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

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

Daniela Vuri §

January 2025

Abstract

We analyze the impact of unexpected health shocks—defined as the sudden diagnosis of cancer, stroke, or heart attack—on the probability of couple dissolution using longitudinal representative data on older individuals (50+). We leverage the longitudinal nature of the HRS and utilize a quasi-experimental research approach that creates counterfactual scenarios for affected households by comparing them to households set to experience the same event in subsequent years. We find that experiencing a health shock significantly increases the probability of couple dissolution by approximately 19% of the mean divorce prevalence. This effect intensifies gradually over time rather than appearing 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, cognitive decline, and financial strain. Our findings suggest that all three mechanisms likely play a role in mediating the relationship between health shocks and the increased probability of divorce.

JEL Codes: I14, I24, J15, Z13, J13.
Keywords: Health shocks, divorce, aging.

*We thank Pedro Albarrán, Marco Bertoni, Kyle Butts, Scott Cunningham, Yarine Fawaz, María Hernández-de-Benito, David Jiménez-Gómez, 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, and the European Health Economics Association Online Seminar Series for their comments.

†Department of Economics, University of Alicante. Email: javier.artero@ua.es. López-Artero acknowledges financial support from Project PID2021-124237NB-I00 (financed by MCIN/ AEI /10.13039/501100011033/ and by FEDER Una manera de hacer Europa) and from ICAR Foundation

‡Department of Economics, University of Alicante and IZA. Email: anna.sanzdegaldeano@gmail.com. Sanz-de-Galdeano acknowledges financial support from Project PID2021-124237NB-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.

§Department of Economics and Finance University of Rome Tor Vergate and IZA. Email: daniela.vuri@uniroma2.it. Vuri acknowledges funding from ICAR foundation.

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 impact on the marital dynamics of older individuals.

This research addresses this gap by focusing on "gray divorces" or "silver splitters", terms that describe the rising divorce rates among those aged 65 and older (Brown and Lin, 2012, 2022). Additionally, a remarkable process of population ageing is taking place in the developed world. For instance, the population of Americans aged 65 and older is expected to grow from 58 million in 2022 to 82 million by 2050 (a 47% increase). During this time, their proportion of the total population is anticipated to rise from 17% to 23% (U.S. Census Bureau, 2023). Given this degree of population aging, which is by no means exclusive to the United States, it is important to study the factors behind the "silver splitters" phenomenon.

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. We use substantial and unanticipated health shocks to establish a causal link between household health and marital dissolution. Additionally, we leverage the longitudinal nature of the HRS and utilize a quasi-experimental research approach that creates counterfactual scenarios for affected households by comparing them to households set to 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.[1] 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" (i.e., comparisons between late and early treatment adopters). 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 adop-

[1] The US does not have a National Health System (NHS) like the UK or Canada. Instead, the US has a fragmented system with a mix of private and public insurance. This might introduce a selection bias due to health insurance choice. However, we limit this problem by focusing on unexpected health shocks, making a causal interpretation of our findings more plausible.

tion 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 intensifies gradually over time rather than appearing 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 (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 suggest that all three mechanisms likely play a role in mediating the relationship between health shocks and the increased probability of divorce. 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 viceversa (Brito and Contreras, 2024), 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 self-reported 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 health or activity limitations and marital dissolution (Vignoli et al., 2024). 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 satisfaction and stability. Carr and Springer (2010) highlighted that chronic illness can lead to both emotional strain and increased dependency, creating potential tensions within long-term partnerships. Umberson et al. (2011) explored how mental health declines can negatively impact spousal interactions, leading to emotional withdrawal and reduced relationship satisfaction. By examining mental health, 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 driving 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,[2] a harmonized version of the HRS developed by the RAND corporation based on the HRS core interviews.[3] 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 for married couples.

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 and serious physical illness on marital dissolution, we exclude couples in which either

[2]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.

[3]The HRS (Health and Retirement Study) is sponsored by the National Institute on Aging (grant number NIA U01AG009740) and is conducted by the University of Michigan.

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.

spouse reported having been diagnosed with cancer, stroke, or a heart attack at baseline, as further discussed in 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 the household's total number of hospital nights exceeds 4 (the median value for hospitalized households) in a given wave, conditional on no reported hospitalizations in the preceding wave.

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

The main dependent variable in our analysis represents the transition from a couple to a non-couple state between successive waves for each household. For simplicity, throughout the paper we will often refer to ”marriages” to mean individuals who are married or declare themselves to be in a couple, and to ”divorces” to mean individuals who are separated or divorced. A couple is defined to dissolve in separation/divorce if either spouse reported being divorced or separated since the prior wave.[4]. The dependent variable is coded as 0 if the original marriage or partnership remains intact across consecutive waves, and is assigned a value of 1 upon the occurrence of couple dissolution. The divorce prevalence we observe in our analytic sample, which amounts to 6.3%, is consistent with findings from other studies utilizing HRS data (Karraker and Latham, 2015).

2.3 Summary statistics

In columns (3) and (4) of Table 2, we present the summary statistics of the main analytic sample that serves as the basis for implementing the Stacked Differences-in-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 col-

[4]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 ever reporting the divorce—indicated by an increase in either partner's total number of marriages.

lege 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 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 1 and 2) with those of the final analytical sample. Reassuringly, the reported statistics indicate that the two samples are fairly 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 Differences-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 (TWFEDD) 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. In this scenario, bias may arise from using "already treated" units as controls,[5]

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

Table 2: Summary Statistics

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

Notes: Summary statistics are presented for two samples: couples observed at baseline (left columns) and the analytical sample (right columns). 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.

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.[6] Hence, we conduct our benchmark analysis both with and without corrective weights, with the former being our preferred specification.

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

[6]Corrective sample weights 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_a^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.

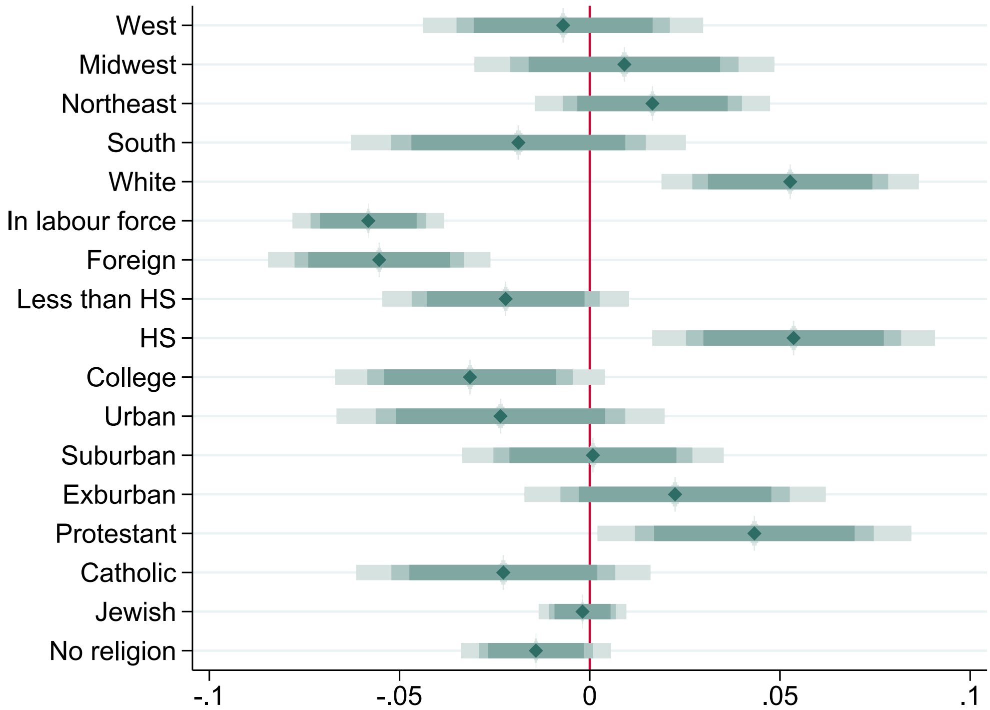
shock(Miller, 2023). Alternatively, the sample can be restricted to households that experience a health shock at some point (Costa-Font and Vilaplana-Prieto, 2022; Fadlon and Nielsen, 2021).

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.

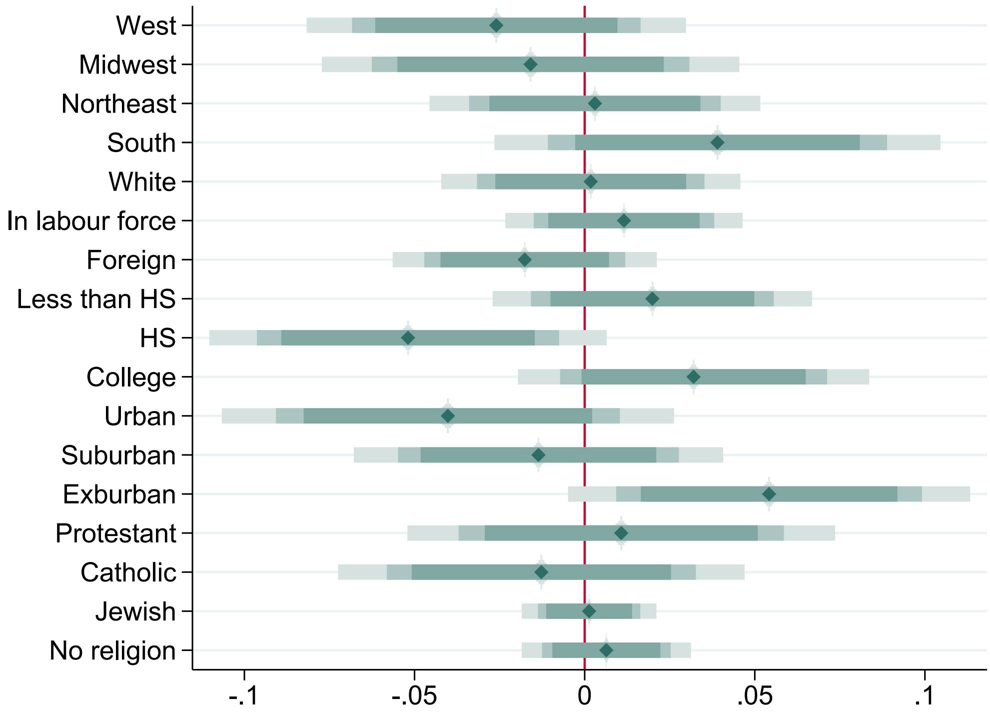
Figure 2, based on our stacked dataset, is analogous to Figure 1, with one key distinction: it reports coefficient estimates for a dummy variable that equals 1 if the individual belongs to a household treated in a particular sub-experiment and 0 otherwise. A visual comparison of Figures 2 and 1 reveals 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 and to be born in the U.S., and significantly less likely to have attained only a high school diploma or less, 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 the 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 5 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 HH 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-experiments. 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. 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. Additionally, 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 units in the stacked dataset, while the control group weights are obtained by dividing the ratio of treated units (people who adopt the treatment in a particular sub experiment divided by the total number of units that adopt the treatment) by the ratio of units serving as controls in a given sub-experiment, and then dividing by the total number of control units. 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:

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

where $Y_{s,t+1,d}$ is the outcome variable in t+1 for households s in sub-experiment d, which is equal to 1 since the moment that divorce happens, $T_{s,d}$ is a variable equal to 1 if the 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 are treated in the sub-experiment), $\theta_{s,d}$ is household s fixed effect in sub experiment d, $\gamma_{t,d}$ is time (wave) fixed effect in sub experiment d, and $\epsilon_{s,t,d}$ 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 x sub-experiment level. Clustering the standard errors at the unit x sub-experiment level would lead to over-rejecting the null hypothesis (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 effect (Table 4, Panel A) is not only statistically significant at the 1% level but also substantial in magnitude. To assess its size, it is helpful to express it relative to the share of divorced couples in our sample, which is approximately 6.3%. Dividing the estimated effect by this share reveals that experiencing a health shock increases the probability of divorce by approximately 19% of the mean divorce prevalence, regardless of whether corrective weights are applied in the estimations.

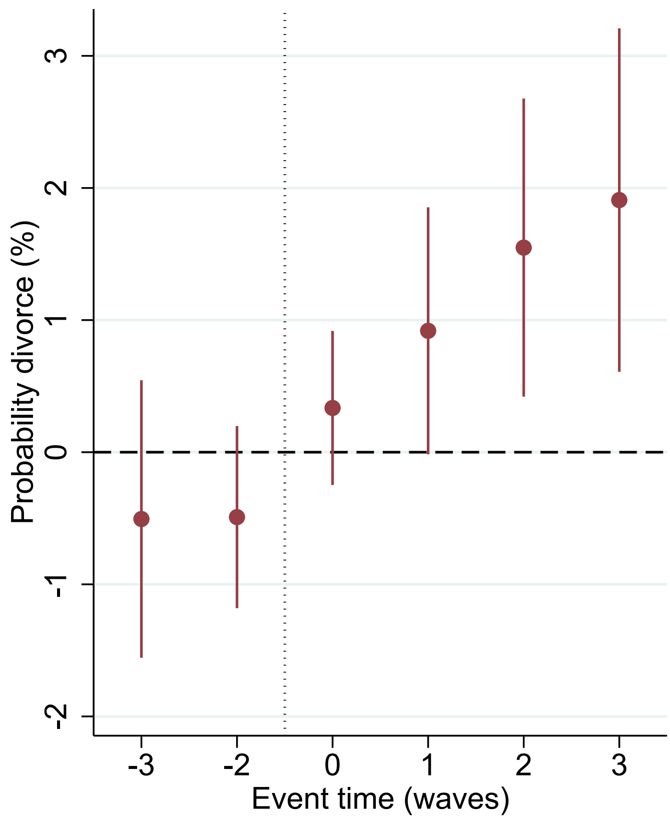
Table 4: The Impact of Health Shocks on the Probability of Divorce

	Without Weights (1)	With weights (2)
Panel A: Post-Treatment Average Effect		
Treated (=1)xPost (=1)	1.156**	1.178***
	(0.465)	(0.451)
Panel B: Event Studies		
Treated (=1)xEvent-time, -3 (=1)	-0.476	-0.506
	(0.543)	(0.535)
Treated (=1)xEvent-time, -2 (=1)	-0.449	-0.491
	(0.357)	(0.351)
Treated (=1)xEvent-time, 0 (=1)	0.334	0.335
	(0.315)	(0.297)
Treated (=1)xEvent-time, 1 (=1)	0.956*	0.919*
	(0.498)	(0.476)
Treated (=1)xEvent-time, 2 (=1)	1.505**	1.549***
	(0.585)	(0.575)
Treated (=1)xEvent-time, 3 (=1)	1.827***	1.909***
	(0.654)	(0.663)
Household x Subexperiment FE	Yes	Yes
Wave x 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 shows the impact of a health shock on the probability of divorce, defined as in section 2.2. Column (2) uses the weights proposed by Wing et al. (2024a) while column (1) does not use weights. Both specifications include household and wave fixed effects. Standard errors are clustered at the household level. Panel (A) is a linear combination of the post-treatment 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] To facilitate visual inspection, Figure 3 presents these event study results.

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



Note: The figure shows the impact of a health shock on the probability of divorce, defined as in section 2.2 using the weights proposed by Wing et al. (2024a). We plot the coefficients for the interaction between the event time dummies and the treatment dummy in Equation (1), with the corresponding 95 percent confidence intervals. The specification includes household and wave fixed effects.

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

[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.

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 frequently 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 consists 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 scores indicate a greater propensity for depressive symptoms.

Since each household includes a husband and a wife, the outcome variable is calculated as the average CESD-8 score of the two partners and then standardized to have a 0 mean and a standard deviation of 1. This standardized measure is used as the dependent variable in Column 1 of Table 5. Additionally, following Botoseneanu et al. (2023), we construct a dummy variable equal to 1 if at least one member of the couple has an (unstandardized) CESD-8 score greater than or equal to 4, and 0 otherwise. If data are missing for one individual in the couple, the dummy takes the value of the observed individual.. This serves as the dependent variable in Column 2 of Table 5.

The results reported in Table 5 convey a clear message: experiencing a health shock has a detrimental effect on mental health. Specifically, in households affected by a health shock, the average CESD-8 score increases by 0.135 standard deviations, while the probability that at least one member of the couple scores 4 or higher on the CESD-8 scale rises by 3.6 percentage points on average (Columns 1, Panel A). The latter represents approximately 27% of the mean prevalence of this indicator in our sample. Moreover, the effecta are persistent over time, and particularly large in the two waves following the health shock. (Column 2, Panel B).[8]

Table 5: The Impact of Health Shocks on Mental Health

	CES-Depression Scale (stand.) (1)	CES-Depression dummy (2)
Panel A: Post-Treatment Average Effect		
Treated (=1)xPost (=1)	0.135***	0.036***
	(0.035)	(0.013)
Panel B: Event studies		
Treated (=1)xEvent-time, -3 (=1)	0.023	0.002
	(0.039)	(0.016)
Treated (=1)xEvent-time, -2 (=1)	-0.002	-0.005
	(0.036)	(0.014)
Treated (=1)xEvent-time, 0 (=1)	0.108***	0.034**
	(0.038)	(0.014)
Treated (=1)xEvent-time, 1 (=1)	0.158***	0.044***
	(0.042)	(0.016)
Treated (=1)xEvent-time, 2 (=1)	0.157***	0.041***
	(0.044)	(0.016)
Treated (=1)xEvent-time, 3 (=1)	0.135***	0.026
	(0.048)	(0.017)
Mean Dependent variable	0	0.131
Number of households	1472	1472
Number of Observations	22197	22197

Notes: This table shows the impact of a health shock on mental health at the household level, using the weights proposed by Wing et al. (2024a). Column (1) is a standardized variable of the mean CES-D levels of the couple, while column (2) is a dummy variable equal to 1 if any member of the couple reports a score in the CES-D scale associated with depression. Both specifications include household and wave fixed effects. Standard errors are clustered at the household level. Panel (A) is a linear combination of the post-treatment coefficients. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4.2.2 Cognitive Impairment

Previous research indicates that severe health shocks can accelerate cognitive decline in old age, often resulting in a substantial deterioration of mental functioning (Schiele and Schmitz, 2023; Alfaro-Acha et al., 2006). This finding suggests that cognitive function may be a relevant mediator in the relationship between health shocks and couple dynamics.

[8]While the treatment effect on the depression probability becomes less precise by event time 3 (Column 2, Panel B), 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.

To measure cognitive function, we first leverage the availability in the HRS of a summary index referred to as COG27. This index aggregates the following measures: i) immediate word recall (range: 0–10), which asks the respondent to repeat 10 words read aloud by the interviewer; ii) delayed word recall (range: 0–10), which asks the respondent to recall the 10 words approximately 10 minutes after they were announced; iii) serial 7's test result (range: 0–5), which measures the number of correct subtractions of seven starting from 100 over five trials;[9] and iv) backward counting from 20 for 10 continuous numbers (range: 0–2). Two points are awarded if successful on the first try, one point if 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 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 any member of the couple scores below 12 and use this binary indicator as the dependent variable in Column 1 of Table 6.

Additionally, the HRS provides information on the diagnosis of memory-related diseases. Starting from wave 10, individuals are separately asked whether they have received an Alzheimer's diagnosis or a dementia diagnosis from a medical professional. From wave 4 onwards, HRS respondents are asked whether they have been diagnosed with any memory-related disease. Using these variables, we construct a dummy variable spanning waves 4 to 15 (the latest available wave), which equals 1 if any member of the couple responds "yes" to any of these questions and 0 otherwise. This binary indicator is used as the dependent variable in Column 2 of Table 6.

The results reported in Table 6 indicate that when a household experiences a health shock, the probability that at least one of its members' cognitive function deteriorates increases significantly and substantially. For example, a health shock increases the

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

Table 6: The Impact of Health Shocks on Cognitive Function

	Cognitive Impairment (1)	Memory Problem Disease (2)
Panel A: Post-Treatment Average Effect		
Treated (=1)xPost (=1)	0.042***	0.036***
	(0.016)	(0.008)
Panel B: Event studies		
Treated (=1)xEvent-time, -3 (=1)	-0.004	-0.003
	(0.016)	(0.006)
Treated (=1)xEvent-time, -2 (=1)	0.003	-0.003
	(0.015)	(0.004)
Treated (=1)xEvent-time, 0 (=1)	0.042**	0.017***
	(0.017)	(0.006)
Treated (=1)xEvent-time, 1 (=1)	0.055***	0.027***
	(0.019)	(0.008)
Treated (=1)xEvent-time, 2 (=1)	0.046**	0.047***
	(0.020)	(0.011)
Treated (=1)xEvent-time, 3 (=1)	0.025	0.052***
	(0.021)	(0.012)
Mean Dependent variable	0.174	0.109
Number of households	1306	1160
Number of Observations	17500	13748

Note: This table shows the impact of a health shock on the cognitive function at the household level using the weights proposed by Wing et al. (2024a). The dependent variable in Column (1) is a dummy equal to 1 if any member of the couple scores below 12 in the COG27 index, while the dependent variable in Column (2) is a dummy variable equal to 1 if any member of the couple has been diagnosed by a medical professional with a memory related disease. Both specifications include household and wave fixed effects. Standard errors are clustered at the household level. Panel (A) is a linear combination of the post-treatment coefficients. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

likelihood that at least one member of the couple exhibits symptoms of cognitive impairment by 4.2 percentage points on average throughout the event window (Panel A, Column 1), which corresponds to approximately 24% of the sample mean for this indicator. This effect is no longer statistically significant at standard levels of testing 3 waves after the health shock, but we cannot reject the null hypothesis that the difference in treatment effects between event times 3 and 0 is zero. [10] The probability of any member of the couple receiving a diagnosis of memory-related disease also increases significantly by 3.6 percentage points (Panel B, Column 1) on average following a health shock (about 33% of the prevalence of memory-related disease in the sample).

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

Moreover, this effect not only persists but also increases steadily and significantly in the years following the health shock (Panel B, Column 2).

4.2.3 Out-of-Pocket Medical Expenses

Finally, we explore whether health shocks induce financial strain by looking at out of pocket expenditures in Table 7. We use out-of-pocket expenditures as a proxy for financial strain, and, as expected, we find that out-of-pocket medical expenses significantly increase after a diagnosis of cancer, stroke or heart attack. Although not directly comparable, this result is consistent with Dobkin et al. (2018), who use an event-study approach and find that out-of-pocket expenses increase among non-elderly (50-59) HRS individuals who suffer non-pregnancy-related hospital admissions (their empirical analogue of an "adverse health shock"). As expected, we find that health shocks significantly increase out-of-pocket medical expenses, with this increase being particularly large in the period immediately following the shock.

Table 7: The Impact of Health Shocks Out-of-Pocket Expenditures

	Out of pocket medical expenditures
Panel A: Post-Treatment Average Effect	
Treated (=1)xPost (=1)	1049.778***
	(297.152)
Panel B: Event studies	
Treated (=1)xEvent-time, -3 (=1)	-34.457
	(234.481)
Treated (=1)xEvent-time, -2 (=1)	114.965
	(217.116)
Treated (=1)xEvent-time, 0 (=1)	1573.784***
	(329.178)
Treated (=1)xEvent-time, 1 (=1)	699.470**
	(350.908)
Treated (=1)xEvent-time, 2 (=1)	912.432**
	(375.683)
Treated (=1)xEvent-time, 3 (=1)	1013.426**
	(424.041)
Mean Dependent variable	2739.857
Number of households	1328
Number of Observations	17703

Note: This table shows the impact of a health shock on households' out-of-pocket medical expenditures using the weights proposed by Wing et al. (2024a). The specification includes household and wave fixed effects. Standard errors are clustered at the household level. Panel (A) is a linear combination of the post-treatment coefficients. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5 Robustness checks

In this section, we test the robustness of our results to variations to the length of the even window and the definition of health shocks.

First, we estimate equation 1 using a different event window length. Specifically, we employ 4 waves before and after the health shocks, as opposed to the 3 waves used in the benchmark specification. The use of a longer event window allows 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. Specifically, units treated in waves 4 and 11, which were included in the main specification, are not part of the sample in Table 8, as the treatment occurred too early or too late in the observation period. Furthermore, the control group comprises households that experienced the treatment more than 4 waves after the reference period of each sub-experiment, rather than 3 waves as specified in the main analysis. This further reduces the number of observations in the final sample. However, the results in Table 8 are quite similar to those reported in Table 4. Again, the pretreatment coefficient suggests the parallel trend and no anticipation assumptions hold. In our preferred specification in column (2) which includes the corrective sampling weights proposed by Wing et al. (2024a), the impact of the health shock increases the probability of divorce by approximately 1.7 percentage points (or 23%). These results are consistent with those obtained using our benchmark event window of three waves.

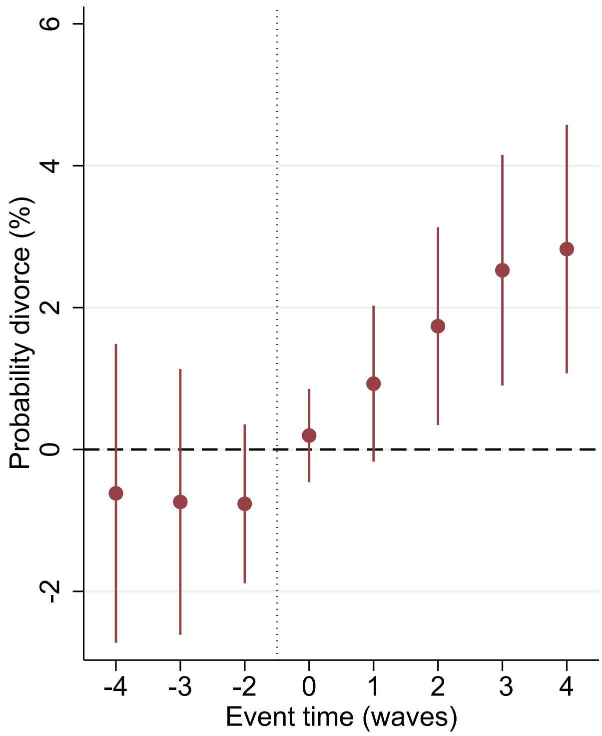
Table 8: The Impact of Health Shocks on the Probability of Divorce - Longer Event Window

	Without Weights (1)	With weights (2)
Panel A: Post-Treatment Average Effect		
Treated (=1)xPost (=1)	1.279**	1.643***
	(0.543)	(0.586)
Panel B: Event studies		
Treated (=1)xEvent-time, -4 (=1)	-0.542	-0.618
	(0.986)	(1.073)
Treated (=1)xEvent-time, -3 (=1)	-0.677	-0.739
	(0.832)	(0.954)
Treated (=1)xEvent-time, -2 (=1)	-0.583	-0.765
	(0.551)	(0.571)
Treated (=1)xEvent-time, 0 (=1)	0.037	0.197
	(0.269)	(0.335)
Treated (=1)xEvent-time, 1 (=1)	0.695	0.928*
	(0.524)	(0.560)
Treated (=1)xEvent-time, 2 (=1)	1.364**	1.738**
	(0.666)	(0.710)
Treated (=1)xEvent-time, 3 (=1)	2.009***	2.526***
	(0.774)	(0.828)
Treated (=1)xEvent-time, 4 (=1)	2.292***	2.825***
	(0.850)	(0.892)
Household x Subexperiment FE	Yes	Yes
Wave x Subexperiment FE	Yes	Yes
Weights	Yes	Yes
Observations	13761	13761
Number of clusters (Households)	830	830
Prevalence of divorce	0.071	0.071
Impact of health shock(percentage)	17.997	23.110

Notes: This table shows the impact of a health shock on the probability of divorce using a longer event window in each sub-experiment (4 waves before and after the health shock in each sub-experiment) to observe its evolution over a longer period of time. Column (2) uses the weights proposed by Wing et al. (2024a) while column (1) does not use weights. Both specifications include household and wave fixed effects. Standard errors are clustered at the household level. Panel (A) is a linear combination of the post-treatment coefficients. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

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

Figure 4: The Impact of Health Shocks on the Probability of Divorce - Longer Event Window



Notes: The figure shows the impact of a health shock on the probability of divorce using a longer event window (4 waves before and after the health shock) and the weights proposed by Wing et al. (2024a). We plot the coefficients for the interaction between the event time dummies and the treatment dummy in Equation 1, with the corresponding 95 percent confidence intervals. The specification includes household and wave fixed effects.

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 (cancer, stroke, or heart attack). Since, unfortunately, we lack data on the reasons for each hospitalization, this measure is expected to yield smaller estimated effects compared to our benchmark measure used in Table 4, which captures more substantial health issues. Reassuringly, however, the results in Table 9 qualitatively align with our main findings: we also find that experiencing a health shock increases the likelihood of couple dissolution under this alternative health shock definition.

Table 9: The Impact of Hospital Admissions on the Probability of Divorce

	Without weights (1)	With weights (2)
Panel A: Post-treatment average effect		
Treated (=1) X Post (=1)	1.122**	1.066**
	(0.548)	(0.540)
Panel B: Event studies		
Treated (=1)xEvent-time, -3 (=1)	0.256	0.085
	(0.647)	(0.698)
Treated (=1)xEvent-time, -2 (=1)	0.217	0.026
	(0.363)	(0.408)
Treated (=1)xEvent-time, 0 (=1)	0.161	0.242
	(0.368)	(0.395)
Treated (=1)xEvent-time, 1 (=1)	1.009*	1.008*
	(0.600)	(0.582)
Treated (=1)xEvent-time, 2 (=1)	1.408**	1.337**
	(0.676)	(0.641)
Treated (=1)xEvent-time, 3 (=1)	1.913**	1.681**
	(0.781)	(0.753)
Household x Subexperiment FE	Yes	Yes
Wave x Subexperiment FE	Yes	Yes
Observations	17857	17857
Number of clusters (Households)	1150	1150
Prevalence of divorce	0.093	0.093
Impact of health shock(percentage)	12.067	11.465

Notes: This table shows the impact of a health shock on the probability of divorce. The health shock is defined as an indicator equal to 1 if the total number of nights the couple spent in the hospital since the last wave exceeds four—the median value of the variable. Column (2) uses the weights proposed by Wing et al. (2024a) while column (1) does not use weights. Both specifications include household and wave fixed effects. Standard errors are clustered at the household level. Panel (A) is a linear combination of the post-treatment coefficients. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

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 sum 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) confirm in simulations. One of the key advantages of this method is its computational efficiency, as it only requires the estimation of a simple TWFE model. Furthermore, the imputation approach is conceptually intuitive, providing a clear and transparent link between the parallel trends and no-anticipation assumptions and the resulting estimator.

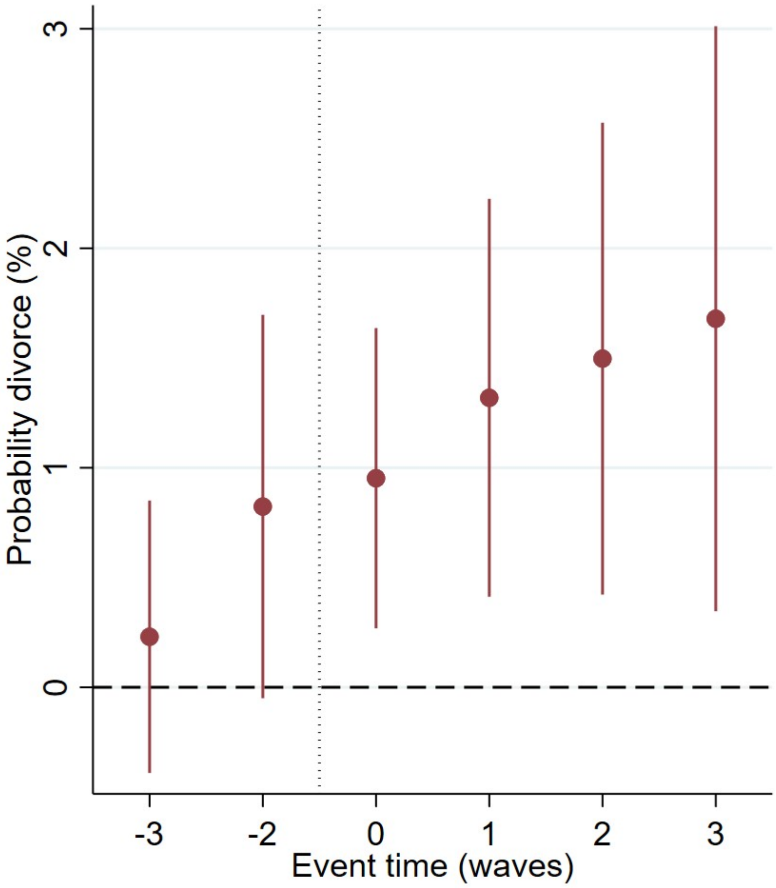
The results of this estimation are presented in Table 10, while the corresponding event study results are depicted in Figure 5. Reassuringly, the results obtained are comparable to our benchmark results.

Table 10: The Impact of Health 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	18.697

Notes: This table shows the estimated impacts 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 5: The Impact of Health Shocks on the Probability of Divorce - Borusyak et al. (2024) Estimator



Notes: This picture plots the estimated impacts of a health shock on the probability of divorce obtained using the estimator proposed by Borusyak et al. (2024)

6 Conclusion

This study provides causal evidence that unexpected health shocks significantly increase the likelihood of marital dissolution among older couples. Using longitudinal data from the Health and Retirement Study (HRS) and employing a stacked difference-in-differences methodology, we find that the probability of divorce rises by approximately 19% following a health shock. This effect intensifies over time, highlighting the long-term strain such events place on marital relationships.

Our findings also shed light on the mechanisms underlying this relationship. Health shocks negatively affect mental health, as evidenced by an increase in depressive symptoms measured using the CES-D scale. They also contribute to cognitive decline, with a higher likelihood of memory-related disease diagnoses, and impose financial strain, as indicated by increased out-of-pocket medical expenses. These mechanisms collectively provide a multidimensional perspective on how health shocks can destabilize marital dynamics in older households.

By highlighting the causal 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.

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," *Available here*.

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. (2024): "The caregiving penalty: Caring for sick parents and the gender pay gap," *Available here*.

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, X. and Zaremba, K. (2024): "Household Penalty," *Available at SSRN 4989822*.

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.

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.

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. (2015): "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.

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.

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, 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. (2024): "Partners' health and silver splits in Europe: A gendered pattern?" *Journal of Marriage and Family*.

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.