Effects of Health Care Policy Uncertainty on Households' Portfolio Choice*

Thomas T. Wiemann† and Robin L. Lumsdaine‡

Abstract

This paper develops a nonparametric identification approach for causal effects of an endogenous macroeconomic variable on microeconomic outcomes. The key assumption is the existence of an exogenous variable that shifts responsiveness to the variable of interest without shifting responsiveness to other macroeconomic time series. We apply the approach to study the effect of health care policy uncertainty (HCPU) on households' portfolio choice using health shocks to capture cross-sectional heterogeneity. Under the additional assumption of risk averse agents, our approach provides an informative bound on the average causal effect of HCPU. The empirical results highlight HCPU as an important determinant of households' financial behavior, and showcase substantial heterogeneity in HCPU effects across varying unexpected changes to health. *JEL* codes: D14, D80, I18, C21, C14.

Keywords: Causal Inference, Partial Identification, Heterogeneous Effects, Health and Retirement Study (HRS), Household Finance

† Department of Economics, University of Chicago

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

*We thank Susan Athey, Teresa Bago d'Uva, Stéphane Bonhomme, Steve Davis, Partha Deb, Jesús Fernández-Villaverde, Davide Furcer, Christian Hansen, Alex Hsu, Peter Hull, Guido Imbens, Alan Isaac, Jonas Lieber, Ashesh Rambachan, Sergio Salgado, Pedro Sant'Anna, and Alexander Torgovitsky for their comments, along with seminar participants at the University of Hawaii, the University of Houston, and the University of Michigan's Institute for Social Research, and participants at the following conferences: the Annual Health Econometrics Workshop (AHEW), Canadian Econometric Study Group (CESG), the International Association for Applied Econometrics (IAAE), the Royal Economic Society (RES), the Society for Financial Econometrics (SoFiE), and the Stanford Institute for Theoretical Economics (SITE) for their comments and suggestions. Wiemann acknowledges 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. 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. Any remaining errors are the responsibility of the authors alone. **Views and opinions expressed are those of the authors only and do not necessarily represent the views of any institutions with which they are affiliated.**

1. Introduction

More than a decade after the 2010 enactment of the Affordable Care Act, health care policy remains a central topic of political debate in the United States. In a comparison of 11 policy categories, Baker et al. (2016) find health care policy to be the second largest source of policy uncertainty in the US, behind only fiscal policy uncertainty. The recent COVID-19 pandemic has driven both economic and health care policy uncertainty to unprecedented levels, nearly double their pre-2020 levels, as countries first struggled to determine an appropriate response (Altig et al., 2020) and now grapple over how to effect an expeditious recovery. Yet the literature on the economic effect of uncertainty about health care policy on households' economic behavior is in its infancy.

A key challenge for identification of causal effects of policy uncertainty, including health care policy uncertainty, is its potential confoundedness with other macroeconomic time series. For example, Bloom (2014) notes that the (general) economic policy uncertainty index of Baker et al. (2016) is 51% higher in recessions. Further, a defining feature of macroeconomic variables such as uncertainty associated with federal policies is their invariance across the economy for any given point in time. As a result, construction of a suitable "control" based on standard alternative assumptions such as parallel trends is not immediately possible. To address these challenges, the empirical policy uncertainty literature frequently compares subpopulations that are expected to differ in their responsiveness to the specific policies but to react similarly to other forms of uncertainty, such as recession-caused economic uncertainty. For example, Giavazzi and McMahon (2012) compare private sector employees with civil servants in their analysis of Germany's 1998 national election, arguing the latter were not directly affected by the candidates' policy differences and hence less exposed to the increased policy uncertainty. Similarly, Baker et al. (2016) consider a regression specification that interacts policy uncertainty with a measure of revenue from government contracts to capture firms' heterogeneous exposure to uncertainty in government spending.

The first contribution of this paper is to formalize the idea that subgroup heterogeneity may be used to aid identification of causal effects of endogenous macroeconomic time series on microeconomic outcomes. We focus our analysis on two causal parameters. First, the expected difference in potential outcomes at two different levels of the macroeconomic variable. This *average causal effect* captures the level-impact of a hypothesized shift in the macroeconomic variable on individuals, holding everything else – including the distribution of unobservables – equal. Second, the *conditional average difference of causal effects* between two fixed subgroups. This parameter describes the observed heterogeneity in responses to a shift in the macroeconomic variable. The two parameters are complementary as they allow for insights into the overall and distributional effects of the macroeconomic variable of interest, respectively. Throughout the identification analysis, we allow for unobserved heterogeneity and make no parametric assumptions on the potential outcomes.

Our first identification result shows that the conditional average difference of causal effects is identified without (conditional) independence assumptions that are often directly placed on the macroeconomic variable of interest. Instead, the key identifying assumption is the existence of a sufficiently exogenous exposure variable that shifts responsiveness to the macroeconomic variable of interest but does not shift responsiveness to other macroeconomic time series. The average causal effect, however, is not generally identified under the same assumptions due to the absence of a "non-treated" group. Our second identification result then shows that a sharp bound on the average causal effect can be derived if sign restrictions on the causal effects across subgroups are available.

The intuition for the proposed approach is straightforward. Identification of the conditional average difference of causal effects is achieved by correlating the macroeconomic variable of interest with values of the coefficients that result from cross-sectional regressions of the outcome on the exposure variable in each time period. The assumptions we introduce permit a causal interpretation by ensuring that changes in the coefficient of the exposure variable across time can be attributed to changes in the macroeconomic variable of interest. By combining this result with a sign-restriction on the effects of macroeconomic variable across all levels of exposure, the bound on the average causal effect can then be constructed by integrating the conditional effects over the marginal distribution of the exposure variable.

Our approach is closely related to the recent literature on difference-in-difference designs, which historically had considered a single treated and control group with a binary treatment but recently have been extended to more general contexts. Particularly related are Callaway et al. (2021), who consider identification of causal effects with continuous treatment, and De Chaisemartin and d'Haultfoeuille (2018) who analyze settings where treatment increases by varying amounts for both groups. Our analysis may be viewed as a difference-in-difference design with continuous treatment in which treatment intensity differs across groups, but where intensity differences are not known *a priori*. We further allow group identities to lie along a continuum. Similarly to De Chaisemartin and d'Haultfoeuille (2018), absence of a non-treated group implies that the level-effect of treatment is generally unidentified, but we show how economic theory can be leveraged to obtain sharp bounds.

In addition, our identification analysis contributes to the existing literature on shift-share designs, where macroeconomic (common) shocks are interacted with a measure of individual-level exposure. Our setting is similar to that of Goldsmith-Pinkham et al. (2020), who consider exogenous exposure to endogenous shocks.[1] In contrast to shift-share approaches, we analyze exposure differences that are not (necessarily) motivated by accounting identities and allow for non-additively separable confounders. Our analysis may thus also be viewed as extending existing exposure designs to settings with unknown functional forms.

The second contribution of this paper is the analysis of the effects of health care policy uncertainty on households' financial decisions. Applying the developed identification design,

[1]In contrast, Adao et al. (2019) and Borusyak et al. (2022) consider endogenous exposure to exogenous shocks.

we use unexpected changes in households' health to capture heterogeneous exposure to health care policy uncertainty. This strategy exploits the idea that worsening health expectations may affect households' responsiveness to medical expenditure risk as induced by health care policy uncertainty. The key assumption is then that unexpected changes in health do not affect households' responses to other macroeconomic variables correlated with health care policy uncertainty (e.g., trade policy uncertainty). Bounds on the average causal effect of health care policy uncertainty on households' relative investment in risky versus safe assets are obtained by leveraging economic theory on background risk, which implies that risk averse households will (weakly) decrease their relative demand for risky assets when faced with an undiversifiable risk such as increased health care policy uncertainty (Pratt and Zeckhauser, 1987; Kimball, 1993; Gollier and Pratt, 1996).

Our empirical analysis uses as a starting point the category-specific health care policy uncertainty index developed by Baker et al. (2016) as the primary variable of interest. The results indicate economically substantive average causal effects of health care policy uncertainty on households' portfolio choice. When health care policy uncertainty increases by 70%, couple (single) households are estimated to increase their safe asset share by *at least* 3.5 (2.7) percentage points.[2] The same increase in health care policy uncertainty also results in an estimated average decrease in couple (single) households' risky asset share of at least 1.5 (2.3) percentage points. The results thus highlight the importance of health care policy uncertainty as a determinant of households' financial behavior.

Complementing the empirical results on the average causal effects of health care policy uncertainty, estimates of the average difference in causal effects across varying levels of health shock severity indicate heterogeneity (with respect to unexpected changes in health)

[2]These interpretations are made precise in Section 4. For illustrative purposes, throughout the paper we refer to an increase of 70% in health care policy uncertainty, corresponding to the increase in the Baker et al. (2016) health care policy uncertainty index from a 2016 average of 110 to the 2017 average of 191. The latter year is associated with extensive political efforts to repeal the Affordable Care Act.

in the effect of health care policy uncertainty on households' relative demand for safe and risky assets. Compared to a couple (single) household that finds itself with fewer severe health conditions than expected, a couple (single) household that experiences no unexpected change in severe conditions is estimated to increase its safe asset share by approximately 3.3 (3.1) percentage points, on average, when health care policy uncertainty increases by 70%. Correspondingly, couple (single) households reduce their share of relative risky assets by approximately 2.9 (2.2) percentage points. Differences in the effect of health care policy uncertainty of similar magnitude are also documented across households that differ by unexpected changes in the number of hospital stays or ability to conduct activities of daily living. In addition to documenting heterogeneity per se, our first stage estimates indicate non-monotonicity in how health shocks change households' responsiveness to health care policy uncertainty. In particular, households that are at their expected health level are estimated to react more strongly to increases in health care policy uncertainty than households that are in substantially better or worse health than expected.

The empirical results on the effects of health care policy uncertainty add to the existing literature on the effects of economic policy uncertainty. The policy uncertainty measure of Baker et al. (2016) that we rely on is a widely used to study the association of policy uncertainty with other macroeconomic variables (e.g., Stock and Watson, 2012) or its effects on firms (e.g., Pástor and Veronesi, 2013; Baker et al., 2016; Gulen and Ion, 2016). We contribute primarily to the burgeoning strand of literature that instead analyzes the effects of policy uncertainty on households, including Giavazzi and McMahon (2012), Luttmer and Samwick (2018), and the ongoing work of Agarwal et al. (2018) and Gábor-Tóth and Georgarakos (2019). To the best of our knowledge, ours is the first study with a focus on health care policy uncertainty, despite the prominence of health care in the US policy landscape.

The paper proceeds as follows: Section 2 defines the two parameters of interest and develops our main identification results in a general framework. Section 3 then relates the identification approach to the empirical context of health care policy uncertainty and describes construction of the empirical variables from data. Section 4 presents estimates of the causal effects of health care policy uncertainty on households' portfolio choice and relates the results to existing empirical research. Section 5 concludes. A series of supplemental appendices describing complementary identification results, data construction and implementation details, and robustness checks is available online.

2. Causal Effects of Macroeconomic Variables on Microeconomic Outcomes

This section defines the parameters of interest, discusses sufficient identifying assumptions, and states our theoretical results.[3] The final subsection briefly discusses estimation and inference.

2.1. Framework

We are interested in the effect of a macroeconomic variable (i.e., a variable that is invariant across individuals at any point in time, e.g., the level of health care policy uncertainty) on a microeconomic outcome (e.g., the share of risky assets an individual invests in). For all individuals $i \in \mathcal{N}$ and time periods $t \in \mathcal{T}$, consider the random variables $(Y_{i,t}, W_t, Z_{i,t}, U_{i,t})$ defined on a common probability space, with joint distribution characterized by the causal model

$$Y_{i,t} = g(W_t, Z_{i,t}, U_{i,t}) \tag{1}$$

where $Y_{i,t}$ denotes the outcome, W_t is the macroeconomic variable of interest, $Z_{i,t}$ is the exposure variable, and $U_{i,t}$ are all determinants of $Y_{i,t}$ other than $(W_t, Z_{i,t})$. Further, $g : \operatorname{supp} W_t \times \operatorname{supp} Z_{i,t} \times \operatorname{supp} U_{i,t} \to \operatorname{supp} Y_{i,t}$ is an unknown structural function, where supp

[3] Although many papers typically omit an explicit definition of the causal parameters discussed here in terms of potential outcomes, the empirical literature on policy uncertainty also aims to estimate average causal effects (see, for example, Giavazzi and McMahon (2012), Pástor and Veronesi (2013), Baker et al. (2016), and Gulen and Ion (2016)). Our new, formal, identification analysis thus may also contribute to a better understanding of empirical results from past and future studies on the microeconomic implications of policy uncertainty.

denotes the support of the corresponding random variable. We emphasize that the causal model in (1) is equivalent to conventional potential outcome notation with $Y_{i,t}(w, z) \equiv g(w, z, U_{i,t})$.[4] We choose to use the latent variable notation instead of potential outcome notation as it simplifies the statement and discussion of the assumptions below.[5]

The key feature of the causal model is that the macroeconomic variable W_t differs only by time but not across individuals. The model further allows the exposure variable $Z_{i,t}$ to differ across time, which fits best with the health shocks used in our empirical application. In other applied settings, however, exposure may be time-invariant. The identification analysis simplifies slightly with time-invariant exposure variables but the main results are unaffected.

The primary parameter of interest is the average causal effect (ACE) on the microeconomic outcome $Y_{i,t}$ of a hypothesized shift in the macroeconomic variable W_t from w to w'. Let $\Delta_{i,t}(w', w, z) \equiv g(w', z, U_{i,t}) - g(w, z, U_{i,t})$ – or equivalently in potential outcome notation $\Delta(w', w, z) \equiv Y_{i,t}(w', z) - Y_{i,t}(w, z)$ – denote the causal effect for the ith individual at time t with fixed exposure value z. The ACE is then defined by

$$\text{ACE}(w', w) \equiv E[\Delta_{i,t}(w', w; Z_{i,t})], \tag{2}$$

where the expectation is with respect to both the observed exposure as well as the unobservables. In cross-sectional settings with a binary treatment, the parameter defined in (2) is frequently referred to as the "average treatment effect." We use the slightly more general terminology "average causal effect" as a macroeconomic variable is rarely thought of as a "treatment."

[4] As $(W_t, Z_{i,t}, U_{i,t})$, determine $Y_{i,t}$ deterministically, this causal model is also referred to as the "all causes" model (see, e.g., Heckman and Vytlacil, 2007).

[5] While the identification analysis of the main text places no parametric assumptions on the potential outcomes, helpful intuition regarding the identifying assumptions and the resulting arguments can be gained from considering a simpler linear causal model with unobserved heterogeneity. Online appendix S.2 presents the identification analysis in this simpler setting.

The second parameter of interest is the conditional average difference of causal effects (DCE) between two fixed subgroups with exposure values at z and z':

$$\text{DCE}(w', w, z', z) \equiv E[\Delta_{i,t}(w', w, z') - \Delta_{i,t}(w', w, z)|Z_{i,t} = z']. \tag{3}$$

In contrast to the ACE which describes the overall expected causal effect, the DCE captures relative differences in causal effects between subgroups at different exposure levels. It thus allows the highlighting of groups of individuals that are affected more (or less) severely than others. Importantly, however, the DCE does not identify levels of effects. That is, the magnitude or sign of the DCE between two subgroups does not allow for conlusions about the magnitude or sign of causal effects on either group separately.

There are two key challenges for identification of the ACE and DCE. First, because the macroeconomic variable W_t is invariant across individuals for any point in time, for each period t it is only ever possible to observe either $W_t = w'$ or $W_t = w$ but never both. Assumptions that allow for extrapolation across time are thus needed for identification. Second, because the expectations in (2) and (3) are with respect to both the exposure variable and the unobservables, they cannot be identified from the data alone. Additional assumptions restricting the joint distribution of observables and unobservables are thus needed.

The approach developed in this paper proposes a new set of identifying assumptions to address the abovementioned challenges. Assumption 1 assumes stationarity of the conditional distribution of unobservables $U_{i,t}$. Stationarity is frequently maintained in the analysis of time series and serves here as an extrapolation device, allowing us to overcome the first identification challenge – that only a single value of the macroeconomic variable W_t is realized at any point in time.[6]

[6]Despite stationarity being a common assumption in the analysis of time series, it is an economically substantive restriction on the data generating process. For example, while Assumption 1 is sufficiently weak to allow for trends in both observable and unobservable variables, structural breaks in the dependence between the macroeconomic variable of interest and other unobserved time series would violate Assumption 1. The supplemental online appendix S.3 presents identification results without the stationarity assumption,

Assumption 1 (Stationarity)

For two time periods $t, t' \in \mathcal{T}$, $U_{i,t'}|W_{t'}, Z_{i,t'} \overset{d}{=} U_{i,t}|W_t, Z_{i,t}$.

Assumption 2 ensures that the exposure variable is sufficiently exogenous. In particular, it requires that for any fixed subgroup, changes in the expected outcome due to changes in the macroeconomic variable of interest are mean-independent of the exposure variable. It thus restricts the conditional distribution of unobservables given the macroeconomic variable to be sufficiently independent of exposure. Note that this is a parallel trends assumption where $Z_{i,t}$ takes the role of treatment, and where instead of restricting trends for only the untreated potential outcomes, trends for *all* subgroups are independent of $Z_{i,t}$.[7] Importantly, Assumption 2 does not restrict the dependence between levels of the outcome and the exposure variable (i.e., there is no exclusion restriction).[8]

Assumption 2 (Parallel Changes)

$\forall w', w \in \operatorname{supp} W_t, z \in \operatorname{supp} Z_{i,t}, E\left[g(w', z, U_{i,t})|W_t = w', Z_{i,t}\right] - E\left[g(w, z, U_{i,t})|W_t = w, Z_{i,t}\right]$ *does not depend on* $Z_{i,t}$.

Our key identifying assumption is Assumption 3. It restricts the unobserved heterogeneity in the effects of the exposure variable on the outcome. In particular, it implies that there should be no dependence between the unobserved heterogeneity in the effects of the exposure variable on the outcome and the macroeconomic variable of interest. To gain intuition for why this assumption is crucial, suppose that the exposure variable affects individual

highlighting that causal inference remains possible but that interpretation must take into account non-stationary unobservable distributions. In particular, focusing on the setting with time-invariant exposure variables, we analyze identification of the causal parameter $E[g(w', Z_i, U_{i,t'}) - g(w, Z_i, U_{i,t})]$ which can be decomposed into the ACE at time t plus the change in the distribution of unobservables from time t to time t': $E[g(w, Z_i, U_{i,t'}) - g(w, Z_i, U_{i,t})]$.

[7]To see the parallel trends analogy more clearly, let $Y^w_{i,t}(z) \overset{d}{=} g(w, z, U_{i,t})|W_t = w$. Then Assumption 2 can equivalently be expressed as $E[Y^{w'}_{i,t}(z) - Y^w_{i,t}(z)|Z_{i,t}] = E[Y^{w'}_{i,t}(z) - Y^w_{i,t}(z)]$.

[8]Assumption 2 is immediately implied when exposure is randomly assigned, i.e., when $Z_{i,t} \perp\!\!\!\perp (U_{i,t}, W_t)$.

responses to two correlated macroeconomic time series, say health care policy uncertainty and general economic uncertainty, of which only the former is observed (i.e., W_t) and the latter is unobserved (i.e., in $U_{i,t}$). Differences in individuals' outcomes with varying exposure levels can then not be uniquely attributed to changes in either of the two macroeconomic variables.

Assumption 3 (Exogenous Differences in the Effects of Exposure)
$\forall w \in \operatorname{supp} W_t, z', z \in \operatorname{supp} Z_{i,t},\ E\left[g(w, z', U_{i,t}) - g(w, z, U_{i,t}) | W_t, Z_{i,t} = z\right]$ *does not depend on* W_t.

Assumption 3 does not restrict all unobserved heterogeneity in the effects of $Z_{i,t}$. Unobserved heterogeneity that is independent of W_t (e.g., idiosyncratic household preferences independent of the current state of the macroeconomy) is allowed for. Importantly, Assumption 3 does not restrict the unobserved heterogeinty in the effects of the macroeconomic variable. To further illustrate Assumption 3 and the types of unobserved heterogeneity it restricts, it is useful to consider examples of functional forms for the structural function g that satisfy or violate the assumption.[9] For ease of discussion, let $U_{i,t} \equiv (U_{i,t}^{(1)}, U_{i,t}^{(2)})$, where the unobservables $U_{i,t}^{(1)}$ are left unrestricted and and the unobservables $U_{i,t}^{(2)}$ are idiosyncratic shocks (i.e., $U_{i,t}^{(2)} \perp\!\!\!\perp (W_t, Z_{i,t})$). Consider first

$$g_1(W_t, Z_{i,t}, U_{i,t}) = \tilde{g}_1(W_t, U_{i,t}) + \tilde{g}_2(W_t, Z_{i,t}, U_{i,t}^{(1)}),$$

where $\tilde{g}_1$ and $\tilde{g}_2$ are arbitrary functions. It is easy to confirm that g_1 satisfies Assumption 3. In contrast, $g_1(W_t, Z_{i,t}, U_{i,t}) + Z_{i,t}U_{i,t}^{(1)}$ does not generally satisfy Assumption 3 due to the final interaction term between the exposure variable and components of $U_{i,t}$ that are not (necessarily) independent of the macroeconomic variable W_t.

Note that both Assumption 2 and Assumption 3 are not invariant to transformations of the outcome. For example, if the assumption holds for the untransformed g, it does not

[9] Online appendix S.2 which considers identification in a linear causal model provides further discussion on the nature of the restriction of unobserved heterogeneity.

generally hold for the composite function $\log \circ g$. This caveat is analogous to the sensitivity of parallel trends assumptions in conventional difference-in-difference designs (see, for example, Roth and Sant'Anna, 2020).

Finally, we assume common support to ensure all the conditional expectation functions are well-defined.

Assumption 4 (Common Support)

$f_{WZ}(w, z) > 0$, $\forall w \in \operatorname{supp} W_t, z \in \operatorname{supp} Z_{i,t}$, where f_{WZ} denotes the joint density.

2.2. *Identification*

Our first identification result is Theorem 1. Here and throughout, we understand identification to mean expressing an unknown parameter as a known function of the distribution of observables only (i.e., identification in the sense of Hurwicz, 1950). Theorem 1 shows that the DCE equals differences in conditional expectation functions of $Y_{i,t}$ given $(W_t, Z_{i,t})$ and therefore implies point identification.

Theorem 1 (Point Identification of the DCE)

Suppose that Assumptions 1 to 4 hold. Then, $\forall w', w \in \operatorname{supp} W_t, z', z \in \operatorname{supp} Z_{i,t}$,

$$\begin{aligned} DCE(w', w, z', z) &= (E[Y_{i,t'}|W_{t'} = w', Z_{i,t'} = z'] - E[Y_{i,t}|W_t = w, Z_{i,t} = z']) \\ &\quad - (E[Y_{i,t'}|W_{t'} = w', Z_{i,t'} = z] - E[Y_{i,t}|W_t = w, Z_{i,t} = z]). \end{aligned}$$

Proof. Provided in Appendix S.1.

Theorem 1 further indicates that the conditional average causal effects are *not* generally point identified by differences in conditional expectation functions. This non-identification of levels is due to the absence of a non-treated group whose effect of treatment would be known to be zero. In particular, as without a non-treated group $E\left[\Delta i, t(w', w, z)|Z_{i,t} = z'\right]$ may take any value on the real line, so may the conditional average causal effect for the subgroup with $Z_{i,t} = z'$. This loss of identification in levels is akin to the loss of identification of the average treatment effect on the treated in fuzzy difference-in-difference designs when treatment proportions increase in all subgroups (De Chaisemartin and d'Haultfoeuille, 2018).[10]

Non-identification of levels from Theorem 1 rests on the presumption that subgroup-specific effects of the macroeconomic variable on the microeconomic outcome may take *any* value. However, such conservativeness is not always necessary. Our main result, stated in Theorem 2, below combines sign restrictions on subgroup-specific causal effects with Assumptions 1 to 4 to show that a sharp bound on the primary parameter of interest – the average causal effect of the macroeconomic variable as defined in Equation (2) – can be identified. We focus on restrictions on the expected difference in potential outcomes $E[\Delta_{i,t}(w', w, z)]$ for all subgroups $z \in \operatorname{supp} Z_{i,t}$.[11]

Theorem 2 (Partial Identification of the ACE)

Suppose that Assumptions 1-4 hold. Take $w' \geq w \in \operatorname{supp} W_t$.

- *If in addition $E[\Delta_{i,t}(w', w, z)] \geq 0, \forall z \in \operatorname{supp} Z_{i,t}$, then*

$$ACE(w', w) \geq \max_{z \in \operatorname{supp} Z_{i,t}} E\Big[DCE(w', w, Z_{i,t}, z)\Big].$$

- *If instead, in addition $E[\Delta_{i,t}(w', w, z)] \leq 0, \forall z \in \operatorname{supp} Z_{i,t}$, then*

$$ACE(w', w) \leq \min_{z \in \operatorname{supp} Z_{i,t}} E\Big[DCE(w', w, Z_{i,t}, z)\Big].$$

Proof. Provided in Appendix S.1.

Restrictions for subgroup-specific average causal effects as required for the bounds in Theorem 2 are frequently plausible in macroeconomic settings. As briefly mentioned in the introduction, for example, Baker et al. (2016) consider the effect of economic policy uncertainty on firm investment, which may be expected to be weakly negative regardless

[10] De Chaisemartin and d'Haultfoeuille (2018) provide bounds for their primary parameter of interest, the "local" average treatment effect, based on support restrictions of the outcome variable. In contrast, the bounds provided in Theorem 2 are based on sign restrictions of subgroup-specific average causal effects.

[11] Alternatively, we may instead restrict the effect only for a subset of subgroups $\mathcal{Z}$. Theorem 2 can then be adapted by replacing $\operatorname{supp} Z_{i,t}$ with $\mathcal{Z}$ in the maximization/minimization in the formulation of the bounds.

of firms' revenue from government contracts. More formally, we draw from literature on background risk to motivate weakly negative effects of health care policy uncertainty on households' relative investment in risky assets in Section 3. In the supplemental online appendix S.9, we further consider the effect of health care policy uncertainty on households' consumption expenditures. Bounds on the subgroup-specific causal effect in the consumption setting are motivated by the literature on precautionary savings (e.g., Zeldes, 1989; Kimball, 1990).[12]

Note that if $E[\Delta_{i,t}(w', w, z)] \geq 0, \forall z \in \operatorname{supp} Z_{i,t}$, Theorem 2 and the definition of the DCE also imply that

$$\mathrm{ACE}(w', w) \geq \max_{z \in \operatorname{supp} Z_{i,t}} E[\Delta_{i,t}(w', w, Z_{i,t}) - \Delta_{i,t}(w', w, z)] \geq 0.$$

Therefore, the sign restriction along with Assumptions 1-4 immediately imply a weakly positive ACE. (The argument for weakly negative subgroup effects is analogous.) This direct bound does not leverage any information in the data, however, and would thus not allow for deeper quantitative insights regarding the magnitude of the macroeconomic variable. In many cases, therefore, it may be only marginally preferable to the most trivial of bounds (i.e., a bound from negative to positive infinity). The difference between the bound provided in Theorem 2 to this zero-valued bound depends crucially on the extent to which $Z_{i,t}$ affects exposure to changes in the macroeconomic variable. If $Z_{i,t}$ does not affect exposure, our bound does not improve on the zero-valued ACE bound. On the other hand, the greater the heterogeneity in responses induced by the exposure variable, the more informative the bound on the average causal effect provided above. For practical usefulness, the choice of

[12]In some settings, more precise information about (some) subgroup-specific effects is available. These may allow for point identification of the ACE under the conditions of Theorem 2. For example, Giavazzi and McMahon (2012) argue in their analysis of Germany's 1998 national election on households' savings that civil servants were not directly affected by the political candidates' policy differences, motivating a non-treated group with known zero-valued effect. Lemma 1 in Appendix S.1 then provides point identification of the corresponding ACE.

a relevant exposure variable is therefore important. We highlight, however, that unlike in weak instrument settings, there are no theoretical concerns about irrelevant $Z_{i,t}$. Existence of *strong* exposure variables is therefore not an assumption we require for identification or estimation.

2.3. Estimation

Estimation of the DCE for a shift of the macroeconomic variable from w to w' and subgroups with exposure values z and z' as identified in Theorem 1 involves estimation of four conditional expectations: $E[Y_{i,t'}|W_{t'} = w', Z_{i,t'} = z']$, $E[Y_{i,t}|W_t = w, Z_{i,t} = z'])$, $E[Y_{i,t'}|W_{t'} = w', Z_{i,t'} = z]$, and $E[Y_{i,t}|W_t = w, Z_{i,t} = z]$. In the absence of functional form assumptions on the structural function g, these conditional expectations can be estimated via nonparametric estimators such as series regression or kernels. Standard asymptotic theory for these nonparametric estimators then conveniently applies to estimators of the DCE.

A point estimate for the bound of the ACE can be constructed by sample averages of DCE estimates over the empirical marginal distribution of the exposure variable and then choosing the maximum (or minimum) value over possible values of the fixed exposure level z. Statistical inference for this estimator is challenging, however, due to the maximization (or minimization) operator. Developing asymptotic theory for estimators of the sharp bound of Theorem 2 lies outside the scope of this paper. We emphasize, however, that simple standard error estimators are available when researchers pre-select a value of the exposure variable, say $\tilde{z}$, which is used in place of the $\arg\max$ (or $\arg\min$) when constructing the bound. In that case, the bound is again simply a difference in conditional expectation estimators, and the same readily-available standard error methods as for the DCE apply. This simplification of statistical inference comes at a loss of sharpness (i.e., the bounds are more conservative than they need to be), but this loss is expected to be minor if researchers have a reasonable guess about which value of the exposure variable $\tilde{z}$ is close to the $\arg\max$ (or $\arg\min$) of the subgroup-specific shift in outcomes – i.e., for the $\arg\max$, a value $\tilde{z} \in \operatorname{supp} Z_{i,t}$ such that

$$E[\Delta_{i,t}(w', w, \tilde{z})] \approx \max_{z \in \operatorname{supp} Z_{i,t}} E[\Delta_{i,t}(w', w, z)].$$

We adopt this approach to statistical inference for the bound on the ACE in our empirical application about the effect of health care policy uncertainty, where a value for $\tilde{z}$ naturally arises from those households with the least severe health shocks.

3. Empirical Setting and Data

This section relates the identification approach developed in the previous section to our empirical analysis of health care policy uncertainty and outlines construction of the empirical variables from data. The corresponding data citations are provided in the supplemental online appendices S.4.

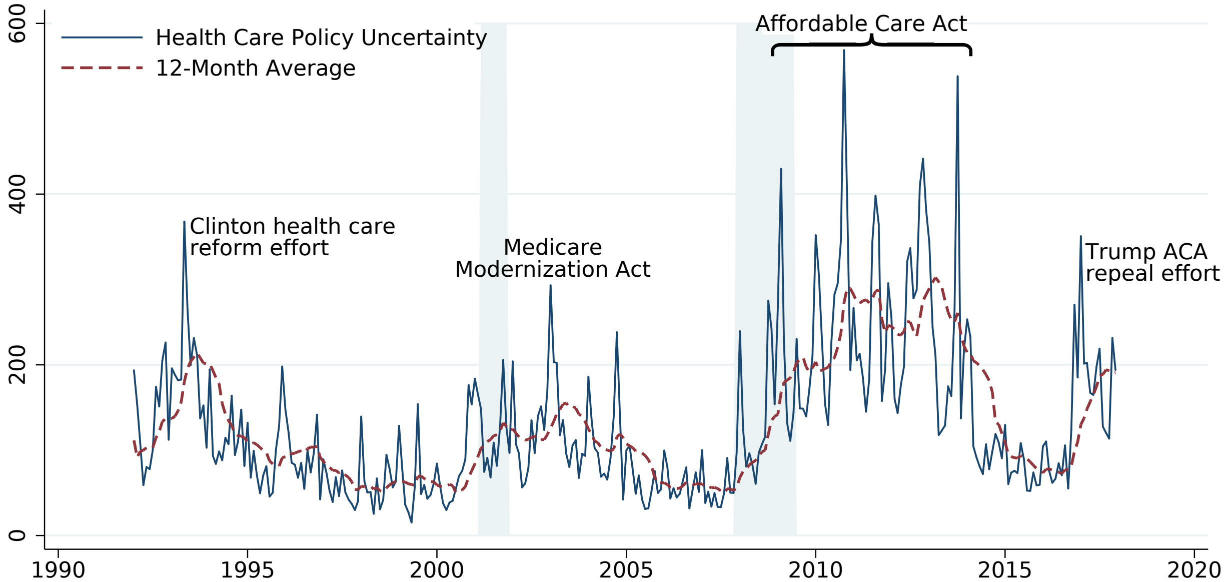
3.1. Health Care Policy Uncertainty

The macroeconomic variable of interest (i.e., W_t) is health care policy uncertainty and we are interested in how this variable affects individuals' portfolio allocations. Our empirical analysis is based on the computer driven, news-based health care policy uncertainty index developed by Baker et al. (2016), equal to the proportion of articles in the Access World News newspaper archive with the word triplets "uncertain", "economic", and "policy" (and their synonyms) as well as at least one term related to health care – e.g., "Medicaid," "health insurance," or "Obamacare."[13] The health care policy uncertainty index is one of several categorical indices developed by Baker et al. (2016) along with their (general) economic policy uncertainty index. These news-based economic policy uncertainty indices have been used in a vast amount of policy uncertainty literature, for example to examine the effect of policy uncertainty on firms (e.g., Pástor and Veronesi, 2013; Baker et al., 2016; Gulen and Ion, 2016) and on other macroeconomic time series (e.g., Stock and Watson, 2012).

[13] 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."

Figure 1 shows the health care policy uncertainty index between 1992 and 2017, and its average over the preceding 12 months. The average is characterized by four substantive increases that can be linked to health care policy efforts in the US.

FIGURE 1
Health Care Policy Uncertainty Index 1992 – 2017



Notes. The Baker et al. (2016) index plotted in this figure 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. The shaded areas represent NBER recession periods. 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, adding also the series' 12-month average and NBER recession bars.

3.2. Portfolio Choice

The empirical analysis aims to assess the effect of health care policy uncertainty on households' portfolio choice. The outcome variables (i.e., $Y_{i,t}$) we focus on are the shares of risky assets (stocks and mutual funds) and safe assets (checking and savings accounts, CDs, government savings bonds and T-bills) as a proportion of total financial wealth.

We construct risky and safe asset share outcome variables from the asset data of the Health and Retirement Study (HRS), a biennial comprehensive, nationally representative longitudinal panel of noninstitutionalized US residents over age 50.[14] 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 portfolio choice (e.g., Rosen and Wu, 2004; Goldman and Maestas, 2013). 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. Due to differences in definitions of important variables, we omit the first wave of the HRS and use the remaining eleven waves (1994-2014).[15] We merge the household data with the 12-month average of the health care policy uncertainty index preceding the end date of the month of each household's HRS interview.

3.3. Health Shocks

Empirical analysis of the effect of health care policy uncertainty on households' portfolio choice poses difficult identification challenges – highlighted in the introduction – that we address with the identification approach developed in the preceding section. The exposure variable (i.e., $Z_{i,t}$) to health care policy uncertainty that we consider is unexpected changes to health (or "health shocks"). Health shocks are an attractive choice for the exposure variable in our empirical application as existing literature suggests meaningful heterogeneity in households' responses to health care policy uncertainty motivated by two key mechanisms.

[14]The two other potential asset categories – bonds and retirement accounts – are not analyzed because the IRA account information in the HRS does not allow for sufficiently detailed risk-classification and only a small fraction of financial wealth is held in bonds. This paper's separate consideration of risky asset investments outside of retirement accounts is in line with existing literature, which further 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, see, for example, Rosen and Wu (2004) and Love and Smith (2010). Further, shares of financial assets are calculated as a proportion of total financial wealth rather than all assets, as non-liquid wealth (e.g., housing wealth) is not readily adjustable (Goldman and Maestas, 2013).

[15]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."

First, because uncertainty in health care policy implies higher medical expenditure risk under worsening health, houeseholds may decrease their exposure to rate of return risk in response to a health shock.[16] Second, unexpected worsening health may lower households' expected lifespan and consumption utility (Smith, 1999), mitigating their responsiveness to any disposable income risk. Because the practical usefuleness of a bound on the average causal effect depends crucially on the relevance of the exposure variable, as highlighted in Section 2, existence of both mechanisms is a promising starting point for the empirical analysis.

The assumptions on the health shocks that suffice for identifying causal effects of health care policy uncertainty on households' relative investment in risky assets are those given in Section 2. In particular, for the DCE, we require first that for any fixed value of the health shocks, changes in the expected outcome are mean-independent of the health shocks (Assumption 2). Changes in the expected outcome for households with a severe health shock *had they not had one* are assumed to be the same as for households with a non-severe health shock. Conversely, changes in the expected outcome for households that did not have a severe health shock *had they had one* are assumed to be the same as for households that actually experienced a severe health shock. This assumption is violated if health shocks are correlated with unobserved household characteristics that can also affect changes in the expected outcome. To address these concerns, we construct health shocks controlling for key household demographics – discussed in detail below. Second, we require that health shocks do not affect households' responses to unobserved variables that are correlated with health care policy uncertainty (Assumption 3). Importantly, the assumption imposes that health shocks do not change how households respond to economic uncertainty (e.g., risk of recession). This restriction is a key concern in the US, where households' health insurance and hence

[16] Empirical evidence for such a mechanism is provided by Atella et al. (2012), who find that health shocks have a negative effect on demand for risky assets in European countries without universal health care but find no evidence of an effect in other European countries.

exposure to medical expenditure risk is often linked to employment. Our analysis of older households aims to address this issue: With the majority of the sample in retirement, most households we analyze are not subject to risk of greater exposure to medical expenditures due to potential job loss. Further, as illustrated by Smith (1999), labor supply will not be substantially altered by health shocks during retirement, and income from Social Security and pensions will remain fixed (i.e., not related to health).

In addition to the assumptions used for the DCE, we require a sign restriction on subgroup-specific causal effects as in Theorem 2 for the bound of the ACE. In the health care policy uncertainty context, we motivate this sign restriction with economic theory on background risk, which formalizes households' portfolio choice in an environment of multiple risks. Pratt and Zeckhauser (1987), Kimball (1993), and Gollier and Pratt (1996) provide definitions of risk aversion which imply that households that are exposed to an undiversifiable risk are less willing to bear other types of risk, including rate-of-return risk. Health care policy uncertainty is such an undiversifiable risk because insufficient health care coverage can magnify the large out-of-pocket medical expenditures that frequently accompany health shocks in the United States. Under the assumption that households in our sample are risk averse regardless of the potential health shock (i.e., that health shocks never make households risk-loving), this implies, in particular, that the expected causal effects of an increase in health care policy uncertainty on relative demand for risky assets are bounded above by zero – i.e., $E[\Delta_{i,t}(w', w, z)] \leq 0, \forall z \in \operatorname{supp} Z_{i,t}$ – and thus allows for the application of Theorem 2. An equivalent argument implies that the expected causal effects of health care policy uncertainty on the relative demand for safe assets are bounded below by zero.

Construction of the health shocks from data requires careful discussion because health is an intrinsically unobserved variable. Fortunately, a variety of proxies have been suggested in the literature. 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?" Self-reported health measures such as this 5-point Likert scale variable are frequently used as proxies for health (e.g., Rosen and Wu, 2004), yet, some doubts remain about the role of potential reporting bias (e.g., Lumsdaine and Exterkate, 2013). As an alternative, some studies have employed more objective measures of health. For example, Wu (2003) and Berkowitz and Qiu (2006) construct 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.[17] Other frequently-used measures are the proxies given by self-reported limitations in Activities of Daily Living (ADLs) and Instrumental Activities of Daily 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, Fan and Zhao (2009) consider a separate variable to capture a respondents' mobility, summing indicators that are equal to one if the respondent has difficulty (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.

Lack of consensus as to which proxy is preferable in the literature motivates a flexible approach. We construct quasi-random health shocks under the assumption that households' health expectations are reasonably approximated by a linear first-order Markov process. This approach is a relaxation of the first-difference construction of health shocks as in Wu (2003) and Berkowitz and Qiu (2006) as it allows for simultaneous consideration of multiple health proxies as well as heterogeneous health expectations conditional on observed household characteristics. In addition, it addresses the need to condition on household demographics related to health. Using a VAR that simultaneously considers all health variables,

[17]'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.

we estimate $Z_{i,t} - AZ_{it-1} = BX_{i,t} + \varphi_{i,t}$, where $Z_{i,t}$ is a 5×1 vector of integers capturing (a) a household's self-reported health measure, (b) the number of limitations reported in the HRS' ADL and IADL measures, (c) the cumulative (from the beginning of the sample to time t) number of severe health conditions, (d) the current number of limitations to mobility, and (e) the number of nights spent in a hospital during the previous two years, and A and B are respectively 5×5 and $5 \times k$ matrices of fixed and unknown coefficients, where k is the number of included household characteristics. The 5×1 vector of residuals, $\varphi_{i,t}$ is then interpreted as an approximation to an unexpected change in the various components of a household's health.

We follow earlier literature and analyze single and couple households separately (e.g., Wu, 2003; Rosen and Wu, 2004; Berkowitz and Qiu, 2006; Love and Smith, 2010), and for couple households define the health measures introduced above as the maximum value across spouses (e.g., Coile and Milligan, 2009; Love and Smith, 2010). Included as controls for household characteristics are a household's age, dummy variables for five educational attainment categories, race categories (described below), and wave-specific wealth and income quartiles. Gender is also included when analyzing single households. For couple households, controls for highest education level and age are constructed as the maximum of both spouses, and six race dummy variables are included, based on all possible combinations of the three categories "White," "Black", and "Other" across the two spouses. The purpose of the inclusion of these controls is to address primary concerns regarding conditional heterogeneity in changes in health. The constructed residuals are orthogonal to the included controls and thus account for possible mechanisms such as a decrease in health with age and associations between health developments and socio-demographic characteristics.

The health data are merged with the HCPU and financial variables. We then remove from the sample those households that have no positive holdings of financial assets, are missing one or more of the above variables, or only occur in a single wave, leaving an analysis sample of 45,384 single and 57,190 couple household-wave observations. The supplemental online appendix S.5 provides detailed documentation of the cleaning steps and the associated reduction in sample at each stage, as well as complementary summary statistics of household characteristics.

4. Results

This section presents estimates of the DCE and estimates of the ACE bounds. Although the identification analysis of these parameters is nonparametric, we focus on estimation of a varying coefficient model (see, e.g., Hastie and Tibshirani, 1993). This choice has several important advantages for our particular application.[18] On one hand, nonparametric estimation of varying coefficients allows for rich and highly non-linear heterogeneity in how health shocks may change the effect of health care policy uncertainty on households' portfolio choice. This flexibility is particularly important for construction of informative bounds; masking this heterogeneity weakly shrinks the bound estimates to zero.[19] On the other hand, it facilitates comparisons with existing empirical literature on effects of policy uncertainty that primarily relies on linear regression estimates. This is due to holding the effects of health care policy unctertainty constant *for a given health shock.* In particular, we consider

$$Y_{i,t} = \beta_0(Z_{i,t}) + \beta_1(Z_{i,t})W_t + U_{i,t}, \tag{4}$$

where β_0 and β_1 are unknown functions which we estimate nonparametrically.

In addition to flexible estimation of β_1, a choice of the exposure variable $Z_{i,t}$ that captures more of the heterogeneity of households' response to health care policy uncertainty will result

[18]Varying coefficient models have recently been popularized in economics by Athey et al. (2019) and Farrell et al. (2021).

[19]Supplemental online appendix S.8 further illustrates the advantages of a flexible varying coefficient approach by contrasting the results with a simple multiplicative-interaction specification. The multiplicative-interaction specification substantially restricts the heterogeneity in the effects of health care policy uncertainty, which masks differences in the subgroup-specific causal effects and thus mitigates the practical usefulness of the constructed bounds.

in a more informative bound on the average causal effect. This observation motivates using the complete set of five distinct health categories (introduced in the preceeding section) when estimating the ACE bound. Unfortunately, visualization of the DCE conditional on five variables is difficult. For illustrative purposes regarding the relative distribution of the causal effect of health care policy uncertainty, we therefore also present results conditioning on a single health category (resulting in a single-dimensional DCE visualized by a line-plot) and two health categories (resulting in a two-dimensional DCE visualized by a heatmap).

When conditioning on a single health category, we estimate β_0 and β_1 using local linear regression. As traditional conventional semiparametric estimators suffer severely under the curse of dimensionality when conditioning on multiple variables, we apply the generalized random forests of Athey et al. (2019) when conditioning on more than one health category. The procedure allows for flexible interaction between types of health shocks; such flexible interaction is *a priori* likely to be a key feature of the DCE, while alleviating the curse of dimensionality through the inherent regularization of random forests.[20] Supplemental online appendix S.6 provides pseudocode and further details on the implementation of our estimation and inference procedures.

4.1. Estimates of the Conditional Average Difference in Causal Effects

Figure 2 presents estimates of the conditional average difference in causal effects based on unexpected changes to severe conditions (top panels) and on unexpected changes to nights spent in hospital (bottom panels).[21] The figures are normalized such that all estimates are relative to the average causal effect of households that are substantially healthier than expected, which we define as having a more favorable unexpected change in the respective

[20]We use the R package of Tibshirani et al. (2020) for random forest-based estimation.

[21]In the interest of space, estimates for only two of the health categories are presented here. Supplemental online appendix S.7 provides analogous figures for the other health categories we consider.

health category than 95% of households in the sample.[22] Note that because larger values for the health categories indicate worse health, negative values correspond to more favorable health. The horizontal axes are normalized to standard deviation units of the corresponding health category for ease of interpretation. The vertical axis shows percentage point differences in the dependent variable given a 100% increase in health care policy uncertainty – i.e., the figure plots the estimated values of the coefficients themselves (solid lines), denoted by $\beta_1(Z_{i,t})$ in the varying coefficient model (4). For illustrative purposes, however, it is more interesting to consider the approximately 70% increase in the index, from the 2016 average of 110 to the 2017 average of 191 (see also Figure 1). The latter year is associated with the extensive political efforts to repeal the Affordable Care Act. The dashed lines indicate bootstrapped 99% uniform confidence bands.[23]

The estimates indicate substantial heterogeneity in the effect of health care policy uncertainty with respect to unexpected changes in households' severe conditions and nights spent in hospital. Consider a single household with no unexpected changes to severe health conditions. Compared to a single household that finds itself with substantially fewer severe health conditions than expected, the first household increases its safe asset share by approximately 3.1 percentage points more, on average, when health care policy uncertainty increases by 70%. For couple households, the analogous increase in relative demand for safe assets amounts to 3.3 percentage points.[24] Similarly, when health care policy uncertainty

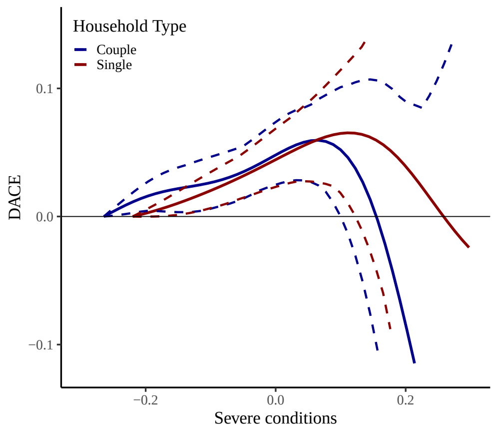
[22]This normalization is akin to the choice of baseline in a linear regression that includes a constant and indicators of categorical variables with one (baseline) category suppressed.

[23]We implement the uniform confidence bands discussed in Montiel Olea and Plagborg-Møller (2019) using a grid of 100 evenly spaced samples from the horizontal axis as an approximation to the infinite-dimensional DCE. The bootstrap samples households rather than individual observations. This block bootstrap procedure thus accounts for dependence within households over time.

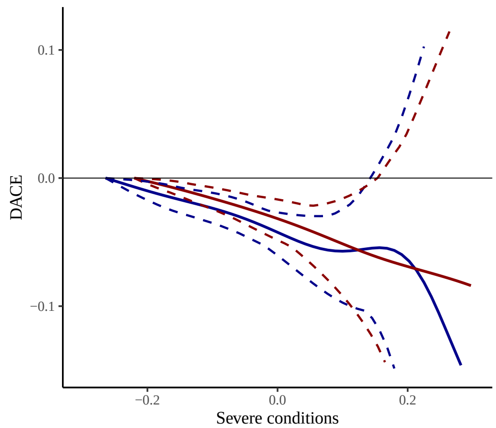
[24]The values are calculated by multiplying the height of the solid DCE curves in the top-left panel of Figure 2 at the origin of the horizontal axis (i.e., when $z = 0$) by 0.7.

Figure 2

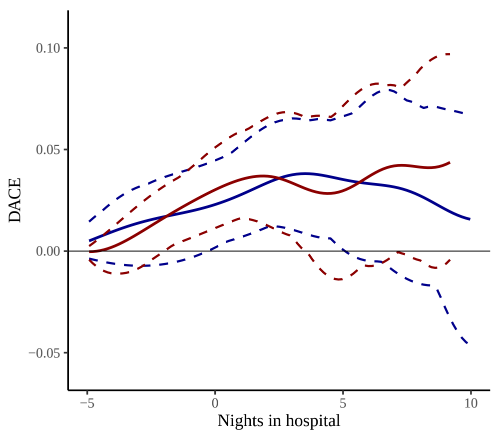
Normalized DCE Estimates for Couple and Single Households



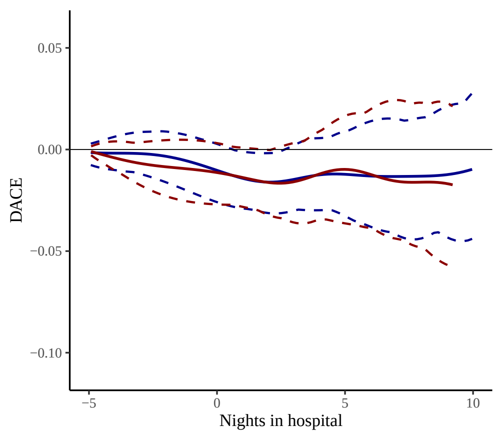
(A) Safe asset share



(B) Risky asset share



(C) Safe asset share



(D) Risky asset share

Notes. These figures show the local linear regression-based estimates of the average difference in causal effects (DCE) of health care policy uncertainty on safe and risky asset share, respectively, conditional on a single health category, for two specific examples: severe conditions (figures A and B) and nights in hospital (figures C and D). The estimates presented are relative to the baseline of households that are in substantially better health than expected, defined here as having a more favorable unexpected change in the corresponding health category than 95% of households in the sample. The plots are based on $B = 1,000$ bootstrap draws, where the solid lines correspond to the mean and the dashed lines correspond to bootstrapped 99% confidence bands. Blue lines correspond to couple households and red lines correspond to single households.

increases by 70%, a single (couple) household with no unexpected changes to severe conditions decreases its risky asset share by approximately 2.2 (2.9) percentage points more than an analogous household with substantially fewer severe health conditions. That the increase in safe asset share is only partially offset by the decrease in risky asset share suggests that households increase their safe asset share not only by decreasing relative investment in stocks, but also by decreasing relative investment in the two excluded financial asset categories – bonds and IRA retirement accounts.

When considering instead unexpected changes in nights spent in hospital, the effects are a little less pronounced. A single (couple) household with no unexpected changes to nights spent in hospital increases its safe asset share by approximately 2.1 (1.6) percentage points when health care policy uncertainty increases by 70%, relative to a household with substantially fewer-than-expected nights in hospital. The analogous effects on relative demand for risky assets are 0.8 and 0.7 percentage points for single and couple households, respectively.

In addition to documenting heterogeneity in the effect of health care policy uncertainty with respect to unexpected changes in health, Figure 2 also suggests nonlinearities. In particular, households that are at their expected health level are estimated to react more strongly to increases in health care policy uncertainty than households that are in either substantially better or worse health than expected. These estimates are in line with economic theory on opposing mechanisms on the effect of health due to increased disposable income uncertainty on one hand and potential loss in consumption utility and/or reduction in life expectancy on the other (see, e.g., Smith, 1999). Interpreting the results in this fashion indicates that the risk induced by higher expected health care costs drives the differences between households that are in better health than expected and those that are at their expected health. Differences between households that are in worse health than expected and households that are at their expected health, on the other hand, may be attenuated by the reduction in life expectancy and utility. Note that the uniform confidence bands in

Figure 2 cannot reject linearity of the normalized DCE on a 99% confidence level.[25] The point estimates are therefore only suggestive of a nonlinear tradeoff between the two health mechanisms.

The DCE results so far condition on a single health category, effectively integrating over all possible interaction effects that may capture additional heterogeneity in households' response to health care policy uncertainty. In an effort to analyze interactions in two dimensions, at least, we compute DCE estimates conditional on both unexpected changes in severe conditions and nights in hospital. As before, we turn our discussion to comparisons of households who have no unexpected changes to health with households who are substantially healthier than expected, which we now define as having a more favorable unexpected change in each of the two health categories than 95% of households in the sample. Figures S.7.2 and S.7.3 in the Appendix S.7 provide visualizations of the estimates and confidence bands in the form of heatmaps for both safe and risky asset share, respectively.

Focusing first on the results for relative demand in safe assets, we find that single (couple) households who have no unexpected changes in both severe conditions and nights in hospital increase their safe asset share by 3.9 (5.3) percentage points when faced with a 70% increase in health care policy uncertainty, compared to households who find themselves with substantially fewer severe conditions and substantially fewer nights in hospital. Similarly, the results for relative demand in risky assets suggest that single (couple) households who have no unexpected changes in both severe conditions and nights in hospital decrease their

[25]The reported uniform confidence bands encompass the set of functional forms of the normalized DCE that cannot be rejected on a 99% confidence level. Whether this set contains a linear functional form can be conveniently checked by assessing whether it is possible to draw a straight line segment starting at the reference point (in our example, the left-most point in the figure) that stays everywhere inside of the confidence bands without intersecting the bands. Curved confidence bands that are sufficiently tight will reject linearity, but wide bands cannot regardless of curvature. For further details on inference with uniform confidence bands see, for example, Montiel Olea and Plagborg-Møller (2019).

risky asset share by 6.0 (4.9) percentage points when faced with a 70% increase in health care policy uncertainty, when compared to households with substantially fewer severe conditions and substantially fewer nights in hospital than expected.

The results indicate the relevance of interactions across different health categories for explaining heterogeneity in the effect of health care policy uncertainty on households' portfolio choice in that the magnitude of the heterogeneous effect is much larger when these interactions are considered. This motivates using the full set of five health categories for estimation of the DCE as a first stage parameter for the construction of a bound for the average causal effect of health care policy uncertainty, to which we turn next.

4.2. Bounding the Average Causal Effect

Table 1 presents estimates of the bounds of the average causal effects of health care policy uncertainty on the relative demand for risky and safe assets, for couple and single households, respectively. The estimates are based on bounding the subgroup-expecific average causal effect by zero for households that are substantially better off than expected, which we again define here as having a more favorable unexpected change in each of the five health categories than 95% of households in the sample. Note that this approach results in more conservative estimates than bounding *all* households' subgroup-specific average causal effects by zero, independent of their health outcomes, which is justified by the assumption of risk averse agents. Brackets provide bootstrapped one-sided 95% confidence intervals.

Focusing first on the safe asset share results in columns (1) and (2), the estimates suggest that couple and single households increase their safe asset share by *at least* 3.5 $(= 0.051 * 70)$ and 2.7$(= 0.039 * 70)$ percentage points, respectively, when faced with a 70% increase in health care policy uncertainty. Further, columns (3) and (4) indicate that couple and single households decrease their risky asset share by *at least* 1.5 and 2.3 percentage points, respectively, when faced with a 70% increase in health care policy uncertainty. With the exception of the bound for couple households on relative demand for risky assets, the estimates are

Table 1
Bounds on the Average Causal Effect

Safe Asset Share		Risky Asset Share	
Couples	Singles	Couples	Singles
(1)	(2)	(3)	(4)
0.051	0.039	-0.022	-0.033
[0.015, ∞)	[0.012, ∞)	(-∞, 0.005]	(-∞, -0.013]

Notes. The table presents estimates of the bound on the average causal effect (ACE) of health care policy uncertainty. Brackets contain one-sided confidence intervals covering of 95% of bootstrapped bounds

significant on a 5% confidence level. Taken together, these results suggest that an increase in health care policy uncertainty shifts households' portfolio choice towards safe assets in the population, and – at least for single households – away from risky assets.

This inference comports with results from earlier studies. Also analyzing HRS data, Rosen and Wu (2004), Edwards (2008), and Love and Smith (2010), for example, respectively find that rating health in the worst category of the subjective health measure is associated with a 1%, 7%, and 1.8% decrease in the risky asset share. Our results thus show that a 70% increase in health care policy uncertainty, as has occurred between 2016 and 2017, shifts the relative demand for risky assets by at least as much as a considerable reduction in households' health, highlighting the importance of health care policy uncertainty as a determinant of households' relative demand for risky assets.

5. Conclusion

In this paper, we develop a new causal identification approach in the context of investigating a macroeconomic variable and its effect on microeconomic outcomes. The key challenge in this setting is that the endogenous macroeconomic variable varies only by time but not cross-sectionally which renders conventional difference-in-difference designs inapplicable. Instead, we consider identification using an exogenous variable that shifts responsiveness to the macroeconomic variable of interest without shifting responsiveness to other macroeconomic time series. We provide sufficient conditions for point identification of the average difference in causal effects between population subgroups and show how sign restrictions derived from economic theory allow for partial identification of the average causal effect. We then apply the proposed methodology to study the effect of health care policy uncertainty on households' portfolio choice. Our empirical results suggest substantial nonlinear heterogeneity in the causal effect of health care policy uncertainty on households' safe and risky asset share with respect to unexpected changes in health. The estimates further indicate that health care policy uncertainty has statistically and economically significant effects on households' financial behavior in the population of older US households.

The empirical analysis shows that an uncertainty increase similar to that associated with efforts to repeal the Affordable Care Act in 2017 decreases the relative demand for stocks and mutual funds by as much as a considerable reduction in health (e.g., Rosen and Wu, 2004). Given the large share of financial assets that older American couples in the HRS data have, the reduction in stock market participation may have direct implications for stock market volatility and the equity premium. Further, health care policy uncertainty appears to disproportionally affect households with unexpected adverse changes to health, which may exacerbate the socio-economic disadvantage associated with bad health in the US (Smith, 1999).

The recent COVID-19 pandemic has resulted in an unprecedented increase in economic and health care policy uncertainty (Altig et al., 2020). As countries begin to emerge from the COVID-19 pandemic, policymakers have sought to develop policies aimed at stimulating the economy, even as they continue to introduce measures to curtail the virus. These dual policy objectives also create a great deal of uncertainty for households. As health care policy likely remains at the center of political debates in the United States and elsewhere for the foreseeable future, further research into the macroeconomic implications of these policy discussions is necessary to assess potentially unintended consequences of political discourse. Our paper provides a flexible identification and estimation approach for analyzing such implications.

References

Adao, Rodrigo, Kolesár, Michal, and Morales, Eduardo (2019). Shift-share designs: Theory and inference. *Quarterly Journal of Economics*, 134(4):1949–2010.

Agarwal, Vikas, Aslan, Hadiye, Huang, Lixin, and Ren, Honglin (2018). Policy uncertainty and household stock market participation. SSRN working paper No 3068792.

Altig, Dave, Baker, Scott, Barrero, Jose Maria, Bloom, Nicholas, Bunn, Philip, Chen, Scarlet, Davis, Steven J, Leather, Julia, Meyer, Brent, and Mihaylov, Emil (2020). Economic uncertainty before and during the covid-19 pandemic. *Journal of Public Economics*, 191:104274.

Atella, Vincenzo, Brunetti, Marianna, and Maestas, Nicole (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, Susan, Tibshirani, Julie, and Wager, Stefan (2019). Generalized random forests. *The Annals of Statistics*, 47(2):1148–1178.

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

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

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

Borusyak, Kirill, Hull, Peter, and Jaravel, Xavier (2022). Quasi-experimental shift-share research designs. *Review of Economic Studies*, 89(1):181–213.

Callaway, Brantly, Goodman-Bacon, Andrew, and Sant'Anna, Pedro HC (2021). Difference-in-differences with a continuous treatment. *Working paper arXiv:2107.02637*.

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

De Chaisemartin, Clément and d'Haultfoeuille, Xavier (2018). Fuzzy differences-in-differences. *Review of Economic Studies*, 85(2):999–1028.

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

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

Farrell, Max H, Liang, Tengyuan, and Misra, Sanjog (2021). Deep learning for individual heterogeneity: An automatic inference framework. Working paper.

Gábor-Tóth, Enikő and Georgarakos, Dimitris (2019). Economic policy uncertainty and stock market participation. Center for Financial Studies Working Paper No 590.

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

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

Goldsmith-Pinkham, Paul, Sorkin, Isaac, and Swift, Henry (2020). Bartik instruments: What, when, why, and how. *American Economic Review*, 110(8):2586–2624.

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

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

Gulen, Huseyin and Ion, Mihai (2016). Policy uncertainty and corporate investment. *The Review of Financial Studies*, 29(3):523–564.

Hastie, Trevor and Tibshirani, Robert (1993). Varying-coefficient models. *Journal of the Royal Statistical Society: Series B*, 55(4):757–779.

Heckman, James J and Vytlacil, Edward J (2007). Econometric evaluation of social programs, part I: Causal models, structural models and econometric policy evaluation. In Heckman, James J and Leamer, E, editors, *Handbook of Econometrics*, volume 6, chapter 70, pages 4779–4874. Elsevier, Amsterdam.

Hurwicz, Leonid (1950). Generalization of the concept of identification. In Koopmans, Tjalling C, editor, *Statistical inference in dynamic economic models*, volume 6, chapter 4, pages 245–257. Wiley, New York.

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

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

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

Lumsdaine, Robin L and Exterkate, Anneke (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.

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

Montiel Olea, José Luis and Plagborg-Møller, Mikkel (2019). Simultaneous confidence bands: Theory, implementation, and an application to SVARs. *Journal of Applied Econometrics*, 34(1):1–17.

Pástor, L'uboš and Veronesi, Pietro (2013). Political uncertainty and risk premia. *Journal of Financial Economics*, 110(3):520–545.

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

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

Roth, Jonathan and Sant'Anna, Pedro HC (2020). When is parallel trends sensitive to functional form? *Working paper arXiv:2010.04814*.

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

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

Tibshirani, J, Athey, S, Friedberg, R, Hadad, V, Hirshberg, D, Miner, L, Sverdrup, E, Wager, S, and Wright, M (2020). `grf`: Generalized random forests R package.

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

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

Effects of Health Care Policy Uncertainty on Households' Portfolio Choice

Supplemental Appendices (for online publication only)

Thomas T. Wiemann and Robin L. Lumsdaine

Table of Contents

S.1. Proofs

This appendix provides the proofs of Theorem 1 and Theorem 2.

A.1. Proof of Theorem 1

Using standard probability calculus, we have

$$
\begin{aligned}
&(E[Y_{i,t'}|W_{t'} = w', Z_{i,t'} = z'] - E[Y_{i,t}|W_t = w, Z_{i,t} = z'])\\
&\quad - (E[Y_{i,t'}|W_{t'} = w', Z_{i,t'} = z] - E[Y_{i,t}|W_t = w, Z_{i,t} = z])\\
&\overset{[1]}{=}(E[g(w', z', U_{i,t'})|W_{t'} = w', Z_{i,t'} = z'] - E[g(w, z', U_{i,t})|W_t = w, Z_{i,t} = z'])\\
&\quad - (E[g(w', z, U_{i,t'})|W_{t'} = w', Z_{i,t'} = z] - E[g(w, z, U_{i,t})|W_t = w, Z_{i,t} = z])\\
&\overset{[2]}{=}(E[g(w', z', U_{i,t'})|W_{t'} = w', Z_{i,t'} = z'] - E[g(w', z, U_{i,t'})|W_{t'} = w', Z_{i,t'} = z])\\
&\quad - (E[g(w, z', U_{i,t})|W_t = w, Z_{i,t} = z'] - E[g(w, z, U_{i,t})|W_t = w, Z_{i,t} = z])\\
&\overset{[3]}{=}(E[g(w', z', U_{i,t})|W_t = w', Z_{i,t} = z'] - E[g(w', z, U_{i,t})|W_t = w', Z_{i,t} = z])\\
&\quad - (E[g(w, z', U_{i,t})|W_t = w, Z_{i,t} = z'] - E[g(w, z, U_{i,t})|W_t = w, Z_{i,t} = z])\\
&\overset{[4]}{=}(E[g(w', z', U_{i,t})|W_t = w', Z_{i,t} = z'] - E[g(w', z, U_{i,t})|W_t = w', Z_{i,t} = z'])\\
&\quad + (E[g(w', z, U_{i,t})|W_t = w', Z_{i,t} = z'] - E[g(w', z, U_{i,t})|W_t = w', Z_{i,t} = z])\\
&\quad - (E[g(w, z', U_{i,t})|W_t = w, Z_{i,t} = z'] - E[g(w, z, U_{i,t})|W_t = w, Z_{i,t} = z'])\\
&\quad - (E[g(w, z, U_{i,t})|W_t = w, Z_{i,t} = z'] - E[g(w, z, U_{i,t})|W_t = w, Z_{i,t} = z])\\
&\overset{[5]}{=}(E[g(w', z', U_{i,t}) - g(w', z, U_{i,t})|W_t = w', Z_t = z'])\\
&\quad - E[g(w, z', U_{i,t}) - g(w, z, U_{i,t})|W_t = w, Z_{i,t} = z'])\\
&\overset{[6]}{=}(E[g(w', z', U_{i,t}) - g(w', z, U_{i,t})|Z_{i,t} = z'] - E[g(w, z', U_{i,t}) - g(w, z, U_{i,t})|Z_{i,t} = z']
\end{aligned}
$$

where [1] substitutes for $Y_{i,t} = g(W_t, Z_{i,t}, U_{i,t})$, [2] rearranges terms, [3] follows from Assumption 1, [4] both adds and subtracts the two terms $E[g(w', z, U_{i,t})|W_t = w', Z_{i,t} = z']$ and $E[g(w, z, U_{i,t})|W_t = w, Z_{i,t} = z']$, [5] follows from Assumption 2, and [6] follows from

Assumption 3. Finally, by rearranging terms and using the definition of the DCE in (3)

$$DCE(w', w, z', z) = E[g(w', z', U_{i,t}) - g(w', z, U_{I,t}) - (g(w, z', U_{i,t}) - g(w, z, U_{i,t}))\,|Z_{i,t} = z']$$

which completes the proof. □

A.2. Proof of Theorem 2

We begin the proof of Theorem 2 with a simple lemma.

Lemma 1

Suppose that Assumption 1-4 hold. Then, $\forall z \in \operatorname{supp} Z_{i,t}$,

$$ACE(w', w) = E\Big[DCE(w', w, Z_{i,t}, z)\Big] + E[g(w', z, U_{i,t}) - g(w, z, U_{i,t})].$$

Proof of Lemma 1

By Theorem 1, we have

$$\begin{aligned}&E[g(w', z', U_{i,t}) - g(w, z', U_{i,t})|Z_{i,t} = z']\\ &\qquad= \mathrm{DCE}(w', w, z', z) + E[g(w', z, U_{i,t}) - g(w, z, U_{i,t})|Z_{i,t} = z'].\end{aligned}$$

Integrating z' over the marginal of $Z_{i,t}$ (which by Assumption 1 is the same as the marginal of $Z_{i,t'}$) and using the law of iterated expectations then results in

$$\begin{aligned}&E[g(w', Z_{i,t}, U_{i,t}) - g(w, Z_{i,t}, U_{i,t})]\\ &\qquad= E\Big[\mathrm{DCE}(w', w, Z_{i,t}, z)\Big] + E[g(w', z, U_{i,t}) - g(w, z, U_{i,t})].\end{aligned}$$

The desired result then follows from the definition of the ACE in Equation (2). □

To complete the proof of Theorem 2, note that when $E[g(w', z, U_{i,t'}) - g(w, z, U_{i,t'})] \geq 0, \forall z \in \operatorname{supp} Z_{i,t}$, we have that

$$\begin{aligned}&\mathrm{ACE}(w', w) = E\Big[\mathrm{DCE}(w', w, Z_{i,t}, z)\Big] + E[g(w', z, U_{i,t}) - g(w, z, U_{i,t})]\\ \Rightarrow\quad &\mathrm{ACE}(w', w) \geq E\Big[\mathrm{DCE}(w', w, Z_{i,t}, z)\Big],\end{aligned}$$

where the first equation follows from Lemma 1, and the inequality follows from $E[g(w', z, U_{i,t'}) -$

$g(w, z, U_{i,t'})] \geq 0, \forall z \in \operatorname{supp} Z_{i,t}$. Because the inequality holds for all $z \in \operatorname{supp} Z_{i,t}$, we also have

$$\mathrm{ACE}(w', w) \geq \max_{z \in \operatorname{supp} Z_{i,t}} E\Big[\mathrm{DCE}(w', w, Z_{i,t}, z)\Big],$$

which is the desired result.

The argument for $E[g(w', z, U_{i,t'}) - g(w, z, U_{i,t'})] \leq 0, \forall z \in \operatorname{supp} Z_{i,t}$ is identical and omitted here for brevity. $\square$

S.2. Identification in a Linear Causal Model

This section presents results on the identification of causal effects of a macroeconomic variable on a microeconomic outcome in a simplified version of the nonparametric causal model discussed in Section 2. We consider the random vector $(Y_{i,t}, W_t, Z_{i,t}, \xi^w_{i,t}, \xi^z_{i,t}, \xi^{wz}_{i,t}, \varepsilon_{i,t})$ whose joint distribution is characterized by the causal model

$$Y_{i,t} = W_t(\beta^w + \xi^w_{i,t}) + Z_{i,t}(\beta^z + \xi^z_{i,t}) + W_t Z_{i,t}(\beta^{wz} + \xi^{wz}_{i,t}) + \varepsilon_{i,t}, \tag{S.2.1}$$

where β^w, β^z, and β^{wz} are unknown fixed coefficients. As in the main text, $Y_{i,t}$ denotes the outcome of individual $i \in \mathcal{N}$ at time $t \in \mathcal{T}$, W_t is the macroeconomic variable of interest, and $Z_{i,t}$ is the exposure variable. $\varepsilon_{i,t}$ are unobserved determinants of the outcome and the random vector $(\xi^w_{i,t}, \xi^z_{i,t}, \xi^{wz}_{i,t})$ captures unobserved heterogeneity.

The main parameter of interest is the expected change in the outcome caused by a marginal change in W_t:

$$\tau_w \equiv \left(\beta^w + E\left[\xi^w_{i,t}\right]\right) + E\left[Z_{i,t}\left(\beta^{wz} + \xi^{wz}_{i,t}\right)\right]$$

If W_t were randomly assigned, this would simply be the coefficient obtained from a regression of $Y_{i,t}$ on W_t. In a macroeconomic setting, W_t often is not randomly assigned. In particular, the approach we pursue allows for correlation between $\varepsilon_{i,t}$ and W_t, making identification of the main effect of W_t challenging. To make progress, we proceed in two steps. First, we provide assumptions to identify the difference in the effect of W_t between two levels of exposure to Z_t, z' and z. This difference is defined by

$$\begin{aligned}\Delta\tau_w(z', z) \equiv& \left(\beta^w + E\left[\xi^w_{I,t} | Z_{i,t} =''\right]\right) +'' \left(\beta^{wz} + E\left[\xi^{wz}_{i,t} | Z_{i,t} =''\right]\right) \\ &- - \left(\beta^w + E\left[\xi^w_{i,t} | Z_{i,t} = z\right]\right) + z\left(\beta^{wz} + E\left[\xi^{wz}_{i,t} | Z_{i,t} = z\right]\right).\end{aligned} \tag{S.2.2}$$

$\tau_w('', z)$ captures differences in the effect of marginal changes in W_t on the outcome from two levels of exposure, $''$ and z. The difference may itself be interesting as it allows for assessment of distributional changes across groups caused by changes in W_t.

The parameter $\tau_w('', z)$ is similar to the average treatment effect on the treated that is often the target in difference-in-difference designs. However, unlike in conventional difference-in-difference designs, there exists no natural non-treated group in the setting considered here., i.e., there does *not* exist a known z such that

$$\tau_w(z) \equiv \left(\beta^w + E\left[\xi^w_{i,t}|Z_{i,t} = z\right]\right) + z\left(\beta^{wz} + E\left[\xi^{wz}_{i,t}|Z_{i,t} = z\right]\right) = 0.$$

Motivated by the background risk literature in the empirical application to health care policy uncertainty, we instead consider the slightly more general setting with known sign-restrictions for subsets the exposure level, that is, where

$$\tau_w(z) \geq 0, \quad \text{or} \quad \tau_w(z) \leq 0, \qquad \forall z \in \mathcal{Z} \subset \operatorname{supp} Z.$$

Our second result shows that sign restrictions of this form then allow for partial identification of the main effect of W_t, τ_w.

We now state assumptions analogous to those in the main text. First, Assumption 1' restricts the distribution of unobserved heterogeneity to be stationary.

Assumption 1' (Stationarity)

$\operatorname{supp} Z_{i,t}$ *is time invariant and for all* $z \in \operatorname{supp} Z_{i,t}$*, (i)* $E[\xi^{wz}_{i,t'}|Z_{i,t'} = z] = E[\xi^{wz}_{i,t}|Z_{i,t} = z]$*. (ii)* $E[\xi^w_{i,t'}|Z_{i,t'} = z] = E[\xi^w_{i,t}|Z_{i,t} = z]$*. (iii)* $E[\xi^z_{i,t'}|Z_{i,t'} = z] = E[\xi^z_{i,t}|Z_{i,t} = z]$*.*

Second, the exposure levels are mean-independent of the unobserved determinants of $\varepsilon_{i,t}$ conditional on the macroeconomic variable.

Assumption 2' (Exogenous Exposure)

$E[\varepsilon_{i,t}|W_t, Z_{i,t}] = E[\varepsilon_{i,t}|W_t]$.

Third, unobserved heterogeneity in the effects of both the macroeconomic variable W_t and the exposure variable $Z_{i,t}$ does not depend on W_t. Assumption 3' is stronger than necessary. In particular, it is possible to allow for unobserved heterogeneity $\xi^w_{i,t}$ that is correlated with W_t itself. However, it must not be correlated with *both* W_t and $Z_{i,t}$. The model in (S.2.1) could thus be amended by the term $W_t\tilde{\xi}^w_{i,t}$ where $E[\tilde{\xi}^w_{i,t}|W_t, Z_{i,t}]$ does not depend on $Z_{i,t}$ but does depend on W_t.

Assumption 3' (Exogenous Heterogeneous Effects)

(i) $E[\xi^{wz}_{i,t}|W_t, Z_{i,t}] = E[\xi^{wz}_{i,t}|Z_{i,t}]$. *(ii)* $E[\xi^{w}_{i,t}|W_t, Z_{i,t}] = E[\xi^{w}_{i,t}|Z_{i,t}]$. *(iii)* $E[\xi^{z}_{i,t}|W_t, Z_{i,t}] = E[\xi^{z}_{i,t}|Z_{i,t}]$.

Finally, we assume common support as in the main text.

Assumption 4' (Common Support)

For any t*,* $f_{WZ}(w, z) > 0$, $\forall w \in \operatorname{supp} W_t, z \in \operatorname{supp} Z_{i,t}$*, where* f_{WZ} *denotes the joint density.*

Assumptions 2'-4' are sufficient to prove point identification of $\tau_w(z', z)$. Theorem 1' below is the analogue of Theorem 1 in the main text.

Theorem 1'

Suppose that Assumptions 1'-4' hold. Then, $\forall w', w \in \operatorname{supp} W_t, z', z \in \operatorname{supp} Z_{i,t}$,

$$\begin{aligned}\Delta\tau_w(z', z) = \Big[& E[Y_{i,t'}|Z_{i,t'} = z', W_{t'} = w'] - E[Y_{i,t}|Z_{i,t} = z', W_t = w] \\ & - (E[Y_{i,t'}|Z_{i,t'} = z, W_{t'} = w'] - E[Y_{i,t}|Z_{i,t} = z, W_t = w])\Big] \Big/ (w' - w).\end{aligned}$$

Proof. Substituting for $Y_{i,t}$ and $Y_{i,t'}$ using the model in (S.2.1) and organizing terms results in

$$
\begin{aligned}
&E[Y_{i,t'}|Z_{i,t'} = z', W_{t'} = w'] - E[Y_{i,t}|Z_{i,t} = z', W_t = w] \\
&\quad - (E[Y_{i,t'}|Z_{i,t'} = z, W_{t'} = w'] - E[Y_{i,t}|Z_{i,t} = z, W_t = w]) \\
=&(w' - w)(z' - z)\beta^{wz} \\
&\quad + w'\Big[\big(E[\xi^{w}_{i,t'}|Z_{i,t'} = z', W_{t'} = w'] + z'E[\xi^{wz}_{i,t'}|Z_{i,t'} = z', W_{t'} = w']\big) \\
&\qquad\qquad - \big(E[\xi^{w}_{i,t'}|Z_{i,t'} = z, W_{t'} = w'] + zE[\xi^{wz}_{i,t'}|Z_{i,t'} = z, W_{t'} = w']\big) \Big] \\
&\quad - w\Big[\big(E[\xi^{w}_{i,t}|Z_{i,t} = z', W_t = w] + z'E[\xi^{wz}_{i,t}|Z_{i,t} = z', W_t = w]\big) \\
&\qquad\qquad - \big(E[\xi^{w}_{i,t}|Z_{i,t} = z, W_t = w] + zE[\xi^{wz}_{i,t}|Z_{i,t} = z, W_t = w]\big) \Big] \\
&\quad + \Big[z' \big(E[\xi^{z}_{i,t'}|Z_{i,t'} = z', W_{t'} = w'] - E[\xi^{z}_{i,t}|Z_{i,t} = z', W_t = w]\big) \\
&\qquad\qquad - z \big(E[\xi^{z}_{i,t'}|Z_{i,t'} = z, W_{t'} = w'] - E[\xi^{z}_{i,t}|Z_{i,t} = z, W_t = w]\big) \Big] \\
&\quad + \Big[(E[\varepsilon_{i,t'}|Z_{i,t'} = z', W_{t'} = w'] - E[\varepsilon_{i,t}|Z_{i,t} = z', W_t = w]) \\
&\qquad\qquad - (E[\varepsilon_{i,t'}|Z_{i,t'} = z, W_{t'} = w'] - E[\varepsilon_{i,t}|Z_{i,t} = z, W_t = w]) \Big].
\end{aligned}
$$

We now proceed term-by-term. Under Assumption 3' (i) and (ii), we have

$$
\begin{aligned}
&w'\Big[\big(E[\xi^{w}_{i,t'}|Z_{i,t'} = z', W_{t'} = w'] + z'E[\xi^{wz}_{i,t'}|Z_{i,t'} = z', W_{t'} = w']\big) \\
&\qquad - \big(E[\xi^{w}_{i,t'}|Z_{i,t'} = z, W_{t'} = w'] + zE[\xi^{wz}_{i,t'}|Z_{i,t'} = z, W_{t'} = w']\big) \Big] \\
=&w'\Big[\big(E[\xi^{w}_{i,t'}|Z_{i,t'} = z'] + z'E[\xi^{wz}_{i,t'}|Z_{i,t'} = z']\big) - \big(E[\xi^{w}_{i,t'}|Z_{i,t'} = z] + zE[\xi^{wz}_{i,t'}|Z_{i,t'} = z]\big) \Big]
\end{aligned} \tag{S.2.3}
$$

and similarly

$$
\begin{aligned}
&w\Big[\big(E[\xi^{w}_{i,t}|Z_{i,t} = z', W_t = w] + z'E[\xi^{wz}_{i,t}|Z_{i,t} = z', W_t = w]\big) \\
&\qquad - \big(E[\xi^{w}_{i,t}|Z_{i,t} = z, W_t = w] + zE[\xi^{wz}_{i,t}|Z_{i,t} = z, W_t = w]\big) \Big] \\
=&w\Big[\big(E[\xi^{w}_{i,t}|Z_{i,t} = z'] + z'E[\xi^{wz}_{i,t}|Z_{i,t} = z']\big) - \big(E[\xi^{w}_{i,t}|Z_{i,t} = z] + zE[\xi^{wz}_{i,t}|Z_{i,t} = z]\big) \Big].
\end{aligned} \tag{S.2.4}
$$

Under Assumption 1' (i) and (ii), the difference between (S.2.3) and (S.2.4) simplifies to

$$(w'-w)\Big[\left(E[\xi^{w}_{i,t}|Z_{i,t}=z']+z'E[\xi^{wz}_{i,t}|Z_{i,t}=z']\right)-\left(E[\xi^{w}_{i,t}|Z_{i,t}=z]+zE[\xi^{wz}_{i,t}|Z_{i,t}=z]\right)\Big].$$

Further, we have

$$\begin{aligned}
&z'\left(E[\xi^{z}_{i,t'}|Z_{i,t'}=z',W_{t'}=w']-E[\xi^{z}_{i,t}|Z_{i,t}=z',W_t=w]\right)\\
&\qquad\qquad -z\left(E[\xi^{z}_{i,t'}|Z_{i,t'}=z,W_{t'}=w']-E[\xi^{z}_{i,t}|Z_{i,t}=z,W_t=w]\right)\\
&\overset{[1]}{=}z'\left(E[\xi^{z}_{i,t'}|Z_{i,t'}=z']-E[\xi^{z}_{i,t}|Z_{i,t}=z']\right)-z\left(E[\xi^{z}_{i,t'}|Z_{i,t'}=z]-E[\xi^{z}_{i,t}|Z_{i,t}=z]\right)\\
&\overset{[2]}{=}0,
\end{aligned}$$

where [1] follows from Assumption 3' (iii), and [2] follows from Assumption 1' (iii).

Finally, note that Assumption 2' implies

$$\begin{aligned}
&(E[\varepsilon_{i,t'}|Z_{i,t'}=z',W_{t'}=w']-E[\varepsilon_{i,t}|Z_{i,t}=z',W_t=w])\\
&\qquad -(E[\varepsilon_{i,t'}|Z_{i,t'}=z,W_{t'}=w']-E[\varepsilon_{i,t}|Z_{i,t}=z,W_t=w])=0.
\end{aligned}$$

Combining results in

$$\begin{aligned}
&E[Y_{i,t'}|Z_{i,t'}=z',W_{t'}=w']-E[Y_{i,t}|Z_{i,t}=z',W_t=w]\\
&\qquad -(E[Y_{i,t'}|Z_{i,t'}=z,W_{t'}=w']-E[Y_{i,t}|Z_{i,t}=z,W_t=w])\\
&=(w'-w)\Big[(\beta^{w}+E\left[\xi^{w}_{i,t}|Z_{i,t}=z'\right])+z'\left(\beta^{wz}+E\left[\xi^{wz}_{i,t}\right]\right)\\
&\qquad\qquad -(\beta^{w}+E\left[\xi^{w}_{i,t}|Z_{i,t}=z\right])+z\left(\beta^{wz}+E\left[\xi^{wz}_{i,t}\right]\right)\Big].
\end{aligned}$$

Division by $(w'-w)$ then completes the proof. □

If in addition the researcher has knowledge of sign restrictions over the subgroup-specific effects, the main effect τ_w is partially identified. Theorem 2' below is the analogue of Theorem 2 in the main text.

Theorem 2'

Suppose that Assumptions 1'-4' hold. Take $w' \geq w \in \operatorname{supp} W_t$.

- *If in addition* $\tau_w(z) \geq 0, \forall z \in \mathcal{Z}$, *then*

$$\tau_w \geq \max_{z \in \mathcal{Z}} E\Big[\Delta\tau_w(Z_{i,t}, z)\Big].$$

- *If instead, in addition* $\tau_w(z) \leq 0, \forall z \in \mathcal{Z}$, *then*

$$\tau_w \leq \min_{z \in \mathcal{Z}} E\Big[\Delta\tau_w(Z_{i,t}, z)\Big].$$

Proof.

Note that

$$\begin{aligned}
\Delta\tau_w(z', z) &= \tau_w(z') - \tau_w(z) \\
&= \left(\beta^w + E\left[\xi^w_{i,t} | Z_{i,t} = z'\right]\right) + z'\left(\beta^{wz} + E\left[\xi^{wz}_{i,t} | Z_{i,t} = z'\right]\right) \\
&\quad - \left(\beta^w + E\left[\xi^w_{i,t} | Z_{i,t} = z\right]\right) + z\left(\beta^{wz} + E\left[\xi^{wz}_{i,t} | Z_{i,t} = z\right]\right).
\end{aligned}$$

Integrating with respect to z' over the marginal distribution of $Z_{i,t}$ and using the law of iterated expectations implies

$$\begin{aligned}
E[\Delta\tau_w(Z_{i,t}, z)] = &\left(\beta^w + E\left[\xi^w_{i,t}\right]\right) + E\left[Z_{i,t}\left(\beta^{wz} + \xi^{wz}_{i,t}\right)\right] \\
&- \left(\beta^w + E\left[\xi^w_{i,t} | Z_{i,t} = z\right]\right) + z\left(\beta^{wz} + E\left[\xi^{wz}_{i,t} | Z_{i,t} = z\right]\right) \\
= &\tau_w - \tau_w(z).
\end{aligned}$$

Consider now the first case in which $\tau_w(z) \geq 0, \forall z \in \mathcal{Z}$. Thus taking any $z \in \mathcal{Z}$,

$$\tau_w = E[\Delta\tau_w(Z_{i,t}, z)] - \tau_w(z) \geq E[\Delta\tau_w(Z_{i,t}, z)].$$

Since this inequality holds for any $z \in \mathcal{Z}$, the first result of Theorem 2' follows. The arguments for the second case in $\tau_w(z)\lambda \leq 0, \forall z \in \mathcal{Z}$ are analogous and omitted here for brevity. □

S.3. Identification without Stationarity

In this appendix we provide an alternative to Theorem 2 that does not maintain the stationarity assumption – i.e., Assumption 1 of the main text – and where the exposure variable Z_i is time-invariant. The target parameter in this setting is

$$\mathrm{ACE}_t(w', w) + E[\tilde{\Delta}_{t',t}(w')],$$

where

$$\mathrm{ACE}_t(w', w) \equiv E[\Delta_{i,t}(w', w, Z_i)] = E[g(w', Z_i, U_{i,t}) - g(w, Z_i, U_{i,t})]$$

and

$$\tilde{\Delta}_{i,t,t'}(w') \equiv g(w, Z_i, U_{i,t'}) - g(w', Z_i, U_{i,t}).$$

While the ACE_t describes the expected change in the outcome caused by a *ceteris paribus* change in W_t at time t, $\tilde{\Delta}_{i,t',t}$ denotes the change in the outcome resulting from a shift in the distribution of the unobservables (but holding the distribution of the macroeconomic variable fixed).[1] Under Assumption 1, we simply have $E[\tilde{\Delta}_{i,t',t}(w)] = 0, \forall w \in \operatorname{supp} W_t$, resulting in the identification results of the main text.

Assumptions 2-4 need to be adapted slightly to account for time-invariant exposure and non-stationary unobservable distributions:

Assumption 5'

$\forall w', w \in \operatorname{supp} W_t, z \in \operatorname{supp} Z_i, E\left[g(w', z, U_{i,t'})|W_{t'} = w', Z_i\right] - E\left[g(w, z, U_{i,t})|W_t = w, Z_i\right]$ does not depend on Z_i.

[1]Note that the $\mathrm{ACE}_t(w', w)$ is nearly identical to the average causal effect defined in (2) except for the subscript that highlights the potential time-varying nature of the ACE.

Assumption 6'

For $\tilde{t} \in \{t', t\}$ *and* $\forall w \in \operatorname{supp} W_t, z', z \in \operatorname{supp} Z_i$, $E\left[g(w, z', U_{i,\tilde{t}}) - g(w, z, U_{i,\tilde{t}}) \middle| W_{\tilde{t}}, Z_i = z\right]$ *does not depend on* $W_{\tilde{t}}$.

Assumption 7'

For $\tilde{t} \in \{t', t\}$, $f_{W_{\tilde{t}} Z_i}(w, z) > 0$, $\forall w \in \operatorname{supp} W_{\tilde{t}}, z \in \operatorname{supp} Z_i$, *where* $f_{W_{\tilde{t}} Z_i}$ *denotes the joint density at time* $\tilde{t}$.

Note that with time-invariant exposure variables, the similarity of Assumption 5' with conventional difference in difference assumptions where Z_i takes the value of the treatment is even more apparent than highlighted in Section 2 with time-varying exposure variables.

Finally, we can no longer obtain partial identification of an average causal effect simply by restricting the subgroup-specific causal effects. Instead, the combination of the subgroup-specific causal effects at time t and the subgroup-specific effects of the change in the distribution of unobservables should be restricted – i.e., we leverage sign restrictions of the form

$$E[\Delta_{i,t}(w', w, z)] + E[g(w', z, U_{i,t'}) - g(w', z, U_{i,t})] \geq 0.$$

When both terms are considered separately, sign-restriction for the first term can be motivated analogously to the sign restrictions leveraged for Theorem 2. Sign restrictions for the second term, however, require additional information about the change in the distribution of unobservables between time periods t and t'.

Under the adapted assumptions and sign restrictions, we can now state an alternative version of Theorem 2 without stationarity:

Theorem 3'

Suppose that Assumptions 5'-7' hold. Take $w' \geq w \in \operatorname{supp} W_t$.

- *If in addition* $E[\Delta_{i,t}(w',w,z)] + E[g(w',z,U_{i,t'}) - g(w',z,U_{i,t})] \geq 0, \forall z \in \operatorname{supp} Z_i$*, then*

$$\begin{aligned} ACE_t(w',w) + E[\tilde{\Delta}_{t',t}(w')] \geq\ & E\Big[E[Y_{i,t'}|W_{t'} = w', Z_i] - E[Y_{i,t}|W_t = w, Z_i]\Big] \\ & - \min_{z \in \operatorname{supp} Z_i} \Big[E[Y_{i,t'}|W_{t'} = w', Z_i = z] - E[Y_{i,t}|W_t = w, Z_i = z]\Big]. \end{aligned}$$

- *If in addition* $E[\Delta_{i,t}(w',w,z)] + E[g(w',z,U_{i,t'}) - g(w',z,U_{i,t})] \leq 0, \forall z \in \operatorname{supp} Z_i$*, then*

$$\begin{aligned} ACE_t(w',w) + E[\tilde{\Delta}_{t',t}(w')] \leq\ & E\Big[E[Y_{i,t'}|W_{t'} = w', Z_i] - E[Y_{i,t}|W_t = w, Z_i]\Big] \\ & - \max_{z \in \operatorname{supp} Z_i} \Big[E[Y_{i,t'}|W_{t'} = w', Z_i = z] - E[Y_{i,t}|W_t = w, Z_i = z]\Big]. \end{aligned}$$

Proof.

The proof of Theorem 3' combines and adapts the proof of Theorems 1 and 2. First,

$$\begin{aligned} & (E[Y_{i,t'}|W_{t'} = w', Z_i = z'] - E[Y_{i,t}|W_t = w, Z_i = z']) \\ & \quad - (E[Y_{i,t'}|W_{t'} = w', Z_i = z] - E[Y_{i,t}|W_t = w, Z_i = z]) \\ & \overset{[1]}{=} (E[g(w',z',U_{i,t'})|W_{t'} = w', Z_i = z'] - E[g(w,z',U_{i,t})|W_t = w, Z_i = z']) \\ & \quad - (E[g(w',z,U_{i,t'})|W_{t'} = w', Z_{i,t'} = z] - E[g(w,z,U_{i,t})|W_t = w, Z_i = z]) \\ & \overset{[2]}{=} (E[g(w',z',U_{i,t'})|W_{t'} = w', Z_i = z'] - E[g(w',z,U_{i,t'})|W_{t'} = w', Z_i = z]) \\ & \quad - (E[g(w,z',U_{i,t})|W_t = w, Z_i = z'] - E[g(w,z,U_{i,t})|W_t = w, Z_i = z]) \\ & \overset{[3]}{=} (E[g(w',z',U_{i,t'})|W_{t'} = w', Z_i = z'] - E[g(w',z,U_{i,t'})|W_{t'} = w', Z_i = z']) \\ & \quad + (E[g(w',z,U_{i,t'})|W_{t'} = w', Z_i = z'] - E[g(w',z,U_{i,t'})|W_{t'} = w', Z_i = z]) \\ & \quad - (E[g(w,z',U_{i,t})|W_t = w, Z_i = z'] - E[g(w,z,U_{i,t})|W_t = w, Z_i = z']) \\ & \quad - (E[g(w,z,U_{i,t})|W_t = w, Z_i = z'] - E[g(w,z,U_{i,t})|W_t = w, Z_i = z]) \\ & \overset{[4]}{=} (E[g(w',z',U_{i,t'}) - g(w',z,U_{i,t'})|W_{t'} = w', Z_i = z']) \\ & \quad - E[g(w,z',U_{i,t}) - g(w,z,U_{i,t})|W_t = w, Z_i = z']) \\ & \overset{[5]}{=} (E[g(w',z',U_{i,t'}) - g(w',z,U_{i,t'})|Z_i = z'] - E[g(w,z',U_{i,t}) - g(w,z,U_{i,t})|Z_i = z'] \end{aligned}$$

where [1] substitutes for $Y_{i,t} = g(W_t, Z_{i,t}, U_{i,t})$, [2] rearranges terms, [3] both adds and subtracts the two terms $E[g(w', z, U_{i,t'})|W_{t'} = w', Z_i = z']$ and $E[g(w, z, U_{i,t})|W_t = w, Z_i = z']$, [4] follows from Assumption 5', and [5] follows from Assumption 6'.

Second, integrating z' over the marginal distribution of Z_i and using the law of iterated expectations implies

$$
\begin{aligned}
&E\Big[E[Y_{i,t'}|W_{t'} = w', Z_i] - E[Y_{i,t}|W_t = w, Z_i]\Big] \\
&\quad - (E[Y_{i,t'}|W_{t'} = w', Z_i = z] - E[Y_{i,t}|W_t = w, Z_i = z]) \\
=& E[g(w', Z_i, U_{i,t'}) - g(w, Z_i, U_{i,t})] - E[g(w', z, U_{i,t'}) - g(w, z, U_{i,t})].
\end{aligned}
$$

Third, we use $E[\Delta_{i,t}(w', w, z)] + E[g(w', z, U_{i,t'}) - g(w', z, U_{i,t})] \geq 0, \forall z \in \operatorname{supp} Z_i$ and the fact that

$$
E[g(w', z, U_{i,t'}) - g(w, z, U_{i,t})] = E[\Delta_{i,t}(w', w, z)] + E[g(w', z, U_{i,t'}) - g(w', z, U_{i,t})] \quad \text{(S.3.1)}
$$

to obtain

$$
\begin{aligned}
E[g(w', Z_i, U_{i,t'}) - g(w, Z_i, U_{i,t})] \geq\ & E\Big[E[Y_{i,t'}|W_{t'} = w', Z_i] - E[Y_{i,t}|W_t = w, Z_i]\Big] \\
& - \min_{z \in \operatorname{supp} Z_i} \Big[E[Y_{i,t'}|W_{t'} = w', Z_i = z] - E[Y_{i,t}|W_t = w, Z_i = z]\Big].
\end{aligned}
$$

Finally, note that

$$
E[g(w', Z_i, U_{i,t'}) - g(w, Z_i, U_{i,t})] = \mathrm{ACE}_t(w', w) + E[\tilde{\Delta}_{t',t}(w')],
$$

which concludes the proof. $\square$

S.4. Dataset Citations

The following provides references to the data sources compiled for the analysis in the main text and the appendix.

Baker, Scott R, Bloom, Nicholas, and Davis, Steven J (2020). Categorical EPU Data. Retrieved on January, 2020, from https://www.policyuncertainty.com/categorical_epu.html.

Health and Retirement Study (2018). 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 (2018). 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.

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.

S.5. Construction of the Analysis Sample

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.[2] An additional 3,707 and 4,095 single and couple household observations, respectively, are omitted from the analysis due to missing one or more of the variables used in estimation. Finally, households with an observation in only one wave (3,332 and 2,439 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 45,384 single and 57,190 couple household-wave observations for the portfolio choice analysis.

Table S.2.1 shows the number of observations that were excluded from the analysis sample due to missing information. The first two columns ("HRS") describe the data used in the main analysis while the third and forth columns ("CAMS") describe the Consumption and Activities Mail Survey data used in the supplemental analysis of consumption in Appendix S.9. The difference in starting observations for each subsample is due to unequal distribution of couple and single households.

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 paper, these numbers point to a similar proportion of observations excluded. Most recently, Gábor-Tóth and Georgarakos (2019) consider the CAMS dataset also employed in

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

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 20,752 observations used for analysis of consumption).

TABLE S.2.1
Tabulation of Deleted Observations

		HRS		CAMS	
		Single	Couple	Single	Couple
Starting		68,507	72,426	17,847	11,811
Risky asset share	-	16,084	8,702		
Consumption	-			4,022	2,837
Education	-	7	6	0	0
Race	-	12	1,393	11	298
Self-reported health	-	46	34	12	6
Severe conditions	-	281	199	66	31
Mobility	-	2,916	2,223	55	14
ADLs	-	1	5	0	0
Hospital nights	-	444	235	73	34
Households recorded in only one wave	-	3,332	2,439	972	475
Final # of observations	=	45,384	57,190	12,636	8,116
Final # of unique households		10,022	10,555	2,983	1,765

Notes. The table provides the number of observations sequentially deleted from the sample due to missing observations in each variable. Additionally, households that are recorded in only one wave or have no total spending are excluded from the analysis. The two bottom rows show the number of final observations and the associated number of unique households, respectively.

Table S.5.2 presents complementary summary statistics. 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. Single households are disproportionately female (more than 71%) and despite being nearly five years older on average, are healthier both in terms of self-reported health and most other indicators (all except ADLs). Couple households tend to have a greater share of risky assets (and a lower share of safe assets), consistent with theories of risk-sharing and diversification.

TABLE S.5.2
Summary Statistics

Single households	Mean	Sd.		Mean	Sd.		Share
Asset Shares			*Hh Characteristics*			*Hh Proportions (%)*	
Risky assets	0.136	0.278	Self-reported health	2.839	1.101	Female	0.713
Safe assets	0.679	0.399	Severe conditions	0.717	0.905	Retired	0.575
IRA	0.171	0.308	Mobility	1.750	1.822	No high school	0.176
Bonds	0.015	0.085	ADLs	0.465	1.169	GED	0.038
Macro. Variables			Hospital nights	2.344	9.130	High school	0.329
HCPU	127.421	102.519	Income ($100,000)	0.437	1.101	Some college	0.247
HCPU_{12}	131.353	75.675	Wealth ($100,000)	2.974	11.692	Above college	0.209
			Years ret.	8.288	10.885	White	0.863
			Age	69.113	11.296	Black	0.100
						Other	0.037

Couple households	Mean	Sd.		Mean	Sd.		Share
Asset Shares			*Hh Characteristics*			*Hh Proportions (%)*	
Risky assets	0.167	0.284	Self-reported health	3.058	1.017	Retired	0.579
Safe assets	0.517	0.409	Severe conditions	0.895	0.908	No high school	0.047
IRA	0.299	0.360	Mobility	1.883	1.747	GED	0.023
Bonds	0.016	0.080	ADLs	0.469	1.199	High school	0.243
Macro. Variables			Hospital nights	2.892	10.929	Some college	0.273
HCPU	128.632	102.413	Income ($100,000)	1.056	2.378	Above college	0.414
HCPU_{12}	133.243	74.815	Wealth ($100,000)	5.408	13.573	White-White	0.886
			Years ret.	6.887	9.188	Black-Black	0.048
			Age	64.943	9.280	Other-Other	0.020
						White-Black	0.005
						White-Other	0.039
						Black-Other	0.003

Notes. Statistics for single households are based on the sample of 45,384 household-wave observations. For couple households, there are 57,190 observations used in the analysis. 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. HCPU_{12} denotes the 12-month average of Baker et al.'s (2016) health care policy uncertainty. Health measures, age, years of retirement, retirement status, and highest obtained education are defined as the maximum across spouses. $ denotes 2010 Dollars.

S.6. Implementation Details

This appendix provides pseudocode for the estimation procedures of the relative distributions of the DCE (e.g., Figure 2) and the bounds on the ACE (see Table 1).

Algorithm 1 provides the pseudocode that serves as a first-step basis for plotting the relative distribution of the DCE. Once the local linear regression coefficients – denoted here as $\hat{\gamma}_k^{(b,s)}$, where b denotes the bootstrap iteration, s corresponds to a particular value on the horizontal axis, and k denotes one of the health categories – are calculated, this relative distribution may readily be calculated. For a particular $k \in \{1, \ldots, \dim Z\}$ and a selected baseline $s^\star \in \{1, \ldots, S\}$, we can calculate $\hat{\gamma}_k^{(b,s)} - \hat{\gamma}_k^{(b,s^\star)}$ for each $b \in \{1, \ldots, B\}$. For each s, this allows for straightforward calculation of bootstrap statistics such as the mean or the 99% confidence bands using Algorithm 2 in Montiel Olea and Plagborg-Møller (2019).

Algorithm 1 BootstrapLLR

Required input: B number of bootstraps; $\{(y_{i,t}, \mathrm{HCPU}_t, Z_{i,t}, X_{i,t})\}_{i\in\mathcal{D},t\in\mathcal{T}}$ the data, where $\mathcal{D}$ is the set of households and $\mathcal{T}$ is the set of time periods; $\{z_k^{(s)}\}_{s\in\{1,\ldots,S\},k\in\{1,\ldots,\dim Z\}}$ a set of values for the Z-variables for which to calculate the local linear regression coefficients;

procedure BootstrapLLR

  for $b \in \{1, \ldots, B\}$ **do**

    $\mathcal{D}_b \leftarrow$ Sample$(\mathcal{D})$ ▷ Block-bootstrap households

    $\{\tilde{Z}_{i,t}^{(b)}\}_{i\in\mathcal{D}_b,t\in\mathcal{T}} \leftarrow$ Residualize$(\{Z_{i,t}, X_{i,t}\}_{i\in\mathcal{D}_b,t\in\mathcal{T}})$ ▷ Residualize Z-variables

    for $k \in \{1, \ldots, \dim Z\}$ **do**

      for $s \in \{1, \ldots, S\}$ **do**

$$\hat{\gamma}_k^{(b,s)} \leftarrow \arg\min_{\gamma} \min_{\alpha,\beta} \sum_{i\in\mathcal{D}_b,t\in\mathcal{T}} K\left(\tilde{Z}_{it,k}^{(b)}, z_k^{(s)}\right)\left(y_{i,t} - \alpha - \tilde{z}_{it,k}^{(b)}\beta - \log \mathrm{HCPU}_t\gamma\right)^2 \tag{S.6.1}$$

Return: $\{\hat{\gamma}_k^{(b,s)}\}_{b\in\{1,\ldots,B\},s\in\{1,\ldots,S\},k\in\{1,\ldots,\dim Z\}}$

Notes. $\dim Z$ denotes the dimension of the vector of Z-variables, Sample denotes the function that samples households from the empirical distribution of households with replacement. Residualize denotes the VAR-based residualization procedure outlined in Section 3. $K(\cdot,\cdot)$ denotes the Gaussian kernel function. The kernel bandwidths are chosen using the Silverman (1989) rule.

The heatmap figures for visualizing the relative distribution of the differences in conditional average causal effects (e.g., Figure S.7.2) can be generated in similar fashion. This is done by replacing the local linear regression specification in equation (S.6.1) below with a generalized random forest of Athey et al. (2019). Note that instead of providing a sequence of values for each health category as in Algorithm 1, computation of the heatmaps with the generalized random forests requires providing a grid of values for each pair of health categories under consideration.

Algorithm 2 provides the pseudocode for estimation of the bounds on the ACE presented in Table 1. From the estimated ACE in each bootstrap iteration, we can calculate bootstrap statistics such as the mean or the 95th percentile of bootstrap draws.

Algorithm 2 BootstrapACE

Required input: B number of bootstraps; $\{(y_{i,t}, \text{HCPU}_t, Z_{i,t}, X_{i,t})\}_{i\in\mathcal{D},t\in\mathcal{T}}$ the data, where $\mathcal{D}$ is the set of households and $\mathcal{T}$ is the set of time periods; $z^\star$ a set of values for the Z-variables that serve as a baseline case for which $E[\Delta_{i,t}(w', w, z^\star)] \geq E[\Delta_{i,t}(w', w, z)]$ or $E[\Delta_{i,t}(w', w, z^\star)] \leq E[\Delta_{i,t}(w', w, z)], \forall z \in \operatorname{supp} Z_{i,t}$ is assumed;

procedure BootstrapACE

    for $b \in \{1, \ldots, B\}$ **do**
        $\mathcal{D}_b \leftarrow \text{Sample}(\mathcal{D})$ ▷ Block-bootstrap households
        $\{\tilde{Z}^{(b)}_{i,t}\}_{i\in\mathcal{D}_b,t\in\mathcal{T}} \leftarrow \text{Residualize}(\{Z_{i,t}, X_{i,t}\}_{i\in\mathcal{D}_b,t\in\mathcal{T}})$ ▷ Residualize Z-variables
        $\hat{g}^{(b)} \leftarrow \text{CausalForest}(\{(y_{i,t}, \log \text{HCPU}_t, \tilde{Z}^{(b)}_{i,t})\}_{i\in\mathcal{D}_b,t\in\mathcal{T}})$
        $\{\hat{\gamma}^{(b)}_{i,t}\}_{i\in\mathcal{D}_b,t\in\mathcal{T}} \leftarrow \hat{g}^{(b)}(\tilde{Z}^{(b)}_{i,t})$
        $\hat{\gamma}^{(b,\star)} \leftarrow \hat{g}^{(b)}(Z^\star)$
        $\widehat{\text{ACE}}^{(b)} \leftarrow \frac{1}{NT}\sum_i \sum_t \hat{\gamma}^{(b)}_{i,t} - \hat{\gamma}^{(b,\star)}$

Return: $\{\hat{\gamma}^{(b,s)}_k\}_{b\in\{1,\ldots,B\},s\in\{1,\ldots,S\},k\in\{1,\ldots,\dim Z\}}$

Notes. Sample denotes the function that samples households from the empirical distribution of households with replacement. Residualize denotes the VAR-based residualization procedure outlined in Section 3. CausalForest corresponds to the implementation provided by Tibshirani et al. (2020).

S.7. Additional Estimation Results

This appendix contains estimation results corresponding to the three other health categories considered in the paper: self-reported health, mobility, and ADLs. For all figures in this appendix, the estimates are relative to the baseline of households that are in substantially better health than expected, defined here as having a more favorable unexpected change in the corresponding health category than 95% of households in the sample. The plots are based on $B = 1,000$ bootstrap draws.

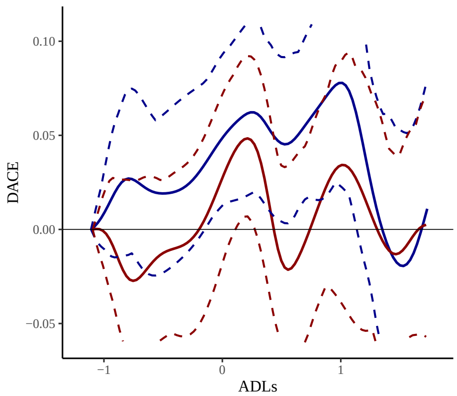
Figure S.7.1 shows the line plots that describe the relative distribution of subgroup-specific average causal effects using single health category. As with Figure 2 in the main text, these plots show the local linear regression-based estimates of the DCE of health care policy uncertainty on safe and risky asset share, respectively, conditional on a single health category. The solid lines correspond to the mean and the dashed lines correspond to the bootstrapped 99% uniform confidence bands, respectively. The figure shows the relative distribution conditional on ADLs (Figures (A) and (B)), mobility (Figures (C) and (D)), and self-reported health (Figures (E) and (F)).

Figures S.7.2 to S.7.9 show heatmaps describing the relative distribution of subgroup-specific average causal effects using two health categories. The center panels correspond to the mean, while the left and the right-most panels provide the bootstrapped lower and upper 99% uniform confidence bands, respectively. As before, the results are normalized such that all estimates are relative to the conditional average causal effect of households that are substantially healthier than expected, which we now define as having a more favorable unexpected change in each of the two health categories than 95% of the households in the sample.[3] Each page of figures presents estimates for couple households in the upper panel (Panel 1) and single households in the lower panel (Panel 2).

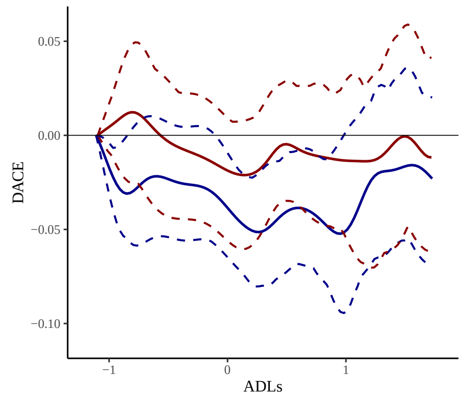
[3]This point corresponds to the bottom-left cell in each of the heatmaps.

Figures S.7.2 and S.7.3 show the random forest-based estimates of the DCE of health care policy uncertainty on safe and risky asset share, respectively, conditional on severe conditions and nights spent in the hospital. Figures S.7.4 and S.7.5 show the analogous plots conditional on severe conditions and ADLs, S.7.4 and S.7.5 show the plots conditional on severe conditions and mobility, and Figures S.7.8 and S.7.9 show the final pair of plots conditional on severe conditions and self-reported health.

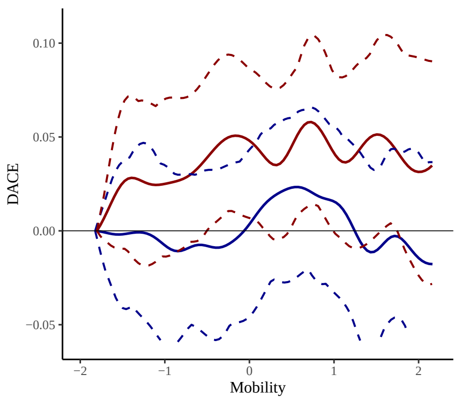
Figure S.7.1
Normalized DCE Estimates for Couple and Single Households (Additional Results)



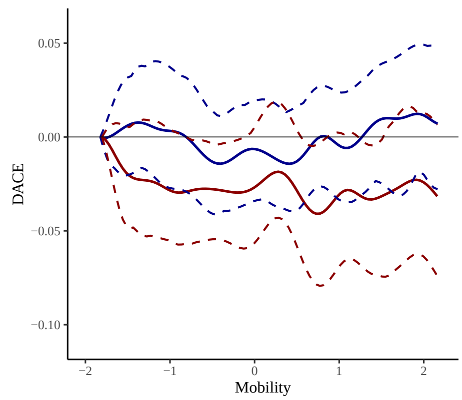
(A) Safe asset share



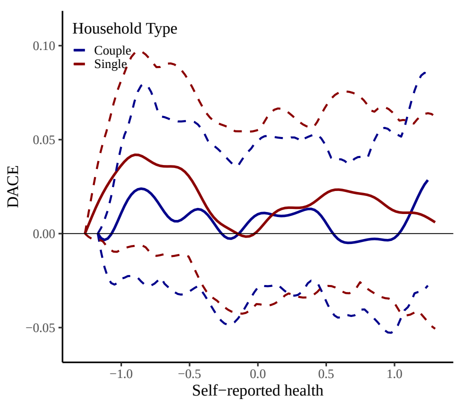
(B) Risky asset share



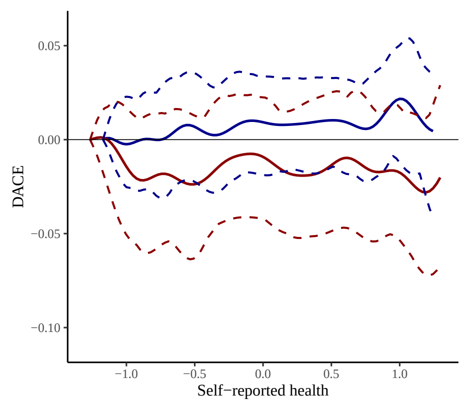
(C) Safe asset share



(D) Risky asset share



(E) Safe asset share

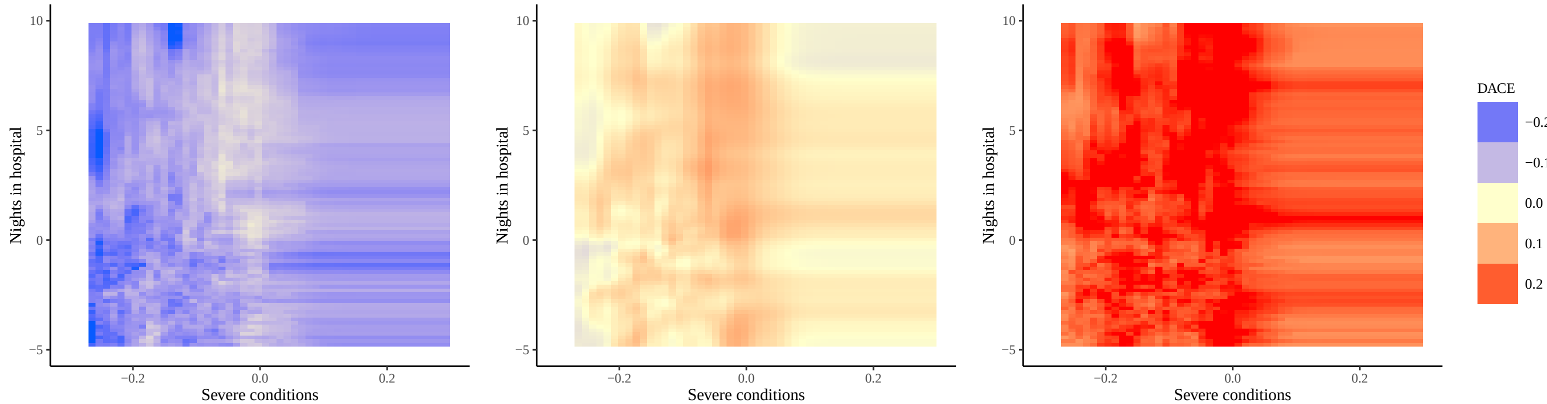


(F) Risky asset share

Figure S.7.2

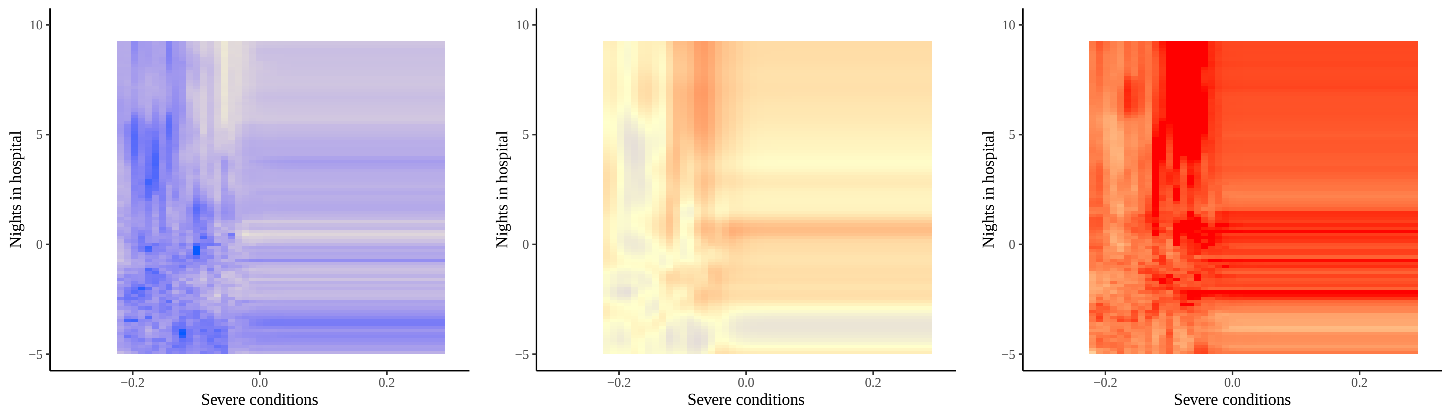
Normalized DCE on Safe Asset Share using Severe Conditions and Nights in Hospital

Panel 1: Couple Households



(A) Lower confidence band (B) Mean (C) Upper confidence band

Panel 2: Single Households

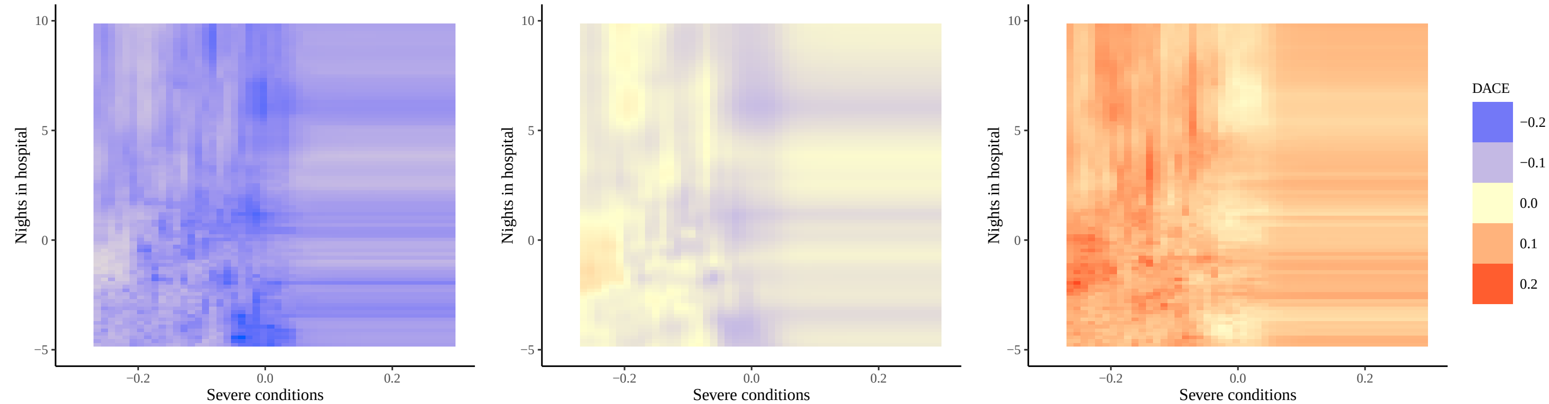


(D) Lower confidence band (E) Mean (F) Upper confidence band

Figure S.7.3

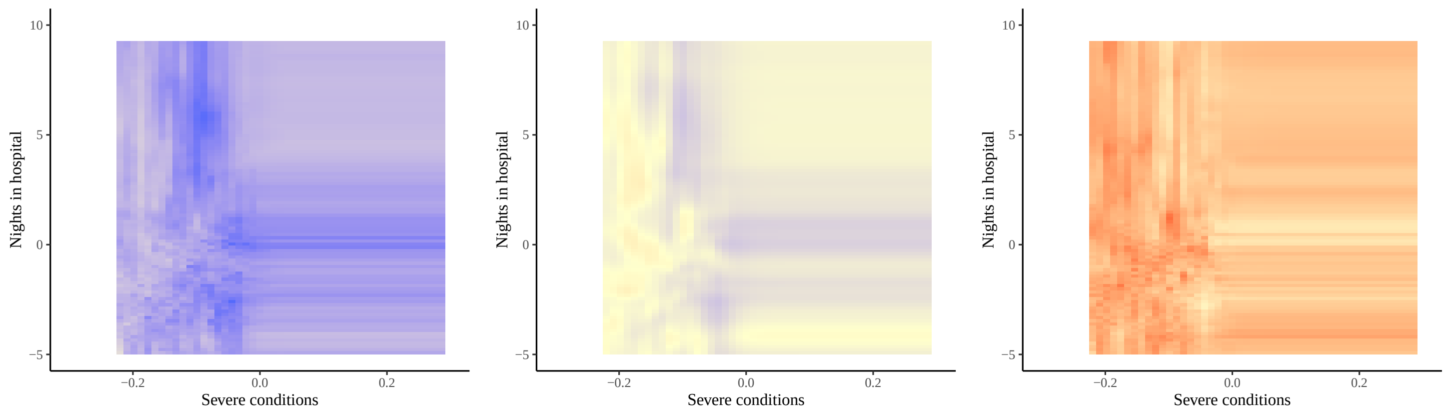
Normalized DCE on Risky Asset Share using Severe Conditions and Nights in Hospital

Panel 1: Couple Households



(A) Lower confidence band (B) Mean (C) Upper confidence band

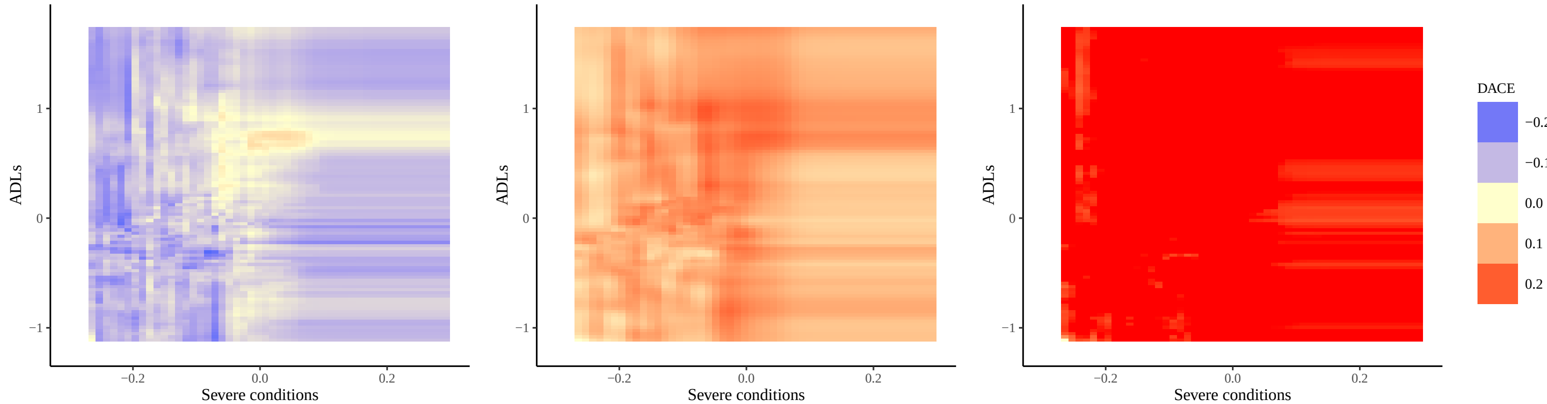
Panel 2: Single Households



(D) Lower confidence band (E) Mean (F) Upper confidence band

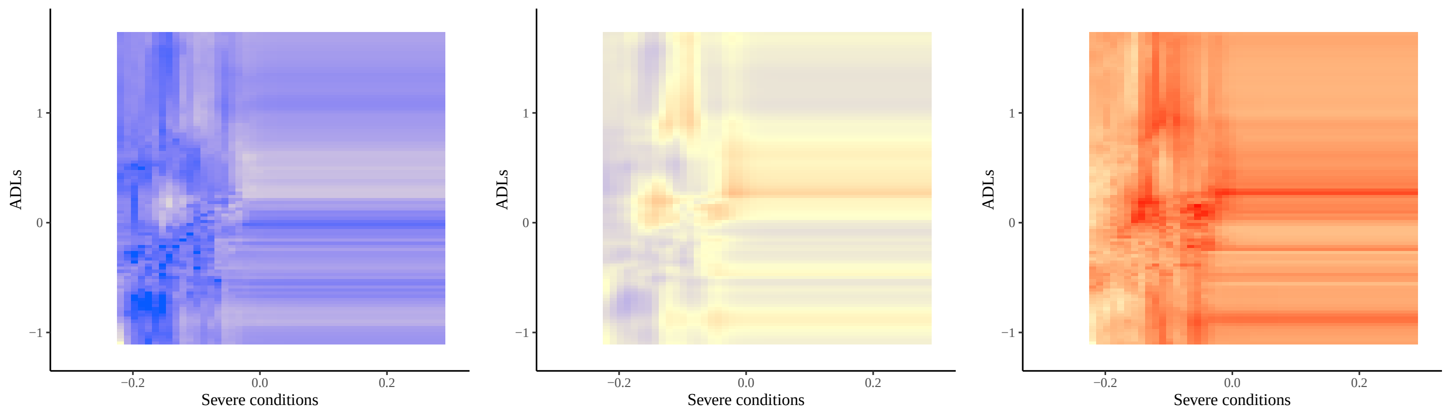
Figure S.7.4
Normalized DCE on Safe Asset Share using Severe Conditions and Activities of Daily Living

Panel 1: Couple Households



(A) Lower confidence band (B) Mean (C) Upper confidence band

Panel 2: Single Households

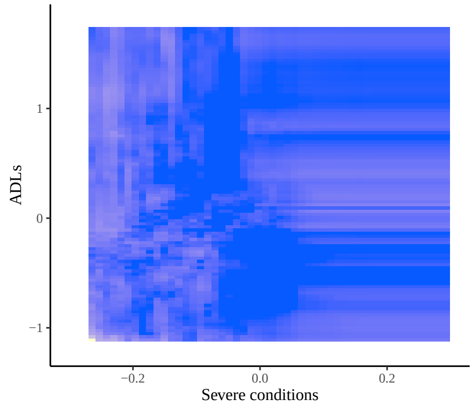


(D) Lower confidence band (E) Mean (F) Upper confidence band

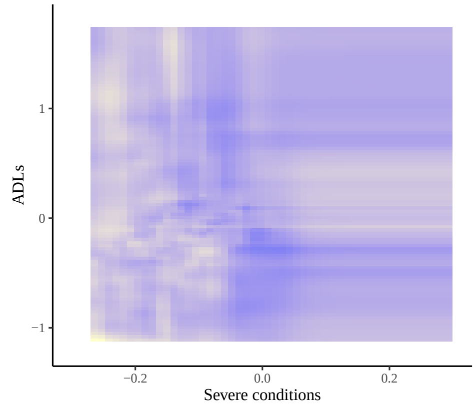
Figure S.7.5

Normalized DCE on Risky Asset Share using Severe Conditions and Activities of Daily Living

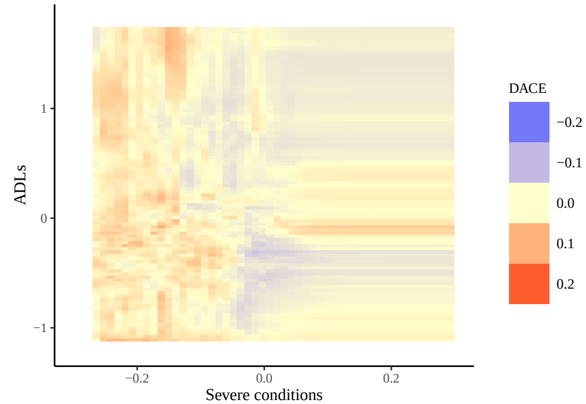
Panel 1: Couple Households



(A) Lower confidence band

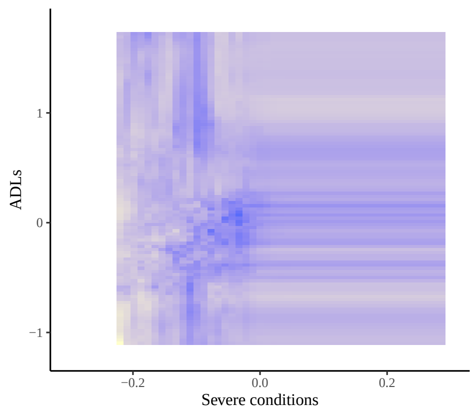


(B) Mean

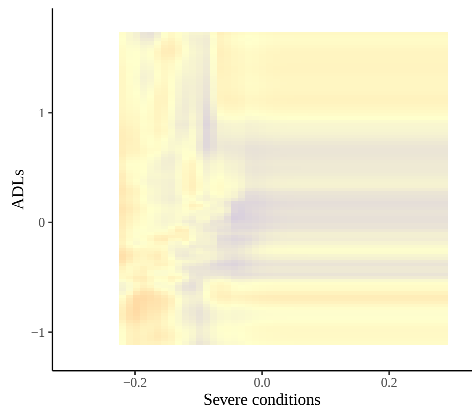


(C) Upper confidence band

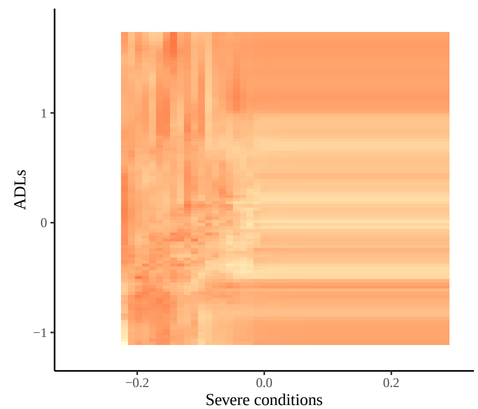
Panel 2: Single Households



(D) Lower confidence band



(E) Mean

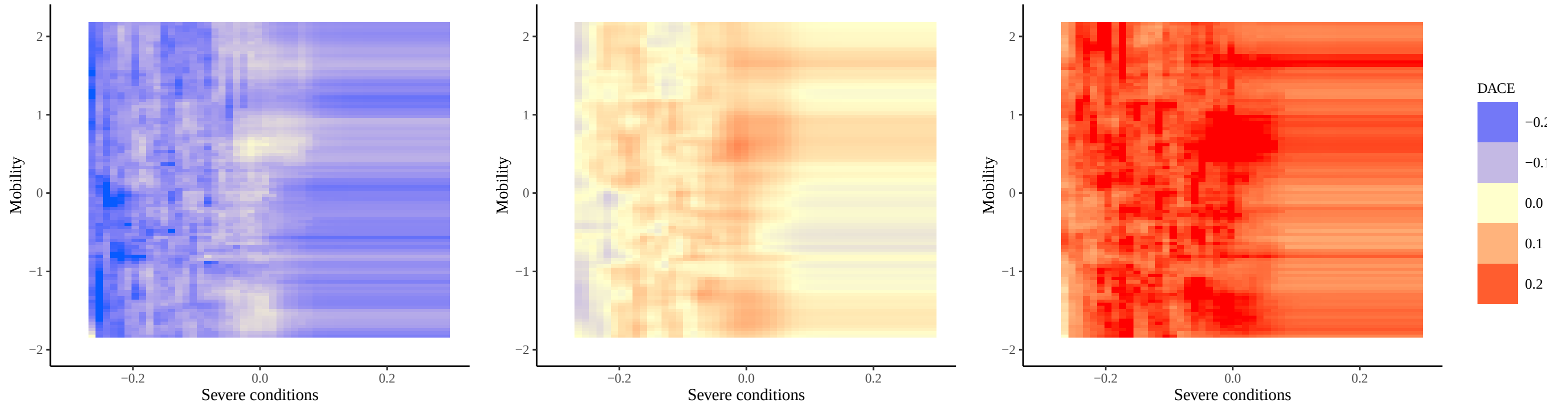


(F) Upper confidence band

Figure S.7.6

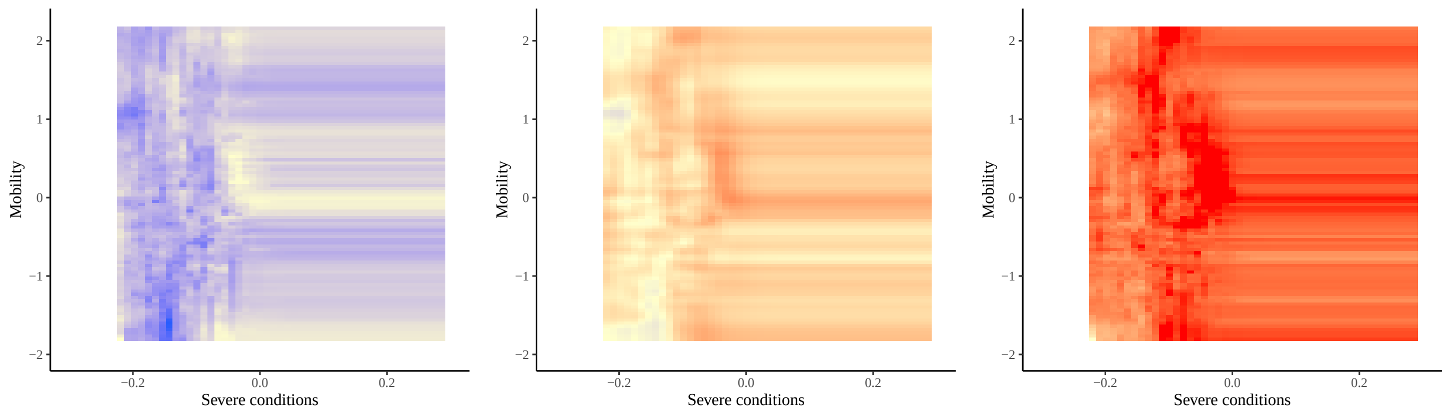
Normalized DCE on Safe Asset Share using Severe Conditions and Constraints to Mobility

Panel 1: Couple Households



(A) Lower confidence band (B) Mean (C) Upper confidence band

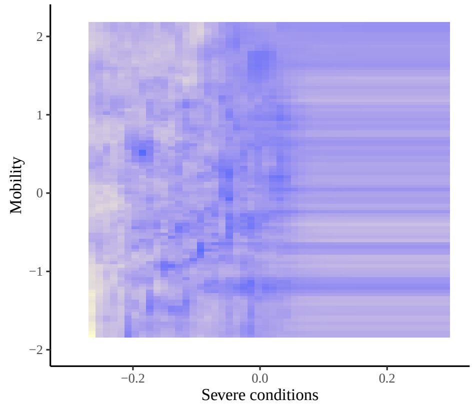
Panel 2: Single Households



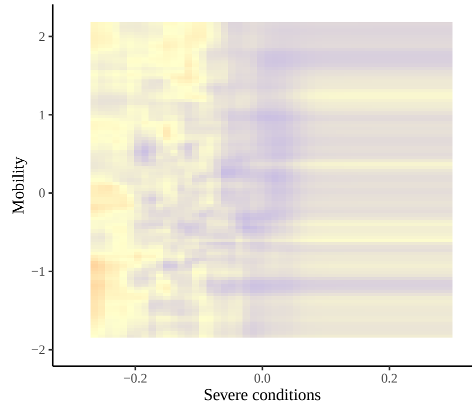
(D) Lower confidence band (E) Mean (F) Upper confidence band

Figure S.7.7
Normalized DCE on Risky Asset Share using Severe Conditions and Constraints to Mobility

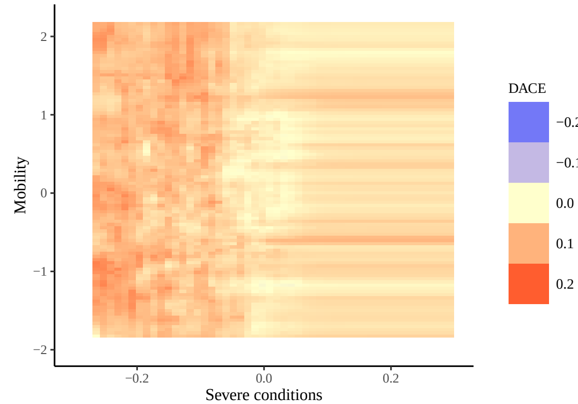
Panel 1: Couple Households



(A) Lower confidence band

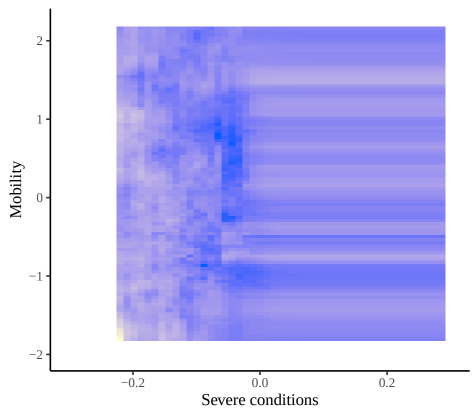


(B) Mean

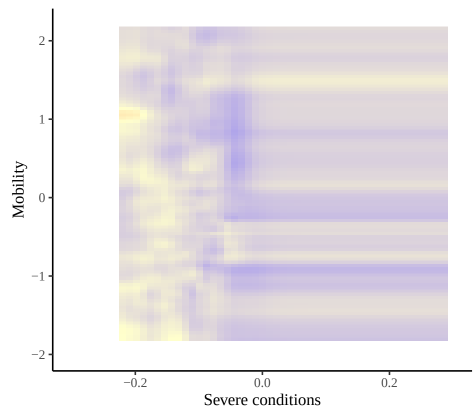


(C) Upper confidence band

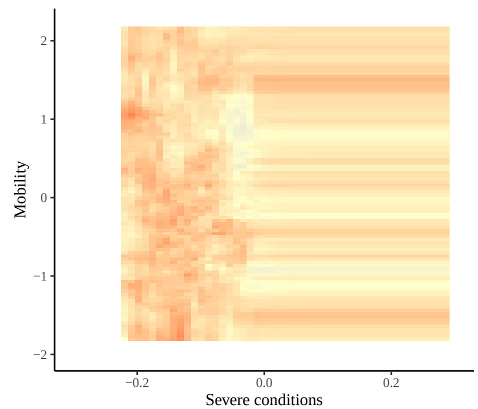
Panel 2: Single Households



(D) Lower confidence band



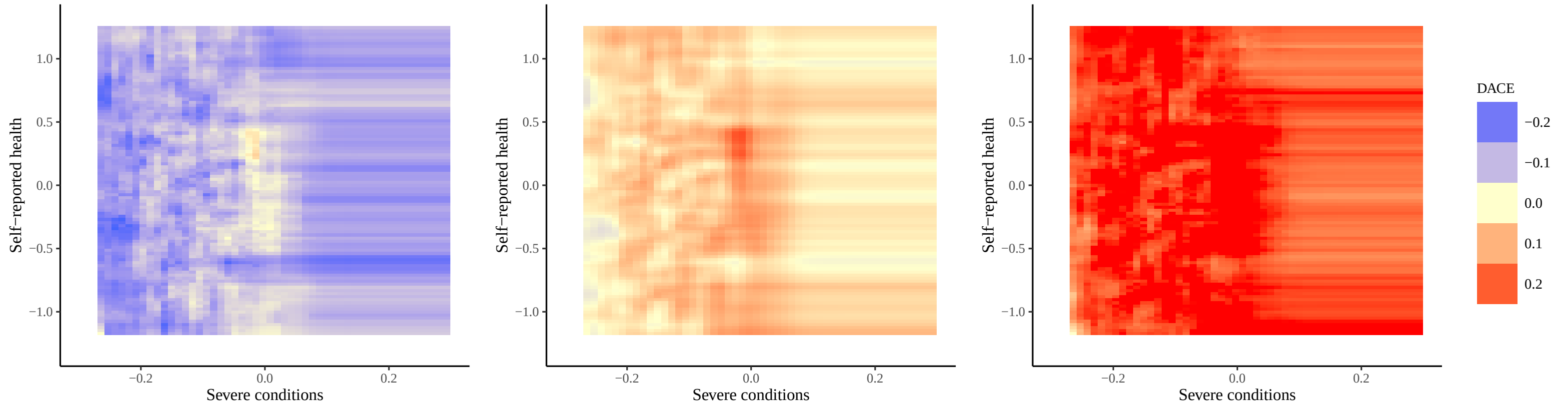
(E) Mean



(F) Upper confidence band

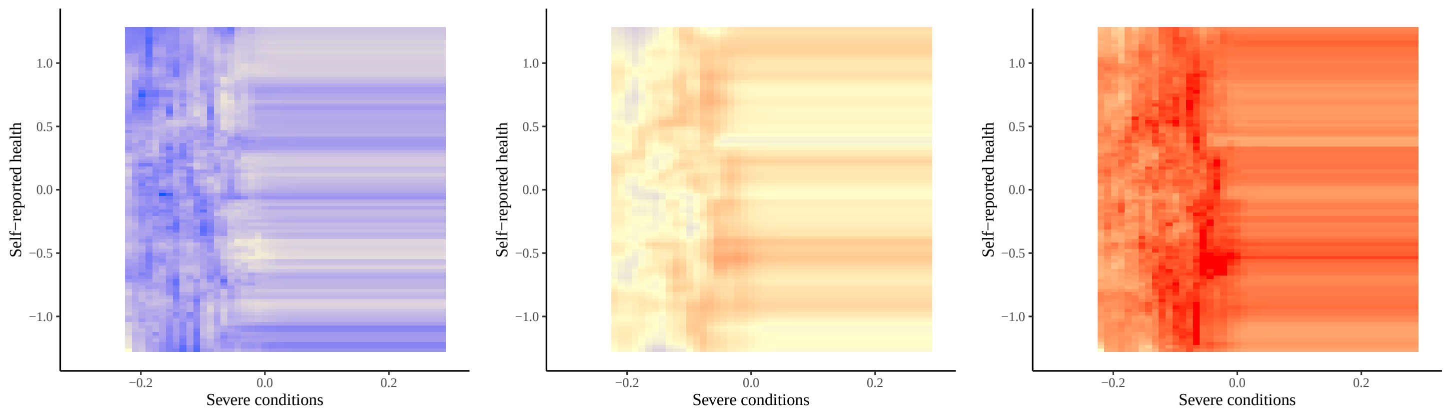
FIGURE S.7.8
Normalized DCE on Safe Asset Share using Severe Conditions and Self-Reported Health

Panel 1: Couple Households



(A) Lower confidence band (B) Mean (C) Upper confidence band

Panel 2: Single Households

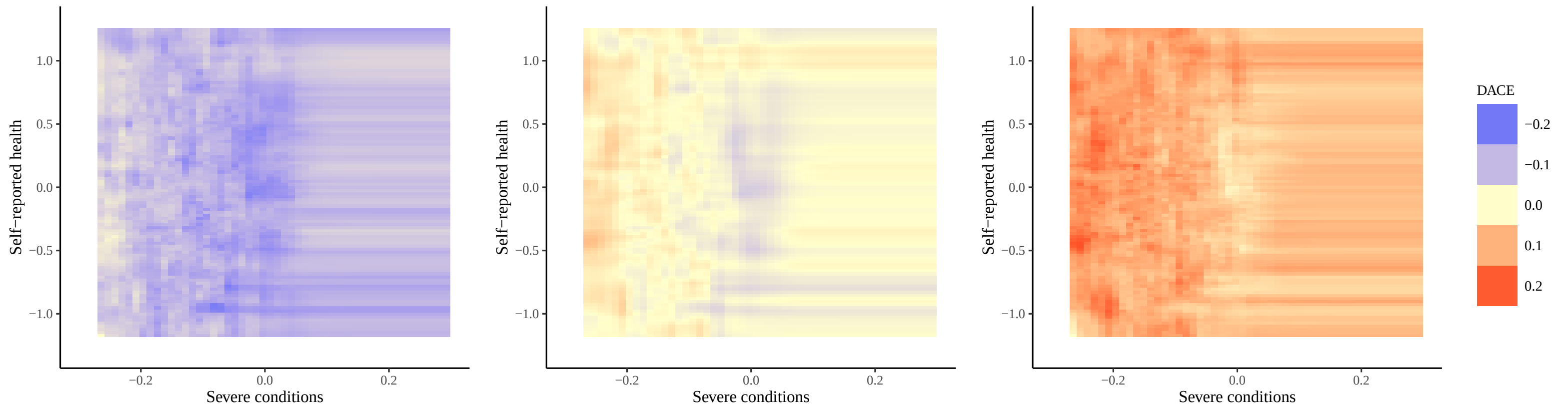


(D) Lower confidence band (E) Mean (F) Upper confidence band

Figure S.7.9

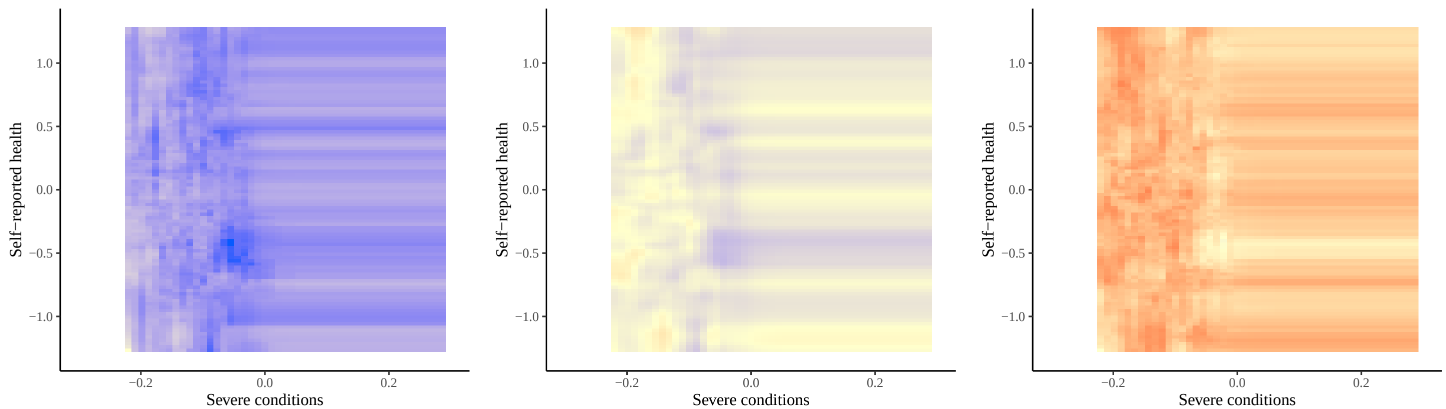
Normalized DCE on Risky Asset Share using Severe Conditions and Self-Reported Health

Panel 1: Couple Households



(A) Lower confidence band (B) Mean (C) Upper confidence band

Panel 2: Single Households



(D) Lower confidence band (E) Mean (F) Upper confidence band

S.8. Multiplicative-Interaction Estimation Results

This appendix evaluates the multiplicative-interaction estimation framework that arises under a linearity assumption on the structural function g, i.e.,

$$g(W_t, Z_{i,t}, U_{i,t}) = \beta_w W_t + \beta_z Z_{i,t} + \beta_{wz} W_t Z_{i,t} + U_{i,t},$$

where $(\beta_w, \beta_z, \beta_{wz})$ are unknown and fixed coefficients. This multiplicative interaction framework imposes strong functional assumptions that can severely limit the extent to which heterogeneous effects of the macroeconomic variable can be captured. We illustrate the substantial limitations of this approach in the context of health care policy uncertainty, highlighting the benefits of the semiparametric alternative we consider in Section 4.

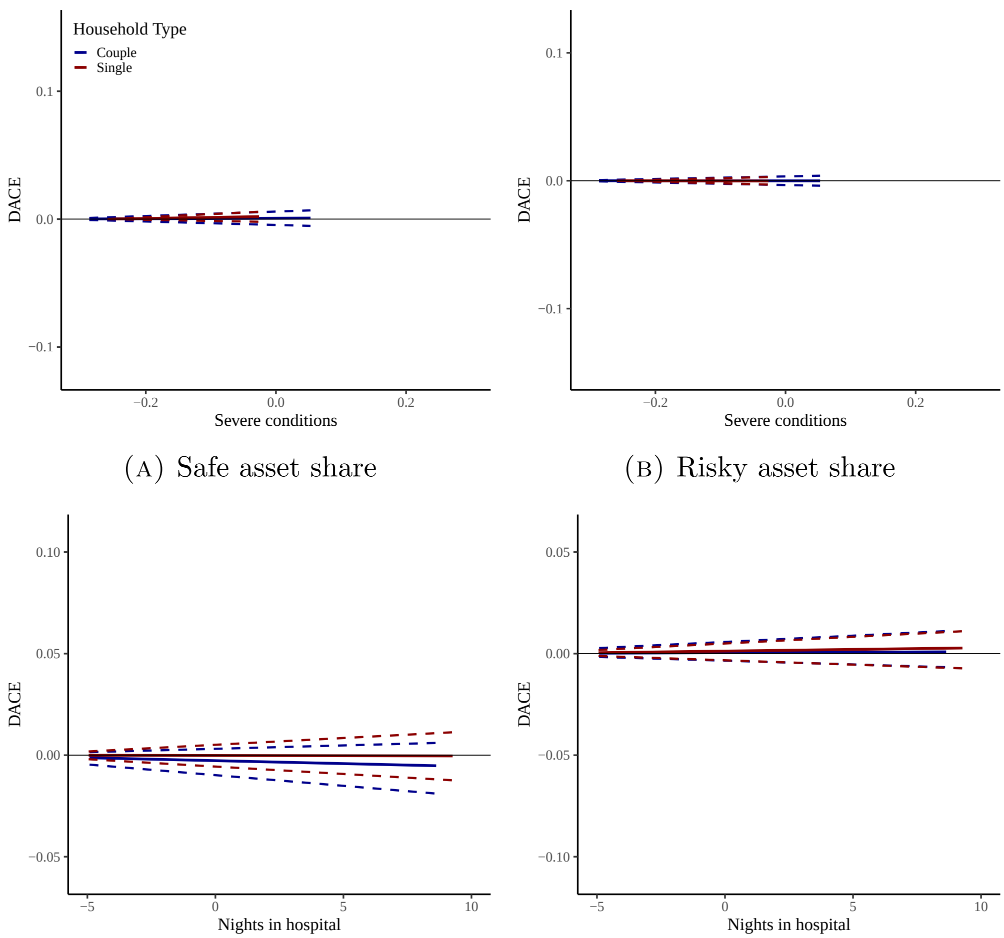
The following figures show the estimated DCE of health care policy uncertainty on safe and risky asset share, respectively, using the multiplicative interaction specification above for all five health categories considered in the paper. As before, the figures are normalized such that all estimates are relative to the subgroup-specific average causal effect of households that are substantially healthier than expected, which we define as having a more favorable unexpected change in the respective health category than 95% of households in the sample.

Figure S.8.1 presents estimates of the DCE based on unexpected changes to severe conditions (top panels) and on unexpected changes to nights spent in hospital (bottom panels) based on the multiplicative-interaction specification. The figure is thus analogous to Figure 2 with the only difference being the functional form assumption. Figure S.8.2 presents analogous estimates for the other three health categories, ADLs (top panels), mobility (center panels), and self-reported health (bottom panels).

In stark contrast to the semiparametric estimates, the restricted estimates of the DCE do not indicate substantial heterogeneity across different values in unexpected changes to health. Note further that because the bounds on the ACE are based on integrating the relative distribution of these effect-differences, the bound derived from these estimates would be close to zero and hence does not carry substantial information.

In the context of our paper, the functional form assumption imposed by a multiplicative interaction specification thus severely limits the possibility for deriving informative bounds on the average causal effect of health care policy uncertainty.

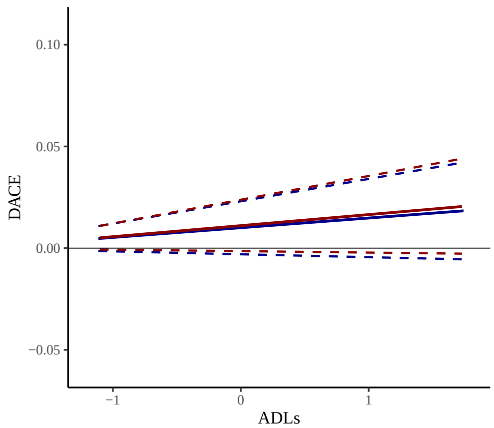
FIGURE S.8.1
Normalized Linear DCE Estimates for Couple and Single Households



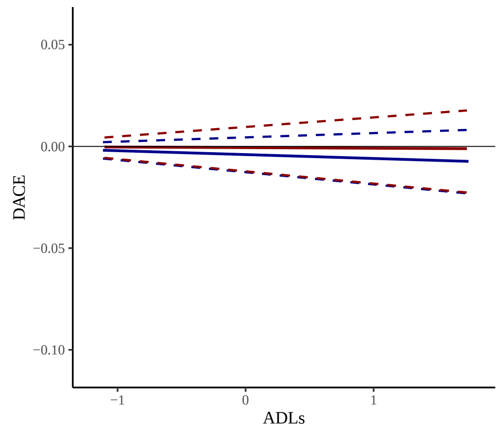
(A) Safe asset share (B) Risky asset share

(C) Safe asset share (D) Risky asset share

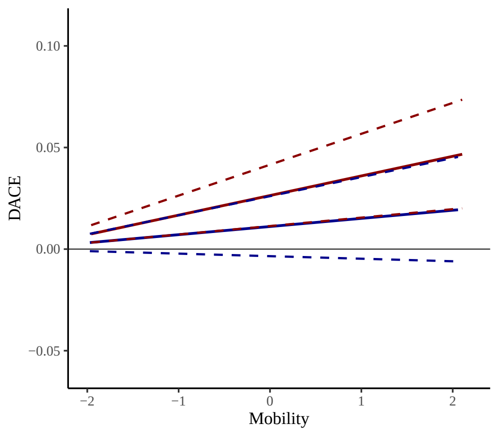
Figure S.8.2
Normalized Linear DCE Estimates for Couple and Single Households (Contd.)



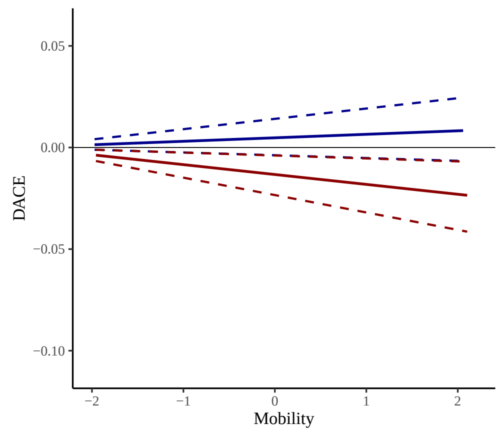
(a) Safe asset share



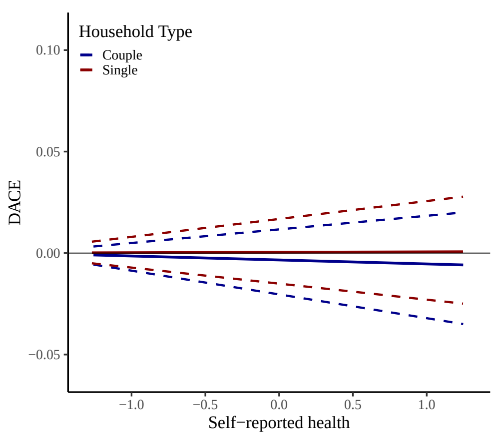
(b) Risky asset share



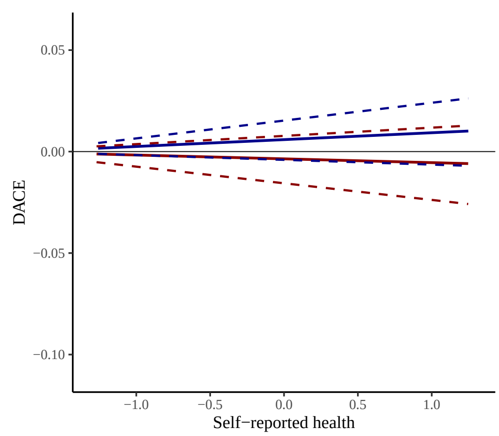
(c) Safe asset share



(d) Risky asset share



(e) Safe asset share



(f) Risky asset share

S.9. Results on Consumption

This appendix provides supplemental results on the effect of health care policy uncertainty on households' consumption expenditures.

To measure spending we use variation in households' consumption expenditures in line with Gruber and Yelowitz (1999). Data for this purpose are 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. We consider both durable and total consumption expenditures, normalized to 2010 dollars.

As with the main analysis sample, we remove those households that are missing one or more of the variables used in estimation, or only occur in a single wave, leaving an analysis sample of 12,636 single and 8,116 couple household-wave observations. Appendix S.5 provides detailed documentation of the cleaning steps and the associated reduction in sample at each stage.

We apply the same methodological approach to identify a bound on the average causal effect of health care policy uncertainty on household consumption expenditures as we did when investigating relative demand in risky assets. Notice, however, that the arguments derived from the background risk literature to motivate a bound on the conditional average causal effects on relative risky asset demand do not immediately transfer to consumption expenditures. Instead, for this application we draw from the extensive literature on precautionary savings (e.g., Zeldes, 1989; Kimball, 1990). Under the assumption of (sufficiently strong) risk aversion, theoretical insights from the precautionary savings literature imply an upper bound of zero on the conditional average causal effect of health care policy uncertainty on consumption expenditures – that is, given an increase in health care policy uncertainty, households, on average, do not increase their consumption.[4]

[4]Note that the assumptions for a precautionary savings mechanism are stronger than for a background

Table S.9.1 provides the bounds on the average causal effect of health care policy uncertainty on consumption expenditures. The effects signs are largely in line with expectations, and hint at a moderate decrease in consumption expenditures when faced with an increase in health care policy uncertainty. However, none of the estimated bounds differ from zero at a 5% significance level. Due to the substantially smaller sample size and measurement complications of the dependent variables, we caution against interpreting these findings as definitive evidence against a precautionary savings effect of health care policy uncertainty.

Table S.9.1
Bounds on the Average Causal Effect on Consumption Expenditures

Durable Consumption		Total Consumption	
Couples (1)	Singles (2)	Couples (3)	Singles (4)
-33.4	-18.5	-2493.1	562.6
$(-\infty,\ 38.2]$	$(-\infty,\ 18.2]$	$(-\infty,\ 425.8]$	$(-\infty,\ 2303.1]$

Notes. The table presents estimates of the bound on the average causal effect of health care policy uncertainty on consumption expenditure. Brackets contain one-sided confidence intervals covering 95% of bootstrapped bounds

risk mechanism. In particular, additional conditions are needed for the third derivatives of the utility function (rather than just the second). See Zeldes (1989) and Kimball (1990).

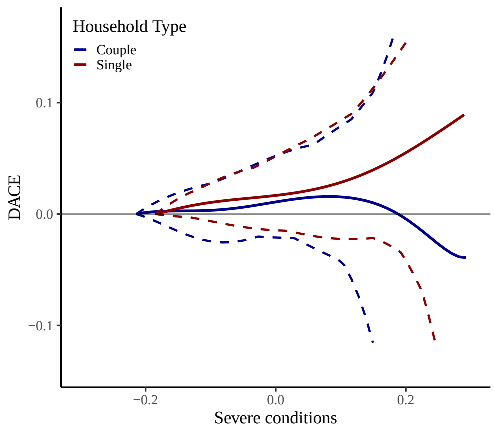
S.10. Robustness Results for Individuals below Age 65

Because most individuals above age 65 in the United States have Medicare, one might suspect that for these individuals, health care policy uncertainty would have less of an effect on their portfolio decisions than for those below age 65.[5] For robustness, in this section the semiparametric estimation results are repeated, this time restricting the estimation to a subsample of households where no household member is 65 years old or older; these are shown in Figure S.10.1 for the two health categories presented in the paper in Figure 2 (severe conditions and nights in hospital) and Figure S.10.2 for the other three health categories (ADLs, mobility, and self-reported health, analogous to Figure S.7.1).[6] The result of this restriction is a fairly large reduction in sample. For couple households, this sample restriction leaves 25,584 of the initial 57,190 observations. For single households, this sample restriction leaves 15,157 of the initial 45,384 observations. A noticeable increase in the bias and variance of the estimates is therefore expected. Nonetheless, the results are largely similar to those using the full sample. In particular, the point estimates still suggest a decreased (increased) relative demand in risky (safe) assets with worsening unexpected changes in health. The local linear regression-based instruments are not significant for most health categories.

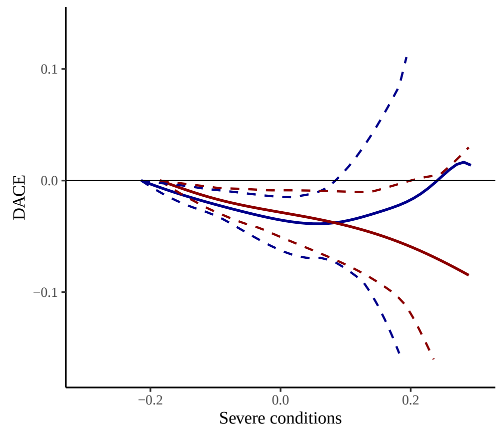
[5]Including individuals over 65 in the sample, as we do in the main paper, should, therefore, attenuate any measured effects.

[6]As before, these plots show the local linear regression-based estimates of differences in average causal effects of health care policy uncertainty on safe and risky asset share, respectively, conditional on a single health category, and the solid lines correspond to the mean and the dashed lines correspond to the bootstrapped 99% uniform confidence bands. The estimates are relative to the baseline of households that are in substantially better health than expected, defined here as having a more favorable unexpected change in the corresponding health category than 95% of households in the sample. The plots are based on $B = 1,000$ bootstrap draws.

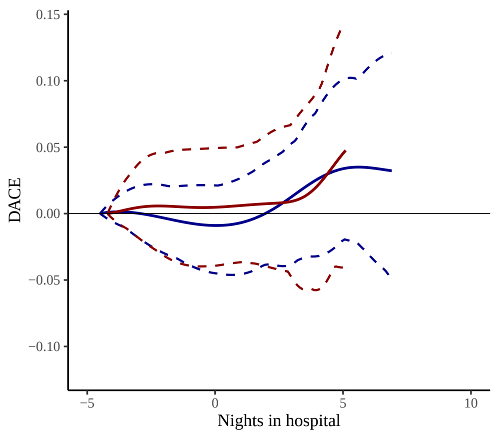
FIGURE S.10.1
Normalized DCE Estimates for Couple and Single Households (pre 65)



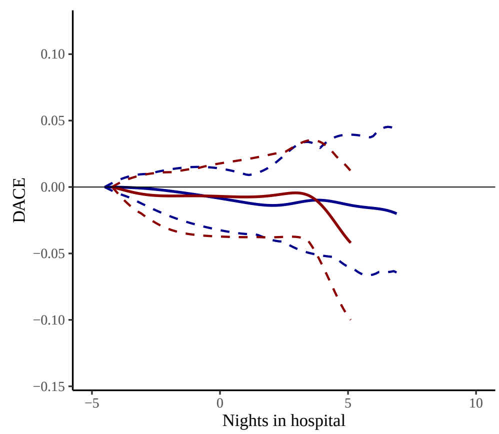
(A) Safe asset share



(B) Risky asset share



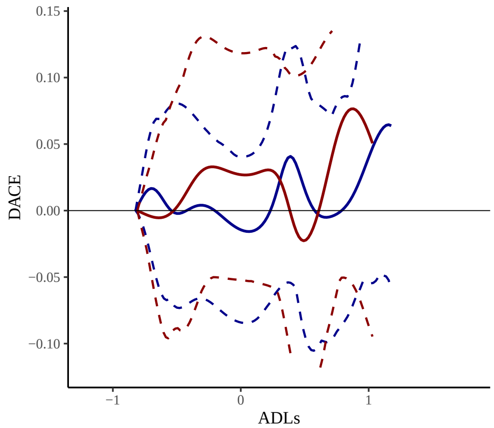
(C) Safe asset share



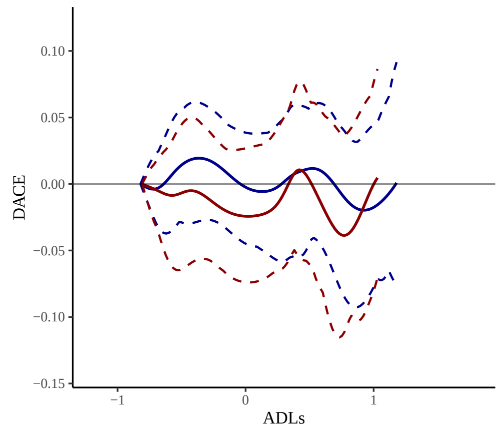
(D) Risky asset share

Figure S.10.2

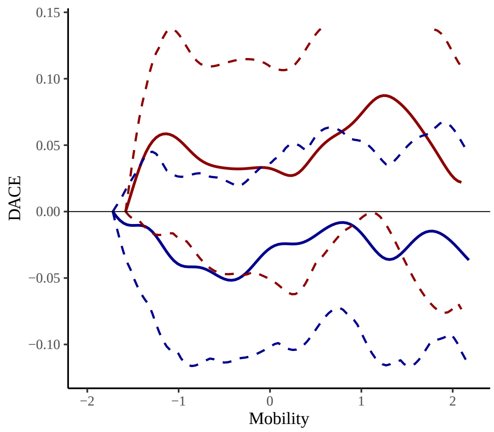
Normalized DCE Estimates for Couple and Single Households (pre 65; contd.)



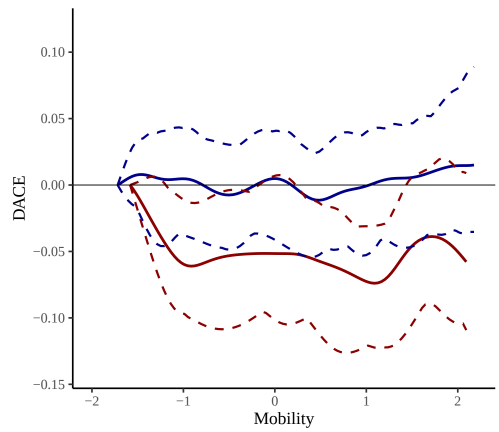
(a) Safe asset share



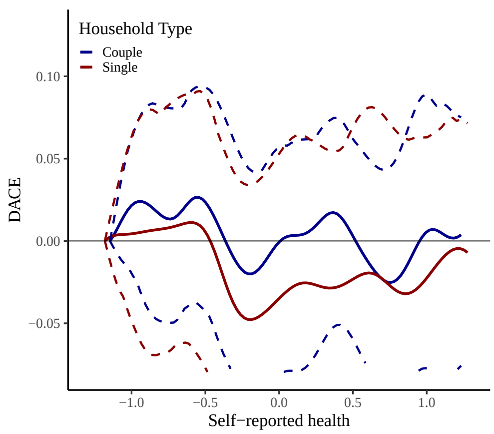
(b) Risky asset share



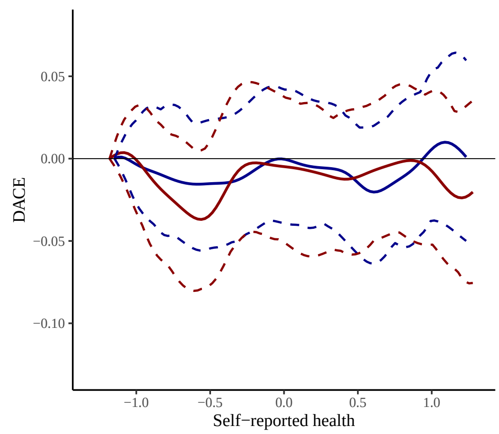
(c) Safe asset share



(d) Risky asset share



(e) Safe asset share



(f) Risky asset share