The Effect of Health Care Policy Uncertainty on Households' Consumption and Portfolio Choice

Thomas T. Wiemann[†] Robin L. Lumsdaine[‡]

PRELIMINARY DRAFT – PLEASE DO NOT QUOTE OR CITE WITHOUT PERMISSION

February 18, 2019

Abstract

While health care reform continues to be a major topic of policy debate in the United States, rising costs of treatment make potential medical expenditures an increasingly important contributor to households' financial risk. To investigate the effect of health care policy uncertainty (HCPU) on households, we develop a simple theoretical model that predicts a negative effect of HCPU on consumption and the relative demand for risky financial assets versus safe assets. The model illustrates that the HCPU effect increases with bad health. We combine rich longitudinal data on older Americans with Baker et al.'s (2016) HCPU index to test these claims, using a latent class model and a model-based recursive partitioning procedure adapted from the machine learning literature. Although the results provide mixed evidence for the HCPU effect on households' total spending, they indicate an important effect on portfolio choice. Further estimates corroborate the theoretical prediction that this effect is increasing in households' health problems.

[†] Department of Statistics, University of Oxford

[‡] Kogod School of Business, American University; Erasmus School of Economics, Erasmus University Rotterdam; National Bureau of Economic Research; Tinbergen Institute

Keywords: Health and Retirement Study (HRS), Consumption and Activities Mail Survey (CAMS), Latent Class Analysis, Model-Based Recursive Partitioning, Household Finance

JEL classification codes: C54, D14, D80, E21, I18

Acknowledgements: This paper builds on thesis work of the first author conducted under the supervision of the second author in partial fulfillment of the degree requirements at Erasmus University Rotterdam. We thank Susan Athey, Teresa Bago d'Uva, Steve Davis, Guido Imbens, and Alan Isaac for comments on an earlier draft, along with seminar participants at the University of Michigan's Institute for Social Research, the University of Houston, and the University of Hawaii and participants at the 2018 Canadian Econometric Study Group Conference for their comments and suggestions. Wiemann would like to acknowledge funding from the Studienstiftung des deutschen Volkes. Lumsdaine is grateful for the generous hospitality during various visits to the University of Portsmouth in contributing to the completion of this project. Any remaining errors are the responsibility of the authors alone.

1 Introduction

Nine years after the 2010 enactment of the Affordable Care Act (ACA), health care policy remains a central topic of political debate in the United States (Johnson, 2017).[1] A recent example is the heatedly fought repeal-and-replace attempt of the ACA in 2017, which was meant to fulfill a central promise of then-Candidate Trump's presidential campaign (Suderman, 2017). It is thus not surprising that Baker et al. (2016) find health care reform to be the second largest source of policy uncertainty in the US, behind only fiscal policy uncertainty. At the same time, medical expenditures are becoming an increasingly important contributor to households' financial risk: between 1990 and 2018, health spending rose from 15% to 21% of total personal consumption expenditures (Bureau of Economic Analysis, 2018). As the welfare effects of policy uncertainty are particularly substantial when the policy has a potentially large impact on households' consumption abilities (Luttmer and Samwick, 2018), health care policy uncertainty is likely to be a major source of policy uncertainty-caused welfare loss. Yet, the welfare effect of uncertainty about health care reform remains largely unassessed. This paper attempts to fill this gap. The results are interesting, not only because they shed light on the economic behavior of households, but also because they point to real macroeconomic consequences of health care policy uncertainty. The latter attaches previously-overlooked costs to the highly polarized political discussions on health care reform in recent years, which seem relevant for both legislatures and voters.

Existing economic literature suggests that households likely react to health care policy uncertainty along two dimensions: consumption and portfolio choice. The first relates to the work on *precautionary savings*. Theoretical buffer-stock models predict that households' consumption decreases when faced with an uncertainty shock, not only because their expected income could be affected, but also to self-insure against income risk (e.g., Zeldes, 1989; Kimball, 1990b; Carroll, 1997; Palumbo, 1999). The second relies on the characterization of policy uncertainty as an uninsurable risk largely beyond one's control. In the presence of such *background risk*, existing theoretical research predicts a decreased willingness of households to endure other types of risk, including rate-of-return risk (e.g., Pratt and Zeckhauser, 1987; Kimball, 1993; Gollier and Pratt, 1996). Both dimensions have been empirically investigated in the context of economic (non-health care specific) policy uncertainty. The findings of Giavazzi and McMahon (2012) and Aaberge et al. (2017) support the claim that households save more in times of political uncertainty, and Agarwal et al. (2018) and Gábor-Tóth and Georgarakos (2018) confirm the predicted decrease in investment in risky financial assets. In related work on uncertainty about Social Security benefits, Delavande and Rohwedder (2011)

[1] Obama (2016) reviews the act as the most important health care reform since the creation of Medicare and Medicaid in 1965. A reference for the law is provided under United States Congress (2010).

provide evidence that higher subjective uncertainty is negatively associated with individuals' investment share in stocks. Being a relatively recent strand of literature, however, research on the effect of policy uncertainty suffers from some qualifications.

A pivotal point of discussion in research on policy uncertainty is the identification of *causal* effects. Among the most convincing and popular identification strategies in the literature is the exploitation of variation in uncertainty exposure (e.g., difference-in-difference designs). To ensure the estimates capture the uncertainty from future policy decisions, researchers aim to identify sub-populations that differ in their reaction to the policies, but are expected to react similarly to other forms of uncertainty such as recession-caused economic uncertainty. For example, in their study of Germany's 1998 national election, Giavazzi and McMahon (2012) argue that civil servants were not directly affected by the candidates' policy differences and are thus a suitable control group for a difference-in-difference design. While the authors present supporting anecdotal evidence, some doubts remain about whether the variation in policy uncertainty exposure is unrelated to variation in exposure to economic uncertainty, and what this implies for inference on causal effects.

Additionally, current literature does not provide explicit theoretical illustrations of possible effect-channels of policy uncertainty on households' economic decisions. Instead, expectations are often based on research on the effects of general, non-policy related economic uncertainty. Especially because the distinction between policy uncertainty effects and economic uncertainty effects is essential for causal inferences in empirical research, a model that illuminates the impact of *policy* uncertainty on households' consumption and financial behavior seems overdue.

This study makes several contributions to complement existing research and address some of its shortcomings. First, to the best of our knowledge, this is the first attempt to assess the impact of health care policy uncertainty on households. For this purpose, we merge longitudinal household data from the 1994-2014 waves of the Health and Retirement Study and the 2001-2015 waves of the supplementary Consumption and Activities Mail Survey with Baker et al.'s (2016) health care policy uncertainty index. Second, this study takes a novel approach in the policy uncertainty literature by exploiting health differences for identification of causal effects. Health variation in exposure to policy uncertainty has several advantages as there is less potential for confoundedness with variation to economic uncertainty as compared to previous studies' strategies. Additional credence for causal identification is given by the high proportion of retirees in the sample which alleviates concerns that health care policy uncertainty may be capturing fears of the loss of health insurance associated with job loss, as well as by explicitly controlling for macroeconomic measures of economic uncertainty.

Third, we develop a simple two-period consumption and portfolio choice model for theoretical illustration of the potential health care policy uncertainty effect.

Finally, we employ a data-driven approach for the identification of heterogeneous causal effects. The methods, adopted from the computational statistics and machine learning literature, are particularly suitable for inference regarding empirically intangible variables such as health, for which numerous proxies exist. In particular, we use two alternatives for assessing whether the magnitude of the policy uncertainty effect is associated with households' health. The first considers a concomitant-variable latent class Tobit model as an extension of a finite mixture model to account for the substantial share of households with no investment in stocks and with an investment share in safe assets of one hundred percent. The second adapts Zeileis and Hornik's (2007) model-based recursive partitioning procedure by incorporating a Tobit model and modifying the splitting-criterion for the analysis of health care policy uncertainty. In doing so, this study contributes to the recent developments on the economic application of machine learning and causal inference (e.g., Athey and Imbens, 2017) by being the first to highlight the broad applicability of model-based recursive partitioning in the presence of asymmetric outcomes and non-binary, potentially confounded variables of interest.[2]

The results show a significant and economically relevant effect of health care policy uncertainty on the total spending of couple households but fail to identify a similar effect for single households. In contrast to this mixed evidence, the empirical results point to an important link between health care policy uncertainty and households' portfolio choice. In particular, households decrease their share in risky financial assets by one to two percentage points when faced with an uncertainty increase similar to the increase observed from 2016 to 2017. Further evidence suggests that the effect of health care policy uncertainty is more substantial for households with worse health. The findings are robust to the inclusion of household- and macro-level controls. As these estimates are comparable in size to a decrease in health (e.g., Rosen and Wu, 2004), this study indicates that health care policy uncertainty is an important determinant of households' financial behavior.

The paper proceeds as follows: Section 2 reviews relevant literature and develops a simple theoretical model on the effects of health care policy uncertainty. Section 3 describes the data and the applied transformations. Section 4 illustrates the empirical strategy. Section 5 presents the results. Section 6 discusses their implications with some suggestions for future research.

[2]Zeileis et al. (2008) are widely cited in the machine learning, statistics, and medical literature. Yet, to the best of our knowledge, the only application in economics is Wagner and Zeileis (2019) – a paper of the author in the *German Economic Review* – where the procedure is discussed in the context of linear regression.

2 Literature Review and Theoretical Framework

This research draws from, and complements, three strings of economic literature: (1) theoretical and empirical work on precautionary savings and portfolio choice in the context of income risk, (2) empirical research on the effects of health and medical expenditure risk on the financial decisions of individuals, and (3) the recent studies on the implications of policy uncertainty for households' economic behavior. This section discusses each in turn and concludes with a simple theoretical model on health care policy uncertainty and households' consumption and portfolio choice.

Precautionary savings have been extensively analyzed in the context of income uncertainty. Savings behavior unaccounted-for by conventional life cycle models was first explained by 'buffer-stock' models (e.g., Zeldes, 1989; Kimball, 1990b; Carroll, 1997). Their main prediction is that consumption is not only related to expected income, as predicted by the life cycle model, but also to higher moments such as income variance. Empirical research, however, has led to ambiguous conclusions on the importance of income uncertainty-caused savings, as estimates range from insubstantial (e.g., Skinner, 1988; Guiso et al., 1992) to economically relevant (e.g., Carroll and Samwick, 1998; Fuchs-Schündeln and Schündeln, 2005). A similarly vast literature on background risk provides insights into households' portfolio choice in an environment of multiple risks. Theoretical models of, for example, Pratt and Zeckhauser (1987), Kimball (1993), and Gollier and Pratt (1996) predict that households exposed to an undiversifiable risk are less willing to bear other types of risk, including rate-of-return risk (Goldman and Maestas, 2013). Supporting empirical evidence is provided by, for example, Guiso et al. (1996), who find that income uncertainty decreases the demand for risky financial assets.

Why might health care policy uncertainty affect the overall consumption and financial behavior of households in a similar manner to income uncertainty? As insufficient health care coverage can magnify the large out-of-pocket medical expenditures that frequently accompany health shocks, health care policy uncertainty could pose a risk to spending needs with potentially similar implications as income risk, a point made by Palumbo (1999). While health care policy uncertainty has not been analyzed *per se*, the claim that uncertainty about future spending needs affects households' consumption and financial behavior finds large support in existing research on health and medical expenditure risk.

Early research on the effect of health risk indicated a substantial negative effect on the demand for risky assets (e.g., Rosen and Wu, 2004; Berkowitz and Qiu, 2006; Edwards, 2008). However, some recent studies challenge these findings. After controlling more thoroughly for unobserved heterogeneity, no or only small effects of health risk can be found (e.g., Fan and

Zhao, 2009; Love and Smith, 2010). Because health might not only affect spending needs but also households' expected lifespan and consumption utility, the interpretation of these results provides some challenges (Smith, 1999; Edwards, 2008). For a more direct estimate on the financial burden of health, other research exploits variation in households' exposure to medical expenditure risk through their insurance coverage. Using exogenous variation in Medicaid eligibility, Gruber and Yelowitz (1999) find that households' medical expenditure risk exposure has a strong negative association with consumption.[3] With a similar identification approach, Goldman and Maestas (2013) find that relative demand for stocks increases as medical expenditure risk decreases. Further supporting evidence is provided by Atella et al. (2012), who find that health shocks have a significantly negative effect on portfolio choice in European countries without universal health care, while discovering no evidence for an effect in other European countries. Recent work by Bogan and Fertig (2013) and Toney and Bogan (2018) find the effects of mental health shocks to be similar, including when such shocks are borne by members of the extended (rather than nuclear) family.

Although health and medical expenditure risk through insurance coverage are closely related to health care policy uncertainty in their implications for consumption and portfolio choice, they are distinct sources of risk. In particular, households can affect their health (e.g., by deleterious behaviors such as smoking) and their insurance coverage (e.g., through private insurance plans), yet, individual households cannot meaningfully influence the national debate on health care reform. On one hand, analyzing macroeconomic health care policy uncertainty alleviates endogeneity concerns caused by the positive associations between households' health and insurance coverage with their wealth or education (e.g., Smith, 1999).[4] On the other hand, it raises challenges due to potential confoundedness with other macroeconomic sources of uncertainty. For example, Bloom (2014) notes that economic policy uncertainty is strongly correlated with economic business cycles.[5] As economic contractions are linked to higher economic uncertainty and lower investment prospects, their omission could lead to over-estimated effects of policy uncertainty. In general, *health care*

[3]In contrast, Starr-McCluer (1996) claims that higher savings are linked to *lower* medical expenditure risk. As risk exposure is assessed by (endogenous) health insurance enrollment, however, it is possible that the findings reflect the positive association of insurance and wealth instead.

[4]This is a substantial issue for identification in previous literature. As higher wealth and education imply higher investment in risky assets and more accumulated wealth, accurate analysis of the isolated effect of medical expenditure risk defined using health and health coverage requires inventive estimation approaches through, for example, quasi-random variation in subsidized health insurance eligibility (e.g., Gruber and Yelowitz, 1999; Goldman and Maestas, 2013) or identification of suitable instruments (e.g., Love and Smith, 2010).

[5]Specifically, Bloom (2014) states that Baker et al.'s (2016) economic policy uncertainty index is 51% higher in recessions. This makes sense intuitively as policies (and the debates surrounding them) are often reactive to recessions – e.g., the bailouts after the Great Recession. Along these lines, the author writes: "public policy that is unclear, hyperactive, or both, may raise [economic] uncertainty." (p162, second paragraph)

policy uncertainty can be expected to be less confounded by economic uncertainty. For example, in contrast to the fiscal policy analogue, *counter-cyclical* health care reform seems rare. Nevertheless, a strong claim for causal effects requires careful consideration of potentially confounding factors.

The recent literature on the effect of policy uncertainty on households provides some insights into how to address the issue of potential confoundedness in the empirical approach. Investigating Germany's 1998 national election, Giavazzi and McMahon (2012) argue that the increase in savings and labor supply they identify is caused by heightened policy uncertainty rather than economic uncertainty, as Germans were likely optimistic about the economy during that time. Unfortunately, similar anecdotal reasoning is difficult to substantiate empirically for the 21-year period considered in this study. Aaberge et al. (2017) corroborate the positive effect of policy uncertainty on households' savings by exploiting a major political shock in China. While the authors account for seasonal confounding by correlating the monthly saving differences in the year of the shock with differences in the subsequent year, it is unlikely that such a monthly effect is generally sufficient to account for confounding economic uncertainty. Alternatives are suggested in the ongoing work of Agarwal et al. (2018) and Gábor-Tóth and Georgarakos (2018), who provide evidence that policy uncertainty decreases relative demand for risky assets. In particular, Agarwal et al. (2018) control for national business cycles by exploiting temporal variation in gubernatorial elections across US states, and Gábor-Tóth and Georgarakos (2018) explicitly account for various measures of economic uncertainty such as the implied volatility index of the S&P 500 index (VIX) as suggested by Bloom (2014). As elections are not generally concerned with only a single political issue such as health care reform, the latter approach seems to be most suitable for this study to address any potential confoundedness with other forms of macroeconomic uncertainty.

As illustrated above, numerous existing studies suggest a negative effect of health care policy uncertainty on households' consumption and relative demand for risky financial assets. Yet, explicit theoretical and empirical evidence has yet to be presented. The simple household model developed in the next section is a first attempt at the former.

2.1 A Simple Household Model of Health Care Policy Uncertainty

In the following, we develop a simple two-period model linking health care policy uncertainty and households' consumption and portfolio choice. As will become evident, it is not the purpose of the proposed model to provide a complete overview of consumption and financial behavior in the presence of uncertainty.[6] Instead, its aim is to provide an illustration of a possible effect-channel of policy uncertainty and to allow for derivation of testable hypotheses.

The model draws from the general idea of Elmendorf and Kimball (2000), who develop a two-period model to investigate the effect of income uncertainty on individuals' consumption and demand for risky financial assets, but differs, in particular, in the definition of uncertainty. Instead of income risk, individuals of this model are subject to risk in spending needs due to uncertainty in health care treatment cost. Further, rather than considering a continuous random state space, we consider a discrete random variable. The latter simplification is similar to that of Delavande and Rohwedder's (2011) two-period model of portfolio choice under Social Security uncertainty.

The following structure applies. At the beginning of period 1, a representative household is endowed with initial wealth W_0 and health that requires H units of treatment. It decides how much of current wealth is consumed immediately (C_1), what share $(1-x)$ of the remaining wealth is invested in the safe asset with return b, and what share (x) is invested in the risky asset. The latter has an expected return of r_1 with probability p_1 and an expected return of r_2 with probability $p_2 = 1 - p_1$. It is assumed that the expected return of the risky asset is larger than that of the safe asset – that is, $p_1 r_1 + p_2 r_2 > b$ – and that $r_1 < b < r_2$. In period 2, in the absence of health care cost uncertainty, the risky asset's return has been determined and the household consumes its realized wealth minus its health cost given by $H * P$, where P is the cost per unit of health care treatment in period 2.

Next, consider the cost per unit of health care treatment that the household faces in period 2 to be uncertain. Specifically, suppose that the per-unit cost is P_1 with probability p_1^H and P_2 with probability $p_2^H = 1 - p_1^H$. For simplicity, it is assumed that the expected value of the cost of treatment is equal to the cost of treatment under certainty – that is, $p_1^H P_1 + p_2^H P_2 = P$ – and that $P_1 > P_2$. The random return of the risky asset and the random per-unit treatment cost are both determined at the beginning of period 2, hence, there exist four exclusive random states $(i, j) \in \{1, 2\} \times \{1, 2\}$ with corresponding probabilities $P\big[\mathcal{S} = (i, j)\big] = p_{ij} = p_i p_j^H$. For simplicity, it is assumed that there is no correlation between returns of the risky asset and per-unit cost of treatment.[7]

[6] For more thorough theoretical investigations of these topics, consider, for example, Zeldes (1989), Bodie et al. (1992), Carroll (1997), Elmendorf and Kimball (2000), or Chacko and Viceira (2005).

[7] Elmendorf and Kimball (2000), using a similar model, discuss the possibility of correlation between the rate-of-return risk and the income risk in some detail. In line with their derivations, a positive correlation between the risk in r and the risk in P magnifies the results of the model developed here.

The household is assumed to choose its initial consumption C_1 and share invested in risky assets x in order to maximize the expected utility given by $E[U(C_1) + U(C_2)]$, where C_2 is consumption in period 2 and $U(\cdot)$ is thrice differentiable in its argument.[8] In line with much theoretical work, we assume that $U(\cdot)$ is monotonically increasing and strictly concave (i.e., $U'(\cdot) > 0$ and $U''(\cdot) < 0$, with superscripts denoting the order of derivative).

Maximizing expected utility, the household solves the problem

$$\max_{(C_1,x)} E[U(C_1) + U(C_2)] = U(C_1) + p_{11}U(C_2^{11}) + p_{12}U(C_2^{12}) + p_{21}U(C_2^{21}) + p_{22}U(C_2^{22}), \quad (1)$$

where $0 \leq C_1 \leq W_0$, $0 \leq x \leq 1$, and $C_2^{ij} = (W_0 - C_1)(1 + xr_i + (1-x)b) - H * P_j$ for $(i,j) \in \{1,2\} \times \{1,2\}$. Note that the household only chooses C_1 and x, which, depending on the period 2 realization of the random states, result in a level of C_2.

It is necessary to impose further structure on the utility function $U(\cdot)$, to allow for inference on the impact of uncertainty in the per-unit cost of health care treatment. First, the utility function is assumed to display *decreasing absolute risk aversion* defined by $\frac{\partial}{\partial z}\left[\frac{-U''(z)}{U'(z)}\right] < 0$. Being a necessary condition for risky asset investment to be positively associated with wealth, this is a standard assumption with large empirical basis (Elmendorf and Kimball, 2000). Second, we assume *decreasing absolute prudence* as defined in Kimball (1990a,b) by $\frac{\partial}{\partial z}\left[\frac{-U'''(z)}{U''(z)}\right] < 0$. Decreasing absolute prudence has the implication that the absolute strength of the precautionary savings motive decreases with wealth. As discussed in Kimball (1990a,b) and Elmendorf and Kimball (2000), this assumption is not particularly constraining on the utility function – it is already satisfied by the commonly used utility functions with decreasing absolute risk aversion – and is a plausible *a priori* condition.[9] The model and the assumptions on the utility structure then imply the following results:

Proposition 1. *When exposed to uncertainty about the health treatment cost, the household reduces consumption in period 1 and reduces its relative risky asset investment (if $H > 0$).*

Proposition 2. *The magnitudes of the period 1 reductions in consumption and the share of investment in the risky asset are positively associated with bad health (i.e., higher H).*

Proof. The proofs of Propositions 1 and 2 are given in Appendix A.

[8] We assume additive separability for simplicity of the derivations for Propositions 1 and 2.

[9] Kimball (1990a) motivates decreasing absolute prudence with the following example (paraphrased and condensed from the original version): Consider a professor with a wealth of $\$10,000$ and Rockefeller with a wealth of $\$10,000,000$. If both have the same preferences except for their differences in initial wealth, who will save more after being told that there is a 50% probability that they will lose $\$5,000$ at the end of the year? As the professor is intuitively expected to save more, this provides some anecdotal evidence for decreasing absolute prudence.

This simple model points to theoretical evidence for a negative effect of health care policy uncertainty on households' consumption and relative demand for risky assets (or, equivalently, a positive effect on savings and relative demand for safe assets). Further, bad health magnifies this effect. Section 5 considers empirical evidence for these theoretical claims.

Finally, it is necessary to discuss some of the model's most important simplifications. First, there is no possibility for increased income in period two through, for example, labor supply adjustment as in Bodie et al. (1992) and Delavande and Rohwedder (2011). Allowing for this additional dimension, however, is likely to strengthen the results derived here. As healthier households are likely to be more flexible in their labor activities than those less healthy, the former can more easily compensate for potential spending need increases in period 2, and should be thus more willing to take other risks (Bodie et al., 1992). Second, it is assumed that households value consumption in periods 1 and 2 equally. If consumption in period 2 is discounted, the effect of uncertainty about the per-unit cost of treatment will be lower in magnitude. Importantly, however, the direction of the effect remains.[10] Similar arguments apply for varying degrees of risk aversion, as different intensities can affect the magnitude but not the direction of the effect when the standard assumption of risk aversion holds. Third, health is assumed to solely impact the household's medical expenditures. As illustrated by Smith (1999), worse health might also have a decreasing effect on a household's utility of consumption and expected lifespan. In this case, worse health has similar effects as lower risk aversion and time preferences for consumption in the present. Because the effect of health risk on households' savings and portfolio choice is theoretically ambiguous (Smith, 1999; Goldman and Maestas, 2013), the effect on households' reaction to health care policy uncertainty might also be ambiguous if the decrease in consumption utility and/or expected lifespan has a comparable effect as the increase in medical expenditures. As empirical research provides mixed results on which effect of health is most pivotal (e.g., Rosen and Wu, 2004; Berkowitz and Qiu, 2006; Love and Smith, 2010), it is difficult to rate this assumption *a priori* without empirical guidance. Last, households are likely to face uncertainty in health as well. For example, Deb and Trivedi (1997, 2002) suggest that households in worse health face higher health risk compared to others. While this potentially magnifies the results of Proposition 2, the total effect is again difficult to judge beforehand given the ambiguous effect of health on consumption utility and lifespan. Some insights into these concerns are gained in the empirical section (Section 5.3).

[10]While a temporal discount factor is common for multi-period models, it seems less necessary for two-period models. In particular, the two models that inspired the model presented here – Elmendorf and Kimball (2000) and Delavande and Rohwedder (2011) – do not employ a temporal discount factor. Although it is not difficult to extend the model, we opt for the simpler option as a result.

3 Data

This section describes the data used and all applied variable transformations. After an overview of the household variables, Baker et al.'s (2016) health care policy uncertainty index is illustrated. The section concludes with summary statistics.

3.1 Household Data

The unit of analysis is the household. The household data come from the Health and Retirement Study (HRS), a comprehensive, nationally representative longitudinal panel of Americans aged over age 50. The older population is a particularly relevant sample of the US population due to its large share of total wealth and financial asset investment, as emphasized by the vast literature using the HRS for analysis of precautionary savings and portfolio choice (e.g., Poterba, 1994; Rosen and Wu, 2004; Goldman and Maestas, 2013; Addoum, 2017). It is also the population for whom health care policy uncertainty is likely to be particularly salient, as individuals in the second half of their lives approach decreasing income, wealth decumulation and declining health. The consumption and financial behavior of these sample households is thus likely to have broader macroeconomic implications. As will be illustrated below, the survey collects detailed information on the health status, wealth, and demographics of respondents.

We omit the first wave of the HRS due to differences in important health variables and use the remaining eleven waves (1994-2014) as well as the eight available waves (2001-2015) of its supplement, the Consumption and Activities Mail Survey (CAMS).[11] This study relies on the versions of the HRS available as the RAND HRS Longitudinal File 2014 and the RAND HRS CAMS Spending Data 2015, which are more user-friendly than the core HRS and contain imputations for missing data including financial asset holdings. The files are merged with the Gateway to Global Aging Harmonized HRS data for additional variables on households not included in RAND HRS.[12]

3.1.1 Portfolio Choice and Financial Variables

For analyzing households' financial behavior and portfolio choice, this study follows the strategy of Rosen and Wu (2004) to collapse financial assets into four categories: safe assets (checking and savings accounts, CDs, government savings bonds and T-bills), risky assets (stocks and mutual funds), bonds, and IRA retirement accounts. The HRS also provides

[11]Specifically, wave one of the HRS does not provide comparable measures to the rest of the waves for the "activities of daily living" and "instrumental activities of daily living" measures.

[12]As a condition of use, we note that: "The HRS is sponsored by the National Institute on Ageing (grant number NIA U01AG009740) and is conducted by the University of Michigan." Citations for the Global Gateway to Aging Harmonized HRS, RAND CAMS, and RAND HRS are provided under Health and Retirement Study (2018a,b,c).

information on the latter's asset composition in the five most recent waves. However, because the empirical approach relies on temporal variation in health care policy uncertainty, omitting the earlier waves is problematic. This study's separate consideration of risky asset investments outside of retirement accounts is in line with existing literature, which points out that IRA assets may be relatively illiquid for some households (e.g., due to costs of adjusting retirement portfolios) and may suffer from measurement error (Rosen and Wu, 2004; Love and Smith, 2010).[13] Further, shares of financial assets are calculated over total financial wealth rather than all assets, as non-liquid wealth (e.g., housing wealth) is not readily adjustable to changes in background risk (Goldman and Maestas, 2013). This study focuses primarily on the shares of risky and safe assets (i.e., the first two of the four categories defined above) as the IRA account information in the HRS does not allow for sufficiently detailed risk-classification and only a fraction of financial wealth is held in bonds.

3.1.2 Consumption and Spending

Existing literature presents various approaches to defining household savings. These include wealth accumulation across survey waves (e.g., Carroll, 1997), the ratio of wealth to permanent income (e.g., Lusardi, 1998), or the difference between income and consumption (e.g., Aaberge et al., 2017). These measures are not easily implemented for the population of households surveyed by the HRS and its supplement CAMS, as a large share is retired (in the wealth decumulation phase of the life cycle) and has relatively low income from Social Security and pensions. Instead, this study exploits variation in households' consumption expenditures in line with Skinner (1988) and Gruber and Yelowitz (1999). Data for this purpose is taken from the CAMS, which documents a variety of spending variables that are not included in the core HRS. In particular, we focus on households' total spending during the previous year.[14] We use total spending in this paper because it seems less plausible that durable spending is discretionary on the extensive margin, in that if the refrigerator breaks, there is little choice but to purchase a new one. On the intensive margin, it is likely that the choice of *how much* to spend on a new durable good is affected by uncertainty, and in that sense durable consumption expenditures are particularly interesting as previous research

[13]Rosen and Wu (2004) also suggest approximating the share of stocks in IRAs using tabulations from the Survey of Consumer Finances (SCF). However, as this does not account for households' potentially heterogeneous reaction to uncertainty, attributing a fixed percentage of IRAs to the risky asset share would bias results on heterogeneous effects. In early results, we followed the (Rosen and Wu, 2004) approach and attributed 11% of IRAs to stocks on the basis of tabulations from the 2013 and 2016 SCFs for respondents aged over 50. Doing so did not affect the conclusions.

[14]In his thesis, Wiemann (2018) used durable goods, which includes household appliances such as TVs and kitchen equipment but excludes car purchases; for reference those results are included in the Appendix. The RAND HRS CAMS data does not report separate car purchases, but instead combines these in a 'transport spending' category including other forms of non-durable spending such as fuel and public transport. Early results with merged durable and transport spending led to similar conclusions (although of higher magnitude), but this strategy was ultimately abandoned due to imprecision in interpretation of the transport category.

describes this category as more volatile and more reactive to economic downturns such as recessions (Romer, 1990; Attanasio, 1999; Crossley et al., 2013).

3.1.3 Health

To adequately assess potential heterogeneity in the effect of health care policy uncertainty conditional on households' health, a suitable measure of health status is needed. Empirically, health is an intrinsically unobserved variable and a variety of proxies have been suggested in previous research (Currie and Madrian, 1999).

Many studies (including the HRS) assess health using survey respondents' answer to a self-reported question of the form "Would you say your health in general is (1) excellent, (2) very good, (3) good, (4) fair, or (5) poor?"[15] Self-reported health measures such as this 5-point Likert scale variable often have been shown to be highly correlated with medically-determined health status, and have thus been argued to be an excellent proxy for health (e.g., LaRue et al., 1979; Ferraro, 1980). Yet, some doubts remain about the role of potential reporting bias (e.g., Bago d'Uva et al., 2008; Lumsdaine and Exterkate, 2013), and whether this is confounded by, for example, wealth or income (Currie and Madrian, 1999).

As an alternative, some studies have employed more objective measures of health. For example, Wu (2003) and Berkowitz and Qiu (2006) suggest using exogenous health shocks given by severe health conditions reported between survey waves. These shocks are defined by a diagnosis of diabetes, lung disease, cancer or malignant tumor growth, or experiencing a stroke or heart problems.[16] Other frequently-used measures are the proxies given by self-reported limitations in Activities of Daily Living (ADLs) and Instrumental Activities of Dailiy Living (IADLs). Five and three of these, respectively, are asked of the respondents in the HRS, capturing possible impairments of respondents to (1) bathe, (2) eat, (3) dress, (4) walk across a room and (5) get in and out of bed, as well as (1) use a telephone, (2) take medication, and (3) handle money. To complement these measures of physical ability, we follow Fan and Zhao (2009) and consider a separate variable to capture a respondents' mobility. In particular, we count difficulty of (1) walking one block, (2) sitting for two hours, (3) getting up from a chair, (4) climbing a flight of stairs, (5) stooping or kneeling, or (6) lifting and carrying 10 lbs.[17] Contrasting Fan and Zhao (2009), this study does not consider self-reported limitations to work since doing so would cause a substantial reduction

[15] Due to the additional dimensionality and computational burden that using indicator variables for each of the five categories would entail, and because our primary interest lies in inferring the direction of association between health and the health care policy effect, we use the original categorical variable.

[16] 'Mild' conditions such as high blood pressure and arthritis also are reported in the HRS. Following the literature, they are not included in the analysis.

[17] The answer choices "Can't do.' and "Don't do." are treated as if the respondent has difficulties with the particular activity.

in the number of usable observations for analysis.[18] In addition, previous literature suggests substantial drawbacks to this variable. For example, Lindeboom and Kerkhofs (2009) point to endogenous, state-dependent reporting bias in self-assessed work limitations.

In summary, the analysis in this paper uses (1) the HRS' categorical self-reported health measure, (2) a count variable of the number of limitations reported in the HRS's *ADL* and *IADL* measures, (3) a count variable of the historical number of severe health conditions, and (4) a count variable of the number of limitations to mobility, as proxy measurements for health status.

Additionally, two variables to assess health care coverage and service utilization are constructed: a dummy variable indicating whether the respondent is uninsured – that is, not covered through private or governmental health insurance – and a count of the number of nights spent in a hospital during the previous two years. Typically, it is difficult to sign the effect of medical expenditure risk on service utilization *a priori* (Zweifel and Manning, 2000), as greater utilization could stem from either more severe conditions or more favorable health care plans with lower service costs (Smith, 1999; Gruber, 2006). To avoid this ambiguity, we use only the number of nights spent in a hospital as the sole measure of utilization rather than also including doctor visits. Because hospital utilization also may reflect income and insurance constraints, we focus on the intensive margin (i.e., the number of nights) rather than the extensive (i.e., whether or not a hospital stay occurred) as the intensive margin more likely captures the presence of health emergencies.[19] While these two measures of health care coverage and service utilization are substantially simplified and cannot fully reflect varying health insurance plans and service utilization in detail, a consideration of this issue similar to that of, for example, Rosen and Wu (2004) and Edwards (2008), seems to be preferred to no consideration, as in much of the literature on the effects of health on portfolio choice (e.g., Wu, 2003; Berkowitz and Qiu, 2006; Love and Smith, 2010).

3.1.4 Final Analysis Sample

Two important data issues arise. The first is the varying household composition across the sample. A common solution is to divide the data into single and couple households (e.g., Wu, 2003; Rosen and Wu, 2004; Berkowitz and Qiu, 2006; Love and Smith, 2010). We follow

[18]Specifically, there is no information on self-reported limitations to work for 5,658 and 3,566 single and couple households, respectively, in the portfolio choice sample from the HRS data. Similarly, for the consumption information in CAMS, these numbers are 1,313 and 465, respectively. In both cases, omitting these observations would correspond to an additional sample size reduction of between 5-10%.

[19]For example, Miller (2012) finds no effect of insurance reform in Massachusetts on non-preventable emergencies but a substantial reduction in non-urgent visits, particularly during hours that doctor's offices are usually open. Litwin and Sapir (2009) also document that older adults in Europe and Israel forego health care utilization due to cost considerations.

this approach rather than including household size as a control variable as in Fan and Zhao (2009) due to intuitively different consequences of health shocks between household types.[20] The second issue relates to the definition of health status for couple households. Rosen and Wu (2004) suggest considering the health of husband and wife separately, Atella et al. (2012) take the average of the health measures across spouses, and Christelis et al. (2010) entirely omit health status of non-respondents. An appealing compromise between these suggestions is given by Coile and Milligan (2009) and Love and Smith (2010), who characterize households on the basis of the least healthy spouse. With the aim of identifying an ill individual (rather than accurately assessing the household's comprehensive health), we follow their suggestion and define the health measures introduced above via the maximum value across spouses for couple households. Similarly, we consider a household to be uninsured if either of the spouses is not covered by private or governmental health insurance.

Apart from the main variables of interest, we compile a set of household-level control variables. Similar to Rosen and Wu (2004) and others, household-level controls include age and the number of years in retirement, as well as dummy variables for five educational attainment categories, race categories (described below), the presence of living children, whether the respondent is retired, and wave-specific wealth and income quartiles. Gender is also included as a control when analyzing single households. For couple households, the controls for the attained education, age, and years of retirement are constructed as the maximum of both spouses. Further, six race dummies for couples are constructed, based on all possible combinations of the three categories "White," "Black," and "Other" across the two spouses.[21] Additional control variables such as labor force status and spouse-demographics were considered, but were ultimately omitted from the analysis. While expansive control specifications are unproblematic for standard econometric analysis given the large sample size, estimation of the latent class model introduced in Section 4.2.1 is computationally infeasible when too many parameters are considered simultaneously. Fortunately, the restriction does not seem particularly binding as the inclusion of the additional covariates did not have any substantive effect in the baseline fixed effect analysis.

Starting with the initial HRS sample of 68,507 single and 72,426 couple households that have a non-zero household sampling weight, we exclude 16,084 and 8,702 observations, respectively, that have no positive holdings of financial assets.[22] An additional 6,341 and 6,547

[20]For example, couple and single households seem to follow systematically different investment strategies (Addoum, 2017), potentially due to differences in risk sharing (Mazzocco, 2004).

[21]With the aim of capturing racial composition of the household only, we disregard which spouse is of which race.

[22]In doing so, we follow Rosen and Wu (2004) who also consider households conditional on holding positive financial assets.

single and couple household observations, respectively, are omitted from the analysis due to missing one or more of the above variables. Finally, households with an observation in only one wave (3,296 and 2,460 observations for single and couple households, respectively) are excluded to ensure that the samples will be identical across the fixed effect and pooled models. This leaves a sample of 42,786 single and 54,717 couple household-wave observations for the portfolio choice analysis. The analysis on precautionary spending reductions relies on the CAMS subsample (a random subset of the HRS) of 17,847 single and 11,811 couple household-wave observations with a non-zero household sampling weight. In this subsample, 5,718 and 3,941 (single and couple) household-wave observations are excluded from the analysis due to missing variables.[23] This leaves 12,129 and 7,870 single and couple household-wave observations, respectively.[24] Appendix B provides a table documenting the amount of deleted observations at each stage.

Summary statistics of household characteristics are presented in Table 1. Nominal financial values are converted to 2010 US dollars using the Consumer Price Index corresponding to the 12 months preceding the respective interview's end-date.[25] The risky asset share is substantially concentrated at the 0-threshold: 62% (75%) of couple (single) households do not record positive investment share in risky assets. In contrast, less than 3% of both household types have a 0-share of safe assets, but 35% (57%) of couple (single) households have all their financial wealth in safe assets. Section 4 illustrates the necessary empirical considerations to address this issue.

3.2 Measuring Health Care Policy Uncertainty

Existing literature presents several measures of policy uncertainty, with much of the current work relying on subjective policy uncertainty (e.g., Delavande and Rohwedder, 2011; Luttmer and Samwick, 2018) or on uncertainty around elections and other significant political events (e.g., Giavazzi and McMahon, 2012; Aaberge et al., 2017; Agarwal et al., 2018). Recently, Baker et al. (2016) take an alternative approach and develop a computer-driven, news-based policy uncertainty index, based on the proportion of articles with word triplets "uncertain", "economic", and "policy" (and their synonyms) in ten major newspapers in the

[23]For single and couple households of the CAMS sample, we exclude 4,022 and 2,837 due to non-positive or missing financial assets, and an additional 733 and 621 due to missing one or more of the above variables. Finally, 963 and 483 observations are omitted as the corresponding households are only reported in one wave.

[24]The panel is unbalanced – that is, many households are not participating in all survey waves. Fortunately, a balanced sample is not required as the identification approach is based on correlating the cross-section and inter-temporal variation of consumption/financial asset shares with inter-temporal variation of health care policy uncertainty. Further, there is no evidence to suggest that surviving households react more strongly to health care policy uncertainty shocks (in line with Love and Smith's (2010) claim in the context of health shocks).

[25]The corresponding data citation is provided under United States Bureau of Labor Statistics (2018).

US. The policy uncertainty index has been shown to be well representative of policy uncertainty in Baker et al. (2016), and has since been used in a vast amount of policy uncertainty literature, in particular to examine the effect of policy uncertainty on firms (e.g., Bloom, 2014) and on macroeconomic tendencies (e.g., Stock and Watson, 2012). To the best of our knowledge, this paper is just the second to apply Baker et al.'s (2016) policy uncertainty indices in a study of households, preceded only by Gábor-Tóth and Georgarakos's (2018) ongoing work.

Fortunately for our examination of health care policy uncertainty, Baker et al. (2016) developed not only a general economic policy uncertainty index but also several categorical indices, which are constructed using the Access World News newspaper archive of about 1,500 US papers. Among these is a health care policy uncertainty index, capturing the frequency of articles that feature the above word triplets as well as one term related to health care such as "Medicaid", "health insurance" or "Obamacare".[26] Figure 1 shows the health care policy uncertainty index between 1992 and 2017, and its average over the preceding 12 months.[27] The average is characterized by four substantive increases that can be linked to health care reform efforts in the US. Note that the health care efforts of the Trump administration are not considered for estimation, as the HRS sample only provides observations up-to and including 2015. We merge the household data with the 12-month average preceding the end date of the month of each individual's/household's HRS interview. The CAMS questionnaire postdates the HRS interviews by several months, but unfortunately, the specific survey month is unknown. With the intention of approximating the date closest to the actual reply of the respondents, each household is matched to the 12-month average preceding November of the respective year as the end of October marks the mid-point of the CAMS' 4-month collection period.

[Figure 1 about here.]

Finally, additional macroeconomic uncertainty measures are merged with the data to address endogeneity concerns and confirm the robustness of the results. In particular, we follow Gábor-Tóth and Georgarakos (2018) and include both the 12-month average of the S&P 500

[26] Additional terms are "health care", "Medicare", "malpractice tort reform", "malpractice reform", "prescription drugs", "drug policy", "food and drug administration", "FDA", "medical malpractice", "prescription drug act", "medical insurance reform", "medical liability", "part d", and "affordable care act." A reference to the index is provided under Baker et al. (2018).

[27] Baker et al. (2016) provide a similar figure of the health care policy uncertainty index until January 2015. We update the authors' figure through December 2017, and add the series' average.

Index growth rate and the Implied Volatility Index (VIX) to account for the first and second moment of financial returns.[28] To capture general economic uncertainty, the 12-month growth rate average of the Conference Board's Coincident Indicator is also included in all specifications. The measure is a weighted average of different macroeconomic variables that typically move in synchronization with the business cycle (e.g., the industrial production index). As such, it not only moves strongly pro-cyclically but also smooths out some of the fluctuations of the individual variables. These features allow researchers to better account for business cycles than individual macroeconomic series would (The Conference Board, 2001).[29] Additional controls including the National Bureau of Economic Research's (NBER's) recession indicator, the Conference Board's Leading Economic Indicator, and Baker et al.'s (2016) fiscal policy uncertainty index were also considered, but are ultimately omitted to keep the coefficient vector concise for reasons discussed above. The results are robust to the inclusion of these additional variables.

Sample summary statistics of the uncertainty indices and macroeconomic variables for the single and couple household subsamples also are included in Table 1.

[**Table 1 about here.**]

4 Methodology

This section describes the empirical approach to analyze the effect of health care policy uncertainty on households' consumption and portfolio choice. Following Honoré's (1992) censored fixed effect model, two heterogeneous effect models – a latent class Tobit model and an adaptation of Zeileis et al.'s (2008) model-based recursive partitioning – are introduced. Throughout, the discussion focuses on the model specifications and their interpretation. Further details, for example describing the specific EM algorithm employed, are provided in Appendix C.

4.1 Censored Fixed Effect Model

As noted in Section 3.1.4, the risky asset share is concentrated at zero and the safe asset share is concentrated at one.[30] Unlike for the analysis of total spending, this asymmetric nature renders standard linear regression estimation unsuitable for the portfolio choice variables. Previous research on portfolio choice provides some guidance on useful techniques for

[28]Both measures are the last price traded (PX_LAST) as taken from Bloomberg (2018a,b), with corresponding data citations in the references.

[29]The corresponding data citation is provided under The Conference Board (2018).

[30]While both the risky asset share and the safe asset share are also bounded by one and zero, respectively, no substantial clustering occurs at these other bounds (i.e., less than 1% of the sample is concentrated at zero for risky assets, and less than 3% at one for safe assets). Following existing literature, we thus consider a one-sided censored regression model.

addressing this bunching. A common solution is a Tobit model (e.g., Wu, 2003; Cardak and Wilkins, 2009), with some examples of random effects extensions (e.g., Rosen and Wu, 2004; Berkowitz and Qiu, 2006). However, as Fan and Zhao (2009) and Love and Smith (2010) note, the consistency of the estimates in the random effects models depends crucially on the assumption that the unobserved heterogeneity is uncorrelated with covariates. As this is likely violated, researchers sometimes instead turn to a fixed effects model. Unfortunately, coefficient estimates of a fixed effects Tobit model optimized via maximum likelihood are inconsistent due to the 'Incidental Parameter Problem' (Honoré, 1992; Greene, 2004).[31]

A suitable alternative is suggested by Honoré (1992), who develops a semiparametric censored fixed effects estimator. Consider a model censored at 0 of the form

$$y_{it} = \max\{0, \beta_{i0} + X_{it}\beta_1 + \varepsilon_{it}\}, \tag{2}$$

where $i \in \{1, \ldots, N\}$ and $t \in \{1, \ldots, T\}$ denote the household and time period, respectively, y_{it} is the censored outcome variable, X_{it} is a set of covariates with corresponding coefficient vector β_1, β_{i0} is the household-fixed effect, and ε_{it} is the error term. Honoré (1992) derives a two-period panel estimator and notes it can be extended in a straightforward manner to a multi-period panel setting by taking into account all time period pairs. In particular, if the pair $(\varepsilon_{it}, \varepsilon_{is})$ is identically distributed as the pair $(\varepsilon_{is}, \varepsilon_{it})$ conditional on $(X_{it}, X_{is}, \beta_{i0})$ for $t, s \in \{1, \ldots, T\}$, then β_1 can be estimated by:

$$\begin{aligned}\hat{\beta}_1 = \arg\min_b \sum_{i=1}^{N} \Bigg(\sum_{t=1}^{T-1} \sum_{s=t+1}^{T} \Bigg(&(\max\{y_{it}, \Delta X_i b\} - \max\{y_{is}, -\Delta X_i b\} - \Delta X_i b)^2 \\ &+ 2I[y_{it} < \Delta X_i b](\Delta X_i b - y_{it})y_{is} + 2I[y_{is} < -\Delta X_i b](-\Delta X_i b - y_{is})y_{it} \Bigg)\Bigg),\end{aligned} \tag{3}$$

where $\Delta X_i = X_{it} - X_{is}$ and I is an indicator function.

The estimator is consistent as $N \to \infty$ rather than $T \to \infty$ as in the maximum-likelihood case (Stephens, 2002) and is hence suitable for a large longitudinal panel such as the HRS. Asymptotic standard errors are computed using the first order condition of equation (3). While notation corresponds to a balanced panel for brevity, the estimator is also suitable for unbalanced panels (Honoré, 2002; Love and Smith, 2010). The model does not require a parametric specification of the error distribution, however, pairwise exchangeability conditional on the explanatory variables is required. In other words, while errors are potentially

[31] Greene (2004) contrasts previous conclusions on the fixed effect Tobit model by finding insubstantial bias in its slope coefficients despite moderate sample size. He corroborates, however, that the model's variance coefficient and standard errors show bias in finite samples.

heteroskedastic across households, they are assumed to be uncorrelated over time. The latter is often not satisfied in panel data,[32] yet, no suitable and consistent estimator of censored regression models seems to exist where this is not a required assumption.[33] For anecdotal evidence that this is not a major concern for the data at hand, we compare conventional standard errors with household-level clustered standard errors of the linear fixed effect model and find no substantive differences (see Appendix C.1 for results). While not without qualifications, Honoré (1992) thus seems to provide the most suitable model for this study.

Unlike with a linear regression model, the resulting estimates cannot readily be interpreted as marginal effects. Recall that for a standard Tobit model, marginal effects would be given by

$$\frac{\partial E[y_{it}|X_{it}, \beta_{i0}, \beta_1]}{\partial x_{k,it}} = \Big[1 - \Phi\Big(-\frac{\beta_{i0} + X_{it}\beta_1}{\sigma}\Big)\Big]\beta_{1k}, \tag{4}$$

where the term in parentheses is the household's probability of a positive y_{it}, and β_{1k} denotes the coefficient corresponding to the kth covariate in X_{it}. Because β_{i0} is 'differenced-out' by Honoré's (1992) estimator, the magnitude of the marginal effects cannot be calculated.[34] While Honoré (2002) notes that β_1 nevertheless captures the marginal effect of covariates on the underlying latent variable y^*, this interpretation is not particularly relevant when interested in spending and investment reductions. Due to this complication, Hochguertel (2003) and Love and Smith (2010) use Honoré's (1992) estimator alongside a Tobit model; the former acting as a robustness check on the more easily interpretable results of the latter. We follow the same approach.

In summary: to empirically test Proposition 1 – health care policy uncertainty negatively affects household consumption and relative demand for risky assets – we regress households' risky and safe asset shares on the log-transformed 12-month health care policy uncertainty average and a set of controls as introduced in Section 3 using both Honoré's (1992) censored fixed effect model as well as a pooled Tobit model.[35] Using a conventional fixed effect linear model, we estimate a similar model for households' log-transformed total spending. The

[32]Bertrand et al. (2004) show that the neglect of error-correlation in linear panel models leads to misspecified standard errors in many empirical studies.

[33]Hochguertel (2003) uses a Tobit model with a parametric variance specification; however, this seems infeasible for our study due to the large number of coefficients in the case of the latent class model introduced in Section 4.2.1.

[34]Note, the sign of the marginal effect and β_{1k} are equal as $\Phi(\cdot) \in [0, 1]$, also for Honoré's (1992) estimator.

[35]Honoré's (1992) estimator is implemented using the Stata code available at his website: `https://www.princeton.edu/~honore/stata/`. A random effects Tobit model as in Rosen and Wu (2004) could also be considered. However, the pooled model leads to consistent estimates of the slope parameters in a random effects model as the effects are additive (Cameron and Trivedi, 2005). Without apparent benefit, this study thus opts for the simpler model. For computational ease, we use one minus the share of safe assets in estimation. Coefficients can be re-adjusted by simple sign-reversal.

economic and statistical significance of the health care policy uncertainty coefficients is used to evaluate the theoretical claims.

4.2 Heterogeneous Effects

The fixed effect models of the previous section control for time-varying household-level variation, time-invariant unobserved heterogeneity, and potentially confounding economic uncertainty using various macro-level controls. Similarly careful empirical specifications are often said to allow for causal inference (e.g., Giavazzi and McMahon, 2012; Baker et al., 2016; Aaberge et al., 2017). To further strengthen the claim that the estimates capture a causal effect of health care policy uncertainty on households' consumption and portfolio choice, we proceed by considering Proposition 2 – that bad health magnifies the effect of health care policy uncertainty.

This identification strategy is similar to that of, for example, Baker et al. (2016), who argue that firms with higher dependency on government contracts should be more affected by policy uncertainty, or Giavazzi and McMahon (2012), who make a similar claim about non-civil servants in Germany. However, for this approach to contribute to the story of causal effects, it is crucial that policy uncertainty exposure is not confounded with exposure to other forms of uncertainty. Civil servants in Germany, for example, are also much less affected by business cycles due to substantial restrictions on their dismissal, which raises questions as to whether the policy uncertainty estimates may be spurious. Compared to previous studies' strategies, variation in households' health seems to be a particularly suitable measure to capture uncertainty exposure across households. As illustrated in Section 2.1, poor health is likely to have an amplifying effect on the impact of health care policy uncertainty, and existing literature does not provide evidence that health affects households' response to, for example, economic downturns. While one channel of potential endogeneity concern may arise because health insurance in the US is often linked to employment, a large share of the sample is retired and is thus not at risk of greater exposure to medical expenditure risk due to potential job loss. Further, as illustrated by Smith (1999), health shocks will not substantially alter labor supply during retirement, and income from Social Security and pensions will remain fixed (i.e., not related to health). As a result, health of older Americans is unlikely to affect their responsiveness to general economic uncertainty. Remaining doubts are addressed by controlling for measures of economic uncertainty as before.

The empirical approach of the existing policy uncertainty literature for analyzing heterogeneity in causal effects is restricted to a multiplicative-interaction specification of the form

$$y_{it} = \beta_{i0} + \beta_1 PU_t + \beta_2 PU_t z_{it} + X_{it}\beta_3 + \varepsilon_{it}, \tag{5}$$

for $i \in \{1, \ldots, N\}$ and $t \in \{1, \ldots, T\}$, where PU_t is the policy uncertainty measure at time t, z_{it} is the exposure variation measure of individual i at time t, and X_{it} are control variables with $z_{it} \in X_{it}$ (e.g., Giavazzi and McMahon, 2012; Baker et al., 2016). While this functional form allows for a simple interpretation of β_2, it comes with some crucial caveats. Most importantly, the multiplicative-interaction model restricts parameter heterogeneity to be of a pre-specified additive form. This is particularly problematic when heterogeneity is expected to stem from complex characteristics such as health, where an array of potential variables is available and it is unclear *a priori* which combinations, interactions and transformations are most appropriate. Because of the numerous possible variations, it is difficult to motivate this approach against the concerns raised by Hainmueller et al. (2018), who show that a substantial share of heterogeneous treatment effect results in leading political science journals seems to be highly model dependent.

To address these qualifications, this study considers two data-driven alternatives for identifying heterogeneous effects: a latent class model and an adaptation of Zeileis et al.'s (2008) model-based recursive partitioning procedure. As neither model requires pre-specification of the functional form of the heterogeneous effect, both are well suited for analyzing possible variation in the effect of health care policy uncertainty with respect to the six health and health care measures. Further, both approaches allow approximation of possible non-linearities in the effect of health care policy uncertainty by segmenting the population into homogeneous sub-populations with distinct coefficients. Despite their similarities, however, each has considerable advantages. As Frick et al. (2014) discuss, latent class models can detect sub-populations when effects are substantially heterogeneous, even when there is only a weak (or no) association between coefficient differences and observed variables. In contrast, the model-based recursive partitioning method requires strong association of sub-populations with observed covariates, but is able to identify parameter heterogeneity in much more detail when this is the case. Because existing literature provides ambiguous guidance about the extent of association of potential sub-populations and observed health variables, it is advantageous to consider both approaches.[36]

Unfortunately, neither Honoré's (1992) censored fixed effect model nor its linear regression equivalent can be employed for either the latent class model or the model-based recursive partitioning procedure for two reasons. First, its solution turns out to be highly computationally intensive. As both heterogeneous effect models require plentiful re-estimation, the necessary computational time of either approach exceeds many days on modern workstations. Second,

[36]As illustrated in Section 3.1, existing literature finds observed health variables to be highly correlated with households' health. However, Deb and Trivedi (1997, 2002) suggest that health care utilization-caused expenditures largely depend on unobserved heterogeneity (and propose a latent class model as a consequence).

and more importantly, the fixed effect estimators do not permit the calculation of likelihood values as the household-fixed effects are 'differenced-out'. As illustrated below, these fixed effects are essential for the calculation of the posterior component-specific likelihoods of the latent class model as well as the split-criterion of model-based recursive partitioning.[37] For the advanced investigations into heterogeneous effects, we thus reconsider pooled regression models without fixed effects. At least to some degree, however, the fact that the sample is split into sub-populations corrects for some unobserved heterogeneity across households.

The following two subsections illustrate the latent class model and the model-based recursive partition procedure with particular focus on their respective contributions to analyzing heterogeneity in the effect of health care policy uncertainty.

4.2.1 Latent Class Model

Latent class models were first introduced to health economics by Deb and Trivedi (1997), and have since been established as important tools for analyzing heterogeneous health care demand (Deb and Trivedi, 2002; Bago d'Uva, 2006).[38] As an advantage over conventional models, they do not require a fixed distinction between households based on observed variables. Instead, latent class models allow the coefficient-heterogeneity to stem from unobserved characteristics such as a households' lifestyle, attitudes to health risk, and long-term health. While the observed health variables might be important determinants of the unobserved characteristics, Deb and Trivedi (1997, 2002) suggest that such proxies may not fully capture heterogeneity from this source. Households who might be characterized as healthy based on the observed variables might nevertheless react strongly to health care policy uncertainty , for example, due to health risk attitude. A latent class model seems particularly tenable for analyzing potential unobserved heterogeneity.[39] A particularly appealing variant is the concomitant-variable latent class model developed by Dayton and Macready (1988). In addition to identifying latent sub-populations and estimating an econometric model within each

[37] Deb and Trivedi (2013) develop a latent class model with fixed effects for normal linear regression models, however, the proposed estimator does not allow for households to transition between classes (i.e., transition from 'healthy' to 'ill' and vice versa).

[38] In the empirical literature on health care utilization, latent class models are often employed as more flexible alternatives to two-part models (e.g., zero-inflated Poisson models). Further examples include Deb and Trivedi (1997), Jiménez-Martín et al. (2002), Atella et al. (2004), and Bago d'Uva and Jones (2009). In a recent study, Gil et al. (2018) extend this methodology by considering a hidden Markov model to investigate dynamic unobserved heterogeneity between households.

[39] From an econometric perspective, latent class models also alleviate concerns about potentially misspecified underlying probability densities. Being more flexible than standard regression models, they can serve as a better approximation to any unknown probability density (Laird, 1978). These features have also motivated the application of latent class models to fields with similar aims of uncovering population heterogeneity as this study, for example, marketing research (e.g., Kamakura and Russell, 1989; Wedel et al., 1993).

segment, it also allows for simultaneous estimation of the association between class membership and observed variables. It has proven valuable in the characterization of heterogeneity in the literature on health care demand utilization (e.g., Bago d'Uva and Jones, 2009), and thus seems to be a promising method for analyzing whether the observed health measures are associated with any heterogeneity in the effect of health care policy uncertainty. For ease of exposition, we omit the term "concomitant-variable" from the model name but emphasize that it is implied in all references to latent class models henceforth in the paper.

In contrast to the standard assumption of one underlying econometric model, a latent class model considers the possibility that the data consists of several unobserved segments, each with similar distributional form but with heterogeneous parameters (Aitkin and Rubin, 1985). Following the latent class interpretation of mixture models (e.g., McLachlan and Peel, 2004), an M-component latent class model can be written as:

$$y_{it} = \begin{cases} f(y_{it}|X_{it}, \theta_1), & \text{if } S_{it} = 1 \\ \vdots & \\ f(y_{it}|X_{it}, \theta_M), & \text{if } S_{it} = M \end{cases} \tag{6}$$

for $i \in \{1, \ldots, N\}$ and $t \in \{1, \ldots, T\}$, where y_{it} is the outcome variable with component-specific densities $f(\cdot|X_{it}, \theta_s)$ with X_{it} being a variable vector and θ_s being the component-specific coefficient vector.[40] $S_{it} = s$ indicates that observation (it) originates from sub-population s. Although the component labels (s) are unobserved, it is possible to estimate the model by considering S_{it} as realizations of the discrete random variable $\tilde{S}$ with corresponding probabilities $P[\tilde{S}_{it} = s] = \pi_{s,it}$ that satisfy $\sum_{s=1}^{M} \pi_{s,it} = 1$. The resulting likelihood function is given by

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{N} \prod_{t=1}^{T} \Big(\sum_{s=1}^{M} \pi_{s,it} f(y_{it}|X_{it}, \theta_s) \Big), \tag{7}$$

where $\boldsymbol{\theta}$ summarizes the component specific coefficients. The mixture densities $f(y_{it}|X_{it}, \theta_s)$ encapsulate a normal density for the analysis of households' total spending and a pooled Tobit density for the analysis of portfolio choice variables to account for the asymmetry of these outcome variables. The pooled Tobit densities are given by

$$f(y_{it}|X_{it}, \theta_s) = \left(\Phi\Big(\frac{-X_{it}\beta_s}{\sigma_s} \Big) \right)^{I[y_{it}=0]} \left(\frac{1}{\sigma_s \sqrt{2\pi}} \exp\Big(\frac{-1}{2\sigma_s^2} \big(y_{it} - X_{it}\beta_s \big)^2 \Big) \right)^{I[y_{it}>0]}, \tag{8}$$

for $s \in \{1, \ldots, M\}$, $i \in \{1, \ldots, N\}$ and $t \in \{1, \ldots, T\}$, where Φ denotes the normal CDF. Here, θ_s summarizes the parameters β_s and σ_s.

[40]It is now also clear why variable parsimony is particularly important in a mixture model setting: each added covariate requires estimation of M additional coefficients.

Apart from identifying potentially heterogeneous θs and the underlying sub-populations, there are two possibilities for additional inference regarding the association between households' health measures and their component membership. First, instead of modelling the *prior* probability of belonging to a particular class as an invariant proportion (e.g., Deb and Trivedi, 1997), we parametrize it as a function of time-variant observed health variables.[41] Here, 'prior' highlights that y_{it} has not been observed yet. Following Dayton and Macready (1988), the prior probability of belonging to component s is modelled as a multinomial logit given by

$$\pi_{s,it} = P[S_{it} = s|Z_{it}] = \frac{\exp\left(\alpha_{s,0} + Z_{it}\alpha_{s,1}\right)}{\sum_{j=1}^{M}\exp\left(\alpha_{j,0} + Z_{it}\alpha_{j,1}\right)}, \tag{9}$$

for $s \in \{1,\ldots,M\}$, $i \in \{1,\ldots,N\}$ and $t \in \{1,\ldots,T\}$, where Z_{it} is a $1 \times K$ vector of concomitant variables with corresponding component-specific coefficient vector $\alpha_{s,1}$. The coefficients of component 1 are set to 0 for identification purposes (i.e., $\alpha_{1,0} = 0$ and $\alpha_{1,1} = 0$).[42] Taking Z_{it} to be the household health variables and letting $z_{k,it}$ with $k \in \{1,\ldots,K\}$ denote its kth element, one can make inferences about the characteristics of the subgroups by examining their partial effects on the prior probability. In a two-component mixture, for example, a positive and significant coefficient $\alpha_{2,1k}$ indicates that households with higher $z_{k,it}$ have a higher probability of belonging to the second component with component-specific coefficients θ_2. Empirical evidence for Proposition 2 is thus found if estimation results in a higher health care policy uncertainty coefficient corresponding to the component with higher prior probabilities of 'ill' households.

Second, it is possible to calculate an observation's *posterior* probability of stemming from the mixture component s – that is, the probability of belonging to a particular component *after* having observed y_{it}. This can be calculated by

$$\tau_{s,it} = P[S_{it} = s|y_{it}, X_{it}, Z_{it}, \boldsymbol{\theta}] = \frac{\pi_{s,it} f(y_{it}|X_{it},\theta_s)}{\sum_{j=1}^{M}\pi_{j,it} f(y_{it}|X_{it},\theta_j)}, \tag{10}$$

for $s \in \{1,\ldots,M\}$, $i \in \{1,\ldots,N\}$ and $t \in \{1,\ldots,T\}$. See, for example, McLachlan and Peel (2004). In contrast to the prior probabilities, the posterior probabilities do not allow for statistical tests on the association between components and features Z_{it} in a straightforward manner without disregarding the uncertainty in determining $\pi_{s,it}$ and θ_s (Kamakura

[41]Bago d'Uva and Jones (2009) parametrize the class membership probabilities as a function of time-*invariant* variables. To allow for households' transition between latent classes depending on their health at each survey, we consider time-variant variables instead.

[42]This follows from the same reasoning as for identification in standard Logit models, where a normalization is also necessary because by design the model describes the distinction between two options in relative terms. Here, setting the coefficients of the first component to zero is common practice.

et al., 1994). Nevertheless, the posterior probabilities enable the characterization of 'typical' component-observations through calculation of the component-specific means. For any variable $v_{k,it}$, this is given by

$$\bar{v}_k^s = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T} \tau_{s,it} v_{k,it}}{\sum_{i=1}^{N}\sum_{t=1}^{T} \tau_{s,it}}, \text{ for } s \in \{1, \ldots, M\}. \tag{11}$$

The component-specific means are particularly useful if no significant association between health variables and components is found, as suggested by Deb and Trivedi (1997, 2002). If the component with the higher health care policy uncertainty coefficient is characterized as 'ill' on the basis of the component-specific means, this would provide empirical support for Proposition 2.

An important issue for latent class models is the number of components, M, that must be fixed *a priori* by the researchers. Because models with different numbers of components are not necessarily nested, conventional likelihood ratio tests are not applicable in this context. When economic theory provides little guidance, studies commonly consider penalized likelihood criteria such as the Akaike Information Criterion (AIC) or Bayes Information Criterion (BIC) for determining an appropriate number of mixture components (McLachlan and Peel, 2004). Analogous to the caveats associated with these criteria in other applications, this approach is not ideal as information criteria do not allow one to statistically test for the appropriate component number. Fortunately, existing literature on the application of mixture models provides some suggestions for selecting the number of mixture components. In particular, Deb and Trivedi (1997, 2002), who analyze heterogeneous demand for health care services, suggest that it seems sensible to assume two underlying sub-populations; one potentially characterized by 'healthy' households, the other by 'ill' households. We follow their example by adopting a two-component latent class model.[43]

Because the logarithm of the likelihood in (7) contains a log of a sum due to the mixture-term, estimation of latent class models using conventional optimization procedures is numerically difficult. As an alternative, we calculate maximum likelihood estimates using Dempster et al.'s (1977) EM algorithm. Van der Heijden et al. (1996) and Karlsson and Laitila (2014) provide the relevant adaptations of the procedure to incorporate estimation of the parametrized class membership probabilities and Tobit densities, respectively. As

[43] For robustness, we also estimated three-component latent class models. While no improvement was found for total spending and risky assets, the information criteria indicated a better fit for safe assets. However, in all cases, the additional sub-population seems to stem from a split of the smaller, health care policy uncertainty-*un*responsive component. The results neither provided more detailed conclusions, nor did the additional component have any bearing on the conclusions from the two-component model. The estimation of the three-component latent class model employed less extensive computational settings, in particular, five EM starting values and 25 bootstrap iterations.

information-based standard errors require very large sample sizes for good asymptotic approximations (McLachlan and Peel, 2004), the presented standard errors are computed using a nonparametric bootstrap procedure with 100 resampling iterations. The algorithm is implemented in R, with code readily available upon request. Appendix C.2 provides an outline of the estimation procedure.

4.2.2 Model-Based Recursive Partitioning

The latent class model is a substantial improvement over an *a priori*-specified heterogeneous effects model such as the multiplicative-interaction model, yet, the model is not itself free of qualifications. Apart from assessing differences between a few sub-populations, for example, more detailed conclusions could be drawn by considering a higher number of potential segments. While such an approach is not straightforward for latent class models due the difficulties in pre-specifying the number of components and the associated computational complexity, recent developments in the interdisciplinary literature on machine learning in economics provide new appealing techniques that address such limitations (Athey, 2017, 2018; Mullainathan and Spiess, 2017). Of particular interest to this study are the developments of partitioning methods (e.g., regression trees) to the topic of heterogeneous causal effects (e.g., Athey and Imbens, 2015, 2017). These methods are useful as they can model non-linear, highly interactive heterogeneous treatment effects while simultaneously permitting statistical inference on the causal structures in a visually intuitive manner. Unfortunately, however, the majority of existing literature on economic applications of machine learning has solely considered the case of binary treatment effects under ideal data conditions (i.e., no censoring/asymmetry). These qualifications render existing methods in economics – among them the causal trees proposed by Athey and Imbens (2015) or the causal forests proposed by Wager and Athey (2018) – unsuitable for applications with continuous parameters of interest and/or with asymmetric, censored outcome variables. A notable exception is the recently developed forest-based local estimation equation estimator of Athey et al. (2019). The current study implements a novel alternative approach.

A machine learning technique on heterogeneous causal effects that is yet to be explored in economics is Zeileis et al.'s (2008) model-based recursive partitioning algorithm. Developed originally for applications involving linear and logistic regression, the procedure was extended to psychometric models (e.g., Strobl et al., 2011, 2015) as well as generalized linear models and maximum likelihood models (Rusch and Zeileis, 2013). We adapt the method to a setting with observed nuisance parameters and – for the analysis of portfolio shares – to asymmetric outcome variables by employing a pooed Tobit model in each partition.

For ease of comparison, the model can be written similarly to the latent class specification of equation (6). In particular, consider the G-partition model given by

$$y_{it} = \begin{cases} f(y_{it}|X_{it}, \theta_1), & \text{if } \{Z_{it}\} \in \mathcal{Z}_1 \\ \quad \vdots & \\ f(y_{it}|X_{it}, \theta_G), & \text{if } \{Z_{it}\} \in \mathcal{Z}_G \end{cases} \tag{12}$$

where a particular observation stems from the gth partition, $g = 1, \ldots, G$, if its features, $\{Z_{it}\}$, are elements of the partition's feature-space, $\mathcal{Z}_g = \mathcal{Z}_{g,1} \times \ldots \times \mathcal{Z}_{g,B}$, with B being the number of features. As before, $f(\cdot|X_{it}, \theta_g)$ denotes the density of a pooled Tobit model given in equation (8) for the analysis of portfolio shares and the density of a pooled normal model for the analysis of total spending.

An important distinction to the latent class model's probabilistic characterization of class membership is that model-based recursive partitioning considers an observation's partition membership to be a *deterministic* function of its feature vector Z_{it}. While thus not permitting the parameter heterogeneity to partly stem from latent characteristics, the procedure can approximate forms of heterogeneous effects that are highly non-linear and interactive. If the parameter differences are strongly associated with the observed health measures, this allows for a level of detail in inference on the underlying data structures that is unmatched by alternative methods. As estimating the model (12) with conventional techniques is computationally infeasible, Zeileis et al. (2008) propose a greedy, forward-searching algorithm for unbiased estimation of the partition-specific parameters θ_g and associated feature-spaces $\mathcal{Z}_g$. In essence, model-based recursive partitioning estimates the G-partition model by (1) fitting the local model $f(\cdot)$ to the data through minimization of the objective function, (2) testing for parameter instability over the set of features Z and selecting the feature associated with the highest instability z_k^*, (3) computing the split point of z_k^* that locally optimizes the objective function, and (4) splitting the data into two sub-partitions and repeating the procedure. Steps (1)-(3) are discussed in some detail below.

Step (1), fitting the local model, corresponds to performing maximum likelihood estimation of a pooled Tobit or normal linear model. Following Zeileis et al. (2008), the corresponding negative log-likelihoods are denoted by Ψ. A relevant side product of this step is each observation's contribution to the score function: $\psi_{it} = \frac{\partial \Psi(y_{it}, X_{it}, \theta)}{\partial \theta}$.

Step (2), testing for parameter instability, assesses whether a sample split with respect to one feature captures potential instability of the model coefficients. For this purpose, the authors suggest testing whether the ψ_{it} fluctuate randomly around their mean (zero) or exhibit systematic deviations over a particular feature. Zeileis et al. (2008) only explicitly consider testing for parameter instability of the entire parameter space. This is problematic as splits could stem from instability of nuisance parameters instead of coefficients of interest. Fortunately, this concern is easily alleviated as the method can be adjusted to test for instability on only a subset of θ in a straightforward manner; instead of ψ_{it}, one can consider $\psi_{it}^{j} = \frac{\partial \Psi(y_{it}, X_{it}, \theta)}{\partial \theta_j}$. Systematic deviations from zero can then be captured by the empirical fluctuation process:

$$W_k(l) = \frac{1}{\hat{\sigma}(\theta_j)\sqrt{n}} \sum_{i=1}^{\lfloor nl \rfloor} \hat{\psi}^{j}_{\eta(z_{j,i})} \qquad (0 \leq l \leq 1), \tag{13}$$

where $\hat{\sigma}(\theta_j)$ and $\hat{\psi}^j$ are the estimated standard error and the estimated score contribution of the coefficient of interest θ_j, n is the number of observations at the current partition, and $\eta(z_{k,i})$ denotes the ordering permutation with respect to the kth feature. Put simply, $W_k(l)$ is the partial sum process of the score contributions ordered by the feature variable, scaled by the number of observations and a suitable standard error estimate.

To assess instability over ordered discrete variables, the authors suggest the test statistic

$$\lambda_{\chi^2}(W_k) = \sum_{c=1}^{C} \frac{1}{n|I_c|} \left|\left| \Delta_{I_c} W_k\left(\frac{i}{n}\right) \right|\right|_2^2, \tag{14}$$

where $\Delta_{I_c} W_k\left(\frac{i}{n}\right)$ is the increment of the empirical fluctuation process over the observations of category $c \in \{1, \dots, C\}$ of the kth feature (with associated indices I_c) – that is, the sum of score contributions of category c. The statistic is then given by the weighted sum of the squared L_2 norm of the increments with asymptotic distribution $\chi^2\big(C-1\big)$ (Zeileis and Hornik, 2007; Zeileis et al., 2008). A substantial advantage of this parameter instability test is that the coefficients and corresponding score contributions only have to be calculated once in each partition. To test for instability across multiple features, the scores are merely reordered. The null hypothesis of stability is rejected in the current partition whenever the minimal p-value corresponding to any feature falls below a preimposed significance threshold.[44] If multiple tests reject stability, the sample is split with respect to the feature with the lowest p-value.

Step (3), finding a suitable threshold to split the sample along the feature z_k^*, is trivial. For a binary split at each node, two rival partitions can be compared in a straightforward manner by comparing the segmented objective function $\sum_{b=1}^{2} \sum_{i \in I_b} \Psi(y_{it}, X_{it}, \theta_b)$. The optimal cut-off is then determined by performing an exhaustive search over all possible partitions.

[44] As concerns about multiple hypothesis testing might arise, Bonferroni-adjusted p-values are considered.

Splitting the sample according to step (3) concludes one iteration of the model-based recursive partitioning algorithm. The procedure is continued until no further parameter instability is found in any nodes or when no further splits are feasible due to sub-sample size. Contrasting the latent class model, no prior choice on the number of partitions, G, is required. We implement model-based recursive partitioning in R using Hothorn and Zeileis's (2015) *partykit* library. Standard errors suggested by Zeileis et al. (2008) are the standard errors of the local models in each partition. Although these are uncorrected for the uncertainty in the estimation of partitions' feature spaces, no correction has yet been developed. Thus the uncorrected standard errors are reported despite this caveat.

Using model-based recursive partitioning, we analyze whether parameter instability in the health care policy uncertainty coefficient can be captured by splitting along the health and health care measures. Resulting estimates provide valuable empirical insights that allow evaluation of Proposition 2. In particular, to assess whether households with worse health are more affected by health care policy uncertainty, the coefficient estimates corresponding to the 'healthy' and 'ill' sample partitions should be compared. As will become clear in Section 5.3, the sample division of the model-based recursive partitioning allows for inference on the parameter heterogeneity that exceeds the detail afforded by a mere binary split.

5 Results

This section applies the models of the previous section to analyze the effect of health care policy uncertainty on households' consumption and financial behavior. After testing for an effect of health care policy uncertainty (Proposition 1), the latent class models and the model-based recursive partitioning procedure are applied to analyze whether potential heterogeneity in the effect of health care policy uncertainty is associated with households' health (Proposition 2).

5.1 Baseline Estimates

The purpose of the baseline regressions is to test Proposition 1 – whether increased health care policy uncertainty causes a decrease in households' consumption and relative demand for risky assets. Table 2 presents the results of the linear fixed effect, the censored fixed effect and the pooled Tobit models. For brevity, only the coefficient corresponding to the log of the 12-month health care policy uncertainty average is included. Complete estimation results can be found in Appendix D.1.

[Table 2 about here.]

The results for total spending in column (1) indicate a negative and significant effect of health care policy uncertainty for couple households. The average marginal effect (-0.039)

indicates a 0.039 percentage decrease in annual total spending for every one percent increase in the 12-month average of the health care policy uncertainty index. For illustrative purposes, however, it is more interesting to consider the approximately 70% increase in the index, from a 2016 average of 110 to the 2017 average of 191. As briefly illustrated in Section 1, the latter year is associated with extensive political efforts to repeal the Affordable Care Act. When faced with such a 70% increase, the results suggest that couple households lower their annual total spending by about 2.73% ($= 70\% * 0.039$). For the average couple household in the sample with a total spending of about \$57,773, this results in an absolute annual spending decrease of \$1,577 on average. While these estimates suggest an economically relevant effect on couple households' spending behavior, no significant effect of health care policy uncertainty on total spending can be found for single households (column 6). Researchers, however, should be careful not to view this as evidence against an effect of health care policy uncertainty on total spending: as noted in Section 3.2, a disadvantage of the CAMS data is the need to approximate the exact survey month. The resulting measurement error in the match between households' total consumption data and the uncertainty they faced might be partially responsible for the ambiguous statistical evidence for Proposition 1.

For couple households, these results are in line with the results of Aaberge et al. (2017) and Giavazzi and McMahon (2012), who find a significant savings-increase for Chinese and German households, respectively, when exposed to higher policy uncertainty caused by political turmoil. While a direct comparison of effect magnitudes to previous studies should be viewed with caution given the differences in sample, context, and empirical approach, the health care policy uncertainty effect on spending found in this study is in line with other estimates: Giavazzi and McMahon (2012), for example, estimate an approximate increase in savings rate (i.e., total savings as percentage of income) of approximately three percentage points caused by policy uncertainty surrounding a national election in Germany.[45]

The estimation results for the portfolio shares support the theoretical predictions of lower relative demand of risky assets versus safe assets in light of increased health care policy uncertainty. Columns (2)-(3) and (7)-(8) imply that an increase in health care policy uncertainty is associated with a decrease in investment share in risky assets for both couple and single households. Using the average marginal effect of the pooled Tobit model, the 70% increase in the annual health care policy uncertainty average decreases the risky asset share of couple and single households by approximately 2.8 ($= 0.7 * 0.040$) and 2.6 ($= 0.7 * 0.037$) percentage

[45]Earlier versions of this study also considered households' spending on durables. Due to the substantial share of households with no spending on durables in the sample, however, it was necessary to employ a Tobit model and use the outcome variable without a log transformation. No economically relevant effect of health care policy uncertainty on durables was found. See Wiemann (2018).

points on average, respectively. As the pooled Tobit model results are larger in magnitude, however, this is likely an upper bound on the true reduction. Similarly, the results in columns (4)-(5) and (9)-(10) indicate a positive association between health care policy uncertainty and the share invested in safe assets. The average marginal effects point to an increase of approximately 2.2 (2.0) percentage points for couple (single) households given the 70% health care policy uncertainty increase discussed earlier. The fixed effect results indicate that the true magnitude of the effect is likely somewhat lower.

These estimates are well aligned with previous literature on the effect of health on portfolio choice. Also analyzing HRS data, Rosen and Wu (2004), Edwards (2008), and Love and Smith (2010), for example, find that rating health in the worst category of the subjective health measure is associated with a 1%, 7%, and 1.8% decrease in risky asset share, respectively. Given a substantial increase in health care policy uncertainty, a decrease in relative demand for risky assets comparable to the effect of bad health seems plausible. This highlights the importance of health care policy uncertainty as a determinant of households' financial behavior.

Estimation of the baseline regression models provides empirical evidence that health care policy uncertainty causes a significant decrease in households' total spending for couples but fails to identify a significant effect for single households. In contrast, the results on the share of risky and safe assets indicate an important link between health care policy uncertainty and portfolio choice for both household compositions. As the regressions control for stock market and business cycle uncertainty, the estimates point to a specific health care policy uncertainty channel rather than a broader economic uncertainty effect. To strengthen this claim, we proceed to analyze whether the effects' magnitudes increase with lower health as suggested by Proposition 2.

5.2 Heterogeneous Effects: Latent Class Model

First insights into the association between households' health and the effect of health care policy uncertainty are gained by applying the latent class models. In particular, we test for statistical significance of the concomitant health variables in the prior probability specification to assess whether component membership corresponding to a higher health care policy uncertainty effect is linked to bad health. The posterior component means provide further suggestive evidence on the characteristics of the identified sub-populations in cases where no statistical significance of the concomitant variables was found.

In this part of the analysis, only couple households are explicitly investigated for heterogeneous effects for brevity of results. Couple households constitute the majority of the HRS sample and concerns that economic behavior of singles cannot readily be adjusted due to health impairment are alleviated by this focus. At the same time, the consumption and portfolio choice of couples is likely to have broader macroeconomic implications, as larger household size is associated with higher total consumption and higher financial wealth. The restriction is also in line with previous research (e.g., Wu, 2003). Corresponding results are presented in Table 3. Coefficient estimates of the concomitant health variables are in the top panel, where the first component's coefficients are normalized to zero for identification. The bottom panel shows the component-specific health care policy uncertainty coefficient. Appendix D.2 provides complete estimation results. The posterior component means of selected variables are shown in Table 4.

[Table 3 about here.]

[Table 4 about here.]

Given that the health care policy uncertainty coefficients of components 1 and 2 (*s.1* and *s.2*, henceforth) in column (1) are both insignificant, we find no evidence for heterogeneous effects on total spending. In contrast, there are substantial differences in the health care policy uncertainty coefficients across components in columns (2) and (3), indicating heterogeneous effects on the risky and safe asset share. However, as latent class estimation does not formulate the heterogeneity as stemming from specific observable characteristics, a discussion of the concomitant variable coefficients and the posterior component average is necessary in order to assess the association between component-specific effects and households' health and health insurance measures.[46]

Despite no evidence of heterogeneity of the health care policy uncertainty effect on total spending, these two measures provide suggestive evidence of a link between households' characteristics and the latent classes. In particular, we find that a household's lack of health insurance seems to be positively associated with the probability of belonging to *s.2* (as seen by the health insurance coefficient's significance at a 10% level and the 13 $(= (0.19 - 0.06) * 100)$ percentage point higher share of uninsured households than in *s.1* as inferred from the corresponding posterior component means reported in the first two lines of Table 4). The remaining coefficients corresponding to the health measures in the prior probability specification are insignificant and indicate no further association between households' health and component membership. While the posterior component means provide some contrasting evidence (e.g., households of *s.2* spend approximately 27% $(= \frac{3.57-2.82}{2.82})$ more nights in a

[46]Because the components are not defined prior to estimation, separate iterations of the bootstrap procedure potentially result in components being numbered in reverse order. To enable calculation of each coefficient and its corresponding standard error estimates, McLachlan and Peel (2004) therefore suggest ordering components by a particularly distinctive characteristic. We have found that numbering the components by their variance is ideal for the data at hand.

hospital than those of *s.1*), we refrain from assigning explicit labels to the components in column (1) for this reason. Refuting Proposition 2, the findings presented here indicate no heterogeneity in total spending that can be linked to the observed health variables.

The results on the risky and safe asset share provide empirical evidence in support of the theoretical predictions. In particular, the estimates in column (2) and (3) of Table 3 indicate that the self-reported health measure is positively associated with *s.2* membership. The results on safe asset shares show that in addition, the number of severe conditions, impairments to mobility, and being uninsured are positively associated with *s.2* membership. Further evidence is provided by the posterior component means. Considering the safe asset results in the bottom two rows of Table 4, the mean observation in *s.2* is characterized by self-reported health and impairments to mobility being 13.6% ($=\frac{3.26-2.87}{2.87}$) and 30.1% ($=\frac{2.16-1.66}{1.66}$) worse compared to the mean observation in *s.1*, respectively. The analogous numbers for risky assets are 6.4% ($=\frac{3.15-2.96}{2.96}$) and 8.1% ($=\frac{2.01-1.86}{1.86}$), respectively. One potential worry is that the positive association between health and membership in the second component is potentially confounded by wealth – that is, that the estimates instead capture exposure differences through total wealth due to its positive association with health (Smith, 1999). Two arguments speak against this concern. First, Fan and Zhao (2009) suggest that health does not affect portfolio choice through changes in financial wealth, as they find no effect of health on financial wealth of older Americans. Second, in addition to the health measures, the prior probability specification includes the household's health insurance status (that is positively associated with wealth). Results indicate a household with worse self-reported health and more impairments to mobility is significantly more likely to belong to component two, keeping health insurance fixed. For example, using the results for the safe asset share, a household with "Very good" self-reported health, covered by health insurance, with either spouse having spent two nights in a hospital, zero impairments to mobility, two severe conditions, and no limitations regarding ADLs, is 24 percentage points less likely to belong to *s.2 a priori* than the same household with "Poor" self-reported health and three impairments to mobility.[47] The strong association of component membership with the various

[47] The values are specified as *a priori* in reference to the probability of belonging to a particular component before observing the outcome and is calculated using the specification defined in equation (9). Specifically, $24 = \left(\frac{\exp(H_1)}{1+\exp(H_1)} - \frac{\exp(H_2)}{1+\exp(H_2)}\right) * 100$, where the argument $H_2 = -0.361 + 0.293 * 2 + 0.049 * 0 + 0.066 * 2 + 0.028 * 0 - 0.001 * 2 + 0.838 * 0$, corresponding to a household with "Very good" self-reported health (the value of the variable is 2), covered by health insurance, with either spouse having spent a total of two nights in a hospital, zero impairments to mobility, two severe conditions, and no limitations regarding ADLs and the argument $H_1 = -0.361 + 0.293 * 5 + 0.049 * 3 + 0.066 * 2 + 0.028 * 0 - 0.001 * 2 + 0.838 * 0$, corresponding to a household that instead has "Poor" self-reported health (the value of the variable is 5) and three mobility impairments, *ceteris paribus*. The analogous *posteriori* probability of belonging to *s.2* is a comparison of the two component-specific likelihoods (equation (10)), that includes the outcome Y in the calculation and does not permit statistical linkage of specific observed features to the probability.

health measures exceeds any potential confoundedness that might exist between total wealth (as captured by better insurance coverage) and health. Thus it is unlikely that wealth is the sole/primary driver of the estimated association with health. This observation provides justification for linking the components of columns (2) and (3) to households' health and alleviates concerns that the association might instead stem from other (non-health-related) household characteristics, in particular, wealth, which are potentially associated with some of the health measures considered. Following Deb and Trivedi's (1997; 2002) terminology, *s.1* and *s.2* are labelled as 'healthy' and 'ill', respectively.

As indicated by the bottom panel of Table 3, observations of the ill component are negatively affected by an increase in uncertainty. Facing a 70% increase in the annual health care policy uncertainty average such as the one from 2016-2017, the mean household in the ill component decreases its investment share in risky assets by about 2 percentage points ($= 0.7 * 0.029$), and increases its share in safe assets by about 2.6 percentage points ($= 0.7 * 0.037$).[48] In contrast, the mean household in the healthy component is not significantly affected by health care policy uncertainty. The magnitude of these estimates is in line with the literature on the effect of health on portfolio choice as illustrated in the previous section. Their importance is further highlighted by the average prior probability of each component, which are usually interpreted as each component's population share (Kamakura and Russell, 1989). Values are presented in Table 3 (line labelled 'Component share'). As approximately 68% to 93% of the sample seems to be strongly negatively affected, health care policy uncertainty is an important determinant of financial behavior for the majority of older American couples. The results are thus not driven by a few unusual observations, but instead reflect a prevalent data pattern. These findings provide strong empirical support for the claim of Proposition 2 that less healthy households reduce their relative demand for risky assets more strongly when exposed to health care policy uncertainty.

5.3 Heterogeneous Effects: Model-Based Recursive Partitioning

Results in the previous section provide support for the claim of Proposition 2 regarding the relative demand for risky assets versus safe assets. However, insights from the latent class model are limited. In addition to allowing for a higher number of sub-populations, model-based recursive partitioning also allows the effect of the health variables to be non-linear and highly interactive, making it most suitable for in-depth analysis of heterogeneity in the effect of health care policy uncertainty.

[48] Contrasting the pooled Tobit model, average marginal effects cannot be calculated as observations are not explicitly assigned to a component. As an alternative, marginal effects for each component's average observation (defined by the posterior component means) are used for interpretation.

[Figure 2 about here.]

Due to the wealth of model-based recursive partitioning results, a structured approach is necessary to draw inferences regarding heterogeneity in the health care policy uncertainty effect. Fortunately, the method allows for intuitive visualization of the estimates through trees. These results are presented in Figure 2. For each tree, the feature with the most significant rejection of parameter stability is given by the top node (i.e., insurance status for total spending and the safe asset share; self-reported health for the risky asset share). From this node, additional partitioning continues along the same lines until a terminal node is reached, that is, no further parameter instability is detected or no further splits are feasible due to insufficient sub-sample size. Therefore, throughout this figure, connecting nodes (rounded corners) state the feature with the most significant rejection of parameter-stability, and terminal nodes (sharp corners) present for each resulting sub-sample the coefficient of the log-transformed annual health care policy uncertainty average, its average marginal effect, and the sub-sample size. Subscripts within the terminal nodes indicate node numbering for the purposes of discussion (below). Complete estimation results are in Appendix D.3.

Note that two branches sharing the same parent node are characterized by significant heterogeneity in their health care policy uncertainty coefficients. This difference can be directly linked to one of the health or health insurance variables, as the empirical fluctuation test of Zeileis and Hornik (2007) investigates the presence along these measures (see step (2) of Section 4.2.2). In addition to these significant parameter differences, bold coefficients of terminal nodes indicate that the health care policy uncertainty effect specific to that particular partition is individually significant.

We begin by analyzing the model-based recursive partitioning results for total spending (Figure 2a). In line with the results of the latent class model of the previous section, the method identifies a statistically significant association between insurance status and partition membership. Adding to the two classes estimated previously, however, further conditional heterogeneity is found with respect to activities of daily living as well as mobility impairments. Specifically, node (1) corresponds to households with health insurance and no ADL or mobility impairments, node (2) corresponds to households with health insurance, no ADL and at least one mobility impairment, node (3) corresponds to households with health insurance and at least one ADL impairment, and node (4) corresponds to households where at least one member is not covered by health insurance. Because the ADL variable measures impairments that exceed those of the mobility variable in severity (e.g., not being able to eat independently is a more severe impairment than not being able to walk a flight of stairs), it is natural to order nodes (1)-(3) as decreasing in health. Under this characterization, the empirical evidence supports the proposition that worse health is associated with a greater decrease in total spending. In particular, households in node (2) decrease their total spending by 2.17% (= $(0.045 - 0.076) * 0.7$)more when faced with a 70% increase in health care policy uncertainty compared to households in node (1). Considering two households in node (1) and (2), respectively, each with an annual total spending of $57,377 (the average annual total spending of couple households), this indicate that the household in worse health (node 2) reduces its annual spending by $1,245 (= 57,377 * 2.16%) more than the household in better health (node 1) when faced with the same 70% increase in health care policy uncertainty. This is in stark contrast to the latent class results, where no statistically significant heterogeneous effects were identified.

Beyond the estimate for total spending, the model-based recursive partitioning results for the risky and safe asset shares provide further evidence that substantial heterogeneity in the effect of health care policy uncertainty is captured by the observed health variables, as indicated by the numerous sample splits in both cases (Figure 2b and 2c).[49]

The moderate size of the tree for total spending allowed for straightforward characterization of partitions as more and less healthy. This is, unfortunately, not always possible because in some cases it is unclear how to rank health across some pairs of health measures. For example, while partition (1) of Figure 2c can be characterized as healthier than partition (2), given the latter's strictly worse self-reported health, the comparison of partitions (4) and (5) is ambiguous. The reason for this is that existing research provides no empirical evidence as to whether, in this case, poor self-reported health or a higher number of mobility impairments (e.g., not being able to walk a block) is a clearer determinant of bad health.[50] While it is possible to speculate, we prefer to avoid this ambiguity in comparing the observed health measures and instead focus on sub-trees that allow for straightforward interpretation.

With this strategy, the results can be analyzed in the context of supporting or refuting Propositions 1 and 2. First, consistent with the results from the other models, the results of the model-based recursive partitioning provide additional evidence supporting Proposition 1. With the exception of node (7) in Figure 2c where the coefficient is not statistically significantly different from zero, nearly all partitions show a striking positive (negative) and statistically significant effect of health care policy uncertainty on the safe (risky) asset share. The result that health care policy uncertainty decreases relative demand for risky assets

[49] In addition, the frequent splits are evidence of a highly non-linear and interactive effect of health and thus provide strong support for the methodological decision not to employ a multiplicative-interaction model in this study.

[50] Love and Smith (2010) employ similar health indices, but provide no evidence on severity in their comparison.

thus seems robust to a variety of model specifications – specifically, a pooled Tobit model, Honoré's (1992) censored fixed effect model, a latent class Tobit model, and a model-based recursive partitioning procedure – that are employed throughout this study and which result in qualitatively identical conclusions.[51] Second, many sub-trees show a more substantial effect of health care policy uncertainty with worsening health or insurance status. For example, nodes (1,2,3,4) in Figure 2c are characterized by coefficients of increasingly positive magnitude, indicating greater effects of health care policy uncertainty on the safe asset investment share as self-assessed health declines. Comparing the effects of a 70% increase in health care policy uncertainty (the annual average), for households characterized by node (1) versus node (4), in particular, an insured household with one or no impairments to mobility and self-reported health of 'very good' or better (node 1) increases its share in safe assets by 4.7 ($= 0.7 * (0.095 - 0.028) * 100$) percentage points less than a comparable household with 'poor' self-reported health (node 4). While this provides evidence that poorer self-assessed health magnifies the effect of health care policy uncertainty, the claim that health has a strictly unidirectional effect cannot be supported in general. In particular, consider the node-pair (6,7) in Figure 2c and the pair (2,3) in Figure 2b. A move from node (2) to node (3) indicates an increase in the households' mobility impairment and from (6) to (7) more severe conditions, respectively, yet, these shifts are associated with a decreasing magnitude of the health care policy uncertainty effect. While a small sample size could be partially responsible for the insignificant effect found for the less healthy partition in the case of safe asset investment (node 7 of Figure 2c), another possible explanation for these results is the potential decrease of consumption utility and expected lifespan with worsening health (Smith, 1999). As illustrated in Section 2.1, if this deteriorating effect of bad health on consumption utility and expected lifespan and the resulting lower risk aversion is more substantial than the impact of bad health on medical expenditures, then the health care policy uncertainty effect is *decreasing* in health issues. The results for the risky investment share seem to suggest that the negative effects of health on consumption utility and expected lifespan exceed the effect of the increase in medical expenditures for households with a particularly large number of impairments to mobility (node 3 of Figure 2b).

In contrast to the results of the latent class model on households' total spending, the model-based recursive partitioning results indicate that worse health is associated with a

[51] The latter two models are explicitly aimed at identifying parameter breaks, but the results demonstrate that the effect of health care policy uncertainty is negative and statistically significant for the large majority of the sample (between 66%-89% in the case of the latent class model and for almost all nodes in the model-based recursive partitioning case) and that it is not just an unrepresentative group of special observations that are driving the effect. This reasoning is in line with Zeileis et al. (2008), who suggest that one of the applications of the model-based recursive partitioning procedure is to serve as a thorough robustness check.

greater spending reduction when faced with health care policy uncertainty. Further, while the results in this section corroborate the negative effect of health care policy uncertainty on households' relative demand for risky assets versus safe assets, the effect of health on the impact of health care uncertainty does not seem to be strictly uni-directional. Nevertheless, the results show that the magnitude of the coefficient of health care policy uncertainty is strongly associated with observed health variables and that a larger effect of health care policy uncertainty seems to be at least partly associated with worse health.

6 Conclusion

This study analyses the effect of health care policy uncertainty on households' consumption and portfolio choice. A simple model is developed to motivate a negative effect on consumption and relative demand for risky financial assets, and to illustrate a potentially magnifying impact of bad health on the effect of health care policy uncertainty. Using the HRS' rich longitudinal data on older Americans and Baker et al.'s (2016) recently developed health care policy uncertainty index, these theoretical claims are tested with linear and censored fixed effect regression models., a latent class model and a model-based recursive partitioning procedure, which we adapt from the machine learning literature. There is mixed empirical evidence for an effect of health care policy uncertainty on households' total spending: we find a significant reduction in total spending for couple households when faced with an uncertainty increase, as well as suggestive evidence that this effect increases with worse health, but fail to identify similar effects for single households. Further results indicate a substantive negative effect of health care policy uncertainty on households' relative demand for risky assets and a magnified impact of bad health on this portfolio choice effect. As these results do not appear to be driven by potentially endogenous household characteristics or other confounding forms of uncertainty, this study indicates that health care policy uncertainty can be an important determinant of households' financial behavior.

The empirical evidence not only suggests that higher health care policy uncertainty is associated with a welfare loss of individual households, but also points to substantive macroeconomic consequences. The estimates suggest that an uncertainty increase similar to the repeal efforts of the Affordable Care Act in 2017 decreases the relative demand for stocks and mutual funds as much as a considerable reduction in health (e.g., Rosen and Wu, 2004). This effect is not driven by a few households, but is a prevalent pattern that is found for about 66% to 89% of older American couples. Given their large share of financial assets, the reduction in stock market participation has direct consequences on stock market volatility (Allen and Gale, 1994), the equity premium (Mankiw and Zeldes, 1991), and wealth inequality (Favilukis, 2013). The latter is particularly grave as health care policy uncertainty disproportionally affects less healthy households, exacerbating the socio-economic disadvantage associated with bad health in the US (Smith, 1999). As health care policy uncertainty and its implications are likely to persist in the foreseeable future, these costs seem relevant for both legislatures and voters.[52]

This study makes first contributions to the analysis of health care policy uncertainty, but many policy questions remain. For example, it is unclear how transparent communication of health care reform plans would affect uncertainty. Although it could have a decreasing effect through reducing misinformation, such news can foreshadow the possibility of policy changes, potentially increasing uncertainty. Further, while limiting legislative bodies' abilities to delay reforms or revoke policies adopted by previous governments might lower uncertainty, it could also impede interventions necessary to address uncertainty. A promising alternative is to reduce the political component of health care policy uncertainty. In particular as health care continues to be a major topic in national election campaigns, assessing how constructive debate can be distinguished from political rhetoric is important. We leave these as avenues for future research.

[52] Even while writing this paper, health care policy uncertainty experienced yet another spike with the US Justice Department joining a lawsuit against some of the Affordable Care Act's provisions, including the rule prohibiting the denial of health care to people with pre-existing medical conditions (e.g., Barnes, 2018).

References

Aaberge, R., Liu, K., and Zhu, Y. (2017). Political uncertainty and household savings. *Journal of Comparative Economics*, 45(1):154–170.

Addoum, J. M. (2017). Household portfolio choice and retirement. *Review of Economics and Statistics*, 99(5):870–883.

Agarwal, V., Aslan, H., Huang, L., and Ren, H. (2018). Policy uncertainty and household stock market participation. Centre for Financial Research working paper.

Aitkin, M. and Rubin, D. B. (1985). Estimation and hypothesis testing in finite mixture models. *Journal of the Royal Statistical Society. Series B (Methodological)*, 47(1):67–75.

Allen, F. and Gale, D. (1994). Limited market participation and volatility of asset prices. *American Economic Review*, 84(4):933–955.

Atella, V., Brindisi, F., Deb, P., and Rosati, F. C. (2004). Determinants of access to physician services in Italy: A latent class seemingly unrelated probit approach. *Health Economics*, 13(7):657–668.

Atella, V., Brunetti, M., and Maestas, N. (2012). Household portfolio choices, health status and health care systems: A cross-country analysis based on SHARE. *Journal of Banking & Finance*, 36(5):1320–1335.

Athey, S. (2017). Beyond prediction: Using big data for policy problems. *Science*, 355(6324):483–485.

Athey, S. (2018). The impact of machine learning on economics: An agenda. In Agrawal, A. K., Gans, J., and Goldfarb, A., editors, *The Economics of Artificial Intelligence: An Agenda*, chapter 21. University of Chicago Press, Chicago.

Athey, S. and Imbens, G. W. (2015). Machine learning methods for estimating heterogeneous causal effects. Stanford University Working Paper.

Athey, S. and Imbens, G. W. (2017). The state of applied econometrics: Causality and policy evaluation. *Journal of Economic Perspectives*, 31(2):3–32.

Athey, S., Tibshirani, J., Wager, S., et al. (2019). Generalized random forests. *The Annals of Statistics*, 47(2):1148–1178.

Attanasio, O. P. (1999). Consumption. In Taylor, J. B. and Woodford, M., editors, *Handbook of Macroeconomics*, volume 1, chapter 11, pages 741–812. Elsevier, Amsterdam.

Bago d'Uva, T. (2006). Latent class models for utilisation of health care. *Health Economics*, 15(4):329–343.

Bago d'Uva, T. and Jones, A. M. (2009). Health care utilisation in Europe: New evidence from the ECHP. *Journal of Health Economics*, 28(2):265–279.

Bago d'Uva, T., O'Donnell, O., and Van Doorslaer, E. (2008). Differential health reporting by education level and its impact on the measurement of health inequalities among older Europeans. *International Journal of Epidemiology*, 37(6):1375–1383.

Baker, S. R., Bloom, N., and Davis, S. J. (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics*, 131(4):1593–1636.

Baker, S. R., Bloom, N., and Davis, S. J. (2018). Health care policy uncertainty index. Retrieved on May 15, 2018, from `http://www.policyuncertainty.com/categorical_epu.html`.

Barnes, R. (2018). Trump administration shifted positions in a lawsuit against Affordable Care Act. It wasn't first legal about-face. *The Washington Post*. Retrieved on June 15, 2018, from `https://www.washingtonpost.com/politics/courts_law/trump-administration-shifted-positions-in-a-lawsuit-against-affordable-care-act-it-wasnt-first-legal-about-face/2018/06/08/c0dd30a8-6b33-11e8-9e38-24e693b38637_story.html?utm_term=.7b5a608a4ec8`.

Berkowitz, M. K. and Qiu, J. (2006). A further look at household portfolio choice and health status. *Journal of Banking & Finance*, 30(4):1201–1217.

Bertrand, M., Duflo, E., and Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics*, 119(1):249–275.

Bloom, N. (2014). Fluctuations in uncertainty. *Journal of Economic Perspectives*, 28(2):153–76.

Bloomberg (2018a). Implied Volatility Index (VIX). Monthly data for the period 01.1988-07.2018 retrieved on August 29, 2018.

Bloomberg (2018b). S&P 500 Total Return Real Time Price (SPXT). Monthly data for the period 01.1988-07.2018 retrieved on August 29, 2018.

Bodie, Z., Merton, R. C., and Samuelson, W. F. (1992). Labor supply flexibility and portfolio choice in a life cycle model. *Journal of Economic Dynamics and Control*, 16(3-4):427–449.

Bogan, V. L. and Fertig, A. R. (2013). Portfolio choice and mental health. *Review of Finance*, 17(3):955–992.

Bureau of Economic Analysis (2018). Table 2.5.5. Personal Consumption Expenditures by Function. Retrieved on May 20, 2018, from `https://www.bea.gov/iTable/iTableHtml.cfm?reqid=19&step=3&isuri=1&1921=survey&1903=74`.

Cameron, A. C. and Trivedi, P. K. (2005). *Microeconometrics: Methods and Applications*. Cambridge University Press, Cambridge.

Cardak, B. A. and Wilkins, R. (2009). The determinants of household risky asset holdings: Australian evidence on background risk and other factors. *Journal of Banking & Finance*, 33(5):850–860.

Carroll, C. D. (1997). Buffer-stock saving and the life cycle/permanent income hypothesis. *The Quarterly Journal of Economics*, 112(1):1–55.

Carroll, C. D. and Samwick, A. A. (1998). How important is precautionary saving? *Review of Economics and Statistics*, 80(3):410–419.

Chacko, G. and Viceira, L. M. (2005). Dynamic consumption and portfolio choice with stochastic volatility in incomplete markets. *The Review of Financial Studies*, 18(4):1369–1402.

Christelis, D., Jappelli, T., and Padula, M. (2010). Cognitive abilities and portfolio choice. *European Economic Review*, 54(1):18–38.

Coile, C. and Milligan, K. (2009). How household portfolios evolve after retirement: The effect of aging and health shocks. *Review of Income and Wealth*, 55(2):226–248.

Crossley, T. F., Low, H., and O'Dea, C. (2013). Household consumption through recent recessions. *Fiscal Studies*, 34(2):203–229.

Currie, J. and Madrian, B. C. (1999). Health, health insurance and the labor market. In Ashenfelter, O. and Card, D., editors, *Handbook of Labor Economics*, volume 3, chapter 50, pages 3309–3416. Elsevier, Amsterdam.

Dayton, C. M. and Macready, G. B. (1988). Concomitant-variable latent-class models. *Journal of the American Statistical Association*, 83(401):173–178.

Deb, P. and Trivedi, P. K. (1997). Demand for medical care by the elderly: A finite mixture approach. *Journal of Applied Econometrics*, 12(3):313–336.

Deb, P. and Trivedi, P. K. (2002). The structure of demand for health care: Latent class versus two-part models. *Journal of Health Economics*, 21(4):601–625.

Deb, P. and Trivedi, P. K. (2013). Finite mixture for panels with fixed effects. *Journal of Econometric Methods*, 2(1):35–51.

Delavande, A. and Rohwedder, S. (2011). Individuals' uncertainty about future social security benefits and portfolio choice. *Journal of Applied Econometrics*, 26(3):498–519.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(1):1–38.

Edwards, R. D. (2008). Health risk and portfolio choice. *Journal of Business & Economic Statistics*, 26(4):472–485.

Elmendorf, D. W. and Kimball, M. S. (2000). Taxation of labor income and the demand for risky assets. *International Economic Review*, 41(3):801–832.

Fan, E. and Zhao, R. (2009). Health status and portfolio choice: Causality or heterogeneity? *Journal of Banking & Finance*, 33(6):1079–1088.

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

Ferraro, K. F. (1980). Self-ratings of health among the old and the old-old. *Journal of Health and Social Behavior*, 21(4):377–383.

Frick, H., Strobl, C., and Zeileis, A. (2014). To split or to mix? Tree vs. mixture models for detecting subgroups. *COMPSTAT 2014–Proceedings in Computational Statistics*, pages 379–386.

Fuchs-Schündeln, N. and Schündeln, M. (2005). Precautionary savings and self-selection: Evidence from the German reunification "experiment". *The Quarterly Journal of Economics*, 120(3):1085–1120.

Gábor-Tóth, E. and Georgarakos, D. (2018). Economic policy uncertainty and stock market participation. Center for Financial Studies working paper.

Giavazzi, F. and McMahon, M. (2012). Policy uncertainty and household savings. *Review of Economics and Statistics*, 94(2):517–531.

Gil, J., Donni, P. L., and Zucchelli, E. (2018). Uncontrolled diabetes and health care utilisation: A bivariate latent Markov model approach. University of York, Health, Econometrics and Data Group (HEDG) Working Papers 18/28.

Goldman, D. and Maestas, N. (2013). Medical expenditure risk and household portfolio choice. *Journal of Applied Econometrics*, 28(4):527–550.

Gollier, C. and Pratt, J. W. (1996). Risk vulnerability and the tempering effect of background risk. *Econometrica*, 64(5):1109–1123.

Greene, W. (2004). Fixed effects and bias due to the incidental parameters problem in the Tobit model. *Econometric Reviews*, 23(2):125–147.

Gruber, J. (2006). The role of consumer copayments for health care: lessons from the rand health insurance experiment and beyond. Report for the Kaiser Family Foundation, accessed on September 19, 2018, at `www.kff.org/insurance/7566.cfm`.

Gruber, J. and Yelowitz, A. (1999). Public health insurance and private savings. *Journal of Political Economy*, 107(6):1249–1274.

Guiso, L., Jappelli, T., and Terlizzese, D. (1992). Earnings uncertainty and precautionary saving. *Journal of Monetary Economics*, 30(2):307–337.

Guiso, L., Jappelli, T., and Terlizzese, D. (1996). Income risk, borrowing constraints, and portfolio choice. *American Economic Review*, 86(1):158–172.

Hainmueller, J., Mummolo, J., and Xu, Y. (2018). How much should we trust estimates from multiplicative interaction models? Simple tools to improve empirical practice. *Political Analysis*. Forthcoming.

Health and Retirement Study (2018a). Harmonized HRS (v.A). Produced by the USC Program on Global Aging, Health, and Policy, with funding and support from the National Institute on Aging (NIA.).

Health and Retirement Study (2018b). RAND HRS CAMS Spending Data 2015 (V2) public use dataset. Produced by the RAND Center for the Study of Aging, with funding from the National Institute on Aging and the Social Security Administration. Santa Monica, CA.

Health and Retirement Study (2018c). RAND HRS Longitudinal File 2014 (V2) public use dataset. Produced by the RAND Center for the Study of Aging, with funding from the National Institute on Aging and the Social Security Administration. Santa Monica, CA.

Hochguertel, S. (2003). Precautionary motives and portfolio decisions. *Journal of Applied Econometrics*, 18(1):61–77.

Honoré, B. E. (1992). Trimmed LAD and least squares estimation of truncated and censored regression models with fixed effects. *Econometrica*, 60(3):533–565.

Honoré, B. E. (2002). Nonlinear models with panel data. *Portuguese Economic Journal*, 1(2):163–179.

Hothorn, T. and Zeileis, A. (2015). partykit: A modular toolkit for recursive partytioning in R. *The Journal of Machine Learning Research*, 16(1):3905–3909.

Jiménez-Martín, S., Labeaga, J. M., and Martínez-Granado, M. (2002). Latent class versus two-part models in the demand for physician services across the European Union. *Health Economics*, 11(4):301–321.

Johnson, J. (2017). Trump's grand promises to 'very, very quickly' repeal Obamacare run into reality. *The Washington Post*. Retrieved on June 15, 2018, from `https://www.washingtonpost.com/politics/trumps-grand-promises-to-very-very-quickly-repeal-obamacare-run-into-reality/2017/07/18/91b5f220-6bd3-11e7-9c15-177740635e83_story.html?noredirect=on&utm_term=.04a354c7ca5d`.

Kamakura, W. A. and Russell, G. J. (1989). A probabilistic choice model for market segmentation and elasticity structure. *Journal of Marketing Research*, 26(4):379–390.

Kamakura, W. A., Wedel, M., and Agrawal, J. (1994). Concomitant variable latent class models for conjoint analysis. *International Journal of Research in Marketing*, 11(5):451–464.

Karlsson, M. and Laitila, T. (2014). Finite mixture modeling of censored regression models. *Statistical Papers*, 55(3):627–642.

Kimball, M. S. (1990a). Precautionary saving and the marginal propensity to consume. NBER Working Paper No. 3403.

Kimball, M. S. (1990b). Precautionary saving in the small and in the large. *Econometrica*, 58(1):53–73.

Kimball, M. S. (1993). Standard risk aversion. *Econometrica*, 61(3):589–611.

Laird, N. (1978). Nonparametric maximum likelihood estimation of a mixing distribution. *Journal of the American Statistical Association*, 73(364):805–811.

LaRue, A., Bank, L., Jarvik, U., and Hetland, M. (1979). Health in old age: How do physicians' ratings and self-ratings compare? *Journal of Gerontology*, 34(5):687–691.

Lindeboom, M. and Kerkhofs, M. (2009). Health and work of the elderly: Subjective health measures, reporting errors and endogeneity in the relationship between health and work. *Journal of Applied Econometrics*, 24(6):1024–1046.

Litwin, H. and Sapir, E. V. (2009). Forgone health care due to cost among older adults in European countries and in Israel. *European Journal of Ageing*, 6(3):167–176.

Love, D. A. and Smith, P. A. (2010). Does health affect portfolio choice? *Health Economics*, 19(12):1441–1460.

Lumsdaine, R. L. and Exterkate, A. (2013). How survey design affects self-assessed health responses in the survey of health, ageing, and retirement in Europe (SHARE). *European Economic Review*, 63:299–307.

Lusardi, A. (1998). On the importance of the precautionary saving motive. *American Economic Review*, 88(2):449–453.

Luttmer, E. F. and Samwick, A. A. (2018). The welfare cost of perceived policy uncertainty: Evidence from social security. *American Economic Review*, 108(2):275–307.

Mankiw, N. G. and Zeldes, S. P. (1991). The consumption of stockholders and non-stockholders. *Journal of Financial Economics*, 29(1):97–111.

Mazzocco, M. (2004). Saving, risk sharing, and preferences for risk. *American Economic Review*, 94(4):1169–1182.

McLachlan, G. and Peel, D. (2004). *Finite Mixture Models*. John Wiley & Sons, New York.

Miller, S. (2012). The effect of insurance on emergency room visits: An analysis of the 2006 Massachusetts health reform. *Journal of Public Economics*, 96(11-12):893–908.

Mullainathan, S. and Spiess, J. (2017). Machine learning: An applied econometric approach. *Journal of Economic Perspectives*, 31(2):87–106.

Obama, B. (2016). United States health care reform: Progress to date and next steps. *Journal of the American Medical Association*, 316(5):525–532.

Palumbo, M. G. (1999). Uncertain medical expenses and precautionary saving near the end of the life cycle. *The Review of Economic Studies*, 66(2):395–421.

Poterba, J. M. (1994). Govemment saving incentives in the United States. In Poterba, J. M., editor, *Public Policies and Household Savings*, volume 1, chapter 1, pages 1–18. University of Chicago Press, Chicago.

Pratt, J. W. and Zeckhauser, R. J. (1987). Proper risk aversion. *Econometrica*, 55(1):143–154.

Romer, C. D. (1990). The Great Crash and the onset of the Great Depression. *The Quarterly Journal of Economics*, 105(3):597–624.

Rosen, H. S. and Wu, S. (2004). Portfolio choice and health status. *Journal of Financial Economics*, 72(3):457–484.

Rusch, T. and Zeileis, A. (2013). Gaining insight with recursive partitioning of generalized linear models. *Journal of Statistical Computation and Simulation*, 83(7):1301–1315.

Skinner, J. (1988). Risky income, life cycle consumption, and precautionary savings. *Journal of Monetary Economics*, 22(2):237–255.

Smith, J. P. (1999). Healthy bodies and thick wallets: The dual relation between health and economic status. *Journal of Economic Perspectives*, 13(2):145–166.

Starr-McCluer, M. (1996). Health insurance and precautionary savings. *American Economic Review*, 86(1):285–295.

Stephens, Jr, M. (2002). Worker displacement and the added worker effect. *Journal of Labor Economics*, 20(3):504–537.

Stock, J. H. and Watson, M. W. (2012). Disentangling the channels of the 2007-09 recession. *Brookings Papers on Economic Activity*, Spring:81–156.

Strobl, C., Kopf, J., and Zeileis, A. (2015). Rasch trees: A new method for detecting differential item functioning in the Rasch model. *Psychometrika*, 80(2):289–316.

Strobl, C., Wickelmaier, F., and Zeileis, A. (2011). Accounting for individual differences in Bradley-Terry models by means of recursive partitioning. *Journal of Educational and Behavioral Statistics*, 36(2):135–153.

Suderman, P. (2017). Why we're still fighting over the health care mandate (op-ed). *The New York Times*. Retrieved on June 15, 2018, from `https://www.nytimes.com/2017/11/15/opinion/republican-tax-plan-health-care.html`.

The Conference Board (2001). Business cycle indicators handbook.

The Conference Board (2018). Coincident Economic Index® (CEI).

Toney, J. and Bogan, V. L. (2018). How extended family mental health issues influence household portfolio allocations. working paper.

United States Bureau of Labor Statistics (2018). Consumer Price Index for All Urban Consumers: All Items. Retrieved on June 20, 2018, from `https://beta.bls.gov/dataViewer/view/timeseries/CUSR0000SA0`.

United States Congress (2010). The Patient Protection and Affordable Care Act, 42 U.S.C. § 18001 et seq. Accessed on July 3, 2018, at `https://www.gpo.gov/fdsys/pkg/PLAW-111publ148/pdf/PLAW-111publ148.pdf`.

Van der Heijden, P. G., Dessens, J., and Bockenholt, U. (1996). Estimating the concomitant-variable latent-class model with the EM algorithm. *Journal of Educational and Behavioral Statistics*, 21(3):215–229.

Wager, S. and Athey, S. (2018). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, 113(523):1228–1242.

Wagner, M. and Zeileis, A. (2019). Heterogeneity and spatial dependence of regional growth in the EU: A recursive partitioning approach. *German Economic Review*, 20(1):67–82.

Wedel, M., DeSarbo, W. S., Bult, J. R., and Ramaswamy, V. (1993). A latent class Poisson regression model for heterogeneous count data. *Journal of Applied Econometrics*, 8(4):397–411.

Wiemann, T. T. (2018). Health care policy uncertainty in the United States and its effect on households' consumption and portfolio choice. Bachelor's thesis.

Wu, C. J. (1983). On the convergence properties of the EM algorithm. *The Annals of Statistics*, 11(1):95–103.

Wu, S. (2003). The effects of health events on the economic status of married couples. *Journal of Human Resources*, 38(1):219–230.

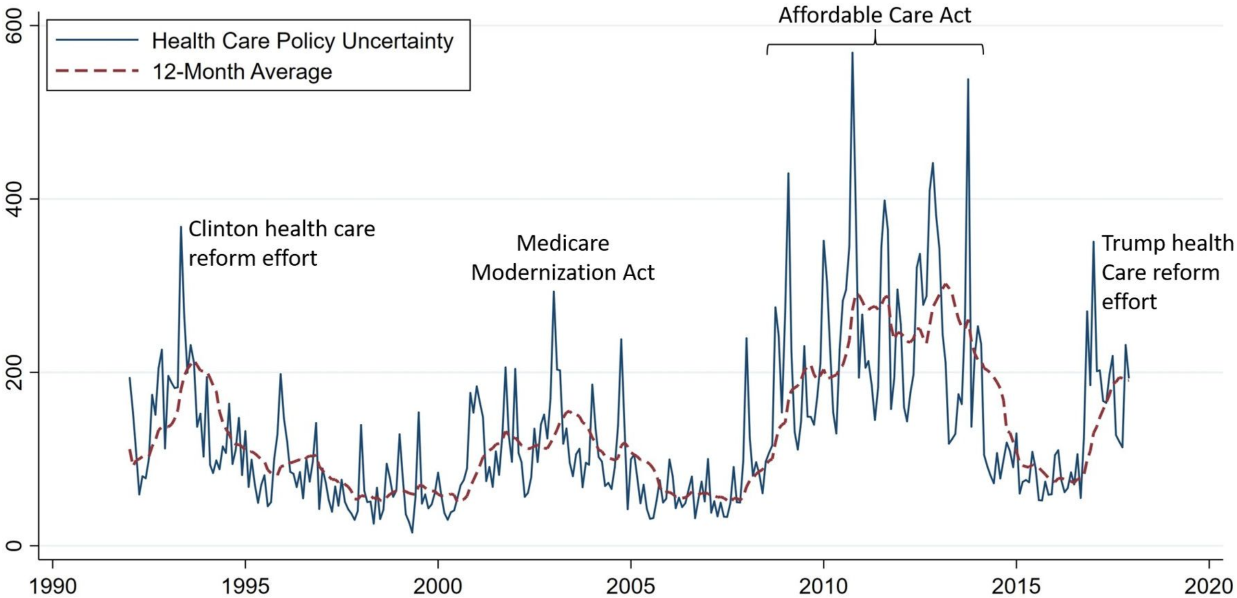
Zeileis, A. and Hornik, K. (2007). Generalized m-fluctuation tests for parameter instability. *Statistica Neerlandica*, 61(4):488–508.

Zeileis, A., Hothorn, T., and Hornik, K. (2008). Model-based recursive partitioning. *Journal of Computational and Graphical Statistics*, 17(2):492–514.

Zeldes, S. P. (1989). Optimal consumption with stochastic income: Deviations from certainty equivalence. *The Quarterly Journal of Economics*, 104(2):275–298.

Zweifel, P. and Manning, W. G. (2000). Moral hazard and consumer incentives in health care. In *Handbook of Health Economics*, volume 1, chapter 9, pages 409–459. Elsevier, Amsterdam.

Figure 1: Health Care Policy Uncertainty Index 1992 - 2017



The index reflects the scaled monthly number of newspaper articles containing the word-triplet "uncertain", "economic", and "policy" (and their synonyms) and one term on health care (e.g., "health insurance"). It is calculated on basis of the Access World News newspaper archive with about 1,500 US papers, normalized to a mean of 100 from 1985 to 2010. The dashed line is the index's average based on the preceding 12 months.

Table 1: Summary Statistics

Single households	Mean	Sd.		Mean	Sd.		Share
Dependent			*Hh Characteristics*			*Hh Proportions (%)*	
Total spend. ($100)	317.757	249.988	Self-reported health	2.826	1.098	Female	0.713
Risky assets (%)	0.137	0.278	Severe conditions	0.707	0.899	Any children	0.823
Safe assets (%)	0.674	0.400	Mobility	1.719	1.807	Retired	0.575
IRA (%)	0.174	0.310	ADLs	0.431	1.108	No high school	0.167
Bonds (%)	0.015	0.085	Hospital nights	2.279	8.892	GED	0.038
Macro. Controls			Uninsured (%)	0.060	0.237	High school	0.329
HPU	127.144	102.650	Income ($100,000)	0.444	1.122	Some college	0.250
HPU_{12}	131.563	75.847	Wealth ($100,000)	3.001	11.909	Above college	0.215
VIX_{12}	20.118	4.756	Years ret.	8.273	10.882	White	0.863
$SP500_{12}$	0.009	0.013	Age	68.784	11.179	Black	0.101
CEI_{12}	0.002	0.001				Other	0.036

Couple households	Mean	Sd.		Mean	Sd.		Share
Dependent			*Hh Characteristics*			*Hh Proportions (%)*	
Total spend. ($100)	573.773	384.518	Self-reported health	3.055	1.014	Any children	0.958
Risky assets (%)	0.168	0.284	Severe conditions	0.894	0.909	Retired	0.579
Safe assets (%)	0.513	0.409	Mobility	1.881	1.744	No high school	0.047
IRA (%)	0.302	0.361	ADLs	0.464	1.185	GED	0.023
Bonds (%)	0.017	0.081	Hospital nights	2.836	10.302	High school	0.243
Macro. Controls			Uninsured (%)	0.067	0.251	Some college	0.273
HPU	128.903	102.990	Income ($100,000)	1.054	2.412	Above college	0.414
HPU_{12}	133.632	74.974	Wealth ($100,000)	5.387	13.627	White-White	0.886
VIX_{12}	19.918	4.829	Years ret.	6.877	9.172	Black-Black	0.048
$SP500_{12}$	0.009	0.013	Age	64.925	9.265	Other-Other	0.020
CEI_{12}	0.002	0.001				White-Black	0.005
						White-Other	0.039
						Black-Other	0.003

Statistics for single households are based on the sample of 42,786 household-wave observations, except for total spending where the CAMS sub-sample of 12,116 household-wave observations is used. For couple households, there are 54,717 observations used in the analysis, with 7,862 contained in the CAMS sub-sample. The HRS-provided household analysis weights are used for calculation. Following Rosen and Wu (2004), safe assets are checking and savings accounts, CDs, government savings bonds and T-bills, and risky assets are stocks and mutual funds. HPU_{12} and VIX_{12} denote the 12-month average of Baker et al.'s (2016) health care policy uncertainty index and the Implied Volatility Index, respectively. $SP500_{12}$ and CEI_{12} denote the 12-month average of the growth rates of the S&P 500 Index and the Conference Board's Coincident Economic Indicator. Health measures, age, years of retirement, retirement status, and highest obtained education are defined on the maximum across spouses. A household is considered uninsured if either spouse in not covered by private of governmental health insurance. $ denotes 2010 Dollars.

Table 2: Fixed Effect and Pooled Tobit Model Results

Couple households	(1) Fixed effects	(2) Fixed effects	(3) Pooled	(4) Fixed effects	(5) Pooled
	log(Total spending)	Risky asset share		Safe asset share	
$\log(\mathrm{HPU}_{12})$	-0.039***	-0.037***	-0.105***	0.022***	0.044***
	(0.012)	(0.006)	(0.005)	(0.004)	(0.005)
$\bar{\partial}\log(\mathrm{HPU}_{12})$			-0.040		0.031
Log-likelihood			-31.662		-41.682
Households	1,729	10,253		10,253	
Observations	7,862	54,717		54,717	
Share censored	0	0.62		0.35	

Single households	(6) Fixed effects	(7) Fixed effects	(8) Pooled	(9) Fixed effects	(10) Pooled
	log(Total spending)	Risky asset share		Safe asset share	
$\log(\mathrm{HPU}_{12})$	-0.017	-0.052***	-0.146***	0.040***	0.061***
	(0.011)	(0.010)	(0.008)	(0.007)	(0.007)
$\bar{\partial}\log(\mathrm{HPU}_{12})$			-0.037		0.028
Log-likelihood			-21.333		-30.915
Households	2,914	9,537		9,537	
Observations	12,116	42,786		42,786	
Share censored	0	0.75		0.57	

Results in column (1) and (6) correspond to a linear fixed effect regression. Results in column (2), (4), (7), and (9) correspond to Honoré's (1992) censored fixed effect estimator. Results in column (3), (5), (8), and (10) correspond to a pooled Tobit model. *, ** and *** denote significance at 10%, 5%, 1% , respectively. Asymptotic standard errors are presented in parentheses. $\bar{\partial}\log(\mathrm{HPU}_{12})$ denotes the average marginal effect corresponding to the coefficient of log of the 12-month average of the health care policy uncertainty index. Note that for the log(Total spending) results using the pooled normal model, marginal effects are constant and hence $\bar{\partial}\log(\mathrm{HPU}_{12})$ is equal to the coefficient estimate $\log(\mathrm{HPU}_{12})$. Total spending is measured in 2010 Dollars. Following Rosen and Wu (2004), safe assets are checking and savings accounts, CDs, government savings bonds and T-bills, and risky assets are stocks and mutual funds. Differences in sample size across columns are due to varying numbers of observations included in the HRS and CAMS datasets as well as varying numbers of couple and single households therein. Section 3.1.4 discusses differences in sample sizes in further detail.

Table 3: Latent Class Model Results – Couple Households

	(1) log(Total spending)		(2) Risky asset share		(3) Safe asset share	
Prior Probability						
Constant	0	−2.329	0	2.026***	0	−0.361***
	-	(2.409)	-	(0.076)	-	(0.037)
Self-reported health	0	0.128	0	0.195***	0	0.293***
	-	(0.295)	-	(0.031)	-	(0.013)
Severe conditions	0	0.067	0	−0.006	0	0.049***
	-	(0.201)	-	(0.028)	-	(0.014)
Mobility	0	0.081	0	−0.008	0	0.066***
	-	(0.113)	-	(0.017)	-	(0.008)
ADLs	0	0.166	0	0.004	0	0.028**
	-	(0.182)	-	(0.026)	-	(0.013)
Nights in hospital	0	−0.003	0	−0.002	0	−0.001
	-	(0.019)	-	(0.001)	-	(0.001)
Uninsured	0	1.336*	0	0.625***	0	0.838***
	-	(0.749)	-	(0.11)	-	(0.057)
Components	*s.1*	*s.2*	*s.1*	*s.2*	*s.1*	*s.2*
$\log(\text{HPU}_{12})$	−0.088	−0.141	0.000	−0.103***	−0.002	0.076***
	(0.110)	(0.101)	(0.006)	(0.005)	(0.001)	(0.006)
$\partial\log(\text{HPU}_{12})\big\rvert_{\bar{X}_C}$	-	-	0.000	−0.029	−0.002	0.037
Component share	0.804	0.196	0.067	0.933	0.318	0.682
Log-likelihood	-6,203.5		-29,669.4		-24,643.6	
Observations	7,862		54,717		54,717	

Results in column (1) correspond to a latent class normal model. Results in column (2) and (3) correspond to a latent class Tobit model. *, ** and *** denote significance at 10%, 5%, 1% , respectively. Estimates are calculated using the EM algorithm illustrated in Appendix C. Bootstrapped standard errors are in parentheses. $\partial\log(HPU_{12})\big\rvert_{\bar{X}_s}$ denotes the marginal effect at the mean observation of each component as identified using the posterior component means. Note that for the log(Total spending) results using the pooled normal model, marginal effects are constant and hence $\partial\log(HPU_{12})\big\rvert_{\bar{X}_s}$ is equal to the coefficient estimate $\log(\text{HPU}_{12})$. Total spending is measured in 2010 Dollars. Following Rosen and Wu (2004), safe assets are checking and savings accounts, CDs, government savings bonds and T-bills, and risky assets are stocks and mutual funds. Differences in sample size across columns are due to varying numbers of observations included in the HRS and CAMS datasets. Section 3.1.4 discusses differences in sample sizes in further detail.

Table 4: Posterior Component Means – Couple Households

log(Total spending)

	Self. health	Severe cond.	Mobil.	ADLs	Hosp. nights	Unins.	Age	No HS	HS	Above coll.	Wealth
s.1	3.10	0.99	1.95	0.42	2.82	0.06	67.71	0.06	0.28	0.38	4.83
s.2	3.45	1.16	2.61	0.93	3.57	0.19	67.46	0.11	0.27	0.32	5.17

Risky asset share

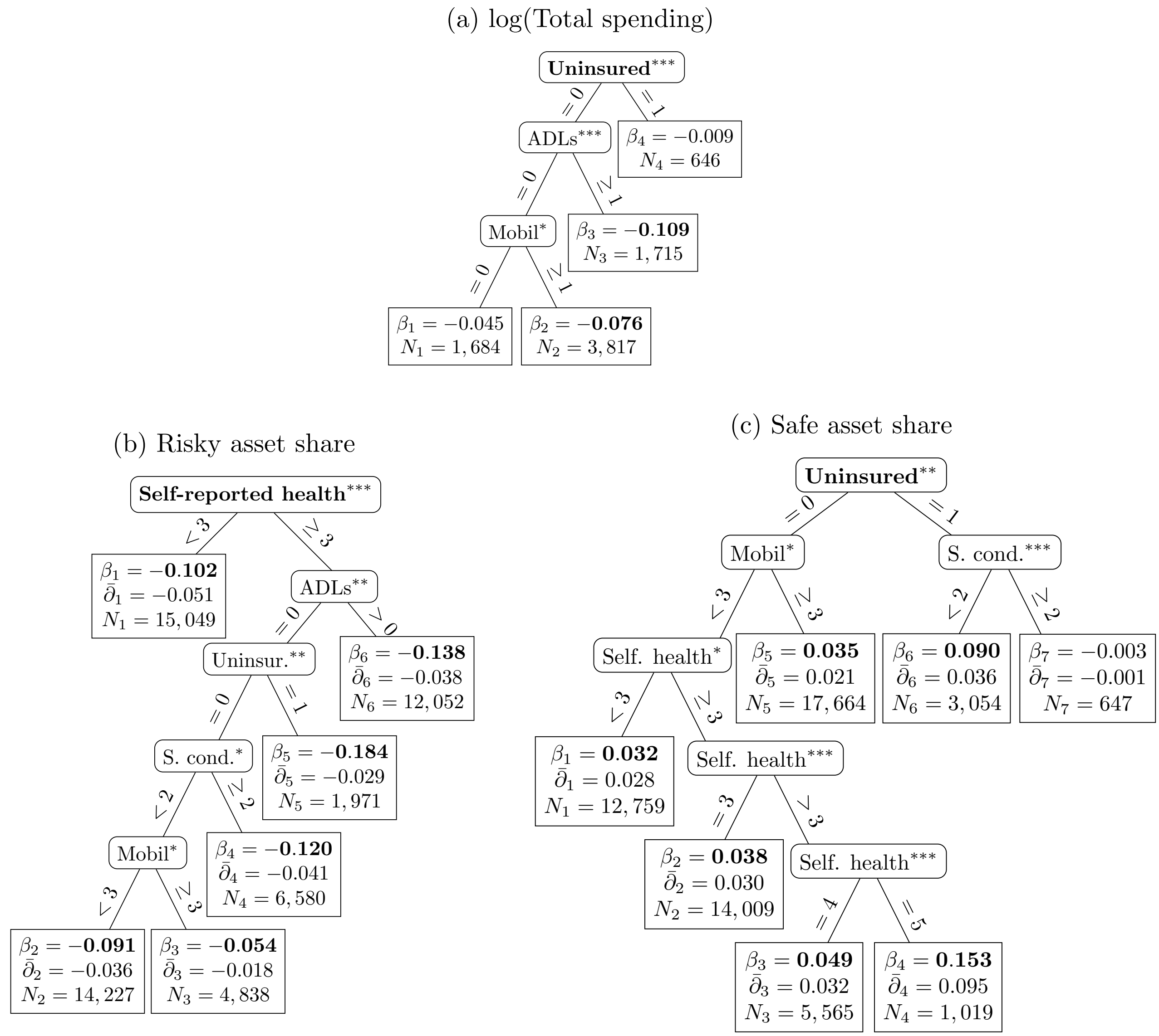
	Self. health	Severe cond.	Mobil.	ADLs	Hosp. nights	Unins.	Age	No HS	HS	Above coll.	Wealth
s.1	2.96	0.91	1.86	0.44	3.07	0.04	67.23	0.02	0.23	0.49	8.22
s.2	3.15	0.97	2.01	0.53	3.20	0.07	66.98	0.09	0.29	0.30	4.85

Safe asset share

	Self. health	Severe cond.	Mobil.	ADLs	Hosp. nights	Unins.	Age	No HS	HS	Above coll.	Wealth
s.1	2.87	0.85	1.66	0.35	2.65	0.04	66.33	0.03	0.22	0.48	6.90
s.2	3.26	1.03	2.16	0.60	3.44	0.08	67.31	0.09	0.30	0.30	4.22

'No HS', 'HS', and 'Above coll.' denote the education categories 'No high school', 'High school', and 'Above college', respectively. Wealth is in \$100,000 2010 Dollars.

Figure 2: Model-Based Recursive Partitioning Tree – Couple Households



Estimates in Figure 2a correspond to a pooled normal model. Results in Figure 2b and 2c correspond to a pooled Tobit model. At each step of the partition, the feature with the most significant rejection of parameter-stability is reported in each connecting node (rounded corners), with ***, **, and * indicating significance at 1%, 5%, and 10% levels of probability, respectively. For example, the feature with the most significant rejection in the first iteration of the procedure is printed in **bold**; readers should read the tree starting at this node. Terminal nodes (sharp corners), state the coefficient on the log-transformed annual health care policy uncertainty average (β), the corresponding average marginal effect ($\bar{\partial}$), and the partition's sample size (N). Note that for the log(Total spending) results using the pooled normal model, $\beta = \bar{\partial}$ and hence there is no need for a separate line to report the marginal effects in Figure 2a. Significance of coefficients at a greater than 95% level of confidence is indicated by **bold** print. Subscripts indicate the node numbering. Following Rosen and Wu (2004), safe assets are checking and savings accounts, CDs, government savings bonds and T-bills, and risky assets are stocks and mutual funds. Complete estimation results are in Appendix D.3.

Appendices

A Proof of Proposition 1 and 2

This section provides a proof for Proposition 1 and 2 of the model developed in Section 2.1.

<u>*Preliminaries.*</u> For ease of notation, define $[C_1^*, x^*]$ to be the maximizing arguments in the case of no uncertainty about P, and define $[\tilde{C}_1^*, \tilde{x}^*]$ to be the maximizing arguments in the case of uncertainty about P. Then, Proposition 1 can be expressed as

$$C_1^* > \tilde{C}_1^* \qquad \text{and} \qquad x^* > \tilde{x}^* \tag{15}$$

and Proposition 2 can be expressed as

$$C_1^*(H_g) - \tilde{C}_1^*(H_g) < C_1^*(H_b) - \tilde{C}_1^*(H_b) \qquad \text{and} \qquad x^*(H_g) - \tilde{x}^*(H_g) < x^*(H_b) - \tilde{x}^*(H_b) \tag{16}$$

where consumption and investment are expressed as a function of required units of treatment (i.e., health) with $H_g < H_b$.

Because the expectation is a positive linear operation, $U(C)$ is strictly concave in C, and C_2 is linear in C_1 and x, proving (15) and (16) is equivalent to proving

$$\begin{aligned} & \frac{\partial E[U(C_1)+U(C_2)]}{\partial C_1} > \frac{\partial \tilde{E}[U(C_1)+U(C_2)]}{\partial C_1} \\ \text{and} \quad & \frac{\partial E[U(C_1)+U(C_2)]}{\partial x} > \frac{\partial \tilde{E}[U(C_1)+U(C_2)]}{\partial x} \end{aligned} \tag{17}$$

where $\tilde{E}$ denotes the expectation under uncertainty about P, as well as

$$\begin{aligned} & \left.\frac{\partial E[U(C_1)+U(C_2)]}{\partial C_1}\right|_{H_g} - \left.\frac{\partial \tilde{E}[U(C_1)+U(C_2)]}{\partial C_1}\right|_{H_g} \\ < \quad & \left.\frac{\partial E[U(C_1)+U(C_2)]}{\partial C_1}\right|_{H_b} - \left.\frac{\partial \tilde{E}[U(C_1)+U(C_2)]}{\partial C_1}\right|_{H_b} \\ \text{and} \quad & \left.\frac{\partial E[U(C_1)+U(C_2)]}{\partial x}\right|_{H_g} - \left.\frac{\partial \tilde{E}[U(C_1)+U(C_2)]}{\partial x}\right|_{H_g} \\ < \quad & \left.\frac{\partial E[U(C_1)+U(C_2)]}{\partial x}\right|_{H_b} - \left.\frac{\partial \tilde{E}[U(C_1)+U(C_2)]}{\partial x}\right|_{H_b} \end{aligned} \tag{18}$$

where the left-hand side corresponds to a household with lower required units of treatment ($H_g < H_b$). Note that the household only chooses C_1 and x, which, depending on the period

2 realization of the random states, result in a level of C_2. Because C_2 is linear in C_1 and x, the application of the chain rule needed in the following steps is straightforward.

For the case of no uncertainty about P, the first order conditions with respect to C_1 and x are given (via application of the chain rule) by

$$\begin{aligned}
0 &= \frac{\partial E[U(C_1) + U(C_2)]}{\partial C_1} \\
&= \frac{\partial U(C_1)}{\partial C_1} - p_1\big[1 + xr_1 + (1-x)b\big]\frac{\partial U(C_2)}{\partial C_2}\bigg|_{C_2 = C_2^1} - p_2\big[1 + xr_2 + (1-x)b\big]\frac{\partial U(C_2)}{\partial C_2}\bigg|_{C_2 = C_2^2}
\end{aligned} \tag{19}$$

$$\begin{aligned}
0 &= \frac{\partial E[U(C_1) + U(C_2)]}{\partial x} \\
&= p_1\big[(W_0 - C_1)(r_1 - b)\big]\frac{\partial U(C_2)}{\partial C_2}\bigg|_{C_2 = C_2^1} + p_2\big[(W_0 - C_1)(r_2 - b)\big]\frac{\partial U(C_2)}{\partial C_2}\bigg|_{C_2 = C_2^2}
\end{aligned} \tag{20}$$

For the case of uncertainty about P, the first order conditions with respect to C_1 and x are given (via application of the chain rule) by

$$0 = \frac{\partial E[U(C_1) + U(C_2)]}{\partial C_1} \tag{21}$$

$$\begin{aligned}
= \frac{\partial U(C_1)}{\partial C_1} \quad & -p_{11}\big[1 + xr_1 + (1-x)b\big]\frac{\partial U(C_2)}{\partial C_2}\bigg|_{C_2 = C_2^{11}} - p_{12}\big[1 + xr_1 + (1-x)b\big]\frac{\partial U(C_2)}{\partial C_2}\bigg|_{C_2 = C_2^{12}} \\
& -p_{21}\big[1 + xr_2 + (1-x)b\big]\frac{\partial U(C_2)}{\partial C_2}\bigg|_{C_2 = C_2^{21}} - p_{22}\big[1 + xr_2 + (1-x)b\big]\frac{\partial U(C_2)}{\partial C_2}\bigg|_{C_2 = C_2^{22}}
\end{aligned}$$

$$0 = \frac{\partial E[U(C_1) + U(C_2)]}{\partial x}$$

$$\begin{aligned}
= \quad & p_{11}\big[(W_0 - C_1)(r_1 - b)\big]\frac{\partial U(C_2)}{\partial C_2}\bigg|_{C_2 = C_2^{11}} + p_{12}\big[(W_0 - C_1)(r_1 - b)\big]\frac{\partial U(C_2)}{\partial C_2}\bigg|_{C_2 = C_2^{12}} \\
+ \quad & p_{21}\big[(W_0 - C_1)(r_2 - b)\big]\frac{\partial U(C_2)}{\partial C_2}\bigg|_{C_2 = C_2^{21}} + p_{22}\big[(W_0 - C_1)(r_2 - b)\big]\frac{\partial U(C_2)}{\partial C_2}\bigg|_{C_2 = C_2^{22}}
\end{aligned} \tag{22}$$

Lemmas. It is useful to begin by deriving two inequalities.

First, for a given positive share invested in the risky asset ($x > 0$) and a given level of consumption in period 1 (C_1), it holds that

$$p_{i1}\frac{\partial U(C_2)}{\partial C_2}\bigg|_{C_2=C_2^{i1}} + p_{i2}\frac{\partial U(C_2)}{\partial C_2}\bigg|_{C_2=C_2^{i2}} \quad > \quad p_i\frac{\partial U(C_2)}{\partial C_2}\bigg|_{C_2=C_2^{i}} \tag{23}$$

$$\Leftrightarrow \quad p_i\left(p_1^H\frac{\partial U(C_2)}{\partial C_2}\bigg|_{C_2=C_2^{i1}} + p_2^H\frac{\partial U(C_2)}{\partial C_2}\bigg|_{C_2=C_2^{i2}}\right) \quad > \quad p_i\frac{\partial U(C_2)}{\partial C_2}\bigg|_{C_2=C_2^{i}} \tag{24}$$

$$\Leftrightarrow \quad p_1^H\frac{\partial U(C_2)}{\partial C_2}\bigg|_{C_2=C_2^{i1}} + p_2^H\frac{\partial U(C_2)}{\partial C_2}\bigg|_{C_2=C_2^{i2}} \quad > \quad \frac{\partial U(C_2)}{\partial C_2}\bigg|_{C_2=C_2^{i}} \tag{25}$$

where the left-hand side corresponds to the case of uncertainty about cost of per-unit treatment and the right-hand side to the certainty case. Equation (25) holds due to the decreasing absolute risk aversion: it follows that the increase in slope between $\frac{\partial U(C_2)}{\partial C_2}\Big|_{C_2=C_2^{i}} < \frac{\partial U(C_2)}{\partial C_2}\Big|_{C_2=C_2^{i1}}$ relative to the decrease in slope between $\frac{\partial U(C_2)}{\partial C_2}\Big|_{C_2=C_2^{i2}} > \frac{\partial U(C_2)}{\partial C_2}\Big|_{C_2=C_2^{i}}$ is larger than the decrease in value between $C_2^i > C_2^{i1}$ relative to the increase in value between $C_2^i < C_2^{i2}$. As C_2^i is given by the linear relation $p_1^H C_2^{i1} + p_2^H C_2^{i2}$, equation (25) follows.

Third, consider the relative increase in expected marginal effect of equation (25) at the two realizations of risky asset returns r_1 and r_2, where $r_1 < r_2$. Because of decreasing absolute prudence, it holds that

$$\frac{p_1^H\frac{\partial U(C_2)}{\partial C_2}\Big|_{C_2=C_2^{11}} + p_2^H\frac{\partial U(C_2)}{\partial C_2}\Big|_{C_2=C_2^{12}}}{\frac{\partial U(C_2)}{\partial C_2}\Big|_{C_2=C_2^{1}}} > \frac{p_1^H\frac{\partial U(C_2)}{\partial C_2}\Big|_{C_2=C_2^{21}} + p_2^H\frac{\partial U(C_2)}{\partial C_2}\Big|_{C_2=C_2^{22}}}{\frac{\partial U(C_2)}{\partial C_2}\Big|_{C_2=C_2^{2}}} \tag{26}$$

where the left-hand side corresponds to the case with low return of the risky asset r_1, the right-hand side corresponds to the case with high return of the risky asset r_2, and the numerators and denominators correspond to the case of uncertainty and no-uncertainty about the per-unit cost of treatment, respectively. Intuitively, this states that the relative effect of the introduction of uncertainty about the per-unit cost of treatment is more substantial at lower levels of consumption (i.e., with low returns of the risky asset r_1) compared to the relative effect at higher levels of consumption (i.e., with high returns of the risky asset r_2).

Proposition 1.1. We first derive that the household decreases consumption in period 1 (C_1) after the introduction of uncertainty about the per-unit cost of treatment. For this purpose, consider the first order conditions (19) and (21). From the inequality (23), it follows that

$$\begin{aligned}
& p_1\big[1+xr_1+(1-x)b\big]\frac{\partial U(C_2)}{\partial C_2}\bigg|_{C_2=C_2^1} + p_2\big[1+xr_2+(1-x)b\big]\frac{\partial U(C_2)}{\partial C_2}\bigg|_{C_2=C_2^2} \\
< \quad & p_{11}\big[1+xr_1+(1-x)b\big]\frac{\partial U(C_2)}{\partial C_2}\bigg|_{C_2=C_2^{11}} + p_{12}\big[1+xr_1+(1-x)b\big]\frac{\partial U(C_2)}{\partial C_2}\bigg|_{C_2=C_2^{12}} \\
+ \quad & p_{21}\big[1+xr_2+(1-x)b\big]\frac{\partial U(C_2)}{\partial C_2}\bigg|_{C_2=C_2^{21}} + p_{22}\big[1+xr_2+(1-x)b\big]\frac{\partial U(C_2)}{\partial C_2}\bigg|_{C_2=C_2^{22}}
\end{aligned} \tag{27}$$

For the first order conditions to hold, $\frac{\partial U(C_1)}{\partial C_1}$ corresponding to the case of uncertainty must be larger than in the case of certainty. As $U(C)$ is increasing and concave, this implies that C_1 corresponding to the case of uncertainty must be smaller than in the case of certainty.

Proposition 1.2. Next, we show that the household decreases its investment share in the risky asset (x) after the introduction of uncertainty about the per-unit cost of treatment. Consider the optimal choice of investment share in the risky asset *without* any uncertainty about the per-unit cost of treatment: x^*. For optimality, the corresponding first order condition (20) at x^* is required to be zero – that is,

$$\begin{aligned}
0 &= \frac{\partial E[U(C_1)+U(C_2)]}{\partial x}\bigg|_{x=x^*}^{certainty} \\
&= p_1\big[(W_0-C_1)(r_1-b)\big]\frac{\partial U(C_2)}{\partial C_2}\bigg|_{C_2=C_2^{1*}} \quad + \quad p_2\big[(W_0-C_1)(r_2-b)\big]\frac{\partial U(C_2)}{\partial C_2}\bigg|_{C_2=C_2^{2*}} \\
&= (W_0-C_1)\left[p_1(r_1-b)\frac{\partial U(C_2)}{\partial C_2}\bigg|_{C_2=C_2^{1*}} \quad + \quad p_2(r_2-b)\frac{\partial U(C_2)}{\partial C_2}\bigg|_{C_2=C_2^{2*}}\right]
\end{aligned} \tag{28}$$

where the partial derivatives with respect to C_2 are introduced via application of the chain rule. This can be compared to the first order condition at the same x^* in case of uncertainty about the per-unit cost of treatment (22):

$$
\begin{aligned}
&\frac{\partial E[U(C_1)+U(C_2)]}{\partial x}\bigg|_{x=x^*}^{uncertainty} \\
= \quad & p_{11}\big[(W_0-C_1)(r_1-b)\big]\frac{\partial U(C_2)}{\partial C_2}\bigg|_{C_2=C_2^{11*}} \\
+ \quad & p_{12}\big[(W_0-C_1)(r_1-b)\big]\frac{\partial U(C_2)}{\partial C_2}\bigg|_{C_2=C_2^{12*}} \\
+ \quad & p_{21}\big[(W_0-C_1)(r_2-b)\big]\frac{\partial U(C_2)}{\partial C_2}\bigg|_{C_2=C_2^{21*}} \\
+ \quad & p_{22}\big[(W_0-C_1)(r_2-b)\big]\frac{\partial U(C_2)}{\partial C_2}\bigg|_{C_2=C_2^{22*}} \\
= \quad & (W_0-C_1)\Bigg[p_1(r_1-b)\bigg(p_1^H\frac{\partial U(C_2)}{\partial C_2}\bigg|_{C_2=C_2^{11*}} + p_2^H\frac{\partial U(C_2)}{\partial C_2}\bigg|_{C_2=C_2^{12*}}\bigg) \\
+ \quad & p_2(r_2-b)\bigg(p_1^H\frac{\partial U(C_2)}{\partial C_2}\bigg|_{C_2=C_2^{21*}} + p_2^H\frac{\partial U(C_2)}{\partial C_2}\bigg|_{C_2=C_2^{22*}}\bigg)\Bigg]
\end{aligned}
\tag{29}
$$

From inequality (23), it follows that we can write the change from the first order condition under certainty about the per-unit treatment cost given by equation (28) versus the introduction of uncertainty in equation (29) by

$$
\begin{aligned}
&\frac{\partial E[U(C_1)+U(C_2)]}{\partial x}\bigg|_{x=x^*}^{uncertainty} \\
= \quad & (W_0-C_1)\Bigg[p_1(r_1-b)\bigg(y\times\frac{\partial U(C_2)}{\partial C_2}\bigg|_{C_2=C_2^{1*}}\bigg) + p_2(r_2-b)\bigg(z\times\frac{\partial U(C_2)}{\partial C_2}\bigg|_{C_2=C_2^{2*}}\bigg)\Bigg]
\end{aligned}
\tag{30}
$$

where y and z are the relative increase in expected slope in the case of low and high returns on the risky asset, respectively. By (26), we know that $y > z$. Hence, we can write $y = (z+\delta)$. This results in

$$
\begin{aligned}
&\frac{\partial E[U(C_1)+U(C_2)]}{\partial x}\Big|_{x=x^*}^{uncertainty} \\
= \quad & (W_0 - C_1)\Big[p_1(r_1-b)\Big(y \times \frac{\partial U(C_2)}{\partial C_2}\Big|_{C_2=C_2^{1*}}\Big) + p_2(r_2-b)\Big(z \times \frac{\partial U(C_2)}{\partial C_2}\Big|_{C_2=C_2^{2*}}\Big)\Big] \\
= \quad & z \times \Big[(W_0 - C_1)\Big(p_1(r_1-b)\frac{\partial U(C_2)}{\partial C_2}\Big|_{C_2=C_2^{1*}} + p_2(r_2-b)\frac{\partial U(C_2)}{\partial C_2}\Big|_{C_2=C_2^{2*}}\Big)\Big] \\
+ \quad & \delta \times p_1(W_0 - C_1)(r_1-b)\frac{\partial U(C_2)}{\partial C_2}\Big|_{C_2=C_2^{1*}} \\
= \quad & z \times \frac{\partial E[U(C_1)+U(C_2)]}{\partial x}\Big|_{x=x^*}^{certainty} + \delta \times p_1(W_0 - C_1)(r_1-b)\frac{\partial U(C_2)}{\partial C_2}\Big|_{C_2=C_2^{1*}} \\
< \quad & 0
\end{aligned} \tag{31}
$$

where the last inequality results from $\frac{\partial E[U(C_1)+U(C_2)]}{\partial x}\Big|_{x=x^*}^{certainty} = 0$, $(r_1 - b) < 0$ and the remaining terms being positive. Because the marginal effect of an additional unit of x is negative at the value x^* in the case of uncertainty about the per-unit cost of treatment, we can conclude that the optimal investment share in the risky asset is lower compared to the case of no uncertainty about the per-unit cost of treatment.

Proposition 2. Finally, we show that the effects stated in Proposition 1 are higher for households with bad health status compared to households with good health status.

Denote the two households' required units of health care treatment by H_b and H_g where $H_b > H_g$ (the higher, the worse). As consumption in period 2 is given by $C_2^{ijb} = (W_0 - C_1)(1 + xr_i + (1-x)b) - H_b * P_j$ and $C_2^{ijg} = (W_0 - C_1)(1 + xr_i + (1-x)b) - H_g * P_j$ for $(i, j) \in \{1, 2\} \times \{1, 2\}$ for bad and good health status, respectively, it holds that $Var(C_2^{ijb}) > Var(C_2^{ijg})$. This is due the difference between the required health expenditures in the two realized per unit cost states is larger for households with bad health status $(H_b * P_1 - H_b * P_2)$ than for households with good health status $(H_g * P_1 - H_g * P_2)$. Therefore, households with bad health are exposed to higher variance in *total* medical expenditures when facing the same level of uncertainty about *per-unit* treatment costs than those with good health. Given that the utility function is characterized by decreasing absolute risk aversion and decreasing absolute prudence, it holds that the difference in magnitudes between the left-hand side and the right-hand side of inequality of (27) and the inequality of (31) are higher in case of bad health status compared to good health status. Consequently, the decrease in consumption in period 1 (C_1) and in investment in the risky asset (x) is stronger for households with worse health.

B Construction of the Analysis Sample

Table B.1 shows the number of observations that were excluded from the analysis sample due to missing information. Additionally, we omit 13 and 8 observations of single and couple households with zero total spending, respectively. The differences in starting observations for each subsample is due to unequal distribution of couple and single households. Further, as noted in Section 3.1.2, the spending data considers the CAMS dataset, which is a subset of data from the HRS sample.

The reductions in sample size are in line with existing studies that use the HRS dataset for household analysis. In particular, Love and Smith (2010) consider the first nine waves of the HRS (1992-2006) with a total sample size of 37,962 and 43,595 for single and couple households, respectively. In comparison with the eleven HRS waves (1994-2014) employed in this study, these numbers point to a similar proportion of observations excluded. Most recently, Gábor-Tóth and Georgarakos (2018) consider the CAMS dataset also employed in this study. The authors consider a pooled single and couple households sample of 19,797 observations for their analysis, similar to the sample we use (with a total of 19,999 observations used for analysis of consumption).

Table B.1: Tabulation of Deleted Observations

		CAMS		HRS	
		Single	Couple	Single	Couple
Starting		17,847	11,811	68,507	72,426
Total spending	-	4,022	2,837		
Risky asset share	-			16,084	8,702
Retired	-	240	115	2,790	2,402
Education	-	0	0	5	4
Race	-	11	296	12	1,329
Any children	-	271	120	840	545
Self-reported health	-	11	6	42	31
Severe conditions	-	59	29	255	192
Mobility	-	53	14	1,906	1,743
Uninsured	-	23	10	147	140
ADLs	-	0	0	1	3
Hospital nights	-	65	31	343	158
# Obs. per household	-	963	483	3,296	2,460
Total spending = 0	-	13	8		
Final Obs	=	12,116	7,862	42,786	54,717

C Further Notes on the Methodology

C.1 Anecdotal Evidence on Standard Errors

Table C.1 presents fixed effect regression results with conventional (in parentheses) and household-level clustered standard errors (in brackets). There do not seem to be substantive differences in standard errors.

Table C.1: Linear Fixed Effect Estimation Results

	(1)	(2)	(3)	(4)	(5)	(6)
	Couple	Single	Couple	Single	Couple	Single
	log(Total spending)		Risky asset share		Safe asset share	
$\log(\text{HPU}_{12})$	-0.039***	-0.017	-0.016***	-0.015***	0.016***	0.02***
	(0.012)	(0.011)	(0.002)	(0.002)	(0.003)	(0.003)
	[0.011]	[0.011]	[0.002]	[0.003]	[0.003]	[0.003]
VIX_{12}	0.005***	0.004**	0.001***	0.000	-0.001***	-0.001**
	(0.002)	(0.002)	(0.000)	(0.000)	(0.000)	(0.000)
	[0.002]	[0.002]	[0.000]	[0.000]	[0.000]	[0.000]
SP500_{12}	2.031***	0.510	-0.053	-0.249**	-0.540***	0.087
	(0.463)	(0.457)	(0.114)	(0.122)	(0.146)	(0.157)
	[0.488]	[0.494]	[0.108]	[0.122]	[0.139]	[0.157]
CEI_{12}	18.209***	12.629***	0.105	0.462	0.079	-1.524
	(5.038)	(4.738)	(1.136)	(1.180)	(1.449)	(1.518)
	[4.737]	[4.436]	[1.105]	[1.157]	[1.422]	[1.515]
2^{nd} Inc. Q.	-0.041	0.063***	0.008	0.007**	-0.003	-0.005
	(0.031)	(0.017)	(0.006)	(0.003)	(0.008)	(0.004)
	[0.035]	[0.017]	[0.005]	[0.003]	[0.009]	[0.005]
3^{nd} Inc. Q.	-0.007	0.091***	0.012*	0.028***	0.001	-0.011**
	(0.031)	(0.020)	(0.006)	(0.004)	(0.008)	(0.005)
	[0.035]	[0.020]	[0.006]	[0.005]	[0.009]	[0.006]
4^{nd} Inc. Q.	0.002	0.121***	0.030***	0.027***	0.005	-0.004
	(0.032)	(0.026)	(0.007)	(0.005)	(0.008)	(0.007)
	[0.036]	[0.026]	[0.006]	[0.006]	[0.009]	[0.008]
2^{nd} Wealth. Q.	0.031	0.062***	0.016***	0.031***	-0.021***	-0.017***
	(0.027)	(0.019)	(0.005)	(0.004)	(0.007)	(0.005)
	[0.030]	[0.020]	[0.005]	[0.004]	[0.008]	[0.006]
3^{nd} Wealth. Q.	0.060**	0.092***	0.072***	0.079***	-0.042***	-0.036***
	(0.030)	(0.024)	(0.006)	(0.005)	(0.007)	(0.006)
	[0.034]	[0.025]	[0.006]	[0.006]	[0.009]	[0.008]
4^{nd} Wealth. Q.	0.092***	0.133***	0.171***	0.166***	-0.050***	-0.056***
	(0.034)	(0.030)	(0.006)	(0.006)	(0.008)	(0.008)
	[0.038]	[0.031]	[0.008]	[0.008]	[0.010]	[0.010]
Retired	-0.040**	-0.052***	-0.013***	-0.012***	-0.043***	-0.023***
	(0.018)	(0.019)	(0.003)	(0.004)	(0.004)	(0.005)
	[0.020]	[0.021]	[0.004]	[0.004]	[0.005]	[0.006]
Years ret.	-0.001	0.001	0.000*	0.000***	0.002***	0.001***
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
	[0.001]	[0.001]	[0.000]	[0.000]	[0.000]	[0.000]
Age	-0.021***	-0.025***	-0.002***	-0.002***	0.004***	0.003***
	(0.002)	(0.002)	(0.000)	(0.000)	(0.000)	(0.000)
	[0.002]	[0.002]	[0.000]	[0.000]	[0.000]	[0.000]
Households	1,729	2,914	10,253	9,537	10,253	9,537
Observations	7,862	12,116	54,717	42,786	54,717	42,786

*, ** and *** denote significance at 10%, 5%, 1%, respectively, as assessed by conventional standard errors in parentheses. Household-level clustered standard errors are in brackets. Time-invariant household variables are subsumed by the fixed effects.

C.2 EM Algorithm

To the best of our knowledge, no version of the EM algorithm exists for specifically estimating concomitant-variable latent class Tobit models. While its derivation follows from other modifications – in particular Van der Heijden et al. (1996) and Karlsson and Laitila (2014) , who develop a variant of the EM algorithm for finite mixture Tobit models and concomitant-variable latent class models, respectively – for completeness we include in this section a description of the specific EM variant employed for this paper. Below we follow McLachlan and Peel (2004) for our introduction of the latent class model as an incomplete-data problem and our discussion of the E- and M-Step of the EM algorithm.[53] The procedure is implemented in R, with code readily available upon request.

The EM algorithm of Dempster et al. (1977) is a suitable alternative for the estimation of mixture models, including the latent class model employed in this paper, as it overcomes the computational complexities of the likelihood in equation (7). The trick is to view the estimation problem as an incomplete-data problem, where a vector of component labels for each observation exists – i.e., we assume the model specified in (6) is true but is unobserved. Let S_i be this missing component membership vector, whose elements $s_{j,i}$ are binary indicators equal to one if y_i stems from mixture component j and zero otherwise. Then, S_i are realized (unobserved) draws from the random vector $\tilde{S}_i$ with multinomial distribution consisting of one draw from M categories with probabilities $[\pi_{1,i}, \ldots, \pi_{M,i}]$ that satisfy $\sum_{j=1}^{M} \pi_{j,i} = 1$. Assuming that we actually *do* observe component membership, the *complete-data* log-likelihood is given by

$$\ell_C(\boldsymbol{\theta}) = \sum_{i=1}^{N} \sum_{j=1}^{M} s_{j,i} \Big(\log\left(\pi_{j,i}\right) + \log\left(f(y_i|X_i, \theta_j)\right) \Big), \tag{32}$$

where $f(y_i|X_i, \theta_j)$ is a pooled normal density in the case of total spending or the pooled Tobit density specified in equation (8) in the case of risky and safe assets, $\pi_{j,i}$ is the concomitant-variable prior probability specification given by equation (9) with arguments Z and parameters α, and $\boldsymbol{\theta}$ summarizes all component specific coefficient vectors (θ_j) and the concomitant-variable coefficients (α). For notational convenience, the panel structure of the data is ignored; N denotes the total amount of household-wave observations. The EM algorithm then proceeds iteratively in two steps; an expectation step and a maximization step. Below, both steps are illustrated, as well as the initialization of the algorithm, its convergence-criterion, and the standard error approximation.

[53] Readers interested in a detailed description of the general EM algorithm for parametric finite mixture models can refer to section 2.8 of McLachlan and Peel (2004).

E-Step. The expectation step requires the computation of the conditional expectation of the elements of the unobserved component label vector $\tilde{s}_{j,i}$ given the data y_i, X_i and Z_i. Iteration $(k+1)$, is given by

$$\begin{aligned} E_{\boldsymbol{\theta}^{(k)}}\left[\tilde{s}_{j,i}|y_i, X_i, Z_i, \boldsymbol{\theta}^{(k)}\right] &= P_{\boldsymbol{\theta}^{(k)}}\left[\tilde{S}_i = j|y_i, X_i, Z_i, \boldsymbol{\theta}^{(k)}\right] \\ &= \frac{\pi_{j,i}^{(k)} f(y_i|X_i, \theta_j^{(k)})}{\sum_{m=1}^{M} \pi_{m,i}^{(k)} f(y_i|X_i, \theta_m^{(k)})} \\ &= \tau_{j,i}^{(k+1)} \end{aligned} \tag{33}$$

where $\tau_{j,i}^{(k+1)}$ is equivalent to the posterior component probability defined in equation (10) when using the coefficient estimate of the kth EM iteration.

M-Step. The maximization step of the $(k+1)$th iteration requires the global maximization of the conditional complete data likelihood with respect to the parameters $\boldsymbol{\theta}$ in order to obtain a new set of coefficient estimates to be used for the next iteration ($\boldsymbol{\theta}^{(k+1)}$). That is,

$$\begin{aligned} \max_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(k)}) &= E_{\boldsymbol{\theta}^{(k)}}\left[\ell(\boldsymbol{\theta})|y, X, Z, \boldsymbol{\theta}^{(k)}\right] \\ &= \sum_{i=1}^{N}\sum_{j=1}^{M} \tau_{j,i}^{(k+1)} \Big(\log\left(\pi_{j,i}\right) + \log\left(f(y_i|X_i, \theta_j)\right)\Big), \end{aligned} \tag{34}$$

where $\boldsymbol{\theta}$ summarizes all component specific coefficient vectors (θ_j) and the concomitant-variable coefficients (α). The conditional likelihood can be split into $(M+1)$ parts, corresponding to the prior probability model and the M component-specific Tobit models, that can be optimized independently. Namely, (34) is equivalent to

$$\begin{aligned} &\max_{\alpha} \sum_{i=1}^{N}\sum_{j=1}^{M} \tau_{j,i}^{(k+1)} \log(\pi_{j,i}) \\ =&\sum_{i=1}^{N}\sum_{j=1}^{M} \tau_{j,i}^{(k+1)} \log\left(\frac{\exp\left(\alpha_{j,0} + Z_{it}\alpha_{j,1}\right)}{\sum_{m=1}^{M} \exp\left(\alpha_{m,0} + Z_{it}\alpha_{m,1}\right)}\right) \end{aligned} \tag{35}$$

where the coefficients of component 1 are set to 0 for identification (i.e., $\alpha_{1,0} = 0$ and $\alpha_{1,1} = 0$), and

$$\max_{\theta_j} \sum_{i=1}^{N} \tau_{j,i}^{(k+1)} \log\left(f(y_i|X_i, \theta_j)\right) \qquad \text{for } j \in \{1, \ldots, M\}. \tag{36}$$

As (35) and (36) are linear transformations of the multinomial logit log-likelihood and the normal or Tobit log-likelihood, respectively, maximization of the three independent parts can be done via conventional maximization techniques.[54] This leads to a new set of coefficient estimates to be used in the next iteration's E-step.

Initialization. To calculate the first E-step, an initial guess of the coefficient vector $\boldsymbol{\theta}^{(0)}$ is needed. McLachlan and Peel (2004) present several alternatives. While a grid search across various initial configurations of $\boldsymbol{\theta}^{(0)}$ seems ideal, this is computationally infeasible due to the high number of parameters considered in this study. Instead, we consider an alternative initialization step through a random guess of the posterior probabilities. That is, $\tau_{j,i}^{(0)}$ are taken as random draws from a uniform distribution between 0.1 and 0.9. These are then considered in a pre-estimation M-step to arrive at an initial guess $\boldsymbol{\theta}^{(0)}$.

For each model, the EM algorithm is executed multiple times to avoid dependency on the specific starting value. This is particularly important as the EM algorithm has been shown to potentially lead to local rather than global maxima (e.g., McLachlan and Peel, 2004). Further, the likelihood of a finite mixture of Tobit models is not globally concave, giving further reason for concern that the EM algorithm terminates in undesirable local maxima. Karlsson and Laitila (2014) suggest the use of several starting points in line with Wu (1983). To alleviate these concerns, we consider 100 random starting values for each model. The instance with the highest likelihood value is selected as the best maximum likelihood estimate.

Convergence criterion. McLachlan and Peel (2004) discuss various convergence criteria of the EM algorithm. One that seems to be commonly applied is that the procedure terminates once $L(\boldsymbol{\theta}^{(k+1)}) - L(\boldsymbol{\theta}^{(k)})$ changes by an arbitrarily small amount – that is, once the values of likelihood (defined in (7)) converge sufficiently. In particular, the EM algorithm employed in this study terminates when

$$L(\boldsymbol{\theta}^{(k+1)}) - L(\boldsymbol{\theta}^{(k)}) < 10^{-6}. \tag{37}$$

Standard errors. In contrast to maximum likelihood estimation, the EM algorithm does not provide standard errors as a by-product. Asymptotic standard errors could be computed using first and second order derivatives of the log-likelihood given by the log of equation (7). However, it should be emphasized that these information-based standard errors are only asymptotically valid and that, in particular for the case of mixture models, the sample size needs to be very large before good standard errors are provided through this method (McLachlan and Peel, 2004).

As an alternative, this study follows the advice of McLachlan and Peel (2004) and adopts a bootstrap procedure to estimate the variance of estimated coefficients. In particular, a nonparametric bootstrap is performed for each model, where random subsamples of 7,862 and 50,000 observations for the analysis of total spending and investment shares, respectively, are

[54] In particular, we employ R's 'optim' function using the BFGS algorithm.

drawn with replacement from the population of the data. Taking the maximum likelihood estimate of the respective model on the full dataset as a starting value as suggested by (McLachlan and Peel, 2004), the latent class models is then estimated on the bootstrap sample using the EM algorithm. The resulting coefficient estimate is denoted by $\hat{\boldsymbol{\theta}}_b$, where b denotes the bootstrap iteration.[55]

The bootstrap coefficient estimate is given by $\bar{\hat{\boldsymbol{\theta}}} = \sum_{b=1}^{B} \frac{\hat{\boldsymbol{\theta}}_b}{B}$, the average over the bootstrap iterations, with corresponding covariance matrix given by $cov(\hat{\boldsymbol{\theta}}) = \sum_{b=1}^{B} \frac{(\hat{\boldsymbol{\theta}}_b - \bar{\hat{\boldsymbol{\theta}}})(\hat{\boldsymbol{\theta}}_b - \bar{\hat{\boldsymbol{\theta}}})^T}{B-1}$. In this study, we calculate estimates for 100 bootstrap samples (i.e., $B = 100$). To allow for bootstrapping standard errors despite potential label switching of the mixture model, the component coefficient vectors are ordered by variance in each bootstrap iteration in line with suggestions by McLachlan and Peel (2004).

[55] The bootstrap algorithm lends itself to straightforward parallelization of the procedure on multi-core workstations to reduce computational time. We parallelize the code using the '*foreach*' package in R.

D Complete Estimation Results

D.1 Baseline Estimates

Table D.1: Complete Estimation Results (Table 2)

Couple Households	(1) Fixed effects log(Total spending)	(3) Fixed effects Risky asset share	(4) Pooled	(5) Fixed effects Safe asset share	(6) Pooled
log(HPU$_{12}$)	-0.039***	-0.037***	-0.105***	0.022***	0.044***
	(0.012)	(0.006)	(0.005)	(0.004)	(0.005)
VIX$_{12}$	0.005***	0.002***	0.002***	-0.002***	-0.002***
	(0.002)	(0.001)	(0.001)	(0.000)	(0.001)
SP500$_{12}$	2.031***	-0.236	-1.013***	-0.793***	-0.131
	(0.463)	(0.280)	(0.320)	(0.203)	(0.275)
CEI$_{12}$	18.209***	-0.056	16.727***	-0.214	-4.733*
	(5.038)	(2.900)	(3.071)	(2.096)	(2.633)
2$^{\text{nd}}$ Inc. Q.	-0.041	0.038	0.068***	-0.005	-0.112***
	(0.031)	(0.030)	(0.020)	(0.020)	(0.016)
3$^{\text{nd}}$ Inc. Q.	-0.007	0.062**	0.164***	-0.000	-0.176***
	(0.031)	(0.029)	(0.019)	(0.019)	(0.015)
4$^{\text{nd}}$ Inc. Q.	0.002	0.094***	0.236***	0.009	-0.235***
	(0.032)	(0.029)	(0.019)	(0.019)	(0.015)
2$^{\text{nd}}$ Wealth Q.	0.031	0.120***	0.136***	-0.059***	-0.235***
	(0.027)	(0.037)	(0.016)	(0.022)	(0.012)
3$^{\text{nd}}$ Wealth Q.	0.060**	0.369***	0.382***	-0.089***	-0.497***
	(0.030)	(0.037)	(0.016)	(0.023)	(0.011)
4$^{\text{nd}}$ Wealth Q.	0.092***	0.592***	0.703***	-0.098***	-0.635***
	(0.034)	(0.038)	(0.016)	(0.023)	(0.011)
Retired	-0.040**	-0.045***	-0.005	-0.062***	-0.112***
	(0.018)	(0.009)	(0.008)	(0.007)	(0.007)
Years ret.	-0.001	0.002***	0.002***	0.003***	0.003***
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
Age	-0.021***	-0.006***	0.001***	0.005***	0.003***
	(0.002)	(0.001)	(0.000)	(0.001)	(0.000)
Any children			0.016		-0.004
			(0.014)		(0.012)
GED			0.075***		-0.068***
			(0.026)		(0.020)
High school			0.215***		-0.267***
			(0.016)		(0.012)
Some College			0.265***		-0.321***
			(0.016)		(0.012)
Above college			0.346***		-0.472***
			(0.016)		(0.012)
White-Black			-0.083**		0.170***
			(0.039)		(0.032)
White-Other			-0.074***		0.153***
			(0.016)		(0.013)
Black-Black			-0.18***		0.259***
			(0.012)		(0.009)
Black-Other			-0.187***		0.309***
			(0.053)		(0.042)
Other-Other			-0.137***		0.199***
			(0.022)		(0.018)
Constant			-0.726***		0.307***
			(0.049)		(0.041)
Observations	7,862	54,717		54,717	

Continued.

Table D.1 (Contd.): Complete Estimation Results (Table 2)

Single Households	(7) Fixed effects	(9) Fixed effects	(10) Pooled	(11) Fixed effects	(12) Pooled
	log(Total spending)	Risky asset share		Safe asset share	
log(HPU_{12})	-0.017	-0.052***	-0.146***	0.040***	0.061***
	(0.011)	(0.010)	(0.008)	(0.007)	(0.007)
VIX_{12}	0.004**	0.000	0.002	-0.001**	-0.001
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
$SP500_{12}$	0.510	-1.078**	-1.018*	0.186	1.267***
	(0.457)	(0.454)	(0.521)	(0.336)	(0.443)
CEI_{12}	12.629***	1.734	9.707*	-4.027	-5.835
	(4.738)	(4.315)	(5.005)	(3.255)	(4.245)
2^{nd} Inc. Q.	0.063***	0.051***	0.179***	-0.015	-0.194***
	(0.017)	(0.017)	(0.013)	(0.012)	(0.010)
3^{nd} Inc. Q.	0.091***	0.105***	0.267***	-0.022*	-0.263***
	(0.020)	(0.018)	(0.014)	(0.013)	(0.012)
4^{nd} Inc. Q.	0.121***	0.086***	0.288***	-0.006	-0.305***
	(0.026)	(0.020)	(0.017)	(0.014)	(0.014)
2^{nd} Wealth Q.	0.062***	0.282***	0.339***	-0.055**	-0.324***
	(0.019)	(0.034)	(0.017)	(0.022)	(0.012)
3^{nd} Wealth Q.	0.092***	0.485***	0.623***	-0.091***	-0.606***
	(0.024)	(0.035)	(0.017)	(0.023)	(0.012)
4^{nd} Wealth Q.	0.133***	0.670***	0.958***	-0.117***	-0.786***
	(0.030)	(0.036)	(0.018)	(0.025)	(0.013)
Retired	-0.052***	-0.071***	-0.026**	-0.049***	-0.121***
	(0.019)	(0.017)	(0.012)	(0.013)	(0.010)
Years ret.	0.001	0.003***	0.002***	0.002***	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)
Age	-0.025***	-0.005***	0.004***	0.007***	0.005***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.000)
Any children			-0.043***		0.033***
			(0.012)		(0.010)
GED			0.078***		-0.082***
			(0.029)		(0.023)
High school			0.147***		-0.252***
			(0.015)		(0.012)
Some College			0.248***		-0.340***
			(0.015)		(0.013)
Above college			0.321***		-0.431***
			(0.016)		(0.014)
Black			-0.271***		0.310***
			(0.015)		(0.012)
Other			-0.129***		0.153***
			(0.026)		(0.021)
Female			0.003		-0.018**
			(0.010)		(0.009)
Constant			-0.884***		0.307***
			(0.062)		(0.052)
Observations	12,116	42,786		42,786	

*, ** and *** denote significance at 10%, 5%, 1% , respectively. Asymptotic standard errors are presented in parentheses. Following Rosen and Wu (2004), safe assets are checking and savings accounts, CDs, government savings bonds and T-bills, and risky assets are stocks and mutual funds. HPU and VIX denote Baker et al.'s (2016) health care policy uncertainty index and the Implied Volatility Index, respectively, matched to the final interview month of the HRS-respondent. SP500 and CEI denote growth rates of the S&P 500 Index and the Conference Board's Coincident Economic Indicator. Subscripts denote 12-month averages. Time-invariant household variables are subsumed by the fixed effects.

D.2 Latent Class Model

Table D.2 : Complete Estimation Results (Table 3)

	(1) log(Total spending)		(2) Risky asset share		(3) Safe asset share	
Prior Probability						
Constant	0	-2.329	0	2.026***	0	-0.361***
	-	(2.409)	-	(0.076)	-	(0.037)
Self-reported health	0	0.128	0	0.195***	0	0.293***
	-	(0.295)	-	(0.031)	-	(0.013)
Severe conditions	0	0.067	0	-0.006	0	0.049***
	-	(0.201)	-	(0.028)	-	(0.014)
Mobility	0	0.081	0	-0.008	0	0.066***
	-	(0.113)	-	(0.017)	-	(0.008)
ADLs	0	0.166	0	0.004	0	0.028**
	-	(0.182)	-	(0.026)	-	(0.013)
Nights in hospital	0	-0.003	0	-0.002	0	-0.001
	-	(0.019)	-	(0.001)	-	(0.001)
Uninsured	0	1.336*	0	0.625***	0	0.838***
	-	(0.749)	-	(0.110)	-	(0.057)
Components	*s.1*	*s.2*	*s.1*	*s.2*	*s.1*	*s.2*
Constant	10.689***	10.099***	0.769***	-0.69***	-0.914	0.631***
	(1.021)	(0.493)	(0.054)	(0.042)	(0.012)	(0.052)
$\log(\text{HPU}_{12})$	-0.088	-0.141	0.000	-0.103***	-0.002	0.076***
	(0.111)	(0.101)	(0.006)	(0.005)	(0.001)	(0.006)
VIX_{12}	0.008	0.003	0.002**	0.002***	-0.001	-0.001*
	(0.017)	(0.014)	(0.001)	(0.001)	(0.000)	(0.000)
SP500_{12}	1.433	-2.748	0.185	-0.988***	-0.105	0.096
	(7.458)	(6.964)	(0.381)	(0.240)	(0.056)	(0.275)
CEI_{12}	20.434	10.697	1.332	14.728***	0.186	-4.667**
	(36.168)	(47.407)	(3.627)	(2.594)	(0.489)	(2.051)
2^{nd} Inc. Q.	-0.064	0.504**	-0.03*	0.064***	-0.002	-0.153***
	(0.324)	(0.214)	(0.023)	(0.019)	(0.004)	(0.019)
3^{nd} Inc. Q.	0.222	0.904***	-0.029*	0.163***	0.011	-0.238***
	(0.534)	(0.223)	(0.022)	(0.018)	(0.004)	(0.019)
4^{nd} Inc. Q.	0.383	1.272***	-0.05**	0.235***	0.019	-0.318***
	(0.653)	(0.255)	(0.022)	(0.018)	(0.004)	(0.019)
2^{nd} Wealth. Q.	0.173	-0.088	-0.01	0.127***	0.007	-0.254***
	(0.302)	(0.147)	(0.019)	(0.018)	(0.005)	(0.018)
3^{nd} Wealth. Q.	0.258	0.207	0.017	0.365***	0.01	-0.575***
	(0.312)	(0.184)	(0.018)	(0.018)	(0.004)	(0.018)
4^{nd} Wealth. Q.	0.375	0.351**	0.021*	0.68***	0.001	-0.767***
	(0.550)	(0.160)	(0.020)	(0.017)	(0.004)	(0.018)

Continued on next page.

Table D.2 (Contd.): Complete Estimation Results (Table 3)

Continued from previous page.						
	(1)		(2)		(3)	
	log(Total spending)		Risky asset share		Safe asset share	
Retired	-0.043	-0.176	-0.021**	0.023***	-0.006	-0.103***
	(0.279)	(0.158)	(0.008)	(0.007)	(0.001)	(0.009)
Years ret.	-0.006	0.001	0.001*	0.002***	0.000	0.002***
	(0.027)	(0.011)	(0.000)	(0.000)	(0.000)	(0.000)
Age	-0.005	0.004	0.000	0.000	0.000	0.001***
	(0.027)	(0.006)	(0.000)	(0.000)	(0.000)	(0.000)
Any children	-0.148	-0.096	0.038*	0.003	0.004	0.004
	(1.096)	(0.150)	(0.020)	(0.011)	(0.003)	(0.015)
GED	0.190	-0.051	-0.001	0.113***	-0.005	-0.049**
	(0.450)	(0.258)	(0.038)	(0.024)	(0.008)	(0.021)
High school	0.188	0.051	0.028*	0.220***	-0.005	-0.229***
	(0.479)	(0.276)	(0.019)	(0.015)	(0.005)	(0.016)
Some College	0.214	0.046	0.019	0.267***	-0.012	-0.274***
	(0.449)	(0.241)	(0.019)	(0.016)	(0.005)	(0.016)
Above college	0.487	0.127	0.022*	0.355***	-0.020	-0.413***
	(0.564)	(0.238)	(0.019)	(0.015)	(0.005)	(0.016)
White-Black	-0.161	-0.371	0.023	-0.077*	-0.012	0.2***
	(0.480)	(0.402)	(0.056)	(0.041)	(0.007)	(0.045)
White-Other	-0.135	-0.305**	0.004	-0.085***	0.001	0.159***
	(0.229)	(0.139)	(0.021)	(0.014)	(0.003)	(0.016)
Black-Black	-0.001	-0.049	0.021*	-0.206***	-0.009	0.280***
	(0.297)	(0.154)	(0.012)	(0.011)	(0.003)	(0.012)
Black-Other	-0.123	-0.253	0.047*	-0.264***	0.007	0.446***
	(0.431)	(0.346)	(0.032)	(0.055)	(0.015)	(0.059)
Other-Other	-0.141	-0.376	-0.035*	-0.126***	0.014	0.201***
	(0.618)	(0.341)	(0.035)	(0.016)	(0.005)	(0.023)
Component share	0.804	0.196	0.067	0.933	0.318	0.682
Log-likelihood	-6,203.503		-29,669.4		-24,643.61	
Observations	7,862		54,717		54,717	

Results in column (1) correspond to a latent class normal model. Results in column (2) and (3) correspond to a latent class Tobit model. *, ** and *** denote significance at 10%, 5%, 1% , respectively. Estimates are calculated using the EM algorithm illustrated in Appendix C. Bootstrapped standard errors are in parentheses. $\partial \log(HPU_{12})|_{\bar{X}_s}$ denotes the marginal effect at the mean observation of each component as identified using the posterior component means. Total spending is measured in 2010 Dollars. Following Rosen and Wu (2004), safe assets are checking and savings accounts, CDs, government savings bonds and T-bills, and risky assets are stocks and mutual funds.

D.3 Model-Based Recursive Partitioning

Table D.3: Complete Estimation Results (Figure 2a)

log(Total spending)	(1)	(2)	(3)	(4)
log(HPU_{12})	-0.045	-0.076***	-0.109***	-0.009
	(0.031)	(0.02)	(0.032)	(0.055)
VIX_{12}	0.001	0.013***	0.008	-0.008
	(0.005)	(0.003)	(0.005)	(0.009)
$SP500_{12}$	0.179	0.568	0.66	-3.514
	(1.155)	(0.764)	(1.391)	(2.526)
CEI_{12}	-10.903	37.513***	16.338	-6.488
	(14.239)	(9.067)	(15.205)	(25.556)
2^{nd} Inc. Q.	-0.066	0.055	0.236***	0.196**
	(0.094)	(0.054)	(0.056)	(0.079)
3^{nd} Inc. Q.	0.196**	0.209***	0.453***	0.4***
	(0.088)	(0.051)	(0.056)	(0.079)
4^{nd} Inc. Q.	0.373***	0.377***	0.71***	0.608***
	(0.085)	(0.052)	(0.058)	(0.086)
2^{nd} Wealth. Q.	0.049	0.149***	0.048	-0.066
	(0.075)	(0.038)	(0.047)	(0.064)
3^{nd} Wealth. Q.	0.173**	0.202***	0.21***	0.19**
	(0.072)	(0.037)	(0.046)	(0.077)
4^{nd} Wealth. Q.	0.26***	0.381***	0.261***	0.429***
	(0.072)	(0.038)	(0.051)	(0.087)
Retired	-0.012	-0.023	-0.065	-0.073
	(0.036)	(0.025)	(0.045)	(0.069)
Years ret.	-0.003	-0.003**	-0.004**	0.002
	(0.002)	(0.001)	(0.002)	(0.005)
Age	-0.004**	-0.01***	-0.004*	-0.005
	(0.002)	(0.001)	(0.002)	(0.005)
Any children	-0.069	0.033	0.069	0.012
	(0.061)	(0.044)	(0.071)	(0.142)
GED	0.443***	0.125**	0.095	-0.036
	(0.136)	(0.063)	(0.076)	(0.117)
High school	0.284***	0.083**	0.168***	0.078
	(0.088)	(0.042)	(0.05)	(0.077)
Some College	0.422***	0.195***	0.296***	-0.014
	(0.089)	(0.043)	(0.052)	(0.08)
Above college	0.586***	0.383***	0.412***	0.078
	(0.088)	(0.044)	(0.054)	(0.086)
White-Black	-0.204	-0.162	-0.048	0.119
	(0.168)	(0.132)	(0.128)	(0.211)
White-Other	-0.105*	-0.067	-0.111	-0.217***
	(0.055)	(0.055)	(0.07)	(0.073)
Black-Black	0.006	-0.127***	0.039	-0.013
	(0.052)	(0.032)	(0.044)	(0.072)
Black-Other	-0.158	-0.247*	-	0.046
	(0.237)	(0.129)	-	(0.225)
Other-Other	-0.162*	-0.091	-0.186**	-0.064
	(0.083)	(0.066)	(0.093)	(0.11)
Constant	10.506***	10.814***	10.316***	10.457***
	(0.234)	(0.137)	(0.199)	(0.395)
Observations	1,684	3,817	1,715	646

The column numbers (1-4) correspond to the node numbers in Figure 2a. *, ** and *** denote significance at 10%, 5%, 1% , respectively. Asymptotic standard errors are presented in parentheses. Following Rosen and Wu (2004), safe assets are checking and savings accounts, CDs, government savings bonds and T-bills, and risky assets are stocks and mutual funds. HPU and VIX denote Baker et al.'s (2016) health care policy uncertainty index and the Implied Volatility Index, respectively, matched to the final interview month of the HRS-respondent. SP500 and CEI denote growth rates of the S&P 500 Index and the Conference Board's Coincident Economic Indicator. Subscripts denote 12-month averages.

Table D.3 (Contd.): Complete Estimation Results (Figure 2b)

Risky asset share	(1)	(2)	(3)	(4)	(5)	(6)
log(HPU_{12})	-0.102***	-0.091***	-0.054***	-0.120***	-0.184***	-0.138***
	(0.008)	(0.010)	(0.020)	(0.015)	(0.052)	(0.013)
VIX_{12}	0.000	0.003***	0.002	0.003	0.003	0.007***
	(0.000)	(0.001)	(0.002)	(0.001)	(0.005)	(0.001)
$SP500_{12}$	-1.303***	-1.003*	0.847	-3.115***	-2.294	0.984
	(0.498)	(0.603)	(1.185)	(1.061)	(2.853)	(0.861)
CEI_{12}	14.749***	22.191***	3.479	26.022***	44.210	-0.228
	(4.791)	(5.864)	(11.365)	(9.929)	(28.123)	(8.276)
2^{nd} Inc. Q.	0.025	0.134***	0.057	-0.103*	0.071	0.109***
	(0.039)	(0.046)	(0.077)	(0.060)	(0.123)	(0.041)
3^{nd} Inc. Q.	0.091**	0.235***	0.199***	-0.002	0.138	0.196***
	(0.036)	(0.044)	(0.074)	(0.057)	(0.117)	(0.040)
4^{nd} Inc. Q.	0.145***	0.299***	0.229***	0.065	0.295**	0.313***
	(0.036)	(0.043)	(0.075)	(0.058)	(0.120)	(0.041)
2^{nd} Wealth Q.	0.135***	0.129***	0.130**	0.150***	0.207**	0.114***
	(0.035)	(0.033)	(0.055)	(0.051)	(0.096)	(0.033)
3^{nd} Wealth Q.	0.343***	0.341***	0.357***	0.402***	0.455***	0.434***
	(0.033)	(0.031)	(0.053)	(0.049)	(0.098)	(0.032)
4^{nd} Wealth Q.	0.653***	0.639***	0.701***	0.757***	0.806***	0.764***
	(0.033)	(0.031)	(0.054)	(0.050)	(0.104)	(0.034)
Retired	0.019	0.019	0.022	-0.032	0.020	-0.072***
	(0.011)	(0.014)	(0.028)	(0.027)	(0.072)	(0.021)
Years ret.	0.003***	0.001	0.002	0.003***	-0.006	0.003***
	(0.000)	(0.000)	(0.001)	(0.001)	(0.006)	(0.000)
Age	-0.001	0.001	0.001	0.002	-0.005	0.004***
	(0.000)	(0.000)	(0.001)	(0.001)	(0.005)	(0.000)
Any children	0.034	0.008	-0.008	0.046	-0.203	0.032
	(0.021)	(0.027)	(0.056)	(0.043)	(0.130)	(0.035)
GED	-0.009	0.109**	0.040	0.003	-0.536*	0.154***
	(0.062)	(0.049)	(0.080)	(0.078)	(0.299)	(0.053)
High school	0.131***	0.209***	0.209***	0.163***	0.211**	0.268***
	(0.038)	(0.032)	(0.055)	(0.043)	(0.103)	(0.033)
Some College	0.160***	0.228***	0.238***	0.247***	0.341***	0.370***
	(0.038)	(0.032)	(0.056)	(0.043)	(0.104)	(0.033)
Above college	0.225***	0.303***	0.368***	0.304***	0.550***	0.456***
	(0.038)	(0.032)	(0.057)	(0.043)	(0.108)	(0.034)
White-Black	-0.081	-0.115*	0.008	0.127	0.227	-0.402***
	(0.070)	(0.069)	(0.146)	(0.112)	(0.189)	(0.133)
White-Other	-0.016	-0.051*	-0.129*	-0.177***	-0.181*	-0.100**
	(0.025)	(0.029)	(0.068)	(0.059)	(0.102)	(0.043)
Black-Black	-0.198***	-0.179***	-0.159***	-0.210***	-0.167*	-0.150***
	(0.023)	(0.021)	(0.040)	(0.037)	(0.085)	(0.028)
Black-Other	-0.110	-0.173	0.080	-0.373	-0.328	-0.411
	(0.097)	(0.082)	(0.189)	(0.233)	(0.274)	(0.203)
Other-Other	-0.101	-0.090	-0.192	-0.233	-0.214	-0.210
	(0.038)	(0.034)	(0.076)	(0.085)	(0.129)	(0.070)
Constant	-0.293	-0.785	-0.952	-0.577	-0.311	-1.115
	(0.089)	(0.103)	(0.190)	(0.146)	(0.478)	(0.113)
Observations	15,049	14,227	4,838	6,580	1,971	12,052

The column numbers (1-6) correspond to the node numbers in Figure 2b.

Table D.3: Complete Estimation Results (Figure 2c)

Safe asset share	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\log(\text{HPU}_{12})$	0.032***	0.038***	0.049***	0.153***	0.035***	0.090***	-0.003
	(0.007)	(0.008)	(0.016)	(0.037)	(0.009)	(0.032)	(0.089)
VIX_{12}	-0.001	-0.003***	-0.002	-0.012***	-0.002*	-0.001	-0.008
	(0.000)	(0.000)	(0.001)	(0.004)	(0.001)	(0.003)	(0.010)
SP500_{12}	-0.089	0.382	-0.380	-1.169	-0.505	-3.027*	1.514
	(0.430)	(0.506)	(0.956)	(2.240)	(0.574)	(1.789)	(5.549)
CEI_{12}	-4.182	-5.744	-4.096	0.124	-1.301	-11.368	13.973
	(4.133)	(4.846)	(9.221)	(21.940)	(5.466)	(17.658)	(54.491)
2^{nd} Inc. Q.	-0.011	-0.095***	-0.093*	-0.450***	-0.115***	-0.158**	-0.053
	(0.033)	(0.035)	(0.051)	(0.139)	(0.028)	(0.068)	(0.156)
3^{nd} Inc. Q.	-0.033	-0.148***	-0.121**	-0.536***	-0.172***	-0.230***	-0.285*
	(0.031)	(0.034)	(0.049)	(0.137)	(0.027)	(0.065)	(0.154)
4^{nd} Inc. Q.	-0.056*	-0.196***	-0.164***	-0.603***	-0.239***	-0.354***	-0.154
	(0.030)	(0.034)	(0.050)	(0.139)	(0.028)	(0.067)	(0.178)
2^{nd} Wealth Q.	-0.142***	-0.171***	-0.316***	-0.227***	-0.277***	-0.222***	-0.230*
	(0.026)	(0.024)	(0.039)	(0.083)	(0.021)	(0.053)	(0.130)
3^{nd} Wealth Q.	-0.354***	-0.375***	-0.577***	-0.395***	-0.575***	-0.502***	0.671***
	(0.025)	(0.023)	(0.038)	(0.081)	(0.020)	(0.055)	(0.143)
4^{nd} Wealth Q.	-0.466***	-0.486***	-0.729***	-0.599***	-0.739***	-0.665***	-1.037***
	(0.025)	(0.023)	(0.039)	(0.084)	(0.021)	(0.059)	(0.185)
Retired	-0.130***	-0.110***	-0.128***	-0.148***	-0.094***	-0.158***	-0.081
	(0.010)	(0.012)	(0.023)	(0.053)	(0.014)	(0.045)	(0.127)
Years ret.	0.003***	0.003***	0.003***	0.003	0.002***	0.008**	0.019*
	(0.000)	(0.000)	(0.001)	(0.002)	(0.000)	(0.004)	(0.009)
Age	0.003***	0.002***	0.003**	-0.002	0.004***	0.006**	0.017**
	(0.000)	(0.000)	(0.001)	(0.002)	(0.000)	(0.003)	(0.008)
Any children	-0.019	-0.003	0.018	-0.014	-0.031	0.129	0.577***
	(0.018)	(0.022)	(0.042)	(0.084)	(0.024)	(0.086)	(0.221)
GED	0.018	-0.042	-0.178***	-0.011	-0.099***	0.077	3.966
	(0.052)	(0.041)	(0.057)	(0.140)	(0.033)	(0.105)	(87.494)
High school	-0.177***	-0.207***	-0.317***	-0.208***	-0.293***	-0.202***	-0.132
	(0.032)	(0.025)	(0.035)	(0.079)	(0.020)	(0.060)	(0.151)
Some College	-0.201***	-0.241***	-0.380***	-0.348***	-0.357***	-0.334***	-0.243
	(0.032)	(0.025)	(0.036)	(0.081)	(0.021)	(0.060)	(0.148)
Above college	-0.330***	-0.388***	-0.484***	-0.354***	-0.514***	-0.560***	-0.470***
	(0.032)	(0.025)	(0.036)	(0.083)	(0.022)	(0.063)	(0.170)
White-Black	0.050	0.234***	0.117	-0.031	0.175**	0.130	0.827**
	(0.061)	(0.056)	(0.112)	(0.181)	(0.070)	(0.141)	(0.407)
White-Other	0.116***	0.150***	0.143***	-0.005	0.143***	0.192***	0.194
	(0.022)	(0.025)	(0.047)	(0.096)	(0.029)	(0.060)	(0.166)
Black-Black	0.314***	0.263***	0.227***	0.297***	0.254***	0.196***	-0.042
	(0.019)	(0.016)	(0.028)	(0.082)	(0.018)	(0.051)	(0.150)
Black-Other	0.276	0.330	0.402	2.628	0.109	0.621	0.941
	(0.081)	(0.068)	(0.151)	(73.553)	(0.092)	(0.215)	(0.506)
Other-Other	0.173	0.145	0.183	0.211	0.135	0.362	1.017
	(0.033)	(0.032)	(0.050)	(0.155)	(0.040)	(0.078)	(0.408)
Constant	-0.160	0.156	0.367	0.732	0.452	0.161	-0.453
	(0.076)	(0.081)	(0.141)	(0.312)	(0.077)	(0.288)	(0.756)
Observations	12,759	14,009	5,565	1,019	17,664	3,054	647

The column numbers (1-7) correspond to the node numbers in Figure 2c.