When the Going Gets Tough: the Impact of Health Shocks on Divorce*

Javier Adrián López Artero [†] Anna Sanz-de-Galdeano[‡]

Daniela Vuri [§]

Abstract

Does illness break marriages? We analyze how unexpected health shocks—sudden diagnoses of cancer, stroke, or heart attack—affect couple dissolution among older adults (50+). Using longitudinal data from the Health and Retirement Study and a quasi-experimental design comparing affected households to those experiencing similar shocks later, we find that health crises increase divorce probability by 19% relative to the baseline rate. This effect builds gradually rather than emerging immediately. We examine three explanatory channels: mental health deterioration, cognitive decline, and financial strain. The evidence indicates these factors play important roles in understanding how health shocks destabilize marriages in later life.

JEL Codes: I14, I24, J15, Z13, J13.
Keywords: Health shocks, divorce, aging, stacked diff-in-diff.

*We thank Pedro Albarrán, Marco Bertoni, Kyle Butts, Scott Cunningham, Yarine Fawaz, María Hernández-de-Benito, David Jiménez-Gómez, Melanie Guldi, Huong Nguyen, and participants at the Applied Economics Seminar Series at the University of Alicante, 2024 Codechella Madrid, the 2024 Summer Institute on Aging at the Venice International University, the American-European Health Economics Study Group - IX Edition (2025), the 2025 World SOLE-EALE-AASLE conference, the 2025 Annual Conference of the European Society for Population Economics (ESPE), the European Health Economics Association Online Seminar Series, the 47th Annual Association for Public Policy Analysis and Management Fall Research Conference, the Economics Seminar Series at the University of Granada, the 50th Simposio de la Asociación Española de Economía, and the 2026 SKILS Workshop for their comments.

Disclosure Statement: López-Artero acknowledges financial support from Project PID2024-160151NB-I00 (financed by MCIN/ AEI /10.13039/501100011033/ and by FEDER Una manera de hacer Europa) and from ICAR Foundation. Sanz-de-Galdeano acknowledges financial support from Project PID2024-160151NB-I00 (financed by MCIN/ AEI /10.13039/501100011033/ and by FEDER Una manera de hacer Europa), from Generalitat Valenciana, Consellería de Innovación, Universidades, Ciencia y Sociedad Digital through project Prometeo CIPROM/2021/068, and from ICAR Foundation. Vuri acknowledges funding from ICAR foundation. None of the authors holds any paid or unpaid positions in relevant for-profit or nonprofit organizations, and there are no disclosures concerning close relatives or partners. No external party had the right to review the manuscript prior to its circulation or submission. The study relies on existing survey data, involves no new data collection on human subjects, and therefore did not require Institutional Review Board (IRB) approval.

Data Availability Statement: The data used in this paper are available online at the official data access service (https://hrsdata.isr.umich.edu/data-products/2020-rand-hrs-fat-file).

[†]Department of Economics, University of Alicante. Email: javier.artero@ua.es.

[‡]Department of Economics, University of Alicante and IZA. Email: anna.sanzdegaldeano@gmail.com.

[§]*Corresponding author*. Department of Economics and Finance, University of Rome Tor Vergata and IZA. Email: daniela.vuri@uniroma2.it.

1 Introduction

Marital stability is a cornerstone of well-being, especially in later life, yet it may be vulnerable to disruptions caused by unexpected health shocks. With longer life expectancies, the overall duration of exposure to the risk of union dissolution among individuals increases, while also raising the prospect of health complications. Health crises, such as the sudden diagnosis of cancer, stroke, or heart attack, can impose significant emotional, cognitive, and financial stress, which in turn may destabilize long-standing partnerships. This phenomenon underscores the need to understand how adverse health events affect marital dynamics.

This paper examines the following question: How do older couples respond to health crises within their households? Do such crises strengthen their bonds, or do they increase the likelihood of couple dissolution? We provide an empirical analysis of these questions, focusing on unexpected health shocks and their impact on marital stability among individuals aged 50 and older. Using longitudinal, representative data from the Health and Retirement Study (HRS), we offer causal evidence that health shocks significantly increase the probability of couple dissolution.

While the literature extensively documents the effects of health shocks on socioeconomic outcomes such as income (Dobkin et al., 2018), wealth (Bonekamp and Wouterse, 2023), consumption (Blundell et al., 2024), out-of-pocket medical expenses (Dobkin et al., 2018; Blundell et al., 2024), residential downsizing (Costa-Font and Vilaplana-Prieto, 2022), and labor supply (Fadlon and Nielsen, 2021), much less is known about their causal effects on marital dynamics at older ages.

Building on this motivation, we focus on the phenomenon commonly referred to as "gray divorce" or "silver splitters," describing the rising divorce rates among adults aged 65 and older (Brown and Lin, 2012, 2022). At the same time, population aging is progressing rapidly across developed countries. In the United States, for example, the number of individuals aged 65 and over is projected to increase from 58 million in 2022 to 82 million by 2050, with their share of the total population rising from 17% to 23% (U.S. Census Bureau, 2023). Against this backdrop, our analysis sheds light on how unexpected health shocks shape marital dissolution at older ages, an aspect that has received comparatively little attention in the literature.

One key challenge in our study is empirically disentangling the effect of health on marital dissolution from the influence of confounding factors and potential reverse causality. To establish a causal link between household health and marital dissolution, we focus on substantial and unanticipated health shocks. In addition, we leverage the longitudinal nature of the HRS and adopt a quasi-experimental design that constructs counterfactual scenarios for affected households by comparing them with households that experience the same event in subsequent years. Because households encounter health shocks —defined as the sudden diagnosis of cancer, stroke, or heart attack— at different points in time, our research design aligns with a staggered adoption differences-in-differences (DID) framework. Traditionally, two-way fixed effects (TWFE) have been the standard method for estimating causal effects in such designs. However, recent studies (Goodman-Bacon, 2021; De Chaisemartin and d'Haultfoeuille, 2020) have highlighted validity concerns with this approach, particularly due to the inclusion of "forbidden comparisons," that is, the use of already-treated units as the control group for units that are just being treated or will be treated later. Since then, new methods have been proposed and applied to enable causal inference in staggered adoption designs (Cengiz et al., 2019; Goodman-Bacon, 2021; De Chaisemartin and d'Haultfoeuille, 2020; Butters et al., 2022; Borusyak et al., 2024; Callaway and Sant'Anna, 2021). The stacked DID is one such approach and involves the following steps: i) creating a separate dataset for each valid sub-experiment that avoids "forbidden comparisons," ii) combining these datasets into a single stacked file, and iii) estimating the average causal effect by applying DID or event study regressions to the stacked dataset. We utilize the stacked DID estimator recently introduced by Wing et al. (2024a) because it offers several advantages beyond avoiding forbidden comparisons: i) they propose a trimming rule to ensure balance in the number of pre- and post- periods for each sub experiment; this rule can be applied to staggered adoption designs to ensure that the average aggregate treatment effect on the treated (ATT) does not suffer from compositional bias; ii) they derive sample weights to correct for differential weighting bias (treatment and control trends are implicitly weighted differently across sub-experiments) and show that a simple weighted least squares estimator identifies a well defined causal effect that they call "the trimmed ATT".

Our analysis reveals several key findings. First, we find that experiencing a health shock significantly increases the probability of couple dissolution by approximately 19% of the mean divorce prevalence. This effect builds gradually over time rather than manifesting immediately after the adverse health event. Additionally, we examine several mechanisms through which health shocks may influence divorce, focusing on three potential channels: mental health deterioration (measured by the CES-Depression scale), cognitive decline (measured by symptoms of cognitive impairment), and financial strain (measured by out-of-pocket medical expenses). Our findings are consistent with these channels playing an important role in explaining why divorce risk rises following a health shock. Our results remain robust across multiple robustness checks, including extending the event window length and employing an alternative health shock definition.

Our paper contributes to a broad literature examining the consequences of health shocks within families (Fadlon and Nielsen, 2019; Blundell et al., 2024). For instance, some studies analyze the effects of children's health shocks on their parents' labor market outcomes (Breivik and Costa-Ramón, 2024) and vice versa (Brito and Contreras, 2023), while others explore the spillover effects of health shocks between spouses (Fadlon and Nielsen, 2021; Riekhoff and Vaalavuo, 2021; Angelini and Costa-Font, 2023; Du and Zaremba, 2024; Arteaga et al., 2024). In particular, this paper contributes directly to the literature that analyzes the effect of poor health or negative health shocks on partner stability. Previous evidence has documented associations between poor mental health and marital instability (Merikangas, 1984; Pevalin and Ermisch, 2004), negative links between self-reported health and marital quality (Booth and Johnson, 1994; Joung et al., 1998), a connection between mental disorders and divorce (Kessler et al., 1998), between post-traumatic stress and marital stability (Negrusa and Negrusa, 2014;

Tauchmann et al., 2023), and between perceived poor self-rated health or activity limitations and marital dissolution (Vignoli et al., 2025). However, most of these studies tend to be descriptive in nature, are unable to offer a causal interpretation of the effect, and/or refer to very specific groups of the population.

Moreover, our paper delves deeper into the potential mechanisms that may drive the impact of health shocks on marital stability. Carr and Springer (2010) highlight that chronic illness can lead to both emotional strain and increased dependency, creating potential tensions within long-term partnerships. Umberson et al. (2011) explore how mental health declines can negatively impact spousal interactions, leading to emotional withdrawal and reduced relationship satisfaction. By examining mental health deterioration, cognitive decline, and financial strain as potential channels linking health shocks to divorce, our study offers a more comprehensive perspective on how adverse health events shape marital outcomes.

Finally, our paper contributes to the literature on the aging population and family demographics (Fuster, 2017; Roberto and Blieszner, 2015). Existing research has primarily focused on relatively young populations, but the impact of poor health on the risk of divorce may become more pronounced as individuals age (Uhlenberg and Myers, 1981). During late middle age and early older adulthood, people often begin to experience the onset of serious health problems. As divorce has become more socially acceptable and divorce rates among older adults continue to rise, understanding the role of health crises in contributing to this trend becomes increasingly important for policymakers and social scientists alike. By shedding light on the interplay between health crises and marital stability, this study contributes to a better understanding of the broader challenges associated with population aging (Cherlin, 2010; Lin et al., 2022). It highlights the need for policymakers to consider the multifaceted impacts of health shocks — not only on individual well-being but also on household dynamics — when designing interventions to support aging populations.

2 Data and Descriptive Evidence

2.1 Database and Sample Selection

This study uses the RAND Health and Retirement Study (HRS) longitudinal file, a harmonized version of the HRS developed by the RAND corporation based on the HRS core interviews.[1] The HRS is an ongoing, nationally-representative, panel study of Americans aged 50 and above, collecting detailed information biennially from 1992 to 2020. It provides information on various domains including demographics, income, health status, insurance coverage, social security benefits, pension plans, and family structure, for both respondents and their spouses, regardless of the spouse's age.

The dataset incorporates 7 distinct entry cohorts, each entering the survey at different waves. The original HRS cohort, comprising individuals born between 1931 and 1941, was first interviewed in 1992 and potentially observed for 15 waves. Subsequent cohorts joined the study as follows: the AHEAD (Assets and Health Dynamics) cohort, born before 1924, the Children of Depression (CODA) cohort, born 1924-1930, and the War Baby (WB) cohort, born 1942-1947, entered in 1998 in wave 4. The Early Baby Boomer (EBB) cohort, born 1948-1953, joined in wave 7 in 2004. The Middle Baby Boomer (MBB) cohort, born 1954-1959, entered in 2010 in wave 10. The Late Baby Boomer (LBB) cohort, born 1960-1965, began participation in 2016 in wave 13. Details on the different cohorts and their entry waves are provided in Table 1.

For analytical purposes, we transform the RAND longitudinal file into a panel dataset, with each observation representing a household at a specific wave. To maintain household-level analysis while avoiding duplicate entries, we retain single observations per household per wave.

Our sample is restricted to individuals who were married or partnered at the time they were first observed in the study. To examine the impact of the onset of unexpected

[1] The HRS is sponsored by the National Institute on Aging (grant number NIA U01AG009740) and is conducted by the University of Michigan. The RAND HRS Longitudinal File is an easy-to-use dataset based on the HRS core data. This file was developed at RAND with funding from the National Institute on Aging and the Social Security Administration.

Table 1: **Cohort distribution**

Wave/Cohorts	**HRS**	**AHEAD**	**CODA**	**WB**	**EBB**	**MBB**	**LBB**
Birth Years	1931-1941	Before 1924	1924-1930	1942-1947	1948-1953	1954-1959	1960-1965
1	1992						
2	1994						
3	1996						
4	1998	1998	1998	1998			
5	2000	2000	2000	2000			
6	2002	2002	2002	2002			
7	2004	2004	2004	2004	2004		
8	2006	2006	2006	2006	2006		
9	2008	2008	2008	2008	2008		
10	2010	2010	2010	2010	2010	2010	
11	2012	2012	2012	2012	2012	2012	
12	2014	2014	2014	2014	2014	2014	
13	2016	2016	2016	2016	2016	2016	2016
14	2018	2018	2018	2018	2018	2018	2018
15	2020	2020	2020	2020	2020	2020	2020

Notes: Distribution of waves and cohorts in the Health and Retirement Study (HRS). The birth years and waves for each cohort are listed, indicating when each cohort entered the study.

and serious physical illness on marital dissolution, we exclude couples in which either spouse reported having been diagnosed with cancer, stroke, or a heart attack at baseline, as further discussed in Section 2.2. Additionally, we exclude couples with missing information prior to marital dissolution, as these gaps prevent accurate identification of the timing of potential health shocks. Given the limited occurrence of divorce among elderly participants, households where both partners are 75 or older at first observation are omitted from the analysis, as well as same-sex marriages due to small sample size.

2.2 Treatment and Outcome Definitions

We use two different measures of health shocks. The first measure, used in our main specification, refers to the onset of three major life-threatening illnesses: cancer, stroke, or heart attack to any member of the couple, provided these conditions are diagnosed by a medical professional since the last interview wave. These three conditions are often used in the health shocks literature due to their unpredictable nature (Smith, 1999, 2005; Trevisan and Zantomio, 2016; Thompson and Conley, 2016), making their onset difficult to anticipate. Moreover, due to their severity as significant health conditions, they are less likely to be subject to misreporting or justification bias compared to milder ailments (Jones et al., 2020).

The second measure, used as a robustness check, relies on hospitalization data from the HRS. This measure aggregates the total nights of hospitalization for both spouses within each household. A health shock is defined as occurring when a household's total hospital nights exceed four (the median value among hospitalized households) in a given wave, conditional on the household reporting no hospitalizations in the preceding wave.

When a household experiences a health shock, the treatment variable is set to 1 from that wave onward. The variable remains 0 if the household stays healthy, as the treatment represents an absorbing state.[2]

The main dependent variable in our analysis captures the transition from couple to non-couple status between successive waves for each household. For simplicity, throughout the paper, we use "marriages" to refer to individuals who are married or declare themselves to be in a couple, and "divorces" to refer to individuals who are separated or divorced. A couple is defined as dissolving in separation/divorce if either spouse reports being divorced or separated since the prior wave.[3] The dependent variable is coded as 0 if the original marriage or partnership remains intact across consecutive waves, and as 1 upon couple dissolution. The divorce prevalence we observe in our analytic sample (6.3%) is consistent with findings from other studies utilizing HRS data (Karraker and Latham, 2015b).

2.3 Summary Statistics

In columns (1) and (2) of Table 2, we present the summary statistics of the main analytic sample that serves as the basis for implementing the Stacked Difference-in-

[2]Our analysis focuses on the impact of the first health shock experienced by a household. If a household experiences multiple health shocks over time (e.g., one spouse experiences a major health crisis followed by the other spouse experiencing a different health shock), we code the treatment based on the timing of the initial shock.

[3]A couple is also considered dissolved due to separation or divorce in the rare case where either partner divorced and remarried between consecutive waves without reporting the divorce—indicated by an increase in either partner's total number of marriages.

Differences estimator described in Section 3.1. As expected, the mean age of wives is lower than that of husbands, reflecting the historical practice of older men marrying younger women (Berardo et al., 1993). Wives are also less likely to have completed college and more likely to have completed only high school. We observe little difference in the rest of the covariates (religion, ethnicity and residence), which is expected due to assortative mating.

As explained in Section 2.2, we begin with a sample of 9,947 married or partnered couples observed at the time of their initial inclusion in the study. We exclude same-sex couples, couples with missing data on relevant variables, couples in which either spouse reported a diagnosis of cancer, stroke, or heart attack at baseline, and couples in which both partners were aged 75 or older at first observation. After these exclusions, our final analytical sample consists of 5,054 couples.

While these selection criteria strengthen our ability to identify the causal relationship between illness onset and marital dissolution, we acknowledge that they may affect the generalizability of our results. To assess the extent of this potential issue, Table 2 compares the background characteristics of the initial sample (columns 3 and 4) with those of the final analytical sample (columns 1 and 2). Reassuringly, the reported statistics indicate that the two samples are largely comparable.

3 Identification Strategy

3.1 Methodology

To examine the causal effect of an unexpected health shock experienced by either spouse in a household on the subsequent likelihood of divorce, we employ a Stacked Difference-in-Differences (Stacked Diff-in-Diff) estimator (Wing et al., 2024a,b). Indeed, as shown by Goodman-Bacon (2021) and De Chaisemartin and d'Haultfoeuille (2020), results from two-way fixed effects difference-in-differences (TWFE) estimators do not have a meaningful causal interpretation when there is staggered treatment adoption (i.e. treatment timing varies across units) and the treatment effect is not constant over time.

Table 2: Summary Statistics

	After Sample Selection		Before Sample Selection	
	Mean	SD	Mean	SD
	(1)	(2)	(3)	(4)
Panel A: Household				
Household region: Midwest	0.25	0.43	0.25	0.43
Household region: Northeast	0.14	0.35	0.14	0.35
Household region: West	0.20	0.40	0.19	0.39
Household region: South	0.41	0.49	0.42	0.49
Household urban area	0.48	0.50	0.47	0.50
Household suburban area	0.22	0.41	0.23	0.42
Household exurban area	0.30	0.46	0.31	0.46
Household number of children	3.25	1.91	3.32	1.99
Panel B: Husband				
Husband Age	64.81	9.34	66.59	10.39
Husband non white race	0.18	0.39	0.18	0.38
Husband born before 1931	0.15	0.36	0.27	0.45
Husband born 1931–1942	0.46	0.50	0.41	0.49
Husband born 1942–1953	0.24	0.43	0.19	0.40
Husband born 1954–1965	0.14	0.34	0.11	0.31
Husband born after 1965	0.01	0.09	0.01	0.08
Husband years of education	12.68	3.47	12.39	3.51
Husband working ft/pt/unemployed	0.45	0.50	0.39	0.49
Husband Protestant	0.60	0.49	0.61	0.49
Husband Catholic	0.28	0.45	0.27	0.44
Husband Jewish	0.02	0.12	0.02	0.13
Husband No religion	0.09	0.28	0.09	0.28
Husband Non-US	0.13	0.33	0.11	0.31
Husband Less than High school completed	0.25	0.43	0.29	0.45
Husband High school completed	0.47	0.50	0.47	0.50
Husband College completed	0.27	0.45	0.24	0.42
Panel C: Wife				
Wife Age	62.49	10.04	64.83	11.38
Wife non white race	0.18	0.38	0.17	0.38
Wife born before 1931	0.06	0.23	0.17	0.38
Wife born 1931–1942	0.40	0.49	0.38	0.49
Wife born 1942–1953	0.34	0.47	0.29	0.45
Wife born 1954–1965	0.17	0.38	0.14	0.35
Wife born after 1965	0.03	0.17	0.02	0.14
Wife years of education	12.67	3.08	12.45	3.00
Wife working ft/pt/unemployed	0.39	0.49	0.33	0.47
Wife Protestant	0.62	0.49	0.63	0.48
Wife Catholic	0.29	0.46	0.28	0.45
Wife Jewish	0.02	0.12	0.02	0.14
Wife No religion	0.06	0.23	0.05	0.22
Wife Non-US	0.14	0.34	0.12	0.32
Wife Less than High school completed	0.21	0.41	0.24	0.43
Wife High school completed	0.58	0.49	0.58	0.49
Wife College completed	0.22	0.41	0.18	0.39
N	5054		9947	

Notes: Summary statistics are presented for two samples: couples observed at baseline (columns 3–4) and the analytical sample (columns 1–2). The analytical sample further excludes same-sex couples, couples with missing information on relevant variables, couples in which either spouse reported a diagnosis of cancer, stroke, or heart attack at baseline, and couples in which both partners were 75 or older at the time of first observation.

In this scenario, bias may arise from using "already treated" units as controls,[4] a practice referred to in the recent literature as "forbidden comparisons". Of course, this would no longer be a concern if the treatment date were the same for all units (as there would be no "forbidden comparisons") or if the treatment impact were constant over time. In our case, treatment adoption is staggered, as households may experience a health shock (treatment) in different waves. Similarly, the assumption that the treatment effect is instantaneous is unrealistic, as households' potential responses to a health shock may take time to manifest.

To address this issue, we use a Stacked Diff-in-Diff estimator that eliminates "forbidden comparisons" by creating distinct "sub-experiments". Each sub-experiment represents a Diff-in-Diff comparison where all treated units (households) receive the treatment in the same wave. After determining the length of the event window, creating each sub-experiment involves selecting appropriate control units (e.g., only units that never received the treatment, or only units that received the treatment but outside the event window of the sub-experiment, or a combination of both types of units), ensuring that "forbidden comparisons" are avoided by construction. Lastly, all the sub-experiments are combined into a "stacked" dataset, which is then used to compare treated households with those that will undergo the treatment in the future ("not yet treated" households). Wing et al. (2024a) point out that the Stacked Diff-in-Diff estimator only recovers the target parameter (the ATT) when the sample is balanced in each sub-experiment and corrective sample weights are used.[5] Hence, we conduct our benchmark analysis both with and without corrective weights, with the former being our preferred specification.

[4]This refers to comparing units treated at different times, where the earlier-treated group serves as a control after its treatment starts.

[5]Corrective sample weights for the trimmed aggregate ATT are equal to 1 for the treated group and for the control group they are constructed as follows: $\frac{N_a^D/N^D}{N_a^C/N^C}$ where N_a^D is the number of groups that first adopt treatment in period of time a, N^D is the total number of groups that adopt treatment at any of the times included in the trimmed set, N_a^C and N^C give the analogous counts for the control groups.

3.2 Balancing Tests

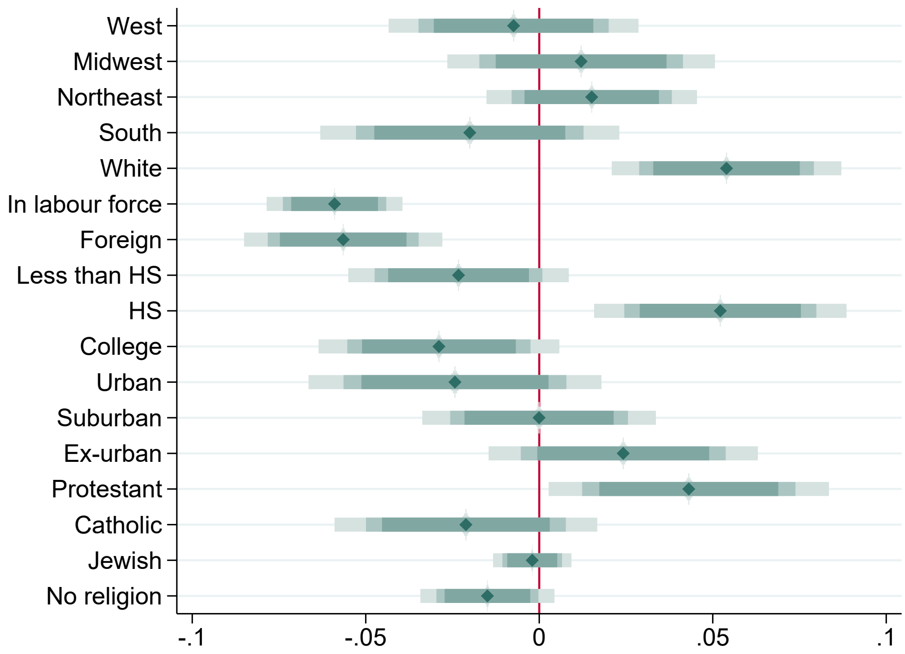
To construct the stacked dataset, it is crucial to identify the appropriate control units, wherein control units include all individuals who experience the treatment outside the event window of a sub-experiment, as well as those who never experience a health shock (Miller, 2023). Alternatively, the sample can be restricted to households that experience a health shock at some point (Costa-Font and Vilaplana-Prieto, 2022; Fadlon and Nielsen, 2021).

To guide the decision on which control group to select, we compare the background characteristics of individuals in households that experience a health shock with those in households that do not (Figure 1). Then, we evaluate the similarity among individuals in households that experience a health shock at different times (Figure 2).

Figure 1 presents the coefficient estimates and confidence intervals from regressions of several background characteristics on an "ever treated" dummy variable, which equals 1 if any spouse reports having experienced a health shock and 0 otherwise. The regressions include controls for age and cohort dummies, with standard errors clustered at the household level. To facilitate the interpretation of the estimated coefficients' magnitudes, continuous background characteristics are standardized to have a mean of 0 and a standard deviation of 1.

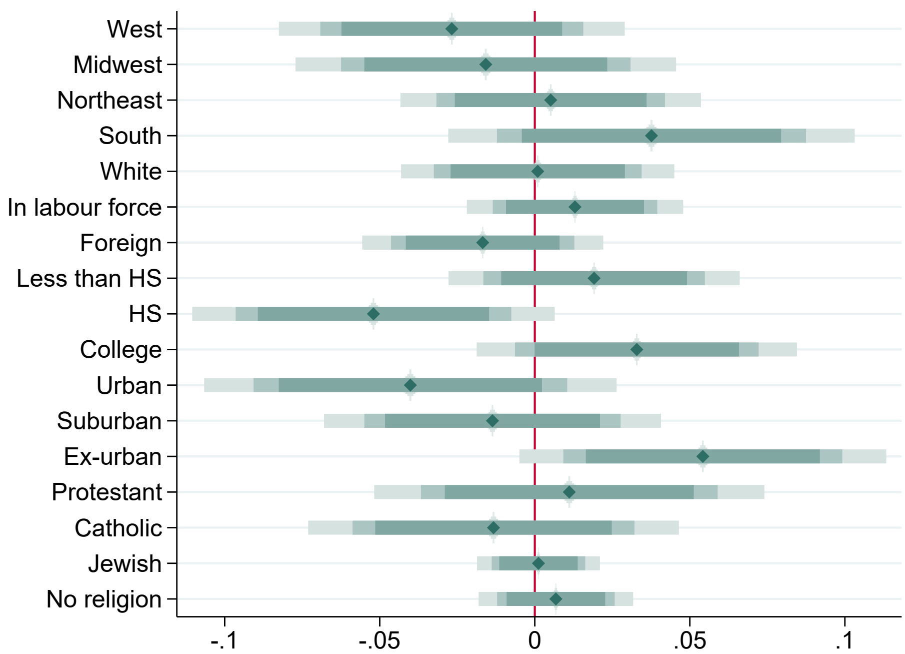
Figure 2, based on our stacked dataset, is analogous to Figure 1, with one key distinction: it reports coefficient estimates for a dummy variable that equals 1 if the individual belongs to a household treated in a particular sub-experiment and 0 otherwise. A visual comparison of Figures 2 and 1 reveals that individuals from households experiencing a health shock at different time points (Figure 2) share more similar characteristics than those from households that never experience a health shock versus those that do (Figure 1). For instance, individuals in households that never experience a health shock are significantly more likely to participate in the labor force, to have attended college, and are less likely to have at most a high school education compared to their counterparts in households where a health shock occurs.

Figure 1: Differences in Characteristics across Households with and without Health Shocks.



Note: The figure shows the coefficients and 95% CI from separate regressions of each variable on an indicator that takes a value of 1 if the household ever suffered a health shock and 0 otherwise, controlling for age and cohort dummies.

Figure 2: Differences in Characteristics within Households that Experience a Health Shock



Note: The figure shows the coefficients and 95% CI from separate regressions of each variable on an indicator that takes the value 1 if the household is in the treatment group in a particular sub-experiment, and 0 otherwise, controlling for age and cohort dummies.

Conversely, Figure 2 shows no statistically significant differences across units that experienced a health shock, using as controls households that received the treatment at a later point. Based on this result, we exclude from our sample households that never experienced a health shock and instead focus on leveraging its timing.

3.3 Stacked Diff-in-Diff Dataset Construction

To construct the stacked dataset, one must make several key decisions. The first involves determining the length of the sub-experiment, that is, how many waves (observations) before and after the adoption of the treatment are included in each sub-experiment. The length of the event window needs to be constant across the different sub-experiments in order to ensure compositional balance (Wing et al., 2024a). We choose to use three waves before and after the health shocks in our benchmark specification and in our analyses of mechanisms. This event window allows us to assess the plausibility of the parallel trends assumption and to observe the evolution of the impact of health shocks over time, while maximizing the number of sub-experiments and households in our sample. In Section 6 we show that our results are robust to increase the event window to four waves before and after the treatment.

Table 3 illustrates how sub-experiments are defined, how many sub-experiments are available in our sample, the treatment period for each sub-experiment (which is the same for all households within a sub-experiment), the periods included in each sub-experiment, and how control units are chosen.

Table 3: **Sub-experiment structure**

Sub-experiment (1)	**First Wave (2)**	**Treated Wave (3)**	**Last Wave (4)**	**Waves in which Control HHs Suffer a Health Shock (5)**
1	1	4	7	8,9,10,11,12,13,14,15
2	2	5	8	9,10,11,12,13,14,15
3	3	6	9	10,11,12,13,14,15
4	4	7	10	11,12,13,14,15
5	5	8	11	12,13,14,15
6	6	9	12	13,14,15
7	7	10	13	14,15
8	8	11	14	15

Note: Column (1) reports the number of sub-experiments available. Columns (2) and (4) indicate the first and last waves observed in each sub-experiment, respectively. Column (3) identifies the wave in which treated households receive the treatment, while Column (5) specifies the wave in which control households undergo the treatment within each sub-experiment.

Specifically, Column 1 lists the number of each sub-experiment, aiming to show how many sub-experiments comprise our final stacked dataset. Columns 2 and 4 indicate the first and last waves included in each sub-experiment, respectively. Column 3 specifies the wave in which all treated households in each sub-experiment experience a health shock, which is the same for all households within a sub-experiment. Lastly, Column 5 identifies the waves in which control households experience a health shock within each sub-experiment. Importantly, control households are those that experience a health shock outside the event window of each sub-experiment, ensuring that only "clean comparison" households are used as controls and avoiding the inclusion of previously treated households as controls. To illustrate this point with a specific example, the first sub-experiment (first row) considers households that experienced a health shock in wave 4 as the treated group, while using as controls those households that suffered a health shock between waves 8 and 15. However, some treatment adoption dates are excluded from the sample. For instance, households with a health shock in waves 2 or 3 are not included as treated or control, as they received the treatment too early and there are insufficient pre-treatment observations. Similarly, those with a health shock in waves 12 through 15 are only used as controls, as the treatment occurred too late, and there are not enough post-treatment observations for these households. Importantly, each sub-experiment includes the *same* periods/waves for both treated and control households, while selecting control households that experienced the health shock *after* the treatment window.

For a household to be included in a given sub-experiment, it must be present in all waves of that sub-experiment[6]. This restriction is necessary in order to ensure that the results are not driven by compositional change as shown by Wing et al. (2024a). This restriction implies that those households experiencing widowhood during the sub-experiment's event window are excluded from that sub-experiment onward. Addi-

[6]Approximately 31.9% of households in the analytic sample experience the death of one spouse at some point during the observation period. Married households are included in the sample until widowhood, at which point they exit the sample. Consequently, mortality shortens the observation window for some households but does not generate differential attrition within sub-experiments or mechanically drive the estimated effects of health shocks on divorce.

tionally, previous research by Wing et al. (2024a) has shown that the basic Stacked Diff-in-Diff estimator fails to correctly identify the Average Treatment Effect on the Treated. This issue comes from the different weights applied to trends in the treated and control groups. To recover the ATT, corrective sample weights can be computed and applied. Specifically, the weights are set to 1 for treated households in the stacked dataset, while the control group weights are obtained by dividing the ratio of treated households (i.e., the number of households that adopt the treatment in a particular sub-experiment divided by the total number of treated households) by the ratio of control households (i.e., the number of households serving as controls in that sub-experiment divided by the total number of control households). These corrective sample weights are used in the regressions throughout the analysis.

3.4 Estimating Equation

After constructing our stacked dataset and removing the "forbidden comparisons," we estimate the following equation:

$$Y_{sd,t+1} = \sum_{\substack{j=-3,\ldots,3 \\ j\neq -1}} [\beta_j * (T_{sd} * 1[t-b=j])] + \sum_{\substack{j=-3,\ldots,3 \\ j\neq -1}} [\lambda_j * 1[t-b=j]] + \theta_{sd} + \gamma_{d,t} + \epsilon_{sd,t} \quad (1)$$

where $Y_{sd,t+1}$ is the outcome variable in period $t+1$ for household s in sub-experiment d, which takes the value 1 from the moment that divorce happens, T_{sd} is a variable equal to 1 if household s is treated in sub-experiment d, $1[t-b=j]$ is a set of indicators for the time difference between the observation wave t and the reference wave b of the sub-experiment (the wave in which treated units receive treatment in the sub-experiment), θ_{sd} are household s fixed effects in sub-experiment d, $\gamma_{d,t}$ are time (wave) fixed effects in sub-experiment d, and $\epsilon_{sd,t}$ are the residuals. Since some households may act as controls in different sub-experiments, we cluster the standard errors at the household level instead of clustering at the household × sub-experiment level. Clustering the standard errors at the household × sub-experiment level would lead to over-rejecting the null hypothesis of no effects of health shocks (Wing et al., 2024a).

4 Results

4.1 Do Health Shocks Increase the Probability of Divorce?

Table 4 presents the OLS estimates of Equation (1), using the stacked dataset described above. To improve readability, all estimated coefficients and standard errors are multiplied by 100. Panel A displays the average post-treatment effect, computed as a linear combination of the post-treatment event coefficients, while Panel B reports the event study coefficient estimates. Our preferred specification, shown in Column 2, applies the corrective weights proposed by Wing et al. (2024a) to estimate the ATT. However, we also report unweighted results in Column 1 to assess whether our main findings hinge on the use of corrective weights.

The estimated post-treatment average effects (Table 4, Panel A) are not only statistically significant but also substantial in magnitude. To assess their size, it is helpful to express them relative to the share of divorced couples in our sample, which is approximately 6.3%. Dividing the estimated effects by this share reveals that experiencing a health shock increases the probability of divorce by 18.4% and 18.8% relative to the mean divorce prevalence when we do not apply corrective weights (Column 1) and when we do so (Column 2), respectively, indicating very similar percentage effects across specifications.

Table 4: The Impact of Health Shocks on the Probability of Divorce

	Without Weights	With Weights
	(1)	(2)
Panel A: Post-Treatment Average Effect		
Treated (=1)× Post (=1)	1.156**	1.178***
	(0.465)	(0.451)
Panel B: Event Study		
Treated (=1)× Event-time, -3 (=1)	-0.476	-0.506
	(0.543)	(0.535)
Treated (=1)× Event-time, -2 (=1)	-0.449	-0.491
	(0.357)	(0.351)
Treated (=1)× Event-time, 0 (=1)	0.334	0.335
	(0.315)	(0.297)
Treated (=1)× Event-time, 1 (=1)	0.956*	0.919*
	(0.498)	(0.476)
Treated (=1)× Event-time, 2 (=1)	1.505**	1.549***
	(0.585)	(0.575)
Treated (=1)× Event-time, 3 (=1)	1.827***	1.909***
	(0.654)	(0.663)
Household × Subexperiment FE	Yes	Yes
Wave × Subexperiment FE	Yes	Yes
Observations	22372	22372
Number of clusters (households)	1325	1325
Prevalence of divorce	0.063	0.063
Impact of health shock(percentage)	18.448	18.804

Notes: This table reports estimates of the effect of a health shock on the probability of divorce, defined as in Section 2.2. Column (2) applies the corrective weights proposed by Wing et al. (2024a), while Column (1) reports unweighted estimates. All coefficients and standard errors are multiplied by 100. Both specifications include household and wave fixed effects. Standard errors are clustered at the household level. Panel A reports the average post-treatment effect, computed as a linear combination of post-treatment event-time coefficients. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

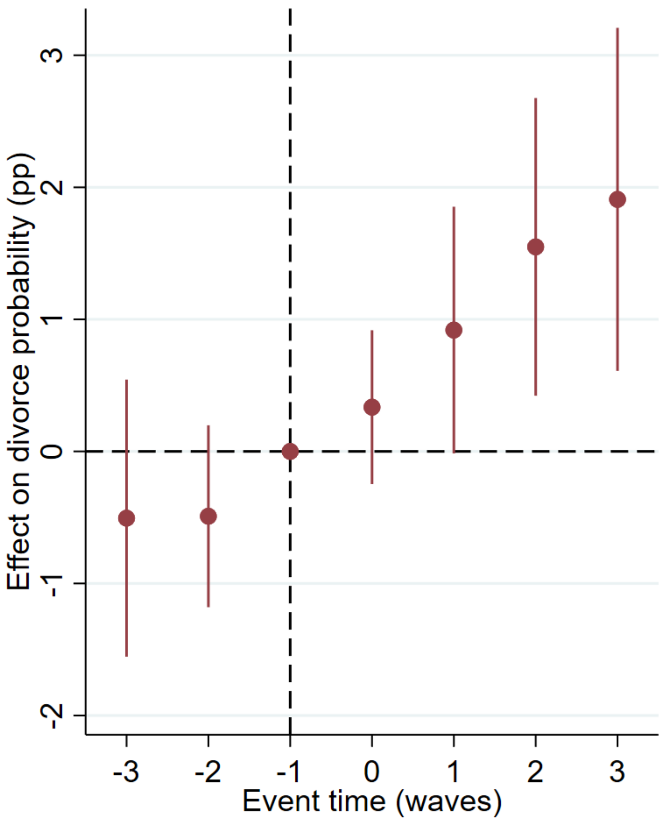
As shown in Panel B of Table 4, the pre-treatment coefficient estimates are not statistically significant, suggesting that, consistent with the unexpected nature of the health shocks, there are no anticipation effects. This finding also supports the plausibility of the parallel trends assumption. In contrast, the post-treatment coefficient estimates increase progressively and become significantly larger over time, indicating that the impact of health shocks on divorce intensifies gradually rather than manifesting fully immediately after the adverse health event.[7]

[7] We formally tested the null hypothesis that the difference in treatment effects between event times 3 and 0 is zero, and clearly rejected it, as the two-sided p-value is 0.0066.

This gradual pattern may reflect several factors, including the time required for marital strain to accumulate, delays inherent in the divorce process itself (such as separation periods or financial arrangements), or decisions to postpone separation until after acute caregiving phases.[8] To facilitate visual inspection, Figure 3 presents these event study results.

Because the graphical representation provides an immediate summary of the underlying dynamics and weighted and unweighted specifications consistently yield very similar estimates across subsequent analyses, the remainder of the paper presents results graphically and focuses on the weighted specification.

Figure 3: The Impact of Health Shocks on the Probability of Divorce



Note: This figure presents event-study estimates of the effect of a health shock on the probability of divorce, defined as in Section 2.2. Estimates are obtained from equation Equation (1) using the corrective weights proposed by Wing et al. (2024a). The plotted coefficients correspond to the interaction between event-time dummies and the treatment indicator. Error bars denote 95 percent confidence intervals. The specification includes household and wave fixed effects.

[8]Given the biennial structure of the HRS, event time 0 corresponds to 24 months after reference wave -1; event time 1 corresponds to 48 months after reference wave -1, and so on.

4.2 Mechanisms

To investigate potential mechanisms by which the occurrence of a health shock may influence the probability of divorce, we focus on three potential channels, namely those related to mental health (Section 4.2.1), cognitive decline (Section 4.2.2), and financial strain (Section 4.2.3).

4.2.1 Mental Health as Measured by the CES-D Scale

To assess individuals' mental health, we utilize the Center for Epidemiologic Studies Depression Scale (CES-D), a widely recognized and validated tool commonly used in psychiatric epidemiology to screen for depression and related disorders (Radloff, 1977). The CES-D version administered in the HRS is a shorter, 8-item version of the original CES-D scale (which originally consisted of 19 items). Each respondent is asked whether they experienced any of the following feelings during the week prior to the interview: depression, everything feels like an effort, restless sleep, loneliness, sadness, inability to get going, happiness, and enjoyment of life. Responses are aggregated into a final score ranging from 0 to 8. The responses to the last two items are reverse-coded (1 minus the variable value) so that higher values consistently indicate a greater propensity for depressive symptoms.

Since each household includes a husband and a wife, mental health outcomes are measured at the household level. The first mental health indicator we use is the average CESD-8 score of the two partners, standardized to have mean zero and unit standard deviation. The second indicator is a dummy variable equal to one if at least one member of the couple reports an (unstandardized) CESD-8 score greater than or equal to four, following Botoseneanu et al. (2023). If data are missing for one individual in the couple, the dummy takes the value of the observed individual.

Figure 4 summarizes the impact of health shocks on these two mental health measures. Taken together, the estimates are consistent with health shocks having detrimental effects on mental health. The top panel shows that experiencing a health shock leads to a statistically significant deterioration in mental health, with the average CESD-

8 score increasing by approximately 0.14 standard deviations. The bottom panel shows a corresponding increase of about 3.6 percentage points in the probability that at least one member of the couple exhibits depressive symptoms, which represents roughly 27% of the mean prevalence of this indicator in the sample (13.1%). Both magnitudes correspond to average effects over the post-treatment period and are statistically significant at the 1% level.

Both panels indicate that the effects are persistent over time and particularly pronounced in the two waves following the health shock.[9]

4.2.2 Cognitive Impairment

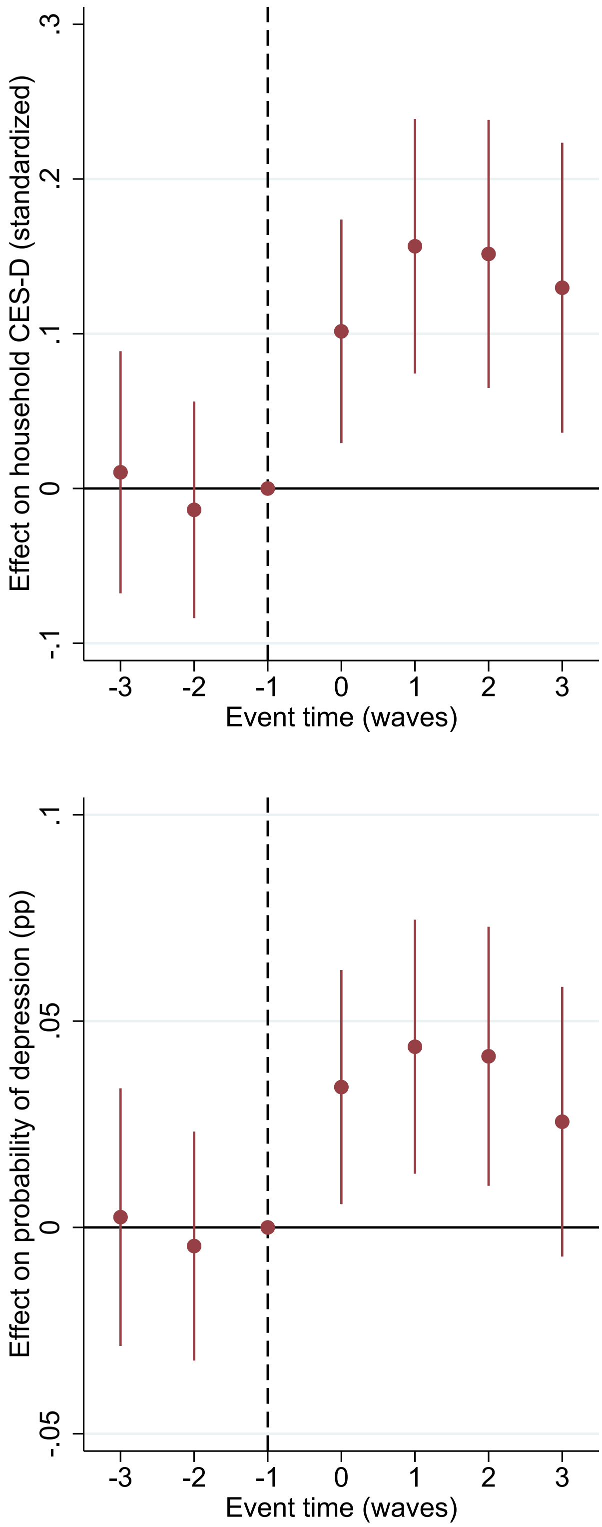
Previous research indicates that severe health shocks can accelerate cognitive decline in later life, often resulting in a substantial deterioration of mental functioning (Schiele and Schmitz, 2023; Alfaro-Acha et al., 2006). This evidence suggests that cognitive function may represent a relevant channel through which health shocks could affect couple dynamics.

To measure cognitive function, we first leverage the availability of a summary index in the HRS referred to as COG27. This index aggregates the following measures: i) immediate word recall (range: 0–10), which asks the respondent to repeat ten words read aloud by the interviewer; ii) delayed word recall (range: 0–10), which asks the respondent to recall the same ten words approximately 10 minutes after they were announced; iii) the serial 7's test result (range: 0–5), which measures the number of correct subtractions of seven starting from 100 over five trials;[10] and iv) backward counting from 20 for 10 continuous numbers (range: 0–2). Respondents receive two points if they are successful on the first try, one point if they are successful on the second try, and zero points if unsuccessful on both attempts. The COG27 index yields a total score ranging from 0 to 27 for each respondent and has been demonstrated to be a reliable predictor

[9] While the treatment effect on the depression probability becomes less precise by event time 3, we cannot reject the null hypothesis that the difference in treatment effects between event times 3 and 0 is zero, as the two-sided p-value is 0.595.

[10] Correct subtractions are based on the previously given number, so even if one subtraction is incorrect, subsequent answers are evaluated against the prior response.

Figure 4: The Impact of Health Shocks on Mental Health



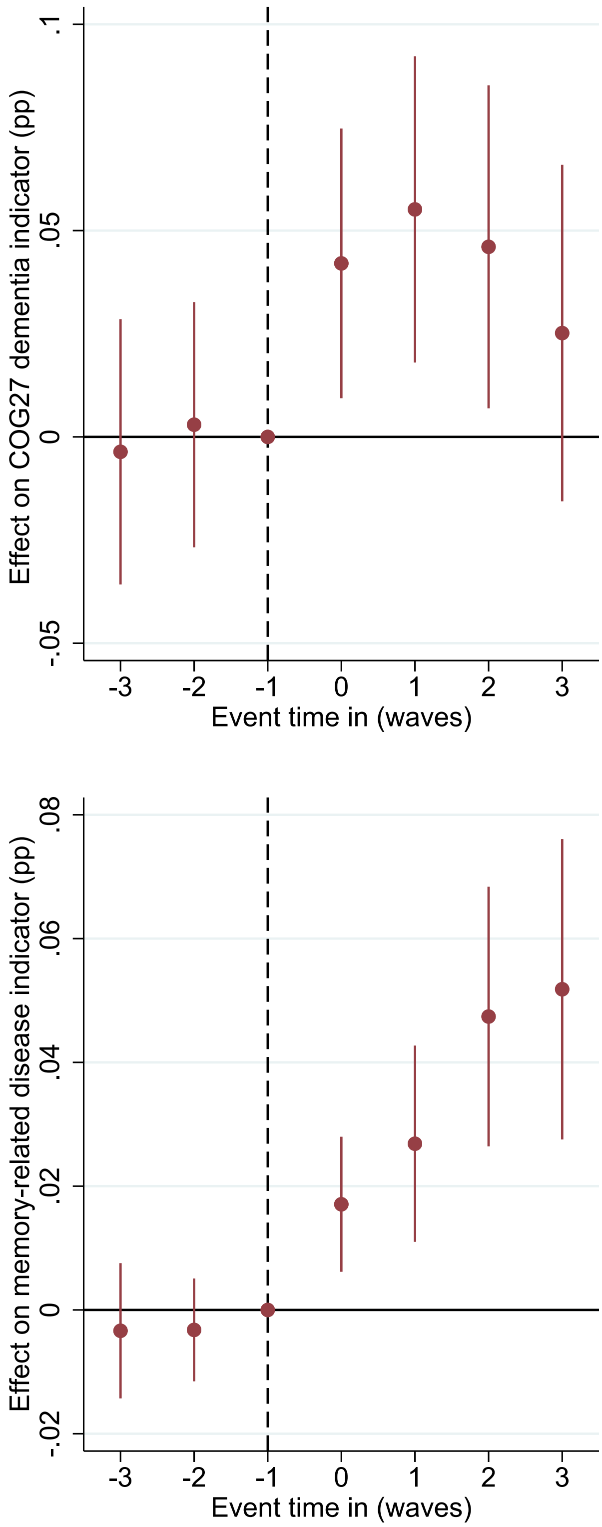
Note: The figure shows the impact of a health shock on mental health at the household level, using the weights proposed by Wing et al. (2024a). The top panel plots the effect on the standardized average CESD-8 score of the couple, while the bottom panel plots the effect on a dummy variable equal to one if at least one member of the couple reports a CESD-8 score associated with depression. The coefficients correspond to the interaction between event-time dummies and the treatment indicator in Equation (1), with 95 percent confidence intervals. The specification includes household and wave fixed effects.

of overall levels of dementia and cognitive impairment in large public surveys (Crimmins et al., 2011). Scores between 0 and 6 on the COG27 index fall within the dementia range, scores between 7 and 11 are associated with cognitive impairment, and scores above 12 are considered within the normal range (Gianattasio et al., 2019). We create a dummy variable equal to 1 if at least one member of the couple scores below 12 and use this binary variable as our first indicator of cognitive function.

Additionally, the HRS provides information on diagnoses of memory-related conditions, although the wording of these questions changes over time. From wave 4 onward, respondents are asked whether a medical professional has ever told them that they have a memory-related disease. Starting in wave 10, this question is refined, and individuals are instead asked separately whether they have received a diagnosis of Alzheimer's disease or dementia. Following the approach in Jeong et al. (2024), we harmonize these measures over time and construct an indicator for memory-related disease diagnoses that spans waves 4 to 15 (the latest available wave). This dummy variable equals one if an individual reports a diagnosis of a memory-related disease in earlier waves or a diagnosis of Alzheimer's disease or dementia in wave 10 and subsequent waves, and zero otherwise. At the household level, the indicator takes the value one if at least one member of the couple reports such a diagnosis. This binary variable is used as our second indicator of cognitive function.

Figure 5 summarizes the impact of health shocks on cognitive function at the household level. The results point to a clear pattern: health shocks significantly increase the likelihood of cognitive deterioration within the couple. The top panel shows that experiencing a health shock raises the probability that at least one member of the couple exhibits cognitive impairment—defined as scoring below 12 on the COG27 index—by approximately 4.2 percentage points on average over the post-treatment period, corresponding to roughly 24% of the sample mean for this indicator (17.4%). While the estimated effect becomes less precise three waves after the health shock, we cannot reject the null hypothesis that treatment effects immediately following the shock and

Figure 5: The Impact of Health Shocks on Cognitive Function



Note: The figure shows the impact of a health shock on cognitive function at the household level, using the weights proposed by Wing et al. (2024a). The top panel plots the effect on a dummy variable equal to one if any member of the couple scores below 12 on the COG27 index, while the bottom panel plots the effect on a dummy variable equal to one if any member of the couple has been diagnosed by a medical professional with a memory-related disease. The coefficients correspond to the interaction between event-time dummies and the treatment indicator in Equation (1), with 95 percent confidence intervals. The specification includes household and wave fixed effects.

three waves later are equal.[11]

The bottom panel shows a similar pattern for clinically diagnosed memory-related diseases. The probability that at least one member of the couple is diagnosed with a memory-related disease increases by approximately 3.6 percentage points on average following a health shock, corresponding to about 33% of the prevalence of this condition in the sample (10.9%). Moreover, unlike cognitive impairment measured through the COG27 index, this effect not only persists but also increases steadily and significantly in the years following the health shock.

4.2.3 Out-of-Pocket Medical Expenses

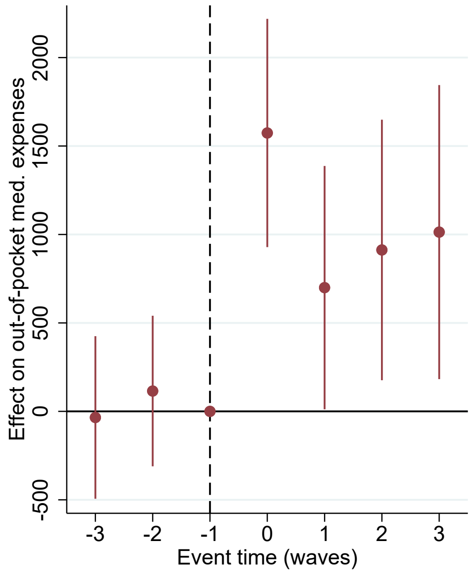
Finally, we examine whether health shocks induce financial strain by focusing on households' out-of-pocket medical expenditures, which we use as a proxy for financial burden. The results indicate that health shocks lead to sizable increases in out-of-pocket medical expenses.

Figure 6 summarizes the dynamic impact of health shocks on out-of-pocket medical expenditures. The figure shows that experiencing a health shock is associated with a large and statistically significant increase in medical expenses, with out-of-pocket expenditures rising by approximately $1,050 on average over the post-treatment period. This increase is particularly pronounced immediately following the shock and remains elevated in subsequent waves.

Although not directly comparable, these findings are consistent with Dobkin et al. (2018), who use an event-study approach and document increases in out-of-pocket expenditures among non-elderly (ages 50–59) HRS individuals following non-pregnancy-related hospital admissions, their empirical analogue of an adverse health shock. Overall, the dynamic patterns shown in Figure 6 suggest that health shocks generate substantial and persistent financial strain through higher medical spending.

[11] The two-sided p-value of the test is 0.397.

Figure 6: The Impact of Health Shocks on Out-of-Pocket Medical Expenditures



Note: The figure shows the impact of a health shock on households' out-of-pocket medical expenditures, using the weights proposed by Wing et al. (2024a). The coefficients correspond to the interaction between event-time dummies and the treatment indicator in Equation (1), with 95 percent confidence intervals. The specification includes household and wave fixed effects.

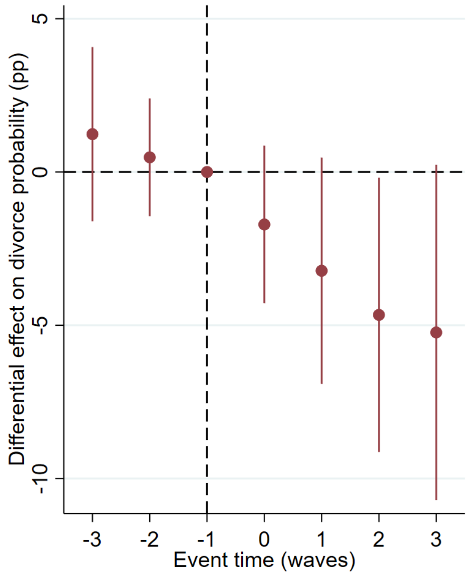
5 Heterogeneity

We now explore whether our results are heterogeneous across several dimensions of interest. To this end, we extend our benchmark model (Equation (1)) to include triple interactions between the treatment indicator, event-time dummies, and an indicator for the subgroup of interest.

5.1 Race

Gray divorces in the U.S. are markedly less frequent among White individuals, especially among those aged 65+ (Brown and Lin, 2022). This is also the case in our analytic sample, with the prevalence of divorce being 5.75% and 9.28% among White and Non-White couples, respectively. We define White couples as those where both members are White. While this divorce gap does not necessarily mean that White couples are more resilient to health shocks, it suggests that baseline relationship dynamics may differ across groups.

Figure 7: The Impact of Health Shocks on Divorce: Heterogeneity between White and Non-White Couples



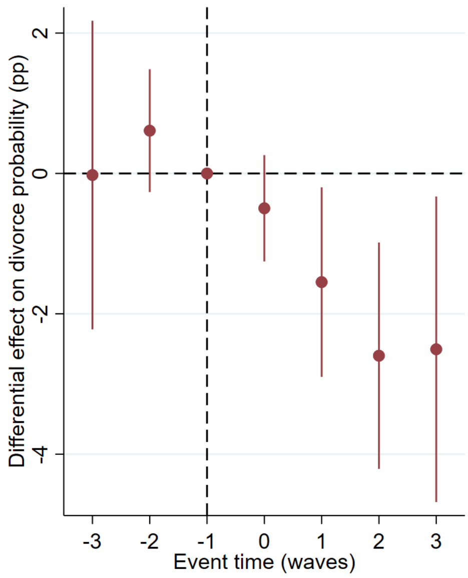
Notes: This figure presents event-study estimates of the differential effect of a health shock on the probability of divorce by race. Estimates are obtained from equation Equation (1) augmented with triple interactions between event-time dummies, the treatment indicator, and an indicator for White couples (equal to one if both spouses are White). The figure reports 95 percent confidence intervals. The specification includes household and wave fixed effects and applies the corrective weights proposed by Wing et al. (2024a).

Our results, graphically summarized in Figure 7, suggest that White couples are indeed significantly less prone than their Non-White counterparts to divorce following a health shock. The estimated effects are only significant at the 10% level in some periods, but the average effect throughout the post-treatment window is not only sizeable but also statistically significant at the 5% level. In particular, in households affected by a health shock, the probability of subsequent divorce rises, on average, by 3.7 percentage points more for Non-White than for White couples (roughly 60% of the sample mean divorce prevalence).

5.2 Religion

Divorce is less frequent among Catholics in the U.S., especially compared to Protestants and the religiously unaffiliated (Smith et al., 2025). In line with this pattern, in our analytic sample, the prevalence of divorce is 3.34% when both partners are Catholic and 7% when at least one partner is not. A natural question is whether marital stability in

Figure 8: The Impact of Health Shocks on Divorce: Heterogeneity between Catholic and Non-Catholic Couples



Note: This figure presents event-study estimates of the differential effect of a health shock on the probability of divorce by religion. Estimates are obtained from Equation (1) augmented with triple interactions between event-time dummies, the treatment indicator, and an indicator for Catholic couples (equal to one if both spouses are Catholic). The figure reports 95 percent confidence intervals. The specification includes household and wave fixed effects and applies the corrective weights proposed by Wing et al. (2024a).

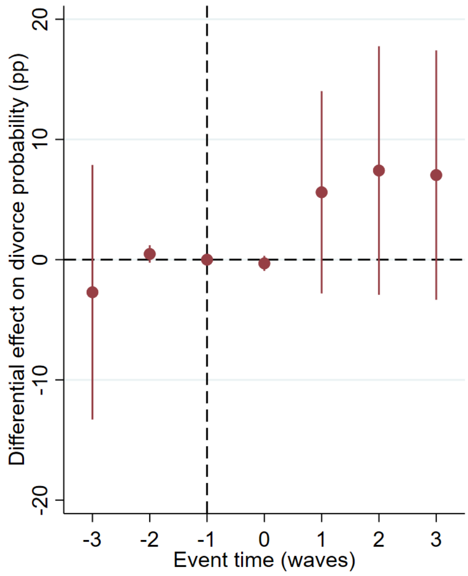
the face of hardship, particularly health shocks, is stronger among Catholic couples. Our results, displayed in Figure 8, suggest that it is. When households suffer a health shock, the probability of divorce is 1.78 percentage points lower (statistically significant) in couples where both partners are Catholic compared to couples where at least one partner is not—a reduction of approximately 30% relative to the mean divorce rate.

5.3 Poverty

In the U.S., divorce is less common among wealthier couples than among poorer ones (Eads and Tach, 2016; Killewald et al., 2023). Consistent with this pattern, divorce prevalence is 17% among households below the poverty line, compared to almost 6% among those above it.

We now investigate whether health shocks may further amplify this gap by having a stronger impact on poor versus non-poor households, defining the former as those whose income in the previous calendar year falls below the U.S. Census Bureau poverty

Figure 9: The Impact of Health Shocks on Divorce: Heterogeneity by Poverty Status



Notes: This figure presents event-study estimates of the differential effect of a health shock on the probability of divorce by poverty status. Estimates are obtained from equation Equation (1) augmented with triple interactions between event-time dummies, the treatment indicator, and an indicator for households below the poverty threshold (based on the U.S. Census Bureau definition). The figure reports 95 percent confidence intervals. The specification includes household and wave fixed effects and applies the corrective weights proposed by Wing et al. (2024a).

threshold. Our results, summarized in Figure 9, do not provide evidence that the impact of health shocks on divorce differs across poverty status: we cannot reject the null hypothesis that the effect is the same for poor and non-poor households at conventional significance levels.

5.4 Gender

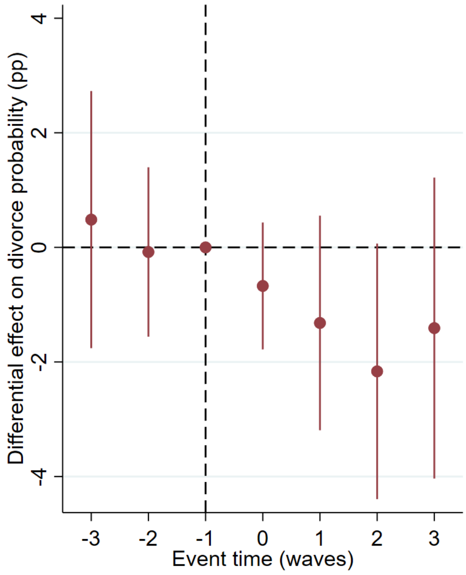
Household responses to health shocks may differ by gender, depending on whether the husband or the wife is affected. In couples with more traditional or unequal gender roles, wives are more likely to assume caregiving responsibilities, which may lead to greater strain when they become ill compared to when husbands do. Such gender norms may also reinforce women's financial dependence and economic vulnerability, thereby limiting their capacity to exit a marriage. In our analytic sample, the prevalence of divorce is higher when wives experience a health shock compared to when husbands do (6.9 percent versus 5.8 percent), which may point in this direction. However, in the literature the evidence on whether health issues have a stronger association with divorce depending on who is affected is mixed. Glantz et al. (2009), using a sample of 515 patients prospectively identified as having either a brain tumour or multiple sclerosis and who were married at the time of diagnosis, find that female gender is the strongest predictor of separation or divorce. Vignoli et al. (2025), drawing on European survey data on adults aged 50 and above, find that divorce risk is higher when wives report poor health or severe activity limitations while husbands remain healthy, but not when husbands experience poor health; in the latter case, divorce risk does not differ significantly from that of couples in good health. Karraker and Latham (2015a), using US survey data for the 50+ population from the HRS, generally fail to reject the null hypothesis that the effect of illness onset differs depending on who experiences it in a statistically significant way.

Our own analysis, summarized in Figure 10, is consistent with this latter set of findings: we cannot detect statistically significant differences in the effect of health shocks on divorce depending on whether the wife or the husband is affected.

While this may contrast with some of the correlational evidence discussed above, those studies rely on multivariate regression analysis and therefore implicitly assume selection on observables. In contrast, in our quasi-experimental framework we do not compare households that ever experience a health shock with those that never do, but rather households that experience a health shock with others that are set to experience one in subsequent years. This matters because households that eventually experience a health shock differ from never-treated households in several relevant dimensions, such as education and labor-force attachment, which are often discussed in relation to gender roles. For instance, as shown in Section 3.2 (Figure 1), even after holding age and cohort effects constant, individuals in households that never experience a health shock are significantly more likely to participate in the labor force and to have attended college compared to their counterparts in households where a health shock occurs.[12] To the extent that our design accounts for these differences, it is therefore not particularly

[12] Research on gender identity indicates that adherence to traditional gender norms is more common among individuals with less education, while those with higher educational attainment are generally more likely to support egalitarian perspectives on the roles of men and women in society. See Du et al. (2021) and Rivera-Garrido (2022), among others.

Figure 10: The Impact of Health Shocks on Divorce: Heterogeneity by Gender



Notes: This figure presents event-study estimates of the differential effect of a health shock on the probability of divorce by gender. Estimates are obtained from equation Equation (1) augmented with triple interactions between event-time dummies, the treatment indicator, and an indicator for whether the wife experiences the health shock. The figure reports 95 percent confidence intervals. The specification includes household and wave fixed effects and applies the corrective weights proposed by Wing et al. (2024a).

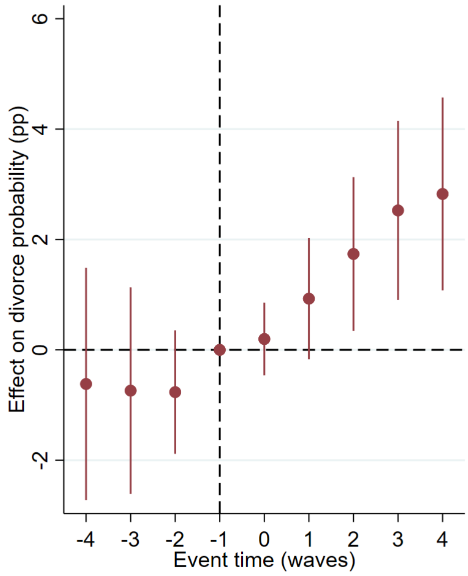
surprising that we do not uncover statistically detectable heterogeneity by gender.

6 Robustness Checks

In this section, we test the robustness of our results to variations in the length of the event window and in the definition of health shocks. We also assess whether our conclusions are sensitive to the use of an alternative difference-in-differences estimator.

First, we estimate Equation 1 using a different event window length. Specifically, we employ four waves before and after the health shock, rather than the three waves used in the benchmark specification. The use of a longer event window allows us to better understand the evolution of the impact over time. However, this approach comes at the cost of excluding certain treatment adoption dates from the stacked dataset, since units treated very early or very late in the panel cannot be followed for the full window. In particular, households treated in waves 4 and 11, which were included in the main specification, are now dropped because the health shock occurs too early or too late in

Figure 11: The Impact of Health Shocks on the Probability of Divorce - Longer Event Window



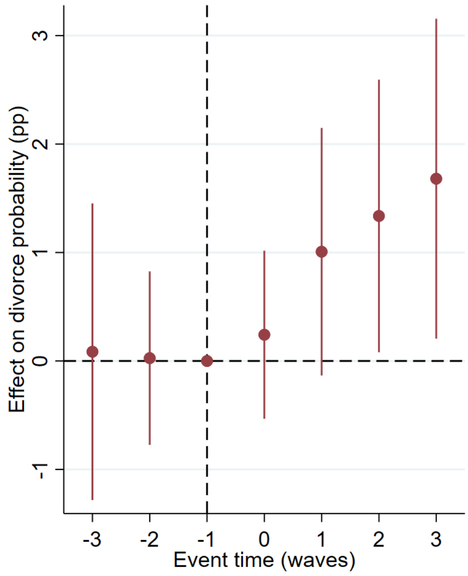
Notes: This figure presents event-study estimates of the effect of a health shock on the probability of divorce using a longer event window (four waves before and after the shock). Estimates are obtained from Equation (1) using the corrective weights proposed by Wing et al. (2024a). The figure reports 95 percent confidence intervals. The specification includes household and wave fixed effects.

the observation period. In addition, the control group now consists of households that experience the health shock more than four waves ahead, instead of three.

Figure 11 shows that our findings are very similar to those obtained using the benchmark three-wave window. The pre-treatment coefficients remain close to zero and statistically insignificant, which is consistent with both the no-anticipation and parallel-trends assumptions. Post-treatment effects again rise gradually and reach values of roughly 1.5–2.5 percentage points several waves after the shock. Taken together, these results indicate that our conclusions are not sensitive to the choice of event window length.

The next robustness check examines an alternative definition of a health shock. Following Dobkin et al. (2018), we define a health shock for a couple as a hospital stay lasting more than four nights—the median value of the variable—calculated by summing the husband's and wife's hospital nights since the previous wave, provided that neither reported an overnight hospital stay in the prior wave. This definition differs from our benchmark measure, which is based on the diagnosis of a specific condition

Figure 12: The Impact of Health Shocks on the Probability of Divorce - Hospital Admissions



Notes: This figure presents event-study estimates of the effect of a health shock on the probability of divorce when health shocks are defined as hospital stays longer than four nights, summed across spouses since the previous wave. Estimates are obtained from Equation (1) using the corrective weights proposed by Wing et al. (2024a). The figure reports 95 percent confidence intervals. The specification includes household and wave fixed effects.

(cancer, stroke, or heart attack). Since, unfortunately, we lack data on the reasons for each hospitalization, this alternative definition likely captures less severe health events, and the resulting effects should therefore be smaller in magnitude than those obtained with our benchmark measure.

Figure 12 shows that the results obtained with this alternative definition are qualitatively consistent with our main findings. The estimated coefficients point to a higher probability of divorce following a hospitalization spell, and the effects become larger over time after the shock. Although the magnitudes are somewhat smaller, as expected, the overall pattern is remarkably similar to the baseline estimates. This reinforces the conclusion that our results are not driven by the specific way in which health shocks are defined, even though the hospitalization-based measure likely captures events that are somewhat less unexpected or less severe than our benchmark definition.

To further assess the robustness of our findings, we estimate the average treatment effect on the treated (ATT) using the methodology proposed by Borusyak et al. (2024), developed for difference-in-differences designs with staggered treatment adoption and heterogeneous causal effects. This approach relies on a transparent imputation procedure. First, period and unit fixed effects are estimated via regression using only untreated observations. Second, these estimated fixed effects are used to impute the untreated potential outcomes and obtain an estimated treatment effect for each treated observation. Finally, a weighted average of these treatment-effect estimates is taken, with weights corresponding to the estimation target (the ATT in our case). The proposed estimator ensures unbiasedness and exhibits favorable efficiency properties, as Borusyak et al. (2024) show in simulations. An additional advantage is its computational simplicity, as it only requires the estimation of a standard two-way fixed-effects model, while offering a clear and transparent link between the identifying assumptions and the resulting estimator.

Table 5 shows that the estimated ATT is very similar to our benchmark estimate, both in percentage-point terms and when expressed relative to the mean divorce rate. The point estimate implies an increase in divorce of about 1.2 percentage points (around 19% of the baseline prevalence), closely matching the effect reported in our benchmark specification.

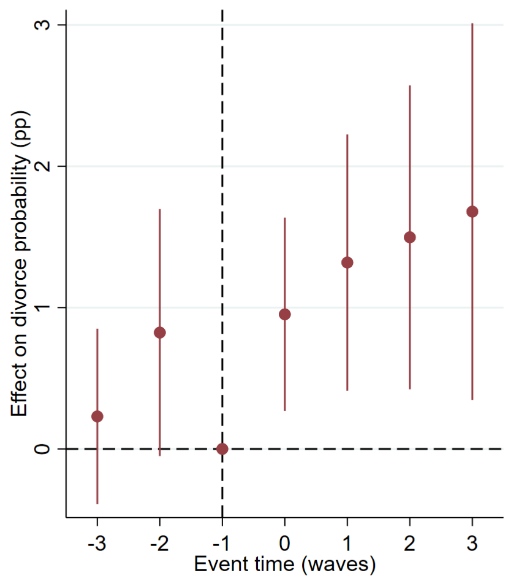
Figure 13 displays the corresponding event-study coefficients. The dynamic pattern is broadly similar to that obtained in our baseline analysis, with post-treatment effects increasing gradually over time and remaining consistently positive. Although one pre-treatment coefficient appears marginally positive, it is statistically indistinguishable from zero and does not reveal any systematic pre-trend, which is consistent with our identifying assumptions. Taken together, the ATT and the event-study coefficients are highly consistent with our benchmark results, suggesting that our main finding does not hinge on the particular estimator used.

Table 5: The Impact of Health Shocks on the Probability of Divorce - Borusyak et al. (2024) Estimator.

	(1)
Average Treatment Effect on the Treated (ATT)	1.224**
	(0.586)
Prevalence of divorce	0.065
Impact of Health Shock (percentage)	18.697

Notes: This table reports estimates of the effect of a health shock on the probability of divorce obtained using the estimator proposed by Borusyak et al. (2024). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure 13: The Impact of Health Shocks on the Probability of Divorce - Borusyak et al. (2024) Estimator



Notes: This figure presents event-study estimates of the effect of a health shock on the probability of divorce obtained using the estimator proposed by Borusyak et al. (2024). 95 percent confidence intervals are shown.

7 Conclusion

This paper examines how unexpected health shocks shape marital stability in later life. Using longitudinal data from the Health and Retirement Study and a Stacked Difference-in-Differences framework (Wing et al., 2024a), we provide causal evidence that serious and unanticipated health events increase the probability of couple dissolution. The estimated effects do not fully materialize immediately after the health shock but instead build gradually over subsequent waves, suggesting that the consequences of health crises unfold along dynamic and cumulative pathways.

Our findings also shed light on the mechanisms underlying this relationship. Following a health shock, households experience persistent deterioration in mental health, worsening cognitive functioning, and marked increases in out-of-pocket medical expenditures. Taken together, these patterns are consistent with health shocks straining emotional, cognitive, and financial resources within couples, increasing the likelihood of relationship breakdown.

To assess the credibility of our findings, we conduct a series of robustness exercises. Our main estimates remain stable when we extend the event window, when we adopt an alternative definition of health shocks, and when we implement an alternative difference-in-differences estimator.

In sum, the evidence reinforces the conclusion that the increase in divorce observed after a health shock reflects a robust empirical pattern. Rather than being isolated medical events, health shocks emerge as key turning points in household dynamics. By highlighting the relationship between health shocks and divorce, this study contributes to a broader understanding of the challenges associated with aging and the rising trend of "silver splitters", underscoring the need for multidimensional policy responses to address these issues effectively. As populations age and health problems become more common at older ages, health shocks will increasingly intersect with family dynamics and marital stability. Policies that strengthen post-diagnosis support—including better health insurance coverage to reduce out-of-pocket medical expenses, mental-health services, caregiver assistance, and financial counseling —may help reduce some of the pressures that ultimately contribute to marital dissolution. In this sense, serious health shocks are not only medical events but also relational events, reshaping household structure and well-being in later life.

References

ALFARO-ACHA, A., SNIH, S. A., RAJI, M. A. ET AL. (2006): "Handgrip strength and cognitive decline in older Mexican Americans," *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 61, 859–865.

ANGELINI, V. AND COSTA-FONT, J. (2023): "Health and wellbeing spillovers of a partner's cancer diagnosis," *Journal of Economic Behavior & Organization*, 212, 422–437.

ARTEAGA, C., VIGEZZI, N. AND GARCIA-GOMEZ, P. (2024): "In Sickness and in Health: The Broad Impact of Spousal Health Shocks," Working paper.

BERARDO, F. M., APPEL, J. AND BERARDO, D. H. (1993): "Age dissimilar marriages: Review and assessment," *Journal of Aging Studies*, 7, 93–106.

BLUNDELL, R., BORELLA, M., COMMAULT, J. ET AL. (2024): "Old age risks, consumption, and insurance," *American Economic Review*, 114, 575–613.

BONEKAMP, J. AND WOUTERSE, B. (2023): "Do different shocks in health matter for wealth?" *Journal of Health Economics*, 87, 102719.

BOOTH, A. AND JOHNSON, D. R. (1994): "Declining health and marital quality," *Journal of Marriage and the Family*, 218–223.

BORUSYAK, K., JARAVEL, X. AND SPIESS, J. (2024): "Revisiting event-study designs: robust and efficient estimation," *Review of Economic Studies*, 91, 3253–3285.

BOTOSENEANU, A., ELMAN, M. R., ALLORE, H. G. ET AL. (2023): "Depressive multimorbidity and trajectories of functional status among older Americans: differences by racial/ethnic group," *Journal of the American Medical Directors Association*, 24, 250–257.

BREIVIK, A.-L. AND COSTA-RAMÓN, A. (2024): "The Impact of Children's Health Shocks on Parents' Labor Earnings and Mental Health," *Review of Economics and Statistics*, 1–45.

BRITO, E. AND CONTRERAS, D. (2023): "The caregiving penalty: Caring for sick parents and the gender pay gap," *Unpublished Ph. d. Thesis). Brown University, Providence, Rhode Island, USA*, 6, 27.

BROWN, S. L. AND LIN, I.-F. (2012): "The gray divorce revolution: Rising divorce among middle-aged and older adults, 1990–2010," *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 67, 731–741.

——— (2022): "The graying of divorce: A half century of change," *The Journals of Gerontology: Series B*, 77, 1710–1720.

BUTTERS, R. A., SACKS, D. W. AND SEO, B. (2022): "How do national firms respond to local cost shocks?" *American Economic Review*, 112, 1737–1772.

CALLAWAY, B. AND SANT'ANNA, P. H. (2021): "Difference-in-differences with multiple time periods," *Journal of econometrics*, 225, 200–230.

CARR, D. AND SPRINGER, K. W. (2010): "Advances in families and health research in the 21st century," *Journal of marriage and family*, 72, 743–761.

CENGIZ, D., DUBE, A., LINDNER, A. ET AL. (2019): "The effect of minimum wages on low-wage jobs," *The Quarterly Journal of Economics*, 134, 1405–1454.

CHERLIN, A. J. (2010): "Demographic trends in the United States: A review of research in the 2000s," *Journal of marriage and family*, 72, 403–419.

COSTA-FONT, J. AND VILAPLANA-PRIETO, C. (2022): "Health shocks and housing downsizing: How persistent is 'ageing in place'?" *Journal of Economic Behavior & Organization*, 204, 490–508.

CRIMMINS, E. M., KIM, J. K., LANGA, K. M. ET AL. (2011): "Assessment of cognition using surveys and neuropsychological assessment: the Health and Retirement Study and the Aging, Demographics, and Memory Study," *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 66, i162–i171.

DE CHAISEMARTIN, C. AND D'HAULTFOEUILLE, X. (2020): "Two-way fixed effects estimators with heterogeneous treatment effects," *American Economic Review*, 110, 2964–2996.

DOBKIN, C., FINKELSTEIN, A., KLUENDER, R. ET AL. (2018): "The economic consequences of hospital admissions," *American Economic Review*, 108, 308–352.

DU, H., XIAO, Y. AND ZHAO, L. (2021): "Education and gender role attitudes," *Journal of Population Economics*, 34, 475–513.

DU, X. AND ZAREMBA, K. (2024): "Household Penalty," *Available at SSRN 4989822*.

EADS, A. AND TACH, L. (2016): "Wealth and inequality in the stability of romantic relationships," *RSF: The Russell Sage Foundation Journal of the Social Sciences*, 2, 197–224.

FADLON, I. AND NIELSEN, T. H. (2019): "Family health behaviors," *American Economic Review*, 109, 3162–3191.

——— (2021): "Family labor supply responses to severe health shocks: Evidence from Danish administrative records," *American Economic Journal: Applied Economics*, 13, 1–30.

FUSTER, V. (2017): "Changing demographics: a new approach to global health care due to the aging population," .

GIANATTASIO, K. Z., WU, Q., GLYMOUR, M. M. ET AL. (2019): "Comparison of methods for algorithmic classification of dementia status in the health and retirement study," *Epidemiology*, 30, 291–302.

GLANTZ, M. J., CHAMBERLAIN, M. C., LIU, Q. ET AL. (2009): "Gender disparity in the rate of partner abandonment in patients with serious medical illness," *Cancer*, 115, 5237–5242.

GOODMAN-BACON, A. (2021): "Difference-in-differences with variation in treatment timing," *Journal of econometrics*, 225, 254–277.

HEALTH AND RETIREMENT STUDY (2024): "Health and Retirement Study, [RAND HRS Longitudinal File 2020 (V2)] public use dataset," Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant numbers NIA U01AG009740 and NIA R01AG073289). Ann Arbor, MI, 2024.

Jeong, Y., Papageorge, N. W., Skira, M. M. et al. (2024): "Genetic Risk for Alzheimer's Disease and Related Dementias: Cognition, Economic Behavior, and Long-Run Planning," *NBER Working Paper*, 32181.

Jones, A. M., Rice, N. and Zantomio, F. (2020): "Acute health shocks and labour market outcomes: evidence from the post crash era," *Economics & Human Biology*, 36, 100811.

Joung, I. M., Van De Mheen, H. D., Stronks, K. et al. (1998): "A longitudinal study of health selection in marital transitions," *Social science & medicine*, 46, 425–435.

Karraker, A. and Latham, K. (2015a): "Authors' explanation of the retraction," *Journal of Health and Social Behavior*, 56, 417–419.

——— (2015b): "In sickness and in health? Physical illness as a risk factor for marital dissolution in later life," *Journal of health and social behavior*, 56, 420–435.

Kessler, R. C., Walters, E. E. and Forthofer, M. S. (1998): "The social consequences of psychiatric disorders, III: probability of marital stability," *American Journal of Psychiatry*, 155, 1092–1096.

Killewald, A., Lee, A. and England, P. (2023): "Wealth and divorce," *Demography*, 60, 147–171.

Lin, I.-F., Brown, S. L. and Mellencamp, K. A. (2022): "The roles of gray divorce and subsequent repartnering for parent–adult child relationships," *The Journals of Gerontology: Series B*, 77, 212–223.

Merikangas, K. R. (1984): "Divorce and assortative mating among depressed patients." *The American Journal of Psychiatry*, 141, 74–76.

Miller, D. L. (2023): "An introductory guide to event study models," *Journal of Economic Perspectives*, 37, 203–230.

Negrusa, B. and Negrusa, S. (2014): "Home front: Post-deployment mental health and divorces," *Demography*, 51, 895–916.

Pevalin, D. J. and Ermisch, J. (2004): "Cohabiting unions, repartnering and mental health," *Psychological Medicine*, 34, 1553–1559.

Radloff, L. S. (1977): "The CES-D scale: A self-report depression scale for research in the general population," *Applied psychological measurement*, 1, 385–401.

RAND Center for the Study of Aging (2024): "RAND HRS Longitudinal File 2020 (V2)," 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, May 2024.

Riekhoff, A.-J. and Vaalavuo, M. (2021): "Health shocks and couples' labor market participation: A turning point or stuck in the trajectory?" *Social Science & Medicine*, 276, 113843.

Rivera-Garrido, N. (2022): "Can education reduce traditional gender role attitudes?" *Economics of Education Review*, 89, 102261.

Roberto, K. A. and Blieszner, R. (2015): "Diverse family structures and the care of older persons," *Canadian Journal on Aging/La Revue canadienne du vieillissement*, 34, 305–320.

Schiele, V. and Schmitz, H. (2023): "Understanding cognitive decline in older ages: The role of health shocks," *European Economic Review*, 151, 104320.

Smith, G. W., Cooperman, A., Alper, B. A. et al. (2025): "Decline of Christianity in the US has slowed, may have leveled off," *Communications*, 202, 4372.

Smith, J. (2005): "Consequences and predictors of new health events," in *Analyses in the Economics of Aging*, University of Chicago Press, 213–240.

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

Tauchmann, H., Simankova, I. and Bünnings, C. (2023): "Health Shocks and Health Behavior: A Long-Term Perspective," .

Thompson, J. and Conley, D. (2016): "Health shocks and social drift: Examining the relationship between acute illness and family wealth," *RSF: The Russell Sage Foundation Journal of the Social Sciences*, 2, 153–171.

Trevisan, E. and Zantomio, F. (2016): "The impact of acute health shocks on the labour supply of older workers: Evidence from sixteen European countries," *Labour Economics*, 43, 171–185.

Uhlenberg, P. and Myers, M. A. P. (1981): "Divorce and the elderly," *The Gerontologist*, 21, 276–282.

Umberson, D., Liu, H., Mirowsky, J. et al. (2011): "Parenthood and trajectories of change in body weight over the life course," *Social science & medicine*, 73, 1323–1331.

U.S. Census Bureau (2023): "2023 National Population Projections Tables: Main Series," U.S. Census Bureau.

Vignoli, D., Alderotti, G. and Tomassini, C. (2025): "Partners' health and silver splits in Europe: A gendered pattern?" *Journal of Marriage and Family*, 87, 1639–1663.

Wing, C., Freedman, S. M. and Hollingsworth, A. (2024a): "Stacked difference-in-differences," Tech. rep., National Bureau of Economic Research.

Wing, C., Yozwiak, M., Hollingsworth, A. et al. (2024b): "Designing difference-in-difference studies with staggered treatment adoption: Key concepts and practical guidelines," *Annual Review of Public Health*, 45.