Demographics, Wealth, and Global Imbalances in the Twenty-First Century

Adrien Auclert* Hannes Malmberg† Frédéric Martenet‡ Matthew Rognlie§

July 2021

Abstract

We use a sufficient statistic approach to quantify the general equilibrium effects of population aging on returns, wealth accumulation, and global imbalances. Combining population forecasts with household survey data from 25 countries, we measure the *compositional effect* of aging: how a changing age distribution affects wealth-to-GDP, holding the age profiles of assets and labor income fixed. We find that this effect is large and heterogeneous across countries. In a baseline OLG model, we show that this statistic, in conjunction with cross-sectional information and two standard macro parameters, pins down general equilibrium outcomes. Through the twenty-first century, population aging will lower returns, increase wealth-to-output ratios, and create large imbalances between countries at different stages of transition. These conclusions extend to a richer model in which bequests, individual savings, and the tax-and-transfer system all respond to demographic change.

*Stanford University, NBER and CEPR. Email: aauclert@stanford.edu.

†University of Minnesota. Email: pmalmber@umn.edu.

‡Stanford University. Email: martenet@stanford.edu.

§Northwestern University and NBER. Email: matthew.rognlie@northwestern.edu.

For helpful comments, we thank Mark Aguiar, Anmol Bhandari, Maricristina De Nardi, Nezih Guner, Fatih Guvenen, Daniel Harenberg, Martin Holm, Gregor Jarosch, Patrick Kehoe, Pete Klenow, Dirk Krueger, Kieran Larkin, Ellen McGrattan, Kurt Mitman, Ben Moll, Serdar Ozkan, Alessandra Peter, Jacob Robbins, Richard Rogerson, Ananth Seshadri, Isaac Sorkin, Kjetil Storesletten, Chris Tonetti, Arlene Wong, Owen Zidar and Nathan Zorzi. We also thank Etienne Gagnon and John Sabelhaus for sharing code and data. This research is supported by the Hellman Faculty Scholar Fund Grant at Stanford.

1 Introduction

The world is in the process of rapid demographic change. Since 1950, the share of the world population above 50 years of age has increased from 15% to 25%, and this share is expected to rise further to 40% by the end of the twenty-first century (Figure 1, Panel A). Macroeconomists agree that demographics has likely been a driver of several important recent trends. An aging population saves more, helping to explain why rates of return have fallen and wealth-to-output ratios have risen (Figure 1, Panels B,C). Insofar as this process is heterogeneous across countries, it can further explain the rise of global imbalances (Figure 1, Panel D).[1]

Beyond this qualitative consensus lies substantial disagreement about magnitudes. For instance, estimates of the effect of demographics on interest rates over the 1970–2015 period range from a moderate decline of 75 basis points (Gagnon, Johannsen and López-Salido 2021) to a large decline of over 300 basis points (Eggertsson, Mehrotra and Robbins 2019). Turning to predictions for the future, economists are starkly divided about the direction of the effect. A number of quantitative models predict falling interest rates going forward (e.g. Gagnon et al. 2021, Papetti 2019). At the same time, an influential hypothesis argues, based on the savings behavior of the elderly, that aging will push savings rates down and interest rates up. This argument, popular in the 1990s as the "asset market meltdown" hypothesis (Poterba 2001, Abel 2001), was recently revived under the name "great demographic reversal" (Goodhart and Pradhan 2020). In the words of ECB chief economist Philip Lane (Lane 2020):

> The current phase of population ageing is contributing to the trend decline in the underlying equilibrium real interest rate [...] While a large population cohort that is saving for retirement puts upward pressure on the total savings rate, a large elderly cohort may push down aggregate savings by running down accumulated wealth.

In this paper, we refute the great demographic reversal and show that, instead, demographics will continue to push strongly in the same direction, leading to falling rates of return and rising wealth-to-output ratios. We find that the key force is the *compositional effect* of an aging population: the direct impact of the changing age distribution on

[1] Appendix A describes the construction of our total return series, depicted in figure 1, panel B. This series declines more since 1980 than the return in Gomme, Ravikumar and Rupert (2011) partly because it includes the returns on safe assets, but primarily because it is calculated as the ratio of income to wealth rather than the ratio of income to measured fixed assets. The very low safe returns in the 1950s are consistent with Jordà, Knoll, Kuvshinov, Schularick and Taylor (2019)

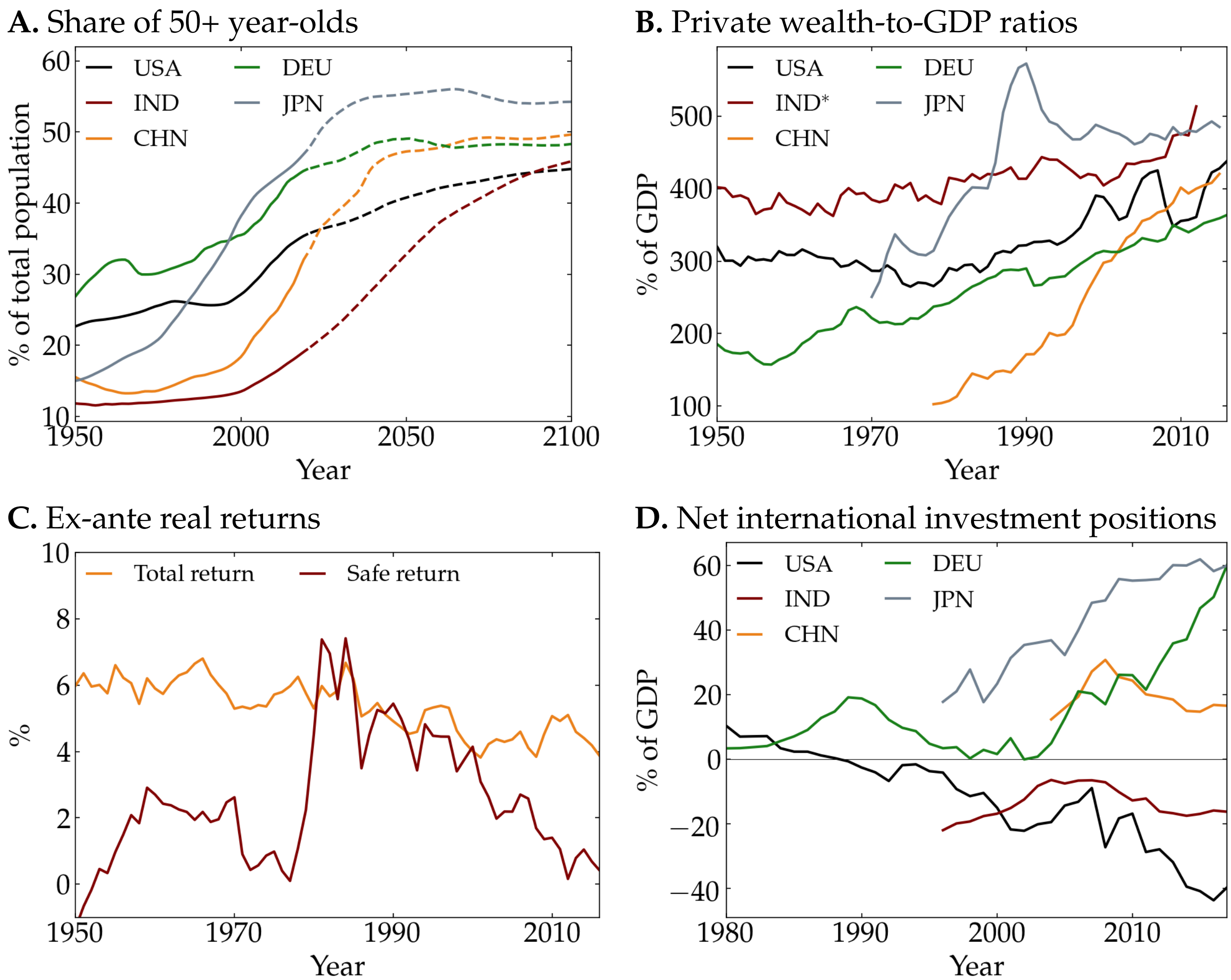


Figure 1: Demographics, wealth, interest rates and global imbalances

Notes: Panel A presents the share of 50+ year-olds from 1950 to 2100 as predicted by the 2019 UN World Population Prospects. Panel B present the evolution of wealth-to-GDP ratios until 2010 from the World Inequality Database (WID). The red line for India shows the national wealth-to-GDP ratio since the WID does not provide data on private wealth. Panel C presents a measure of the US total return on wealth (orange line) and of the US safe rate of return (red line). Details on the construction of the series are in appendix A. Panel D presents the Net International Investment Position taken from the IMF and normalized by GDP.

wealth-to-GDP, holding the age profiles of assets and labor income fixed. In a baseline overlapping generations (OLG) model, this is a sufficient statistic for the actual change in wealth-to-GDP for a small open economy. Further, for a world economy, the compositional effect—when aggregated across countries, and combined with elasticities of asset supply and demand that we obtain with other sufficient statistic formulas—fully pins down general equilibrium outcomes.

To measure the compositional effect, we combine population forecasts with household survey data in 25 countries over the period 2016–2100. We find that it is positive and large everywhere, but also heterogeneous, ranging from 48pp in Hungary to 327pp in India. Since the effect is positive and large, our framework implies that there will be no great demographic reversal: through the 21st century, population aging will continue to push down global rates of return, with our central estimate being -125bp, and push up global wealth-to-output, with our central estimate being a 10% increase, or +45pp. But the heterogeneity will also generate global imbalances, as rapid demographic change in some countries implies especially large compositional effects, leading to a savings glut and growing net foreign asset positions. For instance, we predict that India's net foreign asset position will steadily grow until it reaches 100% of GDP in 2100, while the United States's net foreign asset position will decline to absorb these savings.

Our sufficient statistic framework offers a transparent way to compute the effect of a changing age distribution on key macroeconomic variables. General equilibrium outcomes can be obtained with a limited amount of information: in addition to the data needed for the compositional effect, we only need data on macroeconomic aggregates and assumptions on two standard parameters, the elasticity of intertemporal substitution and the elasticity of substitution between capital and labor. The framework allows us to trace conflicting estimates in the literature to their source: for instance, Gagnon et al. (2021) and Eggertsson et al. (2019) differ not because of their assumptions on demographics, but because their models imply different elasticities of asset demand. It also allows us to understand the error behind the demographic reversal hypothesis, which focuses on flows (a declining savings rate) when the theory tells us that stocks (rising wealth-to-income) are most relevant for the global rate of return. We validate our framework by extending the baseline model to a richer model in which bequests, individual savings, and the tax-and-transfer system all respond to demographic change. We find that the results are always the same qualitatively, and that with one exception—extreme fiscal adjustments that fall entirely either on tax increases or benefit cuts—they are also close quantitatively to those we obtain from the sufficient statistic approach.

Our measure of compositional effects is related to previous attempts at using measures of demographic composition to predict the effects of populating aging on aggregate wealth accumulation and interest rates. An early literature focused on the predictions of life-cycle hypothesis for the effect of changes in the population growth rate on the aggregate savings rate, assuming constant interest rates (e.g. Modigliani 1986, Deaton and Paxson 1995). Closer to our compositional effect exercise, one literature, following Mankiw and Weil (1989) and Poterba (2001), focuses on changing age distributions over fixed asset profiles.[2] A separate literature, following Cutler, Poterba, Sheiner and Summers (1990) and the "demographic dividend" literature (Bloom, Canning and Sevilla 2003), focuses on changing age distributions over fixed income profiles. We consider both effects at once, and show that it is this combination that is the main determinant of general equilibrium outcomes in a fully specified OLG model. Some of our findings echo those of the earlier literature. In particular, Poterba (2001) argued that the asset market meltdown hypothesis was unlikely to be right, since his calculation of projected asset demand showed no change between 2020 and 2050. Our updated calculations for the United States suggest instead that projected asset demand will rise dramatically until the end of the 21st century. The effect from the reversing demographic dividend is also important: in the United States, it contributes a third of the overall increase in our compositional effect measure.

The modern analysis of the causal effect of demographics relies on fully specified structural OLG models. This tradition, which originated in Auerbach and Kotlikoff (1987), has tackled effects of demographics on aggregate wealth accumulation and pension reforms (İmrohoroglu, İmrohoroglu and Joines 1995, De Nardi, Imrohoroglu and Sargent 2001, Kitao 2014), international capital flows (Henriksen 2002, Börsch-Supan, Ludwig and Winter 2006, Domeij and Flodén 2006, Krueger and Ludwig 2007, Backus, Cooley and Henriksen 2014, Bárány, Coeurdacier and Guibaud 2019), and asset returns (Abel 2003, Geanakoplos, Magill and Quinzii 2004, Carvalho, Ferrero and Nechio 2016, Eggertsson et al. 2019, Lisack, Sajedi and Thwaites 2017, Jones 2018, Papetti 2019, Rachel and Summers 2019, Kopecky and Taylor 2020, Gagnon et al. 2021). We contribute to this literature by pointing out a key moment that drives the counterfactuals in these models and can be measured directly in the data. Our framework also helps reconcile the diverging conclusions that this literature has reached regarding the quantitative impact of aging on interest rates, showing how two primitive elasticities — the elasticity of capital-labor substitution and the elasticity of intertemporal substitution in preferences — are the key drivers of quantitative outcomes, once models are matched to micro data and population projections.

Our paper bridges the gap between the reduced-form shift-share literature and the quantitative general equilibrium literature. In doing so, it relates to a literature on sufficient statistics, which is well developed in public finance (Chetty 2009) and trade (Arkolakis, Costinot and Rodríguez-Clare 2012) and is now gaining some traction in macroeconomics. This literature has focused on using cross-sectional information to capture the

[2] There is also a tradition that focuses on changing age distribution over fixed savings *rates* (Summers and Carroll 1987, Auerbach and Kotlikoff 1990, Bosworth, Burtless and Sabelhaus 1991). This calculation is subject to substantial measurement error and may not give the correct sign of the effect on rates of return, as we discuss in section 6.

effect on the impulse response of macroeconomic aggregates either on impact (Auclert 2019, Berger, Guerrieri, Lorenzoni and Vavra 2018, Auclert and Rognlie 2018, Auclert, Rognlie and Straub 2018), or cumulatively (Alvarez, Le Bihan and Lippi 2016, Baley and Blanco 2021). We show how to apply this methodology to forecast the macroeconomic effects of demographics, both in the transition and in the long-run.

The paper proceeds as follows. In section 2, we describe our baseline model and define the compositional effect. We show that the effect of aging on wealth-to-GDP in a small open economy exactly coincides with the compositional effect, and that world equilibrium outcomes can be obtained by combining this effect with elasticities of asset supply and demand; these elasticities, in turn, can also be obtained using sufficient statistic formulas. In section 3, we turn to measurement, gathering data from 25 countries to apply the section 2 framework for 2016–2100. We document large and heterogeneous compositional effects, and calculate the general equilibrium implications. In section 4, we validate our framework by extending the baseline model to capture additional macroeconomic effects of population aging—including those working through bequests and the tax-and-transfer system—and show that the results from sections 2 and 3 are a close fit in nearly all cases. Finally, in section 5 we use our methodology to reconcile disparate findings in the literature, and in section 6 we explain why the demographic reversal hypothesis's focus on savings rates is misleading: although demographic forces will indeed push down net savings rates, this decline—unlike the compositional effect on wealth-to-GDP—does not directly matter for equilibrium rates of return.

2 The compositional effect of demographics

In this section, we set up a benchmark life-cycle model with overlapping generations to study the effects of demographic change. We derive two main theoretical results. First, in a small open economy, demographic change only affects macroeconomic aggregates by changing the age composition of agents. Given a demographic projection, these *compositional effects* can be calculated using data from a single cross section. Second, in an integrated world economy, the long-run effects of demographic change on wealth accumulation, interest rates, and global imbalances can be obtained by simply combining these compositional effects with macroeconomic aggregates, other cross-sectional statistics, and assumptions about two primitive elasticities.

2.1 Environment

Our environment is an integrated world economy with overlapping generations (OLG) of heterogeneous individuals. Time is discrete and runs from $t = 0$ to ∞, agents have perfect foresight, and capital markets are integrated. Apart from the global return on assets, all variables and parameters are allowed to vary across countries, and we drop the country index unless there is a potential ambiguity.

Agents. At any time t, each country is inhabited by N_{jt} individual agents of age j. The total population is $N_t = \sum_j N_{jt}$, and we write $1 + n_{t+1} \equiv N_{t+1}/N_t$ for the growth rate of the population. The demographic primitives are mortality, fertility and migration, which are exogenous to the economy.

An agent in our model is an individual man or woman. Each individual faces a probability ϕ_j of surviving from age j to age $j+1$, so their probability of surviving from birth to age j is $\Phi_j \equiv \prod_{k=0}^{j-1} \phi_k$. For now, we assume that this survival profile is constant over time, and that there is no migration. Hence, the share of the population of age j, which we denote by $\pi_{jt} \equiv \frac{N_{jt}}{N_t}$, varies over time only because of changes in fertility and of the so-called "momentum" effect of demographics.[3]

Agents supply labor exogenously, face idiosyncratic income risk, and can partially self-insure and smooth income over their life cycle by saving in an annuity. The effective labor supply of agents is $\ell(z_j)$, where z_j is a stochastic process. Unless stated otherwise, all individual variables at age j are a function of the whole history of the idiosyncratic shocks z_j, which we denote z^j.

Individuals born in a cohort year k, with a maximal lifespan of T, choose sequences of consumption c and annuities $\mathfrak{a}$ to solve the utility maximization problem

$$\begin{aligned} &\max_{c_{jt}, \mathfrak{a}_{j+1,t+1}} \mathbb{E}_k \left[\sum_{j=0}^{T} \beta_j \Phi_j \frac{c_{jt}^{1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}} \right] \\ \text{s.t.} \quad & c_{jt} + \phi_j \mathfrak{a}_{j+1,t+1} \leq w_t \left((1-\tau)\ell(z_j) + tr(z^j) \right) + (1+r_t)\mathfrak{a}_{jt} \qquad (1) \\ & \mathfrak{a}_{j+1,t+1} \geq 0 \, , \end{aligned}$$

where $t \equiv k + j$ denotes time, w_t is the real wage per efficiency unit of labor at time t, r_t is the return on wealth, τ is the labor tax rate, and $tr(z^j)$ denotes transfers from the

[3] Appendix B.1 shows that these two forces account for the majority of population aging during the US demographic transition between 1952 and 2100. Changing mortality and migration make a more limited contribution, though their importance rises during the latter part of the transition.

government, including social insurance and retirement transfers, for agents of age j with a history z^j. The effective discount factor for age j is $\beta_j \Phi_j$, reflecting discounting due to mortality risk at rate Φ_j and time discounting at an arbitrary age-dependent rate β_j. Deviations from exponential discounting ($\beta_j = \beta^j$ for some β) stand in for age-dependent factors that affect the marginal utility of consumption, such as health status or the presence of children. Hence, this model captures many of the factors that the literature considers essential to understand savings: agents save for life-cycle reasons, for self-insurance reasons, to cover future health costs, and to provide for their children.[4]

The total wealth held by individuals of age j is the product of N_{jt} and the average wealth at age j, $a_{jt} \equiv \mathbb{E}\mathfrak{a}_{jt}$. Aggregate (private) wealth W_t is the sum across age groups:

$$W_t \equiv \sum_{j=0}^{T} N_{jt} a_{jt}. \tag{2}$$

Production. There is a single good used for private consumption, government consumption, and investment. The final output Y_t of this good is produced competitively from physical capital K_t and effective labor input L_t according to an aggregate production function F:

$$Y_t = F(K_t, Z_t L_t),$$

where $Z_t \equiv Z_0(1+\gamma)^t$ captures labor-augmenting technological progress. We assume that F has constant returns to scale and diminishing returns to each factor, with elasticity of substitution η between capital and labor. Effective labor input L_t is a standard linear aggregator

$$L_t = \sum_{j=0}^{T} N_{jt} \mathbb{E}\ell_j, \tag{3}$$

where $\mathbb{E}\ell_j$ denotes average effective labor input per person of age j, capturing variations in experience and hours of work over the life cycle. Capital has a law of motion $K_{t+1} = (1-\delta)K_t + I_t$ without adjustment costs and I_t being aggregate investment. Hence, factor prices equal marginal products: the net rental rate of capital is $r_t = F_K\left(K_t/(Z_t L_t), 1\right) - \delta$, and the wage per efficiency unit of labor is $w_t = Z_t F_L\left(K_t/(Z_t L_t), 1\right)$. We write g_t for the growth rate of the economy:

$$1 + g_{t+1} = \frac{Y_{t+1}}{Y_t}. \tag{4}$$

[4]We assume that children live with one of their parents, whose consumption c_j at age j includes that of the children they care for. Formally, we set $\beta_j = \ell(z_j) = tr(z_j) = 0$ when an individual aged j lives with their parent, so they do not consume or accumulate assets until they start their work life.

With a constant r_t and a stationary population, $g_t \equiv (1+\gamma)(1+n)-1$. Otherwise, g_t also reflects changes in capital intensity and population composition.

Government. The government purchases G_t goods, maintains a constant tax rate on labor income τ, gives individuals state-contingent transfers $tr(z^j)$ indexed to current wages w_t, and finances itself using a risk-free bond with real interest rate r_t. It faces the flow budget constraint

$$G_t + w_t \sum_{j=0}^{T} N_{jt}\mathbb{E}tr_j + (1+r_t)B_t = \tau w_t \sum_{j=0}^{T} N_{jt}\mathbb{E}\ell_j + B_{t+1}, \tag{5}$$

where a positive B_t denotes government borrowing. When demographic change disturbs the balance of aggregate tax receipts and expenditures, the government adjusts G_t to ensure that the debt-to-output ratio $\frac{B_t}{Y_t}$ follows a given, exogenous time path.

Equilibrium. Given demographics, government policy, an initial distribution of assets, and initial levels of bonds and capital across countries such that $F_K - \delta$ is equal to the constant r_0 in each country, an *equilibrium* is a sequence of returns $\{r_t\}$ and country-level allocations such that, in each country, households optimize, firms optimize, and asset demand from households equals asset supply from firms and governments,

$$\sum_c W_t^c = \sum_c (K_t^c + B_t^c).$$

In many cases, it is useful to express stocks relative to GDP. Dividing the above expression by world GDP Y_t, we obtain the equilibrium condition

$$\sum_c \frac{Y_t^c}{Y_t}\frac{W_t^c}{Y_t^c} = \sum_c \frac{Y_t^c}{Y_t}\left[\frac{K_t^c}{Y_t^c} + \frac{B_t^c}{Y_t^c}\right]. \tag{6}$$

A country's net foreign asset position is defined to be its excess of wealth over capital and bonds, $NFA_t^c \equiv W_t^c - (K_t^c + B_t^c)$. Given this definition, (6) says that the average NFA-to-GDP ratio is zero, when countries are weighted by their by GDP.

2.2 A small economy aging alone

We first study a small open economy undergoing demographic change, while all other countries have constant demographic parameters. In this case, the economy faces a global rate of return r which is exogenous and fixed—exogenous because the economy is small, and fixed since all other countries have a fixed demography. This can be seen as a limit when the economy has an arbitrarily small workforce and world GDP weight $\frac{Y_t^c}{Y_t}$, so that its demand and its supply of assets do not affect the world equilibrium condition (6).[5] By studying this case, we can analyze how demographics affect macroeconomic aggregates *directly*, independent of any effects operating through equilibrium adjustments in returns r_t.

We focus on wealth, and our key finding is that demographic change only affects the distribution of the population across age groups, not the wealth levels within any age group. Formally, the economy exhibits what we call *balanced growth by age*, where the full distribution of wealth within every age group grows at a constant rate.

Lemma 1. *For any fixed r, a small open economy eventually reaches a balanced growth path by age on which, for each age j, the full distribution of wealth holdings grows at the same rate γ as technology. In particular, average wealth at age j satisfies*

$$\frac{a_{jt}}{Z_t} = a_j(r). \tag{7}$$

for sufficiently large t and some function $a_j(r)$. If initial asset holdings reflect optimal choices given the fixed r (in which case $a_{j0}/Z_0 = a_j(r)$), the economy starts on this balanced growth path, and equation (7) holds for all t and j.

Proof. In appendix B.2. □

The lemma follows from the observation that demographic change does not change the parameters of individuals' life-cycle problems, once these problems are normalized by productivity. Since constant decision parameters imply constant decision rules, individuals born at different times make the same normalized asset choices given their age, history of shocks, and incoming asset holdings. Over time, normalized asset holdings by age and shock history converge to a constant as the influence of initial asset holdings recedes, and we reach a balanced growth path by age. Since the asset holdings on the balanced growth path reflect the choices of individuals optimizing given r, we will start at the balanced growth path as long as the initial assets also reflect optimization given r. In that case, which we assume from now on, we have $a_{jt} = (1+\gamma)^t a_{j0}$ for all t.

[5] To obtain a fixed interest rate, we assume that all other countries $c' \neq c$ are in demographic steady-state given a set of mortality profiles ϕ_j^c and a common growth rate of newborns n, where the constant growth rate ensures that countries preserve their relative size over time.

Given lemma 1, aggregate wealth per person is then given by

$$\frac{W_t}{N_t} = \sum_j \pi_{jt} a_{jt} = (1+\gamma)^t \sum_j \pi_{jt} a_{j0} \tag{8}$$

Wealth per person changes with the age composition π_{jt} of the population, and otherwise grows at the rate of growth of technology $1+\gamma$.

We next derive output per person. Observe first that a constant global r implies a constant ratio of capital to effective labor $k(r)$, defined by $F_K(k(r),1) = r+\delta$. Aggregate output is then $Y_t = Z_t L_t F(k(r),1)$, where, from (3), aggregate effective labor is $L_t = N_t \sum_j \pi_{jt} \mathbb{E}\ell_j$. Hence, output per person is given by the expression

$$\begin{aligned}\frac{Y_t}{N_t} &= Z_t F(k(r),1) \sum_j \pi_{jt} \mathbb{E}\ell_j \\ &= \frac{F(k(r),1)}{F_L(k(r),1)} (1+\gamma)^t \sum_j \pi_{jt} h_{j0}\end{aligned} \tag{9}$$

where $h_{j0} = w_0 \mathbb{E}\ell_j$ is equal to average labor earnings of individuals of age j, and we have used the fact that the initial wage is $w_0 = Z_0 F_L(k(r),1)$.

Taking the ratio of (8) and (9), we find that W_t/Y_t is proportional to the ratio of $\sum_j \pi_{jt} a_{j0}$ and $\sum_j \pi_{jt} h_{j0}$ (where the constant of proportionality is F_L/F evaluated at $k(r)$). The following proposition summarizes this result.

Proposition 1. *On the balanced growth path by age, the wealth-to-output ratio satisfies*

$$\frac{W_t}{Y_t} \propto \frac{\sum \pi_{jt} a_{j0}}{\sum \pi_{jt} h_{j0}}, \tag{10}$$

where $h_{j0} \equiv \mathbb{E} w_0 \ell_j$ is average pre-tax labor income by age, and $a_{j0} \equiv \mathbb{E} a_{jt}$ is average asset holdings by age.

The proposition implies that all changes in W_t/Y_t reflect the changing age composition π_{jt} of the population, given fixed age profiles a_{j0} and h_{j0}. Equation (10) implies that the log change in wealth to GDP between year 0 and year t is given by

$$\log\left(\frac{W_t}{Y_t}\right) - \log\left(\frac{W_0}{Y_0}\right) = \log\left(\frac{\sum \pi_{jt} a_j}{\sum \pi_{jt} h_j}\right) - \log\left(\frac{\sum \pi_{j0} a_j}{\sum \pi_{j0} h_j}\right) \equiv \Delta_t^{comp}. \tag{11}$$

The key feature of equation (11) is that its right-hand side Δ_t^{comp} can be calculated from demographic projections and cross-sectional data alone, with demographic projections providing π_{jt} and cross-sectional data providing a_{j0} and h_{j0}. We call Δ_t^{comp} the *compositional effect* of aging on W_t/Y_t. Proposition 1 shows that, for a small open economy, the log change in W_t/Y_t is equal to this compositional effect. In the next section, we show that Δ_t^{comp} also plays a key role in an integrated world economy.

2.3 Many countries aging together

We now study the general case when all countries age together, and r_t adjusts to clear the global asset market. Using an asset supply and demand framework, we find that demographic change constitutes a shifter of world asset demand whose magnitude is given by the compositional effect (11) averaged across countries. In the style of sufficient statistics, long-run outcomes can be expressed in terms of the compositional effects together with asset demand and supply sensitivities, and we show that these sensitivities in turn can be given closed form expressions in terms of observables and standard macroeconomic parameters. To simplify, we assume here that net foreign asset positions are zero at an initial date $t = 0$, and further restrict the exogenous path B_t^c/Y_t^c to be a constant. We relax these assumptions in appendix B.4.

Our analysis starts from a first order approximation of the world asset market clearing condition (6). Given the assumptions made so far, this reads:

$$\sum_c \frac{Y_0^c}{Y_0} \Delta \left(\frac{W_t^c}{Y_t^c} \right) = \sum_c \frac{Y_0^c}{Y_0} \Delta \left(\frac{K_t^c}{Y_t^c} \right), \tag{12}$$

where Δ denotes level changes between time 0 and t. In this expression, the left-hand side reflects changes in asset demand, and the right-hand side reflects changes in asset supply—here equal to the change in capital-to-GDP, given that countries maintain a constant bond-to-GDP ratio. Our main results focus on changes between time 0 and the "long-run" LR, in which the world has converged to a demographic steady state. Denote by $\epsilon^{c,d} \equiv \frac{\partial \log(W^c/Y^c)}{\partial r}$ the semielasticity of country c's aggregate asset demand to the rate of return,[6] and by $\epsilon^{c,s} \equiv -\frac{\partial \log((K^c+B^c)/Y^c)}{\partial r} = \frac{\eta}{r_0+\delta} \frac{K_0^c}{W_0^c}$ its semielasticity of asset supply. Then we show in appendix B.3 that, for changes between $t = 0$ and $t = LR$, equation (12) rewrites as

$$\bar{\Delta}_{LR}^{comp} + \bar{\epsilon}^d \cdot (r_{LR} - r_0) \simeq -\bar{\epsilon}^s \cdot (r_{LR} - r_0), \tag{13}$$

where $\bar{\Delta}_{LR}^{comp} \equiv \sum_c \omega^c \Delta_{LR}^{comp,c}$ denotes the compositional effect, averaged across countries

[6]Formally, $\epsilon^{c,d}$ is the derivative with respect to r of the balanced growth level of $\log W/Y$ in a small open economy with exogenous r. This includes both the direct household asset accumulation response to r, and the indirect response from the effect of r on wages. We discuss $\epsilon^{c,d}$ further in the next section.

using initial wealth weights $\omega^c \equiv W_0^c / W_0$, associated with the transition of each of these countries from their initial age distribution to their steady-state age distributions, and $\bar{\epsilon}^d \equiv \sum_c \omega^c \epsilon^{d,c}$, $\bar{\epsilon}^s \equiv \sum_c \omega^c \epsilon^{c,s}$ are the wealth-weighted average semielasticities of asset demand and supply to the interest rate.

Equation (13) shows that demographics only change equilibrium outcomes by shifting out the asset demand curve in line with the compositional effect. In that sense, the compositional effect summarizes the full demographic "shock" to the world equilibrium. Aggregate outcomes are obtained by filtering this shock through the sensitivities $\bar{\epsilon}^d$ and $\bar{\epsilon}^s$. Solving (13) for $r_{LR} - r_0$, we obtain the following proposition.

Proposition 2. *If agents start on a balanced growth path by age, initial net foreign asset positions are zero, and governments maintain debt-to-GDP ratios constant, the long-run change in the rate of return is, up to a first order approximation,*

$$r_{LR} - r_0 \quad \simeq \quad -\frac{1}{\bar{\epsilon}^d + \bar{\epsilon}^s} \bar{\Delta}_{LR}^{comp}, \tag{14}$$

where $\bar{\epsilon}^s = \frac{\eta}{r_0+\delta} \frac{\bar{K}_0}{\bar{W}_0}$ is the average semielasticity of asset supply to r, and $\bar{\epsilon}^d$ is the average semielasticity household asset holdings to r. The wealth-weighted average log change in the wealth-to-output ratio is given by

$$\overline{\Delta_{LR} \log \left(\frac{W}{Y} \right)} \simeq \frac{\bar{\epsilon}^s}{\bar{\epsilon}^s + \bar{\epsilon}^d} \bar{\Delta}_{LR}^{comp} \tag{15}$$

Proof. See appendix B.3. □

Intuitively, the average compositional effect $\bar{\Delta}_{LR}^{comp}$ creates an excess demand for assets at fixed r, which must be absorbed by an increase in the world capital stock and/or a reduction in asset accumulation. If $\bar{\epsilon}^s + \bar{\epsilon}^d$ is large, r falls little, because capital and assets are very sensitive to r. If $\frac{\bar{\epsilon}^s}{\bar{\epsilon}^s + \bar{\epsilon}^d}$ is large, wealth rises a lot, because a large share of the adjustment occurs through increases in the capital stock rather than through a reduction in asset accumulation.

Beyond interest rates and wealth levels, our framework also speaks to global imbalances. To see why, note first that absent an adjustment in r, the net foreign asset position (NFA) of a country would increase one-for-one with its compositional effect. In equilibrium, r must fall to ensure that NFAs are zero on average, so the adjustment in r has to reduce the average NFA by the average compositional effect. Hence, the change in a country's NFA is determined by the difference between its compositional effect and the average compositional effect, subject to an additional adjustment when countries have different sensitivities to changes in r. The following proposition summarizes this result.

Proposition 3. *Given the conditions of proposition 2, the long-run change in country c's net foreign asset position NFA^c satisfies*

$$\log\left(1+\frac{\Delta_{LR}NFA^c/Y^c}{W_0^c/Y_0^c}\right) \simeq \Delta_{LR}^{comp,c} - \overline{\Delta}_{LR}^{comp} + \left(\epsilon^{c,d}+\epsilon^{c,s}-\left(\bar{\epsilon}^d+\bar{\epsilon}^s\right)\right)(r_{LR}-r_0) \quad (16)$$

Proof. See appendix B.3. □

2.4 The asset demand semielasticity ϵ^d

Propositions 2 and 3 show that the compositional effects determine aggregate outcomes given the asset supply semielasticity ϵ^s and asset demand semielasticity ϵ^d for each country. The asset supply semielasticity ϵ^s is only a function of observables and two standard macro parameters: the depreciation rate δ and the elasticity of substitution between labor and capital η.

The asset demand semielasticity ϵ^d is more challenging to obtain. As noted by Saez and Stantcheva (2018), there is a "paucity of empirical estimates" for how long-run asset accumulation responds to changes in the rate of return.[7]

Remarkably, however, in a version of our model without income risk and borrowing constraints, it is possible to express ϵ^d in terms of only the intertemporal elasticity of substitution σ and the observed age profiles of assets and consumption. This result reduces the problem to finding a single structural parameter σ, which has been the topic of an extensive empirical literature.

To explain the intuition, it is helpful to study the case where the technology is Cobb-Douglas and the interest rate equals the growth rate of the economy, $r = g$. In that case, our result takes a simple form:

$$\epsilon^d = \underbrace{\sigma\frac{C}{W}\frac{\mathrm{Var}Age_c}{1+r}}_{\equiv\epsilon^d_{substitution}} + \underbrace{\frac{\mathbb{E}Age_c-\mathbb{E}Age_a}{1+r}}_{\equiv\epsilon^d_{income}}. \quad (17)$$

Here, Age_a and Age_c denote random variables which capture how asset holdings and consumption are distributed across different ages. The random variables are defined over ages j, and have a mass proportional to assets and consumption at each age.[8] Thus, $\mathrm{Var}Age_c$ is large when consumption is spread out across different ages, and $\mathbb{E}Age_c$ −

[7]See section 3.3 for a discussion of empirical estimates.

[8]Formally, we define the probability mass of Age_a at each age j to be $\pi_j a_j/A$, the share of assets in the cross-section held by people of age j, and likewise for Age_c. For the case $g = 0$, this is equivalent to defining the mass as the share of assets held at age j across the life-cycle, but with the cross-sectional definition our result holds more generally.

$\mathbb{E}Age_a$ is large if consumption, on average, occurs at higher ages than when people hold assets.

In appendix B.5, we derive (17), connecting it to the broader logic of life-cycle problems and the cross-sectional outcomes that they produce. In sum, the substitution effect $\epsilon^d_{substitution}$ is the most important.[9] It scales with the intertemporal elasticity of substitution, and is proportional to $\text{Var}Age_c$ since there is more scope for intertemporal substitution if consumption is more spread out over the life-cycle. The income effect ϵ^d_{income} reflects that an increased r increases total income. The size of the increase is proportional to total wealth W, it accrues at an average age of $\mathbb{E}Age_a$, and it is used to increase consumption by a uniform proportion across all ages, implying that the rise in consumption occurs at an average age of $\mathbb{E}Age_c$. Aggregate wealth increases if $\mathbb{E}Age_c$ is higher than $\mathbb{E}Age_a$, because then, on average, the new interest income is saved before it is consumed.

For the more general case, there are two complications. First, when technology is not Cobb-Douglas, the labor share can change with r. Since wealth is ultimately accumulated from labor income, this introduces a new term, $\epsilon^d_{laborshare}$. Second, our result for $r = g$ relied on current values being the same as present values normalized by growth, which is no longer true when $r \neq g$. To accommodate the latter, we define $\hat{r} \equiv \frac{1+r}{1+g} - 1$, and define the present value version of aggregates: $W^{PV} \equiv \sum_j \frac{\pi_j a_j}{(1+\hat{r})^j}$ and $C^{PV} \equiv \sum_j \frac{\pi_j a_j}{(1+\hat{r})^j}$, and Age^{PV}_a and Age^{PV}_c as random variables having probability masses at j proportional to $\frac{\pi_j a_j}{(1+\hat{r})^j}$ and $\frac{\pi_j c_j}{(1+\hat{r})^j}$ respectively. This leads us to the following proposition.

Proposition 4. *Consider a small open economy with a steady-state population distribution π. If individuals face no income risk or borrowing constraints, the long-run semielasticity of the steady state W/Y to the rate of return is given by*

$$\epsilon^d \equiv \frac{\partial \log W/Y}{\partial r} = \epsilon^d_{substitution} + \epsilon^d_{income} + \epsilon^d_{laborshare}. \tag{18}$$

When $\hat{r} = 0$, $\epsilon^d_{substitution}$ and ϵ^d_{income} are given by (17)*. When $\hat{r} \neq 0$,*

$$\epsilon^d_{substitution} = \frac{\sigma}{1+r}\frac{C}{W}\frac{\mathbb{E}Age_c - \mathbb{E}Age^{PV}_c}{\hat{r}} \tag{19}$$

$$\epsilon^d_{income} = \frac{1}{1+r}\frac{\frac{C/C^{PV}}{W/W^{PV}} - (1+\hat{r})}{\hat{r}} \tag{20}$$

[9] In the appendix, we give a numerical illustration where $\frac{C}{Y} = 2/3, \frac{W}{Y} = 4, \sigma = 1/2$, consumption is uniformly distributed between 25 and 80, and assets are held on average at age 60. Then $\epsilon^d_{substitution} \approx 115$ and $\epsilon^d_{income} \approx -34$.

In both cases, $\epsilon^d_{laborshare}$ is given by

$$\epsilon^d_{laborshare} \equiv \frac{(1-s_L)/s_L}{r+\delta}(\eta-1), \qquad s_L \equiv \frac{wL}{Y}. \tag{21}$$

Even if the population distribution π does not start out at a steady state, propositions 2 and 3 still hold under this definition of ϵ^d.

Proof. See appendix B.5. □

The result in proposition 4 is continuous in $\hat{r}$: in the limit $\hat{r} \to 0$, the sum of (19) and (20) becomes (17). Moreover, for small $\hat{r}$ (for instance, $\hat{r} \approx 0.02$, corresponding to $r = 0.04$ and $g = 0.02$), we show in appendix B.5 that (17) remains a close approximation to the actual $\epsilon^d_{substitution}$ and ϵ^d_{income}. Further, the labor share adjustment is relatively small: even if η is varied from 0.5 to 1.5, $\epsilon^d_{laborshare}$ is always an order of magnitude smaller than the rest of ϵ^d.[10]

3 Measurement and implications

Propositions 1–4 deliver a framework for quantifying the impact of demographics on macroeconomic aggregates. In a first step, we measure the compositional effects Δ^{comp}_t in the data. Here, we do this for twenty five countries. We find that these effects are very large and heterogeneous across countries. In a second step, we construct the distributions of the ages of consumption and wealth from the data, and combine these with assumptions on structural elasticities η and σ to back out the semielasticities of asset supply and demand to interest rates. Finally, we apply propositions 2 and 3 to obtain our forecasts for interest rates, wealth, and global imbalances until the end of the twenty-first century.

3.1 Measuring the compositional effect

Equation (11) defines the compositional effect Δ^{comp}_t as the log change in the wealth-to-GDP ratio implied by Proposition 1 between a base year 0 and year t:

$$\Delta^{comp}_t \equiv \log\left(\frac{\sum \pi_{jt} a_j}{\sum \pi_{jt} h_j}\right) - \log\left(\frac{\sum \pi_{j0} a_j}{\sum \pi_{j0} h_j}\right)$$

[10] If $s_L = 2/3$ and $r + \delta = 10\%$, varying η between 0.5 and 1.5 implies a variation in $\epsilon^d_{laborshare}$ from -10 to 10, compared to approximately 80 for $\epsilon^d_{substitution} + \epsilon^d_{income}$ in the numerical illustration in appendix B.5.

This equation shows that Δ_t^{comp} is only a function of demographic projections π_{jt} combined with cross-sectional age profiles of average assets a_{j0} and gross labor income h_{j0} in a base year 0. Here we discuss implementation issues; the next section goes on to measure Δ_t^{comp} in our twenty five countries for a range of years t.

Our choice of base year is dictated by the availability of household surveys. Our base year is 2016 for the United States; for other countries we pick the year closest to 2016 for which there is available data. Appendix table A.1 provides a summary of our survey data sources and the resulting base year by country.

We use population data and projections π_{jt} from the United Nations World Population Prospects. The U.N. define central population projection scenarios as well as a number of alternatives. Fertility is one of the biggest sources of uncertainty in these projections. We take this uncertainty into account by making use of their "high fertility" and their "low fertility" scenarios in addition to their central population projection. Relative to each country's baseline, these scenarios respectively assume 0.5 more and fewer births per mother from 2030 onward.

For labor income profiles, we use data from the Luxembourg Income Study (LIS).[11] The model defines h_{j0} as average pretax labor income of individuals of age j. Using survey weights, we compute the ratio of total labor income earned by individuals of age j–including wages, salaries, bonuses, fringe benefits and self-employment income before social security and labor income taxes–to the number of individuals of age j. This gives us an average that does not condition on working: individuals that are out of the labor force contribute zero towards the average of their age group.

For asset profiles, we use data from a collection of wealth surveys, such as the Survey of Consumer Finance (SCF) in the United States. Since our model features a single asset, a_{j0} corresponds to the average of individual net worth: all assets including housing[12] and defined contribution pension wealth, net of all liabilities including mortgages. For the United States, where an important share of wealth is held in private defined benefit (DB) pension plans, we add an estimate of the age-specific present value of the future stream of payments corresponding to the funded part of these plans.[13] Since our surveys

[11] In the United States, these data are based on the March Current Population Survey (CPS).

[12] The fact that households accumulate assets in part through housing does not change Proposition 1, though it changes the implications for asset returns in Proposition 2.

[13] In practice, we add we add to our age-specific measure of SCF wealth 37.5% of the value individual DB pensions provided by Sabelhaus and Volz (2019), ensuring that the overall amount of defined benefit pension plans in our data is consistent with the aggregate amount of non-federal funded defined benefit asset in the US economy. Unfunded DB liabilities correspond to a future transfer tr_j in the household budget constraint (1), and therefore they do not affect the level of wealth a_{j0} that matters for macroeconomic aggregates. Neither does "social security wealth" (Sabelhaus and Volz 2020, Catherine, Miller and Sarin 2020).

measure wealth at the household level, we obtain individual wealth by splitting up all assets equally between all members of the household that are at least as old as the head or the spouse. In appendix C.2, we show that our results are robust to various alternative ways of projecting household to individual wealth, or to measuring income and wealth at the household level and constructing projections for the number of household heads of age j.

To make our results easy to interpret and compare with existing measures of wealth-to-GDP ratios over time, we will often report implied changes in levels rather than log changes. Observe that when Proposition 1 is satisfied, the change in the level of GDP between period 0 and period t is given by

$$\frac{W_t}{Y_t} - \frac{W_0}{Y_0} = \frac{W_0}{Y_0}\left(e^{\Delta_t^{comp}} - 1\right) \tag{22}$$

To implement this calculation, we need information on the baseline level of the wealth to GDP ratio W_0/Y_0. We pick this to be the ratio of net private wealth to gross domestic product from either the World Inequality Database (WID) or the OECD in 2016. Net private wealth is defined as the sum of housing, business and financial assets, net of liabilities, owned by households and nonprofit institutions serving households.[14]

Choice of base year and age effects. In the small open economy environment of Proposition 1, it is irrelevant which base year 0 we use to construct the cross-sectional age profiles of income and wealth, since these profiles grow at the same constant rate γ: along the balanced-growth-by-age path, in any year s,

$$\log h_{js} - \log h_{j0} = \log a_{js} - \log a_{j0} = s\log(1+\gamma) \quad \forall j. \tag{23}$$

Hence, any alternative choice of s leaves W_t/Y_t unchanged at every t. Given (23), the age effects from a time-age-cohort decomposition in repeated cross-sections, with growth specified to load on time effects, would also recover h_{j0} and a_{j0}.[15]

In practice, it is well known that equation (23) is never exactly satisfied in repeated cross-sections: relative age profiles $\log(h_{js}/h_{j's})$ and $\log(a_{js}/a_{j's})$ tend to vary for any pair

[14]Housing assets include the value of dwellings and land; financial assets include currency, bonds, deposits, equity and investment fund shares, as well as life insurance and private pension funds. In appendix table A.1 we compare private wealth from aggregate data to the aggregated sum of individual survey wealth. In theory these should be equal, by equation (2). In practice, when the two differ, equation (22) implicitly rescales wealth proportionately at each age so that the survey aggregate matches the WID or OECD total.

[15]By contrast, the age effects from the classic Deaton (1997) decomposition, in which growth loads on cohort effects, would recover $h_{j0}(1+\gamma)^{-j}$ and $a_{j0}(1+\gamma)^{-j}$, which are not as useful for our purposes.

of ages (j, j') depending on the base year s (see, e.g., Lagakos, Moll, Porzio, Qian and Schoellman 2017). These shifts could be occurring for two types of reasons.

The first type of reason is not a source of concern for our theory. For instance, these shifts could be due to forces that are exogenous to demographics, such as time variation in productivity growth. There, balanced growth by age remains useful to analyze the causal effect of demographics *holding other forces constant*. They could also be due to general equilibrium variation in interest rates caused by demographics. These are fully consistent with our model, which says that it is useful to counterfactually assume that r is constant—and therefore, that equation (23) holds—as an input into the calculation of general equilibrium effects. These types of effects do imply that the base year matters: to examine how much, in appendix C.2 we compute Δ_t^{comp} using cross-sectional profiles in different baseline years, or age effects extracted from a time-age-cohort decomposition with growth loading on time effects.

The second type of reason is a source of concern for our theory: shifts in age profiles over time could be *directly* caused by demographic change. For instance, agents born at later dates could save more because they have a higher life expectancy, or be working longer due to demographic-induced shifts in health or in the retirement age. These "behavioral" effects create departures from our sufficient-statistic result. Quantifying these forces requires a structural model; we turn to this task in section 4.

3.2 The compositional effect around the world

Figure 2 displays the evolution of the predicted change in wealth-to-GDP from the compositional effect over time, for the twenty five countries in our sample. Between 1950 and 2016, this effect has been positive in all countries, averaging 80 percentage points (pp) of GDP. For comparison, the actual increase in W/Y documented by the WID for countries in which there is data is 220pp; for the United States, the compositional effect was 105pp relative to an actual increase in W/Y of 118pp (see appendix C.3). These numbers show that the compositional effects induced by population aging have, historically, been quantitatively large.

Going forward to the end of the 21st century, we find an effect that is positive, even larger on average, and heterogeneous across countries. The effect ranges from 48pp in Hungary to 237pp in China and 327pp in India; in the United States it is 147pp. Fertility projections matter for the magnitude of this result: in high fertility scenarios in which the population ages less, the effect is brought down to 75pp in the U.S. and 142 in China; in low fertility scenarios in which it ages more, these effects are 245pp and 447pp in these

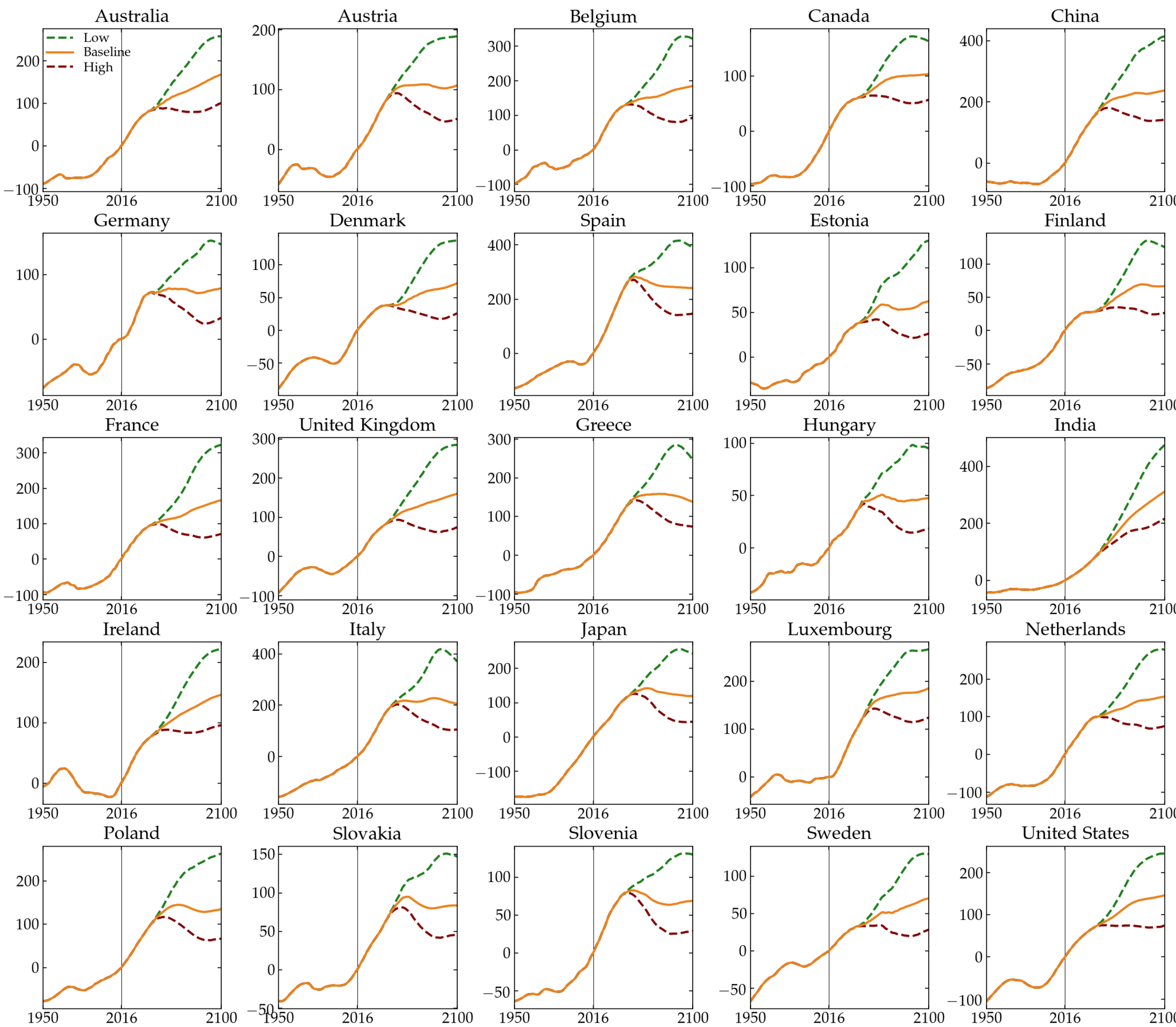


Figure 2: Predicted change in W/Y from compositional effects

Notes: This figure depicts the evolution of the predicted change in the wealth-to-GDP ratio from the compositional effect, calculated using equation (22) for t =1950 to 2100, reported in percentage points. The base year is 2016 (vertical line). The solid orange line corresponds to the medium fertility scenario from the UN, the dashed green line to the low fertility scenario, and the dashed red line to the high fertility scenario.

two countries, respectively.[16]

To understand these results, note from equation (11) that the compositional effect reflects the interaction between changes in the age distribution and the shapes of the wealth and income profiles. Δ_t^{comp} increases as time passes because individuals in older age groups tend to have higher asset holdings and lower labor income than individuals in younger age groups, and because population aging substantially increases the number of individuals in these older age groups.

The case of the United States. We illustrate this logic further by focusing on the case of the United States. The grey bars in Figure 3 show the U.S. population distribution, starting young in 1950, and gradually flattening over time. By 2100, the distribution is almost uniform until age 75, with a large number of very old individuals. Panel A superimposes the 2016 asset profile on these population distributions, and illustrates how demographic change moves individuals into high asset ages. Panel B superimposes the 2016 labor income profile on the population distribution, and illustrates how demographic change first increases aggregate labor income as the baby boomers reach middle-age — the so-called "demographic dividend" (Bloom et al., 2003) — and later decreases aggregate labor income as more individuals reach old age.

This example shows that the interactions between age profiles and age distributions can be subtle. To separate the respective contributions of the asset profile and the labor income profile to the overall composition effect, we perform a first order approximation of equation (11), and write:

$$\Delta_t^{comp} \simeq \underbrace{\frac{\sum (\pi_{jt} - \pi_{j0})\, a_j}{\sum \pi_{j0} a_j}}_{\Delta_t^{comp,a}} - \underbrace{\frac{\sum (\pi_{jt} - \pi_{j0})\, h_j}{\sum \pi_{j0} h_j}}_{\Delta_t^{comp,h}} \tag{24}$$

The first term captures the covariance between the level of asset holdings and the change in the age distribution between year 0 and year t, the second term the covariance between the level of labor income and the change in the age distribution. If, on average, aging between year 0 and year t increases the number of individuals in higher-asset ages, then $\Delta_t^{comp,a}$ is positive, reflecting an increase in wealth per productivity-adjusted-person, per equation (8). If, on average, it also increases the number of individuals in lower-income ages, then $\Delta_t^{comp,h}$ is also positive, reflecting a decline in output per productivity-

[16] By contrast, survey sampling variation matters relatively little, conditional on a fertility scenario. See how small the bootstrapped confidence intervals in figure 4 are relative to the differences in projections across fertility scenarios.

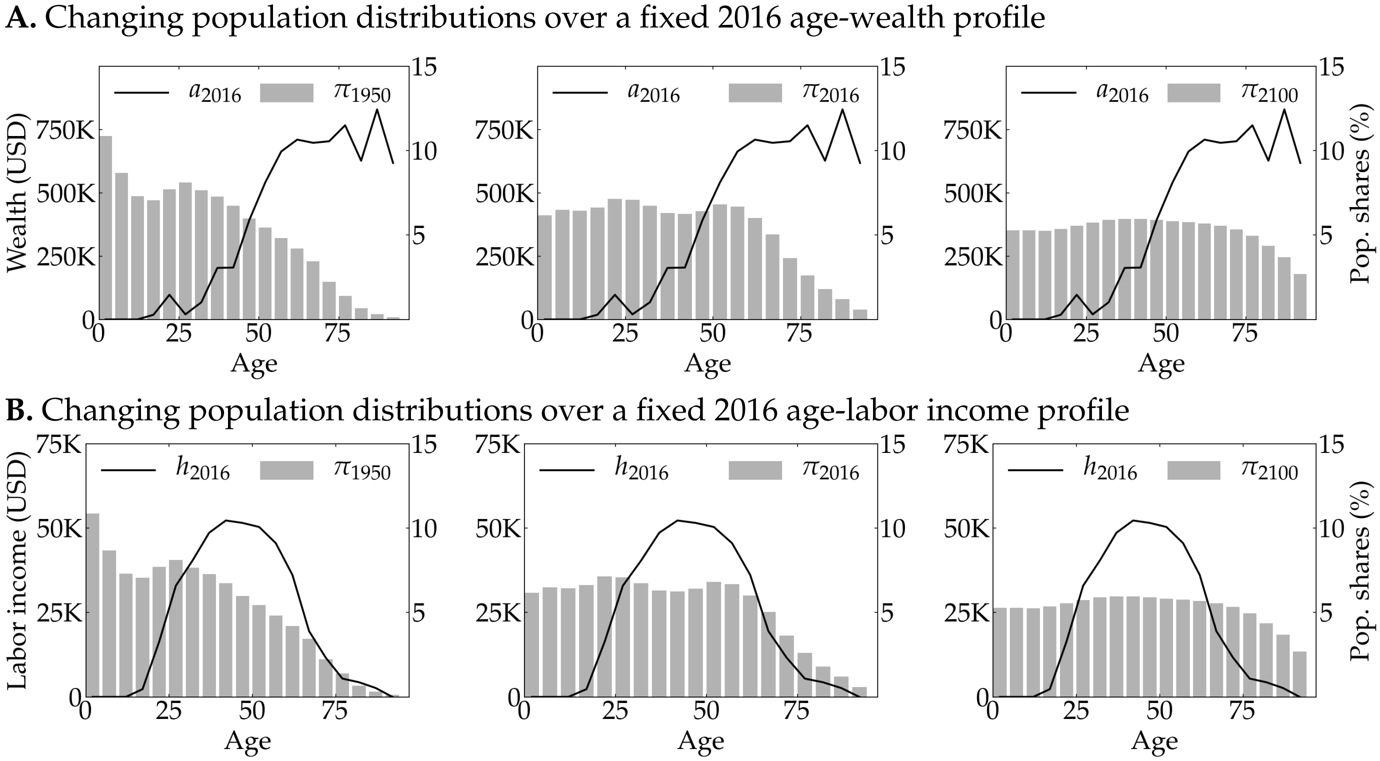


Figure 3: Age-wealth and age-labor income profiles with population age distributions

adjusted-person, per equation (9). In this case, both effects contribute to increase the wealth-to-output ratio.

Figure 4 displays the evolution of these two terms over time, once multiplied by initial wealth to GDP to reflect their contributions to the total change in the level of W/Y.[17] Panel A shows that $\Delta_t^{comp,a}$ always increases as time passes, pushing up on the wealth-to-GDP ratio throughout the sample period. Towards the end of the 21st century, the trend flattens a little. This is because a significant mass of individuals start to reach very old ages where asset accumulation stops. However, the effect does not reverse, because the data shows essentially no asset decumulation at old ages. There is a large literature debating the extent to which life-cycle forces, late-in-life-risks, or bequest motives can account for this fact (see e.g. Abel 2001, Ameriks and Zeldes 2004, De Nardi, French and Jones 2010, De Nardi, French, Jones and McGee 2021). The sufficient statistic result of Proposition 1 allows us to be agnostic about the exact cause of this slow decumulation, within the class of explanations that we allow.[18]

[17]Since $\frac{W_0}{Y_0}\left(e^{\Delta_t^{comp}}-1\right) \simeq \frac{W_0}{Y_0}\Delta_t^{comp} \simeq \frac{W_0}{Y_0}\Delta_t^{comp,a} + \frac{W_0}{Y_0}\Delta_t^{comp,h}$, the two effects approximately sum to the total predicted change from equation (22).

[18]Our benchmark model captures late-in-life risks if β_j increases in old age. It rules out bequests, but when we allow for them in section 4, we find that the compositional effect remains the primary determinant

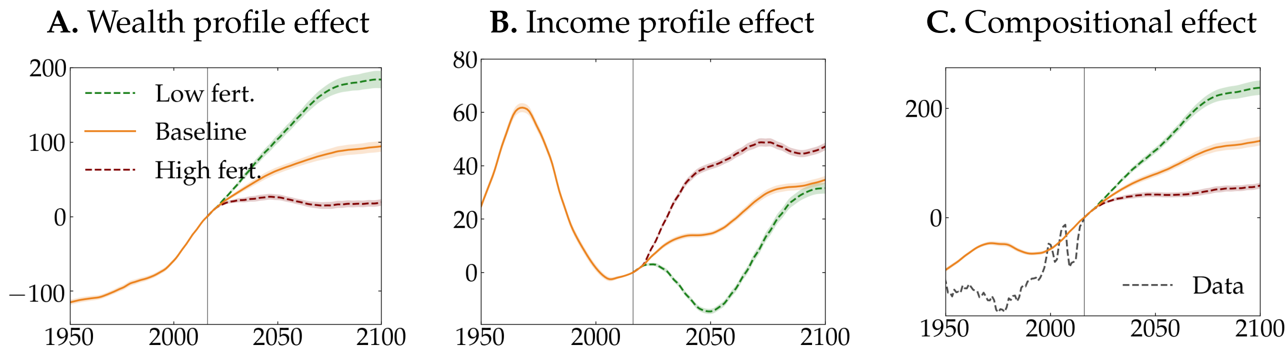


Figure 4: Effects of demographic composition on W and Y: United States 1950-2100

Notes: This figure depicts the evolution of the two terms in the decomposition (24). Panel A presents the contribution from the wealth profile, $\frac{W_0}{Y_0}\Delta_t^{comp,a}$. Panel B presents the contribution from the labor income profile, $\frac{W_0}{Y_0}\Delta_t^{comp,h}$. Panel C presents the overall compositional effect from equation (22), which is approximately equal to the sum of panel A and panel B, overlaid with historical data from the WID. In all graphs, the solid orange line corresponds to the baseline fertility scenario and the dashed green and red lines consider the low and high fertility scenario of the 2019 UN World Population Prospects. A bootstraped 95% confidence interval is computed by resampling observations 10,000 times with replacement.

Overall, panel A confirms and extends important earlier findings by Poterba (2001). Using data from the 1983–1995 waves of the SCF together with population projections until 2050, Poterba (2001) found a flat path for $\Delta_t^{comp,a}$, which he called "projected asset demand", after 2020. Using the more recent SCF waves, and especially updated and more precise population projections, our results instead suggest a substantial increase in $\Delta_t^{comp,a}$ throughout the remainder of the twenty-first century. In addition, Poterba's analysis abstracted away from the effect of aging on the output-to-productivity ratio. Since it is Δ_t^{comp}, rather than $\Delta_t^{comp,a}$, that appears in Propositions 2–4, our results suggest that this is potentially an important omission.

Panel B displays this new effect, measured by $\Delta_t^{comp,h}$. The effect is also positive going forward, contributing about an additional 30 pp to the wealth-to-GDP ratio by the end of the twenty-first century. However, it is negative for the 1970-2010 period. This reflects the demographic dividend. Intuitively, labor income profiles peak earlier and decline earlier than asset accumulation profiles, so aging initially increases output relative to productivity and then pushes down on it. While this effect has been understood for some time (e.g. Bloom, Canning and Sevilla 2003 and Cutler et al. 1990), our contribution is to point out that it is not only important for the effect of demographics on output, but that it also matters for the wealth-to-GDP ratio, and therefore equilibrium returns and global imbalances going forward. Overall, for the United States, the asset accumulation effect contributes

of the effect on W/Y at constant interest rates.

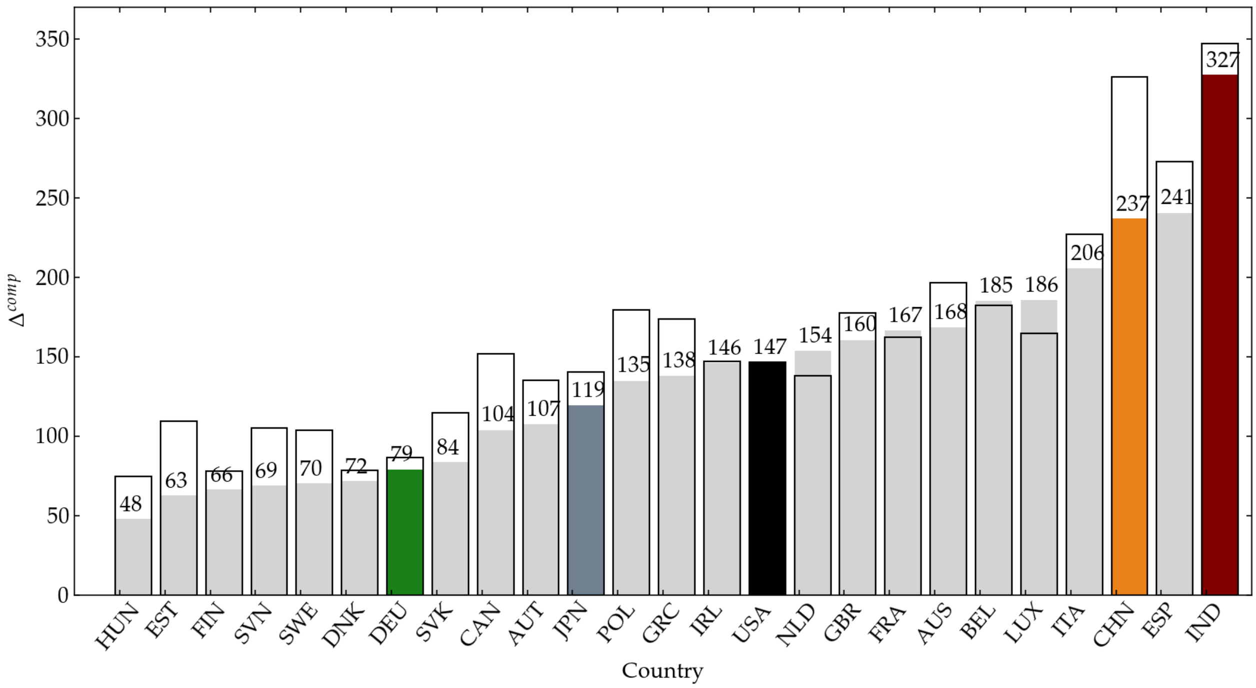


Figure 5: Compositional effects and contribution from demographics alone

Notes: The solid bars show the value of the predicted change in the wealth-to-GDP ratio from the compositional effect between 2016 and 2100 across countries, calculated using equation (22), and reported in percentage points, corresponding to the end point of Figure 2. The transparent bars correspond to the case where Δ^{comp} in equation (11) is calculated using age profiles a_{j0} and h_{j0} from the U.S., but country specific demographics π_{jt}.

about two thirds, and the labor income effect about one third, of the projected increase in W/Y until 2100.

Other countries. The logic behind our findings for Δ_t^{comp} in other countries broadly echo those from the United States. In our online appendix,[19] we reproduce Figures 3 and 4 for all twenty five countries in our sample. While each country has its own peculiarity—for instance, the timing of the demographic dividend is very uneven—in all of them, aging pushes individuals into higher-asset, lower income age groups after 2050.

Figure 5 displays the value of the predicted change in W/Y from composition by 2100 in all countries. It shows that the effect is large everywhere, but also very heterogeneous across countries. In particular, developed countries tend to have smaller compositional effects than developing countries. In principle, this could be because developing countries have different age profiles, or because they experience different demographic transitions. In practice both matter, but the latter is the most important, because these countries have

[19] available at http://web.stanford.edu/~aauclert/demowealth21_country_appendix.pdf

more aging left in front of them, since their demographic transitions were later, and are occurring faster. Figure 5 illustrates this point by showing that there is similarly large heterogeneity if we counterfactually assume that all countries have the same asset and an income profile as the United States. By contrast, appendix figure A.4 shows that countries tend to experience similar compositional effects if they are all assumed to experience U.S. demographics.

Robustness. In appendix C.2, we show that our findings are robust to using alternative ways of allocating household to individual wealth, to calculating compositional effects at the household rather than at the individual level, to using different base years for the cross-sectional income and asset holdings, or to using age profiles that result from a time-age-cohort decomposition. For instance, for the United States, we consider every possible combination of the ten SCF waves between 1989 and 2016, and twelve LIS waves between 1976 to 2016, as well as age effects computed from a time-age-cohort decomposition with growth loading on time. We then calculate Δ_t^{comp} for all of these 143 combinations. The compositional effect varies between 63 and 135 percentage points for 1950 to 2016, and between 58 and 133 percentage points for 2016 to 2100. Hence, while there is some variation across specifications, the effect is always large and positive, both looking back and going forward. This reflects the relative stability of asset and income age profiles over time.

Taking stock. The compositional effect of demographics on wealth-to-GDP has been quantitatively significant historically in every country, and is projected to carry on at a large but heterogeneous pace for the rest of the twenty-first century. These shifts are large enough to significantly affect aggregate wealth accumulation, equilibrium returns, and global imbalances going forward. To address how much, we need to quantify the elasticities of asset supply and demand to interest rates. We turn to this exercise next.

3.3 Asset supply and demand semielasticities

Proposition 2 shows that in the long run, general equilibrium changes in interest rates and wealth-to-output are both proportional to the average long-run compositional effect $\bar{\Delta}_{LR}^{comp}$, modulated by the sensitivities of global asset supply and demand to changes in the rate of return r. We now use our theoretical results to quantify these sensitivities.

Asset supply semielasticity $\bar{\epsilon}^s$. The global asset supply semielasticity reflects how strongly capital responds to changes in the required rate of return.[20] Proposition 2 provides the closed-form solution $\bar{\epsilon}^s = \frac{\eta}{r_0+\delta}\frac{\bar{K}_0}{\bar{W}_0}$, which shows that this response is proportional to the initial global capital-wealth ratio $\frac{\bar{K}_0}{\bar{W}_0}$, the inverse of the user cost of capital $r_0 + \delta$, and the elasticity of substitution between capital and labor η. We quantify these terms as follows. From our calibration of the world economy model in section 4, we take the world capital-wealth ratio to be $\frac{\bar{K}_0}{\bar{W}_0} = 0.78$ and the world user cost of capital to be $r_0 + \delta = 9.7\%$. Given these numbers, Panel A of Figure 6 plots $\bar{\epsilon}^s$ as a function of η in the range of 0 to 1.5, which more than covers the range of typical literature estimates. The resulting $\bar{\epsilon}^s$ varies between 0 and 12; it is equal to 8 under a Cobb-Douglas production function ($\eta = 1$).

Asset demand semielasticity $\bar{\epsilon}^d$. The global asset demand semielasticity reflects how aggregate asset accumulation responds to changes in rates of return. To calculate $\bar{\epsilon}^d$, we use Proposition 4, which expresses this semielasticity in any given country as function of cross-sectional observables from that country, together with the elasticity of intertemporal substitution σ and capital-labor substitution η. In what follows, we assume a constant σ and η across countries. Once we have measured each county-level observable and averaged across countries using 2016 wealth weights ω_c, Proposition 4 tells us that the global asset supply sensitivity $\bar{\epsilon}^d$ can be expressed as a linear function of σ and η as follows:

$$\bar{\epsilon}^d = \sigma \cdot \bar{\epsilon}^{d,\sigma=1}_{substitution} + \bar{\epsilon}^d_{income} + (\eta - 1) \cdot \bar{\epsilon}^{d,\eta=2}_{laborshare} \tag{25}$$

In short, we find that $\bar{\epsilon}^{d,\sigma=1}_{substitution} = 40.5$, $\bar{\epsilon}^d_{income} = -3$ and $\bar{\epsilon}^{d,\eta=2}_{laborshare} = 5.5$. Panel B of Figure 6 displays the resulting aggregate semielasticity of asset supply and demand for $\bar{\epsilon}^d$ for different values of σ and η. Clearly, given that $\bar{\epsilon}^{d,\sigma=1}_{substitution}$ is positive and large and that the other two terms in equation (25) are small, we find $\bar{\epsilon}^d$ to be positive, except if the EIS is extremely low. For a plausible value of the EIS of 0.5 and Cobb-Douglas production, $\bar{\epsilon}^d$ is around 17. This relatively high sensitivity of asset demand to interest rates will have important implications for equilibrium adjustment to demographics.

We now explain how these results are derived, going back to the construction of the objects underlying each. The labor share term in equation (25) is the simplest, since it involves a wealth-weighted average across countries of $(1 - s_L)/s_L \cdot 1/(r_0 + \delta)$. We maintain our assumption of a user cost of capital to be $r_0 + \delta = 9.7\%$, and take country-specific labor shares from our calibration of the world economy model in section 4. Since labor

[20] Recall that we assume that fiscal policy has a constant long-run target for bonds-to-output. If this is endogenous to r, it enters the asset supply sensitivity as well. In models with pure rents, the asset supply sensitivity also includes the response of the capitalized value of these rents to r.

shares are around 2/3 in most countries, $(1 - s_L)/s_L$ is around 0.5, and so the weighted average of $(1 - s_L)/s_L \cdot 1/(r_0 + \delta)$ is around 5.

Next, we turn to the substitution and income effect terms in equation 25. These involve calculating the distributions of the ages of consumption and wealth in each country, as well as its "Present Value" equivalent where the number of agents at age j is discounted at rate $(1 + r)^j$. Proposition 4 requires a stationary distribution, and in simulations we found that it works best at forecasting wealth and interest rates if we use the terminal age distribution to calculate the distributions of ages.

For the distribution of ages of wealth, we take the cross-sectional profile of wealth by age from our surveys, which gives us wealth per person of each age, and then use the country's stationary population distribution implied by the UN population model to determine the share of wealth held by agents of each age. For the distributions of ages of consumption, we could in principle use survey data on consumption directly and then apply the same procedure. Since these tend to be of lower quality than wealth surveys, and since we do not have a survey in each country, we opt instead for backing out the implied profile of consumption from the wealth profile and the income profile, using the budget constraint implied by our model.

Appendix Figure A.5 presents the distribution of ages of consumption and wealth implied by this procedure, for each of our 25 countries. These distributions tend to have a hump shape everywhere–for instance, while average wealth by age tends not to decline too strongly with age, the share of wealth held by 90 year olds is lower because there are fewer 90 year olds, even in a stationary population.[21] They also tend to be relatively spread out across ages, reflecting the fact that all ages account for a significant fraction of both wealth and consumption in the data. Finally, the average age of asset holdings, indicated in a dashed line, tends to be a little higher than the average age of consumption, reflecting the fact that individuals smoothe consumption through their lifetime but accumulate assets especially late in life.

These features of the age distributions of consumption and wealth, together with equation (17), explain our main findings on the relative importance of income and substitution effects for aggregate asset demand.[22] First, the substitution term is approximately a wealth-weighted average of C/W times the variance of age of consumption: since C/W is around $1/6$, and the consumption is very approximately equally distributed uniformly

[21] The share of consumption done by 90 year olds is sometimes elevated in these calculations, due to the way in which we back out the consumption profile from the wealth profile, which involves an annuity formulation with large implied returns at old ages.

[22] Our exact implementation uses equations, where we use a country specific $\hat{r}$, but in practice $\hat{r}$ is small enough in every country that expression (17) gives a useful approximation.

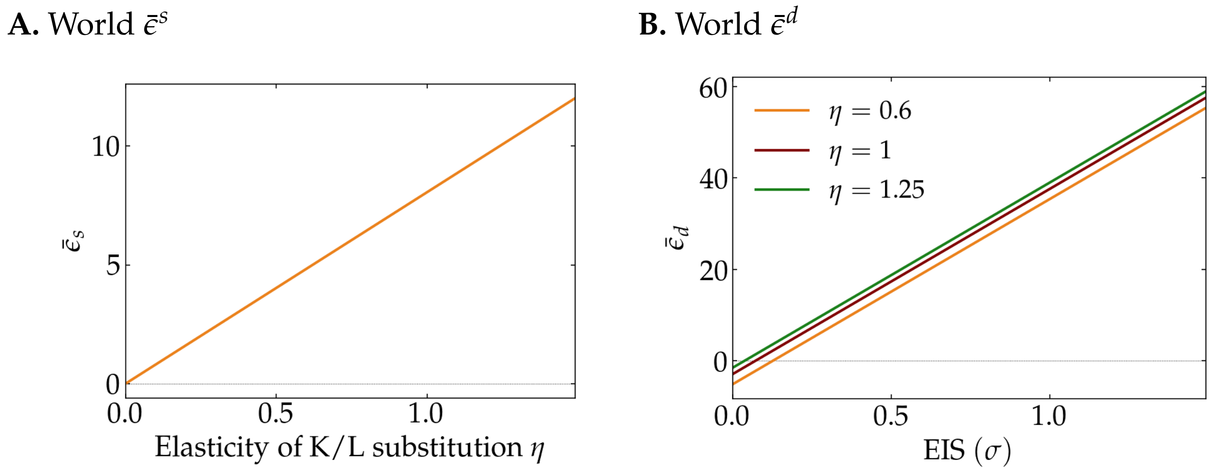


Figure 6: Semielasticities of world asset supply and demand as a function of η and σ

between ages 20 and 80, we obtain a number for $\bar{\varepsilon}^{d,\sigma=1}_{substitution}$ like $(80-20)^2/(12\cdot 6) = 50$. Second, since consumption occurs on average a few years before assets are held, we obtain a small but negative number for $\bar{\varepsilon}^d_{income}$.

Comparison to existing empirical estimates. We briefly contrast our sufficient-statistic based estimate of $\bar{\varepsilon}^d$ to the literature on the asset accumulation response to capital income taxes. Moll, Rachel and Restrepo (2021) provides an overview of this literature and note that the estimates of $\bar{\varepsilon}^d$, which they call the capital-supply elasticity, from Kleven and Schultz (2014), Zoutman (2018), Brülhart, Gruber, Krapf and Schmidheiny (2019) and Jakobsen, Jakobsen, Kleven and Zucman (2020), all lie between 1.25 and 35. Remarkably, this is exactly the range implied by Figure 6 for plausible values of σ. Note in particular that none of the existing empirical estimates are negative, which is consistent with our finding that the income effect tends to be dominated by the substitution effect at all plausible values of σ. Note also that they are all well below infinity, the value implied by all representative-agent models, or models with overlapping generations but dynastic altruism motives (Barro 1974).

3.4 Implications

We are now ready to put together our findings so far to implement Propositions 2 and 3.

We first calculate $\bar{\Delta}^{comp}_{LR}$ from our results in section 3.2. This involves taking the average of the log changes $\bar{\Delta}^{comp,c}$ in each country c, weighted using 2016 wealth weights. For our central population projection, and defining the long-run to be 2100 when the U.N. popu-

Table 1: Change in world interest rate and wealth-to-GDP

	A. $r_{LR} - r_0$			**B.** $\overline{\Delta_{LR} \log \left(\frac{W}{Y}\right)}$		
		σ			σ	
η	0.25	0.50	1.00	0.25	0.50	1.00
0.60	-3.24	-1.59	-0.79	15.6	7.7	3.8
1.00	-2.09	-1.25	-0.70	16.7	10.0	5.6
1.25	-1.71	-1.10	-0.65	17.1	11.1	6.5

lation projections end, this delivers $\bar{\Delta}_{LR}^{comp} = 31.8$ percentage points. This is a substantial displacement of long-run net asset demand from demographic change. Proposition 2 tells us how this is accommodated in equilibrium by a combination of changing returns and additional aggregate wealth accumulation.

Next, we implement formulas (14) and (15). Section 3.3 gives us expressions for the asset and demand semielasticities that enter these formulas directly as a function of η and σ. Since there is a large controversy over the value of these primitive elasticities in the literature, here we limit our task to describing a mapping between η and σ and aggregate outcomes, and commenting on plausible values. We describe this mapping by considering three values for both η and σ. For capital-labor substitution, we consider a low value of $\eta = 0.6$ (Oberfield and Raval, 2021), a high value of $\eta = 1.25$ (Karabarbounis and Neiman, 2014), and the canonical Cobb-Douglas value $\eta = 1$. For the elasticity of intertemporal substitution, we consider $\sigma \in \{0.25, 0.5, 1\}$, which span the range typically considered in the macroeconomics literature.

Table 1 presents our results. The left panel shows the general equilibrium changes in rates of returns, applying equation (14). The right panel shows the average change in log world wealth, in percentage points, applying equation (15).

Our first finding is that equilibrium returns unambiguously fall in response to demographic change. r falls by more when σ or η are low, since these limit the combined responsiveness of asset supply and demand to falling returns. This finding refutes the asset market meltdown and the great demographic reversal hypotheses (Poterba 2001, Goodhart and Pradhan 2020). The combination of a positive $\bar{\Delta}_{LR}^{comp}$ and positive $\bar{\epsilon}^s + \bar{\epsilon}^d$ at any plausible value of σ and η makes this conclusion inevitable: once appropriately normalized by GDP, demographic change must increase asset demand, and only a decline in interest rates can restore the balance of supply and demand in world asset markets.[23] In our central scenario of Cobb-Douglas elasticity and an elasticity of intertemporal subtitu-

[23] In section 6, we explain why thinking of equilibrium in terms of flows rather than stocks can lead one to miss this conclusion.

tion of $1/2$, we obtain a fall in r of 122 basis points by the end of the twenty-first century.

Our second finding is that world wealth increases, though by a lower amount than the 32 percent demographic-induced shift in world asset demand. This finding traces back to our conclusion that $\bar{\epsilon}^d$ must be positive in section 3.3. In principle, if household wealth accumulation was completely insensitive to the decline in the rate of return induced by demographics–or if the income effect dominated, so that they end up accumulating even more assets in response to this decline–then we could see a very large increase in world wealth, accomodated entirely by rising asset supply. In practice, both empirical studies and our sufficient statistic result suggest that the substitution effect ends up dominating quantitatively. In our central scenario, we have $\bar{\epsilon}^s = 8$ and $\bar{\epsilon}^d = 17$, so that $\frac{\bar{\epsilon}^s}{\bar{\epsilon}^d + \bar{\epsilon}^d} \simeq 1/3$, and world wealth rises by an average of 10 log points. Given an average world W/Y of 4.4, this corresponds to a level increase of about 47 percentage point. This number is higher when σ is lower and η is higher, since then the relative contribution of asset supply in world asset market adjustment is higher. Our conclusion that demographics leads to a sizable but not radical increase in the world wealth-to-GDP ratio stands in between predictions by Piketty and Zucman (2014) that continuing declines in the growth rate of population will lead to surging W/Y in the twenty-first century, and the argument by Krusell and Smith (2015) that the representative-agent conclusion of no change in W/Y may be a better prediction given the empirical response of savings rates to changes in the growth rate. We discuss this resolution between points of view further by focusing on savings rate in section 6.

Our final set of results pertain to global imbalances. Here, proposition 3 shows that knowledge of asset supply and demand semielasticities is not as critical: $\epsilon^{c,d}$ and $\epsilon^{c,s}$ only matter to the extent that their sum *differs* across countries, since this then induces heterogeneous responses of NFAs to changes in returns. When these are equal at all points in time, Proposition 3 suggests implementing a projection for $NFAs$ that reads as:[24]

$$\Delta \frac{NFA_t^c}{Y_t^c} \simeq \frac{W_0^c}{Y_0^c} \left(e^{\left(\Delta_t^{comp,c} - \bar{\Delta}_t^{comp}\right)} - 1 \right) \tag{26}$$

Panel A of figure 7 implements this calculation. The solid lines show global imbalances until today for the five large economies discussed in the introduction, and the dashed lines show the projections from equation 26. In the next few decades, we expect to see a widening of existing global imbalances: China's net foreign assets will rise substantially, while the US's will decline. Although these trends flatten mid-century, the

[24] We evaluate the approximation required to use this expression in section 4. Appendix figure A.6 instead applies equation (16) at each point, taking into account the interest rate adjustment and the heterogeneity in elasticities across countries.

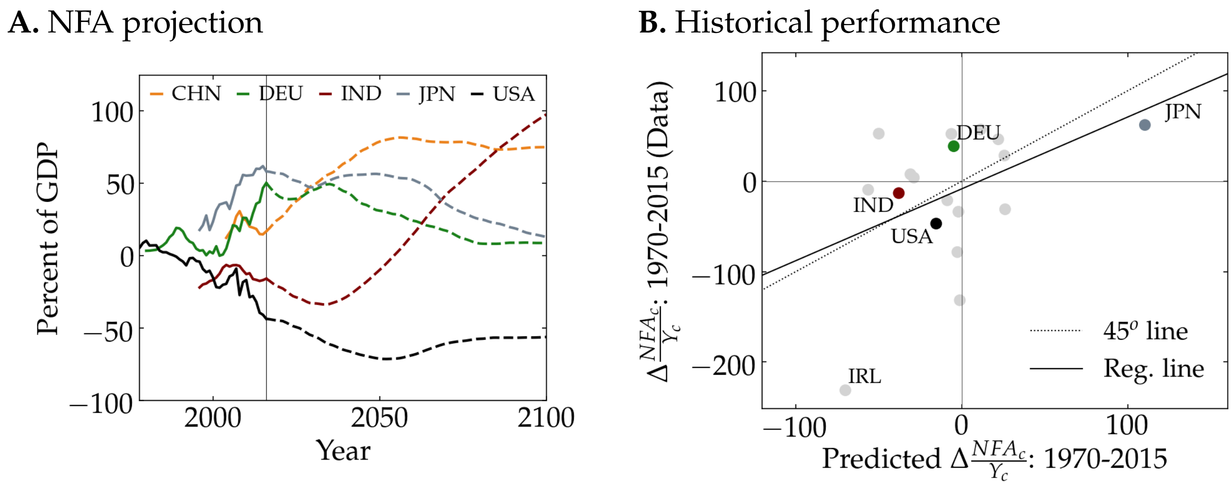


Figure 7: Using the demeaned compositional effect to project NFAs

Notes: Panel A presents the empirical NFA-to-GDP ratio as presented in figure 1 until 2016, and from 2016 on the country-specific demeaned compositional effect until 2100. Panel B compares the shift-share between 1970 and 2015 (x-axis) to the change in NFA from the IMF (y-axis). The black dotted line is a 45^o line.

second half of the 21st century features a conspicuous rise in India's net foreign assets, offset partly by a decline in Germany and Japan, whose demographic transitions at that point are nearly complete. These results traces back to the heterogeneity in compositional effects that we documented in section 3.2, which showed China and India with very large $\Delta_t^{comp,c}$ relative to the world average.

In order to show the potential usefulness of equation (26) in projecting NFAs, we now validate this calculation using historical data. Panel B in Figure 7 shows the projected change in NFAs between 1970 to 2015 plotted against the actual changes in net foreign asset positions during the same time period. For such a simple exercise, the two line up remarkably well, both qualitatively and quantitatively: for instance, Japan had the highest projected rise in its NFA, of around 100pp of GDP, which is actually what occured over this period. Of course, non-demographic forces are also at play (for instance, inflows into Ireland reflecting its status as a tax haven), but this exercise suggest that demographics is in fact an important driver of global imbalances looking backwards (echoing the findings in Backus et al. 2014 and Bárány et al. 2019), and that measuring compositional effects provides a useful approach to forecasting their effects.

4 The compositional effect in a quantitative model

The sufficient statistic analysis in sections 2–3 showed how equilibrium outcomes could be predicted from a small set of moments and parameters. However, while the underlying model in section 2 was rich in some respects, it also abstracted from a number of forces that have been studied in the quantitative demographics literature: for example, bequest motives, changing mortality, changes in the retirement age, and changes in government taxes and transfer schemes.

In this section, we extend the baseline model with these additional features and study how well the sufficient statistic analysis holds up in the extended model. Broadly, we find that the analysis remains a good guide. Qualitatively, the rate of return falls, global wealth increases, and rapidly aging countries accumulate positive net foreign asset positions. Quantitatively, the results are well predicted by the sufficient statistic formulas in propositions 2–4, and these findings are robust to a number of variations in the model. The main exception is when the fiscal adjustment in response to an aging population is one-sided: if the budget is balanced entirely with higher taxes, all aggregate effects of aging are smaller, while if it is balanced entirely with lower benefits, all effects are larger.

4.1 Extending the model

For the household problem, we introduce a number of additional features, inspired by the large literature on quantitative life-cycle models. We also impose a number of parametric restrictions to allow for quantification. Since most of these elements are standard in the literature, we only outline the main features of our model here. Appendix D.1 provides the details.

As in section 2, the model is defined for a set of countries $c \in \mathcal{C}$, which are tied together by an integrated asset market in which net foreign asset positions sum to zero (equation 6). In describing the model, we omit the country superscript c for simplicity. For the production sector, we use the same setup as in section 2, and we also assume that F is a CES production function with elasticity η. The demography component of the model is also as in section 2 with two modifications: survival rates ϕ_{jt} can vary over time and there is an exogenous sequence of migration by age M_{jt}.

To capture the effect of rising longevity and working life, we extend the household model to allow for time-varying age-specific mortality rates ϕ_{jt} and a time-varying retirement policy ρ_{jt}. We also introduce bequests governed by non-homothetic preferences, which helps us explain asset inequality and the limited decumulation of assets at old ages. To reflect the limited share of annuities in aggregate wealth, we remove annuity markets;

households self-insure against mortality risks, with assets remaining at death given as bequests. Last, we assume that there is intergenerational transmission of ability.

The new household problem is

$$\max \mathbb{E}_k \sum_j \beta_j \Phi_{jt} \left[\frac{c_{jt}^{1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}} + \Upsilon Z_t^{\nu-\frac{1}{\sigma}} \left(1-\phi_{jt}\right) \frac{\left(\mathfrak{a}_{jt}\right)^{1-\nu}}{1-\nu} \right] \tag{27}$$

$$\text{s.t.} \quad c_{jt} + \mathfrak{a}_{j+1,t+1} \leq w_t \left((1-\tau_t)\ell_{jt}(z_j)(1-\rho_{jt}) + tr_{jt}(z_j)\right) + (1+r_t)\mathfrak{a}_{j,t} + b_{jt}^r(z_j) \tag{28}$$
$$\mathfrak{a}_{j+1,t+1} \geq -\bar{a} Z_t.$$

Compared to the setup in section 2, the second term in the utility function captures preferences for bequests. Bequest preferences have curvature $\nu \leq \frac{1}{\sigma}$ to allow for non-homothetic preferences, and is scaled with the mortality risk $1-\phi_{jt}$ and a term $Z_t^{\nu-\frac{1}{\sigma}}$ which makes the non-homotheticity consistent with balanced growth. In the budget constraint, $b_{jt}^r(z_j)$ denotes bequests received. The factor $\rho_{jt} \in [0,1]$ denotes a time-varying retirement policy, which captures how much labor individuals are allowed to supply.

Parametrically, we assume that the individual state z_j consists of a permanent component θ, which is Markov across generations, and a transient component ε_j, which is Markov across years. Total labor supply is the product of these two components and a deterministic age profile: $\ell_{jt}(z_j) = \theta\varepsilon_j\bar{\ell}_j$. To determine bequests received $b_{jt}^r(z_j)$, we pool all bequests from parents of each type θ, distributing them across ages j in proportion to a fixed factor F_j, and across types θ' in proportion to the probability that their children have type θ'. Appendix D.1 provides the explicit formula.

For government policy, we assume that transfers reflect the social security system and are given by $tr_{jt}(z_j) = \rho_{jt}\theta d_t$, where d_t models a time-varying replacement rate. The government policy consists of a sequence of retirement policies $\{\rho_{jt}\}$ and a fiscal rule that targets an eventually converging sequence of government debt $\{\frac{B_t}{Y_t}\}$, where the debt sequence is obtained by dynamically adjusting replacement rates d_t, taxes τ_t and consumption G_t (see the appendix for details).

Normalized for productivity growth, the household problem implies a value function

$$\tilde{V}_{jt}(\theta,\epsilon,\tilde{\mathfrak{a}}) = \max_{\tilde{c},\tilde{\mathfrak{a}}'} \frac{\tilde{c}^{1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}} + \Upsilon(1+\gamma)^{1-\nu}\left(1-\phi_{jt}\right)\frac{\left(\tilde{\mathfrak{a}}'\right)^{1-\nu}}{1-\nu} + \frac{\beta_{j+1}}{\beta_j}\phi_{jt}\mathbb{E}\left[\tilde{V}_{j+1,t+1}(\theta,\epsilon',\tilde{\mathfrak{a}}')|\epsilon\right]$$
$$c + (1+\gamma)\tilde{\mathfrak{a}}' \leq \tilde{w}(r_t)\theta\left[(1-\rho_{jt})(1-\tau_t)\bar{\ell}_j\epsilon + \rho_{jt}d_t\right] + (1+r_t)\tilde{\mathfrak{a}} + \tilde{b}_{jt}^r(\theta) \tag{29}$$
$$\mathfrak{a}'(1+\gamma) \geq -\bar{\mathfrak{a}},$$

where a tilde $\sim$ denotes normalization by Z_t, except for $\tilde{V}_{jt} \equiv \frac{V_{jt}}{Z_t^{1-\frac{1}{\sigma}}}$.

4.2 Asset demand and supply in the extended model

Before conducting a full quantification of the extended model, some preliminary theoretical analysis is helpful to structure the comparison with the sufficient statistic analysis in section 2 and 3.

Compared to the basic model in section 2, the extended model differs in that even for a fixed r, demographics affect the wealth-to-output ratio for non-compositional reasons. More precisely, the normalized household problem (29) shows that for a fixed r, there are four reasons for the asset accumulation policy of individuals at a given age to change over time: variation in received bequests $\tilde{b}^r_{jt}(\theta)$, variation in survival rates ϕ_{jt}, variation in tax and benefit policy $\{\tau_t, d_t\}$, and variation in retirement policy ρ_{jt}. These non-compositional forces imply that propositions 2 and 3 do not hold in the extended model, since these propositions relied on the compositional effect being the only shifter of net asset demand.

However, it turns out that the asset demand and supply framework underpinning the propositions still can be applied to the extended model, provided that we replace the compositional effect $\Delta_t^{comp,c}$ with the more general notion of a small-open-economy effect $\Delta_t^{soe,c}$. The latter is defined as the change in the wealth-to-output ratio for a small open economy facing a fixed r over time. Thus, it captures all demographic effects on wealth accumulation—compositional and non-compositional—except for the effects of adjustments in r. We can then prove the following.

Proposition 5. *If the wealth holdings of agents start in a steady state distribution given r_0 and π_0^c, then proposition 2 and 3 hold in the extended model, with $\Delta^{comp,c}$ replaced by $\Delta^{soe,c}$, where $\Delta_t^{soe,c}$ is defined as the change in the wealth-to-output ratio between 0 and t in a small open economy equilibrium with a constant rate of return r_0.*

Proof. See appendix D.2. □

Proposition 5 implies that, to first order, the extended model results can differ from the basic model results either through Δ^{soe} deviating from Δ^{comp} due to non-compositional effects of aging, or the asset demand elasticity $\epsilon^{d,c}$ deviating from the formula in proposition 4.[25] The next two sections conduct a full model calibration and use the proposition to analyze the differences with the sufficient statistic analysis.

[25]Differences in ϵ^s play a minimal role, since $\epsilon^{s,c} \equiv \frac{\eta}{r_0+\delta}\frac{K_0^c}{Y_0^c}$ have the same formula in both the basic and the extended model, and remain exactly the same as long as the model calibration targets r_0 and $\frac{K_0^c}{Y_0^c}$ and uses the same η and δ as the sufficient statistic analysis, which it does.

4.3 Calibration

We calibrate a world economy consisting of the 25 economies from section 3. To obtain parameters for each country, we calibrate a steady-state version of our model to 2016 data. Starting from this steady state, we then simulate the model from 2016 onward given demographic projections.

Steady-state calibration procedure Appendix D.3 spells out the steady-state version of our model, which is for the most part standard.[26] The main calibration results are displayed in table 2. For parameters that are common across countries, we display the world value. Country-specific parameters have a c-superscript, and the US values are displayed for illustration. The detailed calibration procedure is described in appendix D.4, with key points discussed below.

The real rate of return r is the 2016 value from figure 1 in the introduction, with the calculation described in appendix A. For assets, we follow section 3 in taking the wealth-to-output ratio W^c/Y^c from the World Inequality Database (WID). We use data from the IMF to obtain country-specific debt levels B^c/Y^c and net foreign asset positions NFA^c/Y^c, adjusting to ensure that $\sum_c NFA^c = 0$. The capital-output ratio is obtained residually as $K^c/Y^c = W^c/Y^c - B^c/Y^c - NFA^c/Y^c$.[27]

On the production side, we set the elasticity of substitution between labor and capital to unity, $\eta = 1$. Countries have a common labor-augmenting growth rate γ calibrated to the average growth in output per labor unit $\frac{Y_t^c}{L_t^c}$ between 2000 and 2016. The common depreciation rate is calibrated to match aggregate capital consumption from the Penn World Table given the capital stocks calibrated above. Given these parameters, we obtain the investment to output ratio and the labor share in each country from $\frac{K^c}{Y^c}$ and the country-specific growth rate $g^c \equiv (1+n^c)(1+\gamma^c) - 1$.

For government policy, we assume that all countries have a discrete retirement policy, with $\rho_j^c = 0$ for $j < T^{r,c}$ and $\rho_j^c = 1$ for $j \geq T^{c,r}$, where $T^{c,t}$ is the retirement age. The retirement age is calibrated to the effective age of labor market exit, which we define

[26] The main non-standard element is a counterfactual flow of migrants, which is introduced to ensure that the steady state implied by the 2016 birth and death rates can exactly match the observed age distribution in 2016. This method is similar to the one used in Penn Wharton Budget Model (2019), and is one way to address a generic problem in the calibration of steady-state demographic models, which is that observed mortality and population shares might not be consistent with a stationary population distribution. Beyond the initial steady state, this adjustment is not needed, and to simulate the dynamics after 2016, we use the migration flows given in demographic projections.

[27] Note that the implied K/Y for the US is high relative to standard measures of capital stock. Our methodology implicitly assumes that unmeasured capital accounts for this gap. An alternative procedure would be to explain the gap using markups.

Table 2: Calibration parameters

Parameter	Description	U.S.	All	Source
Demographics				
T^w, T	Initial and terminal ages		$20, 95$	
n^c	Population growth rate	0.6%		UN World Population Prospects
π_j^c	Population distribution			UN
ϕ_j^c	Survival probabilities			UN
Returns and assets				
r	Real return on wealth		3.9%	Described in appendix A
W^c/Y^c	Total wealth over GDP	438%		WID
B^c/Y^c	Debt over GDP	106.8%		IMF
NFA^c/Y^c	Net foreign assets	-35.8%		IMF
K^c/Y^c	Capital over GDP	367%		$\frac{W^c}{Y^c} - \frac{B^c}{Y^c} - \frac{NFA^c}{Y^c}$
Production side				
I^c/Y^c	Investment over GDP	30.9%		$\frac{K_c}{Y_c}(\delta + g_c)$
α^c	Constant in prod. fn.	0.356		$(r+\delta)\left(\frac{K^c}{Y^c}\right)^{\frac{1}{\eta}}$
$s^{L,c}$	Labor share	0.64		$1-(r+\delta)\frac{K^c}{Y^c}$
δ	Depreciation rate		5.79%	$\sum_c \delta^c K^c$ (PWT) divided by $\sum_c K^c$
γ	Technology growth		2.03%	World average 2000-16 from $\frac{Y_t}{\sum N_{jt} h_{j0}}$
η	K/L elasticity of subst.		1	Standard
Government policy				
$T^{r,c}$	Retirement age	66		OECD
G^c/Y^c	Consumption over GDP	12.5%		Government budget
$\overline{d}^c$	Social security benefits	71.3%		Benefits-to-GDP from OECD
τ^c	Labor tax rate	31.6%		Balanced total budget
Income process				
χ_ϵ	Idiosyncratic persistence		0.91	Auclert and Rognlie (2018)
υ_ϵ	Idiosyncratic std. dev.		0.92	Auclert and Rognlie (2018)
χ_θ	Intergenerational persist.		0.677	De Nardi (2004)
υ_θ	Intergenerational std. dev.		0.61	De Nardi (2004)
$\underline{a}$	Borrowing limit	0		
Preferences				
σ	EIS		0.5	Standard
$\bar{\beta}^c$	Discount factor process	1.044		See text
ξ^c	Discount factor process	0.00063		See text
Υ^c	Bequests scaling factor	67.95		See text
ν	Bequest curvature		1.32	See text

using information from the OECD and from the labor income profiles.[28] We define the income tax rate τ using OECD data on average tax wedge on personal earnings. Transfers capture the social security system, and satisfies $tr^c(z_j) = \rho_j \theta d^c$, where we calibrate the social security system replacement rate d^c by targeting country-specific benefit-to-GDP ratios net of taxes (OECD, 2019a). Government consumption G^c/Y^c is adjusted to ensure a constant debt-to-output ratio.

For the income process, we use average labor income by age to target the deterministic component of labor supply $\bar{\ell}_j$ for all ages before retirement, $j < T^{r,c}$.[29] For the idiosyncratic term z, the log transient component follows an AR(1) process over the life-cycle, and the log permanent component follows an AR(1) process across generations. The parameters of the AR(1) processes are taken from Auclert and Rognlie (2018) and De Nardi (2004). We assume that the distribution of bequests received across ages F_j is common across countries, and we match it to the age distribution of bequests received in the Survey of Consumer Finances.

The remaining preference parameters are the elasticity of intertemporal substitution σ, the time preference β_j, and the weight and curvature on bequests (Υ, ν). We assume that parameters σ, Υ and ν are common across countries, while the level shifters β_j are allowed to vary across countries according to a quadratic formula $\log \beta_j^c = -j \times \log \bar{\beta}^c + \xi^c (j - 40)^2$, where $\xi^c = 0$ corresponds to exponential discounting. Our calibration first sets σ to 0.5 in line with section 3. To discipline the common Υ and ν, we set them jointly with the parameters of US time discount values β_j^{US} to minimize the squared distance to the US profile of wealth-by-age and bequest-to-GDP ratio, subject to the constraint of precisely matching the US aggregate wealth to GDP ratio.[30] For all other countries, the parameters of β_j^c are set to target the profile of wealth-by-age, again subject to the constraint of exactly matching the wealth-to-output ratio.

Table 3 provides complementary information about the calibration outcomes for the 12 largest economies. The successful fit of the long-run compositional effect reflects the good fit of the labor and wealth profiles. In the appendix, we provide additional information about the calibration, including the fit of labor and wealth profiles and the main

[28]Our main source is the OECD's data on "effective age of labor market exit" (OECD, 2019b). For 7 countries, the age provided by the OECD implies that labor market exit happens after the age at which aggregate labor income falls below implied benefit income. In those cases, we define the latter age as the date of labor market exit. See the appendix for details.

[29]For $j \geq T^{c,r}$, $\bar{\ell}_j$ is calibrated from age-j labor earnings, scaled up by $\frac{LFPR_{T^{r,c}}}{LFPR_j}$ to compensate for labor force participation at j being depressed by retirement. Since $(1-\rho_j)\bar{\ell}_j = 0$ for all $j \geq T^{c,r}$, this value does not matter for steady state, but will matter in simulations where the retirement is increased.

[30]The US bequest-to-GDP ratio is from Alvaredo, Garbinti and Piketty (2017), and in the appendix, we also validate the model to the inequality of bequests taken from Hurd and Smith, 2002).

Table 3: World economy calibration

	Δ_c^{comp}		Components of wealth			Government policy	
Country	Model	Data	$\frac{W_c}{Y_c}$	$\frac{B_c}{Y_c}$	$\frac{NFA_c}{Y_c}$	τ^c	$\frac{Ben_c}{Y_c}$
AUS	30	29	5.09	0.40	-0.46	0.29	0.04
CAN	21	20	4.63	0.92	0.20	0.31	0.04
CHN	47	45	4.20	0.44	0.25	0.30	0.04
DEU	21	20	3.64	0.69	0.58	0.50	0.10
ESP	42	37	5.33	0.99	-0.74	0.39	0.10
FRA	31	30	4.85	0.98	-0.05	0.48	0.13
GBR	27	26	5.35	0.88	0.08	0.31	0.06
IND	65	56	4.16	0.68	-0.08	0.30	0.01
ITA	34	30	5.83	1.31	-0.02	0.48	0.13
JPN	24	22	4.85	2.36	0.66	0.32	0.09
NLD	34	33	3.92	0.62	0.70	0.37	0.05
USA	32	29	4.38	1.07	-0.36	0.32	0.06

parameters for all 25 economies.

4.4 Simulations and results

The steady-state calibration pins down the household parameters, the production parameters, and the initial state of all economies. To study the effect of demographic change, we feed the economy with paths for all demographic variables from the UN World Population Prospects for 2016 to 2100. We are interested in how wealth levels, rates of return, and net foreign asset positions evolve, and how this evolution relates to the compositional findings from section 3.

Formally, we assume that the world economy has reached a stationary equilibrium at 2300 and we solve for the transition dynamics between 2016 and 2300. Our experiments hold preferences and the aggregate production function constant, but government policy instruments change over time as aging creates fiscal shortfalls that need to be compensated. In our main specification, we assume that the retirement age in all countries increases by one month per year over the first 60 years of the simulation (in line with CBO's projection for the US), and that the government operates a fiscal rule that keeps the debt-to-output ratio constant by relying equally on tax increases, benefit cuts, and government consumption reductions.

Changes in r and W/Y. Table 4 reports the simulation results for Δr and $\overline{\Delta \log W/Y}$, together with the corresponding average compositional effect $\bar{\Delta}_{comp}$, average small open

Table 4: Compositional and extended model results: 2016–2100

	Δr	$\Delta \frac{\bar{W}}{\bar{Y}}$	$\bar{\Delta}^{comp}$	$\bar{\Delta}^{soe}$	$\bar{\epsilon}^d$	$\bar{\epsilon}^s$
Pure compositional analysis	-1.25	10.0	31.8		17.3	8.0
Preferred model specification	-1.23	10.3	34.1	30.3	17.1	8.0
Alternative model specifications						
+ Constant bequests	-1.18	10.0	34.1	27.0	14.9	8.0
+ Constant mortality	-1.23	10.9	34.1	27.1	13.9	8.0
+ Constant taxes and transfers	-1.33	11.9	34.1	30.1	14.6	8.0
+ Constant retirement age	-1.49	13.4	34.1	34.1	14.7	8.0
+ No income risk	-1.47	13.2	33.9	33.9	13.8	8.0
+ Annuities	-1.33	11.5	34.2	34.2	17.2	8.0
Alternative fiscal rules						
Only lower expenditures	-1.29	11.0	34.1	32.6	17.9	8.0
Only higher taxes	-0.88	6.7	34.1	19.4	14.6	8.0
Only lower benefits	-1.50	12.9	34.1	39.1	18.4	8.0

Notes: Δr and $\overline{\Delta \log \frac{W}{Y}}$ are reported in percentage points.

economy effect $\bar{\Delta}_{soe}$, and average asset demand and supply semi-elasticities $\bar{\epsilon}^d$ and $\bar{\epsilon}^s$.[31] The preferred model results are presented on the second line, with the first line reproducing the results from the sufficient statistic analysis as a point of comparison.

Overall, the model results are close to the sufficient statistic analysis, with a $\Delta r = -1.23\%$ compared to -1.25% in the sufficient statistic analysis, and $\overline{\Delta \log W/Y} = 10.3\%$ compared to 0.101 in the sufficient statistic analysis. The formulas $\Delta r = -\frac{\bar{\Delta}^{soe}}{\bar{\epsilon}^d + \bar{\epsilon}^s}$ and $\overline{\Delta \log W/Y} = \frac{\bar{\epsilon}^s}{\bar{\epsilon}^d + \bar{\epsilon}^s} \bar{\Delta}_{comp}$ from proposition 5 provide an excellent approximation to the full model results, predicting $\Delta r \approx -1.21\%$ and $\overline{\Delta \log W/Y} \approx 9.7\%$, which are within 3 basis points for the interest rate, and within 0.3 percentage points for the wealth-to-output ratio.

Given the success of the first order approximation formulas, the close correspondence between the full model analysis and the sufficient statistic results reflect three facts: a) that the model calibration successfully approximates the average compositional effect $\bar{\Delta}^{comp}$, b) non-compositional effects of aging are relatively small on average, as indicated by the small differences between $\bar{\Delta}^{comp}$ and $\bar{\Delta}^{soe}$, and c) the model asset demand sensitivity $\bar{\epsilon}^d$ is

[31] Here, $\bar{\Delta}^{comp}$ is calculated as in section 3, and we construct $\bar{\Delta}^{soe}$ by simulating the model for each country given a fixed r_0. For each country, the sensitivities $\epsilon^{d,c}$ and $\epsilon^{s,c}$ are obtained by perturbing r at a small open economy steady state constructed with 2100 demographics, and calculating effect on steady-state W/Y and K/Y.

relatively close to that implied by proposition 4.[32] The small deviations from the sufficient statistic results are also partly offsetting, with the model overstating $\bar{\Delta}^{comp}$ by 2 percentage points, but then having $\bar{\Delta}^{soe}$ 4 percentage points lower than $\bar{\Delta}^{comp}$.

For a), the model's close fit to the compositional effect follows from our calibration strategy. In the initial period, for each country, we directly match average labor income at each age, and calibrate preferences to achieve a good fit to the profile of average wealth across ages. Since our simulation then matches the exact projected change in the age distribution, we closely approximate Δ^{comp}.

To unpack the economic forces behind b) and c), we sequentially shut off the forces that distinguish the full model from the baseline model underlying the sufficient statistic result. This process takes six rows in table 4. The first four leave the initial calibration intact but shut off dynamic changes: first holding constant bequests received, then perceived mortality, taxes and transfers, and retirement age.[33] The last two involve changes to the steady-state calibration itself, first shutting off income risk, and then replacing bequests at death with annuities. By the final row, we have nearly recovered the baseline model, with the only difference being that our calibration is not flexible enough to perfectly hit $\bar{\Delta}^{comp}$.

There is little difference between $\bar{\epsilon}^d$ in the full model and the baseline model, reflecting two offsetting forces. In the full model ϵ^d is pushed higher due to the presence of bequests, which let increased savings in response to r accumulate across generations. At the same time, ϵ^d is pushed lower by the absence of annuities allowing for insurance against mortality risk; since households must self-insure against this risk, their asset accumulation is less sensitive to r. In table 4, after both forces are shut off by removing bequests and introducing annuities, $\bar{\epsilon}^d$ is on net almost unchanged.

In the model, the difference between $\bar{\Delta}^{soe}$ and $\bar{\Delta}^{comp}$ reflects three small, and partially offsetting, demographic forces on net which make households hold less wealth even at fixed r. First, $\bar{\Delta}^{soe}$ is high because bequests grow when lower fertility implies fewer heirs that split the bequest. When we shut off this effect by holding bequests received constant, $\bar{\Delta}^{soe}$ falls from 30 to 27. Second, fiscal adjustments in response to demographic change—a combination of lower expenditures, higher taxes, and lower social security benefits—lead households to save slightly less on net. When we shut off the latter two adjustments

[32]The $\bar{\epsilon}^s$ are identical across all settings since it is only a function of external parameters and moments that are targeted in the calibration.

[33]To make constant bequests received consistent with equilibrium, we assume that government taxes/augments bequests to keep them at the initial level. For mortality, we assume that the mortality rates perceived by individuals ex ante are held constant, while the population still evolves according to objective mortality rates that can change over time. For all these changes to the model, we assume that governments adjust G_t to maintain a constant debt level.

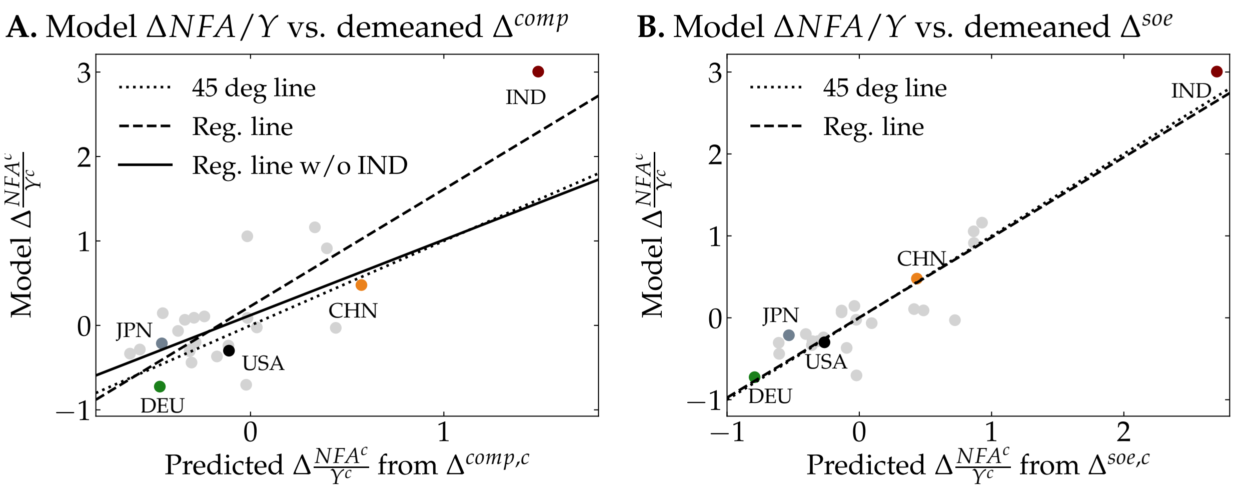


Figure 8: Predicting change in net foreign asset position

Notes: Panel A presents the model-implied change in NFA/Y between 2016 and 2100 on the y-axis, and on the x-axis the change in NFA/Y predicted from the demeaned model compositional effect, $NFA/Y \approx \exp(\Delta^{comp,c} - \bar{\Delta}^{comp}) - 1$, over the same period. The dotted line is a 45 deg line. The dashed line is a regression line, and the solid line is this same regression line when India is excluded. Panel B also shows the model $\Delta NFA/Y$ on the y-axis, but the x-axis presents the change in NFA/Y predicted from the demeaned model small open economy effect, $NFA/Y \approx \exp(\Delta^{soe,c} - \bar{\Delta}^{soe}) - 1$.

and achieve fiscal balance solely through lower expenditures, $\bar{\Delta}^{soe}$ recovers from 27 to 30. Finally, when households' working life is extended and they retire later, they need to save less for retirement. Removing this effect by holding the retirement age constant, $\bar{\Delta}^{soe}$ rises from 30 to 34, and now agrees exactly with the compositional effect $\bar{\Delta}^{comp}$.

This close agreement between $\bar{\Delta}^{soe}$ and $\bar{\Delta}^{comp}$ is robust to many features of the model, but does not hold when fiscal adjustment is very one-sided. If the fiscal shortfall from an aging population is closed entirely with higher taxes, asset accumulation falls, since households have less after-tax income from which to save, and $\bar{\Delta}^{soe}$ declines to 19. If, alternatively, the shortfall is closed entirely by cutting social security benefits, asset accumulation rises, since households must fund more of their own retirement. In our calibration, the relatively small net effect of fiscal adjustments on $\bar{\Delta}^{soe}$ reflects that our fiscal rule uses an even mix of different adjustment margins to close the fiscal shortfall.[34]

Changes to net foreign asset positions. Figure 8 summarizes the model's predictions for the change in net foreign asset position in each country from 2016–2100. Panel A compares the full model findings to the method used in section 3 by plotting the full

[34]These results echo the findings in the pension reform literature about the importance of fiscal adjustment choices for macroeconomic outcomes (see, for example, Feldstein 1974, Auerbach and Kotlikoff 1987, and Kitao 2014).

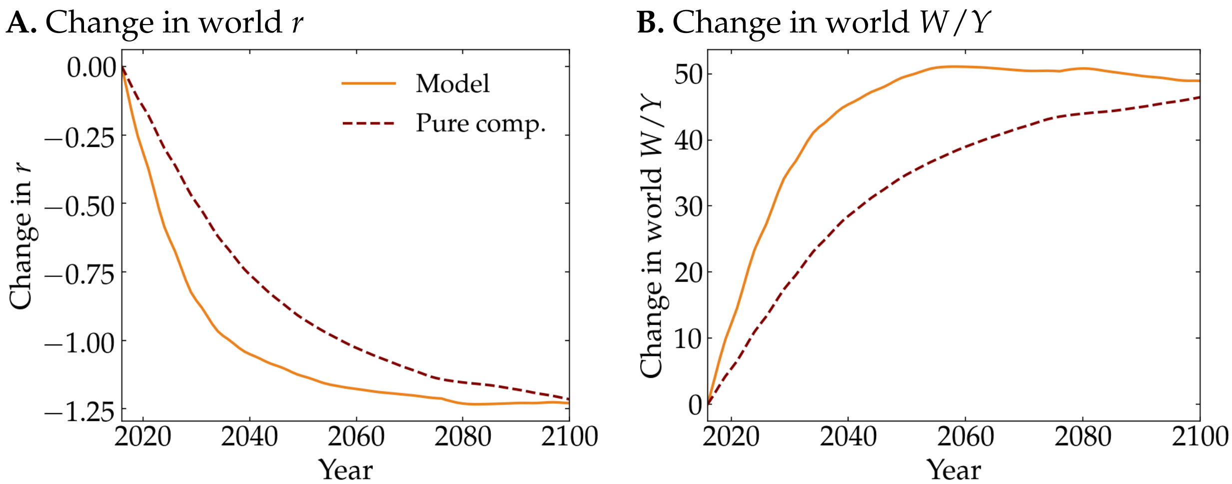


Figure 9: Transition dynamics for rates of return and wealth

Notes: This figure presents the model change in world interest rate and wealth-to-GDP between 2016 and 2100. The solid line corresponds to the model simulations from our preferred model specification and the dashed line to the sufficient statistic formulas $\Delta r = \frac{\bar{\Delta}_t^{comp}}{\bar{\epsilon}^d + \bar{\epsilon}^s}$ and $\frac{W_0}{Y_0}\overline{\Delta \log W/Y} = \frac{W_0}{Y_0}\frac{\bar{\epsilon}^s}{\bar{\epsilon}^d + \bar{\epsilon}^s}\bar{\Delta}_t^{comp}$.

model results on the vertical axis, and the prediction based on demeaned compositional effects $\Delta^{comp,c} - \bar{\Delta}^{comp}$ on the horizontal axis. The compositional predictions are generally quite accurate, and the line of best fit excluding India is close to 45 degrees. In India, however, the model predicts even larger net foreign asset position growth than expected from the compositional effect.

Panel B shows that this discrepancy disappears, and the fit is even closer, when we use the demeaned small open economy effect $\Delta^{soe,c}$ for predictions on the horizontal axis instead. This shows that discrepancies in panel A, including for India, are mostly due to the non-compositional effects $\Delta^{soe,c} - \Delta^{comp,c}$ of aging in our model, rather than non-linearities or heterogeneity in elasticities.

Transition dynamics. Using the calibrated model, we can also solve for the transition dynamics for world r and W/Y, displayed in figure 9. To test the how well the long-run sufficient statistic formulas in propositions 2–3 work at different horizons, we apply them at each date t, using the compositional effects Δ_t^{comp} together with the long-run elasticities ϵ^d and ϵ^s.

As we already know from table 4, the two series nearly coincide by 2100. Their dynamics are also quite similar, but the model predicts a somewhat faster decline in r and rise in W/Y. Both phenomena reflect that the long-run sensitivities ϵ^d overstate the short-run sensitivity of asset accumulation to interest rates. For r, this implies that interest rates

Table 5: Decomposing the change in equilibrium r in existing papers

	Eggertsson et al. (2019)	Gagnon et al. (2021)
Time-period	1970-2015	1970-2015
GE transition		
Δr^{GE}	-3.19%	-0.71%
First order approximation $\Delta r^{ss} = \frac{-\Delta^{soe}}{\epsilon^d + \epsilon^s}$		
Δr^{ss}	-3.46%	-0.49%
Δ^{comp}	31.8 pp	14.6 pp
$\Delta^{soe} - \Delta^{comp}$	16.0 pp	10.9 pp
ϵ^s	5.2	9.5
ϵ^d	8.6	42.2

Notes: This table presents results from Eggertsson et al. (2019) and Gagnon et al. (2021). The first line presents the general equilibrium change in the interest rate reported by the authors. The next lines perform the first-order approximation for the change in the interest rate. We use the authors' codes to compute the compositional effect, the small open economy effect, and the semi-elasticities in their respective models.

have to fall more in the short run to clear the asset market. For W/Y, this implies that asset supply is responsible for more of the adjustment, since the supply adjustment is instantaneous in our model.

5 Interpreting disparate findings in the literature

Proposition 5 shows that the effects of demographics can be interpreted using a demand and supply framework, where changes in r and W/Y reflect changes in net asset demand Δ^{soe} filtered through the semi-elasticities ϵ^d and ϵ^s. If different papers find different effects of demographics, these objects provide a natural diagnostic that can be used to identify the economic origin of the disagreement.

Table 5 reports the result of applying this diagnostic method to Eggertsson et al. (2019) (EMR) and Gagnon et al. (2021) (GJLS). These two papers both use a closed-economy general equilibrium model to study the effect of recent demographic trends on US real interest rates. However, the results are very different, with EMR finding that demography has reduced real interest rates by 3.19 percentage points, while GJLS finds an effect of only -0.71 percentage points.[35] The decomposition in table 5 shows that the steady-state approximation from proposition 5 explains the results well, and that the differences in

[35]Our calculations use replication code for the two papers. For GJLS, we conduct the same experiment as in the paper; for EMR, we isolate the effect of demography by using an experiment which keeps constant markups, the growth rate of TFP, the debt-to-GDP ratio, and the relative price of capital. See appendix E for details on our exercise.

Δr primarily reflect differences in ϵ^d. Indeed, while both papers successfully target the compositional effect and have similar non-compositional effects $\Delta^{soe} - \Delta^{comp}$, EMR has a dramatically lower ϵ^d than GJLS: 8.6 compared to 42.2. In conjunction with a lower supply semi-elasticity, this means that EMR finds a much larger fall in r.

6 Savings

So far in the paper, we have focused our attention on private wealth holdings, decomposing it into the wealth holdings of agents of different age groups per equation (2), and then showing the importance of the compositional effects that this perspective implies for the general equilibrium determination of interest rates, wealth and global imbalances.

An alternative perspective focuses on private savings instead. First, there is a tradition in the literature of computing age-specific savings rates (Summers and Carroll 1987, Auerbach and Kotlikoff 1990, Bosworth et al. 1991). This tends to show that savings rates fall in old ages. Combined with an increasing number of old agents, it is natural to then conjecture than the savings rate will fall with demographics, as Lane (2020) does. From this, Lane (2020) and Goodhart and Pradhan (2020) conclude that demographics will lead to rising rates.

In this section, we first show that the first two conclusions are warranted. We then show that the third is not, since it relies on a flow approach to asset market equilibrium, and ignores the important effect that a falling population growth rate has in translating flows to stocks.

To do this, we return to our model of section 2 to derive a measure of projected savings rates from composition. For the small open economy model, it is simple to prove the equivalent of Proposition 1 for the aggregate savings rate S_t/Y_t: it is given by

$$\frac{S_t}{Y_t} \propto \frac{\sum_j \pi_{jt} s_{j0}}{\sum_j \pi_{jt} h_{j0}}, \tag{30}$$

where s_{j0} is net personal savings by age in the cross-section at date 0.[36] Given (30), changes in the savings rate over time are purely compositional. The analogue to equation (31) is

$$\log\left(\frac{S_t}{Y_t}\right) - \log\left(\frac{S_0}{Y_0}\right) = \log\left(\frac{\sum \pi_{jt} s_{j0}}{\sum \pi_{jt} h_{j0}}\right) - \log\left(\frac{\sum \pi_{j0} s_{j0}}{\sum \pi_{j0} h_{j0}}\right) \equiv \Delta_t^{comp,s}. \tag{31}$$

[36] That is, $s_{j0} = \mathbb{E} s_{j0}$ where $s_{j0} \equiv r_0 a_{j0} + w_0 \left((1-\tau)\ell(z_j) + tr(z^j)\right) - c_{j0}$ is savings for an individual of age j in state (z^j, a_{j0}) at time 0

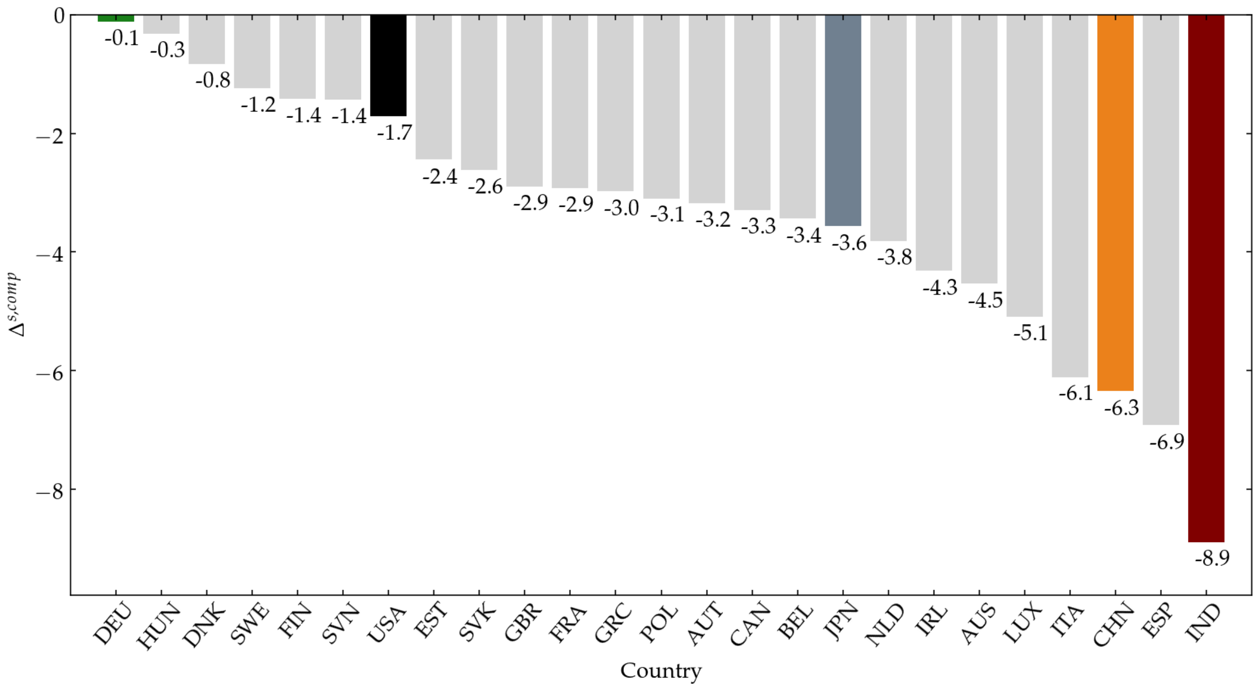


Figure 10: Compositional effects and savings

Notes: Each bar shows the value of the predicted change in the savings-to-GDP ratio from the compositional effect between 2016 and 2100 across countries, calculated using equation (22), reported in level differences.

We implement this calculation in appendix F.[37] Figure 10 summarizes our results, showing the implied change in projected savings rate until 2100. We find that indeed, projected savings from compositional effects fall in every country. This confirms the general view that population aging has progressed sufficiently that the fact that there is an increasing number of elderly agents with lower savings rates is becoming a dominant force.

How is this consistent with our projections of rising wealth-to-GDPs and falling interest rates? It turns out that the flow perspective is misleading to derive general equilibrium consequences. It is Δ_t^{comp}, not $\Delta_t^{comp,s}$, that enters the equations for interest rates in Proposition 2. There is a critical difference between the two, as can be seen from the steady state relationship: in our model, the steady state wealth-to-output ratio W/Y relates to the net savings rate S/Y through the familiar equation

$$\frac{W}{Y} = \frac{S/Y}{g}, \tag{32}$$

[37] There, we show that it is possible to calculate $\Delta_t^{comp,s}$ purely from cross-sectional profiles of assets, so we do not need any direct information on savings rates by age, which are subject to much more measurement error.

so that steady-state statement about savings can be translated into a steady-state statement about wealth, and vice versa. In each of the countries we consider, as demographics causes s to fall, it also causes g to fall due to falling population growth. It turns out that the latter effect dominates, so that W/Y unambiguously rises as S/Y unambiguously falls–and it is the former effect that matters for projecting the effects on interest rates.[38]

7 Conclusion

We use a sufficient statistic approach to quantify the effects of population aging on wealth accumulation, equilibrium interest rates, and capital flows. A simple calculation that rolls forward projected population distributions over fixed age profiles of assets and income constitutes a sufficient statistic for the transition dynamics of the wealth-to-GDP ratio in a special case of our model. Our sufficient statistic approach shows that the macroeconomic effect of aging on aggregate wealth accumulation is large and heterogeneous across countries. This calculation remains a very useful input into the calculation of equilibrium wealth and interest rates away from this special case as it approximates closely the evolution of wealth-to-GDP due to demographic change at fixed interest rate. In an integrated economy, population aging will push the equilibrium rate of return down, and generate large global imbalances.

References

Abel, Andrew B., "Will Bequests Attenuate the Predicted Meltdown in Stock Prices When Baby Boomers Retire?," *Review of Economics and Statistics*, November 2001, *83* (4), 589–595.

__ , "The Effects of a Baby Boom on Stock Prices and Capital Accumulation in the Presence of Social Security," *Econometrica*, March 2003, *71* (2), 551–578.

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

[38]There is an alternative reason to prefer the stock approach, which is that it is less sensitive to measurement error than the flow approach. The flow approach relies on the measurement of personal savings, which is the difference between disposable income and consumption, two large and independently measured quantities. Small relative measurement errors in income and consumption translate into large relative errors in the measurements of savings. For example, in 2019, aggregate US personal savings were 6.1% of GDP – the difference between 76.7% of GDP in disposable personal income and 70.6% of GDP in personal outlays.[39] If actual income were only one percentage point higher, and actual consumption only one percentage point lower, the actual net savings rate would be 8.1% instead of 6.1%. This shows that 1-2% of measurement error in income and consumption can translate to a 30% mismeasurement in the savings rate. With wealth measurement, there is no analogous amplification of error.

Alvarez, Fernando, Hervé Le Bihan, and Francesco Lippi, "The Real Effects of Monetary Shocks in Sticky Price Models: A Sufficient Statistic Approach," *American Economic Review*, October 2016, *106* (10), 2817–2851.

Ameriks, John and Stephen P. Zeldes, "How Do Household Portfolio Shares Vary with Age," *Manuscript*, 2004.

Arkolakis, Costas, Arnaud Costinot, and Andrés Rodríguez-Clare, "New Trade Models, Same Old Gains?," *American Economic Review*, February 2012, *102* (1), 94–130.

Auclert, Adrien, "Monetary Policy and the Redistribution Channel," *American Economic Review*, June 2019, *109* (6), 2333–2367.

_ and Matthew Rognlie, "Inequality and Aggregate Demand," Working Paper 24280, National Bureau of Economic Research, February 2018.

_ , _ , and Ludwig Straub, "The Intertemporal Keynesian Cross," Working Paper 25020, National Bureau of Economic Research, September 2018.

Auerbach, Alan J. and Laurence J. Kotlikoff, *Dynamic Fiscal Policy*, Cambridge: Cambridge University Press, April 1987.

_ and _ , "Demographics, Fiscal Policy, and US Saving in the 1980s and Beyond," *Tax Policy and the Economy*, 1990, *4*, 73–101.

Backus, David, Thomas Cooley, and Espen Henriksen, "Demography and low-frequency capital flows," *Journal of International Economics*, 2014, *92*, S94–S102.

Baley, Isaac and Andrés Blanco, "Aggregate Dynamics in Lumpy Economies," *Econometrica*, 2021, *89* (3), 1235–1264.

Bárány, Zsofia, Nicholas Coeurdacier, and Stéphane Guibaud, "Capital Flows in an Aging World," *Manuscript*, July 2019.

Barro, Robert J., "Are Government Bonds Net Wealth?," *Journal of Political Economy*, November 1974, *82* (6), 1095–1117.

Berger, David, Veronica Guerrieri, Guido Lorenzoni, and Joseph Vavra, "House Prices and Consumer Spending," *Review of Economic Studies*, July 2018, *85* (3), 1502–1542.

Bloom, David, David Canning, and Jaypee Sevilla, *The Demographic Dividend: A New Perspective on the Economic Consequences of Population Change*, RAND Corporation, February 2003.

Börsch-Supan, Axel, Alexander Ludwig, and Joachim Winter, "Ageing, Pension Reform and Capital Flows: A Multi-Country Simulation Model," *Economica*, June 2006, *73* (292), 625–658.

Bosworth, Barry, Gary Burtless, and John Sabelhaus, "The Decline in Saving: Evidence from Household Surveys," *Brookings Papers on Economic Activity*, 1991, *1991* (1), 183–256.

Brülhart, Marius, Jonathan Gruber, Matthias Krapf, and Kurt Schmidheiny, "Taxing Wealth: Evidence from Switzerland," Working Paper 22376, National Bureau of Economic Research, October 2019.

Carroll, Christopher D., "The Method of Endogenous Gridpoints for Solving Dynamic Stochastic Optimization Problems," *Economics Letters*, 2006, *91* (3), 312–320.

Carvalho, Carlos, Andrea Ferrero, and Fernanda Nechio, "Demographics and Real Interest Rates: Inspecting the Mechanism," *European Economic Review*, September 2016, *88*, 208–226.

Catherine, Sylvain, Max Miller, and Natasha Sarin, "Social Security and Trends in Wealth Inequality," SSRN Working Paper No 3546668, Social Science Research Network, Rochester, NY, February 2020.

Chetty, Raj, "Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods," *Annual Review of Economics*, January 2009, *1* (1), 451–488.

Cutler, David M., James M. Poterba, Louise M. Sheiner, and Lawrence H. Summers, "An Aging Society: Opportunity or Challenge?," *Brookings Papers on Economic Activity*, 1990, *1990* (1), 1–73.

Deaton, Angus, *The Analysis of Household Surveys a Microeconometric Approach to Development Policy*, Baltimore, MD: Johns Hopkins University Press, 1997.

Deaton, Angus S. and Christina H. Paxson, "Saving, Inequality and Aging: An East Asian Perspective," *Asia-Pacific Economic Review*, 1995, *1*, 7–19.

Domeij, David and Martin Flodén, "Population Aging and International Capital Flows," *International Economic Review*, July 2006, *47* (3), 1013–1032.

Eggertsson, Gauti B., Jacob A. Robbins, and Ella Getz Wold, "Kaldor and Piketty's Facts: The Rise of Monopoly Power in the United States," Working Paper 24287, National Bureau of Economic Research, February 2018.

—, Neil R. Mehrotra, and Jacob A. Robbins, "A Model of Secular Stagnation: Theory and Quantitative Evaluation," *American Economic Journal: Macroeconomics*, January 2019, *11* (1), 1–48.

Farhi, Emmanuel and François Gourio, "Accounting for Macro-Finance Trends: Market Power, Intangibles, and Risk Premia," *Brookings Papers on Economic Activity*, November 2018, *2018* (Fall).

Feldstein, Martin, "Social Security, Induced Retirement, and Aggregate Capital Accumulation," *Journal of Political Economy*, September 1974, *82* (5), 905–926.

Gagnon, Etienne, Benjamin K. Johannsen, and David López-Salido, "Understanding the New Normal: The Role of Demographics," *IMF Economic Review*, March 2021.

Geanakoplos, John, Michael Magill, and Martine Quinzii, "Demography and the Long-Run Predictability of the Stock Market," *Brookings Papers on Economic Activity*, 2004, *2004* (1), 241–307.

Gomme, Paul, B. Ravikumar, and Peter Rupert, "The Return to Capital and the Business Cycle," *Review of Economic Dynamics*, April 2011, *14* (2), 262–278.

Goodhart, Charles and Manoj Pradhan, *The Great Demographic Reversal: Ageing Societies, Waning Inequality, and an Inflation Revival*, 1st ed., Cham, Switzerland: Palgrave Macmillan, August 2020.

Hall, R. E., "Technical Change and Capital from the Point of View of the Dual," *Review of Economic Studies*, 1968, *35* (1), 35–46.

Henriksen, Espen R., "A Demographic Explanation of U.S. and Japanese Current Account Behavior," *Manuscript*, December 2002.

Hurd, Michael and James P Smith, "Expected Bequests and Their Distribution," Working Paper 9142, National Bureau of Economic Research, September 2002.

İmrohoroglu, Ayşe, Selahattin İmrohoroglu, and Douglas H. Joines, "A Life Cycle Analysis of Social Security," *Economic Theory*, February 1995, *6* (1), 83–114.

Jakobsen, Katrine, Kristian Jakobsen, Henrik Kleven, and Gabriel Zucman, "Wealth Taxation and Wealth Accumulation: Theory and Evidence From Denmark," *Quarterly Journal of Economics*, February 2020, *135* (1), 329–388.

Jones, Callum, "Aging, Secular Stagnation and the Business Cycle," *Manuscript*, February 2018.

Jordà, Òscar, Katharina Knoll, Dmitry Kuvshinov, Moritz Schularick, and Alan M. Taylor, "The Rate of Return on Everything, 1870–2015," *Quarterly Journal of Economics*, August 2019, *134* (3), 1225–1298.

Karabarbounis, Loukas and Brent Neiman, "The Global Decline of the Labor Share," *Quarterly Journal of Economics*, February 2014, *129* (1), 61–103.

Kitao, Sagiri, "Sustainable Social Security: Four Options," *Review of Economic Dynamics*, October 2014, *17* (4), 756–779.

Kleven, Henrik Jacobsen and Esben Anton Schultz, "Estimating Taxable Income Responses Using Danish Tax Reforms," *American Economic Journal: Economic Policy*, November 2014, *6* (4), 271–301.

Kopecky, Joseph and Alan M Taylor, "The Murder-Suicide of the Rentier: Population Aging and the Risk Premium," Working Paper 26943, National Bureau of Economic Research, April 2020.

Krueger, Dirk and Alexander Ludwig, "On the Consequences of Demographic Change for Rates of Returns to Capital, and the Distribution of Wealth and Welfare," *Journal of Monetary Economics*, January 2007, *54* (1), 49–87.

Krusell, Per and Anthony A. Smith, "Is Piketty's "Second Law of Capitalism" Fundamental?," *Journal of Political Economy*, August 2015, *123* (4), 725–748.

Lagakos, David, Benjamin Moll, Tommaso Porzio, Nancy Qian, and Todd Schoellman, "Life Cycle Wage Growth across Countries," *Journal of Political Economy*, December 2017, *126* (2), 797–849.

Lane, Phillip, "International Inflation Co-Movements," *Speech at the Inflation: Drivers and Dynamics 2020 Online Conference*, May 2020, *Federal Reserve Bank of Cleveland/European Central Bank.*

Lisack, Noëmie, Rana Sajedi, and Gregory Thwaites, "Demographic Trends and the Real Interest Rate," *Manuscript*, December 2017.

Mankiw, N. Gregory and David N. Weil, "The Baby Boom, the Baby Bust, and the Housing Market," *Regional Science and Urban Economics*, May 1989, *19* (2), 235–258.

Modigliani, Franco, "Life Cycle, Individual Thrift, and the Wealth of Nations," *American Economic Review*, June 1986, *76* (3), 297–313.

Moll, Benjamin, Lukasz Rachel, and Pascual Restrepo, "Uneven Growth: Automation's Impact on Income and Wealth Inequality," Working Paper 28440, National Bureau of Economic Research, February 2021.

Nardi, M. De, S. Imrohoroglu, and T. J. Sargent, "Saving and Pension Reform in General Equilibrium Models," *Oxford Review of Economic Policy*, March 2001, *17* (1), 20–39.

Nardi, Mariacristina De, "Wealth Inequality and Intergenerational Links," *Review of Economic Studies*, 2004, *71* (3), 743–768.

_ , Eric French, and John B. Jones, "Why Do the Elderly Save? The Role of Medical Expenses," *Journal of Political Economy*, February 2010, *118* (1), 39–75.

_ , _ , John Bailey Jones, and Rory McGee, "Why Do Couples and Singles Save During Retirement?," Technical Report w28828, National Bureau of Economic Research May 2021.

Oberfield, Ezra and Devesh Raval, "Micro Data and Macro Technology," *Econometrica*, 2021, *89* (2), 703–732.

OECD, "The OECD SOCX Manual – 2019 Edition. A guide to the OECD Social Expenditure Database," 2019.

_ , *Pensions at a Glance 2019: OECD and G20 Indicators*, Paris: OECD Publishing, 2019.

Papetti, Andrea, "Demographics and the Natural Real Interest Rate: Historical and Projected Paths for the Euro Area," SSRN Working Paper No 3363512, March 2019.

Penn Wharton Budget Model, "Penn Wharton Budget Model: Dynamic OLG Model," *Technical Report*, May 2019.

Piketty, Thomas, *Capital in the Twenty-First Century*, 1st ed., Cambridge, MA: Belknap Press, March 2014.

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

Poterba, James M., "Demographic Structure and Asset Returns," *Review of Economics and Statistics*, November 2001, *83* (4), 565–584.

Rachel, Lukasz and Lawrence H. Summers, "On Secular Stagnation in the Industrialized World," *Brookings Papers on Economic Activity*, 2019, pp. 1–54.

Sabelhaus, John and Alice Henriques Volz, "Are Disappearing Employer Pensions Contributing to Rising Wealth Inequality?," *FEDS Notes No 2019-02-01*, February 2019.

_ and _ , "Social Security Wealth, Inequality, and Lifecycle Saving," Working Paper 27110, National Bureau of Economic Research, May 2020.

Saez, Emmanuel and Stefanie Stantcheva, "A Simpler Theory of Optimal Capital Taxation," *Journal of Public Economics*, June 2018, *162*, 120–142.

Summers, Lawrence and Chris Carroll, "Why is U.S. National Saving so Low?," *Brookings Papers on Economic Activity*, 1987, *1987* (2), 607–642.

Young, Eric R., "Solving the Incomplete Markets Model with Aggregate Uncertainty Using the Krusell–Smith Algorithm and Non-Stochastic Simulations," *Journal of Economic Dynamics and Control*, January 2010, *34* (1), 36–41.

Zoutman, Floris T., "The Elasticity of Taxable Wealth: Evidence from the Netherlands," *Manuscript*, 2018.

Appendix to "Demographics, Wealth and Global Imbalances in the Twenty-First Century"

A Appendix to Section 1

The total return on wealth r_t for the US from 1950–2016 in panel C of figure 1 is constructed as follows. We take:

- Capital K_t as total private fixed assets at current cost from line 1 of Table 2.1 in the BEA's Fixed Assets Accounts (FA).
- Output Y_t as gross domestic product from line 1 of Table 1.1.5 in the BEA's National Income and Product Accounts (NIPA).
- Wealth W_t as "net private wealth" from the World Inequality Database (WID).
- Net foreign assets NFA_t as the net worth of the "rest of the world" sector from line 147 of Table S.9.a in the Integrated Macroeconomic Accounts (IMA).[40]
- Government bonds B_t as gross federal debt held by the public, from the Economic Report of the President (accessed via fred at FYGFDPUB).
- The safe real interest rate r_t^{safe} as the 10-year constant maturity interest rate—from Federal Reserve release H.15 (accessed via fred at GS10), extended backward from 1953 to 1950 by splicing with the NBER macrohistory database's yield on long-term US bonds (accessed via fred at M1333BUSM156NNBR)—minus a slow-moving inflation trend, calculated as the trend component of annual HP-filtered inflation in the PCE deflator, with smoothing parameter $\lambda = 100$.
- Net capital income $(s_K Y - \delta K)_t$ as corporate profits plus net interest and miscellaneous payments of the corporate sector (sum of lines 7 and 8 in NIPA Table 1.13), plus imputed net capital income from the noncorporate business sector, under the assumption that the ratio of net capital income to net factor income (line 11 minus line 17) in the noncorporate business sector is the same as the ratio of net capital income to net factor income (line 3 minus line 9) in the corporate sector.[41]

We then calculate our baseline total return on wealth series as

$$r_t \equiv \frac{(s_K Y - \delta K)_t + r_t^{safe} B_t}{W_t - NFA_t} \tag{A.1}$$

i.e. as the ratio of net capital income plus real interest income on government debt to domestic assets. This calculation gives the total return on private wealth, excluding changes in asset valua-

[40] This is very similar to the standard net international investment position computed by the BEA, but is chosen because it offers a longer time series.

[41] This imputation is a common way of splitting mixed income within the noncorporate sector between labor and capital, used e.g. by Piketty (2014).

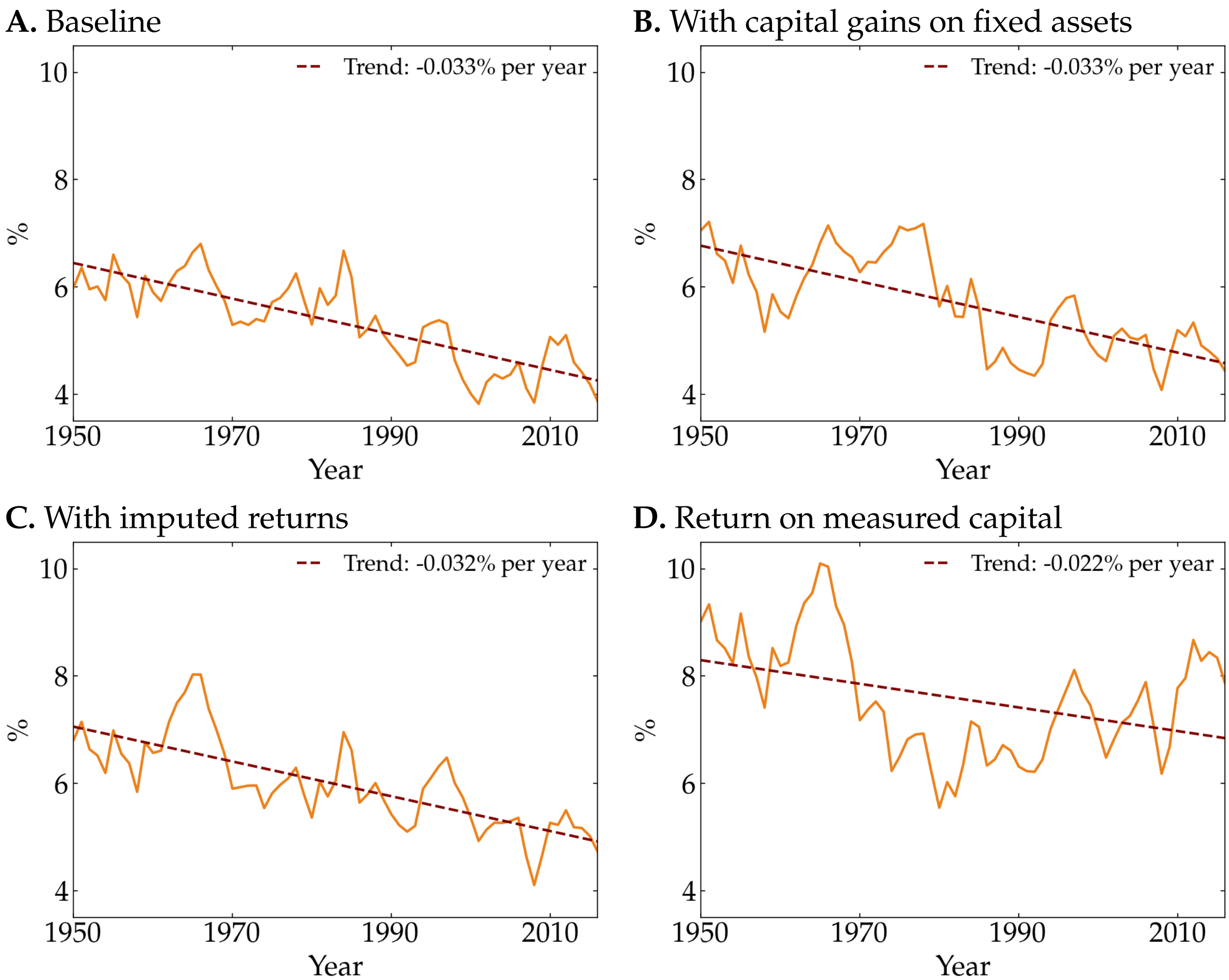


Figure A.1: Alternative ways of constructing total return on wealth in US

Notes: Panel A gives our baseline series for the total return on wealth in the US, as described in the text. Panel B adds capital gains on fixed assets, as measured in the fixed assets accounts. Panel C imputes an additional return on unmeasured wealth $W_t - K_t - B_t - NFA_t$ equal to trend growth. Panel D takes our baseline capital income series and divides it by capital measured in the fixed assets accounts.

tion, under the assumption that the average return on net foreign assets is the same as the average return on private wealth.[42]

This baseline r_t and its trend are displayed in panel A of figure A.1. The other three panels provide alternative ways to calculate r_t.

Panel B adds a slow-moving trend of capital good inflation minus PCE inflation, which we denote by π_{Kt}:

$$r_t \equiv \frac{(s_K Y - \delta K)_t + r_t^{safe} B_t + \pi_{Kt} K_t}{W_t - NFA_t}$$

Average inflation of goods in the capital stock is inferred by taking the ratio of changes in the nominal stock (FA Table 2.1, line 1) and changes in the quantity index (FA Table 2.2, line 1), and as with PCE inflation above, we take the slow-moving trend component using the HP filter with $\lambda = 100$. This accounts for expected capital gains on fixed capital (assuming that the expectation follows the trend).

Panel C assumes that there is some unmeasured return on the portion of wealth $W_t - K_t - B_t - NFA_t$ that cannot be accounted for by capital, bonds, or net foreign assets, which it sets equal to the trend real GDP growth rate g_t:

$$r_t \equiv \frac{(s_K Y - \delta K)_t + r_t^{safe} B_t + g_t (W_t - K_t - B_t - NFA_t)}{W_t - NFA_t}$$

where g_t is again calculated using the HP filter with $\lambda = 100$. If $W_t - K_t - B_t - NFA_t$ is the capitalized value of pure rents in the economy, for instance, its value might be expected to grow in line with output.

Finally, panel D simply divides net capital income by the measured capital stock:

$$r_t \equiv \frac{(s_K Y - \delta K)_t}{K_t}$$

Note that despite these alternative constructions, the 1950–2016 trends in panels A, B, and C of figure A.1 are almost identical: -.033, -.033, and -.032 percentage points, respectively. All show a steady decline.

The return on capital in panel D, on the other hand, is quite different: it has a smaller long-term trend decline, of -.022 percentage points per year, and since roughly 1980 it actually displays a mild increase. This post-1980 pattern of a constant or increasing return on capital has been widely remarked upon in the literature—for instance, Gomme, Ravikumar and Rupert (2011), Farhi and Gourio (2018), Eggertsson, Robbins and Wold (2018). The main source of the disparity between panels A–C and panel D is that the former divide by wealth, while the latter divides only by measured capital. Since our primary object of interest is wealth, we prefer the former convention. Another advantage of using wealth in the denominator is that capital may be imperfectly measured in the fixed assets accounts.

[42] This can be seen by rearranging (A.1) as $r_t = \frac{s_K Y - \delta K + r^{safe} B + rNFA}{W}$, which gives the total return r_t on private wealth if r_t equals the return on NFA_t. We take this route because data on capital income from foreign assets is not comparable to domestic data; for instance, the national accounts only measure dividend payments, not the total net capital income, on foreign equities (other than FDI) held in the US, and also only measure nominal rather than real interest payments on bonds. The trend in r_t, however, is not very sensitive to alternative assumptions on the average rate for NFA_t.

B Appendix to Section 2

B.1 Contribution of changing fertility to aging, 1950-2100

Figure A.2 uses our model of the age distribution of the population in each country to decompose population aging into contributions from fertility, mortality, migration and the so called "momentum" effect. Our measure of population aging is the changes in the share of the population aged 50 or above. Denote by $\Delta\pi$ the change in this share between two periods t_0 and t_1. To isolate the role of primitive forces for $\Delta\pi$, we start with an initial age distribution in year t_0. We obtain the contribution of fertility plus "momentum" by simulating the population distribution holding mortality and migration constant until t_1, and then computing the counterfactual change $\Delta^f\pi$ in the share of the 50+ year old in this scenario. The ratio $\Delta^f\pi/\Delta\pi$ gives us the contribution of fertility and momentum to population aging, which our baseline model of section 2 includes, with the remainder accounted for by mortality and migration, which the baseline model abstracts from. We conduct this exercise over two separate time periods t_0-t_1: 1950-2016 and 2016-2100.

Figure A.2 presents the results, showing $\Delta^f\pi/\Delta\pi$ over these two time periods for the 25 countries in our sample. The top panel shows that, between 1950-2016, fertility and momentum contributed an average of 63.5% of population aging. The bottom panel shows that, between 2016 and 2100, their contributions are projected to shrink a little to an average of 55.9%, but still constitute the majority of the contribution. Hence, our baseline assumption of fixed mortality and migration is a useful first pass at the data, although improving mortality becomes increasingly important to population aging as we look towards the 21st century. Our model of section 4 allows for time variation in mortality and models the savings response to it.

B.2 Proofs of lemma 1 and proposition 1

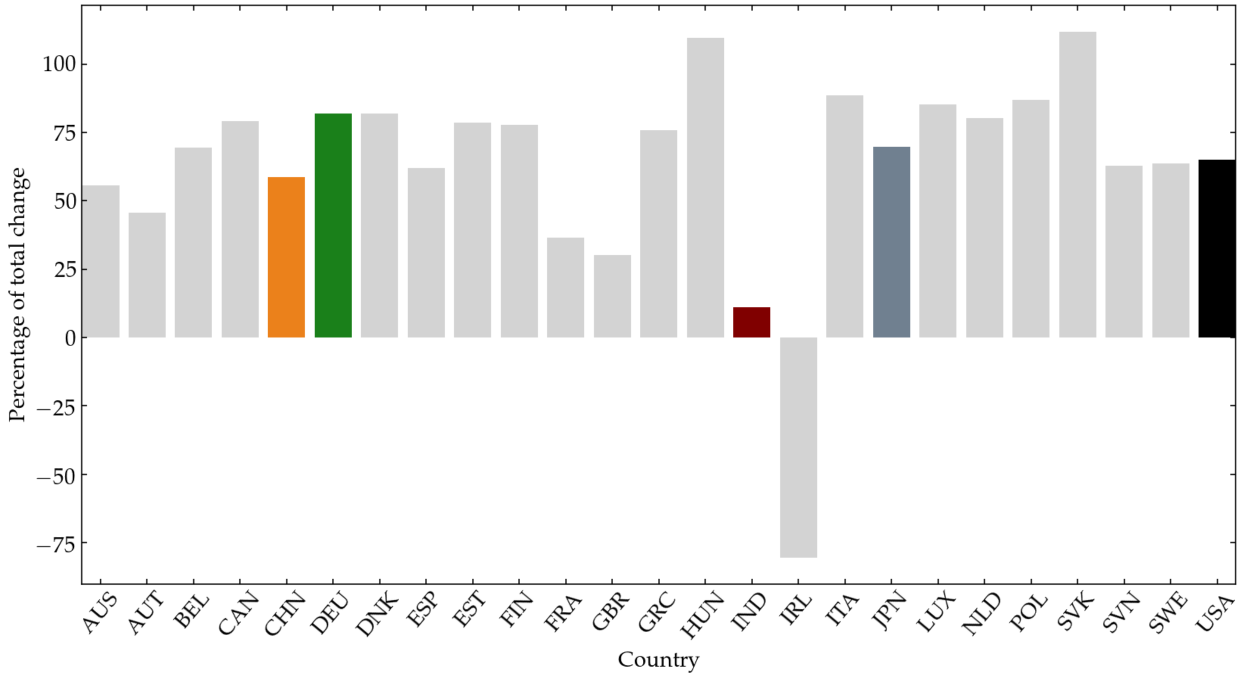
The ratio K_t/Z_tL_t of capital to effective labor is constant over time, pinned down by constant r and the condition $r_t + \delta = F_K(K_t/(Z_tL_t), 1)$. From the condition $w_t = Z_tF_L(K_t/(Z_tL_t), 1)$, w_t is then proportional to Z_t and grows at the constant rate γ. It follows immediately that average pre-tax labor income $h_{jt} \equiv \mathbb{E}w_t\ell_j = (1+\gamma)^t w_0 \mathbb{E}\ell_j$ grows at the constant rate γ.

Letting hats denote normalization of time-subscripted variables by $(1+\gamma)^t$, and defining $\hat{\beta}_j \equiv (1+\gamma)^{j(1-\frac{1}{\sigma})}\beta_j$, the household utility maximization problem (1) becomes

$$\begin{aligned} &\max_{\hat{c}_{jt}, \hat{\mathfrak{a}}_{j+1,t+1}} \mathbb{E}_k \left[\sum_{j=0}^{T} \hat{\beta}_j \Phi_j \frac{\hat{c}_{jt}^{1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}} \right] \\ \text{s.t.} \quad &\hat{c}_{jt} + (1+\gamma)\phi_j \hat{\mathfrak{a}}_{j+1,t+1} \leq w_0 \left((1-\tau)\ell(z_j) + tr(z^j) \right) + (1+r)\hat{\mathfrak{a}}_{jt} \\ &\hat{\mathfrak{a}}_{j+1,t+1} \geq 0 \end{aligned} \tag{A.2}$$

This problem is no longer time-dependent: given the same asset holdings $\hat{a}_j$, state z^j and age j, households optimally choose the same $(\hat{c}_j, \hat{\mathfrak{a}}_{j+1})$ regardless of t. Regardless of their date of birth, every cohort born in this environment will have the same distribution of normalized assets $\hat{a}_j$ at each age j. Hence, once t is high enough that all living agents were born in this environment, there exists a balanced-growth distribution of assets at each age that grows at rate γ. Average assets normalized by productivity satisfy $a_{jt}/Z_t = (\mathbb{E}\mathfrak{a}_{jt})/Z_t = (\mathbb{E}\hat{\mathfrak{a}}_j)/Z_0 \equiv a_j(r)$ for some function

A. 1952-2016 change in the share of 50+ : percentage due to fertility and momentum



B. 2016-2100 change in the share of 50+ : percentage due to fertility and momentum

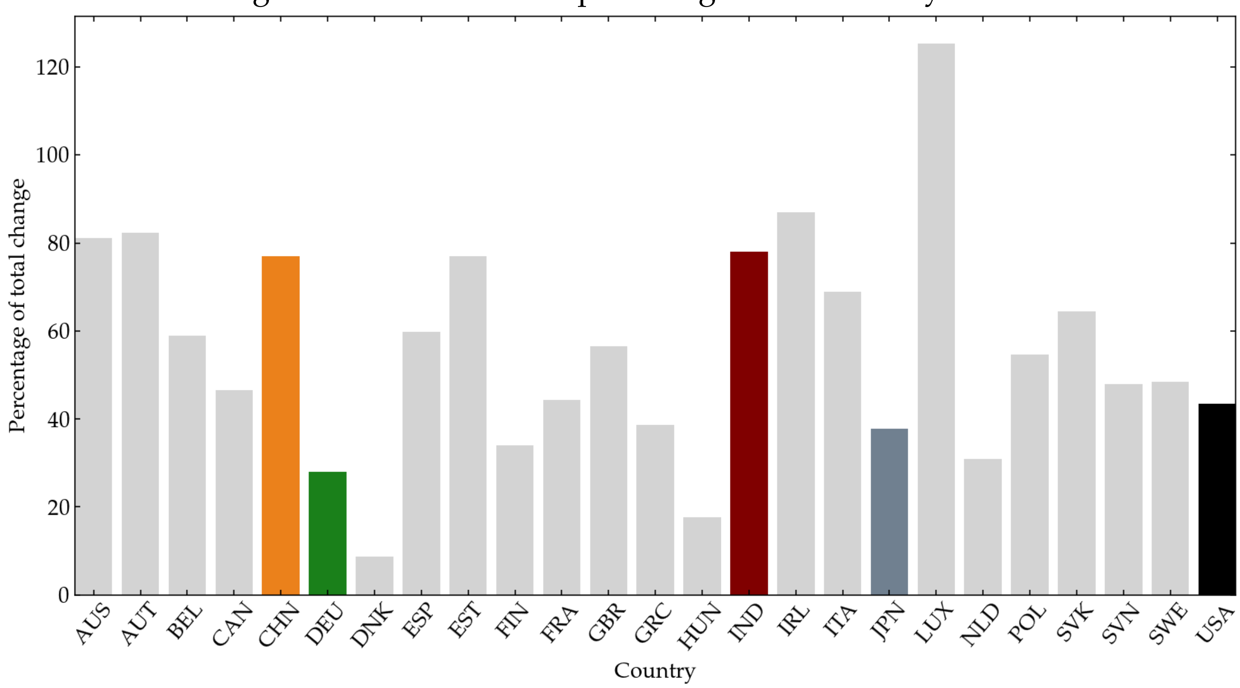


Figure A.2: Contribution of fertility and momentum to population aging

Notes: This figure presents the percentage of the change in the share of 50+ that is due to fertility changes and momentum. It is computed as the ratio between the change in this share under the assumptions of constant mortality rates and migration flows, and under the baseline assumptions for 1952-2016 (panel A) and 2016-2100 (panel B).

$a_j(r)$. If, at date 0, already-living agents start with the joint balanced-growth distribution of assets and states, then this holds immediately.

The ratio of aggregate wealth to aggregate labor at time t is

$$\frac{W_t}{L_t} = \frac{\sum_j N_{jt} a_{jt}}{\sum_j N_{jt} \mathbb{E}\ell_j} = \frac{\sum_j N_{jt}(1+\gamma)^t a_{j0}}{\sum_j N_{jt} h_{j0}/w_0} = (1+\gamma)^t w_0 \frac{\sum_j \pi_{jt} a_{j0}}{\sum_j \pi_{jt} h_{j0}} \tag{A.3}$$

The ratio of output to aggregate labor is

$$\frac{Y_t}{L_t} = \frac{F(K_t, Z_t L_t)}{L_t} = Z_t F\left(\frac{K_t}{Z_t L_t}, 1\right) = Z_t F\left(\frac{K_0}{Z_0 L_0}, 1\right) \tag{A.4}$$

where we use the fact that the capital-to-effective-labor ratio is constant. Dividing (A.3) and (A.4), the wealth-to-output ratio is

$$\frac{W_t}{Y_t} = \frac{w_0}{Z_0 F(K_0/Z_0 L_0, 1)} \frac{\sum_j \pi_{jt} a_{j0}}{\sum_j \pi_{jt} h_{j0}} \tag{A.5}$$

where the first factor is constant with time. We conclude that $\frac{W_t}{Y_t}$ grows in proportion to $\frac{\sum_j \pi_{jt} a_{j0}}{\sum_j \pi_{jt} h_{j0}}$.

B.3 Proofs of propositions 2 and 3

Proof of proposition 2 Within each country c, for a constant rate of return r, lemma 1 shows that there exists a balanced-growth distribution of assets normalized by productivity. Assuming we start with this balanced-growth distribution, then at each t, (A.5) implies

$$\begin{aligned} \frac{W_t^c}{Y_t^c} &= \frac{w_0^c}{Z_0^c F^c(K_0^c/Z_0^c L_0^c, 1)} \frac{\sum_j \pi_{jt}^c a_{j0}^c}{\sum_j \pi_{jt}^c h_{j0}^c} \\ &= \frac{F_L^c(K_0^c/Z_0^c L_0^c, 1)}{F(K_0^c/Z_0^c L_0^c, 1)} \frac{\sum_j \pi_{jt}^c a_{j0}^c}{\sum_j \pi_{jt}^c h_{j0}^c} \\ &= \frac{F_L^c(k^c(r), 1)}{F^c(k^c(r), 1)} \frac{\sum_j \pi_{jt}^c a_{j0}^c}{\sum_j \pi_{jt}^c h_{j0}^c} \equiv \frac{W^c}{Y^c}(r, \pi_t^c) \end{aligned}$$

where $\pi_t^c \equiv \{\pi_{jt}^c\}_j$, and $k(r)$ is the capital-to-effective-labor ratio associated with r, defined implicitly by $F_K^c(k(r), 1) = r + \delta$.

Each country's share of world GDP is then given by

$$\frac{Y_t^c}{Y_t} = \frac{Z_t^c L_t^c y^c(r)}{\sum Z_t^c L_t^c y^c(r)} = \frac{Z_0^c \nu_t^c y^c(r) \sum \pi_{jt}^c \ell_j^c}{\sum Z_0^c \nu_t^c y^c(r) \sum \pi_{jt}^c \ell_j^c} \equiv \frac{Y^c}{Y}(r, \pi_t, \nu_t),$$

where $\nu_t^c \equiv N_t^c / N_t$ and π_t and ν_t denote vectors across all countries, and $y^c(r) \equiv F^c(k^c(r), 1)$.

The capital-to-output ratio in a country can also be written as a function of r, $\frac{K^c}{Y^c}(r) \equiv k^c(r)/F^c(k^c(r), 1)$, and we assume that government policy maintains a constant $\frac{B^c}{Y^c}$ in each country.

We assume that the economy is in balanced growth corresponding to long-run r_0 at date 0, which means that the initial wealth-to-output ratio is $\frac{W^c}{Y^c}(r_0, \pi_0^c)$ and that the initial capital-output ratio is $\frac{K^c}{Y^c}(r_0)$. We also assume that net foreign asset positions in each country are 0 at time 0, i.e.

that

$$\frac{W^c}{Y^c}(r_0, \pi_0^c) - \frac{K^c}{Y^c}(r_0) - \frac{B^c}{Y^c} = 0.$$

In the long run, π_t^c and ν_t^c converge to constants π_{LR}^c and ν_{LR}^c in each country. Suppose that the real return r_t converges to a long-run value r_{LR}. Then the world asset market clearing condition is

$$0 = \sum_c \frac{Y^c}{Y}(r, \pi, \nu) \left[\frac{W^c}{Y^c}(r, \pi^c) - \frac{K^c}{Y^c}(r) - \frac{B^c}{Y^c} \right] \tag{A.6}$$

which holds for both $(r, \pi, \nu) = (r_0, \pi_0, \nu_0)$ and $(r, \pi, \nu) = (r_{LR}, \pi_{LR}, \nu_{LR})$. Subtracting the former from the latter, we have

$$\begin{aligned} 0 &= \sum_c \frac{Y^c}{Y}(r_{LR}, \pi_{LR}, \nu_{LR}) \left[\frac{W^c}{Y^c}(r_{LR}, \pi_{LR}^c) \right. \\ &\quad \left. - \frac{K^c}{Y^c}(r_{LR}) - \frac{B^c}{Y^c} \right] - \sum_c \frac{Y^c}{Y}(r_0, \pi_0, \nu_0) \left[\frac{W^c}{Y^c}(r_0, \pi_0^c) - \frac{K^c}{Y^c}(r_0) - \frac{B^c}{Y^c} \right] \\ &= \sum_c \left[\frac{Y^c}{Y}(r_{LR}, \pi_{LR}, \nu_{LR}) - \frac{Y^c}{Y}(r_0, \pi_0, \nu_0) \right] \left[\frac{W^c}{Y^c}(r_{LR}, \pi_{LR}^c) - \frac{K^c}{Y^c}(r_{LR}) - \frac{B^c}{Y^c} \right] \\ &\quad + \sum_c \frac{Y^c}{Y}(r_0, \pi_0, \nu_0) \left[\frac{W^c}{Y^c}(r_{LR}, \pi_{LR}^c) - \frac{K^c}{Y^c}(r_{LR}) - \frac{B^c}{Y^c} - \left(\frac{W^c}{Y^c}(r_0, \pi_0^c) - \frac{K^c}{Y^c}(r_0) - \frac{B^c}{Y^c} \right) \right] \end{aligned}$$

Note that $\frac{W^c}{Y^c}(r_0, \pi_0^c) - \frac{K^c}{Y^c}(r_0) - \frac{B^c}{Y^c}$ is 0 by the assumption of zero initial NFA. To first order, therefore, the product of $\left[\frac{Y^c}{Y}(r_{LR}, \pi_{LR}, \nu_{LR}) - \frac{Y^c}{Y}(r_0, \pi_0, \nu_0) \right]$ and $\left[\frac{W^c}{Y^c}(r_{LR}, \pi_{LR}^c) - \frac{K^c}{Y^c}(r_{LR}) - \frac{B^c}{Y^c} \right]$ is zero as well. To first order, the above then simplifies to the equivalent

$$\begin{aligned} 0 &= \sum_c \frac{Y_0^c}{Y_0} \left[\frac{W^c}{Y^c}(r_{LR}, \pi_{LR}^c) - \frac{K^c}{Y^c}(r_{LR}) - \frac{B^c}{Y^c} - \left(\frac{W^c}{Y^c}(r_0, \pi_0^c) - \frac{K^c}{Y^c}(r_0) - \frac{B^c}{Y^c} \right) \right] &\text{(A.7)} \\ &= \sum_c \frac{Y_0^c}{Y_0} \left[\frac{W^c}{Y^c}(r_{LR}, \pi_{LR}^c) - \frac{W^c}{Y^c}(r_0, \pi_{LR}^c) + \frac{W^c}{Y^c}(r_0, \pi_{LR}^c) - \frac{W^c}{Y^c}(r_0, \pi_0^c) - \left(\frac{K^c}{Y^c}(r_{LR}) - \frac{K^c}{Y^c}(r_0) \right) \right] \\ &\simeq \sum_c \frac{Y_0^c}{Y_0} \left[\frac{\partial \frac{W^c}{Y^c}(r_0, \pi_{LR}^c)}{\partial r}(r_{LR} - r_0) + \frac{W^c}{Y^c}(r_0, \pi_0^c) \left(\exp(\Delta_{LR}^{comp,c}) - 1 \right) - \frac{\partial \frac{K^c}{Y^c}(r_0)}{\partial r}(r_{LR} - r_0) \right] \\ &\simeq \sum_c \frac{W_0^c}{Y_0} \left[\frac{\partial \log \frac{W^c}{Y^c}(r_0, \pi_{LR}^c)}{\partial r}(r_{LR} - r_0) + \Delta_{LR}^{comp,c} - \frac{1}{\frac{W^c}{Y^c}(r_0, \pi_0)} \frac{\partial \frac{K^c}{Y^c}(r_0)}{\partial r}(r_{LR} - r_0) \right], &\text{(A.8)} \end{aligned}$$

where we write $\frac{Y_0^c}{Y_0}$ and $\frac{W_0^c}{Y_0}$ to denote $\frac{Y^c}{Y}(r_0, \pi_0, \nu_0)$ and $\frac{W^c}{Y}(r_0, \pi_0, \nu_0)$.

Let us also define

$$\begin{aligned} \epsilon^{d,c} &\equiv \frac{\partial \log \frac{W^c}{Y^c}(r_0, \pi_{LR}^c)}{\partial r} \\ \epsilon^{s,c} &\equiv -\frac{1}{\frac{W^c}{Y^c}(r_0, \pi_0)} \frac{\partial \frac{K^c}{Y^c}(r_0)}{\partial r} \\ \omega^c &\equiv \frac{W^c}{W}(r_0, \pi_0, \nu_0) \end{aligned}$$

and divide both sides of (A.8) by $\frac{W}{Y}(r_0, \pi_0, \nu_0)$ to obtain the first-order result

$$\begin{aligned} 0 &\simeq \sum_c \omega^c \left[\Delta_{LR}^{comp,c} + (\epsilon^{d,c} + \epsilon^{s,c})(r_{LR} - r_0)\right] \\ &= \bar{\Delta}_{LR}^{comp} + (\bar{\epsilon}^d + \bar{\epsilon}^s)(r_{LR} - r_0) \end{aligned} \tag{A.9}$$

where we let bars denote averages across countries with initial wealth weights ω^c. The equations (13) and (14) are rearrangements of (A.9).

Now, the change in W^c/Y^c in each country can be written to first order as

$$\Delta_{LR} \log\left(\frac{W^c}{Y^c}\right) = \Delta_{LR}^{comp,c} + \epsilon^{d,c}(r_{LR} - r_0)$$

Summing up both sides with weights ω^c, this becomes

$$\overline{\Delta_{LR} \log\left(\frac{W^c}{Y^c}\right)} = \bar{\Delta}_{LR}^{comp} + \bar{\epsilon}^d(r_{LR} - r_0)$$

and using (A.9) to substitute out for $r_{LR} - r_0$, we obtain (15),

$$\overline{\Delta_{LR} \log\left(\frac{W^c}{Y^c}\right)} = \frac{\bar{\epsilon}^s}{\bar{\epsilon}^s + \bar{\epsilon}^d}\bar{\Delta}_{LR}^{comp} \tag{A.10}$$

Proof of proposition 3 The change in $NFA^c/Y^c = W^c/Y^c - K^c/Y^c - B^c/Y^c$ is given by

$$\begin{aligned} \Delta_{LR}\frac{NFA^c}{Y^c} &= \frac{W_0^c}{Y_0^c}\left[\exp\left(\Delta_{LR}^{comp,c} + (\epsilon^{d,c} + \epsilon^{s,c})(r_{LR} - r_0)\right) - 1\right] \\ &= \frac{W_0^c}{Y_0^c}\left[\exp\left(\Delta_{LR}^{comp,c} + (\epsilon^{d,c} + \epsilon^{s,c})\frac{\bar{\Delta}_{LR}^{comp}}{\bar{\epsilon}^d + \bar{\epsilon}^s}\right) - 1\right] \\ &= \frac{W_0^c}{Y_0^c}\left[\exp\left(\Delta_{LR}^{comp,c} - \bar{\Delta}_{LR}^{comp} + (\epsilon^{d,c} + \epsilon^{s,c} - (\bar{\epsilon}^d + \bar{\epsilon}^s))\frac{\bar{\Delta}_{LR}^{comp}}{\bar{\epsilon}^d + \bar{\epsilon}^s}\right) - 1\right] \\ &= \frac{W_0^c}{Y_0^c}\left[\exp\left(\Delta_{LR}^{comp,c} - \bar{\Delta}_{LR}^{comp} + (\epsilon^{d,c} + \epsilon^{s,c} - (\bar{\epsilon}^d + \bar{\epsilon}^s))(r_{LR} - r_0)\right) - 1\right] \end{aligned}$$

Rearranged, this gives the desired result, which is

$$\log\left(1 + \left(\Delta_{LR}\frac{NFA^c}{Y^c}\right) \Big/ \frac{W_0^c}{Y_0^c}\right) = \Delta_{LR}^{comp,c} - \bar{\Delta}_{LR}^{comp} + (\epsilon^{d,c} + \epsilon^{s,c} - (\bar{\epsilon}^d + \bar{\epsilon}^s))(r_{LR} - r_0)$$

B.4 Relaxing assumptions in propositions 2 and 3

In the more general case, we allow initial NFA's to be non-zero and debt-to-output ratios to vary over time. Below, we show how the formulas are modified in this case, and some discussions of how particular sequences of debt-to-output ratios can mitigate the general equilibrium effects of wages.

Allowing for nonzero initial NFAs. The second, more involved, adjustment is that with nonzero initial NFAs, there is a compositional effect of aging on net asset demand insofar the change in relative GDP across countries is correlated with initial NFAs.

If NFA_0^c is not zero in every country c, we would retain an additional term in (A.7), equal to first order to

$$\sum_c \left[\frac{Y^c}{Y}(r_{LR}, \pi_{LR}, \nu_{LR}) - \frac{Y_0^c}{Y_0}\right] \frac{NFA_0^c}{Y_0^c} = \sum_c \frac{Y_0^c}{Y_0} \Delta_{LR} \log \frac{Y^c}{Y}(r_{LR}, \pi_{LR}, \nu_{LR}) \frac{NFA_0^c}{Y_0^c}$$

When we divide by $\frac{W_0}{Y_0}$ as in our derivation of (A.9), this becomes

$$\sum_c \omega^c \frac{NFA_0^c}{W_0^c} \Delta_{LR} \log \frac{Y^c}{Y}(r_{LR}, \pi_{LR}, \nu_{LR}) \tag{A.11}$$

which will show up as an additional term in (A.9). Since the wealth-weighted average of $\frac{NFA_0^c}{W_0^c}$ is zero by global market clearing, this can be written as a wealth-weighted covariance

$$\text{Cov}_{\omega^c}\left(\frac{NFA_0^c}{W_0^c}, \Delta_{LR} \log \frac{Y^c}{Y}(r_{LR}, \pi_{LR}, \nu_{LR})\right) \tag{A.12}$$

If we define

$$\Delta_L^{demog} R \frac{Y^c}{Y} \equiv \frac{\partial(\log \frac{Y^c}{Y})}{\partial \pi} \Delta_{LR} \pi + \frac{\partial(\log \frac{Y^c}{Y})}{\partial \nu} \Delta_{LR} \nu$$

to be the change in GDP shares caused by demographic change alone, holding r constant, and

$$\bar{\epsilon}^{weight} \equiv \text{Cov}_{\omega^c}\left(\frac{NFA_0^c}{W_0^c}, \frac{\partial(\log \frac{Y^c}{Y})}{\partial r}\right) \tag{A.13}$$

then the modified (A.9) becomes

$$\bar{\Delta}_{LR}^{comp} + \text{Cov}_{\omega^c}\left(\frac{NFA_0^c}{W_0^c}, \Delta_{LR}^{demog} \frac{Y^c}{Y}\right) + (\bar{\epsilon}^d + \bar{\epsilon}^s + \bar{\epsilon}^{weight})(r_{LR} - r_0) = 0 \tag{A.14}$$

and we can solve to obtain

$$r_{LR} - r_0 = \frac{\bar{\Delta}_{LR}^{comp} + \text{Cov}_{\omega^c}\left(\frac{NFA_0^c}{W_0^c}, \Delta_{LR}^{demog} \frac{Y^c}{Y}\right)}{\bar{\epsilon}^d + \bar{\epsilon}^s + \bar{\epsilon}^{weight}}$$

Note that the two departures from our previous result, the covariance in (A.14) and the covariance in the definition (A.13) of $\bar{\epsilon}^{weight}$, both involve wealth-weighted covariances between initial net foreign asset positions as shares of wealth, $\frac{NFA_0^c}{W_0^c}$, and some change in each country's GDP weight (either in response to demographics or endogenously in response to r). A priori, there is no particular reason to have a covariance in either direction here, and indeed we have found that these terms seem quite small in practice, to the point where they are best disregarded in our main analysis.

Our previous simplification for the average change in wealth-to-GDP no longer holds, but we can still write

$$\overline{\Delta_{LR} \log \frac{W^c}{Y^c}} \simeq \bar{\Delta}^{comp} + \bar{\epsilon}^d (r_{LR} - r_0).$$

The change in NFA in each country is

$$\Delta \log\left(1+\frac{\Delta_{LR}NFA^c/Y^c}{W_0^c/Y_0^c}\right) = \Delta^{comp,c} + (\epsilon^{d,c}+\epsilon^{s,c})(r_{LR}-r_0)$$

Change in debt-to-output ratios. Suppose that each country operates a fiscal rule that targets and exogenous sequence $\frac{B_t^c}{Y_t^c}$ which converges to some long-run value $\frac{B_{LR}^c}{Y_{LR}^c}$ in every country. The average change in bonds is a shifter of asset supply, and the new version of (13) is

$$\bar{\Delta}_{LR}^{comp} - \overline{\frac{\Delta_{LR}B^c/Y^c}{W_0^c/Y_0^c}} + \bar{\epsilon}^d(r_{LR}-r_0) \simeq -\bar{\epsilon}^s(r_{LR}-r_0), \tag{A.15}$$

where $\overline{\frac{\Delta_{LR}B^c/Y^c}{W_0^c/Y_0^c}} \equiv \sum_c \omega^c \left(\frac{B_{LR}^c}{Y_{LR}^c} - \frac{B_0^c}{Y_0^c}\right)$ is the average log change in debt-to-output ratios.

We can solve (A.15) to obtain $r_{LR}-r_0$, which is simply the original formula with this shifter in supply subtracted from the compositional effect:

$$r_{LR}-r_0 = \frac{\bar{\Delta}_{LR}^{comp} - \overline{\frac{\Delta_{LR}B^c/Y^c}{W_0^c/Y_0^c}}}{\bar{\epsilon}^d+\bar{\epsilon}^s} \tag{A.16}$$

The average change in wealth-to-GDP now becomes

$$\overline{\Delta_{LR}\log\frac{W^c}{Y^c}} \simeq \frac{\bar{\epsilon}^s}{\bar{\epsilon}^d+\bar{\epsilon}^s}\bar{\Delta}_{LR}^{comp} + \frac{\bar{\epsilon}^d}{\bar{\epsilon}^d+\bar{\epsilon}^s}\overline{\frac{\Delta_{LR}B^c/Y^c}{W_0^c/Y_0^c}} \tag{A.17}$$

which adds the direct impact of increasing debt to (15), and the change in NFA in each country is

$$\log\left(1+\frac{\Delta_{LR}NFA_{LR}^c}{W_0^c/Y_0^c}\right) \simeq \left(\Delta_{LR}^{comp,c} - \frac{\Delta_{LR}B^c/Y^c}{W_0^c/Y_0^c}\right) - \left(\overline{\Delta}_{LR}^{comp} - \overline{\frac{\Delta_{LR}B^c/Y^c}{W_0^c/Y_0^c}}\right) + \left(\epsilon^{c,d}+\epsilon^{c,s} - \left(\bar{\epsilon}^d+\bar{\epsilon}^s\right)\right)(r_{LR}-r_0) \tag{A.18}$$

which now subtracts the change in asset supply from bonds in each country from the compositional effect on asset demand, but is otherwise the same formula as (16).

Neutralizing debt-to-output policy The equations (A.16) and (A.18) show that effects of demographics on interest rates and NFAs can be neutralized if governments conduct a debt policy that absorbs the shift in aggregate asset demand. More precisely, if all governments expand debt in line with their compositional effect

$$\Delta_{LR}^{comp,c} = \frac{\Delta_{LR}B^c/Y^c}{W_0^c/Y_0^c}$$

we obtain $r_{LR}-r_0 \simeq 0$ and $\log\left(1+\frac{\Delta_{LR}NFA_{LR}^c}{W_0^c}\right) \simeq 0$ for every country c. Intuitively, if governments in every country expand debt to perfectly meet the new demand for assets, there is no change in net asset demand, so interest rates stay constant and NFAs do not change. In this case, the change in wealth equals the compositional effect in every country, since there is no general equilibrium feedback reducing the impact of increased asset demand on wealth.

An alternative specification is if each government increases the *level* of its debt-to-output ratio in line with the *average* compositional effect, so that for all c

$$\frac{W}{Y}\bar{\Delta}_{LR}^{comp} = \Delta_{LR}\frac{B^c}{Y^c}$$

In this case, we still have $r_{LR} - r_0 = 0$, but now NFAs change precisely in line the demeaned compositional effect across countries $\Delta^{comp,c} - \bar{\Delta}^{comp,c}$ in line with (A.18).

Strikingly, these findings are also true in the transition, not just in the long run. That is, if the sequence of debt holdings satisfies $\frac{\Delta_t B^c/Y^c}{W_0^c/Y_0^c} = \Delta_t^{comp,c}$ for every t, then interest rates and NFAs are constant over time, and the path of wealth-to-output ratios equals the path of the compositional effect. Moreover, if $\frac{\Delta_t B^c/Y^c}{W_0^c/Y_0^c} = \bar{\Delta}^{comp}$, then the interest rate change is zero at every point in time, and NFAs at every time period for each country is the demeaned compositional effect.

B.5 Proof of proposition 4

TO BE ADDED

C Appendix to Section 3

C.1 Data sources

Demographics. Our population data and projections comes from the 2019 UN World Population Prospects.[43] We gather data between 1950 and 2100 on total number of births, number of births by age-group of the mother, population by 5-year age groups, and mortality rates by 5-year age groups. We interpolate to construct population distributions N_{jt} and mortality rates ϕ_{jt} in every country, every year, and for every age. We compute total population as $N_t = \sum_j N_{jt}$, population distributions as $\pi_{jt} = N_{jt}/N_t$, and population growth rates as $1 + n_t = N_t/N_{t-1}$. Finally, we compute the number of migrants by age M_{jt} as the residual of the population law of motion

$$N_{jt} = (N_{j-1,t-1} + M_{j-1,t-1})\phi_{j-1,t-1}.$$

Age-income profiles We use the LIS to construct the base-year age-income profiles for all the countries we consider. For Australia, the LIS is based on the Survey of Income and Housing (SIH) and the Household Expenditure Survey (HES), for Austria on the Survey on Income and Living Conditions (SILC), for Canada on the Canadian Income Survey (CIS), for China on the Chinese Household Income Survey (CHIP), for Denmark on the Law Model (based on administrative records), for Estonia on the Estonian Social Survey (ESS) and the Survey on Income and Living Conditions (SILC), for Finland on the Income Distribution Survey (IDS) and the Survey on Income and Living Conditions (SILC), for France on the Household Budget Survey (BdF), for Germany on the German Socio-Economic Panel (GSOEP), for Greece on the Survey of Income and Living Conditions (SILC), for Hungary on the Tárki Household Monitor Survey, for India on the India Human Development Survey (IHDS), for Ireland on the Survey on Income and Living Conditions (SILC), for Italy on the Survey of Household Income and Wealth (SHIW), for Japan on the Japan Household Panel Survey (JHPS), for Luxembourg on the Socio-economic Panel "Living in Luxembourg" (PSELL III) and the Survey on Income and Living Conditions (SILC), for Netherlands

[43] https://population.un.org/wpp/

on the Survey on Income and Living Conditions (SILC), for Norway on the Household Income Statistics, for Poland on the Household Budget Survey, for Slovakia on the Survey of Income and Living Conditions (SILC), for Slovenia on the Household Budget Survey (HBS), for Estonia on the Survey on Income and Living Conditions (SILC), for Sweden on the Household Income Survey (HINK/HEK), and for the United Kingdom on the Family Resources Survey (FRS).

Age-wealth profiles. Our wealth data in the United States comes from the 2016 Survey of Consumer Finance. We gather data from other countries as follows. First, we take data from the Luxembourg Wealth Study (LWS)[44] for Australia in 2010, Canada in 2012, Germany in 2012, United Kingdom in 2011, Italy in 2010, and Sweden in 2005. For Australia the LWS is based on the Survey of Income and Housing (SIH) and the Household Expenditure Survey (HES), for Canada on the Survey of Financial Securities (SFS), for Germany on the German Socio-Economic Panel (GSOEP), for Italy on the Survey of Household Income and Wealth (SHIW), for Sweden on the Household Income Survey (HINK/HEK), and for United Kingdom on the Wealth and Assets Survey (WAS). We rescale the survey weights such that they sum up to the correct number of households according to, respectively, the Australian Bureau of Statistics, Statistics Canada, Statistisches Bundesamt, the Office for National Statistics, the Instituto Nazionale di Statistica, and the United Nations Economic Commission for Europe (UNECE). Next, we use the Household Finance and Consumption Survey (HFCS)[45] for Austria in 2010, Belgium in 2010, Estonia in 2014, Spain in 2010, Finland in 2010, France in 2010, Greece in 2010, Hungary in 2014, Ireland in 2014, Luxembourg in 2014, Netherlands in 2010, Poland in 2014, Slovenia in 2014, and Slovakia in 2014. For China, we rely on the 2013 China Household Finance Survey (CHFS).[46] For India, we use the National Sample Survey (NSS).[47] For Japan, we construct a measure of total wealth by age of household head from Table 4 of the 2009 National Survey of Family Income and Expenditure (NFSIE) available on the online portal of Japanese Government Statistics[48]. This table provides average net worth and total number of households by age groups for single person households and households with two or more members, which we aggregate to obtain total household net worth by age group. For Denmark, we use the 2014 table "Assets and liabilities per person by type of components, sex, age and time" produced by Statistics Denmark that provides a measure of average net wealth per person by age group produced from tax data.

Aggregation. Whenever possible, we cross-check that the aggregated-up wealth data from our surveys line up with the wealth-to-GDP ratio computed by the WID or the OECD. Table A.1 provides details on the source of both survey and aggregate data, as well as the wealth-to-GDP ratio computed from the survey, compared to the official statistic.

[44] https://www.lisdatacenter.org/our-data/lws-database/
[45] https://www.ecb.europa.eu/stats/ecb_surveys/hfcs/
[46] http://www.chfsdata.org/
[47] http://microdata.gov.in
[48] https://www.e-stat.go.jp

Table A.1: Wealth-to-GDP ratios from survey data and aggregate data

	Wealth survey data			Aggregate data		
Country	Year	Source	$\frac{W^c}{Y^c}$	Year	Source	$\frac{W^c}{Y^c}$
AUS	2014	LWS	3.59	2016	WID	5.09
AUT	2014	HFCS	2.79	2016	OECD	3.90
BEL	2014	HFCS	3.84	2016	OECD	5.74
CAN	2016	LWS	6.98	2016	WID	4.63
CHN	2013	CHFS	3.27	2016	WID	4.20
DEU	2017	LWS	3.8	2016	WID	3.64
DNK	2014	SD	2.54	2016	WID	3.42
ESP	2014	HFCS	4.96	2016	WID	5.33
EST	2014	HFCS	2.78	2016	OECD	2.64
FIN	2014	HFCS	2.33	2016	WID	2.78
FRA	2014	HFCS	3.30	2016	WID	4.85
GBR	2011	LWS	4.01	2016	WID	5.35
GRC	2014	HFCS	2.65	2016	WID	4.25
HUN	2014	HFCS	1.84	2016	OECD	2.19
IND	2013	NSS	4.01	2016	-	-
IRL	2014	HFCS	3.39	2016	CBI	2.32
ITA	2016	LWS	3.35	2016	WID	5.83
JPN	2009	NSFIE	6.11	2016	WID	4.85
LUX	2014	HFCS	3.80	2016	OECD	3.92
NLD	2014	HFCS	1.80	2016	WID	3.92
POL	2014	HFCS	3.31	2016	OECD	1.50
SVK	2014	HFCS	1.80	2016	OECD	2.17
SVN	2014	HFCS	3.11	2016	OECD	2.82
SWE	2005	LWS	2.00	2016	WID	3.81
USA	2016	SCF	4.38	2016	WID	4.28

Notes: This table summarizes our sources of wealth survey data and aggregate data. Abbreviations are described in the text. The survey-based wealth to GDP ratio W^c/Y^c is computed by aggregating household wealth using survey weights and dividing by GDP per household from the national accounts.

C.2 Robustness

In this section, we show that our results are robust to some of our main assumptions behind the calculation of compositional effects. In the interest of space, we focus here on the United States; conclusions are similar when repeating this exercise in other countries.

Alternative allocation of household to individual wealth. All our surveys measure wealth at the household level. In the main text, we obtain individual wealth by splitting up all assets equally between all members of the household that are at least as old as the head or spouse. The orange line in figure A.3, labeled "baseline", reproduces the projection from the United States under the main fertility scenario (cf figures 2 and 4). The red line shows that allocating all household

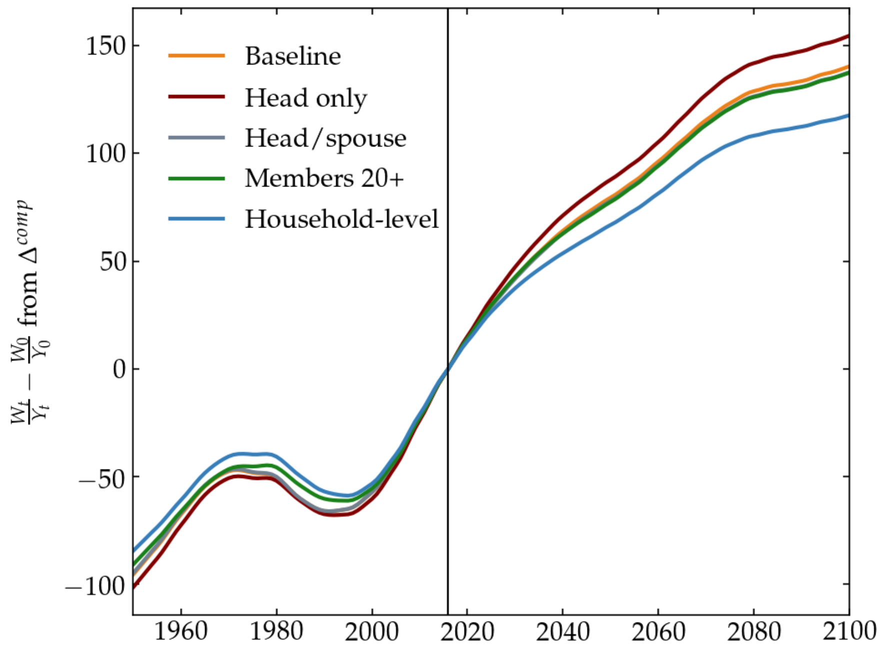


Figure A.3: Predicted change in U.S. W/Y from composition: alternative assumptions

Notes: This figure depicts the evolution of the predicted change in the wealth-to-GDP ratio from the compositional effect, calculated using equation (22) from t =1950 to 2100. The orange line corresponds to our baseline case, where the wealth of households is allocated equally to all members at least as old as the head or the spouse. The red line shows the outcome when wealth is allocated to the head of household only, the gray line to the head and the spouse equally, and the green line to all members aged 20 or more. The blue line presents the outcome when the analysis is conducted at the household-level rather than at the individual level.

wealth to the head increases the compositional effect a little, since heads tend to be older on average; the grey line shows that allocating all wealth equally to head as spouse, as in Poterba (2001), or equally to all household members aged 20 or older. delivers results extremely close to our baseline.

Constructing compositional effects at the household level. All our exercises in the main text of section 3.2, as well as the alternative considered in the previous paragraph, are conducted at the individual level. To gauge the importance of the household vs individual distinction, here we calculate compositional effects at the household level instead.

We first obtain the age-wealth and labor income profiles at the household level, summing the pre-tax labor income of each household member. To convert the age distribution of the population over individuals to an age distribution over households, we use the PSID to estimate a mapping that gives, for each age j, the age of the household head than an average individual of age j lives with.

With this data in hand, we recompute the compositional effect Δ^{comp}. Figure A.3 reports the projected change in W/Y from this exercise under the baseline fertility scenarios. The dashed line reproduces the central individual-level compositional effect from the main text. Overall, the timing of the projected changes in W/Y change slightly, but the overall magnitude remains close.

Table A.2: Sensitivity of the predicted change in W/Y to the choice of base year

Panel A. Predicted change in log W/Y from composition between 1950 and 2016

	h_j year												
a_j year	1974	1979	1986	1991	1994	1997	2000	2004	2007	2010	2013	2016	DH-t
1989	22.8	22.5	22.3	21.7	21.2	20.7	21.1	20.1	19.8	19.1	19.0	19.3	29.3
1992	22.7	22.4	22.2	21.6	21.1	20.6	21.0	20.1	19.8	19.0	18.9	19.2	29.2
1995	24.3	24.0	23.8	23.2	22.7	22.2	22.6	21.7	21.4	20.6	20.5	20.8	30.8
1998	23.6	23.3	23.1	22.6	22.0	21.5	21.9	21.0	20.7	20.0	19.9	20.1	30.2
2001	23.8	23.5	23.3	22.8	22.2	21.8	22.2	21.2	20.9	20.2	20.1	20.3	30.4
2004	25.4	25.1	24.9	24.3	23.8	23.3	23.7	22.8	22.5	21.7	21.6	21.9	31.9
2007	24.8	24.5	24.3	23.8	23.2	22.7	23.1	22.2	21.9	21.2	21.1	21.3	31.4
2010	27.9	27.6	27.4	26.9	26.3	25.8	26.2	25.3	25.0	24.3	24.2	24.4	34.5
2013	26.7	26.4	26.2	25.6	25.1	24.6	25.0	24.0	23.7	23.0	22.9	23.2	33.2
2016	28.1	27.9	27.6	27.1	26.5	26.1	26.5	25.5	25.2	24.5	24.4	24.6	34.7
DH-t	29.2	29.0	28.8	28.2	27.6	27.2	27.6	26.6	26.3	25.6	25.5	25.8	35.8

Panel B. Predicted change in log W/Y from composition between 2016 and 2100

	h_j year												
a_j year	1974	1979	1986	1991	1994	1997	2000	2004	2007	2010	2013	2016	DH-t
1989	26.5	26.1	25.8	25.5	25.5	25.2	25.2	24.7	24.3	24.0	23.6	23.5	27.6
1992	22.5	22.1	21.8	21.5	21.5	21.2	21.2	20.7	20.3	20.0	19.6	19.4	23.6
1995	25.5	25.1	24.8	24.6	24.5	24.2	24.2	23.7	23.3	23.0	22.6	22.5	26.7
1998	22.7	22.4	22.0	21.8	21.7	21.4	21.4	20.9	20.5	20.2	19.8	19.7	23.9
2001	22.5	22.1	21.8	21.5	21.4	21.2	21.1	20.7	20.2	20.0	19.5	19.4	23.6
2004	25.7	25.3	25.0	24.7	24.7	24.4	24.4	23.9	23.5	23.2	22.7	22.6	26.8
2007	24.5	24.1	23.8	23.5	23.5	23.2	23.1	22.7	22.3	22.0	21.5	21.4	25.6
2010	28.7	28.4	28.1	27.8	27.7	27.5	27.4	26.9	26.5	26.3	25.8	25.7	29.9
2013	28.2	27.9	27.5	27.3	27.2	26.9	26.9	26.4	26.0	25.7	25.3	25.2	29.4
2016	30.9	30.5	30.2	29.9	29.9	29.6	29.5	29.1	28.7	28.4	27.9	27.8	32.0
DH-t	28.0	27.7	27.3	27.1	27.0	26.7	26.7	26.2	25.8	25.5	25.1	25.0	29.2

Notes: This table reports values of the predicted change in log W/Y from the compositional effect from equation (22), for alternative base years of the age-wealth and the age-labor income profiles. Panel A considers the period 1950 to 2016, and panel B 2016 to 2100. Every column correspond to an alternative base year for the age-labor income profile, and every row to an alternative base year for the age-wealth profile. The last row and column correspond to the cases where we use the average age effect, with all growth loading on time effects (DH-t).

Alternative choice of base year profiles Table A.2 explores how the magnitude of the compositional effects Δ^{comp} changes when we change the base year 0 we use to construct the age profiles a_{j0} and h_{j0} in equation (11). Panel A considers the implied predicted change in W/Y from equation (22) for the period 1950 to 2016, and panel B does the same for the period 2016 to 2100. In each panel, we use as base year every possible combination of survey years for the age-wealth profile a_{j0} from the SCF (rows) and age-labor income profile h_{j0} from the LIS (columns). In the last row and column label "DH-t", we use the age effects extracted from a time-age-cohort decomposition in the style of Hall (1968) and Deaton (1997), imposing that all growth loads on time effects. As discussed in footnote 15, it is important to load growth on time effects to recover the age profiles that are the correct input into Proposition 1.

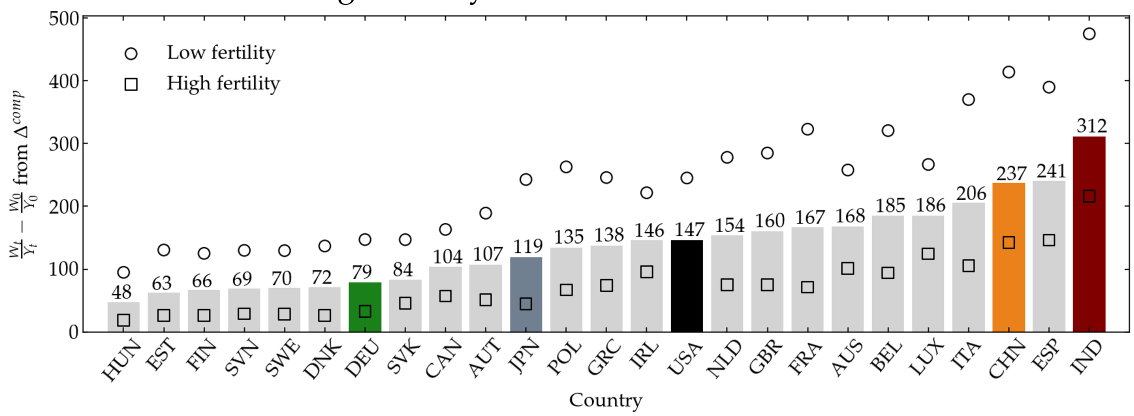
The table indicates that the predicted change in W/Y from composition lies between 64 and 135 pp for 1950-2016 and between 58 and 133 pp for 2016-2100. There are only small variations across columns, indicating a modest importance of the age-labor income profile for the magnitude of Δ^{comp}. The choice of base year the age-wealth profile matters a little more, with later years tending to give rise to larger compositional effects as the age-wealth profile rises increasingly steeply over time.

C.3 Additional results for section 3.2

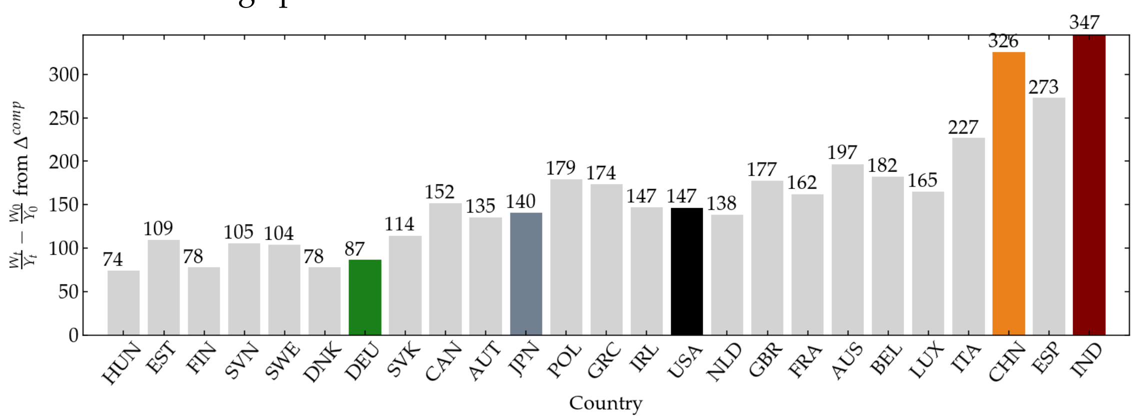
Historical predicted change in W/Y from composition effects vs actual change in W/Y. Table A.3 contrasts, for a range of countries for which the World Inequality Database contains a sufficiently long time series of measured wealth-to-GDP ratios, the measured change in the log of W/Y (labelled "Data") relative to the compositional effect Δ_t^{comp} (labelled "Comp"). The latter is constructed from equation 11 using baseline year age profiles interacted with the actual change in population distributions over the period reported. Both columns are multiplied by 100 to be interpretable as percentage points. The compositional effect predicts an increase in W/Y in every country, consistent with what occurred. For countries like the United States and the Netherlands, the magnitudes also line up closely; for Spain, the compositional effect overpredicts the historical magnitude, while for most other countries the historical increase in W/Y is greater than the compositional effect alone would predict. If demographics was the only force driving wealth-to-GDP ratios then our theory suggests that the rise in W/Y should be less than what is predicted by the compositional effect due to the endogenous response of asset returns; the fact that many countries experienced larger increases suggests that other forces, such as declining productivity growth, have also been at play.

Role of heterogeneity in demographic change vs age profiles. Figure A.4 presents the predicted change in W/Y between 2016 and 2100 from the compositional effect and isolates the contributions from demographic forces and from the age-p ofiles. Panel A repeats the results from section 3.2, ranking countries from lowest to highest compositional effect. It also presents the results under the two UN fertility scenarios. To isolate the contribution from demographic forces, panel B computes the compositional effect where age-profiles in all countries are identical to the US profile. To isolate the contribution from the profiles, panel C computes the compositional effect where population distributions of the US are used in every country. Panels B and C show that both the shape of the profiles and the changes in population distributions matter to the compositional effect, but that the demographic forces play a much more important role in generating shift-shares that are high and heterogeneous across countries.

A. Baseline and low/high fertility scenarios



B. At common age profiles



C. At common demographic change

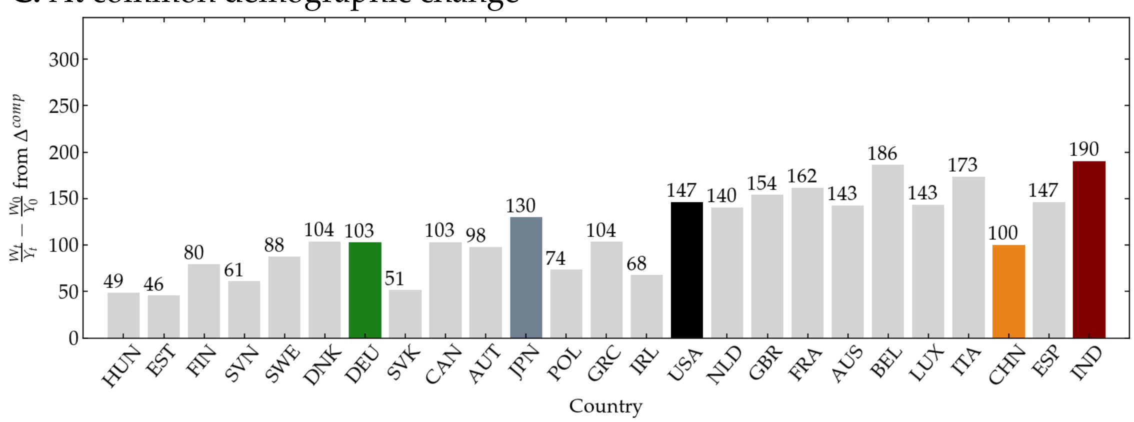


Figure A.4: Predicted change in W/Y from composition between 2016 and 2100

Notes: Panel A presents the change in W/Y between 2016 and 2100 from equation (22) as well as its value using the low fertility (circles) and high fertility (squares) scenarios. Panel B does this calculation again, assuming that all countries have US age profiles of assets and income. Panel C does this calculation again, assuming all countries have the US age distribution in every year.

Table A.3: Historical change in $log(W/Y)$ vs predicted change from Δ^{comp}

Country	Period	Data	Comp.
AUS	1960-2016	59.8	15.6
CAN	1970-2016	82.6	19.2
CHN	1978-2016	140.9	16.8
DEU	1950-2016	67.4	23.7
DNK	1973-2016	80.2	13.8
ESP	1950-2016	19.1	27.6
FIN	2011-2016	9.2	5.5
FRA	1950-2016	109.3	21.4
GBR	1950-2016	37.5	18.9
GRC	1997-2016	17.3	8.7
IND	1950-2016	23.2	10.9
ITA	1966-2016	108.8	23.5
JPN	1970-2016	66.0	42.5
NLD	1997-2016	23.4	21.1
SWE	1950-2016	48.8	19.6
USA	1950-2016	31.6	27.5

C.4 Additional results for sections 3.3 and 3.4

Age profiles of consumption and assets. Figure A.5 presents the age distributions of consumption (orange lines) and asset holdings (red lines), constructed using the procedure described in section 3.3. The consumption profile is backed out of the asset profile and the profile of net income. Net income includes all taxes and transfers; since this measure is not available in most surveys, we back it out of aggregate information on taxes and transfers. In practice, we use net income from our quantitative model of section 4, which is constructed using that information for each country.

Applying equation (16) at each point in time to predict NFAs. Figure A.6 reproduces Figure 7, when we apply equation (16) at each point in time to predict NFAs. Specifically, we apply equation

$$\log\left(1+\frac{NFA_t^c}{W_0^c}\right) \simeq \Delta_t^{comp,c} - \overline{\Delta}_t^{comp} + \left(\epsilon^{c,d} + \epsilon^{c,s} - \left(\bar{\epsilon}^d + \bar{\epsilon}^s\right)\right)(r_t - r_0)$$

where $r_t - r_0$ is, in turn, calculated by applying equation (14) at each point in time,

$$r_t - r_0 \simeq -\frac{1}{\bar{\epsilon}^d + \bar{\epsilon}^s}\bar{\Delta}_t^{comp}$$

where we take $\bar{\epsilon}^d$ and $\bar{\epsilon}^s$ to be the steady state elasticities calculated using our sufficient statistics.[49]

[49] In principle, a more complex sequence-space Jacobian matrix should be used to do these calculations, in practice there does not exist sufficient statistic expression for the Jacobian that underlies $\bar{\epsilon}^d$. Figure 9 shows that this approximation works fairly well in the context of our structural model.

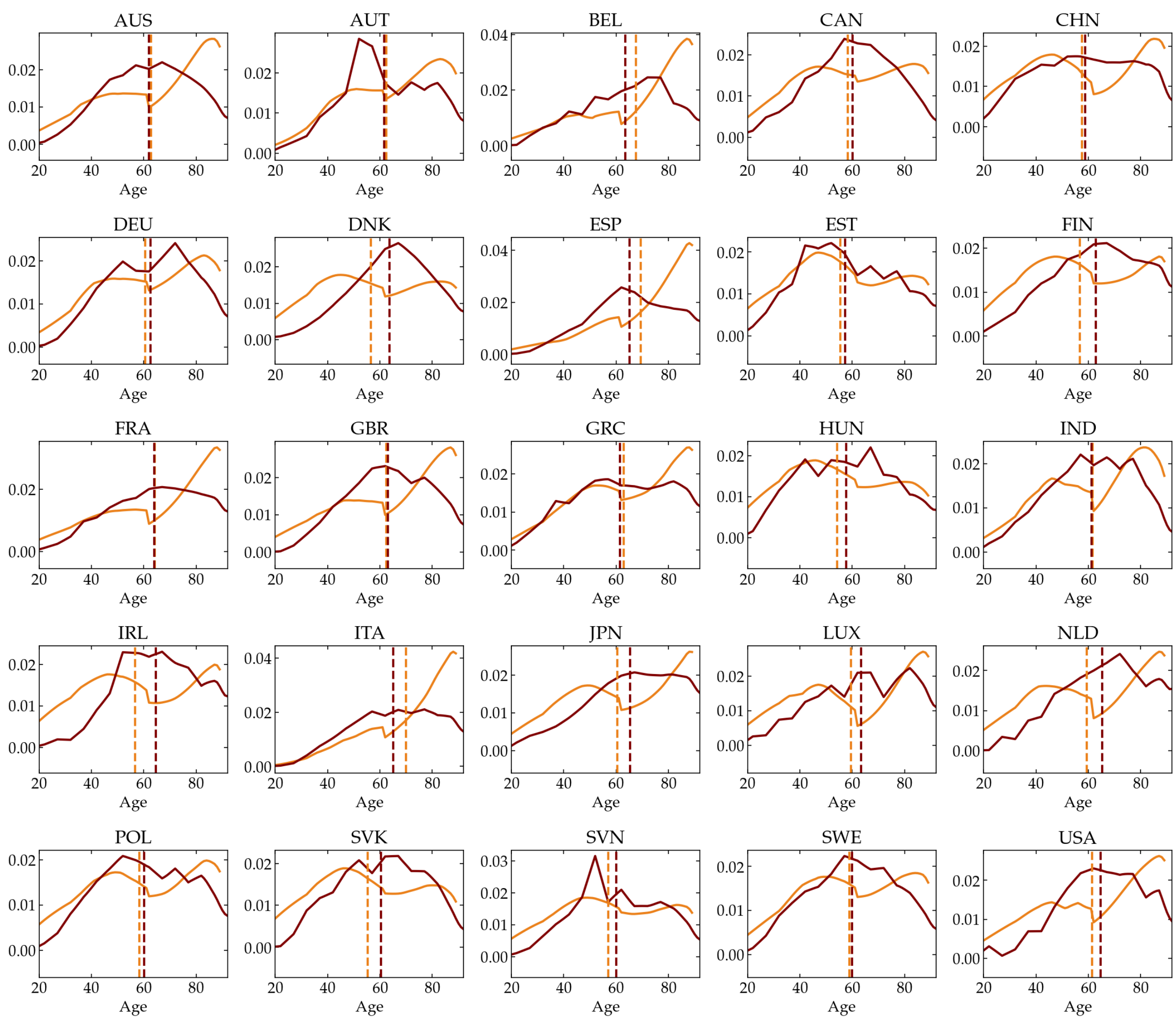


Figure A.5: Distribution of ages of consumption and wealth in each country.

Notes: This figure presents the age distributions of consumption (orange lines) and asset holdings (red lines). The dashed vertical lines depict the average ages of consumption and asset holdings.

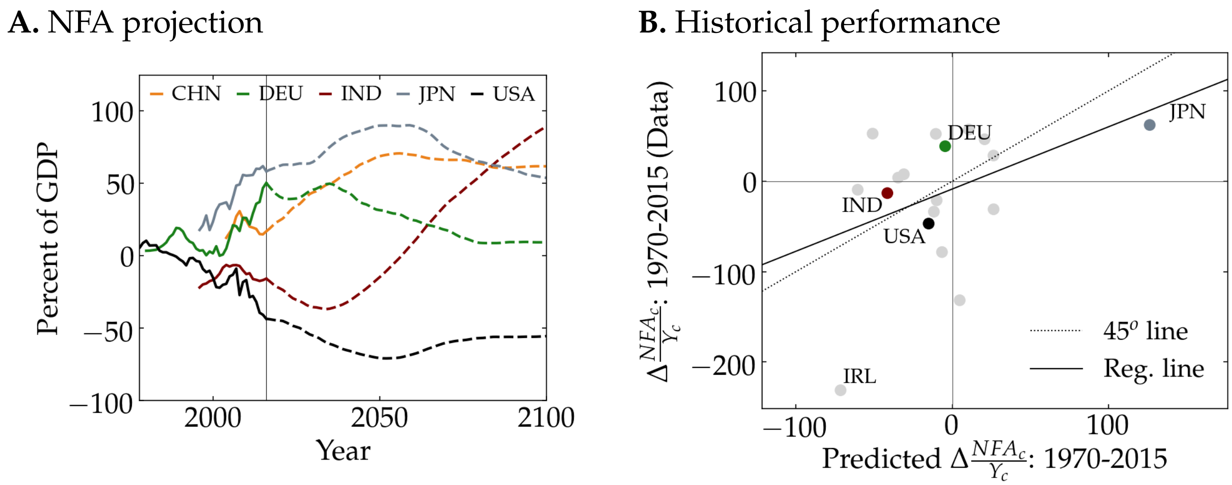


Figure A.6: Using equation (16) to project NFAs

The main findings from Figure A.6 are unchanged, indicating that the interest rate adjustment term does not play a major role when it comes to forecasting NFAs. This is because it only matters to the extent that elasticities of supply and demand differ across countries, and the heterogeneity we calculate from our sufficient statistics is relatively limited.

D Appendix to Section 4

D.1 Full model setup

Here, we describe a the model in section 4 in detail, including the parametric assumptions that we do to allow for the calibration. To facilitate the exposition, we first describe the full model for one country, omitting the country superscript c, and use the model to define a small open economy equilibrium for a fixed sequence $\{r_t\}$. The world equilibrium is subsequently defined as a sequence $\{r_t\}$ that clears the global asset market given the small economy equilibria for all countries.

Demographics. The demographic parameters of the economy are a sequence of births $\{N_{0t}\}_{t\geq -1}$, a sequence of age- and time-specific mortality rates $\{\phi_{jt}\}_{t\geq -1}$ for individuals between age j and $j+1$, a sequence of net migration levels $\{M_{jt}\}_{-1\leq t, 0\leq j\leq T-1}$, as well as an initial number of agents by age $N_{j,-1}$. The assumption that demographic starts at $t=-1$ is done for technical reasons; it allows us to correctly account for migration and bequests received at time $t=0$. Given these parameters, the population variables for $t\geq 0$ evolves according to the exogenous N_{0t} and

$$N_{jt} = (N_{j-1,t-1} + M_{j-1,t-1})\phi_{j-1,t-1}, \quad \forall t\geq 0, j>0 \tag{A.19}$$

for $j>0$. As in section 2, we write $N_t \equiv \sum_j N_{jt}$ for the total population at time t, and $\pi_{jt} \equiv \frac{N_{jt}}{N_t}$ for the age distribution of the population.

Agents' problem. The basic setup is the same as in section 2, with heterogeneous individuals facing idiosyncratic income risk. We restrict the income process so that effective labor supply ℓ_{jt}

is the product of a deterministic term $\bar{\ell}_j$ that varies across ages, a fixed effect θ and a transitory component ϵ, where both the fixed effect and the transitory component have a mean 1. The log transitory component follows a finite-state Markov process with a transition matrix across years $\Pi^{\epsilon}(\epsilon|\epsilon_{-})$ from ϵ_{-} to ϵ, calibrated to have a persistence χ_{ϵ} and a standard deviation v_{ϵ}, while the log permanent component follows a discrete Markov process across generations with a transition matrix $\Pi(\theta|\theta_{-})$ from θ_{-} to θ calibrated to have a persistence χ_{θ} and a standard deviation v_{θ}. The processes are independent, and we write $\pi^{\epsilon}(\epsilon)$ and $\pi^{\theta}(\theta)$ for the corresponding stationary probability mass functions.[50]

We assume that individuals become economically active at age T^w, so that labor income at age j at time t is $w_t(1-\rho_{jt})\theta\epsilon\bar{\ell}_{jt}$, where w_t is the wage per efficiency unit as in section 2, and $\rho_{jt} \in [0,1]$ is a parameter of the retirement system indicating the fraction of labor that households of age j are allowed to supply at time t. After retirement, agents receive social security payments $w_t\rho_{jt}\theta d_t$ in proportion to their permanent type, where d_t encodes a time-varying social security replacement rate.

The state for an individual at age j and time t is given by the fixed effect θ, the transitory effect ϵ, and asset holdings $\mathfrak{a}$, and their value function is given by

$$\begin{aligned} V_{jt}(\theta,\epsilon,\mathfrak{a}) &= \max_{c,\mathfrak{a}'} \frac{c^{1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}} + \Upsilon Z_t^{\nu-\frac{1}{\sigma}}\left(1-\phi_{jt}\right)\frac{(\mathfrak{a}')^{1-\nu}}{1-\nu} + \phi_{jt}\frac{\beta_{j+1}}{\beta_j}\mathbb{E}\left[V_{j+1,t+1}(\theta,\epsilon',\mathfrak{a}')|\epsilon\right] \\ c+\mathfrak{a}' &\leq w_t\theta\left[(1-\rho_{jt})(1-\tau_t)\bar{\ell}_{jt}\epsilon + \rho_{jt}d_t\right] + (1+r_t)\mathfrak{a} + b^r_{jt}(\theta) \\ -\bar{\mathfrak{a}}Z_t &\leq \mathfrak{a}'., \end{aligned} \tag{A.20}$$

which determines decision function $c = c_{jt}(\theta,\epsilon,\mathfrak{a})$ and $a' = a_{j+1,t+1}(\theta,\epsilon,\mathfrak{a})$ for consumption and next-period assets.

The term $\frac{c^{1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}}$ represents the flow utility of consumption, and $\Upsilon Z_t^{\nu-\frac{1}{\sigma}}\left(1-\phi_{jt}\right)(\mathfrak{a}')^{1-\nu}/(1-\nu)$ represents the utility from giving bequests $\mathfrak{a}'$. The bequest utility is scaled by mortality risk $1-\phi_{jt}$, since agents only give bequests if they die, and $\nu \geq \frac{1}{\sigma}$ captures potential non-homotheticicities in bequests, which has been shown to generate more realistic levels of wealth inequality (De Nardi, 2004). The scaling factor $Z_t^{\nu-\frac{1}{\sigma}}$ ensures balanced growth in spite of this non-homotheticity. The term $b^r_{jt}(\theta)$ represents bequests received, and is allowed to vary according to the agent's permanent type.

State distribution. To determine the evolution of states, we assume that the distribution of individuals across θ and ϵ is in the stationary distribution for all ages, times, as well as for arriving and leaving migrants. This implies that the joint distribution across $(\theta,\epsilon,\mathfrak{a})$ is fully characterized by

$$H_{jt}(\mathfrak{a}|\theta,\epsilon) = \mathbb{P}(\mathfrak{a}_{jt} \leq \mathfrak{a}|\theta,\epsilon),$$

where H_{jt} is the conditional probability distribution of assets given θ and ϵ.[51]

[50]Discrete processes are used to facilitate notation. The calibration to the persistence and standard deviation is done using Tauchen's method applied to a Gaussian AR(1) process with a given persistence, standard deviation, and mean.

[51]Formally, given H_{jt}, the joint distribution function $\tilde{H}_{jt}$ of $(\theta,\epsilon,\mathfrak{a})$ can be written $\tilde{H}_{jt}(\theta,\epsilon,\mathfrak{a}) = \sum_{\theta'\leq\theta,\epsilon'\leq\epsilon}\pi^{\theta}(\theta')\pi^{\epsilon}(\epsilon')G_{jt}(\mathfrak{a}|\theta',\epsilon')$.

Over time, the distribution evolves according to

$$H_{j+1,t+1}(\mathfrak{a}|\theta,\epsilon) = \sum_{\epsilon_-} \frac{\Pi^{\epsilon}(\epsilon|\epsilon_-)\pi^{\epsilon}(\epsilon_-)}{\pi^{\epsilon}(\epsilon)} \int_{\mathfrak{a}'} \mathbb{I}(\mathfrak{a}_{j,t}(\mathfrak{a}',\theta,\epsilon) \leq \mathfrak{a}) dH_{j,t}(\mathfrak{a}',\theta,\epsilon) \quad \forall j > T^w, \tag{A.21}$$

where $\mathfrak{a}_{jt}$ is the decision function for assets implied by the agents' problem (A.20). Note that (A.21) implicitly assumes that death is independent of asset holdings and that migrants have the same distribution of assets as residents. At time zero, there is an exogenous distribution of assets $H_{j0}(\cdot|\theta,\epsilon)$ for each age group. As a boundary condition, we assume that individuals do not have any assets before working life starts:

$$H_{j,t}(\mathfrak{a}|\theta,\epsilon) = \mathbb{I}(\mathfrak{a} \geq 0) \quad \forall \theta, \epsilon, j \leq T^w, 0 \leq t, \tag{A.22}$$

where $\mathbb{I}$ is the indicator function.

Bequest distribution We model partial intergenerational wealth persistence by assuming that all bequests from individuals of type θ are pooled and distributed across the types θ' of survivors in accordance with the intergenerational transmission of types. Formally, the total amount of bequests received by agents of type θ of age j at time t is

$$\begin{aligned} N_{jt} b^r_{jt}(\theta) = & F_j \sum_{\theta_-} \left(\frac{\Pi^{\theta}(\theta|\theta_-)\pi^{\theta}(\theta_-)}{\pi^{\theta}(\theta)} \right) \times \\ & \sum_{k=0}^{T} \left[N_{k,t-1} + M_{k,t-1}\right] (1-\phi_{k,t-1}) \times \int_{\mathfrak{a}} \sum_{\epsilon} \pi^{\epsilon}(\epsilon) \mathfrak{a} dH_{kt}(\mathfrak{a};\theta_-,\epsilon) \end{aligned} \tag{A.23}$$

Here, $\sum_k [N_{k,t-1} + M_{k,t-1}](1-\phi_{k,t-1}) \int_{\mathfrak{a}} \sum_{\epsilon} \pi^{\epsilon}(\epsilon) \mathfrak{a} dH_{kt}(\mathfrak{a};\theta_-,\epsilon)$ captures the total amount of bequests given by individuals of type θ_-. The timing is that migrants arrive before the death event and that interest rate accrues after the death event. A share $\frac{\Pi^{\theta}(\theta|\theta_-)\pi^{\theta}(\theta_-)}{\pi^{\theta}(\theta)}$ of these bequests is given to agents of type θ, capturing partial intergenerational transmission by using the probability that an agent of type θ has a parent of type θ_-.

Note that an aging population alters the relative number of agents that give relative to the number of agents that receive bequests, which ceteris paribus increases bequest sizes. The migrants are included, assuming that migrants have the same mortality as the overall population, and that migrants who plan to arrive at t but die between $t-1$ and t augment the bequest pool in the receiving country.

Aggregation Given the decision functions c_{jt} and $\mathfrak{a}_{j+1,t+1}$ and the distribution across states, aggregate consumption and assets satisfy

$$\begin{aligned} W_t &= \sum_{j=0}^{T} N_{jt} \times \int_{\mathfrak{a}} \sum_{\epsilon,\theta} \pi^{\epsilon}(\epsilon)\pi^{\theta}(\theta)[\mathfrak{a} dH_{jt}(\mathfrak{a};\theta,\epsilon) + b^r_{jt}(\theta)] \\ C_t &= \sum_{j=0}^{T} N_{jt} \times \int_{\mathfrak{a}} \sum_{\epsilon'} c_{jt}(\theta_-,\epsilon,\mathfrak{a}) \pi^{\epsilon}(\epsilon) dH_{jt}(\mathfrak{a}|\theta_-,\epsilon). \end{aligned} \tag{A.24}$$

Note that bequests received are included in the definition of today's ingoing assets.

Production. As in section 2, the production is given by a CES aggregate production function, markets are competitive, there are no adjustment costs in capital, and there is labor-augmenting growth at a constant rate γ. These assumptions imply the following equations

$$Y_t = F(K_t, Z_t L_t) \equiv \left(\alpha K_t^{\frac{\eta-1}{\eta}} + (1-\alpha)[Z_t L_t]^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}} \tag{A.25}$$

$$Z_t = (1+\gamma)^t Z_0 \tag{A.26}$$

$$r_t = F_K(K_t, Z_t L_t) - \delta \tag{A.27}$$

$$w_t = Z_t F_L(K_t, Z_t L_t) \tag{A.28}$$

$$K_{t+1} = (1-\delta)K_t + I_t \tag{A.29}$$

$$L_t = \sum_{j=0}^{T} N_{jt}(1-\rho_{jt})\bar{\ell}_{jt}, \tag{A.30}$$

where the last line uses that $\mathbb{E}\theta\epsilon = 1$ to obtain that average effective labor supply is $\bar{\ell}_{jt}$ of individuals of age j.

Government. The government purchases G_t goods and sets the retirement policy ρ_{jt}, the tax rate τ_t, and the benefit generosity d_t. It faces the flow budget constraint

$$G_t + \sum_{j=0}^{T} N_{jt} w_t \rho_{jt} d_t + (1+r_t)B_t = \tau_t w_t \sum_{j=0}^{T} N_{jt}(1-\rho_{jt})\bar{\ell}_{jt} + B_{t+1}, \tag{A.31}$$

where a positive B_t denotes government borrowing. In the aggregation, we use that $\mathbb{E}\theta\epsilon = 1$ within every age group, which means that average benefits and labor income per age-j person are $w_t \rho_{jt} d_t$ and $w_t(1-\rho_{jt})\bar{\ell}_{jt}$ respectively.

The government targets an eventually converging sequence $\left\{\frac{B_{t+1}}{Y_{t+1}}\right\}_{t\geq 0}$. To reach this target, we assume that the government uses a fixed retirement policy ρ_{jt}, and adjusts the other instruments using a fiscal rule defined in term of the "fiscal shortfall" SF_t, defined as

$$\frac{SF_t}{Y_t} \equiv \frac{\bar{G}}{Y} + \frac{\sum_{j=0}^{T}[\rho_{j,t}\bar{d} - \bar{\tau}(1-\rho_{j,t})\bar{\ell}_{jt}]N_{jt}w_t}{Y_t} + (r_t - g_t)\frac{B_t}{Y_t} - (1+g_t)\left[\frac{B_{t+1}}{Y_{t+1}} - \frac{B_t}{Y_t}\right], \tag{A.32}$$

where $g_t = \frac{Y_{t+1}}{Y_t} - 1$. The fiscal shortfall is positive at time t if expenditures minus revenues is too high to reach the debt target when the instruments G, d, and τ are set at some reference levels $\bar{G}$, $\bar{d}$ and $\bar{\tau}$. The fiscal rule consists of three weights on the instruments $\varphi^G, \varphi^\tau, \varphi^d$ that satisfy

$$\varphi^G SF_t = -(G_t - \bar{G}) \forall t \geq 0 \tag{A.33}$$

$$\varphi^\tau SF_t = (\tau_t - \bar{\tau}) \times w_t \sum_{j=0}^{T} N_{jt}\bar{\ell}_{jt}(1-\rho_{jt}) \forall t \geq 0 \tag{A.34}$$

$$\varphi^d SF_t = -(d_t - \bar{d}) \times \left(w_t \sum_{j=0}^{T} N_{jt}\rho_{jt}\right) \forall t \geq 0 \tag{A.35}$$

$$1 = \varphi^G + \varphi^\tau + \varphi^d. \tag{A.36}$$

This specification ensures that the three instruments contribute the respective shares φ^G, φ^τ, and φ^d to closing the fiscal shortfall. Since $\varphi^G + \varphi^\tau + \varphi^d = 1$, the government successfully targets $\frac{B_{t+1}}{Y_{t+1}}$.

Market clearing. The assets in the economy consist of capital K_t, government bonds B_t, and foreign assets NFA_t. The asset market clearing condition is

$$K_t + B_t + NFA_t = W_t. \tag{A.37}$$

Given the other equilibrium conditions, (A.37) can be used to derive the goods market clearing condition[52]

$$NFA_{t+1} - NFA_t = NX_t + r_t NFA_t + W_{t+1}^{mig}. \tag{A.38}$$

Here, $NX_t \doteq Y_t - I_t - C_t - G_t$ is net exports at time t and

$$W_{t+1}^{mig} \doteq \sum_{j=1}^{T} M_{j-1,t} \sum_{\theta,\epsilon} \pi^\theta(\theta)\pi^\epsilon(\epsilon) \left(\int_{\mathfrak{a}} \mathfrak{a} dH_{j,t+1}(\mathfrak{a},\theta,\epsilon) + b_{j,t+1}^\theta(\theta) \right)$$

is the assets at time t that comes from migrants.

Small open economy equilibrium. A small open economy equilibrium is defined for:

- A sequence of interest rates $\{r_t\}_{t=0}^\infty$
- A government fiscal rule $\{B_{t+1}/Y_{t+1}, \rho_{jt}, \varphi^G, \varphi^\tau, \varphi^d, \bar{G}, \bar{\tau}, \bar{d}\}_{t=0}^\infty$
- A sequence of average effective labor supplies $\{\bar{\ell}_{jt}\}_{0\le t, T^w \le j \le T}$
- An initial distribution of assets $\{H_{j0}(a|\theta,\epsilon)\}_{j=0}^T$
- Technology parameters $\{Z_0, \gamma, \delta, \nu, \alpha\}$
- Demographics: initial $\{N_{j,-1}\}_{j=0}^T$ and forcing parameters $\{M_{jt}, \phi_{jt}, N_{0,t+1}\}_{-1\le t, 0\le j\le T}$
- Initial aggregate variables K_0, B_0, A_0

The equilibrium consists of:

- Individual decision functions: $c_{jt}(\theta,\epsilon,\mathfrak{a})$, $\mathfrak{a}'_{jt}(\theta,\epsilon,\mathfrak{a})$
- A sequence of asset distribution functions $\{H_{jt}(a;\theta,\epsilon)\}_{1\le t, T^w\le j\le T}$
- Government policy variables $\{G_t, \tau_t, d_t\}_{t\ge 0}$
- A sequence of wages $\{w_t\}_{t\ge 0}$
- A sequence of bequests received $\{b_{jt}(\theta)\}_{t\ge 0}$
- A sequence of aggregate quantities $\{Y_t, L_t, I_t, K_{t+1}, W_t, C_t, NFA_t\}_{t\ge 0}$

It is characterized by that

[52] Combine the aggregated household budget constraint with the government budget constraint (A.31), the capital evolution equation (A.29), and the asset market clearing condition (A.37).

- r_0 is consistent with $K_0 \implies$ (A.27) holds given K_0 and $L_0 = \sum_j N_{j0}(1-\rho_{j0})\bar{\ell}_{j0}$
- W_0 is consistent with H_{j0}, that is, (A.24) holds
- Individual decision functions solve (A.20).
- The set of H_{jt}'s satisfies the evolution equation (A.21) and the boundary condition (A.22)
- The government policy variables satisfy (A.32)-(A.35).
- Equations (A.25)-(A.30) hold.
- A_t satisfies (A.24) for $t \geq 0$
- $NFA_t = W_t - K_t - B_t$, with B_0 given by the initial condition, and B_{t+1}/Y_{t+1} by the government fiscal rule.
- Bequests received $b_{jt}(\theta)$ satisfy (A.23)

World-economy equilibrium. Consider a set of countries $c \in \mathcal{C}$. A *world-economy equilibrium* is a sequence of returns $\{r_0, \{r_t\}_{t\geq 1}\}$ and a set of corresponding sequences of prices and allocations $\mathcal{S}^c$ for each economy c such that each $\mathcal{S}^c$ is a small open economy equilibrium, and that their NFAs satisfy

$$\sum_{c\in\mathcal{C}} NFA_t^c = 0 \quad \forall t \geq 0 \tag{A.39}$$

D.2 Proof of proposition 5

Let Φ_t^c capture all demographic variables in a country: population shares, fertility, mortality, migration. Given fixed r and B^c/Y^c, long-run government policy only depends on Φ^c. Wages per unit of effective labor only depend on r. Assuming that the steady state of the household problem is unique conditional on demographics, wages, and government policy, we can therefore express it as a function of (r, Φ^c).[53] Let $\frac{W^c}{Y^c}(r, \Phi^c)$ denote the resulting steady-state wealth-to-output ratio.

Output, normalized by technology, only depends on aggregate effective labor supply, which is a function of Φ^c (both directly through the number of people at each age and indirectly through government retirement policy), and the capital-to-effective-labor ratio, which is a function of r. Hence we can write each country's share of global GDP as $\frac{Y^c}{Y}(r, \nu, \Phi)$.

Repeat the proofs of propositions 2 and 3 in appendix B.3 from (A.6) on, replacing π with Φ everywhere, and replacing $\frac{W^c}{Y^c}(r_0, \Phi_{LR}^c) - \frac{W^c}{Y^c}(r_0, \Phi_0^c)$ with $\Delta_{LR}^{soe,c}$ and $\sum_c \omega^c \left(\frac{W^c}{Y^c}(r_0, \Phi_{LR}^c) - \frac{W^c}{Y^c}(r_0, \Phi_0^c)\right)$ with $\bar{\Delta}_{LR}^{soe}$.

[53] Aside from bequests, we have a standard incomplete markets household problem and this would be a standard result. Bequests introduce some complication, since bequests depend on the endogenous distribution of assets, but household asset policy also depends on realized and expected bequests. The solution to the household problem is a fixed point of this process. We assume that the fixed point is unique and a global attractor; in practice, we have found that this assumption is always satisfied.

D.3 Steady-state equations and calibration details

Steady-state equations. Our calibration targets a stationary equilibrium associated with a constant rate of return r. Most elements are standard: we assume constant technology parameters $\{\gamma, \delta, \nu, \alpha, \bar{\ell}_j\}$, a constant bond-to-output ratio $\frac{B}{Y}$, retirement policy ρ_j, tax rate τ, social security generosity d, and government consumption-to-output ratio G/Y. We also assume that there is a fixed distribution of assets $H_j(\tilde{\mathfrak{a}}|\theta, \epsilon)$, where $\tilde{\mathfrak{a}}$ is assets normalized by technology (again, we drop the country superscripts in the description of each country, and reintroduce them when we define the world equilibrium).

The non-standard element is that we introduce a counterfactual flow of migrants to ensure a time-invariant population distribution at the 2016 level, and a match to the 2016 population growth rate. In particular, demography consists of constant mortality rates, a fixed age distribution, a constant population growth rate, and a constant rate of migration by age m_j:

$$\phi_{jt} \equiv \phi_j, \quad \pi_{jt} \equiv \phi_j, \quad N_t = (1+n)^t N_0, \quad m_j \equiv \frac{M_j}{N},$$

and the net migration by age is given by

$$m_{j-1} \equiv \frac{M_{j-1}}{N} = \pi_j \frac{1+n}{\phi_{j-1}} - \pi_{j-1}, \tag{A.40}$$

which ensures that (A.19) holds given a fixed age distribution of population. The notation $\frac{M_{j-1}}{N}$ without a time index is used to indicate the constant ratio $\frac{M_{j-1,t}}{N_t}$. It will be used throughout whenever the ratio of two variables is constant over time.

In normalized form, the consumer problem is

$$\begin{aligned} \tilde{V}_j(\theta, \epsilon, \tilde{\mathfrak{a}}) &= \max_{\tilde{c}, \tilde{\mathfrak{a}}'} \frac{\tilde{c}^{1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}} + \Upsilon(1+\gamma)^{1-\nu}\left(1-\phi_j\right)\frac{\left(\tilde{\mathfrak{a}}'\right)^{1-\nu}}{1-\nu} + \tilde{\beta}\phi_j \mathbb{E}\left[\tilde{V}_{j+1}(\theta, \epsilon', \tilde{\mathfrak{a}}')|\epsilon\right] \\ c + (1+\gamma)\tilde{\mathfrak{a}}' &\leq \tilde{w}_t \theta \left[(1-\rho_j)(1-\tau)\bar{\ell}_j \epsilon + \rho_j d\right] + (1+r)\tilde{\mathfrak{a}} + \tilde{b}_j^r(\theta) \\ -\tilde{\mathfrak{a}} &\leq \mathfrak{a}'(1+\gamma), \end{aligned} \tag{A.41}$$

where a variable with a $\sim$ denotes normalization by Z_t, except for $\tilde{V}_j \equiv \frac{V_{jt}}{Z_t^{1-\frac{1}{\sigma}}}$ and $\tilde{\beta} \equiv \frac{\beta}{Z_t^{1-\sigma}}$. As elsewhere in the paper, we write g for the overall growth rate of the economy

$$1+g \equiv (1+n)(1+\gamma).$$

The consumer problem implies decision functions $\tilde{c}_j(\cdot)$ and $\tilde{a}'_j(\cdot)$, where the latter denotes the choice of next period's normalized assets as a function of the state at age j. From the evolution and boundary conditions of assets (A.21) and (A.22), the stationary distribution of assets satisfies

$$H_j(\tilde{\mathfrak{a}}|\theta, \epsilon) = \begin{cases} \sum_{\epsilon_-} \frac{\Pi^\epsilon(\epsilon|\epsilon_-) \times \pi^\epsilon(\epsilon_-)}{\pi^\epsilon(\epsilon)} \int_{\tilde{\mathfrak{a}}'} \mathbb{I}\left[\tilde{\mathfrak{a}}'_{j-1}(\tilde{\mathfrak{a}}', \theta, \epsilon) \leq \tilde{\mathfrak{a}}\right] dH_{j-1}(\tilde{\mathfrak{a}}'|\theta, \epsilon) & \text{if } j > T^w \\ \mathbb{I}(\tilde{\mathfrak{a}} \geq 0) & \text{if } j = T^w \end{cases},$$

Normalized bequests satisfy

$$\pi_j \tilde{b}_j^r(\theta) = F_j \sum_{\theta_-} \left(\frac{\Pi^\theta(\theta|\theta_-)\pi^\theta(\theta_-)}{\pi^\theta(\theta)} \right) \times \sum_{k=0}^{T} \frac{[\pi_k + m_k](1-\phi_k)}{1+n} \times \int_{\tilde{a}} \sum_{\epsilon} \pi^\epsilon(\epsilon) \tilde{a} dH_k(\tilde{a};\theta_-,\epsilon) \tag{A.42}$$

Aggregate consumption and assets are

$$\frac{C}{NZ} = \sum_{j=0}^{T} \pi_j \sum_{\theta,\epsilon} \pi^\theta(\theta)\pi^\epsilon(\epsilon) \int_{\tilde{a}} c_j(\tilde{a},\theta,\epsilon) dH_j(\tilde{a},\theta,\epsilon)$$

$$\frac{W}{NZ} = \sum_{j=0}^{T} \pi_j \sum_{\theta,\epsilon} \pi^\theta(\theta)\pi^\epsilon(\epsilon) \left(\int_{\tilde{a}} \tilde{a} dH_j(\tilde{a},\theta,\epsilon) + b_j^r(\theta) \right)$$

Finally, since we assume that steady state migrants have the same distribution of assets as regular households, we have

$$\frac{A^{mig}}{NZ} = \sum_{j=1}^{T} m_{j-1} \sum_{\theta,\epsilon} \pi^\theta(\theta)\pi^\epsilon(\epsilon) \left(\int_{\tilde{a}} \tilde{a} dH_j(\tilde{a},\theta,\epsilon) + b_j^\theta(\theta) \right) \tag{A.43}$$

where we recall that m_j is the number of migrants as a share of age group j at time t, and W_j is the total amount of assets of age-j individuals.

The stationary analogues of the production sector equations (A.25)-(A.30) are

$$\frac{Y}{ZN} = F\left[\frac{K}{ZN}, \frac{L}{N}\right] \tag{A.44}$$

$$r + \delta = F_K\left[\frac{K}{ZN}, \frac{L}{N}\right] = \alpha \left(\frac{K}{Y}\right)^{-1/\eta} \tag{A.45}$$

$$\frac{w}{Z} = F_L\left[\frac{K}{ZN}, \frac{L}{N}\right] \tag{A.46}$$

$$(g+\delta)\frac{K}{Y} = \frac{I}{Y} \tag{A.47}$$

$$\frac{L}{N} = \sum_{j=0}^{T} \pi_j (1-\rho_j)\bar{\ell}_j, \tag{A.48}$$

The steady-state government budget constraint is derived from (A.31) given a fixed debt-to-output ratio

$$\frac{G}{Y} + \frac{w \times d \times \sum_j N_j \rho_j}{Y} + (r-g)\frac{B}{Y} = \tau \times \frac{wL}{Y}, \tag{A.49}$$

and the asset market and good market clearing conditions are derived from (A.37) and (A.38):

$$\frac{W}{Y} = \frac{K}{Y} + \frac{B}{Y} + \frac{NFA}{Y} \tag{A.50}$$

$$0 = \frac{NX}{Y} + (r-g)\frac{NFA}{Y} + \frac{A^{mig}}{Y(1+g)}. \tag{A.51}$$

The world asset market clearing condition is

$$\sum_c \omega^c \frac{NFA^c}{Y^c} = 0, \qquad \omega^c \equiv \frac{Y^c}{\sum_c Y^c} \tag{A.52}$$

D.4 Calibration details

Additional calibration information. All demographic data is the UN World Population Prospects, interpolated across years and ages to obtain data for each combination of year and age. For each country, we use the 2016 values for age-specific survival rates ϕ_j^c and population shares π_j^c. The population growth rate is defined as

$$1 + n^c = \frac{N^c_{2016}}{N^c_{2015}}$$

where N^c_{2016} and N^c_{2015} are the population of country c in 2016 and 2015.

Debt-to-output is from the October 2019 IMF World Economic Outlook, and the net foreign asset position from the IMF Balance of Payments and International Investment Positions Statistics, deflated by nominal GDP from the Penn World Table 9.1.

For each country, the labor-augmenting productivity growth γ^c is defined as the average growth rate between 2000 and 2016 in real GDP divided by effective labor supply. For each country, we measure real GDP as expenditure-side real GDP from the Penn World Table 9.1, effective labor supply as $L^c_t = \sum_j N^c_{jt} h^c_j$, with N^c_{jt} taken from the UN World Population Prospects, and h^c_j given by the labor income profiles defined in section 3. We define the world γ as the average of γ^c, weighted by real GDP.

Given γ^c and the elasticity of substitution between capital and labor η, the growth rate of each economy is

$$g^c = (1+n^c)(1+\gamma^c) - 1,$$

and we calibrate the investment-to-output ratios, the share parameter in the production function, and the labor share

$$\frac{I^c}{Y^c} = \frac{K^c}{Y^c}(\delta + g_c)$$

$$\alpha^c = (r+\delta)\left(\frac{K^c}{Y^c}\right)^{\frac{1}{\eta}}$$

$$s^{L,c} = 1 - (r+\delta)\frac{K^c}{Y^c},$$

where the expression for investment and α use (A.47) and (A.45). Note that this calibration ensures that the world asset market clearing condition (A.52) holds for r.

For government policy, we use the average labor wedge to target τ (OECD, 2019a). This measure includes both employer and employee social security contribution, which is consistent with treating w_t as the labor cost for employers. For d, we use data on the share of GDP spent on old age benefits, using data on benefits net of taxes (OECD, 2019b). Our main source for the retirement age is OECD's data on "Effective Age of Labor Market Exit" (OECD, 2019b). For some countries, the age provided by the OECD implies that labor market exit happens after the age at which aggregate labor income falls below implied benefit income. In those cases, we define the latter age as the date of labor market exit. Formally, this is done by calibrating the implied benefit levels for each possible retirement age, and choosing the highest age at which retirement benefits are weakly lower than net-of-tax income. Last, G/Y is calibrated residually to target (A.49) given B/Y, τ, d, and the retirement age.

For individuals, we use Auclert and Rognlie (2018) and De Nardi (2004) to target the standard deviations v_ϵ, v_θ and the persistence parameters $\chi_\epsilon, \chi_\theta$. The processes are discretized using Tauchen's method, using three states for θ and 11 states for ϵ. Both processes are rescaled to ensure that they have a mean of 1.

Additional information about calibration results. Figure A.7 and A.8 show the model fit of age profiles of wage and labor income across all countries. For the labor income profile, the orange depicts labor income $(1-\rho_{j0})\bar{\ell}_j$ in the initial steady state, and the white hollow dots depict $\bar{\ell}_j$ which become relevant as the retirement age increases.

Table A.4 provides the main parameters for all countries, table A.5 provides additional parameters for all countries, and figure A.9 shows the implied yearly discount factor $\frac{\beta^c_{j+1}}{\beta^c_j}$ for all countries and ages.

Last, figure A.10 shows the outcomes for bequests and wealth inequality in the US. Panel A compares the distribution of bequests in the model to the empirical distribution in the data. We measure it as the value of bequests at certain percentiles divided by average bequests. We take the empirical distribution from Table 1 in Hurd and Smith (2002). The legend also reports the resulting model aggregate bequests-to-GDP ratio $\frac{Beq}{Y} = 8.8\%$. Panel B compares the model Lorentz curve to the one obtained in the SCF. We see that our model produces substantial wealth inequality, with the richest 20% holding roughly 70% of wealth. However, it does not go all the way to fit the wealth inequality in the US data.

D.5 Simulating demographic change

Solution method We solve for the perfect foresight transition path between 2016 and 2300 as follows.

In every country, given the initial population distribution $\{N_{j,-1}\}_{j=0}^T$ and the forcing variables $\{M_{jt}, \phi_{jt}, N_{0,t+1}\}_{t\geq -1, 0\leq j\leq T}$ held fixed from 2100 on, we simulate population forward according to (A.19) to obtain population distributions $\{\pi_{j,t}, N_{j,t}\}_{t\geq 0, 0\leq j\leq T}$ and population growth rates $\{n_t\}_{j=0}^T$. Given the effective labor supply profile and the retirement policy, this gives a path for aggregate labor $\{L_t\}_{t=0}^T$ from (A.30).

Next, given a path for the interest rate $\{r_t\}_{t=0}^T$, technological parameters, and aggregate labor, we can obtain the optimal capital-labor ratio from (A.27) and other production aggregates as well as the wage rate $\{\frac{K_t}{L_t}, K_t, Y_t, I_t, w_t\}_{t=0}^T$ follow from (A.25)-(A.29).

Given a government fiscal rule $\{B_{t+1}/Y_{t+1}, \rho_{jt}, \varphi^G, \varphi^\tau, \varphi^d, G_{-1}, \tau_{-1}, d_{-1}\}_{t=0}^T$, we obtain the path for the policies $\{G_t, \tau_t, d_t\}_{t=0}^T$ from (A.33)-(A.35) such that the government budget constraint (A.31) is satisfied for every t.

Then, we solve the household problem as follows. Given a guess for bequests received by type $\{B^r_t(\theta)\}_{t\geq 0,\theta}$, a path of prices $\{r_t, w_t\}_{t=0}^T$, government policy $\{\rho_{jt}, \tau_t, d_t\}_{t=0}^T$, demographic vari-

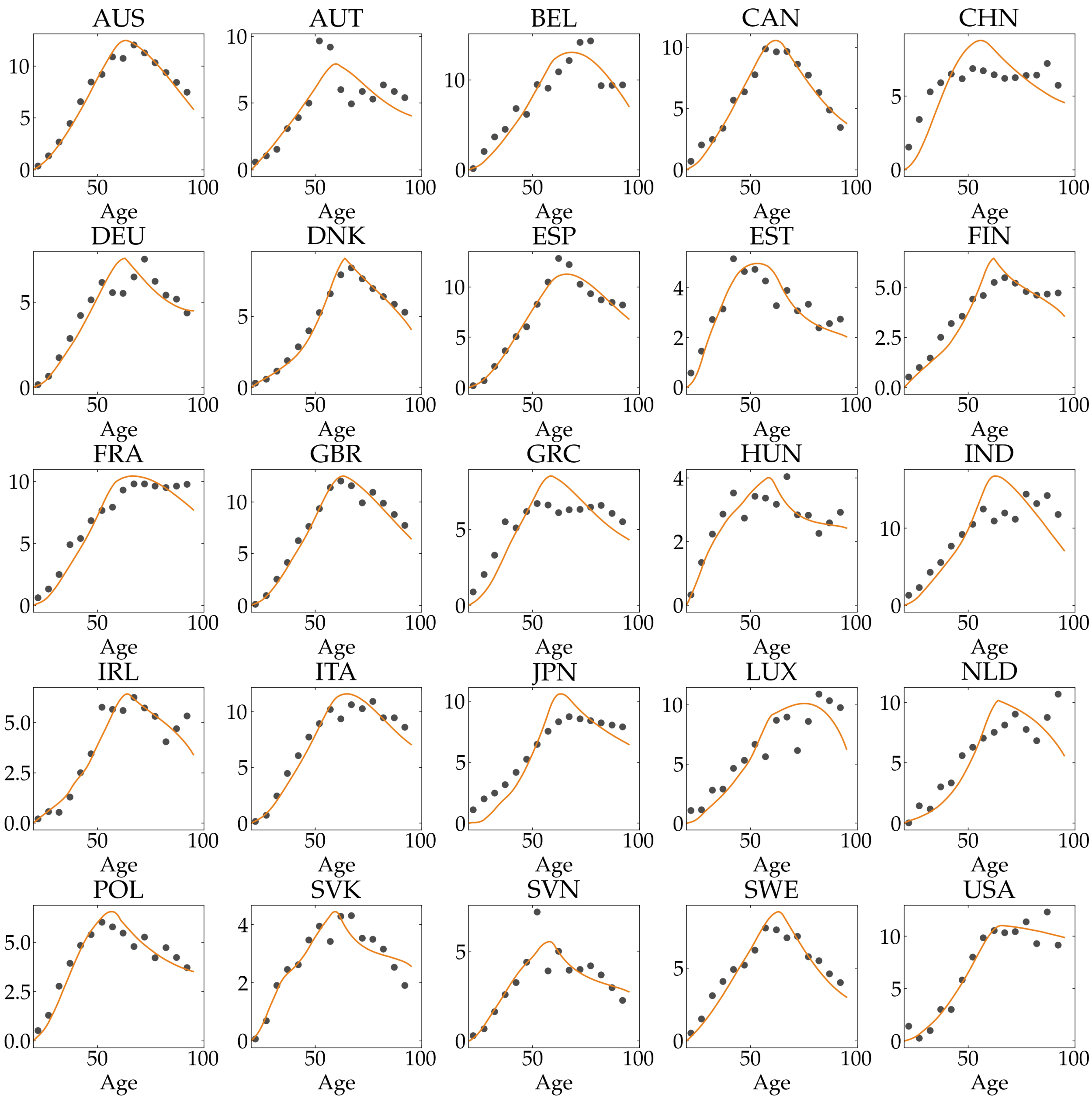


Figure A.7: Calibration outcomes: wealth

Notes: This figure presents the empirical age-wealth profiles (gray dots) and the calibrated model age-wealth profiles in the baseline calibration (orange line) for the 25 countries we consider.

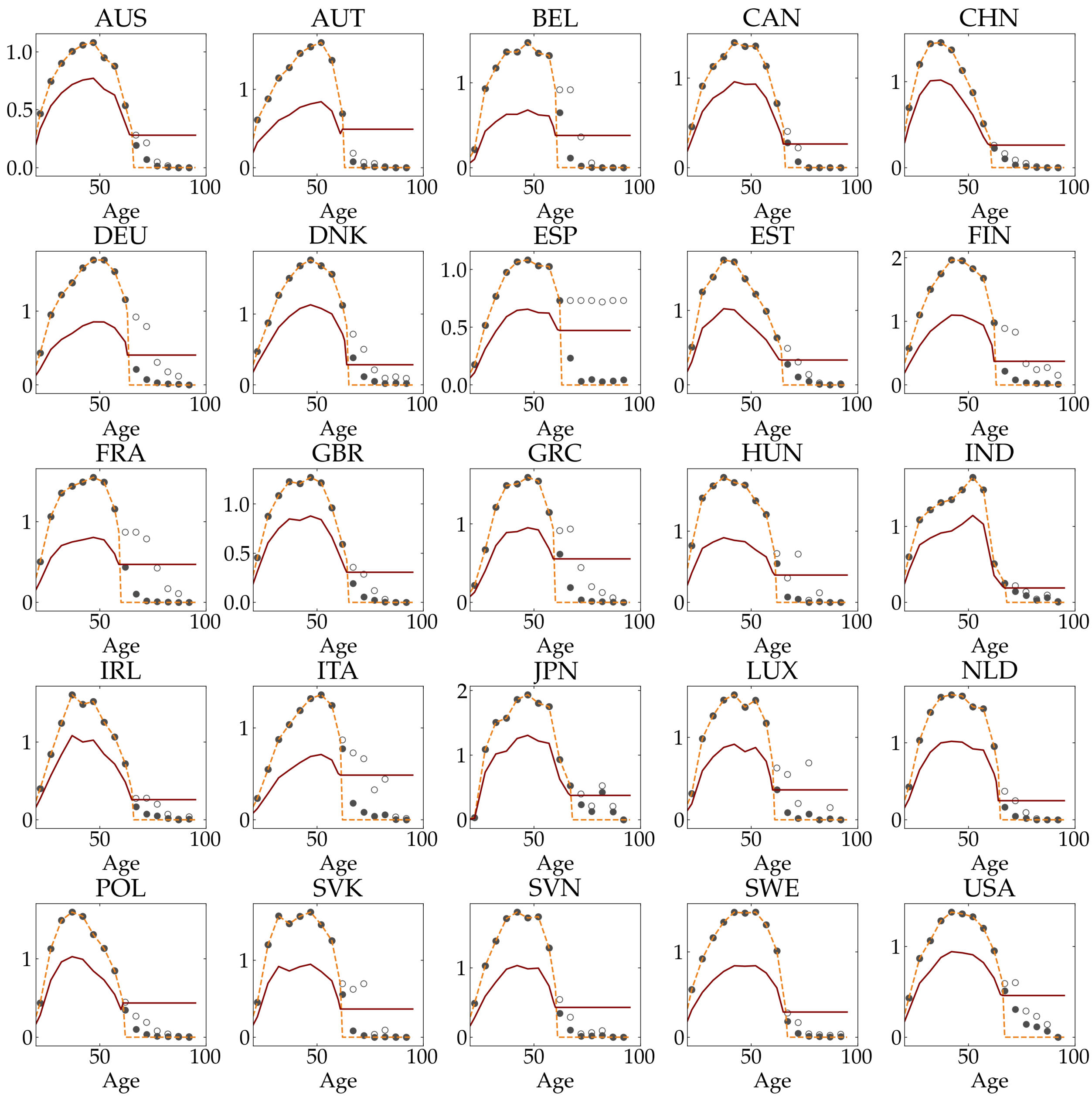


Figure A.8: Calibration outcomes: labor income

Notes: This figure presents the empirical age-labor supply profile from LIS used in section 2 (black dots), as well as the model gross age-labor supply profile (dashed orange line) and the net-of-taxes profile (red line).

Table A.4: World economy calibration

	$\Delta^{comp,c}$		Components of wealth			Government policy	
Country	Model	Data	$\frac{W^c}{Y^c}$	$\frac{B^c}{Y^c}$	$\frac{NFA^c}{Y^c}$	τ^c	$\frac{Ben^c}{Y^c}$
AUS	1.72	1.68	5.09	0.40	-0.46	0.29	0.04
AUT	1.14	1.07	3.90	0.83	0.12	0.47	0.11
BEL	1.91	1.85	5.74	1.06	0.65	0.54	0.09
CAN	1.07	1.04	4.63	0.92	0.20	0.31	0.04
CHN	2.50	2.37	4.20	0.44	0.25	0.30	0.04
DEU	0.82	0.79	3.64	0.69	0.58	0.50	0.10
DNK	0.75	0.72	3.42	0.37	0.46	0.36	0.06
ESP	2.77	2.41	5.33	0.99	-0.74	0.39	0.10
EST	0.64	0.63	2.64	0.09	-0.33	0.39	0.07
FIN	0.65	0.66	2.78	0.63	0.16	0.44	0.09
FRA	1.72	1.67	4.85	0.98	-0.05	0.48	0.13
GBR	1.64	1.60	5.35	0.88	0.08	0.31	0.06
GRC	1.56	1.38	4.25	1.81	-1.25	0.40	0.16
HUN	0.49	0.48	2.19	0.76	-0.54	0.48	0.09
IND	3.75	3.12	4.16	0.68	-0.08	0.30	0.01
IRL	1.49	1.46	2.32	0.74	-1.65	0.33	0.03
ITA	2.33	2.06	5.83	1.31	-0.02	0.48	0.13
JPN	1.32	1.19	4.85	2.36	0.66	0.32	0.09
LUX	2.05	1.86	3.92	0.21	0.64	0.40	0.07
NLD	1.57	1.54	3.92	0.62	0.70	0.37	0.05
POL	1.38	1.35	3.31	0.54	-0.52	0.36	0.10
SVK	0.85	0.84	2.17	0.52	-0.59	0.42	0.07
SVN	0.70	0.69	2.82	0.79	-0.21	0.43	0.11
SWE	0.67	0.70	3.81	0.42	0.08	0.43	0.06
USA	1.62	1.47	4.38	1.07	-0.36	0.32	0.06

Table A.5: World economy calibration

Country	$\bar{\beta}^c$	ξ^c	Y^c	ν^c	α^c	G^c/Y^c
AUS	0.984	0.00022	118.269	1.681	0.500	9.9%
AUT	0.996	-0.00012	118.269	1.681	0.287	22.0%
BEL	0.983	0.00065	118.269	1.681	0.391	22.2%
CAN	1.001	-0.00017	118.269	1.681	0.341	15.5%
CHN	1.024	-0.00003	118.269	1.681	0.341	15.1%
DEU	1.006	-0.00037	118.269	1.681	0.230	27.6%
DNK	1.161	0.00239	118.269	1.681	0.252	20.4%
ESP	0.939	-0.00044	118.269	1.681	0.494	7.6%
EST	1.177	0.00024	118.269	1.681	0.280	21.0%
FIN	1.195	0.00255	118.269	1.681	0.193	25.4%
FRA	1.001	0.00040	118.269	1.681	0.380	15.6%
GBR	1.000	0.00029	118.269	1.681	0.426	10.7%
GRC	1.015	0.00024	118.269	1.681	0.359	6.0%
HUN	1.178	0.00116	118.269	1.681	0.191	28.1%
IND	0.997	0.00041	118.269	1.681	0.347	18.5%
IRL	1.199	0.00284	118.269	1.681	0.314	18.6%
ITA	0.930	-0.00071	118.269	1.681	0.441	11.1%
JPN	1.089	0.00098	118.269	1.681	0.177	12.7%
LUX	1.195	0.00341	118.269	1.681	0.299	20.8%
NLD	1.144	0.00248	118.269	1.681	0.253	21.9%
POL	1.055	0.00057	118.269	1.681	0.319	13.6%
SVK	1.233	0.00199	118.269	1.681	0.218	24.4%
SVN	1.171	0.00076	118.269	1.681	0.219	20.9%
SWE	1.010	-0.00003	118.269	1.681	0.322	23.0%
USA	1.044	0.00063	118.269	1.681	0.356	12.5%

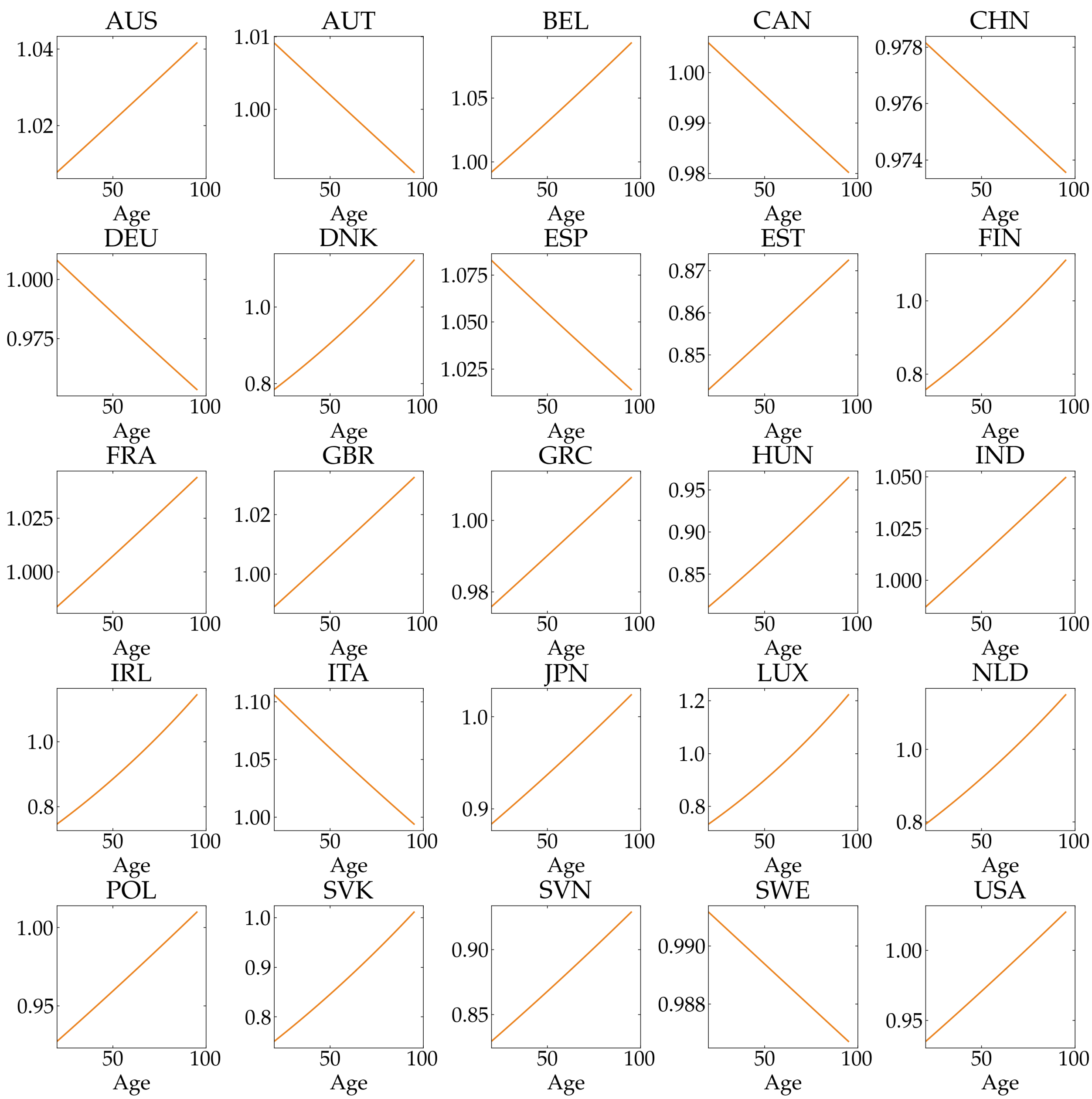


Figure A.9: Age-dependent subjective discount factor β_{j+1}/β_j

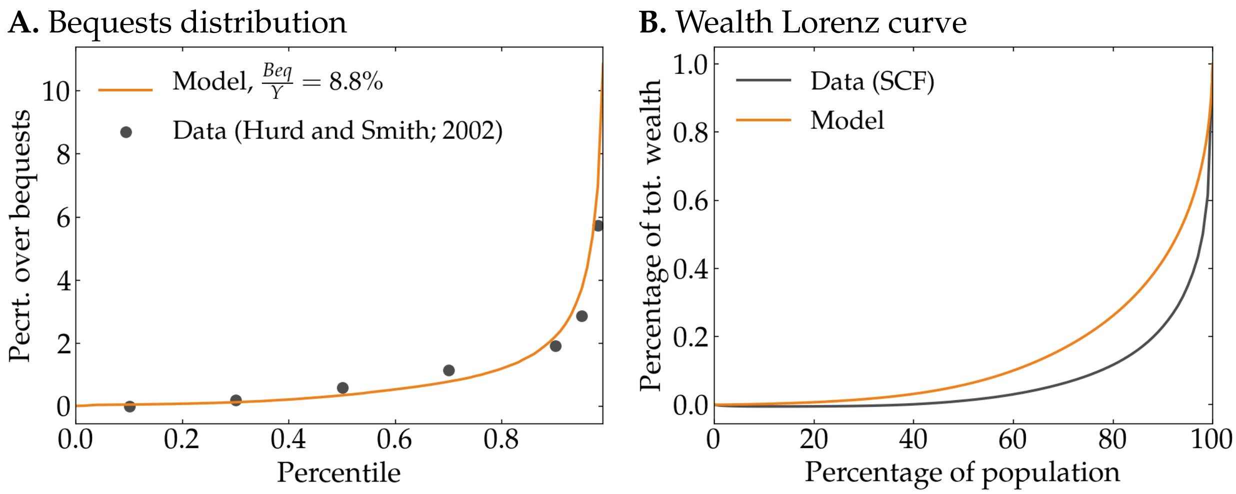


Figure A.10: Distribution of bequests and wealth Lorenz curve

ables $\{n_t, \pi_{j,t}, \phi_{j,t}, nc_{jt}\}_{t\geq 0, 0\leq j\leq T}$, we solve the household problem (A.20) in two steps. First, we use Carroll (2006)'s Endogenous Grid Point Method (EGM) to determine the decision functions $\{c_{jt}(\theta, \epsilon, \mathfrak{a})\}_{t\geq 0, 0\leq j\leq T}$ and $\{a_{j+1,t+1}(\theta, \epsilon, \mathfrak{a})\}_{t\geq 0, 0\leq j\leq T}$. Second, we obtain the distributions following Young (2010). We start from an initial distribution, which we take from the 2016 steady-state, and iterate forward using the asset decision function and the law of motion of the state (θ, ϵ). We then compute aggregates following (A.24).

To solve for the world economy equilibrium, we use a Newton-based method to ensure that bequests received equals bequests given by type θ and that the asset market clearing condition (A.39) is satisfied. We iterate on a 285×1 path for the interest rate by year $\{r_t\}_t$, and a $285 \times 25 \times 3$ path for bequest by year, country and type $\{B^{r,c}(\theta)\}_{t,c,\theta}$ until convergence.

To solve for the small open economy, we hold fixed the path of the interest rate, i.e. $r_t = r_0, \forall t > 0$.

Details on Table 4. Below, we provide details on the results in table 4, starting with the construction of each column, and then the details on the various experiments. The description of the columns applies to the full model analyses; for the pure compositional analysis, some columns have a slightly different interpretation, which is clarified when we discuss this experiment. For all columns, the changes refer to differences between 2016 and 2100. In the left panel, Δr is the change in the rate of return, $\overline{\Delta \log \frac{\bar{W}}{Y}} \equiv \sum_c \omega^c \Delta_{2100} \log\left(\frac{W^c}{Y^c}\right)$ is the average change in the wealth-to-output ratio, weighted by initial shares of wealth.

In the right panel, $\bar{\Delta}^{comp} \equiv \sum_c \omega^c \Delta^{c,comp}_{2100}$ is the average compositional effect between 2016 and 2100, weighted by initial GDP. The term $\bar{\Delta}^{soe} \equiv \sum_c \omega^c \Delta^{c,soe}_{2100}$ is the equivalent average for the small open economy effect. For each country c, $\Delta^{c,soe}$ is defined as the change in $\frac{W^c}{Y^c}$ between 2016 and 2100 in a small open economy equilibrium with a fixed interest rate r_{2016}.

The asset supply and demand sensitivities$\bar{\epsilon}^d = \sum_c \omega^c \epsilon^{c,d}$ and $\bar{\epsilon}^s = \sum^c \omega_c \epsilon^{c,s}$ are the averages of the country sensitivites weighted by initial GDP levels. For each country c, the asset demand sensitivity $\epsilon^{c,d}$ is defined as the semi-elasticity of the steady-state $\frac{W^c}{Y^c}$ with respect to the steady state interest rate r.[54] The asset supply sensitivites are given by $\epsilon^{c,s} = \frac{1}{W^c/Y^c}\frac{\eta}{r+\delta}\frac{K^c}{Y^c}$.

[54] In practice, we calibrate a steady-state to 2100 demographics, and perturb r_{2016} and resolve for a new

The list below describes the pure compositional analysis and the various model experiments. All model experiments feature a retirement age increased by 1 month per year for the first 60 year of the simulation, and all but the one labelled âĂIJstationary distributionâĂİ starts from the steady-state equilibrium calibration.

- **Pure compositional effect.** This row reproduces the exercise in section 3. That is, all changes in r, wealth, and NFAs are defined using proposition **??** and 3 given the initial GDP weights ω^c, the compositional effects $\Delta^{comp,c}$, and the set of sensitivities $\epsilon^{c,d}$ and $\epsilon^{c,s}$. The supply sensitivities are given by $\epsilon^{c,s} = \frac{1}{W^c/Y^c}\frac{\eta}{r+\delta}\frac{K^c}{Y^c}$, where $\frac{K^c}{Y^c}$ is the calibrated capital stock from the steady-state calibration. The demand sensitivities are defined using the expression in proposition 4, using the same construction method as in section 3, but using the calibrated profiles of assets and income to back out the consumption profile and calculate the relevant moments of the asset and consumption profiles.

- **Preferred specification.** The fiscal rule places equal weight on consumption, taxes, and retirement benefits.

- **Constant bequests.** The process $\frac{b^{jt}(\theta)}{w_t}$ of bequests received normalized by wages is kept constant over time. This removes a source of non-compositional increases in asset holdings which comes from an older population implying that people receive more bequests over time. To make a constant sequence of bequests consistent with equilibrium, we assume that it is implemented with an age-type specific lump sum tax/transfer that keeps bequests over wages constant at their 2016 level. To prevent this tax from having second order effects on individual behavior through the government budget constraint, we assume that it is neutralized by variations in government consumption.

- **Constant mortality.** The subjective mortality risk of individuals is kept fixed at their 2016 values, while the population evolution still follows the objective mortality risks.

- **Constant taxes and transfers.** The fiscal rule places all weight on adjustments in government consumption, so that taxes and benefits are constant over time.

- **Constant retirement age.** The retirement age is kept fixed at its 2016 level.

- **No income risk.** The idiosyncratic income risk is switched off and the model is recalibrated.

- **Annuities.** Households get access to annuities and the bequest preference is set to zero: $\Upsilon = 0$.

- **Fiscal rules.** The full adjustment weight is placed on either G, d, or τ.

E Appendix to Section 5

The exercise in table 5 is conducted using replication code for the two papers. To obtain the equilibrium interest rate change Δr^{GE} for GJLS, we run the same experiment as in the paper. For EMR, we consider an exercise which fixes markups, the growth rate of TFP, the debt-to-GDP ratio, and the relative price of capital. This exercise isolates the effect of demographics from the multiple forces considered in the paper.

---

stationary equilibrium, using the resulting perturbation to $\frac{W^c}{Y^c}$ to calculate the derivative.

For the components of the first order approximation, we calculate Δ^{comp} using the model-implied profiles of wealth and labor income at 1970 and the subsequent changes in the age distribution. The small-open economy effect Δ^{soe} is calculated by rerunning the experiment in both papers using a fixed interest rate. The semi-elasticities are calculated based on a 1970 steady state.[55]

F Appendix to Section 6

We first prove the results in the main text. Defining savings for an individual of age j in state (z^j, a_{jt}) at time t as

$$s_{jt} \equiv r a_{jt} + w_t \left((1-\tau)\ell(z_j) + tr(z^j) \right) - c_{jt}$$

and using the budget constraint (1), we see that aggregate savings for agents of age j is given by

$$s_{jt} = \mathbb{E} s_{jt} = \phi_j a_{j+1,t+1} - a_{jt} \tag{A.53}$$

Next, since lemma 1 implies $a_{jt} = a_j(r) Z_t$, we have

$$s_{jt} = \left(\phi_j (1+\gamma) a_{j+1} - a_j(r) \right) Z_t = s_j(r) Z_t$$

Hence, defining aggregate savings as

$$S_t \equiv \sum N_{jt} s_{jt} \tag{A.54}$$

we have that

$$\frac{S_t}{N_t} = \sum \pi_{jt} s_{jt} = \sum \pi_{jt} \underbrace{s_j(r) Z_0}_{s_{j0}} (1+\gamma)^t = \sum \pi_{jt} s_{j0} (1+\gamma)^t$$

Taking the ratio of this expression to equation (9), we obtain the equivalent of Proposition 1,

$$\frac{S_t}{Y_t} = \frac{F_L(k(r), 1)}{F(k(r), 1)} \cdot \frac{\sum \pi_{jt} s_{j0}}{\sum \pi_{jt} h_{j0}} \tag{A.55}$$

which delivers equation (30), from which equation (31) follows immediately.

Next, combining (A.53) and (A.54) and the population dynamics equation $N_{j+1t+1} = \phi_j N_{jt}$, we have

$$S_t \equiv \sum N_{jt} s_{jt} = \sum N_{jt} \phi_j a_{j+1,t+1} - \sum N_{jt} a_{jt} = \sum N_{j+1,t+1} a_{j+1,t+1} - \sum N_{jt} a_{jt} = W_{t+1} - W_t$$

where the last line uses the initial and terminal condition on wealth by age. Hence, the aggregate savings rate is:

$$\frac{S_t}{Y_t} = \frac{W_{t+1} - W_t}{Y_t} = \frac{Y_{t+1}}{Y_t} \frac{W_{t+1}}{Y_{t+1}} - \frac{W_t}{Y_t} = (1 + g_{t+1}) \frac{W_{t+1}}{Y_{t+1}} - \frac{W_t}{Y_t}$$

where g_t is the growth rate of aggregate GDP, the sum of productivity growth, population growth

[55] For EMR, the paper calibrates a 1970 steady-state, which we use. For GJLS, the paper starts the simulation at 1900; to obtain a 1970 steady-state, we rerun their steady-state routine using 1970 demographics.

and changing composition,

$$1+g_{t+1} \equiv \frac{Y_{t+1}}{Y_t} = (1+\gamma)\frac{N_{t+1}}{N_t}\frac{\sum_j \pi_{jt+1}h_{j0}}{\sum_j \pi_{jt}h_{j0}} = (1+\gamma)(1+n_{t+1})\frac{\sum_j \pi_{jt+1}h_{j0}}{\sum_j \pi_{jt}h_{j0}}$$

In steady state, therefore, we have

$$\frac{S}{Y} = g\frac{W}{Y}$$

where $1+g=(1+\gamma)(1+n)$. This is equation (32).

Finally, towards our implementation, we show that S_t/Y_t can be calculated from the cross-sectional profiles of assets a_{jt} and demographic projections alone. We first show that $\Delta_t^{comp,s}$ in equation (31) can be calculated from cross-sectional age profiles of assets $a_{j,0}$. Indeed, we have, starting from $S_t = W_{t+1} - W_t$, we have

$$\begin{aligned}
\frac{S_t}{N_t(1+\gamma)^t} &= \frac{W_{t+1}}{N_t(1+\gamma)^t} - \sum \pi_{jt}a_{j0} \\
&= (1+n_{t+1})(1+\gamma)\sum \pi_{jt+1}a_{j0} - \sum \pi_{jt}a_{j0} \\
&= ((1+n_{t+1})(1+\gamma)-1)\sum \pi_{jt}a_{j0} + (1+n_{t+1})(1+\gamma)\sum(\pi_{jt+1}-\pi_{jt})a_{j0} \\
&= g_{t+1}^{ZN}\sum \pi_{jt}a_{j0} + \left(1+g_{t+1}^{ZN}\right)\sum(\Delta\pi_{jt+1})a_{j0}
\end{aligned}$$

where we have defined $1+g_{t+1}^{ZN} \equiv (1+n_{t+1})(1+\gamma)$. Taking the ration of this expression to equation (9), we have the following expression for the aggregate savings rate:

$$\frac{S_t}{Y_t} = \frac{F_L(k(r),1)}{F(k(r),1)}\left(\frac{g_{t+1}^{ZN}\sum \pi_{jt}a_{j0} + \left(1+g_{t+1}^{ZN}\right)\sum(\Delta\pi_{jt+1})a_{j0}}{\sum \pi_{jt}h_{j0}}\right) \tag{A.56}$$

which is an alternative to equation (A.55).

In principle, to project savings rates from demographic composition, we could equally well implement equation (A.55) or equation (A.56). Summers and Carroll (1987), Auerbach and Kotlikoff (1990), and Bosworth et al. (1991) follow the first route. We prefer to follow the second because it only requires only information that we have already used so far in the paper, and because the computation of age-specific savings rates is subject to a large amount of measurement error.

Figure A.11 displays this predicted change over time in the savings-to-GDP ratio S_t/Y_t from compositional effects, calculated by first computing $\Delta^{comp,s}$ from equation (A.56) and then reporting the projected change in the level $S_0/Y_0\left(e^{\Delta_t^{comp,s}}-1\right)$, as in Figure 2. Our main finding is that the savings rate is projected to *decline* in every country, reflecting the fact that older age groups tend to have lower savings rates than younger age groups. This is fully consistent with an increase in wealth-to-GDP ratios, since the population growth rate is also falling, as section 6 explains.

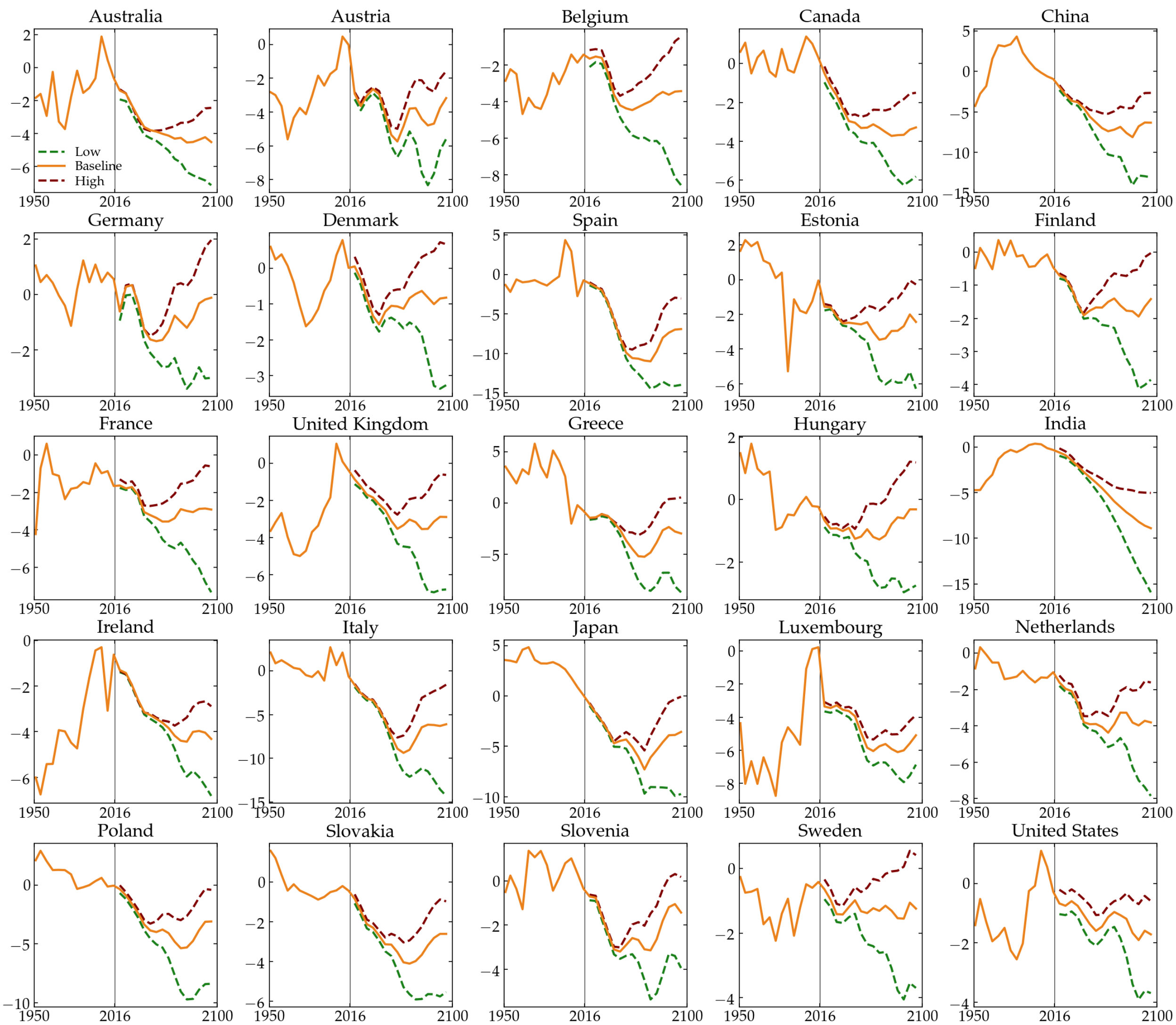


Figure A.11: Predicted change in savings-to-GDP from compositional effects

Notes: This figure depicts the evolution of the predicted change in the savings-to-GDP ratio from the compositional effect for $t = 1950$ to 2100, reported in percentage points. The base year is 2016 (vertical line). The solid orange line corresponds to the medium fertility scenario from the UN, the dashed green line to the low fertility scenario, and the dashed red line to the high fertility scenario.