MICHIGAN RETIREMENT AND DISABILITY RESEARCH CENTER
UNIVERSITY OF MICHIGAN

Promoting research on retirement, disability, and Social Security policy

Medical Spending Risk among Retired Households by Race

Karolos Arapakis, Eric French, John Bailey Jones, and Jeremy McCauley

MRDRC WP 2023-475

UM23-02

Medical Spending Risk among Retired Households by Race

Karolos Arapakis
Boston College

John Bailey Jones
Federal Reserve Bank of Richmond

Eric French
University of Cambridge, IFS

Jeremy McCauley
University of Bristol

November 2023

Michigan Retirement and Disability Research Center, University of Michigan, P.O. Box 1248. Ann Arbor, MI 48104, mrdrc.isr.umich.edu, (734) 615-0422

Acknowledgements

The research reported herein was performed pursuant to a grant from the U.S. Social Security Administration (SSA) funded as part of the Retirement and Disability Research Consortium through the University of Michigan Retirement and Disability Research Center Award RDR18000002-05. The opinions and conclusions expressed are solely those of the author(s) and do not represent the opinions or policy of SSA or any agency of the federal government. Neither the United States government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of the contents of this report. Reference herein to any specific commercial product, process or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply endorsement, recommendation or favoring by the United States government or any agency thereof.

Medical Spending Risk among Retired Households by Race

Citation

Arapakis, Karolos, Eric French, John Bailey Jones, and Jeremy McCauley. 2023. “Medical Spending Risk among Retired Households by Race.” Ann Arbor, MI. University of Michigan Retirement and Disability Research Center (MRDRC) Working Paper; MRDRC WP 2023-475.
https://mrdrc.isr.umich.edu/publications/papers/pdf/wp475.pdf

M
RDRC

Medical Spending Risk among Retired Households by Race *

Abstract

Using data from the Health and Retirement Study linked to administrative Medicare and Medicaid records, along with the Medical Expenditure Panel Survey, we examine how total and out-of-pocket medical expenditures by retired households vary across race, both annually and over their remaining life spans. We find that in a given year all races have similar total expenditures, with any differences attributable to age, education, income, and household structure. Racial inequities in spending largely reflect inequities in other aspects of society, such as educational differences. We also evaluate remaining lifetime medical spending of 65 year olds. Because they have shorter life spans, total remaining lifetime medical spending is lower for individuals in Black households than for those in white or Hispanic households. We also evaluate the different payors of medical spending. For all groups, Medicare and Medicaid pay the majority of all medical expenses. Black and Hispanic households spend less out of pocket than white households and have a greater share paid by Medicaid and Medicare. This largely reflects the greater Medicaid recipiency of racial minorities stemming from their lower economic resources. At age 65, white households will on average incur around $100,000 in out-of-pocket medical spending on deductibles, co-pays, and other liabilities (but excluding insurance premia) over the remainder of their lives, versus $48,000 and $42,000 for Black and Hispanic households. Thus, Black and Hispanic households are better insured by Medicare and Medicaid.

Keywords: medical spending, Medicaid, Medicare, long-term care.

*Preliminary, please do not quote. An earlier draft of this paper was titled "The Distribution of Medical Spending Risk among Retired Households by Race and Gender". The current draft now focuses more heavily on racial differences and therefore we have changed the title. While the current paper still considers gender differences, racial differences are starker and thus we focus on racial differences. We are very grateful to seminar participants at the MRDRC Annual Conference for comments. We are very grateful to Chris Firth, Pauline Sow, Nathan Robino, and Erin Henry for excellent research assistance. Funding from the Social Security Administration through the Michigan Retirement Research Center (MRDRC grant UM23-02), the Economic and Social Research Council (grants (ES/V001248/1) and (ES/T014334/1)) for this work is gratefully acknowledged. The opinions and conclusions expressed are solely those of the authors and do not represent the opinions or policy of SSA, the Center for Retirement Research at Boston College, the Federal Reserve Bank of Richmond, the Federal Reserve System, or any agency of the Federal Government. Any errors are the authors' own.

1 Introduction

We examine how the medical spending risk of older households varies by race. Despite nearly universal enrollment in Medicare, medical spending is a major financial concern among American retirees. Because Medicare does not pay for long hospital and nursing home stays and requires co-payments for many other treatments, most retirees still face the risk of catastrophic, out-of-pocket medical expenses. We distinguish between spending covered by public insurance programs such as Medicare and Medicaid and the out-of-pocket expenses borne by the households themselves. In doing so, we examine the extent to which Medicare and Medicaid insures against the medical spending risk of different races.

We use household-level data from the Health and Retirement Study (HRS) linked to administrative Medicare and Medicaid records, along with the Medical Expenditure Panel Survey (MEPS). Thus, we construct a long panel of all payors for the age 65+ population.

We start by examining racial differences in annual medical spending and whether these differences are mediated through observable factors such as health, income, education, age and marital status. Although Black and Hispanic households spend modestly less than their white counterparts, after controlling for observable variables, their spending is virtually identical. Racial inequities in spending largely reflect inequities in other aspects of society, such as educational differences.

In addition to comparing differences in annual spending, we also measure cumulative lifetime spending. We believe we are the first to measure differences in lifetime medical spending by race. Households care not only about the risk of catastrophic expenses in a single year, but also about the risk of moderate but persistent expenses that accumulate into catastrophic lifetime costs. To account for persistent medical spending, we model the stochastic process for medical spending with the framework developed by Arellano, Blundell, and Bonhomme (2017) for each racial group. This allows for non-linear persistence and non-normal shocks, making it more flexible than what has been used in previous models of medical spending. This approach allows us to better understand the medical spending risk facing older households over the remainders of their lives. Furthermore, we account for racial differences in health and longevity.

At age 65, white households will incur, on average, around \$380,000 of medical expenses over the remainder of their lives, versus \$280,000 and \$400,000 for Black and Hispanic households, respectively. The risk of catastrophic spending is significant for all racial groups. At age 65, white households at the 90th percentile of the spending distribution will incur spending of \$800,000 versus \$650,000 and \$1,000,000 for Black and Hispanic households, respectively. This lower medical spending of Black households largely reflects their shorter life spans.

Medicare, Medicaid, other private and government payors cover the majority, but not all, of medical expenses. To better understand how much households pay out of pocket, we model the co-insurance rates that they face in the way of co-pays, deductibles, and other liabilities. Black and Hispanic households spend less out of pocket than white households and have a greater share paid by Medicaid and Medicare. This largely reflects the greater Medicaid recipiency of racial minorities stemming from lower economic resources. At age 65, white households will, on average, incur around $100,000 in out-of-pocket medical spending on deductibles, co-pays, and other liabilities (but excluding insurance premia) over the remainder of their lives, versus $48,000 and $42,000 for Black and Hispanic households. This means that due to Medicaid eligibility rules, Black and Hispanic households are better insured against high medical expenses.

The rest of the paper is organized as follows. Section 2 contains a literature review. In section 3, we discuss some key features of the data sets that we use in our analysis, the HRS and MEPS, and describe how we construct our measure of medical spending. In section 5 we introduce our model and describe our simulation methodology. We discuss our results in section 6 and conclude in section 8.

2 Previous Literature

There is an emerging literature focusing on racial disparities in health care spending. Most studies find that white individuals spend modestly more than other races (e.g. Escarce and Kapur 2003). In particular, Dieleman et al. (2021) show that among individuals age 65+, healthcare spending of white individuals is 4% higher than for Black individuals and 29% higher than for Hispanic individuals. Other studies find racial differences in particular areas of healthcare spending, such as prescription drug spending (e.g. see Gaskin et al. 2006).

Numerous researchers have documented the health cost risks that older Americans face in any given year, and how it varies by characteristics such as marital status and gender (e.g., McGarry and Schoeni 2005; Goda, Shoven, and Slavov 2013; Cylus et al. 2011). De Nardi et al. (2016) show that total spending is, on average, about $1,100 per year more for older women than for older men, although they attribute this to higher expenditures on nursing homes for women. Goda, Shoven, and Slavov (2013) show that about one-third of the gender difference in total out-of-pocket medical spending can be explained by differences in the rates of widowhood.

Our work extends this literature in several important ways. To the best of our knowledge we construct the most comprehensive panel to date of medical spending for the age 65+ population in the US. Most papers attempting to measure medical spending are limited in which types of spending they observe, and often focus on only out-of-pocket spending (e.g., Feenberg and Skinner 1994, French and Jones 2004, Fahle, McGarry, and Skinner 2016, Hurd, Michaud, and Rohwedder 2017). Many papers also use only cross-sectional data or have short panels (e.g. Escarce and Kapur 2003). In contrast, we use nationally-representative, longitudinal data from the Health and Retirement Study (HRS), which allow us to track households up to 14 years. Our main data set consists of out-of-pocket spending from the public HRS data linked to Medicare and Medicaid spending from administrative records. To this we add data from the Medical Expenditure Panel Survey (MEPS), which allow us to impute private insurance and other payments not measured directly in the HRS. This yields an estimate of medical spending that accounts for all spending, disaggregated by payor, which makes it possible to better understand the roles of both public and private insurance coverage.

We exploit the long panel and state of the art estimation techniques to estimate total costs over remaining life. By focusing on dynamics, we consider not only the risks in a single year, but also the risk of moderate but persistent expenses that accumulate into a catastrophic lifetime cost, and whether these risks differ by race and gender. We build on our earlier work in Arapakis et al. (2021), to model lifetime medical spending shocks using the semi-parametric framework developed by Arellano, Blundell, and Bonhomme (2017). The "ABB" framework allows for non-Gaussian shocks and persistence to the shocks. We extend our previous work by allowing the persistence of the shocks to not only vary by age and the size of the shocks, but also by race and gender. Using such a framework results in a more nuanced understanding of how medical spending risk differs by race and gender. To the best of our knowledge, all papers that have measured racial differences in spending have focused on spending per year, and not remaining lifetime spending. Because, on average, white individuals live longer than Black and Hispanic individuals, focusing on annual spending may understate lifetime differences.

Another key strength of our spending measure is that its link to the rich demographic, medical and financial data found in the HRS, which makes us well placed to measure differences by race and gender. We also build on the current literature by not just estimating how medical spending varies with race and gender but also examining the extent to which these differences can be explained by other observable correlates like education and income.

Our paper is also tied to a larger literature on retirement savings adequacy by race. It is well known that racial minorities often experience worse health than white individuals (Blundell et al., 2022) and thus have greater healthcare needs. Consistent with this view, we find that the worse health experienced by minorities leads to modestly higher predicted medical spending. There is also a growing literature on other dimensions of financial well being. For example, racial minorities have lower income and assets in retirement than their white counterparts (Choukhmane et al., 2022). As we show, the lower income and wealth of minorities increases their Medicaid recipiency, which reduces their burden of medical spending in retirement.

3 Data and Descriptive Statistics

Our medical spending measure begins with HRS data matched with administrative Medicare and Medicaid records.[1] We then use MEPS data to impute spending by payors missing in the HRS, giving us a version of HRS medical spending data that is representative of all payors. We use data from 1999-2012, for which we also have the Medicare and Medicaid records.

3.1 The HRS

The HRS is a representative biennial survey of the population ages 51 and older, and their spouses. Although drawn from the non-institutionalized population when first interviewed, these individuals are tracked and reinterviewed as they enter nursing homes and other institutions.[2] Households are followed until every member dies; attrition for other reasons is low. When a sample member dies, in the next wave an "exit" interview with a knowledgeable party – usually another family member – is conducted. This allows the HRS to collect data on end-of-life medical conditions and spending.

To focus on the Medicare population, we restrict the sample to those ages 65 and older. Consistency with our demographic model also leads us to drop a small number of households who, for example, are "partnered" or whose partner reports conflicting marital status. We also drop households whose head is not white, Black or Hispanic (such as Asian-Americans or Native Americans) because the number of households in this group is too small for meaningful statistical analysis. We further drop households who do not consent to have their data linked to the administrative records. Finally, we restrict ourselves to data in the period 1999-2012, the years for which we have both Medicare and Medicaid data. This leaves us with 4,391 households comprising 45,130 household-year observations. Appendix A documents our sample selection criteria.

The HRS has a wide variety of health indicators. We assign individuals to the nursing home state if they were in a nursing home at least 120 days since the last interview or if they were in a nursing home at least 60 days and died before the next scheduled interview. Our focus is on longer and more expensive stays: as Friedberg et al. (2015) and Hurd, Michaud, and Rohwedder (2017) document, many nursing home stays last only a few weeks and are associated with lower expenses. We assign the remaining individuals a health status of "good" if their self-reported health is excellent, very good or good and a health status of "bad" if their self-reported health is fair or poor.

The HRS collects data on all out-of-pocket medical expenses, including drug costs and nursing home

[1] This section follows closely the description provided in Arapakis et al. (2021).

[2] French and Jones (2004) show that the HRS sample matches very well the aggregate statistics on the share of the older population in a nursing home by 1999, when our sample begins. At that point, roughly 2.8% of men and 6.1% of women in the data had entered nursing homes.

care. The HRS medical spending measure is backward-looking: medical spending at any wave is measured as total out-of-pocket spending over the preceding two years. End-of-life expenditures are included; Fahle, McGarry, and Skinner (2016), who compare the medical spending data from the "core" and exit interviews, show that out-of-pocket spending rises significantly in the last year of life. French, Jones, and McCauley (2017) compare out-of-pocket medical spending data from the HRS, MEPS, and the Medicare Current Beneficiary Survey (MCBS). They find that the HRS matches up well with the MEPS for items that MEPS covers, but that the HRS covers more items.

To control for socioeconomic status, we follow De Nardi et al. (2023) and construct a measure of lifetime earnings or "permanent income" (PI). We first find each household's "non-asset" income, a pension measure that includes Social Security benefits, defined benefit pension benefits, veterans benefits and annuities. Because there is a roughly monotonic relationship between lifetime earnings and pension income, non-asset income is a good proxy for permanent income. We then use a fixed-effects regression to convert non-asset income, which depends on age and household composition as well as lifetime earnings, to create a scalar measure that ranks all households. In particular, we assume that the log of non-asset income for household i at age t follows

$$\ln y_{it} = \alpha_i + \kappa(t, f_{it}) + \omega_{it}, \tag{1}$$

where: α_i is a household-specific effect; $\kappa(t, f_{it})$ is a flexible function of age and family structure f_{it} (i.e., couple, single man, or single woman); and ω_{it} represents measurement error. The percentile ranks of the estimated fixed effects, $\widehat{\alpha}_i$, form our measure of permanent income, $\widehat{I}_i$.

3.2 Administrative Medicare and Medicaid Records

The Centers for Medicare and Medicaid Services (CMS) have confidential administrative spending records for Medicare and Medicaid that we link to the survey responses of consenting HRS respondents. Linking these data to the HRS results in a broad set of spending measures for the years 1999-2012 which we use in our main analysis.

The Medicare records include reimbursement amounts for all services covered by traditional fee-for-service Medicare (Parts A and B) and drug-related spending made under Part D, which began in 2006. The records exclude, however, expenditures made on behalf of individuals enrolled in Medicare Advantage (Part C) plans.[3] As described in the next section, we impute Part C payments using MEPS data. The Medicaid files contain information on enrollment, service use, and spending and includes Medicaid HMO

[3] The fraction of Medicare beneficiaries enrolled in a Medicare Advantage plan has risen from 18% in 1999 to 31% in 2016 (Jacobson et al., 2016).

payments.[4] Appendix B describes the Medicare and Medicaid data in more detail.

3.3 Imputations Using MEPS Data

While the HRS contains accurate measures of out-of-pocket medical spending and can be linked to Medicare and Medicaid records, it does not contain Medicare Part C expenditures or payments made by other payors such as private and smaller public insurers (such as the Veterans Administration and care provided by state and local health departments). To circumvent this issue, we use data from the 1996-2017 waves of the Medical Expenditure Panel Survey (MEPS) to impute these missing payments. The MEPS is a nationally representative survey of non-institutionalized households. MEPS respondents are interviewed up to 5 times over a 2 year period, forming short panels, which we aggregate to an annual frequency.[5]

To impute medical spending not captured in the HRS, we proceed in two steps. First, we use the MEPS data to regress payments not captured in the HRS (e.g., Medicare Part C and "other payors" including private care, Veterans Administration and care provided by state and local health departments) on a set of observable variables found in both data sets: household income, a fourth order age polynomial, labor force participation status, education, marital status, doctor and hospital visits, race indicators, health measures, out-of-pocket spending and interactions. This regression has an R^2 statistic of 0.13 for private insurance payments and 0.02 for other payors. Second, we impute these expenses in the HRS data using a conditional mean-matching procedure, a procedure very similar to hot-decking. This procedure applies the estimated MEPS regression coefficients to the HRS data, yielding predicted values for each HRS household, to which we add residuals drawn from MEPS households with similar levels of predicted spending. We describe our approach in more detail in Appendix C.

3.4 Descriptive Statistics

The first two columns of Table 1 describe the distribution of total medical spending for white, Black and Hispanic households, where household race is defined as the race of the male in two-person households. White households spend on average \$22,100 per year, \$1,400 more than Black households and \$2,400 more than Hispanic households. White households in the top 5% of their distribution spend, on average, \$137,300. This is very similar to the spending of Black and Hispanic households at the tops of their respective distributions. Racial differences in spending are more pronounced in the bottom half of the

[4] Individual states report Medicaid payments different ways, with some states reporting the value of services received for each beneficiary and other states reporting the premium paid to the HMO. In all cases we use the reported amounts and do not make adjustments.

[5] The survey responses are matched to medical spending information provided by health care providers. Although the MEPS does not record certain types of medical expenditures, such as nursing home expenditures, it captures sources of medical spending extremely well. Pashchenko and Porapakkarm (2016) compare MEPS data to the aggregate statistics and show that it captures most types of spending very well.

distribution, where white households on average spend 28% more than Black households and 42% more than Hispanic households.

Table 1: Annual Medical Spending by Race ($2014)

Spending Percentile	Average Spending	Percentage of Total	Percentage paid by: Out-of-pocket	Medicare	Medicaid	Other
Panel A. White Households						
All	22,100	100.0	28.9	59.1	3.9	8.1
95-100	137,300	31.0	28.5	51.8	11.6	8.2
0-50	3,700	8.4	35.3	55.0	1.5	8.2
Panel B. Black Households						
All	20,700	100.0	21.9	61.8	10.6	5.7
95-100	136,000	32.8	8.7	67.2	19.8	4.2
0-50	2,900	7.1	30.2	56.4	6.6	6.8
Panel C. Hispanic Households						
All	19,700	100.0	16.5	61.8	17.5	4.2
95-100	134,000	33.8	7.7	68.8	21.9	1.6
0-50	2,600	6.7	23.0	61.7	10.2	5.0

Notes: "Other" includes private insurance, smaller government insurers, and other forms of spending. "Medicare" includes Medicare Part C, and "Medicaid" includes Medicaid HMOs. Spending percentiles are defined within racial groups. Race is defined as the race of the male in two-person households.

The final four columns of Table 1 present the percentage paid out of pocket, by government sources (Medicare and Medicaid, including both fee-for-service and HMOs), and by "other" payors. There are clear racial differences in who pays for medical spending. Whereas 28.9% of all spending by white households is out of pocket, the percentage falls to 21.9% and 16.5% for Black and Hispanic households, respectively. These differences are even more striking among the top 5%, with white households paying 28.5% out of pocket, while Black and Hispanic households pay 8.7% and 7.7%, respectively. Black and Hispanic households thus face less catastrophic medical expense risk. Relative to white households, Black and Hispanic households rely more heavily on public payors, and Medicaid in particular. In contrast, white households rely more heavily on "other" payors, which are primarily private insurers.

Table 2: Summary Statistics

	By Race			By Household Status		
	White	Black	Hispanic	Single Woman	Single Man	Couple
Age	80.85	78.03	77.18	80.20	78.85	79.74
	(7.80)	(8.42)	(8.36)	(8.24)	(8.21)	(6.31)
Currently Married	0.22	0.07	0.08	0.00	0.00	1.00
	(0.41)	(0.26)	(0.28)	(0.00)	(0.00)	(0.00)
Health Status: Good	0.61	0.46	0.37	0.56	0.55	0.66
	(0.49)	(0.50)	(0.48)	(0.50)	(0.50)	(0.47)
Health Status: Bad	0.31	0.47	0.59	0.36	0.37	0.32
	(0.46)	(0.50)	(0.49)	(0.48)	(0.48)	(0.47)
In a Nursing Home	0.07	0.06	0.04	0.07	0.07	0.02
	(0.26)	(0.24)	(0.19)	(0.26)	(0.26)	(0.14)
Household Member Just Died	0.08	0.07	0.06	0.07	0.10	0.10
	(0.27)	(0.25)	(0.24)	(0.25)	(0.30)	(0.30)
Education (Years)	12.00	9.98	7.95	11.25	11.14	13.47
	(3.64)	(3.83)	(4.02)	(3.48)	(3.92)	(3.04)
Household Income (1000s)	48.93	22.46	17.56	25.75	38.61	111.07
	(84.81)	(29.23)	(19.69)	(32.02)	(101.47)	(118.68)
Household Wealth (1000s)	561	93.7	108	241	380	1,257
	(1,513)	(238)	(267)	(531)	(1,293)	(2,516)
PI Percentile	0.61	0.42	0.32	0.53	0.57	0.69
	(0.24)	(0.27)	(0.26)	(0.26)	(0.28)	(0.21)
Receiving Medicaid	0.17	0.35	0.46	0.25	0.21	0.08
	(0.37)	(0.48)	(0.50)	(0.43)	(0.41)	(0.28)
On Private Insurance	0.26	0.08	0.07	0.21	0.18	0.32
	(0.44)	(0.27)	(0.25)	(0.41)	(0.38)	(0.47)
On Medicare Part C	0.22	0.27	0.31	0.23	0.22	0.25
	(0.41)	(0.44)	(0.46)	(0.42)	(0.41)	(0.44)
Observations	35,171	7,425	2,534	27,783	6,605	8,452

Notes: Health status is self-reported in the HRS: we code reports of "Good", "Very Good", or "Excellent" as "Good" health, and code reports of "Fair" or "Poor" health as "Bad". Health, age and education values for couples are those of the husband. PI percentile is based on a household fixed effect measuring how much income a household earns compared to what their observed characteristics (excluding race) would predict.

The left panel of Table 2 provides more context for these findings, as it presents racial differences in potential drivers of medical spending and its breakdown by payors. On average, white households are older and are more likely to be married, both of which are associated with higher medical spending. White households are in better health, which is associated with lower spending, but are more likely to be in a nursing home, which is associated with higher spending. White households also have higher education, income and wealth. All of this leaves them less likely to qualify for Medicaid, which is means-tested: while 35% of Black households and 46% of Hispanic households receive Medicaid, only 17% of white households receive it.

Table 3: Annual Medical Spending by Household Status ($2014)

Spending Percentile	Average Spending	Percentage of Total	Percentage paid by: Out-of-pocket	Medicare	Medicaid	Other
Panel A. Single Woman						
All	20,300	100.0	26.5	60.1	6.9	6.5
95-100	132,600	32.6	25.6	53.7	14.7	6.0
0-50	3,000	7.5	33.1	56.5	3.5	6.9
Panel B. Single Man						
All	20,700	100.0	26.1	58.1	5.7	10.0
95-100	135,800	32.7	21.9	55.6	15.2	7.3
0-50	2,500	6.1	34.5	51.6	2.6	11.3
Panel C. Married Couple						
All	27,700	100.0	28.9	59.7	1.8	8.6
95-100	151,700	27.3	21.1	60.3	8.9	9.7
0-50	7,200	13.1	36.7	55.1	0.3	7.9

Notes: "Other" includes private insurance, smaller government insurers, and other forms of spending. "Medicare" includes Medicare Part C, and "Medicaid" includes Medicaid HMOs.

Table 3 repeats the analysis of Table 1, this time distinguishing between single women, single men and couples. The spending distributions for single women and men are quite similar, with single men relying slightly more on other payors. The spending of couples is rather different. The average spending of couples, $27,700, is far less than the combined average spending of single women and men, $41,000. This difference is especially pronounced among the top 5%, where married households spend only about 10% more than single ones. In contrast, at the bottom half of the distribution, the spending of couples is more than double than of either type of single. Table 3 also shows that married households rely less heavily on Medicaid than singles.

The right panel of Table 2 shows how the potential spending drivers differ by household status. Couples are richer (in both income and wealth terms) and more educated than singles. Not surprisingly, only 8% of couples receive Medicaid, in contrast to 21-25% of singles. Couples are also less likely to be in a nursing home.[6] This may reflect differences in health or, alternatively, the ability of one spouse to care for the other. Because nursing homes are particularly expensive, the lower usage of couples is a likely reason why the top 5% of couples spend far less on a per capita basis than comparable singles.

4 Decomposing Medical Spending

To quantify how much of the Black-white differences shown in Table 1 are attributable the observable covariates in Table 2, we use the decomposition introduced by Gelbach (2016). In our application, this involves comparing the coefficient on a race indicator in a "baseline" univariate regression to the coefficient on race in a "full" multivariate regression:

$$Y_{i,t} = \beta_{0,base} + \beta_{1,base} Black_i + e_{i,t}, \quad (2)$$

$$Y_{i,t} = \beta_{0,full} + \beta_{1,full} Black_i + X'_{i,t}\beta_{X,full} + u_{i,t}, \quad (3)$$

where: $Y_{i,t}$ is the outcome of interest (either household medical spending or fraction spent out of pocket) for household i at age t; $Black_i$ is an indicator for whether the head of household is Black (instead of white); and $X_{i,t}$ is the vector of covariates that potentially explain racial differences. The covariates include demographics (household structure and age), health, education, income and a set of region dummies. The unexplained parts of $Y_{i,t}$ in the two regressions are captured by $e_{i,t}$ and $u_{i,t}$,

In the Gelbach Decomposition, the difference between the two race coefficients, $\beta_{1,base} - \beta_{1,full}$, is the portion of the white-Black spending gap attributable to the additional observable factors. The decomposition allows us to allocate this difference among each of the observable factors. Because all the factors are added simultaneously to the "full" regression, Gelbach's method decomposes the gap in a way that does not depend on the order in which the factors enter the decomposition.[7] The decomposition for white-Hispanic differences follows the same procedure, but replaces the indicator for Black with an indicator for Hispanic.

Table 4 shows the decomposition for total medical spending. The left panel is for the Black-white spending difference and the right panel is for the Hispanic-white difference. The top row presents the coefficients on race. The first column in each panel (Base) presents the race coefficient from the univariate

[6] In making this comparison, it bears noting that the health indicators for couples are those of the man.

[7] The Gelbach decomposition relies on the formula for omitted variable bias:

$$\beta_{1,base} - \beta_{1,full} = \sum\nolimits_{j \in X} \gamma_{X_j,black}\beta_{X_j,full},$$

where $\gamma_{X_j,black}$ is the coefficient from a univariate regression of the jth element of $X_{i,t}$ on $Black_i$ and $\beta_{X_j,full}$ is the jth element of the coefficient vector $\beta_{X,full}$. (See equation (3).) The effect of the jth covariate on the gap is thus given by $\gamma_{X_j,black}\beta_{X_j,full}$.

regression. This coefficient is simply the mean difference in spending by race; for Black and white households the difference is $\hat{\beta}_{1,base} = -\$1{,}407$. The next column (Full) displays the race coefficient ($\hat{\beta}_{1,full}$) from the regression that includes the vector of covariates in equation (3). This shows that covariates can explain all but $223 of the spending difference between white and Black households.

Table 4: Gelbach Decomposition of Total Medical Spending

	White vs. Black Households			White vs. Hispanic Households		
	Specification			Specification		
	Base	Full	Explained	Base	Full	Explained
Race	-1,407***	-223	-1,184***	-2,439***	-195	-2,244***
	(429)	(417)	(417)	(678)	(662)	(436)
Covariates						
Household Structure and Age	No	Yes	-1,126***	No	Yes	-1,044***
			(147)			(164)
Health	No	Yes	1,345***	No	Yes	1,059***
			(257)			(390)
Education	No	Yes	-1,789***	No	Yes	-3,112***
			(159)			(242)
Income	No	Yes	585***	No	Yes	1,172***
			(128)			(203)
Region	No	Yes	-198***	No	Yes	-319***
			(75)			(102)
N	45,211	45,211	45,211	40,044	40,044	40,044
adj. R^2	0.000	0.224		0.000	0.228	

Notes: Gelbach (2016) decomposition. Standard errors in parentheses. The "explained difference" is $\hat{\beta}_{1,base} - \hat{\beta}_{1,full}$, where $\hat{\beta}_{1,base}$ and $\hat{\beta}_{1,full}$ are the coefficients on race in the "base" and "full" regressions (equations (2) and (3)). This difference is decomposed into the parts explained by each race-related covariate. For example, the part explained by education is $\hat{\beta}_{ed,full}\hat{\gamma}_{ed}$, where $\hat{\gamma}_{ed}$ is the coefficient from an auxiliary regression of education on race and $\hat{\beta}_{ed,full}$ is the coefficient on education in the full regression.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The final column (Explained) in the left panel displays how much of the base Black-white gap is attributable to the covariates, namely $\hat{\beta}_{1,base} - \hat{\beta}_{1,full} = -\$1{,}184$. Household structure and age explain a large proportion of this difference: Black households spend $1,126 less than their white counterparts because they are more likely to be single and younger.[8] Health also plays a large role, but in the opposite direction: health differences imply an additional $1,345 of spending by Black households, as they are more likely to be in bad health, although the percentage in nursing homes is similar. For education, Black

[8] 22% of white households are married versus only 7% of Black households. The average age for white households is 80.7 versus 77.9 for Black households).

households are less likely to be college graduates, which translates into $1,789 of lower spending. The impacts of region and income are statistically significant but relatively small.

The second panel in Table 4 decomposes the $2,439 difference in spending between white and Hispanic households. Again, the covariates we consider explain most of the gap. The same differences that explain the Black-white gap also explain the Hispanic-white gap, although education and income appear to matter more for Hispanic-white gap.

Table 5: Gelbach Decomposition of Out-of-Pocket Spending Fraction

	White vs. Black Households			White vs Hispanic Households		
	Specification			Specification		
	Base	Full	Explained	Base	Full	Explained
Race	-0.070***	-0.036***	-0.034 ***	-0.125***	-0.060***	-0.064***
	(0.004)	(0.004)	(0.001)	(0.006)	(0.006)	(0.003)
Covariates						
Household Structure and Age	No	Yes	0.009***	No	Yes	0.006***
			(0.002)			(0.002)
Health	No	Yes	-0.013***	No	Yes	-0.019***
			(0.002)			(0.003)
Education	No	Yes	-0.009***	No	Yes	-0.019***
			(0.002)			(0.002)
Income	No	Yes	-0.022***	No	Yes	-0.034***
			(0.001)			(0.002)
Region	No	Yes	0.000	No	Yes	0.001
			(0.001)			(0.001)
N	44,506	44,506	44,506	39,412	39,412	39,412
adj. R^2	0.008	0.087		0.012	0.089	

Notes: Gelbach (2016) decomposition. Standard errors in parentheses. The "explained difference" is $\hat{\beta}_{1,base} - \hat{\beta}_{1,full}$, where $\hat{\beta}_{1,base}$ and $\hat{\beta}_{1,full}$ are the coefficients on race in the "base" and "full" regressions (equations (2) and (3)). This difference is decomposed into the parts explained by each race-related covariate. For example, the part explained by education is $\hat{\beta}_{ed,full}\hat{\gamma}_{ed}$, where $\hat{\gamma}_{ed}$ is the coefficient from an auxiliary regression of education on race and $\hat{\beta}_{ed,full}$ is the coefficient on education in the full regression.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In Table 5 we perform a similar exercise as in Table 4, but now consider differences in the fraction of total medical spending that is paid out of pocket. Table 5 shows that the covariates we consider can explain around half of both the Black-white and Hispanic-white gap. The most important covariates in both cases are health and income.

5 The Model

In the previous two sections, we analyzed differences in annual expenditures. We now turn to lifetime medical spending, the discounted sum of the expenditures that older households make over their remaining lifetimes. Because this sum cannot be estimated directly from the data, we rely on simulations. In particular, we estimate the distribution of lifetime medical spending in five steps. In the first step, we estimate the log of total medical spending as a function of age, health, family structure and PI, using Ordinary Least Squares (OLS). In the second step, we estimate the stochastic process for the unexplained component of medical spending – the residuals from the first step regression – using the methodology developed by Arellano, Blundell, and Bonhomme (2017). In the third step, we estimate the mapping from total medical spending to out-of-pocket medical spending. In the fourth step, we estimate a Markov Chain model of health and mortality. In the final step, we use the estimated models to simulate health, mortality, and lifetime medical spending.

5.1 Total Medical Spending

Let $M_{i,t}$ denote total medical spending for household i at time t, and let $m_{i,t}$ denote its logarithm net of the observed variables contained in the vector $X_{i,t}$:

$$\ln M_{i,t} = X'_{i,t}\gamma + m_{i,t}. \tag{4}$$

In practice the vector $X_{i,t}$ includes an age polynomial, health indicators, household structure and death-year indicators, PI percentile, and interactions among the aforementioned variables. We assume that $m_{i,t}$ can be expressed as the sum of the persistent component $\eta_{i,t}$ and the transitory component $\varepsilon_{i,t}$:

$$m_{i,t} = \eta_{i,t} + \varepsilon_{i,t}, \quad \forall i \in \{1, ..., N\}, \forall t \in \{1, ..., T\}. \tag{5}$$

We assume that the persistent component $\eta_{i,t}$ follows a first-order Markov process. The transitory component $\varepsilon_{i,t}$ is normalized to be zero-mean and is assumed to be independent over time and independent of $\eta_{i,t}$; and both components are mean-independent of $X_{i,t}$. We otherwise place no restrictions on either component beyond the technical regularity conditions given in Arellano, Blundell, and Bonhomme (2017). This model nests both the "Gaussian AR(1) plus white noise" specification used by Feenberg and Skinner (1994), French and Jones (2004) and French, Jones, and McGee (2023) and the Markov chain used by Kopecky and Koreshkova (2014). Arapakis et al. (2021) shows that this model fits the data considerably better than either of these more restricted models.

We estimate the process for $\eta_{i,t}$ using the quantile-based framework proposed by Arellano, Blundell, and Bonhomme (2017) and extended by Arellano et al. (2023).[9] To apply this framework, rewrite the conditional distribution for the persistent component $\eta_{i,t}$ as:

$$\eta_{i,t} = Q_\eta(\nu_{i,t} \,|\, \eta_{i,t-1}, a_{i,t}), \quad \nu_{i,t} \overset{iid}{\sim} U[0,1], \tag{6}$$

where $Q_\eta(\nu \,|\, \eta_{i,t-1}, a_{i,t})$ denotes the νth quantile of $\eta_{i,t}$ conditional on its lagged value and age ($a_{i,t}$). The quantile function Q_η maps $\eta_{i,t}$'s conditional rank, $\nu_{i,t}$, into a value of $\eta_{i,t}$ itself. To fix ideas, if we draw $\nu_{i,t} = 0.1$, the realized value of $\eta_{i,t}$ will equal the 10th percentile of the conditional distribution of $\eta_{i,t}$ at age $a_{i,t}$ and lagged value $\eta_{i,t-1}$. As a rank, $\nu_{i,t}$ is distributed uniformly over the $[0,1]$ interval.

In its most unrestricted form, this specification allows for a great degree of flexibility. One way to see this is to construct the persistence measure

$$\phi_\tau(\eta_{i,t-1}, a_{i,t}) = \frac{\partial Q_\eta(\tau \,|\, \eta_{i,t-1}, a_{i,t})}{\partial \eta_{i,t-1}}, \tag{7}$$

with τ denoting the conditional rank of interest. $\phi_\tau(\eta_{i,t-1}, a_{i,t})$ measures the effect of $\eta_{i,t-1}$ on the τth conditional quantile of $\eta_{i,t}$. Persistence can vary by rank (τ), age ($a_{i,t}$) and prior realization ($\eta_{i,t-1}$). In contrast, in the standard AR(1) model, persistence always equals the constant ϕ.

In estimation we parametrically approximate the conditional quantile function by low-order Hermite polynomials. Let $h_k^\eta(\cdot)$ denote the kth Hermite polynomial used in the approximation of $\eta_{i,t}$, with $\{h_k^\eta(\cdot)\}_{k=0}^{K_\eta}$ forming the polynomial basis for the approximation. $Q_\eta(\tau \,|\, \eta_{i,t-1}, a_{i,t})$ is thus a linear combination of the K_η Hermite polynomials, with the coefficients on the polynomials, $\{\beta_k^\eta(\tau)\}_{k=0}^{K_\eta}$ themselves functions of the quantile rank τ.

We thus have

$$Q_\eta(\tau \,|\, \eta_{i,t-1}, a_{i,t}) = \sum_{k=0}^{K_\eta} \beta_k^\eta(\tau) \cdot h_k^\eta(\eta_{i,t-1}, a_{i,t}), \;\; \tau \in (0,1]). \tag{8}$$

The distributions of the initial shock $\eta_{i,1}$ and the transitory shocks $\{\varepsilon_{i,t}\}_t$ are handled in ways analogous to how we handle the persistent component:

$$Q_1(\tau \,|\, a_{i,1}) = \sum_{k=0}^{K_1} \beta_k^1(\tau) \cdot h_k^1(a_{i,1}), \;\; \tau \in (0,1), \tag{9}$$

$$Q_\varepsilon(\tau \,|\, , a_{i,t}) = \sum_{k=0}^{K_\varepsilon} \beta_k^\varepsilon(\tau) \cdot h_k^\varepsilon(a_{i,t}), \;\; \tau \in (0,1). \tag{10}$$

For these distributions we do not condition on $\eta_{i,t-1}$ but only only on age.

Each of the coefficient functions $\left(\{\beta_k^\eta(\tau)\}_{k=0}^{K_\eta}, \{\beta_k^1(\tau)\}_{k=0}^{K_1}, \{\beta_k^\varepsilon(\tau)\}_{k=0}^{K_\varepsilon}\right)$ in equations (8)-(10) is mod-

[9] This section follows closely the notation and language used in Arapakis et al. (2021).

elled with a set of polynomial splines defined over the intervals $\{[\tau_{\ell-1}, \tau_\ell]\}_{\ell=1}^{L}$, along with two low-dimensional tail functions defined over $(0, \tau_1]$ and $[\tau_L, 1)$. It is the parameters for these weighting functions that we must estimate.

As both the persistent and transitory shocks are unobserved, we cannot estimate the parameters of the weighting functions directly using quantile regressions. Furthermore, our data are unbalanced because sample members die. We therefore follow the extension of the *E-M* algorithm described in and applied by Arellano et al. (2023).

- In the *E-step* we find the posterior distribution of the unobserved persistent shocks ($\{\eta_{i,t}\}_t$) implied by the data and the current parameterization of the model. In particular, we use the coefficients of the Hermite polynomials $\left(\{\beta_k^\eta(\tau)\}_{k=0}^{K_\eta}, \{\beta_k^1(\tau)\}_{k=0}^{K_1}, \{\beta_k^\varepsilon(\tau)\}_{k=0}^{K_\varepsilon}\right)$, which fully determine the distributions of the shocks, and a Monte Carlo method to simulate draws from the distributions of the initial shock $\eta_{i,1}$ and the subsequent shocks $\{\eta_{i,t}\}_t$.[10] This part of the procedure is a special case of the Sequential Monte Carlo methods described in greater detail in Creal (2012).

- In the *M-step* we use quantile regressions to update the coefficient functions for the Hermite polynomials, using the distribution of $\{\eta_{i,t}\}_t$ found in the *E-step*. Once the coefficients have been updated, we return to the *E-step* and simulate new draws.

We iterate between the *E* and *M* steps until the parameters converge. See Appendix E for a more detailed description of the methodology.

5.2 Health and Mortality

Let $hs_{i,g,t}$ denote the health of member $g \in \{h, w\}$ in household i at age t. Health has four mutually exclusive possible values: dead; in a nursing home; in bad health; or in good health. We assume that the transition probabilities for an individual's health depend on his or her current health, age, permanent income I, gender g and race, $race_i$.[11] It follows that the elements of the health transition matrix are given by

$$\pi_{q,r}\big(t, I_i, g, race_i\big) = \Pr\left(hs_{i,g,t+1} = r \mid hs_{i,g,t} = q; t, I_i, g, race_i\right), \tag{11}$$

with the transitions covering a one-year interval. Although the HRS interviews every other year, we adopt the approach in De Nardi, French, and Jones (2016), who fit annual models of health to the HRS data for

[10] This approach takes advantage of the Markovian structure of the model and has been shown to perform well in low-dimensional models.

[11] We do not allow health transitions to depend on medical spending. The empirical evidence on whether medical spending improves health, especially at older ages, is surprisingly mixed (De Nardi, French, and Jones 2016). Likely culprits include reverse causality – sick people have higher expenditures – and a lack of insurance variation – almost every retiree gets baseline insurance through Medicare. In addition, we do not allow health transitions to depend on marital status. De Nardi et al. (2023) find that after controlling for income and past health, marital status has little added predictive power.

singles. We estimate health/mortality transition probabilities by fitting the transitions observed in the HRS to a multinomial logit model.[12] See Appendix F for further details.

5.3 Medical Spending Budget Sets

The model developed in this section is for total medical spending. To infer the portion of medical spending that is paid out of pocket by households, we construct budget sets that map total medical spending into out-of-pocket shares, and apply them to the total medical spending values generated by the model.

We estimate these budget sets in two steps for each racial group. First, using a logit model, we estimate the probability of Medicaid receipt as a function of: age, marital status and gender, the health and nursing home status of the husband and wife, the log of total medical spending, and medical spending interacted with PI. Second, using OLS we estimate the share of total medical spending paid out of pocket, using the same variables used to estimate Medicaid recipiency. We estimate separate share functions for households receiving and not receiving Medicaid. Once we have used the logit model to determine whether a particular household is receiving Medicaid, we then apply the appropriate budget set to calculate out-of-pocket spending.

6 Results

6.1 Health and Longevity

We estimate health/mortality transition probabilities by fitting the transitions observed in the HRS to a multinomial logit model, where the variables include a cubic in age, age quadratic in PI, gender, health indicators (good health, bad health, nursing home), interactions, race indicators and race-age interactions. To calculate life expectancy, we then take the joint distribution of all variables, then simulate life histories of health and survival using the estimated logit model.

Table 6 shows the life expectancies implied by our demographic model for white, Black and Hispanic individuals still alive at age 65. The first panel of the table shows life expectancies for white individuals under different configurations of gender, PI percentile, and age-65 health. The healthy live longer than the sick, the rich (higher PI) live longer than the poor, and women live longer than men. For example, a white man at the 10th PI percentile in a nursing home expects to live only 3.8 more years, while a single woman at the 90th percentile in good health expects to live 21.3 more years.

[12] We do not control for cohort effects. Instead, our estimates are a combination of period (cross-sectional) and cohort probabilities. This may lead us to underestimate the life spans expected by younger cohorts as they age. Nevertheless, life spans have increased only modestly over the sample period. Accounting for cohort effects would have at most a modest effect on our estimates.

The second panel of the table shows that at any PI, gender and age-65 health combination, a Black individual expects to live about one year less than his or her white counterpart. For example, a healthy Black man at the 90th percentile will on average live 1.3 years less (17.0 vs. 18.3) than an otherwise identical white man. In contrast, once gender, income and health are controlled for, Hispanic individuals live longer than white individuals.

Table 6: Life expectancy in years, conditional on reaching age 65

	Men			Women			
Permanent Income Percentile	Nursing Home	Bad Health	Good Health	Nursing Home	Bad Health	Good Health	All
White Individuals							
10	3.8	13.3	15.8	5.4	16.2	18.1	16.0
50	4.5	15.8	18.3	6.6	18.9	20.8	19.0
90	4.2	15.7	18.3	6.3	19.2	21.3	19.2
Black Individuals							
10	3.5	12.2	14.6	4.9	15.2	17.1	14.9
50	4.0	14.4	17.0	5.8	17.7	19.8	17.5
90	3.6	14.2	17.0	5.4	17.8	20.1	17.8
Hispanic Individuals							
10	5.0	15.2	16.7	7.3	18.2	19.3	17.0
50	6.0	17.2	18.8	8.7	20.3	21.6	19.4
90	5.5	16.9	18.7	8.2	20.3	21.7	19.3
All Men							
White							17.0
Black							14.8
Hispanic							16.4
All Women							
White							19.8
Black							18.0
Hispanic							19.6

Note: Life expectancies calculated through simulations using estimated health transition and survivor functions.

The next two panels show life expectancies for men and women by race, once we average over health and PI.[13] As Table 2 shows, Black households have lower PI rankings than white households. Accounting for this difference increases the Black-white life expectancy gap to 2.2 years for men and 1.8 years for women. Similarly, a lower average PI level leads Hispanic households to have a lower unconditional life expectancy than white households, even though they live longer at every combination of health and PI.

Another key statistic for our analysis is the probability that a 65 year old will spend significant time

[13] We construct these distributions with bootstrap draws of people aged 63-67 in the HRS.

(a stay of more than 120 days) in a nursing home before he or she dies. Table 7 shows our estimates. We find that 29.0% of white men and 42.5% of white women will have an extended nursing home stay before they die. These probabilities fall modestly, to 27.7% and 38.8% for Black men and women, and more substantially, to 22.3% and 30.9% for Hispanic households. The disaggregated results show that nursing home incidence differs relatively modestly across the PI distribution or by health. Although healthy and high-income people are less likely to be in a nursing home at any given age, they live longer, and older individuals are much more likely to be in a nursing home. For the same reason, some of the higher nursing home use of white households is attributable to their longer life expectancy.

Table 7: Probability of ever entering a nursing home, conditional on being alive at age 65, in percent

	Men	Women
White Individuals	29.0	42.5
Black Individuals	27.7	38.8
Hispanic Individuals	22.3	30.9

Note: Probabilities calculated through simulations using estimated health transition and survivor functions.

6.2 Medical Spending and its Financing

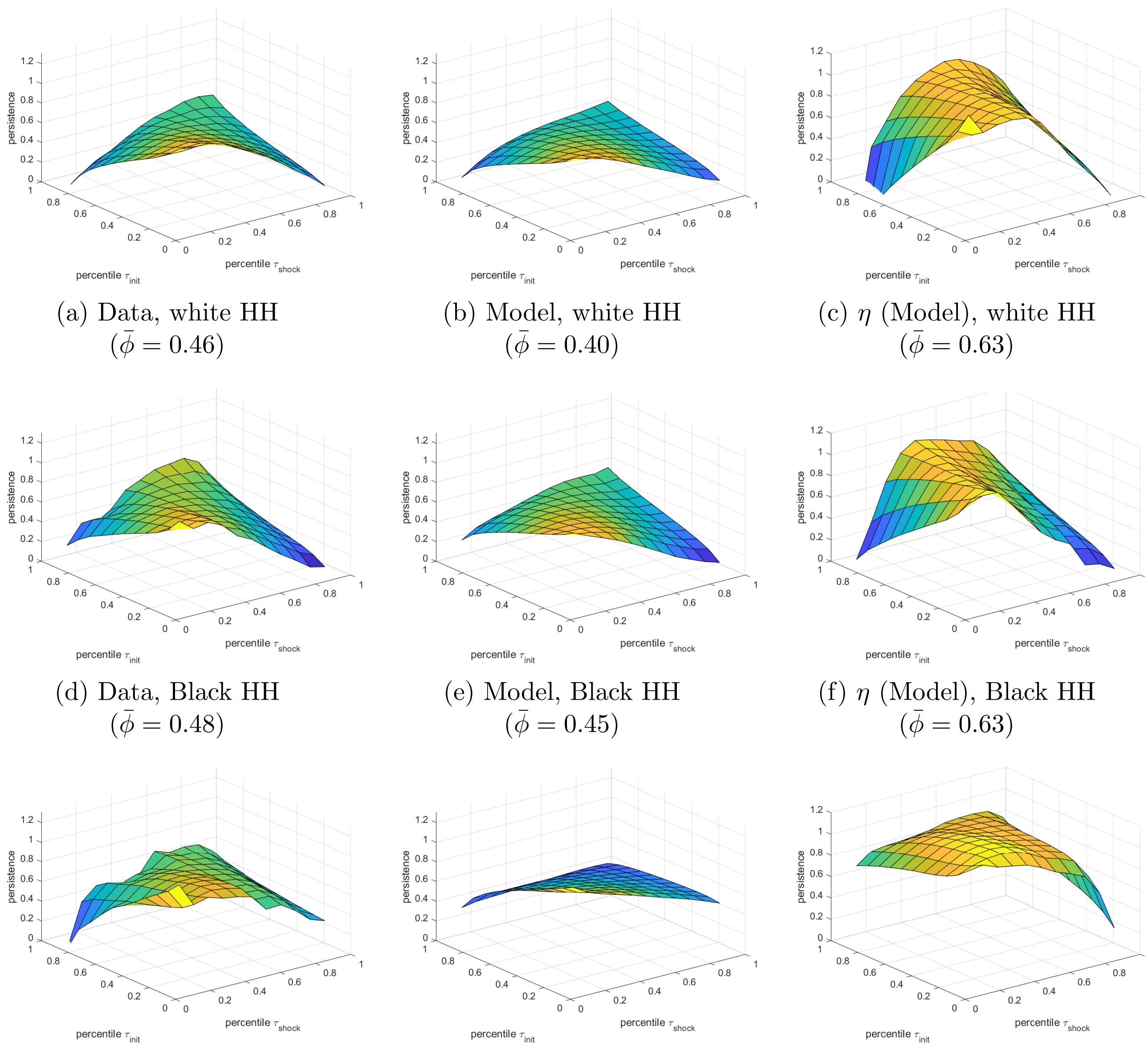
Persistence of Residual Spending:

Both average medical spending and the risk of catastrophic medical spending are important for understanding lifetime financial security. In Table 1 we highlighted racial differences in annual spending, both in averages and in the right tail of medical spending in any given year. But even though catastrophic medical spending can be due to high medical spending in a single year, it is often due instead to multiple years of high spending for chronic conditions such as dementia. To measure this risk, we need good estimates not only of annual medical spending but of the persistence of medical spending over time. This persistence depends in part on the persistence of health, discussed above, but also on the persistence of the medical spending conditional on health and age, namely the total spending residual $m_{i,t}$ defined in equation (4).

To find both the predicted component of medical spending ($X'_{i,t}\gamma$ in (4)) and the residuals, we regress total medical spending on age and PI, household composition, health and nursing home indicators and interactions of these variables.[14] We estimate separate regressions for each race; the R^2 is 0.155 for the full

[14]We distinguish between single man, single women and couples. We include a third-order age polynomial, and a second-order polynomial in PI. We include gender-specific indicators for bad health and being in a nursing home, along with their lags, and gender-specific indicators of whether a household member died in the past year. The nursing home indicators are interacted with household status (single vs. married) and the newly-dead indicators are interacted with PI.

sample, 0.164 for white households, 0.134 for Black households, and 0.137 for Hispanic households. These low R^2s show that medical spending is difficult to predict, even with a rich set of economic, demographic, and health controls. They highlight that most of the variation in medical spending is due to the residual component, making its distribution and persistence crucial to understanding lifetime medical spending risk.



(a) Data, white HH ($\bar{\phi} = 0.46$)

(b) Model, white HH ($\bar{\phi} = 0.40$)

(c) η (Model), white HH ($\bar{\phi} = 0.63$)

(d) Data, Black HH ($\bar{\phi} = 0.48$)

(e) Model, Black HH ($\bar{\phi} = 0.45$)

(f) η (Model), Black HH ($\bar{\phi} = 0.63$)

(g) Data, Hispanic HH ($\bar{\phi} = 0.62$)

(h) Model, Hispanic HH ($\bar{\phi} = 0.64$)

(i) η (Model), Hispanic HH ($\bar{\phi} = 0.88$)

Figure 1: Persistence of Residual Medical Spending, by Race

Figure 1 shows the persistence of $m_{i,t}$, as observed in the data and predicted by the model, along with the persistence of the persistent component of this residual, $\eta_{,it}$, for each race. Formally, persistence ($\phi_\tau(\cdot)$) is the derivative of the conditional quantile function defined in equation (7), averaged across age.

Each graph of $\phi_\tau(\cdot)$ thus shows how responsive today's residual, $m_{i,t}$ (or $\eta_{i,t}$), is to its lagged value, $m_{i,t-1}$ (or $\eta_{i,t-1}$), at each point of its conditional distribution (indexed by the percentile rank τ_{shock}), given its lagged value (indexed by the rank τ_{init}). All panels also present the residual's autocorrelation coefficient, which we denote as $\bar{\phi}$ to highlight that it is the average level of persistence over the distribution. To fix ideas, recall that if $m_{i,t}$ followed an AR(1) process, $\phi_\tau(\cdot)$ would be a constant equal to $\bar{\phi}$, and each panel would show a flat plane. The figure clearly shows, however, that persistence is not constant across the distribution of medical spending.

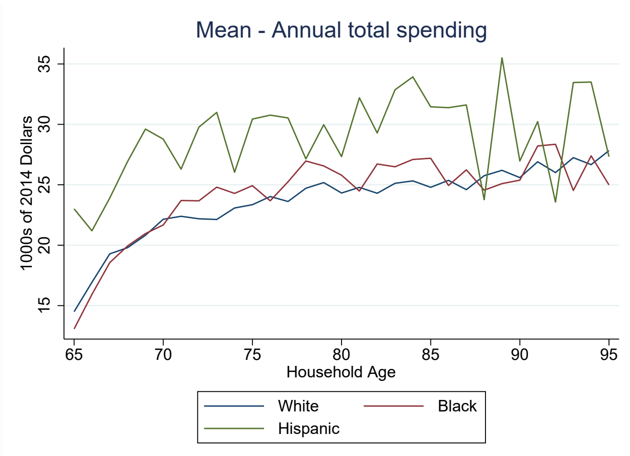
The first row of Figure 1 shows results for white households. Panel (a) shows the estimated persistence of residual spending in the data. Persistence varies greatly with both the lagged values and the contemporaneous shock. It is higher near the middle of the distribution. This means that those in the middle may face moderate but persistent expenses that can accumulate into a catastrophic lifetime cost. In contrast, persistence is quite low (around 0.1) for large medical shocks among those with low initial medical spending. It is much higher, almost 0.5, for those with high initial spending and large positive shocks (the "north" corner of the graph). What this means is that catastrophic medical spending is particularly persistent. Persistence is even higher for those with low initial spending and a low shock (the bottom point on the graph). In short, when medical spending is low (for a given set of correlates), it is unlikely to be high in the near future, and when medical spending is high, it is likely to stay high. The average level of persistence, reported in the title of the panel, is $\bar{\phi} = 0.46$.

Panel (b) shows persistence in the simulated data. The similarity of Panels (a) and (b) shows that the model matches the persistence pattern observed in the data, although it modestly understates the average level of persistence, with $\bar{\phi} = 0.40$ rather than 0.46. Panel (c) shows persistence for the Markov component of medical spending, η. The persistence of η is higher than that of total medical spending, as total medical spending includes the transitory shock ε. Over a large part of its distribution, the persistence of η is close to 0.9, indicating that for chronic health conditions, medical spending is very persistent. It is these shocks, which cause high spending for many periods, that can drain a family's finances.

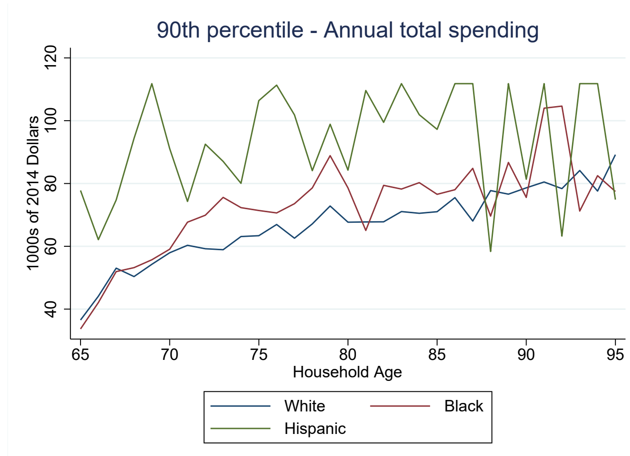
The second row of Figure 1 shows persistence functions for Black households. Comparing the second row with the first reveals that Black and white households have very similar patterns of persistence. In contrast, persistence for Hispanic households is higher than that of white and Black households, suggesting very different dynamics. We also find less evidence of nonlinear persistence for Hispanic households – the persistence of their medical spending is closer to constant. Nevertheless, even among Hispanic households, persistence is higher for those with low spending and a low shock or high spending and a high shock.

Total medical spending over the life-cycle:

To compute the mean and 90th percentile of medical spending, we begin with the age, PI and health histories found in the data. We then simulate health for years outside our sample period and medical spending residuals for households at every age. Applying the estimated medical spending coefficients (γ) to the simulated health histories and medical spending residuals yields total household medical spending for surviving households. Figure 2 presents profiles of medical spending for each race.[15] Medical spending rises rapidly with age for all racial groups. Differences in both the mean (left panel) and the 90th percentile (right panel) of total medical spending by race are modest. This is unsurprising given that Table 1 shows that racial differences in total medical spending are small, and Table 4 shows that these differences are partly explained by white households being older on average than other groups. Medical spending is somewhat higher for Hispanic households. However, this should be taken with caution, as the estimates for Hispanic households are derived from a small sample.



(a) Mean



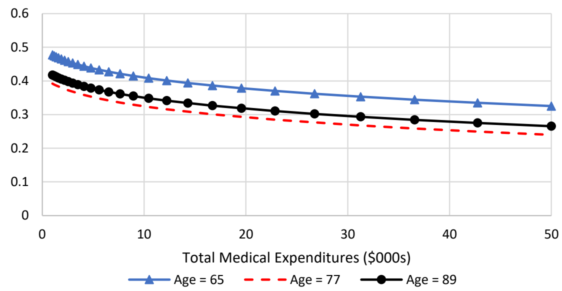
(b) 90th percentile

Figure 2: Mean and 90th percentile of Total Annual Medical Spending, by Race.

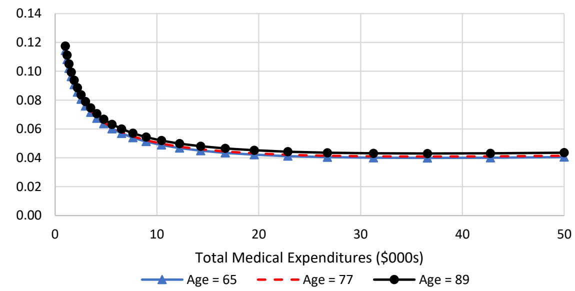
Budget sets:

The fraction of total medical expenses paid out of pocket is modeled as a function age, PI, total medical spending and health. We allow the relationship to differ by race and whether the household receives Medicaid. Figure 3 shows how the out-of-pocket fractions depend on total medical spending and age for different PI groups. The left column of the figure shows results for households not receiving Medicaid. Households with higher total expenses pay a smaller fraction out of pocket, with the fraction dropping by roughly 15 percentage points as total expenses rise from \$1,000 to \$50,000. The fractions are fairly similar across races, although Hispanic households pay a slightly smaller share out of pocket. The right column shows that Medicaid recipients pay a much smaller share, often well below 10%. Black households pay slightly less than white households, and Hispanic households pay even less.

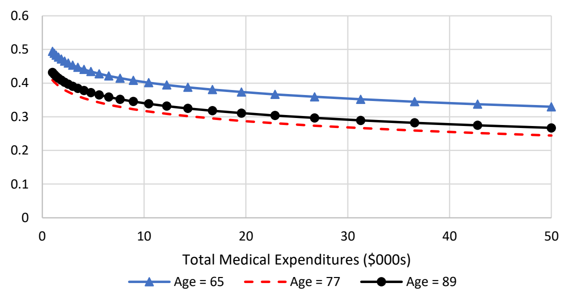
[15]Values are dated by the beginning of the spending interval: for example, expenses for age 70 describe the medical expenses incurred between ages 70 and 72 by people alive at both dates.



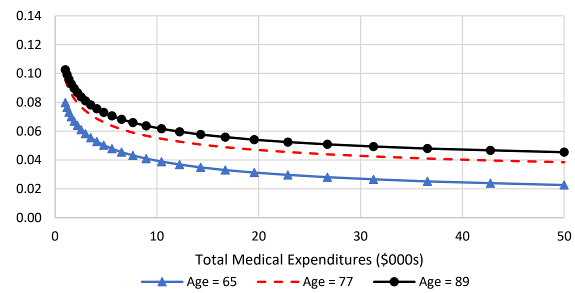
(a) White HH, PI = 0.5, Non-Medicaid



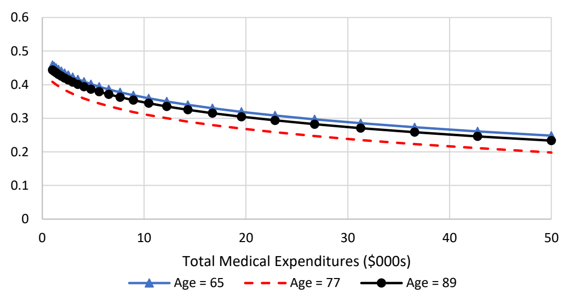
(b) White HH, PI = 0.1, Medicaid



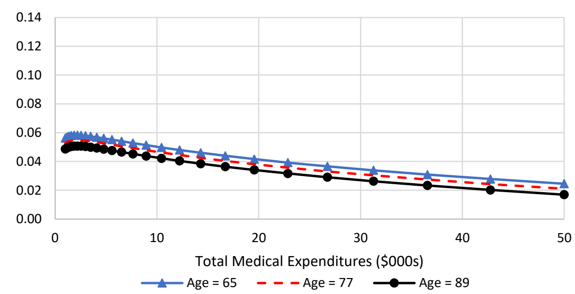
(c) Black HH, PI = 0.5, Non-Medicaid



(c) Black HH, PI = 0.1, Medicaid



(e) Hispanic HH, PI = 0.5, Non-Medicaid



(f) Hispanic HH, PI = 0.1, Medicaid

Figure 3: Fraction of Medical Spending Paid Out of Pocket, by Race

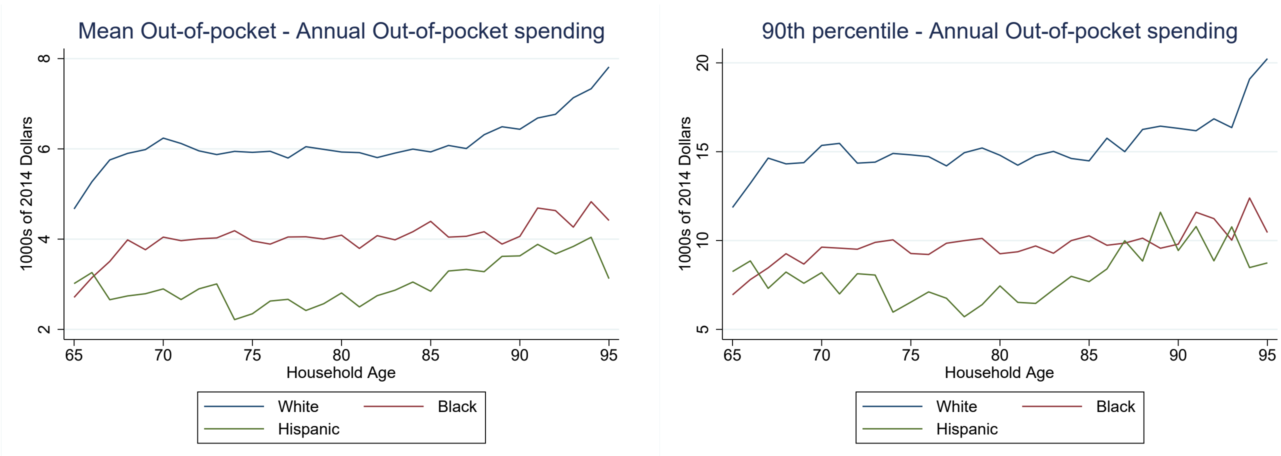
Since Medicaid recipiency is a key driver of the share of spending paid out of pocket, we model recipiency separately for each race as a logistic function of age, PI, total medical spending and health. We consider specifications both with and without lagged Medicaid receipt, which has strong predictive power. Our goal is to capture two key observations from Table 1. First, those with extremely high levels of spending often receive Medicaid, in part because Medicaid covers nursing home care for those unable to afford it. Second, racial minorities tend to have more of their care paid for by Medicaid. The table shows that the probability of receiving Medicaid is negatively related to PI but is increasing in the interaction between medical spending and PI. This means that lower-income households (who are disproportionately Black or Hispanic) are likely to receive Medicaid regardless of their medical needs, while higher-income households receive Medicaid only when spending is high.[16] Having a member in a nursing home raises the

[16]Formally, these differences appear as the Categorically and Medically Needy provisions of Medicaid (De Nardi, French,

probability of Medicaid receipt significantly for white households, modestly for Black households, and not at all for Hispanic households. This too may reflect the tendency of richer households to receive Medicaid only when their expenses are high and in categories, such as nursing home care, that are not covered by other payors such as Medicare.

Out-of-pocket spending over the life cycle:

Combining simulated total spending with the budget sets, Figure 4 shows annual household out-of-pocket spending. The figure highlights the lower out-of-pocket spending of minorities at both the mean and the 90th percentile. This arises from having a higher share of their medical spending paid by Medicaid, especially at older ages.



(a) Mean

(b) 90th percentile

Figure 4: Mean and 90th percentile of annual out-of-pocket medical spending by race.

7 The Distribution of Lifetime Medical Spending

In this section we combine the estimated health and survival model with the medical spending model to simulate remaining lifetime medical spending. The simulations begin with the joint distribution of variables found in our data. Thus our model matches the joint distribution of longevity and medical spending risks faced by households of each race. At each age, we calculate the undiscounted value of remaining lifetime medical spending from that age forward.

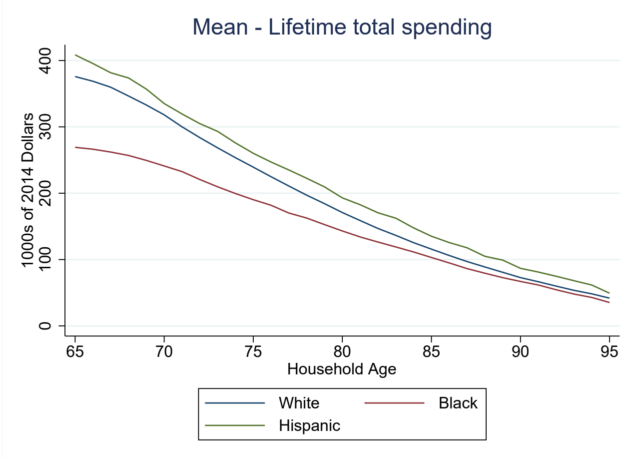
Figure 5 plots remaining lifetime total and out-of-pocket spending by race. The lifetime totals in panel (a) are considerable. At age 65, white households will incur, on average, around \$380,000 of medical spending over the remainder of their lives, versus \$280,000 and \$400,000 for Black and Hispanic households, respectively. Furthermore, the risk of catastrophic spending is significant for all racial groups.

and Jones, 2016).

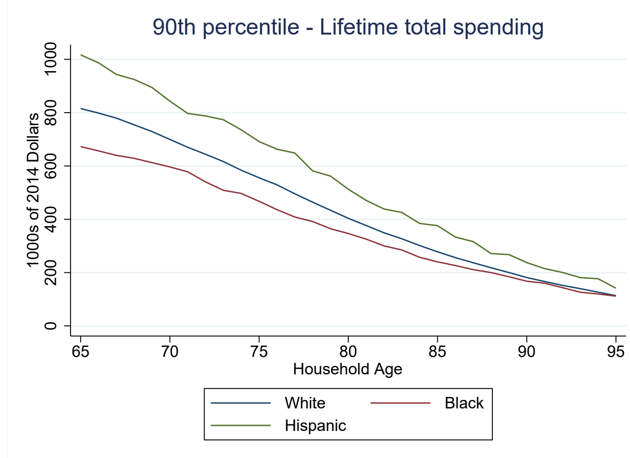
Table 8: Probability of Receiving Medicaid Benefits

	White		Black		Hispanic	
Log Medical Spending	0.270	-0.332	0.330*	0.123	-0.002	0.727
	(0.160)	(0.173)	(0.144)	(0.250)	(0.219)	(0.496)
Log Medical Spending^2	0.003	0.033***	-0.004	0.009	0.025	-0.026
	(0.009)	(0.010)	(0.009)	(0.015)	(0.013)	(0.029)
PI Percentile	-17.715***	-11.347***	-19.187***	-10.034***	-18.113***	-14.267**
	(0.980)	(1.734)	(1.501)	(3.041)	(2.643)	(5.262)
PI Percentile^2	1.645*	-2.312	0.722	-4.483	2.545	-3.811
	(0.696)	(1.333)	(1.120)	(2.429)	(2.059)	(4.118)
Household Age	-2.961***	-4.500***	-0.535	-0.421	1.128	4.078
	(0.533)	(1.213)	(0.633)	(1.378)	(1.151)	(3.287)
Household Age^2/100	3.360***	5.335***	0.425	0.311	-1.472	-5.095
	(0.652)	(1.470)	(0.774)	(1.654)	(1.420)	(4.059)
Household Age^3/10000	-1.262***	-2.087***	-0.088	-0.041	0.631	2.103
	(0.264)	(0.590)	(0.313)	(0.657)	(0.580)	(1.660)
Woman in Bad Health	0.405***	0.291**	0.138	-0.075	0.104	0.260
	(0.053)	(0.109)	(0.072)	(0.174)	(0.122)	(0.278)
Man in Bad Health	0.331***	0.440*	0.180	0.699*	-0.230	-0.509
	(0.098)	(0.179)	(0.146)	(0.332)	(0.227)	(0.538)
Woman in Nursing Home	1.092***	0.815***	0.558***	0.387	-0.029	-0.054
	(0.073)	(0.144)	(0.142)	(0.334)	(0.358)	(0.799)
Man in Nursing Home	1.738***	1.046***	0.520*	0.837	-0.791	-0.789
	(0.128)	(0.249)	(0.249)	(0.540)	(0.508)	(1.104)
Household Dead	0.120	0.136	-0.655***	-0.735*	-1.467***	-0.428
	(0.089)	(0.173)	(0.161)	(0.369)	(0.358)	(0.689)
Couple	-0.935***	0.050	-0.429	-0.892	-0.380	0.134
	(0.181)	(0.340)	(0.233)	(0.575)	(0.328)	(0.796)
Male	0.404***	0.205	0.506***	0.313	-0.618**	-0.537
	(0.088)	(0.169)	(0.132)	(0.312)	(0.210)	(0.507)
Total Medical Spending × PI	1.078***	1.030***	1.323***	1.101***	1.042***	1.469*
	(0.098)	(0.157)	(0.159)	(0.294)	(0.293)	(0.572)
Couple × PI Percentile	0.078	-0.983	1.317*	2.030	-2.114	-5.099
	(0.366)	(0.624)	(0.565)	(1.217)	(1.299)	(3.655)
Medicaid Expenditure Last Year		6.531***		6.481***		5.887***
		(0.121)		(0.178)		(0.251)
Constant	82.9***	122.1***	17.2	10.7	-29.2	-114.4
	(14.4)	(33.1)	(17.2)	(38.0)	(30.9)	(88.3)
N	35,749	31,120	7,453	6,,414	2,476	2,117
pseudo R^2	0.36	0.77	0.28	0.80	0.30	0.78

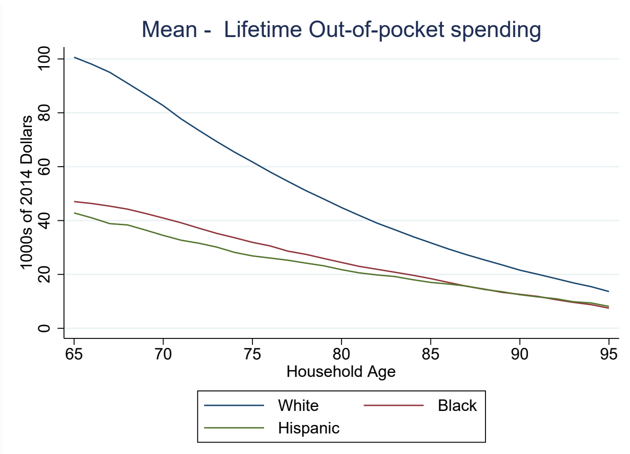
Notes: Logit regression. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1



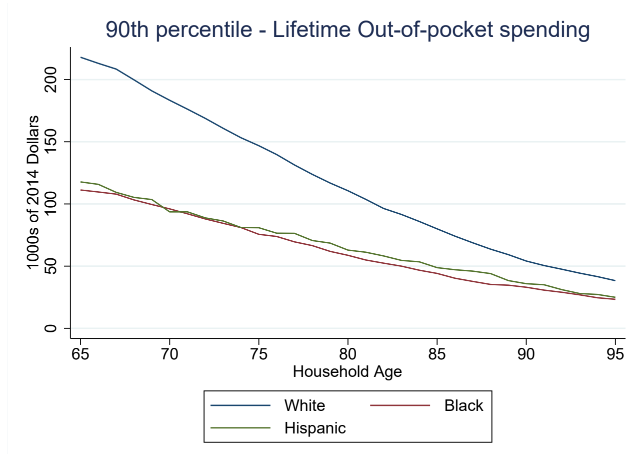
(a) Mean, Total Spending



(b) 90th percentile, Total Spending



(c) Mean, Out-of-Pocket



(d) 90th percentile, Out-of-Pocket

Figure 5: Remaining Lifetime Medical Spending.

As shown in panel (b), at age 65, 10% of white households will incur at least $800,000 of expenses. The 90th percentiles for the Black and Hispanic spending distributions are $650,000 and $1,000,000, respectively.

One might expect the lifetime totals to fall rapidly as households age and near the ends of their lives. This is not the case. A white household alive at age 90 will on average spend more than $70,000 before they die. The amounts for Black and Hispanic households are similar. The slow decline of lifetime spending is due to annual medical spending rising with age, in large part due to the high rates of nursing home entry at older ages. Households that live to older ages have shorter remaining lives but higher annual spending. A number of papers have considered whether most of this spending increase is due to rising mortality and the associated end-of-life expenses: see the discussion in De Nardi et al. (2016). In our framework, we allow medical spending to both jump prior to death and to rise with age among the living.

Our model allows for racial differences in longevity by race and other characteristics which is important for understanding differences in remaining lifetime medical spending of different races. On average, life expectancy for Black individuals at age 65 is almost two years less than white and Hispanic individuals and thus they accrue fewer years of medical spending.

Panel (c) of Figure 5 shows the remaining lifetime out-of-pocket spending predicted by the model. Out-of-pocket spending does not mirror the patterns for total spending seen in panel (a). First, the levels are lower for all races. Second, the levels are especially low for Black and Hispanic households, who pay a smaller share of their medical spending out of pocket. At age 65, white households will on average incur around \$100,000 of out-of-pocket medical spending over the remainder of their lives. The amounts are considerably lower for Black and Hispanic households, at \$48,000 and \$42,000, respectively. Panel (d) shows the lifetime remaining out-of-pocket spending for those at the 90th percentile. White households at the 90th percentile will incur \$220,000 of medical spending, double that of Black and Hispanic households, \$120,000 and \$110,000, respectively.

A number of recent papers have argued that Medicare and Medicaid significantly reduce the out-of-pocket spending risk faced by older households. We find that despite these payors, medical spending risk is high in old age. Black and Hispanic households, however, are better insured than white households by these public programs.

8 Discussion and Conclusions

We document differences in total and out-of-pocket medical spending between white, Black, and Hispanic households. While white households have higher total medical spending on average, racial gaps in total spending are explained by observable covariates such as household structure, health status, and education. white households pay a higher share of their medical expenses out of pocket. This is partially, but not fully, explained by their higher income and better health, which make them less likely to receive Medicaid. This shows that Medicaid provides important insurance against medical spending risk, especially for Black and Hispanic households.

Using an estimated dynamic model of health and medical spending, we simulate lifetime medical spending histories. We show that at age 65, white households will incur significantly more medical spending than Black households over the rest of their lives, which largely reflects the shorter life spans of Black households. Because we model the budget sets households face, we can also calculate how much households pay out of pocket. Here the differences are even more stark. Over the remainders of their lives, white households will on average incur twice as much out-of-pocket medical spending as Black or Hispanic households. This is because Black and Hispanic households are better insured by Medicare and Medicaid.

Our model can be used to evaluate counterfactual policies that impact co-insurance rates, including those designed to insure against nursing home or other catastrophic medical spending.

An important potential next step would be to better model why Black and Hispanic households are much more likely to receive Medicaid. Medicaid recipiency depends on a host of factors that we omit. Perhaps the most important of these is housing and non-housing wealth. Wealth is a difficult variable to model because it depends not only on medical spending, but also other spending choices. Furthermore, households may choose to run down their wealth to qualify for Medicaid. Building in this forward-looking behavior would be a challenging but worthwhile extension.

References

Arapakis, Karolos, Eric French, John Bailey Jones, and Jeremy McCauley. 2021. "On the Distribution and Dynamics of Medical Expenditure Among the Elderly." Working Paper WP 2021-436, Michigan Retirement and Disability Research Center.

Arellano, Manuel, Richard Blundell, and Stéphane Bonhomme. 2017. "Earnings and Consumption Dynamics: A Nonlinear Panel Data Framework." *Econometrica* 85 (3):693–734.

Arellano, Manuel, Richard Blundell, Stéphane Bonhomme, and Jack Light. 2023. "Heterogeneity of Consumption Responses to Income Shocks in the Presence of Nonlinear Persistence." *Journal of Econometrics* :105449.

Blundell, Richard, Jack Britton, Monica Costa Dias, Eric French, and Weijian Zou. 2022. "The Dynamic Effects of Health on the Employment of Older Workers: Impacts by Gender, Country, and Race." *Ann Arbor, MI. University of Michigan Retirement and Disability Research Center (MRDRC) Working Paper* .

Choukhmane, Taha, Jorge Colmenares, Cormac O'Dea, Jonathan Rothbaum, and Lawrence Schmidt. 2022. "Who Benefits from Retirement Saving Incentives in the US? Evidence on Racial Gaps in Retirement Wealth Accumulation." *Unpublished manuscript* .

Creal, Drew. 2012. "A Survey of Sequential Monte Carlo Methods for Economics and finance." *Econometric Reviews* 31 (3):245–296.

Cylus, Jonathan, Micah Hartman, Benjamin Washington, Kimberly Andrews, and Aaron Catlin. 2011. "Pronounced Gender and Age Differences are Evident in Personal Health Care Spending per Person." *Health Affairs* 30 (1):153–160.

David, Martin, Roderick J. A. Little, Michael E. Samuhel, and Robert K. Triest. 1986. "Alternative Methods for CPS Income Imputation." *Journal of the American Statistical Association* 81 (393):29–41.

De Nardi, Mariacristina, Eric French, and John Bailey Jones. 2016. "Medicaid Insurance in Old Age." *American Economic Review* 106 (11):3480–3520.

De Nardi, Mariacristina, Eric French, John Bailey Jones, and Jeremy McCauley. 2016. "Medical Spending on the US Elderly." *Fiscal Studies* 37 (3-4):717–747.

De Nardi, Mariacristina, Eric French, John Bailey Jones, and Rory McGee. 2023. "Why Do Couples and Singles Save during Retirement? Household Differences and their Aggregate Implications."

Dieleman, Joseph L, Carina Chen, Sawyer W Crosby, Angela Liu, Darrah McCracken, Ian A Pollock, Maitreyi Sahu et al. 2021. "US Health Care Spending by Race and Ethnicity, 2002-2016." *JAMA* 326 (7):649–659.

Doucet, Arnaud, Adam M Johansen et al. 2009. "A Tutorial on Particle Filtering and Smoothing: Fifteen Years Later." *Handbook of Nonlinear Filtering* 12 (656-704):3.

Escarce, José J and Kanika Kapur. 2003. "Racial and Ethnic Differences in Public and Private Medical Care Expenditures among Aged Medicare Beneficiaries." *The Milbank Quarterly* 81 (2):249–275.

Fahle, Sean, Kathleen McGarry, and Jonathan Skinner. 2016. "Out-of-Pocket Medical Expenditures in the United States: Evidence from the Health and Retirement Study." *Fiscal Studies* 37 (3-4):785–819.

Feenberg, Daniel and Jonathan Skinner. 1994. "The Risk and Duration of Catastrophic Health Care Expenditures." *Review of Economics and Statistics* 76:633–647.

French, Eric and John Bailey Jones. 2004. "On the Distribution and Dynamics of Health Care Costs." *Journal of Applied Econometrics* 19 (4):705–721.

French, Eric, John Bailey Jones, and Jeremy McCauley. 2017. "The Accuracy of Economic Measurement in the Health and Retirement Study." *Forum for Health Economics and Policy* 20 (2).

French, Eric, John Bailey Jones, and Rory McGee. 2023. "Why Do Retired Households Draw Down Their Wealth So Slowly?" *Journal of Economic Perspectives* forthcoming.

Friedberg, Leora, Wenliang Hou, Wei Sun, Anthony Webb, and Zhenyu Li. 2015. "New Evidence on the Risk of Requiring Long-Term Care." WP 2014-12, Boston College Center for Retirement Research.

Gaskin, Darrell J, Becky A Briesacher, Rhona Limcangco, and Betsy L Brigantti. 2006. "Exploring Racial and Ethnic Disparities in Prescription Drug Spending and Use among Medicare Beneficiaries." *The American Journal of Geriatric Pharmacotherapy* 4 (2):96–111.

Gelbach, Jonah B. 2016. "When Do Covariates Matter? And Which Ones, and How Much?" *Journal of Labor Economics* 34 (2):509–543.

Goda, Gopi Shah, John B Shoven, and Sita Nataraj Slavov. 2013. "Does Widowhood Explain Gender Differences in Out-of-pocket Medical Spending among the Elderly?" *Journal of Health Economics* 32 (3):647–658.

Hurd, Michael D., Pierre-Carl Michaud, and Susann Rohwedder. 2017. "Distribution of Lifetime Nursing Home use and of Out-of-Pocket Spending." *Proceedings of the National Academy of Sciences* 114 (37):277–297.

Jacobson, Gretchen, Giselle Casillas, Anthony Damico, Tricia Neuman, and Marsha Gold. 2016. "Medicare Advantage 2016 Spotlight: Enrollment Market Update." *The Henry J. Kaiser Family Foundation Website* URL `https://files.kff.org/attachment/Issue-Brief-Medicare-Advantage-2016-Spotlight-Enrollment-Market-Update`.

Koenker, Roger and Gilbert Bassett Jr. 1978. "Regression quantiles." *Econometrica: Journal of the Econometric Society* :33–50.

Kopecky, Karen and Tatyana Koreshkova. 2014. "The Impact of Medical and Nursing Home Expenses and Social Insurance Policies on Savings and Inequality." *American Economic Journal: Macroeconomics* 6 (3):29–72.

McGarry, Kathleen and Robert F Schoeni. 2005. "Widow (er) Poverty and Out-of-pocket Medical Expenditures near the End of Life." *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences* 60 (3):S160–S168.

Nicholas, Lauren Hersch, Kenneth M Langa, Theodore J Iwashyna, and David R Weir. 2011. "Regional Variation in the Association between Advance Directives and End-of-life Medicare Expenditures." *JAMA* 306 (13):1447–1453.

Pashchenko, Svetlana and Ponpoje Porapakkarm. 2016. "Medical spending in the US: facts from the Medical Expenditure Panel Survey Data Set." *Fiscal Studies* 37 (3-4):689–716.

A Sample Selection

We drop households for various reasons. These include disagreement between household members, problematic mortality transitions, households with multiple members of the same sex, and refusal to provide Medicare and Medicaid spending records. Table 3 below denotes the starting public HRS sample size and the sample size after every drop.

Drop reason	Post-drop sample size
	26,598
One spouse dead and the other claims never married or divorced	26,596
One spouse claims married and the other something else	25,343
Two or more members of the same sex	24,897
One spouse claims never married an the other is not missing	24,818
One spouse claims divorced and the other widowed	24,545
People who are partnered	23,466
People who "died", then came back to life	23,161
Transitions with widowed or both widowed	23,055
Transitions to got married or divorced	22,719
Household split etc. based on sub-household identifier	22,695
Households that appear once and have missing marital status	22,683
No domestic partner in sample but married in some wave	21,903
No living member the year household joined HRS	21,713
Households dead or under 65 in the estimation period (1999-2012)	9,969
2+ years post death Medicare and Medicaid spending	6,783
Incomplete Medicare and Medicaid records	4,391

Notes: Sample size refers to the number of unique households.

Table 9: Sample Selection

B Our Medicare and Medicaid Data

Medicare

We link restricted Medicare fee-for-service (Parts A and B), and Part D data for the years 1999-2012 (2006 was the first year of Medicare Part D and thus our Part D data begins then) to our HRS survey data for respondents who consent to allow their Medicare data to be linked to their survey responses (approximately 64.7% percent of persons in our study population). These records have enrollment information and data on reimbursement amounts for inpatient, skilled nursing facility, home health, and hospice claims (Medicare Part A), as well as outpatient, carrier (non-institutional medical care providers such as individual or group practitioners, non-hospital labs, and ambulances), and durable medical equipment claims for Medicare Part B.

We use the Beneficiary Annual Summary File (BASF), which summarizes information from the micro-level claims records. The BASF contains annual information for each individual on the number of months of enrollment in Medicare Part A, Part B, and non-fee-for-service plans. The BASF has information on Medicare fee-for-service (FFS) claims. Almost all claims for services used by non-FFS Medicare patients are not observed in these data, so all analyses exclude an individual in a given year if they were enrolled in a non-FFS Medicare plan for more than half the year.

Medicare Part D is the prescription drug benefit. We calculate the Medicare Part D payment using the Part D event files. For the Part D Medicare contribution we subtract from the gross drug cost the payments paid by the beneficiary, family, or friends.

Medicaid

As with the Medicare data, we are able to link restricted Medicaid data (CMS Medicaid Analytic eXtract, or "MAX" files) for those in the HRS who gave permission, allowing us to measure Medicaid expenditures for the Medicaid beneficiaries in our data set for the years 1999-2012. The MAX files contain personal summaries (which contain eligibility, enrollment, and demographic information) and claims data across four service categories (inpatient, long-term care, prescription drugs, and other services). Other services include a variety of services (e.g., physician services and lab work) that do not fit under the other three service categories. The inpatient, long-term care, prescription drugs, and other services files contain the primary variable of interest, "Medicaid Payment Amount," which is the total amount of money paid by Medicaid for a particular service. We sum over all the claims for all the different service categories for a particular individual in each year.

C Imputing Missing Medical Expenditures

Our goal is to measure all medical spending: the variable M_{it} in equation (4) of the main text is defined to include out-of-pocket spending, Medicare and Medicaid payments, and private and other public (such as Veterans Administration benefits, and care provided by local and state health departments) insurance payments. While the HRS includes information on out-of-pocket spending and can be linked to Medicare and Medicaid payments, it does not include Medicare Part C, private, or other public insurance payments. In this appendix, we describe how we use data from the Medical Expenditure Panel Survey (MEPS) to impute these payments in the HRS. Although the MEPS has extremely high quality information on all payors for all household members, it lacks the long panel dimension of the HRS. Our imputation procedures allow us to exploit the best of both data sets.

Our imputation procedure has two steps. First, we use the MEPS to infer private and other public insurance payments, conditional on variables observed in both data sets. Second, we impute private and other public insurance payments in the HRS data using a conditional mean matching procedure (which is a procedure very similar to hot-decking).

First Step of Imputation Procedure

We use the MEPS to infer payments of other payors, conditional on the observable variables that exist in both the MEPS and the HRS datasets.

Let i index individuals in the HRS and j index individuals in the MEPS. Define M_{it}^{obs} as out of pocket, Medicaid, and Medicare (Part A, B, and D, but not Part C) payments observed in both the HRS and MEPS data sets, M_{it}^{miss} as the components of medical spending missing in the HRS but observed in the MEPS, and $M_{it} = M_{it}^{miss} + M_{it}^{obs}$ as total medical spending. To impute M_{it}^{miss}, which is missing in the HRS, we follow De Nardi et al. (2023) and use a predictive mean-matching regression approach. There are two steps to our procedure. First, we use the MEPS data to regress M_{it}^{miss} on observable variables that exist in both data sets. Second, we impute M_{it}^{miss} in the HRS data using a conditional mean-matching procedure, a procedure very similar to hot-decking.

First, for every member of the the MEPS sample, we regress the variable of interest M^{miss} on the vector of observable variables z_{jt}, yielding $M_{jt}^{miss} = z_{jt}\beta + \varepsilon_{jt}$. Second, for each individual j in the MEPS, we calculate the predicted value $\widehat{M^{miss}}_{jt} = z_{jt}\hat{\beta}$, and for each member of the sample we calculate the residual $\widehat{\varepsilon}_{jt} = M_{jt}^{miss} - \widehat{M^{miss}}_{jt}$. Third, we sort the predicted value $\widehat{M^{miss}}_{jt}$ into deciles and keep track of all values of $\hat{\varepsilon}_{jt}$ within each decile. We use this procedure separately to impute Medicare Part C benefits, private payments, and other payments.

In practice, we include in z_{jt} a fourth-order age polynomial, marital status, gender, self-reported health (=1 if self reported health is good, very good, or excellent), race, visiting a medical practitioner (doctor, hospital or dentist), out-of-pocket medical spending, education of head (high school, some college, college), death of an individual, and total household income. We estimate this regression two times: once for the privately insured, and once for other payors.

Because the measure of medical spending in the HRS is medical spending over two years, we divide HRS out-of-pocket medical spending by 2 and assume that medical spending is equal across the two years.

Second Step of Imputation Procedure

For every observation in the HRS sample with a positive Medicaid indicator, we impute $\widehat{Med}_{it} = z_{it}\widehat{\beta}$, using the values of $\widehat{\beta}$ estimated from the MEPS. Then we impute ε_{it} for each observation of this subsample by finding a random observation in the MEPS with a value of $\widehat{Med}_{jt}$ in the same decile as $\widehat{Med}_{it}$, and setting $\hat{\varepsilon}_{it} = \hat{\varepsilon}_{jt}$. The imputed value of Med_{it} is $\widehat{Med}_{it} + \hat{\varepsilon}_{it}$.

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. Second, to improve goodness of fit we use a large number of "z" variables. We find that adding extra variables are very important for improving goodness of fit when imputing payments. Because hot-decking uses a full set of interactions, this would result in a large number of hot-decking cells relative to our sample size. Thus, in this context, hot decking is too data intensive.

D Validating the Administrative Medical Spending Data

Here, we examine in greater detail the accuracy of the administrative medical spending data, as well as the out-of-pocket spending found in the Assets and Health Dynamics of the Oldest Old (AHEAD) cohort of the HRS, comparing them to data from the MCBS. See De Nardi, French, and Jones (2016) and De Nardi et al. (2016) for more details of the MCBS data and (for example) Nicholas et al. (2011) for details of the HRS linked data.

The MCBS is a nationally representative survey of Medicare beneficiaries, consisting of Disability Insurance recipients and Medicare recipients aged 65 and older. The survey contains an over-sample of beneficiaries older than 80 and disabled individuals younger than 65. Respondents are asked about health status, health insurance, and health care spending (from all sources). The MCBS data are matched to Medicare records, and medical spending data are created through a reconciliation process that combines information from survey respondents with Medicare administrative files. As a result, the survey is thought to give extremely accurate data on Medicare payments and fairly accurate data on out-of-pocket and Medicaid payments. As in the HRS survey, the MCBS survey includes information on those who enter a nursing home or die. Respondents are interviewed up to 12 times over a 4 year period. We aggregate the data to an annual level. In both samples, we applied only modest sample selection restrictions. The key sample selection issue shown in Table 9 is that, in the HRS, we drop households with missing or erroneous Medicare or Medicaid records.

Here we compare distributions of total, out-of-pocket, Medicare, and Medicaid payments between the MCBS are the HRS data. Medical spending the HRS is measured at an individual level (rather than household) to be comparable with the MCBS. The comparison can be seen in Table 10. Medical spending is higher in our HRS sample than in the MCBS sample. Furthermore, this higher level of spending is driven by higher out-of-pocket spending, Medicare, and Medicaid spending. These differences potentially are an advantage of the HRS data since, as noted in De Nardi et al. (2016), the MCBS clearly understates aggregate Medicare and especially Medicaid spending, potentially due to the issue that the MCBS does not have administrative data on Medicaid spending, and thus relies heavily on imputation.

The next set of benchmarking exercises that we perform is for out-of-pocket medical spending, Medi-

	Total Spending				OOP			
Total Spending	HRS		MCBS		HRS		MCBS	
Percentiles	Average Exp.	Pct. Total	Average Exp.	Pct. Total	Average Exp.	Pct. Total	Average Exp.	Pct. Total
All	17,091	100	14,120	100	3,825	100	2,740	100
95-100%	114,238	33.4	97,880	34.6	45,643	59.6	26,930	49.1
90-95%	59,000	17.3	48,890	17.3	8,619	11.3	6,700	12.2
70-90%	26,870	31.4	20,540	29.1	3,480	18.2	2,920	21.3
50-70%	9,025	10.6	7,750	11	1,394	7.3	1,360	9.9
0-50%	2,502	7.3	2,250	8	178	3.6	420	7.6

	Medicare				Medicaid			
Total Spending	HRS		MCBS		HRS		MCBS	
Percentiles	Average Exp.	Pct. Total	Average Exp.	Pct. Total	Average Exp.	Pct. Total	Average Exp.	Pct. Total
All	11,343	100	7,720	100	1,896	100	1,320	100
95-100%	85,268	37.6	67,560	43.7	33,773	89.1	24,980	94.7
90-95%	41,731	18.4	28,370	18.4	4,092	10.8	1,360	5.2
70-90%	17,251	30.4	10,280	26.6	230	0.2	10	0.1
50-70%	5,031	8.9	2,980	7.7	0	0	0	0
0-50%	1,076	4.7	550	3.5	0	0	0	0

Table 10: Individual Medical Spending Percentiles: HRS versus MCBS

Income Quintile	Total Income	Annuity Income	Out-of-pocket Expenses	Medicaid Recipiency	Total Income	Out-of-pocket Expenses	Medicaid Recipiency
	HRS/AHEAD Data				MCBS Data		
1	7,740	4,820	2,550	60.9	6,750	4,050	69.9
2	10,290	8,270	4,270	28.1	10,020	5,340	41.8
3	15,500	10,900	5,050	11.0	13,740	6,470	15.5
4	19,290	14,390	6,360	5.6	19,710	7,300	8.0
5	33,580	26,300	7,000	3.0	44,150	8,020	5.4

Notes:1996-2010, for those age 72 and older in 1996.

Table 11: Income, Out-of-pocket Spending, and Medicaid Recipiency Rates: HRS versus MCBS

caid recipiency and income between the AHEAD cohort of the HRS and MCBS. For both the HRS and MCBS, we restrict the sample to singles (over the sample period) who meet the HRS/AHEAD age criteria (at least 70 in 1994, 72 in 1996, ...) and who are not working over the sample period. Because the MCBS sample lacks spousal information, for this analysis, we focus only on singles. We use De Nardi, French, and Jones's (2016) measure of permanent income and construct a measure of permanent income, which is the percentile rank of total income over the period we observe these individuals (the MCBS asks only about total income). The first four columns of Table 11 show sample statistics from the full HRS/AHEAD sample while the final three columns of the table show sample statistics from the MCBS sample. The first statistics we compare are income. Total income in the HRS/AHEAD data (including asset and other non-annuitized income) lines up well with total income in the MCBS data, although income in the top quintile of the MCBS is higher than in the HRS/AHEAD. Next, we compare out-of-pocket medical spending in the MCBS and HRS/AHEAD. Out-of-pocket medical spending (including insurance payments) averages \$2,550 in the bottom PI quintile and \$7,000 in the top quintile in the HRS/AHEAD. In comparison, the same numbers in the MCBS data are \$4,050 and \$8,020. Overall, out-of-pocket medical spending in the MCBS and HRS/AHEAD are similar, which may be surprising given that the two surveys each have their own advantages in terms of survey methodology.[17] The share of the population receiving Medicaid transfers is also very similar in the HRS/AHEAD and MCBS. 61% and 70% of those in the bottom PI quintile are on Medicaid in the the HRS/AHEAD and MCBS, respectively. In the top quintile, 3% of people are on Medicaid in the HRS/AHEAD whereas 5% are in the MCBS.

[17]There are more detailed questions underlying the out-of-pocket medical expense questions in the HRS, including the use of "unfolding brackets". Respondents can give ranges for medical expense amounts, instead of a point estimate or "don't know" as in the MCBS. The MCBS has the advantage that forgotten medical out-of-pocket medical expenses will be imputed if Medicare had to pay a share of the health event.

E Estimation methodology

We estimate the model using the extension of the *E-M* algorithm employed by Arellano, Blundell, and Bonhomme (2017). Recall from equations (8)-(10) that the functions $Q_\eta(\tau \,|\, \eta_{i,t-1}, a_{i,t})$, $Q_1(\tau \,|\, a_{i,1})$ and $Q_\varepsilon(\tau \,|\, , a_{i,t})$ are constructed from Hermite polynomials $\big(\{h_k^\eta(\cdot)\}_{k=0}^{K_\eta}, \{h_k^1(\cdot)\}_{k=0}^{K_1}, \{h_k^\varepsilon(\tau)\}_{k=0}^{K_\varepsilon}\big)$, using the coefficient functions $\{\beta_k^\eta(\tau)\}_{k=0}^{K_\eta}$, $\{\beta_k^1(\tau)\}_{k=0}^{K_1}$, and $\{\beta_k^\varepsilon(\tau)\}_{k=0}^{K_\varepsilon}$. The coefficient functions are in turn modeled with a set of polynomial splines defined over the intervals $\{[\tau_{\ell-1}, \tau_\ell]\}_{\ell=1}^{L}$, along with two tail functions for $(0, \tau_1]$ and $[\tau_L, 1)$. It is the parameters for these weighting functions that we must estimate.

Define θ as the vector of all parameters (the β parameters) in equations (8)-(10). The procedure to estimate θ is as follows. Starting with the vector $\hat{\theta}^{(0)}$, we iterate between the following two steps until $\hat{\theta}^{(j)}$ converges:

1. *Stochastic E-Step*: For each observation i, draw S values of $\eta_i^{(s)} = (\eta_{i1}^{(s)}, ..., \eta_{iT}^{(s)})$ from $f_i(\,.\,; \hat{\theta}^{(j)})$ $\big($derived from $Q_1^{(j)}(\cdot)$, $Q_\eta^{(j)}(\cdot)$ and $Q_\varepsilon^{(j)}(\cdot)\big)$.

2. *M-step*: Find

$$\underset{\beta_{\ell 0}^\eta, \ldots, \beta_{\ell K_\eta}^\eta}{\operatorname{argmin}} \sum_{i=1}^{N} \sum_{s=1}^{S} \sum_{t=2}^{T} \rho_{\tau_\ell}\left(\eta_{it}^{(s)} - \sum_{k=1}^{K_\eta} \beta_{\ell k}^\eta \, h_k^\eta(\eta_{i,t-1}^{(s)}, a_{it}) \right), \quad \ell = 1, ..., L.$$

   We use $\rho_\tau(\cdot)$ to denote Koenker and Bassett Jr's (1978) quantile "check" function. To identify the full set of splines, this function is minimized at each point ℓ on the grid over τ. The coefficients for $\varepsilon_{i,t}$ and $\eta_{i,1}$ likewise solve

$$\underset{\beta_{\ell 0}^\varepsilon, \ldots, \beta_{\ell K_\varepsilon}^\varepsilon}{\operatorname{argmin}} \sum_{i=1}^{N} \sum_{s=1}^{S} \sum_{t=1}^{T} \rho_{\tau_\ell}\left(m_{i,t} - \eta_{it}^{(s)} - \sum_{k=1}^{K_\varepsilon} \beta_{\ell k}^\varepsilon \, h_k^\varepsilon(a_{it}) \right), \quad \ell = 1, ..., L,$$

$$\underset{\beta_{\ell 0}^1, \ldots, \beta_{\ell K_1}^1}{\operatorname{argmin}} \sum_{i=1}^{N} \sum_{s=1}^{S} \rho_{\tau_\ell}\left(\eta_{i1}^{(s)} - \sum_{k=1}^{K_1} \beta_{\ell k}^1 \, h_k^1(a_{i1}) \right), \quad \ell = 1, ..., L.$$

   There are also moment conditions related to the tails of the distribution: See Arellano, Blundell, and Bonhomme (2017). These estimates give us $\hat{\theta}^{(j+1)}$.

For longer panels, settings with unbalanced data, or when estimating more complicated models the *E-step* can perform poorly when using standard samplers (e.g., Metropolis-Hastings). We therefore employ the sequential Monte-Carlo (SMC) approach implemented by Arellano et al. (2023). Comprehensive surveys of these methods can be found in Doucet, Johansen et al. (2009) and Creal (2012).

We will use the Gaussian analogues to equations (8), (9) and (10) as importance distributions.

Step 1: SMC Stochastic E-Step to sample from $f(\eta_{i,1}, ..., \eta_{i,T} | Y_i^T, a_i^T)$. For $i = 1, ..., N$:

At $t = 1$:

1. Sample S particles $\eta_1^{(s)} \sim g(\eta_1 | y_1)$, where $g(\cdot)$ is the closed form posterior from the Gaussian model.
2. Compute the weights $w_1(\eta_1^{(s)})$ and apply a self-normalization to obtain $W_1^{(s)} \propto w_1(\eta_1^{(s)})$.
3. If $Var(W^{(s)})$ exceeds some threshold, re-sample $\{W_1^{(s)}, \eta_1^{(s)}\}$ to obtain S equally weighted particles.

At $t > 1$:

1. Sample S particles $\eta_t^{(s)} \sim g(\eta_t | \eta_{t-1}, y_t)$, where $g(\cdot)$ is the closed-form posterior from the Gaussian model.
2. Compute the weights $w(\eta_{1:t}^{(s)})$ and apply a self-normalization to obtain $W_t^{(s)} \propto w_t(\eta_{1:t}^{(s)})$.
3. If $Var(W^{(s)})$ exceeds some threshold, re-sample $\{W_t^{(s)}, \eta_t^{(s)}\}$ to obtain S equally weighted particles.
4. If $t = T$, sample P particles to be used in the *M-Step*. (We set $P = 1$).

Step 2: M-Step

1. Update quantile regressions for equations (8), (9) and (10) .
2. Update Laplace parameters for the tail functions.
3. Update parameters for Gaussian proposal distributions.

F Demographic Transition Probabilities in the HRS

Let $hs_{i,j,t} \in \{0, 1, 2, 3\}$ denote death ($hs_{i,j,t} = 0$) and the 3 mutually exclusive health states of the living (nursing home = 1, bad = 2, good = 3, respectively) of household member j, household i, time t. Let $x_{i,j,t}$ be a vector that includes a constant, age, permanent income, gender, and powers and interactions of these variables, and indicators for previous health and previous health interacted with age. Our goal is to construct the likelihood function for the transition probabilities.

Using a multivariate logit specification, we have, for $q \in \{1, 2, 3\}$, $r \in \{0, 1, 2, 3\}$, we rewrite equation (11) as

$$\begin{aligned}
\pi_{q,r,t} &= \Pr(\, hs_{i,g,t+1} = r |\, hs_{i,g,t} = q; x_{i,g,t}) \\
&= \gamma_{qr} \Big/ \sum_{s \in \{0,1,2,3\}} \Big(\gamma_{qs}, \\
\gamma_{qs} &\equiv 1, \quad s = 0 \\
\gamma_{qs} &= \exp\left(x_{i,g,t}\beta_s\right), \quad s \in \{1, 2, 3\},
\end{aligned}$$

where $\{\beta_s\}_{s=1}^3$ are coefficient vectors for each future state s and $x_{i,g,t}$ is the explanatory variable vector which depends on the current state q.

The formulae above give 1-period-ahead transition probabilities, whereas what we observe in the HRS data set are 2-period ahead probabilities, $\Pr(hs_{i,g,t+2} = r|\, hs_{i,g,t} = q; x_{i,g,t})$. The two sets of probabilities are linked, however, by

$$\begin{aligned}\Pr(hs_{i,g,t+2} = r|\, hs_{i,g,t} = q; x_{i,g,t}) &= \sum_s \Pr(hs_{i,g,t+2} = r|\, hs_{i,g,t+1} = s; x_{i,g,t}) \Pr(hs_{i,g,t+1} = s|\, hs_{i,g,t} = q; x_{i,g,t}) \\ &= \sum_s \pi_{sr,t+1}\pi_{qs,t},\end{aligned}$$

imposing $\pi_{00,t+1} = 1$. This allows us to estimate $\{\beta_k\}$ directly from the data using maximum likelihood.