

CONSUMPTION, SAVINGS, AND THE DISTRIBUTION OF PERMANENT INCOME*

Ludwig Straub
MIT

January 2, 2018

Job Market Paper

Please check <http://economics.mit.edu/grad/straub/research> for updates.

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 models with precautionary savings motives—consumption is a *linear* function of permanent income. This implies that macroeconomic aggregates are neutral with respect to shifts in the distribution of permanent income. Motivated by this neutrality result, I develop novel approaches to test for linearity in U.S. household panel data, which consistently estimate the elasticity of consumption to permanent income in common precautionary savings models. The estimates suggest an elasticity of 0.7, soundly rejecting linearity. To quantify the effects of this deviation from neutrality, I extend a canonical precautionary savings model to include non-homothetic preferences across periods, capturing the idea 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 caused: (a) a decline in real interest rates of around 1%; (b) an increase in the wealth-to-GDP ratio of around 30%; (c) wealth inequality to rise almost as rapidly as it did in the data.

*I am indebted to Iván Werning, Jonathan Parker, Robert Townsend, and Alp Simsek for their continuous guidance and support throughout this project. I am very grateful to Daron Acemoglu, Marios Angeletos, Adrien Auclert, Sebastián Fanelli, Ernest Liu, Matthew Rognlie, and Nathan Zorzi for numerous helpful conversations. I also would like to thank Martin Beraja, Vivek Bhattacharya, Ricardo Caballero, Arnaud Costinot, Emily Gallagher, Daniel Greenwald, Adam Guren, Greg Howard, Chen Lian, Christopher Palmer, Andrés Sarto, and Olivier Wang for many useful comments, and the seminar participants at the MIT macro and finance lunches. I thank the Macro-Financial Modeling Group for financial support. All errors are my own. *Email:* straub@mit.edu

1 Introduction

U.S. labor income inequality has increased substantially over the past few decades (Katz and Murphy, 1992; Autor et al., 2008), with the top 10% now earning over 35% of all labor income (Piketty and Saez, 2003). A significant share of this increase appears to be the result of rising dispersion in the fixed-effect component of labor income, which captures the returns to skill or ability and which I henceforth refer to as *permanent income*.¹ Indeed, Guvenen et al. (2017) argue 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 predicted to be entirely 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 widely 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 challenge the existing neutrality paradigm, both empirically and quantitatively. I have two main findings. First, I propose ways to consistently estimate the permanent income elasticity of consumption; I find estimates around 0.7, significantly below 1, indicating a concave consumption function in permanent income. Second, I incorporate non-homothetic preferences into a canonical precautionary-savings model to match this elasticity and study the quantitative implications: the model suggests that 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 going forward).

The first contribution of this paper is to propose new ways to test the linearity of consumption in permanent income, building on previous work by Friedman (1957), Mayer (1972), and Dynan et al. (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 yields the *permanent income elasticity of consumption*, ϕ , which is equal to 1 under the null hypothesis that consumption is

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

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

³There are exceptions to this, including Hubbard et al. (1994, 1995) and De Nardi (2004). See the discussion below.

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 future quasi-differenced 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 around 0.7, statistically and economically 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.

This finding is best interpreted as follows: if working-age household A always earns twice as much in after-tax income as working-age household B, household A will not spend 100% more, but rather only 70% more. One may wonder whether this constitutes a significant source of non-neutrality. As I illustrate with a simple back-of-the-envelope calculation that assumes consumption c to be a power function of income y , $c \sim y^\phi$, the difference is sizable: the shift in the U.S. income distribution between 1980 and 2014 implies translates into a reduction in aggregate consumption by approximately 4%.

The second contribution of this paper is to investigate the implications of a concave consumption function in permanent income quantitatively. To this end, I build a non-neutral version of a quantitative life-cycle model with idiosyncratic income risk and incomplete markets in the tradition of Deaton (1991), Huggett (1996) and Gourinchas and Parker (2002).⁴ Aside from standard sources of non-neutrality, such as nonlinear tax-and-transfer and social security systems (Hubbard et al., 1994, 1995; Scholz et al., 2006), the key elements of my model are two kinds of non-homothetic preferences. The first element follows the seminal work of De Nardi (2004) and assume that bequests are treated as a luxury good. The second key element concerns “life-cycle” non-homothetic preferences, over consumption across periods, which turn out to be the most important source of non-neutrality in the model. Such 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, which a poorer household cannot afford to do. 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.⁵ I calibrate the strength of the life-cycle non-homotheticity to match

⁴See also Imrohoroglu et al. (1995), Ríos-Rull (1996) and Carroll (1997) among others.

⁵The 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”.

the elasticity ϕ when estimating the same regressions on artificial panel data simulated from the model.⁶ 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) wealth distribution of 2014 quite well—except at the very top, for the fractiles within the top 1%.⁷

I use the calibrated economy as a laboratory to study the implications of rising permanent income inequality. In partial equilibrium (PE), keeping the interest rate fixed, I find that a shift from a steady state with the level of permanent income inequality of 1970 to that of 2014 shows a considerable increase in aggregate wealth, of just above 130% of GDP. As a point of comparison, the U.S. net foreign asset position has declined by “only” 18% of GDP since the 1997–98 Asian financial crisis, which is often attributed to the recent “global savings glut”. This is a first hint that rising income inequality may not have been neutral in the past few decades.

I then simulate the general equilibrium transitional dynamics from the 1970 steady state to recent levels of permanent income inequality, which I assume to remain constant after 2014.⁸ This exercise allows the model to speak directly to the forces behind three important recent macroeconomic trends: (a) the decline in real (natural) interest rates since the 1980s (Laubach and Williams, 2003, 2015), (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% through 2017, which explains 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 to decline, eventually falling by another 1%. 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. In particular, this means the most able or skilled workers entering today will amass much larger fortunes over their lifetimes than previous generations. This effect causes a large and predictable decline in interest rates going forward.

Concerning the second trend, the endogenous interest rate response limits the rise in aggregate wealth to around 30% of GDP through 2017 (again, roughly one-third of the rise in the data), with an eventual total increase of 55%.

Finally, the model offers an explanation for the size and speed of the increase in the top 10%

⁶In that sense, my calibration shares similarities with indirect inference (Gourieroux et al., 1993; Smith, 1993, 2008; Guvenen and Smith, 2014), in that one of the IV regressions is treated as “auxiliary model” along which the model is matched to the data.

⁷It is well known that a substantial fraction of the top 1% and especially the top 0.1% are business owners, who are not modeled here, which could explain the deviation. See Quadrini and Ríos-Rull (1997) and Cagetti and De Nardi (2006) for models of entrepreneurship and the wealth distribution.

⁸The 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 F). The improvements were developed jointly with Adrien Auclert and Matt Rognlie.

wealth share and around two-thirds of the increase in the top 1% wealth share.⁹ In sum, the model suggests that rising permanent labor income inequality alone can account for a significant share of three major macroeconomic trends.

Literature. My paper combines several strands of a vast literature at the intersection of inequality, consumption dynamics, and macroeconomics.¹⁰

First, my paper contributes to the large empirical literature that tests the permanent income hypothesis (PIH), starting with [Friedman \(1957\)](#) himself. 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.¹¹ 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.¹²

The main exception to this is the work of [Dyner et al. \(2004\)](#), henceforth DSZ.¹³ Their paper computes savings rates, either as consumption-based measures $(Y - C)/Y$ (CEX) or as wealth difference based measures $\Delta A/Y$ (SCF, PSID), and documents two main facts. First, savings rates increase across current income quintiles. Second, savings rates still increase in income quintiles if income is instrumented by lagged or future income or education. My empirical exercise follows their lead, innovating along several dimensions. First, I focus solely on consumption and not on wealth differences, which are problematic because it is generally difficult to disentangle ex-ante savings behavior from ex-post returns or transfers. Second, I show that the relationship between log consumption and log income is roughly log-linear, which allows a focus on a single elasticity parameter ϕ . Finally, and most importantly, I use a panel data set with consumption and income (the PSID since 1999), which allows for the development of 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

⁹The dynamics of the wealth distribution have recently been investigated theoretically by [Gabaix et al. \(2017\)](#), and numerically by [Hubmer et al. \(2016\)](#), [Kaymak and Poschke \(2016\)](#), and [Aoki and Nirei \(2017\)](#) in incomplete-markets models.

¹⁰For recent surveys and books, see among others [Bertola et al. \(2005\)](#); [Krusell and Smith \(2006\)](#); [Heathcote et al. \(2009\)](#); [Güvenen \(2011\)](#); [Quadrini and Ríos-Rull \(2015\)](#); [De Nardi et al. \(2015\)](#); [Piketty and Zucman \(2015\)](#); [De Nardi and Fella \(2016\)](#); [Attanasio and Pistaferri \(2016\)](#); [Piketty \(2017\)](#); [Benhabib and Bisin \(2017\)](#).

¹¹See e.g. [Friedman \(1957\)](#), [Mayer \(1966\)](#), [Evans \(1969\)](#), and [Mayer \(1972\)](#).

¹²This empirical work started using Euler equation-based tests ([Hall, 1978](#); [Flavin, 1981](#); [Hall and Mishkin, 1982](#)) and now also includes well-identified empirical studies ([Johnson et al., 2006](#); [Parker et al., 2013](#)).

¹³For similar approaches see [Bozio et al. \(2013\)](#) for the UK and [Alan et al. \(2015\)](#) for Canada. [Gustman and Steinmeier \(1999\)](#) and [Venti and Wise \(2000\)](#) propose a look at the relationship between retirement wealth and lifetime income. As part of my analysis in Section 5, I show that this relationship is not well suited to inform the degree of non-neutrality in the presence of income shocks.

improves upon simple instruments such as lagged or future income (which lead to downward-biased results under the assumptions of a canonical neutral 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.

The second main contribution of this paper is the analysis of rising permanent income inequality 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 significantly increase wealth inequality compared to a homothetic model, albeit by not quite as much as the data.¹⁴ My paper follows their lead and argues that one needs (considerably) more non-neutrality than what the existing model elements generate. 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 relatively well.

In addition, my paper builds on a recent literature that studies the quantitative effects of income inequality in (mostly) neutral economies. Here, [Auclert and Rognlie \(2017\)](#) show how greater inequality has aggregate effects that crucially depend on the types of incomes (transitory, persistent or permanent) that become more unequal. They also consider implications for economies at the zero lower bound, where there can be a feedback loop between aggregate demand and endogenous income risk. [Heathcote et al. \(2010\)](#) 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.¹⁵ The key distinguishing feature of my paper is that I focus on rising *permanent* income inequality, arguably among the most important drivers of rising income inequality in the U.S..¹⁶ And, to study its consequences, it is important to model an economy that gets the degree of non-neutrality right.

The interest in modeling non-homothetic consumption-savings behavior is shared by a number of earlier, mostly deterministic, papers. [Uzawa \(1968\)](#) was the earliest well-known efforts to do this using recursive [Koopmans \(1960\)](#) utility and [Epstein and Heynes \(1983\)](#) and [Epstein \(1987\)](#) subsequently extended this work. [Lucas and Stokey \(1984\)](#) used such preferences to study many-agent neoclassical growth models without degenerate wealth distributions (see also [Obstfeld \(1990\)](#)). Interestingly, however, these papers end up focusing on the opposite of the empirical case, namely

¹⁴See also [Huggett and Ventura \(2000\)](#), [Scholz et al. \(2006\)](#) and [De Nardi and Yang \(2014\)](#). [Kumhof et al. \(2015\)](#) study the interaction of income inequality, debt and defaults in a two-agent economy with non-homothetic preferences.

¹⁵[Holm \(2017\)](#) studies the differential effects of monetary policy during times with increased persistent income risk.

¹⁶See [Güvenen et al. \(2017\)](#) for direct evidence on the importance of permanent income inequality, in line with [Sabelhaus and Song \(2010\)](#), who document a reduction in transitory and persistent income volatility. Suggestive evidence also appears in [DeBacker et al. \(2013\)](#) and [Kopczuk et al. \(2010\)](#).

economies where richer agents save less than poorer agents, since this is precisely the case in which multi-agent deterministic infinite-horizon models have a stationary distribution.¹⁷ In contrast to these papers, my economy admits a non-degenerate stationary wealth distribution despite richer agents saving more, which is possible because my model also includes borrowing constraints and idiosyncratic risk.

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 et al. \(2008\)](#) and [Kaplan and Violante \(2010\)](#).¹⁸ 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. They use this method to show that household consumption “under-reacts” to permanent income shocks. [Kaplan and Violante \(2010\)](#) extend the framework to allow for shocks with persistence less than unity and show that those shocks generally lead to larger degrees of under-reaction due to self-insurance. Compared to these papers, the focus in this paper is on the relationship between the level of consumption and the level of permanent income, rather than on changes. I show in extensions of both the empirical analysis and my model that the degree of partial insurance is largely orthogonal to the curvature in consumption as a function of permanent income. However, this paper very much shares the spirit of these papers in that they identify important moments in similar panel data on consumption and income and use them to inform microfounded consumption-savings models.

Layout. I begin in Section 2 by demonstrating in a stylized, two-agent framework how a concave consumption function in permanent income can be modeled and what its likely effects are. Section 3 introduces a canonical precautionary-savings model and explains under what assumptions this model is neutral with respect to changes in the permanent income distribution. I test for neutrality in Section 4. Extending the canonical model, Section 5 then relaxes the neutrality assumptions—mainly by introducing non-homothetic preferences—and highlights the main properties of the non-homothetic model. The effects of rising income inequality in partial and general equilibrium are simulated in Section 6. Section 7 concludes and discusses potential avenues for future research. The appendix contains all proofs, as well as additional empirical and quantitative results.

¹⁷In a few more theoretically-oriented papers, however, richer agents do save more. [Cole et al. \(1992, 1995, 1998\)](#), [Robson \(1992\)](#), [Ray and Robson \(2012\)](#) (implicitly or explicitly) model utility over one’s wealth rank. [Carroll \(2000\)](#) models utility over wealth directly. In [Becker and Mulligan \(1997\)](#) agents can invest in raising their discount factor. Recent models of savings behavior with disagreements between current and future selves can generate similar patterns, see e.g. [Harris and Laibson \(2001\)](#) and [Cao and Werning \(2016\)](#).

Finally, a number of papers studies the non-homothetic preferences at the intersection of macroeconomics and development (see, e.g. [Moav \(2002\)](#); [Galor and Moav \(2004\)](#)).

¹⁸For other papers in this literature see e.g. [Cochrane \(1991\)](#), [Townsend \(1994\)](#), [Attanasio and Davis \(1996\)](#), [Attanasio and Pavoni \(2011\)](#), [Blundell et al. \(2016\)](#), [Arellano et al. \(2017\)](#).

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 neutral and non-neutral models. For this purpose, Section 2.1 introduces the utility maximization problem of a single dynasty, exemplifying how the consumption function can be concave in permanent income. Section 2.2 then combines two such dynasties and studies the partial equilibrium implications of greater permanent income inequality. Section 2.3 closes the economy by adding a standard neoclassical supply side and characterizes the general equilibrium implications of rising inequality. All figures in this section are constructed using a standard calibration which I explain in Appendix B.¹⁹

2.1 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 $t = 0, 1, 2, \dots$ the current generation solves

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

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

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,

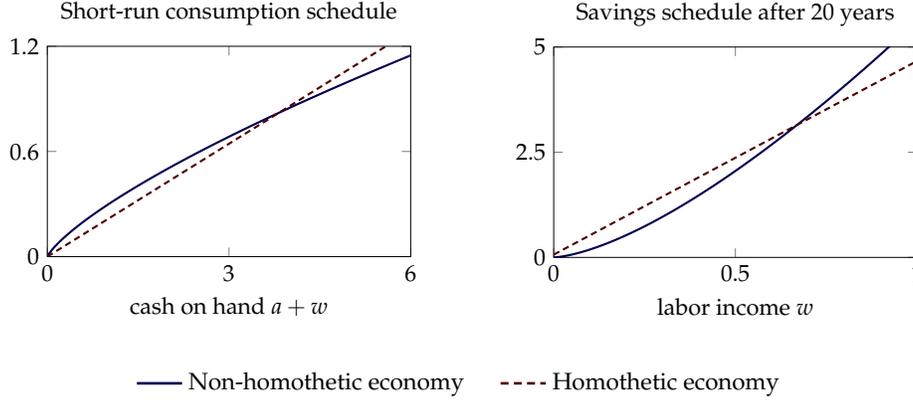
$$u(c) = \frac{(c/z)^{1-\sigma} - 1}{1-\sigma} \quad U(a) = \frac{(a/z)^{1-\Sigma} - 1}{1-\Sigma}, \quad (3)$$

but the two (inverse) elasticities $\sigma, \Sigma > 0$ are allowed to differ. In this formulation $z > 0$ is a normalization parameter that allows the model to retain aggregate scale invariance.²⁰ Heterogeneity in the inverse elasticities σ, Σ represents the single deviation from a standard homothetic consumption-savings model. It allows the model to capture the fact that richer dynasties may

¹⁹ Atkinson (1971) and Benhabib et al. (2011) study related OLG economies with a non-homothetic bequest motive.

²⁰ In a model with growth, it would be natural to assume z grows at the same speed as the economy. This captures the idea that for savings behavior it is not the absolute level of one's income that matters, but the income relative to the aggregate economy.

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.

have a greater propensity to save in the following way: the utility maximization problem (1) can be thought of as a simple decision problem between two goods, consumption c_t and savings a_{t+1} . In this decision 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.²¹ With utilities as power functions, this is the case if $\Sigma < \sigma$, so that the utility over savings (bequests) is more linear than the utility over consumption. This can also be seen from the Euler equation,

$$c_t/z = (\beta R)^{-1/\sigma} (a_{t+1}/z)^{\Sigma/\sigma}, \quad (4)$$

which shows that consumption c_t adjusts by less than savings a_{t+1} when $\Sigma/\sigma < 1$.²²

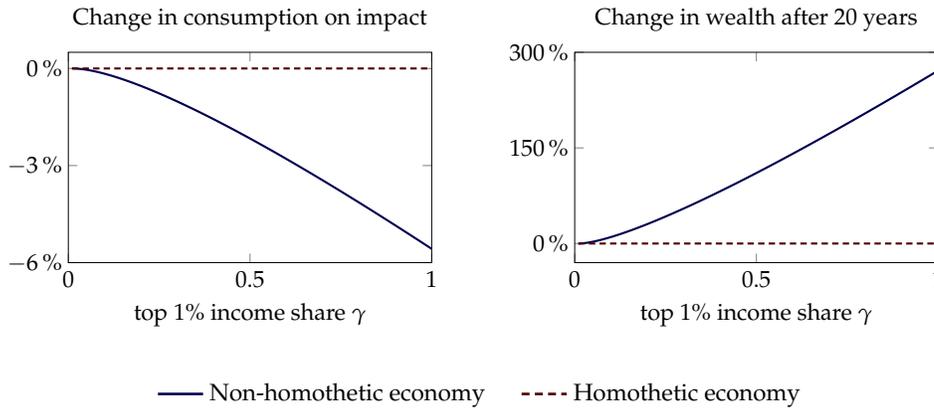
Figure 1 illustrates two key outcomes of the utility maximization problem (1). Panel (a) shows the optimal *short-run* consumption choice c_t as a function of cash on hand. I call it “short-run” as it takes current assets as given. Panel (b) shows the optimal *long-run* asset position as a function of permanent income w , where long-run means after 20 years. In both panels the agent starts with the average wealth and income position in the economy. The panels show two cases: 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.

As a side note, the consumption schedule turns out to be well approximated by a simple power

²¹See also Strotz (1955) and Blinder (1975) for early deterministic life-cycle models with non-homothetic utility over bequests. See Andreoni (1989) for further work on joy-of-giving preferences and Abel and Warshawsky (1988) for the relationship with perfect altruism.

²²It 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 in current income or assets, but not permanent income.

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



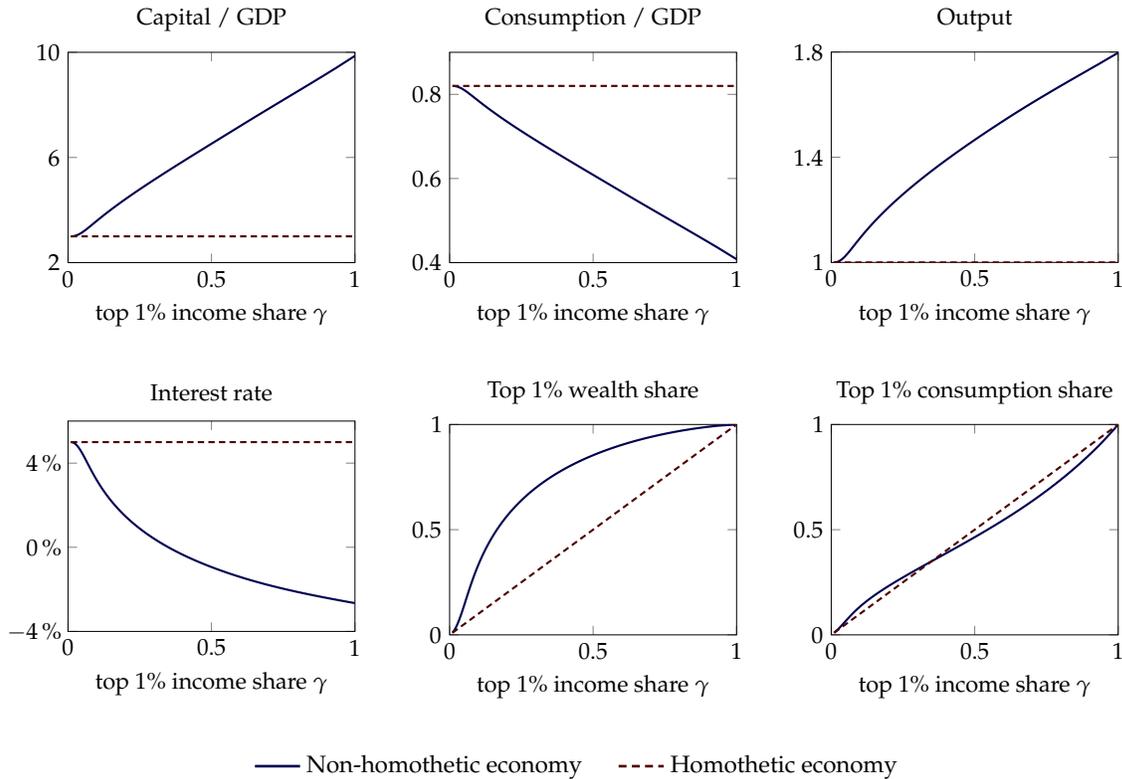
function $c_t \approx k(a_t + w)^\phi$ for large values of cash on hand, where the exponent is given by the ratio of the elasticities, $\phi = \Sigma/\sigma$. The elasticity ϕ takes a central role in this paper, as it succinctly characterizes the degree of concavity in consumption as a function of permanent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

2.2 Partial 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. Thus, 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, $w^r \geq w^p$. Assume that the population share of the rich dynasty is $\mu \in (0, 1)$. 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 . In all figures below, I take μ to be 1%. Denote by $W = w^r + w^p$ total labor income, which is assumed to be constant in this subsection, so that inequality γ uniquely defines w^r and w^p .

Imagine that the economy is initially perfectly equal, $\gamma = \mu$, and consider an unanticipated increase in inequality, $\gamma > \mu$. Figure 2 shows what happens to short-run consumption and long-run savings in this scenario. Given the curves in Figure 1 the result is unsurprising, yet powerful: in the homothetic model, where $\Sigma = \sigma$, nothing happens to either consumption or savings. This is a direct consequence of the linearity in Figure 1 and makes this economy a simple example of an economy where the permanent income distribution is *neutral*. In the non-homothetic economy, aggregate consumption falls on impact, and long-run savings rise.

Figure 3: Stylized model: Neutrality and non-neutrality in general equilibrium.



2.3 General equilibrium effects of greater inequality

The results in the previous subsection raise the question of what happens in general equilibrium. Closing the model requires to specify the supply side of this economy, which is assumed to be given by a Cobb-Douglas aggregate production function,

$$Y = F(K, L^r, L^p) = AK^\alpha (L^r)^{(1-\alpha)\gamma} (L^p)^{(1-\alpha)(1-\gamma)},$$

where $A > 0$, K denotes capital (assumed to depreciate at rate $\delta > 0$), and L^p and L^r denote labor supplied by the poor and rich dynasties.

Again the same experiment is conducted: Starting at perfect equality, $\gamma = \mu$, what happens in this economy when γ is increased? The six panels in Figure 3 show the long-run outcomes for the homothetic and the non-homothetic economies. As anticipated, the homothetic economy is neutral, so that none of the aggregate quantities K, C, Y or interest rates are affected. Wealth inequality increases, but only at the same rate as income inequality, reflecting the proportionality of assets and income in the model. Similarly, consumption inequality increases at the same rate, as well.

By contrast, in the non-homothetic economy, the capital stock and output increase with greater inequality, while interest rates fall. Interestingly, wealth inequality rises faster than one-for-one with

inequality, but this does not translate into greater consumption inequality: the level of consumption inequality is similar to that in the homothetic model. This is mainly due to the endogenous interest rate decrease, which reduces the rich dynasty's capital income.

2.4 Takeaway for the rest of this paper

The results in this section show that homothetic preferences tend to induce a linear consumption function in permanent income, while non-homothetic preferences induce a concave consumption function. This allows models endowed with the former to be *neutral* and models endowed with the latter to be *non-neutral*. These ideas foreshadow the rest of this paper: the general model in Section 3 nests the stylized homothetic model and proves a general neutrality result; the concavity parameter ϕ is estimated in the data in Section 4; then, I extend the general model to include non-homothetic preferences to match ϕ in Section 5. Finally, Section 6 investigates the partial and general equilibrium properties of rising income inequality.

3 General Model and Neutrality Result

In the previous section I demonstrated that a simple homothetic dynastic economy is *neutral* with respect to changes in the permanent income distribution: all aggregates are invariant in partial and general equilibrium, while measures of inequality change linearly in permanent income inequality. This section generalizes these results and proves that they carry over to a large class of models. The general model introduced in this section also lays the foundation for what is to come: the conditions identified here motivate the empirical analysis in Section 4, and the quantitative model in Section 5 is a version of the general model.

3.1 Setup

Time is discrete, $t \in \{0, 1, \dots\}$, 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.²³ 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 type $s \in \mathcal{S}$ as innate skill or ability.²⁴ Agents with skill s are endowed with on average a single efficiency unit of skill s and make up a constant share $\bar{\mu}_s \in [0, 1]$ of the population. To allow for overlapping generations, I assume that there is a constant inflow and outflow of agents at rate $\bar{\delta} \geq 0$,

²³Endogenous bequests are important in a realistic model of wealth inequality. See e.g. [Castaneda et al. \(2003\)](#), [De Nardi \(2004\)](#), [Benhabib et al. \(2011\)](#).

²⁴This model abstracts from endogenous investment into human capital. See [Heathcote et al. \(2010\)](#) for a model along those lines.

where zero is included. An agent's age is indexed by $k \in \mathbb{N}$. With an OLG structure, $\bar{\delta} > 0$, 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 .²⁵ 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, $Y = (r + \delta)K + \sum_{s \in \mathcal{S}} w_s L_s$. The main comparative statics exercise in this section will be 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 (relevant only in the OLG case, if $\bar{\delta} > 0$) 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. 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. 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 and with some skill $s \in \mathcal{S}$. In the infinite-horizon case, agents are already alive at date $t = 0$. The 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. In the life-cycle case, 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.

²⁵The results in this section generalize to arbitrary neoclassical production functions and arbitrary shifts in the distribution of labor income.

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

$$\begin{aligned}
V_{k,s}(a, z, \varphi) &= \max u_k(c) + \beta(1 - \delta_k)\mathbb{E}_{z,\varphi} V_{k+1,s}(a' + b', z', \varphi') + \beta\delta_k U(a') & (5) \\
c + \frac{1}{1+r}a' &\leq a + \Theta_k(z)w_s - T_k(\Theta_k(z)w_s) \\
(b', \varphi') &\sim v(\cdot|s, k, \varphi) \\
a' &\geq 0.
\end{aligned}$$

3.2 Equilibrium

Denote the state space by $\mathbb{S} \equiv \mathcal{S} \times \{1, \dots, K_{death}\} \times \mathbb{R}_+ \times \mathcal{Z} \times \{0, 1\}$ 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 $\mathcal{S} \times \{1, \dots, K_{death}\} \times \mathbb{R}_+$ 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 (5), where the conditional bequest distribution $v(\cdot|s, k, \varphi)$ is given by

$$v(\mathbf{B}, \varphi'|s, k, \varphi) = \begin{cases} 1_{\{0,1\}}(\mathbf{B}, \varphi') & \text{if } \varphi = 1 \\ (1 - \delta_{k+k_{born}})1_{\{0,0\}}(\mathbf{B}, \varphi') + \frac{1}{\bar{\mu}_s} \sum_{s'} P_{ss'} \chi(s', k + k_{born}, \mathbf{B}) & \text{if } \varphi = 0 \end{cases} \quad (6)$$

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, $L_s = \bar{\mu}_s$, (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

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

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

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

3.3 Assumptions for neutrality

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 $k \in \{1, \dots, K_{death}\}$.*

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 for the neutrality results is that bequests play no redistributive role, that is, no rich person leaves any wealth to a less-skilled offspring.

Assumption 3 (Perfect skill persistence). *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, $\tau_b = 1$.*

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 rule out boundary cases with ill-defined equilibrium wealth distributions.

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.

Having made these important but common assumptions, I can now characterize the micro implications of steady state equilibria in this economy.

3.4 Linearity and aggregate neutrality

I start by showing a helpful 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 $(s, k, \mathbf{A}, z, \varphi) \in \mathcal{S}$, any state $(s, k, a, z, \varphi) \in \mathcal{S}$ and any skill $s' \in \mathcal{S}$ it holds in any equilibrium that*

$$\begin{aligned} \mu(s, k, \mathbf{A}, z, \varphi) &= \frac{\bar{\mu}_s}{\bar{\mu}_{s'}} \times \mu\left(s', k, \mathbf{A} \frac{w_{s'}}{w_s}, z, \varphi\right) \quad \text{and} \quad \chi(s, k, \mathbf{A}) = \frac{\bar{\mu}_s}{\bar{\mu}_{s'}} \times \chi\left(s', k, \mathbf{A} \frac{w_{s'}}{w_s}\right) \\ \mathbf{c}_{k,s}(a, z, \varphi) &= \frac{w_s}{w_{s'}} \times \mathbf{c}_{k,s'}\left(a \frac{w_{s'}}{w_s}, z, \varphi\right) \quad \text{and} \quad \mathbf{a}_{k,s}(a, z, \varphi) = \frac{w_s}{w_{s'}} \times \mathbf{a}_{k,s'}\left(a \frac{w_{s'}}{w_s}, z, \varphi\right). \end{aligned}$$

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 $aw_{s'}/w_s$. An interesting implication of this is that the distribution of MPCs is the same for each skill s .²⁶ This is an immediate consequence of differentiating the equation for $\mathbf{c}_{k,s}(a, z, \varphi)$ in Lemma 1 with respect to a . I state and prove this result formally in Appendix C.2.

Lemma 1 can be used to derive testable predictions based on micro-level consumption behavior. Proposition 1 does this for the relationship between individual consumption and permanent income.

Proposition 1 (Linear consumption function). *Under Assumptions 1–4, in any equilibrium, each agent i with age k has a linear consumption function in permanent income, that is, in logs*

$$\log c_{ik} = \text{const}_k + \log w_{s(i)} + \epsilon_{ik}, \tag{8}$$

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

$$\log y_{ik} = \widetilde{\text{const}}_k + \log w_{s(i)} + \tilde{\epsilon}_{ik}, \tag{9}$$

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

²⁶There is some evidence that MPCs decrease with education (not conditioning on assets a), see Jappelli and Pistaferri (2006, 2014).

Proposition 1 motivates a simple log-linear specification that I use as a basis for testing linearity in Section 4. The next result directly follows from Lemma 1 as well.

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

$$\begin{aligned} \text{Var}_{i,k} \log c_{ik} &= \text{const} + \text{Var}_s \log w_s \\ \text{Var}_{i,k} \log (a_{ik} + y_{ik}) &= \text{const} + \text{Var}_s \log w_s, \end{aligned}$$

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

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 the average permanent income $\mathbb{E}_s w_s$,*

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

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 the distribution of permanent incomes is irrelevant for aggregate consumption, savings, investment, tax revenues, bequests, asset prices, and the interest rate.

The intuition behind this result is straightforward given the discussion of the previous subsection. 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.

3.5 Discussion

These results are an example of an exact aggregation result. In essence, the Engel curves for consumption in different time periods are linear in permanent income (and symmetric across agents). This allows one to treat the economy *as if* there were only a single skill type earning the average permanent income. The focus on permanent incomes—that is, individual fixed effects—distinguishes

this result from previous aggregation results: 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.

The linearity result has been stated in a fairly general way, but not as general as possible. Similar results hold with progressive tax systems,²⁷ endogenous labor supply, habit formation, aggregate risk, or non-separable preferences (e.g. Epstein-Zin preferences).

There are a few limitations to the linearity result, however. First and foremost, this 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](#) below, I explore this effect quantitatively and find it to be negligible, even for large changes in the income distribution, such as a sudden movement from the U.S. distribution in 1970 to the U.S. distribution in 2014.

Second, the result also no longer applies when initial assets or borrowing constraints are nonzero *and* do not scale with an agent's permanent income level w_s .²⁸ Again, I explored these departures numerically, and they have only marginal effects on the validity of the linearity proposition. I discuss other limitations and how they affect my tests of the proportionality hypothesis in [Section 4.4](#).

Finally, I assume joy-of-giving preferences. If, instead, one assumes altruistic preferences, *and* relaxes [Assumption 3](#), introducing imperfect skill persistence, the result no longer holds exactly. In that case, altruistic preferences introduce two reasons for non-neutrality: first, parents treat bequests as a luxury good since the higher their own permanent income, the relatively lower their child's permanent income is expected to be, inducing parents to save more. In an economy with a joy-of-giving utility U , this would correspond to bequests being treated as a luxury good, and will be an integral part of the quantitative model in [Section 5](#). Second, altruism also allows rising permanent income inequality to *directly* affect savings behavior due to greater precautionary savings. This second feature cannot be informed by the empirical exercise in this paper, and is not a feature of models with an ad-hoc joy-of-giving utility U .²⁹

²⁷This works as long as post-tax incomes are a power function of pre-tax incomes, which holds relatively well in U.S. data. See, e.g. [Benabou \(2000, 2002\)](#) and [Heathcote et al. \(2017\)](#), as well as [Appendix D.7](#) of this paper.

²⁸Especially for borrowing constraints, it seems natural that they would scale in one's permanent income level.

²⁹For evidence on the strength of this channel, see the recent work by [Boar \(2017\)](#). For changes in income inequality in an infinite-horizon economy, which can be thought of agents linked by altruistic bequest motives, see [Auclert and Rognlie \(2017\)](#). For quantitative OLG models with altruism, see [Castaneda et al. \(2003\)](#); [Cagetti and De Nardi \(2006, 2009\)](#).

4 Testing for Neutrality

I introduced a general, neutral model in the previous section. In this section I test for neutrality in the data. Inspired by Proposition 1, the goal is to estimate the following system of equations,

$$\log c_{it} = \mathbf{X}'_{it}\beta + \phi \log w_{s(i)} + \epsilon_{it} \quad (10a)$$

$$\log y_{it} = \tilde{\mathbf{X}}'_{it}\tilde{\beta} + \log w_{s(i)} + \eta_{it} + \psi_{it} + v_{it}. \quad (10b)$$

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 v_{it} is measurement error.³⁰ The key question is whether ϕ , the *permanent income elasticity of consumption*, is equal to 1 or not.

To answer this question, I pursue three separate approaches: an OLS approach, where each household's permanent income level is computed as a symmetric average over log residualized incomes (Section 4.2); and two IV approaches, where each household's current income level is instrumented with two different predictors of permanent income (Section 4.3). In Section 4.4 I discuss at length various potential weaknesses of my empirical designs and provide robustness checks (see also Appendix D.1). Finally, I discuss the economic significance of the estimated value of ϕ using a simple partial equilibrium consumption function framework in Section 4.5.

It is important to stress that my analysis is not about assessing the relationship between *changes* in consumption and *changes* in permanent or persistent income. Instead, my analysis focuses on the dependence of the consumption *level* on a given *level* of permanent income.³¹

Throughout this section, I will denote by $\hat{y}_{it} \equiv \log y_{it} - \tilde{\mathbf{X}}'_{it}\hat{\beta}$ 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. Its initial sample consisted of 5,000 households, of which 3,000 (the “Survey Research Center”, or SRC, sample) were chosen as a representative sample of the U.S. population at the time. Since 1968, researchers have followed both these households and their children's “split-off” households. 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 (Beckett et al., 1988; Andreski et al., 2014).

³⁰All variables will be formally defined below.

³¹Temporary responses to income shocks and the degree of insurance against such shocks are the subject of extensive research, recent papers include Arellano et al. (2017), Blundell et al. (2008), Blundell et al. (2016), Guvenen and Smith (2014), Heathcote et al. (2014), and Kaplan and Violante (2010), among many others.

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).³² 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 will show in Section 4.4 below, moving to the more comprehensive (since 2005) consumption measure has only a minor effect on my results. In my baseline measure, I include all available expenditure categories, including durable goods. 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.³³ I include (non-housing) durable goods since, under analogous assumptions to the ones in Section 3.4, durable goods purchases would scale linearly in permanent income. Still, I also show results for non-durable consumption expenditure below.

Income data. As the income variable in my baseline regressions, I use post-tax household labor income. This is the right income concept to use for my exercise since, as the aforementioned literature on various channels of partial insurance convincingly argues, there exist various channels which may be operative, at least temporarily, in response to changes in permanent incomes. Among the most important such channels are the tax-and-transfer system and the labor supply of family members, both of which are accounted for by using post-tax household labor income. The income measure consists of labor income of all family members minus taxes (computed using NBER's TAXSIM program). I discuss below robustness with respect to alternative income measures—most prominently, with respect to using an after-tax total income measure which includes all forms of capital income as well as private and public transfers.

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 extreme 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, will only be 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 Section 4.4. Throughout, I use PSID's post-1999 longitudinal sample weights.

³²Unlike the CEX, however, the post-1999 PSID consumption data does not suffer from a downward trend relative to the PCE (see, e.g. Blundell et al. (2016)).

³³My 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.

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.³⁴ The results are robust to several other sets of controls, see Section 4.4.

4.2 A first look 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 neutrality 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

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

where T is the odd number of incomes being averaged.³⁵ When $T = 1$, \bar{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 (8) is then

$$\hat{c}_{it} = \mathbf{X}'_{it}\beta + \phi\bar{y}_{it}^T + \epsilon_{it}. \quad (11)$$

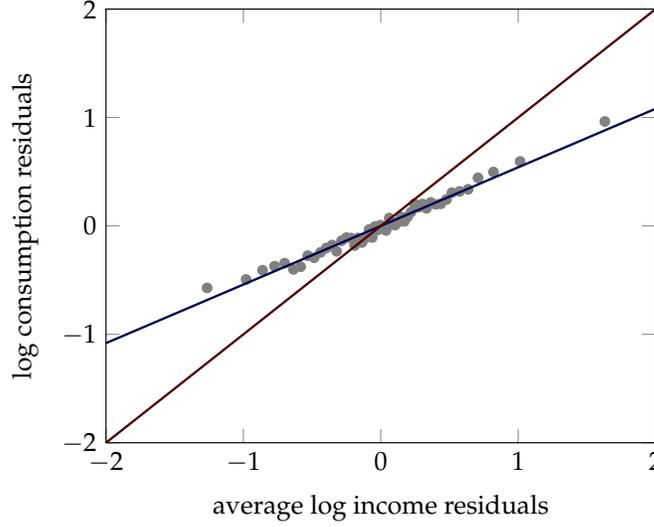
For long averages, $T \rightarrow \infty$, under a suitable law of large numbers for the income processes η_{it}, ψ_{it} , the average income residuals \bar{y}_{it}^T are measuring \hat{w}_i without noise. In that case, the neutral model in Section 3.4 would correspond to $\phi = 1$, implying a linear consumption function in permanent incomes. Since T is finite, however, one can generally expect elasticities ϕ below 1, even if the assumptions of the neutral 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 4 shows the results of a non-parametric version of (11) spanning 9 years ($T = 5$). Specifically, it shows the results of a regression of \hat{c}_{it} on controls and 50 bins of \bar{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 (11). Columns 1–3 show the results of specification (11) 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

³⁴The location dummy is constructed as the interaction of an urban-rural dummy and dummies for the nine Census divisions.

³⁵See Kopczuk et al. (2010) for similar symmetrically-averaged income measures.

Figure 4: 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.

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 (11) 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 economy 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,

$$\hat{c}_{it} = \phi \hat{w}_i + \mathbf{X}'_{it} \beta + \epsilon_{it} \quad (12a)$$

$$\hat{y}_{it} = \hat{w}_i + \eta_{it} + \psi_{it} + v_{it}, \quad (12b)$$

where $t \in t_0(i) + \{0, 1, \dots, K_{death} - 1\}$. I make the following baseline assumptions on the model (12), all of which are satisfied in the neutral model of Section (3). First, all random variables in (12) are iid across households i . Second, measurement error v_{it} is iid over time and uncorrelated with consumption, $Cov(\epsilon_{it}, v_{it}) = 0$. Third, permanent income \hat{w}_i and the controls \mathbf{X}_{it} are uncor-

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

	log household consumption			
	(1) $T = 1$	(2) $T = 5$	(3) $T = 9$	(4) c and y FE
log Income	0.396 (0.01)	0.547 (0.02)	0.645 (0.03)	0.561 (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.

related with the income shocks η_{it} , ψ_{it} , measurement error v_{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, $Cov(\epsilon_{it}, \psi_{it-\tau}) \geq 0$ 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.

The two biases of OLS regressions. Having introduced these assumptions, it is possible to investigate the biases of OLS regressions. As an example, 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

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

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 standard attenuation bias due to the presence of measurement error in income.

It turns out that the consumption smoothing bias is rather hard to overcome. Simple instruments, such as future or lagged incomes are able to eliminate attenuation through measurement error, but due to the presence of persistent income shocks η_{it} , the consumption smoothing bias remains.³⁶ I now propose novel IV strategies that eliminate both biases, using additional assumptions on the autocorrelation structure of the persistent income shocks η_{it} .

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

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

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_{it}^{η} , that is, $\text{Cov}(\epsilon_{it+\tau}^{\eta}, \epsilon_{it}) = 0$ 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

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

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} . When η_{it} is an AR(1) process, the instrument is simply given by quasi-differenced future incomes,

$$z_{it+\tau} = \hat{y}_{it+\tau} - \rho\hat{y}_{it+\tau-1}.$$

Intuitively, this IV strategy combines two ideas. 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

³⁶For instance, some of the specifications Mayer (1972) and Dynan et al. (2004) used instruments along these lines.

Table 2: IV specifications for time-averages and group-averages approach.

	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 (0.02)	0.675 (0.04)	0.686 (0.04)	0.699 (0.05)	0.717 (0.05)	0.741 (0.06)	0.732 (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.

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.³⁷ I consider this case now. In particular, assume that

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

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

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

For an arbitrary process η_{it} , this variable is not an exogenous instrument and not even one whose asymptotic bias can be signed. 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 neutrality.³⁸

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

³⁷See the recent survey by [Meghir and Pistaferri \(2010\)](#). 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](#)).

³⁸In 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.

incomes $z_{i,t+\tau}$ are observable and use all available instruments for each observation. The persistence parameter ρ is estimated in the data at a similar horizon (15 years) as the maximum span of the instruments $z_{i,t+\tau}$ I use in the estimation (see Appendix D.4 for details). This procedure yields a moderate annual persistence of $\rho = 0.90$.³⁹ 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 the next section), 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 .⁴⁰

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: a specification that controls for cash on hand; group-level specifications, similar to those in the seminal work of Attanasio and Davis (1996); and specifications using the imputed consumption measure in Blundell et al. (2008). The regression results for each of the following robustness exercises appear in Table 3. To facilitate comparison with my previous results, the first row shows baseline estimates taken from Table 1, column 3 and Table 2, columns 4 and 7.

Heterogeneous time or bequest preferences. One of the most immediate concerns one may have about the specifications in Section 4.3 is endogeneity through 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, I rerun the baseline OLS specification from Table 1, column 2 as well as the IV specification from Table 2, column 3,⁴¹ with additional controls that proxy for factors influencing savings preferences: the household head’s education, race, sex. Additionally, I include the household head’s reported bequest intention, which was elicited in the 2007 wave of the PSID.⁴²

³⁹This persistence estimate lies between the typical estimates of 0.95 – 1.00 for “restricted income profile” income processes and of 0.80 – 0.85 for “heterogeneous income profile” estimates.

⁴⁰Values of ρ beyond around 0.94 (annual), however, do generate significant weak instrument problems.

⁴¹The results are very similar for the other OLS and IV specifications.

⁴²The 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).

Table 3: Robustness checks.

	OLS with $T = 9$	IV with $\rho = 0.90$	IV with initial income
1. Baseline	0.64 (0.03)	0.70 (0.04)	0.73 (0.05)
2. Education and preference controls	0.58 (0.03)	0.69 (0.07)	0.78 (0.08)
3. Controlling for positive business wealth	0.55 (0.02)	0.68 (0.05)	0.72 (0.05)
4. HIP – education specific trends	0.59 (0.03)	0.68 (0.06)	0.73 (0.06)
5. Comprehensive consumption	0.66 (0.03)	0.71 (0.04)	0.77 (0.05)
6. Education as IV		0.73 (0.03)	
7. Total post-tax income	0.69 (0.03)	0.71 (0.04)	0.72 (0.04)
8. Nondurable consumption	0.58 (0.03)	0.59 (0.04)	0.62 (0.05)
9. Aguiar-Bils relative expenditures	0.62 (0.08)	0.72 (0.10)	0.71 (0.15)

Note. This table lists OLS and IV estimates for 8 different specifications. Row 1 shows the baseline specifications. 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. Row 5 uses the comprehensive post-2005 consumption measure. Row 6 uses education as instrument for current income residuals. Row 7 uses total post-tax household income as income measure. Row 8 uses non-durable consumption as consumption measure. Row 9 uses data on relative expenditures on luxuries vs necessities to achieve robustness to non-classical measurement error as in Aguiar Bils (2015). 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.

While these are merely proxies for preferences, one can expect that if preference heterogeneity is an important confounding factor in my specifications, some of it should be picked up by my controls.

The second row in Table 3 shows that both the OLS and IV estimates are very similar to the baseline estimates in magnitude, and the IV estimates differ insignificantly from the baseline estimates. These results seem counterintuitive at first. It is worth reiterating that they hold conditional on permanent income: unconditionally, households with a college-educated household head do save significantly more. However, as my results suggest, this is almost entirely driven by the fact that such households can “afford” to save due to higher earnings, and not because they are more frugal than their less educated counterparts.⁴³

Observe that even if unobserved preference heterogeneity were confounding my estimates, this

⁴³These results do not imply that preferences over the *type* of consumption are entirely independent of education and the other aforementioned proxies. For instance, compared to households headed by a high school dropout, I find that college-educated households tend to spend less on consumption categories such as transportation or housing, but spend significantly more on education, health and food (for a given level of permanent income).

would not resurrect the neutrality result in Section 3. I demonstrate this point in Appendix D.9 in a setting where unobserved heterogeneity in preferences is the sole cause of $\hat{\phi} < 1$.

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 et al., 2016). 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.

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.

Heterogeneous income profiles and advance information. A recent literature argues that income age profiles may have heterogeneous slopes (Guvenen, 2007, 2009; Guvenen and Smith, 2014; Guvenen et al., 2016), 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 have expected—but not by a whole lot, at least for this rough measure of heterogeneous income profiles. In fact, the IV estimate is insignificantly different from the baseline estimate.

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 $-(\rho - (\rho^*))\eta_{it+\tau-1}$, 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.

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 for each of the new consumption categories on post-2005 data (see Appendix D.3 for details).⁴⁴ The fifth row of Table 3 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. Row 6 of Table 3 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 7 of Table 3, I use a total post-tax, post-transfer income measure, which includes capital income. The results are similar.

Nondurable consumption. Row 8 of Table 3 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 (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

⁴⁴See Blundell et al. (2008) or Boar (2017) for recent examples that employ similar imputation strategies.

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 9 of Table 3 and are comparable with the baseline estimates.

4.5 Is the degree of concavity economically significant?

The evidence in the previous sections suggests that the permanent income sensitivity of consumption ϕ is around 0.70, and significantly below 1 in a statistical sense. This raises the obvious question of whether an estimate of 0.70 is also *economically* significantly different from 1. While this question is central to my quantitative general equilibrium analysis in Section 5, a simple “back of the envelope” calculation—inspired by the simple model in Section 2—already informs this question.

To operationalize this approach, suppose there is a continuum of one-period-lived households $i \in [0, 1]$, with permanent income levels y_i and an ad-hoc consumption function $\mathcal{C}(y_i) = \kappa y_i^\phi$ with some $\kappa > 0$. I now characterize how aggregate consumption $C \equiv \int \mathcal{C}(y_i) di$ depends on the distribution of permanent incomes $\{y_i\}$, holding aggregate income $Y \equiv \int y_i di$ constant.

Let $F(y)$ and $f(y)$ denote the initial cdf and pdf of the distribution of $\{y_i\}$ and $G(y)$ and $g(y)$ the new cdf and pdf, assumed to have the same mean as $F(y)$. The object of interest is the percentage change in consumption,

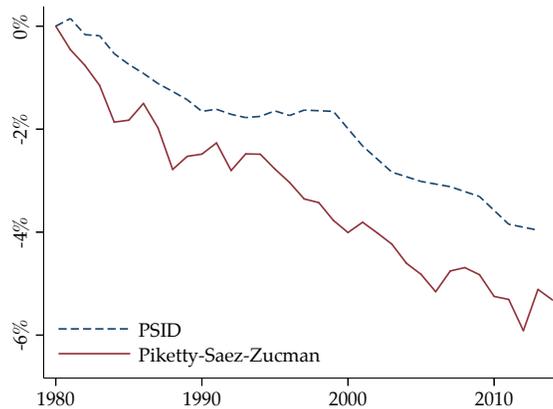
$$\frac{\Delta C}{C} = \frac{\int_0^\infty (g(y) - f(y)) \mathcal{C}(y) dy}{\int_0^\infty \mathcal{C}(y) f(y) dy}. \quad (13)$$

I evaluate expression (13) using two different data sets on the income distribution since 1980. The first is the PSID dataset used above. Here I take the income distribution to be the ± 4 year average \bar{y}_{it}^4 of residual incomes \hat{y}_{it} (net of year, age, and household size dummies), where all income observations within a nine-year window are included.⁴⁵ The second dataset is the distribution of post-tax national income by equal-split adults as computed by [Piketty et al. \(2016\)](#). For details on the calculation of $\Delta C/C$ see Appendix D.10.

Figure 5 plots the percentage annual consumption decline relative to 1980 using the two datasets. Using the PSID data I find an annual consumption shortfall due to the rise in inequality of around 3% in 2013. Using the administrative-level data, the shortfall is almost twice as large, and has reached around 5.5% in recent years. Despite the difference in magnitudes, both are on a steady downward path, with no sign of slowing down. Interestingly, 5% is the estimate obtained by Alan Krueger using an entirely different (but similarly ad-hoc) approach in a speech as Chair of the Council of Economic Advisors ([Krueger, 2012](#)). As another simple check, one can explicitly compute $\Delta C/C$ when F and G are Pareto distributions. As I illustrate in Appendix D.8, using the measured decline in the Pareto tail coefficient since 1980, this yields a decline of 5.4%, almost identical to the

⁴⁵This includes 9 income observations until 1997, and 5 income observations during the biennial survey years thereafter.

Figure 5: Annual consumption shortfall due to rising income inequality according to ad-hoc model.



Note. The figure shows simulated shortfalls in aggregate consumption based on an ad-hoc consumption function with income sensitivity $\phi = 0.7$. As income data, the dashed line uses post-tax income residuals from the PSID, averaged within households in a ± 4 year window; the solid line uses the post-tax income distribution from [Piketty et al. \(2016\)](#).

one in [Figure 5](#).⁴⁶

4.6 Taking stock

The evidence strongly rejects the linearity of consumption in permanent income. Indeed, the elasticity ϕ seems closer to 0.7 than it is to 1. There are many reasons that could be behind this, which generally fall into one of two categories: heterogeneity in transfers or bequests received and nonlinearity in Engel curves over consumption at different times.

Two cases where the first category can generate ϕ 's smaller than one are: a concave social security system; and if it were the case that poorer agents receive larger private transfers or bequests so that these households can consume a larger fraction of their measured permanent income. While the first case follows almost mechanically, and will be a core ingredient in the quantitative model, the second seems fairly implausible. In fact, it most likely goes the opposite way in the data, as it is usually richer children that inherit disproportionately larger estates.

Among explanations involving nonlinear Engel curves two of the most prominent are: non-homotheticities in consumption and bequests; and differential income risks faced by the rich vs. the poor, or their respective children. Here, non-homotheticities broadly encompass all reasons why higher-income households have a larger marginal preference for consumption or bequests later in life. This includes many things poor agents generally cannot afford, for instance, spending on college education (or gifts) for kids or grandkids, charitable giving, or expensive medical treatments. These types of expenditures are typically back-loaded and occur late in one's work life or in retirement.⁴⁷

⁴⁶Interestingly, even if all the variation in the data were driven by unobserved preference heterogeneity, the impact of rising income inequality is still very large. See [Appendix D.9](#).

⁴⁷Another channel in this category is a dependence of mortality on income which also induces richer households to

Rising income risk with permanent incomes could also be a powerful force that allows richer agents to save more. Recent evidence by [Guvenen et al. \(2016\)](#) presents a mixed picture of this issue: on the one hand, second-moment income risk generally declines with income (except at the very top of the income distribution), while left-skewness generally increases with income (except at the very top). Finally, it could also be the case that households save to insure their children against income risk, as documented in a recent paper by [Boar \(2017\)](#). Since an offspring's income risk in *levels* is likely to scale with his or her income level, this explanation is, for the purposes of this paper, similar to bequests being treated as luxury goods by parents.

I present evidence for the nonlinearity in Engel curves (the second channel) in Appendix [D.6](#): there, I show that estimates of the elasticity ϕ generally increase in age, pointing towards a more back-loaded consumption pattern; and that transfers to children—mostly to support educational expenses and home purchases—are very skewed towards high permanent income households and occur late in life, with more than half of all transfers occurring after the household head turns 70 years (conditional on survival).

5 A Non-Homothetic Life-Cycle Economy

I now investigate the quantitative implications of my empirical results. To do so, I modify the neutrality framework of Section [2](#) to allow for several forces pushing for a concave consumption function in permanent income. I study the steady state implications of the model in this section and simulate historical transitional dynamics in Section [6](#).

At the outset of this quantitative investigation it is by no means obvious in what ways the neutrality model of Section [2](#) should be modified in order to be consistent with the type of consumption and savings pattern that I documented. Indeed, several forces are discussed in Section [4.6](#) that could plausibly push the estimate ϕ into the direction observed in the data. One way to account for these forces is to model all of them separately. Yet, that quickly pushes the limits of what is currently feasible computationally and what can be calibrated independently. I therefore take a more narrow approach, in which I modify the model in as few ways as possible that still do justice to the kinds of forces at play. I focus mainly on two: a social safety net that is concave in lifetime incomes and non-homothetic preferences over consumption and bequests. As I will discuss, especially the latter is important in generating the significant deviations from the linearity of the consumption function that I found in the data.^{[48](#)}

This section proceeds as follows. In Section [5.1](#), I introduce the non-homothetic life cycle model, which then is calibrated in Section [5.2](#). In Section [5.3](#), I show the calibration results and discuss the importance of the non-homothetic savings motive. I compare my empirical estimation of ϕ in the

save more for consumption late in life.

⁴⁸It turns out not to matter for the positive predictions of the model whether some part of what is picked up as non-homothetic preferences here is in reality induced by policies, rather than preferences, such as college tuition subsidies that decline with income. Both generate the same non-homothetic savings behavior.

non-homothetic life cycle model with a variety of other precautionary-savings models in Section 5.4.

Throughout this section, I will compare all results from my non-homothetic life cycle model with an (almost) neutral homothetic model that I introduce below as well. In Appendix E, I provide details of the model computations used for this section, as well as additional results.

5.1 Model

The quantitative model is a version of the (general equilibrium) life-cycle precautionary-savings model introduced in Section 3.1, with the following modifications.

Timing. Each period or age corresponds to one year.

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 $k + 1$.⁴⁹ 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 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 . Two initial conditions are imposed for agents. First, I assume that agents start initially with zero assets and a persistent income shock η_{t_0} drawn from a normal distribution with variance σ_η^2 .⁵⁰ Second, I assume agents inherit their parents’ skill with probability $p_{inherit} \in [0, 1]$ and with probability $1 - p_{inherit}$ are assigned a random skill according to their population shares.

Government. The government levies income taxes and makes social security payments. In particular, retired agents earn social security payments $T^{socsec}(\bar{y})$ which are modeled with a piecewise linear schedule according to the Old Age and Survivor Insurance component of the Social Security system (see also Huggett and Ventura (2000) or De Nardi and Yang (2014))

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

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 state variable—I predict each agent’s lifetime income based on that agent’s last working-age income.

⁴⁹The assumption of a zero death probability before 65 is made for computational simplicity.

⁵⁰Since agents accumulate assets as they grow older, one could easily regard this economy as one where agents are “born” at age 30, rather than at age 25, with a non-trivial initial asset position (that was accumulated over the 5 previous years).

All agents pay income taxes according to a progressive income tax system. As in [Benabou \(2000\)](#) and [Heathcote et al. \(2017\)](#), I assume that pre-tax earnings y^{pre} are taxed according to

$$T^{inctax}(y^{pre}) = y^{pre} - \tau^{inctax} (y^{pre})^{1-\lambda}, \quad (14)$$

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

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

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, breaking Assumption 1. In particular, I assume preferences over consumption at age k are given by

$$u_k(c) = \frac{(c/z)^{1-\sigma_k}}{1-\sigma_k}, \quad (15)$$

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

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

where $\kappa > 0$ and $\underline{a} > 0$. These two assumptions constitute the most important deviations from the neutrality result of Section 3.4 and therefore merit an extensive discussion. Both assumptions seek to capture some of the core forces behind the empirical findings in Section 4.

Age-dependent elasticities σ_k . The key idea behind these preferences is to change income elasticities of spending across periods. To achieve this in the most straightforward and transparent way, I use “addilog” preferences that were pioneered by [Houthakker \(1960\)](#) among others. 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 elasticities ε_k are 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, and are most attractive if an agent has a sufficiently large income. 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 and assume that σ_k is lower for higher ages, capturing the fact that higher income agents seem to spend relatively more in the future, in accordance with the evidence in Section 4 and Appendix D.6. Moreover, I assume that the ratios σ_k/σ_{k+1} are constant until retirement, with a constant that will be determined by the calibration.

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 turns out to be a feature, not a bug: since homothetic models are well-known to predict *increasing* curvature in the value function with age, the non-homothetic model is able to generate curvatures much more in line with the evidence, e.g. from the age structure in portfolio allocation.⁵¹ Second, the preferences imply an elasticity of intertemporal substitution (EIS) that rises with income or wealth, in line with estimates by [Blundell et al. \(1994\)](#) 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 life-cycle economy goes back to the seminal work of [De Nardi \(2004\)](#).

Finally, analogous to the preference specification in Section 2, I assume there is a normalization constant $z > 0$ in both $u_k(c)$ and $U(a)$.

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 bequests are estimated to be in the vicinity of 5% of GDP (see Section 5.2 below), limiting their quantitative role.⁵²

Net foreign assets and net exports. Finally, I allow agents to hold foreign assets. In particular, I incorporate the net foreign asset position in a way that allows the U.S. to earn a larger return on its assets compared to its liabilities. In particular, this means that a net foreign asset position of NFA implies a net income stream that is larger than $\frac{r}{1+r}NFA$. I call the difference “external excess return”. The balance of payments at the steady state is then given by

$$0 = ExtExcessReturn + NX + \frac{r}{1+r}NFA.$$

While this part of the model is not crucial for any results, it captures returns to U.S. wealth holdings more accurately.

⁵¹[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.

⁵²The 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.

5.2 Calibration

Wherever possible, I calibrate the model 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 roughly the same time during which I estimated the concavity parameter ϕ .⁵³

Birth death skills. I assume there are $\mathcal{S} = 3$ skills with population shares of $\bar{\mu}_s$ of 0.90, 0.09, and 0.01, capturing the bottom 90%, the next 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 similar using different skill groups.⁵⁴ I calibrate mortality rates $\{\delta_k\}$ by age to the data from the Center for Disease Control for 2011.⁵⁵

Production. The gross capital share α is computed for the U.S. non-financial corporate sector for 2014. to avoid issues regarding the treatment of mixed incomes.⁵⁶ This yields $\alpha = 36.7\%$. The calibration of income shares γ_s is explained below. Depreciation is set to a value of $\delta = 0.07$ in order to match the capital-output ratio in 2014 of $K/Y = 3.05$, using a post-tax interest rate of $r = 3\%$ (see below). 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 redistributive effects without mechanical changes in Y .

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 30% of average household labor income, corresponding to around \$10,000 per adult in 2014 dollars. This is a conservative estimate following Guvenen et al. (2016). Other choices yield almost identical results. I estimate the income tax progressivity λ from PSID data on pre-tax and post-tax incomes, produced by NBER's TAXSIM program (see details in Appendix D.7). For 2013, this yields $\lambda = 0.159$. 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.⁵⁷ 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, giving approximately

⁵³As it turns out, however, the 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.

⁵⁴While 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%. A model without the 1% generates similar aggregate predictions but less wealth inequality at the top.

⁵⁵To 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).

⁵⁶In particular, I compute the gross corporate labor share as the compensation of employees divided by the gross value added net of taxes of the U.S. non-financial corporate sector using data from the BEA's National Income and Product Accounts. See also Rognlie (2015) and Barkai (2016).

⁵⁷Around 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: Elasticity of per-period utility function u by age.

Age group	25 – 44 years	45 – 64 years	65+ years
Average σ	7.2	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.

$\tau^{cap} = 40\%$. Government spending G is set to be the residual in the government budget constraint. In the calibration, this gives a value of $G/Y = 14\%$.

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

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

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 v_{it} is measurement error. Since v_{it} and ψ_{it} are indistinguishable, I follow [Heathcote et al. \(2010\)](#) and assume the variance of v_{it} is equal to $\sigma_v^2 = 0.02$.⁵⁸ 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. The labor income shares $\{\gamma_s\}$ are calibrated to match the bottom 90%, the next 9% and top 1% income shares using updated data from [Piketty and Saez \(2003\)](#). The inheritance probability of parental skill, $p_{inherit}$, is calibrated to match the slope between between parental and child income ranks measured in [Chetty et al. \(2014\)](#). This gives $p_{inherit} = 0.35$.

Preferences. As in the discussion in Section 2.1, I choose a simple parametric form for the age profile in consumption elasticities, namely a simple exponential decay, $\sigma_k/\sigma_{k+1} = \sigma^{slope} > 0$, during one’s working life and flat thereafter.⁵⁹ In addition to the slope, I pick the median elasticity $\bar{\sigma}$. I choose a standard parameter, $\bar{\sigma} = 2.5$. 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 et al. \(2017\)](#) 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: $\beta = 0.89$, $\sigma^{slope} = 0.94$, $\kappa = 16$, $\underline{a} = 1.6$. The parameter z , which is irrelevant in the homothetic case ($\sigma^{slope} = 1$), is set to 30% of

⁵⁸I also include this measurement error term in any simulated income data below.

⁵⁹I 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.

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.⁶⁰ Table 4 shows the age-dependence of σ_k and average levels of σ_k for three stages in one’s life cycle.

Interest rate r . I choose a level of $r = 3\%$ for the after-tax interest rate (before taxes this corresponds to $r^{pre} \approx 5\%$), giving rise to a private wealth to GDP ratio of around 4.2, in line with recent U.S. levels of household net worth over GDP (Piketty and Zucman, 2014).

NFA. The net foreign asset position, relative to GDP, in 2014 is -27.4% . Net exports relative to GDP are given by -2.9% .

Table 5 summarizes the calibration.

Homothetic benchmark. In addition to the non-homothetic life-cycle model I described, I also calibrate a “homothetic version” of this economy. For that, I set $\sigma_k = \bar{\sigma}$ 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 $\kappa = 1.3$. 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 neutrality benchmark in Section 3.⁶¹

5.3 Calibration results

I explore the calibrated non-homothetic and homothetic models in a number of dimensions.⁶²

Estimating ϕ in simulated data. I simulate artificial data from the non-homothetic and homothetic models and estimate the exact same regressions as those designed to test linearity in Section 4 (see Appendix E.2 for details). Table 6 shows the results, alongside the results I found in Section 4. In columns 1–2, it is visible that in both the data and the two models, the estimate without income averaging ($T = 0$) is significantly attenuated relative to the ones with income averaging, as expected from the discussion in Section 4.3. Interestingly, the $T = 4$ estimate for the homothetic model is already significantly different from the other two, 0.89 compared to 0.64 and 0.71 respectively, but still significantly below 1.

Three IV specifications for stationary η_{it} appear in columns 3–5, where—just as in Section 4.3—I estimate the regressions with quasi-differenced future incomes as instruments, treating ρ as a

⁶⁰It turns out that numerically, z is fairly aligned with β , so different values of z mainly shift the discount factor around but do not materially affect any other results.

⁶¹It is not exactly neutral since there is imperfect skill persistence and the income tax schedule is progressive, that is, Assumptions 2 and 3 are violated.

⁶²Additional calibration results can be found in Appendix E.3. There, I discuss model implications for: risk aversion over the life cycle; MPCs out of current income; life cycle profiles for consumption, income and wealth; and the joint distribution of labor income and wealth.

Table 5: Calibrated parameters of the baseline non-homothetic life cycle model.

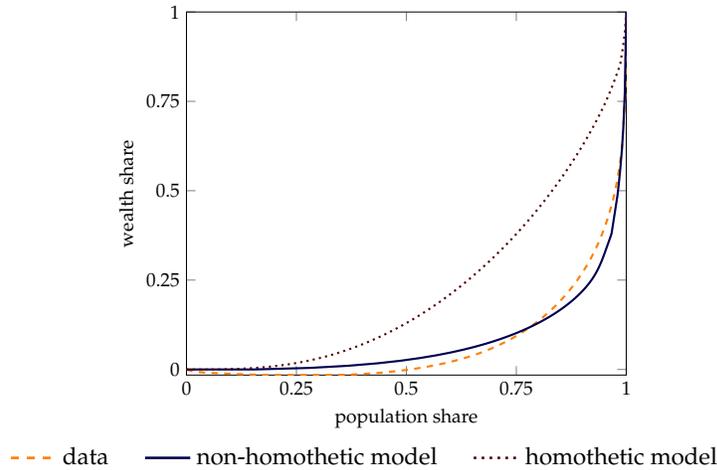
Parameter	Description	Value	Origin
<i>Birth, death, skills</i>			
S	Number of permanent types	3	see text
$\{\bar{\mu}_s\}$	Population shares by type	$\{0.9, 0.09, 0.01\}$	see text
$\{\delta_k\}$	Mortality rates by age		CDC, 2011
<i>Production</i>			
α	Capital share	0.37	NIPA, 2014
$\{\gamma_s\}$	Labor income shares	$\{0.65, 0.24, 0.11\}$	Piketty and Saez (2003), updated to 2014
δ	Depreciation	0.07	match $K/Y = 3.05$ (NIPA, 2014)
A	Total factor productivity	0.66	normalize $Y = 1$
<i>Government</i>			
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
$\bar{\lambda}$	Income tax progressivity	0.16	PSID, 2013
τ^{inctx}	Average income tax	0.30	NIPA, see text
τ^{cap}	Capital tax	0.40	NIPA, see text
G/Y	Government spending / GDP	0.14	gov. budget constraint
<i>Productivities</i>			
ρ	Income shock persistence	0.90	PSID
σ_η^2	Var. of innovations to persistent shock	0.028	PSID
σ_ψ^2	Var. of transitory income shocks	0.055	PSID
σ_v^2	Var. of measurement error in incomes	0.02	literature, see text
$p_{inherit}$	Prob. of intergen. skill transmission	0.35	Chetty et al (2014)
<i>Preferences</i>			
z	Scale term in utility function	0.30	30% of average income
β	Discount factor	0.89	match interest rate $r = 0.03$
σ^{slope}	Ratio of elasticities σ_{k+1}/σ_k	0.94	match $\phi = 0.699$
κ	Weight on bequest motive	16.06	match bequests / GDP = 0.05
\underline{a}	Intercept in bequest utility	1.62	30% share with beq. $\leq 6.25\%$ avg. income
$\bar{\sigma}$	Elast. of intertemp. substitution, median age	2.5	literature
<i>Other</i>			
NFA/Y	Net foreign asset position over GDP	-0.27	2011 US NFA, Lane and Milesi-Ferretti (2007)
NX/Y	Net exports over GDP	-0.029	US. net exports (NIPA 2014)

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

	OLS		IV			
	$T = 0$	$T = 4$	$\rho = 0.89$	$\rho = 0.9$	$\rho = 0.91$	initial income
Data	0.40	0.64	0.69	0.70	0.72	0.73
Non-homothetic	0.53	0.71	0.69	0.70	0.71	0.72
Homothetic	0.68	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.

Figure 6: 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 the distribution is. The data on wealth inequality is from [Saez and Zucman \(2016\)](#).

parameter chosen by the econometrician. The columns show a relatively good match between the regression results in the non-homothetic economy and the data. This is partly by construction: I calibrated the non-homothetic model to match the data estimate at the assumed persistence $\rho = 0.90$. The results in the homothetic economy are very close to 1. This is reassuring and a confirmation that the assumptions in Section 4.3 are indeed reasonable in the context of a life-cycle precautionary savings model. The result with the initial income IV specification (Column 6) is similar: as explained in Section 4.3, there is an upward bias in the estimated elasticity ϕ , which is visible in both models. The non-homothetic model’s estimate is relatively close to the data, while the distance to the homothetic model is considerable.

Wealth inequality. The fact that richer people save relatively more naturally generates more wealth inequality than would occur in a neutral economy. But how much more? Figure 6 shows the

Table 7: Top wealth shares.

	Share of wealth in top					
	0.1%	1%	5%	10%	20%	50%
Data (Saez-Zucman, 2016)	20.5%	38.6%	60.6%	73.0%	86.4%	100.1%
Non-homothetic model	10.1%	38.8%	67.8%	78.1%	87.0%	97.3%
Homothetic model	2.5%	11.8%	26.0%	37.5%	55.0%	87.0%

Note. This table shows the plot in Figure 6 in numbers.

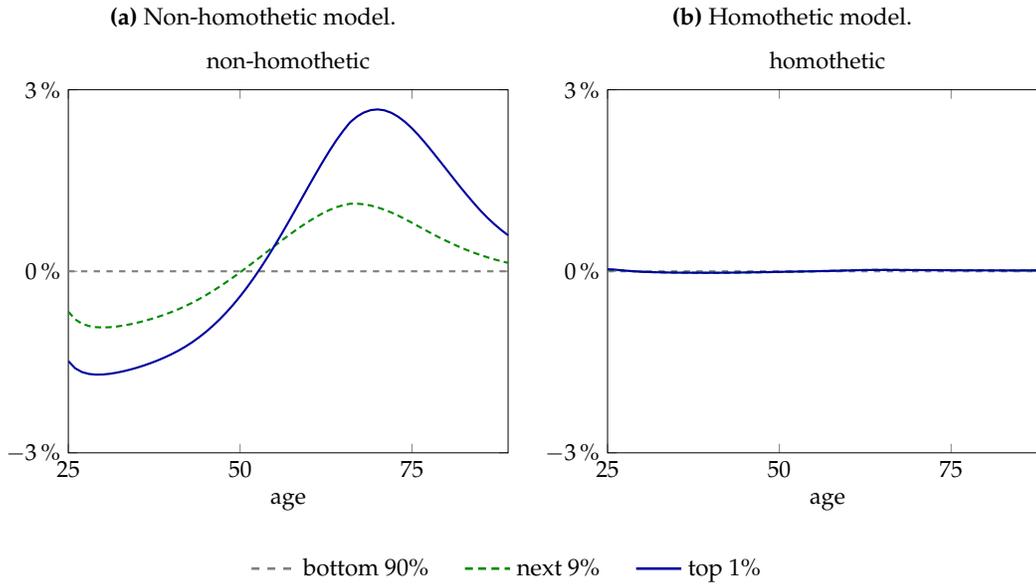
Lorenz curve for wealth or, in other words, what fraction of wealth the bottom x percent share of the population holds, where x is anywhere between 0% and 100%. As is visible, the non-homotheticity generates a significant amount of wealth inequality that matches that in the data quite successfully overall, despite wealth inequality not having been a calibration target. Table 7 confirms this but highlights that the mechanism does not seem to capture the very top. This is partly due to the fact that there is no top 0.1% skill group in the model, but may also reflect the need for other forces to explain tail inequality, such as entrepreneurship and more generally return heterogeneity, as in the models of [Quadrini \(2000\)](#), [Cagetti and De Nardi \(2006\)](#), and [Benhabib et al. \(2015, 2017\)](#), among others.

Consumption profiles. To illustrate the non-homotheticity, Figure 7 plots expenditure shares over the life cycle. The shares are constructed as the fraction of present value of consumption at a given age k , $(1+r)^{-k}\bar{c}_{k,s}$, in the total present value of income. For simplicity, Figure 7 shows the shares relative to the shares of the bottom 90% skill group. This forcefully illustrates the non-homotheticity. While in the homothetic model (Panel (b)), agents of all skills choose the exact same expenditure shares (i.e. all skill groups are on the same linear Engel curve), this is not the case in the non-homothetic model (Panel (a)): there, agents with higher permanent incomes choose to shift their present-value consumption profiles considerably more towards higher ages.

5.4 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. \(2016\)](#), or one with an extremely productive state, as in [Castaneda et al. \(2003\)](#) and [Kindermann and](#)

Figure 7: Expenditure shares (of present value income) by skill group.



Note. This graph shows age profiles of expenditure shares for each skill group in two models. The expenditure shares are computed as fractions of present value of income. The graph is normalized by subtracting the expenditure shares of the bottom 90% skill group. The plot illustrates that early consumption has a lower permanent income elasticity than late consumption in the non-homothetic model. See Appendix E.3.7 for regular *current* value consumption profiles.

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 E.4.

Table 8 shows the results. The first three rows in that table are the data, the non-homothetic economy and the homothetic economy. The alternative models considered in Table 8 split into three blocks, starting with models with alternative preference or transfer assumptions, then models with income processes that do not satisfy the assumptions of the econometric model in Section 4.3, and a few other models. The columns represent 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. If the true ϕ is equal to 1, the model is exactly neutral, as in Section 3.4. The second and third columns show the two IV specifications. I added two more specifications that have been considered by the literature and that are discussed below.

Alternative preferences or transfers. The first model in group 2 is an entirely neutral economy, with

Table 8: Comparison across models.

	true ϕ	IV, $\rho = 0.9$	IV, ini	BPP	Ret. wealth slope
<i>1. Data and main models</i>					
Data		0.70	0.73		
Non-homothetic	0.70	0.70	0.72	0.36	0.18
Homothetic	0.99	0.99	1.00	0.38	0.18
<i>2. Alternative preferences or transfers</i>					
Homothetic w/out bequests	1.00	1.00	1.02	0.36	0.24
Homothetic w/ social security	0.95	0.95	0.96	0.33	0.20
Homothetic u but luxury bequests	0.95	0.95	0.96	0.33	0.20
<i>3. Alternative income process</i>					
AR(1) with $\rho = 0.95$	1.00	0.93	1.01	0.52	0.24
Permanent-transitory	1.00	0.92	1.01	0.65	0.24
Heavy-tailed	1.00	0.90	1.00	0.38	0.20
Extreme productivity state	1.00	0.91	1.00	0.55	0.06
Heterogeneous income profiles	1.00	1.10	1.03	0.26	0.24
<i>4. Other</i>					
Partial insurance	1.00	1.00	1.04	0.25	0.24
Random discount factors	1.00	1.00	1.12	0.37	0.30

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 E.4 for details.

$\phi = 1.00$ 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} . One can see that the AR(1) IV estimate of ϕ shrinks somewhat, although only by around 0.05. The third model in group 2 is a model with a luxury bequest motive only (i.e. $\sigma_k = const$), where κ, \underline{b} are calibrated as before to match the total bequest flow in the economy and a 30% share of bequests below 6.25% of average income (see De Nardi (2004)). Again, while this helps to push down the stationary IV estimate of ϕ , it does so modestly, by around 0.05.⁶³ The initial income IV estimates are similar.

Alternative income processes. Group 3 in Table 8 considers alternative income processes. The first model considers an AR(1) persistent component η_{it} with persistence $\rho = 0.95$ the innovation variance σ_η^2 of which is calibrated to match the same overall income dispersion as the benchmark economies with $\rho = 0.90$. The second and third models consider permanent-transitory income shocks and heavy-tailed income shocks. For the former, I use the parameters in Kaplan and Violante (2010), and for the latter I adapt the income process in Guvenen et al. (2016). Both push down the stationary IV estimate, but not below 0.90. The initial income IV in those models is consistently

⁶³This 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 savings rate dispersion among the younger or middle-aged agents.

above 1. The fourth model is a simple adaption of the “extreme productivity state” income process of [Kindermann and Krueger \(2017\)](#) (see [Castaneda et al. \(2003\)](#) for the seminal work in this regard). The AR(1) IV estimate is somewhat lower. Finally, the last model in this group considers the role of heterogeneity in income profiles in a homothetic economy, calibrated to match a life-cycle increase within-cohort log income dispersion of 0.2. This tends to push up the AR(1) IV estimates, but leaves the initial income IV unchanged. It shall be noted, however, that some models with heterogeneous income profiles include a learning component ([Guvenen, 2007](#); [Guvenen and Smith, 2014](#)), which may push these estimates down somewhat.

Partial insurance economy. The first model in group 4 is a partial insurance economy, where an innovation ϵ_{it}^{η} to η_{it} is mitigated by a transfer $-\theta\epsilon_{it}^{\eta}$ ([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 8, 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.

BPP regressions. For all models, I also compute the partial insurance coefficients with respect to persistent shocks as in [Blundell et al. \(2008\)](#), henceforth BPP. One can see that they are generally relatively low, indicating a substantial amount of self-insurance with respect to permanent shocks. Interestingly, introducing a non-homothetic savings motive leaves the BPP coefficient almost identical, despite changing the permanent income elasticity parameter ϕ .

Retirement wealth slope. In an effort to inform the relationship between savings and permanent income, some authors have pointed towards the lifetime income-slope of the ratio of retirement wealth and lifetime income (see e.g. [Gustman and Steinmeier \(1999\)](#) and [Venti and Wise \(2000\)](#)). Putting aside any general issues with wealth-based savings statistics one may have,⁶⁴ the last column computes these slopes in the models. Interestingly, in all models the slope is positive, around 0.15 – 0.25 (in units of average income), even in entirely neutral models. One reason for this is that agents that receive positive income shocks late in the life cycle will not just earn relatively more due to high life cycle productivity, but also save more out of the additional income increase. This generates a positive correlation between lifetime incomes and savings when entering retirement.

⁶⁴The presence of capital gains, receipt of bequests, or heterogeneity in returns on wealth makes them relatively unattractive from the perspective of the consumption approach of this paper

Interestingly, the non-homothetic economy does not generate an outstandingly large slope here, in fact, the slope is similar to the other economies, suggesting that this slope parameter, while intrinsically interesting by itself, may not be well-suited to inform the linearity of consumption as a function of permanent income.

6 The Non-Neutral Effects of Rising Permanent Income Inequality

The previous section demonstrates the ability of the non-homothetic model to capture the empirical degree of concavity of consumption as a function of permanent income. This section explores the quantitative implications of rising permanent income inequality in the non-homothetic economy, in two steps: first, in Section 6.1, I keep the interest rate fixed and compute the partial equilibrium implications of a rise in permanent income inequality that mirrors what the U.S. has experienced since 1970. I do this mainly by comparing the 1970 and 2014 steady states with each other. In Section 6.2, I compute the general equilibrium transitional dynamics induced by the shift in the distribution of permanent income.

6.1 Comparing steady states in partial equilibrium

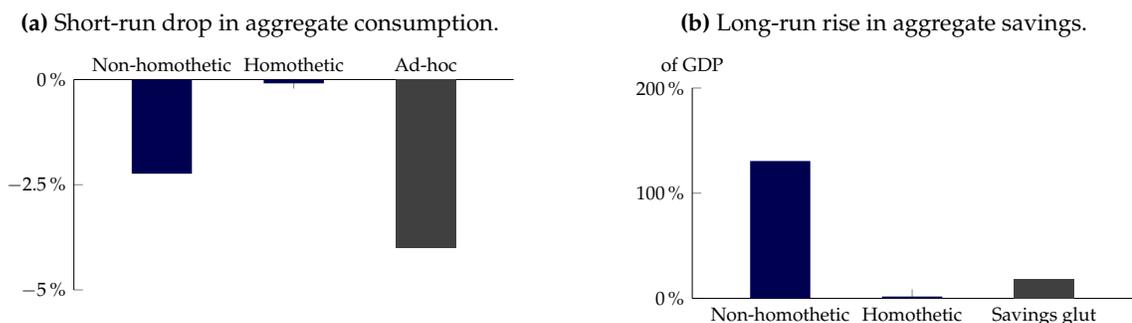
To operationalize this partial equilibrium experiment, I fix interest rates at their 2014 level but assume the distribution of labor income shares $\{\gamma_s\}$ matches the 1970 income distribution (Piketty and Saez, 2003).⁶⁵ Then, I assume from one day to the next that labor income shares $\{\gamma_s\}$ jump back to their 2014 levels. I focus on two outcomes in this subsection: the immediate shortfall in consumption—to allow for a comparison with the ad-hoc exercise in Section 4.5—as well as the long-run accumulation of wealth that is implied by the new steady state with greater inequality. I compare the results with the response in the homothetic economy (which is almost neutral in the long-run).

Short-run effect on consumption. I calculate the percentage decrease in aggregate consumption after the unanticipated shift in income inequality. In the ad-hoc exercise in Section 4.5, the consumption response was around -4% . Panel (a) of Figure 8 shows these results. In the non-homothetic model, initial consumption drops by around -2.5% in total, somewhat below the ad-hoc estimate. In the homothetic model, the number is also negative, around -0.1% , since as discussed in Section 3, the homothetic model is not neutral in the short-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 whopping 130% of GDP in additional wealth, whereas the homothetic economy has not accumulated any extra wealth. As rough order of magnitude: the U.S. net foreign asset position declined by 18% of GDP since the end of the Asian

⁶⁵As in Section 4.5, I assume for simplicity that all changes to the income distribution are driven by permanent income changes, in line with the evidence in Sabelhaus and Song (2010) and Guvenen et al. (2014, 2017).

Figure 8: 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. “Savings glut” refers to the increase in savings by foreigners in U.S. assets since 1999.

crisis in 1999. This reflects greater savings by foreigners in U.S. assets and is sometimes referred to as evidence of a “global savings glut”. Remarkably, the savings-glut driven increase in aggregate savings pales in comparison with the effects of rising income inequality.⁶⁶

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 8 is very persistent: permanently richer agents have greater savings rates, so an increase in their incomes snowballs into a large pile of additional savings.

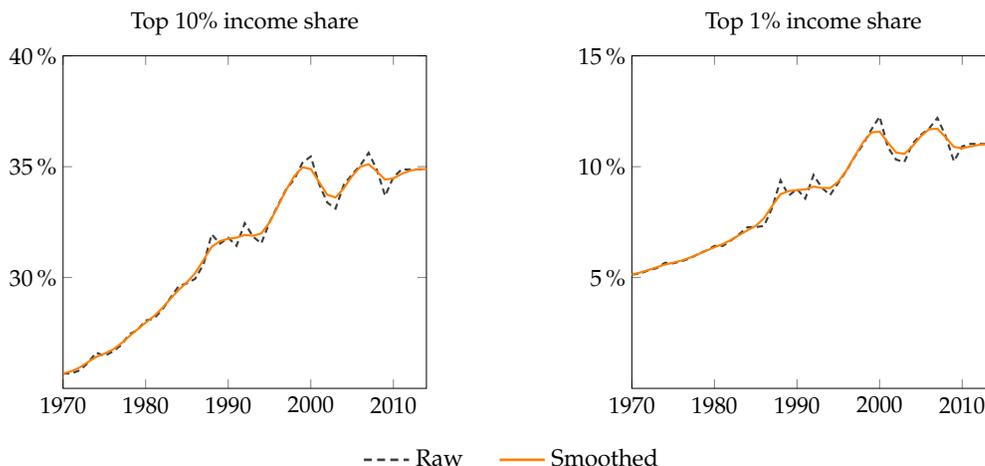
6.2 General equilibrium transitional dynamics

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. In particular, it raises two questions: how large are the effects in general equilibrium? And how long does it take for the effects to show? This section answers these questions by considering the general equilibrium (GE) transitional dynamics.

This exercise requires to overcome significant computational challenges, however. The state space S of the non-homothetic economy 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. I deal with these challenges by reducing the state space of the model to 400k states, and by improving existing algorithms along a number of dimensions. I discuss these improvements in Appendices E and F and formally define a non-stationary equilibrium in

⁶⁶The observed decline in the U.S. net foreign asset position is a general equilibrium outcome. Still, through the lens of this model, this is the right comparison: since I assume the NFA is interest inelastic in this model, any PE decline in the NFA is exactly equal to the eventual GE decline. Therefore, the 18% of GDP can be regarded as a PE decline in the NFA.

Figure 9: 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 \(2003\)](#).

Appendix E.

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 9). 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 all assets in equal proportions to their wealth.⁶⁷

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

Figure 10 depicts the model-implied transitional dynamics for these three outcomes. First, the real interest rate declines by around 1% through 2017, 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 to decline, eventually falling by another 1%. 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. In particular, this means the most able or skilled workers—one may think of recent computer science graduates—entering today will amass much larger fortunes over their lifetimes than previous generations. This effect causes a large and predictable decline in interest rates going forward.

⁶⁷While assets are interchangeable in this economy, they are subject to different valuation changes upon realization of a shock, such as rising income inequality.

Figure 10: Transitional dynamics from rising permanent income inequality.



Note. The figures show the response 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).

Second, the endogenous interest rate response limits the rise in aggregate wealth to around 30% of GDP through 2017 (again roughly one third of the rise in the data), with an eventual total increase of 55%. Here, the increase in wealth is accommodated by an increase in the quantity of capital since the model works with a standard neoclassical supply side. In a robustness exercise, I simulated the transitional dynamics for a Lucas-tree economy with a fixed capital stock K , in which any increase in wealth is necessarily accommodated by the price of capital. The resulting increase in aggregate wealth was even stronger in this case, around 50% of GDP through 2017.

Finally, the model explains almost the entire size and speed of the increase in the top 10% wealth share and around half of the increase in the top 1% wealth share. In sum, the model suggests that rising permanent labor income inequality alone can already account for a significant share of these three macroeconomic trends.

7 Conclusion

The ultimate goal of this paper is to study the implications of rising inequality in permanent income for macroeconomic aggregates, as well as for the wealth distribution. Motivated by a “neutral” model in which the distribution of permanent income is irrelevant, I propose new ways to test the proportionality of consumption and permanent income in the data, finding an elasticity of $\phi \approx 0.7$. I show 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.

There are several 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, especially when the likely consequences of targeted policies are to be informed by the model.

Second, the interaction between technology and income inequality deserves further study. In particular, by way of lowering equilibrium interest rates, income inequality can indirectly induce investment into new technologies such as automation. This, in turn, may change the labor income distribution and the capital share going forward. Such a feedback loop could make shocks to income inequality very persistent, or even permanent, in terms of convergence to a new steady state with high income and wealth inequality and low interest rates.

Finally, the non-homothetic savings behavior studied in this paper has implications beyond income inequality, for instance for connection between private debt and aggregate debt. In a canonical, homothetic, infinite horizon economy, higher levels of private debt, e.g. induced by more relaxed borrowing constraints, do not affect steady state aggregate demand: while higher debt entails larger debt payments, forcing borrowers to cut back on consumption, savers perfectly balance this shortfall. This is because both borrowers and savers react to such a permanent transfer in the same way. In contrast, in a non-homothetic economy, savers do not perfectly balance the shortfall. Private debt overhang therefore weighs negatively on aggregate demand. As long as interest rates are not against the zero lower bound, this necessitates a reduction in interest rates, which stimulates more private debt. In [Straub \(2017\)](#) I explore this nexus of private debt, aggregate demand and interest rates.

References

- Aaberge, Rolf, Anthony B. Atkinson, S. Königs, and C. Lakner**, "From Classes to Copulas: Wages capital, and top incomes," 2013.
- Abel, Andrew B. and Mark Warshawsky**, "Specification of the Joy of Giving: Insights from Altruism," *The Review of Economics and Statistics*, 1988, 70 (1), 145–149.
- Abowd, John M and David Card**, "On the Covariance Structure of Earnings and Hours Changes," *Econometrica*, 1989, 57 (2), 411–445.
- 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, Anthony B. Atkinson, Thomas Piketty, and Emmanuel Saez**, "The Top 1 Percent in International and Historical Perspective," *Journal of Economic Perspectives*, 2013, 27 (3), 3–20.
- , **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.
- Andreoni, James**, "Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence," *Journal of Political Economy*, 1989, 97 (6), 1447–1458.
- 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.
- Arellano, Manuel, Richard Blundell, and Stephane Bonhomme**, "Earnings and Consumption Dynamics: A Nonlinear Panel Data Framework," *Econometrica*, 2017, 85 (3), 693–734.
- 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 Nicola Pavoni**, "Risk Sharing in Private Information Models With Asset Accumulation: Explaining the Excess Smoothness of Consumption," *Econometrica*, 2011, 79 (4), 1027–1068.
- **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," 2017.
- Autor, David H., 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.
- Barkai, Simcha**, "Declining Labor and Capital Shares," *Job Market Paper*, 2016.
- Becker, Gary S. and Casey B. Mulligan**, "The Endogenous Determination of Time Preference," *The Quarterly Journal of Economics*, 1997, 112 (3), 729–758.
- Beckett, 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 Distribution and Social Contract," *The American Economic Review*, 2000, 90 (1), 96–129.
- , "Tax and Education Policy in a Heterogeneous Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?," *Econometrica*, 2002, 70 (2), 481–517.
- Benhabib, Jess, Alberto Bisin, and Mi Luo**, "Wealth distribution and social mobility in the US: A quantitative approach," *NBER Working Paper Series*, 2015, 21721, 1–32.
- , —, and —, "Earnings inequality and other determinants of wealth inequality," *American Economic Review*, 2017, 107 (5), 593–597.
- , —, and **Shenghao Zhu**, "The Distribution of Wealth and Fiscal Policy in Economies With Finitely Lived Agents," *Econometrica*, 2011, 79 (1), 123–157.
- and —, "Skewed Wealth Distributions: Theory and Empirics," 2017.
- 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," 2017.
- 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.
- and —, "Estate Taxation, Entrepreneurship, and Wealth," *American Economic Review*, 2009, 99 (1), 85–111.
- Cao, Dan and Iván Werning**, "Dynamic Savings Choices with Disagreements," 2016, (April), 1–100.
- 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.
- **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 Geography of Intergenerational Mobility in the United States,” *Quarterly Journal of Economics*, 2014, 129 (4), 1553–1623.
- Cochrane, John H.**, “A Simple Test of Consumption Insurance,” *Journal of Political Economy*, 1991, 99 (5), 957–976.
- 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,” 2016.
- , —, **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.
- Epstein, Larry G.**, “A simple dynamic general equilibrium model,” *Journal of Economic Theory*, 1987, 41 (1), 68–95.

- and **J. Allan Heynes**, “The Rate of Time Preference and Dynamic Economic Analysis,” 1983.
- Evans, Michael K.**, *Macroeconomic Activity: Theory Forecasting and Control*, Joanna Cotler Books, 1969.
- Fagereng, Andreas, Luigi Guiso, Davide Malacrino, and Luigi Pistaferri**, “Heterogeneity and Persistence in Returns to Wealth,” 2016.
- 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,” 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* 1957.
- Gabaix, Xavier, Benjamin Moll, Jean-Michel Lasry, and Pierre-Louis Lions**, “The Dynamics of Inequality,” *Econometrica*, 2017.
- Galor, Oded and Omer Moav**, “From Physical to Human Capital Accumulation,” *The Review of Economic Studies*, 2004, 71 (4), 1001–1026.
- Gourieroux, C., A. Monfort, and E. Renault**, “Indirect Inference,” *Journal of Applied Econometrics*, 1993, 8, S85–S118.
- Gourinchas, Pierre-Olivier and Jonathan A. Parker**, “Consumption over the life cycle,” *Econometrica*, 2002, 70 (1), 47–89.
- Gustman, Alan L. and Thomas L. Steinmeier**, “Effects of Pensions on Savings: Analysis with Data from the Health and Retirement Survey,” 1999.
- 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?,” 2016.
- , **Greg Kaplan, Jae Song, and Justin Weidner**, “Lifetime Incomes in the United States Over Six Decades,” 2017.
- , **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.
- and **Frederic S. Mishkin**, “The Sensitivity of Consumption to Transitory Income: Estimates from Panel Data on Households,” *Econometrica*, 1982, 50 (2).
- Harris, Christopher and David Laibson**, “Dynamic Choices of Hyperbolic Consumers,” *Econometrica*, 2001, 69 (4), 935–957.

- 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 —, “Consumption and labor supply with partial insurance: An analytical framework,” *American Economic Review*, 2014, 104 (7), 2075–2126.
- , —, 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,” 2017.
- Houthakker, Hendrik S.**, “Additive Preferences,” *Econometrica*, 1960, 28 (2), 244–257.
- Hryshko, Dmytro and Iouri 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.
- Imrohoroglu, Aye, Selahattin Imrohoroglu, and Douglas H Joines**, “A life cycle analysis of social security,” *Economic Theory*, 1995, 6 (1), 83–114.
- 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.
- 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, Bari 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* 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.
- Koopmans, Tjalling C.**, “Stationary Ordinal Utility and Impatience,” *Econometrica*, 1960, 28 (2), 287–309.
- Kopczuk, Wojciech, 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, Alan B.**, “The Rise and Consequences of Inequality in the United States,” Technical Report 2012.
- 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.
- 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,” 2015.
- Lucas, Robert E. Jr. and Nancy L. Stokey**, “Optimal growth with many consumers,” *Journal of Economic Theory*, 1984, 32, 139–171.
- 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**, “The Propensity to Consume Permanent Income,” *American Economic Review*, 1966, 56 (5), 1158–1177.
- , *Permanent Income, Wealth, and Consumption* 1972.
- Meghir, Costas and Luigi Pistaferri**, “Earnings, Consumption and Lifecycle Choices,” 2010.
- Moav, Omer**, “Income Distribution and Macroeconomics: The Persistence of Inequality in a Convex Technology Framework,” *Economics Letters*, 2002, 75, 0–9.
- Obstfeld, Maurice**, “Intertemporal dependence, impatience, and dynamics,” *Journal of Monetary Economics*, 1990, 26 (1), 45–75.
- Ömer T. Acikgöz**, “On the existence and uniqueness of stationary equilibrium in Bewley economies with production,” *Journal of Economic Theory*, 2017, (October).

- Parker, Jonathan A., Nicholas S. Souleles, David S. Johnson, and Robert McClelland**, "Consumer spending and the economic stimulus payments of 2008," *American Economic Review*, 2013, 103 (6), 2530–2553.
- Piketty, Thomas**, *Capital in the twenty-first century*, Harvard University Press, 2017.
- **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 Ratios 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.
- **, Emmanuel Saez, and Gabriel Zucman**, "Distributional national accounts: methods and estimates for the United States," *NBER Working Paper*, 2016, (December), 1–53.
- 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**, "Understanding the U.S. Distribution of Wealth," *Federal Reserve Bank of Minneapolis Quarterly Review*, 1997, 21 (2), 22–36.
- **and —**, *Inequality in macroeconomics*, 1 ed., Vol. 2, Elsevier B.V., 2015.
- Ray, Debraj and Arthur Robson**, "Status, Intertemporal Choice, and Risk-Taking," *Econometrica*, 2012, 80 (4), 1505–1531.
- Ríos-Rull, José-Víctor**, "Life-Cycle Economies and Aggregate Fluctuations," *Review of Economic Studies*, 1996, 63, 465–489.
- Robson, Arthur**, "Status, the Distribution of Wealth, Private and Social Attitudes to Risk," *Econometrica*, 1992, 60 (4), 837–857.
- 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 "Cooley, T.F. (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.
- Scholz, John Karl, Ananth Seshadri, and Surachai Khitatrakun**, "Are Americans Saving " Optimally " for Retirement?," *Journal of Political Economy*, 2006, 114 (4), 607–643.
- Smith, Anthony A. Jr.**, "Estimating nonlinear timeseries models using simulated vector autoregressions," *Journal of Applied Econometrics*, 1993, 8, S63–S84.

—, “Indirect Inference,” *The New Palgrave Dictionary of Economics*, 2008, 2, 1–9.

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.

Straub, Ludwig, “Private Debt Traps,” 2017.

Strotz, Robert H., “Myopia and inconsistency in dynamic utility maximization,” 1955.

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

Townsend, Robert M., “Risk and Insurance in Village India,” *Econometrica*, 1994, 62 (3), 539–591.

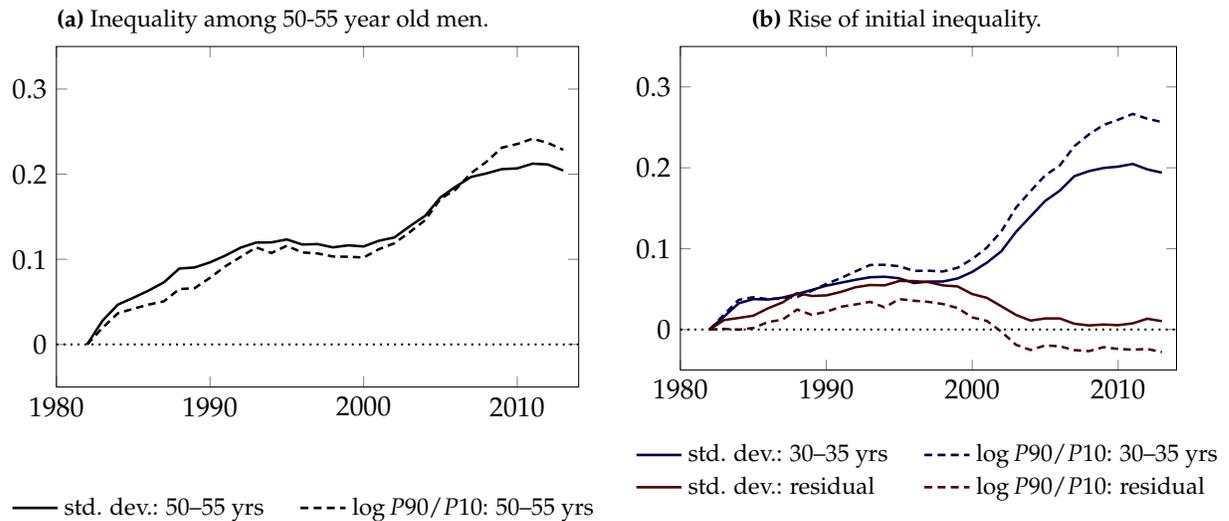
Uzawa, Hirofumi, “Market Allocation and Optimum Growth,” *Australian Economic Papers*, 1968.

Venti, Steven F. and David A. Wise, “Choice, Chance and Wealth Dispersion at Retirement,” 2000, (February).

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.

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. \(2017\)](#).

Appendix

A The importance of permanent income inequality

In a recent paper, [Guvenen et al. \(2017\)](#) 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.⁶⁸ 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

⁶⁸I 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.

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 $\sigma = 1.5$, 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 scaling parameter z is chosen so that the economy is in equilibrium. This procedure implies $\beta = 7.2$ and $z = 3.2$. (Notice that β has the role of what usually is $\beta/(1 - \beta)$, which explains why it is so large.)

Figure 1. Combining the Euler equation (4) with the budget constraint (2), the consumption policy function $c(a + w)$ can be found by solving the implicit equation

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

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 $c = (1 - R^{-1})a + w$,

$$a = z (\beta R)^{1/\Sigma} \left(\frac{(1 - R^{-1})a + w}{z} \right)^{\sigma/\Sigma}. \quad (16)$$

Figure 2. At the equality steady state, per-head wealth levels of poor and rich are equal, $a^r = a^p \equiv a = RK$. Total PE consumption after a shift in income inequality is then given by $C(\gamma) = \mu c(a + \gamma W/\mu) + (1 - \mu)c(a + (1 - \gamma)W/(1 - \mu))$, where $\gamma \in [\mu, 1]$. Aggregate wealth after 20 years is constructed as the sum of both types' asset positions after 20 years, which in turn are obtained by iterating the asset policy function given a distribution of wages.

Figure 3. 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 (16) 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 3 is standard.

C Omitted proofs

C.1 Proof of Lemma 1

The key idea behind the proof of this result is that the dynamic programming problem (5) 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

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

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

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

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 (6). Notice that for any measurable set $\mathbf{A} \subset \mathbb{R}_+$ it holds that

$$\tilde{v}(\mathbf{A}, \varphi' | s', k, \varphi) = \tilde{v}\left(\frac{w_s}{w_{s'}} \mathbf{A}, \varphi' | s, k, \varphi\right) \quad (17)$$

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 (5) with anticipated bequest distribution \tilde{v} . Denote the corresponding value function by \tilde{V} . It must be that

$$\tilde{V}_{k,s'}(a, z, \varphi) = \left(\frac{w_{s'}}{w_s}\right)^{1-\sigma} \tilde{V}_{k,s}\left(\frac{w_s}{w_{s'}} a, z, \varphi\right) \quad (18)$$

for the following reasons: if \tilde{V} were not the (unique) solution to the convex programming problem (5), one could easily use (18) to construct a second solution. \tilde{V} solves (5) due to Assumptions 1 and 2, the fact that the agent starts with zero assets, and equation (17). Building on (18), it is immediate that the unique policy functions satisfy

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

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 (7) given \tilde{v} and $\tilde{\mathbf{a}}$ yields

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

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.

C.2 MPCs in the neutral economy

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

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

As it turns out, the distribution of MPCs is the same across skill groups.

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

$$\mathbf{MPC}_{k,s}(a, z, \varphi) = \mathbf{MPC}_{k,s'} \left(\frac{w_{s'}}{w_s} a, z, \varphi \right). \quad (19)$$

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

Proof. Equation (19) is a direct consequence of the definition of MPCs and Lemma 1,

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

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,

$$\mu(k, \mathbf{A}, z, \varphi | s) = \mu \left(k, \mathbf{A} \frac{w_{s'}}{w_s}, z, \varphi | s' \right). \quad (20)$$

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 (19) and (20). □

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.

C.3 Proof of Proposition 1

From Lemma 1 it follows that

$$\log \mathbf{c}_{k,s}(aw_s, z, \varphi) - \log w_s = \log \mathbf{c}_{k,s'}(aw_{s'}, z, \varphi) - \log w_{s'}. \quad (21)$$

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 ,

$$\mu(w_s \mathbf{A}, z, \varphi | s, k) = \mu(w_{s'} \mathbf{A}, z, \varphi | s', k). \quad (22)$$

Using that notation, define

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

which is well-defined by (22). By definition of const_k , one can then write

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

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

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

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

Equation (9) follows by definition of income as

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

where again $\log \Theta_k(z)$ can be further decomposed into an average and a mean zero term. This gives (9).

C.4 Proof of Proposition 2

Using (21), one can decompose

$$\log \mathbf{c}_{k,s}(a, z, \varphi) = \log w_s + \mathcal{C} \left(k, \frac{a}{w_s}, z, \varphi \right). \quad (23)$$

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

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

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

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

For the same reason, one can write

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

with some other constant const' . Together,

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

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

C.5 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

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

With decomposition (23), this becomes

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

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,

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

Table 9: More robustness checks.

	OLS with $T = 9$	IV with $\rho = 0.90$	IV with initial income
1. Baseline	0.64 (0.03)	0.70 (0.04)	0.73 (0.05)
2. Incl. household heads > 65 years	0.68 (0.03)	0.71 (0.04)	0.73 (0.04)
3. Controlling for liquid assets	0.63 (0.03)	0.74 (0.05)	0.70 (0.05)
4. Households with original sample heads	0.61 (0.04)	0.66 (0.07)	0.76 (0.06)

Note. This table lists OLS and IV estimates for 5 different specifications. Row 1 shows the baseline specifications. Row 2 shows results for households with heads of all ages. Row 3 controls for a cubic in liquid assets relative to income. Row 4 runs the baseline OLS and IV specifications on the subsample with only the households headed by original PSID sample heads. 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.

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

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

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

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. \(2008\)](#).

Additional household-level specifications. All results are shown in [Table 9](#). As comparison, the first row contains the baseline estimates from [Table 3](#). The second row 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.⁶⁹ The results are similar as in the baseline case. The third row adds a

⁶⁹Since labor income is ill-defined for most households above 65, I use total household income, including income from capital for this regression.

Table 10: Group-level specifications.

	log group consumption	
	(1) OLS	(2) IV
log group income	0.642 (0.04)	0.720 (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.

cubic 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 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. The fourth row re-estimates the two 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.

Group-level specifications. In Table 10, I show estimates for a number of 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 (11). Specifically, let $\{I_g\}$ be a set of 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

$$\bar{c}_{gt} = \mathbf{X}'_{gt} \beta + \phi \bar{y}_{gt} + \epsilon_{gt}. \quad (24)$$

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

Table 11: 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 (0.06)	0.645 (0.03)	0.569 (0.06)	0.566 (0.06)	0.562 (0.06)	0.557 (0.07)	0.551 (0.07)	0.784 (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.

impute various measures of consumption expenditure for the PSID from 1980 to 1992.⁷⁰ Among the measures of consumption expenditure imputed is total consumption expenditure, which I will be using below. The results and conclusions are very similar with the other measures the authors impute, namely a measure of nondurable consumption expenditure and measures that include the service flow of durables rather than the expenditure.

This consumption data is especially 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 11](#) 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 11](#) 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.

⁷⁰I thank the authors for making their data available online.

D.2 Robustness to non-classical measurement error

In a recent paper, [Aguiar and Bills \(2015\)](#) 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 Bills \(2015\)](#) 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 Bills \(2015\)](#).

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 Bills \(2015\)](#) in assuming: that true good-level expenditures can be represented by a log-linear demand system,⁷¹

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

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 households 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

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

with $\check{\beta}_g = (\beta_g - \beta_{g_0}) / (\gamma_g - \gamma_{g_0})$ and $\check{\epsilon}_{igt} = (\epsilon_{igt} - \epsilon_{ig_0t}) / (\gamma_g - \gamma_{g_0})$. To implement this regression, I use the estimated expenditure elasticities in [Aguiar and Bills \(2015, Table 2, column I\)](#) for the following categories which are (more or less) 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 3.

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

⁷¹This is not possible globally, see the discussion in [Aguiar and Bills \(2015\)](#).

according to the reduced 70% measure, and by \mathbf{c}_{it} household i 's consumption according to the comprehensive post-2005 measure. Label by $k = 1, \dots, 6$ 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.

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 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

Table 12: Estimated AR(1) + iid process.

ρ	σ_α^2	σ_η^2	σ_ψ^2
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.

distinguished from the transitory income shock. I follow [Heathcote et al. \(2010\)](#) and assume that $\sigma_v^2 = 0.02$. This leaves me with four parameters to estimate: σ_α^2 , ρ , 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 \(1984\)](#).⁷² 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

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

for any $s \geq 0$. As mentioned before, σ_v^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

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

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 $\mathbf{G}(\sigma_\alpha^2, \rho, \sigma_\eta, \sigma_\psi^2)$. The minimum distance estimates of $(\sigma_\alpha^2, \rho, \sigma_\eta, \sigma_\psi^2)$ are then the solution to

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

where \mathbf{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 $(\sigma_\alpha^2, \rho, \sigma_\eta, \sigma_\psi^2)$, whose asymptotic standard errors I compute using a block-bootstrap with 500 iterations that is clustered at the household level. [Table 12](#) shows the results.

⁷²See [Meghir and Pistaferri \(2010\)](#) for a recent survey over the literature.

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 $s > 0$,

$$\log(m_{s+1} - m_s) = \text{const} + s \log \rho. \quad (25)$$

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 below.

Needless to say, there is an active debate about the “right” income process. The simple estimation strategy in (25) 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 Section D.6, an estimated ρ around 0.90 fails to generate the almost linear shape of the income dispersion over the life-cycle.

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 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.⁷³ 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.

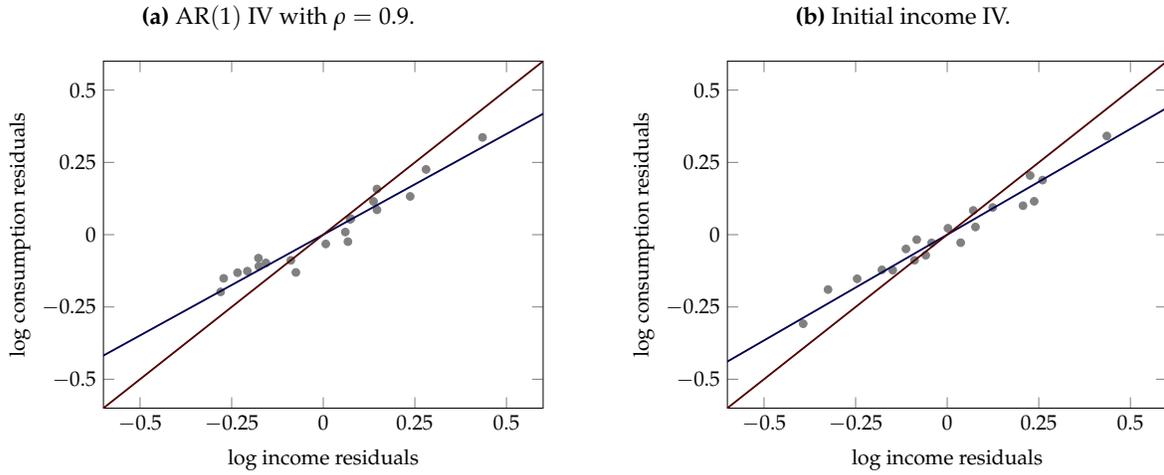
D.6 Evidence on the mechanism

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. As I explain in Section 4.6, however, there are several other plausible channels.

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.

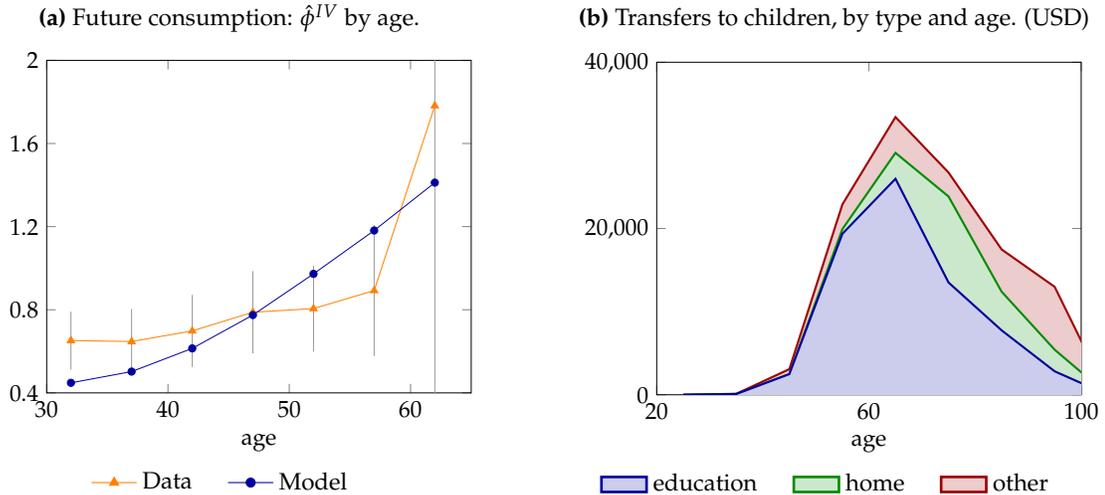
⁷³See the definitions in Section 4.3.

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.

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



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.

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 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.

D.7 Estimation of the income tax progressivity

My life cycle model in Section 5 requires an estimate of tax progressivity. To this end, I follow Benabou (2000) and Heathcote et al. (2017) and assume in a given year t , total post-tax incomes are a power function of pre-tax incomes,

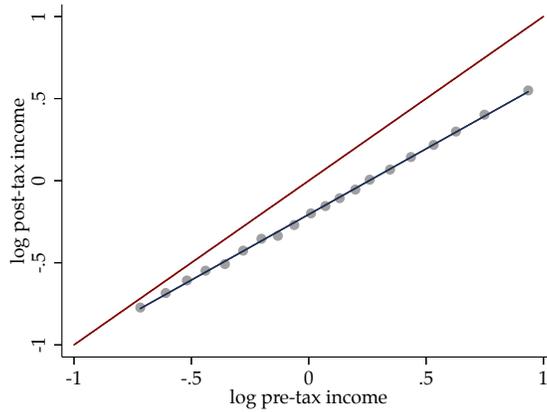
$$y^{posttax} = \tau_t^{inctax} (y^{pretax})^{1-\lambda_t} \quad (26)$$

where τ_t^{inctax} is a constant and $\lambda_t \in [0, 1]$ is the tax progressivity parameter (see (14)). Clearly, the closer λ_t is to 1, the more extremely high incomes are taxed. On its face, the power function seems like an arbitrary assumption, yet as shown in Figure 14 the fit is remarkable (see also Heathcote et al. (2017)): Figure 14 is a binned scatter plot for the year 2011, plotting $\log y^{posttax}$, as computed using NBER's TAXSIM program, against $\log y^{pretax}$ in 20 bins. The sample as in Section 4.1, only that I restrict it to households whose head is less than 65 years old (i.e. working age). I further exclude households with very low pre-tax household incomes, namely less than 50% of the average pre-tax income in that year.⁷⁴

As part of their data preparation, Heathcote et al. (2017) adjust pre-tax earnings by adding employers' shares (50%) of social security and medicare taxes. To be consistent with their adjustment, I scale $\log y^{pretax}$ equivalently in all years so that my estimate for λ over the same survey years 2001–2007 as theirs yields the same estimate of $\lambda = 0.181$. In the model of Section 5 I use the estimated elasticity for 2013, which is $\lambda = 0.16$.

⁷⁴Both in the data and in my model in Section 5, there are other support mechanisms that are active for lower incomes and distort the power law (26) upwards, pushing $y^{posttax}$ above the predicted value from the power law.

Figure 14: Estimating income tax progressivities.



Note. The figure shows a binned scatter plot with 20 bins of log total post-tax household incomes against total log total pre-tax household incomes for PSID households with a working-age head in 2011. Household incomes below 50% of the average income have been excluded. The blue line is the linear fit, the red line is the 45° line.

D.8 Analytical example for the economic significance of the concavity

To illustrate the economic significance of the estimated income elasticity of consumption, $\phi = 0.7$, I derive a simple example for the case where the income distribution follows an exact Pareto distribution. Denote by $F(y)$ the cdf and by $f(y)$ the pdf of a Pareto distribution with tail parameter $\xi > 0$ and lower bound $\underline{y} > 0$. Since the mean of the income distribution is kept constant under comparative statics, aggregate consumption is given by

$$C = \int_{\underline{y}}^{\infty} f(y)(y/\bar{y})^{\phi} dy.$$

Using that $f(y) = \xi y^{-\xi-1} \underline{y}^{\xi}$ this can be shown to yield

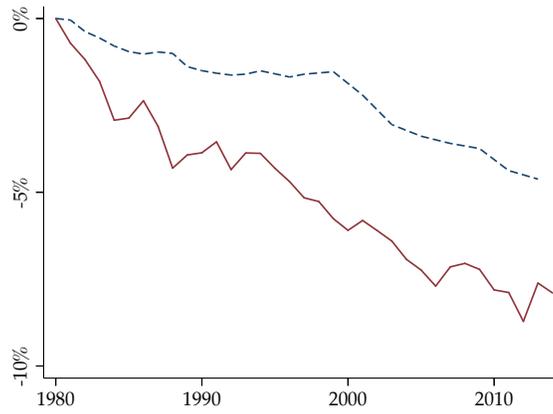
$$C = \left(\frac{\xi - 1}{\xi} \right)^{\phi} \frac{\xi}{\xi - \phi}. \quad (27)$$

As I explain in Section D.10 below, I estimate Pareto tails using post-tax income data from [Piketty et al. \(2016\)](#) and the PSID. For example, for the data from [Piketty et al. \(2016\)](#), I find a decline in the Pareto tail parameter from $\xi \approx 2.17$ in 1980 to $\xi \approx 1.66$ at the end of the sample period. This yields a percentage decline in C of $\Delta C/C \approx 5.4\%$, very much in line with [Figure 5](#).

D.9 Ad-hoc model with pure preference heterogeneity

In this section, I consider the extreme case where the entire variation in [Section 4](#) is driven by unobserved preference heterogeneity. I argue—using the ad-hoc model of [Section 4.5](#)—that even in this case, under plausible assumptions, the effects of rising income inequality are significant.

Figure 15: The ad-hoc model with pure preference heterogeneity.



Note. This figure is the analog of Figure 5 for the case of an ad-hoc linear consumption function with preference heterogeneity. As income data, the dashed line uses post-tax income residuals, averaged within households in a ± 4 year window; the solid line uses the post-tax income distribution from Piketty et al. (2016).

Preference heterogeneity is therefore no guarantee for neutrality.

Thus, suppose the empirical relationship between consumption and permanent income is entirely driven by unobserved preference heterogeneity and not by a concave consumption function in permanent income. In particular, each agent i has a consumption function $C_i(y) = k_i y$ with idiosyncratic shifter $k_i > 0$. So, higher i agents not only have larger incomes but are also more patient, that is, k_i is declining in i . To generate the empirical relationship between $C_i(y_i)$ and y_i , one would need $k_i = k y_i^{\phi-1}$ for some common $k > 0$. Using the same notation for cdfs F and G as before, the change in consumption is now given by

$$\frac{\Delta C}{C} = \frac{\int_0^{\infty} (G^{-1}(F(y)) \cdot y^{-1} - 1) C(y) f(y) dy}{\int_0^{\infty} C(y) f(y) dy}.$$

How large is this quantitatively? Figure 15 illustrates this for the same two income datasets that were used in Section 4.5.⁷⁵ It is evident that the assumption of pure preference heterogeneity does not necessitate zero effects from changes in the income distribution. In fact, as Figure 15 illustrates, the effects may be even larger.

D.10 Construction of Figures 5 and 15

Figures 5 and 15 show $\Delta C/C$ for the concave permanent consumption function case and the pure preference heterogeneity case. $\Delta C/C$ is computed by computing a measure of consumption C for each year that holds aggregate income constant. Then, $\Delta C/C$ is just the percentage change of the consumption measure in any given year relative to 1980.

⁷⁵For details on the calculation of $\Delta C/C$ see Appendix D.10.

Figure 5. Denote the cdf of the income distribution in a given year by $F(y)$, and its mean by \bar{y} . As before, suppose F admits a Lebesgue measurable density $f(y)$. The formula for C is then given by

$$C = \int_0^\infty f(y)(y/\bar{y})^\phi dy.$$

Since I use data on income shares to compute C , I now explain how to rewrite this expression entirely in terms of top income shares

$$\bar{s}(q) \equiv \bar{y}^{-1} \int_{F(q)}^\infty f(y)y dy \quad (28)$$

which is a number in $[0, 1]$ for each quantile $q \in [0, 1]$. I call it “top income share” because it is the income share of all individuals in the top $1 - q$ of the population, or equivalently, one minus the Lorenz curve for income. From its definition (28) it follows that $\bar{s}'(q) = -\bar{y}^{-1}F^{-1}(q)$ and thus C can be expressed as

$$C = \int_0^1 (-\bar{s}'(q))^\phi dq.$$

For a distribution with Pareto tail with tail parameter ξ and lower bound \underline{y} , $F(y) = 1 - (y/\underline{y})^{-\xi}$ for large y , and

$$\bar{s}(q) = \frac{1}{\bar{y}} \frac{\xi}{\xi - 1} \underline{y}^\xi \left(\underline{y}(1 - q)^{-1/\xi} \right)^{1-\xi} = \frac{\xi}{\xi - 1} \frac{\underline{y}}{\bar{y}} (1 - q)^{(\xi-1)/\xi} \quad (29)$$

for q close to 1. Notice that \bar{y} is only equal to $\frac{\xi}{\xi-1}\underline{y}$ if F is an exact Pareto.

In the data, I only observe $\bar{s}(q)$ on an equally spaced grid with spacing $\Delta q = 0.01$. Denote the grid points by q_i for $i = 0, \dots, 100$. I approximate C by numerically approximating the integral until a cutoff quantile q_I , $I \in \{0, \dots, 100\}$ and approximating the tail after q_I by a Pareto distribution, such that

$$C \approx \sum_{i=0}^{I-1} \left(\frac{\bar{s}(q_i) - \bar{s}(q_{i+1})}{\Delta q} \right)^\phi \Delta q + \left(\frac{\underline{y}}{\bar{y}} \right)^\phi \frac{(1 - q_I)^{1-\phi/\xi}}{1 - \phi/\xi},$$

where the second term is the Pareto tail integral. I set $I = 90$, but the result is very robust to other reasonable choices of I . I estimate ξ and \underline{y}/\bar{y} by regressing $\log \bar{s}(q_i)$ on $\log(1 - q_i)$ for $i = I, I + 1, \dots, 100$, which according to (29) recovers $(\xi - 1)/\xi$ as slope parameter and $\log \left(\xi / (\xi - 1) \underline{y} / \bar{y} \right)$ as intercept.

Figure 15. In the case with pure preference heterogeneity, the expression for C depends both on the income distribution of the year for which C is being calculated and the income distribution that was present at the time ϕ was being measured. I denote by $F(y)$ the income distribution at the time of measurement of ϕ and by $G(y)$ the income distribution in the year for which C is being calculated. As before, I observe the top income shares $\bar{s}_F(q)$ and $\bar{s}_G(q)$ for both distributions, I measure their respective Pareto tail coefficients ξ_F, ξ_G and ratios $\underline{y}_F/\bar{y}_F, \underline{y}_G/\bar{y}_G$ as outlined above, and I pick a threshold quantile q_I .

In this case, C can be written as

$$C = \int_0^1 (-\bar{s}'_F(q))^{\phi-1} (-\bar{s}'_G(q)) dq$$

which is approximately

$$C \approx \sum_{i=0}^{I-1} \left(\frac{\bar{s}_F(q_i) - \bar{s}_F(q_{i+1})}{\Delta q} \right)^{\phi-1} (\bar{s}_G(q_i) - \bar{s}_G(q_{i+1})) + \left(\frac{y_F}{\bar{y}_F} \right)^{\phi-1} \left(\frac{y_G}{\bar{y}_G} \right) \frac{(1-q_I)^{1+(1-\phi)/\xi_F-1/\xi_G}}{1+(1-\phi)/\xi_F-1/\xi_G}.$$

I choose $I = 90$ as before and assume the year where ϕ was measured in the data is the midpoint of the sample in which I measured it, which is 2007. This concludes the construction of Figure 15.

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. In Section E.3, I discuss a number of additional model implications. Finally, Section E.4 goes over the details of the alternative models mentioned in Table 8.

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. Here, I allow for non-stationary environments so as to nest the transitional dynamics exercise of Section 6.

E.1.1 Household maximization

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

$$V_{k,s,t}(a, z, \varphi) = \max_{\{c, a'\}} u_k(c) + \beta(1 - \delta_k) \mathbb{E}_{z, \varphi} V_{k+1, s, t+1}(a' + b', z', \varphi') + \beta \delta_k U(a') \quad (30)$$

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

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

$$a' \geq 0.$$

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

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

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

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

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

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

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,

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

where I normalized the skill endowments by group to 1 since in L_s is inelastically supplied at $\bar{\mu}_s$. I assume that the representative firm maximizes profits subject to adjustment costs $\zeta(\cdot)$,

$$J_t(K_-) = \max_{d, K, I} \left\{ d + \frac{1}{1+r_t} J_{t+1}(K) \right\} \quad (31)$$

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

$$K = K_- - \delta K_- + I$$

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 $W_t = \sum_s \bar{\mu}_s w_{s,t}$.

Adjustment costs. I choose a standard quadratic adjustment cost function $\zeta(x) = \frac{1}{4\delta} \frac{x^2}{2}$ as for instance in [Auclert and Rognlie \(2017\)](#).

External excess returns. As explained above, there can be non-trivial external excess returns in the model, which need to be included in total wealth held by domestic agents. I define this recursively as

$$J_t^{ext} = ExtExcessReturn_t + \frac{1}{1+r_t} J_{t+1}^{valuation},$$

where excess returns follow from the balance of payments

$$0 = ExtExcessReturn_t + NX + \frac{r_t}{1+r_t} NFA.$$

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 (30) 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 (31) given $\{r_t, w_{s,t}\}$,
(c) the government sets government spending G_t according to its budget constraint

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

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

(d) the goods market clears,

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

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

$$A_t \equiv \int a d\mu_t = (1+r_t)B + J_t(K_t) + J_t^{ext} + NFA,$$

(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

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

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 background materials by [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,

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

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 $\{v_t(b', \varphi' | s, k, \varphi)\}_t$. 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,⁷⁶ and the transitory shock on 3 states. In addition to 65 age states, 3 skill groups, and a number of inheritance states to keep track whether a household has made an inheritance already, this leaves me with 8712 idiosyncratic states, and ≈ 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

⁷⁶The [Tauchen \(1986\)](#) method is generally less precise than the [Rouwenhorst \(1995\)](#) method for persistent income processes ([Kopecky and Suen, 2010](#)).

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 four parts: the net foreign asset position NFA , government bonds B , financial wealth in equities $v = (F_K K - \delta K)/r$, and financial wealth in claims on external excess returns. 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,

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

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.⁷⁷ After simulating the data, all income observations are multiplied by the measurement error term $\exp\{v_{it}\}$, 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).

In addition to those three specifications, I also run additional specifications for comparison across models in Table 8.

BPP specification. I replicate a (simplified) version of the approach in Blundell et al. (2008), henceforth BPP, as follows. Conditional on simulated data on log labor income residuals \hat{y}_{it} and log

⁷⁷I also experimented with many iterations of smaller samples to ensure there are no small sample biases.

consumption *residuals* \hat{c}_{it} (both after partialing out age dummies), BPP estimate the specification⁷⁸

$$\begin{aligned}\hat{y}_{it} &= \eta_{it} + \psi_{it} \\ \eta_{it} &= \eta_{it-1} + \epsilon_{it}^{\eta} \\ \psi_{it} &= \epsilon_{it}^{\psi} + \theta \epsilon_{it-1}^{\psi} \\ \Delta \hat{c}_{it} &= \phi^{BPP} \epsilon_{it}^{\eta} + \chi \epsilon_{it}^{\psi} + \zeta_{it}.\end{aligned}$$

Here, $\{\epsilon_{it}^{\psi}, \epsilon_{it}^{\eta}, \zeta_{it}\}$ are iid shocks, mutually independent of each other. In words, BPP assume a permanent-transitory income process, with the transitory component being modeled as $MA(1)$ process, and evaluate how much of the permanent income innovation ϵ_{it}^{η} is “passed through” to consumption; this is captured by the coefficient ϕ^{BPP} ; and how much of the transitory shock ψ_{it} is passed through; this is captured by χ . Here, I focus only on ϕ^{BPP} . Following the reasoning in BPP’s Appendix C and extending the logic to a $MA(1)$ process for transitory income shocks, ϕ^{BPP} is identified in the above model as the IV estimate of a regression of $\Delta \hat{c}_{it}$ on $\Delta \hat{y}_{it}$, using $\hat{y}_{it+2} - \hat{y}_{it-3}$ as instrument.

Retirement wealth slope. Finally, I implement a specification based on retirement wealth that has been used to test for savings non-homotheticity before, see e.g. [Gustman and Steinmeier \(1999\)](#); [Venti and Wise \(2000\)](#). To this end, I measure each agent’s asset position at retirement a_i^{ret} , and compute each agent’s (annualized) lifetime earnings,

$$y_i^{PV} = \sum_{k=1}^{K_{ret}} R^{-k} y_{ik} \quad y_i^{PV,a} = \frac{1}{\sum_{k=1}^{K_{ret}} R^{-k}} y_i^{PV}.$$

Here, y_{ik} denotes agent i ’s post-tax and post-transfer income, $y_{ik} = \max\{\underline{y}, T^{inctax}(y_i^{pre})\}$, at age k . Finally, I estimate

$$\frac{a_i^{ret}}{y_i^{PV}} = \alpha + \gamma y_i^{PV,a} + v_{it}$$

and report γ in Table 8.

E.3 Additional steady state results

E.3.1 The schedule of elasticities σ_k

Table 13 shows the age-dependent elasticities σ_k for each age k .

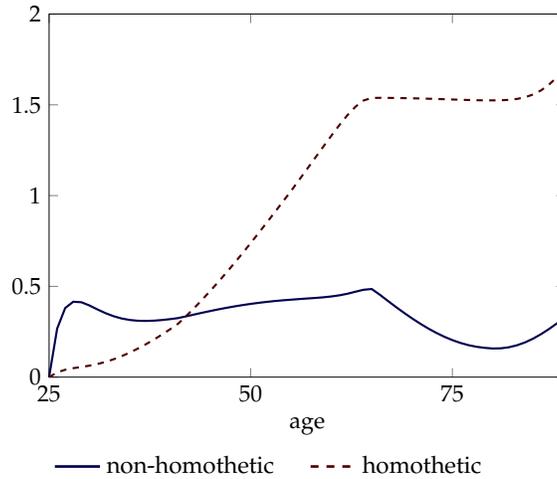
⁷⁸The notation was adjusted to match the one used in this paper. Since the economy in which I estimate this specification is stationary, I only consider the case in which ϕ and χ as well as all shock variances are constant.

Table 13: The age profile of elasticities.

Age	σ_k	Age	σ_k	Age	σ_k	Age	σ_k
25	11.7204	45	3.7369	65	1.6402	85	1.2581
26	11.0686	46	3.5389	66	1.5972	86	1.2581
27	10.453	47	3.3561	67	1.5576	87	1.2581
28	9.8716	48	3.1872	68	1.5211	88	1.2581
29	9.3226	49	3.031	69	1.4875	89	1.2581
30	8.8041	50	2.8865	70	1.4567		
31	8.3144	51	2.7527	71	1.4285		
32	7.852	52	2.6288	72	1.4028		
33	7.4153	53	2.514	73	1.3795		
34	7.0029	54	2.4075	74	1.3585		
35	6.6134	55	2.3088	75	1.3397		
36	6.2456	56	2.2172	76	1.3229		
37	5.8982	57	2.1323	77	1.3082		
38	5.5702	58	2.0534	78	1.2955		
39	5.2604	59	1.9802	79	1.2847		
40	4.9678	60	1.9123	80	1.2758		
41	4.6915	61	1.8493	81	1.2687		
42	4.4306	62	1.7909	82	1.2634		
43	4.1842	63	1.7367	83	1.2599		
44	3.9514	64	1.6866	84	1.2581		

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

Figure 16: 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 (red, dashed) and non-homothetic (blue, solid) models.

E.3.2 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 $\bar{\mu}_k(s, a, z, \varphi)$ be the distribution over (s, a, z, φ) conditional on age k . Then, define the average risk aversion at age k as

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

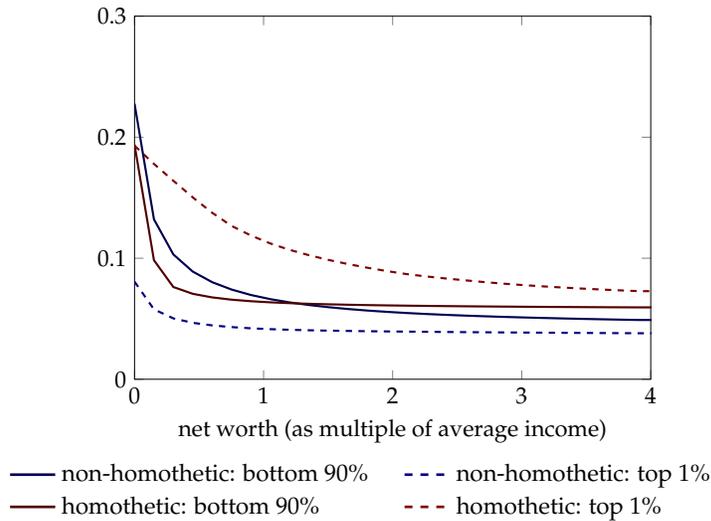
Figure 16 shows the path of risk aversions Σ_k for the homothetic and the non-homothetic economies.

As can be seen, risk aversion in the homothetic economy increases considerably over the life cycle. If risky assets were 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.⁷⁹

⁷⁹See [Wachter and Yogo \(2010\)](#) for an important contribution that makes this point in a non-homothetic life-cycle economy.

Figure 17: 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).

E.3.3 The distribution of MPCs

As Section C.2 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 the non-homothetic economy implies. Figure 17 shows MPCs at the median income state (when $\eta = \psi = 0$), 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).

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).

E.3.4 Savings rates by permanent income group

One way to look at the influence of non-homothetic preferences on agents' behavior is through savings rates. Table 14 shows average work-life (25-64 years) savings rates, defined as $\frac{y_{it} + ra_{it} - c_{it}}{y_{it} + ra_{it}}$, for each of the three skill groups. Vast differences are visible in the non-homothetic model: the average

Table 14: Work-life savings rates by permanent income.

	Savings rate in		
	top 1%	next 9%	bottom 90%
Non-homothetic model	57%	30%	1%
Homothetic model	17%	16%	16%

Note. This table shows savings rates by permanent income, for the non-homothetic and the homothetic economies. The savings rates are defined as $(y + ra - c) / (y + ra)$ and are averaged over the entire work-life (25–64 years) and all idiosyncratic income and asset states.

Table 15: Joint distribution of income and wealth.

	Fraction of top 1% (labor income) in top 1% (wealth)	Fraction of 1% (wealth) in top 1% (labor income)
Data for 2000 (Alvaredo et al., 2015)	0.27	0.28
Model (calibrated to 2014)	0.34	0.42

Note. This table shows the joint distribution of top income and top wealth shares, both in the non-homothetic baseline economy and in the data.

savings rate in the top 1% is 57%, while the next 9% are able to save 30%. The bottom 90% only save on average 1%. The homothetic economy, by contrast, is almost neutral⁸⁰ and essentially shows no difference in savings behavior by permanent income.

E.3.5 Joint distribution of labor and capital income

The reason behind the success in generating wealth inequality in this model is purely non-homothetic savings behavior. A way to gauge whether this channel is too strong in the model relative to the data is to study implications for the joint distribution of income and wealth, which was recently highlighted in [Aaberge et al. \(2013\)](#) and [Alvaredo et al. \(2013\)](#).

In [Table 15](#), I compute the fraction of agents in the top 1% of the labor income distribution that are also in the top 1% of the wealth distribution (and vice versa), and compare it to the data in [Aaberge et al. \(2013\)](#). It is clear that top incomes and top wealth levels are too aligned relative to the data, but not by very much. In fact, the data is for the year 2000, and it could well be that the income and wealth distributions have become even more aligned in recent years.⁸¹ In addition, there is of course a lot more sources of variation in the data than in the model—for instance heterogeneity in returns on wealth—that could explain why the income-wealth alignment is somewhat lower in the data.

⁸⁰Recall that it does allow for bequests and imperfect skill transmission, and therefore is not entirely neutral.

⁸¹In fact, this is exactly what happens during the model's transitional dynamics.

E.3.6 Lorenz curves

The Lorenz curves for pre-tax incomes, consumption and wealth are shown and compared to the data in Figure 18, panel (a) (see Appendix E.3 for details). 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 5.3, the model is successful in matching the overall distribution of wealth, especially compared to the homothetic model, which is shown as the dotted red line.

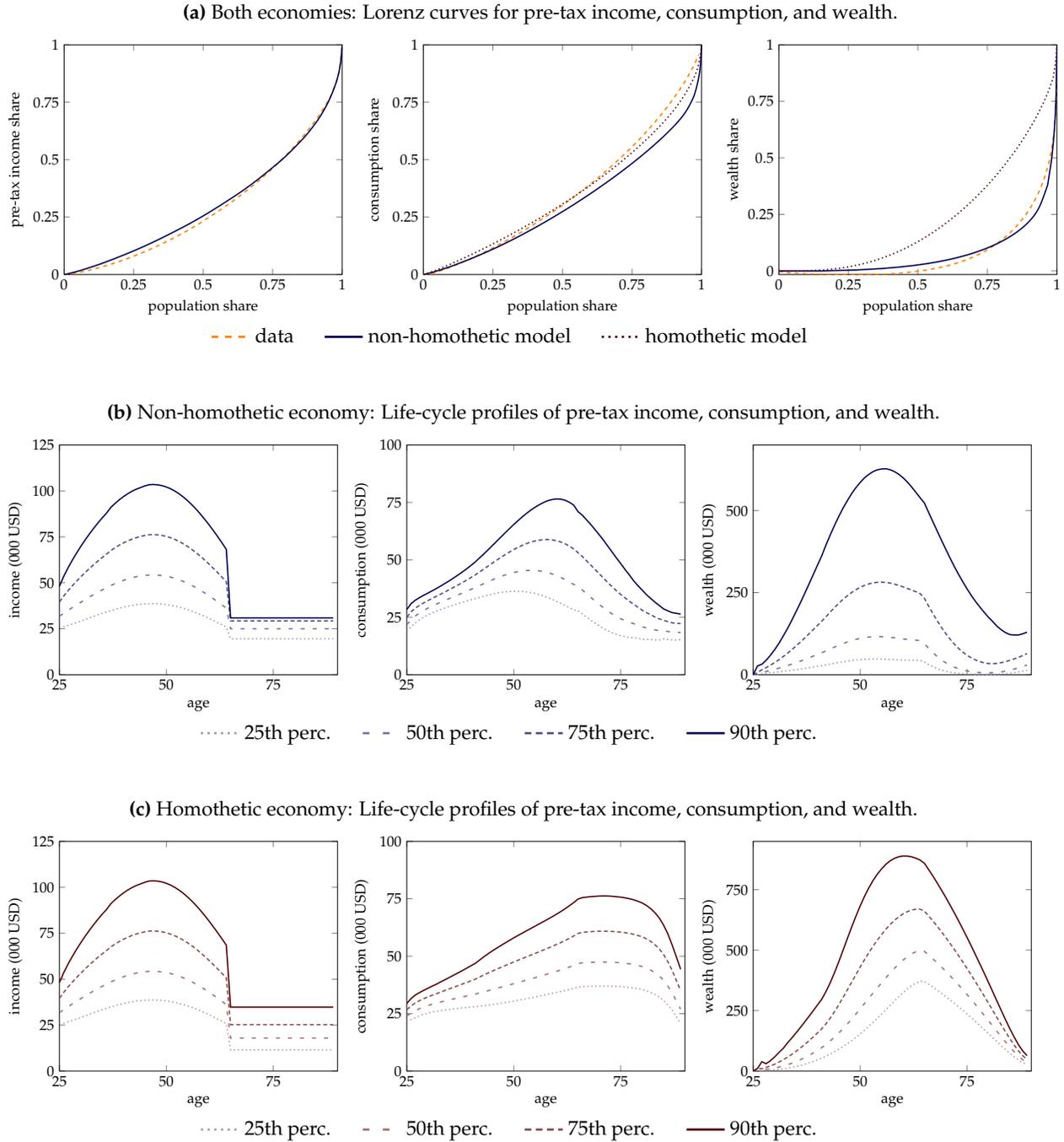
The data for the Lorenz curve for consumption is computed in the same way as the Lorenz curve for income, from the PSID. Strikingly, it shows a much smaller difference between the non-homothetic and homothetic economies—much like in the simple model of Section 2 (see e.g. Figure 3). The non-homothetic model does, however, predict larger consumption inequality at the top end compared to the data. One reason for this could be that certain expenses (such as for kids' education, see Figure 13) are counted as expenditure in the model, but are not in the data.

E.3.7 Life cycle profiles

Panels (b) and (c) of Figure 18 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 dissaved 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

Figure 18: Key life-cycle characteristics of the non-homothetic and homothetic model.



Note. This figure shows key characteristics of two economies: a non-homothetic economy, where rich households save disproportionately more than poor households in relation to their incomes, and a standard homothetic economy, where the savings rates are equal. See Appendix E.3 for details.

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 18 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, $Q_k^{\text{lognormal}}(p)$. 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 $Q_k^{\text{lognormal}}(p_0)$. The dollar figures were computed using an average household income of \$70,000 (USD in 2014).

E.3.8 Life cycle dispersion

While the plots in Figure 18 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 et al. (2004), Guvenen (2007) and Huggett et al. (2011), among many others.

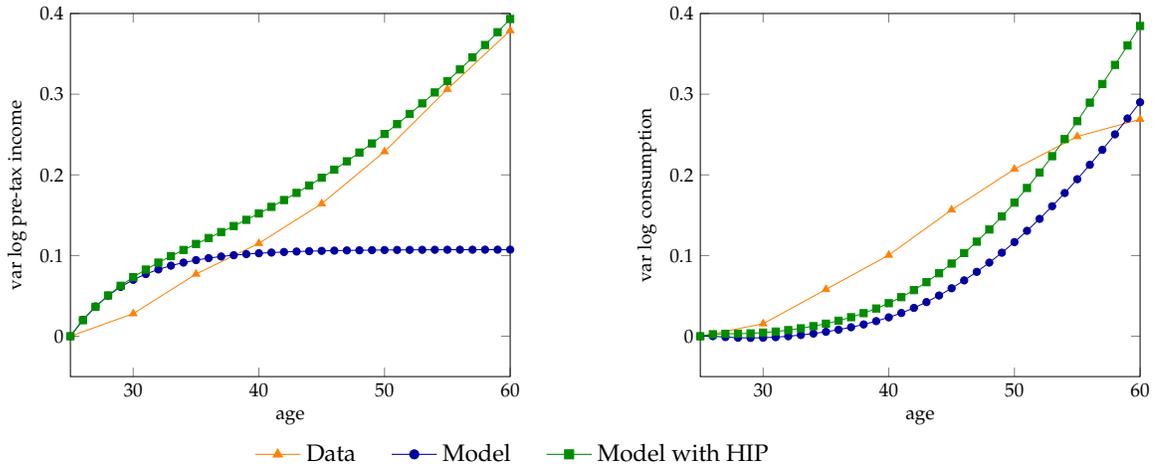
There are two common ways to match both the rise in income dispersion and consumption dispersion: either a persistent AR(1) income process, see e.g. Storesletten et al. (2004); or an income process with less persistence but ex-ante heterogeneous income profiles (HIP), which, together with learning about 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 economy that *in addition* has heterogeneity in income profiles (modeled as in Guvenen (2007, 2009)), to better capture the increase in dispersion over the life cycle.⁸²

Figure 19 shows the age profiles of income and consumption dispersions in both the baseline non-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.

⁸²The addition of heterogeneity in income profiles slightly pushes up the estimates of the permanent income elasticity of consumption shown in Section 5.3, by around 0.03. (The HIP model could be re-calibrated to match the empirical estimate for ϕ .)

Figure 19: 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 \(1994\)](#).

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.

E.4 Background on the models used for comparison in Section 5.4

I compare the non-homothetic and homothetic life cycle models to a variety of other life cycle and infinite horizon models in Table 8. Here, I explain for each model how it was calibrated and list its parameters. 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 16.

All alternative models are extensions of the homothetic framework that was introduced at the end of Section (5.2). In addition, to ensure that the model is perfectly neutral, I assume the

Table 16: Calibrated parameters the comparison models.

	Re-calibrated parameters		
	β	κ	\underline{a}
<i>1. Main models</i>			
Non-homothetic	0.886	16.059	1.616
Homothetic	0.990	1.340	
<i>2. Alternative preferences or transfers</i>			
Homothetic w/out bequests	0.997		
Homothetic w/ social security	0.998	0.392	
Homothetic u but luxury bequests	0.999	19.967	9.208
<i>3. Alternative income process</i>			
AR(1) with $\rho = 0.95$	0.998		
Permanent-transitory	1.000		
Heavy-tailed	0.991		
Extreme productivity state	0.975		
Heterogeneous income profiles	0.998		
<i>4. Other</i>			
Partial insurance	1.003		
Random discount factors	0.956		

government levies a 100% tax rate on bequests.⁸³ This entirely neutral framework is the model underlying the first model in group 2 in Table 8.

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 5.2. 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 $\sigma_{\eta}^2 = 0.01$, $\sigma_{\psi}^2 = 0.05 - \sigma_v^2$, $\sigma_{\alpha}^2 = 0.15$.

The third model is a model with heavy-tailed income shocks. Here, I incorporate some elements of the income process in Guvenen et al. (2016), and assume that the income process is given by

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

⁸³One could have also introduced a perfect market for annuities to ensure neutrality. I expect the results to be very similar.

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

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

The initial standard deviations of η_{it}^l are given by σ_{ini}^l . I take the parameters $\rho^l, \sigma^l, \mu^l, \sigma_{ini}^l$ directly from [Guvenen et al. \(2016\)](#), Column 3 of Table II. Since I do not incorporate the entire (more complicated) income process in [Guvenen et al. \(2016\)](#), I rescale the innovation $\eta_{it}^l - \rho_l \eta_{it-1}^l$ to match the increase in the life cycle variance of 0.6 in [Guvenen et al. \(2016\)](#). 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. \(2016\)](#).

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, $\sigma_\psi^2 = 0$; (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 \(2017\)](#).⁸⁴ 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_β^2 . I assume $\rho = 0.82$ as in [Guvenen \(2009\)](#) and $\sigma_\beta^2 = 0.00012$ to match the life cycle increase in the variance of income of 0.3.

⁸⁴The results are similar if the transition matrix for uneducated agents is being used here.

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

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

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.

F 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 8](#)). 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,

$$\delta A_t \equiv -A_t + (1 + r_t)B + J_t(K_t) + J_t^{ext} + NFA.$$

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.⁸⁵

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 $t = t_1, \dots, t_l$. 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 $\delta v_t \equiv \frac{\Delta(J_0(K_0) + J_0^{ext})}{\Delta r_t}$, and compute the “pure” response (without revaluation) $\mathbf{J}_t^{pure} \equiv \mathbf{J}_t - \delta v_t \mathbf{J}^{reval}$.
- (c) For any time t that is not in $\{t_1, \dots, t_l\}$, 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 $t_\ell, t_{\ell'} \in \{t_1, \dots, t_l\}$, 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

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

⁸⁵The algorithms presented here were jointly developed with Adrien Auclert and Matt Rognlie.

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

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

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

(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 $\|\delta\mathbf{A}(\mathbf{r}^{(n)})\|_\infty$.