill-complementary investments. In the context of the current model, this is particularly

[48] I do this by assuming the normalization factor o grows proportionately with productivity, see Appendix G.4 for details.

[49] Note that growth is irrelevant for gross saving rates.

Figure 10: Amplification: non-homotheticity and capital-skill complementarity.



Note. These figures illustrate the interaction between interest rates and inequality. The black and green lines are computed by solving for optimal capital choices K for a given interest rate r, then using K to pin down labor income inequality γ. The blue line is computed by iterating over the household problem to solve for r such that household capital demand equals K, for each given level of income inequality γ. Intersections between the black / green lines and the colored lines correspond to steady states of the respective models. The dashed lines represent an exogenous shift in skill-biased technical change.

interesting, as a non-homothetic household side generates an interaction which runs the other way around, from income inequality to interest rates.

To explore this two-way feedback mechanism, I modify the production function in the spirit of Krusell et al. (2000),

$$Y = A\left(\gamma_1 L_1^{\alpha_1} + (1-\gamma_1)\left(\gamma_2 K^{\alpha_2} + (1-\gamma_2)L_2^{\alpha_2}\right)^{\alpha_1/\alpha_2}\right)^{1/\alpha_1}. \tag{13}$$

I calibrate the economy just like before, so that in a 2014 steady state the labor income distribution, the capital stock and the interest rate are the same. The elasticities α_1, α_2 are taken from Krusell et al. (2000).[50] Moreover, since the production function only has two types of skills, I work with a two-skill version of the non-homothetic model, where unskilled labor L_1 makes up 90% of the population and skilled labor L_2 makes up 10%. I denote by $\gamma \equiv \frac{w_2 L_2}{w_1 L_1 + w_2 L_2}$ the share of labor income earned by the top 10%.

Why does a production function along the lines of (13) lead to a conceptually very different relationship between interest rates and income inequality γ than a simple Cobb-Douglas production function? The black and green lines in Figure 10 illustrate this. The left panel (black line) shows that income inequality γ is entirely independent of interest rates r in a Cobb-Douglas economy; after all,

[50] This yields: $\gamma_1 = 0.39, \gamma_2 = 0.95, \alpha_1 = 0.4, \alpha_2 = -0.5$. As usual, A is used for normalization of Y to 1 in steady state.

Table 9: Effects of rising income inequality for greater ϕ.

	Implied ϕ	Change from 1970 to 2014 levels of income inequality (*pp.*) Wealth / GDP	Interest rate	Top 10% wealth share
Non-homothetic model	0.70	53	-2.17	24
	0.80	37	-1.40	23
	0.90	25	-0.91	18
Non-homothetic status model	0.70	47	-1.89	24
	0.80	35	-1.32	20
	0.90	20	-0.72	15
Homothetic model	1.00	1	-0.04	6

Note. This table shows how the effects of rising income inequality depend on the level of the targeted curvature of the consumption function (in permanent income) ϕ. $\phi = 0.80$ and $\phi = 0.90$ denote model versions that have been recalibrated to match the same targets other than ϕ.

γ is just a parameter of the production function. With capital-skill complementarity (right panel, green line), however, lower interest rates lead to greater capital accumulation, which raises skill premia and thus income inequality γ.

The dashed solid blue lines in Figure 10 show what happens when the production functions are combined with a non-homothetic household side. It is constructed by solving for the equilibrium interest rate that would be implied when the top 10% income share among households is fixed at γ, instead of endogenously pinned down by the production function (13). In line with our results before, an increase in income inequality significantly pushes down interest rates in a non-homothetic model. Intersections in Figure 10 correspond to steady states in the economy with capital-skill complementarity.

Both capital-skill complementarity and non-homothetic savings behavior predict downward-sloping relationships between income inequality and interest rates. This may lead to amplification of shocks. To illustrate this, consider an "inequality shock" (e.g. due to skill-biased technical change) which exogenously shifts outward the production curve in the $r - \gamma$ diagram in Figure 10. While in the left panel, this leads to a modest increase in income inequality and reduction in interest rates, in the right panel, both effects are greatly amplified. In practice, such amplification would manifest itself in a sustained, and mostly *endogenous*, increase in income inequality over many decades, as well as a sustained endogenous decline in interest rates.

6.7 Robustness

The analysis so far was based on a non-homothetic model that was calibrated to match a permanent income elasticity of consumption of $\phi = 0.70$. Yet, this estimate clearly has error bands around it. This section explores the sensitivity of the macro results with respect to ϕ, and with respect to an alternative model (the status model, see Appendix G.1). To do so, I re-calibrate both non-homothetic models to match either $\phi = 0.80$ or $\phi = 0.90$, as alternative choices of ϕ. The remaining calibration is unchanged (still to match data from 2014). Then, I compare a steady state with the 1970 income distribution to the one with the 2014 income distribution.

Table 9 shows the results. The non-homothetic status model with $\phi = 0.70$ is predicts similar effects from changes in income inequality to the ones implied by the baseline non-homothetic model with $\phi = 0.70$. Moreover, even the economies with $\phi = 0.80$ and $\phi = 0.90$ predict sizable effects of changes in income inequality, although the effects with $\phi = 0.90$ are noticeably more modest. This highlights once more the tight connection between the curvature of the consumption function and the aggregate effects of income inequality.

7 Conclusion

The ultimate goal of this paper was to study the implications of rising inequality in permanent income for macroeconomic aggregates, as well as for the wealth distribution. Motivated by a linear model in which the distribution of permanent income is irrelevant, I proposed new ways to measure the relationship between consumption and permanent income in the data, finding an elasticity of around $\phi \approx 0.7$. I then showed that a quantitative model that incorporates non-homotheticities in life-cycle spending can match this estimate, while standard models cannot. This has important consequences: the recent rise in permanent-income inequality has pushed interest rates down and generated a rise in wealth inequality almost as rapid as that in the data. Moreover, there is an important interaction with capital-skill complementarity, which gives rise to amplification. An initial increase in skill-biased technical change may increase inequality, thus lowering interest rates, which then again may lead to skill-biased capital investments.

There are two avenues related to this project that are well worth exploring in future research. First, in the quantitative model of Section 5, non-homothetic preferences drive almost the entire observed concave relationship between consumption and permanent income. This allowed me to conduct counterfactual analyses with the right overall level of concavity, but avoided the question of which forces at a microeconomic level account for how much of the observed concavity. Knowing the sources of concavity, however, is important.

Second, it would be very fruitful to integrate the type of non-homothetic preferences studied in this paper with models that include multiple assets and return heterogeneity (as in Kaplan and Violante 2014). In such a model, increases in income inequality would endogenously increase the demand for the high-return asset, bidding up its price and amplifying the impact on wealth inequality. Moreover, the model would be able to speak to the return differential between the two assets. A richer model along these lines would be especially worthy of further research.

References

Abbott, Brant and Giovanni Gallipoli, ""Permanent Income" Inequality," 2019.

Abowd, John M and David Card, "On the Covariance Structure of Earnings and Hours Changes," *Econometrica*, 1989, *57* (2), 411–445.

Acemoglu, Daron and Pascual Restrepo, "The Race Between Man and Machine: Implications of Technology for Growth, Factor Shares and Employment," *American Economic Review*, 2017, *108* (6), 1488–1542.

Aguiar, Mark and Mark Bils, "Has Consumption Inequality Mirrored Income Inequality?," *American Economic Review*, 2015, *105* (9), 2725–2756.

Aiyagari, S. Rao, "Uninsured Idiosyncratic Risk and Aggregate Saving," *Quarterly Journal of Economics*, 1994, *109* (3), 659–684.

Alan, Sule, Kadir Atalay, and Thomas F. Crossley, "Do the Rich Save More? Evidence from Canada," *Review of Income and Wealth*, 2015, *61* (4), 739–758.

Altonji, Joseph G. and Lewis M. Segal, "Small-Sample Bias in GMM Estimation of Covariance Structures," *Journal of Business and Economic Statistics*, 1996, *14* (3), 353–366.

Alvaredo, Facundo, Bertrand Garbinti, and Thomas Piketty, "On the Share of Inheritance in Aggregate Wealth: Europe and the USA, 1900 - 2010," *Economica*, 2017, *84* (334), 239–260.

Ameriks, John and Stephen P. Zeldes, "How do household portfolio shares vary with age," 2004.

Andreski, Patricia, Geng Li, Mehmet Zahid Samancioglu, and Robert F. Schoeni, "Estimates of annual consumption expenditures and its major components in the PSID in comparison to the CE," *American Economic Review*, 2014, *104* (5), 132–135.

Aoki, Shuhei and Makoto Nirei, "Zipf's law, pareto's law, and the evolution of top incomes in the United States," *American Economic Journal: Macroeconomics*, 2017, *9* (3), 36–71.

Atkinson, Anthony B., "Capital Taxes, the Redistribution of Wealth and Individual Savings," *Review of Economic Studies*, 1971.

Attanasio, Orazio P. and Luigi Pistaferri, "Consumption Inequality," *Journal of Economic Perspectives*, 2016, *30* (2), 3–28.

— **and Martin Browning**, "Consumption over the Life Cycle and over the Business Cycle," *American Economic Review*, 1995, *85* (5), 1118–1137.

— **and Steven J. Davis**, "Relative Wage Movements and the Distribution of Consumption," *Journal of Political Economy*, 1996, *104* (6), 1227–1262.

Auclert, Adrien and Matthew Rognlie, "Inequality and Aggregate Demand," 2018.

— **, Bence Bardoczy, Matthew Rognlie, and Ludwig Straub**, "Using the Sequence-Space Jacobian to Solve and Estimate Heterogeneous-Agent Models," 2019.

Autor, David H., Frank Levy, and Richard J. Murnane, "The Skill Content of Recent Technological Change: An Empirical Exploration," *Quarterly Journal of Economics*, 2003, (November), 1279–1334.

— **, Lawrence F. Katz, and Melissa S. Kearney**, "Trends in U.S. Wage Inequality: Revising the Revisionists," *The Review of Economics and Statistics*, 2008, *90* (2), 300–323.

Bach, Laurent, Laurent E. Calvet, and Paolo Sodini, "From Saving Comes Having? Disentangling the Impact of Savings on Inequality," 2018.

Becketti, Sean, William Gould, Lee Lillard, and Finis Welch, "The Panel Study of Income Dynamics after Fourteen Years: An Evaluation," *Journal of Labor Economics*, 1988, *6* (4), 472–492.

Benabou, Roland, "Unequal Societies: Income Dstribution and Social Contract," *The American Economic Review*, 2000, *90* (1), 96–129.

Benhabib, Jess, Alberto Bisin, and Mi Luo, "Earnings inequality and other determinants of wealth inequality," *American Economic Review*, 2017, *107* (5), 593–597.

__ , __ , and __ , "Wealth distribution and social mobility in the US: A quantitative approach," *American Economic Review*, 2018, *21721*, 1–32.

__ and __ , "Skewed Wealth Distributions: Theory and Empirics," *Journal of Economic Literature*, 2017, *56*, 1261–91.

Bertola, Giuseppe, Reto Foellmi, and Josef Zweimuller, *Income Distribution in Macroeconomic Models* 2005.

Blinder, Alan S, "Distribution Effects and the Aggregate Consumption Function," *Journal of Political Economy*, 1975, *83* (3), 447–475.

Blundell, Richard, Luigi Pistaferri, and Ian Preston, "Consumption Inequality and Partial Insurance," *American Economic Review*, 2008, *98* (5), 1887–1921.

__ , __ , and Itay Saporta-Eksten, "Consumption Inequality and Family Labor Supply," *American Economic Review*, 2016, *106* (2), 387–435.

__ , Martin Browning, and Costas Meghir, "Consumer Demand and the Life-Cycle Allocation of Household Expenditures," *Review of Economic Studies*, 1994, (61), 57–80.

Boar, Corina, "Dynastic Precautionary Savings," 2018.

Bozio, Antoine, Carl Emmerson, Cormac O'Dea, and Gemma Tetlow, "Savings and wealth of the lifetime rich: Evidence from the UK and US," 2013.

Cagetti, Marco and Mariacristina De Nardi, "Entrepreneurship, Frictions, and Wealth," *Journal of Political Economy*, 2006, *114* (5), 835–870.

Carroll, Christopher D., "Buffer-Stock Saving and the Life Cycle / Permanent Income Hypothesis," *Quarterly Journal of Economics*, 1997, *CXII* (1), 1–55.

__ , "Why Do the Rich Save So Much?," *Does Atlas Shrug? The Economic Consequences of Taxing the Rich.*, 2000, *D*, 465–84.

__ , "The method of endogenous gridpoints for solving dynamic stochastic optimization problems," *Economics letters*, 2005, pp. 312–320.

__ , "Precautionary saving and the marginal propensity to consume out of permanent income," *Journal of Monetary Economics*, 2009, *56* (6), 780–790.

__ and Miles S Kimball, "On the Concavity of the Consumption Function," *Econometrica*, 1996, *64* (4), 981–992.

Castaneda, Ana, Javier Díaz Giménez, and José-Víctor Ríos-Rull, "Accounting for the U.S. Earnings and Wealth Inequality," *Journal of Political Economy*, 2003, *111* (4), 818–857.

Chamberlain, Gary, "Chapter 22 Panel data," *Handbook of Econometrics*, 1984, *2*, 1247–1318.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez, "Where is the Land of Opportunity? The Geopgraphy of Intergenerational Mobility in the United States," *Quarterly Journal of Economics*, 2014, *129* (4), 1553–1623.

Cole, Harold L., George J. Mailath, and Andrew Postlewaite, "Social Norms, Savings Behavior, and Growth," *Journal of Political Economy*, 1992, *100* (6), 1092–1125.

__ , __ , **and** __ , "Incorporating Concern for Relative Wealth into Economic Models," *Federal Reserve Bank of Minneapolis Quarterly Review*, 1995, *19* (3), 12–21.

__ , __ , **and** __ , "Class systems and the enforcement of social norms," *Journal of Public Economics*, 1998, *70* (1), 5–35.

Constantinides, George M. and Darrell Duffie, "Asset Pricing with Heterogeneous Consumers," *Journal of Political Economy*, 1996, *104* (2), 219–240.

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

__ **and Fang Yang**, "Bequests and heterogeneity in retirement wealth," *European Economic Review*, 2014, *72*, 182–196.

__ **and Giulio Fella**, "Saving and Wealth Inequality," *Review of Economic Dynamics*, 2017, *26*, 280–300.

__ , __ , **and Fang Yang**, "Piketty's Book and Macro Models of Wealth Inequality," *NBER Working Paper No. 21730*, 2015, pp. 2–48.

Deaton, Angus, "Saving and Liquidity Constraints," *Econometrica*, 1991, *59* (5), 1221–1248.

__ **and Christina Paxson**, "Intertemporal Choice and Inequality," *The Journal of Political Economy*, 1994, *102* (3), 437–467.

DeBacker, Jason, Bradley Heim, Vasia Panousi, Shanthi Ramnath, and Ivan Vidangos, "Rising inequality : Transitory or permanent ? New evidence from a panel of U.S . tax returns," *Brookings Papers on Economic Activity*, 2013, (1), 67–142.

Dynan, Karen E., Jonathan Skinner, and Stephen P. Zeldes, "Do the Rich Save More?," *Journal of Political Economy*, 2004, *112* (2), 397–444.

Fagereng, Andreas, Luigi Guiso, Davide Malacrino, and Luigi Pistaferri, "Heterogeneity and Persistence in Returns to Wealth," 2019.

__ **, Martin Blomhoff Holm, Benjamin Moll, and Gisle J. Natvik**, "Saving Behavior Across the Wealth Distribution," 2019.

Farhi, Emmanuel and François Gourio, "Accounting for Macro-Finance Trends: Market Power, Intangibles, and Risk Premia," 2019.

Favilukis, Jack, "Inequality, stock market participation, and the equity premium," *Journal of Financial Economics*, 2013, *107* (3), 740–759.

Fisher, Irving, *The Theory of Interest*, New York: Macmillan, 1930.

Flavin, Marjorie A., "The Adjustment of Consumption to Changing Expectations About Future Income," *Journal of Political Economy*, 1981, *89* (5), 974–1009.

Friedman, Milton, *A Theory of the Consumption Function*, Princeton University Press, 1957.

Gabaix, Xavier, Benjamin Moll, Jean-Michel Lasry, and Pierre-Louis Lions, "The Dynamics of Inequality," *Econometrica*, 2017.

Gomez, Matthieu, "Asset Prices and Wealth Inequality," 2019.

Gourinchas, Pierre-Olivier and Jonathan A. Parker, "Consumption over the life cycle," *Econometrica*, 2002, *70* (1), 47–89.

Guvenen, Fatih, "Learning your earning: Are labor income shocks really very persistent?," *American Economic Review*, 2007, *97* (3), 687–712.

__ , "An empirical investigation of labor income processes," *Review of Economic Dynamics*, 2009, *12* (1), 58–79.

__ , "Macroeconomics with Heterogeneity: A Practical Guide," *Economic Quarterly*, 2011, *97* (3), 255–326.

__ **and Anthony A. Jr. Smith**, "Inferring Labor Income Risk and Partial Insurance from Economic Choices," *Econometrica*, 2014, *82* (6), 2085–2129.

__ **, Fatih Karahan, Serdar Ozkan, and Jae Song**, "What do Data on Millions of U.S. Workers Reveal About Life-Cycle Earnings Risk?," 2019.

__ **, Greg Kaplan, Jae Song, and Justin Weidner**, "Lifetime Incomes in the United States Over Six Decades," 2018.

__ **, Serdar Ozkan, and Jae Song**, "The Nature of Countercyclical Income Risk," *Journal of Political Economy*, 2014, *122* (3), 621–660.

Hall, Robert E., "Stochastic Implications of the Lifecycle Permanent Income Hypothesis: Theory and Evidence," 1978.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante, "Quantitative Macroeconomics with Heterogeneous Households," *Annual Review of Economics*, 2009, *1*, 319–354.

__ **,** __ **, and** __ , "The Macroeconomic Implications of Rising Wage Inequality in the United States," *Journal of Political Economy*, 2010, *118* (4), 681–722.

__ **,** __ **, and** __ , "Optimal Tax Progressivity: An Analytical Framework," *Quarterly Journal of Economics*, 2017, *forthcomin.*

Hendricks, L, "Bequests and retirement wealth in the United States," 2001.

Holm, Martin Blomhoff, "Monetary policy transmission with income risk," 2018.

Houthakker, Hendrik S., "Additive Preferences," *Econometrica*, 1960, *28* (2), 244–257.

Hryshko, Dmytro and Iourii Manovskii, "How Much Consumption Insurance in the U.S.," 2017.

Hubbard, R. Glenn, Jonathan Skinner, and Stephen P. Zeldes, "The importance of precautionary motives in explaining individual and aggregate saving," *Carnegie-Rochester Confer. Series on Public Policy*, 1994, *40*, 59–125.

__ **,** __ **, and** __ , "Precautionary Saving and Social Insurance," *Journal of Political Economy*, 1995, *103* (2), 360–399.

Hubmer, Joachim, Per Krusell, and Anthony A. Jr. Smith, "The Historical Evolution of the Wealth Distribution: A Quantitative-Theoretic Investigation," *NBER working paper*, 2016, *WP 23011* (NBER), 37.

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

__ **and Gustavo Ventura**, *Understanding why high income households save more than low income households*, Vol. 45 2000.

__ **,** __ **, and Amir Yaron**, "Sources of lifetime inequality," *American Economic Review*, 2011, *101* (7), 2923–2954.

Jappelli, Tullio and Luigi Pistaferri, "Intertemporal Choice and Consumption Mobility," *Journal of the European Economic Association*, 2006, *00235* (March), 75–115.

— **and** — , "Fiscal policy and MPC heterogeneity," *American Economic Journal: Macroeconomics*, 2014, *6* (4), 107–136.

Johnson, David S., Jonathan A. Parker, and Nicholas S. Souleles, "Household expenditure and the income tax rebates of 2001," *American Economic Review*, 2006, *96* (5), 1589–1610.

Kaplan, Greg and Giovanni L. Violante, "How Much Consumption Insurance Beyond Self-Insurance?," *American Economic Journal: Macroeconomics*, 2010, *2* (4), 53–87.

— **and** — , "A Model of the Consumption Response to Fiscal Stimulus Payments," *Econometrica*, 2014, *82* (4), 1199–1239.

Karabarbounis, Loukas and Brent Neiman, "Accounting for Factorless Income," *NBER Macroeconomics Annual*, 2018.

Katz, Lawrence F. and Kevin M. Murphy, "Changes in Relative Wages, 1963-1987: Supply and Demand Factors," *The Quarterly Journal of Economics*, 1992, *107* (1), 35–78.

Kaymak, Bariş and Markus Poschke, "The evolution of wealth inequality over half a century: The role of taxes, transfers and technology," *Journal of Monetary Economics*, 2016, *77*, 1–25.

Keynes, John Maynard, *The General Theory of Employment, Interest and Money*, Palgrave Macmillan, 1936.

Kindermann, Fabian and Dirk Krueger, "High Marginal Tax Rates on the Top 1%? Lessons from a Life Cycle Model with Idiosyncratic Income Risk," 2017.

Kopczuk, Wojciech, "What Do We Know about the Evolution of Top Wealth Shares in the United States?," *Journal of Economic Perspectives*, 2015, *29* (1), 47–66.

— **, Emmanuel Saez, and Jae Song**, "Earnings Inequality and Mobility in the United States: Evidence From Social Security Data Since 1937," *Quarterly Journal of Economics*, 2010.

Kopecky, Karen A. and Richard M. H. Suen, "Finite state Markov-chain approximations to highly persistent processes," *Review of Economic Dynamics*, 2010, *13* (3), 701–714.

Krueger, Dirk and Alexander Ludwig, "On the optimal provision of social insurance: Progressive taxation versus education subsidies in general equilibrium," *Journal of Monetary Economics*, 2016, *77*, 72–98.

— **and Fabrizio Perri**, "Does Income Inequality Lead to Consumption Equality? Evidence and Theory," *Review of Economic Studies*, 2006, *73* (June), 163–193.

Krusell, Per and Anthony A. Jr. Smith, "Income and Wealth Heterogeneity, Portfolio Choice, and Equilibrium Asset Returns," *Macroeconomic Dynamics*, 1997, *1* (02), 387–422.

— **and** — , "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 1998, *106* (5), 867–896.

— **and** — , "Quantitative macroeconomic models with heterogeneous agents," *Econometric Society Monographs*, 2006, *41* (March), 298.

— **, Lee E. Ohanian, Jose-Victor Rios-Rull, and Giovanni L. Violante**, "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis," *Econometrica*, 2000, *68* (5), 1029–1053.

Kuhn, Moritz, Moritz Schularick, and Ulrike I. Steins, "Income and Wealth Inequality in America," 2019.

Kumhof, Michael, Romain Rancière, and Pablo Winant, "Inequality , Leverage and Crises," *American Economic Review*, 2015, *105* (3), 1217–1245.

Laubach, Thomas and John C Williams, "Measuring the Natural Rate of Interest," *Review of Economics and Statistics*, 2003, *85* (4), 1063–1070.

— **and** — , "Measuring the Natural Rate of Interest Redux," *Business Economics*, 2016, *51* (2), 57–67.

MaCurdy, Thomas E., "The Use of Time Series Process to Model the Error Structure of Earnings in a Longitudinal Data Analysis," *Journal of Econometrics*, 1982, *18*, 83–114.

Mayer, Thomas, *Permanent Income, Wealth, and Consumption* 1972.

Meghir, Costas and Luigi Pistaferri, "Earnings, Consumption and Lifecycle Choices," *Handbook of Labor Economics*, 2011.

Muth, John F, "Optimal Properties of Exponentially Weighted Forecasts," *Journal of the American Statistical Association*, 1960, *55* (290), 299–306.

Olivi, Alan, "Revealed Preferences and Beliefs from Consumption-Savings Decisions," 2019.

Ömer T. Acikgöz, "On the existence and uniqueness of stationary equilibrium in Bewley economies with production," *Journal of Economic Theory*, 2017, (October).

Philippon, Thomas and Germán Gutiérrez, "Investmentless Growth: An Empirical Investigation," *Brookings Papers on Economic Activity*, 2017.

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

— **and Gabriel Zucman**, "Capital is Back: Wealth-Income Rations in Rich Countries 1700–2010," *Quarterly Journal of Economics*, 2014, *129* (3), 1255–1310.

— **and** — , "Wealth and inheritance in the long run," *Handbook of Income Distribution*, 2015, *2*, 1303–1368.

Pischke, Jörn-Steffen, "Individual Income, Incomplete Information, and Aggregate Consumption," *Econometrica*, 1995, *63* (4), 805–840.

Poterba, James and Todd Sinai, "Tax expenditures for owner-occupied housing: Deductions for property taxes and mortgage interest and the exclusion of imputed rental income," *American Economic Review*, 2008, *98* (2), 84–89.

Primiceri, Giorgio E. and Thijs van Rens, "Heterogeneous life-cycle profiles, income risk and consumption inequality," *Journal of Monetary Economics*, 2009, *56* (1), 20–39.

Quadrini, Vincenzo, "Entrepreneurship, Saving, and Social Mobility," *Review of Economic Dynamics*, 2000, *3* (1), 1–40.

— **and José-Víctor Ríos-Rull**, *Inequality in macroeconomics*, 1 ed., Vol. 2, Elsevier B.V., 2015.

Rognlie, Matthew, "Deciphering the fall and rise in the net capital share," *Brookings Papers on Economic Activity*, 2015, (Spring), 1–56.

Rouwenhorst, K.G., "Asset pricing implications of equilibrium business cycle models," in T.F. Cooley, ed., *Frontiers of Business Cycle Research*, 1995, pp. 294–330.

Sabelhaus, John and Jae Song, "The great moderation in micro labor earnings," *Journal of Monetary Economics*, 2010, *57* (4), 391–403.

Saez, Emmanuel and Gabriel Zucman, "Wealth Inequality in the United States Since 1913: Evidence from Capitalized Income Tax Data," *Quarterly Journal of Economics*, 2016, *131* (May), 519–578.

— **and Stefanie Stantcheva**, "A Simpler Theory of Optimal Capital Taxation," *Journal of Public Economics*, 2017, *forthcomin.*

Scholz, John Karl, Ananth Seshadri, and Surachai Khitatrakun, "Are Americans Saving " Optimally " for Retirement?," *Journal of Political Economy*, 2006, *114* (4), 607–643.

Storesletten, Kjetil, Christopher I. Telmer, and Amir Yaron, "Consumption and risk sharing over the life cycle," *Journal of Monetary Economics*, 2004, *51* (3), 609–633.

Strotz, Robert H., "Myopia and inconsistency in dynamic utility maximization," *The Review of Economic Studies*, 1955, *23* (3), 165–180.

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

Wachter, Jessica A. and Motohiro Yogo, "Why do household portfolio shares rise in wealth?," *Review of Financial Studies*, 2010, *23* (11), 3929–3965.

Zeldes, Stephen P., "Consumption and Liquidity Constraints: An Empirical Investigation," *Journal of Political Economy*, 1989, *97* (2), 305–346.

Appendix
(for online publication)

A The importance of permanent income inequality

In a recent paper, Guvenen et al. (2018) investigate the dynamics of inequality in lifetime incomes, and inequality by age and cohort. In particular, they show that most of the increase in income inequality is due to rising inequality in initial incomes. Together with other evidence that rising inequality was not due to transitory shocks (Kopczuk et al. 2010; Sabelhaus and Song 2010), this points to a rise in permanent income inequality, or—which turns out to be quite similar—a rise in the initial variance of the persistent component of income.

In this section, I use their data on male earnings from the Continuous Work History Subsample of the U.S. Social Security Administration's Master Earnings File.[51] Figure 11 shows two measures of income inequality, the standard deviation and the $\log(P90/P10)$ ratio. It decomposes the rise in income inequality at the life-cycle earnings peak (50-55 year olds), which is shown in Panel (a), into two pieces, which are shown in Panel (b): inequality that was already present for 30-35 year olds (blue), and a residual (red). The figure illustrates that until the 1990s, both types of inequality increased, even though initial inequality increased more rapidly. Since the 1990s, however, initial inequality seemed to have accounted for more than the observed rise in inequality at ages 50-55.

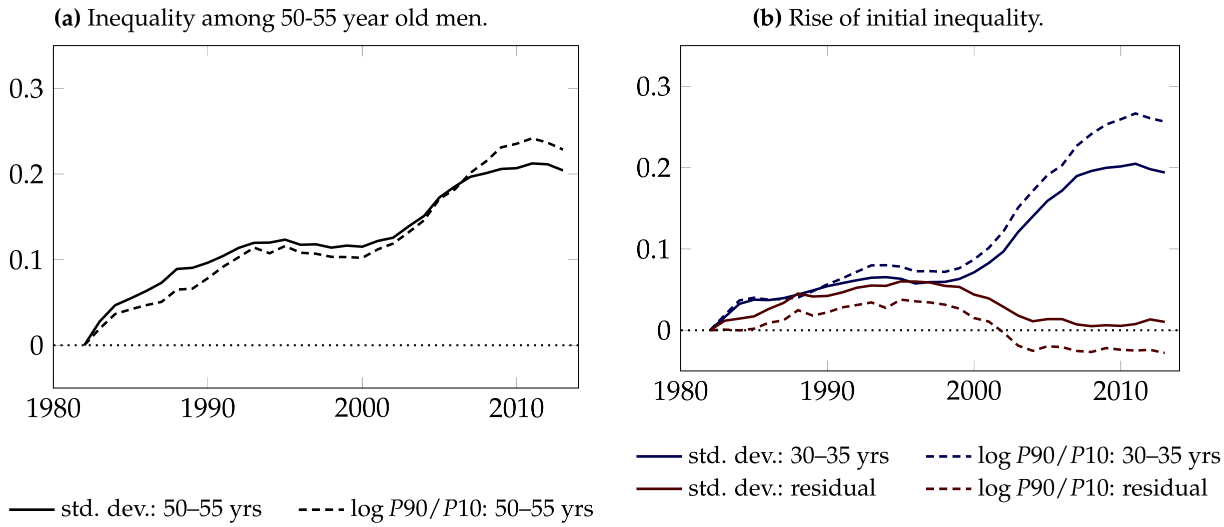
While this evidence will almost surely not be the last word on the importance of permanent income inequality, it does suggest, however, that permanent factors were an important driver of income inequality in recent decades, worthy of a thorough investigation.

B Calibration of the simple model

Calibration. To generate the plots, I calibrate the equality steady state of the two models (homothetic and non-homothetic) as follows. As mentioned in Section 2, the share of rich agents is given by $\mu = 1\%$. The capital share is taken to be $\alpha = 0.33$, the interest rate $r = 0.05$, and depreciation $\delta = 0.06$, giving a capital stock of $K = 3$. The curvature of the per-period utility function is $\sigma = 2.5$, which is also equal to Σ in the homothetic economy. In the non-homothetic economy, $\Sigma = 0.7\sigma$, foreshadowing my empirical results, and the normalization parameter o is chosen so that the economy is in equilibrium. This procedure implies $\beta = 7.2$ and $o = 3.2$. (Notice that β has the role of what usually is $\beta/(1-\beta)$, which explains why it is so large.)

[51] I thank Fatih Guvenen for posting data on inequality statistics by age and cohort online at https://fguvenendotcom.files.wordpress.com/2017/03/gksw2017_figuredata_v1.xlsx.

Figure 11: The importance of rising permanent income inequality.



Note. Panel (a) shows the evolution income inequality at the peak of the life cycle, for 50-55 year old men. Panel (b) decomposes this trend into inequality that is already present for 30-35 year old men, and a residual. The former is inequality in initial incomes, while the latter can be regarded as changes to the life-cycle increase in income inequality. The two measures of inequality plotted here are the standard deviation as well as the $\log(P90/P10)$ ratio, which is divided by 2 to make it roughly comparable to the standard deviation. The initial year is the first for which all measures are available in the data of Guvenen et al. (2018).

Figure 1. Combining the Euler equation

$$c_t/o = (\beta R)^{-1/\sigma} (a_{t+1}/o)^{\Sigma/\sigma}, \tag{14}$$

with the budget constraint (2), the consumption policy function $c(a+w)$ can be found by solving the implicit equation

$$c/o + R^{-1} (\beta R)^{1/\Sigma} (c/o)^{\sigma/\Sigma} = (a+w)/o.$$

The savings schedule after 20 years, $a_{20}(w)$, starting at the perfect equality steady state where per head assets are equal to K, can be found by iterating the asset policy function $a(a+w) = R(a+w-c(a+w))$. The steady state savings schedule $a_\infty(w)$ can be found as the solution to the Euler equation, after replacing c by its steady state value of $c = (1-R^{-1})a + w$,

$$a = o\,(\beta R)^{1/\Sigma} \left(\frac{(1-R^{-1})a + w}{o} \right)^{\sigma/\Sigma}. \tag{15}$$

Figure 2. In general equilibrium, the interest rate R is endogenously determined, as the value of total assets in the economy, $R^{-1}A$, must equal the capital stock K. K is determined by

$$\alpha \frac{Y}{K} = R - 1 + \delta.$$

This equation and (15) jointly describe steady state assets A and the steady state interest rate R. In principle, this system can have multiple solutions, but in this calibration, there is a unique solution. The calculation of the other quantities in Figure 2 is standard.

C Omitted proofs

C.1 Definition of equilibrium

Denote the state space by $\mathbb{S} \equiv \mathcal{S} \times \{1, \ldots, K_{death}\} \times \mathbb{R}_+ \times \mathcal{Z} \times \{0, 1\}$ endowed with the Borel σ-algebra $\mathcal{B}_\mathbb{S}$ on $\mathbb{S}$. I define a steady state equilibrium as follows.

Definition 1 (Steady-state equilibrium)**.** A steady state equilibrium in the benchmark economy is a vector of aggregate quantities $\{Y, K, L_s\}$, a probability distribution μ defined over $(\mathbb{S}, \mathcal{B}_\mathbb{S})$ and a measure of bequests χ defined over $\mathcal{S} \times \{1, \ldots, K_{death}\} \times \mathbb{R}_+$ with the Borel σ-algebra, a set of policy functions $\{\mathbf{c}_{k,s}(a, z, \varphi), \mathbf{a}_{k,s}(a, z, \varphi)\}$, a set of prices $\{r, w_s\}$ such that: (a) the policy functions solve the optimization problem (4), where the conditional bequest distribution $v(\cdot|s, k, \varphi)$ is given by

$$v(\mathbf{B}, \varphi'|s, k, \varphi) = \begin{cases} 1_{\{0,1\}}(\mathbf{B}, \varphi') & \text{if } \varphi = 1 \\ (1 - \delta_{k+k_{born}})1_{\{0,0\}}(\mathbf{B}, \varphi') + \frac{1}{\overline{\mu}_s} \sum_{s'} P_{ss'}\chi(s', k + k_{born}, \mathbf{B}) & \text{if } \varphi = 0 \end{cases} \tag{16}$$

where $\mathbf{B} \subset \mathbb{R}_+$ is measurable and the notation 1_X denotes the indicator function for a given set X, (b) the representative firm maximizes profits $F(K, L_s) - (r + \delta)K - \sum w_s L_s$, (c) the government budget constraint

$$G + rB \leq \int_{(s,k,a,z,\varphi)} T_k(\Theta_k(z_k)w_s)d\mu + \tau^b \int_{(s,k,b)} bd\chi$$

is satisfied, (d) the goods market clears, $Y = \delta K + \int \mathbf{c}_{k,s}(a, z, \varphi)d\mu$, (e) all markets for efficiency units of each skill clear, $L_s = \overline{\mu}_s$, (f) the asset market clears,

$$\frac{1}{1+r}A \equiv \frac{1}{1+r}\int_{(s,k,a,z,\varphi)} a\, d\mu = B + K,$$

(g) the bequest distribution is consistent with the distribution over states, $\chi(s, k, (1 - \tau^b)\mathbf{A}, z, \varphi) = \delta_k \mu(s, k, \mathbf{A}, z, \varphi)$, where $\mathbf{A} \subset \mathbb{R}_+$ measurable, and (h) aggregate flows and bequests are consistent

$$\mu(s, k+1, \mathbf{A}, z', \tilde{\varphi}) = \sum_{\varphi} \int_{(b',\varphi') \text{ s.t. } \varphi'=\tilde{\varphi}} \int_{(s,k,a,z,\varphi) \text{ s.t. } \mathbf{a}_{k,s}(a,z,\varphi)+b' \in \mathbf{A}} \Pi_{zz'} d\mu dv(\cdot|s, k, \varphi)$$

$$\mu(s, 1, \mathbf{A}, z, \varphi) = \pi_z \overline{\mu}_s 1_{\{0\}}(\varphi)1_{\{0\}}(\mathbf{A}).$$

C.2 Technical assumption for linearity

Assumption 4 (Unique wealth distribution given r)**.** *Given any interest rate $r > 0$ and permanent incomes $\{w_s\}$, there exists at most a single wealth distribution μ for which (a) and (g) of Definition 1 can be* *satisfied.*

This assumption essentially rules out the special case of no income risk and an infinite horizon, where it is well known that there does not exist a unique wealth distribution. Note that it still allows for multiple steady-state equilibria to exist (as in Acikgöz 2017) as long as each equilibrium interest rate is associated with a unique wealth distribution.

C.3 A useful lemma

I start by showing a useful auxiliary result which states that all steady state policy functions and asset distributions scale with permanent income.

Lemma 1. *Under Assumptions 1–4, for any measurable set $(s, k, \mathbf{A}, z, \varphi) \subset \mathbb{S}$, any state $(s, k, a, z, \varphi) \in \mathbb{S}$ and any skill $s' \in \mathcal{S}$ it holds in any equilibrium that*

$$\mu(s, k, \mathbf{A}, z, \varphi) = \frac{\overline{\mu}_s}{\overline{\mu}_{s'}} \times \mu\left(s', k, \mathbf{A}\frac{w_{s'}}{w_s}, z, \varphi\right) \quad \textit{and} \quad \chi(s, k, \mathbf{A}) = \frac{\overline{\mu}_s}{\overline{\mu}_{s'}} \times \chi\left(s', k, \mathbf{A}\frac{w_{s'}}{w_s}\right)$$

$$\mathbf{c}_{k,s}(a, z, \varphi) = \frac{w_s}{w_{s'}} \times \mathbf{c}_{k,s'}\left(a\frac{w_{s'}}{w_s}, z, \varphi\right) \quad \textit{and} \quad \mathbf{a}_{k,s}(a, z, \varphi) = \frac{w_s}{w_{s'}} \times \mathbf{a}_{k,s'}\left(a\frac{w_{s'}}{w_s}, z, \varphi\right).$$

Lemma 1 has two crucial implications: distributions (over assets and bequests) and policy functions (for consumption and assets) scale in permanent income. For instance, fix an age k, an income state z, and a bequest indicator φ. Lemma 1 shows that an agent with skill s and asset position a consumes exactly $w_s/w_{s'}$ times as much as an agent with skill s' and asset position $aw_{s'}/w_s$.

Proof. The key idea behind the proof of this result is that the dynamic programming problem (4) is separate for each skill s, by Assumption 3. Take two skills s, s'. I will argue by contradiction that if the scaling properties in Lemma 1 do not hold for s and s', one can construct an alternative wealth distribution $\tilde{\mu}$, contradicting Assumption 4.

Assume that the scaling property of the bequest distribution χ in Lemma 1 does not hold, that is, there exist two skills s, s', an age k_0, and a measurable set $\mathbf{A}_0 \subset \mathbb{R}_+$ so that

$$\chi(s, k_0, \mathbf{A}_0) \neq \frac{\overline{\mu}_s}{\overline{\mu}_{s'}} \times \chi\left(s', k_0, \mathbf{A}_0\frac{w_{s'}}{w_s}\right).$$

Fix s, s' and define a new bequest distribution $\tilde{\chi}$ that is equal to χ for all skills except s', where I define

$$\tilde{\chi}(s', k, \mathbf{A}) \equiv \frac{\overline{\mu}_{s'}}{\overline{\mu}_s}\chi\left(s, k, \mathbf{A}\frac{w_s}{w_{s'}}\right)$$

for any age k and measurable set $\mathbf{A} \subset \mathbb{R}_+$. Similarly, define the anticipated bequest distribution $\tilde{v}$ that corresponds to bequest distribution $\tilde{\chi}$ as in (16). Notice that for any measurable set $\mathbf{A} \subset \mathbb{R}_+$ it holds that

$$\tilde{v}(\mathbf{A}, \varphi'|s', k, \varphi) = \tilde{v}\left(\frac{w_s}{w_{s'}}\mathbf{A}, \varphi'|s, k, \varphi\right) \tag{18}$$

due to Assumption 3, according to which either the skill transition matrix is the identity, $P = I$, or there are no bequests.

Next, consider the dynamic programming problem (4) with anticipated bequest distribution $\tilde{v}$. Denote the corresponding value function by $\tilde{V}$. It must be that

$$\tilde{V}_{k,s'}(a,z,\varphi) = \left(\frac{w_{s'}}{w_s}\right)^{1-\sigma} \tilde{V}_{k,s}\left(\frac{w_s}{w_{s'}}a,z,\varphi\right) \tag{19}$$

for the following reasons: if $\tilde{V}$ were not the (unique) solution to the convex programming problem (4), one could easily use (19) to construct a second solution. $\tilde{V}$ solves (4) due to Assumptions 1 and 2, the fact that the agent starts with zero assets, and equation (18). Building on (19), it is immediate that the unique policy functions satisfy

$$\tilde{\mathbf{c}}_{k,s'}(a,z,\varphi) = \frac{w_{s'}}{w_s}\tilde{\mathbf{c}}_{k,s}\left(a\frac{w_s}{w_{s'}},z,\varphi\right)$$

and

$$\tilde{\mathbf{a}}_{k,s'}(a,z,\varphi) = \frac{w_{s'}}{w_s}\tilde{\mathbf{a}}_{k,s}\left(a\frac{w_s}{w_{s'}},z,\varphi\right).$$

Finally, constructing $\tilde{\mu}$ as in (17) given $\tilde{v}$ and $\tilde{\mathbf{a}}$ yields

$$\tilde{\mu}(s',k,\mathbf{A},z,\varphi) = \frac{\overline{\mu}_{s'}}{\overline{\mu}_s} \times \mu\left(s,k,\mathbf{A}\frac{w_s}{w_{s'}},z,\varphi\right).$$

This proves that, if the scaling property for the bequest distribution χ does not hold, one can construct a (different) wealth distribution $\tilde{\mu}$ which satisfies (a) and (g) of Definition 1, for the same interest rate r and the same skill prices $\{w_s\}$. This contradicts Assumption 4. Therefore, the scaling property must hold for χ. Following the same steps as above, however, also establishes similar scaling properties of $\chi, v, V, \mathbf{a}, \mathbf{c}$ and μ. This concludes our proof of Lemma 1. $\square$

C.4 MPCs in a linear economy

Define the *MPC out of current income* of an agent in state $(s,k,a,z,\varphi) \in \mathbb{S}$ as

$$\mathbf{MPC}_{k,s}(a,z,\varphi) \equiv \frac{\partial}{\partial a}\mathbf{c}_{k,s}(a,z,\varphi).$$

Under linearity, the distribution of MPCs is the *same* across skill groups.

Corollary 1 (MPCs under linearity). *Under Assumptions 1–4, for any state $(s,k,a,z,\varphi) \in \mathbb{S}$ it holds that*

$$\mathbf{MPC}_{k,s}(a,z,\varphi) = \mathbf{MPC}_{k,s'}\left(\frac{w_{s'}}{w_s}a,z,\varphi\right). \tag{20}$$

In particular, the distribution of MPCs is the same within all skill groups.

Proof. Equation (20) is a direct consequence of the definition of MPCs and Lemma 1,

$$\mathbf{MPC}_{k,s}(a,z,\varphi) = \frac{\partial}{\partial a}\mathbf{c}_{k,s}(a,z,\varphi)$$
$$= \frac{w_s}{w_{s'}}\frac{w_{s'}}{w_s}\frac{\partial}{\partial a}\mathbf{c}_{k,s'}\left(a\frac{w_{s'}}{w_s},z,\varphi\right) = \mathbf{MPC}_{k,s'}\left(\frac{w_{s'}}{w_s}a,z,\varphi\right).$$

To prove the claim on the distribution of MPCs, define first the following conditional distribution

$$\mu(k,a,z,\varphi|s) \equiv \frac{1}{\bar{\mu}_s}\mu(s,k,a,z,\varphi).$$

Notice that by Lemma 1,

$$\mu(k,\mathbf{A},z,\varphi|s) = \mu\left(k,\mathbf{A}\frac{w_{s'}}{w_s},z,\varphi|s'\right). \tag{21}$$

The claim is that the distribution of the random variable $(k,a,z,\varphi) \mapsto \mathbf{MPC}_{k,s}(a,z,\varphi)$ under probability distribution $\mu(k,a,z,\varphi|s)$ is the same for each skill $s \in \mathcal{S}$. This immediately follows from the combination of (20) and (21). □

The result in Corollary 1 is interesting because it documents that without conditioning on cash-on-hand a, the distribution of MPCs is unaffected by permanent income. Aside from potential endogeneity issues, this is at odds with the evidence in Jappelli and Pistaferri (2006, 2014), which documents that MPCs tend to decline in education, in line with the predictions of a non-homothetic model (see appendix F.2).

C.5 Proof of Proposition 1

Equation (5) is directly shown in Lemma 1. It follows that

$$\log \mathbf{c}_{k,s}(aw_s,z,\varphi) - \log w_s = \log \mathbf{c}_{k,s'}(aw_{s'},z,\varphi) - \log w_{s'}. \tag{22}$$

Define the conditional distribution given age k and skill s as

$$\mu(\mathbf{A},z,\varphi|s,k) \equiv \frac{\mu(s,k,\mathbf{A},z,\varphi)}{\mu(s,k,\mathbb{R}_+,\mathcal{Z},\{0,1\})}.$$

As direct consequence of the scaling property of μ,

$$\mu(w_s\mathbf{A},z,\varphi|s,k) = \mu(w_{s'}\mathbf{A},z,\varphi|s',k). \tag{23}$$

Using that notation, define

$$const_k \equiv \int (\log \mathbf{c}_{k,s}(aw_s,z,\varphi) - \log w_s)\, d\mu(\cdot|s,k)$$

which is well-defined by (23). By definition of $const_k$, one can then write

$$\log \mathbf{c}_{k,s}(aw_s, z, \varphi) = const_k + \log w_s + \epsilon(k, a, z, \varphi).$$

Translating the behavior into the associated stochastic process, one can then write

$$\log c_{it} = const_{k(i,t)} + \log w_{s(i)} + \epsilon_{it}$$

where by construction of $\epsilon(k, a, z, \varphi)$, ϵ_{it} has zero mean conditional on age k and skill s.

Equation (7) follows by definition of income as

$$\log y_{k,s}^{pre}(z) = \log \Theta_k(z) + \log w_s$$

where again $\log \Theta_k(z)$ can be further decomposed into an average and a mean zero term. This gives (7).

C.6 Proof of Proposition 2

Using (22), one can decompose

$$\log \mathbf{c}_{k,s}(a, z, \varphi) = \log w_s + \mathcal{C}\left(k, \frac{a}{w_s}, z, \varphi\right). \tag{24}$$

By the law of total variance, the variance of the left hand side under probability measure μ can be computed as

$$\mathrm{Var}_\mu \left[\log \mathbf{c}_{k,s}(a, z, \varphi)\right] = \mathbb{E}_{\overline{\mu}}\left[\mathrm{Var}_{\mu(\cdot|s)}\left[\log \mathbf{c}_{k,s}(a, z, \varphi)\right]\right] + \mathrm{Var}_{\overline{\mu}}\left[\mathbb{E}_{\mu(\cdot|s)}\left[\log \mathbf{c}_{k,s}(a, z, \varphi)\right]\right].$$

The first term does not involve s since the distribution of $\mathcal{C}\left(k, \frac{a}{w_s}, z, \varphi\right)$ under $\mu(\cdot|s)$ is the same for all s. Define

$$const \equiv \mathbb{E}_{\overline{\mu}}\left[\mathrm{Var}_{\mu(\cdot|s)}\left[\log \mathbf{c}_{k,s}(a, z, \varphi)\right]\right].$$

For the same reason, one can write

$$\mathbb{E}_{\mu(\cdot|s)}\left[\log \mathbf{c}_{k,s}(a, z, \varphi)\right] = \log w_s + const'$$

with some other constant $const'$. Together,

$$\mathrm{Var}_\mu \left[\log \mathbf{c}_{k,s}(a, z, \varphi)\right] = const + \mathrm{Var}_{\overline{\mu}}\left[\log w_s\right].$$

A similar argument shows the result for log wealth (cash on hand), concluding our proof of Proposition 2.

C.7 Proof of Proposition 3

Denote by $C(\{w_s\})$ aggregate consumption as function of skill prices $\{w_s\}$. Using the law of iterated expectations, one can express C as

$$C(\{w_s\}) = \mathbb{E}_{\overline{\mu}}\left[\mathbb{E}_{\mu(\cdot|s)}\mathbf{c}_{k,s}(a,z,\varphi)\right].$$

With decomposition (24), this becomes

$$C(\{w_s\}) = \mathbb{E}_{\overline{\mu}}\left[w_s\mathbb{E}_{\mu(\cdot|s)}\exp\mathcal{C}\left(k,\frac{a}{w_s},z,\varphi\right)\right]$$

where again $\mathbb{E}_{\mu(\cdot|s)}\exp\mathcal{C}\left(k,\frac{a}{w_s},z,\varphi\right) \equiv \kappa_C$, which is independent of w_s. Therefore,

$$C(\{w_s\}) = \kappa_C \sum_s \overline{\mu}_s w_s$$

and similarly for total assets as function of skill prices $A(\{w_s\})$, and total bequests $A^{beq} \equiv \int_{(s,k,b)} b d\chi$.

To show that a change in labor income shares $\{\gamma_s\}$ leaves all aggregate quantities unchanged, notice that

$$\overline{\mu}_s w_s = (1-\alpha)\gamma_s Y$$

and so $C = \kappa_C(1-\alpha)Y$ and $A = \kappa_A(1-\alpha)Y$.

These derivations show that neither goods market clearing, nor asset market clearing, nor the government budget constraint are affected by a change in labor income shares $\{\gamma_s\}$. Therefore, there exists an equilibrium where also the aggregate interest rate and the aggregate capital stock stay the same.

D Additional empirical results

D.1 Additional robustness exercises

This section provides additional OLS and IV specifications at the household level, as well as a set of group-level specifications and specifications that use the imputed consumption measure by Blundell et al. (2008).

Comprehensive consumption. As mentioned in Section 4.1 when describing the dataset, the PSID introduced a somewhat more comprehensive consumption measure in 2005 by adding six consumption categories. Since my OLS and AR(1) IV specifications rely on future income data, the consumption data that is effectively used spans the years 1999–2005 and thus barely overlaps with the new measure. To have a meaningful comparison between the two consumption measures, I impute the missing comprehensive measure before 2005 by estimating a set of demand equations

Table 10: More robustness checks.

	OLS with $T = 9$	IV with $\rho = 0.90$	IV with initial income
1. Baseline	0.64	0.70	0.73
	(0.03)	(0.04)	(0.05)
2. Comprehensive consumption	0.66	0.71	0.77
	(0.03)	(0.04)	(0.05)
3. Education as IV		0.73	
		(0.03)	
4. Total post-tax income	0.69	0.71	0.72
	(0.03)	(0.04)	(0.04)
5. Nondurable consumption	0.58	0.59	0.62
	(0.03)	(0.04)	(0.05)
6. Aguiar-Bils relative expenditures	0.62	0.72	0.71
	(0.08)	(0.10)	(0.15)
7. Incl. household heads > 65 years	0.68	0.71	0.73
	(0.03)	(0.04)	(0.04)
8. Controlling for liquid assets	0.63	0.74	0.70
	(0.03)	(0.05)	(0.05)
9. Households with original sample heads	0.61	0.66	0.76
	(0.04)	(0.07)	(0.06)
10. Dropping 10% lowest incomes each year	0.67	0.74	0.74
	(0.03)	(0.04)	(0.05)
11. No location controls	0.68	0.72	0.75
	(0.03)	(0.04)	(0.05)

Note. This table lists OLS and IV estimates for 9 different specifications. Row 1 shows the baseline specifications. Row 2 uses the comprehensive post-2005 consumption measure. Row 3 uses education as instrument for current income residuals. Row 4 uses total post-tax household income as income measure. Row 5 uses non-durable consumption as consumption measure. Row 6 uses data on relative expenditures on luxuries vs necessities to achieve robustness to non-classical measurement error as in Aguiar Bils (2015). Row 7 shows results for households with heads of all ages. Row 8 controls for a cubic in liquid assets relative to income. Row 9 runs the baseline OLS and IV specifications on the subsample with only the households headed by original PSID sample heads. Row 10 drops observations with labor income below 10% of the average per year. Row 11 does not control for location in the regression. All specifications control for year, age, household size and location dummies. All IV specifications have first stage F statistics above 10. Standard errors are corrected for heteroskedasticity and clustered by household.

for each of the new consumption categories on post-2005 data (see Appendix D.3 for details).[52] The second row of Table 10 shows the results. The OLS and the AR(1) IV estimates are close to their baseline counterparts, and the estimate for the initial income IV is somewhat larger.

Education as instrument. The third row of Table 10 shows IV results when education is used as the instrument. This is a common instrument in the previous literature (Dynan et al. 2004; Bozio et al. 2013). The estimate for ϕ is slightly larger than that of my baseline IV specification.

Total post-tax income. My specifications so far used a measure of post-tax labor income, not including capital income. In row 4 of Table 10, I use a total post-tax, post-transfer income measure, which includes capital income. The results are similar.

Nondurable consumption. Row 5 of Table 10 shows results for nondurable consumption only (defined as the sum of food expenditure, rent, property taxes, home insurance expenditure, utilities, transportation, education, childcare and health-related expenditures). The somewhat lower estimates suggest that including durable consumption is important since permanent-income richer households seem to spend a disproportionate amount on those expenditures.

Non-classical measurement error. In a recent paper, Aguiar and Bils (2015) argue that there is non-classical measurement error in the interview survey of the Consumer Expenditure Survey (CEX), in the sense that richer respondents are more likely to underreport their spending. Aguiar and Bils (2015) show that this matters for the measurement of consumption inequality. There are two reasons why this may be less of a concern for this paper: Compared to the CEX, there is no evidence of a downward trend in the coverage of the PSID consumption expenditure data (Blundell et al. 2016); and, in the present study, the estimate for ϕ is only affected if consumption is relatively more underreported than income.

Nevertheless, I use several strategies to deal with this issue. First, I present estimates using three different consumption measures: the post-1999 PSID measure, the comprehensive post-2005 measure, and the BPP imputed measure (for details, see Appendix D.1). Second, I re-estimated the three main specifications using the PSID's post-1999 wealth data, finding a permanent income elasticity of wealth of approximately $2-3$ (depending on the exact specification and measure of wealth). This is significantly greater than 1 (neutrality) and is in line with the analogous estimate of 2 in the non-homothetic model of Section 5. And finally, I applied the method proposed by Aguiar and Bils (2015) to this paper. Using their estimates for expenditure elasticities, I back out total expenditure measures along the income distribution from the ratios of luxury expenditures to expenditures on necessities (for details, see Appendix D.2). The resulting estimates are shown in row 6 of Table 10 and are comparable with the baseline estimates.

Other household-level specifications. Row 7 shows the baseline specification, only on the larger sample of households whose head is between 30 and 80 years old, rather than between 30 and 65 years old.[53] The results are similar as in the baseline case. Row 8 adds a cubic polynomial in the ratio of liquid assets to income as controls to $\mathbf{X}_{it}$. I measure liquid assets as the sum of a household's

[52]See Blundell et al. (2008) or Boar (2018) for recent examples that employ similar imputation strategies.

[53]Since labor income is small for most households above 65, I use total household income, including income from capital for this regression.

Table 11: Group-level specifications.

	log group consumption	
	(1) OLS	(2) IV
log group income	0.642	0.720
	(0.04)	(0.05)
Year FE	Yes	Yes
Age FE	Yes	Yes
Observations	256	104
R-squared	0.73	0.79

Note. This table lists OLS and IV specifications at the group level. Groups are constructed using the interaction of 5-year birth cohorts and education dummies. Column 1 shows simple OLS estimates of log group consumption on log group income residuals. Column 2 shows IV estimates using future income as instrument. Standard errors are corrected for heteroskedasticity.

wealth in cash and stocks, net of credit card, medical, legal and other debts as well as net of loans from relatives. The estimates are somewhat larger but insignificantly different. Row 9 re-estimates the baseline specifications on the subsample of households whose heads are original PSID sample heads, rather households headed by heads that joined the PSID by way of marriage. Hryshko and Manovskii (2017) argue that these groups of households differ in terms of their income shocks. The results here show that they do not differ significantly in terms of the concavity of their consumption function. Row 10 drops observations with labor income below 10% of the average per year. Row 11 does not control for location in the regression. Both imply similar results.

Group-level specifications. In Table 11, I show estimates for two simple group-level specifications. Following the seminal work of Attanasio and Davis (1996), I define groups as the interaction of 5-year birth cohorts and education groups of household heads (no high school, high school, less than four years of college, four years of college). For columns 1–2, I use a group-averaged version of (9). Specifically, let $\{I_g\}$ be the set of the mutually exclusive groups of households and define for each group their average log consumption, $\bar{c}_{gt} \equiv |I_g|^{-1} \sum_{i \in I_g} \hat{c}_{it}$, and their average income residuals, $\bar{y}_{gt} \equiv |I_g|^{-1} \sum_{i \in I_g} \hat{y}_{it}$. I estimate

$$\bar{c}_{gt} = \mathbf{X}'_{gt}\beta + \phi \bar{y}_{gt} + \epsilon_{gt}. \tag{25}$$

The IV approach in column 2 uses 2-year ahead group income $\bar{y}_{gt+1}$ as instrument for $\bar{y}_{gt}$. The results are close to the ones found in the household-level approaches in Section 4.

Imputed consumption measure by Blundell et al. (2008). In an important paper, Blundell et al. (2008) estimate a demand equation for food consumption expenditure in data from the CEX and use it to impute various measures of consumption expenditure for the PSID from 1980 to 1992.[54] Among the measures of consumption expenditure they impute is total consumption expenditure, which is the measure I am using below. The results and conclusions are very similar with the other measures the

[54] I thank the authors for making their data available online.

Table 12: OLS and IV estimates using the imputed consumption data from Blundell et al. (2008).

	OLS		IV					
	(1) $T = 1$	(2) $T = 9$	(3) $\rho = 0.88$	(4) $\rho = 0.89$	(5) $\rho = 0.9$	(6) $\rho = 0.91$	(7) $\rho = 0.92$	(8) ini.
log Income	0.555	0.645	0.569	0.566	0.562	0.557	0.551	0.784
	(0.06)	(0.03)	(0.06)	(0.06)	(0.06)	(0.07)	(0.07)	(0.13)
Year FE, Age FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hh.size FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	703	2050	8072	8072	8072	8072	8072	13244
1st stage F			39.0	33.6	28.5	23.7	19.4	77.6
R-squared	0.38	0.60	0.76	0.76	0.76	0.76	0.76	0.79

Note. This Table lists OLS and IV estimates when using the imputed consumption measure from Blundell et al. (2008). Columns 1 and 2 show OLS estimates for the case of no income averaging ($T = 1$) and averaging over $T = 9$ income observations. Columns 3-7 show the estimates from IV specifications with various assumed income shock persistences ρ. Column 8 shows estimates with age 25 income as instrument. Standard errors are clustered by household and corrected for heteroskedasticity.

authors impute.

Their consumption data is helpful in the context of this paper, for at least three reasons: first, the data is from an entirely different source, namely the CEX, not the PSID, which I have been using for my analysis. Second, similar to the comprehensive post-2005 consumption measure in the PSID, the imputed consumption data includes total expenditure across all categories. And third, the imputed data is from an entirely different time—from 1980 to 1992—whereas my analysis had to be restricted to the time after 1999 due to data limitations.

In Table 12 I reestimate my main OLS and IV specifications. First, while the estimates are somewhat more noisy, it is reassuring that they all lie around 0.7, thus confirming my previous results. In fact, the estimates in columns 1-7 of Table 12 are generally around $0.05 - 0.15$ *lower* than their counterparts using the PSID consumption data after 1999 (Tables 1 and 2). The upper bound estimate using the initial income IV is around 0.05 larger than its post-1999 counterpart and considerably noisier.

Additional issues. *Learning.* Intuitively, it seems plausible that households entering their working lives might be imperfectly informed about their permanent income, having to learn it over time (Muth 1960; Pischke 1995). Could this explain my findings? A simple back-of-the-envelope calculation using the permanent-transitory income process in Kaplan and Violante (2010) shows that learning about permanent income is quick: agents learn around 95% of the variation in permanent income after just 5 years in the labor market.[55] This suggests that while (Bayesian) learning might play a role for the curvature of the consumption function in the first couple of years, it is unlikely to

[55] This also resonates with the result found by Guvenen (2007) that learning about permanent income, α^i in his notation, is very quick.

remain important thereafter.

Partial insurance. One may wonder how the presence of partial insurance arrangements affect my results. First, notice that many insurance channels, such as government tax and transfers as well as total family labor supply, are already taken into account in the post-tax income measure that I use. Still, there may be informal insurance among households that I cannot observe. In a recent paper, Guvenen and Smith (2014) present an explicit model of partial insurance, where an agent hit by an innovation ϵ_{it}^{η} to the persistent shock η_{it} receives an unobserved transfer $-\theta\epsilon_{it}^{\eta}$ that mitigates the shock, $\theta \in (0,1)$. This does not affect the two IV strategies, since neither relies on any specifics of the income shock other than the persistence parameter.

Misspecification of ρ. A similar logic carries over to the case where the econometrician specifies the wrong ρ in the construction of the AR(1) IV. For instance, if the persistence parameter chosen by the econometrician ρ is too large relative to the true persistence, say ρ^*, the IV estimate $\hat{\phi}^{IV}$ tends to be biased upwards. In that case, the ρ-difference $\hat{y}_{it+\tau} - \rho\hat{y}_{i,t+\tau-1}$ contains the term $-\left(\rho - (\rho^*)\right)\eta_{it+\tau-1}$, which is positively correlated with the error term ϵ_{it}, as agents with high income shocks η_{it} tend to save more and consume less out of current income. The bias goes in the opposite direction if $\rho > \rho^*$. This explains why the estimates in Table 2 tend to be smaller for smaller ρ's.

D.2 Robustness to non-classical measurement error

In a recent paper, Aguiar and Bils (2015) argue that there is non-classical measurement error in the interview survey of the Consumer Expenditure Survey (CEX), in the sense that richer respondents are more likely to underreport their spending. Aguiar and Bils (2015) show that this matters for the measurement of consumption inequality. In this paper, the estimate for ϕ can be affected by such measurement error whenever high-income households underreport their spending relatively more compared to low-income households, and relatively more compared to their income. To test this concern, I propose an estimation strategy similar to the one proposed in Aguiar and Bils (2015).

Denote by c_{igt} the observed expenditure of household i on good g at time t, and by c_{it} its total observed expenditure across all goods. Denote by c_{igt}^* and c_{it}^* the true expenditures. As before, assume that true total expenditures are governed by

$$\log c_{it}^* = \phi \log w_i + \mathbf{X}_{it}\beta + \epsilon_{it}.$$

In addition, I follow Aguiar and Bils (2015) in assuming: that true good-level expenditures can be represented by a log-linear demand system,[56]

$$\log c_{igt}^* = \gamma_g \log c_{it}^* + \tilde{\mathbf{X}}_{it}\tilde{\beta}_g + \epsilon_{igt},$$

where time fixed effects are part of $\tilde{\mathbf{X}}_{it}$; and that measurement error at the level of goods, $\log c_{igt} - \log c_{igt}^*$, is orthogonal to true total expenditures $\log c_{it}^*$. In other words, when high-income house-

[56] This is not possible globally, see the discussion in Aguiar and Bils (2015).

holds underreport their spending, they do so proportionally across categories. In the following, fix some good g_0 (it is irrelevant which good).

Under these three assumptions, one obtains that

$$\frac{\log c_{igt} - \log c_{ig_0t}}{\gamma_g - \gamma_{g_0}} = \phi \log w_i + \mathbf{X}_{it}\beta + \tilde{\mathbf{X}}_{it}\breve{\beta}_g + \breve{\epsilon}_{igt}$$

with $\breve{\beta}_g = (\beta_g - \beta_{g_0}) \,/\, (\gamma_g - \gamma_{g_0})$ and $\breve{\epsilon}_{igt} = (\epsilon_{igt} - \epsilon_{ig_0t}) \,/\, (\gamma_g - \gamma_{g_0})$. To implement this regression, I use the estimated expenditure elasticities in (Aguiar and Bils, 2015, Table 2, column I) for the following categories which are (roughly) comparable across the CEX and the PSID: housing, food at home, transportation (incl. other transportation), utilities, health expenditures, food away from home, education, child care. As controls $\tilde{\mathbf{X}}_{it}$, I use time dummies, 5-year age brackets and household size dummies. The results are in row 9 of Table 10.

D.3 Imputation procedure for the comprehensive consumption measure

I use the following approach to impute the comprehensive consumption measure to years before survey year 2005 (calendar year 2004). Denote by c_{it} household i's total consumption in year t according to the reduced 70% measure, and by $\mathbf{c}_{it}$ household i's consumption according to the comprehensive post-2005 measure. Label by $k = 1, \ldots, 6$ the six new consumption categories introduced in 2005, and by c_{it}^k household i's consumption in those categories. Thus, for all years $t \geq 2005$,

$$\mathbf{c}_{it} = c_{it} + \sum_{k=1}^{6} c_{it}^k.$$

I impute the value of c_{it}^k for every k separately. In particular, for a given k, I model c_{it}^k as

$$\log c_{it}^k = \mathbf{P}_t \eta^k + \gamma^k \log c_{it} + \mathbf{X}_{it}\beta^k + \epsilon_{it}^k$$

where $\mathbf{X}_{it}$ includes a variety of household controls, such as a cubic polynomial in age, and dummies for 5-year cohort bracket, household size, race, and sex; $\mathbf{P}_t$ is a vector of log prices, one for each consumption category (including those already in c_{it}). To avoid measurement error in $\log c_{it}$, I instrument $\log c_{it}$ using future consumption $\log c_{it+2}$. Having estimated $\eta_t^k, \gamma^k, \beta^k$ I compute predictions $\tilde{c}_{it}^k$ for each household i, consumption category k and time $t < 2005$. The imputed comprehensive measure of consumption expenditure is then

$$\tilde{\mathbf{c}}_{it} \equiv \begin{cases} c_{it} + \sum_{k=1}^{6} \tilde{c}_{it}^k & t < 2005 \\ \mathbf{c}_{it} & t \geq 2005 \end{cases}.$$

D.4 Estimation of the income process

In this section, I explain the estimation strategy I use to estimate the income process used in this paper. Throughout, I measure "income" as pre-tax labor income, as measured in the PSID. As sample, I use all years in the PSID from survey year 1981, the first year without significant income top-coding, to survey year 1997, which is the last year for which there is annual income information. I include all households between ages 30 and 55, to avoid issues due to early retirement, and exclude households with pre-tax labor income below 25% of any given year's average pre-tax labor income (for the survey year 2013, this threshold amounts to a household income of approximately \$15k). To avoid small sample bias, the sample is restricted to households for whom there are at least 15 not necessarily consecutive years of income information. This leaves me with a sample of 2,817 households. In line with the models, I estimate an income process for total household labor income (per household member).

Estimating the income process. Consider the following standard model for log incomes $\log y_{it}$,

$$\log y_{it} = f(\mathbf{X}_{it}, \beta) + \alpha_i + \eta_{it} + \psi_{it} + \nu_{it},$$

where $f(\mathbf{X}_{it}, \beta)$ are a set of controls, α_i are income fixed effects, η_{it} is an AR(1) process whose persistence is ρ_η and whose innovations have variance σ_η^2 , ψ_{it} is a transitory income shock with variance σ_ψ^2, and ν_{it} is a measurement error term with variance σ_ν^2. I take the controls $f(\mathbf{X}_{it}, \beta)$ to be a cubic polynomial in age, dummies for household size, and year dummies. Measurement error cannot be distinguished from the transitory income shock. I follow Heathcote et al. (2010) and assume that $\sigma_\nu^2 = 0.02$. This leaves me with four parameters to estimate: σ_α^2, ρ_η, σ_η^2, and σ_ψ^2.

I employ a stationary minimum distance estimation (MDE) procedure that is standard by now and was first developed in Chamberlain (1984).[57] First in the procedure, I residualize incomes by partialing out the demographic and life-cycle controls $f(\mathbf{X}_{it}, \beta)$. Denote the income residual for individual i at age k by $\hat{y}_{ik}$, where the lowest age is normalized to $k = 1$, so that the maximum age is $K = 26$. The autocovariances of $\hat{y}_{ik}$ are then given by

$$Cov(\hat{y}_{ik}, \hat{y}_{ik+s}) = \sigma_\alpha^2 + \sigma_\eta^2 \sum_{j=0}^{k-1} \rho^{2j} + 1_{\{s>0\}} \left(\sigma_\psi^2 + \sigma_\nu^2 \right),$$

for any $s \geq 0$. As mentioned before, σ_ν^2 is set exogenously. I use 15 time periods for estimation, so s ranges between 0 and 14. Implicit in this formulation is the assumption that the initial variance of the persistent income process is also given by σ_η^2. Denote the right hand side by $g_{k,k+s}(\sigma_\alpha^2, \rho, \sigma_\eta, \sigma_\psi^2)$, and symmetrically set $g_{k+s,k}$. Define the empirical covariance matrix of income by

$$G_{k,k'} = \frac{1}{|\mathcal{I}_{k,k+s}|} \sum_{i \in \mathcal{I}_{k,k+s}} \hat{y}_{i,k} \hat{y}_{i,k'}, \quad k, k' \in \{1, \dots, K\}.$$

[57] See Meghir and Pistaferri (2011) for a recent survey over the literature.

Table 13: Estimated AR(1) + iid process.

ρ	σ^2_α	σ^2_η	σ^2_ψ
0.904	0.131	0.026	0.052
(0.064)	(0.017)	(0.011)	(0.015)

Note. This table shows the estimated parameters of a process for log income residuals. The process consists of a permanent component, an AR(1) component and a transitory shock. Standard errors are block-bootstrapped with 500 iterations and clustered at the household level.

The MDE estimator minimizes the distance between g and G. To implement it, stack all $K(K+1)/2$ unique values of $G_{k,k'} - g_{k,k'}$ into a vector, denoted by $\mathbf{G}(\sigma^2_\alpha, \rho, \sigma_\eta, \sigma^2_\psi)$. The minimum distance estimates of $(\sigma^2_\alpha, \rho, \sigma_\eta, \sigma^2_\psi)$ are then the solution to

$$\min_{\sigma^2_\alpha, \rho, \sigma_\eta, \sigma^2_\psi} \mathbf{G}(\sigma^2_\alpha, \rho, \sigma_\eta, \sigma^2_\psi)' W \mathbf{G}(\sigma^2_\alpha, \rho, \sigma_\eta, \sigma^2_\psi)$$

where W is a weighting matrix. I use an identity weighting matrix, which was shown to be less prone to small sample bias by the simulations in Altonji and Segal (1996). This procedure yields consistent and asymptotically normal estimates for $(\sigma^2_\alpha, \rho, \sigma_\eta, \sigma^2_\psi)$, whose asymptotic standard errors I compute using a block-bootstrap with 500 iterations that is clustered at the household level. Table 13 shows the results.

Alternative way to estimate the persistence. The most relevant parameter in this estimation is the persistence ρ. In the same stationary setup as the one used above, one can identify ρ directly as follows. Define $m_s \equiv \mathbb{E}[\hat{y}_{ik}\hat{y}_{ik+s}]$ where the expectation includes k, and $s > 0$. Then, for any $s > 0$,

$$\log(m_{s+1} - m_s) = const + s \log \rho. \tag{26}$$

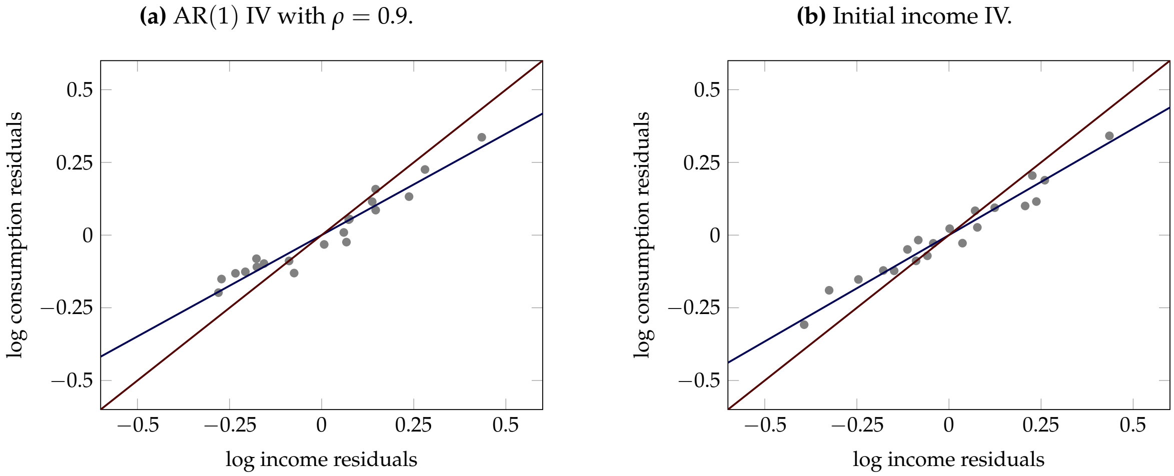
I used this simple linear relationship to confirm that in a variety of other settings and samples, the persistence parameter ρ estimated this way lies around 0.90 (or even lower).

Needless to say, there is an active debate about the "right" income process. The simple estimation strategy in (26) can be viewed as being focused entirely on estimating the ρ that best matches the *relative* magnitudes of the moments m_s. As is well-known and as I point out in Appendix D.6, the downside of matching the relative magnitudes of moments m_s is that an estimated ρ around 0.90 fails to generate the almost linear shape of the income dispersion over the life-cycle. It is worth emphasizing, that if one would like to match income dispersions over the life cycle, one can equally well calibrate a non-homothetic model like the one in Section 5 with a simple random-walk income process, using the initial-income IV estimates. The results are available upon request.

D.5 Scatter plots for the IV specifications

To show that there is also no particular source of nonlinearity that drives the two IV results, this section shows the non-parametric results for the AR(1) IV specification ($\rho = 0.9$), as well as the

Figure 12: Non-parametric estimates for the IV second stages.



Note. This figure shows the results of a regression of log consumption on controls and predicted permanent income bins, which are constructed using a first-stage regression of income on controls and the instrument(s). Panel (a) does this for the AR(1) IV with persistence parameter $\rho = 0.9$ and panel (b) does this for the initial income IV.

initial income IV specification.

To this end, I estimate a set of predicted permanent incomes $\hat{w}_{it}^{p}$ by projecting current income residuals $\hat{y}_{it}$ on the controls $\mathbf{X}_{it}$ in the consumption equation as well as the set of instruments used—$\{z_{i,t+\tau}\}_{\tau>1}$ for the AR(1) IV and z_i for the initial income IV.[58] I then run a non-parametric regression of log consumption $\hat{c}_{it}$ on controls $\mathbf{X}_{it}$ and dummies for 20 bins of the predicted permanent incomes $\hat{w}_{it}^{p}$. The results for the two IV specifications are shown in Figure 12. No obvious nonlinearity or outliers drive the result.

D.6 Evidence on the mechanism

One may wonder what happens with the additionally saved funds of permanently richer households. This section provides evidence for two channels: future consumption and intergenerational transfers to children.

Future consumption. To investigate whether permanent-income richer households increase their future consumption relatively more than permanently poorer households, I re-run the IV estimation in Section 4.3 by 5-year age groups. Since the initial income IV specification does not require any future labor incomes, one can re-estimate it for every 5-year age group from 30 to 65 years. The orange triangles in Figure 13(a) shows the results. There is a significant increase in the elasticity, with later ages consuming relatively more.

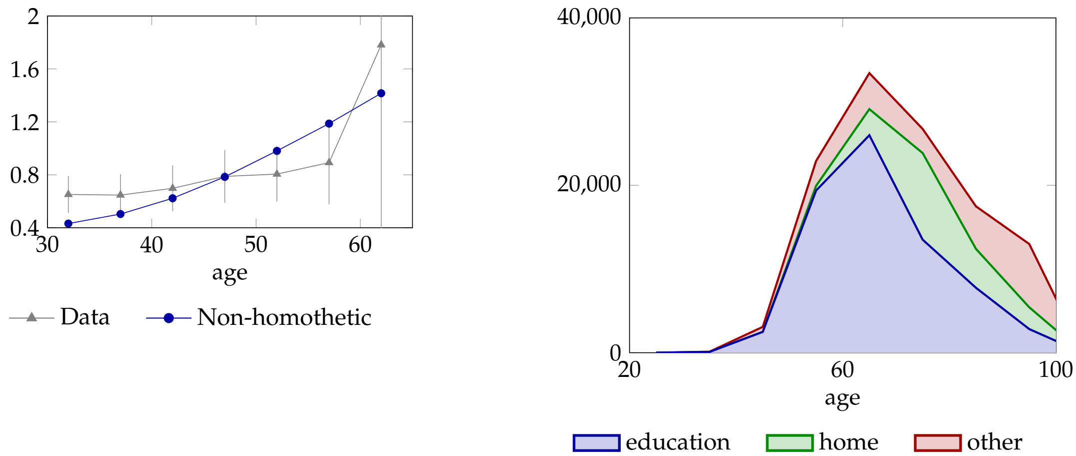
Intergenerational transfers. In 2013, the PSID surveyed households about the size and type of intergenerational transfers they received from their parents, or made to their children. Using log

[58]See the definitions in Section 4.3.

Figure 13: Evidence on the mechanism: Why do rich people save more?

(a) Future consumption: $\hat{\phi}^{IV}$ by age.

(b) Transfers to children, by type and age. (USD)



Note. Panel (a) shows the estimated permanent income elasticity of consumption using the initial income instrument, by 5-year age groups (blue squares, with grey error bars). In black are the estimates from the quantitative model of Section 5. Panel (b) shows transfers to children, averaged by age, for three types of transfers.

total transfers made as left-hand-side variable instead of $\hat{c}_{it}$ in the initial income IV regressions one finds that transfers (those made during one's work-life) have a permanent income elasticity of around 1.8 and therefore are a key reason why permanently richer agents save. At what ages are transfers given, and what are they used for? Figure 13(b) shows the average annual transfer given by age, and for three types of transfers: school-related transfers, such as college tuition payments, house purchase related transfers, and other types of transfers. It is clearly visible that the transfers are sizable and happen late in life (and therefore need to be saved towards).

Comparison with the model. The blue circles in Panel (a) show the results when replicating the initial IV regressions by 5-year age bins on simulated data from the quantitative non-homothetic model of Section 5. It can be seen that in the model, the regressions predict a greater slope in age than what seems to be in the data. Part of that could be explained by the presence of other expenses (such as the ones in Panel (b)), which are not part of consumption expenditure in the data, but are treated as such by the model.

E Computational Appendix

In this section I explain the methods that were used to simulate the models in Section 5. I start by laying out the definition and computation of the steady state of the non-homothetic model in Section E.1. Then, Section E.2 provides details on the regression analyses on model-simulated data that are shown in Table 6.

E.1 Simulation of the model of Section 5.1

I first set up the household maximization problem, then I explain how I solve it. I allow for non-stationary environments so as to nest the transitional dynamics exercise of Section 6. Throughout, I focus on utility functions $u_k(c)$ over consumption; solving a model with status in the utility function is similar.

E.1.1 Household maximization

Given parameters and given an interest rate path $\{r_t\}$, a household solves the following dynamic programming problem,

$$V_{k,s,t}(a, z, \varphi) = \max_{\{c,a'\}} u_k(c) + \beta(1 - \delta_k)\mathbb{E}_{z,\varphi} V_{k+1,s,t+1}(a' + b', z', \varphi') + \beta\delta_k U(a') \tag{27}$$

$$c + \frac{1}{1 + r_t} a' \leq a + y_{k,s,t}(z)$$

$$(b', \varphi') \sim v_t(\cdot|s, k, \varphi)$$

$$a' \geq 0.$$

Here, $V_{k,s,t}(a, z, \varphi)$ is the agent's value function at time t, $w_{s,t}$ is the agent's skill price at time t, $v_t(\cdot|s, k, \varphi)$ is the endogenous distribution of bequests at time t, b' is the random bequest which is received at the parent's death, and the agent's post-tax, post-transfer income is given by

$$y_{k,s,t}(z) = \max \left\{ \underline{y}, \Theta_k(z) w_{s,t} - T_t^{inctax}(\Theta_k(z) w_{s,t}) \right\}$$

before retirement, $k \leq K_{ret}$, and by

$$y_{k,s,t}(z) = \max \left\{ \underline{y}, T^{socsec}\left(\overline{y}_{s,t}(z), W_t\right) - T_t^{inctax}\left(T^{socsec}(\overline{y}_{s,t}(z), W_t)\right) \right\}$$

after retirement; $\overline{y}_{s,t}(z)$ is the predicted average pre-tax income, conditional on ending up in state z when moving into retirement,

$$\overline{y}_{s,t}(z) = \frac{1}{K_{ret}} \sum_{k=1}^{K_{ret}} \mathbb{E}\left[\Theta_k(z_k) w_{s,t} | z_{K_{ret}} = z\right].$$

Government. The government levies a time-dependent nonlinear income tax schedule,

$$T_t^{inctax}(y^{pre}) = y^{pre} - \tau_t^{inctax}\left(y^{pre}\right)^{1-\lambda_t},$$

where, given a path for tax progressivity $\{\lambda_t\}$, $\{\tau_t^{inctax}\}$ is chosen to yield the same aggregate tax income in each period.

Production. The production function is allowed to be time-dependent,

$$Y_t = F(t, K_t, \{L_s\}_{s\in\mathcal{S}}) = AK_t^\alpha \prod_s (L_s/\overline{\mu}_s)^{(1-\alpha)\gamma_{s,t}},$$

where I normalized the skill endowments by group to 1 since in L_s is inelastically supplied at $\overline{\mu}_s$. I assume that the representative firm maximizes profits,

$$J_t(K_-) = \max_{d,I} \left\{ d + \frac{1}{1+r_t} J_{t+1}(K_- - \delta K_- + I) \right\} \tag{28}$$

$$d = F_t(K_-, \{L_s\}) - \sum_s w_{s,t} L_s - I$$

The skill prices are given by

$$w_{s,t} = \partial F(t, K_t, \{L_s\})/\partial L_s = (1-\alpha)\gamma_{s,t} Y_t$$

and the average wage is given by $W_t = \sum_s \overline{\mu}_s w_{s,t}$.

E.1.2 Definition of equilibrium

Definition 2. A *competitive equilibrium* consists of a path of aggregate quantities $\{Y_t, K_t, I_t\}$, paths of distributions $\{\mu_t\}$ of agents and $\{\chi_t\}$ of bequests, both defined over the state space $\mathbb{S}$, paths for policy functions $\{\mathbf{c}_{k,s,t}(a,z,\varphi), \mathbf{a}_{k,s,t}(a,z,\varphi)\}$, paths of prices $\{r_t, w_{s,t}\}$ such that: (a) each agent solves the optimization problem (27) given $\{r_t, w_{s,t}\}$, where the conditional bequest distribution $v_t(\cdot|s,k,\varphi)$ is given by

$$v_t(\mathbf{B}, \varphi'|s,k,\varphi) = \begin{cases} 1_{\{0,1\}}(\mathbf{B}, \varphi') & \text{if } \varphi = 1 \\ (1-\delta_{k+k_{born}}) 1_{\{0,0\}}(\mathbf{B}, \varphi') + \sum_{s'} P_{ss'} \chi_t(s', k+k_{born}, \mathbf{B}) & \text{if } \varphi = 0 \end{cases},$$

(b) the representative firm sets $\{K_t, I_t\}$ to solve the profit-maximization problem (28) given $\{r_t, w_{s,t}\}$, (c) the government sets government spending G_t according to its budget constraint

$$G_t = \int T_t^{inctax}(\Theta_k(z_k) w_s) d\mu_t(s,k,a,z,\varphi) + \tau^b \int b dv_t(s,k,b) - r_t B,$$

$$G_t = \int_{(s,k,a,z,\varphi)} T_k(\Theta_k(z_k) w_{s,t}) d\mu_t + \tau^b \int_{(s,k,a,z,\varphi)} b d\chi_t - r_t B$$

(d) the goods market clears,

$$Y_t = I_t + \int \mathbf{c}_{k,s,t}(a,z,\varphi) d\mu_t,$$

(e) all markets for efficiency units of each skill clear, $L_s = \overline{\mu}_s$, (f) the asset market clears,

$$A_t \equiv \int a\, d\mu_t = (1+r_t)B + J_t(K_t),$$

(g) the bequest distribution is consistent with the distribution over states, $\chi_t(s,k,\mathbf{A},z,\varphi) = \delta_k \mu_t(s,k,\mathbf{A},z,\varphi)$, where $\mathbf{A} \subset \mathbb{R}_+$ measurable, and (h) aggregate flows and bequests are consistent

$$\mu_{t+1}(s,k+1,\mathbf{A},z',\tilde{\varphi}) = \sum_{\varphi} \int_{(b',\varphi') \text{ s.t. } \varphi'=\tilde{\varphi}} \int_{(s,k,a,z,\varphi) \text{ s.t. } \mathbf{a}_{k,s,t}(a,z,\varphi)+b' \in \mathbf{A}} \Pi_{zz'} d\mu_t d\upsilon_t(\cdot|s,k,\varphi).$$

E.1.3 Computing the household's optimal decisions

I use a version of the method of endogenous grid points, modified to allow for the receipt of bequests. Since this method is fairly standard, I do not explain the basics and instead refer the interested reader to the materials in Carroll (2005).

At age T, the household solves a simple maximization problem between consumption and bequests, which I solve to find the consumption policy $\mathbf{c}_{K_{death},s,T}(a,z,\varphi)$. I then iterate backwards using the Euler equation and the method of endogenous grid points,

$$\hat{\mathbf{c}}_{k,s,t}(a',z,\varphi)^{-\sigma} = \beta(1+r_t) \sum_{z'} \Pi_{zz'} \int_{(b',\varphi')} \left((1-\delta_k) \mathbf{c}_{k+1,s,t+1}(a+b',z,\varphi')^{-\sigma} + \delta_k U'(a') \right) d\upsilon_t(b',\varphi'|s,k,\varphi)$$

where the new policy function is then computed by inverting the budget constraint,

$$\mathbf{c}_{k,s,t}(a',z,\varphi) + \frac{1}{1+r_t} a' = a + y_{k,s,t}(z)$$

to obtain the asset policy function $a' = \mathbf{a}_{k,s,t}(a,z,\varphi)$, and the consumption policy function

$$\mathbf{c}_{k,s,t}(a,z,\varphi) = a + y_{k,s,t}(z) - \frac{1}{1+r_t} \mathbf{a}_{k,s,t}(a,z,\varphi).$$

After solving for the consumption and savings policy functions at all ages and times, I iterate forward the savings policy functions to compute the sequence of distributions across all idiosyncratic states $\{\mu_t(s,k,a,z,\varphi)\}$. Finally, I compute the endogenous bequest distributions $\{\upsilon_t(b',\varphi'|s,k,\varphi)\}_t$. I iterate over this entire process until the endogenous bequest distributions have converged.

In this type of framework with stochastic bequests, a key time factor in the simulation is the convolution of a given bequest distribution with a given asset distribution. To ease this issue, I implemented my own non-uniform fast Fourier transform (FFT) algorithm that, different from conventional FFT algorithms still allows me to work on non-uniform grids, as long as those grids are piecewise uniform. This modification sped up the steady state computation considerably.

E.1.4 Solving for the general equilibrium steady state

In my steady state analysis, I implement this method with 151 asset states. I discretized the persistent part of the income process using the Rouwenhorst (1995) method on 11 states,[59] and the transitory shock on 3 states. In additional to 65 age states, 3 skill groups, and a number of

[59] The Tauchen (1986) method is generally less precise than the Rouwenhorst (1995) method for persistent income processes (Kopecky and Suen 2010).

inheritance states to keep track whether a household has made an inheritance already, this leaves me with 8712 idiosyncratic states, and $\approx$ 1.3 million asset–idiosyncratic state pairs. My implementation of the above algorithm allows me to solve the economy given a bequest distribution (and given an interest rate r) in approximately 20 seconds. It takes 60 seconds to iterate until the bequest distribution converges (tolerance 10^{-6}). All times are measured on a 2009 MacBook Pro. The code was implemented in Matlab and C.

To solve for the general equilibrium steady state in this economy, I compute the household's maximization problem, and aggregate the agents' wealth levels to get total asset demand A_{demand}. The total asset supply A_{supply} in the economy splits into two parts: government bonds B and financial wealth in equities $v = (F_K K - \delta K)/r$. I use Matlab's *fzero* command to solve for the equilibrium interest rate r that equalizes A_{demand} and A_{supply}.

E.2 Estimation of regressions on model-simulated data

In Table 6 I show regressions that were performed on simulated data from the model. To implement those regressions, I closely follow my empirical analyses in Section 4. I construct the same measure of post-tax labor income,

$$y_{it}^{posttax} = y_{it}^{pre} - \frac{y_{it}^{pre}}{y_{it}} T^{inctax}(y_{it}^{pre}),$$

that is, I subtract from pre-tax labor income the share of taxes accounted for by labor income, rather than transfers $\underline{y}$. In practice, $y^{pre} = y$ for almost all agents, except when an agent's income is below the minimum income threshold, below which the agents receives income transfers. The sample of agents I focus on is the same as in the data, namely agents between ages 30 and 65. The data I use to run specifications is a panel of 75,000 agents whose income draws are determined by Monte-Carlo simulations and whose behavior is given by the consumption policy functions implied by the model.[60] After simulating the data, all income observations are multiplied by the measurement error term $\exp\{\nu_{it}\}$, and then residualized by partialing out age effects.

I implement the following specifications:

OLS specifications. Average income residuals are computed as symmetric averages of T observations, spaced out over $2T - 1$ years to mimic the biennial nature of the PSID sample I use.

AR(1) IV specifications. As in Section 4, I use quasi-differenced future incomes as instruments, and I include all observations with at least three such instruments. Again, I pretend the panel was biennial, as is the relevant subsample of the PSID that I use.

Initial income IV specification. The initial income IV specification is constructed using initial income at age 25 as instrument. For comparability for the previous two specifications, I also restrict the consumption data to be between years 30 and 57 (as I do in the data).

[60] I also experimented with many iterations of smaller samples to ensure there are no small sample biases.

Table 14: The age profile of elasticities.

Age	σ_k	Age	σ_k	Age	σ_k	Age	σ_k
25	11.6273	45	3.7291	65	1.6437	85	1.2626
26	10.9839	46	3.5326	66	1.6009	86	1.2626
27	10.3761	47	3.351	67	1.5614	87	1.2626
28	9.8019	48	3.1832	68	1.525	88	1.2626
29	9.2595	49	3.028	69	1.4915	89	1.2626
30	8.7471	50	2.8844	70	1.4607		
31	8.263	51	2.7514	71	1.4326		
32	7.8057	52	2.6281	72	1.407		
33	7.3738	53	2.5139	73	1.3837		
34	6.9657	54	2.408	74	1.3628		
35	6.5803	55	2.3098	75	1.344		
36	6.2161	56	2.2186	76	1.3273		
37	5.8721	57	2.134	77	1.3126		
38	5.5472	58	2.0555	78	1.2999		
39	5.2402	59	1.9826	79	1.2891		
40	4.9502	60	1.915	80	1.2802		
41	4.6763	61	1.8522	81	1.2731		
42	4.4175	62	1.794	82	1.2678		
43	4.173	63	1.74	83	1.2643		
44	3.9421	64	1.69	84	1.2626		

Note. This table shows the age profile of elasticities σ_k.

E.3 The schedule of elasticities σ_k

Table 14 shows the age-dependent elasticities σ_k for each age k.

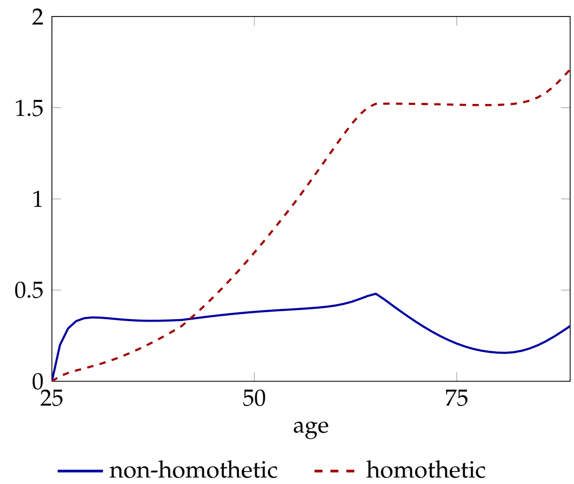
F Additional steady state results

F.1 Risk aversion over the life cycle

To elucidate the implications of non-homothetic preferences for risk aversion, and how it changes over the life cycle, I compute for each age the average curvature of the value function. Let $\overline{\mu}_k(s, a, z, \varphi)$ be the distribution over (s, a, z, φ) conditional on age k. Then, define the average risk aversion at age k as

$$\Sigma_k \equiv - \int \frac{a \partial_{aa} V_{k,s}(a, z, \varphi)}{\partial_a V_{k,s}(a, z, \varphi)} d\overline{\mu}_k(s, a, z, \varphi).$$

Figure 14: Risk aversion over the life cycle.



Note. The figure shows the age profiles of risk aversion, measured as average curvature in the value function, in the homothetic model (red, dashed) and the non-homothetic model (blue, solid).

Figure 14 shows the path of risk aversions Σ_k for the homothetic model and the non-homothetic model.

As can be seen, risk aversion in the homothetic economy increases considerably over the life cycle. If risky assets where available in this economy, this would suggest that young agents have by far the largest shares of risky assets in their portfolios, while older agents hold mostly risk-less assets. The intuition for this is generally that labor income can be viewed as "bond" with a limited amount of risk, which young agents are well endowed with, much more than older agents.

As shown in Ameriks and Zeldes (2004), this is not the case in the data. Younger agents tend to own roughly equally safe portfolios as their older agents. Interestingly, this is exactly what the non-homothetic model generates: there, Σ_k increases only mildly with age, rationalizing why portfolios do not become more risky over the life-cycle.[61]

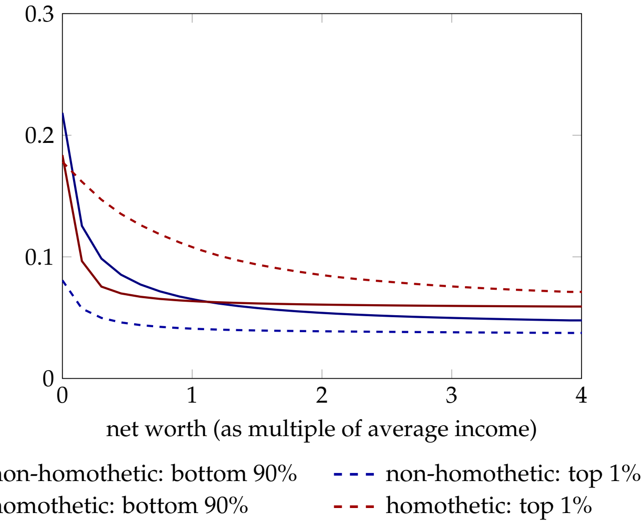
F.2 The distribution of MPCs

As Section C.4 illustrates, in a neutral model, the distribution of MPCs is the same within different skill groups. In the data, it seems to be the case that MPCs actually (unconditionally) decline in measures of permanent income, such as education (Jappelli and Pistaferri 2006, 2014).

Interestingly, this is precisely what a non-homothetic economy implies. Figure 15 shows MPCs at the median income state (when $\eta = \psi = 0$), averaged over all asset states and ages, for the

[61]See Wachter and Yogo (2010) for an important contribution that makes this point in a non-homothetic life-cycle economy.

Figure 15: Average MPCs by skill at the median income state.



Note. The figure shows the MPC at the median income state, averaged over all asset states and ages, for the bottom 90% (solid) and the top 1% (dashed). The colors represent the homothetic economy (red) and the non-homothetic economy (blue).

bottom 90% (solid) and the top 1% (dashed). The colors represent the homothetic economy (red) and the non-homothetic economy (blue).

In the homothetic economy, MPCs of the top 1% lie *above* those of the bottom 90% conditional on assets, and would be exactly equal unconditionally, since the top 1% tend to have larger asset positions. In the non-homothetic economy, MPCs of the bottom 90% are higher than their homothetic counterparts, for small asset positions, while the MPCs of the top 1% are lower.

This illustrates that non-homothetic preferences increase the spread in the distribution of MPCs, and can rationalize why high permanent income agents may have lower MPCs (both unconditionally or conditional on assets).

F.3 Lorenz curves

The Lorenz curves for pre-tax incomes, consumption and wealth are shown and compared to the data in Figure 16, panel (a). To construct the Lorenz curve for pre-tax income in the data, I add back the estimated cubic age profile to the residualized incomes of the 2011 and 2013 waves of the PSID. This gives me a distribution of incomes that is comparable across years and households of various sizes but does not strip out the age efficiency profile. The curves align fairly closely with pre-tax income in the model.

The data for the empirical Lorenz curve for wealth is taken from World Top Income Database for the U.S. in 2014 (Saez and Zucman 2016). Wealth is here measured as net personal wealth, that is assets (including housing) net of debts, capitalizing capital income data from income tax returns. In the model, I use the wealth distribution, i.e. the distribution of a_{it}'s, of all households (of all ages). Here, as stressed in Section 5.4, the model is relatively successful in matching the overall distribution of wealth, especially compared to the homothetic model, which is shown as the dotted red line.

The data for the Lorenz curve for consumption is computed in the same way as the Lorenz curve for income, from the PSID. Strikingly, it shows a much smaller difference between the non-homothetic and homothetic economies. The non-homothetic model does, however, predict larger consumption inequality at the top end compared to the data. One reason for this could be that significant later-in-life expenses (such as the transfers in Figure 13) are counted as expenditure in the model, but as transfers in the data.

F.4 Life cycle profiles

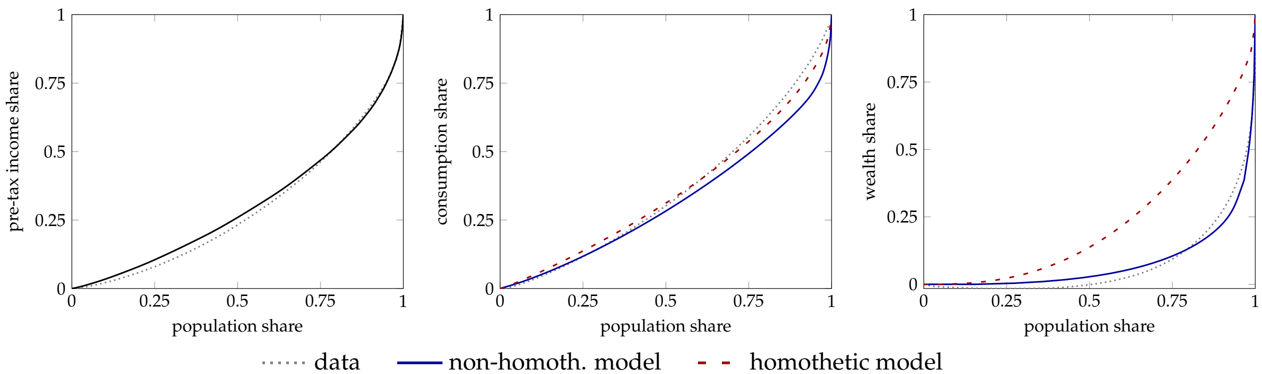
Panels (b) and (c) of Figure 16 show the life cycle profiles of post-tax income, consumption and wealth in the two models, for the 25th, 50th, 75th and 90th percentiles. While the income profiles are the same during the working life, the retirement system in the homothetic case is a simple linear rule for social security. Comparing the life cycle profiles of consumption, the key difference is that consumption profiles are generally higher for the lower percentiles (25th and 50th) and shifted more towards higher ages for the higher percentiles (75th and 90th). This behavior is at the heart of the non-homothetic life cycle model: richer agents tend to save more out of their income, which limits their consumption expenditure early on and increases both consumption expenditure and bequests later in life.

The consequences for wealth accumulation and within-cohort wealth dispersion are especially striking: The top wealth percentiles increase much more rapidly during the work life, and fall less rapidly during retirement. Again, this is the product of two key model ingredients. First, due to the non-homotheticity in preferences, the wealth of the rich increases more rapidly during the work-life. Second, due to the non-homotheticity in bequests, this additional wealth is dis-saved much more slowly, which feeds back into larger bequests for the children of rich parents. Since disproportionately many of these children are high-skilled themselves, this again implies a steeper increase in wealth during the work-life of an average high-skilled agent.

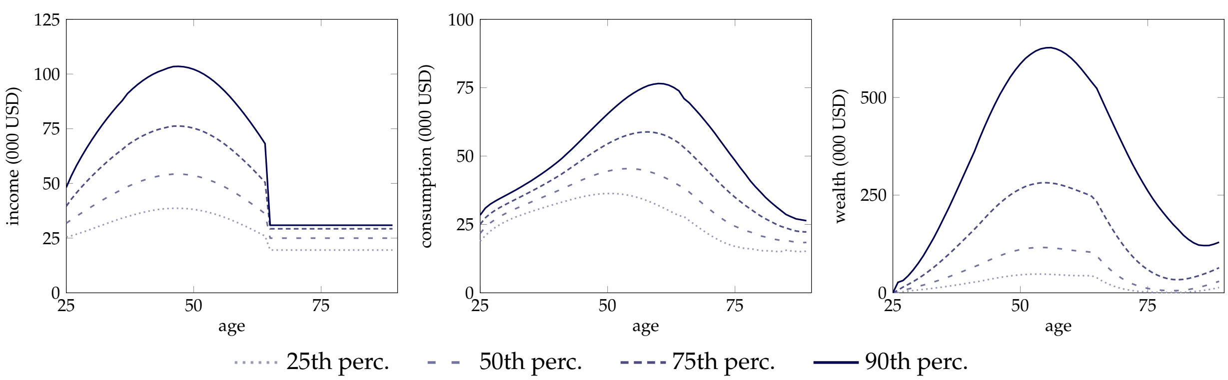
Construction of the life cycle profiles. I explain the computation of the life cycle plots in Figure 16 by percentiles using the age profile of wealth as an example. For every age k, I compute the asset distribution of agents at that age. Call its quantile function $Q_k(p)$. Since the distribution is discrete, I approximate $Q_k(p)$ for each age k locally around a given percentile p_0 with a log-normal distribution, $Q_k^{lognormal}(p)$. Since I do not consider the very top percentiles, the log-normal distribution fits very well. For a given percentile p_0, the plot shows the age profile of the fitted log-normal percentiles $Q_k^{lognormal}(p_0)$. The dollar figures were computed using an average household income of $\$70,000$ (USD in 2014).

Figure 16: Key life-cycle characteristics of the non-homothetic and homothetic models.

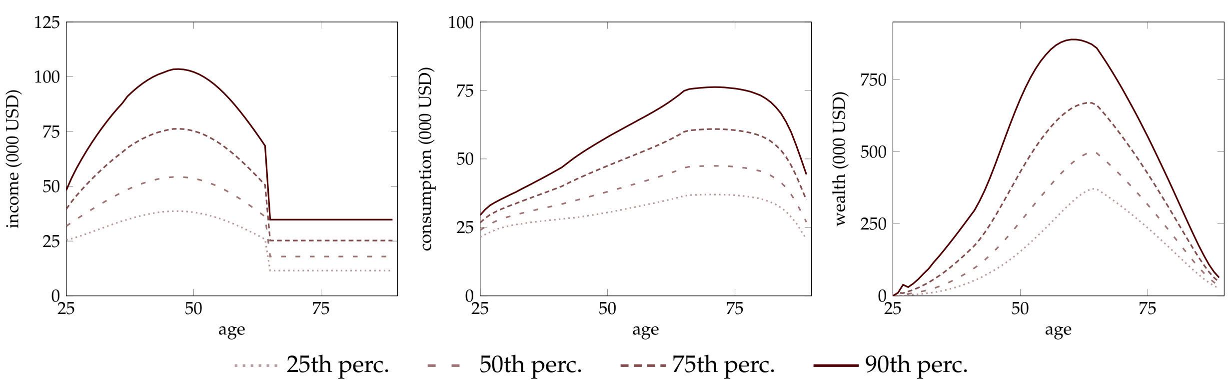
(a) Lorenz curves for pre-tax income, consumption, and wealth.



(b) Non-homothetic σ economy: Life-cycle profiles of pre-tax income, consumption, and wealth.

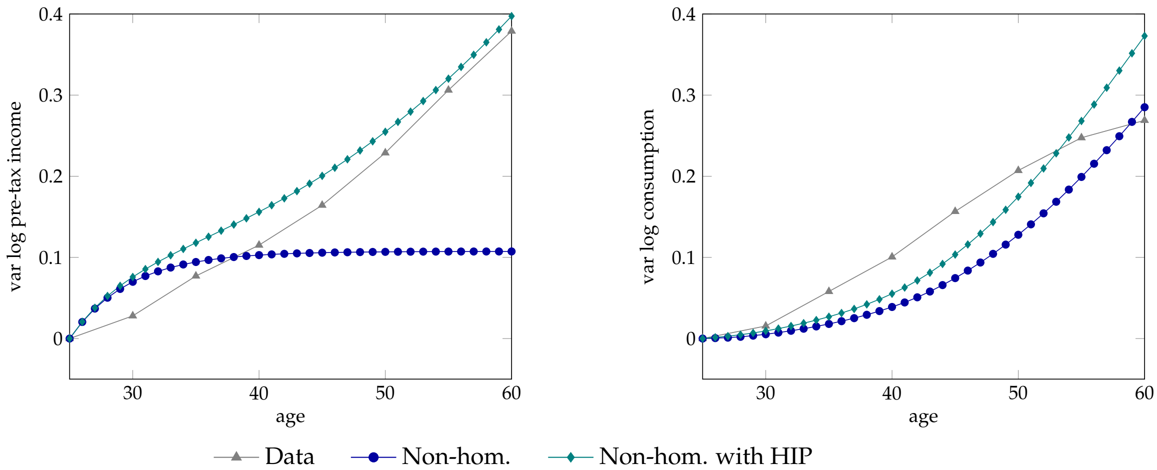


(c) Homothetic economy: Life-cycle profiles of pre-tax income, consumption, and wealth.



Note. This figure shows key characteristics of two economies: the non-homothetic economy, where rich households save disproportionately more than poor households in relation to their incomes, and a standard homothetic economy, where the saving rates are equal. See Appendix F for details.

Figure 17: Age profiles of income and consumption dispersion.



Note. This figure shows the income and consumption dispersion over the work-life (normalized to 0 at age 25) for two models: the baseline non-homothetic model and an extension with heterogeneous income profiles that is designed to match the increase in income dispersion in the data. The data on the income dispersion comes from the PSID and on the consumption dispersion comes from the CEX. Both were constructed using cohort fixed effects following Deaton and Paxson (1994).

F.5 Life cycle dispersion

While the plots in Figure 16 are instructive to illustrate the non-homotheticity, there is a second more standard way to illustrate the degree to which the dispersion in consumption and income increases over the work-life. This way of looking at life-cycle consumption and savings behavior goes back to Deaton and Paxson (1994) who used it to think about the degree to which agents have access to informal insurance arrangements. More recently, age profiles of dispersion in income and consumption have been examined by Storesletten, Telmer and Yaron (2004), Guvenen (2007) and Huggett, Ventura and Yaron (2011), among many others.

There are two common ways to match both the rise in income dispersion and consumption dispersion: either a very persistent $AR(1)$ income process, see e.g. Storesletten et al. (2004); or an income process with less persistence but ex-ante heterogeneous income profiles (HIP), which, together with learning about the slope, allows to match both life-cycle dispersion plots (Guvenen 2007).

The baseline non-homothetic model introduced in Section 5 falls in neither of those two categories. The income process has a persistence parameter between these two categories, but without heterogeneity in income profiles. I now explore the dispersion profiles in the baseline non-homothetic economy, as well as in an extension to that model with heterogeneity in income profiles (modeled as in Guvenen 2007, 2009), to better capture the increase in dispersion over the life cycle.[62]

Figure 17 shows the age profiles of income and consumption dispersions in both the baseline non-

[62] The addition of heterogeneity in income profiles slightly pushes up the estimates of the permanent income elasticity of consumption shown in Section 5.4, by around 0.03. (The HIP model could then be re-calibrated to match the empirical estimate for ϕ.)

homothetic economy and the HIP extension. It can be seen that the baseline economy predicts too small an increase in the income dispersion and about the right increase in consumption dispersion. Also, the consumption dispersion is convex in age, rather than approximately linear, which is likely driven by the simplistic assumption that σ_k falls exponentially in age. The HIP extension improves the fit for income dispersions, since by construction, it matches the overall increase in income dispersion. It also leads to greater consumption dispersion, but not by a lot.

Data construction. The income dispersion profile was computed using PSID waves from survey year 1969 to survey year 2013. I follow the construction in Huggett et al. (2011). Income is pre-tax labor income, and the sample consists of all male-headed households whose real income (in 1968 USD) is between $\$1,500$ and $\$1.5m$, and who supplied between 520 and 5820 hours of work in the respective year. The consumption dispersion profile was computed using data from the CEX that was downloaded from Fabrizio Perri's website (see the data appendix to Krueger and Perri 2006). I use total consumption expenditure and exclude all observations with consumption expenditure below 5% of the average.

G Alternative models

G.1 Non-homothetic status model

In this section, I consider an alternative non-homothetic model, one building on the idea that rich individuals might derive utility from having a certain status in the economy. I define status $x = a/A$ as one's wealth position a relative to average wealth A in the economy. The working-life utility function in this model is given by

$$u(c, x) = \frac{c^{1-\sigma}}{1-\sigma} + U(x)$$

where $x = a/A$ is one's status, and $U(x) = \kappa \frac{(x+\underline{x})^{1-\sigma}}{1-\sigma}$ is the utility over status. If the intercept $\underline{x}$ were zero, this would be a homothetic model (everyone cares equally about status), so $\underline{x} > 0$ allows the model to be non-homothetic.[63] There is no bequest utility.

I jointly calibrate $\beta, \kappa, \underline{x}$ to match the same interest rate and the same permanent income elasticity of consumption. Moreover, to make the two non-homothetic models more or less comparable, I choose as third target the top 1% wealth share of the non-homothetic model. This gives $\beta = 0.94$, $\kappa = 4700$ and $\underline{x} = 15$. As can be seen in Table 9, the non-homothetic status model predicts very similar effects from rising income inequality.

G.2 Background on the models used for comparison in Section 5.5

Table 7 compares the non-homothetic and homothetic life cycle models to a variety of other life cycle and infinite horizon models. Here, I explain each model and how it was calibrated.

[63] Notice that the model is naturally balanced-growth compatible if $\sigma = 1$ (log preferences), even if $\underline{x} > 0$.

Explanation. *Alternative preferences or transfers.* The first model in group 2 is an entirely neutral economy, with $\phi = 1.00$ and serves as the framework for further extensions. The second model in group 2 extends this model by including the concave actual social security schedule T^{socsec}, which modestly reduces the IV estimates of ϕ by around 0.05. Similarly, when only a luxury bequest motive is added (i.e. $\sigma_k = const$, and $\kappa, \underline{b}$ calibrated as before), the IV estimates of ϕ shrink by around 0.05.[64]

Alternative income processes. Group 3 in Table 7 considers alternative income processes: an AR(1) process with $\rho = 0.95$; a random-walk income process (calibrated as in Kaplan and Violante 2010); a heavy-tailed income process (adapted from Guvenen et al. 2019); an "extreme productivity state" income process (in the spirit of Castaneda et al. 2003 and adapted from Kindermann and Krueger 2017); a heterogeneous-income profile (HIP) income process as in Guvenen (2009). For all models except the last, the AR(1) IV estimate exhibits a downward bias of up to 0.1 while the initial income IVs are consistently upward biased. The last model generates an upward bias.

Partial insurance economy. The first model in group 4 is a partial insurance economy, where an innovation ϵ^{η}_{it} to η_{it} is mitigated by a transfer $-\theta\epsilon^{\eta}_{it}$ (Guvenen and Smith 2014). In line with my reasoning in Section 4.4, the AR(1) IV estimate is equal to 1. In addition, even the OLS specification with $T = 4$ barely changes compared to the homothetic benchmark economy. This underscores that the issue of partial insurance against income shocks is largely separate from the concavity in consumption as a function of permanent income.

Discount factor shocks. In the last model in Table 7, I simulate an economy with discount factor shocks. In particular, I adapt the parametrization of these shocks from Hubmer et al. (2016). While this technique is extremely useful in matching wealth inequality, it is not well-suited to match the estimated concavity in the consumption function. Indeed, the IV estimate is at 0.99, very close to 1. Why is this? While there are very persistent savers in this economy, the discount factor shocks are assumed to be entirely independent of one's income or consumption choices. This is where my model departs: by modeling non-homotheticity in consumption decisions, this essentially induces an "effective" discount factor that depends positively on consumption. Thus, non-homotheticity positively aligns income and savings decisions a lot more than a random-β environment would.

Calibration. Generally, my goal was to use parameters as standard as possible, and as consistent across models as possible. After parameter choices were made, the discount factor was calibrated in all models to yield the same post-tax equilibrium interest rate of 3%. This is important to avoid, for instance, that agents are constantly against their borrowing constraint (e.g. if β is too low relative to r), which would make finding $\phi = 1$ unsurprising. All model parameters are listed in Table 15. All alternative models are extensions of the homothetic framework that was introduced at the end of Section (5.3). In addition, to ensure that the model is perfectly neutral, I assume the government

[64] This is not to say this motive is not important, however. Indeed, it generates significant improvements in matching wealth inequality. However, it appears to do so mostly by slowing old-age dissaving, rather than creating saving rate dispersion among younger or middle-aged agents.

Table 15: Calibrated parameters the comparison models.

	Re-calibrated parameters		
	β	κ	$\underline{a}$
1. Main models			
Non-homothetic	0.89	15.8	1.72
Homothetic	0.991	1.47	0
2. Alternative preferences or transfers			
Homothetic w/out bequests	0.998		
Homothetic w/ social security	0.995	1.28	0
Homothetic u but luxury bequests	0.994	2680	107
3. Alternative income process			
AR(1) with $\rho = 0.95$	0.999		
Permanent-transitory	1		
Heavy-tailed	0.992		
Extreme productivity state	0.978		
Heterogeneous income profiles	0.999		
4. Other			
Partial insurance	1		
Random discount factors	0.956		

levies a 100% tax rate on bequests.[65] This entirely neutral framework is the model underlying the first model in group 2 in Table 7.

Alternative preferences or transfers. In the second model in group 2, the government uses a realistic social security schedule $T^{socsec}(y, W)$, given by the one used in Section 5.3. The third model in group 2 extends the framework by a non-homothetic bequest motive, and calibrates κ and $\underline{a}$ to match both the total bequest flow (relative to GDP) in the economy, as well as the share of agents with bequests below 6.25% of average income.

Alternative income processes. The first model in group 3 is an extension of the neutral framework where the persistence of the income process is changed from $\rho = 0.90$ to $\rho = 0.95$, keeping the stationary variance $\sigma_\eta^2/(1-\rho^2)$ of the persistence component the same.

The second model is an economy with a permanent-transitory income shock, that is, with $\rho = 1$. I use the parameters in Kaplan and Violante (2010), which are $\sigma_\eta^2 = 0.01$, $\sigma_\psi^2 = 0.05 - \sigma_\nu^2$, $\sigma_\alpha^2 = 0.15$.

The third model is a model with heavy-tailed income shocks. Here, I incorporate some elements of the income process in Guvenen et al. (2019), and assume that the income process is given by

$$\hat{y}_{it} = \eta_{it}^1 + \eta_{it}^2 + \psi_{it},$$

[65] One could have also introduced a perfect market for annuities to ensure neutrality. I expect the results to be very similar.

where η^l_{it} is an AR(1) for $l = 1, 2$, with innovations that are mixed normals,

$$\eta^l_{it} - \rho_l \eta^l_{it-1} \sim \begin{cases} -p^l \mu^l & \text{with prob. } 1 - p^l \\ \mathcal{N}((1-p^l)\mu^l, (\sigma^l)^2) & \text{with prob. } p^l \end{cases}.$$

The initial standard deviations of η^l_{it} are given by σ^l_{ini}. I take the parameters $\rho^l, \sigma^l, \mu^l, \sigma^l_{ini}$ directly from Guvenen et al. (2019), Column 3 of Table II. Since I do not incorporate the entire (more complicated) income process in Guvenen et al. (2019), I rescale the innovation $\eta^l_{it} - \rho_l \eta^l_{it-1}$ to match the increase in the life cycle variance of 0.6 in Guvenen et al. (2019). This implies $\sigma^{l=2} = 0.20$ and $\mu^{l=2} = -0.24$. Finally, I compute p^l as an average over the age-dependent probabilities in Table III in Guvenen et al. (2019).

The fourth model is a simple adaption of the extreme-productivity state in Kindermann and Krueger (2017). In particular, I assume that: (a) there is no transitory component, $\sigma^2_\psi = 0$; (b) the persistent component η_{it} is given by a 7-state Markov chain, with transition matrix

$$\Pi = \begin{bmatrix} 0.957899 & 0.028954 & 0.000328 & 0.000002 & 0 & 0.012817 & 0 \\ 0.007239 & 0.958063 & 0.021717 & 0.000164 & 0 & 0.012817 & 0 \\ 0.000055 & 0.014478 & 0.958118 & 0.014478 & 0.000055 & 0.012817 & 0 \\ 0 & 0.000164 & 0.021717 & 0.958063 & 0.007239 & 0.012817 & 0 \\ 0 & 0.000002 & 0.000328 & 0.028954 & 0.957899 & 0.012817 & 0 \\ 0 & 0 & 0.028087 & 0 & 0 & 0.969688 & 0.002225 \\ 0 & 0 & 0 & 0 & 0 & 0.267852 & 0.732148 \end{bmatrix}$$

and productivity levels

$$\eta = \begin{bmatrix} 0.2112 & 0.4595 & 1 & 2.1761 & 4.7353 & 7.3949 & 1284.3139 \end{bmatrix}$$

that corresponds to the transition matrix of educated agents in Kindermann and Krueger (2017).[66] In particular, there is an extreme productivity state (state 7) in which agents earn a large multiple of the earnings in all other states. The persistent component α_i has 4 states, $\pm\sigma \pm p$ where $\sigma^2 = 0.1517$ and p (the college premium) is equal to $\log(1.8)$ Krueger and Ludwig (2016).

For the fifth model in group 5, the income process is changed to allow for heterogeneity in ex-ante known income profiles,

$$\hat{y}_{ik} = \alpha_i + \beta_i(k-1) + \eta_{it} + \psi_{it}$$

where $k \geq 1$ is the agent's age, and β_i is drawn from a normal distribution with variance σ^2_β. I assume $\rho = 0.82$ as in Guvenen (2009) and $\sigma^2_\beta = 0.00012$ to match the life cycle increase in the variance of income of 0.3.

Other models. The first model in group 4 is a partial insurance economy, where, following

[66]The results are similar if the transition matrix for uneducated agents is being used here.

Table 16: Wealth inequality in the heterogeneous-β model.

	Share of wealth in top				
	1%	5%	10%	20%	50%
Data (Saez-Zucman, 2016)	38.6%	60.6%	73.0%	86.4%	100.1%
Non-homothetic model	38.0%	67.1%	77.7%	86.7%	97.2%
Heterogeneous β model	53.9%	82.1%	90.8%	94.6%	99.0%

Guvenen and Smith (2014) agents partially insure themselves against innovations to η_{it}, so that actual income $\log y_{it}^{actual}$ differs from observed income $\log y_{it}$ and is given by

$$\log y_{it}^{actual} = \log y_{it} + \theta \left\{ \mathbb{E}_{t-1} \left[\log y_{it} \right] - \log y_{it} \right\}.$$

For example, if an agent experiences an unexpected decline in income, the term in curly brackets is positive. Thus, if, say, $\theta = 0.5$, that agent would receive a transfer that dampens the impact of this shock by 50%. θ is estimated by Guvenen and Smith (2014) to be equal to 0.451 and this is the parameter I use in my simulation.

The second model in group 4 is a random discount factor economy. Random discount factors are commonly used in infinite horizon economies (interpretable as dynastic economies), so I consider them in an infinite horizon economy as well. This means, I assume that there is no death risk, $\delta_k = 0$ for all k, no life cycle earnings structure, $\Theta_k(z) = 1$ for all k, and no retirement, $K_{ret} = K_{death} = \infty$ (and therefore also no social security). I introduce shocks to discount factors by allowing the discount factor between periods t and $t+1$, denoted by β_t, to evolve according to the AR(1) process

$$\beta_t = \rho^\beta \beta_{t-1} + (1 - \rho^\beta)\mu^\beta + \sigma^\beta \epsilon_t^\beta,$$

where I pick $\rho^\beta = 0.992$ and $\sigma^\beta = 0.0019$ as in the recent paper by Hubmer et al. (2016). To estimate the regression on an infinite horizon economy, I simulate the model from some time t onwards, and pretend that all agents had age $k = 1$ at that time.

G.3 Heterogeneous-β model

An alternative model that is consistent with the evidence in Section 4 is a model with persistent differences across skill groups that affect savings behavior. The simplest such difference is one in discount factors β. To see whether this could indeed rationalize the evidence and what its implications would be, I simulate a version of the homothetic model in Section 5.1 which assumes that discount factors for the three skills are allowed to be different, $\beta_s, s = 1 \ldots 3$.[67] To calibrate the three β's, I assume that $\beta_3 - \beta_2 = \beta_2 - \beta_1$. This leaves me with two degrees of freedom, the

[67] It is important to keep in mind that typical models in the tradition of Krusell and Smith (1997) model the process for discount factors as being independent of incomes, the opposite of the skill-specific discount factor assumption here.

Table 17: Changes in income inequality in the heterogeneous-β model.

	Implied ϕ	Change from 1970 to 2014 levels of income inequality (*pp.*)		
		Wealth / GDP	Interest rate	Top 10% wealth share
Non-homothetic model	0.70	89	-2.17	24
Non-homothetic status model	0.70	80	-1.89	24
Heterogeneous β model	0.70	88	-2.13	3

average β as well as the gap between discount factors of neighboring skill groups. The two degrees of freedom are pinned down matching the equilibrium interest rate r as well as the estimate of $\phi = 0.70$. This yields an average discount factor of 0.91 and a gap of 0.16. Thus, while the model is able to match the ϕ estimate, clearly a significant amount of dispersion in discount factors is necessary. I next discuss two implications of the heterogeneous-β model.

Wealth inequality. Due to the large dispersion in discount factors necessary to match the data, a significant amount of wealth inequality is generated—in fact, too much relative to the data (Table 16).

Effects of income inequality. Table 17 shows the effects from a change in steady-state income inequality, from its 1970 levels to 2014 levels. As can be seen, the effects on private wealth (over GDP) and interest rates are very similar to the ones in the two non-homothetic models. The effect on wealth inequality, however, is significantly smaller. This is likely related to the fact that the heterogeneous-β model already vastly over-predicts wealth inequality.

G.4 Growth in the non-homothetic model

In this section, I sketch a simple extension of the non-homothetic model with balanced growth. (The non-homothetic status model can easily be made compatible with balanced growth, for instance by assuming $\sigma = 1$.) I distinguish two different sources of growth.

Population growth. Assume that every agent has $1 + g > 1$ children, and that any bequests are shared equally among them. Let $L_\ell \equiv L_0(1 + g)^\ell$ be the total population size of the cohort born at $t = \ell$, where without loss L_0 is assumed to be such that the total population were equal to 1 if $g = 0$. Any member of cohort $t = \ell$ solves the standard Bellman equation (4), except with a smaller bequest $(1 + g)^{-1}b'$, that is, the Bellman equation becomes

$$V_{k,s}(a, z, \varphi, \ell) = \max\ u_k(c) + \beta(1 - \delta_k)\mathbb{E}_{z,\varphi}V_{k+1,s}(a' + (1 + g)^{-1}b', z', \varphi', \ell) + \beta\delta_k U(a')$$

with the same constraints as in (4).

As the size of each cohort grows with L_ℓ, so will the density of agents in each state (k, s, a, z, φ). Since each such agent behaves exactly in the same way, one can collectively describe their total consumption $\mathbf{c}$, savings $\mathbf{a}$, and bequests $\mathbf{b}$ as the solution to

$$V_{k,s}(\mathbf{a}, z, \varphi, \ell) = \max\ u_k\left(\frac{\mathbf{c}}{L_\ell}\right) + \beta(1-\delta_k)\mathbb{E}_{z,\varphi}V_{k+1,s}\left(\mathbf{a}' + \mathbf{b}', z', \varphi', \ell\right) + \beta\delta_k U\left(\frac{\mathbf{a}'}{L_\ell}\right)$$

with budget constraint

$$\mathbf{c} + \frac{1}{1+r}\mathbf{a}' \leq \mathbf{a} + L_\ell \Theta_k(z) w_s - L_\ell T_k(\Theta_k(z) w_s).$$

To map this model into the framework in Section 3, one can normalize the economy by the size of the total population at time t, $\mathbf{L}_t = \mathbf{L}_0(1+g)^t$, and express agents' choices in normalized terms, that is, $\hat{c} \equiv \mathbf{c}/\mathbf{L}_t$, $\hat{a} \equiv \mathbf{a}/\mathbf{L}_t$, $\hat{b} \equiv \mathbf{b}/\mathbf{L}_t$. This yields a Bellman equation

$$V_{k,s}(\hat{a}, z, \varphi) = \max\ u_k\left(\frac{\hat{c}}{o_k}\right) + \beta(1-\delta_k)\mathbb{E}_{z,\varphi}V_{k+1,s}\left(\hat{a}' + \hat{b}', z', \varphi'\right) + \beta\delta_k U\left(\frac{\hat{a}'}{(1+g)^{-1}o_k}\right)$$

with budget constraint

$$\hat{c} + \frac{1+g}{1+r}\hat{a}' \leq \hat{a} + o_k\Theta_k(z)w_s - o_k T_k(\Theta_k(z)w_s).$$

Here, I define $o_k \equiv \frac{L_\ell}{\mathbf{L}_{\ell+k}} = \frac{L_0}{\mathbf{L}_0}(1+g)^{-k}$ to be the relative size of the cohort with age k. This fits the framework in Section 3 after rescaling Θ_k and T_k, and using the effective interest rate $(1+r)/(1+g)$ and normalized utility function $u_k(\cdot/o_k)$. This works irrespective of whether u_k or U (or both) are non-homothetic.

Productivity growth. To include growth in productivity $A_t \equiv A_0(1+g)^t$ in a possibly non-homothetic economy, it is important that the non-homotheticity grows with the economy (or else saving rates mechanically rise over time). This is endogenously the case in the non-homothetic status model, as status is a measure of one's wealth relative to that of others. In the baseline non-homothetic model, this has to be hard-wired by assuming simultaneous preference trends that let the normalization factor in the utility function grow over time, $o_t \equiv o_0(1+g)^t$. In that case, when normalizing the agents' choices by A, $\hat{c} \equiv c/A$, $\hat{a} \equiv a/A$, and $\hat{b} \equiv b/A$, one finds

$$V_{k,s}(\hat{a}, z, \varphi) = \max\ u_k\left(\frac{\hat{c}}{o}\right) + \beta(1-\delta_k)\mathbb{E}_{z,\varphi}V_{k+1,s}\left(\hat{a}' + \hat{b}', z', \varphi'\right) + \beta\delta_k U\left(\frac{\hat{a}'}{(1+g)^{-1}o}\right)$$

with budget constraint

$$\hat{c} + \frac{1+g}{1+r}\hat{a}' \leq \hat{a} + \Theta_k(z)\hat{w}_s - \hat{T}_k(\Theta_k(z)\hat{w}_s).$$

For the budget constraint, I also normalized wages $\hat{w}_s \equiv w_s/A$ and assumed the tax function is indexed, i.e. $T_{k,t}(y) \equiv A_t\hat{T}_k(y/A_t)$. Thus, as long as o_t trends upwards over time with the growth in the economy, the normalized model is again exactly of the form of the framework in Section 3

H Simulation of the transitional dynamics in Section 6

The transitional dynamics are computed as a (non-stationary) equilibrium, defined in Section E.1. As defined there, the equilibrium has perfect foresight.

Simulation. I simulate the transitional dynamics starting at a steady state with the 1970 income distribution, for $T = 400$ years into the future. After 2014, the labor income shares that are fed into the model are assumed to remain constant (see Figure 6). To avoid memory problems when simulating long transitional dynamics, it is necessary to reduce the state space, down from 1.3 million states. I achieve this by reducing the income states to 10 (5 for the persistent shock $\times$ 2 for the transitory shock), so that in total there are $400k$ states in this reduced version of the model. I verified that the steady state predictions of the reduced version are comparable to the larger model.

Algorithm. The computation of an equilibrium candidate *given* an interest rate path $\{r_t\}$ is as follows:

1. Assume a path for the measure of bequests $\{\chi_t\}$.
2. Given $\{\chi_t\}$, iterate backwards in time using the Euler equation to find the paths for policy functions $\{\mathbf{c}_{k,s,t}(a,z,\varphi), \mathbf{a}_{k,s,t}(a,z,\varphi)\}$.
3. Given the policy functions $\{\mathbf{a}_{k,s,t}(a,z,\varphi)\}$, iterate forward in time to compute the path of equilibrium distributions $\{\mu_t\}$ as well as the path of bequest measures $\{\chi_t\}$.
4. Repeat Steps 1–3 until the path of bequest measures $\{\chi_t\}$ converged.

Each such simulation produces a path for the aggregate asset imbalance,

$$\delta A_t \equiv -A_t + (1+r_t)B + J_t(K_t).$$

Below, I use the notation $\delta\mathbf{A}$ for the vector of asset imbalances $\{\delta A_t\}$ and $\mathbf{r}$ for the vector of interest rates $\{r_t\}$. I use the following algorithm to simulate the transitional dynamics equilibrium:[68]

1. *Compute an approximated Jacobian* $\mathbf{J}$*, whose columns are denoted by* $(\mathbf{J}_t)$*:*
   (a) Compute the "revaluation impulse" $\mathbf{J}^{reval}$ as the slope of the response $\delta\mathbf{A}$ that is obtained if the initial value of capital $J_0(K_0) + J_0^{ext}$ is increased by a small amount. This step is important because any future interest rate change affects the value of current non-bond assets from capital, and from trading external assets.
   (b) Denote by $\mathbf{e}_t$ the t-th unit vector. Compute $\mathbf{J}_t \equiv \frac{1}{\Delta r}\delta\mathbf{A}(\mathbf{r}^{final} + \Delta r \times \mathbf{e}_t)$ for a small step Δr for a small number of times $t = t_1, \ldots t_l$. I use $l = 8$ and $\{t_l\} = \{1,2,3,4,5,10,55,100\}$. Store the "revaluation effect" that each such simulation produces, that is, store $\delta v_t \equiv \frac{\Delta\left(J_0(K_0)+J_0^{ext}\right)}{\Delta r_t}$, and compute the "pure" response (without revaluation) $\mathbf{J}_t^{pure} \equiv \mathbf{J}_t - \delta v_t \mathbf{J}^{reval}$.

[68] See Auclert, Bardoczy, Rognlie and Straub (2019) for a new algorithm building on a similar idea of pre-computing the Jacobian.

(c) For any time t that is not in $\{t_1, \ldots, t_l\}$, compute $\mathbf{J}_t^{pure}$ as follows. Here, $(\mathbf{J}_t^{pure})_s$ denotes the s-th element of the vector.

i. Extrapolate $\mathbf{J}_t^{pure}$ assuming a log-linear decay at both sides. After this step, $(\mathbf{J}_t^{pure})$ is treated as a two-sided vector, where $(\mathbf{J}_t^{pure})_s$ is well defined for any $s \in \mathbb{Z}$. Extrapolate $\mathbf{J}^{reval}$ in the same way.

ii. When $t < t_l$ and t lies between $t_\ell < t_{\ell'}$ with $t_\ell, t_{\ell'} \in \{t_1, \ldots t_l\}$, interpolate $\mathbf{J}_t^{pure}$ horizontally, that is, for all $s \in \mathbb{Z}$,

$$(\mathbf{J}_t^{pure})_{t+s} \equiv \exp\left\{\frac{t_{\ell'} - t}{t_{\ell'} - t_\ell} \log(\mathbf{J}_{t_\ell}^{pure})_{t_\ell+s} + \frac{t - t_\ell}{t_{\ell'} - t_\ell} \log(\mathbf{J}_{t_{\ell'}}^{pure})_{t_{\ell'}+s}\right\}.$$

iii. When $t > t_l$, shift $\mathbf{J}_{t_l}^{pure}$ horizontally,

$$(\mathbf{J}_t^{pure})_{t+s} \equiv (\mathbf{J}_{t_l}^{pure})_{t_l+s}.$$

(d) Compute the columns of the approximated Jacobian as

$$\mathbf{J}_t = \mathbf{J}_t^{pure} + \delta v_t \times \mathbf{J}^{reval}$$

which gives a vector $\mathbf{J}_t \in \mathbb{R}^T$ for each $t \in \{0, \ldots, T-1\}$.

2. Use a nonlinear version of Krylov subspace method *GMRES* to solve for the equilibrium interest rate $\mathbf{r}$, starting with guess $\mathbf{r}^{(0)} = \mathbf{r}^{final}$:

(a) Given guess $\mathbf{r}^{(n)}$, $n \geq 0$, evaluate $\delta\mathbf{A}(\mathbf{r}^{(n)})$.

(b) Compute $\tilde{\mathbf{r}}^{(n+1)} = \mathbf{r}^{(n)} - \mathbf{J}^{-1} \cdot \delta\mathbf{A}(\mathbf{r}^{(n)})$.

(c) Compute $\mathbf{r}^{(n+1)} \equiv \sum_{m=1}^{n+1} \lambda_m \tilde{\mathbf{r}}^{(m)}$ where $\sum_m \lambda_m = 1$ and the weights $\{\lambda_m\}$ minimize the norm

$$\Big\| \sum_{m=0}^{n} \lambda_{m+1} \delta\mathbf{A}(\mathbf{r}^{(m)}) \Big\|.$$

(d) Go back to step 2(a) until $\|\delta\mathbf{A}(\mathbf{r}^{(n)})\|_\infty$ is sufficiently small.

This procedure takes around 80 minutes to converge on a 2009 MacBook Pro, with a tolerance of 10^{-7} for $\|\delta\mathbf{A}(\mathbf{r}^{(n)})\|_\infty$.