# Does Religious Participation Buffer the Financial Consequences of Health Shocks in Later Life?

August 2026


## Abstract

This paper uses severe later-life health shocks as a boundary test of religious participation as informal insurance: do broad pre-shock measures of religious-service attendance and private belief predict smaller financial and psychological deterioration? Using nine biennial waves of the Health and Retirement Study from 2006 through 2022, the analysis aligns respondents around the first newly reported cancer, stroke, or strictly defined heart attack and compares their outcome changes with those of eligible respondents who remain untreated in the same cohort-specific event window. At the first-report interview, affected respondents had a 0.516-unit increase in inverse-hyperbolic-sine out-of-pocket spending, a 9.6-percentage-point increase in catastrophic spending, and a 0.440-unit decline in transformed household wealth relative to untreated comparisons and the omitted interview approximately two years earlier. At affected-group reference medians, these coefficients correspond illustratively to about $879 higher out-of-pocket spending and $44,500 lower wealth (nonlinear back-transformations; Table 8). Spending and catastrophic risk remained elevated at subsequent interviews. Wealth also remained lower in the principal specification, although the estimate approximately four years later was attenuated after baseline-trajectory adjustment. Condition-specific analyses show especially persistent wealth losses after stroke; heart-attack estimates are harder to interpret because the financial outcomes diverge before first report. Excluding every treated first event involving heart attack leaves the pooled financial estimates at least as large and restores clean two-lead financial pre-event diagnostics. Neither weekly attendance nor standardised private belief produces a statistically robust pooled or condition-specific moderation pattern. Later attendance estimates remain conditional on outcome observation because weekly attendees have modestly higher follow-up availability.



Keywords: religion; health shocks; medical spending; household wealth; depression; ageing; Health and Retirement Study

JEL codes: D14; I12; J14; Z12; Z13

**Acknowledgements**

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. The study used data from the Health and Retirement Study (HRS), conducted by the University of Michigan. The authors acknowledge the HRS participants and research team for making these longitudinal data available for scientific research. The authors declare that they have no financial relationships, personal relationships, or other potential conflicts of interest that could have influenced the research reported in this article.

## 1. Introduction

A serious health event is not only a medical shock. It can simultaneously increase treatment expenditure, reduce earnings or household production, require caregiving, and force households to draw down liquid and non-liquid assets. The resulting burden may persist after the immediate episode of illness. For older adults, these consequences arise in a setting where labour-market adjustment is limited, medical expenditure is highly skewed, and the protection offered by insurance does not necessarily cover the broader loss of financial security or psychological well-being. The central economic question is therefore not simply whether illness increases medical spending, but how households absorb the multidimensional risk created by illness.

Religious institutions are plausible components of that adjustment process. Congregations can supply information, transportation, caregiving, charitable assistance, and access to social networks. Religious belief may also supply meaning and emotional resources during periods of uncertainty. These features motivate a familiar proposition in the economics of religion: religious participation may partly insure consumption or well-being against adverse shocks (Chen, 2010; Dehejia et al., 2007; Iyer, 2016). Yet the same mechanisms do not imply an unambiguous effect on observed medical expenditure. Better information and practical support could enable treatment and raise out-of-pocket spending, while informal assistance could reduce catastrophic burden or asset depletion. Private belief could operate mainly through psychological coping rather than through household finances.

This paper asks whether older adults with different levels of religious participation experience different medical, financial, and psychological trajectories after an acute health shock. It is framed as a boundary test of religious participation as informal insurance: broad attendance and belief measures may indicate access to social and coping resources, but they do not directly measure transfers, caregiving, transportation, information, or congregational capacity. The empirical question is therefore whether these pre-event indicators forecast smaller post-shock losses, not whether religious communities ever provide assistance.

The empirical analysis uses nine biennial Health and Retirement Study (HRS) waves from 2006 through 2022. The source panel contains 407,106 respondent-wave observations for 45,234 respondents. The data are organised around each respondent's first recorded health event. Changes among affected respondents are compared with changes among eligible respondents who remain untreated throughout the same event window, with the interview approximately two years before the first report serving as the omitted reference period. The main outcomes are transformed out-of-pocket medical spending, an indicator that medical spending exceeds 10% of positive household income, transformed household wealth, and depressive symptoms. The principal moderator is weekly-or-more religious-service attendance recorded before the event; a second analysis uses a standardised private-belief index in the smaller psychosocial sample.

The findings have a clear asymmetry. Relative to untreated comparison respondents and the omitted pre-event interview, serious health events are followed by large financial changes. At the first-report interview, transformed out-of-pocket spending is 0.516 units higher, the probability that medical spending exceeds 10% of income is 9.6 percentage points higher, and transformed household wealth is 0.440 units lower. Spending and catastrophic-spending risk remain elevated at the next two interviews. Wealth remains lower in the principal specification, although the approximately four-year estimate becomes smaller and statistically imprecise after baseline-characteristic-related trajectories are allowed to differ. Depressive symptoms also

increase, but their pre-event movement makes the psychological estimates descriptive rather than clean event-study effects.

Religious participation, however, does not produce a statistically robust protective pattern. Fully saturated models include the required event-time-by-religion lower-order interactions. The joint post-event tests for weekly attendance and private belief are nonsignificant for each financial outcome, and the confidence intervals remain wide. The appropriate conclusion is therefore not that religion has exactly no effect, but that the present design does not identify a stable average association between broad pre-shock religion measures and smaller post-shock losses.

The paper contributes in four ways. First, it connects the economics of religion with research on the economic consequences of health shocks. Second, it evaluates organisational participation and private belief separately, rather than treating religiosity as a single construct. Third, it triangulates across medical spending, catastrophic burden, household wealth, and depressive symptoms. Fourth, it tests whether pooled estimates conceal diagnosis-specific patterns by distinguishing cancer, stroke, strict heart attack, and multiple concurrent first reports. The contribution is not a claim that religion never provides support. It is evidence about the limits of broad pre-shock religion measures as predictors of average resilience after severe later-life health events.

## 2. Related literature and conceptual framework

### *2.1 Religion as an economic institution*

Modern economics treats religion not only as a set of beliefs but also as an institution that organises networks, collective action, charitable activity, and benefits shared within a group (Iyer, 2016). This perspective creates a direct link between religious participation and household risk management. Dehejia et al. (2007) find that involvement with religious organisations is associated with smoother consumption or happiness following income shocks, while Chen (2010) interprets increased religious involvement during the Indonesian financial crisis as a response to unmet insurance needs. Ager and Ciccone (2018) document greater religious-community membership in areas exposed to common agricultural risk. Most directly, Auriol et al. (2020) randomly provided commercial funeral insurance to Pentecostal church members in Ghana and found that insured participants reduced religious giving, consistent with formal insurance substituting for perceived religious insurance. Together, these studies establish that religion can have an economic insurance function, although attendance is not itself a direct measure of transfers or assistance.

Risk can also change religious behaviour. Bentzen (2019) finds that exposure to natural disasters is followed by higher religiosity, with religious coping emerging as an important explanation. This creates a central identification problem: a shock may mobilise religious participation or private coping rather than merely reveal the protective effect of religion. Participation is also chosen and may reflect health, preferences, family structure, personality, and access to non-religious networks. Experimental evidence confirms that religious content can alter religiosity and selected economic outcomes, but it does not establish that ordinary differences in service attendance have the same meaning or effect (Bryan et al., 2021). The present paper therefore measures attendance and belief before the recorded health event, accounts for stable respondent differences, and examines whether outcomes were already changing before the event.

### 2.2 Health shocks and household economic risk

Health shocks generate both direct and indirect economic losses. Gertler and Gruber (2002) show that households may be unable to protect consumption fully against severe illness, while Dobkin et al. (2018) document sharp increases in out-of-pocket expenditure and broader financial distress after hospital admission. In later life, medical-spending risk is persistent and highly unequal across households (Jones et al., 2018). Blundell et al. (2024) further show that consumption and out-of-pocket medical spending among older U.S. households respond to transitory health shocks, highlighting the interaction between health states, spending needs, and available insurance. Health news can also alter inter vivos transfers and bequest decisions (Kværner, 2023). Beyond money, health shocks change labour supply, household expenditure, and spousal caregiving time, making illness a household time-allocation shock as well as a medical event (Zilio et al., 2025). This literature motivates the paper's emphasis on medical spending, catastrophic burden, wealth, and psychological well-being rather than expenditure alone.

The distinction matters for interpretation. A reduction in medical spending after a shock may indicate financial protection, but it may also indicate forgone care. Conversely, an increase in spending may reflect improved access to necessary treatment. Changes in wealth can capture the cumulative financing of care and other income or caregiving disruptions that a spending measure misses. A credible test of religious protection must therefore examine the joint pattern of expenditure, catastrophic burden, and asset depletion rather than attach a protective interpretation to any single outcome.

### 2.3 Religion, social support, and psychological coping

Research in social epidemiology and psychology provides complementary mechanisms. Church-based relationships may supply emotional and practical support that differs from non-religious social ties (Krause, 2002). Religious coping may provide meaning, comfort, and a framework for responding to stress, but its effects depend on the form of coping and the context of the event (Pargament & Abu Raiya, 2007). Meta-analytic evidence links religiousness to depression on average while also emphasising variation across measures and stressful circumstances (Smith et al., 2003). HRS evidence likewise shows that public attendance and private religious importance need not have the same association with later-life outcomes (Idler et al., 2017).

Longitudinal evidence also warns against interpreting attendance-health associations mechanically. Depression may cause people to withdraw from religious activity, producing reverse causality (Maselko et al., 2012), and physical, psychological, and social circumstances can predict later religious-service attendance among older HRS respondents (Cowden et al., 2022). Attendance and the subjective importance of religion can also have different relationships with later-life depressive symptoms (Orr et al., 2019). More recent economic studies of pandemic-related shocks find evidence consistent with religiosity or the ability to participate in religious activities moderating mental-health deterioration in some settings (Bahal et al., 2023; Iyer et al., 2025). These findings reinforce the possibility that any buffering relationship depends on the type of shock, the institutional setting, and the dimension of religious engagement. The present paper therefore separates organised participation from private belief and measures both before the recorded health event.

### 2.4 Hypothesised channels, expected patterns, and contribution

The framework yields three channels. A healthcare-access channel may raise post-shock medical spending if congregational participation improves information, transport, treatment adherence, or provider navigation. An informal-insurance channel predicts lower catastrophic burden and smaller wealth losses if religious networks provide transfers, food, caregiving, transport, or help with ordinary expenses. A coping channel predicts smaller increases in depressive symptoms among respondents with stronger pre-shock belief. Organisational attendance and private belief need not operate through the same channel.

A coherent protective pattern would therefore combine lower catastrophic risk, smaller wealth loss, and/or less psychological deterioration, even if medical spending is unchanged or higher. The analysis does not observe congregational transfers, caregiving, navigation, treatment adherence, or religious coping directly, so these channels organise the expected signs rather than constitute estimated mediation mechanisms.

Existing studies show that religious organisations may provide insurance-like support, risk can itself change religiosity, serious illness reshapes household resources, and religious participation is associated with social and psychological resources in later life. These strands have rarely been integrated in a design that measures religion before an acute health event and follows medical spending, catastrophic burden, wealth, and depressive symptoms over the same event window. This paper addresses that gap as a boundary test: it asks whether broad measures of organisational participation and private belief forecast smaller post-shock losses while recognising that the underlying assistance mechanisms are not directly observed.

## 3. Data and measurement

### 3.1 Health and Retirement Study

The analysis combines information from the HRS Core and Fat Files, the RAND HRS Longitudinal File, RAND detailed imputations, and the Psychosocial and Lifestyle Questionnaire. The structure, sampling design, and principal data products of the HRS are described in Sonnega (2025). The study uses nine biennial survey waves from 2006 through 2022. The harmonised source panel contains 407,106 respondent-wave observations for 45,234 respondents before event-window and outcome-specific restrictions. In the principal transformed-spending model, stacking generates 464,209 respondent-wave records from 27,994 distinct respondents.

Private religious belief is drawn from the Psychosocial and Lifestyle Questionnaire. Since 2006, this questionnaire has generally been given to a randomly selected half of eligible respondents, with the same respondent usually completing it every four years (Smith et al., 2023). The belief models consequently contain fewer observations than the attendance models.

### 3.2 Incident health shocks and event time

The health event is the first observed newly reported cancer diagnosis, stroke, or heart attack that satisfies the study's strict incident-event rules. The heart-attack definition requires a newly reported event together with the available doctor-confirmation and timing information, reducing reliance on retrospective ever-had-a-heart-attack reports. The first strict composite event is classified into mutually exclusive categories using

condition-specific first-report years: cancer only, stroke only, strict heart attack only, or multiple conditions first reported concurrently. Among the 5,319 treated respondent-cohort events in the principal transformed-spending and wealth model, 2,398 are cancer only, 1,687 stroke only, 1,034 strict heart attack only, and 200 multiple concurrent events.

The timing is measured relative to the interview in which the condition is first reported, rather than by a common calendar year. Respondents experience health events in different years, so calendar years cannot serve as a shared event clock. Event 0 is the first-report interview. Because HRS interviews are usually about two years apart, event -1 is approximately two years before the first report and is the reference period; event -2 is approximately four years before; event +1 is approximately two years after; and event +2 is approximately four years after. These intervals are approximate because the diagnosis may have occurred at any point between interviews. Tables and figures retain the event-time notation, while the narrative translates it into ordinary time. Cohort-by-calendar-year fixed effects account for conditions common to respondents observed in the same calendar year.

### *3.3 Religion measures*

The main religion measure is an indicator for attending religious services weekly or more often. Weekly attendance is used because it represents routine organisational participation and gives a readily interpretable contrast between regular and less regular participants. Attendance is recorded in the pre-shock reference wave. Measuring it before the recorded event reduces a direct contemporaneous relationship between illness and attendance, although symptoms that began before diagnosis may already have reduced the ability to attend. Year-specific tabulations show a smooth distribution across the underlying four attendance categories. Weekly-or-more attendance is 39.4% in 2012, 39.5% in 2014, and 37.6% in 2016, with a gradual decline from 41.6% in 2006 to 35.8% in 2020. Categorical attendance specifications were also examined and did not produce a consistent protective pattern.

The second religion measure is a four-item mean index of private religious belief recorded before the shock in the Leave-Behind sample. The items concern belief in divine oversight, belief in a larger plan, the integration of religious beliefs into daily life, and strength or comfort from religion. To prevent repeated stacked records from receiving extra influence in the scale transformation, the index is standardised across 44,229 unique respondent-anchor-year observations. The raw index has a mean of 4.943 and a standard deviation of 1.453; the reported coefficients therefore describe the difference associated with a one-standard-deviation increase in pre-shock belief. The belief models are smaller because the psychosocial questionnaire is administered to rotating subsamples.

### *3.4 Outcomes*

Out-of-pocket medical spending is expressed using the inverse hyperbolic sine, a logarithm-like transformation that accommodates observed zero spending and reduces the influence of extremely large amounts. Zeros are not recoded as missing or dropped: at the reference interview, 16.1% of observed out-of-pocket values are zero in both the affected and eligible-comparison groups. Records with genuinely missing spending remain outcome-specific missing observations. For readability, the paper refers to this outcome as transformed out-of-pocket spending. A positive coefficient means that spending increased relative to the

comparison period. Because the transformation changes the scale, the coefficients should not be interpreted mechanically as exact percentage changes.

Catastrophic medical spending is defined as out-of-pocket medical spending above 10% of household income. The indicator is defined only when household income is positive and nonmissing; records with zero, negative, or missing income are set to missing for this outcome rather than being classified as catastrophic by construction. Because the outcome is coded as zero or one, its coefficients are percentage-point changes. The ratio can remain sensitive to very low positive income, so it is treated as supporting evidence and interpreted jointly with transformed spending and wealth rather than on its own. The present manuscript does not infer household insolvency from crossing the 10% threshold.

Household wealth is expressed using the same inverse-hyperbolic-sine transformation because wealth can be zero or negative. The paper refers to this as transformed household wealth. A negative coefficient means that wealth declined relative to the comparison period, but the coefficient is not an exact percentage change. Wealth is measured for the household but attached to respondent records, which creates possible dependence between spouses and is discussed in the limitations.

The financial variables are analysed in the monetary units supplied by the harmonised HRS/RAND files. Household income is measured in reported dollars; across unique respondent-anchor-year records, its median is approximately $40,167 and its interquartile range is approximately $19,220 to $80,760. Cohort-by-calendar-year fixed effects absorb price-level movements common to treated and comparison respondents in the same survey year. For readability, Table 8 back-transforms the principal inverse-hyperbolic-sine coefficients at representative affected-group reference values using sinh[asinh(y*) + beta] - y*. These are illustrative original-scale equivalents rather than average marginal effects, exact percentage changes, inflation-adjusted causal effects, or population-representative dollar estimates.

Psychological well-being is measured with the Center for Epidemiologic Studies Depression (CES-D) symptom score. Higher scores indicate more depressive symptoms. The CES-D lead is statistically significant, indicating that symptoms were already changing before the event was recorded. CES-D is therefore a secondary descriptive outcome; the paper's primary evidence concerns medical spending and wealth.

### *3.5 Analytical samples and sample progression*

The source panel contains 45,234 respondents before incident-event, event-window, and outcome-specific restrictions. Strict incident coding identifies 5,394 first health-event waves before model-specific exclusions. The principal transformed-spending and wealth models retain 27,994 respondents and 5,319 treated respondent-cohort events; the corresponding attendance models retain 27,977 respondents and 5,300 treated events, while the private-belief models retain 20,436 respondents and 1,965 treated events. Appendix Table B1 summarises the progression from the source panel to the main analytical samples. Because the outcome-specific samples overlap, the final rows are not additive.

### *3.6 Reference-wave characteristics*

Table 1 describes characteristics measured at event -1 - the interview approximately two years before first report - for respondent-cohort units in the exact principal transformed-spending and wealth estimation sample. Using an absolute standardised mean difference of 0.10 as a descriptive imbalance flag (Austin, 2009), affected units differ from eligible comparisons in age, private insurance, self-rated health, ADL limitations, chronic conditions, baseline CES-D, and household size. Age, self-rated health, and chronic-condition count have absolute SMDs above 0.25. Baseline medical spending and wealth differences are smaller, while weekly attendance and standardised private belief are closely aligned. Inverse control-reuse weighting does not remove the age and health differences but leaves the religion measures closely aligned. These are differences in reference-wave levels rather than direct tests of parallel outcome trends: respondent-by-cohort fixed effects absorb stable level differences, while Model C allows event-time trajectories to vary with the observed pre-event characteristics.

## 4. Empirical strategy

### *4.1 Stacked event-study specification*

The analysis creates a separate event cohort around each first observed health event. Within a cohort, outcomes for affected respondents are followed from two interviews before to two interviews after the first-report interview - approximately four years on either side in the biennial HRS. Comparison respondents must have no previously recorded qualifying event, be observed in the corresponding cohort window, and remain free of a first reported cancer, stroke, or qualifying heart attack throughout that window. No respondent is used as a control after treatment. Eligible respondents can serve as comparisons in more than one cohort, and a small share later experience an event in a more distant cohort. The preferred specification includes respondent-by-cohort fixed effects and cohort-by-calendar-year fixed effects, with standard errors clustered by respondent. Household-clustered inference is reported as a robustness check for the financial outcomes and their religion interactions. The respondent-by-cohort effects absorb characteristics that remain constant within a respondent's participation in a given cohort, while the cohort-by-calendar-year effects absorb common time shocks within treatment cohorts. This cohort-aligned design follows the logic of stacked event studies and responds to the broader concern that pooled event-study estimators with staggered treatment timing can mix effects across cohorts and periods when treatment effects are heterogeneous (Callaway & Sant'Anna, 2021; Sun & Abraham, 2021; Borusyak et al., 2024).

The saturated moderation model is:

$$Y_{ist} = \alpha_{is} + \lambda_{st} + \Sigma_{k\neq -1}\, \beta_k(E^k_{ist} \times Shock_{is}) + \Sigma_{k\neq -1}\, \delta_k(E^k_{ist} \times R_i,pre) + \Sigma_{k\neq -1}\, \gamma_k(E^k_{ist} \times Shock_{is} \times R_i,pre) + \varepsilon_{ist}.$$

Here, i identifies the respondent, s identifies the treatment cohort, and t identifies calendar time. Event time k is measured relative to the first reported health event, with event -1 omitted. The beta coefficients measure the affected-versus-comparison difference at each event time relative to the omitted pre-event interview; they are therefore difference-in-differences parameters rather than raw within-person changes. The delta terms allow respondents with different pre-shock religion measures to follow different event-time profiles even in the absence of treatment. The gamma coefficients are the moderation parameters of interest: they show whether the affected-versus-comparison trajectory differs by weekly attendance or by standardised private belief. The estimation includes all lower-order and triple-interaction terms required by the saturated moderation model.

The event-study estimates remain associational. Fixed effects remove stable respondent differences within a cohort comparison, but they do not remove time-varying confounding, symptoms that begin before formal diagnosis, differential survival, or selective post-event observation. The health-shock coefficients are therefore described as event-associated trajectories rather than as unqualified causal effects, and the religion coefficients are interpreted as differential associations rather than causal effects of participation.

### *4.2 Baseline-characteristic adjustment*

Table 5 distinguishes three specifications. Model A estimates the preferred event-study model on the full outcome-specific sample. Model B re-estimates the same model on respondents with complete pre-shock covariates, thereby separating complete-case selection from statistical adjustment. Model C uses exactly the Model B sample and adds interactions between event time and pre-shock age, income, wealth, private insurance, self-rated health, limitations in activities of daily living, chronic conditions, partnership status, and household size. These event-time interactions allow trajectories to differ with baseline characteristics; entering only time-invariant covariate levels would add no information because they are absorbed by respondent-by-cohort fixed effects.

### *4.3 Diagnostics, sample accounting, control reuse, and post-event availability*

The preferred window reports one pre-event coefficient: event -2, approximately four years before first report, relative to event -1, approximately two years before first report. A nonzero lead indicates that the outcome was already changing before the condition was recorded and weakens a causal interpretation. A supplementary expanded window adds event -3, approximately six years before first report, and jointly tests events -3 and -2. This improves the observable pre-trend assessment, but it cannot rule out a level shift or nonlinear deterioration beginning between events -2 and -1. Moreover, event -1 is not guaranteed to be a biologically untreated baseline because the condition may begin after that interview but before it is first reported at event 0. The financial outcomes form the primary evidence; CES-D is secondary because its pre-event coefficients are statistically significant.

Tables 2–4 report stacked observations, distinct respondents, and treated respondent-cohort events. The principal transformed-spending model contains 464,209 observations, 27,994 respondents, and 5,319 treated events. The corresponding attendance model contains 462,084 observations and 5,300 treated events, while the belief model contains 175,074 observations and 1,965 treated events. These counts distinguish the large stacked record totals from the number of independent respondents and treated events.

The principal event-study models are not weighted by HRS sampling or attrition weights. The estimand therefore concerns the observed eligible analytical panel rather than a nationally representative population average. The inverse control-reuse weights used in robustness analyses address repeated contribution to the stacked comparison pool; they are not survey weights. A single cross-wave sampling weight is not retained in the analytical stacks, and the rotating Leave-Behind belief sample further complicates a common population-weighted specification.

Control reuse is substantial. Among 27,025 respondents who serve as controls at least once, the mean number of control cohorts is 4.37, the median is four, and the maximum is eight. About 86.1% appear in

more than one control cohort and 46.3% appear in at least five. The principal estimates are unweighted with respect to reuse. A sensitivity analysis assigns each control respondent a weight equal to the inverse of the number of cohorts in which that person serves as a control; treated observations retain unit weight. Appendix Table B5 additionally tests whether weekly attendance or standardised private belief is concentrated among frequently reused controls. For that diagnostic, religion is attached at the first control appearance with a nonmissing measure, and reuse is grouped as one, two to four, or five to eight control cohorts.

Appendix Table B6 examines whether post-event analytical availability differs by pre-shock religiosity among treated respondents. Administrative right-censoring is handled by restricting the approximately two-year diagnostic to shock cohorts through 2020 and the approximately four-year diagnostic to cohorts through 2018. Availability is defined as a nonmissing outcome in the analysis data at the scheduled event time. The available Tracker extract contains aggregate linkage summaries but not respondent-level death and exit statuses linked to religion, so the diagnostic cannot distinguish death from non-death nonresponse.

### *4.4 Statistical inference and precision*

For each religion measure and financial outcome, the analysis reports event-specific respondent-clustered confidence intervals and a joint Wald test of the interaction coefficients at event 0, event +1, and event +2. It also reports the equally weighted average of those three post-event coefficients, its delta-method standard error, and minimum detectable effects at 80% and 90% power. Multiplicity is organised by inferential role. The primary confirmatory family contains the six pooled financial joint tests - three outcomes for each of the two religion measures - and Holm adjustment across all six yields adjusted p values of 1.000. The condition-specific analysis is secondary and applies Holm adjustment across the nine exclusive-disease financial tests within each moderator. As a conservative sensitivity, a single Holm adjustment across all 24 pooled and exclusive-disease financial joint tests yields a minimum adjusted p value of .595. Robustness re-estimations are not counted as additional hypotheses because they re-estimate the same substantive contrasts under alternative samples, weights, or covariance estimators.

Because the study did not prespecify a defensible smallest effect of substantive interest, equivalence tests are not used to support the main conclusions. Interpretation instead relies on pointwise confidence intervals, joint post-event tests, multiplicity-adjusted p values, and minimum detectable effects.

### *4.5 Condition-specific sensitivity analysis*

To assess whether the pooled estimates conceal diagnosis-specific trajectories, the binary shock indicator is replaced by the four mutually exclusive first-event categories. The main condition-specific model estimates separate event-time differences for each diagnosis against the common eligible untreated comparison pool. The moderation models add the complete lower-order and triple interactions between event time, condition category, and weekly attendance or standardised private belief. Formal Wald tests assess whether the cancer, stroke, and heart-attack post-event paths are equal and whether each diagnosis-specific set of post-event moderation coefficients is jointly zero. The multiple-concurrent category is reported descriptively because it

contains only 200 principal-sample events and 55-58 treated events in the belief models. A secondary sensitivity compares cancer-only events with an acute vascular group comprising stroke-only, heart-attack-only, and concurrent stroke-heart events, excluding events that concurrently include cancer. Because the strict heart-attack subgroup fails the expanded financial pre-event tests, an additional pooled sensitivity excludes every treated first event involving heart attack, including concurrent first reports, while retaining the common eligible comparison pool.

## 5. Results

### *5.1 Financial consequences of health shocks*

Throughout this section, the first-report interview refers to the HRS interview in which cancer, stroke, or a qualifying heart attack is first recorded. The next interview is described as approximately two years later and the second subsequent interview as approximately four years later. These intervals are approximate because the health event occurred sometime between interviews.

Table 2 and Figures 1-3 show a clear financial pattern. Relative to eligible respondents who remained untreated and to the omitted interview approximately two years before first report, affected respondents experienced a sharp increase in medical spending and catastrophic-spending risk together with a decline in household wealth. The spending response was largest at the first-report interview and then diminished, whereas the principal wealth difference remained pronounced across the observed follow-up period.

At the first-report interview, transformed out-of-pocket spending increased by 0.516 units (SE 0.048), catastrophic-spending risk increased by 9.6 percentage points (SE 0.8 percentage points), and transformed household wealth declined by 0.440 units (SE 0.080). All three estimates are statistically significant at the 1% level. The spending and wealth coefficients are on an inverse-hyperbolic-sine scale and therefore should not be interpreted mechanically as exact percentage changes.

The financial deterioration persisted beyond the first-report interview. Approximately two years later, transformed spending remained 0.208 units higher, catastrophic-spending risk 5.3 percentage points higher, and transformed wealth 0.467 units lower than the affected-versus-comparison difference at the omitted pre-event interview. The adverse differences were still present approximately four years later, although the spending response had receded from its initial peak. The longer-run wealth result is the least stable component of the principal pattern and becomes statistically weaker after allowing baseline characteristics to predict event-time trajectories (Section 5.4).

To make the transformed coefficients more interpretable, Table 8 reports nonlinear back-transformations at affected-group reference values. At the reference medians, the first-report estimates correspond illustratively to about $879 higher out-of-pocket spending and $44,500 lower household wealth. These are magnitude aids rather than average marginal effects: the implied dollar difference varies with the baseline value, as shown in Appendix Table B15.

The expanded pre-event diagnostic uses interviews approximately six and four years before first report, with the interview approximately two years before first report as the reference. For transformed spending, catastrophic-spending risk, and transformed wealth, the two leads are neither individually nor jointly

significant (joint p values .234, .440, and .552). This strengthens the interpretation of a financial break around first report, but it cannot rule out deterioration within the final two-year interval or establish that the omitted interview is a biologically untreated baseline. CES-D fails the same diagnostic and is therefore treated separately.

### *5.2 Psychological changes around the health event*

Depressive symptoms follow a different trajectory. CES-D is 0.111 points lower approximately four years before first report than at the omitted interview approximately two years before it, indicating that psychological deterioration was already under way before the condition was recorded. CES-D then rises by 0.214 points at first report and remains elevated at the two subsequent interviews (Table 2; Appendix Figure A1).

The post-event increase is substantively consistent with psychological strain surrounding serious illness, but the significant pre-event movement prevents interpretation as a clean change beginning at the first-report interview. CES-D is therefore retained as a descriptive secondary outcome rather than used to anchor the identification claims.

### *5.3 Does religious participation change the financial consequences?*

The central question is whether the affected-versus-comparison financial deterioration was smaller among respondents who attended religious services weekly or held stronger private religious beliefs before the event. The fully saturated specifications allow religion groups to differ at each event time rather than imposing a common post-event effect. Across the financial outcomes, however, the interaction estimates change sign across interviews and their pointwise confidence intervals include zero.

For weekly attendance, the first-report interactions are small relative to the main health-shock effects: 0.075 for transformed spending, 1.7 percentage points for catastrophic spending, and effectively zero for transformed wealth. Approximately two years later, the spending and wealth interactions become negative, but the estimates remain imprecise. Joint tests across first report and the two follow-up interviews do not reject no attendance moderation for spending (p=.428), catastrophic risk (p=.386), or wealth (p=.492). The wealth confidence interval approximately two years later spans both economically meaningful protection and adverse association, illustrating why the result is inconclusive rather than evidence of exact equivalence (Tables 3 and 7).

Private belief produces the same broad conclusion. The event-specific financial interactions are close to zero relative to their standard errors, and the joint post-event p values range from .559 to .853 across the three outcomes. The belief analysis uses the smaller Psychosocial Questionnaire sample, so its confidence intervals are wider and smaller protective associations are especially difficult to exclude (Tables 4 and 7).

Taken together, the estimates do not show the coordinated pattern expected under broad informal insurance: neither weekly attendance nor stronger private belief consistently reduces catastrophic exposure and wealth loss across the same post-event horizons. The evidence therefore does not identify a robust average buffer, but the confidence intervals and power calculations also caution against interpreting the estimates as proof that all religion-related support is absent.

### *5.4 Robustness and statistical precision*

Table 5 separates complete-case selection from adjustment for differential baseline-related trajectories. Restricting the analysis to complete cases changes the estimates little, suggesting that the covariate-complete subsample is not driving the main pattern. When event-time profiles are allowed to vary with pre-shock age, income, wealth, insurance, health, partnership status, and household size, the first-report transformed-spending coefficient rises from 0.516 to 0.572 and remains elevated at both follow-ups. The wealth trajectory attenuates, however, and the approximately four-year estimate loses statistical significance. Persistent medical-cost exposure is therefore more robust than the longest-run wealth estimate.

The stacked design repeatedly uses some respondents as controls, but inverse control-reuse weighting moves both the principal health-shock estimates and the attendance interactions negligibly (Appendix Table B4). This result matters because weekly attendance is modestly more common among heavily reused controls; the weighted specification shows that the reuse imbalance is too small to explain the moderation null.

Comparator contamination is also limited. Most control respondents are never treated in the observed stack, and those who are treated later are at least eight years from their recorded event when they first serve as controls. Removing every control appearance from respondents who are ever treated leaves the outcome trajectories essentially unchanged (Appendix Table B9).

Because spending, income, and wealth are household-linked measures, the models were also re-estimated with household-clustered inference. The resulting standard errors are virtually identical to respondent-clustered values for the main effects and religion interactions, so the inferential conclusions do not depend on the clustering level (Appendix Table B7).

The direct reuse diagnostic confirms a modest attendance gradient (Cohen's d=0.19) but no meaningful belief gradient (d=0.02). Together with the weighted re-estimation, this reduces the plausibility that differential control reuse mechanically washes out a protective religion interaction (Appendix Tables B4-B5).

Differential follow-up availability is a more consequential qualification. Weekly attendees are approximately 3.5-4 percentage points more likely to have observed spending and wealth outcomes at the two post-event horizons after cohort adjustment, whereas private belief is unrelated to availability. Because respondent-level death and other exit statuses cannot be separated in the analysis extract, the direction of the resulting selection is unknown. Later attendance interactions must therefore be read as conditional on outcome observation (Appendix Table B6).

The power calculations make the null result more interpretable. At 80% power, the minimum detectable average post-event effects for weekly attendance are 0.256 transformed-spending units, 3.9 percentage points of catastrophic spending, and 0.440 transformed-wealth units. The corresponding belief MDEs are 0.179, 3.0

percentage points, and 0.313. Thus, the design can reject large average buffers but remains unable to rule out smaller relationships, especially for wealth and in the smaller belief sample (Table 7).

Multiplicity adjustment does not change this conclusion. Holm correction across the six pooled financial joint tests yields adjusted p values of 1.000. A more conservative adjustment across all pooled and exclusive-disease financial joint tests produces a minimum adjusted p value of .595, so no isolated result becomes persuasive when the full moderation exercise is considered.

### *5.5 Condition-specific financial and religion-moderation patterns*

The pooled financial deterioration is not driven by a single diagnosis. Cancer-only events are followed at first report by a 0.526-unit increase in transformed spending, an 8.8-percentage-point increase in catastrophic risk, and a 0.517-unit decline in transformed wealth. These adverse differences persist, although the approximately four-year cancer wealth estimate is smaller and borderline. Stroke-only events show a similarly large immediate spending and catastrophic-risk response, but the distinctive feature is wealth: the transformed wealth difference widens from -0.417 at first report to -0.685 approximately two years later and -0.711 approximately four years later (Appendix Table B10).

Strict heart-attack estimates are less interpretable because transformed spending, catastrophic risk, and wealth all fail the expanded pre-event diagnostic (joint p values .020, .001, and .017). The coefficients therefore describe a trajectory surrounding first report rather than a clean onset contrast. The 200 multiple-concurrent events are also imprecisely estimated. Although the cancer, stroke, and heart-attack profiles look substantively different, formal joint tests do not reject equality of their post-event paths for the three financial outcomes (Appendix Tables B10 and B13).

The heart-attack exclusion sensitivity shows that this pre-event problem does not drive the pooled findings. Removing all 1,161 first events involving heart attack leaves 4,158 treated events. The first-report coefficients remain large - 0.565 for transformed spending, 9.7 percentage points for catastrophic risk, and -0.485 for transformed wealth - while the two-lead financial pretrend tests become comfortably insignificant. Attendance and belief joint tests remain nonsignificant throughout (Appendix Table B16).

Condition-specific moderation also fails to reveal a coherent religious buffer. No exclusive-disease financial attendance test survives Holm correction; the closest is stroke-related wealth, with an unadjusted joint p value of .102. For private belief, the strongest apparent result is heart-attack catastrophic spending, but it is preceded by a significant pre-event interaction and does not survive multiplicity adjustment. Combining stroke and heart attack into an acute vascular group yields the same substantive conclusion (Appendix Tables B11-B13).

## 6. Discussion

### *6.1 Health shocks and later-life financial resilience*

The increase in medical spending and catastrophic risk, together with lower household wealth, is consistent with evidence that serious illness remains only partially insured (Gertler & Gruber, 2002; Dobkin et al., 2018; Jones et al., 2018; Blundell et al., 2024). Following all three outcomes in one event-study framework clarifies the welfare interpretation: higher spending alone may indicate access to necessary care, but higher spending combined with catastrophic burden and wealth loss indicates reduced financial resilience.

The persistence of the financial outcomes supports a household-level interpretation in which illness affects caregiving, consumption, transfers, labour supply, and asset decisions as well as medical bills (Kværner, 2023; Zilio et al., 2025). The longest-run wealth result is more sensitive to baseline-trajectory adjustment than the immediate and two-year effects (Section 5.4).

Diagnosis-specific results clarify the pooled average: cancer generates broad deterioration, stroke the most persistent wealth loss, and heart attack the least credible onset contrast because of pre-event divergence. The pooled estimand is retained, but the exclusion sensitivity shows that the headline findings do not depend on the problematic heart-attack subgroup (Section 5.5).

### *6.2 Religious participation and the informal-insurance hypothesis*

The absence of a consistent attendance-related buffer qualifies, rather than overturns, evidence that religious organisations can perform insurance-like functions. Dehejia et al. (2007) associate religious participation with smoother consumption or happiness after income shocks, Chen (2010) interprets increased religious involvement during financial crisis as a response to unmet insurance needs, and Ager and Ciccone (2018) link exposure to agricultural risk with stronger religious-community membership. Most directly, Auriol et al. (2020) show in Ghana that access to formal funeral insurance reduces religious giving, consistent with some perceived substitutability between formal insurance and church-based protection. Those studies establish that religious communities can matter economically, but they do not imply that ordinary service attendance will protect households against every type of shock.

The present setting differs in several theoretically important respects. Cancer, stroke, and heart attack can generate large, prolonged, and medically specialised needs that may exceed the financial capacity of informal networks. Older Americans also operate within an institutional environment shaped by Medicare, private insurance, family transfers, and formal healthcare providers. Religious assistance may therefore be complementary to formal systems - for example through transport, information, meals, or caregiving - without being large enough to alter household wealth or the probability that medical spending exceeds 10% of income. In this sense, the findings identify a boundary condition for the informal-insurance hypothesis: participation in a religious organisation is not, by itself, a sufficient indicator that the organisation supplies resources at the scale or in the form needed to offset a major health shock.

The pattern across outcomes also matters for theory. A healthcare-access channel could increase out-of-pocket spending if congregational contacts help respondents obtain treatment, whereas an informal-insurance channel should reduce catastrophic burden or asset depletion. The attendance estimates do not reveal a

coherent combination of higher treatment-related spending and lower financial hardship, nor a consistent decline in both spending and wealth loss. This weakens the interpretation that weekly attendance captures a broad average insurance mechanism in these data. It does not exclude assistance to particular members or congregations, especially because the analysis does not observe transfers, caregiving, transport, network position, membership duration, or congregational resources directly.

The diagnostics weaken control reuse as an explanation for the attendance null because inverse-reuse weighting barely changes the interactions. Differential observation remains more consequential: weekly attendees are modestly more likely to have follow-up outcomes, so later attendance estimates compare selected observed groups rather than unconditional trajectories (Section 5.4).

Pooling also does not appear to hide a large buffer confined to one diagnosis: no condition-specific attendance or belief pattern survives pre-event checks and multiplicity adjustment (Section 5.5). Smaller diagnosis-specific relationships, especially in the belief sample, remain difficult to exclude.

### *6.3 Private belief and psychological coping*

The belief results similarly place limits on a simple coping interpretation. Religious-coping theory does not predict uniform protection across all stressors; the consequences of belief depend on whether coping is positive or negative, whether the event is interpreted as controllable, and whether religious meaning is supported by social relationships (Pargament & Abu Raiya, 2007). Smith et al. (2003) likewise find that associations between religiousness and depression vary with the form of religiosity and the stressful context. More recent evidence suggests that religiosity or access to religious participation moderated mental-health deterioration during the COVID-19 crisis in some settings (Bahal et al., 2023; Iyer et al., 2025). Against that background, the absence of a stable belief-related reduction in CES-D after serious illness should be interpreted as context-specific evidence rather than a general rejection of religious coping.

Two features of the present analysis reinforce that caution. First, depressive symptoms were already changing before the health event was first reported, making it difficult to isolate a psychological response beginning at diagnosis. Second, longitudinal research shows that health and depression can themselves shape later religious participation, and that public attendance and private religious importance have distinct pathways (Maselko et al., 2012; Orr et al., 2019; Cowden et al., 2022). Measuring belief before the first report reduces, but cannot eliminate, these selection and timing problems. The results therefore show that the available pre-event belief index does not predict a robust average difference in post-event CES-D trajectories; they do not show that meaning, prayer, pastoral care, or other forms of coping are unimportant to particular individuals.

### *6.4 Interpreting the qualified boundary result*

Taken together, the findings refine the claim that religion can provide resilience. The relevant question is whether broad pre-shock attendance and belief measures identify support that is activated, sufficiently large, and suited to a particular health event. Condition-specific financial trajectories differ, but they do not translate into coordinated attendance- or belief-related protection.

The qualified boundary result is therefore specific: the available religion measures do not reliably identify an average financial or psychological buffer among respondents whose outcomes remain observed, while confidence intervals and power calculations leave room for smaller relationships.

### *6.5 Implications for policy and research*

The policy implication is not that religious organisations are irrelevant. They may remain useful partners in outreach, treatment navigation, caregiving, and emotional support. The evidence instead cautions against assuming that these roles substitute for formal health and financial protection. The large increases in catastrophic spending and the accompanying wealth losses underscore the continuing need for institutions that protect older households from both direct medical costs and the wider economic consequences of illness. Future research should link health events to direct measures of congregational transfers, in-kind assistance, caregiving, network centrality, denomination, and organisational capacity. Such data would allow the healthcare-access, informal-insurance, and coping channels to be tested directly rather than inferred from attendance and belief alone.

## 7. Limitations

First, religious participation is not randomly assigned. Measuring it before the recorded shock and including respondent-by-cohort fixed effects reduce some concerns, but symptoms that begin before diagnosis may change attendance, and other time-varying characteristics may influence both religious participation and outcomes. The moderation estimates remain associational.

Second, the timing design cannot fully establish a clean untreated baseline. The expanded analysis observes two leads, approximately six and four years before first report, and does not reject their joint equality to zero for the financial outcomes. Even so, two leads cannot rule out a level shift or nonlinear pre-trend beginning between the four-year lead and the omitted interview approximately two years before first report. Because the condition is first reported only at the next biennial interview, biological onset or early symptoms may also fall within that final interval. The financial estimates are therefore event-associated differences rather than unqualified causal effects; CES-D is treated descriptively because both observed leads indicate pre-event movement.

Third, mortality, survey nonresponse, and outcome missingness select the observed post-event sample. Out-of-pocket spending and wealth are observed 3.7 percentage points more often for weekly attendees at the approximately two-year horizon and 4.1 percentage points more often at the approximately four-year horizon; cohort-adjusted differences are 3.5 and 3.9 percentage points. Private belief is not significantly related to post-event availability. Because respondent-level death and exit statuses linked to religion are unavailable in the analysis extract, the attendance difference cannot be decomposed into mortality and non-death nonresponse. The later moderation estimates are therefore conditional on outcome observation, and the direction of any resulting selection bias is unknown.

Fourth, household income and wealth are attached to respondent records and may be duplicated across spouses. The household component of HHIDPN was used to re-estimate inference at the household level. Household-clustered standard errors are virtually identical to the respondent-clustered values for the main financial coefficients and the religion interactions, so within-household dependence does not explain the

reported significance or null moderation pattern. Nevertheless, the household outcomes should still be interpreted as shared household resources rather than independent spouse-specific quantities.

Fifth, attendance and belief are imperfect proxies for the mechanisms of interest. The HRS does not directly identify congregational transfers, caregiving, transportation, food support, pastoral counselling, or the duration and quality of membership. The analytical data used for the moderation models retain the constructed belief index but not all constituent items, so internal reliability could not be recomputed for the final estimation sample.

Sixth, condition-specific analyses improve on the pooled specification but do not eliminate heterogeneity concerns. Cancer, stroke, and strict heart attack are classified separately at the first composite event, and the models identify especially persistent stroke-related wealth losses. The heart-attack financial outcomes exhibit significant pre-event divergence, so those subgroup coefficients are descriptive. Excluding every event involving heart attack leaves the pooled financial estimates at least as large and produces clean two-lead financial pre-event tests, which shows that heart-attack contamination does not drive the headline results. The multiple-concurrent group remains small, however, and the belief sample contains only 944 cancer, 607 stroke, 356 heart-attack, and 58 multiple treated events in the transformed-spending model. Smaller diagnosis-specific moderation relationships therefore remain difficult to exclude.

Seventh, Table 8 provides illustrative original-scale equivalents, but these are not constant-dollar average treatment effects. The nonlinear back-transformation depends on the selected baseline value, and the source variables remain in the monetary units supplied by the harmonised HRS/RAND files. Cohort-by-calendar-year fixed effects absorb common survey-year price movements within the event-study comparison, but a separate inflation harmonisation and spending-component decomposition were not reconstructed. The analysis therefore cannot determine whether higher spending reflects hospital, prescription, dental, long-term-care, home-care, or other services.

Eighth, the principal event-study estimates are unweighted by HRS sampling and attrition weights. They should therefore be interpreted as associations for the observed eligible panel, not as nationally representative population effects. Inverse control-reuse weights address the stacked design but do not correct differential sampling, mortality, survey attrition, or Leave-Behind questionnaire selection. Finally, Table 1 reports reference-wave characteristics for the exact principal estimation sample. The affected and comparison groups differ in age, chronic conditions, self-rated health, insurance, and some household characteristics, although baseline religion measures are closely aligned. These level differences are absorbed by respondent-by-cohort fixed effects, and Table 5 allows trajectories to vary with the observed baseline characteristics, but the design still depends on the absence of unobserved differential changes within the final pre-event interval.

## 8. Conclusion

First newly reported cancer, stroke, and strictly defined heart-attack events are followed by substantial medical and financial deterioration relative to eligible untreated comparisons. Condition-specific analyses show broad financial deterioration after cancer and especially persistent wealth losses after stroke. Heart-attack subgroup trajectories are more difficult to interpret because the financial outcomes diverge before first

report, but excluding every first event involving heart attack leaves the pooled financial estimates at least as large and the financial pre-event diagnostics clean.

The evidence does not identify a corresponding average or condition-specific protective relationship for pre-shock religious participation. Fully saturated models, joint post-event tests, multiplicity adjustments, inverse control-reuse weighting, a never-treated-control restriction, household-clustered inference, and a cancer-versus-acute-event sensitivity do not reveal a statistically robust attendance- or belief-related buffering pattern. Later attendance estimates require additional caution because weekly attendees have modestly higher post-event outcome availability.

The paper's conclusion is therefore substantive but bounded by measurement and selection: the available attendance and belief indicators do not reliably predict smaller average post-shock financial losses among respondents whose outcomes remain observed. Future research should connect longitudinal health events to direct measures of congregational transfers, caregiving, transportation, network position, and institutional capacity, while distinguishing effects on treatment access from effects on household financial protection.

## References


Austin, P. C. (2009). Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. Statistics in Medicine, 28(25), 3083-3107. https://doi.org/10.1002/sim.3697

Ager, P., & Ciccone, A. (2018). Agricultural risk and the spread of religious communities. *Journal of the European Economic Association, 16*(4), 1021–1068. https://doi.org/10.1093/jeea/jvx029

Auriol, E., Lassébie, J., Panin, A., Raiber, E., & Seabright, P. (2020). God insures those who pay? Formal insurance and religious offerings in Ghana. *The Quarterly Journal of Economics, 135*(4), 1799–1848. https://doi.org/10.1093/qje/qjaa015

Bahal, G., Iyer, S., Shastry, K., & Shrivastava, A. (2023). Religion, Covid-19 and mental health. *European Economic Review, 160*, Article 104621. https://doi.org/10.1016/j.euroecorev.2023.104621

Bentzen, J. S. (2019). Acts of God? Religiosity and natural disasters across subnational world districts. *The Economic Journal, 129*(622), 2295–2321. https://doi.org/10.1093/ej/uez008

Blundell, R., Borella, M., Commault, J., & De Nardi, M. (2024). Old age risks, consumption, and insurance. *American Economic Review, 114*(2), 575–613. https://doi.org/10.1257/aer.20220555

Borusyak, K., Jaravel, X., & Spiess, J. (2024). Revisiting event-study designs: Robust and efficient estimation. *The Review of Economic Studies, 91*(6), 3253–3285. https://doi.org/10.1093/restud/rdae007

Bryan, G., Choi, J. J., & Karlan, D. (2021). Randomizing religion: The impact of Protestant evangelism on economic outcomes. *The Quarterly Journal of Economics, 136*(1), 293–380. https://doi.org/10.1093/qje/qjaa023

Callaway, B., & Sant'Anna, P. H. C. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics, 225*(2), 200–230. https://doi.org/10.1016/j.jeconom.2020.12.001

Chen, D. L. (2010). Club goods and group identity: Evidence from Islamic resurgence during the Indonesian financial crisis. *Journal of Political Economy, 118*(2), 300–354. https://doi.org/10.1086/652462

Cowden, R. G., Nakamura, J. S., Chen, Z. J., Case, B., Kim, E. S., & VanderWeele, T. J. (2022). Identifying pathways to religious service attendance among older adults: A lagged exposure-wide analysis. *PLOS ONE, 17*(11), e0278178. https://doi.org/10.1371/journal.pone.0278178

Dehejia, R., DeLeire, T., & Luttmer, E. F. P. (2007). Insuring consumption and happiness through religious organizations. *Journal of Public Economics, 91*(1–2), 259–279. https://doi.org/10.1016/j.jpubeco.2006.05.004

Dobkin, C., Finkelstein, A., Kluender, R., & Notowidigdo, M. J. (2018). The economic consequences of hospital admissions. *American Economic Review, 108*(2), 308–352. https://doi.org/10.1257/aer.20161038

Gertler, P., & Gruber, J. (2002). Insuring consumption against illness. *American Economic Review, 92*(1), 51–70. https://doi.org/10.1257/000282802760015603

Idler, E., Blevins, J., Kiser, M., & Hogue, C. (2017). Religion, a social determinant of mortality? A 10-year follow-up of the Health and Retirement Study. *PLOS ONE, 12*(12), e0189134. https://doi.org/10.1371/journal.pone.0189134

Iyer, S. (2016). The new economics of religion. *Journal of Economic Literature, 54*(2), 395–441. https://doi.org/10.1257/jel.54.2.395

Iyer, S., Larcom, S., Shah, J., & She, P.-W. (2025). Do religious people cope better in a crisis? Evidence from the UK pandemic lockdowns. *Journal of Economic Behavior & Organization, 237*, Article 107136. https://doi.org/10.1016/j.jebo.2025.107136

Jones, J. B., De Nardi, M., French, E., McGee, R., & Kirschner, J. (2018). *The lifetime medical spending of retirees* (NBER Working Paper No. 24599). National Bureau of Economic Research. https://doi.org/10.3386/w24599

Krause, N. (2002). Church-based social support and health in old age: Exploring variations by race. *The Journals of Gerontology: Series B, 57*(6), S332–S347. https://doi.org/10.1093/geronb/57.6.S332

Kværner, J. (2023). How large are bequest motives? Estimates based on health shocks. *The Review of Financial Studies, 36*(8), 3382–3422. https://doi.org/10.1093/rfs/hhac093

Maselko, J., Hayward, R. D., Hanlon, A., Buka, S., & Meador, K. (2012). Religious service attendance and major depression: A case of reverse causality? *American Journal of Epidemiology, 175*(6), 576–583. https://doi.org/10.1093/aje/kwr349

Orr, J., Tobin, K., Carey, D., Kenny, R. A., & McGarrigle, C. (2019). Religious attendance, religious importance, and the pathways to depressive symptoms in men and women aged 50 and over living in Ireland. *Research on Aging, 41*(9), 891–911. https://doi.org/10.1177/0164027519860270

Pargament, K. I., & Abu Raiya, H. (2007). A decade of research on the psychology of religion and coping: Things we assumed and lessons we learned. *Psyke & Logos, 28*(2), 742–766. https://doi.org/10.7146/pl.v28i2.8398

Smith, J., Ryan, L., Larkina, M., Sonnega, A., & Weir, D. (2023). *Psychosocial and lifestyle questionnaire 2006–2022*. University of Michigan. https://hrs.isr.umich.edu/publications/biblio/129032

Smith, T. B., McCullough, M. E., & Poll, J. (2003). Religiousness and depression: Evidence for a main effect and the moderating influence of stressful life events. *Psychological Bulletin, 129*(4), 614–636. https://doi.org/10.1037/0033-2909.129.4.614

Sonnega, A. (2025). *Using Health and Retirement Study data: A guide for new users*. Survey Research Center, Institute for Social Research, University of Michigan. https://hrs.isr.umich.edu/sites/default/files/new-user-guide/HRS-New-User-Guide-July-2025.pdf

Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics, 225*(2), 175–199. https://doi.org/10.1016/j.jeconom.2020.09.006

Zilio, F., Hickey, R., McDonald, J. T., Sun, E. C., & Zhang, Y. (2025). Health shocks and household allocation of time and spending. *Review of Economics of the Household*. Advance online publication. https://doi.org/10.1007/s11150-025-09804-2

## Tables and Figures

**Table 1. Reference-wave characteristics of affected and eligible untreated respondent-cohort units**

| Characteristic | Affected | Eligible controls | SMD | Inverse-reuse weighted SMD |
|---|---|---|---|---|
| Age (years) | 70.14 (10.53) | 67.04 (10.75) | 0.291 | 0.386 |
| Household income ($000s) | 35.0 [17.7, 67.9] | 40.8 [19.5, 81.9] | -0.050 | -0.025 |
| Household wealth ($000s) | 125.0 [15.8, 362.5] | 140.0 [22.1, 388.0] | -0.039 | 0.007 |
| Private insurance | 48.5% | 55.3% | -0.135 | -0.147 |
| Self-rated health (1=excellent, 5=poor) | 3.12 (1.09) | 2.84 (1.06) | 0.257 | 0.204 |
| ADL limitations | 0.56 (1.25) | 0.41 (1.10) | 0.135 | 0.093 |
| Chronic-condition count | 2.43 (1.48) | 2.01 (1.44) | 0.291 | 0.296 |
| Married/partnered | 61.5% | 62.6% | -0.022 | -0.020 |
| Household size | 2.19 (1.17) | 2.31 (1.29) | -0.092 | -0.129 |
| Baseline IHS OOP spending | 6.78 (3.24) | 6.71 (3.19) | 0.024 | 0.061 |
| Baseline catastrophic OOP | 24.2% | 20.4% | 0.091 | 0.085 |
| Baseline CES-D | 1.77 (2.15) | 1.46 (1.98) | 0.153 | 0.111 |
| Weekly attendance | 38.4% | 39.4% | -0.021 | 0.020 |
| Private belief (standardized) | -0.01 (1.01) | -0.00 (1.00) | -0.004 | 0.008 |

Notes. Values are means (standard deviations), percentages for binary variables, or medians [interquartile ranges] for household income and wealth. The reference interview is event -1, approximately two years before first report. SMD denotes the standardised mean difference; for income and wealth, SMDs use inverse-hyperbolic-sine transformed levels. The exact principal transformed-spending/wealth estimation sample contains 5,319 affected and 114,215 eligible comparison respondent-cohort units after singleton removal. Inverse-reuse weights equal one for affected units and the inverse of the number of control cohorts for comparison respondents. Missingness is low except for private belief, available for 1,965 affected and 41,457 comparison units.

**Table 2. Estimated affected-versus-comparison outcome differences around the first-report interview**

| | (1) Transformed OOP spending | (2) OOP above 10% of income | (3) Transformed household wealth | (4) CES-D symptoms |
|---|---|---|---|---|
| Approximately 4 years before first report | 0.078<br>(0.048) | 0.006<br>(0.008) | -0.047<br>(0.085) | -0.111***<br>(0.030) |
| First-report interview | 0.516***<br>(0.048) | 0.096***<br>(0.008) | -0.440***<br>(0.080) | 0.214***<br>(0.030) |
| Approximately 2 years after first report | 0.208***<br>(0.056) | 0.053***<br>(0.008) | -0.467***<br>(0.098) | 0.147***<br>(0.034) |
| Approximately 4 years after first report | 0.162**<br>(0.064) | 0.038***<br>(0.010) | -0.379***<br>(0.116) | 0.128***<br>(0.040) |
| **Observations** | **464,209** | **456,099** | **464,209** | **439,817** |
| **Respondents** | **27,994** | **27,774** | **27,994** | **26,683** |
| **Treated events** | **5,319** | **5,293** | **5,319** | **5,001** |
| Respondent-by-cohort fixed effects | Yes | Yes | Yes | Yes |
| Cohort-by-calendar-year fixed effects | Yes | Yes | Yes | Yes |

Notes. Coefficients are affected-versus-comparison differences relative to the omitted interview approximately two years before the condition was first reported. The other rows correspond to approximately four years before, the first-report interview, approximately two years after, and approximately four years after. Spending and wealth use the inverse hyperbolic sine and are not exact percentage changes. Standard errors clustered by respondent are in parentheses. CES-D is secondary because depressive symptoms were already changing before the first report.

**Table 3. Do affected-versus-comparison outcome differences vary by weekly religious attendance?**

| | (1)<br>Transformed OOP spending | (2)<br>OOP above 10% of income | (3)<br>Transformed household wealth | (4)<br>CES-D symptoms |
|---|---|---|---|---|
| Weekly-attendance difference: approximately 4 years before | 0.077<br>(0.099) | 0.005<br>(0.016) | -0.226<br>(0.169) | 0.035<br>(0.059) |
| Weekly-attendance difference: first-report interview | 0.075<br>(0.098) | 0.017<br>(0.016) | -0.002<br>(0.162) | -0.034<br>(0.061) |
| Weekly-attendance difference: approximately 2 years after | -0.104<br>(0.115) | -0.001<br>(0.017) | -0.252<br>(0.199) | -0.075<br>(0.068) |
| Weekly-attendance difference: approximately 4 years after | -0.046<br>(0.131) | 0.026<br>(0.020) | -0.242<br>(0.230) | 0.013<br>(0.081) |
| **Observations** | **462,084** | **454,101** | **462,084** | **437,921** |
| **Respondents** | **27,977** | **27,758** | **27,977** | **26,669** |
| **Treated events** | **5,300** | **5,275** | **5,300** | **4,985** |

Notes. Coefficients indicate whether the affected-versus-comparison difference varies between weekly and less-frequent attendees, relative to the omitted interview approximately two years before the first report. The fully saturated model includes timing-by-shock, timing-by-attendance, and timing-by-shock-by-attendance terms. Weekly attendance is measured before the event. Respondent-by-cohort and cohort-by-calendar-year fixed effects are included. Standard errors are clustered by respondent.

**Table 4. Do affected-versus-comparison outcome differences vary by private religious belief?**

| | (1)<br>Transformed OOP spending | (2)<br>OOP above 10% of income | (3)<br>Transformed household wealth | (4)<br>CES-D symptoms |
|---|---|---|---|---|
| Belief difference: approximately 4 years before | -0.044<br>(0.072) | -0.008<br>(0.012) | -0.123<br>(0.118) | 0.004<br>(0.049) |
| Belief difference: first-report interview | 0.007<br>(0.071) | -0.013<br>(0.012) | -0.029<br>(0.120) | -0.028<br>(0.046) |
| Belief difference: approximately 2 years after | -0.059<br>(0.082) | -0.001<br>(0.013) | -0.175<br>(0.131) | -0.092*<br>(0.050) |
| Belief difference: approximately 4 years after | -0.006<br>(0.093) | -0.000<br>(0.015) | -0.053<br>(0.157) | -0.004<br>(0.060) |
| **Observations** | **175,074** | **172,889** | **175,074** | **169,902** |
| **Respondents** | **20,436** | **20,332** | **20,436** | **19,905** |
| **Treated events** | **1,965** | **1,957** | **1,965** | **1,893** |

Notes. Coefficients indicate how the affected-versus-comparison difference varies with a one-standard-deviation increase in pre-shock belief, relative to the omitted interview approximately two years before the first report. The fully saturated model includes timing-by-shock, timing-by-belief, and timing-by-shock-by-belief terms. Belief is standardised across 44,229 unique respondent-anchor-year records. Respondent-by-cohort and cohort-by-calendar-year fixed effects are included. Standard errors are clustered by respondent.

**Table 5. Sensitivity to complete-case restriction and baseline covariate adjustment**

| Timing relative to first report | **Model A<br>Full sample** | **Model B<br>Complete case** | **Model C<br>Event-interacted covariates** |
|---|---|---|---|
| **Panel A. Transformed OOP spending** | | | |
| Approximately 4 years before first report | 0.078<br>(0.048) | 0.098**<br>(0.048) | 0.094*<br>(0.048) |
| First-report interview | 0.516***<br>(0.048) | 0.527***<br>(0.048) | 0.572***<br>(0.048) |
| Approximately 2 years after first report | 0.208***<br>(0.056) | 0.219***<br>(0.056) | 0.292***<br>(0.056) |
| Approximately 4 years after first report | 0.162**<br>(0.064) | 0.170***<br>(0.064) | 0.264***<br>(0.064) |
| **Observations** | **464,209** | **458,888** | **458,888** |
| **Respondents** | **27,994** | **27,925** | **27,925** |
| **Treated events** | **5,319** | **5,252** | **5,252** |
| **Panel B. OOP above 10% of income** | | | |
| Approximately 4 years before first report | 0.006<br>(0.008) | 0.007<br>(0.008) | 0.009<br>(0.008) |
| First-report interview | 0.096***<br>(0.008) | 0.097***<br>(0.008) | 0.100***<br>(0.008) |
| Approximately 2 years after first report | 0.053***<br>(0.008) | 0.053***<br>(0.009) | 0.056***<br>(0.009) |
| Approximately 4 years after first report | 0.038***<br>(0.010) | 0.039***<br>(0.010) | 0.044***<br>(0.010) |
| **Observations** | **456,099** | **450,980** | **450,980** |
| **Respondents** | **27,774** | **27,706** | **27,706** |
| **Treated events** | **5,293** | **5,226** | **5,226** |
| **Panel C. Transformed household wealth** | | | |
| Approximately 4 years before first report | -0.047<br>(0.085) | -0.029<br>(0.085) | -0.076<br>(0.085) |
| First-report interview | -0.440***<br>(0.080) | -0.436***<br>(0.080) | -0.371***<br>(0.081) |
| Approximately 2 years after first report | -0.467***<br>(0.098) | -0.458***<br>(0.098) | -0.340***<br>(0.099) |
| Approximately 4 years after first report | -0.379***<br>(0.116) | -0.360***<br>(0.116) | -0.204*<br>(0.116) |
| **Observations** | **464,209** | **458,888** | **458,888** |
| **Respondents** | **27,994** | **27,925** | **27,925** |
| **Treated events** | **5,319** | **5,252** | **5,252** |

Notes. Model A is the full-sample event-study model. Model B applies the complete baseline-covariate restriction but does not add covariate interactions. Model C uses the same complete-case sample and interacts event time with pre-shock age, income, wealth, private insurance, self-rated health, ADL limitations, chronic conditions, partnership status, and household size. Coefficients are affected-versus-comparison differences relative to the omitted pre-event interview. All models include respondent-by-cohort and cohort-by-calendar-year fixed effects. Standard errors are clustered by respondent. * p < .10. ** p < .05. *** p < .01.

**Table 6. Pointwise 95% confidence intervals for post-shock religion interactions**

| Religion measure and outcome | First-report interview | Approximately 2 years later | Approximately 4 years later |
|---|---|---|---|
| **Weekly attendance** | | | |
| Transformed OOP spending | 0.075 [-0.117, 0.268] | -0.104 [-0.330, 0.121] | -0.046 [-0.304, 0.211] |
| OOP above 10% of income | 0.017 [-0.014, 0.047] | -0.001 [-0.035, 0.033] | 0.026 [-0.013, 0.064] |
| Transformed household wealth | -0.002 [-0.321, 0.316] | -0.252 [-0.642, 0.138] | -0.242 [-0.694, 0.209] |
| CES-D symptoms | -0.034 [-0.153, 0.086] | -0.075 [-0.209, 0.059] | 0.013 [-0.146, 0.172] |
| **Standardised private belief** | | | |
| Transformed OOP spending | 0.007 [-0.133, 0.146] | -0.059 [-0.219, 0.101] | -0.006 [-0.189, 0.176] |
| OOP above 10% of income | -0.013 [-0.037, 0.011] | -0.001 [-0.027, 0.025] | -0.000 [-0.030, 0.029] |
| Transformed household wealth | -0.029 [-0.265, 0.207] | -0.175 [-0.430, 0.081] | -0.053 [-0.361, 0.255] |
| CES-D symptoms | -0.028 [-0.118, 0.061] | -0.092 [-0.190, 0.007] | -0.004 [-0.122, 0.115] |

Notes. Cells report the interaction estimate and pointwise 95% confidence interval. Attendance compares weekly-or-more attendance with less frequent attendance. Belief coefficients refer to a one-standard-deviation increase. Catastrophic-spending coefficients are proportions. Pointwise intervals do not establish practical equivalence.

**Table 7. Joint post-event tests and minimum detectable effects for financial outcomes**

| Moderator | Outcome | Average post-event effect | SE | 95% CI | Joint p | Holm p | 80% MDE | 90% MDE |
|---|---|---|---|---|---|---|---|---|
| Weekly attendance | Transformed OOP | -0.025 | 0.091 | [-0.205, 0.154] | 0.428 | 1.000 | 0.256 | 0.297 |
| Standardised belief | Transformed OOP | -0.019 | 0.064 | [-0.144, 0.105] | 0.853 | 1.000 | 0.179 | 0.207 |
| Weekly attendance | Catastrophic OOP | 0.014 | 0.014 | [-0.013, 0.041] | 0.386 | 1.000 | 0.039 | 0.045 |
| Standardised belief | Catastrophic OOP | -0.005 | 0.011 | [-0.026, 0.016] | 0.689 | 1.000 | 0.030 | 0.034 |
| Weekly attendance | Transformed wealth | -0.166 | 0.157 | [-0.474, 0.142] | 0.492 | 1.000 | 0.440 | 0.509 |
| Standardised belief | Transformed wealth | -0.085 | 0.112 | [-0.305, 0.134] | 0.559 | 1.000 | 0.313 | 0.363 |

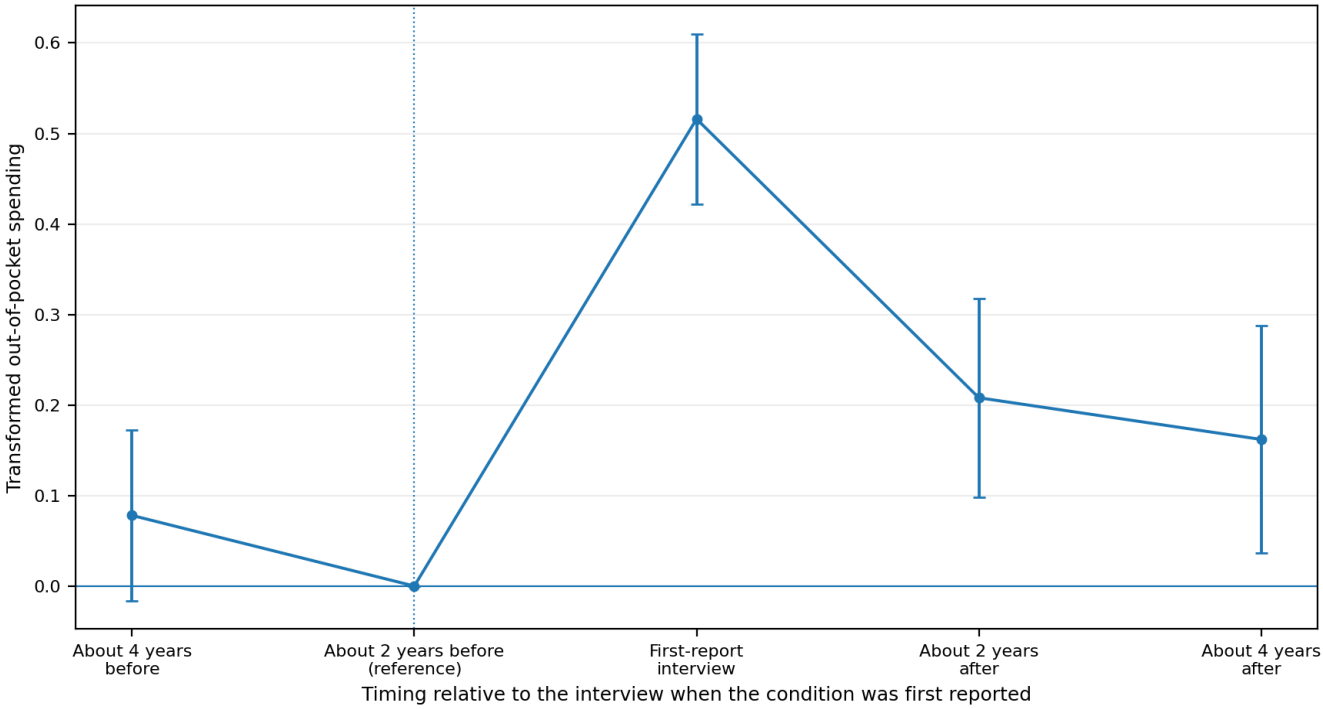
Notes. The average post-event effect is the equally weighted mean of the event 0, +1, and +2 interaction coefficients. "Joint p" tests whether all three coefficients equal zero. Holm p values adjust all six pooled financial joint tests across the two moderator families. MDEs are two-sided minimum detectable effects at alpha = .05. Catastrophic-spending effects and MDEs are proportions; 0.039 corresponds to 3.9 percentage points.

**Table 8. Illustrative original-scale magnitudes of the pooled financial coefficients**

| Outcome | Timing | IHS coefficient | Reference value | Illustrative difference | 95% coefficient-based range | Interpretation |
|---|---|---|---|---|---|---|
| Out-of-pocket spending | First report | 0.516 | $1,300 median OOP | $879 higher | $683 to $1,094 higher | Illustrative back-transformation |
| Out-of-pocket spending | ~2 years later | 0.208 | $1,300 median OOP | $301 higher | $134 to $488 higher | Illustrative back-transformation |
| Out-of-pocket spending | ~4 years later | 0.162 | $1,300 median OOP | $228 higher | $47 to $433 higher | Illustrative back-transformation |
| Household wealth | First report | -0.440 | $125,000 median wealth | $44,529 lower | $30,917 to $56,172 lower | Illustrative back-transformation |
| Household wealth | ~2 years later | -0.467 | $125,000 median wealth | $46,673 lower | $30,115 to $60,342 lower | Illustrative back-transformation |
| Household wealth | ~4 years later | -0.379 | $125,000 median wealth | $39,411 lower | $17,558 to $56,820 lower | Illustrative back-transformation |

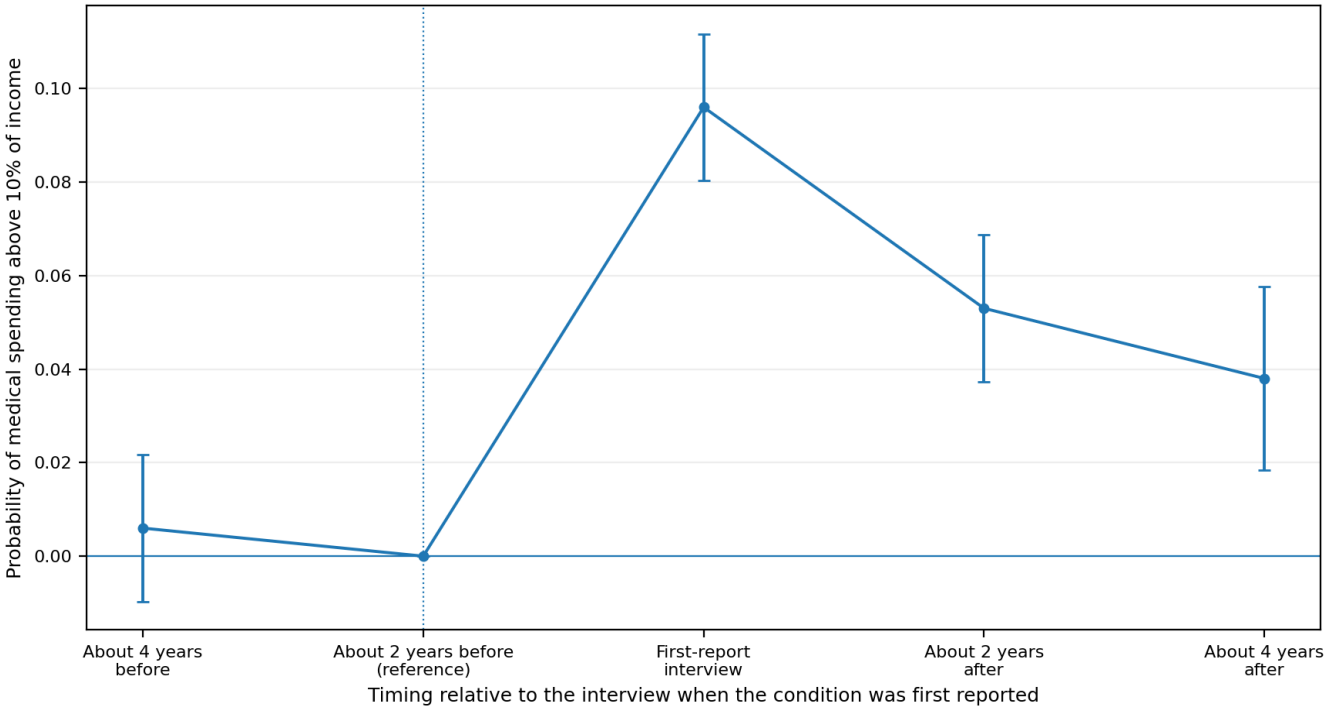
Notes. Original-scale illustrations use sinh[asinh(y*) + beta] - y*, where y* is the affected-group event -1 median. The range back-transforms the pointwise 95% confidence interval for beta. These values are not average marginal effects, exact percentage changes, inflation-adjusted causal effects, or population-representative dollar estimates. Because the transformation is nonlinear, results differ at other baseline values; Appendix Table B15 reports sensitivity across the affected-group reference quartiles.

**Figure 1. Estimated affected-versus-comparison difference in transformed out-of-pocket spending**



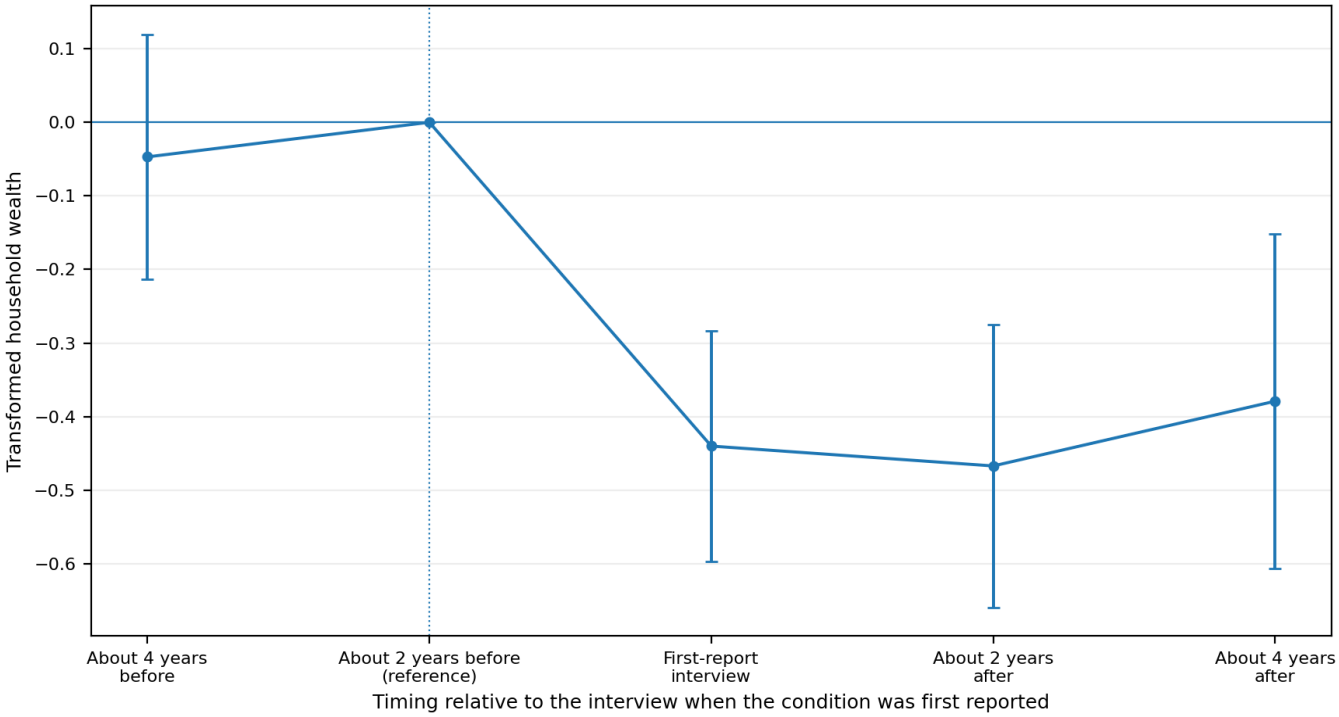

Notes. Points report estimated affected-versus-comparison differences and bars report pointwise 95% confidence intervals. The reference interview, approximately two years before the first report, is plotted at zero. The horizontal axis translates the biennial event-time scale into approximate elapsed years. Coefficients are transformed spending units.

**Figure 2. Estimated affected-versus-comparison difference in catastrophic-spending risk**




Notes. Points report estimated affected-versus-comparison differences and bars report pointwise 95% confidence intervals. The reference interview, approximately two years before the first report, is plotted at zero. The horizontal axis translates the biennial event-time scale into approximate elapsed years. Coefficients are proportions; multiply by 100 for percentage points.

**Figure 3. Estimated affected-versus-comparison difference in transformed household wealth**




Notes. Points report estimated affected-versus-comparison differences and bars report pointwise 95% confidence intervals. The reference interview, approximately two years before the first report, is plotted at zero. The horizontal axis translates the biennial event-time scale into approximate elapsed years. Negative coefficients indicate lower transformed household wealth.