Essays on Saving and Intergenerational Transfers over the Life-Cycle

Rory Michael McGee

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

of

University College London.

Department of Economics

University College London

December 9, 2019

I, Rory Michael M^{c}Gee, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.

No part of this thesis has been presented before to any university or college for submission as part of a higher degree. Chapters 2 and 4 are sole-authored research papers of my own. Chapter 3 was undertaken as joint work with Mariacristina De Nardi, Eric French and John Bailey Jones.

Abstract

This thesis contains three self-contained chapters studying the microeconomics of household savings decisions and intergenerational transfers.

Chapter 1 provides an introduction. Chapters 2 and 3 discuss savings behaviour during retirement. Chapter 4 discusses intergenerational transfers over the life-cycle.

Chapter 2 aims to understand the reasons households save, the assets they save in, and why they leave bequests. Combining panel data from the UK with variation in the tax costs of moving home due to exogenous policy changes helps to better identify the different reasons people might save. I combine this with a structural model of retirement savings which explicitly models housing, liquid assets, multiple sources of idiosyncratic risk, and house price dynamics. Aggregate house price changes have large impacts on retirees' retained wealth and a large proportion of a generation's good luck is passed on to their descendants. I take the current structure of disregard eligibility for programs that insure retirees as given and find that for every pound it costs the government, increasing the disregards for liquid assets provides more insurance value than increasing the disregards for houses.

Chapter 3 documents stark differences between the savings behaviour of single retirees and couples using US data. An estimated model of retirement savings that matches these differences shows precautionary savings against medical spending risk explains singles' slow deaccumulation of wealth. For couples, who accumulate assets, this role is smaller and the interaction between this risk, the desire to leave a bequest, and a desire to insure the surviving spouse explains their savings.

In Chapter 4, I explore the motives behind lifetime intergenerational transfers using a life-cycle model in which children provide informal care to their parents in exchange for monetary transfers. Calibrating this to US data provides a lower bound on the insurance value provided by these transfers, which I show is economically significant.

Finally, Chapter 5 concludes.

Impact statement

There are a number of ways that the contents of this thesis could be put to beneficial use for the common good.

By documenting the most important reasons that older households retain wealth well at older ages, Chapters 2 and 3 should be of interest to those concerned with the wellbeing of older households and the consequence of policy reforms. Studying which features of their economic environment are important should be of interest to academics economists, government civil servants, policy makers more broadly, and academic researchers beyond economics.

In particular, Chapter 2 decomposes which features of government policies which insure retirees against the risk of catastrophic long term care costs provide the largest benefit. While principally focussed on the UK, the policy environment shares a number of features with the US policy landscape. These findings should be of interest to anyone understanding the costs and benefits of different policy solutions designed to address an ageing population.

Finally, Chapter 4 explores how families might provide each other with sources of insurance against different types of risk. Informal care is often cited as a potential solution to the increasing needs of the elderly. The results of this chapter should help illuminate some of the potential implications of pursuing these policies.

As I write, these policies are part of an active debate surrounding the future of social care funding and health insurance in both the UK and the US. It is my hope that I will be able to both publish this research in widely read academic outlets and continue to contribute to the ongoing policy debate.

Acknowledgements

This PhD was made possible by a great deal of support and encouragement. I would like to thank my supervisors Professors Mariacristina De Nardi and Eric French. Without your guidance, advice, and patience I do not think I would have finished this thesis. You have guided me through all the intellectual contributions in this thesis and have maintained faith in me throughout the programme.

I would also like to extend special thanks to Richard Blundell, Cormac O'Dea, and John Bailey Jones. From each of you I have learnt how to be a better economist, as a researcher and as a representative of the profession. It has been a privilege to learn from each of you.

I have spent the majority of my PhD as a research scholar at the Institute for Fiscal Studies. I am indebted to the IFS for providing me with the funding and facilities to continue my studies, as well as for providing a stimulating, supportive and collegial research environment. They have taught me a great deal about policy relevant research and communicating important academic findings to a wider audience.

They say you can't choose your family and this is only slightly less true in academia. In Peter Spittal and Gonzalo Paz Pardo I was lucky to find two academic siblings who were always willing to share their time, feedback, code, worries and encouragement. They both always find new insights to offer, even after they have seen endless versions of the same paper.

Many other friends and colleagues have also at various times gone above and beyond the call of duty to inspire and challenge me, patiently explain concepts and methods to me, and most importantly, provide me with support and encouragement at the times I most needed it. This category includes far too many individuals to list individually. You know who you are. I will be forever grateful to every one of you.

Finally, I would like to thank my family. My long-suffering parents and siblings have provided me with years of support and encouragement. All of you have helped me to get to where I am today. To Phoebe, I owe you the largest thanks. You have made each day of this process feel a little easier and this has been more valuable than you could imagine.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

How and why do households save? In which assets and for what reasons? How does this change when their economic environment changes? To what extent do family links provide insurance or contribute to inequality? How does the design of the tax and transfer system change the answers to these questions?

This thesis examines questions such as these in three self-contained chapters which study the microeconomics of household savings decisions and intergenerational transfers over the life-cycle. They do this by combining economic theory, rigorous microeconometric methods, and a detailed approach to modelling the risks and decisions households face over their life-cycle.

The first substantive chapter, Chapter 2, begins from the idea that retirees born in different years experience the same house price shocks at different ages. These run-ups or falls in prices can have a large effect on their retirement savings, the wealth they leave to future generations, and how they value different asset-tested government programs that provide insurance. Many of these programs determine eligibility by separately testing housing and other assets - creating important interactions between changes in house prices and the insurance these programs provide.

Answering these questions requires accurately quantifying precautionary savings against long term care expenses, longevity risk, and changes in family size as well as the desire to leave a bequest and how retirees choose between housing and other assets. This is a considerable empirical challenge because the same savings can be used to self-insure against each of these risks or to leave a bequest. To identify the different reasons people might save, I combine panel data from the UK, including self-reported bequest intentions, with variation in the tax costs of moving home due to exogenous policy changes and to house price shocks that push the value of a house over thresholds in the system. I use state of the art econometric tools to identify household preference heterogeneity and combine this with a rich structural model of retirement savings which explicitly models housing and other more liquid assets, lifespan and long term care expense uncertainty, and house price dynamics.

With this framework I show that aggregate house price changes, like those in the UK over recent decades, have large impacts on retirees' retained wealth that last for many years after the price appreciation stops or slows down. The effect of these house price run-ups is a key part of explaining the "retirement saving puzzle". I compute the marginal propensity to consume and bequeath out of housing wealth and income, finding that a large proportion of a generation's good luck is passed on to future generations.

Taking the current structure of disregard eligibility for programs like Medicaid that insure retirees I find that for every pound it costs the government increasing the disregards for liquid assets provides far more value than increasing the disregards for houses. These findings have important implications for why people save and the design of social insurance.

In Chapter 3, we document the stark differences between the savings behaviour of single retirees and those who are still in couples. Couples tend to accumulate assets while they are both still alive, whereas single households deaccumulate their wealth. For a couple entering retirement, the wife usually outlives her husband and there is a large risk of entering a nursing home. This chapter seeks to understand how these life transitions affect the savings decisions of couples and singles? When the first member of a household dies medical spending rises, but also continues to be higher for the surviving single person. At the same time their pension income falls. To understand these differences, we build a model of retired couples and singles facing uncertain longevity and medical expenses in which couples and singles can have different bequest motives. Both might care about heirs, but couples might also care about their surviving spouse.

We use the Aging and Health Dynamics Among the Oldest Old data and the method of simulated moments to estimate our model and disentangle the importance of these life transitions and different saving motives for couples and singles. For singles, we find that precautionary savings against medical spending risk explains the slow deaccumulation of their wealth. For couples, who are on average much wealthier, the role of medical expenses is much smaller and the interaction between this risk, the desire to leave a bequest, and a desire to insure the surviving spouse together explain almost all of their savings.

The third substantive chapter, Chapter 4, focusses on both the retirement period and earlier life-cycle decisions. Specifically, I attempt to understand how the extended family provides insurance and how valuable is this insurance? Many households receive bequests when their parents die, but there also sizeable inter vivos transfers made between parents and their adult children while both are still alive. Understanding the motivations that generate these transfers is important to understand how they interact with public policy and other life cycle decisions.

I use the Panel Study of Income Dynamics to show that transfers of time are more frequent than monetary transfers. In order to directly compare their value with monetary transfers I use both the individual's market wage and the cost of market substitutes which I assume are supplied at the minimum wage. Under either of these measures, time gifts are in fact slightly larger than monetary gifts.

I focus on the exchange motive for inter vivos transfers where parents receive valuable time gifts in exchange for monetary gifts and build a dynamic model of this intergenerational interaction. The results suggest that a single motive for inter vivos transfers cannot capture the full distribution of inter vivos transfers, but can explain a substantial amount. I then use the model to quantify a lower bound for the insurance value of these transfers in the absence of direct altruism between parents and children.

Finally, Chapter 5 concludes. Each substantive chapter contains its own brief conclusion summarizing the results for this specific piece of research and the contribution to this area of research. Chapter 5 briefly recaps these chapters and discusses potential avenues for future research building on the work contained in this thesis.

Chapter 2

Old Age Savings and House Price Shocks

2.1 Introduction

Across the OECD the over 65 population has grown by over 40% in the last 25 years and is projected to continue to rise. These older households hold lots of wealth, even at advanced ages, primarily in the form of houses, and often leave bequests. Houses, however, are complex assets. First, they are risky. Some cohorts experience substantial house value increases, while others drops and stagnation. Second, houses provide a consumption flow, are illiquid, and are often exempt from tests on assets that determine eligibility for important government insurance programs.

This paper aims at understanding the role of housing in retirement. I estimate a dynamic life-cycle model of consumption, savings, and portfolio choice for older households using panel data on household choices, beliefs, and the risks they face. To the best of my knowledge, this is the first paper that allows for extensive and intensive margins of housing adjustment, a rich set of risks in retirement, including aggregate risks, and heterogeneity in bequest motives. Estimating the model with rich micro data and exploiting quasi-experimental variation in tax regimes over time allows me to separately identify and quantify the demand for housing and its illiquidity separately from other savings motives. I use this rich framework to understand portfolio rebalancing in retirement and the effect of housing and changes to house prices on the dynamics of saving among the elderly. Furthermore, I use the estimated model to quantify the intergenerational transmission of house price shocks and evaluate the welfare impacts of the distinction between housing and liquid wealth in the means testing of government Long Term Care (LTC) insurance.

The estimated model that incorporates differences in assets and preference heterogeneity in bequest motives. Households choose between investing in housing, risk-free liquid wealth, or consuming today. Focusing on the role of preference heterogeneity and the characteristics that make these asset classes different requires that I accurately model the set of risks facing households. Each member of the household faces uncertainty over their health status, mortality, costs from long term care needs, and uncertainty over aggregate house price levels. These rich sources of risks also determine a household's desire for liquidity, consumption, and savings for a bequest over different horizons, and consequently their demand for different assets.

Separately identifying precautionary savings motives, the desire to remain in one's home, and a bequest motive when there are many reasons households may hold wealth, let alone with heterogeneous bequest motives, presents a considerable empirical challenge.[1] To do so, I employ a new strategy that combines exogenous variation in housing and estate taxes with data on wealth composition and a measure of household subjective expectations within a structural framework. Incorporating subjective probabilities of leaving an inheritance into the estimation provides an additional source of information about the heterogeneous preferences of individual households and the full path of future saving behaviour over longer horizons.

I study the financial behaviour of retirees in England during a period which saw numerous reforms to both estate and residential property transaction tax schedules as well as large changes in house prices. Transaction taxes create an implicit tax on home equity adjustment, or downsizing, and variation in the tax schedule over time generates quasi-experimental variation in the size of this implicit tax. I examine how these implicit taxes distort the home moving decisions of retirees. Using this plausibly exogenous variation, I estimate the decrease in household mobility and adjustments of home equity in response to increases in the transaction tax bur-

[1] See, for instance, De Nardi et al. (2010, 2016a); Ameriks et al. (2011, 2018); Lockwood (2018).

den. I use data covering the period between 2002 and 2014, simulating households through the various reforms to their budget constraint and realizations of aggregate house prices. This is key to identifying the different costs households face when adjusting their portfolio. I validate the model against the reduced form evidence and show that it reproduces household responses to changes in the financial incentives to adjust their housing stock - a key model mechanism. Retirees in the UK and the US have similar medical expenditure risk driven by LTC related needs, similar life expectancies, and are covered by similar public programs. Thus, my results are also useful to understand retirement savings and their implications in many countries.

The model matches two key facts that I document in the data. First, that primary residences constitute the majority of household wealth and that in a time of house price appreciations, the deaccumulation of housing wealth through downsizing is masked. Because housing is less liquid and historically has experienced long periods of large appreciation it is especially attractive for those wishing to leave as a bequest. Second, new evidence that different households report systematically different expectations about leaving bequests, outcomes which are likely to reflect differences in the amount these households plan to leave. This reinforces important evidence supporting that bequest motives exist and are heterogeneous across households (Laitner and Juster, 1996; Kopczuk and Lupton, 2007).

Armed with a carefully estimated model that matches these and other important facts, I quantify household responses to changes in different sources of retirement wealth and to what extent house price run-ups are shared with younger generations. I find that over a third of house price shocks at age 70 are passed on to future generations as a bequest while almost 70% of a liquid wealth shock is passed on as bequests (both of these results are larger than existing estimates such as those in Altonji and Villanueva, 2003). Increases in housing and liquid wealth have opposite effects on liquidity constraints. This changes the incentives to access wealth held in housing leading to very different downsizing behaviour and bequests.

My estimated model parameters indicate substantial heterogeneity in bequest motives. Around half the population have zero, or close to zero, estimated bequest motives, while the remaining population have positive and quantitatively important bequest motives. These differences in bequest preferences affect how households deaccumulate wealth in retirement and, in particular, how strongly they respond to changes in their portfolio.

I use my estimated model to evaluate the extent to which asset disregards (in housing and other assets) affect government provided insurance against the risk of large LTC expenses. I take the current structure of disregards and eligibility for these (Medicaid-like) benefits in the UK as given. I simulate retirees through a set of reforms that eliminate disregards for specific assets to isolate how the insurance provided by these programs is affected by the specific design of the asset testing. Comparing across these different scenarios, I find that for every pound it costs the government increasing the disregards for liquid assets provides more value than increasing the housing disregards. Asset tests that determine eligibility treat housing and other assets differently in the UK system, as well as Medicaid in the US, and is a feature of many tax and transfer systems. My results are informative about this common feature of asset-tested social insurance programs.

Related Literature This paper contributes to four important strands of the literature. Firstly, it contributes to an established literature exploring the so-called "retirement saving puzzle", that households do not deaccumulate their wealth during retirement, and quantifying various savings motive for the elderly. I make a significant contribution to a second, highly related, literature analysing the distribution of household bequest motives. There is a large literature on the role of housing wealth in savings decisions to which this paper is closely related. Finally, incorporating quasi-experimental variation and self reported subjective probabilities into the identification of a large structural model contributes to the nascent literature attempting to provide more robust and transparent identification.

I incorporate both the important precautionary and bequest savings channels from earlier work on the retirement savings puzzle, combining it with rich heterogeneity in assets choices and preferences. In estimating the different savings motives in retirement, I combine self reported probabilities of leaving an inheritance (a widely available survey instrument) with a rich asset structure and quasi-experimental variation. Ameriks et al. (2018) instead combine panel data on the liquid component of household portfolios with specially designed *strategic survey question* on bequests and long term care- in effect, stated choice. I exploit the self insurance information in the composition of household portfolios between liquid and illiquid assets. Inkmann and Michaelides (2012), De Nardi et al. (2016a), and Lockwood (2018) all estimate quantitatively important and prevalent bequest motives as a feature that rationalizes household under-utilization of insurance products (life insurance, Medicaid participation and long term care insurance respectively). The allocation of wealth across assets with different self insurance capacities uses similar variation in household precautionary incentives and provides a potential alternative explanation for this underutilization.[2]

I provide a link between the precautionary savings focused retirement savings literature and the literature exploring heterogeneity in household bequest motives by combining an estimation of heterogeneous bequest motives with a state of the art structural model. I estimate a latent distribution of household bequest motives while also allowing for a richer environment of empirically relevant risks in retirement and a more flexible approach to estimating the bequest motive. I allow both the overall strength of the bequest motive and the extent to which they are a luxury to vary across households in addition to the extensive margin of the linear bequest motive in Hurd (1989a) and Kopczuk and Lupton (2007). Put differently, I allow for variation even among households who do have a bequest motive. Ameriks et al. (2016) allow for variation in the strength of the bequest motive out of financial wealth estimated directly from variation in responses to strategic survey questions,

[2]Since Yaari (1965), several studies have explored the role of idiosyncratic risk in old age with Hurd (1989a) suggesting that mortality risk is the primary empirical driver of savings in retirement or, as in Palumbo (1999), that medical expenses faced by retired households are necessary to explain their limited deaccumulation. Studies in this tradition argue that risk averse households maintain wealth and exhibit slow deaccumulation because of high levels of precautionary savings and that bequests are accidental. De Nardi et al. (2010) examine the role of precautionary savings motives using a structural model and find that longevity and medical expenditure risk dominate for the majority of single households. While much of this literature focuses on the United States, evidence from Dutch (Alessie et al., 1999), Norwegian (Kvaerner, 2017), Swedish (Nakajima and Telyukova, 2018b) or other cross country comparisons (Blundell et al., 2016b) emphasises similar motives. See De Nardi et al. (2016b) for an extensive review of this literature

but without choice data and do not study the non-financial component of household portfolios.

To understand the role of housing wealth in retirement separately from other assets, I present new descriptive evidence on the lifetime frequency and size of housing adjustments by retirees. In the short run, households retain capital gains and their housing wealth tracks house price movements. Similarly to Fagereng et al. (2019), failing to distinguish between active and passive saving when asset prices change can substantially overstate household savings rates. Passive saving dominates in the short run, however, adjustments to their housing wealth are large and common over the entire retirement period. This paper explores how the housing wealth effect[3] interacts with different sources of idiosyncratic risks as well as its implications for future generations. I extend the housing decision faced by households by modelling adjustments to their housing stock on the intensive margin and capture the active rebalancing of the portfolios held by retirees. Nakajima and Telyukova (2018a) and Cocco and Lopes (2019), who both model retirees' decisions to remain in their own home, have the closest asset structure to this paper among studies focussed on wealth in retirement.[4]

Finally, this paper contributes to a literature focussed on the credible identification of structural models. Exploiting exogenous policy variation in estate taxation and housing transaction taxes in the estimation of bequest motives (building on Voena (2015) and Blundell et al. (2016a)) complements the instrumental variable approach proposed by Lee and Tan (2017). Additionally, the identification approach in this paper uses elicited self reported probabilities of leaving a bequest in the future as dependent variables and also to classify unobserved preference heterogeneity. van der Klaauw and Wolpin (2008) and van der Klaauw (2012) demonstrate the

[3]An increase in household expenditures in response to an increase in home values. An inexhaustive list of contributions explicitly exploring the size and heterogeneity of the expenditure response using micro data include Mian et al. (2013); Kaplan et al. (2016); Aladangady (2017); Berger et al. (2018); Guren et al. (2018).

[4]Earlier attempts to understand the effect of house price changes on the consumption and savings decisions of retirees includes Skinner (1993). A parallel literature in household finance, including Love (2010) and Hubener et al. (2016), finds that marital status and household demographics are important determinants of household portfolio allocations over the life cycle.

value of using non-choice data as outcome variables in identifying dynamic discrete choice structural models while Pantano and Zheng (2013) shows how they can be used to identify household level fixed effects. Hendren (2013) uses subjective probabilities to infer differences in household private information or unobserved risk types.

The remainder of this paper proceeds as follows. Section 2.2 describes the data used in this paper and presents descriptive results. The quasi-experimental tax variation is described in detail in Section 2.3. Section 2.4 describes the model and Section 2.5 detail the identification and estimation of the model. Estimation results are given in Section 2.6 and 2.7 discusses the implications of the results. Section 2.8 empirically evaluates means tested long term care benefits. Finally, Section 2.9 concludes.

2.2 Data & Key Facts

This section first discusses the dataset used in the paper, the English Longitudinal Study of Ageing (ELSA). Then, it outlines the key facts on the evolution of wealth in retirement and, in particular, housing wealth that this paper aims to understand.

2.2.1 Data

ELSA is a biennial longitudinal survey that contains a representative sample of the non-institutionalized English population aged 50 and over. ELSA is an ageing survey modelled on the US Health and Retirement Study (HRS). It collects detailed panel data on demographics, earnings, health, wealth levels and portfolios through a combination face to face interviews and supplementary questionnaires. ELSA begins in 2002/03 and I use data collected in the first 7 waves.

To construct the sample, I keep only households where the head is above the age of 65 (the state pension age for men) and who do not report large labour income (those in excess of pension credit levels, a means tested benefit which tops up household income for those out of work and eligible for state pensions) and abstract from labour decisions around retirement. Following De Nardi et al. (2018), I allow for household composition changes only at death and drop households when either

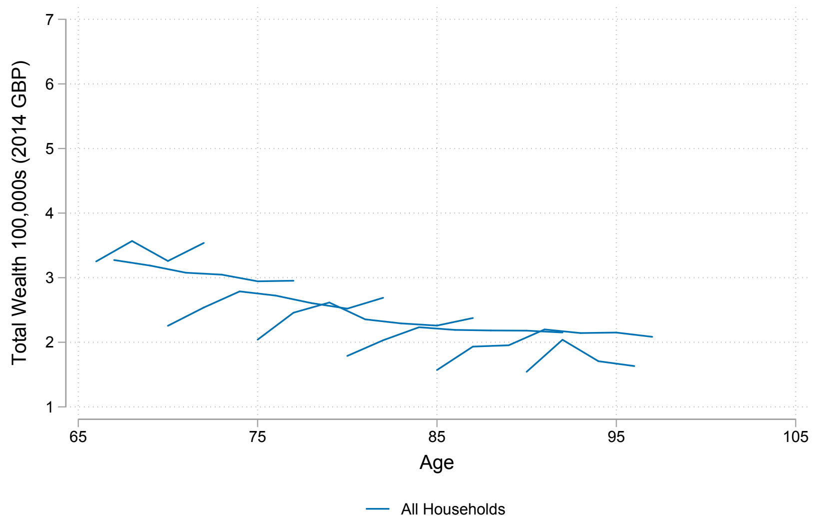


Figure 2.1: Mean Total Wealth by Cohort

a new individual enters or leaves the household before death - this drops all households who either divorce or remarry during the sample period. I make use of the original sample cohort and households that are included as new entrants because, even if attrition is differential, these new households are included to maintain the representative sample as the survey ages.

In the next part of this section, I show how the savings of retired households and the major components of their portfolio have evolved over time. I present statistics for within group means by cohort (and, as discussed below, stratified by additional data) where I top-code wealth moments at the within group 95th percentile and drop cells with fewer than 15 observations to mitigate the impact of outliers. A subset of the moments presented here also comprise the moments used when estimating the model (discussed in more detail in section 2.5).

2.2.2 Key Facts

Figure 2.1 plots the mean total wealth, or net worth, of households by age for several five-year birth cohorts in ELSA. Total wealth is the sum of housing wealth and liquid wealth. Housing wealth consists of the value of their primary residence. Following Berger et al. (2018) mortgage debt is included in liquid (or non-housing) wealth. Liquid wealth additionally includes savings and current accounts, bonds/gilts, premium bonds, shares, trusts, and other physical assets less credit card debt, private debt and any other outstanding loans or debts.[5]

Comparing across birth cohorts, mean total wealth appears to decline moderately with age. By the time that they reach their late 90s, the oldest birth cohort hold about 40% less wealth than the youngest birth cohort at age 70. However, within each birth cohort mean wealth remains at high levels for much of their retirement and exhibits little signs of deaccumulation For all but the youngest birth cohort, total wealth is higher at the end of the sample period than at the start. This lack of deaccumulation at the end of life is inconsistent with a basic life cycle model where households accumulate wealth during working life and draw down their wealth in retirement.

Similar patterns in total wealth are widely studied in the US (such as resuls in De Nardi et al., 2018, who document the savings profiles of elderly US couples and singles) and typically exhibit more deaccumulation. Medical costs, longevity risk, bequests, and housing decisions have been discerned as important factors explaining these asset holdings.

One striking feature of the UK data is the presence of time effects, which have been little studied in the context of the US.[6] The x-axis plots the average age within birth cohort - consequently, within cohort ageing is equivalent to plotting a time dimension. The steep growth in total wealth followed by a levelling out occurs for the same calendar years in each birth cohort. For the youngest two cohorts who age into the sample in later calendar years only the flat portion of the profile after the initial increase is observed. As I document below, the rapid rise and peak in household total wealth broadly follows the aggregate trend in house prices around the 2008 financial crisis.

As a cohort ages, it is increasingly comprised of rich people due to mortal-

[5] Specifically savings accounts include savings accounts with a bank or building society as well as TESSA, all forms of ISA, PEPs, National Savings Accounts and life insurance savings. Following Berger et al. (2018) mortgage debt, secondary residences and other properties are included in liquid (or non-housing) wealth. I drop retirees who directly own businesses.

[6] As noted in Schulhofer-Wohl (2018) many studies attempt to cleanse time effects from moments in the data

ity differences between the rich and poor.[7] To mitigate this composition effect and highlight cross sectional differences in the level and portfolio composition of wealth, I pursue three complementary approaches. First, I present results for total wealth grouped by permanent income quantiles. This controls for the lifetime income levels of the households. To calculate permanent income I follow the approach in De Nardi et al. (2018) and exploit the approximately monotonic relationship between lifetime resources and pension income in the UK. Each household is then ranked by their position in the permanent income distribution and the measure is invariant to household demographics (I describe this in more detail in Appendix A.1). I generate three permanent income groups: the top 25% of households, the second quartile, and the bottom 50% of households. I merge together the bottom two quartiles as conditional on their initial home ownership status the two groups are extremely similar. However, there are substantial differences in initial home ownership rates in the two bottom quartiles.

Second, to understand changes in their portfolio I present results for the different forms of wealth, housing and non-housing wealth, by PI, and by the initial home ownership status of households. This not only controls by income levels and housing tenure, but also shows how households save in different savings vehicles. Initial home ownership is defined in the first wave a household is sampled and this definition keeps the household composition constant in the analysis. Third, in Appendix A.2 I present results for a balanced panel consisting of those households who enter the sample in the first wave and survive until the final wave. This eliminates composition bias, but by stricter selection requirements leads to a selection bias. Nevertheless, the results in Appendix A.2 indicate that the stylized facts are not driven by changing household composition.

In figure 2.2, I reproduce the mean total wealth for birth cohorts also separated by their permanent income. Comparing figures 2.1 and 2.2 suggest that the means in figure 2.1 mask considerable heterogeneity, with a strong permanent income gradient. The top quartile of households have more than twice the assets of the bottom

[7] Attanasio and Emmerson (2003) document this composition difference in the UK

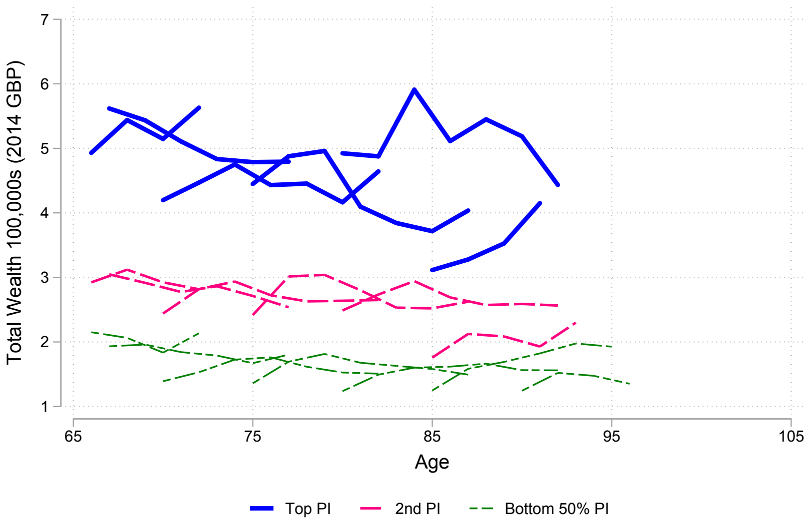


Figure 2.2: Mean Total Wealth by Cohort

half of the distribution. While those in the 2nd quartile are around £200,000 poorer on average than their top quartile counterparts, they are also around £100,000 richer than retirees in the bottom half of the lifetime resources distribution.[8] Although there are signs of wealth deaccumulation by households, it remains minimal. Furthermore, within permanent income group, the evidence of cross cohort deaccumulation (or pervasive cohort effects) is much smaller. In figure 2.2, mean total wealth rises by approximately £50,000 in the first three waves and gradually declines through the remainder of the sample. This initial increase occurs across PI groups for all birth cohorts in the sample in 2002/03, however, conditional on PI there is a larger decrease in wealth after the increase for those with higher PI levels. Consequently, between the start and the end of the sample, total wealth remains almost flat for all cohorts.[9] This limited deaccumulation of resources across and within cohorts suggests that even those who survive until older ages will bequeath much of their initial wealth held at age 70. Using comparable US data from the HRS, Poterba et al. (2017) similarly document a high degree of persistence in early

[8]Excluding the 1915 birth cohort, for whom the difference is in the region of £50,000. There is also considerable heterogeneity in the size of these gaps by Permanent Income with the gap exceeding £100,000 for a number of cohort age combinations

[9]This suggests that composition bias drives some of the flattening of savings profiles and the cross cohort gradient in figure 2.1. Appendix A.2 discusses composition bias in more detail.

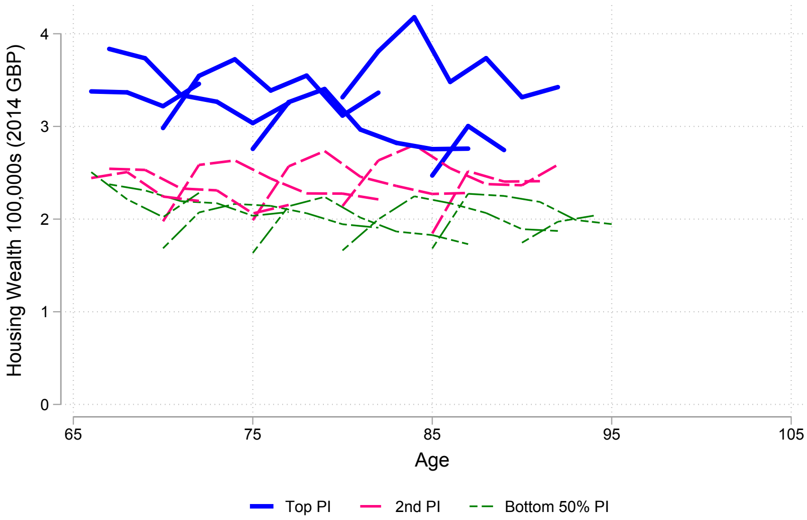
retirement wealth and wealth at death.

2.2.3 Different Forms of Wealth

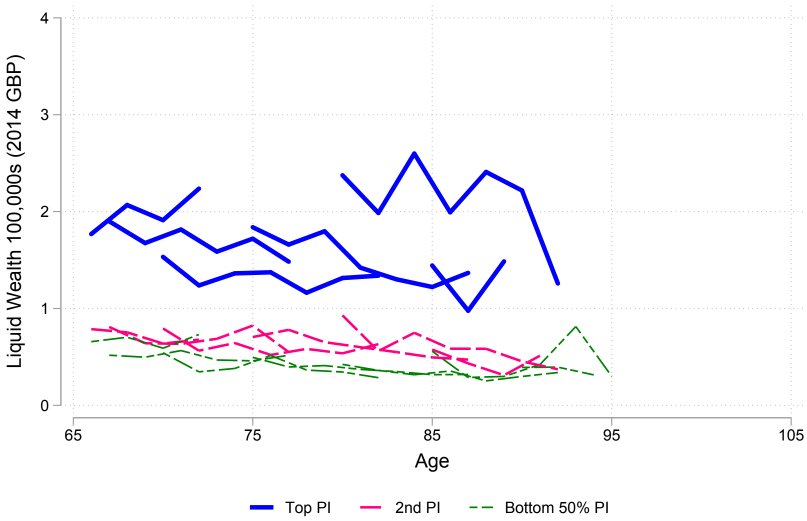
Figures 2.1 and 2.2 show that the savings of households in retirement differ by household lifetime resources. To understand the importance of household portfolios, I continue to document savings heterogeneity by a) the type of wealth owned by households and b) the home ownership status of households. In Figure 2.3 I plot the mean wealth of initial owners by birth cohort and permanent income.

The top panel shows the mean housing wealth of households (their primary residence) and the bottom panel shows the corresponding mean liquid wealth for the same groupings. For all cohorts and PI groups, housing wealth displays evidence of the aggregate trend in the total wealth profiles, However, liquid wealth does not exhibit the same degree of cyclicality. As with the total wealth profiles in figure 2.2, there is a permanent income gradient (here while also conditioning on initial home ownership status) for both housing and liquid wealth. Peak mean owner occupied housing wealth is above £350,000 for the top permanent income group while on average it is approximately £260,000 for the second quartile. The differences between PI groups are compressed when compared with total wealth (although the base is smaller) with the average gap between the second quartile and the bottom half of the distribution dropping to £40,000.

The gap between the liquid wealth of the top quartile and the second permanent income quartile is of similar magnitude to the absolute gap in housing wealth because housing wealth is a smaller proportion of the portfolio of richer households. However, the difference between the liquid wealth of the second quartile and the bottom 50% is much smaller. In levels, liquid wealth shows some evidence of deaccumulation which is concentrated at older ages. The largest change in the top two permanent income groups are £110,000 and £55,000, but for most cohorts the drop is below £25,000. For the largest reductions, this is equivalent to a 30-35% reduction in the liquid wealth stock (rising to 50% for the largest drop), but has a small effect on the total wealth base. In contrast, there is even less deaccumulation of housing wealth. At most, mean household housing wealth decreases by £60,000



(a) Mean Housing Wealth



(b) Mean Liquid Wealth

Figure 2.3: Wealth by Cohort & PI (Initial Owners)

(the top PI group and second youngest birth cohort who age into the sample during house price depreciation), but the majority of cohorts retain similar levels of housing wealth or even grow in real terms between the beginning and the end of the sample.

Returns or appreciation in asset prices driven by aggregate trends affect housing, but similar patterns are not present in liquid wealth. Furthermore, housing provides a consumption flow and functions as a store of wealth as well as being subject to large adjustment costs. Explicitly modelling these assets is important for understanding household savings and household demand for self insurance. However, Figure 2.3a is still insufficient to disentangle the active and passive saving in housing wealth. Later in this section I show direct evidence of housing transitions and the change in the value of housing wealth to provide evidence on the active saving component of deaccumulation.

Finally, I turn to initial renters and their liquid wealth. After retirement, initial renters tend to belong to lower PI percentiles. For this reason I pool all renters together. At the mean, initial renters are poorer than their home owning counterparts. They hold around 25% of the liquid wealth of the corresponding bottom PI groups. However, unlike homeowners who have on average £200,000 in housing wealth and £50,000 in liquid wealth, initial renters are cash and income poor. The liquid wealth for initial renters is approximately stable, in contrast with Nakajima and Telyukova (2018a) who show that the assets of elderly US households who transition from owner occupation to renting decline.

The key fact established in this section is that, even stratifying by wealth type and important financial characteristics, neither housing wealth or liquid wealth exhibit large declines. Furthermore, households retain capital gains in housing which exposes housing to aggregate fluctuations in house prices which affect different birth cohorts at different ages. Similar results using ELSA data are found in Blundell et al. (2016b) and Crawford (2018) who projects that the median household will spend down less than £10,000 of their financial wealth in retirement. In the next part of this section I further explore the housing transitions of older households.

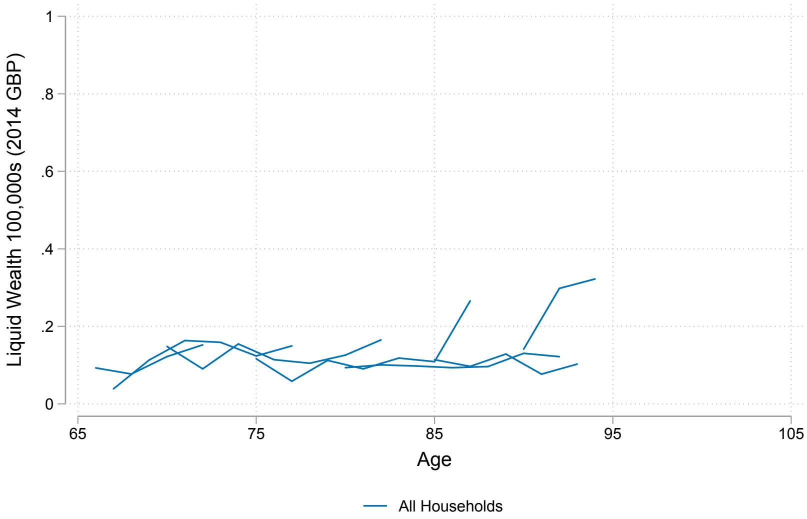


Figure 2.4: Mean Liquid Wealth by Cohort (Initial Renters)

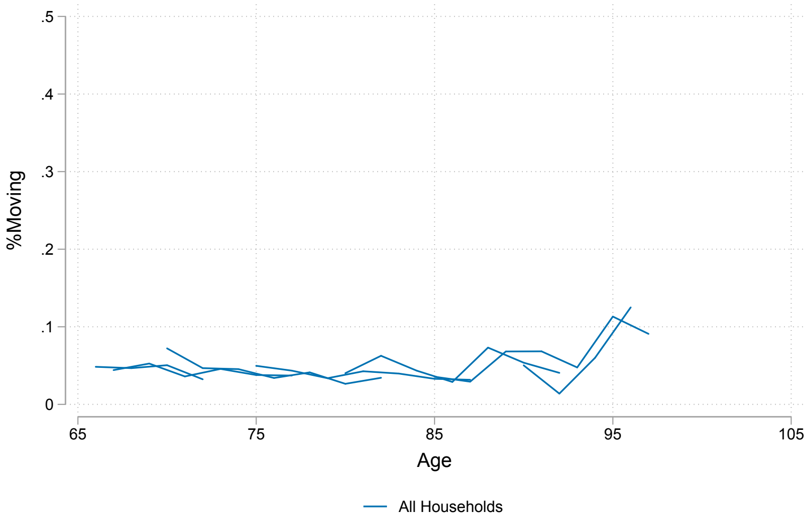


Figure 2.5: Moves in the last 2 years

2.2.4 The Role of Housing Transitions in Retirement

To understand the economic importance of household mobility and the active portfolio rebalancing of older households, I focus on the frequency and size of decisions to adjust housing wealth.[10]

Figure 2.5 shows the frequency with which initial homeowners move proper-

[10] An alternative exercise deflating housing wealth by house price change is presented in appendix A.3

	Downsizers		
	All	Remaining Owners	All Upsizers
Mean Housing Change[a]	-136	-104	55.4
Mean Relative Change[b]	0.48	0.72	1.34
Fraction of Movers (%)	76.8	50.9	23.2
N	219	145	66

Source: Author's own calculation from ELSA. [a] £1000s in 2014 prices, [b] Relative change is defined as the ratio of the new price to the old price at time of sale.

Table 2.1: Average Housing Wealth Change by Move Type

ties in two year periods (the frequency of the ELSA data). There is limited age and cohort variation with an average of 4.5% of households moving each wave. Households over 65 move infrequently in any given wave and on average only once. There is little evidence, in contrast with Angelini et al. (2014), that retired households front-load their movement decisions earlier in retirement and almost 50% of households have moved by age 90.

The home moving decision is one of the key financial decisions households make during retirement. Table 2.1 provides statistics on the mean level and relative change in housing wealth for three different categories: all downsizers, those who downsize excluding household's who transition to renting, and those who upsize. Separating households by those who downsize and those who upsize controls for differences in the type of move. Within downsizers, who are over 75% of moves, I also separate out transitions to renters to control for the largest differences in the fraction withdrawn. Conditional on downsizing, the average household releases 52% of the current value of their house or over £135,000.

Among downsizers the largest relative changes are those who completely downsize, Downsizers who remain owner occupiers release over £100,000 of equity, or approximately 40% of the average wealth level, and 30% of their housing wealth. Finally, household's who upsize are the smallest group, but represent over

20% of move. In levels they make the smallest change to the mean level of their housing wealth. Nevertheless, the £55,000 change they make is still an economically significant increase and on average they increase their housing wealth by one third. These are large changes in the portfolio composition of households who move house.

2.2.5 Subjective Bequest Probabilities

ELSA includes a number of survey questions that directly elicit the subjective expectations of respondents. I make use of a standard survey instrument that asks the probability of leaving a bequest larger than £150,000 with answers on a 101 point scale between 0 and 100. Household responses covary strongly with wealth and in Appendix A.4 I provide an example question, further discussion of subjective probability questions in ELSA, an alternative validation approach, and within birth cohort age profiles.

This subsection establishes two empirical facts. First, that subjective bequest probabilities contain informational content over and above demographic or economic variables. This demonstrates the advantage of including these measures as moments when estimating the model. Second, that different households report systematically different expectations about leaving bequests even after controlling for an extensive set of observable characteristics. I interpret this as indirect evidence supporting the hypothesis that bequest preferences are heterogeneous.

Subjective bequest probabilities contain information about a household's expected future path of savings over and above their current observable characteristics. To formalize the intuition, I estimate a series of quantile regressions for the partial correlation of future wealth and current subjective probabilities (controlling for additional observables $X_{i,t}$). The conditional quantile function $Q(\cdot|\cdot)$, for a given quantile τ, is given by:

$$Q_{Wealth_{i,t+1}}(\tau|\cdot) = \beta(\tau)Pr(Bequest \geq £150,00)_{i,t} + \delta(\tau)X_{i,t} \tag{2.1}$$

Figure 2.6 displays the results of these estimated partial correlations, $\beta(\tau)$,

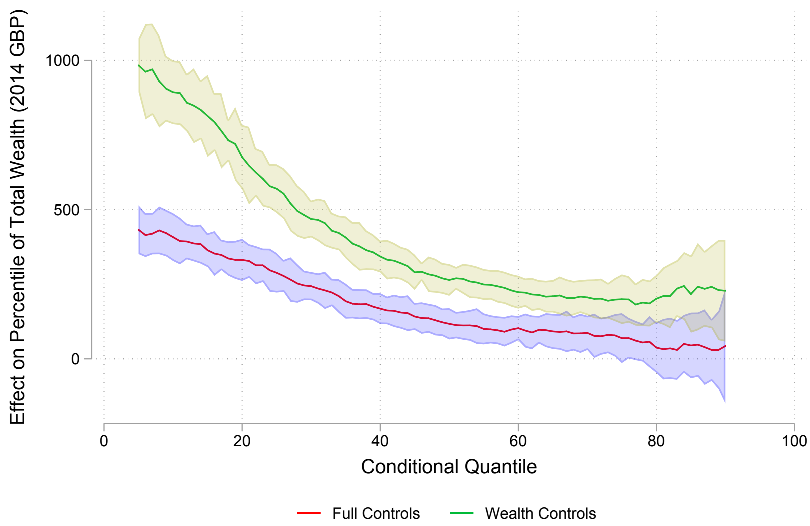


Figure 2.6: Information Content of Subjective Bequest Probabilities

Estimated partial correlations. Full controls include: current period wealth; polynomials in age and permanent income; household demographics; the health of each individual in the household; the sample wave and homeownership status

graphically for two alternative specifications of the conditional quantile function, in the first I additionally control for current period wealth; polynomials in age and permanent income; household demographics; the health of each individual in the household; the sample wave and homeownership status. In the second specification I control only for sample wave and current period wealth. The value of the coefficient at each point of the x-axis is the partial correlation (at a given conditional quantile of total wealth) of a 1 percentage point increase in the probability of leaving a large bequest.

The results from the quantile regression show that individual level variation in the subjective probability of leaving a large bequest is a statistically and economically significant predictor of future wealth holdings for all but the top of the wealth distribution. Under both specifications there is a decreasing pattern. In the main specification, at the conditional 5th percentile of future wealth a percentage point increase in the probability is associated with a £425 increase in tomorrow's wealth, while at the median it has fallen to approximately £100. Part of the decline is driven by difference in observables across the wealth distribution. However, the absence of the effect for the richest households is also an artefact of the survey design: these households hold assets well in excess of the £150,000 threshold and report that they are likely to leave a large bequest (reducing variation in the independent variable). Comparing the alternative specification, the estimated effect approximately halves in size when a full set of controls is included. Interpreting the systematic fall in the estimated effect across the distribution of total wealth highlights that while observable characteristics (or the state variables in a household problem) may explain a large fraction of the link between subjective beliefs and wealth there is a significant proportion that is unexplained. Exploiting this additional source of variation is useful when estimating a large structural model of retirement savings decisions.

2.2.5.1 Bequest Preference Index

There is considerable evidence that households differ in their preferences for leaving an inheritance. This is important because it may affect their savings decisions over the life-cycle, the extent to which they realize capital gains from house price changes by downsizing, and how they value resources that are not spent during their lifetime.

To better measure preference heterogeneity in the population (and how it is correlated with observable characteristics), I construct a single index of a household's likelihood of leaving a large bequest. I exploit the panel component of the ELSA data and combine multiple self reported probabilities (measured in different waves of the survey) into a single time invariant index that captures persistent and systematic differences across households.

Formally, I estimate a regression where the continuous measure of the self reported probability is the dependent variable. The object of interest is a household specific fixed effect which proxies the permanent differences in bequest preferences.[11] To the extent that all relevant state variables in the household problem are controlled for, fixed effects do not proxy systematically different expectations or the average realization of uncertainty. Instead, the fixed effect absorbs varia-

[11] In constructing this index I control for a number of contemporaneous household characteristics. See Appendix A.4 for a full description.

	OLS	Observables	Household Fixed Effect	Residual
Fraction of Total Variation	0.497	0.426	0.307	0.267

Table 2.2: Decomposing Variation in Subjective Probabilities

tion in the idiosyncratic preference for bequests and other time invariant features. Table 2.2 reports the share of variance in the subjective probability of leaving a bequest greater than £150,000 attributable to: a rich set of covariates, their bequest preference index (a transformation of the intraclass correlation) and a residual. In addition, I report the share of the variance attributed to the same set of covariates in a linear regression (the R-Squared) for comparison.

These results reinforce the link between wealth accumulation and self reported bequest probabilities displayed in Figure 2.6. First, observable household characteristics explain a large proportion of the individual variation in reported bequest probabilities, and, second, a significant fraction of the variation is attributed to systematic and persistent difference across households. As discussed above, these persistent differences may reflect variation in preferences for bequests. For this reason I model heterogeneity in household preference and exploit the bequest preference index in estimating the distribution of latent preference types discussed in more detail in Section 2.5.

2.3 Tax Reform as a source of Quasi-Experimental Variation

The UK tax system provides substantial variation in incentives for retired households due to features of estate and housing transaction taxes. When households cross thresholds in these tax systems their incentives can change substantially.

To help identify the structural model described in the next section I leverage variation over time in tax policy and the cross-sectional differences in household incentives produced by these thresholds. Changes to estate taxation and residential property transaction taxes provide a source of quasi-experimental variation that change the returns to leaving a bequest, the returns to holding different assets, and the cost of transforming housing wealth into liquid wealth at different points in the wealth distribution over time. In addition to the effect of exogenous policy changes, the large appreciation in house prices provides a second source of variation in the form of 'bracket creep' effects when price appreciation moves households into higher tax brackets. This section begins by describing the important reform to inheritance taxation before also describing the changes to transaction taxes. Finally, I show the reduced form effect of transaction tax reforms on the decisions of retired households using a regression discontinuity research design.

2.3.1 Inheritance Tax in the Sample Period

Despite its name, UK Inheritance Tax is levied on the estate of an individual who dies and not on the recipient of a bequest. Where an individual leaves the entirety of their estate to a spouse or civil partner (in the case that there are privately owned assets) there is no inheritance tax levied on the estate.

During the sample period, Inheritance Tax is charged at a constant rate of 40% of the estate above an exemption threshold and the Inheritance Tax exemption threshold is indexed to RPI. In 2010, this threshold was £325,000. A major reform was implemented on the 9th of October 2007. From this date, the tax exemption threshold increased by any unused proportion of a deceased spouse or civil partner's nil-rate band (even if the first partner died before 9 October 2007). Suppose the husband in the household died in 2003 and left £50,000 to their heirs and the wife died in 2010. The effective exemption threshold for the wife would be £600,000 because she is entitled to the full amount of her own exemption threshold (£325,000) and the unused proportion of her husband's nil-rate band (£325,000 less the £50,000 already bequeathed).

This effectively doubled the exemption threshold for the majority of older households.[12] Figure 2.2, which shows the mean wealth by birth cohort and PI,

[12] For the UK, Crawford and Mei (2018) report that nearly all wealth is left to a surviving partner

	Threshold by Rate				
Effective from	1%	3%	4%	5%	7%
28 March 2000	£60,000	£250,000	£500,000	*Not in use*	*Not in use*
17 March 2005	£120,000	£250,000	£500,000	*Not in use*	*Not in use*
23 March 2006	£125,000	£250,000	£500,000	*Not in use*	*Not in use*
03 September 2008[a]	£175,000	£250,000	£500,000	*Not in use*	*Not in use*
01 January 2010	£125,000	£250,000	£500,000	*Not in use*	*Not in use*
06 April 2011	£125,000	£250,000	£500,000	£1,000,000	£2,000,000

All thresholds and rates refer to transactions of residential property. During the time period there are additional exemptions for disadvantaged areas. [a] denotes the "Stamp Duty Holiday" where the 0% rate threshold was temporarily extended

Table 2.3: Rates and Thresholds for Stamp Duty Land Tax

shows that the mean wealth holdings in the top 50% of the lifetime resource distribution are near or above the original exemption rate.

2.3.2 Housing Transaction Taxes in the Sample Period

The Stamp Duty Land Tax (SDLT) was introduced in the UK in 2003, replacing the pre-existing Stamp Duty, and constitutes a transaction tax levied on all residential properties in the UK. During the sample period, the tax takes the form of a percentage rate charged on the whole purchase price if the price is above a particular threshold (because SDLT varies the average tax rate this creates discontinuous jumps, or *notches*, in the tax incentives). In 2005 the threshold for the lowest rate, 1%, doubled and increased again in 2006. In 2011, new higher rates were introduced at 5% and 7% for all properties above £1 million and £2 million. In addition to these changes, in 2008 the UK government introduced the 'Stamp Duty Holiday' a temporary (15 month) increase to the lower threshold from £125,000 to £175,000 expiring on December 31st 2009. This change is studied in both Besley et al. (2014) and Best and Kleven (2018). Table 2.3 summarizes the changes over the duration of my sample.[13]

How do transaction taxes affect retired households? Households who have

when one exists.

[13] In addition between the 1st of January 2010 and the 24th of March 2012 first time buyers enjoyed an additional exemption for residential properties costing less than £250,000

large amounts of wealth tied up in their home face these transaction costs if they choose to downsize and withdraw equity (or re-optimize because of reduced demand for housing services). Relative to a world without the transaction tax this creates significant disincentives. Consider a household owning a £400,000 house and wishing to downsize to a £300,000 house. Absent transaction costs (and any fixed costs of adjustment or changes due to collateral) the home owner would release £100,000 of equity. Under the SDLT policy the transaction tax levied on the new purchase is £9,000 which has an implied tax rate of 9% on the equity withdrawal. Besley et al. (2014) suggest that the incidence falling on sellers is 40%[14] and in this example the total cost paid is equivalent to a 10.2% tax on the £100,000 released (40% of £12,000 and 60% of £9,000).[15]

2.3.3 The Impact of SDLT on Home Moving Decisions

In estimating the structural model outlined in the next section, this paper directly incorporates variation in tax schedules over time. The reforms to UK transaction taxes generate additional variation over time in the incentives households face when choosing whether or not to sell their house. How strongly older households decisions vary with the changes to these incentives is an empirical question.

I show the response to transaction taxes in reduced form evidence by analysing the mobility decisions of older households around stamp duty thresholds using a regression discontinuity research design.[16] A reduction in home mobility is a reduction in the extensive margin of home equity adjustment. The empirical specification draws on Hilber and Lyytikäinen (2017) who analyse the moving decisions of working age UK households and my results suggest that the effect of transaction

[14]Kopczuk and Munroe (2015) present alternative estimates of transaction tax incidence using New Jersey Mansion taxes and find that it is entirely incident on the seller. In contrast, Slemrod et al. (2017) use notches from Washington, DC and estimate equal incidence on buyers and sellers.

[15]For a a household with a house worth £250,000 wishing to downsize to a house worth £200,000, and now release 20% of the equity in their home, the effective tax rate on the equity released is 13.2% (40% of £7,500 and 60% of £6,000 divided by the £50,000 base)

[16]A recent literature has shed light on the effect of housing transaction taxes and their impact on transaction volumes (Best and Kleven, 2018), sale prices (Besley et al., 2014; Kopczuk and Munroe, 2015; Slemrod et al., 2017) and mobility decisions (Hilber and Lyytikäinen, 2017). However, a potential concern is that these findings are driven by younger households who have a higher baseline mobility rate - if that were true there would be no additional identifying power from the SDLT reforms when studying older households.

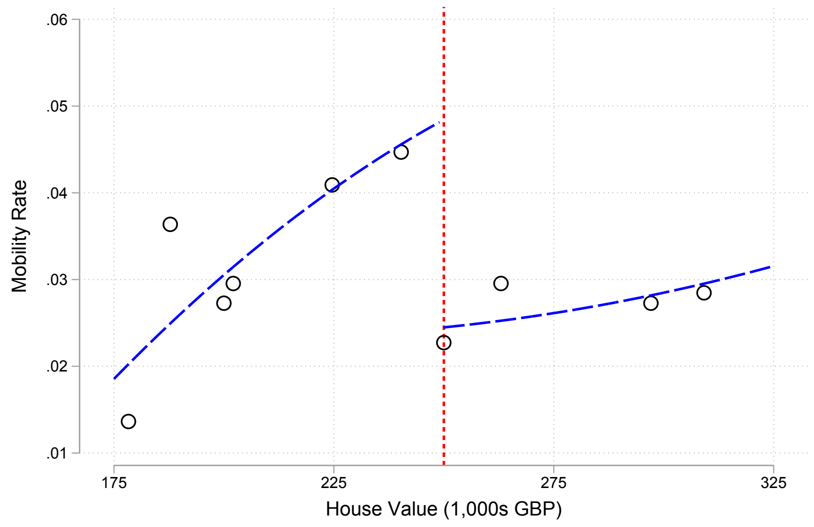


Figure 2.7: Mobility Rate and House Values

Circles are mobility rates for deciles of house value distribution within the sample window. The blue dashed shows the predicted fit of a regression of moving on a treatment for exceeding the £250,000 threshold and a quadratic in home value using a 30% window around the threshold. Further details are provided in Table 2.4.

taxes is of the same order of magnitude for old and young households.

I focus on a *notch* in the tax system at the £250,000 threshold because it remains constant throughout the sample period. For sale values that exceed this threshold, there is an increase in the average tax rate paid on the transaction from 1 to 3% and a discontinuous increase in the SDLT burden of £5,000. The outcome variable of interest, $Move_{i,t}$, is a dummy variable denoting a household's mobility between waves t and $t+1$ with treatment defined as a house value greater than or equal to £250,000:

$$Move_{i,t} = \beta_0 + \beta_1 Treat_{i,t} + f(HouseValue_{i,t}) + \delta X_{i,t} + u_{i,t} \tag{2.2}$$

The vector of control variables, $X_{i,t}$, includes a polynomial in household age, a polynomial in permanent income, household demographics, and wave and region indicators. I present results approximating the flexible function of house value $f(\cdot)$ in the conditional expectation function under three separate specifications: linear and quadratic with common slopes and a non-parametric local linear estimator where the slope differs across the discontinuity. I limit the analysis to a maximum 30% interval around the discontinuity in the tax schedule to avoid contamination from the effects of the 'Stamp Duty Holiday' (see Table 3).

The key identifying assumption is that, conditional on the covariates, $u_{i,t}$ is uncorrelated with the treatment indicator $Treat_{i,t}$. In regression discontinuity frameworks, this is satisfied if other covariates vary smoothly and the forcing variable ($HouseValue_{i,t}$) cannot be manipulated. Two features of the data reduce the concern of manipulation: first, home moves are measured in the following wave so that the reported home value is predetermined and, second, self reported home valuation is not the actual sale price used to calculate the SDLT burden. Following Kolesár and Rothe (2018) I report standard errors clustered at the household level and provide alternative confidence intervals with guaranteed coverage properties in appendix A.5.[17]

Figure 2.7 plots the results graphically and shows evidence of a decrease in mobility for those households who exceed the £250,000 threshold. Table 2.4 presents results from the regression analysis, varying the order of the polynomial in house value and the band around the stamp duty threshold included in the regression. The first row shows the results for a linear specification with a common slope on either side of the discontinuity. Consistent with the graphical evidence, the negative effect of an increase in the transaction tax burden on home moves is largest for narrow windows around the threshold. However, it is negative and statistically significant across all of the reported bandwidths. The second row displays results for a quadratic specification. For the smallest band around the cut-off value the result is of a similar magnitude to the linear specification, but estimated with less precision. Increasing the size of the band around the cut-off has a larger effect on the precision of the estimate than in the linear specification and has only a modest effect on the point estimate which is stable across different windows around the discontinuity in the SDLT schedule.

[17] Appendix A.5 additionally provides results for higher order polynomials.

	Band around cuttoff				
Order of polynomial	10%	15%	20%	25%	30%
Common Slope					
Linear	-0.0445**	-0.0475***	-0.0207*	-0.0200*	-0.0250**
	(0.0181)	(0.0160)	(0.0114)	(0.0116)	(0.0110)
	-729.5	*-879.4*	*-2118*	*-2112*	*-2873*
Quadratic	-0.0365	-0.0450**	-0.0270**	-0.0218*	-0.0265**
	(0.0241)	(0.0193)	(0.0133)	(0.0130)	(0.0112)
	-727.9	*-877.5*	*-2118*	*-2110*	*-2873*
Non-parametric					
Local Linear	-0.0278	-0.0341	-0.0356*	-0.0303**	-0.0287**
	(0.0334)	(0.0247)	(0.0189)	(0.0155)	(0.0141)
N	1224	1559	3023	3233	3979

All regressions additionally control for wave fixed effects, a polynomial in age, household demographics, a polynomial in permanent income and region dummies. Following Kolesár and Rothe (2018), Standard Errors are clustered by household. The Akike Information Criterion is shown in italics * $p < 0.10$, ** $p < 0.5$, *** $p < 0.01$

Table 2.4: The Effect of Transaction Taxes on Household Mobility

Finally, the third row displays a specification using a non-parametric local linear estimator. For all bands around the discontinuity the non-parametric method yields similar point estimates to the parametric approach; however, for small bands around the cut-off these results are imprecisely estimated. For larger bands around the discontinuity in the SDLT schedule the results are precisely estimated and the negative effect of an increase in the transaction tax burden is statistically significant. The treatment effect of exposure to higher transaction taxes is negative in all specifications and the magnitude of the effect is robust to alternative estimation windows and methods for approximating the conditional expectation function.

The estimates of exposure to higher transaction taxes, which range from a 2 to 4 percentage point reduction in mobility, suggest that the effect of an increase in transaction taxes for older households is also economically significant. Despite their lower baseline mobility, these results imply older households have similar reductions to working age households (Hilber and Lyytikäinen, 2017) in relative terms.

2.4 A Model of Household Savings After Retirement

In this section, I introduce a model of the savings decision of retired households that is able to generate the key empirical results in Section 2.2. The model features realistic risks in retirement and includes a rich model of housing decisions that matches the institutional features of the UK. Households face idiosyncratic and exogenous risk in health status, mortality, long term care expenditures, and (for couples) the size of the household. In addition, households are exposed to aggregate risk in the form of a common and stochastic process for house prices. The model advances the retirement savings literature in two main dimensions. First, I allow for each household to have heterogeneous preferences over bequests and, second, I allow for households to adjust their stock of housing on the intensive margin as well as the extensive margin of homeownership.

A household begins retirement as either a single or couple. For couples, if their spouse dies the surviving member continues as a single. Single retirees cannot remarry. Household size affects the income available to households, their utility from consumption as well as health transitions, mortality and medical expense risk.

Households may be either homeowners or renters (with housing wealth equal to zero). Each period, a household chooses their consumption, home equity, and the stock of financial assets for the next period. Financial assets are risk free, perfectly liquid and yield constant interest r. There is no borrowing.[18] The housing stock depreciates at rate δ and has a price p_h which households take as given. In addition, renters (who may choose to purchase a house) must choose a level of housing services purchased at rental price r^h.

[18] In the current version of the model I rule out collateralized and uncollateralized borrowing. The reasons for this are threefold. First, the majority of older households have paid off their mortgage or have positive liquid wealth balances. Second, after retirement many households fail to meet the income requirements of traditional forward mortgages and in the data very few retired households take out new mortgages to upsize. Third, although reverse mortgage products do exist (as considered in Nakajima and Telyukova, 2017; Cocco and Lopes, 2019) they are rarely used in the UK context. I provide further discussion of the UK and US reverse mortgage markets in appendix A.8. Finally, the UK market is tightly controlled on negative equity where the total value of the mortgage is still required to be paid in full.

At the beginning of each period, each household observes their current age, permanent income, who is alive in the household, cash on hand, housing wealth, health, medical expense shock and the level of the aggregate house price. Decisions are made after shocks are observed and new shocks arrive at the end of the period after decisions have been made.

When describing the model, I suppress the index i for an individual household - in the interest of clarity I make one exception: the coefficients of their heterogeneous bequest motive.

2.4.1 Demographics

A household is either a single man, single woman, or a couple. The state variable f is the household structure describing their demographics.

$$f = \begin{cases} \text{Single Man} \\ \text{Single Woman} \\ \text{Couple} \end{cases} \tag{2.3}$$

2.4.2 Preferences

Preferences are time separable, with a constant discount factor β. Households maximize expected utility, the per-period utility function is given by:

$$u(s_j, c_j, h_j) = \frac{s_j(\frac{1}{\alpha_{s_j}} c_j^{\sigma} h_j^{1-\sigma})^{1-\gamma} - 1}{1-\gamma} \tag{2.4}$$

where c_j is the consumption of non-durable goods at age j and h_j is the level of housing consumption at age j. The term s is a deterministic function of family status f and is the size parameter equal to the number of adults in the household and α_s is the consumption equivalence scale for total consumption. In this specification, γ is the coefficient of relative risk aversion and σ is the weight of non-durable consumption relative to housing services.[19]

[19] I impose a within period Cobb-Douglas aggregator for total consumption as many studies, such as Davis and Ortalo-Magné (2011), find that the expenditure shares on housing and consumption are constant

Utility from bequests for a household i, $\phi^i(b)$ is net of taxes and takes the form of a warm glow bequest motive as in De Nardi (2004) or Andreoni (1989). The functional form for $\phi^i(b)$ is given by:

$$\phi^i(b) = \frac{\phi_1^i(\phi_2^i + b)^{(1-\gamma)} - 1}{1-\gamma} \tag{2.5}$$

where ϕ_1^i controls the relative weight of bequests and total consumption for household i, while ϕ_2^i controls the curvature. Therefore ϕ_2^i controls the extent to which bequests are a luxury good. For positive ϕ_2^i marginal utility of small bequests is bounded, while the marginal utility of large bequests declines more slowly than consumption. One interpretation of the household specific preferences[20] is that ceteris paribus ϕ_1^i represents heterogeneity in altruism (or varies with the weight on the utility of future generations), while ϕ_2^i represents the human and financial wealth of the next generation. This specification is also consistent with other interpretations of the bequest motive such as pure egoism or strategic bequest motives (Bernheim et al., 1985).

2.4.3 Income, Health Status and Mortality

Income Households earn a return r on their financial assets a. Non-asset pension income y is deterministic and depends on age j, current family structure f and permanent income I:

$$y = y(j, f, I) \tag{2.6}$$

In addition to Estate Tax and Stamp Duty levied on housing transactions which are described in detail above, taxes are levied on pension income and income from financial assets.

Health Status Health status can take one of three values for each living household member

[20]This interpretation is consistent with Abel and Warshawsky (1988) formulation of so called 'Joy of Giving' bequest motives as a reduced form of altruism

$$m = \in \{good, bad, ADL, dead\} \tag{2.7}$$

and transitions according to an age, family structure and permanent income dependent Markov process. Following Ameriks et al. (2018) I use difficulties with Activities of Daily Living (ADLs) to define the worst health state. For couples, m denotes a pair with a health status for each member - for notational convenience I continue to use m to denote the nine valued health status for the couple.

Mortality Individuals within households face exogenous mortality risk which depends on age, family structure, health status and permanent income.

2.4.4 Medical Spending

In the literature focusing on US retirees, out of pocket medical expenditure risk is an important driver of precautionary savings. In the UK the NHS provides comprehensive coverage for acute and chronic medical expenses. However, long term care risks pose considerable out of pocket risk- with lifetime care costs for 10% of individuals exceeding £100,000.[21]

I define mx_j as the flow of all out-of-pocket medical expenses incurred between j and $j-1$. Medical expenses are exogenous[22] and depend upon the current health status of the household, the last period health status of the household, household permanent income, family structure, age, and an idiosyncratic component, $\varepsilon_{mx,j}$:

$$\ln mx_j(\cdot) = \mu_{mx}(\cdot) + \sigma_{mx}(\cdot) \times \varepsilon_{mx,j} \tag{2.8}$$

$$\mu_{mx}(\cdot) = \mu_{mx}(m_{j-1}, m_j, I, f_{j-1}, f_j, j) \tag{2.9}$$

$$\sigma_{mx}(\cdot) = \sigma_{mx}(m_{j-1}, m_j, I, f_{j-1}, f_j, j) \tag{2.10}$$

$$\varepsilon_{mx,j} \sim N(0,1) \tag{2.11}$$

[21]The UK figure is drawn from the Dilnot Report. This is lower than the figures for US lifetime medical spending of retirees reported in Jones et al. (2018), but there are two major differences: comparable US numbers include all medical spending (including Medicaid expenditure and hospital stays) and are reported at the household level, not for individuals

[22]This is consistent with Ameriks et al. (2018) who find limited income elasticity of endogenous medical spending for individuals with ADL needs using flexible health state dependent utility of spending and an estimated structural model.

2.4.5 Housing

It is costly to move home. I model these transaction costs with three features that capture the different types of costs faced by different households: the formal transaction tax (SDLT), a proportional transaction cost, and an age varying fixed cost of adjustment. These costs reflect real transaction or moving costs as well as any psychic costs (expressed in their financial value) associated with the housing search and moving between homes.

The total value of a house, h, is $p_h h$. If a household wishes to purchase a new house (including renters who hold $h = 0$ housing) they must pay the transaction cost with the total cost of adjusting housing $Q(\cdot,\cdot,\cdot)$ taking the following form:

$$Q(h_{t+1}, h_t, p_{h,t}, j) = \begin{cases} 0, & \text{if } h_{t+1} = h_t \\ p_{h,t}h_{t+1} - p_{h,t}h_t(1-\pi) + F_j & \\ +(1-\kappa)\cdot\tau_h(p_{h,t}h_{t+1}) + \kappa\cdot\tau_h(p_{h,t}h_t), & \text{if } h_{t+1} \neq h_t \end{cases} \tag{2.12}$$

A household that does not adjust their housing stock has an adjustment cost of 0. If they adjust their stock of housing, the total cost consists of three parts. First, the change in housing evaluated at today's price and net of a proportional transaction cost π. The proportional component of the transaction cost allows the costs to vary between houses of different values or sizes.[23] Second, the age dependent fixed costs of moving, F_j. These account for age variation in these costs such as those driven by declining physical ability in old age. Finally, the transaction tax (SDLT), $\tau_h(\cdot)$, which has incidence on the seller $\kappa \in [0,1]$. Accounting for the incidence which falls on buyers and sellers captures a salient feature of the UK tax landscape and incorporates potential sale price manipulation around SDLT thresholds in a reduced form way.

To parametrize the incidence of transaction taxes I use estimates from Besley

[23] Real costs may be proportional because the complexity of realtor fees, legal agreements or surveying varies with the property value. Likewise hassle or psychic costs may vary because larger houses are associated with moving (or disposing of) more possessions or require different search intensities due to differential market thickness.

et al. (2014) and model the evolution of the tax system over time - this is key to identifying the rich cost structure in the model. The age dependent fixed costs of moving is an age invariant fixed cost and quadratic in age:

$$F_j = F_0 + F_1 j + F_2 \frac{j^2}{100} \tag{2.13}$$

In addition to costs when they move, homeowners must pay the maintenance cost δ every period which is proportional to the total value of their house. Renters (who may choose to purchase) rent housing services this period at a fraction r^h of the sale price.[24] Consequently, renters are exposed to housing market volatility and the model can generate precautionary savings for renters, as well as a precautionary owning motive for housing rich and income poor households.[25]

2.4.6 House Prices

House prices are stochastic and their log evolves as follows:

$$\ln(p_{h,t+1}) = \mu_h + \rho_h \ln(p_{h,t}) + \varepsilon_{h,t+1}, \quad \varepsilon_{h,t+1} \sim N(0, \sigma_h) \tag{2.14}$$

This is a standard AR(1) process with drift μ_h which reflects the trend growth in house prices. This formulation, including the nested random walk case with $\rho_h = 1$, is common in the literature[26] and fits the data at both individual and aggregate levels well (see Nagaraja et al., 2011; Berger et al., 2018, respectively). I model aggregate house price movements rather than the exposure of households to idiosyncratic shocks. Thus, the price level is common to all households at calendar time t.

2.4.7 Recursive Formulation

Let a_t denote the liquid wealth balance of households at time period t and let r denote the return on these balances. Total post tax income is $\tau_y(r a_t + y_t(\cdot), \tau)$ with

[24] This rental price includes the rental premium and any additional utility from home ownership. A higher value of r^h implies that it is costlier for renters to rent a home providing equivalent levels of housing consumption.

[25] This is consistent with the empirical evidence in Sinai and Souleles (2005) suggesting that some owner occupiers use housing to hedge against volatility in rental markets.

[26] See, for example, Mitman (2016) and Berger et al. (2018) as well as Campbell and Cocco (2007) and Attanasio et al. (2012) in the UK context

vector τ summarizing the tax system. I follow Deaton (1991) and redefine the problem in terms of cash-on-hand:

$$x_t = a_t - \delta p_{h,t} h_t + \tau_y(r a_t + y_t(\cdot), \tau) + tr_t(\cdot) - mx_t, \tag{2.15}$$

The law of motion for cash on hand next period is given by

$$\begin{aligned} x_{t+1} = & \, x_t - r^h \widetilde{h_t} p_{h,t} - c_t - \delta p_{h,t+1} h_{t+1} - Q(h_{t+1}, h_t, p_{h,t}, j) - mx_{t+1} \\ & + y\big(r(x_t - c_t - r^h \widetilde{h_t} p_{h,t} - Q(h_{t+1}, h_t, p_{h,t}, j)) + y_{t+1}(\cdot), \tau\big) + tr_{t+1}(\cdot) \end{aligned} \tag{2.16}$$

Here $\widetilde{h_t}$, the rental choice is, 0 for all households who own a house in period t ($h_t > 0$). The first line is the amount of savings brought forward into the period, net of housing maintenance costs, and the second line is income after taxes and transfers.

The tax function accounts for means tested transfers excluding the state coverage of Long Term Care expenses. Following Hubbard et al. (1994, 1995), and De Nardi et al. (2010) I assume that the government provides means-tested transfers, $tr_t(\cdot)$, that bridge the gap between a minimum consumption floor and the household's resources when households are exposed to long term care costs. Define the resources available next period after tax, but *before* government transfers with

$$\begin{aligned} \widetilde{x_{t+1}} = & \, x_t - r^h \widetilde{h_t} p_{h,t} - c_t - mx_{t+1} - Q(h_{t+1}, h_t, p_{h,t}, j) - \delta p_{h,t+1} h_{t+1} \\ & + \tau_y\big(r(x_t - c_t - r^h \widetilde{h_t} p_{h,t} - Q(h_{t+1}, h_t, p_{h,t}, j)) + y_{t+1}(\cdot), \tau\big) \end{aligned} \tag{2.17}$$

Consistently with the state coverage of long term care expenses which depends on total resources, housing, health status and family status, government transfers next period are

$$
\begin{aligned}
tr_{t+1}(\widetilde{x}_{t+1}, f_{t+1}, h_{t+1}, m_{t+1}, p_{h,t+1}) = \max\big\{0, c_{min}(f_{t+1}, h_{t+1}, m_{t+1}) - \\
(\widetilde{x}_{t+1} - a_{D,t+1}(f_{t+1}, \widetilde{x}_{t+1}, m_{t+1}) \\
- h_{D,t+1}(f_{t+1}, h_{t+1} p_{h,t+1}, m_{t+1}))\big\}, \quad (2.18)
\end{aligned}
$$

Where $a_{D,t+1}(\cdot)$ and $h_{D,t+1}(\cdot)$ are an asset and housing disregard respectively. In the UK, means testing occurs at the individual level and implies that housing owned by couples and 50% of their assets are excluded from means testing at the household level when only one member of the household enters a long term care facility.[27] This is a key feature of the economic environment for retirees which introduces large distortions into their self insurance decision.

The law of motion for cash on hand next period can thus be rewritten as

$$x_{t+1} = \widetilde{x}_{t+1} + tr_{t+1}(\widetilde{x}_{t+1}, f_{t+1}, h_{t+1}, m_{t+1}, p_{h,t+1}). \quad (2.19)$$

To ensure that savings are always non-negative, I require total expenditures do not exceed total resources:

$$c_t + Q(h_{t+1}, h_t, p_{h,t}, j) + r^h \widetilde{h}_t \leq x_t, \quad \forall t. \quad (2.20)$$

I define liquid savings as cash on hand net of total expenditure:

$$a_{t+1} = x_t - c_t - Q(h_{t+1}, h_t, p_{h,t}, j) - r^h \widetilde{h}_t \quad (2.21)$$

Finally, I assume bequests are only possible when the final surviving member of the household has died. Bequests are exposed to long term care costs and constrained to be non negative. When calculating the current value of the estate, houses are liquidated at current prices and the total value of the estate is taxed. This implies

[27] A household which cannot afford to reach the minimum level of consumption even after liquidating their assets is forced to sell their house and expend all of their assets. They begin the next period as a renter and receive transfers which provide them the consumption floor.

the following formulation for the after tax value of their consolidated wealth:

$$b_t = \tau_b(max\{Q(0,h_{t+1},p_{h,t}) + a_{t+1} - mx_{t+1}, 0\}) \tag{2.22}$$

Where estates face the adjustment cost of selling, but with no other purchase.

I now provide the recursive formulation of the household problem. As is conventional, I use a prime to denote next period variables. The state variables of a household are given by $\Omega = (i, j, f, I, m, h, x, p_h)$. These variables are: the idiosyncratic bequest motive (index i), age (j), family structure (f)[28], permanent income (I), health status (m), housing (h), cash on hand (x), and the aggregate house price level (p_h). Throughout I use $h = 0$ to denote renters. First, the recursive problem for homeowners is:

$$\begin{aligned} V_j^i(f,I,m,h,x,p_h) = max_{\{c,h',a'\}} \Big\{ & u(s,c,h) + \\ & \beta \cdot surv(j,I,m,f) \times \\ & E[V_{j+1}^i(f',I,m',h',x',p_h') \,\big|\, \Omega, h', a'] \\ & + \beta(1 - surv(j,I,m,f))E[\phi^i(b) \,\big|\, \Omega, h', a'] \Big\} \end{aligned} \tag{2.23}$$

subject to equations (2.3)-(2.14) and(3.9)-(2.20) and bequests are constrained by (2.22). Households choose consumption c, savings in financial assets (before long term care costs) a' and the new housing stock h'. Household's take expectations over individual mortality, the size of the family structure tomorrow f', household health m', the transitory component of medical expenses ε_{mx}, and the level of house prices, p_h'. Due to medical expense uncertainty, households also take an expectation over realized cash on hand tomorrow x' and the possibility that they are compelled to sell their house to finance long term care costs.

Second, the recursive problem for the renter ($h = 0$) is:

[28] Note that household size, s, is a deterministic function of family structure

$$V_j^i(f,I,m,h=0,x,p_h) = max_{\{c,a',\tilde{h},h'\}}\Big\{u(s,c,\tilde{h})+ \\ \beta\cdot surv(j,I,m,f)\times \\ E[V_{j+1}^i(f',I,m',h',x',p_h')\,\big|\,\Omega,h',a'] \\ +\beta(1-surv(j,I,m,f))E[\phi^i(b)\,\Big|\,\Omega,h',a']\Big\} \tag{2.24}$$

subject to equations (2.3)-(2.14) and(3.9)-(2.20) and that bequests are constrained by equation (2.22).

The problem for the renter differs from the homeowners problem in equation (2.23) in two simple ways. First, the value of their existing housing stock is zero and, second, they must purchase a level of housing services $\tilde{h}$ at associated rental price r^h. Together, these two differences imply a renter specific budget constraint.

2.5 Estimation

I adopt a two-step estimation strategy. In the first step I estimate (or calibrate using existing evidence) those parameters that can be cleanly identified outside of the model. In the second step I estimate the remaining model parameters with the method of simulated moments (MSM) taking the first step parameters as given. I find the parameter values that minimize the distance between the simulated life cycle profiles and the profiles in the data where the distance criteria is measured by the GMM criterion function.

Heterogeneity in individual preferences for bequests is estimated in the two step procedure, as part of the first step I *classify* households into latent groups and in the second step I estimate the preference parameters for these groups. Crucially, the two step approach retains tractability in the estimation. This section proceeds as follows: first, I elaborate on the *classification* step. Second, I discuss the remaining first stage estimation. Third, I detail the moment conditions I choose to match and how I construct the moments in the simulated data. Finally, I provide a discussion of the model's identification which serves an explanation for how these moment conditions were selected. The results of the estimation procedure are discussed in Section 2.6.

2.5.1 Preference Heterogeneity

Equation 2.5 specifies the form of preference heterogeneity, allowing both the relative weight and curvature of the bequest motive to vary across households. In practice, I discretize preference heterogeneity by assuming that households can belong to one of a finite number of types (or classes) which differ in their preferences over bequests.

I do not observe each household's latent type - instead , for each household, I estimate type membership outside the model. Subsequently, I treat these types as given when estimating the preference parameters for each group with the remaining preference parameters in the second stage. I follow Bonhomme et al. (2019), who discretize the types of possible firms in a first stage, in adopting a two stage estimator and use a k-means clustering approach to determine type membership.[29] This approach treats the bequest preference parameters as non-linear group fixed effect that is assumed to be time and policy invariant.

Letting z_i denote a vector of household characteristics, the k-means clustering problem used in the classification step (for a given number of clusters K) is defined as:

[29] Lentz et al. (2019) further allow for an iterated classification procedure in a dynamic sorting model of workers and firms. As discussed in Bonhomme et al. (2017) (which establishes asymptotic properties of two step grouped fixed effect estimators as approximations to underlying continuous distributions in non-linear panel data models) these methods are particularly attractive because they maintain tractability by using a data driven approach to reduce the state space (alleviating the curse of dimensionality). The use of unsupervised learning techniques to cluster households is gaining traction in economics as a data driven method for establishing latent household types- additional applications include: identifying the life cycle employment paths of entrepreneurs (Humphries, 2018), the work disincentives and investment productivity of mothers (Mullins, 2018), the decomposition of wage inequality across workers and firms (Lentz et al., 2019), and the associative matching between workers and firms Dauth et al. (2018).

$$\min_{\mathscr{K},\{\bar{z}_k\}_{k=1}^K} \sum_{k=1}^{K} N_k \sum_{k(i)=k} || z_i - \bar{z}_k ||^2 \quad = \min_{\mathscr{K},\{\bar{z}_k\}_{k=1}^K} SSE, \quad \text{with} \bar{z}_k = \frac{1}{N_k} \sum_{k(i)=k} z_i \tag{2.25}$$

Where the classification is given by:

$$\mathscr{K} = \{k(i)\}_{i=1}^{n} \tag{2.26}$$

I fix the number of types, K, at five and treat this as known throughout the analysis.[30] Given the vector of household characteristics, the classification step minimises the within cluster sum of squared errors (equivalent to selecting clusters to minimise the population sum of squared errors). Informed by the existing empirical literature and economic theory, cluster assignment is an unrestricted function of three household characteristics: their estimated permanent income rank, their total wealth at their initial observation, and their value of the bequest preference index which I estimate in Section 2.2.5. These variables allow for a flexible, yet parsimonious, grouping of households. Each of these measures is individually (and jointly) correlated with potential unobserved differences in the desire of households to leave an inheritance because they are outcomes of choices made by households. Retirement wealth is determined by early life choices (Venti and Wise, 1998) and this may reflect differential saving rates or portfolio investments made by households with different desires to leave an inheritance. As I focus on the retirement period I abstract from these decisions, but allow the types to depend on the observed level of household wealth that may be informative about differences in their preferences. Put differently, allowing for an arbitrary correlation between household wealth when they enter the sample and their preferences allows for the possibility that part of these differences are explained by choices made in early life that I do not explicitly model. Allowing type membership to vary with a household's lifetime

[30] Heuristic methods for k-means clustering identify five as the optimal number of clusters. These heuristics are described in more detail in Appendix A.6.

Group	Number of Households	Share	Mean Characteristics: Permanent Income (Percentile Rank)	Mean Characteristics: Total Wealth (2014 GBP)
Type 1	847	18.23%	66.9	157,647
Type 2	1,584	34.09%	21.8	67,913
Type 3	573	12.33%	82.8	998,855
Type 4	980	21.09%	72.6	312,900
Type 5	663	14.27%	26	323,902

Table 2.5: Distribution of Latent Household Types

income allows those who are lifetime poor (rich) and maintain large levels of wealth in retirement to be distinguished from those who do not as well as the potential that past work effort is determined by the desire to leave a bequest. Finally, while the bequest preference index may be a noisy measure of household preferences it is informative about systematic expected differences in future bequests across households.

Table 2.5 gives the distribution of households over these discrete types.[31] Additionally, I present summary statistics for permanent income and wealth by household type. This provides a concise characterization of the estimated types. The marginal distributions of household characteristics are shown in appendix A.6 to provide a comprehensive characterization.

Type 1 is comprised of households with above average lifetime incomes, but comparatively low levels of wealth. In contrast, Type 2 households are on average lower lifetime income with low levels of wealth in retirement. Initial renters are classified into Type 1 and 2 with high lifetime income renters falling into the first type. Type 4 and 5 are distributed in a similar pattern with Type 4 households high income and high wealth. Type 5 households are relatively high wealth households drawn from the lower part of the lifetime income distribution. The remaining Type 3 households are comprised of the richest households in lifetime income and re-

[31]While any labelling of households is ad hoc, I label types numerically such that the average value of the bequest preference index is increasing across types.

tirement wealth. Household types are ranked in order of their bequest preference index: Type 1 and 2 have values that are on average below the mean (and vice versa for Types 4 and 5) with the largest systematic differences in household subjective probabilities of leaving a bequest occurring in Types 1 and 5.

2.5.2 First Stage Estimates

In the first stage I calibrate the remaining parameters that control the utility function, housing market and budget constraint using values from the literature and estimate parameters that are cleanly identified outside the model. These are: the deterministic profile for household income as a function of the state variables (equation 2.6) and income tax function; the estimated health status transition probabilities and mortality (including transitions in family structure); and the time series for aggregate house prices (equation 2.14). Table 2.6 summarizes the value and source of the first stage estimates. I report all values as annual.

The Utility Function I parametrize that the consumption equivalence scale, as a function of household size, using the OECD scale. The remaining parameters in the per-period utility function and bequest utility are estimated in the second stage.

The Housing Market The annual depreciation cost offset by maintenance, δ, is set at 2% and the rental cost is set at 4.05% of the sale price (Cocco and Lopes, 2019; Etheridge, 2017, respectively). The rental cost captures the implied gross rental yield for private landlords as well as the implicit utility premium from home ownership. I estimate the time series profile for aggregate house prices using the HM Land Registry UK house price index series.[32] House prices are highly persistent, have a significant upwards trend, and large innovations at an annual frequency. The incidence of the transaction cost that falls on the seller is taken from the estimates in Besley et al. (2014).

The Budget Constraint I calibrate the annual rate of return on the risk free asset, r, as 3% following Bozio et al. (2017). I estimate non-asset pension income directly from the data in ELSA and describe this procedure in more detail in Appendix A.1.

[32] The UK has a number of house price series see Chandler and Disney (2014) for a review.

Parameter	Description	Value	Source
	Utility Function		
α_s	Consumption Equivalence Scale	1.5	OECD Modified Scale
	Housing Market		
δ	Housing Maintenance Costs	0.02	Cocco and Lopes (2019)
r_h	Rental Cost	0.0405	Etheridge (2017)
ρ_h	House Price AR(1) persistence	0.977	HM Land Registry
μ_h	House Price Drift	0.019	HM Land Registry
σ_h	House Price S.D. Innovations	0.095	HM Land Registry
κ	Incidence of SDLT on Seller	0.4	Besley et al. (2014)
	Budget Constraint		
r	Risk Free Return	0.03	Bozio et al. (2017)
$y(\cdot)$	Deterministic Income Profile		ELSA
τ_y	Income Tax Function		TAXBEN
c_{min}	LTC consumption floor (Singles)	£4,956	Lockwood (2018)
c_{min}	LTC consumption floor (Couples)	£7,434	Lockwood (2018)
	Mortality and Demographic Transitions		
$surv(\cdot)$	Survival Probabilities		ELSA
$Pr(m^g_{j+1}\|\cdot)$	Health status		ELSA
	LTC Costs		
$\mu_{mx}(\cdot)$	Mean		HRS
$\sigma_{mx}(\cdot)$	Conditional variance		HRS

All values are annual and expressed in 2014 prices.

Table 2.6: 1st Stage Parameter Estimates

PI Percentile	Men: Good Health Life Expectancy	Men: Good Health ADL Years	Men: Bad Health Life Expectancy	Men: Bad Health ADL Years	Women: Good Health Life Expectancy	Women: Good Health ADL Years	Women: Bad Health Life Expectancy	Women: Bad Health ADL Years
Singles								
10th	13.65	2.02	11.23	2.85	18.21	2.57	16.38	3.82
50th	16.91	2.32	14.91	2.92	20.02	2.65	19.14	3.82
90th	19.57	1.58	17.83	1.87	20.93	1.91	20.03	2.66
Couples								
10th	13.31	2.15	10.95	3.02	19.18	3.92	17.99	5.60
50th	16.79	2.36	15.10	3.21	21.19	4.07	20.65	5.48
90th	19.29	1.57	17.40	2.19	21.89	2.65	20.88	3.56

Conditional on surviving to age 66. ADL difficulties defined as 2 or more. For couples the. calculation assumes both spouses have the same health at age 66

Table 2.7: Life Expectancy & Expected Duration of ADL difficulties

The income tax function is a modified version of a common log-linear functional form[33] where after tax income is given by:

$$\tilde{y} = \bar{y} + \lambda_y y^{1-\tau_y}$$

where λ_y controls the level of taxation, τ_y controls the progressivity and $\bar{y}$ captures features of state assistance for older households which correspond to an income floor. I estimate this separately for couples and singles using ELSA data combined with tax and benefit entitlements calculated using TAXBEN.[34] Means tested coverage of social care costs is not included in the data used to estimate the tax function. The consumption floor is set to replicate the utility value of receiving public assistance for long term care needs and includes any disutility from receiving state care. For couples this value is equivalized.

Mortality and Demographic Transitions I estimate survival probabilities and health status transition probabilities directly from the ELSA data using a multinomial logit approach and allow transitions to depend on age, family size, health

[33] The earliest example dates to Feldstein (1969)
[34] Appendix A.11 provides additional details.

status and permanent income. ELSA has data on six different ADL measures for each individual and I define the ADL state as those who have difficulties with at least two of these measures. ADL measures capture a range of needs associated with institutional long term care use and care in the community. Summary statistics are reported in table 2.7. The ELSA ADL data and the specifics of these measures are discussed in more detail in Appendix A.9.

Long Term Care Costs Micro-data on the out of pocket costs faced by households in the UK is scarce, however, for some smaller costs Banks et al. (2019) document that figures reported in the Health and Retirement Study (HRS) line up closely with those reported in ELSA.[35] I use the HRS to estimate the cost of care in the community and long term care costs for UK households. Specifically, I construct an equivalent health status measure and estimate the parameters of the medical spending process described in equations 2.8-B.7. After estimating the process for the US, I impose that the mean and variance of medical spending is zero in all health states other than the ADL state to better replicate the institutional environment of the UK.

2.5.3 Moment Conditions

In the second stage I estimate the remaining parameters taking the first stage estimates as given. The remaining parameters are:

$$\theta = (\{\phi_1^k, \phi_2^k\}_{k=1}^5, \beta, \gamma, \pi, F_0, F_1, F_2) \tag{2.27}$$

The following moment conditions comprise my estimator:

1. For initial homeowners, I match mean liquid wealth and housing wealth by age, permanent income, and cohort because liquid and housing wealth imply different ability to self insure over short and long run horizons and because the mortality and LTC expenditure risks households face vary with their age and permanent income. For renters I match mean liquid wealth by age and cohort because renters are predominantly drawn from low lifetime incomes.

[35] These data are only available for wave 8 of ELSA.

2. To exploit the additional information about future wealth trajectories contained in self reported bequest probabilities, I match mean subjective bequest probabilities by age, permanent income, and cohort for initial homeowners. For renters I match mean subjective bequest probabilities by age and cohort.

3. Moving is costly, but liberates liquidity from housing. To identify the extent of these costs and help pin down household's demand for liquidity versus housing services I also match the fraction of moves by age and cohort for initial homeowners.[36]

Using within group means as moment conditions requires that the model is able to fit well the full distribution of wealth holdings in the population because I condition the moments on permanent income and a household's total wealth rank is closely correlated with their lifetime permanent income. I provide more details of the model's identification and the motivation for selecting these moments in section 2.5.4. Finally, I top code wealth moments in the data and in the simulations at the 95th percentile. This mitigates the impact of the very wealthy and other potential sources of measurement error. In practice I use 7 five year birth cohorts and 3 PI groups. The data versions of these moments are discussed in Section 2.2 and I perform identical operations to calculate the simulation equivalents. To calculate model moments for subjective bequest probabilities I compute the model implied objective probabilities for each individual.

The MSM approach is standard.[37] I simulate 50,000 sample households where their initial state variables (including sample wave of entry) are drawn from the joint distribution of state variables in the ELSA data. These simulated households are simulated for the duration that the equivalent ELSA household remains in the sample and receive the same realization of shocks that their ELSA donor receives. Consequently, households are sampled for a window that includes the same reforms

[36] In the data very few renters choose to purchase homes, but many move between different rental accommodation. The model has no conceptual equivalent of rental to rental moves

[37] See for example Gourinchas and Parker (2002); De Nardi et al. (2010). The model outlined above has no closed form solution and instead I use numerical methods which are described in Appendix A.7

to Inheritance Tax and Transaction Taxes (SDLT) that they experience in the data (at most 5 different regimes). Due to the frequency of these reforms, I do not explicitly target pre and post periods - instead, the behavioural response to the tax system are embedded in the moments I target: the home moving decisions, wealth deaccumulation, and bequest expectations of households by cohort and wave. I use a diagonal weighting matrix which takes the diagonal of the asymptotically optimal weighting matrix because the full asymptotically optimal weighting matrix is known to behave poorly in small samples (Altonji and Segal, 1996).

2.5.4 Identification

In this class of retirement savings models it is difficult to separately identify bequest and precautionary motives using only data on the levels of household wealth (see De Nardi et al., 2016b, for a discussion).[38] Instead, I combine data on wealth composition, subjective bequest probabilities and exogenous policy reforms (discussed in detail in section 2.3) to identify parameters of the utility function, the additional costs of moving home and heterogeneous bequest motives. These policy reforms shift the returns and risks associated with holding different assets as well as the returns to adjusting home equity.[39] In complex non-linear models, all moments potentially influence all model parameters, however, I provide intuition for why particular moments are more informative about certain parts of the model.

Parameters in the period utility function In the joint estimation, I estimate three parameters of the utility function: the intertemporal discount factor, β, the coefficient of relative risk aversion, γ, and the non-housing consumption share, σ. Both the intertemporal discount factor (β) and the coefficient of relative risk aversion (γ) affect the slopes of consumption and wealth profiles. To separately identify them I exploit variation in the risks household face at different ages and levels of the permanent income distribution: those with low PI face lower longevity and higher LTC costs while alive. In contrast, those with higher permanent income survive

[38]When targeting moments at the average, a small change in the risk aversion parameter (and consequently the precautionary savings motive) compensates the change in fit from eliminating the bequest motive.

[39]The assumption that the preference parameters of the structural model are not affected by the policy changes or house prices is a form of the exclusion restriction.

for longer and experience lower LTC costs. The differences in the level of liquid wealth (and the consequently the portfolio share) across the PI and age distribution identifies the risk aversion of households and the discount factor.

Renters are particularly responsive to the consumption share of housing because it determines their expenditure share on rent. This parameter drives variation in their marginal utility of expenditure as house prices change throughout the sample.[40] Matching the liquid wealth of renters identifies σ through variation in the house prices they face.

The Cost of Moving Households in the model move only when the benefits of adjusting their housing stock are larger than the costs. Households who hold similar housing stocks have potentially large differences in their financial incentives due to the design of the tax system. Tax policy and aggregate house prices generate variation in the returns to holding housing assets and, importantly, the cost of adjusting them over time and at different points in the distribution. However, the remaining costs of adjustment at each age, π and F_j, is policy invariant (an exclusion restriction) and is identified by the frequency of home moving by birth cohort and age.

Consider two otherwise identical households who are born in different birth cohorts. At the same age they face different house prices and different tax incentives for adjusting their housing stock. Differences in the home moving rate between ages and across cohorts (who face exogenously different costs and returns) as well as across the wealth distribution can be used to identify the parameters of π and F_j analogously to a difference in difference research design.[41]

Bequest Motives Intuitively, households with lower survival probabilities are more responsive to their bequest motives. Furthermore, the opportunity cost of saving for a bequest differs across assets. Matching the savings behaviour (in different assets)

[40] The substitutability of housing and non-housing consumption also determines how homeowners are differentially insured by the flow consumption of their housing - a form of housing services annuity. The higher the degree of substitutability between housing and non-housing consumption, the fewer liquid assets a household needs to self insure. However, renters provide an alternative to identify this substitution.

[41] In practice, the full dynamic model controls for differences across households such as differences in portfolio, health, PI or differences in bequest motives and contamination from multiple reforms that would otherwise confound this approach. Furthermore, it conditions on the full sequence of future mortality, demography, and long term care risks.

of households at different ages (and PI levels) exploits age (and PI) variation in their survival probability and saving motives, however, in practice it is difficult to separately identify the demand for housing from the demand for bequests.

The reform to estate taxation shifts the return to saving for a bequest independently from the return to saving for future lifetime consumption (including the flow consumption from housing): the extent to which households adjust their savings decisions in response to this tax reform helps pin down the weight on bequests. To leverage information on the full path of future wealth deaccumulation for all households I match subjective bequest probabilities which provides a separate source of identification for bequest motives over and above their observed current choices.[42] I match measures of household subjective bequest probabilities where conditioning on age, PI, and homeownership captures variation in the life expectancy, risks, and portfolio structure that also influence the bequests that households expect to leave.[43]

Heterogeneity in Bequest Motives Systematic differences in expected bequests across household types reflect systematic differences in their expected wealth paths. Likewise, systematic differences in household responses to policy changes reflect systematic differences in their incentives. The structural model provides a parametric interpretation for these differences. The model environment is rich enough to incorporate many other sources of observable heterogeneity across households including the effects of aggregate shocks, the tax environment, their endogenous portfolio choices, realisations of idiosyncratic shocks and any differences in initial conditions. This is crucial to identify structural differences in bequest motives rather than other characteristics that lead households to save (or expect to save) differently.[44] Finally, it is important to stress that the differences across households in

[42]Similarly, van der Klaauw and Wolpin (2008) argue that subjective household expectations contain household level information on optimal future behaviour conditional on current state variables.

[43]The different components of identification also exploit variation at different parts of the wealth distribution. It is only households who either exceed the exemption threshold in 2007 or expect to on death that are affected by the reform to estate taxation (the "treated" households). When preferences are not (necessarily) homogeneous and correlated with wealth, this is insufficient variation to recover the full distribution of bequest preferences. As shown in section 2.2, the subjective bequest probabilities are more informative for the bottom 50% of the wealth distribution.

[44]Household types are estimated *ex ante* which allows the model to replicate the composition of types in each of the moments (by cohort, PI, age and ownership). Furthermore, each of the groupings that define moments closely map to each of the types.

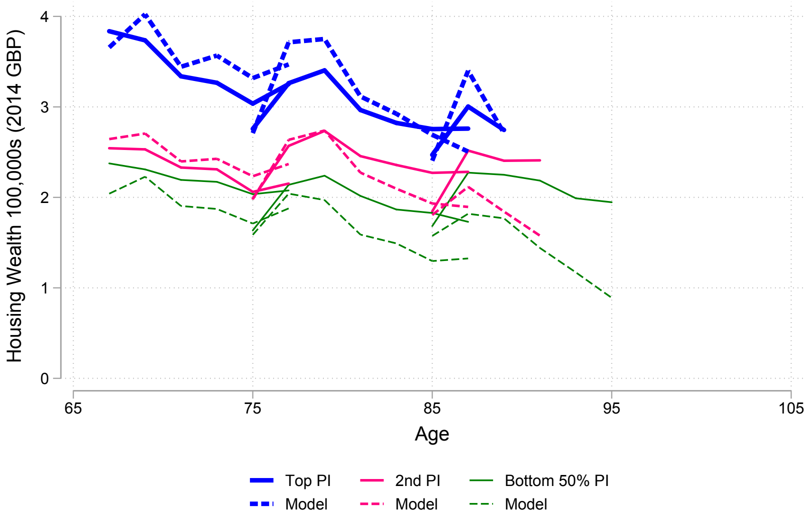
each of the five types may be rationalized solely by differences in their observables (for instance expected longevity and liquidity of their portfolio) and the estimation approach in this paper imposes no *a priori* restrictions on the differences in bequest motives or their relative magnitudes.

2.5.5 Econometric Concerns

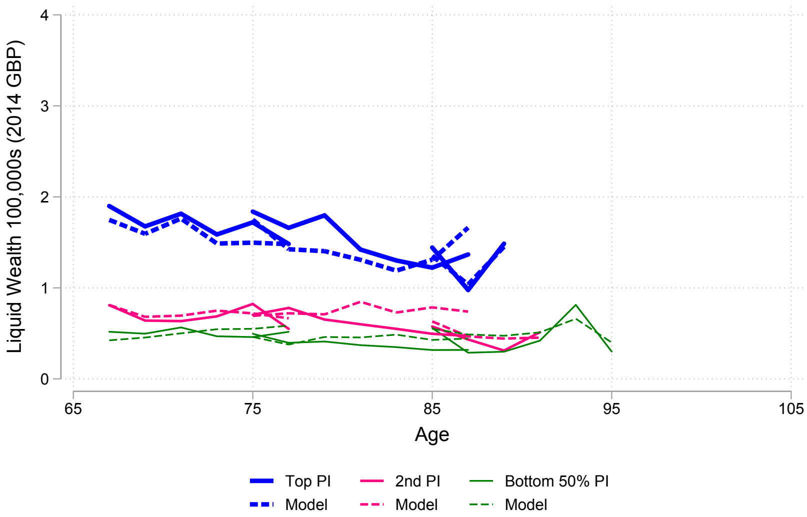
When simulating the model there are two important sources of non-stationarity: cohort effects and time effects. I now describe the approach to mitigate the effects of non-stationarity concerns.

Different cohorts have been exposed to differential income growth, asset prices and asset growth. Consequently their wealth holdings at the same age can differ substantially - in other words cross sectional moments may attribute differences between cohorts to differences in the savings rates of households and imply substantially different estimated preference parameters. As outlined above, aggregate house prices vary substantially across time and households in the data are exposed to different policy regimes at different points in time. By sampling household initial conditions, controlling flexibly for household permanent income, and simulating the sequence of observed aggregate shocks and policy reforms faced in their retirement I replicate differences across cohorts and time periods. I then construct moment conditions by cohort to eliminate the source of bias in parameter estimates. Formally, this paper makes two assumptions to address the age-time-cohort problem: that cohorts vary in their composition and initial characteristics (but not the other features of their economic environment) and, second, that relevant time effects are captured by the policy reforms and changes to the aggregate house prices. This is a structural approach to the age-time-cohort problem that explicitly accounts for differences across households and leverages policy reforms for identification - for a related semi-structural approach that purges data of age and cohort effects see Schulhofer-Wohl (2018).

A related problem is that household mortality is negatively correlated with lifetime income which means that surviving members of a cohort have higher wealth. To address “mortality bias” (and an analogous problem of sample attrition) in the



(a) Mean Housing Wealth



(b) Mean Liquid Wealth

Figure 2.8: Model Fit - Wealth Profiles (Initial Owners)

simulations each household is given the sequence of mortality and attrition shocks that are observed in the data household from which their initial conditions are drawn. The selection in the simulated unbalanced panel mirrors that in the data following the approach suggested in De Nardi et al. (2010).

2.6 Estimation Results

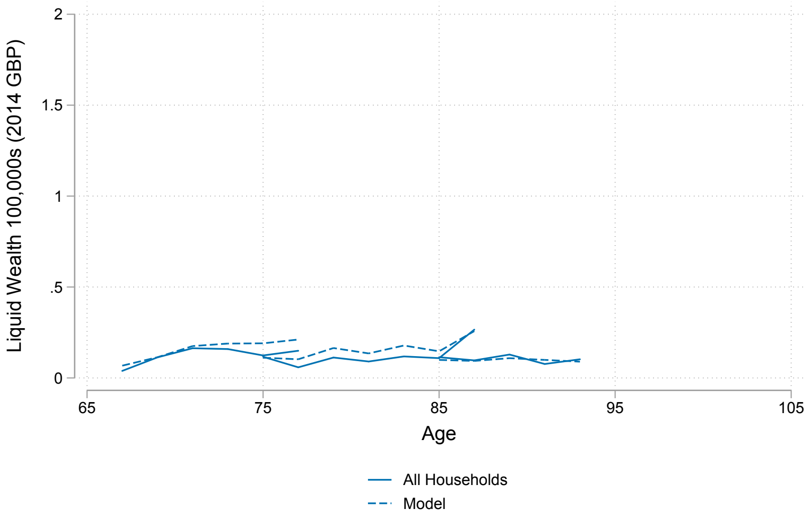
Results from the second stage estimation and their standard errors are provided in Table 6. Figures 2.8 - 2.10 display the corresponding data and simulated moments. For clarity I display only an alternating subset of the birth cohorts in each graph. I validate the model against moments that were excluded from the estimator and present the additional birth cohorts in appendix A.11. The model is able to capture key features of the data well across birth cohorts and permanent income levels. It correctly predicts the sharp increase in housing wealth and gradual decline throughout the sample. It also matches the relative magnitudes of housing and liquid wealth. Furthermore, it is able to match the permanent income gradient in wealth and the differences in wealth levels between owners and renters. The liquid wealth levels of initial renters in the model are very close to their data counterparts. The model underpredicts the housing wealth of lower income owners while modestly overpredicting their liquid wealth holdings. This is because it understates the persistence of the increase in housing wealth for the lower parts of the permanent income distribution.

The model endogenously generates the illiquidity of housing, producing a moving rate that approximately matches the infrequent decision to move home observed in the data. Despite the higher moving rate, on average households move only once during their retirement.

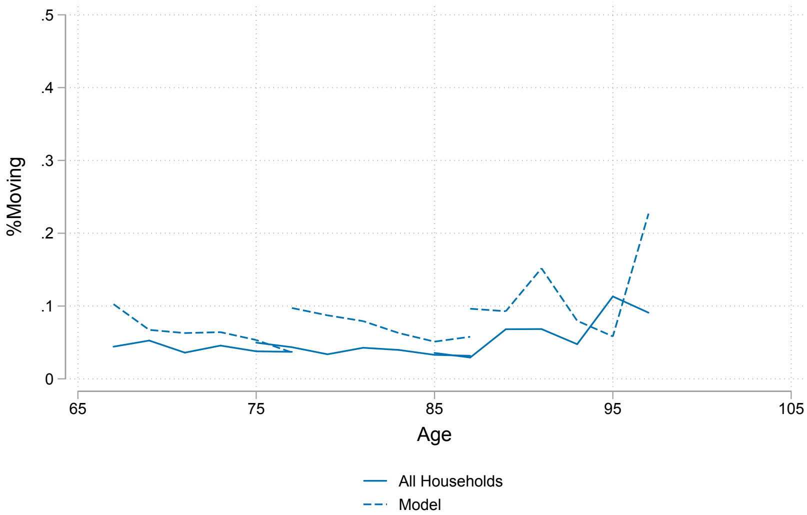
Finally, Figure 2.10 plots the data and simulated profiles for the subjective bequest probabilities. The simulated profiles match the permanent income and homeownership gradient while also matching differences across cohorts and the within cohort gradient. If anything, on average the model slightly overpredicts the probability of leaving a bequest in the future.[45]

The first panel of Table 6 reports the estimated parameters which are assumed to be common across households. The value of the time discount factor, β, on an

[45]Note that this feature of the simulated data helps explain why the fit of the model does not increase when moving costs are increased. While this would improve the model fit for the moving rate and some of the wealth moments for owners, the induced increase in non-liquid housing wealth would increase the probability of leaving a bequest and increase the error between these simulated and data moments.

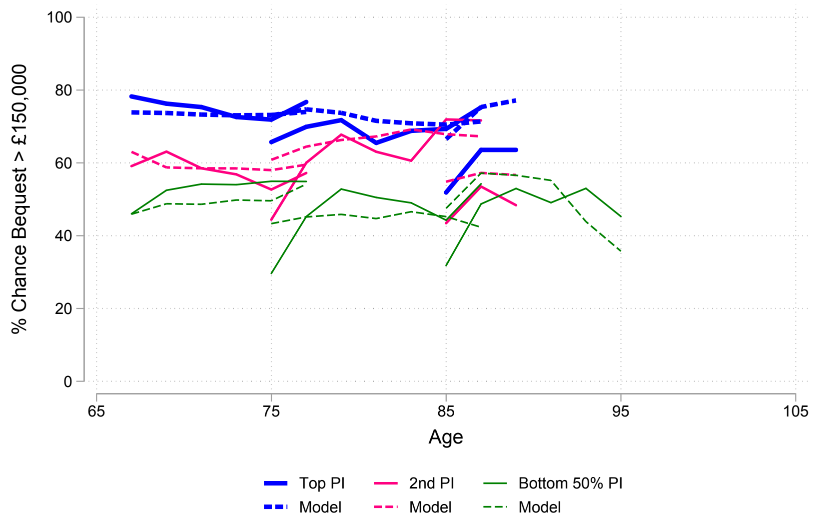


(a) Mean Liquid Wealth (Initial Renters)

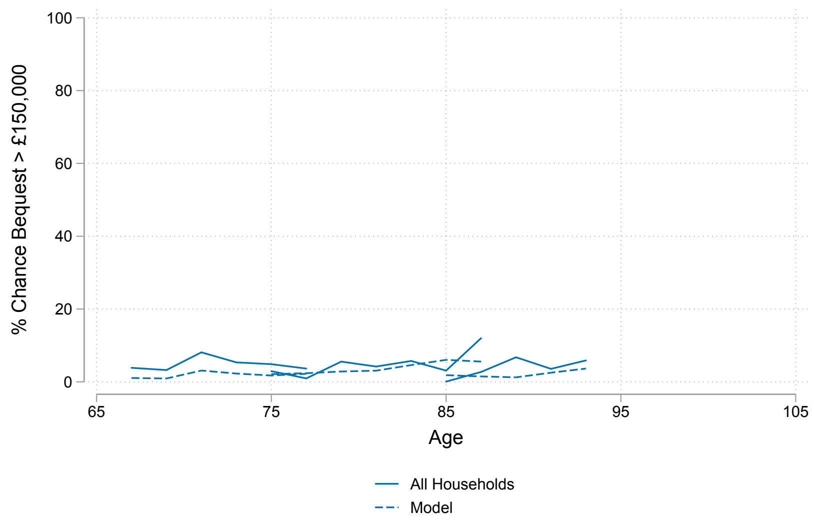


(b) Moves in the last 2 years

Figure 2.9: Model Fit - Home Moving Rates and Renters



(a) Initial Owners



(b) Initial Renters

Figure 2.10: Model Fit - Subjective Bequest Probabilities

Common Parameters			
	Preferences		
β	γ	σ	
Time Discount Factor	CRRA	Consumption Weight	
0.999	3.82	0.636	
(0.007)	(0.0475)	(0.0083)	
	Transaction Costs		
π	F_0	F_1	F_2
Proportional	Fixed (Age Polynomial)		
0.0926	12,176	27.1	33.7
(0.0032)	(214)	(1.73)	(1.35)

Bequest Parameters by Type					
Type 1		*Type 2*		*Type 3*	
Φ_1^1	Φ_2^1	Φ_1^2	Φ_2^2	Φ_1^3	Φ_2^3
Weight	Shifter	Weight	Shifter	Weight	Shifter
0.173	299,727	0.0118	100,708	0.0255	12,997
(18.23)	(2.54E+08)	(19.85)	(1.77E+07)	(0.0037)	(2,928)
Type 4				*Type 5*	
Φ_1^4	Φ_2^4			Φ_1^5	Φ_2^5
Weight	Shifter			Weight	Shifter
0.315	7,355			0.223	4,746
(0.0286)	(366)			(0.0284)	(451)

Standard Errors are in parenthesis. These are calculated using the standard formula for the asymptotic variance and correct for simulation error.

Table 2.8: Estimated Parameters

annual basis is close to 1 which is higher than many life cycle estimates. However, the retired households I model in this framework face substantial mortality risk in each period which implies a much lower effective discount factor.

The coefficient of relative risk aversion, γ, is in line with typical life cycle estimates. Together with estimated weight on non-housing consumption, σ, it implies an intertemporal elasticity of substitution for consumption of 0.36 which falls within the range of mean intertemporal elasticity of substitution in a recent meta-analysis of 169 published articles (Havránek, 2015, which also corrects for publication bias). Although estimated values for retired households are typically lower than working age households, the weight on non-housing consumption is lower than values in excess of 0.75 used in the literature on housing decisions (see, for example, Berger et al., 2018; Nakajima and Telyukova, 2017, 2018a). The value estimated in this paper implies the housing expenditure share for renters is 0.364 which almost exactly reproduces the mean housing expenditure share for retired renters of 0.341 derived from expenditure data.[46]

The second row of Table 6 report the coefficients for the proportional transaction cost and the age polynomial for the fixed cost. Taken together, these coefficients imply that housing assets have substantial adjustment costs and are broadly in line with values estimated in other lifecycle settings. Typically, proportional transaction costs on housing π are calibrated or estimated between 5% to 6% (e.g. Bajari et al., 2013), however, Cocco (2013) argues that the total costs often reach between 8% and 10%. The fixed cost estimates imply little age variation in the fixed costs and reflect additional financial costs such as realtor fees, solicitors fees, property surveys or the hiring of movers that are not modelled explicitly as well as hassle costs expressed in their financial value.

I estimate housing adjustment costs without also estimating a multiplicative utility premium for homeowners[47] which generates incentives to remain in both

[46]This is calculated from the Living Costs and Food Survey (the national UK expenditure survey) using over 65 households in England during the same sample period. I use the sum of non-durable expenditures and housing expenditures to calculate total consumption. Using total expenditure instead gives a share of 0.314

[47]This is a feature in a number of papers including Cocco and Lopes (2019); Nakajima and Telyukova (2018a, 2017); Bajari et al. (2013). An estimated rental cost which has equivalent impli-

owner occupied housing and a larger home by distorting the marginal rate of substitution between housing and non-housing consumption. Instead, the estimated transaction costs produce these incentives. While these costs would be smaller in a model with a homeownership utility premium the size of the trade-offs facing households would remain the same - it is the magnitude of these trade-offs rather than their composition that is key for the quantitative results in the paper.

The lower panel displays the type specific estimated bequest parameters. Consider first the heterogeneity in the weight of the bequest motive (ϕ_1). Consistent with Venti and Wise (1998), these results suggest that there is an association between initial retirement wealth and the strength of a household's bequest motive. Conditional on PI those with lower levels of retirement wealth (i.e Type 1 instead of Type 4 and Type 2 instead of Type 5) have weaker estimated bequest motives.[48] For Types 3- 5, bequests are modest luxuries. The next section provides additional discussion of the estimated bequest motives.

2.6.1 Validation against Quasi-Experimental Evidence

Examining moments of the data that were not explicitly targeted in estimation provides a test of the model's goodness of fit. In this subsection, I focus on the quasi-experimental effect of housing transaction taxes on home mobility estimated using a regression discontinuity design in section 2.3. To compare the model and the data I estimate an identical equation in the model and the data. I report these results in Table 2.9

To compare the model and the data I rescale the treatment effect by average mobility at the threshold because, as discussed above, the model produces higher level of mobility in the elderly population than in the data. Consequently, Table 2.9 displays the relative change in mobility as a household crosses the £250,000 (treatment effects without normalization are reported in appendix A.11). Comparing the

cations is also used in Berger et al. (2018); Etheridge (2017)

[48]For Type 1 and Type 2 households the standard errors on both bequest parameter estimates suggest that they are not identified. The next section shows that for these households the combined effect of these parameters is to generate zero effective bequest motive and, consequently, the point estimate of each individual parameter is uninformative. Conditional on no bequest motive these parameters are not separately identified. The absence of the bequest motive is the result of the estimation and it is not imposed ex-ante.

	Band around cut-off				
Order of polynomial	10%	15%	20%	25%	30%
Data					
Linear	-0.912**	-0.973***	-0.4251*	-0.411*	-0.512**
	(0.372)	(0.329)	(0.234)	(0.237)	(0.225)
Quadratic	-0.748	-0.923**	-0.553**	-0.446*	-0.512**
	(0.494)	(0.396)	(0.272)	(0.266)	(0.230)
Model					
Linear	-0.522	-0.394	-0.347	-0.402	-0.410
Quadratic	-0.556	-0.404	-0.312	-0.399	-0.410

All regressions additionally control for wave fixed effects, a polynomial in age, household demographics, a polynomial in permanent income and region dummies. Following Kolesár and Rothe (2018), Standard Errors are clustered by household. * $p < 0.10$, ** $p < 0.5$, *** $p < 0.01$

Table 2.9: The Relative Effect of Transaction Taxes on Household Mobility: Model and Data

upper panel with the lower panel shows that across the linear and quadratic specifications, the model generates a reduction in mobility of 40% which is economically and statistically comparable to the responses in the data.

One of the key trade-offs in the model is the households' willingness to transform housing wealth into liquid assets. In particular, understanding the implications of changes of house prices and the demand for public insurance hinges on the magnitude of this trade-off. Housing transaction volumes are known to be pro-cyclical (See Ortalo-Magné and Rady, 2004) and the size of the extensive and intensive margin responses to changes in housing wealth are a crucial part of the model implications discussed in the next section. Using quasi-experimental evidence to validate model responses to changes in the transaction cost and changes in the implicit cost of home equity withdrawal demonstrate that these responses are quantitatively important and realistic.

2.7 Model Implications

The previous section discusses values of the estimated parameters, however, in a large model it is often difficult to interpret the size of these parameters. This section attempts to address this issue by examining the bequest motives and quantifying the mechanisms that drive the savings of the elderly.

2.7.1 Interpreting the Size of the Bequest Motive

It is widely accepted that estimated parameters for the weight and curvature of the bequest function are difficult to interpret and, when the weight and curvature vary, lead to ambiguous ranking of the role of bequest motives across studies. Instead, following De Nardi et al. (2010), it has become standard to express these numbers in terms of a static allocation problem. Suppose a single individual knew they were going to die next period and faced no further uncertainty, then how much would they consume today and how much would be left as a bequest? When bequests are luxuries, the level at which they become operable can be loosely interpreted as the annuity value of consumption below which households have zero marginal propensity to bequeath. Figure 2.11 reports the bequest share in this thought experiment for different values of total wealth.[49]

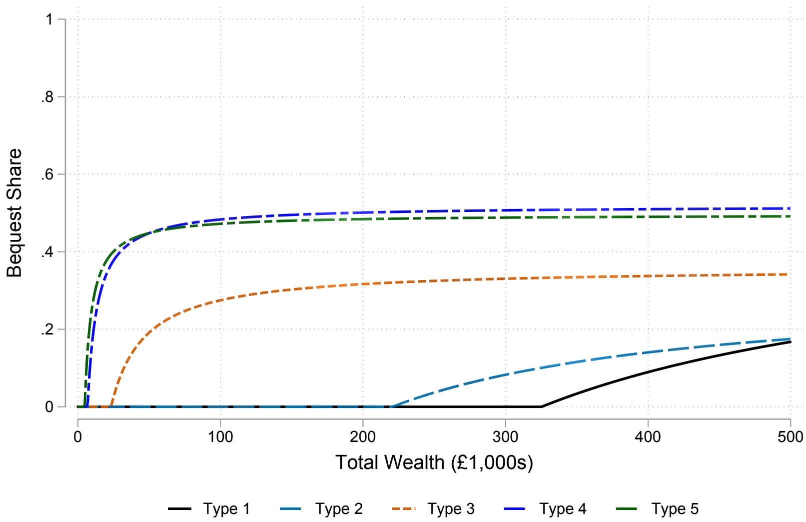
The top panel shows bequest allocations for each of the estimated types. The bottom panel calculates the implied bequest share for a variety of related studies which estimate bequest motives in retirement. Beginning with the estimated parameters, Type 1 (who have a high weight, but the highest curvature) and Type 2 have effectively 0 estimated bequest motives. Indeed, they would need to achieve an annuity value of consumption above £300,000 and £200,000 respectively before the bequest motive is active. This closely corresponds to the latent type in Kopczuk and Lupton (2007) who are assumed to have no bequest motive although the fraction of households without a bequest motive is larger due to additional features (such as housing and long term care risk) which rationalize the savings decisions of older households.[50]

Turning to Type 4 who have the largest weight. This type has the strongest bequest motive, both in terms of the allocated share, which approaches 52%, and that it is operative for much of the wealth distribution. The estimated weight for

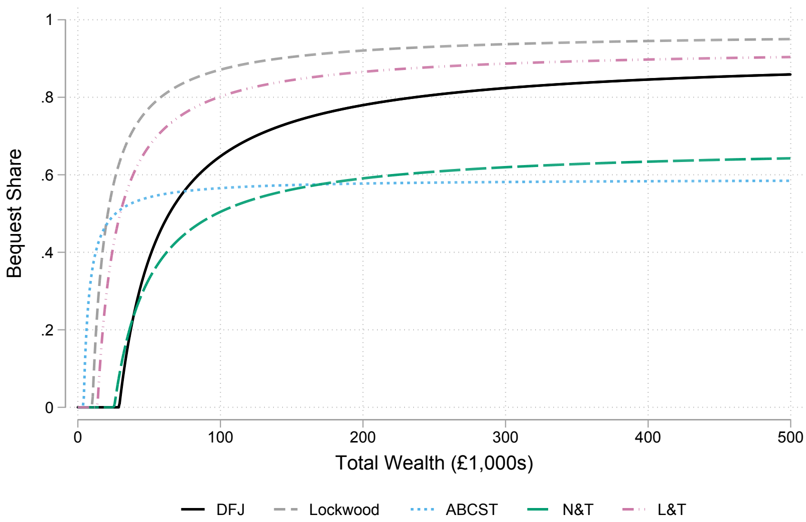
---

[49] Full details of this calculation and how studies are made comparable are given in appendix A.10

[50] Interpreting these allocations through the lens of the annuity value of consumption implies extrapolating to large values beyond the support of the type specific distribution.



(a) Bequest Allocation (Estimates)



(b) Bequest Allocation (Literature)

Figure 2.11: Bequest Allocation

Types 1-4 in panel (a) refer to preference types estimated in section 2.6 and allocations are calculated with estimated preference parameters reported above. In panel (b) the same statistics are calculated using results from De Nardi et al. (2010) (DFJ), Lockwood (2018), Ameriks et al. (2018) (ABCST), Nakajima and Telyukova (2018a) (N&T), and Lee and Tan (2017) (L&T). Further details of the calculation provided in Appendix A.10.

Type 5 households are slightly weaker implying small differences in the degree of luxuriousness and allocations. However, in total the estimates imply very similar allocations among those making positive bequests.

In contrast, Type 3 has more modest bequest motives with asymptotic marginal propensity to bequeath of 35%. Conditional on their high level of retirement wealth, the model is able to explain the continued wealth holdings of the richest households through observable differences in the state variables of these households rather than preferences. It is important to stress that there is considerable overlap in the support of household wealth between types which means that the heterogeneity in bequest motives is not only capturing an underlying luxury good.

Comparing these results with the results in the right hand panel reveals three important differences from existing estimates in the literature. First, the bequest shares implied by my estimates are more conservative. The closest estimates in the literature are Nakajima and Telyukova (2018a), who model the homeownership decision, and Ameriks et al. (2018) who model the financial wealth of a wealthier population and match strategic survey responses. Taken together, this suggests that other estimates may capture either the illiquid nature of housing or its consumption flow in their estimates of the bequest motive. Second, the estimates for a heterogeneous bequest motive suggest a degree of variation across households that is as large as the variation in homogeneous bequests produced by different estimates.

Third, despite allowing the degree to which bequest motives are luxuries to vary across households the estimated bequest shares show less curvature within type relative to the literature.[51] It is possible that differences in the portfolio of households or heterogeneity in the strength of the bequest motive (that are correlated with wealth) are no longer proxied by the degree to which bequests are a luxury. An alternative explanation for the lack of curvature is an issue of common support. For example, while Types 1 and 2 contains many renters, the households in Type 4 and 5 are typically richer. Likewise very few households in Type 1 are rich enough to make a positive bequest allocation. Without variation across the entire support of

[51] While the presence of dispersion in bequest motives echoes findings in Kopczuk and Lupton (2007), the functional form here is more flexible

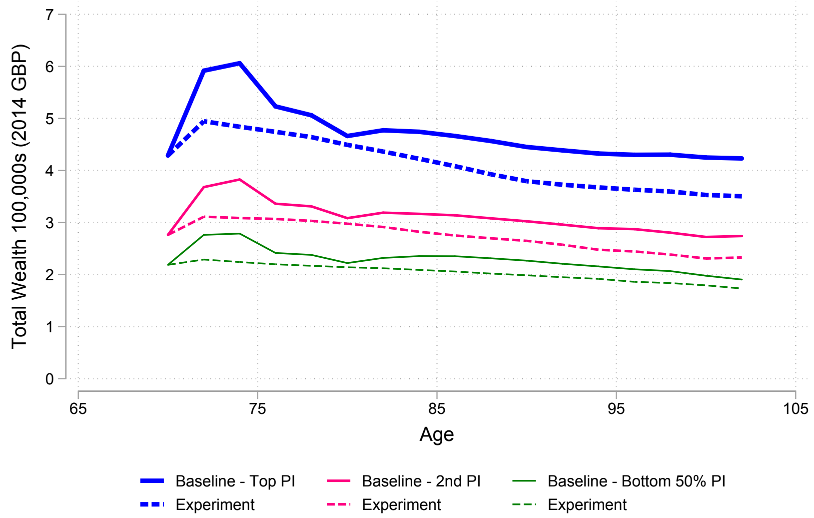


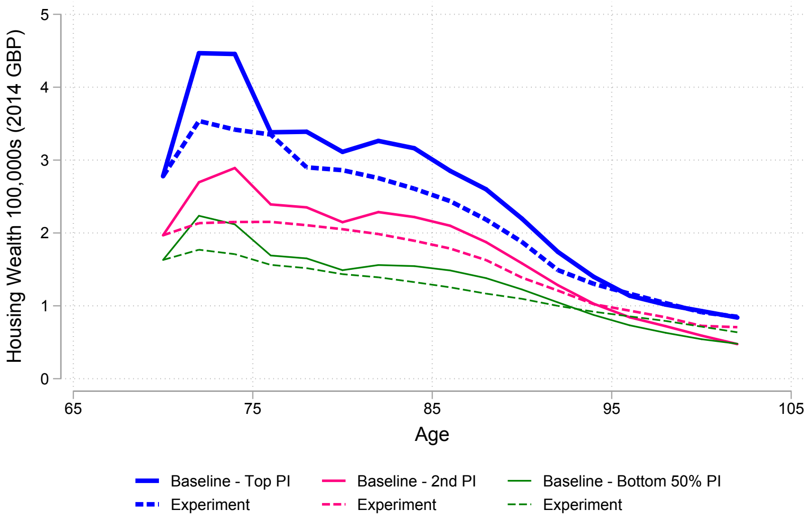
Figure 2.12: Experiment 1- Total Wealth

the wealth distribution it is hard to precisely identify the extent to which bequests are a luxury among households.

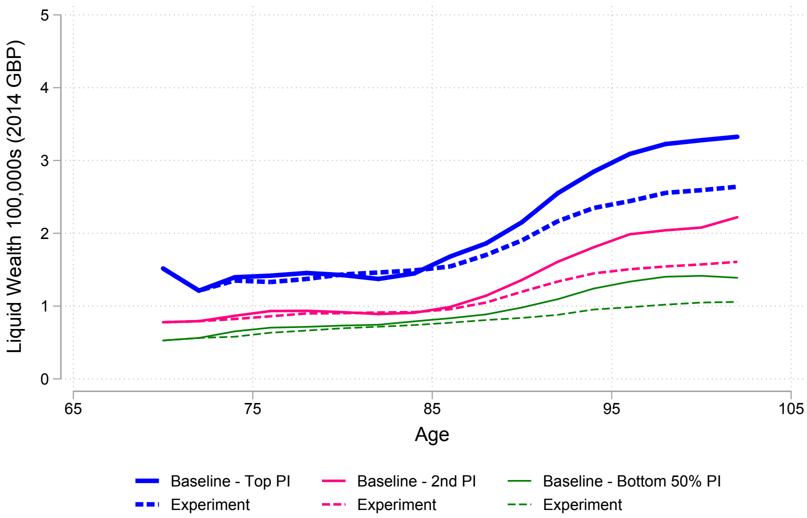
2.7.2 Decomposing the Role of Housing

To determine the quantitative importance of housing in retirees savings, I use the estimated parameters and change one feature of the model at a time. For each of these different environments, I compute the new household policy functions, simulate the model and compare the resulting asset accumulation profiles to the asset profiles generated by the baseline model. I display asset profiles for households who are age 68 in the first wave of ELSA. Throughout, I focus on the wealth of initial homeowners because the wealth of renters is negligible. Appendix A.12 provides additional decomposition exercises.

First, I fix house prices at their 2002 level, but hold household expectations constant. Figure 2.12 plots the simulated profiles of total wealth for initial owners. Compared to the wealth profiles in the baseline simulated economy, total household wealth decreases at all ages. The effect is largest at younger ages where the baseline profiles include the rapid house price appreciation of the early 2000s, but also has long lasting effects at older ages where the cumulative effect of downsizing behaviour is largest- on average total wealth at age 96 falls by 15%. Figure 2.13



(a) Mean Housing Wealth



(b) Mean Liquid Wealth

Figure 2.13: Experiment 1- Portfolio

breaks the total wealth into its two components.

Housing wealth decreases when house prices are held constant (top panel). This is due to the mechanical effect of eliminating house prices as well as the behavioural response as simulated households re-optimize, however these two effects almost cancel out by late into their retirement. The bottom panel shows the corresponding effect on their liquid wealth. Households continue to maintain large levels of wealth as buffers against future shocks. The decline in house prices makes these

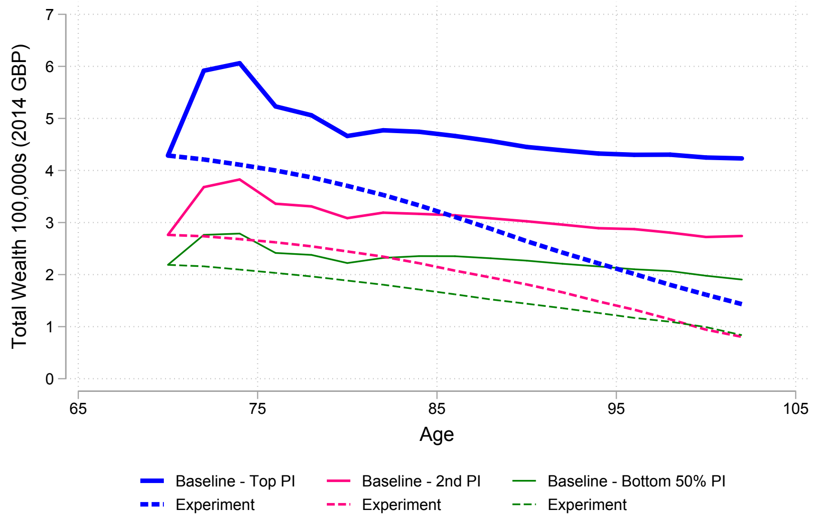


Figure 2.14: Experiment 2-Total Wealth

households poorer and they substitute from consumption to liquid savings to offset the wealth effect of decreased house prices. This explains why liquid wealth is almost constant through the start of the sample. Averaged across the households' remaining life span liquid, wealth balances decrease due to the effect at older ages (7.5%) while housing wealth declines by a larger fraction (12.5%). Relative to the baseline, when households don't experience periods of house price appreciation, they reduce the frequency with which they move home (by 25%) because there is less equity in their home to cash out. This transmits into a reduction in consumption by around 5% over their remaining lifetime, but additionally a reduction in bequests of over 10%.

Next, to fully understand the role of housing, I eliminate the remaining features of housing that make it different from liquid wealth: the difference in returns, the consumption flow, the illiquidity and different exposure to long term care costs. This is equivalent to a single asset model and the total wealth profile of households shown in Figure 2.14.

Eliminating house price fluctuation has a significant effect on the savings behaviour of retirees, but eliminating the remaining features of housing assets has an even larger effect. In contrast to the baseline, households deaccumulate the majority of their wealth throughout retirement. Averaging over the periods a household is alive, total wealth decreases by 28% while the size of the average bequest left decreases by almost 40% and the effect on those who survive past 95 is even larger. The results of this experiment show that the portfolio composition and the differences between asset classes have a first order effect on the levels and age profiles of household savings as well as the bequests they leave. Understanding and accounting for differences in asset classes is important in evaluating policy proposals that affect the elderly.

2.7.3 Responses to Unanticipated Shocks

The previous experiments decompose savings behaviour, highlighting housing wealth and how illiquid assets affect the trajectory of wealth in retirement. To further understand these important mechanisms and drivers of saving, I turn to an additional set of experiments which simulate household responses to an unanticipated increase in either their after tax income or an increase in the aggregate house price they face.

The income shock is equivalent to a tax rebate (See e.g. Parker et al., 2013; Kaplan and Violante, 2014; Misra and Surico, 2014) targeted at age 70 households and generates a one-time 10% increase in their after tax income. The house price shock also occurs at age 70 and raises the level of house prices by 10%. However, after age 70 the future house prices follow the AR(1) process describe in equation 2.14 and thus the effect of the unanticipated shock is persistent.

To understand how households respond to changes in their portfolio and total wealth, I report two measures: the aggregate Marginal Propensity to Consume (MPC) for annual non-housing consumption as well as the aggregate Marginal Propensity to Bequeath (MPB).[52] The MPC is measured in the period the shock arrives and captures the contemporaneous non-housing consumption response to these changes. The MPB is measured at death and provides a summary statistic of how this wealth is used over their remaining life cycle. This includes financing

[52] I construct the aggregate measures by computing a household level measure and explicitly aggregating households in the model. The results I present here depend on the distribution of state variables in this cohort. The average MPBs reported below also integrate over uncertainty in their remaining lifetime including mortality risk, house price changes and medical expense risk.

Shock	Marginal Propensity to Consume	Bequeath
Income	0.151	0.71
House Price	0.029	0.38

Simulated responses for a single birth cohort to a one-time 10% increase in income and a one-time 10% increase in house prices. In both simulations the shock arrives at age 70 and the annual MPC is measured contemporaneously.

Table 2.10: Household Responses to Unanticipated Shocks

other expenses such as housing adjustment costs and Long Term Care expenses. Table 2.10 reports the average Marginal Propensity to Consume (MPC) as well as the Marginal Propensity to Bequeath (MPB).

Turning first to the MPCs. At the arrival of the shock, the contemporaneous MPC out of a transitory income shock is larger than the housing wealth shock and both estimates are within the range of estimates in the respective literatures.[53] They imply that for an additional £1 of wealth at age 70, an age 70 household consumes an additional 15 pence when they experience an income shock and 3 pence when they experience an increase in house prices. Furthermore, the MPB out of the two shocks are very different. Almost twice as much of the one time income shock is transmitted to bequests[54] than the house price shock.

Households respond differently to the two shocks over time - particularly those who are liquidity constrained (or likely to be during their remaining lifetime). In response to the income shock, low wealth households with large housing portfolio shares are less likely to downsize. These households no longer liquidate housing wealth and no longer pay the cost of adjustment because the income shock alleviates liquidity constraints in some states of the world. This effectively increases their lifetime saving.

In contrast, an unanticipated increase in housing wealth driven by price appre-

[53]For example, Aladangady (2017) finds an MPC out of housing wealth of 0.047 on the dollar.

[54]The order of magnitude is consistent with estimates of the MPB out of social security income reported in Lee and Tan (2017) and substantially larger than the MPB estimated in Altonji and Villanueva (2003)

ciation actually reduces the savings of these same households over their remaining lifespan. These households increase the frequency with which they move home to access otherwise trapped home equity because the financial return to downsizing has grown (the utility effect of changing housing consumption is constant). On average there is a similar change in their available liquid resources (cash on hand) under both experiments, but in the case of the house price shock this is driven by an increase in the intensive and extensive margins of downsizing. The key driver of the difference in MPBs is the response of the constrained households and how they trade off housing for liquidity.[55]

It is important to emphasize that these results do not suggest that houses are less likely to be bequeathed than financial wealth nor that homeowners are less likely to leave a bequest.

Finally, I explore how responses differ across household preference types. To isolate the effect of preferences, I resolve and simulate the model shutting down the correlation between types and initial conditions. This means mortality, health and medical expense uncertainty is held constant across the types as well as the wealth, permanent income and portfolio composition. Table 2.11 reports the results of this experiment for the income shock. Results for the house price shock and reintroducing the correlation between initial conditions and preferences are presented in Appendix A.13.

Focussing on the variation in MPCs across preference type (to either shock): there is is a correlation with preference type and realized bequests. Recall that Types 1 and 2 have essentially no estimated bequest motive while Types 4 and 5 have the strongest. In addition to heterogeneity in household constraints, differences in preferences generate substantial difference in household responses. Finally, it is important to summarize the evidence of the strength of the bequest motive and it's heterogeneity. That Types 1 and 2 still bequeath a positive share (on average) from these unanticipated shocks suggests that a significant proportion of bequests occur

[55] A second mechanical effect also explains differences in household response. house prices are *persistent* not *permanent*: the mechanical effect of the house price shock at age 70 on average dies out of the households remaining life. This means that even if they do not adjust their behaviour in response to the house price shock, its transmission to bequests will be smaller.

Shock		Marginal Propensity to Consume	Marginal Propensity to Bequeath
Income	Type 1	0.295	0.526
	Type 2	0.294	0.526
	Type 3	0.274	0.566
	Type 4	0.076	0.785
	Type 5	0.043	0.838

Simulated responses for a single birth cohort to a one-time 10% increase in income. The shock arrives at age 70 and the MPC is measured contemporaneously. Preference parameters for each type are taken from the estimation results above. To separate the role of preferences the correlation between preference type and initial conditions is set to 0.

Table 2.11: Household Responses to Unanticipated Shocks by Preference Type

in the absence of bequest motives. This is consistent with the effect of eliminating bequest motive in the previous experiments. However, bequest motives have an important effect on portfolio choice and generate substantially different pass through of income shocks to bequests. This is inconsistent with the view that bequests are entirely accidental. Instead, it supports the argument that many bequests are *incidental*: households with bequest motives value the large bequests that arise incidentally from self-insuring late-life risks by holding stocks of liquid wealth and maintaining large housing wealth positions.

2.8 Valuing the Means Testing of Long Term Care Programs

Many countries provide means tested benefits for households who have large long term care expenditures, but limited private resources. Perhaps the most prominent of these programs is Medicaid in the US, which shares a number of institutional features with its UK counterpart. In the UK and the US, means testing occurs at the individual level. The asset component of means testing includes the value of a primary residence if they are single and move into a residential facility, but excludes their home when they are in a couple.[56] Similarly, for those in a couple only half of their assets must be used to finance the long term care needs of an individual before state assistance is provided. In practice this provides insurance for a spouse against the risk associated with their partner and also creates differences in the extent of social insurance for Long Term Care needs across the wealth and income distribution as well as between couples and singles.

In this section, I simulate changes to the generosity and design of this means testing. I compare the resulting increases (or decreases) to the resulting gains (losses) in consumer welfare. To measure the costs associated with these reforms I compute the present discounted value of payment changes (including implicit changes covered by disregarded assets) and assume it costs the government £1 to provide £1 of payments.[57] I compare the compensating variation under the alternative policy environment with the actuarial value of the alternative policy. I consider three reforms (in the current version of the paper these reforms are not revenue neutral): first, I increase the value of transfers for those receiving the ADL consumption floor, second, I eliminate only the financial wealth disregard that applies to couples, and, third, I eliminate the financial wealth and housing asset disregards that apply to couples. In the baseline these disregards apply at 100% of the housing and 50% of the financial wealth of couples where only one member has ADL needs.[58]

The compensating variation to each reform is the immediate cash on hand payment that would leave the retiree as well off as before the reform. This is an exante measure that is forward looking - it incorporates mechanical effects and the behavioural responses to the reform. Specifically, this is computed at age 68 and

[56] Strictly, in both cases houses are countable assets (in the language of Medicaid) while "homes" are excluded. Although there is some variation between US states, a house continues to qualify as a "home" if either a spouse (community spouses in the language of Medicaid) or dependent relative continues to live there or a nursing home stay is deemed temporary with intent to return.

[57] This benchmark assumes the government does not use other methods to make transfers more or less attractive to potential claimants.

[58] In the model the 50% rule applies each period to avoid keeping track of an additional state variable rather than as an effective lifetime cap. This potentially understates the value of the financial disregard. However, for most households the effect is negligible.

	Increase in PDV of payments			Compensating	Ratio
	Total	Transfers	Disregard	Variation	CV/PDV
Initial Renters	389	371	18	-565	1.46
Initial Owners	157	109	48	-540	3.44
Initial Owners					
Top PI quartile	56	22	34	-170	3.3
2nd PI quartile	137	94	43	-793	5.80
Bottom 50% PI	244	182	62	-639	2.62
Single Men	92	92	n/a	-173	1.89
Single Women	428	42	n/a	-97	2.30
Couples	369	288	82	-1044	2.83
Health					
Good	126	94	33	-310	2.45
Bad	227	187	41	-899	3.95
ADL	568	493	76	-1360	2.39

Note: Columns (1)-(3): £increase in the present discounted value of government transfers as of age 68. Column (4): £value of transfer needed to compensate people for the expansion of the consumption floor. Column (5): Ratio of columns (4 and 3)

Table 2.12: The Costs and Benefits of Increasing the ADL consumption floor by 10%

defined as $\chi_{68} = \chi^i_{68}(f_{68}, I, m_{68}, h_{68}, x_{68}, p_h)$ where:

$$\chi_{68} = V^i_{68}(f_{68}, I, m_{68}, h_{68}, x_{68}, p_h|Base) = V^i_{68}(f_{68}, I, m_{68}, h_{68}, x_{68} + \chi_{68}, p_h|Reform) \tag{2.28}$$

Where $V^i_{68}(\cdot)$ is the age 68 value function computed for a given set of state variables. When reporting results by group I average across all members and define groups by initial status.

To understand the insurance provided by the consumption floor I first analyse a 10% increase in this floor. This corresponds to an increase in the consumption floor from £4,956 (£7,434) to £5,452 (£8,177) per year for singles (couples). Table 2.12 reports the results from this exercise.

The discounted present value of total payments is reported in the first column and columns (2) and (3) separate this into direct transfers and implicit payments through disregards. Payments increases for initial owners and initial renters. The total increase in present discounted value of payments is relatively small because the onset of ADL conditions is typically some time after age 68 and a low probability event and, although there is some change to implicit payments through disregards, the fiscal burden is primarily driven by an increase in transfer receipt. To understand how this is driven by the income and wealth gradient among households, I decompose the sample by permanent income for those who are initial owners. On average, initial owners receive a smaller increase in transfers which also decrease in the level of their lifetime income. Separating the population by their family status indicates that the majority of these payments are made to those in couples or who are single women at age 68.

The fourth column presents the compensating variation and the final column presents the ratio of compensating variation to the change in total payments. Despite differences in the present discounted value of payments, the compensating variation between renters and owners is similar. However, the final column reveals that although the poorest households receive the largest increase in payments they also have a lower private value per £1 that is spent. Across all households the ratio of the compensating variation and payments is greater than 1- demonstrating that the consumption floor provides valuable insurance to retired households. Both the compensating variation and the ratio increase in income, but are not monotonic. High income households have higher lifetime levels of consumption. Consequently, large nursing home expenses precipitate a larger drop in consumption. Second, although they are less likely to have ADL needs at the beginning of the sample, they face higher life expectancy and are more likely to have high long term care costs when they survive long into retirement.[59]

However, for those at the very top of the distribution the value falls. The total value of these households' portfolio is almost twice as high as the next PI group which makes them less likely to qualify for the consumption floor and reduces the ratio of compensating variation to payments. Furthermore, they are more likely to be in a couple and have assets covered by the couple specific disregard. Despite

[59]De Nardi et al. (2016a) present a similar argument when evaluating Medicaid expansion for singles.

	Increase in PDV of payments			Compensating	Ratio
	Total	Transfers	Disregard	Variation	CV/PDV
Initial Renters	-342	576	-918	595	1.74
Initial Owners	-852	159	-1012	5678	6.66
Initial Owners					
Top PI quartile	-921	49	-970	9225	10.02
2nd PI quartile	-1012	157	-1169	6308	6.23
Bottom 50% PI	-730	216	-946	3107	4.23
Single Men	n/a	n/a	n/a	n/a	n/a
Single Women	n/a	n/a	n/a	n/a	n/a
Couples	-1474	528	-2002	9040	6.13
Health					
Good	-491	88	-579	4169	8.488
Bad	-620	346	-966	4441	7.16
ADL	-1938	863	-2801	5108	2.64

Note: Columns (1)-(3): £increase in the present discounted value of government transfers as of age 68. Column (4): £value of transfer needed to compensate people for the elimination of the disregards. Column (5): Ratio of columns (4 and 3)

Table 2.13: The Costs and Benefits of Eliminating the Financial Asset Disregard for Couples

the large benefits for richer households, it is worth emphasising that this does not account for the financing of this expansion: under a progressive tax system these households who have higher incomes are likely to bear a greater share of the costs.

Table 2.13 presents results from eliminating only the disregards for the financial assets of couples while maintaining the housing disregard. By construction, this reform does not affect singles. The changes in the total value of government transfers are much larger than the reform to the consumption floor. While the reduction in payments is driven by a reduction in the disregarded assets, for many households it is offset by a substantial increase in payments. In contrast to the expansion of the consumption floor, the effects are monotonic in lifetime income. As before, those with the highest consumption experience the largest drops when exposed to large long term care costs. However, this reform substantially increases their exposure to the risk of high medical expenses associated with their spouse -

increasing the probability that ex post they rely on their housing wealth to finance future consumption. It is worth remarking on two important effects for households who experience "smaller" and "larger" long term care costs. For these "smaller" shocks, households may not find it optimal to adjust their housing stock and instead exhibit excess sensitivity in their non-housing consumption - as discussed in Chetty and Szeidl (2007) and Kaplan and Violante (2014), this magnifies the welfare costs of these shocks. In the baseline, the disregard on financial assets provides insurance against these fluctuations. Under "larger" shocks, households absorb these shocks into non-housing consumption and also housing consumption because they downsize to liquidate housing wealth. This ex-post reliance means that they pay large adjustment costs to liquidate wealth and provide a buffer for their future. This exacerbates the wealth effect from losing their financial wealth, lowering the utility received from future bequests, but also in turn exposes newly liquidated wealth to long term care costs. When it is costly to adjust housing and households place a large value on the consumption from housing this also raises the value of liquid buffers that allow them to avoid adjusting their housing stock. This intuition helps explain why the valuation of each £1 is high.

Finally, Table 2.14 presents results from eliminating both disregards for couples. For initial renters who benefit primarily from the financial component the valuations are similar. For owners the total value of government payments is substantially larger as the new policy environment affects a larger fraction of their portfolio. As outlined above, under the baseline policy many couples find saving to self insure relatively cheap. Liquid wealth is only partly exposed to these costs, but can be used to insure any future long term care risks for the spouse as well as provide self insurance against longevity risk. In the event that a household is fortunate and does not face high long term care expenses it may always be left as a bequest. At the same time they can enjoy the utility of a large house and consumption today without requiring high levels of liquidity for self insurance. This policy additionally eliminates the housing disregard and increases the exposure of households to long term care risk. Household valuations, column (4), increase substantially when

	Increase in PDV of payments			Compensating	Ratio
	Total	Transfers	Disregard	Variation	CV/PDV
Initial Renters	-347	2985	-3332	663	1.91
Initial Owners	-1750	470	-2220	7681	4.34
Initial Owners					
Top PI quartile	-1924	74	-1998	12200	6.34
2nd PI quartile	-2043	452	-2495	8526	4.18
Bottom 50% PI	-1496	701	-2197	4514	3.02
Single Men	n/a	n/a	n/a	n/a	n/a
Single Women	n/a	n/a	n/a	n/a	n/a
Couples	-2857	2190	-5047	12214	4.27
Health					
Good	-972	297	-1269	5281	5.43
Bad	-1048	1580	-2628	5789	5.52
ADL	-3786	3657	-7443	9353	2.47

Note: Columns (1)-(3): £increase in the present discounted value of government transfers as of age 68. Column (4): £value of transfer needed to compensate people for the elimination of the disregards. Column (5): Ratio of columns (4 and 3)

Table 2.14: The Costs and Benefits of Eliminating Disregards for Couples

compared to the experiment that eliminates financial disregard. However, the per £1 valuation falls because the policy is almost twice as expensive for each initial owner.

In contrast to higher wealth households and home owners, for initial renters the ratio of their private valuation to £1 spent by the government is approximately constant across experiments. While households place a high value on the state provided insurance for long term care needs (see above), the design of means testing which imposes a 100% effective tax rate on their household wealth substantially mutes these benefits for the richer households. In particular, these results suggest that households place a high value on policies that insure their liquid wealth (and help them avoid liquidity constraints), but have a lower per £1 valuation for policies that insure their housing wealth.

2.9 Conclusion

In this paper, I develop and estimate a structural model of retirement savings decisions with realistic risks, housing, and heterogeneity in bequest preferences. Combining data on wealth composition and exogenous policy changes facilitates the separate identification of different motives for holding wealth. Estimation reveals that households exhibit large differences in the weight they place on leaving wealth for future generations and this is closely correlated with the wealth they accumulate across their lifetime. Accounting for differences in preferences and the liquidity composition of households' portfolios reduces the estimated level of risk aversion and the role of the precautionary savings channel in explaining the retirement savings puzzle.

Simulating a number of counter-factual economic environments isolates the role of different model features in driving retirement savings. Housing explains a substantial fraction of the level of wealth holdings in retirement. Understanding the portfolio composition of households and how they trade-off liquidity and housing is key. Model validation shows that these mechanisms reproduce reduced form estimates identified from quasi-experimental variation. This trade-off drives differences in the response to unanticipated income and housing wealth shocks. Demand for extracting liquidity from their home is increased when house prices increase, but reduced by an unanticipated liquid wealth shocks. This has opposite effects for marginal downsizers and creates a lower aggregate marginal propensity to bequeath from house price shocks. The estimated response to shocks differ substantially from the estimates in Altonji and Villanueva (2003), suggesting that a large proportion of retirement wealth is eventually bequeathed. As suggested in Gan et al. (2015), this may mitigate some of the concerns of financing an ageing population because older generations share their good fortune with future generations.

Finally, I address the role of means testing in the tax and transfer system with a particular focus on how means testing treats different asset classes. I concentrate on publicly provided Long Term Care insurance which features substantial means testing and an exemption for housing that varies across married and single individuals in the UK and the US. When households must first exhaust their private assets, this creates an implicit 100% marginal tax rate. Furthermore, differences in asset classes mean that differences in the design of means testing program may reinforce or discourage self insurance behaviour. When households like to live in their home, this amplifies precautionary savings motives and the demand for liquidity so that, for every pound it costs the government, increasing the disregard for liquid assets provides more value than increasing the disregard for housing.

Chapter 3

Couples' and Singles' Savings After Retirement

3.1 Introduction

There is considerable agreement on the reasons why single retirees hold assets. Recent work has shown that much of their saving behavior can be explained by the possibility of expensive medical conditions at older ages (De Nardi et al., 2009, 2010), or by bequest motives that act as luxury goods (De Nardi, 2004). A number of recent studies seek to disentangle these two motives by including other features of the data, such as survey responses (Ameriks et al., 2011, 2018), long-term care insurance purchases (Lockwood, 2018), or Medicaid usage (De Nardi et al., 2016a).

Less attention has been given to the saving of retired couples. Couples tend to be richer than singles and face a more complex set of risks. To what extent do members of a couple care about their own consumption and medical expenses, as opposed to leaving bequests to their surviving spouse and other heirs? What happens to savings when one of the members of the couple dies, and why?

In the Assets and Health Dynamics of the Oldest Old (AHEAD) dataset, about 50% of individuals aged 70 or older are in a couple, while about 50% are single. Because retired couples are not only numerous, but also richer, they hold a large share of the elderly's wealth. Better understanding why they save matters, including in the assessment of policy reforms.

Couples and singles differ in important ways. Being in a couple allows its members to pool their longevity and medical expense risks, but also exposes each member to their spouse's risks, including the income loss and high medical expenses that often accompany a spouse's death (Braun et al., 2017). A couple might also care about leaving bequests to the surviving spouse, in addition to other heirs.

We start our analysis of couples by establishing facts about their retirement savings and the medical expenses and longevity risks that they face. We then develop a model of couples and singles during retirement, in a framework that incorporates observed heterogeneity in life expectancy and medical expenses, and that explicitly models means-tested social insurance. We allow for bequest motives to other heirs for both couples and singles, and for altruism towards the surviving spouse for couples. We estimate our model using the Method of Simulated Moments (MSM). Finally, we use the model to evaluate the extent to which the savings of retirees are driven by bequest motives towards the other spouse, bequest motives towards other heirs, and medical expenses.

In the first part of our analysis we document that while the savings of retired singles tend to stay roughly constant or fall over time, those of couples increase as long as both spouses are alive. We also find that assets drop sharply at the time of death of each spouse and that, by the time the second spouse dies, a large fraction of the wealth of the original couple has vanished. Depending on the specification, assets decline by around \$100,000 at the time of an individual's death. A large share of this drop, but not all of it, is explained by the high medical expenses at the time of death. For example, out of pocket medical spending plus death expenses are approximately \$20,000 during the year of death (whereas medical spending is \$8,000 per year for similarly aged people who do not die).

Our second-step estimates allow us to evaluate to what extent the risk sharing and economies of scale of a couple help insure against longevity and medical-expense risk and to what extent couples get a better or worse deal from publicly provided health insurance. They will also allow us to estimate bequest motives towards the surviving spouse, in contrast to bequests to children or others. Finally, it enables us to study in what ways the responses of a couple will differ from the responses of a single person when, for example, public health insurance becomes more limited, or its quality worsens, or some of its means-testing criteria become tighter, or the generosity of widow(er)s benefits change.

3.2 Related Literature

Much of the previous structural literature only studies singles. In a previous paper De Nardi et al. (2010) show that post-retirement medical expenses and government-provided insurance are important to explaining the saving patterns of U.S. single retirees at all income levels, including high permanent-income individuals who keep large amounts of assets until very late in life. These savings patterns are due to two important features of out-of-pocket medical expenses. First, out-of-pocket medical and nursing-home expenses can be large. Second, average medical expenditures rise very rapidly with age and permanent income. Medical expenses that rise with age provide the elderly with a strong incentive to save, and medical expenses that rise with permanent income encourage the rich to be more frugal. In other work, we showed that heterogeneous life expectancy is important to matching the savings patterns of retired elderly singles (De Nardi et al., 2009).

Poterba et al. (2011) show that relatively little dissaving occurs amongst retirees whose family composition does not change, but that assets fall significantly when households lose a spouse. We find similar results. Furthermore, as in French et al. (2006), we document significant drops in assets when the last member of the household passes away.

Previous literature had shown that high income individuals live longer than low income individuals (Attanasio and Emmerson, 2003; Deaton and Paxson, 2001, see). This means that high income households must save a larger share of their lifetime wealth if they are to smooth consumption over their retirement. Differential mortality rates thus provide a potential explanation for why high income households have higher savings rates than low income households. We extend the analysis along this dimension by explicitly modeling the interaction of life expectancy for individuals in couples and the differential life expectancy for couples and singles.

Even in presence of social insurance such as Medicare and Medicaid, households face potentially large out-of-pocket medical and nursing home expenses (see French and Jones, 2011, 2004; Palumbo, 1999; **?**). The risk of incurring such expenses might generate precautionary savings, over and above those accumulated against the risk of living a very long life (Laitner et al., 2018).

Hubbard et al. (1995) argue that means-tested social insurance programs such as Supplemental Security Income and Medicaid provide strong incentives for low income individuals not to save. De Nardi et al. (2016c) finds that these effects extend to singles in higher permanent income quintiles as well. In this paper, to allow for these important effects, we model means-tested social insurance explicitly for both singles and couples.

Bequest motives could be another reason why households, and especially those with high permanent income, retain high levels of assets at very old ages (Dynan et al., 2002; Ameriks et al., 2011)). De Nardi (2004) and Castañeda et al. (2003) argue that bequest motives are necessary to explain why the observed distribution of wealth is more skewed and concentrated than the distribution of income (Díaz-Giménez et al., 1997).De Nardi et al. (2016a) shows that bequest motives help fit both assets and Medicaid recipiency profiles for singles. We allow for a richer structure of bequest motives, in that couples might want to leave resources to the surviving spouse, children and other heirs, while singles might want to leave bequests to children and other heirs.

Previous quantitative papers on savings have used simpler models that omit one or more of these features. Hurd (1989a) estimates a structural model of bequest behavior in which the time of death is the only source of uncertainty. Palumbo (1999) focuses on the effect of medical expenses and uncertain lifetimes, but omits bequests. Dynan et al. (2002, 2004) consider the interaction of mortality risk, medical expense risk and bequests, but use a stylized two-period model. Moreover, none of these papers model household survival dynamics, assuming instead that households (while "alive") always have the same composition. In contrast, we explicitly model household survival dynamics: when the first household member dies, assets are optimally split among the surviving spouse and other heirs. Although Hurd (1999) extends his earlier model to include household survival dynamics, he omits medical expense risk, and, in contrast to his earlier work, he does not estimate his model.

Our work also complements the one on retirement behavior by couples by Blau and Gilleskie (2008), Casanova (2011), and Gallipoli et al. (2013).

Including couples and simultaneously considering bequest motives, social insurance, uncertain medical expenses, and uncertain life expectancy is important for at least two reasons. First, Dynan et al. (2004) argue that explaining why the rich have high savings rates requires a model with precautionary motives, bequest motives and social insurance. Second, simultaneously considering multiple savings motives allows us to identify their relative strengths. This is essential for policy analysis. For example, the effects of estate taxes depend critically on whether rich elderly households save mainly for precautionary reasons, or mainly to leave bequests (see for example Gale and Perozek, 2001).

3.3 Data and key facts

We use data from the Asset and Health Dynamics Among the Oldest Old (AHEAD) dataset. The AHEAD is a sample of non-institutionalized individuals, aged 70 or older in 1993. These individuals were interviewed in late 1993/early 1994, and again in 1996, 1998, 2000, 2002, 2004, 2006, 2008, 2010, 2012, and 2014.[1] Appendix B.1 describes the details of our sample selection and data work.

We only consider retired households to abstract from the retirement decisions and focus on the determinants of savings and consumption. Because we only allow for household composition changes through death, we drop households where an individual enters a household or an individual leaves the household for reasons other than death. Fortunately, attrition for reasons other than death is a minor concern in our data.

We break the data into 5 cohorts. The first cohort consists of individuals that

[1] We do not use 1994 assets, nor medical expenses, due to underreporting (Rohwedder et al., 2006).

were ages 72-76 in 1996; the second cohort contains ages 77-81; the third ages 82-86; the fourth ages 87-91; and the final cohort, for sample size reasons, contains ages 92-102. Even with the longer age interval, the final cohort contains relatively few observations. In the interest of clarity, we exclude this cohort from our graphs, but we use all cells with at least 10 observations when estimating the model.

We use data for 8 different years; 1996, 1998, 2000, 2002, 2004, 2006, 2008, 2010, 2012 and 2014. We calculate summary statistics (e.g., medians), cohort-by-cohort, for surviving individuals in each calendar year—we use an unbalanced panel. We then construct life-cycle profiles by ordering the summary statistics by cohort and age at each year of observation. Moving from the left-hand-side to the right-hand-side of our graphs, we thus show data for four cohorts, with each cohort's data starting out at the cohort's average age in 1996.

Since we want to understand the role of income, we further stratify the data by post-retirement permanent income (PI). Hence, for each cohort our graphs usually display several horizontal lines showing, for example, median assets in each PI group in each calendar year. These lines also identify the moment conditions we use when estimating the model.

Our PI measure can be thought of as the level of income if there were two people in the household at age 70. We measure post-retirement PI using non-asset non-social insurance annuitized income and the methods described in appendix B.2 . The method maps the relationship between current income and PI, adjusted for age and household structure. The income measure includes the value of Social Security benefits, defined benefit pension benefits, veterans benefits and annuities. Since we model means-tested social insurance from SSI and Medicaid explicitly through our consumption floor, we do not include SSI transfers. Because there is a roughly monotonic relationship between lifetime earnings and the income variables that we use, our measure of post-retirement PI is also a good measure of lifetime permanent income.

3.3.1 Savings and social insurance

The following figures display median assets, conditional on birth cohort and income tercile, for different configurations of couples and singles.

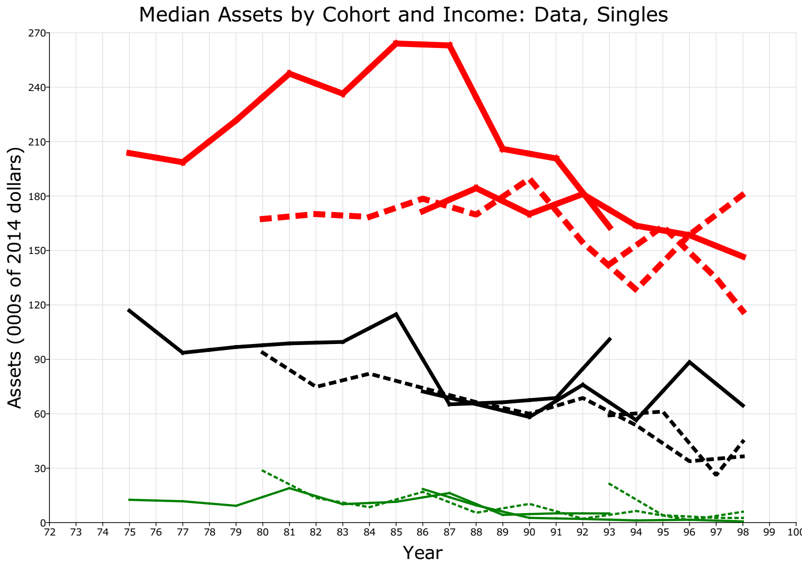


Figure 3.1: Median assets for current singles. Each line represents median assets for a cohort-income cell, traced over the time period 1996-2014. Thicker lines refer to higher permanent income groups. Solid lines: cohorts ages 72-76 and 83-88 in 1996. Dashed lines: ages 77-82 and 89-96 in 1996.

Figure 3.1 displays the net worth (or assets, from now on, since the retirees have very little debt) of the unbalanced panel of people who are currently single, that is, either enter the sample because single at the start of the period, or join it later as they become single (the picture would look very similar if we were to limit the sample to only people who were initially single only, see Appendix B.4). The figure displays the important observation that the savings of elderly singles depends on their income (which is predetermined at retirement). Elderly singles with the lowest lifetime incomes reach retirement with little to no assets and accumulate no additional assets as they age. Rather, they rely on annuitized income and government insurance to fund their consumption and their medical expenses. In the middle of the income distribution, these elderly singles start their retirement with some assets, but then they run them down. At the top of the income distribution, instead, assets of the survivors stay roughly constant as they age and survive. Previous work

(De Nardi et al., 2010, 2016a) has shown that the age profiles of medical expenses in old age, which increase with age and income, and the availability of means-tested social insurance (such as Medicaid and SSA) are key drivers of these patterns and that bequest motives appear to explain little of the lack or slow decumulation of the higher income singles.

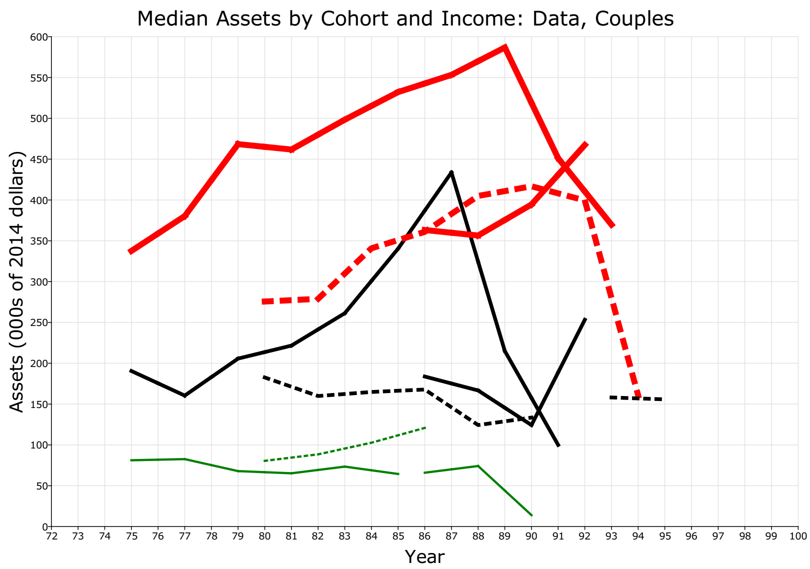


Figure 3.2: Median assets for households intact couples. Each line represents median assets for a cohort-income cell, traced over the time period 1996-2014. Thicker lines refer to higher permanent income groups. Solid lines: cohorts ages 72-76 and 83-88 in 1996. Dashed lines: ages 77-82 and 89-96 in 1996.

Figure 3.2 displays the assets of "intact couples," that is couples that start out and continue as such. Thus, if one or both components die, we drop them from the figure. This figure displays some important facts. First, conditional on the same income tercile, even couples in the lowest income terciles reach retirement with some assets. Second, the asset profiles of the lower to middle income surviving couples display no decumulation, even at the lowest levels of income. Third, the highest income intact couples actually have increasing savings in age for much of their retirement, with drops in their assets only starting to appear after age 85.

Figure 3.3 reports median assets for the population of those who are initially in a couple. Some of these couples lose one spouse during the period in which we observe them, in which case we report the assets of the surviving spouse. Relative to

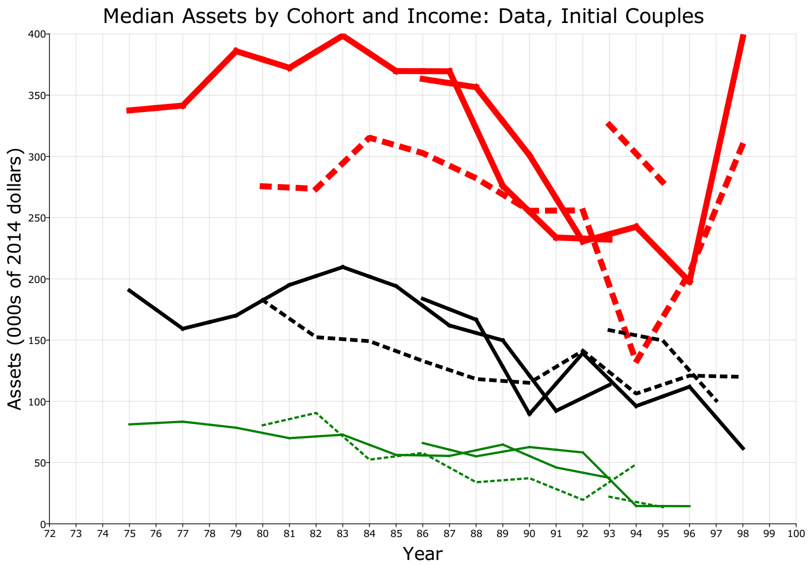


Figure 3.3: Median assets for households initial couples. Each line represents median assets for a cohort-income cell, traced over the time period 1996-2014. Thicker lines refer to higher permanent income groups. Solid lines: cohorts ages 72-76 and 83-88 in 1996. Dashed lines: ages 77-82 and 89-96 in 1996.

intact couples, initial couples in the higher PI terciles decumulate their assets more quickly. Some of this decumulation is due to the loss of assets that occurs when one of the spourses dies.

To formalize and quantify this observation, we turn to documenting the savings profiles of couples when one of their members dies. To do so, we build event studies around the death of a household member by comparing households where a death occurs with control households who do not experience a death.[2]

Figure 3.4 reports the asset levels for these two groups of couples around the death of a household member. We normalize the death to occur at year 0 and follow the households for three waves before the death and two waves after the death. Comparing the two groups, we see some initial differences in their wealth level (note that this is not necessarily the period we match on) suggesting long terms differences between the groups. However, these differences grow if we compare

[2]We match households to a household that did *not* experience a death at the same date or within the window shown here. Households who die may differ on a number of observable characteristics-for example lower income households are likely to die earlier. Thus, for each household we take a matched control household from the set of households who have a PI percentile within 15 percentage points and where their 1996 wealth differs by no more than 20% (or $20,000 in levels).

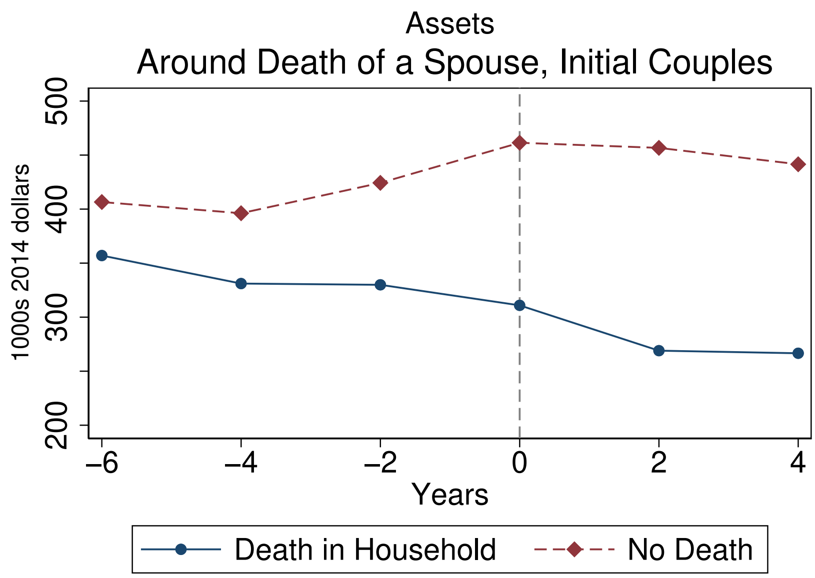


Figure 3.4: Assets around death of a household member for initial couples. Each line represents the mean assets for a group of initial couples around the death (or no death) of a member of the household. Death date is normalised to occur at year 0. Solid lines: Couples who experience a death. Dashed lines: Matched couples who do not experience a death.

households at the death of the first household member where they have grown to \$140,000 (from \$40,000 initially) due to declines among those who die and the increasing profile for households where both members survive.

The simpler exercise of merely tabulating the wealth decline when one member of the household dies suggests smaller estimates of the wealth decline at the time of death of a spouse. This procedure suggests a decline of wealth of \$30,000 at the death of a spouse. A large share of this drop, but not all of it, is explained by the high medical expenses at the time of death. For example, out of pocket medical spending plus death expenses are approximately \$20,000 during the year of death (whereas medical spending is \$6,000 per year for similarly aged people who do not die).

To measure the decline in wealth at the time of death of the final member of the family, we exploit the exit surveys, which include information on the heirs' reports of the value of the estate.[3] In addition, we also use data from post-exit interviews,

[3] Some couples lose both members between successive AHEAD waves, but this is a relatively infrequent occurrence.

which are follow up surveys of heirs, to better measure wealth held in the estate. The results suggest even larger declines in wealth when the final member of the household dies. For those at the top of the income distribution, the death of a single person results in a wealth decline of approximately \$110,000, whereas the decline is closer to \$70,000 for those at the bottom of the income distribution.

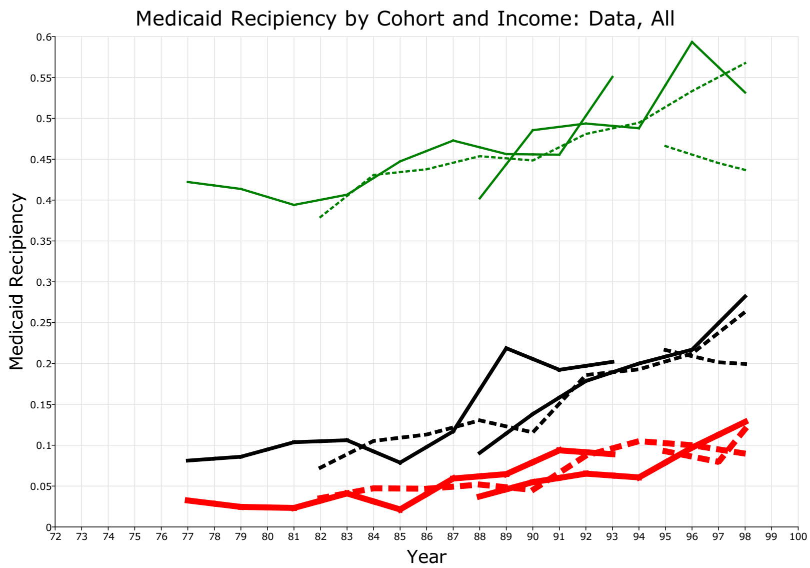


Figure 3.5: Medicaid recipiency by age, income, and cohort. Each line represents the fraction of Medicaid recipients for a cohort-income cell, traced over the time period 1996-2014. Thicker lines refer to higher permanent income groups. Solid lines: cohorts ages 72-76 and 83-88 in 1996. Dashed lines: ages 77-82 and 89-96 in 1996.

Because previous work has shown that means-tested programs are very important determinants of savings (see our related literature section) and because our goal is understanding the savings of the retirees, it is important to take into account the extent to which these retirees benefit from mean-tested government insurance programs that help pay for their medical expenses, such as Medicaid. Figure 3.5 plots Medicaid recipiency rates conditional on age, income, and cohort for all of the retirees in our sample. Several patterns are worth pointing out. First, the lines are flipped compared to the asset graphs. It is the lowest income retirees that are more likely to end up on Medicaid. Second, and consistently with the asset-tested nature of these programs (high assets people are not eligible) and the asset profiles that we have described over the retirement period, Medicaid recipiency of increases fast earlier in their retirement for retirees with intermediate values of income, while it only tends to increase at very advanced ages for higher income people.

3.3.2 Health and mortality

Income Percentile	Men Nursing Home	Men Bad Health	Men Good Health	Women Nursing Home	Women Bad Health	Women Good Health	All
Singles							
10	3.03	6.92	8.68	4.07	11.29	13.18	10.17
50	3.02	7.78	10.29	4.05	12.29	14.86	11.51
90	2.91	8.11	10.94	3.80	12.51	15.37	11.99
Couples							
10	2.73	7.83	9.82	3.95	12.10	14.05	11.30
50	2.77	9.39	12.18	3.99	13.74	16.27	13.37
90	2.74	10.39	13.50	3.88	14.59	17.28	14.45
Single Men							8.99
Married Men							11.52
Single Women							13.88
Married Women							15.77
Oldest Survivor							17.90
Probability that Oldest Survivor is Woman							63.7%

Table 3.1: Life expectancy in years, conditional on reaching age 70.

Table 3.1 shows life expectancies for singles and couples at age 70, respectively.[4] Rich people, women, married people, and healthy people live much longer than their poor, male, single, and sick counterparts.

More specifically, a single man at the 10th permanent income percentile in a nursing home expects to live only 3.0 more years, while a single woman at the 90th percentile in good health expects to live 15.4 more years. The far right column of

[4] We estimate health transitions and mortality rates simultaneously by fitting the transitions observed in the HRS to a multinomial logit model. We allow the transition probabilities to depend on age, sex, current health status, marital status, permanent income, as well as polynomials and interactions of these variables. Using the estimated transition probabilities, we simulate demographic histories, beginning at age 70, for different gender-PI-health combinations. All tables use the appropriate distribution of people over state variables to compute the number which is object of interest.

the top two panels shows average life expectancy conditional on permanent income, averaging over both genders and over all health states. It shows that singles at the 10th percentile of the permanent income distribution live on average 10.2 years, whereas singles at the 90th percentile live on average 12.0 years. It also highlights that a 70 year old man at the 10th permanent income percentile in a nursing home married to a 70 year old woman with the same health expects to live only 2.7 more years, while a 70 year old married woman at the 90th percentile in good health (married to a 70 year old man also in good health) expects to live 17.3 more years.

Conditional only on gender, people in couples live about 2 years longer than singles: single 70 year old women live on average 13.9 years versus 15.8 for married women but, conditional on PI and health, the differences in longevity are much smaller. Thus, married people live longer than singles, but much of the difference is explained by the fact that married people tend to have higher PI and to be in better health.

The bottom part of Table 3.1 shows the number of years of remaining life of the oldest survivor in a household when both the man and the woman are 70. On overage the last survivor lives and additional 17.9 years. In that case, woman is the oldest survivor 63.7% of the time.

Table 3.2 shows that single men and women face on average a 26.4% and 37.2% chance of being in a nursing home for an extended stay, respectively, while married men and women face on average a 19.5% and 36.3% chance of being in a nursing home for an extended stay. Married people are much less likely to transition into a nursing home at any age, but married people often become single as their partner dies and as they age. Furthermore, married people tend to live longer than singles, so they have more years of life to potentially enter a nursing home. Permanent income has only a small effect on ever being in a nursing home. Those with a high permanent income are less likely to be in a nursing home at each age, but they tend to live longer.

Income Percentile	Men Bad Health	Men Good Health	Women Bad Health	Women Good Health	All
Singles					
10	23.6	25.3	35.8	37.9	32.8
50	22.8	24.8	35.5	38.2	32.5
90	20.3	22.8	32.2	35.8	30.1
Couples					
10	17.3	19.2	34.4	37.0	28.7
50	16.6	18.8	34.1	37.3	28.7
90	14.6	16.8	31.4	34.5	26.3
Single Men					26.4
Married Men					19.5
Single Women					37.2
Married Women					36.3

Table 3.2: Probability of ever entering a nursing home, conditional on being alive at age 70.

3.3.3 Income

The goal of this subsection is to illustrate how income changes for couples when one of the two spouses dies. We model income as a function of permanent income.[5] Figure 3.6 presents predicted income profiles for those at the 20^{th} and 80^{th} percentiles of the permanent income distribution. For each permanent income level, we display three scenarios, all commencing from the income of a couple. Under the first scenario, the household remains a couple until age 100. Under the second one, the man dies at age 80. Under the third one, the woman dies at age 80.

Couples' average annual income ranges from about \$14,000 to over \$30,000 per year in the 20^{th} and 80^{th} percentiles of the PI distribution, respectively. As a point of comparison, median wealth holdings for these two groups are \$70,000 and

[5] More specifically, income is a third order polynomial in age, dummies for family structure, family structure interacted with an age trend, and a fifth order polynomial in permanent income. The estimates use a fixed effects estimation procedures, where the fixed effect is a transformation of initial permanent income. Hence the regression results can be interpreted as the effect of changing age or family structure for the same household: see appendix B.2 for details of the procedure.

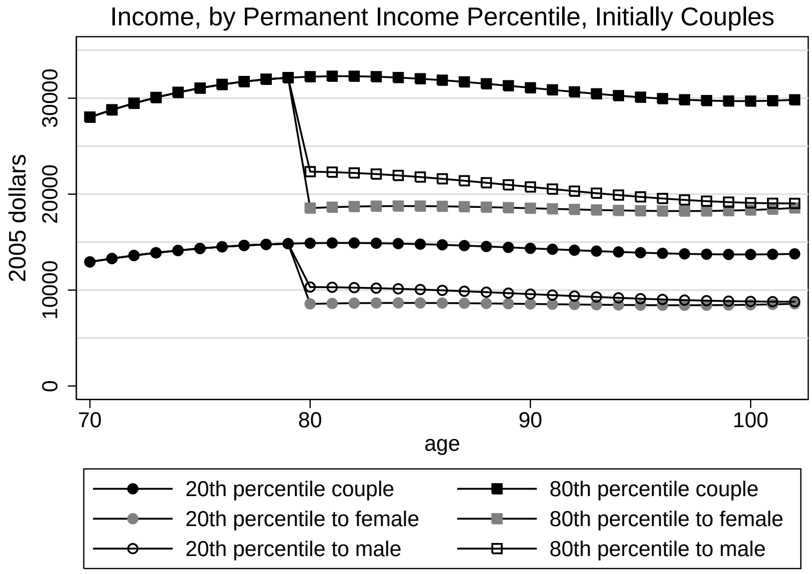


Figure 3.6: Income, conditional on permanent income and family structure. Figure assumes all households begin as couples, then either stay in a couple, or switch to being to a single men or single woman at age 80.

$330,000 at age 74, respectively.

Our estimates suggest that couples in which the husband dies at age 80 suffer a 40% decline in income, while couples in which the wife dies at 80 suffer a 30% decline in income. These income losses at the death of a spouse reflect the fact that although both Social Security and defined benefit pensions have spousal benefits, these benefits replace only a fraction of the deceased spouse's income.

More specifically, people can receive benefits either based on their own history of Social security contributions (in which case they are a "retired worker"), or based on their spouse's or former spouse's history (in which case they receive the "spouse" or "widow" benefit). A married person who never worked can receive 50% of their spouse's benefit if their spouse is alive and is a "retired worker". The same person can receive up to 100% of their spouse's benefit if their "retired worker" spouse has died. Thus the household benefit can receive 100%+50%=150% of the former worker spouse's benefit when alive and 100% of the former worker spouse's benefit when either spouse has died. Thus, after the death of a household member the household would maintain (100%/150%)=67% of the original Social Security benefit and would experience a 33% drop in benefits.

In contrast, a person who earned the same amount as their spouse will not receive a spousal or widow's benefit. In this case, both spouses in the couple will receive 100% of their own "retired worker" benefit, which is based off of their own earnings history. After the death of a spouse, the household benefit will be (100%/(100%+100%)= 50% of its level when both were alive.

To perform our calculations, we make several assumptions, including that both spouses begin receiving benefits at the normal retirement age. In practice, there are many modifications to this rule, including those to account for the age at which the beneficiary and spouse begin drawing benefits.[6] Our regression estimates capture the average drop in income at the time of death of a spouse, averaging over those who retire at different ages and have different claiming histories.

3.3.4 Medical spending

The goal of this subsection is to illustrate how medical expenses change for couples when one of the two spouses dies. In our model, we explicitly model Medicaid, a means-tested program that helps insure medical expenses of the needy. Thus, the appropriate measure of medical expense risk that we have to measure in the data are out-of-pocket medical costs (net of insurance co-pays that are not based on one's resources), plus Medicaid payments.

The AHEAD data contains information on out of pocket medical spending, but not on Medicaid payments. Fortunately, the Medicare Current Beneficiary Survey (MCBS) has extremely high quality information on Medicaid payments plus out of pocket medical spending. It uses a mixture of both administrative and survey data and is described in greater detail in De Nardi et al. (2016c). One drawback of the MCBS, however, is that although it has information on marital status and household income, it does not have information on the medical spending or health of the spouse.

To exploit the strength of both datasets, we use the conditional mean matching

[6] See https://socialsecurity.gov/planners/retire/yourspouse.html for more details of calculation of spousal benefits.

procedure described in Appendix B.3 to obtain our measured out-of-pocket medical expenditures plus Medicaid payments for the AHEAD data from the MCBS data. The procedure preserves both the mean as well distribution of medical spending, conditional on age, income, out of pocket spending and other variables.

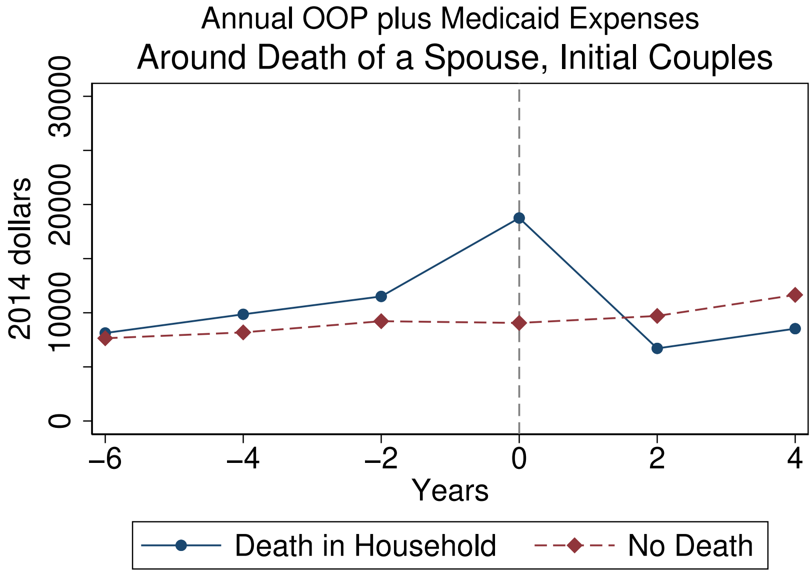


Figure 3.7: Annual medical expenses around death of a household member for initial couples. Each line represents the mean out-of-pocket medical expenditures plus Medicaid payments. for a group of initial couples around the death (or no death) of a member of the household. Death date is normalised to occur at year 0. Solid lines: Couples who experience a death. Dashed lines: Matched couples who do not experience a death.

Figure 3.7 displays our measure of total medical expenses for couples who experience a death in the household and for a comparison group who do not. This uses the same matched comparison group and event study approach used in Figure 3.4. When a member of the household dies expenses increase slowly as average health declines and there is a sharp increase in medical expenses at the wave of death. These households have annual medical expenses that is approximately $12,000 larger than their comparison group and the difference between the groups totals $30,000 in the 6 years before death. Although Figure 3.7 displays the mean value across each of these households, there is considerable variance across households.

The medical expenses of the two groups are almost identical 6 years ahead of a death and for the continuing single person there is a large reduction in the waves following the death of a household member. Despite this reduction, the surviving spouse still has higher medical expenses per person when compared to the couple who do not experience a death in the household. Similarly, both medical expenses for singles rise in the preceding years before increasing sharply at the wave of death while their underlying health declines.

3.4 The model

Consider a retired household with family structure f_t (either a single person or a couple), seeking to maximize its expected lifetime utility at household head age t, $t = t_r, t_r + 1 ..., T + 1$, where t_r is the retirement age, while T is the maximum potential lifespan.

We use w to denote women and h to denote men. Each person's health status, hs^g, $g \in \{h, w\}$, can vary over time. The person can be in a nursing home, in bad health, or in good health.

For tractability, we assume a fixed age gap between the husband and the wife in a couple, so that one age is sufficient to characterize the household. To be consistent with the data frequency, and to reduce computation time, our time period is two years long. We assume that wives are two years younger than their husbands.

3.4.1 Preferences

Households maximize their utility by choosing savings, bequests and current and future consumption. The annual discount factor is given by β. Each period, the household's utility depends on its total consumption, c, and the health status of each member. The within-period utility function for a single is given by

$$u(c, hs) = (1 + \delta(hs))\frac{c^{1-\nu}}{1-\nu}, \tag{3.1}$$

with $\nu \geq 0$. When $\delta(.) = 0$, health status does not affect utility.

The flow utility function for couples is:[7]

$$u^c(c, hs^h, hs^w) = [1 + \delta(hs^h) + 1 + \delta(hs^w)]\frac{(c/\eta)^{1-\nu}}{1-\nu}, \quad (3.2)$$

where $1 < \eta \leq 2$ determines the extent to which couples enjoy economies of scale in the transformation of consumption goods to consumption services.

When a household member dies, the estate can be left to the surviving spouse (if there is one) or to other heirs, including the household's children. Estates are subject to estate taxes, but the exemption level during the time period in our sample is above the actual assets of the vast majority of the households in our sample. For this reason, we abstract from explicitly modelling estate taxation. When the final member of the household dies bequests are net of their medical expenses.

We indicate with b the part of the estate that does *not* go to the surviving spouse, and assume that the deceased member of the household derives utility $\theta_j(b)$ from leaving that part of the estate to heirs other than the spouse. The subscript j indicates whether there is a surviving spouse or not, and whether one or two people have just died. In particular, $\theta_0(b)$ gives the utility from bequests for a single person with no surviving spouse, $\theta_1(b)$ gives the utility from bequests when there is a surviving spouse, and $\theta_2(b)$ gives the utility from bequests when both spouses die at the same time. The bequest function takes the form

$$\theta_j(b) = \phi_j \frac{(b + k_j)^{1-\nu}}{1-\nu}, \quad (3.3)$$

where k_j determines the curvature of the bequest function, and ϕ_j determines its intensity. It was first used to help explain savings over the life cycle by De Nardi (2004) and can support several interpretations of the "bequest motive:" dynastic or "warm glow" altruism (as in Becker and Tomes, 1986; Abel and Warshawsky, 1988; Andreoni, 1989); strategic motives (as in Bernheim et al., 1985; Brown, 2006); or some form of utility from wealth itself (as in Carroll, 1998; Hurd, 1989b).

[7]Mazzocco (2003) shows that under full commitment, the behavior of a couple can be characterized by a unique utility function if the husband and wife share identical discount factors, identical beliefs and Harmonic Absolute Risk Aversion utility functions with identical curvature parameters.

3.4.2 Timing

The timing is the following: the household begins the period with a given marital status, which is just widowed man or woman, single man, single woman, or couple.

The people who are just widowed first optimize over bequests left upon the death of the second spouse. After distributing spousal bequests, they become single men or women with a set of given state variables.

Single men, single woman, or couples use the cash on hand they have at the beginning of the period to consume and save. After that, mortality shocks, health shocks, and medical shocks happen, income net of taxes is received, and government transfers take place. As a result, the amount of cash on hand that is carried over into the next period is also realized.

Thus, in case of death, households where the final member(s) has just died leave all their resources to their heirs. When one of the members of the couple dies, the survivor, before taking any other decisions, splits cash on hand at the beginning of the period into bequests left after the death of the first spouse and cash on hand that he or she actually gets to use to consume and save. Our timing assumes, consistently with reality, the medical costs associated with death are collected before any bequests can be made and that Medicaid pays bill that are incurred even when the patient dies.

3.4.3 Technology and sources of uncertainty

We assume that non-asset income at time t, $y_t(\cdot)$, is a deterministic function of the household's permanent income, I, age, family structure, and gender if single.

$$y_t(\cdot) = y(I, t, f_t, g). \tag{3.4}$$

We do not include received bequests as a source of income, because very few households aged 65 or older receive them.

There are several sources of uncertainty:

1) Health status uncertainty. The transition probabilities for a person's future health status depend on that person's permanent income, age, gender, current health

status and current marital status:

$$\pi_t(\cdot) = \Pr(hs_{t+1}|I,t,g,hs_t,f_t). \tag{3.5}$$

2) Survival uncertainty. Let $s(I,t,g,hs_t,f_t)$ denote the probability that an individual of gender g is alive at age $t+1$, conditional on being alive at age t, having time-t health status hs, enjoying household permanent income I, and having family structure f_t.

3) Medical expense uncertainty at the household level. We define m_{t+1} as the flow of all out-of-pocket medical expenses, including insurance premia, and medical expenses covered by the consumption floor incurred between periods t and $t+1$. We assume that medical expenses depend upon the health status of each family member (wife w and husband h), household permanent income, family structure, gender if single, age, and an idiosyncratic component, ψ_{t+1}.

$$\begin{aligned} \ln m_{t+1} = {} & m(hs_t^h, hs_t^w, hs_{t+1}^h, hs_{t+1}^w, I, g, t+1, f_t, f_{t+1}) \\ & + \sigma(hs_t^h, hs_t^w, hs_{t+1}^h, hs_{t+1}^w, I, g, t+1, f_t, f_{t+1}) \times \psi_{t+1}. \end{aligned} \tag{3.6}$$

Notice that we allow it to depend on the household's composition and health status at the beginning of the period (which was realized at the very end of the previous period) and at the end of this period (which will be carried over into next period). This allows us to capture the jump in medical spending that occurs in the period a family member dies – that is f_t changes – and to incorporate the impact of two subsequent health realizations on medical spending. The latter helps better represent the cost of prolonged periods of bad health or nursing home stays. Out timing thus implies that the value of m_{t+1} depends is not known by the household when it decides how much to consume between periods t and $t+1$.

Following Feenberg and Skinner (1994) and French and Jones (2004), we assume that ψ_{t+1} can be decomposed as

$$\psi_{t+1} = \zeta_{t+1} + \xi_{t+1}, \quad \xi_{t+1} \sim N(0, \sigma_\xi^2), \tag{3.7}$$

$$\zeta_{t+1} = \rho_m \zeta_t + \varepsilon_{t+1}, \quad \varepsilon_{t+1} \sim N(0, \sigma_\varepsilon^2), \tag{3.8}$$

where ξ_{t+1} and ε_{t+1} are serially and mutually independent. We discretize ξ and ζ, using quadrature methods described in Tauchen (1986).

3.4.4 Recursive formulation

Let a_t denote assets at period t and let r denote the constant pre-tax rate of return earned on those assets. Total post-tax income is given by $\Upsilon(r a_t + y_t(\cdot), \tau)$, with the vector τ summarizing the tax structure.

Define the available resources *before* government transfers as

$$\widetilde{x}_t = a_t + \Upsilon\big(r a_t + y_t(\cdot), \tau\big) - m_t, \tag{3.9}$$

Following Hubbard et al. (1994, 1995), we assume that the transfers bridge the gap between a minimum consumption floor and the household's financial resources. Consistently with the main Medicaid and SSI rules, we can then express government transfers as

$$tr_t(\widetilde{x}_t, f_t, a_t, y_t) = \max\Big\{0, c_{min}(f_t) - \big(\widetilde{x}_t - a_d(f_t, a_t) - y_d(f_t, y_t, a_t)\big)\Big\}, \tag{3.10}$$

$$a_d(f_t, a_t) = \min\{a_t, a_d(f_t)\},$$

$$y_d(f_t, y_t, a_t) = \min\{y_t + r a_t, y_d(f_t)\},$$

where $a_d(f_t)$ and $y_d(f_t)$ are the statutory disregard limits for assets and income, respectively. We allow the guaranteed consumption level c_{min}, the asset disregard a_d, and the income disregard y_d to all vary with family structure.

To save on state variables we follow Deaton (1991) and express the household's financial resources in terms of cash-on-hand:

$$x_t = a_t + \Upsilon(r a_t + y_t(\cdot), \tau) - m_t + tr_t(\widetilde{x}_t, f_t, a_t, y_t). \tag{3.11}$$

Households choose consumption and the remaining wealth carried forward to period $t+1$ is

$$a_{t+1} = x_t - c_t, \tag{3.12}$$

Next period's resources before transfers and cash-on-hand can thus be all expressed in terms of this period's cash on hand and are, respectively,

$$\widetilde{x}_{t+1} = (x_t - c_t) + \Upsilon\big(r(x_t - c_t) + y_{t+1}(\cdot), \tau\big) - m_{t+1}, \tag{3.13}$$

$$x_{t+1} = x_t - c_t + \Upsilon\big(r(x_t - c_t) + y_{t+1}(\cdot), \tau\big) - m_{t+1} + tr_{t+1}(\widetilde{x}_{t+1}, f_{t+1}, x_t - c_t, y_{t+1}). \tag{3.14}$$

The value function for a single individual of gender g can thus be written in terms of cash on hand

$$\begin{aligned} V_t^g(x_t, hs_t, I, \zeta_t) = \max_{c_t} \Big\{ & u(c_t, hs_t) + \beta s(I, t, g, hs_t, f_t) \\ & \times E_t\Big(V_{t+1}^g(x_{t+1}, hs_{t+1}, I, \zeta_{t+1})\Big) \\ & + \beta(1 - s(I, t, g, hs_t, f_t)) E_t \theta_0(x_{t+1}) \Big\}, \end{aligned} \tag{3.15}$$

subject to equations (3.4)-(3.8), (3.13), (3.14), and the following constraint that ensures that consumption is at least as high as its floor and that savings are non-negative

$$c_{min}(f_t) \leq c_t \leq x_t, \quad \forall t. \tag{3.16}$$

The newly widowed (single) individual of gender g has to first distribute bequests after the death of the first spouse before he or she can make savings and consumption decisions as a single person

$$V_t^{ng}(x_t, hs_t, I, \zeta_t) = \max_{b_t} \Big\{ \theta_1(b_t) + \omega V_t^g(x_t - b_t, hs_t, I, \zeta_t) \Big\}, \tag{3.17}$$

subject to equation

$$b_t \in [0, x_t - (c_{min}(f_t) + a_d(f_t) + y_d(f_t))] \tag{3.18}$$

which prohibits the surviving spouse from using bequests to become eligible for government transfers. The term ω measures the weight a couple gives to the surviving spouse, as opposed to the intact couple or other heirs, when it takes expectations over future outcomes.

The value function for couples can be written as

$$\begin{aligned} V_t^c(x_t, hs_t^h, hs_t^w, I, \zeta_t) = \max_{c_t} \Big\{ & u^c(c_t, hs_t^h, hs_t^w) \\ & + \beta s(I,t,w,hs_t,1) s(I,t,h,hs_t,1) E_t \Big(V_{t+1}^c(x_{t+1}, hs_{t+1}^h, hs_{t+1}^w, I, \zeta_{t+1}) \Big) \\ & + \beta s(I,t,w,hs_t,1)(1 - s(I,t,h,hs_t,1)) E_t \Big(V_t^{nw}(x_{t+1}^w, hs_{t+1}, I, \zeta_{t+1}) \Big) \\ & + \beta (1 - s(I,t,w,hs_t,1)) s(I,t,h,hs_t,1) E_t \Big(V_{t+1}^{nh}(x_{t+1}^h, hs_{t+1}, I, \zeta_{t+1}) \Big) \\ & + \beta (1 - s(I,t,w,hs_t,1)(1 - s(I,t,h,hs_t,1) E_t \theta_2(x_{t+1}) \Big\}, \end{aligned} \tag{3.19}$$

subject to equations (3.4)-(3.8), (3.13), (3.14), and (3.16). The dating of the continuation value for widows, $V_t^w(\cdot)$, reflects that wives are one model period (two years) younger than their husbands.

3.5 Estimation

We adopt a two-step strategy to estimate the model. In the first step, we estimate or calibrate those parameters that, given our assumptions, can be cleanly identified outside our model. In particular, we estimate health transitions, out-of-pocket medical expenses, and mortality rates from raw demographic data. In our baseline estimates we fix the annual discount factor at 0.97 and impose that the marginal utility of consumption is independent of health (health preference shifters are equal to zero). We set the consumption floor for couples to be a multiple of the consumption floor for singles, that we estimate. Finally, we assume the weight on the surviving spouse is equal to the weight on the couple and that the utility from bequests when both individuals in a couple die is the same as when the final household member dies.

In the second step, we estimate the rest of the model's parameters (risk aversion, consumption equivalence scale, the consumption floor for singles, and bequest parameters)

$$\Delta = (\nu, \eta, c_{min}(f_t = s), \phi_0, \phi_1, k_0, k_1)$$

with the method of simulated moments (MSM), taking as given the parameters that were estimated in the first step. In particular, we find the parameter values that allow simulated life-cycle decision profiles to "best match" (as measured by a GMM criterion function) the profiles from the data.

Because our underlying motivations are to explain why elderly individuals retain so many assets and to explain why individuals with high income save at a higher rate, we match the median and the 75th percentile of assets by cohort, age, and permanent income. Because we wish to study differences in savings patterns of couples and singles, we match profiles for the singles and couples separately.

Because Medicaid is an important program insuring medical expenses and consumption of the poorest, we validate our model against Medicaid recipiency. Additionally, we reproduce our event study design for changes in wealth when a member of the household dies as an additional validation exercise. Neither of these moments are targets in the estimation.

In particular, the moment conditions that comprise our estimator are given by

1. Median asset holdings by PI-cohort-year for the currently singles who are still alive when observed.

2. Median asset holdings by PI-cohort-year for those who are currently couples with both members currently alive.

3. 75^{th} percentile asset holdings by PI-cohort-year for the currently singles who are still alive when observed.

4. 75^{th} percentile asset holdings by PI-cohort-year for those who are currently couples with both members currently alive.

   When there is a death in a couple, the surviving spouse is included in the current singles' profile of the appropriate age, cohort, and permanent income cell; in keeping with our assumption that all singles differ only in their state variables.[8]

The cells for net worth are computed as follows.[9] Household i has family structure f_t; which indexes households that are currently couples and those that either are currently singles. We sort type-f households in cohort c by their permanent income levels, separating them into $Q = 3$ terciles. Suppose that household i's permanent income level falls in the qth permanent income interval of households in its cohort and family structure.

For Medicaid recipiency, we construct cells for all individuals, whether they are married or not, but conditioning by PI and cohort and we match Medicaid recipiency by cell.

The mechanics of our MSM approach are as follows. We compute life-cycle histories for a large number of artificial households. Each of these households is endowed with a value of the state vector $(t, f_t, x_t, I, hs_t^h, hs_t^w, \zeta_t)$ drawn from the data distribution for 1996, and each is assigned the entire health and mortality history realized by the household in the AHEAD data with the same initial conditions. This way we generate attrition in our simulations that mimics precisely the attrition relationships in the data (including the relationship between initial wealth and mortality).

We discretize the asset grid and, using value function iteration, we solve the model numerically. This yields a set of decision rules, which, in combination with the simulated endowments and shocks, allows us to simulate each individual's assets, medical expenditures, health, and mortality. We compute assets from the artificial histories in the same way as we compute them from the real data. We use these profiles to construct moment conditions, and evaluate the match using our GMM criterion. We search over the parameter space for the values that minimize the cri-

[8]This assumption appears to hold remarkably well in the data, in that conditional on our state variables, the savings and Medicaid recipients of recently singles are very similar to those of the people we were already single in the past. This is perhaps not surprising given that the vast majority of single people in our sample were married at some point.

[9]Simulated agents are endowed with asset levels drawn from the 1996 data distribution, and thus we only match asset data 1998-2014. As was done when constructing the figures from the HRS data, we drop simulated cells with fewer than 10 observations from the moment conditions.

terion. Appendix B.6 contains a detailed description of our moment conditions, the weighting matrix in our GMM criterion function, and the asymptotic distribution of our parameter estimates.

When estimating the life-cycle profiles, and subsequently fitting the model to those profiles, we face two well-known problems. First, in a cross-section, older households were born in an earlier year than younger households and thus have different lifetime incomes. Because lifetime incomes of households in older cohorts will likely be lower than the lifetime incomes of younger cohorts, the asset levels of households in older cohorts will likely be lower also. Therefore, comparing older households born in earlier years to younger households in later years leads to understate asset growth. Second, households with lower income and wealth tend to die at younger ages than richer households. Therefore, the average survivor in a cohort has higher lifetime income than the average deceased member of the cohort. As a result, "mortality bias" leads the econometrician to overstate the average lifetime income of members of a cohort. This bias is more severe at older ages, when a greater share of the cohort members are dead. Therefore, "mortality bias" leads to overstate asset growth.

We use panel data to overcome these first two problems. Because we are tracking the same households over time, we are obviously tracking members of the same cohort over time. Similarly, we do separate sets of simulations for each cohort, so that the (initial) wealth and income endowments behind the simulated profiles are consistent with the endowments behind the empirical profiles. As for the second problem, we explicitly simulate demographic transitions so that the simulated profiles incorporate mortality effects in the same way as the data, both for couples and singles.

3.6 Estimation results

Section 3.3 includes results from our first-step estimation, that we use as inputs for our structural model, and the outputs that we require our model to match. In this section, we report our second-step parameter estimates for our structural model, the

ν: RRA coefficient	4.346
	(0.1533)
η: consumption equivalence scale	1.326
	(0.2194)
ϕ_0: bequest intensity, single	171,900,000
	(76,590,000)
κ_0: bequest curvature, single (in 000s)	7,017
	(1,310)
ϕ_1: bequest intensity, surviving spouse	186,900
	(47,420)
κ_1: bequest curvature, surviving spouse (in 000s)	645
	(100)
ϕ_2: bequest intensity, both spouses	171,900,000
	(NA)
κ_2: bequest curvature, both spouses (in 000s)	7,017
	(NA)
$c_{min}(f=2)$: annual consumption floor, couples	3,457
	(NA)
$c_{min}(f=1)$: annual consumption floor, singles	2,304
	(299.3)

Notes: Standard errors in parentheses. NA refers to the fixed ratio of couples' and singles' consumption floors and other parameters fixed in estimation

Table 3.3: Estimated parameters from our second-step estimation

model fit, and we discuss the model's identification.

3.6.1 Second-step Results

Table 3.3 presents our estimated preference parameters. Our estimate of ν, the coefficient of relative risk aversion, is 4.3 is similar to earlier estimates for retired singles. De Nardi et al. (2010); Ameriks et al. (2018)), and Lockwood (2018) all estimate models of the savings decisions of retired singles and estimate $\nu > 4$.

Our estimates imply that achieving the same level of consumption services as a couple requires substantially lower per person expenditures. Fernández-Villaverde and Krueger (2007) argue for similar economies of scale using an average across a variety of existing estimates. Furthermore, Hong and Ríos-Rull (2012) estimate large economies of scale using data on Life Insurance purchases in a model of life cycle consumption and savings decisions. Intuitively, η is identified from the extent to which couples save to self insure future risks relative to singles which is conceptually similar. Our estimated value, $\eta = 1.33$, is identical to their estimate.

The point estimates of θ_0 and κ_0 imply that, in the period before certain death, the bequest motive becomes operative once consumption exceeds \$45,300 per year. For individuals in this group, the marginal propensity to bequeath, above the threshold level, is 98 cents out of every additional dollar. Taken together, this implies that end of life bequest motives are luxury goods and important for wealthy households. We constrain θ_2 and κ_2, the corresponding estimates for couples to equal our estimated parameters for singles. However, differences in the size of the household and the consumption equivalence scale mean that for couples facing certain death bequest motives are more luxurious and weaker.

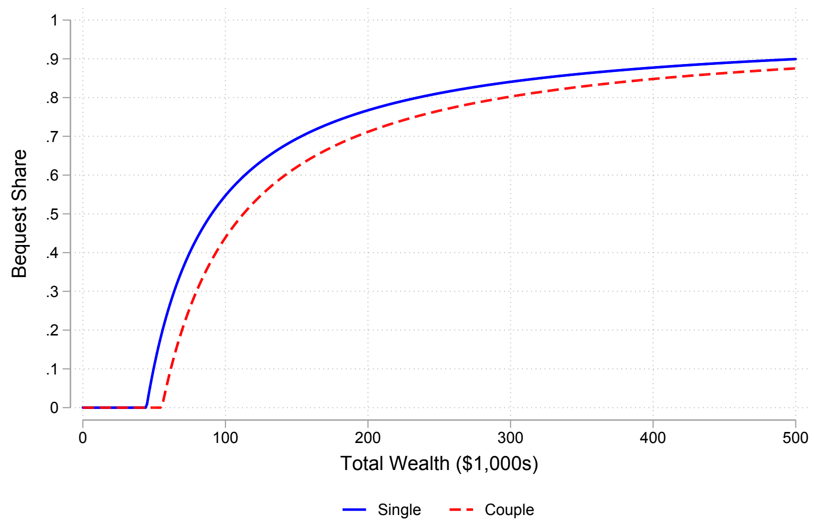


Figure 3.8: Expenditure share allocated to bequests for singles (solid) and couples (dashed) facing certain death in the next period.

Figure 3.8 plots the allocation of available resources that would be left as bequests for households facing certain death. Our point estimates imply a large marginal propensity to bequeath, above the threshold level, and although this is the asymptotic allocation of resources to bequests, for much of the distribution of wealth bequest allocations are smaller. In comparison, our estimates of the bequest motive when the first household member dies are weaker.

The consumption floor for single households is estimated at $2,300 per year. In addition we set the annual income disregard at $600 which means in practice the consumption floor is $2,900 because disregarded income is almost always spent by households who receive the consumption floor. In addition, single households who qualify for the consumption floor are entitled to an asset disregard of $3000 which raises the effective value of the consumption floor.

We now turn to highlighting how well the model fits the some key aspects of the data and to highlighting some of the model's implications for saving at older ages.

3.6.2 Parameter Identification

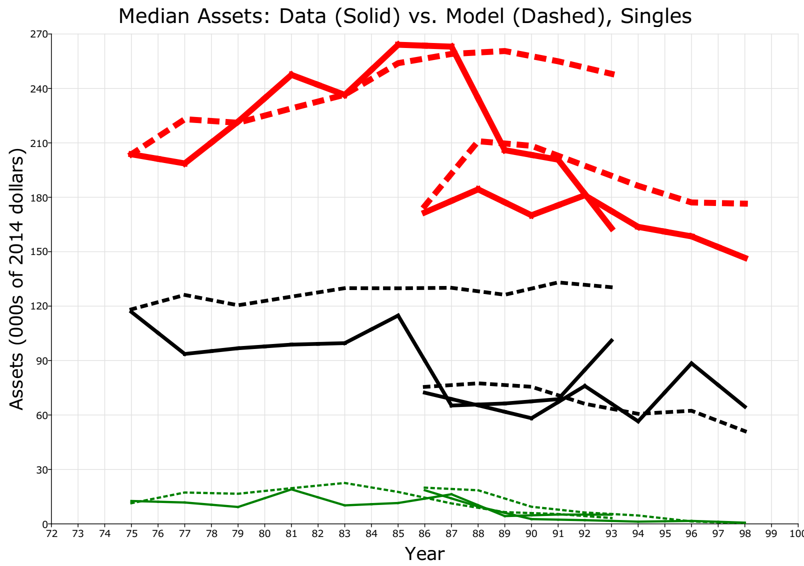
3.6.3 Model Fit

The model does a good job of matching its data targets, and matches other features of the data as well. We compare the median and 75th percentile of assets for couples and singles generated by the model (dashed line) to the data (solid line) for each PI tercile. For clarity we show targeted moments for only two of our birth cohorts in each panel.

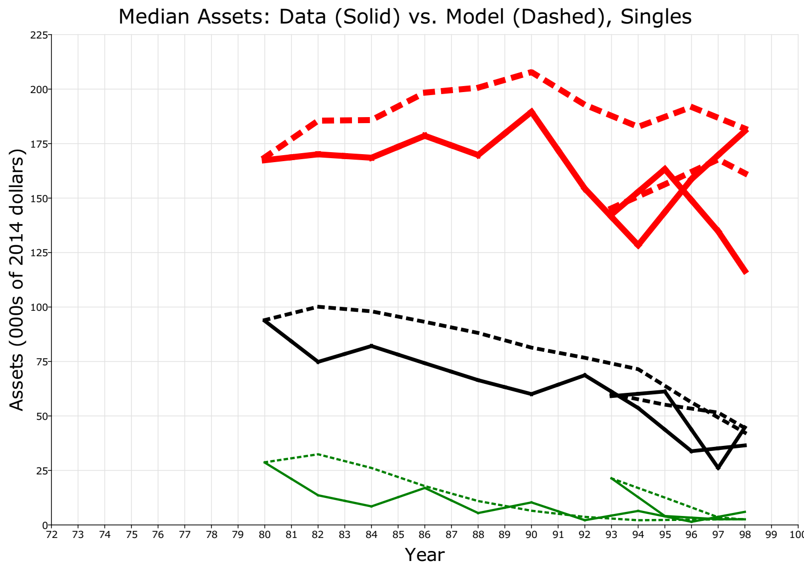
Figure 3.9 plots median net worth by age, PI and birth cohort for the population of current singles in the model (dashed line) and compares it to those in the data (solid line). The model matches broad difference in the way that median assets evolve with age by PI. The model tends to overpredict assets for the higher PI singles although it does produce differences across cohorts and age within these PI groups.

Figure 3.10 shows the equivalent graph for intact couples. Here too the model is successful at reproducing the key features of the data. The model predicts that while both members of the couple remain alive they are more likely to accumulate wealth and matches how this varies with PI, cohort and age. While the model does not produce the rapid accumulation and decline of wealth between ages 83 and 89 for the middle PI tercile of the youngest cohort it closely tracks the data for all other cohort and PI groups as well as matching the age profile for this group at other ages.

Turning to Figure 3.11, we plot the 75th percentile of the asset distribution



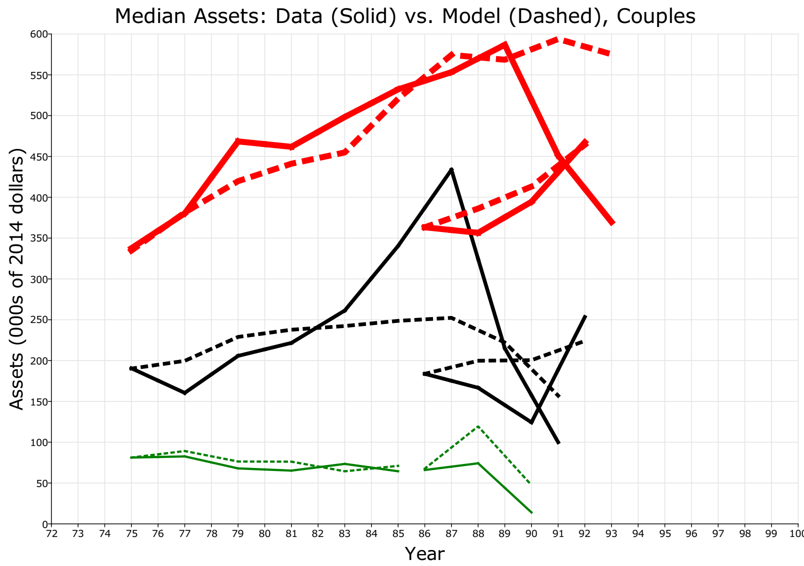
(a)



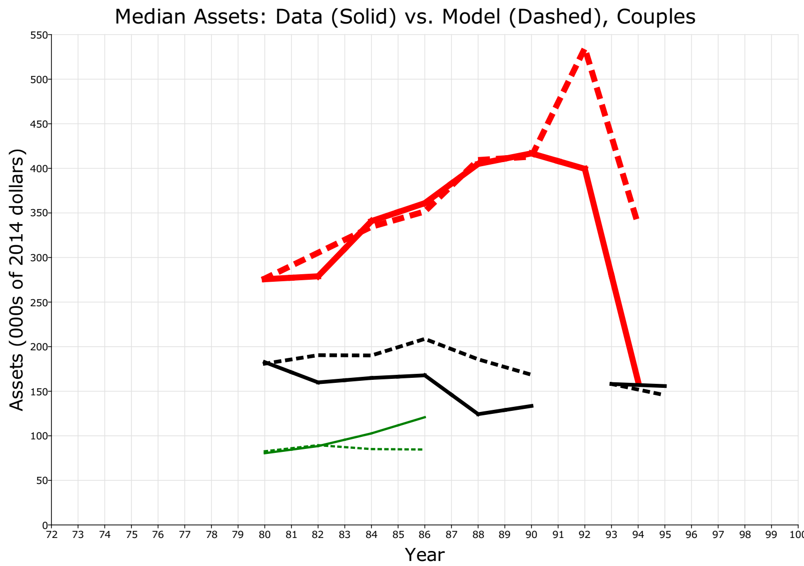
(b)

Figure 3.9: Median assets for single households. Solid lines: cohorts ages 72-76, 83-88 in 1996 in panel (a) and ages 77-82 and 89-96 in 1996 in panel (b). Dashed lines: model generated data.

for our model generated (dashed) sample of singles and compare it with the same moment from the data (solid). Compared to the median asset profiles we report in Section 3.3, the 75th percentile is both higher in levels for all PI groups and shows more evidence of accumulation. Furthermore, compared to the median wealth of intact couples, for the same PI group and birth cohort the 75th percentile of wealth



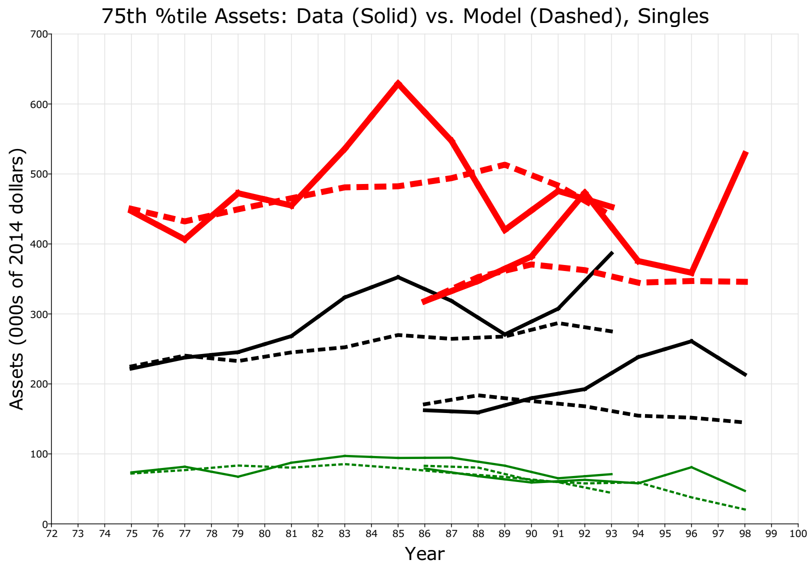
(a)



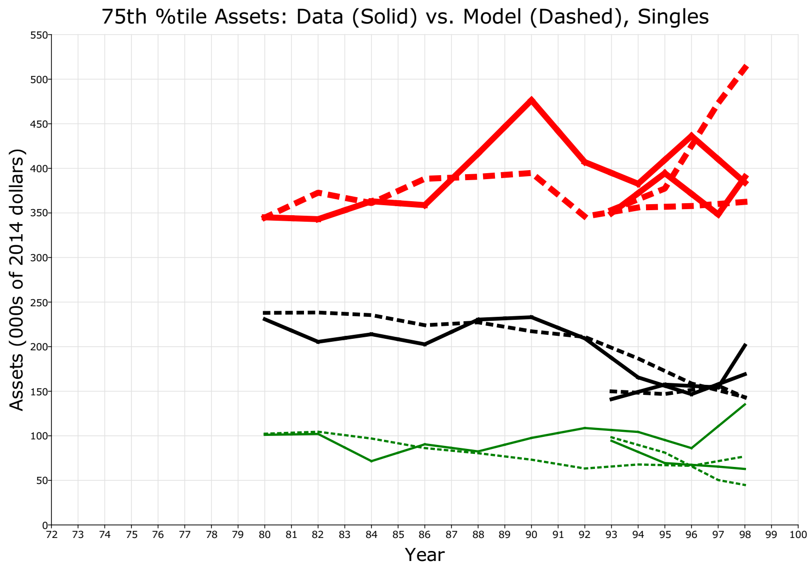
(b)

Figure 3.10: Median assets for households initial couples. Solid lines: cohorts ages 72-76, 83-88 in 1996 in panel (a) and ages 77-82 and 89-96 in 1996 in panel (b). Dashed lines: model generated data.

held by singles is large - although the accumulation is more modest. Using these PI, age and birth cohort conditional moments across couples and singles helps to identify households precautionary savings motive separately from bequest motives and the degree to which they are luxury goods. Again, the model performs well matching differences in savings patterns by PI and age, including the tendency of



(a)

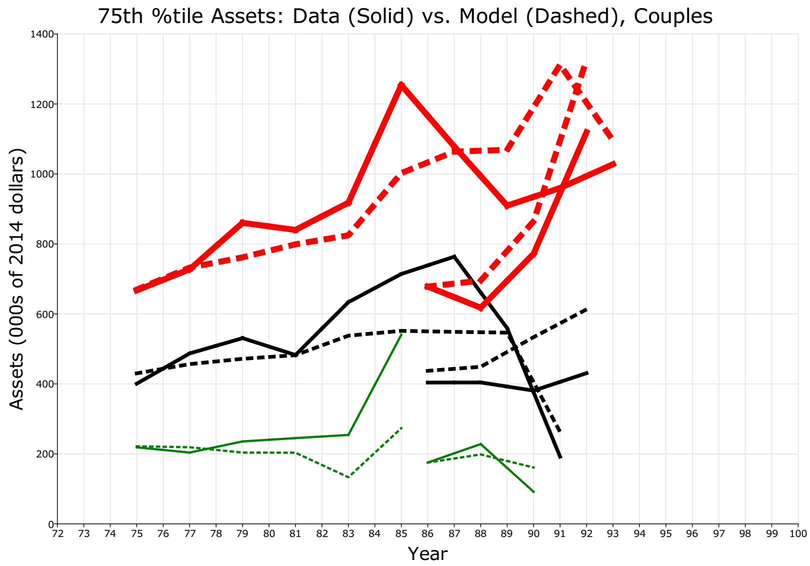


(b)

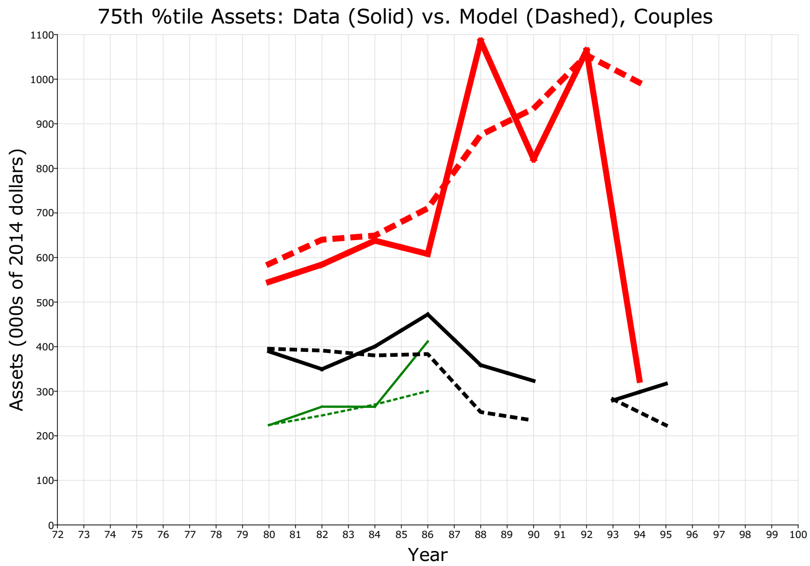
Figure 3.11: 75th percentile assets for single households. Solid lines: cohorts ages 72-76, 83-88 in 1996 in panel (a) and ages 77-82 and 89-96 in 1996 in panel (b). Dashed lines: model generated data..

higher-PI singles to accumulate assets towards the top of the distribution. In contrast to the median assets of single households, if anything the model predicts too little wealth accumulation.

Finally, Figure 3.12 displays the equivalent data moment for our sample of intact couples. These wealth holdings are larger than the sample of singles and



(a)



(b)

Figure 3.12: 75th percentile assets for households initial couples. Solid lines: cohorts ages 72-76, 83-88 in 1996 in panel (a) and ages 77-82 and 89-96 in 1996 in panel (b). Dashed lines: model generated data.

for the top PI tercile they exceed over $1,000,000 by their early 90s. In general, these savings profiles show more accumulation in levels than the 75th percentile of single households' assets. There are large differences in the rate and timing of this accumulation by Permanent Income. In addition to matching the accumulation at the median by age, PI and cohort the model does a good job of generating the growth in savings for this higher level of asset holdings. It does a reasonable job of matching the flatter profiles of lower PI groups while still being able to match the eventual wealth holdings above $1,000,000 for the top terciles

3.6.4 Model Validation

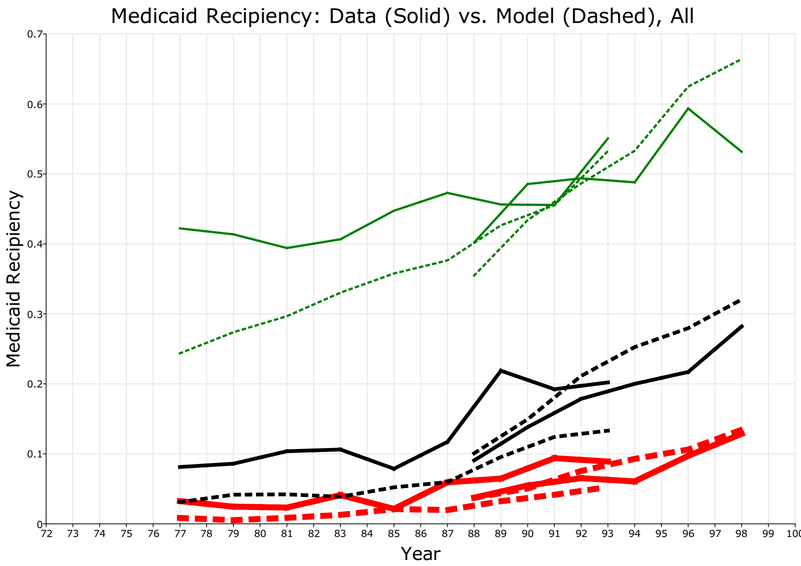
In addition to comparing how the model matches moments used in estimation, we also compare our model with moments from the data which are *not* part of the GMM criterion we use to estimate the model.

Figure 3.13 compares the Medicaid recipiency profiles generated by the model (dashed line) to those in the data (solid line) for the same birth cohorts. The graphs show that the model matches the general patterns of Medicaid usage. In particular, the model matches difference in the level of usage by age and household permanent income. The model tends to underpredict usage at younger ages, especially by the poor, and to overpredict usage by older households, especially by the middle third of lifetime incomes.

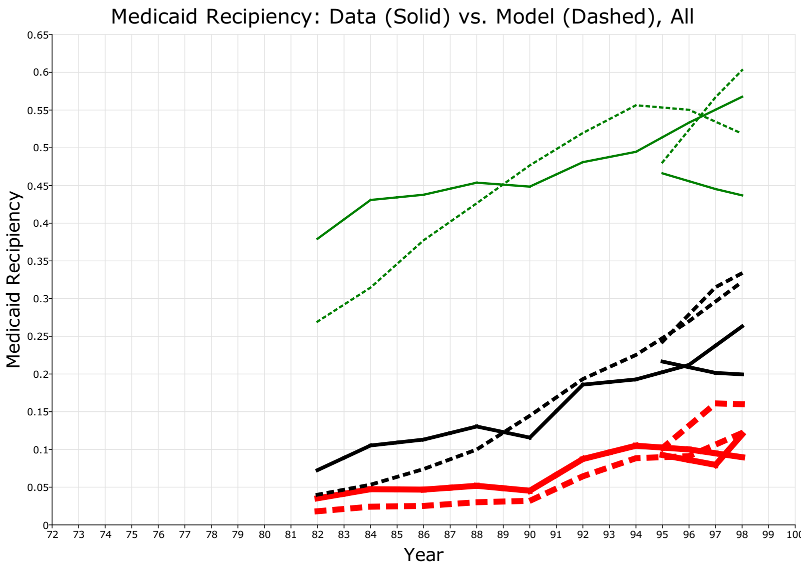
Matching the key features of Medicaid recipiency is important for policy analysis and understanding how medical expense risk drives the savings of elderly couples and singles. It means the model is able to capture the current level of social insurance for households as well as matching the risk of catastrophic medical spending which causes higher income households to rely on Medicaid.

Our second validation exercise returns to the savings profiles of couples when one of their member dies. We use the same event study design as Section 3.3 and construct a model counterpart using the set of matched households in the data.[10] We display the results for the model generated data in panel (a) of Figure 3.14 and reproduce Figure 3.4 for the reader's convenience. Although this is not targeted, the model generated data reproduces the level of wealth held by couples around the death of a spouse. The model overpredicts the level of wealth in the periods before death for simulated couples where one member of the household does, however, it

[10]Formally, in addition to drawing initial conditions, health and mortality trajectories for simulated households we assign simulated households to either of these groups based on the data equivalent. For the simulated households who do not experience a death we additionally assign the same counterfactual death date as the data equivalent.



(a)

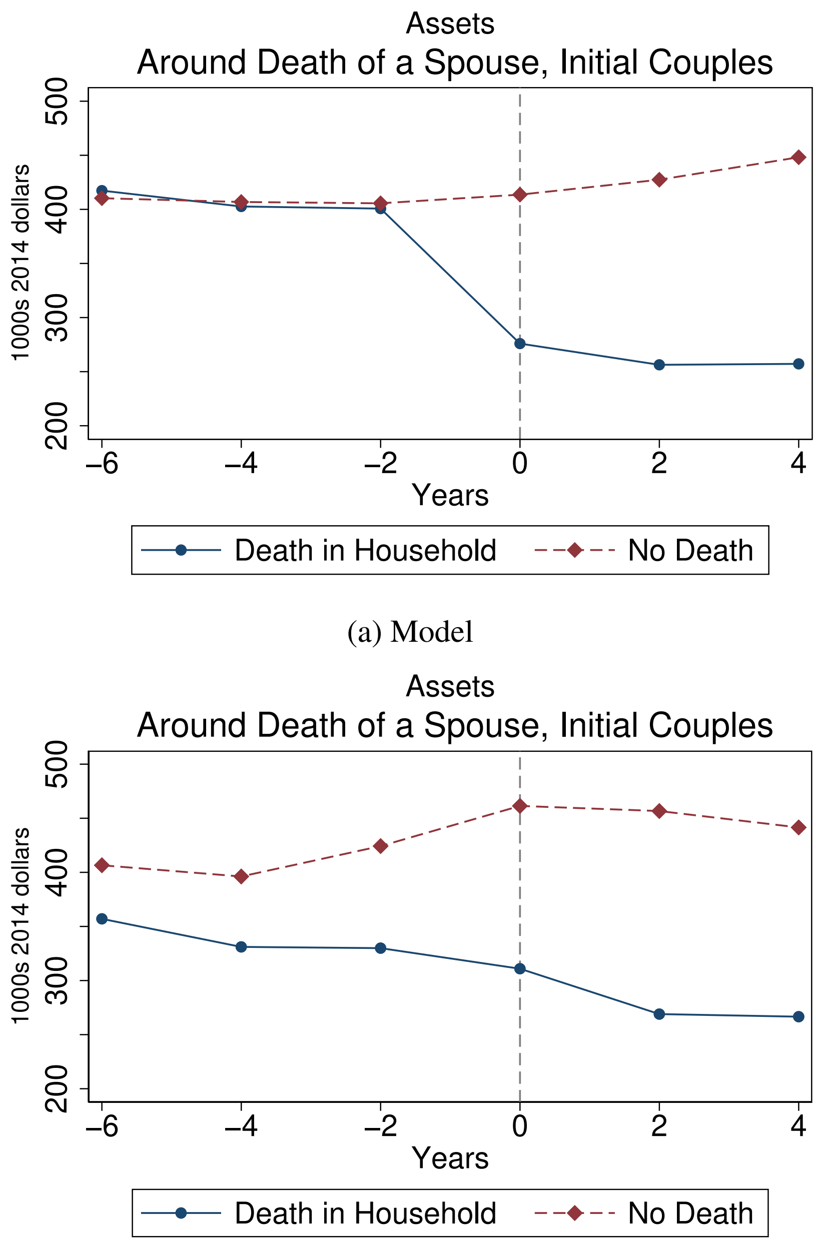


(b)

Figure 3.13: Medicaid recipiency by age, income, and cohort. Solid lines: cohorts ages 72-76, 83-88 in 1996 in panel (a) and ages 77-82 and 89-96 in 1996 in panel (b). Dashed lines: model generated data.

reproduces the differences in the trajectory of savings between both groups as well as the total difference in wealth at the death of a spouse. Furthermore, 4 years after the household member dies both our simulations and the AHEAD data predict a difference of around $140,000.

This exercise is an important test of our model because it demonstrates that our



(a) Model

(b) Data

Figure 3.14: Assets around death of a household member for initial couples in the model generated data and AHEAD data. Each line represents the mean assets for a group of initial couples around the death (or no death) of a member of the household. Death date is normalised to occur at year 0. Solid lines: Couples who experience a death. Dashed lines: Matched couples who do not experience a death. Panel (b) reproduces Figure 3.4.

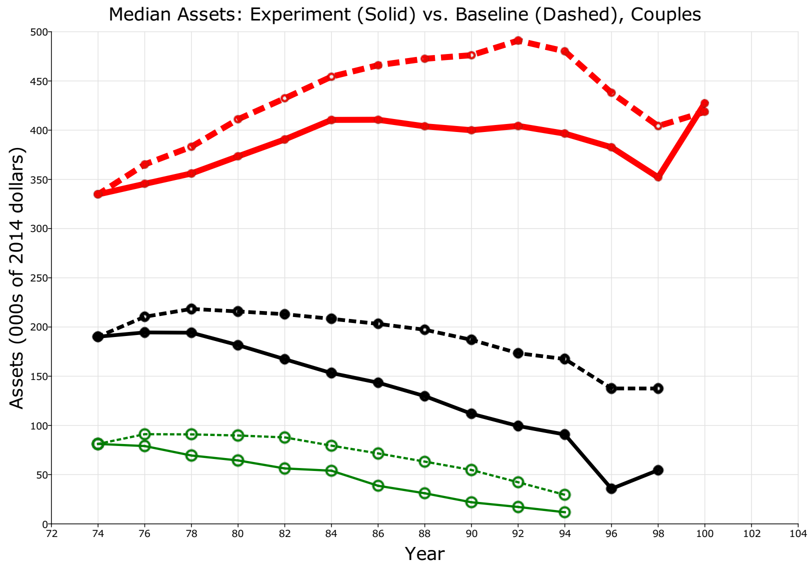
estimated framework reproduces the dynamics of household saving around these important life transitions. Furthermore, like our validation against Medicaid participation, this is not a target in our estimation procedure. Unlike our Medicaid participation, this uses a subset of the data in our target moments and demonstrates that our estimated model is consistent with panel data on the evolution of household wealth even if we disaggregate our moments and focus on specific subsamples.

3.7 What Determines Savings for Singles and Couples?

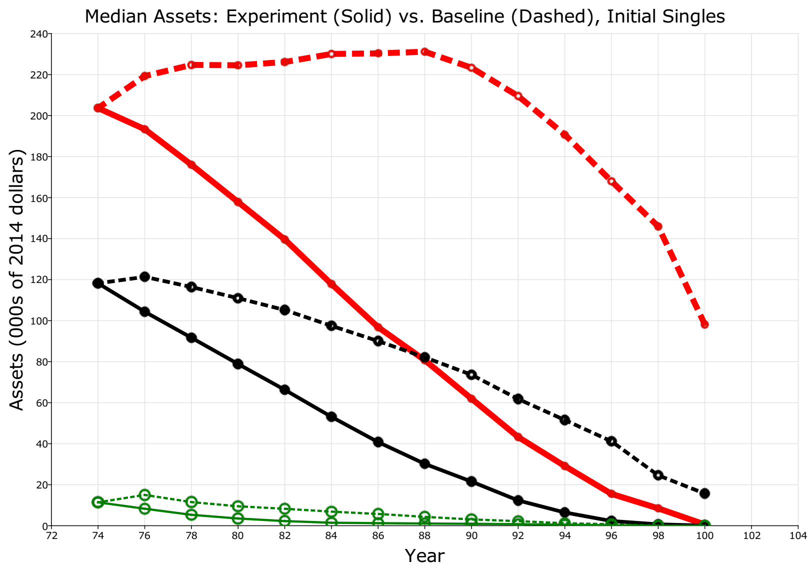
To determine the importance of the key mechanisms in our model and how the strengths of these mechanisms differ across retired couples and singles, we fix the estimated parameters at their benchmark values and then change one feature of the model at a time. For each of these different economic environments, we compute the optimal saving decisions, simulate the model, and compare the resulting asset accumulation profiles to the asset profiles generated by the baseline model. We display median asset profiles for the AHEAD birth year cohort whose members were aged 72–76 (with an average age of 74) in 1996. We present results separately for intact couples and the sample of initial singles.[11] We allow for differences in the realized mortality and the evolution of health between singles and couples which we simulate using our estimates of the processes described in Section 3.4 (except for the particular aspect changed in each experiment). As we document in Section 3.3, there are important differences between the mortality, health, and medical spending risks faced by couples and singles.

First, we ask whether the medical expenditures we measure in the data are important drives of the savings behavior for couples as well as singles. To answer this question, we zero out all out-of-pocket medical expenditures for everyone and look at the corresponding profiles for intact couples and initial singles. Figure 3.15 shows that the role of medical expenditures differs substantially when we compare across couples and singles. Medical expenses are a big determinant of the savings

[11] We do not show the results for the sample of all singles. These results are a combination of changes in the asset accumulation of initial singles, changes in the asset accumulation of couples which affect the wealth of new widow(er)s, and changes in the asset accumulation for those who have partners who have died.



(a) Couples



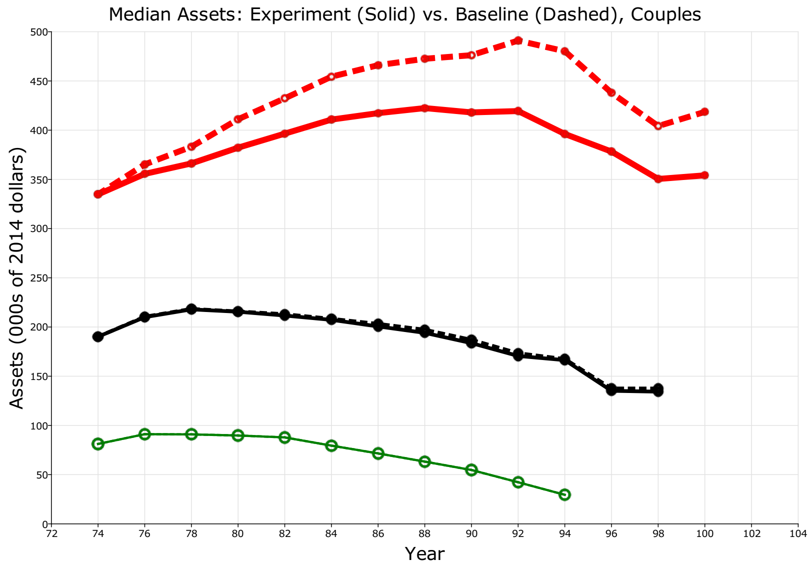
(b) Initial Singles

Figure 3.15: Median assets by permanent income tercile: baseline model (dashed lines) and model with no medical expenses (solid lines). Panel a shows couples and panel b shows initial singles

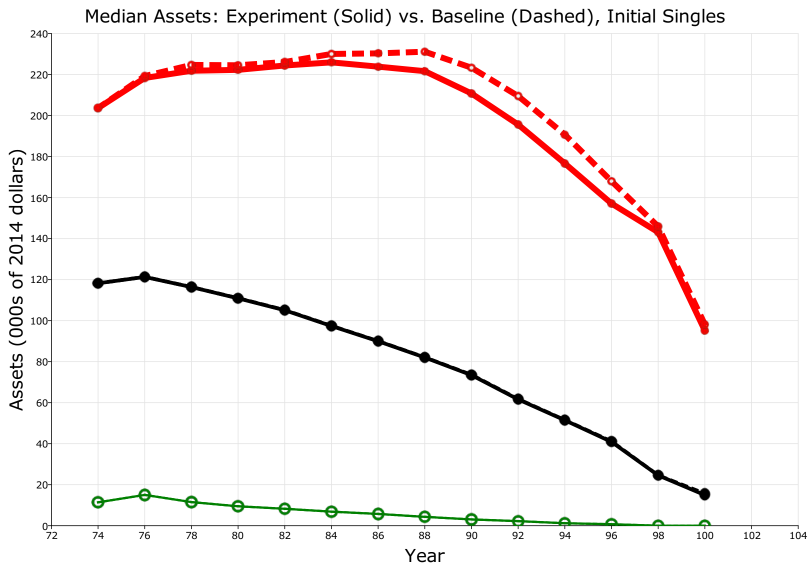
behavior of elderly singles (panel b), however, they are less important for elderly households in couples (panel a). Elderly singles reduce their current consumption in order to maintain higher levels of wealth and pay for the high medical expenses that occur later in life. This is true across the permanent income distribution - for singles at the top of the PI distribution these expenses are larger and they are relatively less insured by the government provided consumption floor. For the highest PI tercile, eliminating medical expenditures leads them to deaccumulate all of their assets by age 100. In our baseline, these households accumulate wealth, increasing their savings by around $30,000 in their 70s, and still hold almost $100,000 dollars at the same age.

In contrast, the effect on couples is much smaller suggesting that medical expenditures are only part of the explanation for why they continue to hold assets at advanced ages. Focussing on the middle PI tercile (who hold approximately the same level of initial wealth as the top PI singles), we show that if there were no medical expenditures their wealth would decline, but much less dramatically than singles. Relative to the baseline, they would reduce their savings by $50,000 at age 86, but continue hold assets late into life. By age 98 they would still have over $50,000 in savings. Unlike singles, the effects are smaller for those at the top of the PI distribution. The top PI group would still accumulate assets in their 70s and hold on to them well into their 90s. For couples, and in particular the richest couples, reductions in current consumption to pay for high out-of-pocket medical expenditures are small relative to other savings motives. When the government finances all medical expenditures, lifetime resources (to spend on consumption or to leave a bequest from) increase. This makes it more attractive to leave a bequest and as bequests are a luxury good this matters for richer couples. Our results show that while medical expenditures are one of the most important drivers of singles behaviors, they are only part of the picture for couples.

We next consider the role of our estimated bequest motives. We shut down, holding all other preference parameters constant, the utility from leaving a bequest ($\phi_0 = \phi_1 = \phi_2 = 0$). Figure 3.16 shows the results. For singles, this has close to zero effect on their median savings. Instead, for couples we find an effect for households in the top third of the PI distribution who still accumulate assets for much of the period, but do so by around $50,000 less. However, relative to their initial wealth they still hold around $15,000 more at age 100. That this only has an effect for the richest couples is perhaps unsurprising given that our estimated



(a) Couples



(b) Initial Singles

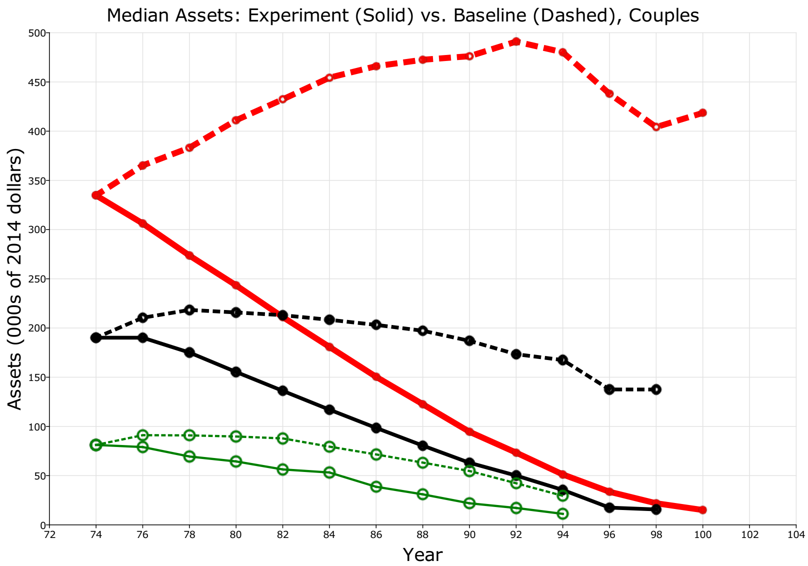
Figure 3.16: Median assets by permanent income tercile: baseline model (dashed lines) and model with no bequest motives (solid lines). Panel a shows couples and panel b shows initial singles

bequest motives are luxury goods. For these higher PI couples, removing the utility from leaving a bequest has a very similar effect on their saving as considering a world without medical expenditures. A household who holds assets in order to finance medical expenditures in the future may die prematurely without needing to pay for costly expenses. Likewise, they may live a long life, but not experience declining health or a large medical expenditure shock. In our baseline model, the saving to pay for this contingency is not wasted because the household would enjoy utility from the bequest that they leave in these cases. For households with an active bequest motive, this lowers the opportunity cost of self insurance which comes from foregoing lifetime consumption.[12] Our results suggest that for richer couples the interactions of bequest motives and medical expenditures may be important.

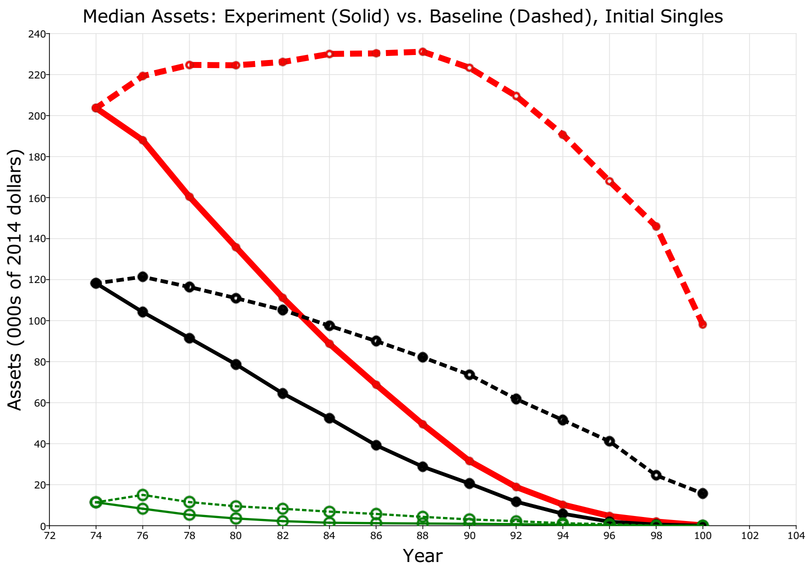
One reason why eliminating medical expenditures may only have a smaller effect on couples is that when their lifetime wealth increases, as it effectively does in this scenario, they choose to leave that wealth as a bequest to their heirs. Alternatively, they may hold on to a large proportion of it because they care about their spouse. We start by exploring this first effect and eliminate the utility from leaving a bequest in a world where the government pays for all medical expenditures - in effect, combining our first two experiments. We show the results from this experiment in Figure 3.17. Comparing the asset profile of couples in this experiment with our previous two we see large differences. The change is most dramatic for the top tercile of permanent income. Instead of accumulating assets in retirement, they increase their consumption and deaccumulate their wealth holding less than $40,000 by age 96 instead of the nearly $450,000 dollars they hold in the baseline. For the middle tercile couples, we find they increase their consumption - halving their assets by age 86 and retain only small amounts by age 96.

Given we find such small effects from eliminating the bequest motive it may not be clear why the effect of these motives can be jointly large. As we describe above, when households no longer have to finance medical expenditures they can use this money to increase their lifetime consumption, leave bequests, or for a couple leave it to the surviving member of a household. In addition, our results show that, for richer couples there are large interactions between precautionary savings motives and the bequest motive. Precautionary saving to finance medical expenditures is complimentary with saving for a bequest. In this experiment, the op-

[12] A similar argument applies to financing consumption against longevity risk and for the decision to self insure instead of purchasing formal Long Term Care insurance (see Lockwood, 2018, for an example).



(a) Couples



(b) Initial Singles

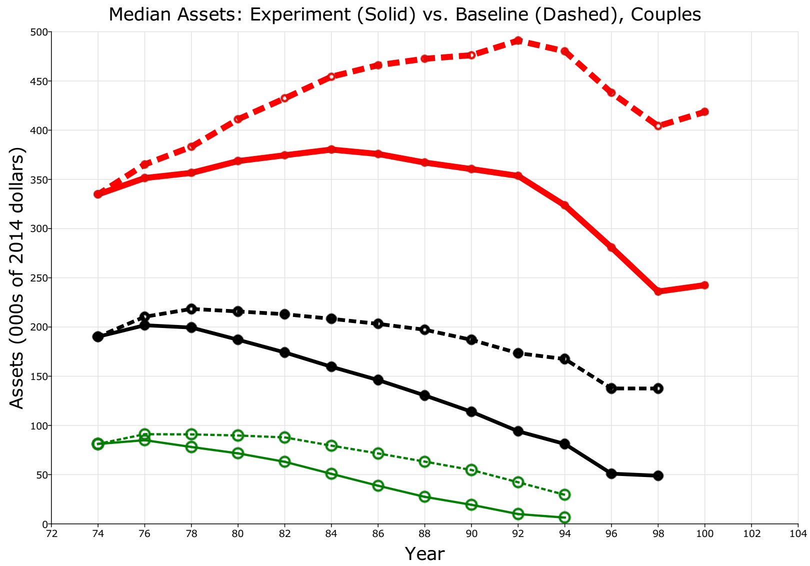
Figure 3.17: Median assets by permanent income tercile: baseline model (dashed lines) and model with no bequest motives and no medical expenses (solid lines). Panel a shows couples and panel b shows initial singles

portunity cost of dying with positive assets is large and households increase their consumption. Singles continue to hold assets to self insure against longevity risk. Couples also face longevity risk, but also changes to their income and the marginal utility of consumption when the first member of the household dies.

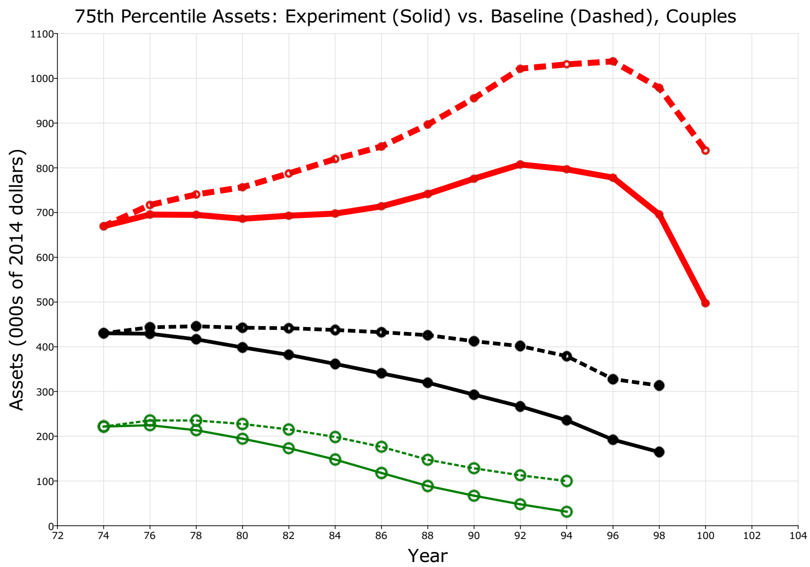
For the bottom tercile of the permanent income distribution, there is no difference between the behavior in this experiment and the experiment where we eliminate only medical expenditures. Even without medical expenditures, they still are not lifetime rich enough to leave a bequest so the additional change to their preferences has no bite. This suggests their lower levels of consumption and slow deaccumulation of assets is driven by a mixture of medical expenditures and a desire to maintain wealth for their spouse. Similarly, we see the overall speed of deaccumulation is slower for couples than it is for singles and even by age 98 the top two thirds of the permanent income distribution retain \$10,000 in assets. We explore the effect of their surviving spouse in the next experiment.

Finally, in Figure 3.18 we turn to understanding the role of the surviving spouse. We shut down, by setting the weight on the surviving spouse $\omega = 0$ in equation 3.17, altruism towards the surviving spouse for both members of the couple. By construction, this has no effect on the initial singles and so we additionally show the effect on the 75th percentile of the asset distribution for couples (panel b). In relative terms, shutting down altruism towards the surviving spouse has a similar effect at the median and the 75th percentile of assets although it is larger at the median. For the median and 75th percentile profiles, deaccumulation as couples age is much larger across the PI distribution. At the median the lowest PI tercile deaccumulates nearly all of their wealth by age 94, in contrast to the \$40,000 they hold in our baseline model. The middle PI tercile hold almost \$100,000 dollars less by age 98 and for the top tercile this rises to \$175,000. These results are similar order of magnitude to our first experiment where we eliminate medical expenses. Despite larger difference in absolute terms, the proportional effect of eliminating spousal altruism declines with both permanent income and the level of initial wealth.

Taken together, our results suggest that there are important differences in the key savings motives between couples and singles. For singles we find that medical expenses are the single most important factor and the precautionary savings motive explains a large fraction of their slow deaccumulation at advanced ages. In contrast, for couples we do not find a single savings motive that has the same quantitative importance. Instead, the interaction between bequest motives, medical expenses,



(a) Couples



(b) Couples - 75th Percentile Assets

Figure 3.18: Assets by permanent income tercile: baseline model (dashed lines) and model with no weight on the surviving spouse (solid lines). Panel a shows the median assets for couples and panel b shows the 75th percentile of assets for couples

and the desire to insure the surviving spouse is important.

3.8 Conclusions

Over one-third of total wealth in the United States is held by households over age 65. This wealth is an important determinant of their consumption and welfare.

As the U.S. population continues to age, the elderly's savings will only grow in importance.

Retired U.S. households, especially those who are part of a couple and have high income, decumulate their assets at a slow rate and often die with large amounts of assets, raising the questions of what drives their savings behavior and how their savings would respond to policy reforms.

We develop a model of optimal lifetime decision making and estimate key properties of the model. We find that singles live less long than people who are part of a couple, but are more likely to end up in a nursing home in any given year. For that reason, singles also have higher medical spending, per person, than people who are part of a couple. We also find that assets drop sharply with the death of a spouse. By the time the second spouse dies, a large fraction of the wealth of the original couple has vanished, with the wealth falls at the time of death of each spouse explaining most of the decline. A large share of these drops in assets is explained by the high medical expenses at the time of death. This suggests that a large fraction of all assets held in retirement are used to insure oneself against the risk of high medical and death expenses.

We use our estimated model to decompose and understand the importance of different savings motives for couples and singles. Our results for singles, using our rich model of health dynamics and medical spending, reinforce earlier findings - the risk of large medical expenditures at old ages explains the majority of their continued wealth holdings. Our results for couples, however, differ substantially. Couples, who are both richer and face different risks, do not save because of a single primary motive. Instead the interaction between bequest motives, medical expenses, and their desire to insure the surviving spouse all play an important quantitative role.

Chapter 4

Implicit Exchange, Intergenerational Transfers and Insurance within the Family

4.1 Introduction

Households face a variety of different risks over their life cycle. When markets are incomplete how do households weather different shocks? Do they self insure these risks as individual households relying on accumulated wealth or labour supply to insure themselves against idiosyncratic risk? Or do they rely on support from outside the household from the state, from friends, or from their wider family? By circumstance of birth, households often have pre-existing networks where risks are only partially correlated - namely their family. Understanding how and why households transfer resources across their family network is key to understanding the costs and benefits of providing insurance through the tax and transfer system. Furthermore, the level of intergenerational insurance has important implications for public policy including the effects of social security interventions on the young and old, how wealth inequality occurs and persists as well as Ricardian equivalence and how tax effects occur across generations.

This paper aims to evaluate the exchange (or strategic) motive for intergenerational transfers in a dynamic context. I use this to understand the lifetime value of inter vivos transfers. I develop a life cycle model of exchange motivated transfers among the extended family. Households bargain with their parents and children, swapping monetary transfers in exchange for gifts of time, supply labour, and save in risk free assets. I take parameters directly from the literature and show the model reproduces life cycle choices and key measures of transfers in the data. To my knowledge, this is the first paper to assess the exchange motive in a dynamic setting. Using this quantitative model, I show that the potential to self insure through the family (in addition to standard means of self insurance in liquid savings) has large benefits for households.

This has three main contributions. First, it extends the established static exchange framework to a dynamic life cycle model with‘overlapping generations. I present a computationally tractable model in which households can bargain over the surplus from inter-familial insurance by exchanging money and time. In considering the role of both financial transfers and gifts of time this paper explores the varied distribution of time gifts in the data, in addition to the provision of long term care. Second, I calibrate the model to assess the exchange motive's ability to generate life cycle choices and patterns of transfers consistent with data collected in the PSID. Finally, I use the implications of household choices under the exchange and altruistic motives to calculate a lower bound on the welfare value of family insurance. The extended family has shed light on key questions in the educational attainment of adults; the intergenerational persistence of inequality; the generational burden of long term care for the elderly and the investments into children. An important feature of many of these studies is the role of the insurance provided by the extended family.

Although there is evidence on the frequency and size of familial insurance, predicting household response to policy depends directly on the type and strength of the motive for intergenerational transfers. A large reduced form literature devoted to intergenerational transfers attempts to distinguish between various motives. A key debate surrounds attempts to separate two prominent motives for inter vivos transfers. An altruistic motive, where families care directly about their descendants, and an exchange motive, where families engage in mutually advantageous insurance. There is substantial evidence that families do not provide full insurance or efficiently share risks. However, attempts to test the candidacy of different motivations have proved inconclusive; the results of empirical tests have not cleanly identified a single motivation.

Assessing the motives for why families make transfers is important. Different theories of inter vivos transfer behaviour predict different responses to policy interventions as well as different burdens of policy reform. An altruistic theory predicts that public transfers can crowd out private transfers leaving the total effect for the poor or old ambiguous as their private transfers reduce. Under the exchange theory increasing the outside option through social security increases leverage and may magnify the effect of the social security measure. Additionally, the scope, frequency and value of the familial insurance under different policies explicitly depends on the motives the extended family has for providing the insurance. Despite this, the prevailing modelling choice in both the macroeconomic and applied microeconomic literatures is to assume households either exist without an extended family or incorporate the extended family by assuming altruistic connections. This paper departs from this norm, to highlight the role of the exchange motive in a dynamic setting and shows that it can produce a number of key aggregate data feature much like altruistically motivated transfers, To do this it sets out a model of the joint determination of financial and time gifts.

This paper is related to a large literature that studies intergenerational transfers. The first strand of the literature attempts to test the motive for families providing intergenerational transfers and the extent of risk sharing within the family. Instead of using the degree of risk sharing as a test of the altruistic motive, I assume the exchange motive and use a quantitative model to infer a lower bound to the lifetime value of this potential insurance. Altonji et al. (1997) centre their test on redistribution of resources between generations. In a static model, if transfers from parents to children are non-zero, redistributing one dollar of income from the parent's income to the child's income will result in a reduction of the transfer from parent to child by one dollar (under perfect altruism or perfect risk sharing). Using data from the 1988 PSID they find redistribution in the region of 0.04-0.13 cents and reject the hypothesis that the extended family provide perfect insurance against income risk.

A similar approach is used by Altonji et al. (1992) and Choi et al. (2016) who test the joint distribution of consumption between family members- finding little evidence in support of altruistic risk sharing. Fluctuations in dynastic income do not affect ones own consumption, however, there is some evidence of varying marginal propensity to consume from transitory income by extended family wealth. Both Cox (1987) and Cox and Rank (1992) present results that reject perfect familial insurance. Using a restriction on the sign of the derivative of transfers with respect to income they present limited evidence that rejects even partial altruism at the family level. Similarly, using PSID data Attanasio et al. (2015) directly consider the extended family as a partial insurance mechanism and find that over 60% of shocks could be insured by the family network. However, there is no evidence that shocks are insured. Finally, exploiting longitudinal data McGarry (2016) stresses the importance of the life cycle and life cycle events in determining the receipt and value of transfers. This highlights the importance of dynamic concerns considered here and of considering alternative motives for intergenerational transfers.

A second branch of the literature incorporates intergenerational transfers into dynamic models. Many of these papers focus solely on monetary transfers. Early work by Laitner (1988, 1992) incorporated altruism into a dynamic framework and generated endogenous patterns of transfers as luxury goods as well as analysing the long run effects of transfers on a dynasty. By calibrating a dynamic general equilibrium model with transfers to target the annual flow of bequests and trusts Nishiyama (2002) attempts to quantify the level of altruism between parents and children in a very limited life cycle framework. Despite assuming altruistically motivated households both of these papers generates strategic interaction between parents and children due to the complicated incentives arising from transfers in multiple periods (and the possibility of rent seeking behaviour by children who may not have direct preferences over their parent's utility) as well as different partial altruism generating disagreement over the optimal transfer.[1] Kaplan (2012) incorporates transfers from parents to children into a model of residence choice and self insurance. However,

[1] This is also discussed within the family decision making literature. See, for example, Del Boca and Flinn (2012)

in this model parents are unable to save or receive transfers from their children and it only attempts to model early adulthood. Abbott et al. (2018) allow for partially altruistic parents in a study of educational choice and under this assumption predict significant crowding out within a general equilibrium environment. Both papers find an important role for intergenerational transfers as an insurance opportunity for children. However, despite targeting average financial transfers Kaplan finds minimal levels of altruism (around 4% of the weight on own utilty) when compared to the the much larger level (approximately 25%) found in Abbott et al. as well as failing to account for the rich heterogeneity in transfer amounts.

In recent years, a number of papers have focussed on time transfers in dynamic models of intergenerational transfers, specifically focussing on the question of Long Term Care. Fahle (2015) has a model with no saving for the child, a coresidence decision and provision of long term care. Similarly, Ko (2018) focusses on how strategic interactions with children affects the demand for Long Term Care Insurance in a model with overlapping generations. Mommaerts (2016) estimates significant risk sharing, but limited evidence of partial altruism and shows that this rationalizes under-utilization of Long Term Care Insurance, but does not model early life cycle transfers. Barczyk and Kredler (2018) have a continuous time model with partial altruism and savings for both households, but treats labour supply and the opportunity cost of transfers as exogenous. A smaller literature of structural models estimates incorporates intergenerational transfers, but estimates the transfers exogenously as in Keane (2002). The literature has, however, neglected to study the dynamics of family transfers of money and time with both savings for all households and endogenous opportunity costs of time transfers.

A final branch of the intergenerational transfer literature specifically studies time gifts and the provision of informal long term care by children. This paper documents time gifts more generally and extends the modelling environment to include other uses for these gifts while allowing for health and informal care to be substitutes. Byrne et al. (2009) study the choice between formal health care and the provision of informal care by the family, they find large variation in the burdens of informal care that is correlated with the opportunity costs of potential caregivers. Studying informal care giving in the National Longitudinal Study Norton et al. (2013) find evidence consistent with financial transfers as rewards for the provision of informal care. Similarly, using data from the HRS, Groneck (2016) finds evidence of bequests as compensation or rewards for informal care, notably this effect requires a codified will.

An outline of the paper follows. In the next section I present descriptive statistics that summarize the distribution of family transfers between adult US households as well as evidence of the importance of transfers for life cycle decisions. Section 4.3 describes the life cycle model and Section 4.5 describes the calibration process. Section 4.6 discusses the results of the calibration exercise. Section 4.7 evaluates a lower bound for the welfare value of family transfers. Section 4.8 concludes.

4.2 Evidence of Family Transfers

In this section I document evidence of family transfers from the 2013 Roster and Transfer Supplement (hereafter RTS) to the PSID. It collects data on living parents and adult children (including spousal parents who don't have the 'PSID gene') to document information on family members including age, educational attainment and family structure (e.g. marriage/partner status and children). This provides information on the extended families of households who do not posses the 'PSID gene' and are omitted from the standard PSID sample. A full discussion of the 'PSID gene' and procedure used by the PSID to sample across generations is provided in appendix C.4.1. In addition to rostering the family, the RTS also contains information on a variety of transfers made between the extended family.

The RTS draws two distinctions in surveying the PSID sample. First, the RTS asks respondents separately about gifts of time and money. Second, the RTS draws a distinction between short term and long term gifts. Short term gifts are transfers and gifts made within the last calendar year (2012 for the PSID 2013 wave). Long term gifts are gifts given or received during the entire adult life, since turning 18. In this section I summarize key facts characterizing inter-vivos transfers in the US.

Example questions are provided in appendix C.4.3. To avoid conflating gifts or transfers that may arise due to a co-residence arrangement (such as transfers made to pay a mortgage or rent) these questions are asked of separate households. In order to fully characterize the history of gift giving, questions that refer to long term gifts allow for the current co-residence of the parent and child.

Questions asking individuals about their financial and personal relationships with other family members may evoke a variety of emotional responses which can lead to concerns about the accuracy of the survey responses. In addition to approximation error or transposition error, further inaccuracies may exist due to recall bias or spousal secrecy. When one spouse receives or gives a gift without informing their partner the responses given to the RTS will not be an accurate description of their transfer behaviour. A large and well documented problem is that social pressures may lead to under or over-reporting for households who are at different parts of the transfer distribution. Respondents may wish to overstate their generosity while understating their dependence on help from others.

Gale and Scholz (1994) report that respondents in the Survey of Consumer Finances (SCF) substantially under-report the amounts they have received relative to reported gifts in the same survey. In part, this is attributable to the $3,000 threshold for including gifts- two children receiving $2,500 each would not have associated transfers in the SCF, but the total gift of $5,000 is reported for the parent. Two features of the RTS survey design mitigate concerns of reporting errors relative to other survey data. First, the PSID attempts to contact both parents and children. Knowing that the survey has the opportunity to verify the information provided likely reduces the incentive to intentionally misreport (although unintentional inaccuracy is not addressed). Second, while deliberate under-reporting or recall errors are still possible, the lower threshold of $100 in the RTS eliminates under-reporting created by the design of the survey. However, the sampling choices in the PSID (which does not oversample the wealthiest individuals in the United States) may bias measures of the average transfer downwards. If the transfer behaviour of households is increasing in wealth, this suggests the results in this paper characterise a lower bound

	Any Activity	As a Giver	As a Receiver	As Both
All	80.7%	64.9%	61.6%	45.7%

Responses from PSID Heads calculated from the 2013 PSID & RTS.

Table 4.1: Transfer Activity after Age 18

for transfer activity within the US.

Table 4.1 provides a cross sectional picture of the frequency of transfer activity in the United States. I pool all PSID households in the 2013 PSID and construct frequencies for transfer activity using the long term gift questions in the RTS. In order to calculate the frequency of transfer activity since turning 18, for each category of gift[2] households are assigned an indicator which denotes if they are a giver, a receiver or neither. The neither categorisation includes those households for which the respondent refuses to answer or answers “Don’t know”. Over 80% of households report engaging in transfers after the age of 18.

However, the exact nature of these transfers represent a considerable empirical puzzle given the limited evidence of efficient risk sharing in the literature discussed above. The high frequency of transfer activity is economically significant, but at odds with credible evidence on the lack of insurance provided by the extended family. Furthermore, transfer activity may be understated because the pooled cross section approach does not account for the age of the household. If households are more likely to give or receive transfers as they age then this result will not accurately capture the lifetime frequency of transfer activity. Similarly, transfers may be understated if older households are less likely to recall or report transfers received in their youth.

When accounting for both gifts of money and time, the frequency of giving and receiving are similar (columns 2 and 3). Understanding the decisions and expecta-

[2]In the raw data, in addition to distinguishing between long term and short term gifts, households are asked about transfers for a subset of specific expenses. I pool together all givers or receivers across all categories of gift giving.

	Fraction Non-zero	Mean (> 0)	% Family Income	10th Perc.	50th Perc.	90th Perc.
Money Given	28.93%	5236 $	7.00%	258.1 $	1548 $	11355 $
Money Received	15.59%	3364 $	12.81%	154.8 $	825.8 $	7329 $
Time Given (hrs/wk)	33.92%	9.380	n/a	0.385	2.885	24
Time Received (hrs/wk)	21.10%	10.26	n/a	0.269	2.692	28

Responses from PSID Heads calculated from the 2013 PSID & RTS. Dollar amounts in 2015 US$. I calculate the net transfers between pairs of households in the case that both households report positive giving.

Table 4.2: Transfers in the Last Calendar Year

tions of both the givers and receivers is necessary to correctly estimate the value of these transfers of the life cycle. Nearly half the current adult US population has both given and received transfers from a family member within their life time. As with the effect of household ageing on the results in the first column this does not reflect the true probability of being both a giver and receiver. However, it does suggest that this is a common form of intergenerational interaction. Reciprocally transferring with another generation or moving from giving to receiving as they age are important features of the life cycle.

Table 4.2 reports the mean values for transfers given in the past year. Approximately 1 in 5 households give short term gifts or loans above the $100 threshold to their immediate family members. The rate of receiving these monetary transfers for adult children under 30 is greater than that of the total population while older households are more likely to give. Strictly positive monetary gifts are unequally distributed.[3] However, as participation only accounts for a fraction of the population the unconditional distribution has a much larger positive skew. The mean

[3]This inequality in gifts is similar to the dispersion in US wealth when measuring inequality by the Gini coefficient. However, it is important to note that two very different distributions may have the same coefficient despite large differences in their overall distribution.

monetary transfer given is 7% percent of family income and these transfers have a high variance. The impact of a transfer is likely to be greater for those under 30 (where both the mean size of transfer received and proportion of the population receiving is larger than the total population) as those under 30 typically have lower incomes, lower assets and are more likely to be liquidity constrained. In a standard buffer-stock saving model, receiving money early in life would help households insure against idiosyncratic risk and relax their borrowing constraints. Relative to their wealth, the size of transfers for this group is larger consistent with the evidence on the size of monetary transfers received. Consequently, family links may mitigate their exposure to idiosyncratic risk or, alternatively, increase inequality within a given cohort.

Households who give positive financial transfers report larger amounts, but these transfers are smaller proportions of family incomes for the giving households in proportional terms even though they are on average smaller gifts. The value of any gift to the recipient is larger than the cost to the giver which is consistent with a view that considers the extended family as a source of insurance.

While previous attempts in the literature are able to recreate some features of an implied monetary transfer distribution(e.g. Laitner (1992) or Nishiyama (2002)), these approaches do not consider gifts of time which occur more frequently. The average gift of time in 2012 is around 20 percent of the average US working year. This is substantial, but it is difficult to directly compare with monetary transfers. Using a household's market wage as a measure of their opportunity cost, the average time gift represents a larger share of income than the average monetary gift.[4] The prevalence of time gifts emphasises that any model of intergenerational transfers consistent with the RTS data would need to consider the motivations for gifts of time as well as monetary gifts.

As with the distribution of monetary gifts, transfers of time also exhibit large skew. The skewed distribution of these transfers highlights a shortcoming of the

[4]This assumes the correct opportunity cost is the market wage of the giver or receiver. A sensible alternative is to use the opportunity cost of a market substitute. Valuing these gifts using the minimum wage delivers similar results.

approaches used in the intergenerational transfer literature. Matching the means of these distributions ignores the large amount of heterogeneity at the household level. A credible theory of household behaviour must account for the role of household heterogeneity in determining transfers (such as the desire to insure income risk across generations) as well as the rich heterogeneity in the distribution of observed transfers.

The skewed distribution of these time gifts suggests a number of different qualitative classifications of time gifts. A small fraction of households report amounts consistent with 24 hours a day, 365 days a year and approximately the top 2% of positive time gifts constitute a minimum of 35 hours a week. At the top of the distribution, these amounts are consistent full time informal care for a family member. Full time informal care need not be constant activity, such as bathing, cleaning or cooking, but may require individuals be on call or constantly present. These requirements will impose considerable costs on the care giver far beyond the 'active' time given. In contrast, households who provide very small amounts of time may be more likely to be providing one of gifts for typically menial tasks. While those who provide large amounts of time that equate to small weekly amounts provide support for large projects or a different form of informal care. Other forms of informal care includes assistance with activities (e.g. groceries, bills, cleaning) on a weekly basis that are not necessarily strict substitutes for nursing home residence, but instead they allow households to maintain their independence as they age or their abilities decline. It is this distribution of time transfers related, but not equivalent to the Long Term Care decision that motivates the choice of utility from time gifts in the model proposed in this paper.

4.3 A Life Cycle Model of the Exchange Motive for Inter-Vivos Transfers

The model is a discrete time, incomplete market, overlapping generations model with an infinitely lived government and intergenerational bargaining. Households maximise lifetime utility by choosing assets, consumption, labour supply transfers between the family in each period. Table C.1 in Appendix C.1 summarises all the parameters described in this section.

4.3.1 Demographics and Life Cycle

The decision-making unit in the model is the household. All households contain a married couple who start their working life already married. Households die stochastically. The economy consists of J overlapping generations. Let $j = 1,...,J$ denote the age of an agent and η_j denote the probability of surviving from age j to $j+1$. η_j is assumed to be 1 before the household retires ($j < j^{RT}$) and to be 0 conditional on reaching the final period ($j = J$). A household's child is 'born' an adult at age $j = j^{CH}$ at which point the child enters the economy as a $j = 1$ household. The birth of a child is assumed to be exogenous and each household has only one child household who enters the economy as a married couple.

In the first period in which their child enters the economy parents ($j = j^{CH}$) and children ($j = 1$) can bargain over a transfer of assets in exchange for a flow of (delayed) informal care services. This period consists of two subperiods. In the first subperiod it is assumed that all income uncertainty has resolved and the bargaining process takes place under complete information. In the second, households choose their labour supply (if working), savings, and consumption as in all other periods. The child is able to perfectly commit to supplying services in later life. Dividing each period into two subperiods simplifies exposition and, as presented here, eliminates tied transfers. With full commitment, tied transfers which must be used to purchase a particular consumption good or service are always weakly suboptimal in the bargaining environment of the model.

Parents and children commit to this flow of services and there is no renegotiation. They bargain over the transfer and flow of services subject to the constraint that the net flow of services supplied by children has to be weakly positive. In other words, children are not receivers of services.[5]

[5] An alternative approach would be to incorporate flows of time from parent to child as childcare and explicitly model the child care needs of households with children below working age. This would also reconcile additional data on the frequency with which grandparents provide care for their grandchildren.

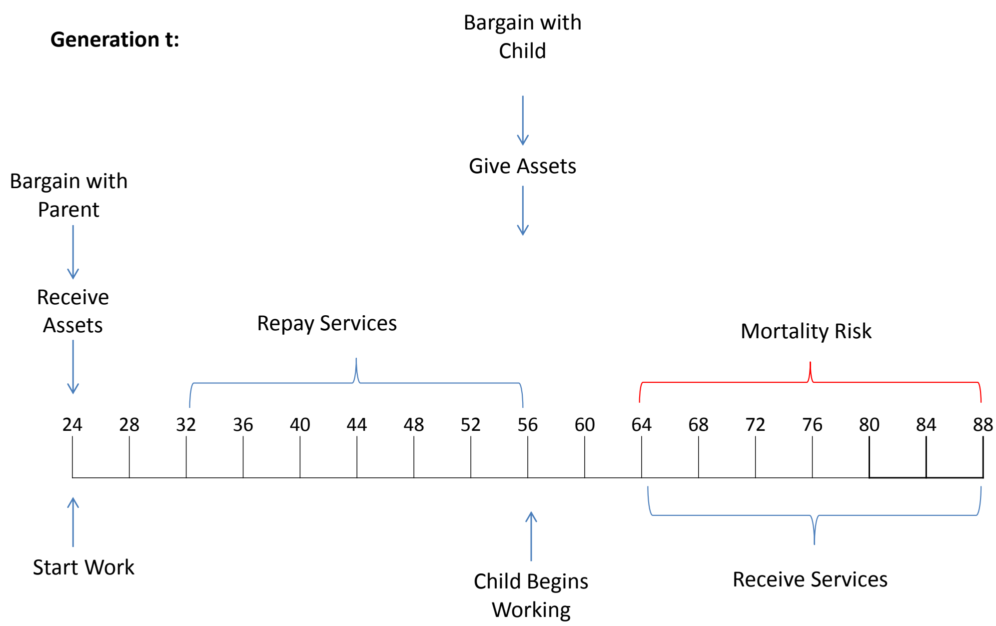


Figure 4.1: The life cycle of a household

Services are repaid by the child during the parents retirement: ages $j \geq j^{RT}$ for the parent and ages $j^{RT} - j^{CH} \leq j \leq J - j^{CH}$ for the child. The previous generation are assumed to die before the next generation enters the economy; it is not possible to simultaneously be a parent and child. Consequently, all services are repaid to the prior generation before the next generation is born.

Figure 4.1 presents the structure of the life cycle for a given generation and Figure 4.2 presents the structure of the overlapping generations for specific choices of the retirement and birth ages used in the calibration. Households enter the model at 24, retire at 64 and die with certainty by the age of 88 with each period lasting four years. Both of these figures assume that the households live to their maximum life span. If their parents die before age 88 households are assumed to no longer repay services. During retirement parents receive care services, face health, mortality, and medical expenditure shocks and receive a pension. Retired households do not supply labour.

There are five stages in the life cycle of a household, three distinct stages as

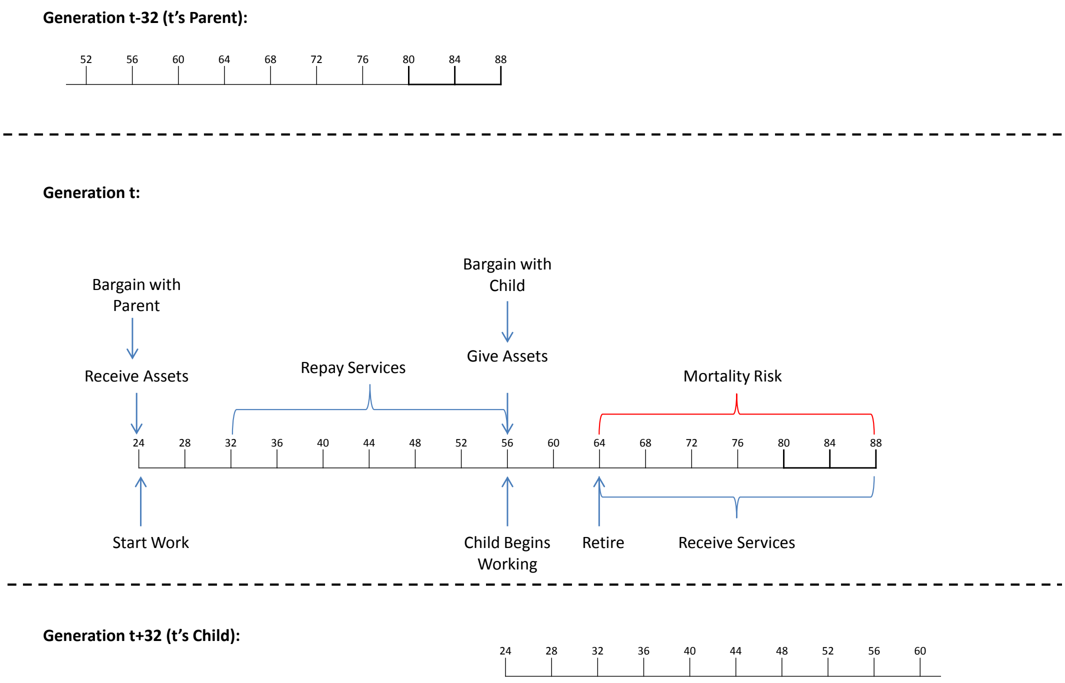


Figure 4.2: The model demographics

a child and a further two as a living parent. For the child, who is always working, the stages are: the periods before their parent retires and services begin to be repaid ($1 \leq j < j^{RT} - j^{CH}$), the periods where they repay services, and the periods they are orphaned (if any). Once a parent dies informal care services are no longer repaid. For the parent: the first stage is when the parent is working before receiving services ($j < j^{RT}$), and the second is retirement ($j \geq j^{RT}$).

4.3.2 Preferences

Preferences are assumed to be time separable, with a constant discount factor β. Utility from consumption is monotone increasing and multiplicatively separable from consumption. Child households receive dis-utility from supplying services, s, to their parents and parents derive utility from receiving services.

Services received are assumed to be imperfect leisure substitutes and the cost to a household of supplying services are time costs in forgone leisure. This is consistent with the measure of services as time gifts in the RTS and captures the endogenous response to changes in their opportunity costs.

Letting s_j denote the net supply of services in period j and h_j denote the labour supply in period j, then leisure in period j for a child household, l_j^k, is given by:

$$l_j^k = 1 - h_j^k - \kappa \cdot I_{h_j^k > 0} - s_j \cdot I_{j^P \geq j^{RT}} \tag{4.1}$$

Where κ is the fixed cost of working. This assumes that the leisure cost of supplying time gifts is independent of the parent's health status, but that services are only repaid when the child's parent is past retirement age.

For the parent household in period $j < j^{RT}$ leisure is given by:

$$l_j^p = 1 - h_j^p - \kappa \cdot I_{h_j^p > 0} \tag{4.2}$$

Which is identical to the leisure of the child before the parent retires. For periods $j \geq j^{RT}$, the living parent is retired (and receiving services). Let m_j denote the health status of a parent household at age j and their leisure is given by:

$$l_j^p = 1 + (\delta s_j - \delta m_j) \tag{4.3}$$

Where δ controls the leisure cost of ill health and $\delta \in (0,1)$ is required so that total leisure remains positive.

This formulation allows for health dependent utility. When household receives no time this is a re-normalisation of the Palumbo (1999) preference shifter for ill health and when $\delta = 0$ utility is independent of both child time and health status[6]. Households in ill health without time from their children face higher marginal utility of consumption and (compared to households with higher amounts of child time) generate higher expenditures- a mechanical effect consistent with the literature on informal long term care and endogenous medical expenditure (Ameriks et al., 2018). Care services may effectively bundle labour, which has a market substi-

[6]The modified Palumbo functional form for total leisure implies total leisure for parents is convex in services. Convexity of period utility in leisure for parents and concavity for children is sufficient for uniqueness of the bargaining game outcome with the utility function used throughout. The necessary and sufficient conditions for uniqueness are presented in Appendix C.2

tute, and affection[7] which implies a more complicated outside option.

The within-period utility function is assumed to be isoelastic and Cobb-Douglas between consumption and leisure in time j, utility is given by:

$$u(c_j, l_j) = \frac{(c_j^{\sigma}(l_j)^{1-\sigma})^{1-\gamma} - 1}{1-\gamma} \tag{4.4}$$

In this specification γ is the coefficient of relative risk aversion and σ is the weight of consumption in relation to leisure. The complementarity of leisure and consumption depends on the value of γ. For any γ greater than 1 leisure and consumption are substitutes. For all reasonable parametrizations of γ this condition is satisfied; in the discussion below consumption and leisure are treated as substitutes.

Utility from bequests, b, is net of taxes and takes the form of a warm glow bequest motive as in De Nardi (2004). $\phi(b)$ is given by:

$$\phi(b) = \frac{\phi_1(\phi_2 + b)^{\sigma(1-\gamma)} - 1}{1-\gamma} \tag{4.5}$$

Where ϕ_1 controls the relative weight of bequests and consumption while ϕ_2 controls the curvature. Therefore, ϕ_2 controls the extent to which bequests are a luxury good. For positive ϕ_2 marginal utility of small bequests is bounded, while marginal utility of large bequests declines more slowly than consumption.

4.3.3 Earnings, Health Status and Medical Expenses

Earnings When employed, households earn their productivity, whose log (equation (4.6)) is the sum of a deterministic component $f(j)$ and a stochastic component (u).

$$ln\ e_j = f(j) + u_j \tag{4.6}$$

u is the sum of an AR(1) component ψ and a transitory white noise component υ

[7] Simulated or genuine affection which may be an important part of other time transfers.

$$u_j = \psi_j + \upsilon_j \tag{4.7}$$

$$\psi_j = \rho \psi_{j-1} + \varepsilon_j, \quad \forall j > 1 \tag{4.8}$$

$$\varepsilon_j \sim N(0, \sigma_\varepsilon^2), \quad \forall j > 1 \tag{4.9}$$

I assume the white noise component υ is measurement error with a constant variance.

A considerable literature (see for example Solon, 1992; Chetty et al., 2014), have documented large long run correlations between the incomes of parents and their children. As households in the model are forward looking and predict their future interactions with the next generation it is important to account for the joint distribution of these initial incomes. Upon entering the model at age $j = 1$, I assume that idiosyncratic productivity can be decomposed into a persistent element inherited across generations and a white noise component. For child household k and parent p this is given by:

$$\psi_1^k = \rho_{inherit} \psi_{j_{CH}-1}^p + v_{k1} \tag{4.10}$$

$$v_{k1} \sim N(0, \sigma_{k1}^2) \tag{4.11}$$

Health Status Before retirement, health status is drawn from a degenerate distribution and is assumed to be "good". After retirement health status can take one of two values

$$m_j = \begin{cases} 1 & \text{if in bad health} \\ 0 & \text{if in good health} \end{cases}$$

and transitions according to an age dependent Markov process.

Medical Expenses In retirement households are exposed to out of pocket medical expenses, M_j. The approach to modelling out of pockets medical expenses draws on De Nardi et al. (2010) by allowing for persistence in medical expenses. This captures an important risk facing households at older ages and generates a precautionary savings motive over the life cycle. Log medical expenses (conditional on health status and) are given by the sum of a deterministic component and a person specific component:

$$ln\ M_j = H(m, j) + \sigma_M(m, j) \cdot \zeta_j \tag{4.12}$$

French and Jones (2004) find the person specific component is volatile and persistent even after accounting for health status. This is well represented by the sum of a Gaussian AR(1) process μ and a normally distributed transitory process ν.

$$\zeta_j = \mu_j + \nu_j \tag{4.13}$$

$$\mu_j = \rho_M \mu_{j-1} + \omega_j \tag{4.14}$$

$$\omega \sim N(0, \sigma_\omega^2) \tag{4.15}$$

$$\nu \sim N(0, \sigma_\nu^2) \tag{4.16}$$

Both health status and, consequently, medical expenses are exogenous.

4.3.4 Taxes, Government Transfers and Budget Constraints

Taxes & Government Transfers The government taxes labour income, capital income, and estates to provide the flat pension to retirees, p, and a consumption floor, $\underline{c}$.

Following Bénabou (2002) I approximate the federal income tax and social security programs with the following flexible parametric tax function. Total after tax income on labour income, y, is given by:

$$\tau_l(y) = \tau_{l,1}(y)^{1-\tau_{l,2}} \tag{4.17}$$

Where $\tau_{l,1}$ controls the average level of taxation and $\tau_{l,2}$ controls the progressivity of the combined income tax and social security schedule.

Capital income tax is a flat rate τ_a. Estates larger than a given exemption level, ex_b, are taxed at rate τ_b on the excess above ex_b.

It is assumed that government transfers provide a consumption floor $\underline{c}$ and transfers to retired households, b_M, are given by:

$$b_M = max\Big\{0, \underline{c} + m_j - \Big[1 + r(1-\tau_a)\Big]a - p\Big\} \tag{4.18}$$

Thus, government transfers make up the difference between total assets and minimum consumption for households with a large medical expenditure shock. This requires households to first exhaust their stock of private assets.

Budget Constraints Households are unable to borrow and can save in a single risk free asset which accrues returns at rate r. Letting T denote the net asset transfer from parents to children the budget constraints for retired households are given by:

$$c_j + a_{j+1} \leq a_j(1 + r(1-\tau_a)) + p - T_j - M_j + b_{M,j} \tag{4.19}$$

While working age parents and children face:

$$c_j + a_{j+1} \leq a_j(1 + r(1-\tau_a)) + \tau_l(e_j h_j)) - T_j \tag{4.20}$$

4.3.5 The Household's Recursive Problem

The economy is populated by J overlapping generations. Let $j = 1, ..., J$ denote the age of the household. Households with age $j < j^{RT}$ are working, while households with age $j \geq j^{RT}$ are retired. Retirement is exogenous. Bargaining takes place in the first subperiod of j.

During the second subperiod of each period a working agent chooses consumption c, labour supply h and their stock of risk free assets for the next period a'.

Letting p and k denote parent and child respectively, the dynasty's state variables are given by $X = (\tilde{a}^k, \tilde{a}^p, s, e^k, e^p, m^p, M^p, j)$.

These variables are, the child in the dynasty's assets carried from the previous period ($\tilde{a}^k$), the parent in the dynasty's assets carried from the previous period ($\tilde{a}^p$), the agent's committed flow of services (s), the child's productivity (e^k), the parent's productivity (e^p), the parent's health status (m^p), the parent's out of pocket medical expenses (M^p) and the agent's age (j). As health status and medical expenditure risk only occurs in retirement the child's status and expenditure can be dropped without loss of generality. Similarly, the age gap is exogenous and one age from the dynasty can be dropped without loss of generality.

The states for a household are given by $x = (a, s, e, m, M, j)$ for both parent and children (for whom m and M is degenerate). In the recursive problem of the household the relevant asset state variable is a the after bargaining asset level. The states of the dynasty without a surviving parent household are those of the surviving child household. The sequence of second subperiod problems can be solved treating the bargaining rule as an exogenous transition between states. Equilibrium requires that expectations for this transition are consistent with optimal behaviour. It is this sequence of recursive problems presented here.

To maintain computational tractability, I restrict the household information set after the bargaining period to include only the state variables of the individual household and not their counterpart in the dynasty.[8] As each pair of households only bargain once at the start of the child's life this restriction only binds for the distribution of bequests. I restrict bequest expectations such that each period in which their parent may die, conditional on their parent dying, a couple expects a bequest drawn from the distribution of bequeathable assets (endogenously determined by the asset choices of parent households in retirement) conditional on the known level of s and age.

The outcome of the bargaining game is determined uniquely by the ages, asset positions, health status, medical expenses and incomes of parents and children (i.e. the states of the full dynasty). In this formulation each parent child pair only bargain once and a number of these states are degenerate. I solve for a Markov Perfect Nash bargaining solution. Equilibrium outcomes T the monetary transfer from parent to child and s the updated services transferred from child to parent are a deterministic function of the states of a parent and child:

[8] The alternative is possible, but under the current solution algorithm at a prohibitive computational cost

$$\begin{pmatrix} T \\ s \end{pmatrix} = \Upsilon(X)$$

Taking the outcomes of the bargaining game as a vector-valued deterministic function of the states of a parent and child, $\Upsilon(X)$, the actions of their counterparts as given, and consequently the distribution of counterparts conditional on s as given, an agents optimal decision rules are functions for consumption $c(x)$, labour supply $h(x)$ and next period's assets $a'(x)$ that solve the dynamic problem below.[9]

(i) At age $j < j^{CH}$ (after the bargaining game) the child household enters the second subperiod with after transfer assets a, care obligations s, idiosyncratic productivity e health status m (which is always equal to good health) and medical expenses M (always equal to 0 before retirement). The child's problem is given by:

$$\begin{aligned} V(a,s,e,m,M,j) = \max_{c,h,a'} \Big\{ & u(c,l) + \beta \eta_{j^P} E_j \Big[V(a',s,e',m,M,j+1) \Big] \\ & + \beta (1-\eta_{j^P}) E_j \Big[V_{orphan}(a'+b',e',j+1) \Big] \Big\} \end{aligned} \quad (4.21)$$

s.t.

$$c + a' \leq a(1 + r(1-\tau_a)) + \tau_l(e_j h_j) \quad (4.22)$$

$$l = 1 - h - \kappa \cdot I_{h>0} - s \cdot I_{jP > j^{RT}} \quad (4.23)$$

$$a' \geq 0, \quad c \geq 0, \quad l \geq 0, \quad h \geq 0 \quad (4.24)$$

The total expectation for next period is decomposed into an expectation conditional on the survival of the household's parent household and a second expectation conditional on the death of their parent household. The first expectation is taken over the next period's realisation of idiosyncratic productivity. The second expectation, when the parent household dies, is an expectation

[9] Throughout I use $j^P = j + j^{CH}$ to denote the parent's age. The age gap between generations is exogenous which means this is a deterministic function of the child's age and a redundant state variable.

over the realisation of the household's idiosyncratic productivity and the level of inheritance received (conditional on s).

In $j = 1$, a is equal to the initial transfer as the only assets a child has access to are the transfers from their parent household.

(ii) At age $j^{RT} - j^{CH} < j < j^{CH} - 1$ the orphaned agent enters the period with asset level a and idiosyncratic productivity e all other states in the vector x are equal to zero and dropped from the equations for notational convenience. Their parent is dead so they do not have an opportunity to bargain, pay no services and receive no transfers or bequests. Their problem is given by:

$$V_{orphan}(a,e,j) = \max_{c,h,a'} \left\{ u(c,l) + \beta E_j \left[V_{orphan}(a',e',j+1) \right] \right\} \tag{4.25}$$

s.t.

$$l = 1 - h - \kappa \cdot I_{h>0} \tag{4.26}$$

$$c + a' \leq a(1 + r(1 - \tau_a)) + \tau_l(e_j h_j)) \tag{4.27}$$

$$a' \geq 0, \quad c \geq 0, \quad l \geq 0, \quad h \geq 0 \tag{4.28}$$

$$\tag{4.29}$$

(iii) At age $j = j^{CH} - 1$ (after the bargaining game) the child enters the second subperiod with after transfer assets a, care obligations s, idiosyncratic productivity e and health status m (which is always equal to good health). Their parent dies with certainty at the end of the period and at the start of the next period they become a parent. The child's problem is given by:

$$V(a,s,e,m,j) = \max_{c,h,a'} \Big\{ u(c,l) + \beta E_j \Big[V(a' + b' - T', s', e', m', M', j+1) \Big]$$

s.t.

$$c + a' \leq a(1 + r(1-\tau_a)) + \tau_l(e_j h_j) \quad (4.30)$$

$$l = 1 - h - \kappa \cdot I_{h>0} - s \cdot I_{j^P > j^{RT}} \quad (4.31)$$

$$a' \geq 0, \quad c \geq 0, \quad l \geq 0, \quad h \geq 0 \quad (4.32)$$

$$T' = \Upsilon_1(X') \quad (4.33)$$

$$s' = \Upsilon_2(X') \quad (4.34)$$

The expectation is taken over the next period's realisation of idiosyncratic productivity, the bequest they inherit and the outcome of the initial bargaining game which they enter as the parent.[10] Assets for their child in the next period are deterministic and equal to 0 (before any transfers made between households). The (new) child may inherit their parent's productivity as described in Section 4.3.3 above.

(iv) At age $j = j^{CH} - 1$ the orphaned agent enters the period with asset level a and income y. Their parent is dead so they pay no services and expect no bequests. Their child is born in the next period. Their problem is given by:

[10]From the ex-ante position one period ahead this is uncertain due to the realization of the parent's idiosyncratic uncertainty and the child's initial income draw. Ex-post, after uncertainty is realized, the bargaining game is deterministic.

$$V_{orphan}(a,e,j) = \max_{c,h,a'} \Big\{ u(c,l) + \beta E_j \Big[V(a'-T',s',e',m,M,j+1) \Big]$$

s.t.

$$c + a' \leq a(1 + r(1-\tau_a)) + \tau_l(e_j h_j) \tag{4.35}$$

$$l = 1 - h - \kappa \cdot I_{h>0} - s \cdot I_{j^P > j^{RT}} \tag{4.36}$$

$$a' \geq 0, \quad c \geq 0, \quad l \geq 0, \quad h \geq 0 \tag{4.37}$$

$$T' = \Upsilon_1(X') \tag{4.38}$$

$$s' = \Upsilon_2(X') \tag{4.39}$$

The expectation is taken over the next period's realisation of idiosyncratic productivity and the outcome of the initial bargaining game which they enter as the parent. As their parent is dead, they expect no bequest.

(v) At age $j^{CH} \leq j < j^{RT}$ (after the bargaining game) the parent household enters the period with after transfer assets a, care obligations (to be received) s, idiosyncratic productivity e and health status m (which is always equal to good health before retirement) The parent's problem is given by:

$$V(a,s,e,m,M,j) = \max_{c,h,a'} \Big\{ u(c,l) + \beta E_j \Big[V(a',s,e',m,M,j+1) \Big] \Big\} \tag{4.40}$$

s.t.

$$c + a' \leq a(1 + r(1-\tau_a)) + \tau_l(e_j h_j) \tag{4.41}$$

$$l = 1 - h - \kappa \cdot I_{h>0} \tag{4.42}$$

$$a' \geq 0, \quad c \geq 0, \quad l \geq 0, \quad h \geq 0 \tag{4.43}$$

The interest rate on assets is denoted by r, and the evolution of e is described above. m' and M' are the health status of the parent and medical expenses which follow degenerate distributions before retirement. Expectations are taken over next period idiosyncratic productivity.

(vi) At age $j \geq j^{RT}$ the parent household enters the period with after transfer assets a, care obligations s and health status m (which is always equal to good health before retirement) The agent is retired and their child has an obligation to provide them with s services now and in the future. Retired agents do not participate in the labour force and have a pension income p.

$$V(a,s,0,m,j) = \max_{c,a'} \Big\{ u(c,l) + \beta \eta_j E_j \Big[\Big[V(a',s,0,m,M,j+1) \Big] \Big\} \quad (4.44)$$

$$+ \beta (1-\eta_j) \phi(b) \Big\} \quad (4.45)$$

s.t.

$$l = 1 + \delta(s-m) \quad (4.46)$$

$$c + a' \leq a(1 + r(1-\tau_a)) + p - M + b_M \quad (4.47)$$

$$a' \geq 0, \quad c \geq 0 \quad (4.48)$$

$$b = \tau_b(a') \quad (4.49)$$

The evolution of m' and M' are described in Section 4.3.3 above. Conditional on survival, retired agents form expectations over health status and medical expenses. Conditional on death, households gain utility from leaving bequests (net of taxes).

4.3.6 The Bargaining Game

In the first period in which their child is in the economy ($j = j^{CH}$) parents and children can bargain over a transfer of assets in exchange for a flow of (delayed) informal care services. This period consists of two subperiods. In the recursive formulation of the previous section agents take the outcome of the bargaining game and transition over stochastic states as an exogenous mapping from the end of period states onto states for the next periods. This subsection describes the bargaining game in further detail.

At the beginning of the bargaining subperiod it is assumed that exogenous states have resolved. Parents and children then play one of a set of games indexed by the states of the dynasty. Each game in this set is a bargaining game where the payoffs of each player is determined by their value function evaluated at the post bargaining states of the dynasty. Following the Markovian nature of the household problem I adopt Markov Perfect Equilibrium (MPE) as the solution concept of this bargaining game.

The bargaining rule

$$\begin{pmatrix} T' \\ s' \end{pmatrix} = \Upsilon(X)$$

is then the solution to the following weighted surplus maximisation for a dynasty with a child of age j:

$$max_{\{T',s'\}} \Big(V(a^k+T',s',e^k,m^k,M^k,j) - V(a^k,s,e^k,m^k,M^k,j)\Big)^{\lambda} \\ \times\Big(V(a^p-T',s',e^p,m^p,M^p,j^P) - V(a^p,s,e^p,m^p,M^p,j^P)\Big)^{1-\lambda} \quad (4.50)$$

$$\text{s.t.} \quad (4.51)$$

$$a^p - T' \geq 0 \quad (4.52)$$

$$a^k + T' \geq 0 \quad (4.53)$$

$$0 \leq s' < 1 \quad (4.54)$$

Where the parameter λ is the bargaining weight of the child in the process. Both the static model of the exchange motive for inter-vivos transfers and models of informal Long Term Care frequently use a Nash bargaining solution to specify the transfer decisions. For simplicity of interpretation, I maintain this assumption in the dynamic model considered here. Furthermore, for given value functions, the Nash bargaining game has a unique equilibrium and, consequently, deterministic bargaining rule under restrictions discussed in Appendix C.2. Unlike many previous papers within this literature I choose to directly model the opportunity cost of transfers of time and the intensive margin of time gifts including it as an outcome that is directly bargained over.

It is assumed that this game takes place under complete information after income shocks have been realised, and households can perfectly commit to the outcome. In part the assumption of complete information is an informational assumption and in part a timing assumption. The timing assumption assumes that the bargaining takes place after stochastic states have resolved, e.g. parent and child productivity, and known to their respective households. The informational assumption is that the pair of productivities, assets and health state is revealed to both households in the bargaining game, before either player chooses their actions. Under complete information the outside option and payoffs can be computed for and by both parties.

The effect of the game for parents is to increase the motive for asset accumulation through two mechanisms. First, value functions are increasing and concave in wealth. This implies parents accumulate assets to increase their outside option in the bargaining game and by increasing their threat point, holding other variables constant, the outcome of the game is better for the parent. Simultaneously, the costs associated with a transfer in foregone future consumption decrease due to the concavity of the value functions. More assets are accumulated to improve their threat point and 'purchase' services from their child. Second, it magnifies the uncertainty the parent faces between periods as the outcomes of bargaining depends on the realisation of both the parent's income and the child's income. No household can be made worse off through bargaining as they split surplus only if it exists, but parents can use their wealth in order to buy leisure services from their child and smooth the marginal utility of consumption between states of the world. This interacts with the precautionary saving motive that exists due to health status, medical expense, income, and mortality uncertainty. In addition to using wealth to smooth their maginal utility of consumption, households can also use their early life cycle wealth to smooth their leisure and preference shocks.

For the newly born child, the start up gift of assets undermines the desire to save as it relaxes their borrowing constraint. In the incomplete market environment of this model an injection of start up assets places them closer to the optimal buffer stock and the possible receipt of bequests in the future diminishes their savings motive. However, for children in future periods the bargaining interaction also increases the saving motive (all other things being equal) as they will become parents and the desire to increase one's own threat point increases the desire to save. Furthermore, bequests are a luxury good and monetary transfers from their parent may increase their lifetime resources sufficiently. During periods of repayment the marginal utility of consumption increases as their total leisure falls, again providing further motivation for asset accumulation.

4.3.7 The Equilibrium of the Intergenerational Game

The equilibrium of the intergenerational game constitutes a recursive stationary equilibrium of the overlapping generations economy (a formal definition is provided in Appendix C.3). This requires that the expectations of bequests (conditional on state s, parent age and mortality) are consistent with the endogenous distribution of parent households; the distribution of expected counterpart households when children are born are consistent with the endogenous distribution of households; and are mutually consistent with the optimal choices made by households (including the markov perfect endogenous bargaining rule).

4.4 Discussion of the Intergenerational Links

To compute the equilibrium, I solve the model using backwards induction and iterate on the distribution of households in the economy. However, a number of the assumptions of the model and their impact on household decisions can be discussed informally here.

4.4.1 The Assumptions of the Bargaining Game

The first key assumption of the bargaining game is that it is played under complete information. Under asymmetric information the participants would be unable to identify which game they are playing beyond a subset induced by their information partition. In a strategic interaction with asymmetric information the Nash bargaining representation of the solution employed above would no longer hold and alternative solution concepts may generate multiplicity.

In each bargaining subperiod, assuming a markov perfect strategy limits the complexity of the strategy available to each player. However, when the game is played once as parent and once as child this assumption because there is no repeated element to the game.[11] Second, the model with a single bargaining interaction in each generation cannot capture the full dynamic elements of family transfers (these are well documented in McGarry, 2016) and thus limits the insurance potential of family transfers. This exposes households to a greater degree of idiosyncratic risk in the bargaining period because of the additional uncertainty from their counterpart's income and the outcome of the bargaining game. Finally, the repeated interaction entails that the optimal policies of the household are necessarily forward looking. They take the future bargaining outcomes as given, but optimally choose their savings accounting for the effect it has on future intergenerational interactions.

The assumption of perfect commitment is key to the quantitative results presented below. Although the bargaining outcome is constrained by the participation constraint, I eliminate reneging as an option for the household and can thus sustain a larger set of outcomes. These outcomes would not necessarily be deviation proof (when compared to the repeated game with limited or no commitment) in future periods once future states have realised. Exploring the role of commitment and repeated bargaining is left as an exercise for the future.

4.4.2 The Exchange Motive

The central assumption of this paper is to limit the role of altruism in determining intergenerational transfers. The evidence for various motives for intergenerational transfers is discussed in the introduction. In practice, with the limitations of existing panel data sets that incorporate multiple generations, it is not possible to formally identify the different motives from each other without restrictive assumptions (for example Groneck, 2016, finds evidence of implicit rewarding through bequests that

[11] Formally, in this the MPE set is equal to the set of admissible strategies that are also subgame perfect equilibria.

could be interpreted in support of either altruism or exchange). The approach taken in this paper is to assume all transfers can be represented in an exchange framework.[12]

Households may still appear to behave as if they are altruistic towards one another, smoothing lifetime marginal utilities between generations rather than seemingly self interested exchange, and to the extent that households are altruistic it is captured in the bargaining weight. Increasing the bargaining weight on the child implies the child receives more of the total surplus from insurance between generations - in a pareto weight model with partial altruism it is as if the pareto weight on the child has increased. For simplicity of exposition, consider a pair of households (a parent and their child) who exist in a static setting. These two households have a large surplus to divide between them. In the standard partial altruistic model the weights on parent and children determine which of the many possible allocations is implemented (as in the marriage and within household bargaining literature). In contrast, in the exchange framework this is determined by the bargaining weight.

Note, however, that the implications of a dynamic model of the exchange motive imply different incentives and feasible allocations. There are two ways in which this framework can provide an important lower bound. First, the value households derive from transfers in this framework is only the insurance component of the transfers. Assuming that there is some amount of, possibly heterogeneous, altruism across families that is not captured in this model would increase the value from these interactions. The value an altruistic family place on transfers must exceed the welfare calculations found in this paper as the total value of transfers is the sum of insurance value and the altruistic glow. Second, for the same reasons, exchange motives provide a lower bound on the frequency of transfers. Returning to the example above, assume there is no insurance surplus to divide between the households. It may be that for sufficiently altruistic families the utility they receive from

[12] This stance is extreme. Given that rejections of risk sharing abound in the empirical literature, it is no less extreme than the alternative assumption of one-sided (or two-sided) altruism. Furthermore, a number of studies suggest a need for understanding multiple transfer motives simultaneously (see McGarry, 2016, for example) and by building on the exchange motive in a dynamic context, this takes a necessary step towards that ultimate aim.

seeing their counterpart made better off outweighs the cost to their own utility (with concave utility functions altruistic parents and children smooth the utility between generations). Despite no surplus, an altruistic motive generates transfers which are necessarily welfare improving. These transfers are ruled out by the pure exchange framework. In this sense, the results in this paper are conservative estimates for the role of inter vivos transfers by providing a lower bound on the value and frequency of transfers.

In the spirit of developing a lower bound I make an additional restriction on the utility function of households relative to the widely implemented static model. The utility of parents is increasing in attention (or services rendered from child to parent) which may be costly for the child to provide. Conflicting interests between generations is by a restriction on the utility function. In the static model it is sufficient for there to exist a region in which $\frac{\partial U_p}{\partial s}$ is positive while $\frac{\partial U_k}{\partial s}$ is negative. As this restriction on the derivatives is not global the static model of the exchange motive allows for altruistic or egoistic behaviour, but posits that this exhibits a bliss point. In keeping with developing a lower bound, I restrict preferences so that services are always costly to the child, i.e. the period utility function satisfies $\frac{\partial U_k}{\partial s} \leq 0$.

4.4.3 The Role of Bequests

This paper assumes that bequests are not directly part of the intergenerational bargaining engaged in between households. Parents care directly about after tax bequest sizes which is consistent with evidence of changes in estate size in response to changes in estate taxes in Kopczuk (2013) and the evidence of equal division amongst children (see Wilhelm, 1996; Tomes, 1988). The warm glow bequest utility has two interpretations. Firstly, it is compatible with an egoistic motivation, where households care directly about the size of the transfer they leave or gift. Alternatively, Abel and Warshawsky (1988) provide microfoundations for the egoist model of bequests. Under a number of simplifying assumptions, they show that without inter vivos transfers a 'joy of giving' bequest motive represents a reduced form of the altruistic motive. This reduced form expression provides a relationship between the parametrisation of the 'joy of giving' motive and altruism where small measures of altruism imply large estimates for the 'joy of giving' motive. The specification in this paper is compatible with the empirical evidence of equal division and compatible with the partial level of altruism absorbed in the bargaining weight (see the discussion of the exchange motive above).[13]

These roles for the bequest motive are standard within models of life cycle savings, however, there are a number of ways in which it interacts with the exchange motive. To the extent that time services are substitutes for consumption there is a mechanical effect on inheritances left- for the same level of wealth (at which bequests are positive) those parents in receipt of more time leave larger bequests. This captures that bequests may provide a delayed reward for transfers of time. Allowing the strategic use of bequests under credible disinheritance would provide parents with the last mover advantage as in the discussions of Becker (1976) and Hirshleifer (1977). A last mover advantage following from the strategic use of bequests is also assumed in a number of static expositions of the exchange motive, such as Bernheim et al. (1985) where it justifies perfect commitment and the lack of concern over enforcement. This would further expand the set of deviation proof equilibria outside the set of outcomes considered here, but at the cost of considerably increasing the complexity and computational difficulty.

4.4.4 Parent Health is Independent of Child Time

In the model parent health status is independent of child time. There are two ways within the overall framework of the model that this could be relaxed. Parent health transitions are assumed to be entirely exogenous. An alternative would allow for the amount of time the child gives to the parent to affect the transitions between health status. This would increase the possible surplus to split in the dynasty by increasing the value to the parent from receiving time (including the change to the transition probabilities as well as the direct utility benefit). Excluding this possibility is consistent with the lower bound approach used elsewhere in this paper and

[13] Bernheim and Severinov (2003) suggest an alternative explanation for equal bequests. They consider a model in which the division of bequests provide a signal about a parent's altruistic preferences and in which children directly care about their perceptions of parental affection. Multiple equilibria, including the primogeniture customary in Asia, are sustainable with social norms providing equilibrium selection.

does not require taking a stance on the efficacy of time investments in old age health production.

The second way in which health status is independent of child time is in the period utility of the parent and child. It is assumed that the time received from children is still valuable to the parent (and costly to the child) when the parent is in good health. As discussed above, the distribution of time gifts is highly skewed and time gifts seem to fall into two broad categories. There are small gifts of time associated with menial, household and administrative tasks and larger gifts of informal LTC. In order to generate these smaller (and more frequent) time gifts it is necessary that time gifts are valuable in both health states (the actual utility for a given level s and consumption c depends on m).

4.5 Calibration

Model parameters are calibrated in two ways. The first group of parameters are estimated outside the model, directly from the data, and the second group are set taking parameters from the literature. Sections 4.5.2 and 4.5.3 describe how these two sets of parameters are set in greater detail. I discuss the data source and sample selection criteria in Section 4.5.1.

4.5.1 Data and Sample

The primary data source used in this paper is the Panel Study of Income Dynamics (PSID). This paper uses data taken from the main PSID survey (waves 1968-2013) as well as the 2013 Rosters and Transfer Data (henceforth, RTS) supplement to the 2013 PSID. The PSID is a longitudinal study collecting biennial (annual before 1997) data on households in the US and collects information on their descendants in separate households. Appendix C.4.1 describes the rules used by the PSID to follow the descendants of original 1968 families.

Primarily, I use data from the PSID waves 1999 through to 2013. The PSID provides detailed information on labour market outcomes and family composition since 1968, however, in 1999 the PSID introduced a new set of questions to the standard PSID questionaire capturing richer expenditure data (and again in 2005) on both household consumption (See Andreski et al., 2014, for a systematic comparison with the Consumer Expenditure Survey) and medical spending. Furthermore, for this period of interest the PSID collects information on the wealth and asset holdings of households in every wave. I check the robustness of my estimates (where the data exists in earlier waves) to the choice of waves.

The model treats the marriage market and household formation as exogenous, the empirical analysis only considers households who are married or permanently cohabiting. Single individuals are not included in this sample irrespective of whether they were never married, widowed or divorced. However, households who marry for the first time (or divorce and remarry alternative partners) are included in the sample. If interactions with earlier and later generations systematically affect the type of cohabitations that are formed (or persist) this is not captured in the model.

I exclude households for a number of additional reasons. I require households to respond to the survey at least 3 times-this implies they are observed over at least a 4 year period (the length of a time period in the model). I exclude households younger than 24 and older than 92 in the main analysis.[14]. This leaves a sample of 6,361 separate households observed. Where further restrictions are imposed in the estimation of specific parameters they are described in detail below.

Table 4.3 summarizes the characteristics of the sample of households I analyse. Table 4.4 shows the proportion of married households who satisfy each of the conditions for exclusion (note that some households may satisfy multiple conditions for exclusion).

4.5.2 Estimated Parameters

This subsection details the estimation procedures for parameters outside the model

Earnings process

Deterministic Component In order to estimate the deterministic component of full time earnings (in equation 4.6), I first estimate the parameters of a fixed effect re-

[14]I include households between 20 and 100 when estimating polynomials in ages in order to avoid issues associated with endpoints

Characteristic	Mean	Median
Average Income[a]	89,873	68,000
Average Net Worth[a,b]	360,726	51,737
Working Head	84.9%	1
Working Spouse	68.4%	1
Hours Head	2010	2080
Hours Spouse	1300	1560
Age of Head	49.9	49
Age of Spouse	48.4	47
Poor Health Head	4.0%	0
Poor Health Spouse	3.68%	0
Family Size	3.02	3
Sample Size	6,361	

Notes: All values are the author's own calculations. All variables are calculated from the PSID & RTS using the PSID sample weights and reported to 3 significant figures unless otherwise noted. All variables are the mean values conditional on being selected into the sample; [a]Values are in 2015 US Dollars and rounded to the nearest dollar value; [b] Total Family Wealth excluding Home Equity and Pension Wealth .

Table 4.3: Mean Characteristics of the Sample

	Number	Proportion
Age restriction	55 (459)	0.37%
Married households	3,501 (25,709)	23.6%
Spouse observations	3,430 (29,801)	23.1%

Notes: Column Number reports the number of households excluded by applying each restriction. The number of individual-year observations are given in parenthesis below. Some households may satisfy multiple conditions for exclusion. Some households satisfy multiple conditions in such a manner that they remain in the sample after imposing one condition, but with insufficient year observations and are deleted in the combined application of the sampling rules.

Table 4.4: Reasons for exclusion from the sample

gression on a second order polynomial in age, self reported health status, the wave specific unemployment rate (to account for aggregate fluctuations) an indicator for part time work (less than 500 hours annually) and a set of indicators for family size.

$$ln\ e_{i,j} = \hat{f}_0 + \hat{f}_1 j + \hat{f}_2 j^2 + \hat{f}_3 Urate_{i,j} + \hat{f}_4 m_{i,j} + \hat{f}_5 \mathbf{1}_{pt_{i,j}=1} + \sum_z \hat{f}_6^z \mathbf{1}_{famsize_{i,j}=z} + u_{i,j} \tag{4.55}$$

I estimate this equation only on a subsample of households who work over 300 hours in a year and discard outlying high earners (an hourly rate above $500 in 2015 dollars). I calculate the wage rate for a given couple as the total combined wages of the couple and exclude self employed households.[15]

Stochastic Component Having removed the deterministic age trend, I estimate the parameters of the stochastic process detailed in equation 4.7 using the variance co-variance matrix of the estimated residuals $\hat{u}_{i,j}$. I assume that all transitory components are measurement error and jointly estimate the persistence and variances using a minimum distance method.

Although I estimate the intergenerational transmission of productivity separately, in estimating this process I allow for the variance of the initial innovation to differ from later innovations. This (highly persistent) initial draw captures some of the variation in fixed effect which is omitted from the model. The variance for the innovation to the initial draw is lower than the 4-year variance of the stochastic process, however as detailed below I explicitly allow for inherited productivity and thus the total variance of any initial draw is considerably larger.

Intergenerational Persistence As discussed in Section 4.3.3 I allow for the intergenerational transmission of idiosyncratic productivity. For the persistence of this process I impose the estimates from Chetty et al. (2014) who estimate a linear regression slope between 0.25 and 0.35 and use the variance of the initial draw calculated above.

Estimates for the stochastic component of idiosyncratic productivity and the intergenerational persistence are presented in Table 4.5. Parameter values are re-

[15] As discussed in French (2005), these estimates are not estimates of the true process for offered wages because they fail to account for selection and may be biased upwards.

Parameter	Variable	Estimate
ρ	Autoregressive coefficient of wages	0.884
$\rho_{inherit}$	Integenerational persistence in wages	0.344[a]
σ^2_{k1}	Variance of initial innovation	0.429
σ^2_{ε}	Variance of subsequent innovations	0.480

Notes: All values are the author's own calculations and calculated from the PSID & RTS using the PSID sample weights and reported to 3 significant figures unless otherwise noted. Values of the process are reported for the four year intervals used in the model; [a] Taken from the baseline log-log regression reported in Chetty et al. (2014)

Table 4.5: Estimates of the Earnings Process

ported for the 4 yearly processes used in the model.

Health Status

The PSID asks for the self reported health status of both the head and the spouse. Respondents are asked to report whether they are in one of "Excellent, Very good, Good, Fair, Poor, Don't Know or N/A" health.[16] In the model there is a single health status for the couple and to construct an equivalent variable from the PSID data I define couples to be in Bad Health if either the head or the spouse reports "Poor" health in response to the question.

In order to estimate the transitions, I assume that health status today can be expressed as a linear probability model and is determined by an individual fixed effect and a quadratic in age (interacted with lagged health status). I estimate this equation in first differences and correct for measurement error by using lagged health statuses as an instrument for lagged health status change.[17]

Medical Expenses

The procedure for calculating out of pocket medical expenses is similar to the approach used to estimate the earnings processes above. I calculate the total out of pocket medical expenditure by aggregating questions which ask households about

[16] Very few households report either of the last two classifications

[17] Lagged health status interacted with age are used as an instrument for lagged health status change interacted with health.

medical care expenditures at the household level. Details of this aggregation can be found in Appendix C.4.2.

When estimating the processes for medical expenses Lockwood (2018) restricts his sample to households with over \$100,000 in assets. The requirement that households have significant holdings of wealth (excluding their home) attempts to minimize the bias introduced by the fact that Medicaid pays for much of the care received by people with little wealth. Unlike the HRS, the PSID typically has a much larger sample of (relatively) poorer households and imposing this liquid asset threshold eliminates approximately half the PSID sample. Furthermore, as the PSID waves are biennial eliminating on current wealth ignores the effect that significant expenditures in the previous two years may have on present wealth. Additionally, the PSID questions specifically prime households, as discussed in Appendix C.4.2 which discusses the construction of medical expenses, and break down totals including copayments and insurance covered payments. Therefore, I do not drop individuals from my estimation on the basis of liquid wealth.

Deterministic Component In order to estimate the deterministic component of log out of pocket medical expenses, I first estimate the parameters of a fixed effect regression on a second fourth polynomial in age, self reported health status and the interaction of the polynomial with health status.

$$ln\ M_{i,j} = \sum_{z=0}^{4} \{H_z j^z + H_{z+5} j^z \cdot \mathbf{1}_{m_i,j=1}\} + \sigma_M(m_{i,j}, j) \cdot \zeta_{i,j} \tag{4.56}$$

In order to isolate the medical expenditure of retired household this is estimated only on the over 65 subsample.

Stochastic Component Having removed the deterministic age trend, I first estimate the parameters of the age and health status conditional variance $\sigma_M(\cdot,\cdot)$. From these estimates I recover the remaining residual ζ.

To estimate the parameters of the stochastic process I calculate the variance covariance matrix of the estimated residuals and jointly estimate the persistence and variances using a minimum distance method. It is not possible to separately identify measurement error from the transitory component of expenses. A number

Parameter	Variable	Estimate
ρ_M	Autoregressive coefficient of medical expenses	0.798

Notes: All values are the author's own calculations and calculated from the PSID & RTS using the PSID sample weights and reported to 3 significant figures unless otherwise noted. Values of the process are reported for the four year intervals used in the model.

Table 4.6: Estimates of the Out of Pocket Medical Expenses Process

of papers[18] emphasise the importance of a transitory component of out of pocket medical expenses and therefore I assume that there is no measurement error in this constructed variable.

Estimates for the stochastic components of log medical expenditure are reported in in Table 4.6. Parameter values are reported for the 4 yearly processes used in the model.

4.5.3 Imposed Parameters

The remaining parameters are set with reference to the literature. They are detailed in this section. Table 4.7 summarises these choices.

Preferences

β **Discount Rate** I set the annual value of β to 0.98 following Attanasio et al. (2008). Although this at the higher end of the estimates found in the literature, it is frequently used in models analysing the role of women in the labour market. Since this paper considers the labour supply of both partners I use a value consistent with estimates of female labour supply. This may lead to an overestimate of the value of familial insurance (as households care more about the future), however, this is offset by the conservative value of γ discussed below.

σ **Consumption Weight** σ is set at 0.68.

[18] See, for example French and Jones (2004) who find the process for medical expenses is well modelled by the sum of a persistent and transitory component and Lockwood (2018) which assumes acute medical expenses are entirely transitory.

δ **Taste for Ill Health** I set δ as a renormalisation of the taste shifter for ill health estimated in De Nardi et al. (2010). I normalise their value so that for retired households who are in ill health and receive no gifts of time from their children the implied linear change to the value of consumption is equivalent.

As this paper also includes a bequest motive I use the value estimated in their corresponding specification including both bequest motives and health dependent preferences. In this specification they estimate a larger effect of health status on consumption preferences. This implies that time gifts are more valuable to parents, but the competing bequest motive mitigates the desire to perform inter-vivos transfers.

κ **Fixed Cost of Working** κ is set to 0.05 which implies that households lose 5% of their leisure time at the extensive margin. This fixed cost is sizeable, but smaller than the values estimated in French (2005). This paper explicitly accounts time gifts received by parents and incorporates leisure costs to their children. The opportunity cost of time given to one's parent is affected by labour force participation. Those households of working age have their own heterogeneous distribution of promised time and, consequently, a heterogeneous distribution of leisure endowment. This generates substantial heterogeneity in reservation wages for each household when combined with the fixed leisure cost of participation. However, due to the endogenous determination of time transfers, the results in this paper are sensitive to the choice of κ.

ϕ **Bequest Utility Parameters** I take the estimates of bequest utility parameters from Lockwood (2018). These estimates are similar to results found in Ameriks et al. (2011) and De Nardi et al. (2016a) implying both a quantitatively important bequest motive and that bequests are luxury goods. Those rich enough have a large marginal propensity to bequeath out of wealth[19].

γ **Coefficient of Relative Risk Aversion** The modal choices of γ in related papers which study intergenerational transfers follow the bounds adopted in the wider structural literature. When utility in consumption and leisure is separable (see Ka-

[19] Lockwood (2018) also estimates the risk aversion of retirees over bequests.

plan, 2012; Fahle, 2015; Abbott et al., 2018) values used are between 1 and 2 as which are consistent with values uses in life cycle estimation and studies of risk sharing through marriage (e.g. Voena, 2015) and is between 3 and 4 when the utility function is non-separable in consumption and leisure, as it is in this paper (See, for example, French, 2005, from wider life cycle literature). A notable exception is Barczyk and Kredler (2018) who use a value of 3.8 when utility is separable.

A key contribution in this paper is to provide a lower bound on the insurance value of inter-vivos transfers. This approach is conditional on a given set of parameter choices. I use a value of 3 for γ as this is the lower bound where utility is non-separable. This choice implies household's are less risk averse and in turn implies the ability to share risks between generations is less valuable- a conservative choice when quantifying the pure exchange value of transfers.

The Bargaining Weight

This paper explicitly excludes altruistically motivated transfers between households. To the extent that households are altruistic and this affects the allocation between generations it is captured in the bargaining weight. Increasing the bargaining weight on the child household implies that (when there is a positive surplus) parents receive a smaller share of the surplus, this is as if they have larger altruistic preferences over their child. The role of the bargaining weight is key in determining the size and frequency of transfers.

I set λ, the bargaining weight on the child, to 0.5. This implies that parents and children divide the bargaining surplus equally. Both Fahle (2015), who uses the same value, and Barczyk and Kredler (2018) use Nash bargaining weights to determine part of the transfers in equilibrium. While Barczyk and Kredler do not use equal weights, for practical purposes their paremeterization is similar.[20]

[20] They estimate substantial altruism from parent to child and little child altruism towards their parent. In addition they estimate a very large bargaining weight on the parent. However, the large weight on the parent in bargaining mitigates the small weight placed by the child directly on the parents utility. Direct comparison between the bargaining weights is not possible, but the offsetting effects of the bargaining weight and two-sided altruism brings the two parametrisations closer.

Taxes and Government Transfers

Income Taxes I use the estimates of the federal income tax and social security programs estimated and reported in Guner et al. (2014). I use the estimated tax and transfer schedule for married households with two children.[21] This is consistent with the demography of the households considered within the model. In practical terms this implies lower average tax rates for households.

The average level of taxation, $\tau_{l,1}$, is set to 0.925 and the progressivity, $\tau_{l,2}$, is set to 0.070

Capital Gains and Bequests The capital gains tax, τ_a, is set at 20% (See Kotlikoff et al., 1999). There are two approaches possible when calibrating the parameters of the estate tax. The first is to use the exemption level and tax rates specified in the tax code. However, this ignores the considerable legal and illegal avenues for reducing a household's estate tax burden. The second approach is to estimate estate tax functions that account for the degree of tax reduction. I use estimates from De Nardi and Yang (2016) which match the fraction of estate tax revenue to output and the fraction of estates that pay estate taxes. The estimated exemption level is $756,000 (in 2000 dollars) and the marginal tax rate is 21.43%.

This exemption level is close to the exemption level of the time, but the the marginal tax rate is considerably lower than the 56% marginal tax rate in place which suggest considerable scope for avoidance or evasion.

Pensions and the Consumption Floor The average monthly benefit for social security beneficiaries in June 2015 was $1,135 per person. To parametrize the annual benefit for a couple I use this value to calculate annual benefits at the household level. The consumption floor is set at $2,665. This is taken from De Nardi et al. (2010) and their specification including an operative bequest motive.

Remaining parameters

I set the annual market rate of return on assets r to 3 percent. Consumption is adjusted by an economy of scale assuming 2 children in each household after the

[21] Guner et al. (2014) also report tax rates for all married households and include social security. These results are provided in an online appendix and are almost identical to the estimates used for married households with two children

Parameter	Value	Reference
Discount factor (β)	0.98	Attanasio et al. (2008)
Consumption Weight (σ)	0.68	
Taste for ill health (δ)		De Nardi et al. (2010) & Text
Fixed cost of working (κ)	0.05	See Text
Weight on Bequest (ϕ_1)		Lockwood (2018)
Curvature of Bequest (ϕ_2)		Lockwood (2018)
CRRA (γ)	3	See Text
Child's Bargaining Weight(λ)	0.5	See Text
Household income tax (τ_l)		Guner et al. (2014)
Bequest tax (τ_b)	21.43%	De Nardi and Yang (2016)
Exemption level on bequests (ex_b)	$756,00	De Nardi and Yang (2016)
Capital gains tax (τ_a)	20%	Kotlikoff et al. (1999)
Universal pension (p)	$27,240	SSA
Consumption Floor ($\underline{c}$)	$2,665	De Nardi et al. (2010)
Market return on assets (r)	3%	De Nardi et al. (2010)
Consumption Scale (c)		McClements Scale
Survial Probabilities(η)		Actuarial Tables

Table 4.7: Pre-set parameters of the model

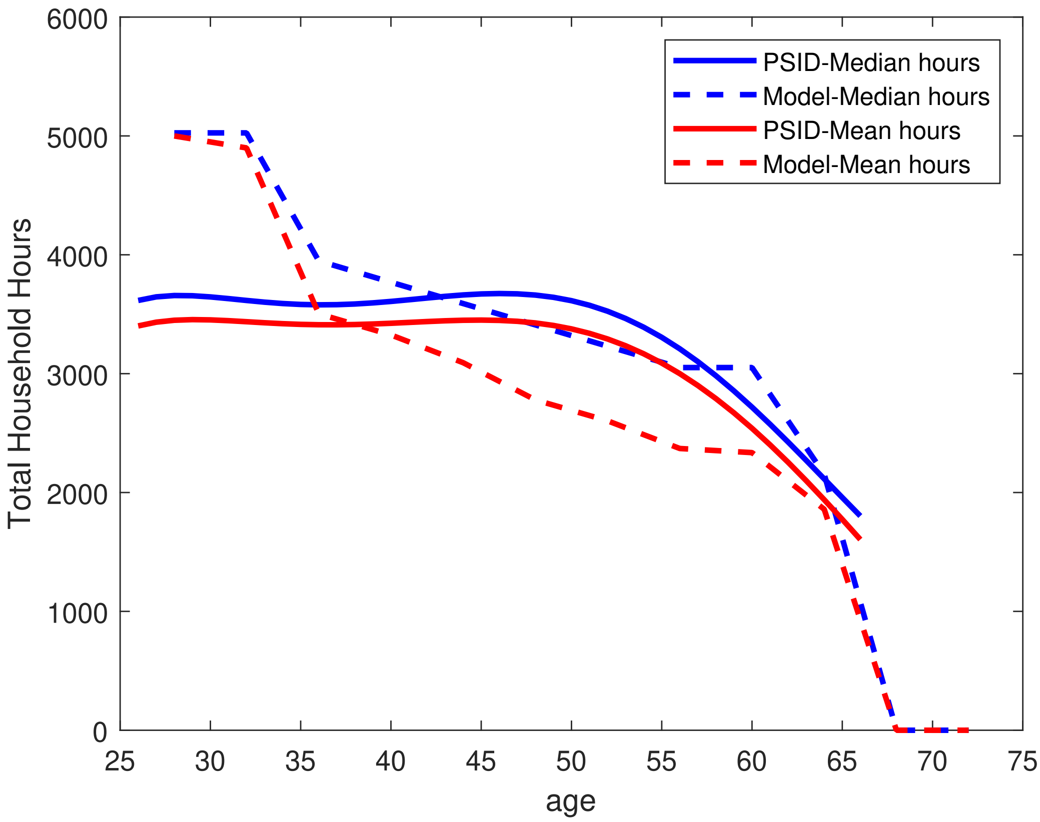
exogenous fertility date and before their children enter the model as adults.[22] Survival probabilities (independent of health status) are calculated at each age from life tables. Table 4.7 summarises the sources of parameters used in the model.

4.6 Results

In this section, I show how the calibrated model performs relative to moments from the data. I begin with life cycle profiles for key choices in the model, before comparing the model generated distribution of inter vivos transgers

Figures 4.3, and 4.4 show, respectively, a comparison of the profiles of labour supply and household wealth in both the data and the simulations of the calibrated model. I exclude home equity from the PSID asset profiles to accurately capture the self insurance opportunities available to households and follow the procedure in

[22]Each parent is assumed to have one male child and one female child who effectively marry identical partners on entering the model. This leaves one descendent household per parent household

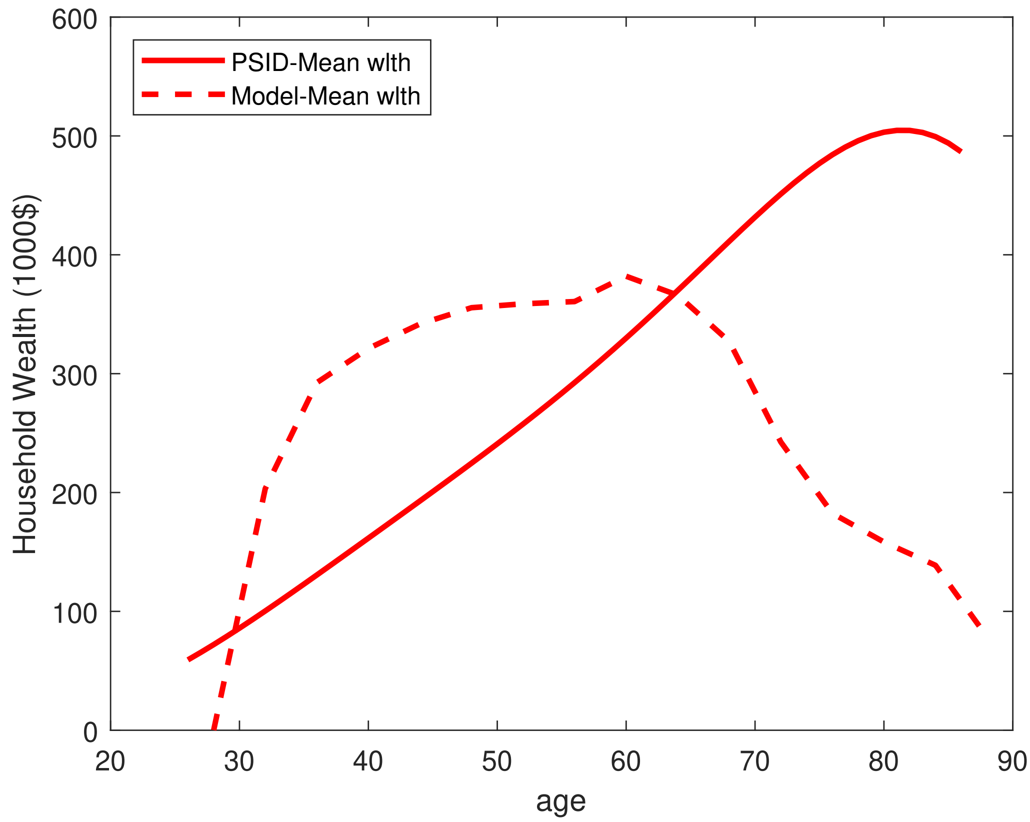


Notes: Mean and Median household labour supply from the model (dashed) and PSID (solid)

Figure 4.3: Life Cycle Profiles of Household Labour Supply

Fernández-Villaverde and Krueger (2007) to purge time and cohort effects from the data. The model has an exogenous retirement age and I drop data for hours worked above this age.

Figure 4.3 presents the mean and median age profile for total household hours worked for both the model (dashed) and the data (solid). There are two important features of the PSID data: first, the data is negatively skewed at all ages with median hours worked exceeding mean hours worked and, second, while the profile is close to flat betweem 25 and 50, hours of work begin to decline in the early 50s. This accelerates with age. The model profiles broadly replicate the level of labour supply over the the life cycle at both and produce a similarly skewed distribution. However, the model generates a larger gap between (recall that no moments are explicitly



Notes: Mean household wealth from the model (dashed) and PSID (solid)

Figure 4.4: Life Cycle Profiles of Household Wealth

targetted in estimation or calibration) the mean and median hours worked because it produces too little participation on the extensive margin. This also leads it to understate the median hours worked. Finally, the model substantially overpredicts hours worked in a household's late 20s. This is because of the upward bias in the wage offer distribution (see above) and, importantly, because households anticipate the future promised care services to their parents and have a precautionary labour supply motive. Once parents age into retirement there is a large reduction in labour supply.

At first glance, the model generated asset profile seems much further from it's data counterpart. The model generated mean asset profile exhibits a familiar hump shaped life cycle pattern which peaks around age 60. In contrast, the data profile rises more slowly until age 80, before starting to decline. A key reason for this is that

	PSID		$\lambda = 0.5$	
	Fraction Non-zero	Mean (> 0)	Fraction Non-zero	Mean (> 0)
Time Given (hrs/wk)	33.9%	9.4	28.5%	12
Money Given	28.9%	$5,236	9%	$6,790

Notes: Responses from PSID Heads calculated from the 2013 PSID & RTS. Dollar amounts in 2015 US$. All transfers are reported as occurring for a single generation, the total frequency of time is multiplied by the fraction surviving through retirement.

Table 4.8: Transfers in the Last Calendar Year

simulated households face a maximum life span that is considerably shorter than their data counterparts. In the data approximately a third of the population outlive age 88 although the average age at death is in line with the model. However, despite differences between the life cycle saving paths of the simulated model households and the data households it suggests similar demand for savings over the life cycle. This thesis explicitly addresses the retirement period in other chapters.

Table 4.8 provides evidence of the generated transfer distribution. Although no moments presented are targetted in estimation, the fit of the implied transfer distribution can be seen as an explicit test of the exchange motive's credibility in a dynamic setting under reasonable paremterizations.

The amount of family insurance generated by the model is on a similar order of magnitude to those in the data. Time gifts occur with approximately the correct frequency for any given generation, however, the design of the bargaining game means that cross sectional time transfers exceed gifts of money. Furthermore, the dynamic model of family exchange is is not rich enough to generate the full dynamic life cycle pattern of these transfers. In part this is driven by the inability of households to engage in repeated bargaining. Households commit to small transfers of time for up to the next 30 years of their lifetime, while parents offset this with a monetary transfer that moves children away from the borrowing constraints.

That the model does not fully replicate the frequency and size of transfers is not a rejection of the exchange motive. However, it does suggest that a richer model which includes more flexible dynamic transfers and, possibly, multiple motives for providing transfers may be required to match the rich features of the transfer data.

4.7 Evaluating the Value of Family Transfers

The calibrated model used in this paper provides a lower bound on the frequency and value of transfers to the family. It only generates transfers that are consistent with gains from risk sharing and excludes those transfers where improving the utility of other generations within the family outweighs the cost to own utility.

For this reason it is possible to use the calibrated model to generate a lower bound estimate on the value of the intra-familial insurance provided by inter vivos transfers. The model in this paper generates a stationary distribution of households, their wealth and transfer choices. In order to evaluate the welfare value of transfers I employ an ex ante welfare criterion before initial ability is realised. I follow Conesa et al. (2009) in calculating the expected lifetime utility of a worker (before initial productivity draw) in the stationary equilibrium. In order to compare across the different specifications of the model, I report ex ante utility as a percentage of the baseline ex ante utility. In addition, when examining the effect of eliminating intergenerational transfers I also calculate the compensating variation that would make households indifferent between living in the baseline and the counterfactual experiment.

I calculate these values for five different specifications of the model. The first is the baseline calibration of the model presented above. The second specification repeats the calibration exercise, but without any intergenerational persistence in the earnings process. The third and fourth specifications are calibrations of the model with no bequest motive and no bequests (i.e. the government taxes all estates at 100% and raises a surplus) respectively, both of these specifications include inter vivos transfers. The fifth specification is as the third, but also excludes inter vivos transfers. Table 4.9 reports the results of these experiments.

	Baseline (1)	No Inherited Productivity (2)	Accidental Bequests (3)	No Bequests (4)	No Intergenerational Transfers (5)
Average Lifetime Utility	100%	99.3%	95.7%	95.6%	95.6%

Table 4.9: Welfare Comparisons

Comparing columns 1 and 2, eliminating the intergenerational persistence in income is welfare decreasing. In part this is because of the insurance value of intergenerational transfers. The largest surpluses occur for a combination of wealthy parent with unproductive children. Those parent households who have had a sequence of large productivity shocks (or bequests) and accumulated large amounts of assets, but have a large negative productivity shock in the period directly before bargaining are made worse off. The probability they are matched with low productivity children declines and the bargaining surplus shrinks. In part, this is offset by the opposite effect for parents and children, high productivity households are now more likely to be paired with low productivity children and the surplus increases. When transfers are solely determined by the exchange motive (as presented here), currently wealthy and high productivity households are less likely to insure their children because they have higher outside option. Redistributing children with low productivities towards these productivities lowers aggregate welfare.

Columns 3 and 4 compare the average ex ante utility under a model with no bequest motive, but accidental bequests and no bequests. These results suggest that the average household has large lifetime expected utility gains from receiving bequests as well as the expect utility from leaving bequests in retirement. Without bequests households are on average poorer as there is a larger motive to run down assets in retirement.

The final column eliminates inter vivos transfers from the model. The ex ante utility loss associated with eliminating inter vivos transfers is comparable to the elimination of bequests. Although the financial component of inter vivos transfers are much smaller in magnitude they provide very similar roles. First, for the retired parent household bequests and the receipt of time gifts have similar effects. They both provide insurance against future uncertainty. The large amounts of wealth held by households who do leave bequests can also be used to insure them against medical shocks, similarly the time gifts made by children are substitutes for their parent's consumption. Second, the effect of financial transfers at the start of life are significant. They bring households away from borrowing constraints affording them self insurance and allow them to earn compound interest on this stock.

Finally, comparing the baseline economy and the economy of specification five I calculate a compensating variation and find that households would require a one time transfer of $19,000 at the start of their life time in order to compensate them for their inability to engage in inter vivos transfers. While this can be considered a lower bound, it is still implies households derive considerable benefits from these transfers.[23]

4.8 Conclusion

This paper develops and solves a life cycle model of exchange motivated transfers amongst the extended family. I propose a model that is computationally tractable whilst incorporating both the rich earnings risk faced by working households as well as the various risks to health, mortality and wealth in retirement. This builds on previous work to allow for a quantitative assessment of the exchange motive to generate patterns of inter vivos transfers

I use the implications of household choices under the exchange and altruistic motives to calculate a lower bound on the welfare value of family insurance. A large number of households in the model engaged in positive transfers and the ex-ante insurance value of transfers is sizeable. The model does not feature altruistic motives for inter vivos transfers and this result is driven by those households where there are large gains from trade.

[23] In particular, the annuity value of a small transfer early in the life cycle is large.

The wider inter vivos transfer literature has found it difficult to distinguish between competing motives for transfer behaviour. The central assumption of this paper is to assume that households make transfers and provide insurance when exchange is mutually beneficial. Taking stock, this paper presents new evidence on the empirical and economic significance of transfers of time. The large welfare effect, even excluding an altruistic glow, from eliminating these transfers highlights the importance of these transfers for many households. Combining multiple motives for making transfers and considering the combination of exchange as well as altruism towards other generations may be necessary to replicate the full distribution and life cycle patterns of inter vivos transfers. Similarly, the results in this paper rest on the assumption of full commitment. Other papers in the inter vivos transfer and informal long term care literatures have incorporated no commitment, while studies of the decisions made by married households (including the decision to divorce) favour a limited commitment framework. Analysing the robustness of the results to these modelling choices may shed light on the importance of different motives, while credibly identifying alternative explanations within the same framework may prove a fertile avenue for future research.

Chapter 5

Conclusion and Thoughts on Future Research

Each of the three substantive chapters of this thesis contains its own conclusion. Therefore, rather than re-stating the conclusions in full, I briefly summarize these pieces of research and use this section to discuss possible future research on these topics. The suggestions for future research discussed here build on the findings of this thesis and could be researched using techniques and data similar to those used in the rest of the dissertation.

Chapter 2 highlights the role of housing in the portfolio of retired households. It shows that this has important quantitative implications for the generational transmission of wealth as well as how retired households value the transfers they receive from means-tested government programs that insure them against the risk of catastrophic Long Term Care expenses.

Inheritance (or estate) taxation is an important policy instrument and an existing literature highlights the efficiency and equity trade-offs. Results in Chapter 2 highlight the role of differences in preferences for leaving a bequest as well as differences in how different types of wealth is drawn down in retirement. This means that some households will respond more to changes in estate taxation than others. Incorporating these insights into a model of the full life-cycle is necessary to draw conclusions about optimal inheritance taxation because these policies are important for wealth accumulation in earlier life. It is also crucial to consider the generation who receive inheritances. The wealth effect and welfare impacts among those who receive these inheritances may change substantially in a world where it allows younger households to access leveraged returns on housing or alleviates their liquidity constraints. If this exacerbates inequalities due to other differences across households (like heterogeneous work ability) this may also change the equity argument for an inheritance tax. This work would aim to understand a number of key features of inheritance taxation in a rich quantitative framework. For example, the extent to which targeting these taxes by allowing for exemptions in different asset classes is (or is not) optimal? How important is the role of unobservable preferences for leaving an inheritance? How should we set the overall level of inheritance taxation?

A second key insight in Chapter 2 is that the interaction between household portfolio composition and the design of the tax and transfer system can have a first order effect on retirees. This raises the important question of how should means tested benefits treat different assets over the life-cycle? The tax and transfer system provides insurance against temporary and permanent income shocks over the course of the life-cycle as well as other risks. Many parts of these systems feature means testing assessed against household incomes and assets. In the real world, the implementation of these asset tests is complex with some exempting housing assets (for example, disability insurance and food stamps in the US or pension benefits in Australia) or assuming a return on financial wealth and current account balances that counts towards an income test (universal credit in the UK). This can distort the user cost of housing, subsidizing homeownership and have large implicit redistribution. While there is a large body of work focussing on how policies such as home mortgage tax deductions or transaction taxes subsidize the user cost of housing, we still know relatively little about the incentives created by differences in means-testing across multiple asset classes over the life-cycle. Furthermore, this may provide important answers to help explain differences in portfolio composition across households and in future work I would seek to understand what effect it has on the benefits that the tax and transfer system provides.

Chapter 3, which also focusses on retired households, highlights important differences between single retirees and those who are in couples. Empirically, there are important differences in the trajectory of their retirement wealth, the distribution of contemporaneous risks and risks that they may face in the future. We show that accounting for these differences implies single retirees and couples save for very different reasons. For singles the majority of their saving is explained by precautionary savings against medical expense risk, but for couples bequest motives and a desire to insure their surviving spouse are also important.

This chapter brings these important difference between Couples and Singles to the fore. Given the stark differences in their observed behaviour and the key drivers it raises important questions about the design of government policies, of which there are many, that treat couples and singles differently or provide insurance across life transitions. A natural follow up to the work in Chapter 3 is to ask how well designed these policies are? In particular, do they strike the right balance between couples and singles or could we tilt benefits away from one easily identifiable group to another without imposing any additional costs to the government. While we would hope to provide answers at a general level this also speaks to the design of specific policies such as the survivor benefits provided by the Social Security Administration.

The final substantive chapter, Chapter 4, documents the importance of time gifts made between generations. This is an initial attempt at understanding the role of the exchange motive in a dynamic context. An obvious next step would be to fully estimate a model of the exchange motive, however, it may also be feasible to develop a structural model of multiple motives for inter vivos transfers and bequests. Empirically identifying these motive remains a considerable challenge, but it would be interesting to see to what extent progress can be made in this direction.

Appendix A

Appendix to Chapter 2

A.1 Computing Household Level Permanent Income

Following De Nardi et al. (2018) I infer household level measures of permanent income that is invariant to the household structure. Individual non-labor income is the sum of state pension income, private pension income, annuity income, war pensions, widows pensions and any other declared non-labor income. It excludes both employment and self-employment income. Other than state pensions it does not include benefit income (in the model benefit income is computed as part of the tax function). For singles household income is the same as individual income and for couples it is the sum across husband and wife.

I assume log household income for household i at age j follows:

$$\begin{aligned}\ln y_{i,j} = \beta_0 + \mathbf{1}[f_{i,j} = single\ man] \cdot (\beta_{sman} + \beta_{age\times sman} \cdot j) \\ + \mathbf{1}[f_{i,j} = single\ woman] \cdot (\beta_{swoman} + \beta_{age\times swoman} \cdot j) \\ + \beta_{age} \cdot j + \beta_{age^2} \cdot j^2 + \beta_{age^3} \cdot j^3 \\ + \beta_{PI} \cdot I_i + \beta_{PI^2} \cdot I_i^2 + \beta_{PI^3} \cdot I_i^3 + \beta_{PI^4} \cdot I_i^4 + \beta_{PI^5} \cdot I_i^5 + e_{i,j}\end{aligned} \tag{A.1}$$

where as in the main text $f_{i,j}$ represents family status for household i at age j and I_i their time invariant permanent income. In practice I estimate the following fixed effect regression to obtain consistent estimates of the coefficients on age, family structure and their interaction:

$$\begin{aligned}\ln y_{i,j} = \beta_0 + \mathbf{1}[f_{i,j} = single\ man] \cdot (\beta_{sman} + \beta_{age \times sman} \cdot j) \\ + \mathbf{1}[f_{i,j} = single\ woman] \cdot (\beta_{swoman} + \beta_{age \times swoman} \cdot j) \\ + \beta_{age} \cdot j + \beta_{age^2} \cdot j^2 + \beta_{age^3} \cdot j^3 + \gamma_i + e_{i,j}\end{aligned} \quad \text{(A.2)}$$

For each household the estimated vector of coefficients is used to compute the mean residual (or the estimate of their fixed effect) $\hat{\gamma}_i$. $\hat{I}_i$ is computed as the percentile rank of $\hat{\gamma}_i$. The final step is to perform the following regression:

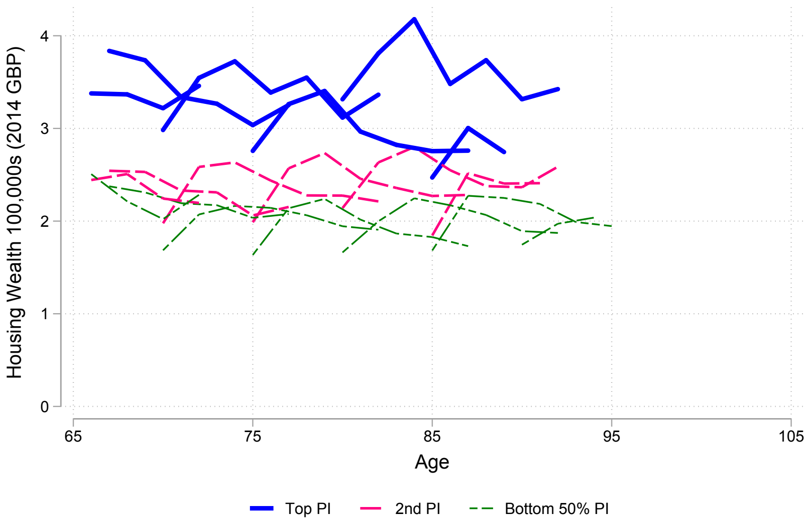
$$\hat{\gamma}_i + \hat{e_{i,j}} = \beta_{0,PI} + \beta_{PI} \cdot \hat{I}_i + \beta_{PI^2} \cdot \hat{I}_i^2 + \beta_{PI^3} \cdot \hat{I}_i^3 + \beta_{PI^4} \cdot \hat{I}_i^4 + \beta_{PI^5} \cdot \hat{I}_i^5 + u_{i,j} \quad \text{(A.3)}$$

Which recovers the mapping from the permanent income index to the log of household income. For further exposition I refer the reader to De Nardi et al. (2018).

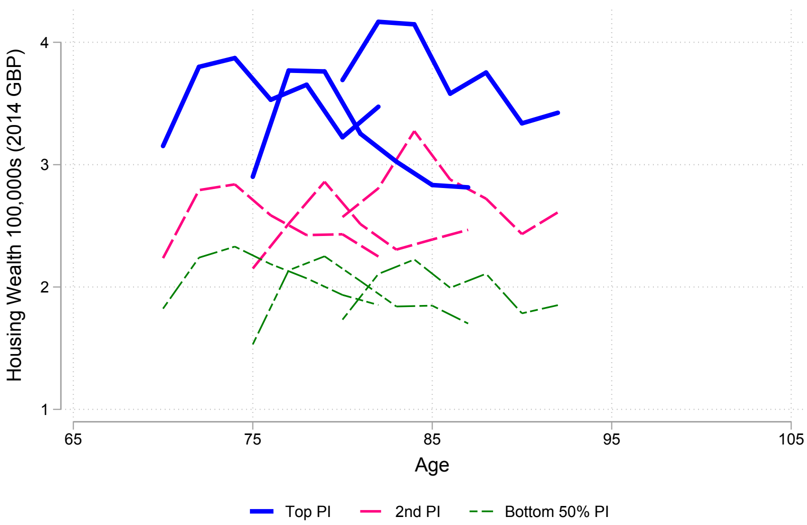
A.2 Attrition and Composition Bias

In this appendix I present results comparing the main sample with an alternative subsample - households who enter in the first wave of ELSA and remain in the sample until the final wave of my sample. These are the unbalanced and balanced panels respectively. By considering the same sample of households this eliminates composition bias as the year on year changes in asset levels are the changes experienced by those households and not changes in the composition of the sample. However, this imposes stricter selection criteria - those who remain in the sample for its full length are, on average, richer than those who exit the sample. Consequently, this selected sample is initially richer than the unbalanced panel and the results are not representative of the elderly population. This is true within both birth cohort and PI grouping.

Figure A.1 presents the housing wealth of these two samples while figure A.2b presents the liquid wealth of these households. The left hand panels reproduces results for the unbalanced samples from figures 2.3a and 2.3b.



(a) Mean Housing Wealth (Unbalanced)



(b) Mean Housing Wealth (Balanced)

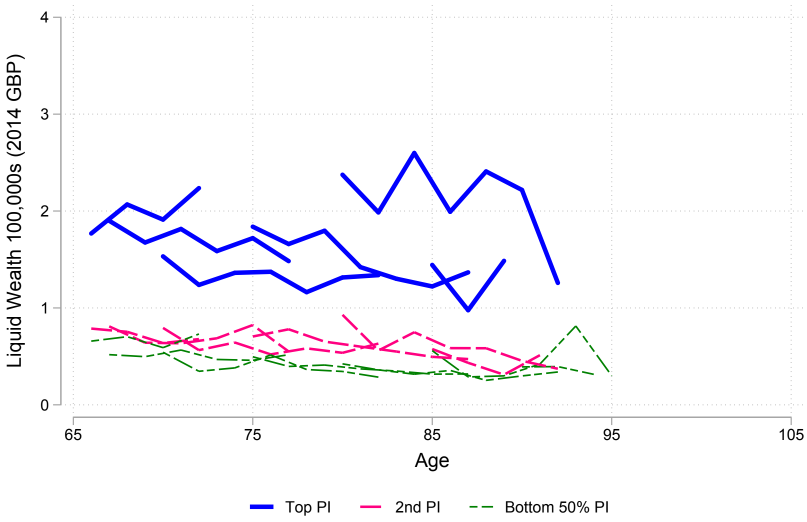
Figure A.1: Housing Wealth by Cohort & PI (Initial Owners)

Comparing the unbalanced panel to the balanced panel in figure A.1 there are two important differences. First, the youngest and oldest cohorts are omitted from the balanced panel as they do not satisfy sample selection in the first wave or survive until the final wave. Second, within PI groups and birth cohorts, the balanced panel has more wealth than the unbalanced panel. There is little evidence of more deccumulation in the housing wealth of the balanced panel than the unbalanced panel for any of the PI groups. For some groups, relative to the start of the sample, there is a small amount of overall deccumulation of housing wealth, however, conditioning on PI and birth cohort does almost completely removes the attrition bias. The housing wealth profiles in both the unbalanced and balanced panels are qualitatively and quantitatively similar, confirming two of the key findings above.

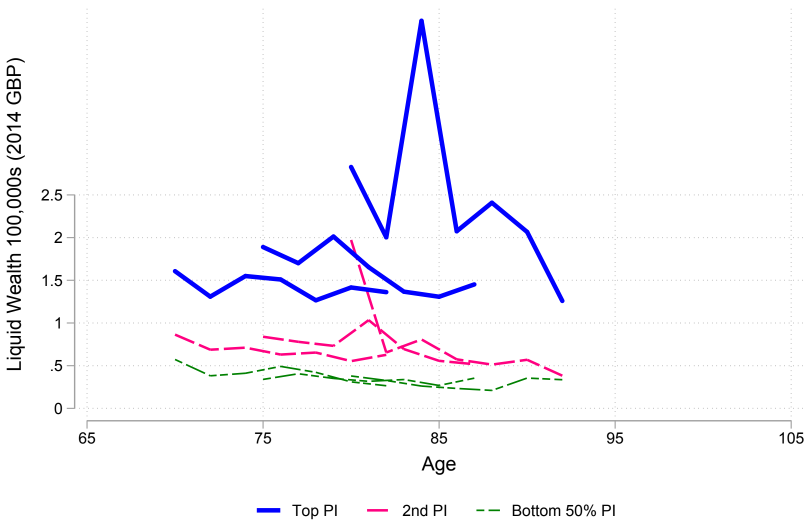
First, the overall feature of these profiles is that deccumulation in housing wealth is slow. Second, the housing wealth of households rises and falls in line with aggregate trends. These key facts are present in both the unbalanced and balanced sample, which shows that they are not driven by compositional changes in the unbalanced sample.

Similarly, when comparing the liquid wealth of the balanced and unbalanced panels (figure A.2) the top two PI groups in the balanced panel are richer than their unbalanced panel counterparts. The liquid wealth of the bottom group is similar in both the balanced and unbalanced panel. There is evidence of more deccumulation in the balanced panel. The top two PI groupings in the balanced panel deccumulate more liquid wealth than in the unbalanced panel. Despite more deccumulation of liquid wealth by the top two PI groups in the balanced panel, the pattern of slow liquid wealth deccumulation in retirement is present in both the unbalanced and balanced samples.

The key findings in this section, that households deccumulate wealth slowly and have wealth profiles that are driven by aggregate trends, are not caused by compositional bias. These results are consistent with similar exercises for the US in De Nardi et al. (2010, 2018) and for the UK and US in Blundell et al. (2016b).

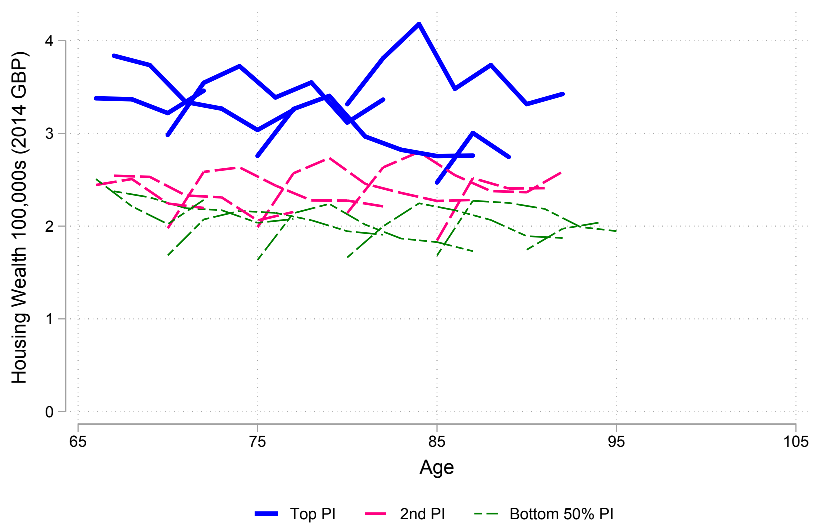


(a) Mean Liquid Wealth (Unbalanced)

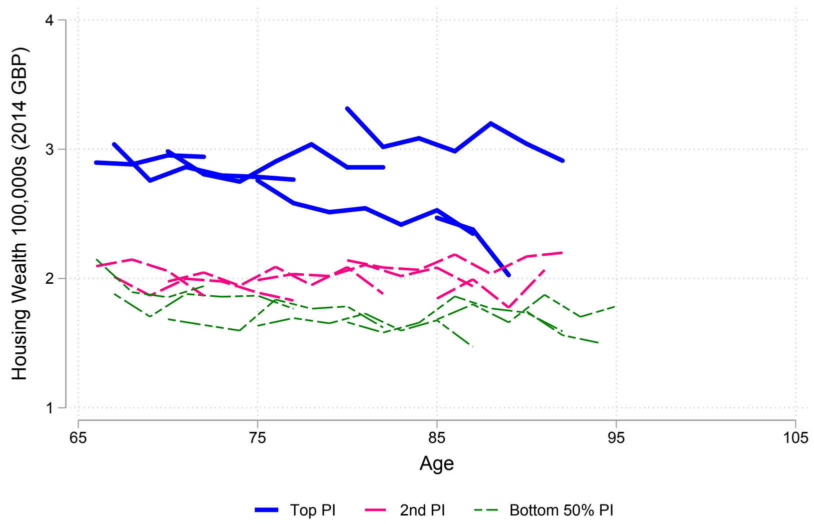


(b) Mean Liquid Wealth (Balanced)

Figure A.2: Liquid Wealth by Cohort & PI (Initial Owners)



(a) Mean Housing Wealth



(b) Mean Housing Wealth (HPI Deflated)

Figure A.3: Housing Wealth by Cohort & PI (Initial Owners)

A.3 House Price Deflated Wealth Profiles

For this population standard forward mortgages are not widely available and in this sample period reverse mortgages are very rare, therefore, many households withdraw housing wealth by downsizing.[1] As discussed in the section 2.2, volatility in house prices causes figure 2.2 to conflate active and passing saving.

[1] I provide further discussion of the UK and US reverse mortgage markets in appendix A.8 and discuss collateralized and uncollateralized borrowing in section 2.4 where I describe the model.

Households care about their realized wealth, however, this masks the amount of deccumulation that occurs for individual households. As an attempt to disentangle the active and passive saving decisions of households I construct counter-factual housing wealth profiles that hold constant aggregate trends.

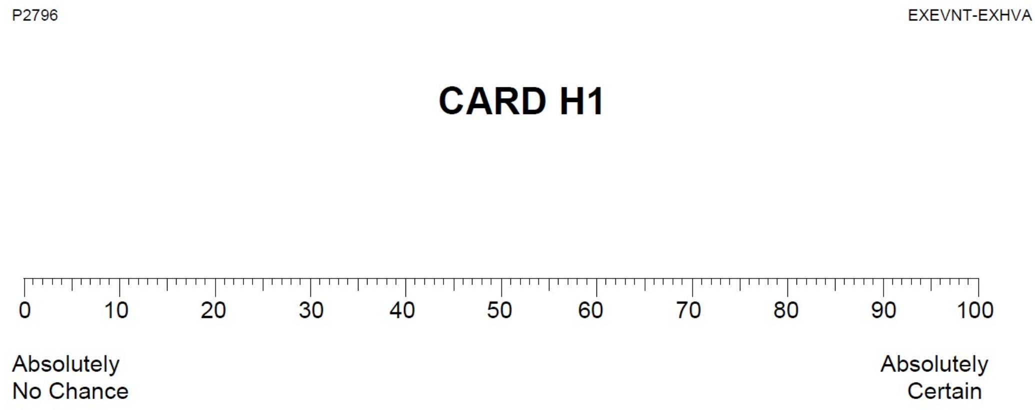
In panel A.3a I reproduce the housing wealth profile of the initial owners. While figure A.3b shows the housing wealth profile (separated by birth cohort and permanent income) for the same households, but deflating the value of their housing by a Housing Price Index.[2]

Absent price changes, on average households slowly deccumulate housing wealth. Comparing the two panels of figure A.3 shows that the change in asset prices obscures the decision by households to on average lower their housing wealth. Although there is some heterogeneity across cohorts and PI, mean housing wealth (the top panel) exhibits only small changes between the start of the sample period an the end. The deflated profiles show that the effects of passive saving explain the initial increase in housing wealth and the maintained high levels despite the active dis-saving decision of households at the mean (an infrequent decision).

A.4 Subjective Bequest Probabilities: Further Details

As in the HRS, ELSA respondents are asked to provide their subjective probabilities for a number of events. In addition to the questions about the distribution of bequests they may leave, households in ELSA are asked about their subjective survival probabilities; that they will have to enter a nursing home or pay for long term care and the probability that they have sufficient financial resources to meet their needs in retirement. ELSA subjects are given the following prompt before being

[2] I use the UK HPI provided by HM Land Registry which uses transaction data (include cash purchases) to construct house prices. A similar exercise, without differentiating by permanent income or initial home ownership status, is presented in Blundell et al. (2016b). Household downsizing decisions in year t are made with the actual wealth they have at t. By deflating the profiles, I implicitly assume that a household who release 25% of their true value at time t would also release 25% of their deflated value - this allows for the mechanical effect of the counter-factual price, but no behavioural effect. This is a strong (and unreasonable) assumption, but it highlights the first order effect of house price growth on the housing wealth profile.



Source: English Longitudinal Study of Ageing

Figure A.4: Interviewer Show Card

asked any of these questions

> ‘Now I have some questions about how likely you think various events might be. When I ask a question I’d like you to give me a number from 0 to 100, where 0 means that you think there is absolutely no chance an event will happen, and 100 means that you think the event is absolutely certain to happen.’

and provided with the show card in figure A.4. Furthermore, they are asked a practise question of weather events in order to familiarize them with the concept. I provide an example of the subjective bequest expectation below.

> ’Including property and other valuables that you might own, what are the chances that you will leave an inheritance totalling £150,000 or more?’

Despite the battery of subjective probability questions it is still possible that respondents may not understand the question or that the reported probability is un-informative and does not reflect the individual’s actual subjective expectation. The ideal test of the assumption that they are informative would be to use realized and counter-factual bequests for each household. However, this is impossible as I do not observe bequests under the counter-factual.

The results in the main text demonstrate that subjective probabilities correlate with household wealth. In order to further validate the information content of the subjective probability measures I perform two validation exercises. First, I make use of an approach in Hurd and Smith (2002) to validate similar HRS questions which follows from a recursive decomposition. Second, I test the significance of lagged subjective probabilities to predict future wealth (de)accumulation.

For an age j individual i the subjective probability $y_{i,j}$ satisfies:

$$y_{i,j} = Pr(bequest \geq £150,00) = \tag{A.4}$$

$$E_j\Big[\sum_{t=j}^{T}[Pr(wealth_{i,j+1} \geq £150\text{k}) \cdot \prod_{h=j}^{t} surv_h \cdot (1 - surv_{t+1})\Big] \tag{A.5}$$

$$y_{i,j} = \mathbf{1}\{wealth_{i,j+1} \geq £150\text{k}\} \times (1 - surv_{i,j+1}) + surv_{i,j+1} \times E_t[y_{i,j+1}] \tag{A.6}$$

In this recursive decomposition the first component is the probability I die before the next period and have at least £150,00 in total wealth. The second component is the probability I survive to tomorrow multiplied by my expectation of tomorrow's subjective probability. Letting D_t denote the set of households who die between t and $t+1$, for small time intervals the average over the population at time t satisfies:

$$\bar{y}_{i,t} \approx \sum_{i \in D_t} \mathbf{1}\{wealth_{i,t+1} \geq £150\text{k}\} + \sum_{i \notin D_t} y_{i,t+1} \tag{A.7}$$

The right hand side of this equation can be directly calculated in ELSA up to wave 5 as death data is not available after this date. Finally, the gap between these two objects is given by:

$$Error_t = \bar{y}_{i,t} - \Big(\sum_{i \in D_t} \mathbf{1}\{wealth_{i,t+1} \geq £150\text{k}\} + \sum_{i \notin D_t} y_{i,t+1}\Big) \tag{A.8}$$

Table A.1 reports the wave specific error (measured in percentage points). The average error in household predictions is small when pooling across eaves and is largest when house prices undergo the largest changes in the sample. This demon-

Wave	Error
1	-11.04
2	-5.57
3	-1.37
4	-2.22
All	-5.41

Table A.1: Average Forecast Error

strates that self reported probabilities are consistent with the observed aggregate wealth decisions of households.

A.4.1 Constructing the Bequest Preference Index

The estimating equation from which I recover the bequest preference index described in Section 2.2.5 is given by:

$$y_{i,t} = \beta X_{i,t} + f(W_{i,t}^{liquid}, W_{i,t}^{housing}) + \lambda_t + \gamma_i + u_{i,t} \tag{A.9}$$

I control for total wealth by allowing for within wave quintile specific effects and control separately for home ownership and the share of total wealth in housing. These quantile specific effects impose limited restrictions on the underlying function and allow me to recover the effect of total wealth holdings and portfolio composition. I control for contemporaneous characteristics of the household with time t period controls for age, household income, gender, marital status, health for all household members, subjective survival probabilities, and vital statistics of their parents as well as wave fixed effects.[3] Finally, the object of interest is a household specific fixed effect γ_i. Household fixed effects are additionally rezidualized on time invariant permanent income and birth cohort dummies.

As discussed in the main text, I place no direct interpretation on the coefficients

[3]Including wave specific effects is important in this analysis because the survey question is defined in reference to a nominal threshold that is fixed for all waves and consequently progressively less informative in real terms.

Panel A: Higher Order Polynomials

	Band around cuttoff				
Order of polynomial	10%	15%	20%	25%	30%
Common Slope					
Cubic	-0.0548*	-0.0377	-0.0575***	-0.0533***	-0.0287**
	(0.0332)	(0.0289)	(0.0180)	(0.0176)	(0.0140)
	-730.9	*-877.8*	*-2124*	*-2118*	*-2871*
Quartic	-0.0383	-0.0305	-0.0519**	-0.0549***	-0.0293**
	(0.0398)	(0.0294)	(0.0207)	(0.0180)	(0.0148)
	-727.7	*-882.1*	*-2127*	*-2116*	*-2869*
N	1224	1559	3023	3233	3979

Panel B: Bounded Second Derivative Inference

Smoothness (K)	0.001	0.01	0.02	0.1	0.1
Local Linear	-0.0290	-0.0290	-0.0290	-0.0290	-0.0290
BSD CI	[-0.0571,-0.000892]	[-0.0571,-0.000875]	[-0.0572,-0.000826]	[-0.0587,0.000694]	[-0.0539,-0.00405]
Implied Bandwidth	30%	30%	30%	30%	30%
Significance Level	5%	5%	5%	5%	10%
Eff. Sample Size	943	943	943	943	943

All regressions additionally control for wave fixed effects, a polynomial in age, household demographics, a polynomial in permanent income and region dummies. Following Kolesár and Rothe (2018), Standard Errors in panel A are clustered by household. The Akike Information Criterion is shown in italics. In panel B the implied bandwidth is the one that minimizes the length of the resulting CI for a given choice of K. * $p < 0.10$, ** $p < 0.5$, *** $p < 0.01$

Table A.2: The Effect of Transaction Taxes on Household Mobility

in this regression and view as a statistical exercise designed to capture systematic differences across households. For the reasons discussed there, I view the estimated household specific fixed effects $\hat{\gamma}_i$ as noisy, but informative measures of underlying differences in household preferences.

A.5 Additional Reduced Form Results

A.5.1 Additional RDD results

Table A.2 presents additional results from the regression discontinuity estimation of the effect of transaction taxes on the mobility of older household described in section 2.3.

Panel A shows the robustness of the common slope estimates to a higher order polynomial. First, using both the cubic and quartic estimators the estimated treatment effect remains negative for all bandwidths and is significant for larger bandwidths around the discontinuity in the SDLT schedule. The results are consistent with the lower order polynomial and local linear estimates presented in the main text. Gelman and Imbens (2017) caution over using higher-order global polynomials as they may lead to noisy estimates, sensitivity to the degree of the polynomial and poor coverage of confidence intervals. While the results here demonstrate some of these problems, nevertheless the resulting point estimates are similar to those obtained with lower order polynomials and non-parametric methods (particularly over the largest estimation window). The estimated treatment effects are not driven by the choice of approximation to the conditional expectation function.

A key issue in many of the applications of regression discontinuity analysis is the underlying discrete support of the forcing variable - this is also a concern in this context. In Table 2.4, tests of statistical significance use standard errors clustered at the household level. As recommended by Kolesár and Rothe (2018) they are not clustered at the values in the support of the forcing variable (Lee and Card, 2008, motivates adjusting standard errors in this manner). However, the confidence intervals in Table 2.4 undercover the true average treatment effect when model misspecification bias is large (typically when large bandwidths are used or the discrete support leads to insufficient observations in a small neighbourhood of the threshold). Panel B uses an alternative method in Kolesár and Rothe (2018) to construct confidence intervals that corrects for misspecification bias and has guaranteed coverage properties.

Implementing this alternative method requires that the researcher chooses a smoothness constant K (which is equivalent to a bound on the second derivative of the conditional expectation function) with a value of $K = 0$ indicating that the conditional expectation function is known to be linear. In each column the bias corrected point estimate is estimated using a local linear estimator and with the bandwidth chosen optimally for a fixed smoothness constant (and smoothness class).

Column 1-5 report confidence intervals constructed with $K \in \{0.001, 0.01, 0.02, 0.1\}$ which represent a range of smoothness parameters ranging from an 'optimistic' to a 'pessimistic' choice (the central values are chosen to sandwich a lower bound

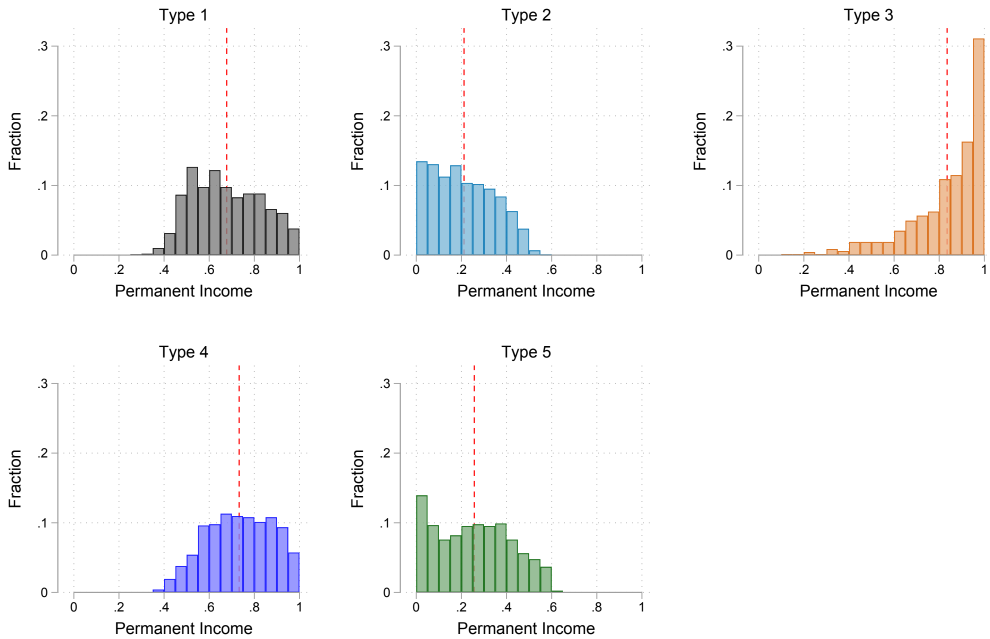


Figure A.5: Permanent Income Across Groups

estimate for K of 0.012). The resulting confident intervals are reasonably tight and are close to the clustered standard errors reported in Table 2.4. Results remain statistically different from 0 at the 5% significance level for all but the most pessimistic value of the smoothness value. Furthermore, even for this extreme case a 90% confidence interval excludes 0.[4]

A.6 Additional Clustering Details

A.6.1 Marginal Distribution of Household Characteristics

The marginal distributions of the household characteristics z_i by household type are displayed in figures A.5 to A.7 with mean values denoted by the dashed vertical line. These marginal distributions give a succinct description of how the *k-means* clustering algorithm partitions household's based on their characteristics.

Figure A.5 plots the marginal distribution of permanent income by clusters. Both Type 1 and Type 4 are primarily drawn from the upper half of the PI distribu-

[4]Although this is not analogous to a one-sided test it is also the case that the null hypothesis of a weakly positive treatment effect is rejected at the 5% significance level.

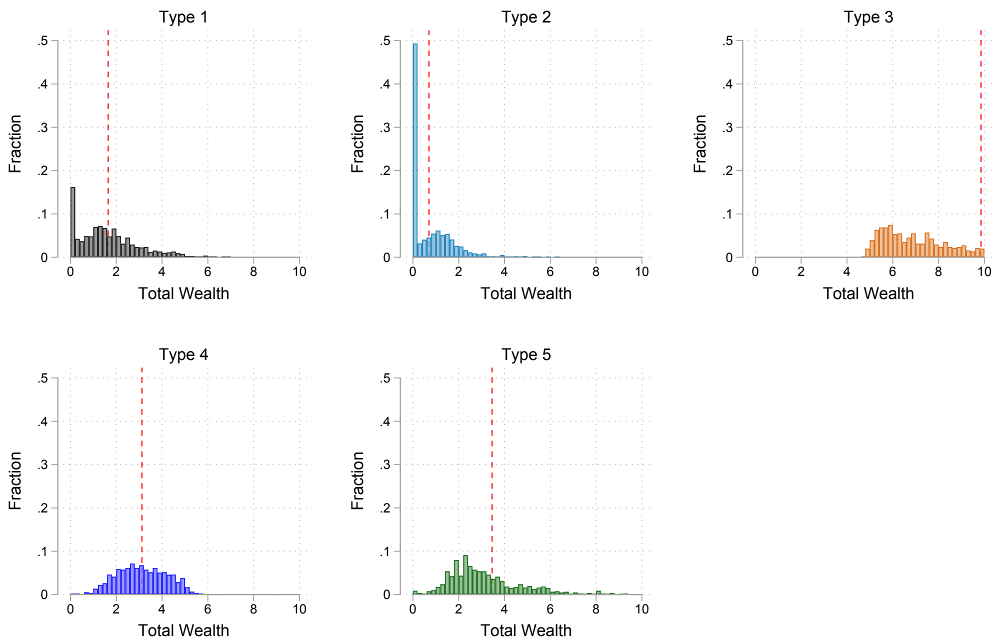


Figure A.6: Total Wealth Across Groups

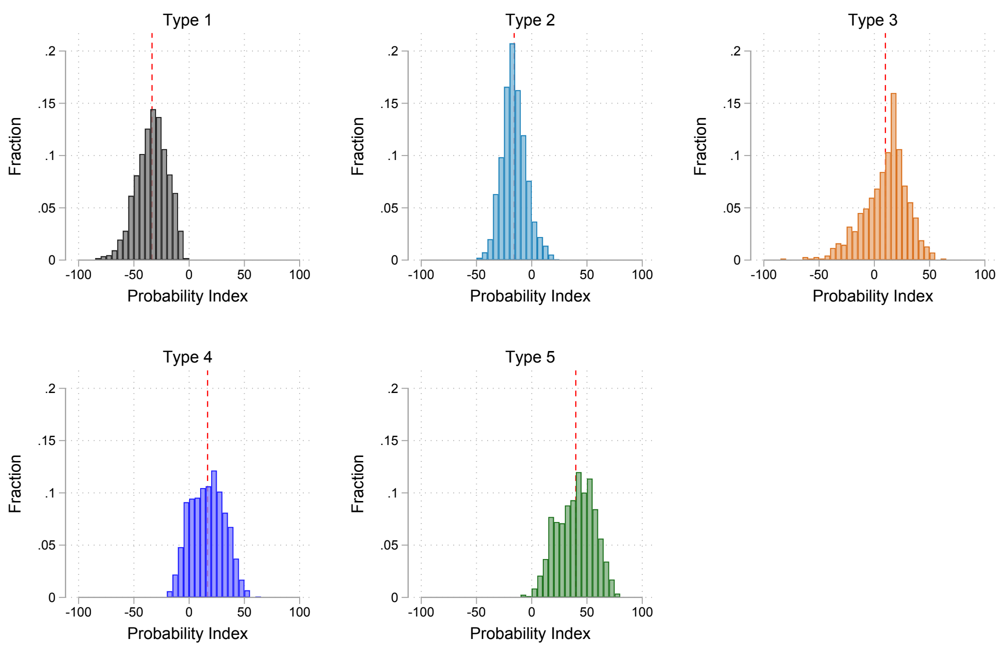


Figure A.7: Bequest Probability Index Across Groups tion. Type 3 is also concentrated in the upper half of the PI distribution and is drawn from primarily the richest households. In contrast, Types 2 and 4 have the majority of households are drawn from the bottom half of the PI distribution. Despite their different clusters, Type 1 and 4 have similar marginal distributions of permanent income and the same is true for the marginal distributions for Types 2 and 5.

Types 1 and 2 (figure A.6) have the lowest level of assets, with the majority of renters contained in Type 2. Types 4 and 5 hold more wealth, with Type 3 having the highest average wealth (and a long tail which is omitted from figure A.6 for ease of comparison). Despite similar PI distributions Types 1 and 4 differ in their average wealth holdings by almost £150,000. Similarly, Types 2 and 5 (who also had similar PI distributions) differ by £250,000. Despite largely being drawn from opposing halves of the PI distribution, the marginal distribution of wealth held by Types 4 and 5 is similar. Finally, figure A.7 plots the marginal distribution of the bequest probability index - I adopt a naming convention that means the household groups are labelled by their bequest probability index. On average Type 1 and Type 2 households systematically report lower probabilities of leaving an inheritance than observationally similar households while those in Type 4 and Type 5 report systematically higher probabilities. Type 3 households also report systematically higher probabilities of leaving an inheritance than observationally similar households due to the negative skew in the distribution, however Type 3 contains a large mass of both households reporting systematically higher and lower probabilities. Although the clustering algorithm assigns each household to a fixed group, the support of z_i for each Type is not exclusive and there is considerable variation in household characteristics within types.

A.6.2 The Number of Clusters in the K-Means Clustering

The approach to estimating household level latent preference types in this paper draws on the two step procedure in Bonhomme et al. (2019). When using the k-means clustering approach, the researcher is left with two degrees of freedom: a) which variables to use to cluster the households (denoted by the vector z) and b) the number of clusters.[5] The choice of variables to cluster on is motivated by the economic problem agents face and is discussed in more detail in Section, this appendix details the choice of the number of clusters.

In order to select the number of clusters used in the analysis, I follow standard data-based methods used in the machine learning literature (See Hastie et al., 2009, for an overview of both clustering methods and data-based heuristics used in k-means clustering). First, restating the clustering problem indexed by a given number of clusters K:

$$\min_{\mathscr{K}_K,\{\bar{z}_k\}_{k=1}^K} \sum_{k=1}^{K} N_k \sum_{k(i)=k} || z_i - \bar{z}_k ||^2 = \min_{\mathscr{K}_K,\{\bar{z}_k\}_{k=1}^K} SSE_K \tag{A.10}$$

$$\bar{z}_k = \frac{1}{N_k} \sum_{k(i)=k} z_i$$

Where the classification (for a given number of clusters K) is given by:

$$\mathscr{K}_K = \{k(i)\}_{i=1}^{n} \tag{A.11}$$

In order to select the optimal number of clusters, separate solutions are obtained to the problem in equation for number of clusters $K \in \{1,...,K_{max}\}$. It is well documented that the measure of within cluster dissimilarity (the Sum of Squared Errors) is decreasing in the number of clusters k which precludes the use of cross-validation techniques. Instead, a number of heuristics which use the following intuition are proposed: suppose there is a true number of clusters in the data K^{true}. Then for $K < K^{true}$ the algorithm assigns a subset of the true groups to each cluster, consequently, increasing the number of clusters allows the algorithm to assign groups in a subset to a new cluster. To the extent that these subgroups are strict, increasing the number of clusters when $K < K^{true}$ is associated with a large decrease in the measure of within cluster dissimilarity. In contrast, when $K > K^{true}$ one of the clusters

[5] It is also necessary to specify the initialization of the clusters, however, I use a multi-start algorithm where the initial assignment of clusters across households is drawn from 10,000 random seeds

partitions a true group into two clusters and the decrease in the measure of within cluster dissimilarity must be smaller. This logic forms the basis of using within cluster dissimilarity measures to select the optimal number of clusters - the optimal number of clusters is located at the kink in the marginal decrease in the Sum of Squared Errors.[6]

Figure A.8 displays results for two commonly used heuristic methods for identifying this kink point. Figure A.8a plots the Sum of Squared Errors (SSE_K) against the number of clusters. The 'Elbow statistic' identifies the kink point from this graph by visual inspection. Using this metric suggests a kink point at $K = 5$.

Figure A.8b instead plots the Gap statistic Tibshirani et al. (2001) which is an automated way of identifying the kink point. The gap statistic is defined as:

$$Gap_n(K) = E_n^*[log(SSE_K)] - log(SSE_K) \tag{A.12}$$

Where E_n^* denotes a bootstrapped expectation drawn from a uniform sampling over the data.[7]. Let B denote the number of bootstrap replications and $sd_B(K)$ denote the standard deviation of the $log(SSE_K)$ replications. Then

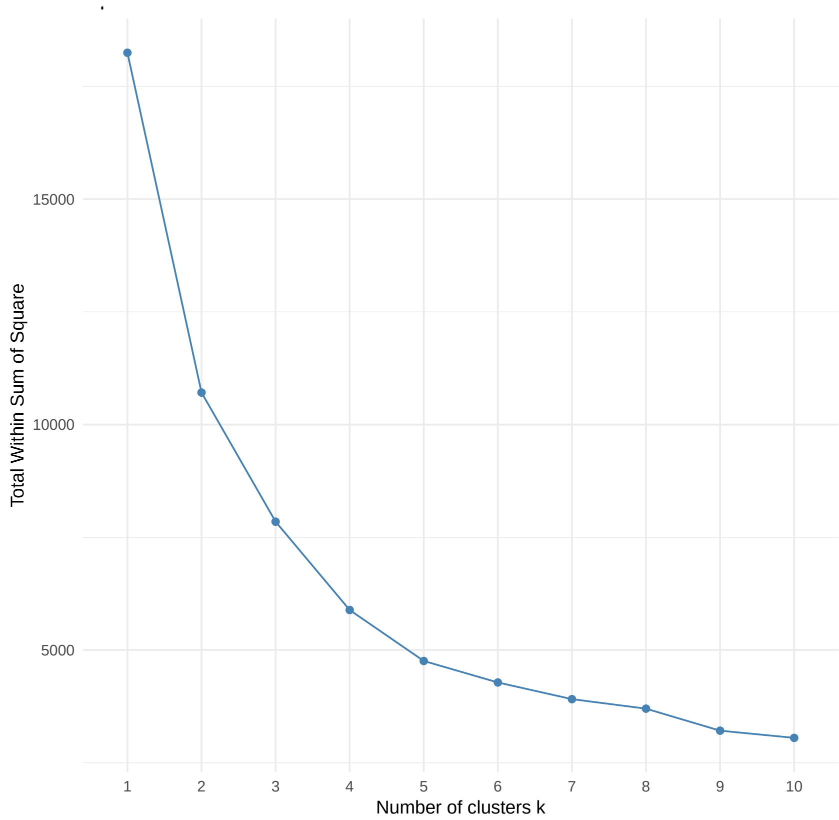
$$\hat{K}^{true} = \arg\min_K \{K : Gap_n(K) \geq Gap_n(K+1) - s_{K+1}\} \tag{A.13}$$

$$s_K = sd_B(K)\sqrt{(1+1/B)} \tag{A.14}$$

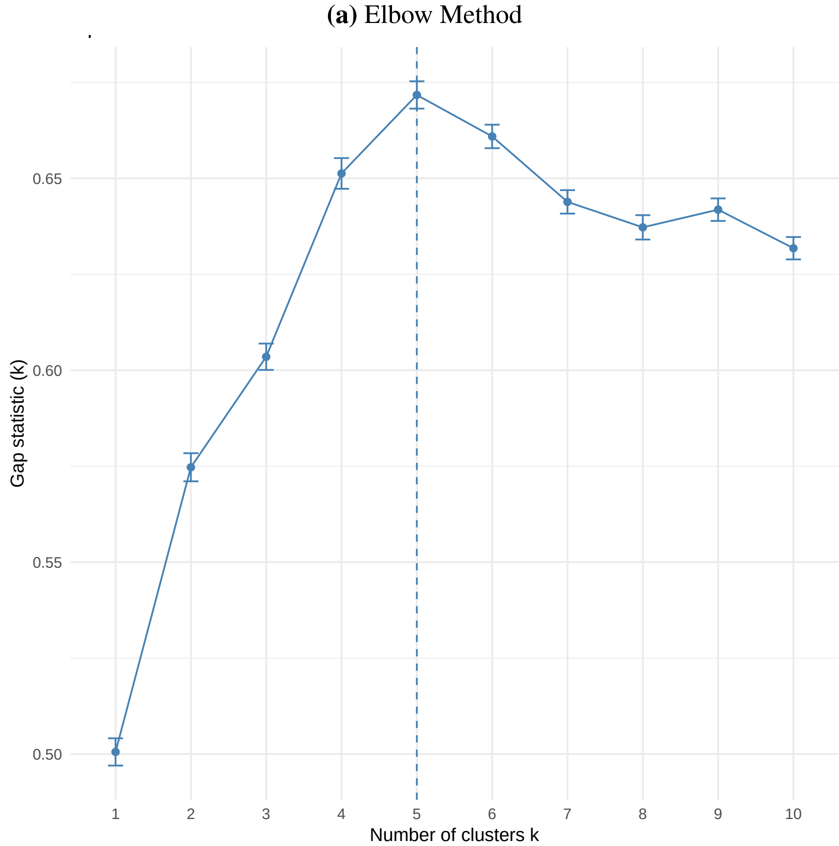
In words, the Gap statistic identifies the optimal number of clusters as the smallest K such that the increase in the Gap statistic is less the the simulation error (displayed in the error bars in figure A.8b). The optimal number clusters as determined by the Gap statistic is $K = 5$.

One potential concern in using a two-step approach to estimating latent heterogeneity is that the sample size in each cluster may be small which may deliver imprecise estimates in the second stage without enough variation to credibly iden-

[6]Alternatively, where the marginal increase in the Explained Sum of Squares begins to asymptote
[7]In practice, I use 50 bootstrapped replications to calculate the expected $log(SSE_K)$



(a) Elbow Method



(b) Gap Statistic

Figure A.8: Optimal Number of Clusters tify differences across the population. Instead, the optimal number of clusters determined by the heuristic methods generates groups with large sample sizes.

A.6.3 Alternative Clustering Procedure

The estimation procedure in this paper estimates preference parameters taking the results of the assignment procedure as given. Although the k-means algorithm is a popular technique to partition the data into clusters and is widely used in economic applications of group fixed effects, a number of alternative methods for partitioning data into clusters are used in other applications - these alternative methods often return different partitions of the underlying data. It is not feasible to estimate the model for a variety of assignment procedures and, instead, I present a comparison of the partitions under an alternative k-medoids assignment procedure (holding fixed the number of clusters).

Formally, k-medoids clustering is defined as:

$$\min_{\mathscr{K}_K, \{\bar{z}_k\}_{k=1}^K} \sum_{k=1}^{K} N_k \sum_{k(i)=k} || z_i - \bar{z}_k || \tag{A.15}$$

$$\bar{z}_k = \text{median}_{k(i)=k}(z_i)$$

While the k-means procedure minimizes the sum of squared Euclidean distances, the k-medoids procedure instead minimizes a sum of pairwise distances. Consequently, it is more robust to large outliers and measurement error (that produces large distances).

A comparison between the estimated clusters under a k-means and a k-medoids procedure are displayed in Table A.3.

A.7 Numerical Procedure

The dynamic programming problem described in Section 2.4 does not admit a closed form analytic solution.

I solve the model using backwards induction. At each age I compute the optimal savings, housing and consumption decision for all possible combinations of

		k-medoids					
		Type 1	Type 2	Type 3	Type 4	Type 5	Share
k-means	Type 1	97.76	0	0.35	1.89	0	18.23%
	Type 2	10.48	89.27	0	0	0.25	34.09%
	Type 3	0	0	93.72	6.28	0	12.33%
	Type 4	1.73	0.1	0	95.71	2.45	21.09%
	Type 5	0	0.6	0.15	1.06	98.19	14.27%
Share		21.73%	30.54%	11.64%	21.45%	14.52%	

Table A.3: Comparison of Latent Household Types with Alternative Clustering Procedure

the state variables. I use the policy functions to compute the value function and iterate backwards. Given the optimal household decisions for a given set of parameter values and household initial conditions I simulate forward households through the different policy regimes drawing values of ξ and ζ from their distribution using Monte Carlo methods. I then construct moment conditions from this simulated data in the exact same way as in the data - this forms the basis of my estimation procedure.

This appendix discusses the implementation of each of these procedures in more detail.

A.7.1 Discretization

The model has four discrete state variables: age, health status, family structure and idiosyncratic bequest motive. There are four additional state variables that must be discretized: permanent income, housing, cash on hand, and the aggregate house price level as well as the additional transitory medical expense shock. Permanent income is placed on an unequally spaced grid with 6 elements, where the grid points are concentrated towards the extremes of the distribution.

Housing, which is both a state and a choice variable, is discretized[8] using a

[8] The presence of numerous notches in the transaction tax mean that the budget set is non-convex. The flow utility of housing depends on the initial state variable rather than the choice of next period housing (excepting renters) Together these features necessitate the practical choice to discretize both the state variable and housing choice.

single point to denote current renters and 14 additional points for homeowners - the first 12 points of this grid are placed at the median of the 12 quantiles of the 2002 housing wealth distribution of my ELSA sample (conditional on being below £1,250,000 which covers over 99% of the sample) with two additional points placed at £1,250,000 and £2,500,000. Cash on hand is placed on a grid with 42 points placed on an exponential scale. I use a small number of cash on hand points for the available resource because the solution method (described below) involves calculating an exact solution to the Euler Equation at each point. The log of the aggregate house price level is placed on a grid with 6 elements using the method of Tauchen (1986). Finally, the transitory component of medical expenses is placed on a grid with 3 elements using the method of Tauchen (1986).[9]

Consumption and next period liquid wealth are not placed on a grid. Instead, individuals can choose any feasible level of consumption and next period liquid wealth. In total, the value function and policy functions are calculated for 23,619,600 combinations of state variables for each age and policy regime.

A.7.2 Computing the Solution to the Household's Problem

In order to tractably solve this problem while maintaining a high level of accuracy I model the choice of housing as a discrete choice and follow the modified version of the endogenous grid-point method (EGM) algorithm for discrete continuous dynamic choice models in Iskhakov et al. (2017)[10]. The EGM algorithm was first introduced in economics by Carroll (2006) who demonstrated improvements in both speed and accuracy in a buffer stock savings model.

However, the model presented here introduces non-convexities through the consumption floor and housing choice (which I discretize) as well as the kinks in the transaction tax and tax on estates. In this paper, cash on hand is not deterministic and I adapt their method by controlling for household savings (the deterministic component of cash on hand) as the end of period state variable. Holding fixed the housing choice combining the Euler equation for consumption with the predeter-

[9]Results with 3 or 5 points for the transitory shock are indistinguishable

[10]Fella (2014) also considers a version of the EGM algorithm for non-smooth non-convex problems

mined level of saving delivers the optimal consumption policy, savings and the cash on hand state variable given the housing stock. A variation of the same EGM approach is discussed in **?** for a retirement framework without a housing choice, but featuring public care aversion.

However, housing is a choice variable. The housing stock today effects the marginal utility of consumption today and the level of housing tomorrow effects expected marginal utility of consumption tomorrow and total available resources tomorrow. In practice, at each set of state variables today I compute the implied consumption and savings decision using the household Euler equation for every choice of housing tomorrow. This is referred to as the EGM step which returns the housing choice conditional policy functions[11] for a given set of state variables Ω:

$$c^*(\Omega|h') \tag{A.16}$$

$$a'^*(\Omega|h') \tag{A.17}$$

Which can then be combined to give the housing choice conditional value function:

$$\begin{aligned} V(\Omega|h') = u(s, c^*(\Omega|h'), h, m) + \\ \beta \cdot surv(j, I, m) E[V(\Omega') \,|\, \Omega, h', a'^*(\Omega|h')] \\ +\beta(1 - surv(j, I, m)) E[\phi^i(b) \,|\, \Omega, h', a'^*(\Omega|h')] \end{aligned} \tag{A.18}$$

Given the optimal rules conditional on the choice of housing tomorrow, I compute the conditional value function for every possible choice of housing tomorrow and take the maximum across all house choices. The value function is then given by:

[11] The rental expenditure for current renters follows from the within period marginal rate of substitution.

$$V(\Omega) = \max_{h' \in \mathscr{H}} \{V(\Omega|h')\} \tag{A.19}$$

A.7.2.1 Calculating the Housing Choice Conditional Optimal Policy and Value Function

The Euler equation for the homeowner gives:

$$\begin{aligned} & u_c(s,c,h',m) \geq \\ & \beta \cdot surv(j,I,m) E[\frac{\partial}{\partial \tilde{a}'} V^i_{j+1}(f',I,m',h',x',\zeta',p'_h) \Big| \Omega, c', h'] \\ & \quad + \beta(1 - surv(j,I,m)) E[\frac{\partial b}{\partial \tilde{a}'} \frac{\partial}{\partial b} \phi^i(b) \Big| \Omega, c', h'] \end{aligned} \tag{A.20}$$

Away from the borrowing constraint the Euler equation holds with equality and the EGM approach is to calculate the right hand side of this expression (for a given savings choice) and calculate the implied optimal consumption by inverting the marginal utility of consumption. Given consumption and savings, cash on hand today is found by rearranging the budget constraint.

Conditional on tomorrow's housing choice I follow this procedure and I document the calculation of the expected future marginal utility of saving. For saving below the consumption floor, the marginal utility of saving is 0 and I follow Hubbard et al. (1995) in replacing the consumption floor with an indicator function in the Euler equation.

$$\begin{aligned} \frac{\partial}{\partial \tilde{a}'} V^i_{j+1}(f',I,m',h',x',\zeta',p'_h) = & \frac{\partial x'}{\partial \tilde{a}'} \frac{\partial}{\partial x'} V^i_{j+1}(f',I,m',h',x',\zeta',p'_h) \\ = & (1 + r\tau'_y(ra' + y)) \frac{\partial}{\partial x'} V^i_{j+1}(s',I,m',h',x',\zeta',p'_h) \\ = & (1 + r\tau'_y(ra' + y)) u_c(s', c^\star(\Omega'), h^\star(\Omega'), m') \\ & \times \mathbf{1}[c_{min}(s',h') < \tilde{x}' + \delta h'] \end{aligned} \tag{A.21}$$

This is then inserted into the right hand side of equation A.20. Similarly, the effect of increasing saving on the marginal utility of bequests is given by the following upper envelope of the choice-specific value functions:

$$\begin{aligned}\frac{\partial b}{\partial \tilde{a}'}\frac{\partial}{\partial b}\phi^i(b) =& \frac{\partial b}{\partial \tilde{a}'}\phi^i_b(b) = \frac{\partial b}{\partial a'}\phi^i_b(b)\\ =& \mathbf{1}[Q(0,h',p'_h)+a'>0]\cdot\phi^i_b(Q(0,h',p'_h)+a')\end{aligned} \tag{A.22}$$

Substituting the results in equations A.21 and A.22 and inverting equation A.20 gives:

$$\begin{aligned}c^\star(\Omega|h') = u_c(s,h',m)^{-1}(&\\ &\beta\cdot surv(j,I,m)E[(1+r\tau'_y(ra'+y))u_c(s',c^\star(\Omega'),h^\star(\Omega'),m')\\ &\qquad\times\mathbf{1}[c_{min}(s',h')<\tilde{x}'+\delta h']\,|\,\Omega,\tilde{a}',h']\\ &\quad+\beta(1-surv(j,I,m))E[\mathbf{1}[Q(0,h',p'_h)+a'>0]\\ &\qquad\times\phi^i_b(Q(0,h',p'_h)+a')\,|\,\Omega,\tilde{a}',h'])\end{aligned} \tag{A.23}$$

Where the expected value on the right hand side is left in terms of a' conditional on the chosen level of savings $\tilde{a}'$.

When the continuation value, the sum of the discounted expected value function and expected utility from bequests, is globally concave then the the FOC outlined above will be necessary and sufficient. However, when this continuation value is not concave the FOC is necessary, but not sufficient. The continuation value of the model studied in this paper is not globally concave due to the presence of the consumption floor and the discrete housing decision which introduce kinks in the value function.[12] Consequently, the optimal policies delivered by the EGM step do not necessarily correspond to the optimal policies of the model.

[12]Typically kinks which occur due to next period non-concavities are referred to as *primary* kinks while kinks that perpetuate backwards from future period non-concavities are referred to as *secondary* kinks. The presence of further uncertainty in future periods helps to smooth out some of the secondary kinks, but the approach used here accounts for both types of kinks.

In order to ensure that the globally optimum consumption value is selected from the multiple solutions to the Euler Equation I construct the (housing choice specific) upper envelope over segments of the (housing choice specific) value function in regions of the endogenous cash on hand grid where multiple solutions are detected. This procedure follows the method described in Iskhakov et al. (2017).

The DC-EGM method specifies a grid for the post-decision savings state and returns the housing choice-specific optimal policies and value functions on an endogenous grid. Consequently, an extra step is needed before it is possible to compare them in the upper envelope calculation in A.19. I refer to this step as *regularization* and interpolate each of the housing choice-specific value functions and policy functions onto a pre-specified exogenous cash on hand grid that is common across housing choices. In the *regularization* step, when interpolating the value function for households who choose to locate at the borrowing constraint for the next period I use the analytic solution for their value function (given the computed expected value function associated with the borrowing constraint) next period.

A.7.3 Simulation

I simulate 50,000 simulated households who have initial conditions observed in the data. Household's are simulated through the path of observed shocks (health, mortality and aggregate house prices) together with the entire profile of unobserved shocks. This procedure perfectly replicates any compositional changes in the sample as they age and die.

Given household states the optimal choices in the simulation are calculated starting from the age of first observation and moving forward. As in the solution of the dynamic programming problem the optimal policy is found by first conditioning on next period housing choice and then by maximising over these conditional value functions. In doing so, the problem must now be evaluated at points which are outside the grid chosen in the solution. This is achieved by discretizing the choice of available consumption (into 100 possible choices for the percentage value of current resources) and using linear interpolation.

In the simulation, I assume that households face a maximum of five different tax regimes during the sample period (depending on their entry into the sample and duration). The transition across these tax regimes is treated as a zero probability event with the exception of the tax regime which corresponds to the so called "Stamp Duty Holiday". This means the model is solved separately for each of these tax regimes and households are simulated through the tax regimes they experience. In the case of the "Stamp Duty Holiday" in line with the nature of the policy, households perfectly forecast the reversion to the previous regime after one period.

A.7.4 2nd Stage Estimation

Given the estimated first stage parameters, the second stage estimation selects the parameter vector $\hat{\theta}$ which minimises the GMM criterion function described in equation A.24.

$$\hat{\theta} = \arg\min_{\theta \in \Theta} G(\theta)' W G(\theta) \tag{A.24}$$

In life cycle models of the type featured here the GMM criterion function may have multiple local minima and without analytic derivatives. It is not possible to formally establish that any optima is a global minimum and in practice Simulated Method of Moments estimates are found by employing multiple starting points for a derivative free optimisation algorithm.

As discussed above, computation of the simulated moments is costly due to the large state space of the model and multiple policy regimes. Consequently, I adopt a methodology similar in spirit to Guvenen and Smith (2014) where minimization of the GMM criterion function proceeds in two steps by combining a form of iterated grid search in the first step with a derivative free optimizer in the third step.

In the first step, I first compute 3,000 candidate parameter values and evaluate the objective function at each parameter value. The candidate values are selected by drawing from a 14-dimensional (the number of parameters to be estimated) *Sobol sequence* which is a low discrepancy quasi-random sequence. The candidate parameter values are then ranked based on the value of the objective function at these parameters. I then use the 30 highest ranked candidate parameter vectors (the 1%

with the smallest value of the objective function) to generate a new hypercube on the parameter space. I take the minimum and maximum parameter value in each dimension of the parameter space and use these as new lower and upper bounds in each of the dimensions. This produces the smallest hypercube which surrounds the polytope defined by the convex hull of the 30 highest ranked candidate parameter vectors. In practice, this greatly reduces the overall admissible parameter space without necessarily producing tight bounds on any individual parameter. I compute a further 3,000 candidate parameters on this new hypercube and iterate on this procedure.[13]. Sampling points from this new hypercube (substantially larger than the convex hull) slows the rate at which regions of the parameter space are discarded and is similar to the averaging of the best estimate and new draws from a Sobol sequence in the *Tik-Tak algorithm* described in Arnoudy et al. (2019). I iterate this step 5 times. The second step uses the top 1% sample from the final set of first stage evaluation as starting values for a derivative free optimiser.

I use the BOBYQA algorithm for numerical optimization Powell (2009), a trust region based method, in the second stage. Typically, BOBYQA uses fewer evaluations of the objective function than other derivative free methods (for example the Nelder-Mead Simplex method Nelder and Mead, 1965). By combining the BOBYQA method with the multiple starting points selected above it appears that the parameters obtain the global minimum.

At each stage of the estimation I parallelize both the calculation of the dynamic programming problem and simulation. In the first stage I also parallelize the evaluations of the candidate parameter vector and in the second stage the derivative free optimisations from different starting points using the facilities of the University College London Computer Science High Performance Computing Cluster.

A.7.5 Computing the Standard Errors

I use the standard formula for the asymptotic variance of the MSM estimator including adjustment for simulation error. I do not adjust for the first stage estimates

[13] Gavazza et al. (2018) use a Sobol sequence to construct candidate parameters in a form of iterated grid search

- in particular, this treats the cluster determined groups as known ex-ante.

To calculate the Jacobian of the moment conditions with respect to the parameters I use numerical differentiation.

A.8 Reverse Mortgages in the UK: Further Details

In the model presented in section 2.4 I assume that retired households do not have access to "Reverse Mortgages" or other home equity release products. As discussed above, in line with the results in Nakajima and Telyukova (2017), it is likely that a small number of households in the model would find it optimal to use a reverse mortgage as a means of equity extraction. In general, older households do not have access to the same lines of credit as they typically have small incomes and large asset positions. Consequently, financial products that allow households to access wealth stored in their home without moving out of the home (home equity release products) have become increasingly popular and the focus of a burgeoning empirical literature.

In the UK market, home equity release products are issued in small volume: between 1991 and 2011 270,000 home equity plans were created for over 55 homeowners at an average of 13,500 a year (SHIP 20th anniversary report). This represents a small fraction of the over 65 population. In the UK, home equity withdrawals typically take the form of lifetime mortgages (as with HECM which capture over 90% of the US market, see Cocco and Lopes, 2019) where households receive a lump sum or line of credit today, but no repayment is due until either the house is sold or the homeowners die. As with standard forward mortgages, these products have up front fixed costs which are rolled into the principle and accrue interest. However, a key difference is that households do not pay down the debt. Consequently, the compounding of fees on the initial fixed costs and earlier interest payments leads to debts that grow over time. Cocco and Lopes document a number of UK and US differences.

There are a number of reasons demand for the product in the UK may be lower (including differences in the burden of medical expenses). Cocco and Lopes document that the average loan to value (LTV) ratio of UK products is much lower than their US counterparts (the LTV ratio is typically around 20 percentage points) and that UK products are characterized by lower up front costs, but higher interest rates. The higher rates offered in the UK are higher than both standard forward mortgage products and their US counterparts. Furthermore, in this context it is important to remember that they also apply to the lower initial fixed costs. In the US, the Federal Housing Administration (FHA) underwrites HECM products (aiming to break even on average) while this is not true in the UK. It may be that this leads to differences in the adverse selection and moral hazard behaviour of product holders (although Davidoff and Wetzel, 2014, suggest that this effect is small in the US context).

It is well documented (Davidoff et al., 2017) that the limited financial literacy of potential product users contributes to the low take up of reverse mortgage products and that those who have indirect experience through peers are more likely to use reverse mortgage products. In the UK context, two major mortgage retailers offered home equity release products between 1996 and 1998 - these products were shared appreciation schemes (where home owners are insured against falls in house prices, but own progressively less and less of the equity in their home when prices rise). The shared appreciation schemes were not subject to standard financial regulation and subsequently were the target of negative press coverage and an ongoing class action lawsuit. In addition to cross country differences in the financial literacy of households it may be that the spillover effects of exposure are different in the UK context.

This appendix documents differences between the UK and US demand for reverse mortgages and their institutional context as well as suggesting some potential explanations. However, as with all financial products the product offered will reflect demand and these cross country differences. See Cocco and Lopes for an overview.

A.9 Activities of Daily Living

In each wave of ELSA each household member is asked to whether they have any difficulties with Activities of Daily Living (ADL). They are asked about difficulties in six different categories of activities:

1. difficulty dressing, including putting on shoes and socks
2. difficulty walking across a room
3. difficulty bathing or showering
4. difficulty eating, such as cutting up food
5. difficulty getting in and out of bed
6. difficulty using the toilet, including getting up or down

Individual's may interpret these questions differently, but they are intended to capture the minimum range of daily activities typically performed by the adult population. As such, they are proxies for an individual's ability to live independently. ELSA also asks individual's about a further set of Instrumental Activities of Daily Living (IADL), including managing money and preparing meals among others, that are further proxies of an individual's ability to live alone. Both ADLs and IADLs are hierarchical, for example the last remaining ADL is typically eating, and are correlated with each other. I use experiencing two difficulties with ADLs as a threshold because this captures both mortality and medical expenditure effects (See Ameriks et al., 2018; Robinson, 1996) in a parsimonious manner.

An alternative approach is to combine these measures with further information on health conditions and health behaviours to build a frailty index (This is the approach pursued in Braun et al., 2019) as a unidimensional proxy for underlying health. There are three reasons I do not use a frailty index approach. First, both ADLs and self reported health status are inputs into standard frailty indexes and highly correlated with additional components. Consequently, the distribution of frailty indexes conditional on the parsimonious three state health status show little overlap in their distributions (they do not formally first order stochastically dominate one another). The second important reason for using an ADL based approach is that the standard inputs used in constructing frailty indices, see Searle et al. (2008)

for example, are not available in all waves. Data on some comorbidities and features of physical health are collected in ELSA by nurse visits and only available in waves 2, 4 and 6. Similarly, health behaviours are missing from some waves and information about particular conditions differs across waves. This makes constructing both a time and epidemiological consistent frailty index problematic. Finally, Using a frailty index would introduce an exogenous, stochastic and continuous state variable. Accurately capturing the dynamics of frailty within a parsimonious state space is an interesting research question within itself and any attempt to discretize the frailty state space requires imposing potentially ad hoc thresholds. It is the author's belief that using an established ADL measure in this context captures the benefits of a more complicated health model while maintaining both tractability and transparency without exceeding the limits of the data.

A.10 Computing the Bequest Share

As discussed in Section 2.7, it is common to report the marginal propensity to consume out of final period wealth and threshold levels of annuity consumption above which households have a positive marginal propensity to bequeath. In order to facilitate this comparison I compute the solution to the following static allocation problem for a single renter:

$$\max_{c,h^r} u(c,h^r) + \beta\phi(b) \tag{A.25}$$

s.t.

$$c + r^h p_h h^r + b = x \tag{A.26}$$

To simplify the discussion, I solve the problem expressed in terms of expenditures, e, under unit house prices. The indirect utility function is given by:

$$u^e(e) = \frac{e^{1-\gamma}}{1-\gamma} \times \left[\sigma^\sigma \left(\frac{1-\sigma}{r^h}\right)^{1-\sigma}\right]^{1-\gamma} = \frac{e^{1-\gamma}}{1-\gamma} \times \bar{u} \tag{A.27}$$

And the following maximisation problem defines the allocation between within period expenditures and bequests:

$$\max_{e \leq x} u^e(e) + \beta\phi(x - e) \tag{A.28}$$

The solution to this static allocation problem, where a single household knows they will die with certainty next period and faces no medical expenses, characterises the marginal propensity to bequeath and the threshold level of annuity consumption above which bequests are operative.

The first order condition for an interior solution equates marginal utility of expenditure today with the discounted marginal utility of leaving a bequest. Using the budget constraint and the first order condition delivers the following expression for the marginal propensity to expend at an interior solution:[14]

$$MPE = \frac{\bar{\phi}}{1+\bar{\phi}} \quad \text{where} \quad \bar{\phi} = \left(\frac{\beta\phi_1}{\bar{u}}\right)^{-\frac{1}{\gamma}} \tag{A.29}$$

The threshold value of final period wealth above which bequest motives become operative (or the annuity value of consumption) is given by:

$$c_{beq} = \bar{\phi} \times \phi_2 \tag{A.30}$$

Finally, to generate the results in Figure 2.11 the author must take a stand on how to bring their differing frameworks in line with the model estimated in this paper. This is requires two assumptions. First, I assume that (in line with the model in this paper) bequest utility is discounted by the time preference β and liquid assets do not earn a return when the value of consolidated wealth that enters bequest utility is calculated. Second, of the papers from the related literature presented in

[14] Alternatively, (e.g. De Nardi et al., 2010) this is reported as

$$MPE = \frac{1}{1+\tilde{\phi}} \quad \text{where} \quad \tilde{\phi} = \bar{\phi}^{-1}$$

	$\bar{y}$	λ_y	τ_y	R^2
Singles	556	99.1	0.468	0.90
	(737)	(24.8)	(.0213)	
Couples	5,083	7.62	0.213	0.927
	(819)	(2.32)	(0.0262)	

Table A.4: Tax Function Parameter Estimates

Figure 2.11, only Nakajima and Telyukova (2018a) estimate a model with housing. Consequently, for all other papers I use the non-housing share of consumption from the baseline estimates to calculate $\bar{u}$ while continuing to use their own estimate of the coefficient of relative risk aversion, γ. These assumptions generate minor differences between the MPC reported in a given study and the MPE used in this paper, but do not change any of the qualitative implications for the strength of the bequest motives estimated here when compared to other estimates in the literature. The final step adjusts the curvature of the bequest motive to homogenize the studies for a 2014 price level.

A.11 Additional Estimation Details

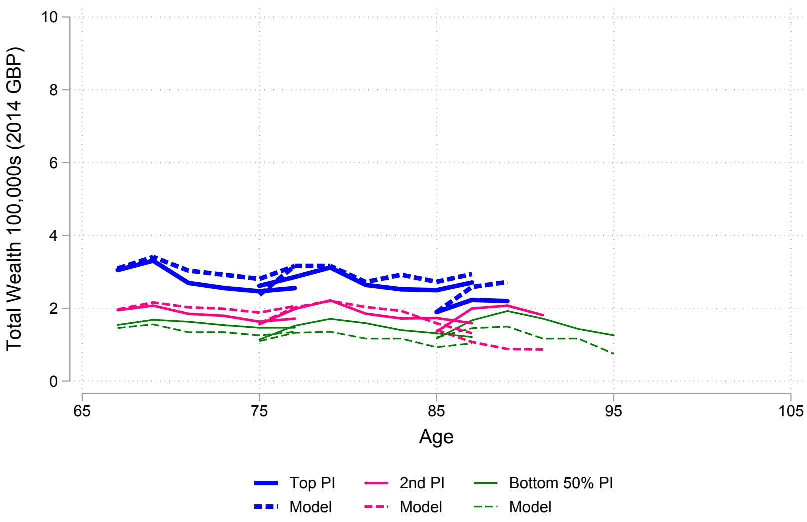
A.11.1 The Income Tax Function

$$\tilde{y} = \bar{y} + \lambda_y y^{1-\tau_y}$$

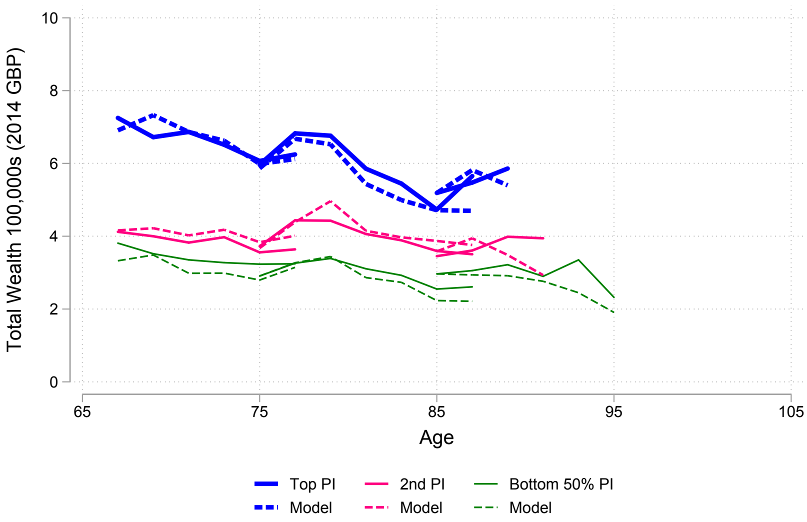
I combine data from TAXBEN, a microsimulation model of the UK tax and benefit system (for further details see Waters, 2017), with individual household data for my ELSA sample in order to estimate the tax function. The tax function includes both taxes and benefits including those that are means tested, for example pension credit, and provided to all older household, such as winter fuel allowance (which motivates the inclusion of an additional constant in after tax income).

A.11.2 Model Fit for Untargeted Moments

I present the model fit for two additional sets of moments that were not included in the estimator. This form of out validation is common in exercises that estimate structural models.



(a) 25th Percentile



(b) 75th Percentile

Figure A.9: Model Fit - Total Wealth Profiles (Initial Owners)

These results are presented in Figure A.9. I choose the 25th percentile and 75th percentile of total wealth holdings amongst initial owners to demonstrate that the model matches within PI heterogeneity.[15] The cross-sectional distribution of households is important ingredient in determining the quantitative implications of the model. The model is able to capture differences across the conditional wealth distribution and captures higher order moments of the wealth distribution well despite not being constrained to match these moments.

A.11.3 Model Fit for Alternative Birth Cohorts

Figures A.10 - A.12 display the corresponding data and simulated moments for the alternative set of birth cohorts not included in 2.8 - 2.10.

The model produces a similar level of fit to the wealth holdings as for the cohorts presented in the main text and captures the same features of the data. The model is unable to correctly match the liquid wealth and housing wealth holdings of the second oldest cohort in their early 90s. However, these two data moments are imprecisely estimated due to volatility in the reported wealth of the richest households. Under alternative sample selection the model does not exhibit this divergence from the data.

As with the other birth cohorts the model generates a moving rate that is higher than the observed moving rate. For the youngest cohorts this moving rate is closer than other birth cohorts, but for the oldest cohort the model struggles to replicate the high levels at age 95 without generating large deviations from the data at earlier ages. Again, due to the smaller number of households sampled at these ages these moments are more imprecisely estimated in the data.

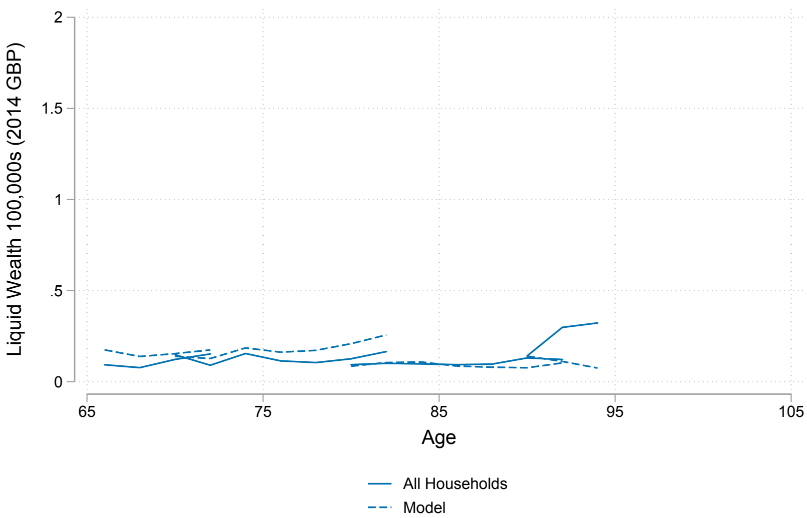
Finally, figure A.12 plots the data and simulated profiles for the subjective bequest probabilities. The results for these moments are very similar to the results presented in the main text.

[15] I have already that the model matches well the relative portfolio shares of households

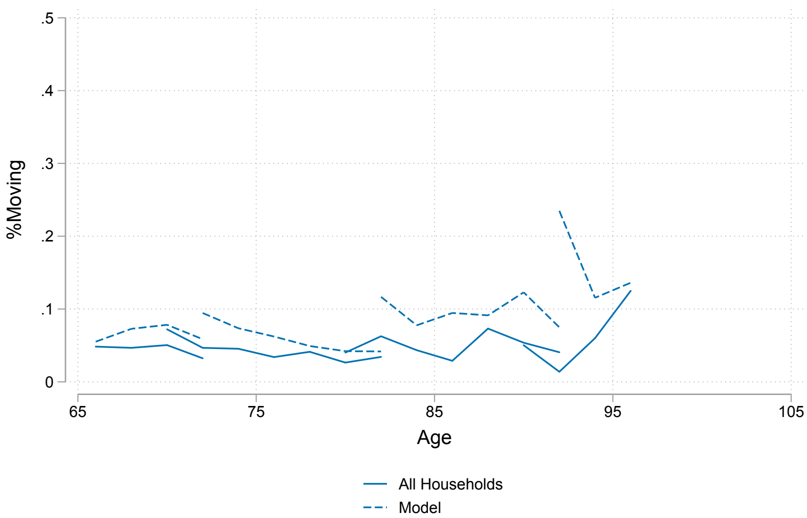
(a) Mean Housing Wealth

(b) Mean Liquid Wealth

Figure A.10: Model Fit - Wealth Profiles (Initial Owners)

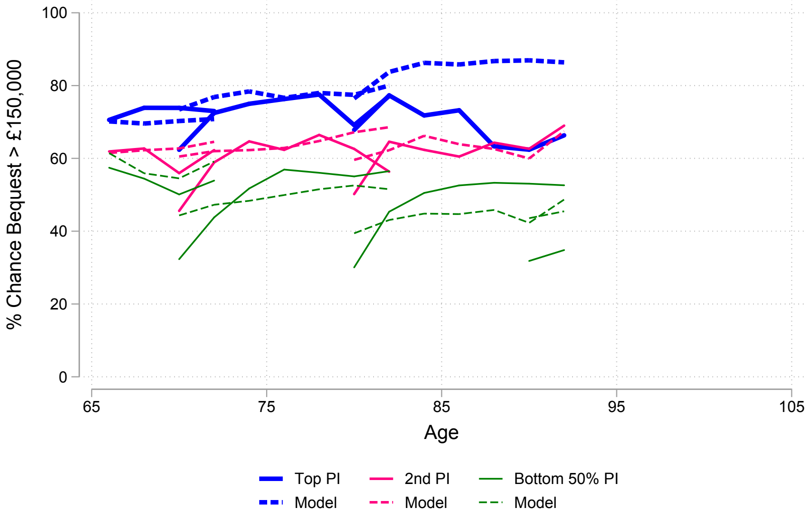


(a) Mean Liquid Wealth (Initial Renters)

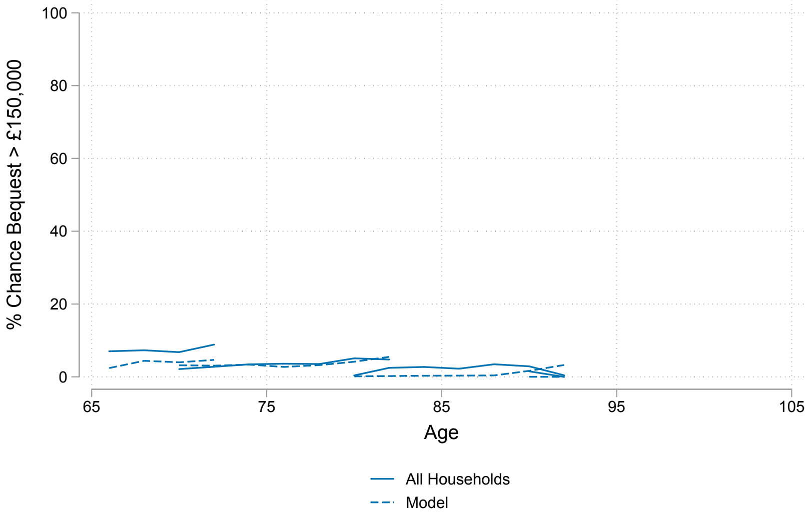


(b) Moves in the last 2 years

Figure A.11: Model Fit - Home Moving Rates and Renters



(a) Initial Owners



(b) Initial Renters

Figure A.12: Model Fit - Subjective Bequest Probabilities

	Band around cut-off				
Order of polynomial	10%	15%	20%	25%	30%
Data					
Linear	-0.0445**	-0.0475***	-0.0207*	-0.0200*	-0.0250**
	(0.0181)	(0.0160)	(0.0114)	(0.0116)	(0.0110)
Quadratic	-0.0365	-0.0450**	-0.0270**	-0.0218*	-0.0265**
	(0.0241)	(0.0193)	(0.0133)	(0.0130)	(0.0112)
N	1224	1559	3023	3233	3979
Model					
Linear	-0.0737	-0.0556	-0.0491	-0.0568	-0.0578
Quadratic	-0.0785	-0.0570	-0.0441	-0.0563	-0.0578

All regressions additionally control for wave fixed effects, a polynomial in age, household demographics, a polynomial in permanent income and region dummies. Following Kolesár and Rothe (2018), Standard Errors are clustered by household. * $p < 0.10$, ** $p < 0.5$, *** $p < 0.01$

Table A.5: The Effect of Transaction Taxes on Household Mobility: Model and Data

A.12 Decomposing Savings Motives

To determine the quantitative importance of different savings motives, I use the estimated parameters and change one feature of the model at a time. For each of these different environments, I compute the new household policy functions, simulate the model and compare the resulting asset accumulation profiles to the asset profiles generated by the baseline model. I display asset profiles for households who are age 68 in the first wave of ELSA. Throughout, I focus on the wealth of initial homeowners because the wealth of renters is negligible.

First, I fix house prices at their 2002 level as in the main text, but also change household expectations. In this counter-factual, households now know that there are no returns on housing and that houses bear no risk. Figure A.13 plots the simulate profiles of total wealth for initial owners. Compared to the wealth profiles in the baseline simulated economy, total household wealth decrease at all ages. The effect is largest at younger ages where the baseline profiles include the rapid house price appreciation of the early 2000s, but also has effects at older ages. Relative to the experiment holding prices constant in the main text, there are even larger cross

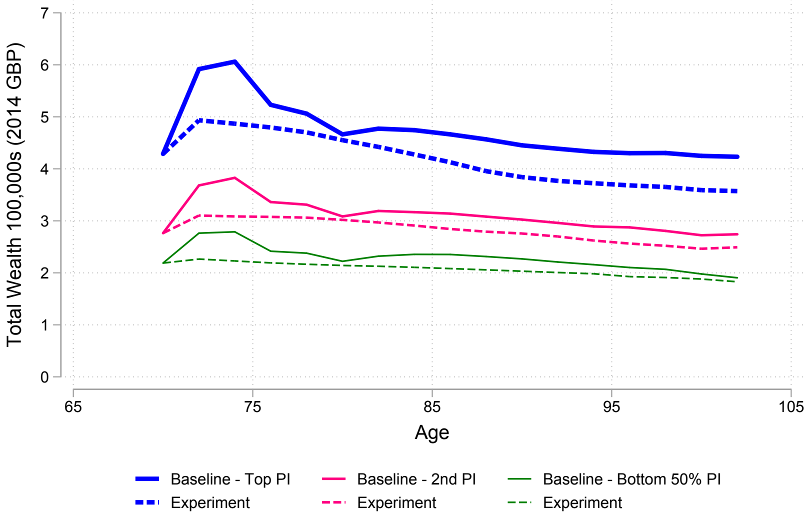
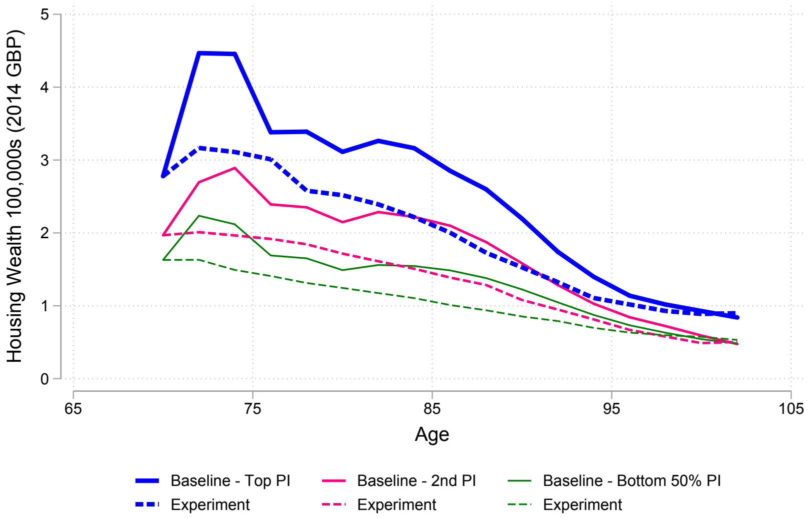


Figure A.13: Experiment A1- Total Wealth

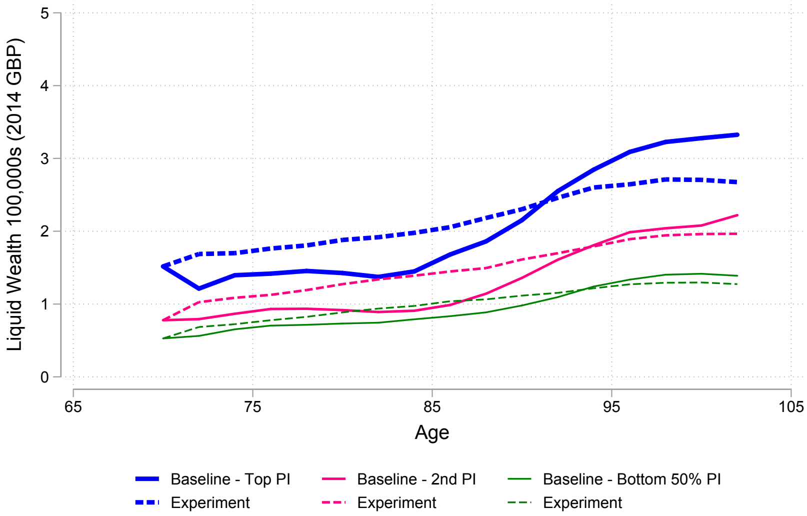
sectional differences with much of the effect on total wealth concentrated at the top of the permanent income distribution. Figure A.14 breaks the total wealth into its two components.

Housing wealth decreases when house prices are held constant. The top panel shows housing wealth. There is a large reduction in the housing wealth held by households due to the mechanical effect of eliminating house prices as well as the behavioural response as simulated households re optimize. The bottom panel shows the corresponding effect on their liquid wealth. For the top 50% of lifetime incomes, liquid wealth increases for much of their retirement which dampens the overall deccumulation of wealth. Households continue to maintain large levels of wealth as buffers against future shocks and substitute from consumption to saving in order to offset the wealth effect of decreased house prices. Averaged across the households' remaining life span liquid wealth balances increase due to the effect on high PI households at younger ages (17%) while housing wealth declines (15%). Relative to the baseline, when households don't experience periods of house price appreciation, they reduce the frequency with which they move home (by 18%). This transmits into lower consumption by the household and smaller bequests

In the next experiment I eliminate both bequest motives and their heterogeneity. As the estimated bequest motive for the Type 1 household is close to zero



(a) Mean Housing Wealth



(b) Mean Liquid Wealth

Figure A.14: Experiment A1- Portfolio

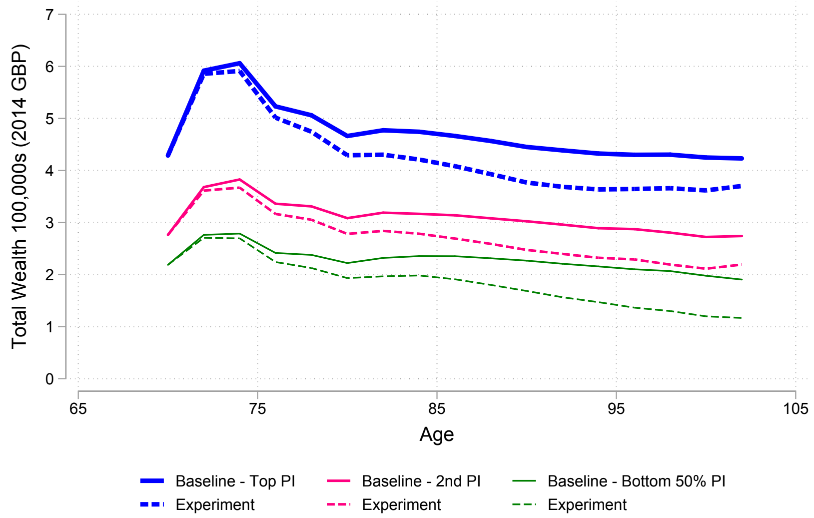


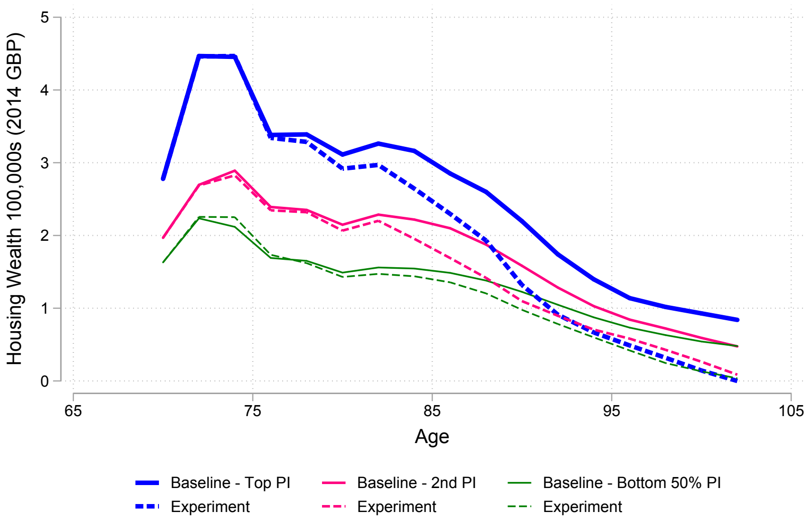
Figure A.15: Experiment A2- Total Wealth

across the support of Type 1 wealth this is approximately equivalent to giving every household they Type 1 bequest motive. Figure A.15, shows that the difference in total wealth between the experiment and the baseline is small (9%) , however, for survivors this grows with age. The effect on total wealth obscures the effect of bequest motives on household portfolios.

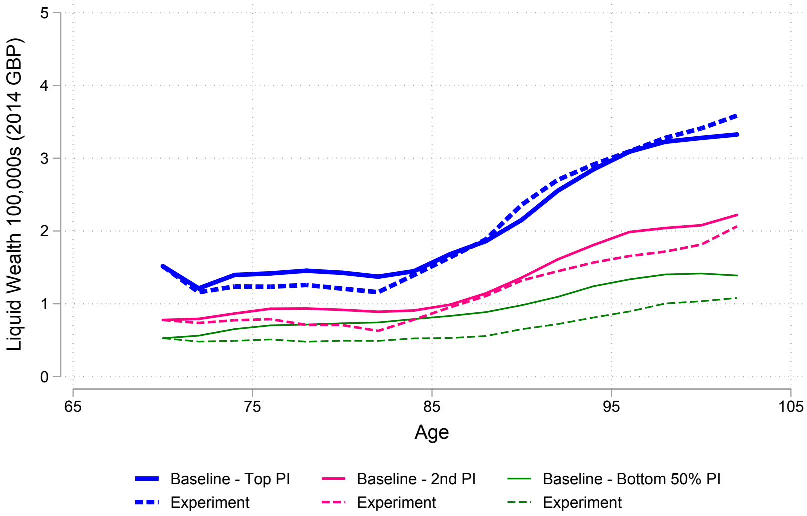
Eliminating bequest motives has a large effect on the composition of household assets and the extent of this effect varies with age. For older households, eliminating the bequest motive makes saving in a house less attractive. However, they still face substantial costs to adjusting their housing wealth and consuming more today. The size of these adjustment costs means that there is an effect on both the intensive and extensive margins. Those who move home in the baseline economy release more equity and especially those who move at older ages.

Next, I look at the interaction of housing and bequest motives. To understand the interaction of these saving motives, I return to the second experiment in the main text which eliminated the housing asset and additionally eliminate bequest motives. This is shown in figure A.17.

As with the previous experiment eliminating housing, the wealth trajectory in retirement is very different from the baseline economy. Eliminating bequest motives has a small effect on the profile of this wealth deccumulation during early



(a) Mean Housing Wealth



(b) Mean Liquid Wealth

Figure A.16: Experiment A2- Portfolio

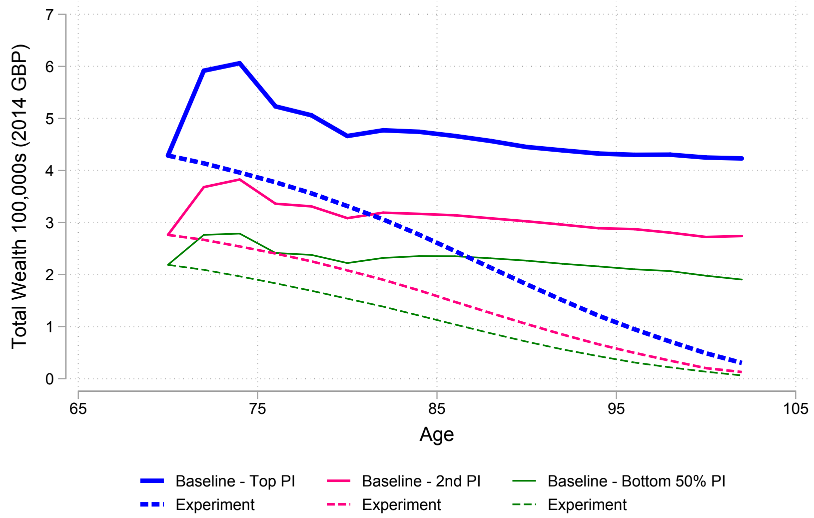


Figure A.17: Experiment A3- Total Wealth

retirement, but a large effect for the oldest households. Under this experiment, the wealth of those who survive to age 100 is approximately 75% lower for the richest households who have the largest bequest motives, while for lower PI groups the effect is closer to 60%. Because these differences grow with age the average reduction in total wealth during the whole retirement period is 20% lower when compared with the experiment where only housing is eliminated. In contrast, the average reduction in bequests between the two experiments is close to 50%.

Taken together, these results suggest that bequest motives interact with the portfolio choices of households. Housing has a larger effect on wealth in retirement, but bequest motives are still important and have a large effect on how households allocate their wealth across different asset classes as they age as well as the wealth holdings of households in the absence of this portfolio choice.

In addition to eliminating bequest motives I now show how large the effect of increasing bequest motives is by endowing each household with the strongest estimated bequest motives. This has a large effect on wealth holdings increasing average wealth for those who survive to age 96 by over £100,000 or over 40% of wealth in the baseline economy.

In this counter-factual, households retain higher levels of housing wealth. For older households, eliminating the bequest motive makes saving in a house more

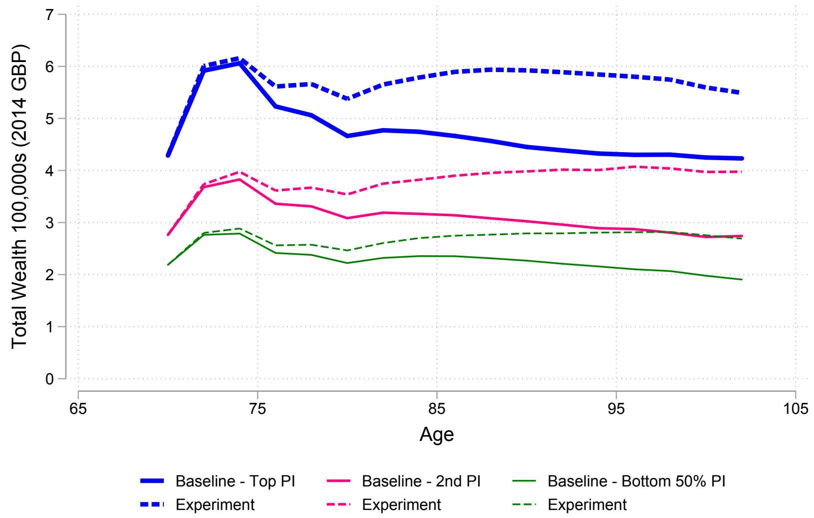


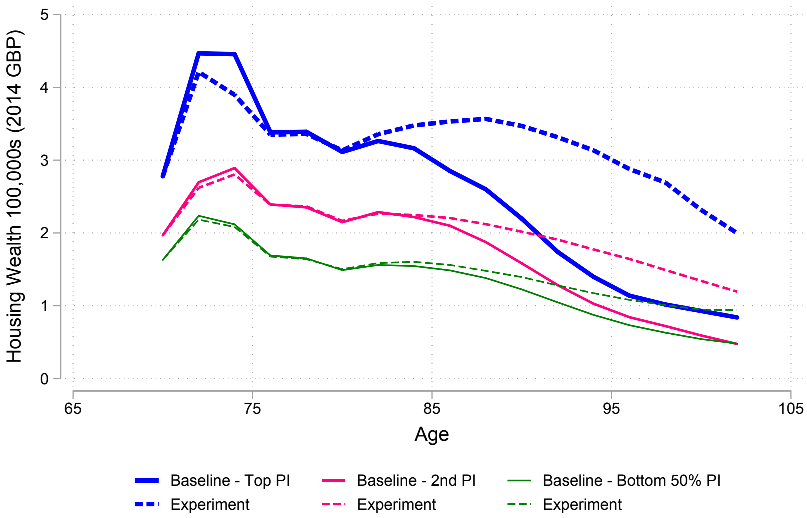
Figure A.18: Experiment A4- Total Wealth

attractive and paying the adjustment costs to capture trapped equity even less attractive. There is a decline on both the intensive and extensive margins of housing adjustment. However, increasing bequest motives also has a large effect on their liquid wealth savings even as they release less equity from their home. Relative to bequests lifetime consumption is less attractive and households increase their savings in all forms of wealth which they finance by decreasing their lifetime consumption.

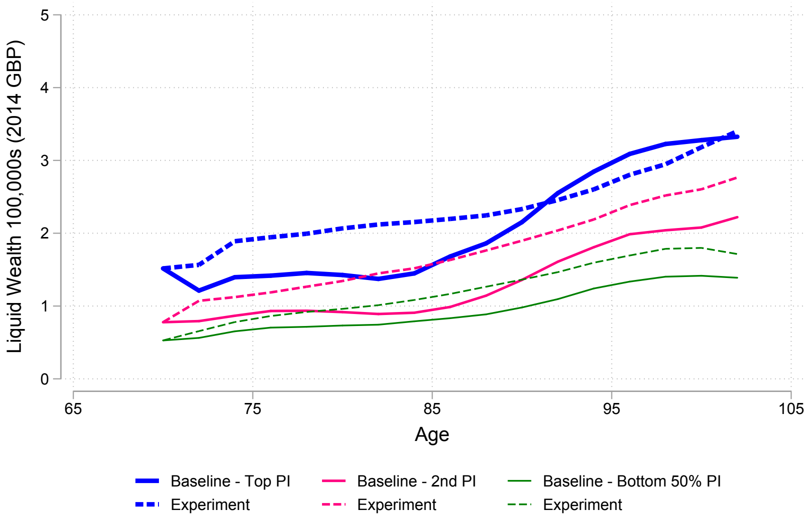
As in Lockwood (2018), bequest motives are important for households portfolio choice (in Lockwood - the decision over LTC insurance products), however, here the effect on total wealth is stronger. The role of heterogeneous bequest motives is also highlighted in this experiment- the effect of eliminating or increasing bequest motives varies with the level of household permanent income and the level of wealth. The intensity of bequest motives also varies along these dimensions and experiment 3 highlights that effect of bequests is not homogeneous.

Finally, I look at the interaction of housing and bequest motives. To understand the interaction of these saving motives, I return to the second experiment which eliminated the housing asset and additionally eliminate bequest motives. This is shown in figure A.20.

This single asset version of the model misses the cyclical role of house prices



(a) Mean Housing Wealth



(b) Mean Liquid Wealth

Figure A.19: Experiment A4- Portfolio

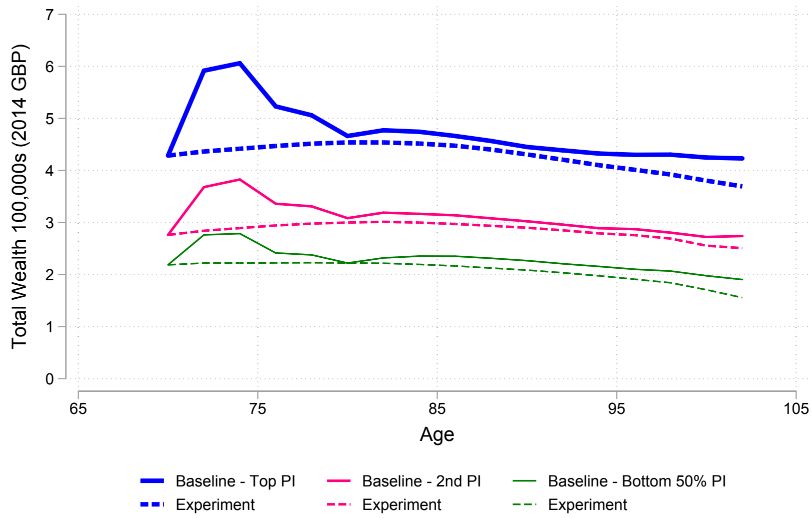


Figure A.20: Experiment A5- Total Wealth

	Ratio Relative to Benchmark Model							
	Total Wealth	At Age 96	Housing Wealth	Liquid Wealth	Cons	Bequests	Home Moves	Δ Home Value
Two Asset Model								
2002 Prices	0.893	0.859	0.880	0.926	0.962	0.899	0.776	0.879
2002 Prices & No Uncertainty	0.899	0.886	0.762	1.167	0.938	0.920	0.826	1.415
No Bequests	0.905	0.791	0.920	0.873	1.063	0.837	1.461	1.177
Max Bequests	1.160	1.406	1.081	1.315	0.860	1.298	0.623	0.693
One Asset Model								
Base	0.724	0.429	0.000	2.074	1.593	0.619	n/a	n/a
No Bequests	0.606	0.128	0.000	1.735	1.730	0.396	n/a	n/a
Max Bequests	0.922	0.946	0.000	2.639	1.357	1.004	n/a	n/a

Table A.6: Contribution of Alternative Model Mechanisms to Aggregate decisions

in total wealth, but is able to reproduce near identical levels of wealth. This shows that one of the advantages of accurately modelling housing is to eliminate one potentially large source of bias in estimating bequest motives. Without housing the model would require a substantially larger role for bequests or precautionary savings motives.

Tables A.6 summarizes these results and shows how key economic aggregates change for this cohort of retirees. These are displayed as shares relative to the baseline economy (the solid lines in each graph). Additional results are provided in

	Change Relative to Benchmark	
	Total Wealth	At Age 96
Two Asset Model		
2002 Prices	-28,175	-35,611
2002 Prices & No Uncertainty	-26,479	-28,944
No Bequests	-24,988	-53,046
Max Bequests	41,969	102,880
One Asset Model		
Base	-72,314	-144,657
No Bequests	-103,395	-220,857
Max Bequests	-20,538	-13,807

Table A.7: Contribution of Alternative Model Mechanisms to Aggregate decisions

Table A.7 which summarizes the average deviation in total wealth across the whole retirement period and for those who survive to age 96. This is provided in levels.

A.13 Additional responses to unanticipated shocks

In this appendix, I show additional results for the household responses to unanticipated shocks discussed in Section 2.7.3 of the main text.

I begin by expanding on Table 2.11 by also including the results for House Prices in Table A.8.

Focussing on the variation in MPCs and MPBs to the house price shock across preference type reveals a very different picture. The obvious correlation in the case of the income shock has all but disappeared. The differences in the MPB out of a house price shock by preference type are smaller than in the case of the income shock because of the importance of the interaction between housing and bequests - bequests dampen household incentives to adjust portfolios. However, the sign of the MPB with respect to the strength of the Bequest motive is less clear.

Shock		Marginal Propensity to Consume	Marginal Propensity to Bequeath
Income	Type 1	0.295	0.526
	Type 2	0.294	0.526
	Type 3	0.274	0.566
	Type 4	0.076	0.785
	Type 5	0.043	0.838
House Price	Type 1	0.023	0.413
	Type 2	0.023	0.413
	Type 3	0.023	0.425
	Type 4	0.028	0.359
	Type 5	0.025	0.339

Simulated responses for a single birth cohort to a one-time 10% increase in income and a one-time 10% increase in house prices. In both simulations the shock arrives at age 70 and the MPC is measured contemporaneously. Preference parameters for each type are taken from the estimation results above. In order to separate the role of preferences the correlation between preference type and initial conditions is set to 0.

Table A.8: Household Responses to Unanticipated Shocks by Preference Type

Household decision rules feature regions of inaction[16] and when households have operative bequest motives this shrinks the regions for which they find it optimal to move. When house prices appreciate and the returns to downsizing increase for households, those who are less likely to move in the baseline (these Type 4 and 5 households) may have a larger extensive margin effect due to the combination of shock size and preference parameters. This generates the non-monotonic response by preference types because more households cross the boundary of their inaction region on the housing adjustment margin. Extensive inspection reveals that this non-monotonicity is not a general feature of the model and only occurs for particular parameter values and thus it is not emphasized in the main text as the intuition is similar to the difference in the aggregate response to house prices and income shocks. However, this is instructive in that it highlights the potential for asymmetric responses to these shocks that arise from non-linearities in household decision rules.

[16](S,s) policy rules are common in applications with durable goods, such as housing

Shock		Marginal Propensity to Consume	Marginal Propensity to Bequeath
Income	Type 1	0.177	0.758
	Type 2	0.468	0.673
	Type 3	0.195	0.363
	Type 4	0.069	0.721
	Type 5	0.074	0.972
House Price	Type 1	0.061	0.307
	Type 2	0.028	0.392
	Type 3	0.0125	0.469
	Type 4	0.012	0.352
	Type 5	0.025	0.396

Simulated responses for a single birth cohort to a one-time 10% increase in income and a one-time 10% increase in house prices. In both simulations the shock arrives at age 70 and the MPC is measured contemporaneously. Preference parameters for each type are taken from the estimation results above.

Table A.9: Household Responses to Unanticipated Shocks by Preference Type and Heterogeneity in Initial Conditions

To elaborate on the interaction between these shocks and individual heterogeneity, I reintroduce the correlation between preference type and initial conditions. These results are documented in Table A.9.

These results highlight that households with the same preferences such as Types 1 and 2 (or close to in the case of Types 3 and 4) behave very differently when differences in their distribution of initial state variables are accounted for. Indeed, the non-monotonicity is now also present in the household response to income shocks. The model estimated in this paper has rich household level heterogeneity which generates policy rules that are non-linear in multiple dimensions. As highlighted by Kaplan and Violante (2014) and Attanasio (2000) the combination of non-linear decision rules and idiosyncratic shocks can generate varied properties for aggregate dynamics.

Appendix B

Appendix to Chapter 3

B.1 Sample selection and data handling

To keep the dynamic programming problem manageable, we assume a fixed difference in age between spouses, and we take the average age difference from our data. In our sample, husbands are on average 3 years older than their wives. To keep the data consistent with this assumption, we drop all households where the wife is more than 4 years older or 10 years younger than her husband.

We begin with 6047 households. After dropping 401 households who get married, divorced, were same sex couples, or who report making other transitions not consistent with the model, 753 households who report earning at least $3,000 in any period, 171 households with a large difference in the age of husband and wife, and 87 households with no information on the spouse in a household, we are left with 4,634 households, of whom 1,388 are couples and 3,246 are singles. This represents 24,274 household-year observations where at least one household member was alive.

A key advantage of the AHEAD relative to other datasets is that it provides panel data on health status, including nursing home stays. We assign individuals a health status of "good" if self-reported health is excellent, very good or good, and are assigned a health status of "bad" if self-reported health is fair or poor. We assign individuals to the nursing home state if they were in a nursing home at least 120 days since the last interview (or on average 60 days per year) or if they spent at least 60 days in a nursing home before the next scheduled interview and died before that scheduled interview.

The AHEAD has information on the value of housing and real estate, autos, liquid assets (which include money market accounts, savings accounts, T-bills, etc.), IRAs, Keoghs, stocks, the value of a farm or business, mutual funds, bonds, “other” assets and investment trusts less mortgages and other debts.

We do not include pension and Social Security wealth for four reasons. First, we wish to to maintain comparability with other studies (Hurd, 1989a; Attanasio and Hoynes, 2000, for example). Second, because it is illegal to borrow against Social security wealth pension and difficult to borrow against most forms of pension wealth, Social Security and pension wealth are much more illiquid than other assets. Third, their tax treatment is different from other assets. Finally, differences in Social Security and pension are captured in our model by differences in the permanent income measure we use to predict annual income.

One important problem with our asset data is that the wealthy tend to underreport their wealth in virtually all household surveys (Davies and Shorrocks, 2000). This will lead us to understate asset levels at all ages. However, Juster et al. (1999) show that the wealth distribution of the AHEAD matches up well with aggregate values for all but the richest 1 % of households. Given that we match medians (conditional on permanent income), underreporting at the very top of the wealth distribution should not seriously affect our results.

B.2 Inferring Permanent Income

We assume that log income follows the process

$$\ln y_{it} = \kappa_1(t, f_{it}) + h(I_i) + \omega_{it} \tag{B.1}$$

where $\kappa_1(t, f_{it})$ is a flexible functional form of age t and family structure f_{it} (i.e., couple, single men, or single woman) and ω_{it} represents measurement error. The variable I_i is the household's percentile rank in the permanent income distribution. Since it is a summary measure of lifetime income at retirement, it should not change during retirement and is thus a fixed effect over our sample period. However, income could change as households age and potentially lose a family member. Our procedure to estimate equation (B.1) is to first estimate the fixed effects model

$$\ln y_{it} = \kappa_1(t, f_{it}) + \alpha_i + \omega_{it} \tag{B.2}$$

which allows us to obtain a consistent estimate of the function $\kappa_1(t, f_{it})$. Next, note that as the number of time periods over which we can measure income and other variables for individual i (denoted T_i) becomes large,

$$plim_{T_i \to \infty} \frac{1}{T_i} \sum_{t=1}^{T_i} [\ln y_{it} - \kappa_1(t, f_{it}) - \omega_{it}] = \frac{1}{T_i} \sum_{t=1}^{T_i} [\ln y_{it} - \kappa_1(t, f_{it})] = h(I_i). \tag{B.3}$$

Thus we calculate the percentile ranking of permanent income I_i for every household in our sample by taking the percentile ranking of $\frac{1}{T_i} \sum_{t=1}^{T_i} [\ln y_{it} - \widehat{\kappa}_1(t, f_{it})]$, where $\widehat{\kappa}_1(t, f_{it})$ is the estimated value of $\kappa_1(t, f_{it})$ from equation (B.2). Put differently, we take the mean residual per person from the fixed effects regression (where the residual includes the estimated fixed effect), then take the percentile rank of the mean residual per person to construct I_i. This gives us a measure of the percentile ranking of permanent income I_i. However, we also need to estimate the function $h(I_i)$, which gives a mapping from the estimated index I_i back to a predicted level of income that can be used in the dynamic programming model. To do this we estimate the function

$$[\ln y_{it} - \widehat{\kappa_1}(t, f_{it})] = h(I_i) + \omega_{it} \tag{B.4}$$

where the function $h(I_i)$ is a flexible functional form.

In practice we model $\kappa_1(t, f_{it})$ as a third order polynomial in age, dummies for family structure, and family structure interacted with an age trend. We model $h(I_i)$ as a fifth order polynomial in our measure of permanent income percentile.

Given that we have for every member of our sample t, f_{it}, and estimates of I_i and the functions $\kappa_1(.,.), h(.)$, we can calculate the predicted value of $\ln \widehat{y}_{it} = \widehat{\kappa_1}(t, f_{it}) + \widehat{h}(\widehat{I_i})$. It is $\ln \widehat{y}_{it}$ that we use when simulating the model for each household. A regression of $\ln y_{it}$ on $\ln \widehat{y}_{it}$ yields a R^2 statistic of .74, suggesting that our predictions are accurate.

B.3 Imputing Medicaid plus Out of Pocket Medical Expenses

Our goal is to measure the data generating process of the sum of Medicaid payments plus out of pocket expenses: this is the variable $\ln(m_{it})$ in equation (3.6) of the main text. If the household is drawing Medicaid benefits, then the household will spend less than $\ln(m_{it})$ on out of pocket medical spending (Medicaid picking up the remainder).

The AHEAD data contains information on out of pocket medical spending, but not on Medicaid payments. Fortunately, the Medicare Current Beneficiary Survey (MCBS) has extremely high quality information on Medicaid payments plus out of pocket medical spending. One drawback of the MCBS, however, is that although it has information on marital status and household income, it does not have information on the medical spending or health of the spouse. Here we explain how to exploit the best of both datasets.

We use a three step estimation procedure. First, we use the MCBS to infer Medicaid payments for recipients, conditional on the observable variables that exist in both datasets. Second, we impute Medicaid payments in the AHEAD data using conditional mean matching procedure (which is a procedure very similar to hotdecking). Third, we estimate the data generating process for out of pocket plus Medicaid medical spending in the HRS.

First Step Estimation Procedure:

We use the MCBS to infer Medicaid payments for recipients, conditional on the observable variables that exist in both the MCBS and the HRS datasets. Define oop_{it} as the out of pocket medical expenses and m_{it} is the sum of out of pocket plus Medicaid payments that we wish to plug in the model, and so Med_{it} is the dollar value of Medicaid payments.

To impute Med_{it} in (which is missing in the HRS), we follow David et al. (1986) and French and Jones (2011) and use the following predictive mean matching regression approach. First, for every member of the the MCBS sample with a positive Medicaid indicator (i.e., a Medicaid recipient), we regress the variable of interest Med on the vector of observable variables z, yielding $Med = z\beta + \varepsilon$. Second, for each individual j in the MCBS we calculate the predicted value $\hat{Med}_{jt} = z_{jt}\hat{\beta}$, and for each member of the sample we calculate the residual $\hat{\varepsilon}_{jt} = Med_{jt} - \hat{Med}_{jt}$. Third, we sort the predicted value $\hat{Med}_{jt}$ into deciles and keep track of all values of $\hat{\varepsilon}_{jt}$ within each decile.

In practice we include in z_{jt} nursing home status, the number of nights spent in a nursing home, an age polynomial, total household income, marital status, self-reported health, race, visiting a medical practitioner (doctor, hospital or dentist), out of pocket medical spending, education and death of an individual. Because the measure of medical spending in the HRS is medical spending over two years, we take two year averages of the MCBS data to be consistent with the structure of the HRS. A regression of Med_{jt} on z_{jt} yields a R^2 statistic of .67, suggesting that our predictions are accurate.

Second Step Estimation Procedure:

Next, for every individual i in the HRS sample with a positive Medicaid indicator, we impute $\hat{Med}_{it} = z_{it}\hat{\beta}$ using the values of $\hat{\beta}$ estimated using the MCBS. Then we impute ε_{it} for each member of the HRS sample by finding a random individual j in the MCBS with a value of $\hat{Med}_{jt}$ in the same decile as $\hat{Med}_{it}$ in the HRS, and set $\varepsilon_{it} = \hat{\varepsilon}_{jt}$. The imputed value of Med_{it} is $\hat{Med}_{it} + \varepsilon_{it}$.

As David et al. (1986) point out, our imputation approach is equivalent to hot-decking when the "z" variables are discretized and include a full set of interactions. The advantages of our approach over hot-decking are two-fold. First, many of the "z" variables are continuous, and it seems unwise to discretize them. Second, we use a large number of observable variables "z". We find that adding extra variables are very important for improving goodness of fit when imputing medicaid payments. Even a small number of variables generate a large number of hot-decking cells, as hot-decking uses a full set of interactions. Thus, in this context, hot decking is too data intensive.

We predict Medicaid expenditures for 3,756 Medicaid eligible households (i.e., Medicaid beneficiaries) in the HRS and report the results of this imputation exercise in Table B.1. Households where the last surviving member died between the previous and current waves of the sample have the largest imputed Medicaid payments. In contrast, couples have the smallest Medicaid payments per individual, but Medicaid expenditures at the household level are larger than for either single men or women. The imputed Medicaid payments for both spouses in a couple are approximately equal and the results for couples are not driven by the expenditures of only the husband or wife. In Table B.1 new widows and widowers who's spouse has died between waves of the HRS sample are included in the rows for single men and single women- on average the Medicaid payments of dead spouses are less than 20% of the single household's Medicaid expenditures.

Family Structure	Number of Medicaid Eligible Households	Mean	Standard Deviation
Dead	1,040	21,800	31,500
Couples	287	15,300	25,700
Single Men	351	11,600	21,500
Single Women	2,078	12,300	21,800

Table B.1: Imputed Medicaid payments for Medicaid beneficiaries in HRS.

The distribution of imputed Medicaid expenditures in the HRS is close to the distribution observed in the MCBS. Mean imputed Medicaid expenditures, Med_{it}, for Medicaid recipients in the HRS is \$14,050. This is lower than the corresponding value of \$16,000 in the MCBS. This small difference is due to the distribution of observable variables, z, in both the MCBS and HRS. In particular, the number of nights in a nursing home is 20% lower in the HRS than in the MCBS, due in part to the fact that the initial HRS sample is for the non-institutionalized population who were not in a nursing home. **Third Step Estimation Procedure:**

Our value of m_{it} in the HRS is thus $oop_{it} + Med_{it}$, where Med_{it} was calculated in the second step. Define $X_{it} = (hs_t^h, hs_t^w, I, g, t, f_t, f_{t-1})$. Recall that equation (3.6) of the main text is given by $\ln(m_{it}) = m(X_{it}) + \sigma(X_{it}) \times \psi_t$.

Using HRS, we use fixed effects estimation procedures to estimate

$$\ln(m_{it}) = \alpha_i + m(X_{it}) + \varepsilon_{it}. \tag{B.5}$$

which gives $\hat{m}(X_{it})$. Next we construct the estimated residuals $(\widehat{\varepsilon_{it} + \alpha_i}) = \ln(m_{it}) - \hat{m}(X_{it})$ and estimate the regression (using OLS and no fixed effects):

$$(\widehat{\varepsilon_{it} + \alpha_i})^2 = h(X_{it}) + \zeta_{it}. \tag{B.6}$$

This allows us to recover the variance of ε_{it} conditional on X_{it}.

In order to infer the variance of medical expenditures conditional on X_{it}, note that from equations (3.6) and (B.5)

$$\sigma(X_{it}) \times \psi_{it} = \ln(m_{it}) - m(X_{it}) = (\varepsilon_{it} + \alpha_i). \tag{B.7}$$

so we can obtain the conditional variance by noting that $E[\psi_{it}^2] = 1$ and so

$$[\sigma(X_{it})]^2 = E[(\varepsilon_{it} + \alpha_i)^2 | X_{it}] \tag{B.8}$$

where we estimate $E[(\varepsilon_{it} + \alpha_i)^2 | X_{it}] = h(X_{it})$ in equation (B.6).

B.4 Other savings graphs

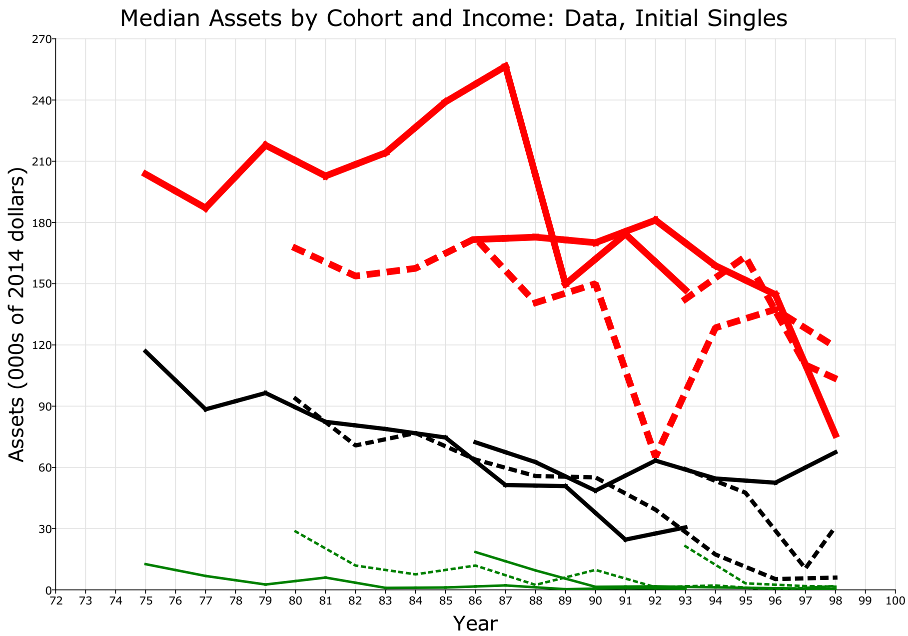


Figure B.1: Median assets for households initially single. Each line represents median assets for a cohort-income cell, traced over the time period 1996-2014. Thicker lines refer to higher permanent income groups. Solid lines: cohorts ages 72-76 and 83-88 in 1996. Dashed lines: ages 77-82 and 89-96 in 1996.

Figure B.1 displays the asset profiles for households that were single when first observed in the AHEAD, classified according. Median assets are increasing in permanent income, with the 74-year-olds in the highest PI income of the singles holding about $210,000 in median assets, while those at the lowest PI quintiles holding essentially no assets. Over time, those with the highest PI tend to hold onto significant wealth well into their nineties, those with the lower PIs never save much, while those in the middle PIs display quite some asset decumulation as they age.

B.5 Outline of the computation of the value functions and optimal decision rules

We compute the value functions by backward induction. We start from the singles, find their time T value function and decision rules by maximizing equation (3.17) subject to the relevant constraints, and $V^g_{T+1} = \theta_0(x_t - c_t)$, $g = h, w$. This yields the value function V^g_T and the decision rules for time T. We then find the decision rules at time $T-1$ by solving equation (3.17) with V^g_T. Continuing this backward induction yields decision rules for time $T-2, T-3, ..., 1$.

We find the decisions for couples by maximizing equation (3.19), subject to the relevant constraints and the value function for the singles, and setting V^c_{T+1} to the appropriate bequest motive value. This yields the value function V^c_T and the decision rules for time T. We then find the decision rules at time $T-1$ by solving equation (3.19) using $V^c_T, V^g_T, g = h, w$. Continuing this backward induction yields decision rules for time $T-2, T-3, ..., 1$.

We discretize the persistent component and the transitory components of the health shock, and interest rate into Markov Chain following Tauchen and Hussey (1991). We assume a finite number of permanent income categories. We take cash-on-hand to lay into a finite number of grid points.

Given each level of permanent income of the household, we solve for decision rules for each possible combination of cash-on-hand, income, health status, and persistent component of the health shock. We use linear interpolation within the grid and linear extrapolation outside of the grid to evaluate the value function at points that we do not directly compute.

For the singles, for simplicity of notation, here for the most part we drop any reference to the missing spousal variables everywhere, and just use similar notation to the one for couples (except for the number of arguments in the function). In the code, we have different names, say for example for the survival of single people and couples.

B.6 Moment Conditions and the Asymptotic Distribution of Parameter Estimates

Our estimate, $\hat{\Delta}$, of the "true" preference vector Δ_0 is the value of Δ that minimizes the (weighted) distance between the estimated life cycle profiles for assets found in the data and the simulated profiles generated by the model. For each calendar year $t \in \{1,...,T\}$, we match median and the 75th percentile of assets for 3 permanent income terciles in 5 birth year cohorts, both for singles and couples, leading to a total of 60T moment conditions. Sorting households into terciles also requires us to estimate the 1/3rd and 2/3rd permanent income quantiles for each birth year cohort and initial family type. This produces a total of 20 nuisance parameters, which we collect into the vector γ. Each of these parameters has its own moment condition.

The way in which we construct these moment conditions builds on the approach described in French and Jones (2007). Useful references include Buchinsky (1998) and Powell (1994). Consider first the permanent income quantiles. Let $q \in \{1,2,...,Q-1\}$ index the quantiles. Assuming that the permanent income distribution is continuous, the π_q-th quantile of permanent income for initial family type f of cohort c, $g_{\pi_q}(c,f)$, is defined as

$$\Pr\left(I_i \leq g_{\pi_q}(c,f)|c,f\right) = \pi_q. \tag{B.9}$$

In other words, the fraction of households with less than g_{π_j} in permanent income is π_j. Using the indicator function, the definition of π_j-th conditional quantile can be rewritten as

$$E\left(\left[1\{I_i \leq g_{\pi_q}(c,f)\} - \pi_j\right] \times 1\{c_i = c\} \times 1\{f_i = f\}\right) = 0, \tag{B.10}$$

for $c \in \{1,2,...,C\}$, $f \in \{\text{single},\text{couple}\}$, $q \in \{1,2,...,Q-1\}$.

The more important set of moment conditions involves the permanent income-conditional age-asset profiles. Suppose that household i's permanent income level falls in the qth permanent income interval of households in its cohort, i.e.,

$$g_{\pi_{q-1}}(c,f) \leq I_i \leq g_{\pi_q}(c,f). \tag{B.11}$$

We assume that $\pi_0 = 0$ and $\pi_Q = 1$, so that $g_{\pi_0}(c,f) \equiv -\infty$ and $g_{\pi_Q}(c,f) \equiv \infty$. Let $a_{cfqt}(\Delta,\chi)$ be the model-predicted median observed asset level for group $cfqt$. Recall that the median is just the $1/2$ quantile.[1] Assuming that observed assets have a continuous conditional density, we arrive at the following moment condition:

$$E\Big(1\{\tilde{a}_{it} \leq a_{cfqt}(\Delta_0,\chi_0)\} - 1/2 \,|\, c,f,q,t, \text{household observed at } t\Big) = 0. \tag{B.12}$$

Equation (B.12) is adjusted to allow for "missing" as well as deceased households, as in French and Jones (2004). Using indicator function notation, we can convert this conditional moment equation into an unconditional one:

$$\begin{aligned} E\Big(&[1\{\tilde{a}_{it} \leq a_{cfqt}(\Delta_0,\chi_0)\} - 1/2] \times 1\{c_i = c\} \times 1\{f_i = f\} \\ &\times 1\{g_{\pi_{q-1}}(c,f) \leq I_i \leq g_{\pi_q}(c,f)\} \times 1\{\text{household observed at } t\} \mid t\Big) = 0, \end{aligned} \tag{B.13}$$

for $c \in \{1,2,...,C\}$, $f \in \{\text{single}, \text{couple}\}$, $q \in \{1,2,...,Q\}$, $t \in \{t_1, t_2 ..., t_T\}$.

Suppose we have a data set of I independent households that are each observed at T separate calendar years. Let $\varphi(\Delta,\gamma;\chi_0)$ denote the $(60T+20)$-element vector of moment conditions described immediately above, and let $\hat{\varphi}_I(.)$ denote its sample analog. Letting $\widehat{\mathbf{W}}_I$ denote a $(60T+20)\times(60T+20)$ weighting matrix, the MSM estimator $(\hat{\Delta}', \hat{\gamma}')'$ is given by

$$\underset{\{\Delta,\gamma\}}{\arg\min}\ \frac{I}{1+\tau}\ \hat{\varphi}_I(\Delta,\gamma;\chi_0)'\widehat{\mathbf{W}}_I\hat{\varphi}_I(\Delta,\gamma;\chi_0),$$

where τ is the ratio of the number of observations to the number of simulated observations.

[1] We follow the same approach when calculating the 75th percentile of assets. To eliminate repetition we only describe the procedure for the median.

In practice, we estimate χ_0 as well, using the approach described in the main text. Computational concerns, however, compel us to treat χ_0 as known in the analysis that follows. Under regularity conditions stated in Pakes and Pollard (1989) and Duffie and Singleton (1993), the MSM estimator $\hat{\theta}$ is both consistent and asymptotically normally distributed:

$$\sqrt{I}\left(\begin{pmatrix}\hat{\Delta}\\ \hat{\gamma}\end{pmatrix}-\begin{pmatrix}\Delta_0\\ \gamma_0\end{pmatrix}\right)\rightsquigarrow N(0,\mathbf{V}),$$

with the variance-covariance matrix $\mathbf{V}$ given by

$$\mathbf{V}=(1+\tau)(\mathbf{D}'\mathbf{W}\mathbf{D})^{-1}\mathbf{D}'\mathbf{W}\mathbf{S}\mathbf{W}\mathbf{D}(\mathbf{D}'\mathbf{W}\mathbf{D})^{-1},$$

where: $\mathbf{S}$ is the variance-covariance matrix of the data;

$$\mathbf{D}=\left[\left.\frac{\partial\varphi(\Delta,\gamma;\chi_0)}{\partial\Delta'}\right|_{\Delta=\Delta_0}\quad\left.\frac{\partial\varphi(\Delta,\gamma;\chi_0)}{\partial\gamma'}\right|_{\gamma=\gamma_0}\right] \tag{B.14}$$

is the $(60T+20)\times(7+20)$ gradient matrix of the population moment vector; and $\mathbf{W}=\text{plim}_{\to\infty}\{\widehat{\mathbf{W}}_I\}$. Moreover, Newey (1985) shows that if the model is properly specified,

$$\frac{I}{1+\tau}\hat{\varphi}_I(\hat{\Delta},\hat{\gamma};\chi_0)'\mathbf{R}^{-1}\hat{\varphi}_I(\hat{\Delta},\hat{\gamma};\chi_0)\rightsquigarrow\chi^2_{60T-7},$$

where $\mathbf{R}^{-1}$ is the generalized inverse of

$$\begin{aligned}\mathbf{R}&=\mathbf{PSP},\\ \mathbf{P}&=\mathbf{I}-\mathbf{D}(\mathbf{D}'\mathbf{W}\mathbf{D})^{-1}\mathbf{D}'\mathbf{W}.\end{aligned}$$

The asymptotically efficient weighting matrix arises when $\widehat{\mathbf{W}}_I$ converges to $\mathbf{S}^{-1}$, the inverse of the variance-covariance matrix of the data. When $\mathbf{W}=\mathbf{S}^{-1}$, $\mathbf{V}$ simplifies to $(1+\tau)(\mathbf{D}'\mathbf{S}^{-1}\mathbf{D})^{-1}$, and $\mathbf{R}$ is replaced with $\mathbf{S}$. But even though the optimal weighting matrix is asymptotically efficient, it can be severely biased in small samples. (See, for example, Altonji and Segal (1996).) We thus use a "diagonal" weighting matrix, as suggested by Pischke (1995). The diagonal weighting scheme uses the inverse of the matrix that is the same as **S** along the diagonal and has zeros off the diagonal of the matrix.

We estimate **D**, **S** and **W** with their sample analogs. For example, our estimate of **S** is the $(60T+20)\times(60T+20)$ estimated variance-covariance matrix of the sample data. When estimating preferences, we use sample statistics, so that $a_{cfqt}(\Delta,\chi)$ is replaced with the sample median for group $cfqt$. When computing the chi-square statistic and the standard errors, we use model predictions, so that the sample medians for group $cfqt$ is replaced with its simulated counterpart, $a_{cfqt}(\hat{\Delta},\hat{\chi})$.

One complication in estimating the gradient matrix **D** is that the functions inside the moment condition $\varphi(\Delta,\gamma;\chi)$ are non-differentiable at certain data points; see equations (B.10) and (B.13). This means that we cannot consistently estimate **D** as the numerical derivative of $\hat{\varphi}_I(.)$. Our asymptotic results therefore do not follow from the standard GMM approach, but rather the approach for non-smooth functions described in Pakes and Pollard (1989), Newey and McFadden (1994) (section 7) and Powell (1994).

In finding **D**, it proves useful to partition $\varphi(.)$ into the $60T$-element vector $\varphi_\Delta(.)$, corresponding to the moment conditions described by equation (B.13) and the $20-element$ vector $\varphi_\gamma(.)$, corresponding to the moment conditions described by equation (B.10). Using this notation, we can rewrite equation (B.14) as

$$\mathbf{D}=\begin{bmatrix}\frac{\partial\varphi_\Delta(\Delta_0,\gamma_0;\chi_0)}{\partial\Delta'} & \frac{\partial\varphi_\Delta(\Delta_0,\gamma_0;\chi_0)}{\partial\gamma'}\\ \frac{\partial\varphi_\gamma(\Delta_0,\gamma_0;\chi_0)}{\partial\Delta'} & \frac{\partial\varphi_\gamma(\Delta_0,\gamma_0;\chi_0)}{\partial\gamma'}\end{bmatrix}, \tag{B.15}$$

and proceed element-by-element.

It immediately follows from equation (B.10) that

$$\frac{\partial\varphi_\gamma(\Delta_0,\gamma_0;\chi_0)}{\partial\Delta'}=\mathbf{0}. \tag{B.16}$$

To find $\frac{\partial \varphi_\gamma(\Delta_0,\gamma_0;\chi_0)}{\partial \gamma'}$, we rewrite equation (B.10) as

$$\Pr(c_i = c \,\&\, f_i = f) \times \left[F\left(g_{\pi_q}(c,f)|c,f\right) - \pi_j\right] = 0, \tag{B.17}$$

where $F\left(g_{\pi_q}(c,f)|c,f\right)$ is the c.d.f. of permanent income for family type-f members of cohort c evaluated at the π_j-th quantile. Differentiating this equation shows that $\frac{\partial \varphi_\gamma(\Delta_0,\gamma_0;\chi_0)}{\partial \gamma'}$ is a diagonal matrix whose diagonal elements are given by

$$\Pr\left(c_i = c \,\&\, f_i = f\right) \times f\left(g_{\pi_q}(c,f)|c,f\right). \tag{B.18}$$

In practice we find $f\left(g_{\pi_q}(c,f)|c,f\right)$, the conditional p.d.f. of permanent income evaluated at the π_j-th quantile, with a kernel density estimator.

To find $\frac{\partial \varphi_\Delta(\Delta_0,\gamma_0;\chi_0)}{\partial \Delta'}$ and $\frac{\partial \varphi_\Delta(\Delta_0,\gamma_0;\chi_0)}{\partial \gamma'}$, it is helpful to rewrite equation (B.13) as

$$\Pr\left(c_i = c \,\&\, f_i = f \,\&\, \text{household observed at } t\right) \times \int_{g_{\pi_{q-1}}(c,f)}^{g_{\pi_q}(c,f)} \left[\int_{-\infty}^{a_{cfqt}(\Delta_0,\chi_0)} f(\tilde{a}_{it}|c,f,I_i,t)\, d\tilde{a}_{it} - \frac{1}{2}\right] f(I_i|c,f)\, dI_i = 0, \tag{B.19}$$

It follows that the rows of $\frac{\partial \varphi_\Delta(\Delta_0,\gamma_0;\chi_0)}{\partial \Delta'}$ are given by

$$\Pr\left(c_i = c \,\&\, f_i = f \,\&\, g_{\pi_{q-1}}(c,f) \le I_i \le g_{\pi_q}(c,f) \,\&\, \text{household observed at } t\right) \times f\left(a_{cfqt}|c,f,g_{\pi_{q-1}}(c,f) \le I_i \le g_{\pi_q}(c,f),t\right) \times \frac{\partial a_{cfqt}(\Delta_0,\gamma_0;\chi_0)}{\partial \Delta'}. \tag{B.20}$$

Proceeding similarly, it can be shown that each row of $\frac{\partial \varphi_\Delta(\Delta_0,\gamma_0;\chi_0)}{\partial \gamma'}$ has the following two non-zero elements:

$$\Pr\left(c_i = c \,\&\, f_i = f \,\&\, \text{household observed at } t\right) \times \begin{pmatrix} -f\left(g_{\pi_{q-1}}(c,f)|c,f\right)\left[F\left(a_{cfqt}|c,f,g_{\pi_{q-1}}(c,f),t\right) - 1/2\right] \\ f\left(g_{\pi_q}(c,f)|c,f\right)\left[F\left(a_{cfqt}|c,f,g_{\pi_q}(c,f),t\right) - 1/2\right] \end{pmatrix}', \tag{B.21}$$

with

$$f\left(g_{\pi_0}(c,f)|c,f\right)\left[F\left(a_{cfqt}|c,f,g_{\pi_0}(c,f),t\right)-1/2\right] \equiv 0, \quad \text{(B.22)}$$

$$f\left(g_{\pi_Q}(c,f)|c,f\right)\left[F\left(a_{cfqt}|c,f,g_{\pi_Q}(c,f),t\right)-1/2\right] \equiv 0. \quad \text{(B.23)}$$

In practice, we find $F\left(a_{cfqt}|c,f,I,t\right)$, the conditional c.d.f. of assets evaluated at the median a_{cfqt} by finding kernel estimates of the mean regression of $1\{a_{it} \leq a_{cfqt}\}$ on I (holding c,f and, t fixed).

Appendix C

Appendix to Chapter 4

C.1 Parameter Definitions

Table C.1 contains the definitions of parameters used in the model. See table 4.7 in Section 4.5 for the sources of parameters used in the model.

C.2 Uniqueness of Bargaining Outcomes

In this Appendix I discuss conditions for the uniqueness of the bargaining rule. The bargaining objective function is given by:

$$max_{\{T',s'\}} \Big(V(a^k+T',s',e^k,m^k,M^k,j)-V(a^k,s,e^k,m^k,M^k,j)\Big)^{\lambda}$$

$$\times\Big(V(a^p-T',s',e^p,m^p,M^p,j^P)-V(a^p,s,e^p,m^p,M^p,j^P)\Big)^{1-\lambda} \quad (C.1)$$

$$\text{s.t.} \quad (C.2)$$

$$a^p-T'\geq 0 \quad (C.3)$$

$$a^k+T'\geq 0 \quad (C.4)$$

$$0\leq s'<1 \quad (C.5)$$

For a dynasty with states indexed by X. Without loss of generality, define:

$$W^{X,p}(T',s')=V(a^p-T',s',e^p,m^p,M^p,j^P) \quad (C.6)$$

	Preference Parameters
β	Discount factor
σ	Consumption Weight in util. func.
δ	Taste for ill health
κ	Fixed cost of working
ϕ_1	Weight on Bequest
ϕ_2	Curvature of Bequest
γ	Coeff. of Rel. Risk. Aversion (util.)
	Earnings Process
$\{f_i\}_{i=0}^2$	Params. of det. component
$\{\hat{f}_i\}_{i=0}^2$	Biased (due to selection) $\{f_i\}_{i=0}^2$
u	Stochastic component
ψ	Autoregressive component in stoch. earn.
υ	Measurement error in stoch. earn.
ρ	Autoregressive param. in stoch. earn.
ε	Innovation to Autoregressive component
$\rho_{inherit}$	Integenerational persistence
v_{k0}	Initial innovation
σ_{k0}^2	Variance of initial innovation
σ_{ε}^2	Variance of subsequent innovations
	Medical Expenses
$\{H_i\}_{i=0}^2$	Params. of det. component
$\{\hat{H}_i\}_{i=0}^2$	Biased (due to selection) $\{H_i\}_{i=0}^2$
$\{\sigma_{M_i}\}_{i=0}^2$	Params. of variance
ζ	Stochastic component
μ	Autoreg. component in stoch. expenses.
ν	Transitory. component in stoch. expenses.
ρ_M	Autoreg. param. in stoch. expenses.
ω	Innovation to Autoregressive component
σ_{ν}^2	Variance of Transitory. component
σ_{ω}^2	Variance of Autoregressive component
	Taxes and Govt. Transfers
τ_l	Household income tax & benefits func.
τ_b	Bequest tax function
ex_b	Exemption level on bequests
τ_a	Capital gains tax
τ	All taxes
p	Universal pension
$\underline{c}$	Consumption Floor
b_M	Government transfers for consumption floor
	State Variables
$\mathbf{X}$	Vector of all state variables(dynasty)
$\mathbf{x}$	Vector of all state variables(household)
j	Age
a	Assets
s	Committed services
e	Idiosyncratic productivity
m	Health Status
M	Medical Expenses
	Household Choices
c	Consumption
a'	Next period assets
h	Labour supply
	Dynasty Choices
T'	Financial transfer
s'	Updated services
	Bargaining Game
λ	Bargaining weight on the child
$\Upsilon(\cdot)$	Deterministic Bargaining Rule

Table C.1: Parameter Definitions

The parent's payoff function for any choice of T' and s' given the dynasty is indexed by states X (in other words they are playing the game indexed by dynastic states X). Similarly, define:

$$W^{X,k}(T',s') = V(a^k + T', s', e^k, m^k, M^k, j) \tag{C.7}$$

As the child's payoff function for any choice of T' and s' given the dynasty is indexed by states X.

It is sufficient to show that the two payoff functions $W^{X,k}(T',s')$ and $W^{X,p}(T',s')$ are concave in arguments T' and s' in order to show uniqueness of the bargaining rule. Relaxing the assumptions on the functional form of the utility form the derivations are performed using assuming utility is independent of the transfer size, but depends flexibly on consumption, leisure, health status and services $u^i(c,l,s,m)$. However, I impose the following maintained assumptions:

$$u^i_c > 0, \quad u^i_l > 0 \quad i \in \{k,p\} \tag{C.8}$$

$$u^p_k < 0 \quad u^p_s > 0 \tag{C.9}$$

These are the standard assumptions of monotonicity in consumption and leisure as well as the assumption that parents enjoy the time of their child, but that it is costly for the child. As in the main text I maintain the normalisation that positive values of T are associated with flows away from parents to children.

I now turn to evaluation of the Hessian matrix for each of the indirect utility functions, $W^{X,p}$ and $W^{X,k}$. By standard application of the envelope theorem it follows that:

$$W^{X,p}_{T',T'} = -V_{a,a}(a^p - T', s', e^p, m^p, M^p, j^{CH}) = -u_{c,c} \tag{C.10}$$

$$W^{X,p}_{T',s'} = -V_{a,s'}(a^p - T', s', e^p, m^p, M^p, j^{CH}) = -u_{c,s'} \tag{C.11}$$

and

$$W^{X,k}_{T',T'} = V_{a,a}V(a^k + T', s', e^k, m^k, M^k, 1) = u_{c,c} \tag{C.12}$$

$$W^{X,k}_{T',s'} = V_{a,s'}V(a^k + T', s', e^k, m^k, M^k, 1) = u_{c,s} \tag{C.13}$$

Intuitively, there are two effects from a change in s'. The first is the combination of delayed effects (conditional on parent survival) of a reduction in total leisure endowment for the child (or increase for the parent). The second is any marginal change on future outcomes of bargaining. These are trivially zero for the parent who never bargains again and must also be zero for the child. Expectations over future bargaining outcomes can only change if the marginal change to the state variable s' has a first order change on optimal policies for the household. As optimal policies satisfy household first order conditions (and the non-negativity of leisure is never binding) these indirect effects are necessarily zero. It then follows that:

$$W^{X,p}_{s',s'} = -V_{s',s'}(a^p - T', s', e^p, m^p, M^p, j^{CH}) = \sum_{i=0}^{J-1} \beta^i u_{s',s'} \cdot Pr(alive at\, j^{CH} + i) \tag{C.14}$$

and

$$W^{X,k}_{s',s'} = V_{s',s'}V(a^k + T', s', e^k, m^k, M^k, 1) = \sum_{i=0}^{j^{CH}-1} -\beta^i u_{s',s'} \cdot Pr(Parent alive at\, j^{CH} + i) \tag{C.15}$$

Using Sylvester's criterion to establish concavity of the bargaining objective function, concavity of the bargaining objective function requires:

$$W^{X,p}_{s',s'} W^{X,p}_{T',T'} - (W^{X,p}_{s',T'})^2 < 0 \tag{C.16}$$

and

$$W^{X,k}_{s',s'}W^{X,k}_{T',T'} - (W^{X,k}_{s',T'})^2 < 0 \tag{C.17}$$

Substituting the results above, it is sufficient for period utility to be concave in consumption and services. Note that a similar argument holds for the extension of the model to repeated bargaining and renegotiation.

C.3 Definition of Recursive Stationary Equilibrium

As discussed in Section 4.3.7, the equilibrium of the intergenerational game played between parents and their children induces a recursive stationary equilibrium of the OLG economy. In contrast to the macroeconomic literature, prices are exogenous and the equilibrium is determined by the bargaining rule. A formal definition of the recursive stationary equilibrium is provided here.

A recursive stationary equilibrium is given by

$$\begin{cases} \text{allocations } c(x), h(x), a'(x), \\ \text{a rule for the outcome of the bargaining process } \Upsilon \\ \text{and a constant distribution of people over the state variables } x\text{: } \theta(x) \end{cases}$$

such that the following hold:

(i) Given the interest rate, the wage, government tax rates and transfers, the no borrowing constraint, the bargaining outcome rule Υ and distributions over counterparts, the functions $c(x)$, $h(x)$ and $a'(x)$ solve the above described maximization problem for a household with state variables x.

(ii) The probability distribution $\theta(x)$ is the stationary distribution associated with $[c(x), h(x), a'(x), \Upsilon, \{Q_{m,j}\}_{j=1,J}\{Q_{M,j}\}_{j=1,J}, \{Q_{e,j}\}_{j=1,j^{RT}-1}, p, \tau_a, \tau_b, \tau_l, ex_b,$ r, $\underline{c}]$,

(iii) The distribution of expected bequests is consistent with the distribution of counterparts (conditional on s) implied by $\theta(x)$ and the policy rules of parent households $a'(x)$

(iv) The distribution of expected bargaining partners is consistent with the all possible counterparts (conditional on s) implied by $\theta(x)$

(v) The rule for the bargaining outcomes, Υ is consistent with the distribution of consistent with the Markov Perfect Equilibrium of the bargaining process.

C.4 Additional Data Details

C.4.1 PSID Split Off Families

The original 1968 PSID drew two concurrent samples which were combined into the core PSID sample. The first was sampled to represent US demographics while the second allowed the oversampling of those with lower incomes. Since 1968 the sample size of the PSID has been expanded in two ways. There have been two additional samples added to the core 1968 sample, covering Latinos and immigrants, chosen in order to maintain the representativeness of the overall US demographics (which had changed between 1968 and the early 21st century). The sample has also increased by following those families which are 'split off'. Family Heads and their 'Wife' sampled in 1968 are said to have the 'PSID gene' which indicates that they are followed and reinterviewed by future waves of the PSID. Additionally, any child (biological or adoptive) of somebody with the 'PSID gene' inherits the 'PSID gene' and are also followed for future waves. When sample members with the 'PSID gene' in any family move out and establish a household this 'split off' household is sampled in addition to their original household. However, the sample includes information on additional individuals without the 'PSID gene' as questions are asked about all members of the household, such as the 'Wife' of a 1968 head's child or children in law more generally. In the case of divorce where a member of the household without the gene retains custody of the children this parent remains in the sample until the children leave that household.[1] Thus, over time, the PSID has expanded and in some cases has data on up to four generations of a sample family.

[1]For further details see the PSID documentation

C.4.2 Medical Expenditure Data

Since 1999 the PSID has asked respondent households a number of questions about their medical expenses in the last 2 years (the survey window). Households are asked for total medical expenses in this period, but this includes payments covered by Medicare, Medicaid and health insurance. Respondents are also prompted to include these costs in addition to copayment costs, see the following example from 1999:

> Could you give me your best estimate of the total cost of all medical care for [you/you and your family living there] in 1997 and 1998 combined? That is, the total of your out-of-pocket costs you just gave me, plus the costs covered by Medicare, Medicaid, or other health insurance?
>
> –(Question H82 reproduced from the 1999 wave PSID codebook)

The amount of medical expenses covered by health insurance, Medicare and Medicaid is not of interest in estimating the out of pocket medical expenses. Instead, I make use of questions which explicitly ask for out of pocket expenses in a number of medical categories. In order to construct the relevant measure of out of pocket medical expenses I aggregate the reported out-of-pocket amounts for doctor, outpatient surgery, dental bills; nursing home and hospital bills; prescriptions, in-home medical care, special facilities, and other services.

This aggregation raises two possible concerns:

First, I do not consider the role of formal nursing homes in the paper. As PSID respondents are not permanent residents of nursing homes these are temporary stays after a particular hospital visit or ailment.

Second, these expenses refer to the entire family unit and may include spending on hospital bills or prescriptions for dependants. However, this constructed variable is only used in constructing estimates for retired households. I explicitly restrict the sample to retired households and these households are less likely to have dependent family members.

Respondents who respond “Don’t Know”, “NA” or refuse a response to any individual category of medical expenditure are assigned a value of 0 expenses in this category. Those who respond in this way to all categories are dropped from the estimation of the process for logged medical expenses.

C.4.3 Transfer Data

The RTS draws a distinction between long term and short term gifts. Long term gifts are gifts received by a child since they turned 18. Short term gifts are gifts received in the last calendar year. The RTS is collected as part of the 2013 PSID, thus short term gifts are for the calendar year 2012. In order to accommodate this difference, the questions asked are slightly different. The following is an example of a short term question reproduced from the RTS codebook:

> In 2012, did [your/HEAD’s] (or spouse’s)] parent(s)/child(ren) give you any money, loans or gifts of $100 or more?

If the answer to the above question is “yes” then a follow up question is asked about the amount of the gift. In addition to the monetary amount households are asked to report the transfers as a unit of time. An example of a long term question is reproduced from the RTS codebook:

> (Since/ Next, since) turning 18, [have you/ has HEAD (or Wife/“Wife”)] received any of the following from [your/HEAD’s (Wife/“Wife’s”)] parents:
>
> Financial help for other expenses?

Again if the answer to the above question is “yes” then a follow up question is asked about the amount of the gift. For questions about the incidence of gifts a specific parent is referenced, however for the monetary amount reported no distinction on the source is reported. Separate questions ask about the transfers of time and money for short term gifts. For long term gifts in addition to money and time transfers there are separate questions for general financial help (as above); help with a child’s home; and help with education.

Bibliography

Abbott, Brant, Giovanni Gallipoli, Costas Meghir, and Giovanni L. Violante, "Education Policy and Intergenerational Transfers in Equilibrium," *Journal of Political Economy*, December 2018, pp. 000–000.

Abel, Andrew B. and Mark Warshawsky, "Specification of the Joy of Giving: Insights from Altruism," *The Review of Economics and Statistics*, February 1988, *70* (1), 145–149.

Aladangady, Aditya, "Housing Wealth and Consumption: Evidence from Geographically-Linked Microdata," *American Economic Review*, November 2017, *107* (11), 3415–3446.

Alessie, Rob, Annamaria Lusardi, and Arie Kapteyn, "Saving after retirement: evidence from three different surveys," *Labour Economics*, June 1999, *6* (2), 277–310.

Altonji, Joseph G. and Ernesto Villanueva Villanueva, "The Marginal Propensity to Spend on Adult Children *," in "in" 2003.

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

_ , Fumio Hayashi, and Laurence J. Kotlikoff, "Is the Extended Family Altruistically Linked? Direct Tests Using Micro Data," *The American Economic Review*, December 1992, *82* (5), 1177–1198.

_ , _ , and _ , "Parental Altruism and Inter Vivos Transfers: Theory and Evidence," *Journal of Political Economy*, December 1997, *105* (6), 1121–1166.

Ameriks, John, Andrew Caplin, Steven Laufer, and Stijn Van Nieuwerburgh, "The Joy of Giving or Assisted Living? Using Strategic Surveys to Separate Public Care Aversion from Bequest Motives," *The Journal of Finance*, 2011, *66* (2), 519–561.

_ , Joseph Briggs, Andrew Caplin, Matthew D. Shapiro, and Christopher Tonetti, "The Long-Term-Care Insurance Puzzle: Modeling and Measurement," Technical Report w22726, National Bureau of Economic Research, Cambridge, MA October 2016.

_ , _ , _ , _ , and _ , "Long-Term Care Utility and Late in Life Saving," Working Paper 20973, National Bureau of Economic Research November 2018.

Andreoni, James, "Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence," *Journal of Political Economy*, 1989, *97* (6), 1447–1458.

Andreski, Patricia, Geng Li, Mehmet Zahid Samancioglu, and Robert Schoeni, "Estimates of Annual Consumption Expenditures and Its Major Components in the PSID in Comparison to the CE †$," *American Economic Review*, May 2014, *104* (5), 132–135.

Angelini, Viola, Agar Brugiavini, and Guglielmo Weber, "The Dynamics of Homeownership among the 50+ in Europe," *Journal of Population Economics*, July 2014, *27* (3), 797–823.

Arnoudy, Antoine, Fatih Guvenen, and Tatjana Kleineberg, "Benchmarking Global Optimizers," Working Paper July 2019.

Attanasio, Orazio P., "Consumer Durables and Inertial Behaviour: Estimation and Aggregation of (S, s) Rules for Automobile Purchases," *The Review of Economic Studies*, 2000, *67* (4), 667–696.

_ and Carl Emmerson, "Mortality, Health Status, and Wealth," *Journal of the European Economic Association*, 2003, *1* (4), 821–850.

__ **and Hilary Williamson Hoynes**, "Differential Mortality and Wealth Accumulation," *Journal of Human Resources*, 2000, *35* (1), 1–29.

__ **, Costas Meghir, and Corina Mommaerts**, "Insurance in extended family networks," Working Paper 21059, National Bureau of Economic Research April 2015.

__ **, Hamish Low, and Virginia Sanchez-Marcos**, "Explaining Changes in Female Labor Supply in a Life-Cycle Model," *American Economic Review*, September 2008, *98* (4), 1517–1552.

__ **, Renata Bottazzi, Hamish W. Low, Lars Nesheim, and Matthew Wakefield**, "Modelling the demand for housing over the life cycle," *Review of Economic Dynamics*, January 2012, *15* (1), 1–18.

Bajari, Patrick, Phoebe Chan, Dirk Krueger, and Daniel Miller, "A Dynamic Model of Housing Demand: Estimation and Policy Implications," *International Economic Review*, 2013, *54* (2), 409–442.

Banks, James, Richard Blundell, Peter Levell, and James P. Smith, "Life-Cycle Consumption Patterns at Older Ages in the United States and the United Kingdom: Can Medical Expenditures Explain the Difference?," *American Economic Journal: Economic Policy*, August 2019, *11* (3), 27–54.

Barczyk, Daniel and Matthias Kredler, "Evaluating Long-Term-Care Policy Options, Taking the Family Seriously," *The Review of Economic Studies*, April 2018, *85* (2), 766–809.

Becker, Gary S., "Altruism, Egoism, and Genetic Fitness: Economics and Sociobiology," *Journal of Economic Literature*, September 1976, *14* (3), 817–826.

__ **and Nigel Tomes**, "Human Capital and the Rise and Fall of Families," *Journal of Labor Economics*, 1986, *4* (3), 1–39.

Bénabou, Roland, "Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?," *Econometrica*, March 2002, *70* (2), 481–517.

Berger, David, Veronica Guerrieri, Guido Lorenzoni, and Joseph Vavra, "House Prices and Consumer Spending," *The Review of Economic Studies*, 2018.

Bernheim, B. Douglas and Sergei Severinov, "Bequests as Signals: An Explanation for the Equal Division Puzzle," *Journal of Political Economy*, August 2003, *111* (4), 733–764.

_ , Andrei Shleifer, and Lawrence H. Summers, "The Strategic Bequest Motive," *Journal of Political Economy*, December 1985, *93* (6), 1045–1076.

Besley, Timothy, Neil Meads, and Paolo Surico, "The incidence of transaction taxes: Evidence from a stamp duty holiday," *Journal of Public Economics*, November 2014, *119*, 61–70.

Best, Michael Carlos and Henrik Jacobsen Kleven, "Housing Market Responses to Transaction Taxes: Evidence From Notches and Stimulus in the U.K.," *The Review of Economic Studies*, January 2018, *85* (1), 157–193.

Blau, David M. and Donna Gilleskie, "The Role of Retiree Health Insurance in the Employment Behavior of Older Men," *International Economic Review*, 2008, *49* (2), 475–514.

Blundell, Richard, Monica Costa Dias, Costas Meghir, and Jonathan Shaw, "Female Labor Supply, Human Capital, and Welfare Reform," *Econometrica*, September 2016, *84* (5), 1705–1753.

_ , Rowena Crawford, Eric French, and Gemma Tetlow, "Comparing Retirement Wealth Trajectories on Both Sides of the Pond," *Fiscal Studies*, March 2016, *37* (1), 105–130.

Boca, Daniela Del and Christopher Flinn, "Endogenous household interaction," *Journal of Econometrics*, January 2012, *166* (1), 49–65.

Bonhomme, Stéphane, Thibaut Lamadon, and Elena Manresa, "Discretizing unobserved heterogeneity," Technical Report March 2017.

_ , _ , and _ , "A Distributional Framework for Matched Employer Employee Data," *Econometrica*, 2019, *87* (3), 699–739.

Bozio, Antoine, Guy Laroque, and Cormac O'Dea, "Discount rate heterogeneity among older households: a puzzle?," *Journal of Population Economics*, April 2017, *30* (2), 647–680.

Braun, R. Anton, Karen A. Kopecky, and Tatyana Koreshkova, "Old, Sick, Alone and Poor: A Welfare Analysis of Old-Age Social Insurance Programs," *The Review of Economic Studies*, 2017, *84* (2), 580–612.

_ , _ , and _ , "Old, Frail, and Uninsured: Accounting for Features of the U.S. Long-Term Care Insurance Market," *Econometrica*, May 2019.

Brown, Meta, "Informal Care and the Division of End-of-Life Transfers," *The Journal of Human Resources*, 2006, *41* (1), 191–219.

Buchinsky, Moshe, "Recent Advances in Quantile Regression Models: A Practical Guideline for Empirical Research," *Journal of Human Resources*, 1998, *33*, 88–126.

Byrne, David, Michelle S. Goeree, Bridget Hiedemann, and Steven Stern, "Formal Home Health Care, Informal Care, and Family Decision Making*," *International Economic Review*, November 2009, *50* (4), 1205–1242.

Campbell, John Y. and João F. Cocco, "How do house prices affect consumption? Evidence from micro data," *Journal of Monetary Economics*, April 2007, *54* (3), 591–621.

Carroll, Christopher D., "Why Do the Rich Save So Much?," 1998. NBER Working Paper 6549.

__ , "The method of endogenous gridpoints for solving dynamic stochastic optimization problems," *Economics Letters*, June 2006, *91* (3), 312–320.

Casanova, Maria, "Happy Together: A Structural Model of Couples' Retirement Choices"," 2011. Mimeo.

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

Chandler, Daniel and Richard Disney, "Measuring house prices: a comparison of different indices," Technical Report, Institute for Fiscal Studies May 2014.

Chetty, Raj and Adam Szeidl, "Consumption Commitments and Risk Preferences," *The Quarterly Journal of Economics*, May 2007, *122* (2), 831–877.

__ , Nathaniel Hendren, Patrick Kline, Emmanuel Saez, and Nicholas Turner, "Is the United States Still a Land of Opportunity? Recent Trends in Intergenerational Mobility," Working Paper 19844, National Bureau of Economic Research January 2014.

Choi, HwaJung, Kathleen McGarry, and Robert F. Schoeni, "Consumption and the extended family," *Economics Letters*, March 2016, *140*, 34–38.

Cocco, João F., "Evidence on the Benefits of Alternative Mortgage Products," *The Journal of Finance*, 2013, *68* (4), 1663–1690.

__ **and Paula Lopes**, "Aging in Place, Housing Maintenance, and Reverse Mortgages," *The Review of Economic Studies*, 2019.

Conesa, Juan Carlos, Sagiri Kitao, and Dirk Krueger, "Taxing Capital? Not a Bad Idea after All!," *American Economic Review*, March 2009, *99* (1), 25–48.

Cox, Donald, "Motives for Private Income Transfers," *Journal of Political Economy*, June 1987, *95* (3), 508–546.

_ **and Mark R. Rank**, "Inter-Vivos Transfers and Intergenerational Exchange," *The Review of Economics and Statistics*, May 1992, *74* (2), 305–314.

Crawford, Rowena, "The use of wealth in retirement," Technical Report June 2018.

_ **and Pierfrancesco Mei**, "An overview of the ELSA 'end of life' data," Technical Report June 2018.

Dauth, Wolfgang, Sebastian Findeisen, Enrico Moretti, and Jens Suedekum, "Matching in Cities," Technical Report w25227, National Bureau of Economic Research, Cambridge, MA November 2018.

Davidoff, Thomas and Jake Wetzel, "Do Reverse Mortgage Borrowers Use Credit Ruthlessly?," SSRN Scholarly Paper ID 2279930, Social Science Research Network, Rochester, NY July 2014.

_ **, Patrick Gerhard, and Thomas Post**, "Reverse mortgages: What homeowners (don't) know and how it matters," *Journal of Economic Behavior & Organization*, January 2017, *133*, 151–171.

Davies, James B. and Antony F. Shorrocks, "The Distribution of Wealth," in Anthony B. Atkinson and François Bourguignon, eds., *Handbook of Income Distribution*, Handbooks in Economics, vol. 16. Amsterdam; New York and Oxford: Elsevier Science, North-Holland, 2000, pp. 605–675.

Davis, Morris A. and François Ortalo-Magné, "Household expenditures, wages, rents," *Review of Economic Dynamics*, April 2011, *14* (2), 248–261.

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

_ **and Christina Paxson**, "Mortality, Education, Income, and Inequality among American Cohorts," in David A. Wise, ed., *Themes in the Economics of Aging*, NBER Conference Report Series. Chicago and London: University of Chicago Press, 2001, pp. 129–165.

Díaz-Giménez, Javier, Vincenzo Quadrini, and José-Víctor Ríos-Rull, "Dimensions of Inequality: Facts on the U.S. Distributions of Earnings, Income and Wealth," *Federal Reserve Bank of Minneapolis Quarterly Review*, 1997, *21* (2), 3–21.

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

_ and Fang Yang, "Wealth inequality, family background, and estate taxation," *Journal of Monetary Economics*, February 2016, *77*, 130–145.

_ , Eric French, and John B. Jones, "Life Expectancy and Old Age Savings," *American Economic Review: Papers and Proceedings*, 2009, *99* (2), 110–115.

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

_ , _ , and John Bailey Jones, "Medicaid Insurance in Old Age," *American Economic Review*, November 2016, *106* (11), 3480–3520.

_ , _ , and _ , "Savings After Retirement: A Survey," *Annual Review of Economics*, 2016, *8* (1), 177–204.

_ , _ , John B. Jones, and Jeremy McCauley, "Medical Spending of the U.S. Elderly," *Fiscal Studies*, 2016, *37* (3-4), 717–747.

_ , _ , _ , and Rory McGee, "Couples and Singles' Savings After Retirement," Working Paper 2018.

Duffie, Darrell and Kenneth J. Singleton, "Simulated Moments Estimation of Markov Models of Asset Prices," *Econometrica*, 1993, *61* (4), 929–952.

Dynan, Karen E., Jonathan Skinner, and Stephen P. Zeldes, "The Importance of Bequests and Life-Cycle Saving in Capital Accumulation: A New Answer," *American Economic Review*, 2002, *92* (2), 274–278.

_ , _ , and _ , "Do the Rich Save More?," *Journal of Political Economy*, 2004, *112* (2), 397–444.

Etheridge, Ben, "House Prices and Consumption Inequality," Working Paper November 2017.

Fagereng, Andreas, Benjamin Moll, Martin Blomhoff Holm, and Gisle Natvik, "Saving Behavior Across the Wealth Distribution: The Importance of Capital Gains," Technical Report July 2019.

Fahle, Sean, "Harnessing the Potential of Family Caregivers: A Solution to the U.S. Long-Term Care Crisis?," Working Paper May 2015.

Feenberg, Daniel and Jonathan Skinner, "The Risk and Duration of Catastrophic Health Care Expenditures," *Review of Economics and Statistics*, 1994, *76*, 633–647.

Feldstein, Martin S., "The Effects of Taxation on Risk Taking," *Journal of Political Economy*, September 1969, *77* (5), 755–764.

Fella, Giulio, "A generalized endogenous grid method for non-smooth and non-concave problems," *Review of Economic Dynamics*, April 2014, *17* (2), 329–344.

Fernández-Villaverde, Jesús and Dirk Krueger, "Consumption over the Life Cycle: Facts from Consumer Expenditure Survey Data," *The Review of Economics and Statistics*, 2007, *89* (3), 552–565.

French, Eric, "The Effects of Health, Wealth, and Wages on Labour Supply and Retirement Behaviour," *The Review of Economic Studies*, April 2005, *72* (2), 395–427.

_ and John Bailey Jones, "On the distribution and dynamics of health care costs," *Journal of Applied Econometrics*, January 2004, *19* (6), 705–721.

_ and _ , "The Effects of Health Insurance and Self-Insurance on Retirement Behavior," 2007. Michigan Retirement Research Center Working Paper 2007-170.

_ and _ , "The Effects of Health Insurance and Self-Insurance on Retirement Behavior," *Econometrica*, May 2011, *79* (3), 693–732.

_ , Mariacristina De Nardi, John B. Jones, Olyesa Baker, and Phil Doctor, "Right before the End: Asset Decumulation at the End of Life," *Economic Perspectives*, 2006, *30* (3), 2–13.

Gale, William G. and John Karl Scholz, "Intergenerational Transfers and the Accumulation of Wealth," *The Journal of Economic Perspectives*, October 1994, *8* (4), 145–160.

_ and Maria G. Perozek, "Do Estate Taxes Reduce Saving," in John Shoven, ed., *Rethinking Estate and Gift Taxation*, Brookings Institution Press, 2001.

Gallipoli, Giovanni, Jutong Pan, and Sarah Turner, "Household Responses to Individual Shocks: Disability and Labour Supply"," 2013. Mimeo.

Gan, Li, Guan Gong, Michael Hurd, and Daniel McFadden, "Subjective mortality risk and bequests," *Journal of Econometrics*, October 2015, *188* (2), 514–525.

Gavazza, Alessandro, Simon Mongey, and Giovanni L. Violante, "Aggregate Recruiting Intensity," *American Economic Review*, August 2018, *108* (8), 2088–2127.

Gelman, Andrew and Guido Imbens, "Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs," *Journal of Business & Economic Statistics*, August 2017, *0* (0), 1–10.

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

Groneck, Max, "Bequests and Informal Long-Term Care: Evidence from HRS Exit Interviews," *Journal of Human Resources*, March 2016.

Guner, Nezih, Remzi Kaygusuz, and Gustavo Ventura, "Income taxation of U.S. households: Facts and parametric estimates," *Review of Economic Dynamics*, October 2014, *17* (4), 559–581.

Guren, Adam, Alisdair McKay, Emi Nakamura, and Jón Steinsson, “Housing Wealth Effects: The Long View,” Working Paper w24729, National Bureau of Economic Research, Cambridge, MA June 2018.

Guvenen, Fatih and Anthony A. Smith, “Inferring Labor Income Risk and Partial Insurance From Economic Choices,” *Econometrica*, 2014, *82* (6), 2085–2129.

Hastie, Trevor, Robert Tibshirani, and Jerome Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.*, 2nd ed., Springer, 2009.

Havránek, Tomáš, “Measuring Intertemporal Substitution: The Importance of Method Choices and Selective Reporting,” *Journal of the European Economic Association*, 2015, *13* (6), 1180–1204.

Hendren, Nathaniel, “Private Information and Insurance Rejections,” *Econometrica*, September 2013, *81* (5), 1713–1762.

Hilber, Christian A. L. and Teemu Lyytikäinen, “Transfer taxes and household mobility: Distortion on the housing or labor market?,” *Journal of Urban Economics*, September 2017, *101*, 57–73.

Hirshleifer, Jack, “Shakespeare vs. Becker on Altruism: The Importance of Having the Last Word,” *Journal of Economic Literature*, June 1977, *15* (2), 500–502.

Hong, Jay H. and José-Víctor Ríos-Rull, “Life Insurance and Household Consumption,” *The American Economic Review*, 2012, *102* (7), 3701–3730.

Hubbard, R. Glenn, Jonathan Skinner, and Stephen P. Zeldes, “Expanding the Life-Cycle Model: Precautionary Saving and Public Policy,” *The American Economic Review*, 1994, *84* (2), 174–179.

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

Hubener, Andreas, Raimond Maurer, and Olivia S. Mitchell, "How Family Status and Social Security Claiming Options Shape Optimal Life Cycle Portfolios," *The Review of Financial Studies*, April 2016, *29* (4), 937–978.

Humphries, John Eric, "The Causes and Consequences of Self-Employment over the Life Cycle," 2018, p. 83.

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

Hurd, Michael D., "Mortality Risk and Bequests," *Econometrica*, 1989, *57* (4), 779–813.

_ , "Mortality Risk and Bequests," *Econometrica*, 1989, *57* (4), 779–813.

_ , "Mortality Risk and Consumption by Couples," 1999. Working Paper 7048, National Bureau of Economic Research.

Inkmann, Joachim and Alexander Michaelides, "Can the Life Insurance Market Provide Evidence for a Bequest Motive?," *Journal of Risk and Insurance*, September 2012, *79* (3), 671–695.

Iskhakov, Fedor, Thomas H. Jørgensen, John Rust, and Bertel Schjerning, "The endogenous grid method for discrete-continuous dynamic choice models with (or without) taste shocks," *Quantitative Economics*, July 2017, *8* (2), 317–365.

Jones, John Bailey, Mariacristina De Nardi, Eric French, Rory McGee, and Justin Kirschner, "The Lifetime Medical Spending of Retirees," Working Paper 24599, National Bureau of Economic Research May 2018.

Juster, F. Thomas, James P Smith, and Frank Stafford, "The measurement and structure of household wealth," *Labour Economics*, June 1999, *6* (2), 253–275.

Kaplan, Greg, "Moving Back Home: Insurance against Labor Market Risk," *Journal of Political Economy*, June 2012, *120* (3), 446–512.

_ **and Giovanni L. Violante**, "A Model of the Consumption Response To Fiscal Stimulus Payments," *Econometrica*, 2014, *82* (4), 1199–1239.

_ **, Kurt Mitman, and Giovanni L Violante**, "Non-durable Consumption and Housing Net Worth in the Great Recession: Evidence from Easily Accessible Data," Working Paper 22232, National Bureau of Economic Research May 2016.

Keane, Michael P, "Financial Aid, Borrowing Constraints, and College Attendance: Evidence from Structural Estimates," *American Economic Review*, May 2002, *92* (2), 293–297.

Ko, Ami, "An Equilibrium Analysis of the Long-Term Care Insurance Market," Working Paper 2018.

Kolesár, Michal and Christoph Rothe, "Inference in Regression Discontinuity Designs with a Discrete Running Variable," *American Economic Review*, August 2018, *108* (8), 2277–2304.

Kopczuk, Wojciech, "Incentive Effects of Inheritances and Optimal Estate Taxation," *The American Economic Review*, May 2013, *103* (3), 472–477.

_ **and David Munroe**, "Mansion Tax: The Effect of Transfer Taxes on the Residential Real Estate Market," *American Economic Journal: Economic Policy*, 2015, *7* (2), 214–257.

_ **and Joseph P. Lupton**, "To Leave or Not to Leave: The Distribution of Bequest Motives," *The Review of Economic Studies*, January 2007, *74* (1), 207–235.

Kotlikoff, Laurence J., Kent Smetters, and Jan Walliser, "Privatizing Social Security in the United States—Comparing the Options," *Review of Economic Dynamics*, July 1999, *2* (3), 532–574.

Kvaerner, Jens, "How Large Are Bequest Motives? Evidence Based on Shocks to Mortality," SSRN Scholarly Paper ID 2985465, Social Science Research Network, Rochester, NY June 2017.

Laitner, John, "Bequests, Gifts, and Social Security," *The Review of Economic Studies*, April 1988, *55* (2), 275–299.

_ , "Random earnings differences, lifetime liquidity constraints, and altruistic intergenerational transfers," *Journal of Economic Theory*, December 1992, *58* (2), 135–170.

_ **and F. Thomas Juster**, "New Evidence on Altruism: A Study of TIAA-CREF Retirees," *The American Economic Review*, 1996, *86* (4), 893–908.

_ **, Dan Silverman, and Dmitriy Stolyarov**, "The Role of Annuitized Wealth in Post-retirement Behavior," *American Economic Journal: Macroeconomics*, July 2018, *10* (3), 71–117.

Lee, David S. and David Card, "Regression discontinuity inference with specification error," *Journal of Econometrics*, February 2008, *142* (2), 655–674.

Lee, Siha and Kegon T. K. Tan, "Bequest Motives and the Social Security Notch," January 2017.

Lentz, Rasmus, Suphanit Piyapromdee, and Jean-Marc Robin, "On Worker and Firm Heterogeneity in Wages andEmployment Mobility: Evidence from DanishRegister Data," February 2019.

Lockwood, Lee M., "Incidental Bequests and the Choice to Self-Insure Late-Life Risks," *American Economic Review*, September 2018, *108* (9), 2513–2550.

Love, David A., "The Effects of Marital Status and Children on Savings and Portfolio Choice," *The Review of Financial Studies*, 2010, *23* (1), 385–432.

Mazzocco, Maurizio, "Household Intertemporal Behavior: A Collective Characterization and Empirical Tests," 2003. Mimeo.

McGarry, Kathleen, "Dynamic aspects of family transfers," *Journal of Public Economics*, May 2016, *137*, 1–13.

Mian, Atif, Kamalesh Rao, and Amir Sufi, "Household Balance Sheets, Consumption, and the Economic Slump," *The Quarterly Journal of Economics*, November 2013, *128* (4), 1687–1726.

Misra, Kanishka and Paolo Surico, "Consumption, Income Changes, and Heterogeneity: Evidence from Two Fiscal Stimulus Programs," *American Economic Journal: Macroeconomics*, October 2014, *6* (4), 84–106.

Mitman, Kurt, "Macroeconomic Effects of Bankruptcy and Foreclosure Policies," *American Economic Review*, August 2016, *106* (8), 2219–2255.

Mommaerts, Corina, "Long-Term Care Insurance and the Family," Working Paper November 2016.

Mullins, Joseph, "Improving Child Outcomes Through Welfare Reform," Working Paper 2018.

Nagaraja, Chaitra H., Lawrence D. Brown, and Linda H. Zhao, "An autoregressive approach to house price modeling," *The Annals of Applied Statistics*, March 2011, *5* (1), 124–149.

Nakajima, Makoto and Irina A. Telyukova, "Reverse Mortgage Loans: A Quantitative Analysis," *The Journal of Finance*, April 2017, *72* (2), 911–950.

_ **and Irina A Telyukova**, "Home Equity in Retirement," Working Paper 2018.

_ **and Irina A. Telyukova**, "Medical Expenses and Saving in Retirement: The Case of U.S. and Sweden," April 2018.

Nelder, J. A. and R. Mead, "A Simplex Method for Function Minimization," *The Computer Journal*, January 1965, *7* (4), 308–313.

Newey, Whitney K., "Generalized Method of Moments Specification Testing," *Journal of Econometrics*, 1985, *29* (3), 229–256.

_ **and Daniel L. McFadden**, "Large Sample Estimation and Hypothesis Testing," in Robert Engle and Daniel L. McFadden, eds., *Handbook of Econometrics, Volume 4*, Elsevier, Amsterdam, 1994.

Nishiyama, Shinichi, "Bequests, Inter Vivos Transfers, and Wealth Distribution," *Review of Economic Dynamics*, October 2002, *5* (4), 892–931.

Norton, Edward C., Lauren Hersch Nicholas, and Sean Sheng-Hsiu Huang, "Informal Care and Inter-vivos Transfers: Results from the National Longitudinal Survey of Mature Women," Working Paper 18948, National Bureau of Economic Research April 2013.

Ortalo-Magné, François and Sven Rady, "Housing transactions and macroeconomic fluctuations: a case study of England and Wales," *Journal of Housing Economics*, December 2004, *13* (4), 287–303.

Pakes, Ariel and David Pollard, "Simulation and the Aysmptotics of Optimization Estimators," *Econometrica*, 1989, *57* (5), 1027–1057.

Palumbo, Michael G., "Uncertain Medical Expenses and Precautionary Saving Near the End of the Life Cycle," *The Review of Economic Studies*, 1999, *66* (2), 395–421.

Pantano, Juan and Yu Zheng, "Using Subjective Expectations Data to Allow for Unobserved Heterogeneity in Hotz-Miller Estimation Strategies," SSRN Scholarly Paper ID 2129303, Social Science Research Network, Rochester, NY December 2013.

Parker, Jonathan A, Nicholas S Souleles, David S Johnson, and Robert McClelland, "Consumer Spending and the Economic Stimulus Payments of 2008," *American Economic Review*, October 2013, *103* (6), 2530–2553.

Pischke, Jorn-Steffen, "Measurement Error and Earnings Dynamics: Some Estimates from the PSID Validation Study," *Journal of Business & Economics Statistics*, 1995, *13* (3), 305–314.

Poterba, James M., Steven F. Venti, and David A. Wise, "The Composition and Drawdown of Wealth in Retirement," *Journal of Economic Perspectives*, 2011, *25*, 95–117.

_ , _ , and _ , "Longitudinal Determinants of End-of-Life Wealth Inequality," Working Paper 23839, National Bureau of Economic Research September 2017.

Powell, James, "Estimation of Semiparametric Models," in Robert Engle and Daniel L McFadden, eds., *Handbook of Econometrics, Volume 4*, Elsevier, Amsterdam, 1994.

Powell, M J D, "The BOBYQA algorithm for bound constrained optimization without derivatives," Working Paper August 2009.

Robinson, James, "A long-term care status transition model," Technical Report, Georgia State Universit 1996.

Rohwedder, Susann, Steven J. Haider, and Michael Hurd, "Increases in Wealth among the Elderly in the Early 1990s: How Much is Due to Survey Design?," *Review of Income and Wealth*, 2006, *52* (4), 509–524.

Schulhofer-Wohl, Sam, "The age-time-cohort problem and the identification of structural parameters in life-cycle models," *Quantitative Economics*, July 2018, *9* (2), 643–658–658.

Searle, Samuel D., Arnold Mitnitski, Evelyne A. Gahbauer, Thomas M. Gill, and Kenneth Rockwood, "A standard procedure for creating a frailty index," *BMC Geriatrics*, September 2008, *8* (1), 24.

Sinai, Todd and Nicholas S. Souleles, "Owner-Occupied Housing as a Hedge against Rent Risk," *The Quarterly Journal of Economics*, 2005, *120* (2), 763–789.

Skinner, Jonathan, "Is Housing Wealth a Sideshow?," Working Paper, National Bureau of Economic Research, Cambridge, MA November 1993.

Slemrod, Joel, Caroline Weber, and Hui Shan, "The behavioral response to housing transfer taxes: Evidence from a notched change in D.C. policy," *Journal of Urban Economics*, July 2017, *100*, 137–153.

Solon, Gary, "Intergenerational Income Mobility in the United States," *The American Economic Review*, 1992, *82* (3), 393–408.

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

Tibshirani, Robert, Guenther Walther, and Trevor Hastie, "Estimating the number of clusters in a data set via the gap statistic," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, May 2001, *63* (2), 411–423.

Tomes, Nigel, "Inheritance and Inequality Within the Family: Equal Division Among Unequals, or Do the Poor Get More?," in Denis Kessler and Andre Masson, eds., *Modelling the accumulation and distribution of wealth*, Clarendon Press, 1988, pp. 79–104.

van der Klaauw, Wilbert, "On the Use of Expectations Data in Estimating Structural Dynamic Choice Models," *Journal of Labor Economics*, 2012, *30* (3), 521–554.

_ **and Kenneth I. Wolpin**, "Social security and the retirement and savings behavior of low-income households," *Journal of Econometrics*, July 2008, *145* (1–2), 21–42.

Venti, Steven F. and David A. Wise, "The Cause of Wealth Dispersion at Retirement: Choice or Chance?," *The American Economic Review*, 1998, *88* (2), 185–191.

Voena, Alessandra, "Yours, Mine, and Ours: Do Divorce Laws Affect the Intertemporal Behavior of Married Couples?," *American Economic Review*, August 2015, *105* (8), 2295–2332.

Waters, Tom, "TAXBEN: The IFS tax and benefit microsimulation model," Technical Report November 2017.

Wilhelm, Mark O., "Bequest Behavior and the Effect of Heirs' Earnings: Testing the Altruistic Model of Bequests," *The American Economic Review*, September 1996, *86* (4), 874–892.

Yaari, Menahem E., "Uncertain Lifetime, Life Insurance, and the Theory of the Consumer," *The Review of Economic Studies*, April 1965, *32* (2), 137–150.