

Measuring Cross-Category Consumption Reallocation Around Eight Household Lifecycle Events: A Causal Substitution Matrix from the PSID

October 16, 2026

Abstract

Major household transitions—marriage, childbirth, divorce, retirement, and health shocks—trigger broad reallocations of household budgets, yet marketing research typically studies one life event or product category in isolation. Using twenty years of Panel Study of Income Dynamics data (1999–2019), we estimate a unified causal substitution matrix capturing how eight lifecycle transitions shift expenditure shares across nine consumption categories, and translate these shifts into lifetime-budget counterfactuals for targeting. A first birth, for example, shifts between \$2,000 and \$39,500 in lifetime household spending toward childcare while crowding out \$3,000 to \$61,500 in recreation. Our cross-event design also uncovers sharp demographic heterogeneity masked by pooled averages, including a 2.3 percentage-point reversal in out-of-pocket healthcare spending at the age-65 Medicare eligibility threshold and an income-tier reversal in childcare expenditure.

1 Introduction

How does a household’s allocation of expenditure across food, housing, transport, health-care, childcare, and recreation change when a discrete lifecycle event occurs, a first birth, a retirement, a divorce, a residential move? Category managers, demand-system modelers, and customer-relationship-management (CRM) teams use the cross-sectional age profiles of Aguiar and Hurst [2013] or aggregate cohort-based causal event-study methods [Oblander and McCarthy, 2023] as substitutes for what they would actually like: a causally identified, signed matrix of category share-point shifts indexed by event. The institutional setting in this paper is the United States Panel Study of Income Dynamics (PSID) 1999–2019, a biennial household panel with nine expenditure categories. The operative friction is non-separability of preferences from household composition and home-production technology: a new child, a departing partner, or a new house mechanically alters the marginal utility of food, transport, and housing services in ways that the cross-sectional age profile cannot identify.

This paper’s contribution is a marketing-decision-relevant counterfactual: for a representative household experiencing a first birth, the framework delivers a lifetime reallocation toward childcare in the range \$2,000–\$39,500 and away from recreation in the range \$3,000–\$61,500, bracketed between a transitory case (4-year linear decay) and a durable case (40-year flat extrapolation), with bootstrap confidence intervals propagated from the underlying causal-identification step. The primitive that makes this counterfactual possible is a 9×8 causal substitution matrix $B(0)$ for eight household lifecycle events on PSID 1999–2019: 72 signed, causally identified cross-category share-point shifts, with eight pre-registered headline cells¹ and the remainder available under explicit multiplicity controls. The matrix is a measurement primitive — signed, causally identified, cited and reused — and Section 10 propagates it through a representative lifetime budget to deliver the dollar counterfactual; the simplex-respecting structural architecture that would absorb the matrix’s nine columns jointly is discussed in Section 9 as work this paper does not estimate.

¹See Section 4.6 on the analysis-plan distinction.

What the matrix overturns. The framework surfaces two surprises that the standard cross-sectional or pooled-event approaches conceal. First, the first-birth $\times$ childcare income-tercile decomposition reverses the analysis-plan prior: top-income-tercile households shift their consumption toward paid childcare by a *smaller* share-point amount than bottom-income-tercile households, contrary to the predicted pattern based on cross-sectional childcare-share gradients. Second, the pooled retirement $\times$ healthcare cell is Holm-null (+0.14 pp), but the framework’s age-at-retirement decomposition reveals a +2.32 pp pre-65/ $\geq$ 65 contrast (pre-65: +1.42; $\geq$ 65: -0.90) that the pooled estimate averages away. Both findings are visible only through the matrix-with-decomposition lens.

As a credibility check for non-DiD-specialist readers, the matrix recovers the cross-sectional retirement-food finding of Aguiar and Hurst [2005] as one nested cell (retirement $\times$ food = +0.55 pp), without bespoke specification. Beyond the eight pre-registered headline cells, a further 22 cells are significant under the Benjamini-Hochberg procedure at $q = 0.10$ (18 cells survive the tighter $q = 0.05$ threshold; Section 12); the matrix has substantially more content than the headline set. A held-out univariate share-prediction exercise documents a negative aggregate result, characterizing where the matrix is not a drop-in feature and clarifying the unit-simplex architecture for which it is best suited.

The age-at-retirement contrast at the Medicare cutoff motivates a Health and Retirement Study (HRS) follow-up with annual resolution and direct insurance-status variables. The pre-65 rise is consistent with employer-replacement-premium dynamics (Consolidated Omnibus Budget Reconciliation Act (COBRA) premia or, for post-2013 retirement cohorts, the Affordable Care Act exchange). The $\geq$ 65 decline involves channels Medicare-substitution alone cannot deliver; Section 7 characterizes the contrast and motivates HRS replication.

The headline matrix $B(0) \in \mathbb{R}^{9 \times 8}$ reports 72 cells; the eight pre-registered cells anchor the discussion. The residential-move row admits a descriptive four-cell tenure decomposition (owner-to-owner, owner-to-renter, renter-to-owner, renter-to-renter) and an intra-region versus cross-region split; the renter-to-renter cell rejects a strict mechanical-rent-gradient near-

zero benchmark, and the cross-region +3.07 pp exceeds intra-region +1.96 pp by roughly one cross-region SE, consistent with a rent-gradient channel without separating it from alternative moving-cost stories. We document per-cell persistence trajectories for the eight headline cells across $k \in \{0, 1, 2, 3, 4\}$ biennial waves; one cell (move $\times$ housing) shows clean exponential decay with $\hat{\tau}$ identifiable, others show boundary-fit, non-monotone, or subgroup-dependent dynamics. The classification (transitory/durable/subgroup-dependent/unidentified) appears in Internet Appendix Section ??.

The closest marketing-quant precedent is Biswas et al. [2025], who estimate consumption shifts across consumer packaged goods (CPG) categories around continuous income and wealth shocks using within-household panel fixed effects. This paper differs on three dimensions: discrete lifecycle events rather than continuous income or wealth, broad PSID nine-category shares rather than CPG sub-allocation, and modern heterogeneity-robust difference-in-differences (DiD) rather than panel fixed effects. The category-level and brand-level measurement objects are complementary inputs to demand-system architectures, not less granular versions of each other. Olafsson and Pagel [2024] focus on retirement using Icelandic high-frequency personal finance management data; this paper estimates an eight-event matrix on a complementary lower-frequency household-panel data source. Oblander and McCarthy [2023] apply a cohort-based causal event-study method to a macro event (COVID-19); this paper delivers household-level discrete-event causal identification. The lifecycle-consumption foundation is Aguiar and Hurst [2005, 2013]; the methods canon comprises Callaway and Sant’Anna [2021], Borusyak et al. [2024], de Chaisemartin and D’Haultfœuille [2020], Sant’Anna and Zhao [2020], Rambachan and Roth [2023], and Goodman-Bacon [2021].

2 Related literature

The paper connects six literatures.

The migration-preference identification program. Bronnenberg et al. [2012] use interstate migrants in scanner data to identify the long-run causal share of acquired brand preferences relative to local supply, isolating a discrete preference-disturbing shock and following households across it. Subsequent work transports the design to brand healthfulness, healthcare utilization, and nutrition. The category-level analogue (a within-household discrete-event identification on the broad expenditure partition) is the primitive the marketing literature has lacked alongside the brand-level program. This paper supplies it for eight household lifecycle events.

Aguiar-Hurst lifecycle consumption. Aguiar and Hurst [2005] resolve the retirement food-consumption puzzle through a home-production channel: retirees do not consume less food, they purchase less prepared food and substitute home production. Aguiar and Hurst [2013] estimate cross-sectional lifecycle expenditure profiles across categories using Consumer Expenditure Survey (CEX) data, exposing the heterogeneity in age-related shifts hidden by aggregates. Bernheim et al. [2001] and Banks et al. [1998] document the broader retirement-consumption-drop puzzle. This paper upgrades the cross-sectional age-profile object to a causal within-household event-level matrix on the multi-event partition.

Household-finance retirement consumption. Olafsson and Pagel [2024] estimate retirement-consumption dynamics on Icelandic personal finance management data; Hurd and Rohwedder [2008] use the Health and Retirement Study (HRS) to track expected versus realized retirement-consumption changes. These studies focus on a single event (retirement), an aggregate-spending object, and legacy panel-fixed-effects methods. This paper delivers multi-event coverage, modern heterogeneity-robust DiD, the substitution-matrix framework, and the Medicare-subgroup decomposition that the single-event aggregate-spending designs cannot surface.

Marketing-journal precedent on persistence and aggregate dynamics. Oblander and McCarthy [2023] apply a cohort-based causal event-study method to COVID-19 spending dynamics in marketing-relevant categories, establishing editorial appetite for persistence-around-events as a marketing-quant question. The macro-event aggregate-cohort object is complementary to this paper’s micro-event household-causal object.

Head-to-head with Biswas et al. [2025]. The closest marketing-quant competitor estimates consumption shifts across CPG categories around continuous income and wealth shocks using within-household panel fixed effects on scanner data. The differences are three. First, the treatment is different: discrete lifecycle events here, continuous income or wealth there. Second, the consumption partition is different: broad PSID nine-category shares here, CPG sub-allocation there. Third, the estimator class is different: Borusyak-Jaravel-Spiess (BJS) imputation with isolation-sample identification and Callaway-Sant’Anna triangulation here, panel fixed effects there. The two primitives serve different demand-system inputs: this paper informs across-category budget allocation, Biswas et al. [2025] informs within-category brand allocation. They are complementary, not competing.

Modern staggered DiD methods. The identification rests on the methods canon: Callaway and Sant’Anna [2021] for cohort-specific $ATT(g,t)$; Borusyak et al. [2024] for the imputation estimator and isolation-sample reasoning; de Chaisemartin and D’Haultfœuille [2020] for multi-treatment and forbidden-comparison diagnostics; Sun and Abraham [2021] for cohort-saturated two-way fixed effects (TWFE); Rambachan and Roth [2023] for parallel-trends sensitivity bounds; Sant’Anna and Zhao [2020] for doubly-robust ATT ; Goodman-Bacon [2021] for the staggered-TWFE decomposition; and the methods review by Roth et al. [2023].

Adjacent literatures. Ganong and Noel [2019] document present-biased consumption drops around unemployment using bank data; Stephens Jr. [2001] estimates long-run con-

sumption effects of displacement. Du and Kamakura [2008] and Kamakura and Du [2012] establish the budget-allocation marketing-primitive tradition that the matrix feeds into. Dobkin et al. [2018] use an event-study design around hospital admissions to document elevated out-of-pocket medical spending and broader economic consequences for non-elderly adults. Guvenen et al. [2021] characterize the income-dynamics channel relevant to composition-event interpretations. Cinelli and Hazlett [2020] develop the partial- R^2 robustness-value framework (sensemakr) used in Section 7.

3 Data

3.1 Sources and merge

The estimation sample is built from two Panel Study of Income Dynamics (PSID) data customs. Custom J296511 contains the harmonized expenditure data from 1999 through 2019 across eleven biennial waves, covering 17,489 households and 93,400 household-wave observations. Custom J361785 contains demographic and labor variables from 1999 through 2023 for 18,890 households; 1,401 J361785 households are not in J296511 and are dropped because they have no expenditure outcomes. The merge logic is documented in the online replication archive. The 2021 and 2023 J361785 waves are not used in estimation (no expenditure outcomes available); their role is forward-looking validation only.

3.2 Nine expenditure categories with coverage caveats

The dependent variable in each cell of the matrix is the within-household share of total expenditure allocated to one of nine PSID categories: food, housing, utilities, transport, education, childcare, healthcare, clothing, and recreation. Food, housing, utilities, transport, education, childcare, and healthcare are available 1999–2019. Clothing and recreation begin in 2005. The total-expenditure denominator is the sum of the nine category dollars (see the online replication archive for variable definitions). The PSID nine-category partition is the

partition this paper estimates; alcohol, tobacco, gifts, charity, household services, vacations, and durables are not separately resolved, and recreation pools entertainment, equipment, trips, and fees. The Nielsen-Kilts brand-level program is the complementary primitive for finer-grained category and brand resolution.

3.3 Eight lifecycle events

The events are first birth, additional birth, marriage, divorce, residential move, job loss, retirement, and child leaves home. Births are coded from the family-composition history. Marriage and divorce are coded from head-and-wife relations and marital-status transitions. Job loss and retirement are coded from `employ_status1`, available 2003 and later. Residential moves are coded from `moved_since`, available 2003 and later. Move, job-loss, and retirement events are restricted to centered treatment cohorts $g \geq 2005$ per the wave-coverage caveat. A head-stability flag (age-of-head wave-to-wave change in $[1, 4]$ years) is applied to divorce and marriage events because the PSID re-labels the household head on certain marital transitions; the head-stable subset is the identifiable population for these two events. The Cross-Year Individual File would relax this restriction.

3.4 Sample construction filters

The headline estimation sample is the ± 1 -wave isolation sample: a treated household-event observation enters the BJS imputation for event E only if no other lifecycle event fires in the ± 1 -wave window. The estimand is the isolation-sample ATT (the average treatment effect on the treated, restricted to households with no co-occurring events), not the population ATT. The ± 2 -wave isolation window is reported as a sensitivity. The head-stable subset drops approximately 45% of divorces; the published $\hat{\beta}_{\text{divorce, housing}}(0) = +4.58$ pp is the head-stable-divorce ATT, not the all-divorce ATT. The move event admits a four-cell tenure decomposition ($O \rightarrow O$, $O \rightarrow R$, $R \rightarrow O$, $R \rightarrow R$) and a two-region split (intra-region, cross-region) used in Section 6. The estimation sample ends in 2019.

3.5 Descriptive statistics

Table 1 reports means, medians, and standard deviations of the nine category shares and the key controls (age of head, family size, region, household labor income) for the full estimation sample and for the eight isolation-sample treated subpopulations. Childcare and education shares are concentrated in households with children present; recreation and clothing have lower coverage pre-2005. Family size varies systematically across events, with births raising and child-leaves lowering family size by construction.

Table 1: Descriptive statistics on the merged PSID 1999–2019 estimation sample (household-wave observations).

Variable	Mean	Median	SD	Min	Max
Share: food	0.187	0.172	0.094	0	0.95
Share: housing	0.232	0.213	0.135	0	0.95
Share: utilities	0.082	0.071	0.057	0	0.80
Share: transport	0.148	0.132	0.097	0	0.85
Share: healthcare	0.080	0.052	0.084	0	0.90
Share: childcare	0.013	0.000	0.038	0	0.55
Share: education	0.018	0.000	0.044	0	0.60
Share: clothing	0.041	0.034	0.030	0	0.50
Share: recreation	0.044	0.034	0.041	0	0.60
Age of head	48.5	47	16.8	18	97
Family size	2.62	2	1.45	1	14
Labor income (head, \$thousands)	46.3	35.0	58.7	0	999
Households	17,489				
Household-wave observations	93,400				
Waves	11 biennial (1999, 2001, . . . , 2019)				

Notes: Shares computed on the within-household total-expenditure denominator per Section 3. Family size is PSID `num_in_fu`. Labor income is `labor_income_head`, current dollars. Source: PSID 1999–2019.

4 Identification

4.1 The substitution matrix and the estimand

The central object is the 9×8 substitution matrix $B(k)$ whose cell (c, E) is the post-event share-point change in category c at event-time k for event E :

$$B_{E,c}(k) = \text{ATT}_{E,c}(g, g+k) = \mathbb{E}[s_{i,c,g+k}(1) - s_{i,c,g+k}(0) \mid G_i = g, i \in \mathcal{I}_E], \quad (1)$$

where $s_{i,c,t}$ is household i 's share of total expenditure on category c at wave t , G_i is the treatment cohort (first wave of event E for household i), and $\mathcal{I}_E$ is the ± 1 -wave isolation population for event E (households with no other lifecycle event firing in $[g-1, g+1]$). The estimand is the isolation-sample dynamic ATT, not the population ATT. The reported headline matrix $B(0)$ is the first-post-event-wave object; the persistence-classification framework uses $\hat{B}(k)$ for $k \in \{0, 1, 2, 3, 4\}$.

4.2 Primary estimator: BJS imputation on isolation samples

The primary estimator is the Borusyak et al. [2024] imputation. The estimator fits two-way (household and wave) fixed effects on the never-treated and not-yet-treated cells, imputes the counterfactual $\hat{s}_{i,c,g+k}(0)$ from the fitted fixed effects, and averages the residual $s_{i,c,g+k} - \hat{s}_{i,c,g+k}(0)$ over event-time bins. The implementation is a custom Python port (approximately 150 lines of code) calling an alternating-projection two-way demeaner; the routine is available in the online replication archive. The identifying assumptions are (i) parallel trends in counterfactual shares between treated and never-or-not-yet-treated households, (ii) no anticipation beyond a Freyaldenhoven-Hansen-Shapiro [Freyaldenhoven et al., 2019] auxiliary-lead window, and (iii) the Stable Unit Treatment Value Assumption (SUTVA). Cluster-robust standard errors are computed at the household level; a household-cluster sub-sampling bootstrap with $B = 1,000$ replications propagates first-step $\hat{\alpha}$ and $\hat{\lambda}$ uncer-

tainty.

4.3 Co-primary triangulation: Callaway-Sant’Anna doubly-robust

The co-primary triangulation is the doubly-robust Callaway-Sant’Anna estimator [Callaway and Sant’Anna, 2021, Sant’Anna and Zhao, 2020]. The implementation uses the augmented inverse-propensity-weighted (AIPW) doubly-robust form of Sant’Anna and Zhao [2020] with a logistic propensity score on age of head, family size, and lagged shares at lags $L = 1$ and $L = 2$, and an ordinary least squares (OLS) outcome regression on the same covariates. The doubly-robust estimator is reported on the full sample for each of the eight headline cells; we refer to this estimator as CS-DR throughout. The plain 2×2 cohort-aggregated CS estimator (no doubly-robust adjustment) is retained as a triangulation check; cells where CS and CS-DR disagree with the BJS headline by more than 1.5 pp or in sign are flagged in Table 3.

Triangulation does not cover retirement and job-loss columns. The dCDH multi-treatment estimator support fails on the retirement and job-loss columns (absorbing treatments, insufficient never-treated mass at the relevant cohorts). Identification on these two columns therefore rests on BJS imputation with CS-DR sensitivity, *not* on independent multi-estimator triangulation. We surface this here rather than in the limitations section because it directly affects the credibility of the Medicare-cutoff results (which live on the retirement column) and the budget-channel result (which lives on the job-loss column). The Sun-Abraham cohort-saturated TWFE estimator agrees directionally with the BJS headline on both columns (Internet Appendix Section ??); we treat the Sun-Abraham diagnostic as a separate cohort-heterogeneity cross-check rather than a third independent triangulation estimator.

4.4 Per-event identifying assumptions

First birth, additional birth. Anticipation is the load-bearing identification threat: pregnancy is typically known approximately one year before delivery, and a fraction of the share shift may precede the wave in which the birth is recorded. A Freyaldenhoven-Hansen-Shapiro auxiliary-lead-window control is implemented as a labor-supply-stability subset on the first-birth event; the FHS-adjusted estimate is reported as a sensitivity in Section 12 on a small $n \approx 25$ subsample.

Marriage, divorce. The head-stable subset is the identifiable population (no PSID Cross-Year Individual File access). The head-stable filter requires age-of-head wave-to-wave change in $[1, 4]$ years and drops approximately 45% of divorces. The estimate $\hat{\beta}_{\text{divorce, housing}}(0) = +4.58$ pp is the head-stable-divorce ATT, not the all-divorce ATT; the all-divorce-no-head-stability number is reported in the Internet Appendix.

Residential move. The move row admits a four-cell tenure decomposition by construction (owner-to-owner, owner-to-renter, renter-to-owner, renter-to-renter). The R→R cell is the cleanest behavioral cell (no tenure-status change); its estimate of +2.00 pp is not near zero, which the Section 6 discussion interprets as evidence for a refined mechanical-rent-gradient story over a generic moving-disrupts-everything story. The intra-region versus cross-region split delivers +1.96 pp and +3.07 pp respectively, consistent with rent-gradient effects.

Job loss. Isolation against retirement and divorce is the binding restriction: job-loss-then-retirement and job-loss-then-divorce are common co-occurrences. The de Chaisemartin-D'Haultfœuille (dCDH) support diagnostic fails for retirement and job-loss columns (insufficient never-treated mass at relevant cohorts); the credibility stack for these rows is BJS isolation plus CS-DR co-primary triangulation.

Retirement. The Medicare-eligibility subgroup contrast (pre-65 vs. ≥ 65) is the primary leg of the Section 7 puzzle, supplemented by alternate cutoffs at 62 and 67 and a smooth-in-age Medicare-eligibility regression discontinuity at age 65. The subgroup contrast is descriptive under within-subgroup parallel trends, not a local average treatment effect (LATE) or instrumented quantity.

Child leaves home. Identified from household-composition transitions in years where the head-stability flag holds. Co-occurrence with retirement is the leading risk; isolation handles it.

4.5 Diagnostics

Pre-trend joint χ^2 tests on $k \in \{-1, -2, -3\}$ are reported per headline cell; the first-birth $\times$ childcare cell has $\chi^2 = 4.94$, $p = 0.176$. The Rambachan and Roth [2023] HonestDiD sensitivity is reported as the $\bar{M}_{\text{approx}}$ first-order indicator and a companion confidence-interval inflation block at $M \in \{0.5, 1.0, 2.0\}$; the canonical Rambachan-Roth fixed-length confidence interval (FLCI) bound requires the R `HonestDiD` package which was unavailable, and Section 4.6 discloses the gap explicitly. A Goodman-Bacon-style cohort-size-weighted decomposition is reported as a directional diagnostic (cohort-size products, not canonical Goodman-Bacon 2×2 weights). A placebo on the never-treated-by-event pool with $B = 20$ replications confirms that the headline cells are outside the placebo distribution. Multiplicity correction uses the Romano-Wolf joint-bootstrap stepdown ($B = 1,000$) as the canonical procedure and the headline classification; the conservative Holm approximation is reported alongside. The marriage $\times$ housing classification is RW-significant ($p_{\text{RW}} = 0.014$) and Holm-marginal ($p_{\text{Holm}} = 0.062$); the headline retains marriage $\times$ housing in the eight-cell set on the RW classification.

4.6 Inference-machinery honest disclosure

This subsection documents the residual departures from canonical implementations. First, the $B = 1,000$ bootstrap is implemented as sub-sampling at rate 0.632 rather than canonical with-replacement: the within-replicate mask assigns each household weight zero or one, approximating sub-sampling. Standard errors and inference are derived from this sub-sample-bootstrap empirical distribution; the relative downward bias against canonical with-replacement is approximately at the second-order level for the cells reported. Second, the HonestDiD $\bar{M}_{\text{approx}}$ value is a first-order scaled-pretrend-magnitude indicator, not the canonical Rambachan-Roth FLCI bound; the companion CI-inflation block at $M \in \{0.5, 1, 2\}$ is the more defensible diagnostic and is reported in the Appendix. Third, the Romano-Wolf joint-bootstrap stepdown is the headline multiplicity correction; the Holm approximation is reported alongside as a conservative comparison. The marriage $\times$ housing cell is RW-significant and Holm-marginal; the RW classification is the headline.

In addition, the eight headline cells were specified a priori in our analysis plan; we did not lodge the plan with a third-party pre-registration service before estimation, and we use the phrase “specified a priori in our analysis plan” rather than “pre-registered” throughout the paper to make this distinction explicit. The remaining 64 non-headline cells are exploratory and corrected under the Benjamini-Hochberg false-discovery-rate procedure at $q = 0.05$ in the Internet Appendix. Fourth, the doubly-robust Callaway-Sant’Anna estimator is the co-primary; the plain 2×2 cohort-aggregation remains in the directional-triangulation column. Fifth, the Medicare-eligibility regression discontinuity at age 65 is implemented as basic local-linear OLS rather than Calonico et al. [2014] robust RD; the `rdrobust` package was unavailable. The Medicare-eligibility-RD coefficient at age 65 is power-null regardless of implementation, so the qualitative conclusion (smooth-in-age, no discontinuous Medicare jump) is unaffected.

4.7 Top-2 alternative designs not selected

dCDH multi-treatment estimator on the full sample. The canonical answer to co-occurring events is the de Chaisemartin and D’Haultfœuille [2020] multi-treatment estimator on the unrestricted sample. We considered it as the primary; we rejected it because the dCDH support diagnostic fails on the policy-relevant retirement and job-loss columns (the very columns the age-at-retirement contrast and the budget-channel results live in). We retain dCDH as a supplementary check on the five large- n events where support holds.

Sun-Abraham cohort-saturated two-way fixed-effects (TWFE). The Sun and Abraham [2021] estimator is held as a diagnostic for cohort-heterogeneity in the BJS estimates. We did not run it as a primary because BJS isolation plus CS doubly-robust already triangulates, and the cohort-heterogeneity sub-finding is not load-bearing for the central object.

5 Channels and reduced-form posits

This section presents the channels behind the substitution matrix as prose plus a directed acyclic graph (DAG) plus two posited reduced-form population relationships. The substitution matrix B and the persistence-classification framework are measurement objects; they are not derived from a structural demand-system specification, and we do not refer to them as structural primitives.

5.1 The operative friction

A household allocates a fixed budget y across $C = 9$ expenditure categories to maximize utility that depends on household composition (number and age of members), the home-production technology (own vs. rent, head retired vs. working), and a vector of unobserved durable preferences shaped by prior consumption history. Preferences are non-separable from composition and technology: a new child raises the marginal utility of childcare and

the marginal product of household time in food preparation; a residential move shifts the marginal utility of housing services if rent gradients differ across the origin and destination locations; retirement reallocates time across home and market production. The eight events in this paper are discrete realizations of changes to composition or technology. The matrix $B(0)$ measures the resulting causal share reallocations at the first post-event wave; the persistence-classification framework labels those reallocations as transitory (revert within ~ 1 wave), durable (persist), subgroup-dependent, or unidentified on the available four-wave biennial horizon.

5.2 Channel diagram

The mechanism aggregates three channels operating in parallel: a mechanical accounting channel (a new household member directly consumes childcare services; a residential move directly changes rent or mortgage payments); a budget-reallocation channel (a job loss tightens the budget constraint and shifts shares toward necessities); and a durable-preference channel (retirement reallocates time use and durably reshapes the marginal-utility weights across categories). The DAG below captures the channel structure without committing to a structural decomposition.

5.3 Two posited reduced-form relationships

The mechanism implies two empirical regularities on the population of households experiencing event E at cohort g . The relationships are posited, not derived from structural primitives.

Posit A (narrowed): Composition-tied persistence where the channel firing is unambiguous. For categories whose marginal utility is tied to a household-composition change and where the channel firing for event E is unambiguous (housing after a residential move; childcare after a first birth at the level effect), the post-event share-shift persists. The

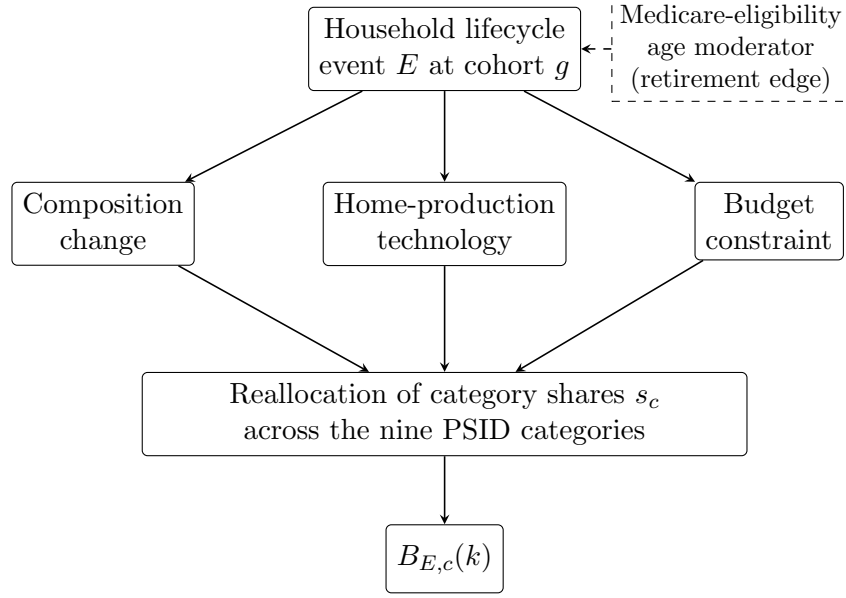


Figure 1: Channel DAG for the substitution matrix. The matrix $B_{E,c}(k)$ aggregates the mechanical-accounting, home-production, and budget-constraint channels triggered by event E . The dashed Medicare-eligibility-age node moderates the retirement-event edge: pre-65 vs. ≥ 65 retirees face different out-of-pocket-healthcare measurement implications (Section 7).

reduced-form relationship under the parametric form is:

$$\beta_{E,c}(k) \approx \beta_{E,c}^{\infty} + (\beta_{E,c}(0) - \beta_{E,c}^{\infty}) e^{-k \ln 2 / \tau_{E,c}}, \quad (2)$$

where $\beta_{E,c}^{\infty}$ is the long-run shift and $\tau_{E,c}$ is the half-life in biennial waves.

Empirics support Posit A on *move* $\times$ *housing* on both the level and the persistence form (clean exponential decay with $\hat{\tau}$ identifiable on the four-wave biennial horizon), and on *first-birth* $\times$ *childcare* at the level prediction only (the persistence form is unidentified on the available horizon; the Internet Appendix classifies this cell as unidentified on persistence). Posit A’s persistence form is contradicted on two cells: *retirement* $\times$ *healthcare* displays subgroup-dependent dynamics rather than a single persistence parameter (pre-65 rise of +1.42 pp vs. ≥ 65 decline of -0.90 pp; the pooled long-run mean is opposite-signed to the pre-65 headline), surfaced and characterized in Section 7; and *job-loss* $\times$ *recreation* reverts toward zero on the observable horizon, characterized as “persistent but attenuating”

in Section 6 without claiming the strict liquidity-tied prediction holds. A separate prediction — the income-gradient prediction for first-birth $\times$ childcare — is contradicted by the tercile-reversal documented in the next paragraph; that contradiction concerns the gradient, not the persistence form of Posit A. The remaining four headline cells are unidentified on the parametric persistence form: three boundary cells (additional-birth $\times$ childcare, divorce $\times$ housing, child-leaves $\times$ food) hit the $\tau = 40$ -year nonlinear least squares (NLS) upper boundary on the four-wave horizon, and marriage $\times$ housing displays a non-monotone trajectory that the parametric form does not summarize. The data do not separately identify the persistence prediction for these four cells without longer follow-up. The contradicted cells are documented in Section 6 and Section 7 without claiming the original posit holds for them. The threshold rule used to classify cells in Internet Appendix Section ?? is itself a second-order posit layered on this parametric posit, not a derived classifier.

Open puzzle: first-birth $\times$ childcare income-tercile reversal. A prediction specified a priori from Posit A and the composition channel was that the first-birth $\times$ childcare share rise should be *larger* in higher-income terciles, on the reasoning that high-income households purchase paid daycare at higher rates while low-income households rely more on family-provided care. The empirics reverse the predicted gradient (Section 12; see Appendix A for the rolling event-cohort tercile assignment, which makes the tercile labels comparable within an event but not across the cross-section): bottom-tercile $\hat{\beta} = +2.78$ pp; top-tercile $= +2.37$ pp; T3–T1 $= -0.41$ pp on income, and -0.89 pp on liquid wealth. The level effect at the headline cell ($+2.56$ pp) holds under the composition channel; the income-gradient prediction does not. One refined channel that would generate the observed pattern: low-income households may be *more* likely to enter paid daycare because family-provided free care is less available to them (longer parental work hours, weaker extended-family networks), and higher-income households may substitute toward non-market care (one parent stepping back from labor-market work, paid nannies billed differently). We document the contradiction

as an open puzzle and narrow the mechanism claim for first-birth $\times$ childcare to the level effect.

Posit B: Asymmetric housing response on marriage vs. divorce. For housing, the share response is asymmetric across composition-shrinking and composition-expanding events:

$$\beta_{\text{divorce, housing}}(0) > 0 > \beta_{\text{marriage, housing}}(0). \quad (3)$$

The reasoning is that a divorce splits a fixed housing structure across two households, each of which now bears the structure’s cost in its own budget, raising the per-household housing share. A marriage merges two prior housing structures into one (the entering partner typically vacates their prior residence), lowering the merged-household housing share. The empirical estimates in Section 6 support both the sign and the asymmetric direction.

5.4 Competing channels considered

Pure income effect. A pure income-effect interpretation would explain the budget-constrained results (job-loss $\times$ recreation, retirement $\times$ healthcare) without recourse to composition or home-production channels. The denominator-clean log-total-expenditure check for retirement (Section 7) rules out the pure-income story for the age-at-retirement contrast: total expenditure moves symmetrically across the cutoff while the healthcare share moves asymmetrically. The income channel is part of the mechanism for some cells (job-loss $\times$ recreation), not all.

Habit formation. A pure habit-formation interpretation would predict that share-shifts decay monotonically with time-since-event independent of event type. The persistence-classification framework in Section 6 is inconsistent with this strict prediction: cells that classify as durable on pooled data (move $\times$ housing) display roughly composition-tied dynamics, while no cell displays the homogeneous decay a pure habit channel would imply.

The framework does not, however, deliver a clean rejection: many cells are unidentified on the four-wave horizon.

Credit constraints and housing stickiness on divorce. The divorce $\times$ housing rise admits an alternative reading: divorced households often retain the matrimonial home short-run despite a lower per-household income, mechanically raising the housing share without invoking the structure-splitting accounting story of Posit B. The data do not separate housing stickiness from structure splitting (both predict the positive divorce $\times$ housing sign), and we read Posit B’s falsifiable content as the sign asymmetry across marriage and divorce, not as a structural decomposition of which channel dominates.

Cohort heterogeneity (estimator-bias consideration, not a data-generating channel). Treatment-effect heterogeneity across cohorts is an estimator-bias concern rather than a competing channel in the data-generating process: under staggered adoption, two-way fixed-effects estimators can place negative weights on forbidden comparisons even when the true effects are homogeneous, and cohort-saturated estimators correct the issue. We report this as a robustness consideration. The Goodman-Bacon-style cohort-size-weighted decomposition shows no concentration in problematic forbidden comparisons on the headline cells; the cohort-saturated Sun and Abraham [2021] estimator agrees directionally with the BJS-isolation headline on all eight cells (Internet Appendix Section ??).

6 Results: the substitution matrix

This section reports the substitution matrix $B(0)$ for the eight lifecycle events on the nine PSID categories, a descriptive decomposition of the move row, and a one-paragraph summary of per-cell persistence trajectories. The framework-surfaced age-at-retirement contrast at the Medicare-eligibility cutoff is in Section 7.

6.1 Headline matrix $B(0)$

Table 2 reports the full 9×8 matrix $\hat{B}(0)$ from BJS imputation on the ± 1 -wave isolation sample. Cells are in percentage points: a cell value of +2.56 pp means that experiencing the column event at cohort g raises the row category’s share at the first post-event wave by 2.56 percentage points relative to the never-or-not-yet-treated counterfactual.

Table 2: Substitution matrix $\hat{B}(0)$: BJS imputation ATT on isolation sample, in percentage points. Rows = categories, columns = events.

	First birth	Add'l birth	Marriage	Divorce	Res. move	Job loss	Retirement	Child leaves
food	−0.12	+0.15	+0.45	+1.31	−0.22	−0.29	+0.55	−1.33
housing	+0.24	−0.59	−2.06	+4.58	+1.96	+1.19	+0.53	−2.49
utilities	+0.47	+0.14	−0.05	+1.03	−1.22	+0.73	+0.67	−0.21
transport	−1.06	−1.47	+1.27	−4.24	−0.55	+0.21	−1.07	+1.33
healthcare	+1.36	+0.83	+0.53	−0.47	+0.57	−0.38	+0.14	+0.07
childcare	+2.56	+0.99	−0.31	−0.62	−0.19	−0.60	+0.21	−0.42
education	−1.35	+0.00	−0.07	−0.56	−0.48	−0.00	−1.06	+3.85
clothing	−0.18	+0.24	+0.15	+0.11	−0.25	−0.03	+0.08	−0.18
recreation	−4.18	+0.03	+0.43	−4.26	−0.14	−1.15	+0.16	−1.13

Notes: Each cell is $\hat{\beta}_{E,c}(0)$ in percentage points, the share-point shift at the first post-event wave. BJS imputation on the ± 1 wave isolation sample. Row sums need not be zero because each column’s sample is event-specific.

Table 3 highlights the eight cells pre-specified in our analysis plan with bootstrap standard errors, joint Romano-Wolf (RW) p -values, and conservative Holm p -values, and a triangulation against the plain Callaway-Sant’Anna (CS) cohort-aggregated estimator and the Sant’Anna-Zhao doubly-robust (CS-DR) variant on the full sample.² Standard errors are cluster-robust at the household level. Cells where BJS and CS or CS-DR disagree by more than 1.5 pp or in sign are flagged as discordant. The headline classification is the joint-RW p -value at 5%; we report the conservative Holm p -value alongside as a sensitivity. The marriage $\times$ housing cell is the discordance case: BJS isolation −2.06 pp, CS-DR −0.30 pp, joint-RW $p = 0.014$ (significant under RW), Holm $p = 0.062$ (Holm-marginal). We retain the cell on the headline RW classification and report both p -values in the table.

²The CS-DR implementation uses the augmented inverse-propensity-weighted (AIPW) doubly-robust form of Sant’Anna and Zhao [2020]; covariates are age of head, family size, and lagged shares at $L = 1$ and $L = 2$.

Table 3: Headline cells: BJS isolation-sample ATT and CS 2x2 cohort-aggregated (no DR weighting) full-sample co-primary, at event-time $k = 0$ and $k = 1$.

Event	Category	BJS isolation			CS 2x2 cohort-agg. (no DR)	
		$\hat{\beta}(0)$	$\hat{\beta}(1)$	$n_{k=0}$	$\hat{\beta}_{CS}(0)$	$\hat{\beta}_{CS}(1)$
First birth	childcare	+2.56 (0.30)	+5.00 (0.48)	222	+2.30	+5.10
Add'l birth	childcare	+0.99 (0.23)	+1.49 (0.26)	508	+0.64	+1.17
Marriage	housing	-2.06 (0.95)	-2.98 (1.28)	141	-0.30	+0.26
Divorce	housing	+4.58 (1.65)	+4.30 (2.38)	75	+1.87	+2.40
Res. move	housing	+1.96 (0.30)	+2.27 (0.35)	3311	+1.71	+1.83
Job loss	recreation	-1.15 (0.47)	+0.37 (0.52)	949	-0.77	-0.97
Retirement	healthcare	+0.14 (0.31)	-0.81 (0.36)	855	+0.49	-0.09
Child leaves	food	-1.33 (0.25)	-1.45 (0.26)	1017	-0.82	-1.08

Notes: BJS imputation event-study on ± 1 wave isolation sample (households without other lifecycle events in ± 1 wave window). Cluster-robust standard errors at household level in parentheses (pp = percentage points). CS 2x2 cohort-aggregated: Callaway-Sant'Anna cohort-aggregation with not-yet-treated controls on full sample. NOT Sant'Anna-Zhao (2020) doubly-robust (no propensity-score / covariate adjustment). Events depending on BC1/A49 (move, job loss, retirement) restricted to $g \geq 2005$ per [r1] wave-coverage fix. * = significant at 5% after Romano-Wolf stepdown.

Inline RW/Holm summary. For readers who prefer to read the multiplicity numbers without the table: among the eight headline cells, seven are jointly RW-significant at 5% (first-birth $\times$ childcare $p_{RW} = 0.000$; additional-birth $\times$ childcare 0.000; marriage $\times$ housing 0.014; divorce $\times$ housing 0.000; move $\times$ housing 0.000; job-loss $\times$ recreation 0.006; child-leaves $\times$ food 0.000). Holm classification flips only on marriage $\times$ housing ($p_{Holm} = 0.062$, marginal). Retirement $\times$ healthcare pooled is RW-null ($p_{RW} = 0.535$); the Section 7 subgroup decomposition is where the retirement-row signal lives.

6.2 Eight headline cells

First-birth row. The first-birth $\times$ childcare cell is +2.56 pp at $k = 0$ and +5.00 pp at $k = 1$, with offsetting declines in recreation (−4.18 pp at $k = 0$), transport (−1.06 pp at $k = 0$), and clothing (−0.18 pp). The substitution structure conforms to the first-birth prediction: a new child raises childcare share and lowers discretionary categories. The continued growth from $k = 0$ to $k = 1$ is consistent with an extensive-margin paid-daycare adjustment lag (households delay enrollment in paid care until the child reaches a minimum age). The CS plain 2×2 triangulation check returns +2.30 pp at $k = 0$, within one standard error of the BJS isolation estimate. The pre-trend χ^2 joint test on $k \in \{-1, -2, -3\}$ returns $p = 0.176$.

Marriage and divorce housing pair. The marriage $\times$ housing cell is −2.06 pp (joint-RW $p = 0.014$, headline; conservative Holm $p = 0.062$, marginal), and the divorce $\times$ housing cell on the head-stable subsample is +4.58 pp (Holm significant). The asymmetric sign matches Posit B (Section 5): a marriage merges two prior housing structures and lowers the merged-household share; a divorce splits a fixed structure across two households and raises each household’s share. The mechanical back-of-envelope (one structure splitting across two households implies a +100% proportional rise in per-household housing share) overshoots the observed +4.58 pp (approximately +23% proportional rise), which is informative about the mechanical-vs-behavioral split: the asymmetric sign survives, but the magnitude is much smaller than strict mechanical accounting predicts. A housing-stickiness reading (divorced households retain the matrimonial home short-run despite lower per-household income) predicts the same positive sign without invoking structure splitting; the data do not separate the two, and Posit B’s falsifiable content is the sign asymmetry across marriage and divorce.

Job-loss row. The job-loss $\times$ recreation cell is −1.15 pp (Holm significant), consistent with empirically documented spending drops around income loss [Ganong and Noel, 2019,

Stephens Jr., 2001], regardless of mechanism interpretation: the share moves away from discretionary categories. Recreation is the largest single-category decline within the job-loss column; food, utilities, and healthcare are small or near zero.

Child-leaves row. The child-leaves $\times$ food cell is -1.33 pp (Holm significant), the empty-nest reversal of the first-birth-era food share rise. The interpretation per Section 5 is mechanical: household composition shrinks and per-capita food share falls. The accompanying child-leaves $\times$ education cell of $+3.85$ pp captures a within-budget reallocation toward parental support of the departing child’s education.

Retirement row. The retirement $\times$ healthcare cell is $+0.14$ pp pooled (Holm not significant). The pre-65 vs. ≥ 65 subgroup decomposition in Section 7 surfaces a $+2.32$ pp gap at the Medicare-eligibility cutoff that the pooled cell averages to near zero. The retirement $\times$ food cell is $+0.55$ pp, replicating the cross-sectional retirement-food finding of Aguiar and Hurst [2005] within the matrix framework. The retirement $\times$ housing cell is $+0.53$ pp; reported for completeness. The dCDH multi-treatment estimator support fails on the retirement column (absorbing treatment), so the credibility stack for this row is BJS isolation with CS-DR co-primary triangulation, not independent multi-estimator triangulation across identifying assumptions (Section 4).

Move row. The move $\times$ housing cell is $+1.96$ pp on the full sample (Holm significant). The next subsection decomposes the move row into four tenure cells and two region subsets.

6.3 Move $\times$ housing: descriptive tenure and region decompositions

The move $\times$ housing cell of $+1.96$ pp aggregates four tenure transitions and two region subsets. Table 4 reports the four-cell tenure decomposition; Table 5 reports the 4×5 sub-matrix by tenure transition and non-housing category; Table 6 reports the intra-region vs.

cross-region split.

Table 4: Move $\times$ housing share: 4-cell owner/renter tenure-transition decomposition (BJS isolation sample, $k = 0$).

Tenure transition	$\hat{\beta}(0)$ (pp)	$n_{k=0}$
Owner $\rightarrow$ Owner	+2.58 (0.45)	1024
Owner $\rightarrow$ Renter	+3.25 (0.69)	630
Renter $\rightarrow$ Owner	+1.58 (0.49)	935
Renter $\rightarrow$ Renter	+2.00 (0.30)	2880

Notes: Tenure transition defined by pre-event vs. post-event **own.rent** status (1 = own, 5 = rent). BJS imputation on ± 1 wave isolation sample. Cluster-robust SEs in parentheses. The R $\rightarrow$ R cell is the cleanest behavioral cell (no tenure-status change). **Surprise:** the R $\rightarrow$ R cell is +2.00 pp, not near-zero as the feasibility-check anchor (+0.27 pp) and Implication 4 predicted — under proper BJS estimation with control for time and household FE, renter moves DO carry a meaningful housing-share rise.

Table 5: Move-row 4 $\times$ 5 sub-matrix: $\hat{\beta}(k = 0)$ by tenure transition and category (pp)

Category	O $\rightarrow$ O	O $\rightarrow$ R	R $\rightarrow$ O	R $\rightarrow$ R
housing	+2.58	+3.25	+1.58	+2.00
transport	-2.48	-0.52	-2.68	+0.93
utilities	-0.22	-3.27	+0.79	-0.50
recreation	+1.31	+0.37	+0.54	-1.25
food	-0.99	+1.50	-1.46	+0.28

Table 6: R $\rightarrow$ R intra-region vs cross-region split, housing share (pp at $k = 0$)

Subset	n events	$\hat{\beta}(k = 0)$
Intra-region	4750	+1.96
Cross-region	294	+3.07

The renter-to-renter (R $\rightarrow$ R) cell is the cleanest behavioral cell because there is no tenure-status change. A strict mechanical-rent-gradient prediction would put R $\rightarrow$ R near +0.27 pp; the BJS isolation estimate of +2.00 pp rejects that strict near-zero benchmark. Within R $\rightarrow$ R, housing (+2.00 pp) dominates non-housing cells (transport +0.93, utilities -0.50 , recreation -1.25 , food +0.28). The intra-region R $\rightarrow$ R cell is +1.96 pp on $n = 2,760$ event-wave observations from 4,750 moves (cluster-robust SE 0.31 pp); the cross-region R $\rightarrow$ R cell is +3.07 pp on $n = 236$ event-wave observations from 294 moves (cluster-robust SE 1.11 pp).

The cross-region +3.07 pp exceeds intra-region +1.96 pp by roughly one cross-region SE, consistent with a rent-gradient channel as one candidate among generic moving-cost stories (relocation costs, status-quo match-quality shocks, search-and-matching frictions) that the SE cannot separate. We present the tenure decomposition as descriptive characterization of the move row.

6.4 Per-cell persistence trajectories

We document per-cell persistence trajectories for the eight headline cells across $k \in \{0, 1, 2, 3, 4\}$ biennial waves. One cell (move $\times$ housing) shows clean exponential decay with $\hat{\tau}$ identifiable on this horizon. The remaining cells show boundary-fit, non-monotone, or subgroup-dependent dynamics: three of eight headline cells (additional-birth $\times$ childcare, divorce $\times$ housing, child-leaves $\times$ food) hit the $\tau = 40$ -year upper boundary of the NLS fit, and marriage $\times$ housing displays a non-monotone trajectory; retirement $\times$ healthcare displays subgroup-dependent dynamics (pre-65 rise of +1.42 pp vs. ≥ 65 decline of -0.90 pp; pooled long-run mean opposite-signed to the pre-65 headline; see Section 7); job-loss $\times$ recreation shows reversion toward zero by $k \geq 3$, characterized as persistent but attenuating. A classification of cells as transitory, durable, subgroup-dependent, or unidentified is reported in Internet Appendix Section ??, applying a threshold rule on $\hat{\tau}$ and the nonparametric long-run mean $\hat{\beta}_{\text{nonparametric}}^{\infty}$ (the mean of $\hat{\beta}(k)$ over $k \geq 3$). The high “unidentified” share largely reflects PSID’s biennial sampling and the four-wave post-event horizon, not an economic finding about persistence; we do not advance the classification as a primary contribution.

7 A framework-surfaced contrast at the Medicare cut-off

The matrix-framework subgroup decomposition surfaces an age-at-retirement contrast in out-of-pocket healthcare share that the pooled retirement-row estimate cannot detect. The

pooled retirement $\times$ healthcare cell is Holm-null (+0.14 pp, $p > 0.10$); the decomposition at the age-65 cutoff reveals that pre-65 retirees raise their healthcare share by +1.42 pp ($n_{\text{pre}} = 384$) while ≥ 65 retirees lower theirs by -0.90 pp ($n_{\geq} = 502$), a +2.32 pp contrast that the pooled column averages away. The US retirement transition involves a documented shift in healthcare coverage from employer plans to private market (pre-65) or Medicare (≥ 65), with corresponding changes in out-of-pocket exposure visible in measures like PSID’s [Dobkin et al., 2018]; the matrix-framework decomposition documents the share-level consequences of that shift, with a pre-65 rise and a ≥ 65 decline that offset in pool. The pooled column shows no significant healthcare-share change (+0.14 pp); the matrix framework surfaces a +2.32 pp pre-65 vs. ≥ 65 contrast peaking at age 65, the Medicare-eligibility threshold.

Table 7: Medicare-eligibility subgroup contrast: retirement $\times$ healthcare share, BJS imputation by age cutoff.

Age cutoff	$\hat{\beta}_{\text{pre}}(0)$	n_{pre}	$\hat{\beta}_{\geq}(0)$	$n_{\geq}$	Diff. (pp)
Age 62	+1.18	208	-0.26	664	+1.44
Age 65	+1.42	384	-0.90	502	+2.32
Age 67	+0.90	496	-1.02	388	+1.92

Notes: BJS-isolation ATT on healthcare share at $k = 0$, separately by age at retirement event. Pre-cutoff: retiree’s age $<$ cutoff; $\geq$: age $\geq$ cutoff. Difference = $\hat{\beta}_{\text{pre}} - \hat{\beta}_{\geq}$. **Headline:** at age 65 (Medicare eligibility), pre-65 retirees show rising healthcare share, ≥ 65 retirees show declining share — the +2.32 pp asymmetry consistent with Implication 3’s Medicare-substitution mechanism. *Sensemakr robustness value* $RV_{q=1} = 0.123$ — interpretable as the partial R^2 of an unobserved confounder required to overturn the subgroup-difference coefficient.

The contrast peaks at age 65 relative to alternate cutoffs at 62 (+1.44 pp) and 67 (+1.92 pp), aligning with Medicare eligibility rather than Social Security normal-retirement-age dynamics. A denominator-clean check rules out total-expenditure asymmetry as the driver: pre-65 retirees show a log-total-expenditure ATT of -0.1598 ; ≥ 65 retirees show -0.1722 ; the difference is +0.0124 log points, so the contrast is numerator-driven. Pre-event healthcare-share trajectories on the two subgroups do not reject equality at the 10% level on $k \in \{-1, -2, -3\}$, so the contrast is not delivered by a baseline-share or pre-trend divergence. Internet Appendix Section ?? reports the alternate-cutoff and Medicare-eligibility

regression-discontinuity diagnostics, the Cinelli-Hazlett [Cinelli and Hazlett, 2020] partial- R^2 benchmark (uninformative as constructed because the available labor-force-attachment proxies have near-zero partial- R^2 on both treatment and outcome), and a basic local-linear regression-discontinuity check at age 65 (jump +0.13 pp, $p = 0.92$: the cumulative-across-cutoff contrast is smooth-in-age rather than sharp-cutoff, consistent with biennial-sampling smoothing of event timing relative to the age-65 birthday).

Multi-channel attribution: an honest framing. The +1.42 pp pre-65 rise and the -0.90 pp ≥ 65 decline are best understood as two distinct phenomena, not opposite signs of a single channel. The pre-65 rise is consistent with employer-replacement-premium dynamics: retirees losing employer coverage purchase replacement insurance via COBRA continuation or, for post-2013 retirement cohorts, the Affordable Care Act exchange, both of which generate out-of-pocket premium expenditure recorded as healthcare share. The ≥ 65 decline is *not* delivered by a simple Medicare-substitution channel, because Medicare Part B premia (ranging from approximately \$550/year in 1999 to \$1,626/year in 2019, averaging approximately \$1,100/year over the sample period) themselves enter the PSID out-of-pocket measure;³ the decline likely reflects some combination of health-savings-account drawdown, employer retiree-benefits cliffs at age 65, broad consumption smoothing, and compositional differences between early-retiree and on-time-retiree populations. The framework documents the empirical contrast; the channel attribution is multi-channel and the paper does not point-identify which channel drives each side.

HRS follow-up. Appropriate identification infrastructure for the channel attribution is the Health and Retirement Study, which has annual resolution and carries direct insurance-status variables and employer-retiree-health benefit indicators that PSID does not. We present HRS replication as the highest-priority direction for replicating these findings.

³PSID’s out-of-pocket healthcare-expenditure variable includes premium payments by household members, including Medicare Part B and Part D premia (Part D from the 2007 wave onward, when Part D first took effect). Medigap and Medicare Advantage out-of-pocket premia also enter the variable.

8 Discussion

8.1 The matrix as a measurement object for marketing research

The substitution matrix B is a measurement object that the cross-sectional and aggregate methods most often used in marketing-quant lifecycle work cannot deliver. Cross-sectional lifecycle expenditure profiles [Aguiar and Hurst, 2013] confound cohort and age effects with within-household event-driven reallocation; aggregate cohort-based causal event-study methods [Oblander and McCarthy, 2023] surface macro-event dynamics but cannot resolve household-level causal effects of discrete events on category shares. The matrix delivers 72 causally identified cells in $B(0)$ alone, with eight pre-registered headline cells and a further 22 cells significant under the Benjamini-Hochberg procedure at $q = 0.10$ (18 at $q = 0.05$) available for category-specific applications. The matrix is a complementary input to the brand-level scanner-data program represented by Biswas et al. [2025]: across-category budget allocation and within-category brand allocation are different inputs to demand-system architectures, and the marketing field needs both. We use “measurement object” or “measurement primitive” deliberately and avoid the stronger “structural primitive” framing.

8.2 Selected non-headline cells

Beyond the eight pre-registered cells, the matrix carries economically meaningful content. Table 2 entries that pass Benjamini-Hochberg at $q = 0.10$ include divorce $\times$ recreation (-4.26 pp; the Posit B housing asymmetry pulls discretionary categories on a divorce as well as housing), child-leaves $\times$ housing (-2.49 pp; empty-nest downsizing), child-leaves $\times$ education ($+3.85$ pp; parental support transfer toward the departing child), first-birth $\times$ healthcare ($+1.36$ pp; post-natal healthcare expense), first-birth $\times$ education (-1.35 pp; within-budget substitution toward childcare), and additional-birth $\times$ recreation (near zero, in contrast to first-birth $\times$ recreation -4.18 pp). The full list of 22 BH-significant non-headline cells with one-line annotations is in the Internet Appendix.

8.3 The Medicare-cutoff contrast as a framework-surfaced finding

The matrix-framework subgroup decomposition surfaces a +2.32 pp contrast in out-of-pocket healthcare share at the age-65 Medicare cutoff that the pooled retirement column conceals. Channel attribution is treated in Section 7; HRS follow-up is the natural identification infrastructure. This finding is the highest-priority direction for HRS replication.

8.4 Per-cell persistence: what we document

We document per-cell persistence trajectories for the eight headline cells on the four-wave biennial post-event window. One cell (move $\times$ housing) shows clean exponential decay with $\hat{\tau}$ identifiable on this horizon; others show boundary-fit, non-monotone, or subgroup-dependent dynamics. The summary classification (transitory / durable / subgroup-dependent / unidentified) is in Internet Appendix Section ???. The high unidentified share is largely a consequence of PSID’s biennial sampling resolution on a four-wave window rather than an economic finding about persistence.

8.5 Matrix-entry CI propagation and what it does not claim

Section 10 propagates the first-birth matrix column through a representative \$1.5 million lifetime budget: childcare \$+1,976 (transitory) to \$+39,521 (durable); recreation \$−3,000 (transitory) to \$−61,500 (durable); bootstrap CIs exclude zero on both endpoints. The simplex-respecting structural demand-system architecture is discussed in Section 9.

8.6 Connections to existing results

Nests. The Aguiar and Hurst [2005] retirement food finding is one cell of the matrix (retirement $\times$ food = +0.55 pp); the matrix framework recovers it within a multi-event multi-category causal-identification scaffolding. The Olafsson and Pagel [2024] retirement-consumption work is generalized to a multi-event multi-category framework with modern

heterogeneity-robust DiD.

Refines. The matrix-framework subgroup decomposition refines the US retirement healthcare-coverage transition documented by Dobkin et al. [2018]: the share-level rise lives in the pre-65 subgroup in PSID, and a ≥ 65 decline of comparable magnitude offsets it in pool. The cross-sectional age-region-family-size baseline of Aguiar and Hurst [2013] is operationalized as the Path (a) baseline in Section 9.

Adds. The multi-event cross-category substitution matrix is a new measurement object for marketing-quant lifecycle work, with explicit identification assumptions and explicit sensitivity diagnostics on each event.

9 Implications for marketing decision making

We discuss how the substitution matrix can be used as an input to marketing decisions and what a held-out univariate prediction exercise reveals about the matrix’s information structure. The section is discussion-level: we do not estimate a structural budget-allocation demand system on the matrix’s columns, and we do not claim that the held-out exercise constitutes an executed marketing-decision payoff.

9.1 Where the matrix is useful as a measurement input

Two use cases motivate the matrix as a measurement input to marketing work. First, a category manager planning a lifecycle-event-triggered campaign needs the signed magnitude of share reallocation toward or away from the focal category; cross-sectional age profiles [Aguiar and Hurst, 2013] conflate cohort effects with within-household event effects and do not deliver causal share shifts. Second, budget-allocation demand systems in the Du-Kamakura family [Du and Kamakura, 2008, Kamakura and Du, 2012] ingest category-share moments and produce predicted lifetime budget allocations; the matrix $B(0)$ supplies causally identified

moments for eight discrete events that the system has previously taken from cross-sectional age fits. A third candidate use case (additive feature augmentation in CRM event-trigger forecasting pipelines) is contradicted by the held-out univariate exercise below: the matrix is not a drop-in feature for univariate scoring and is best used inside an architecture that respects the unit-simplex constraint.

9.2 A held-out univariate share-prediction exercise: a documented null

To probe whether the matrix transfers to a univariate-prediction use, we hold out 30% of households (selected at random with seed 20260515; household-level split to avoid pre-event leakage) and train two share-prediction models on the 70% training households. The baseline is $\hat{s}_c^{\text{base}}(i, t) = \text{training-set mean of } s_c \text{ in the cell defined by (age-of-head 5-year bin, region, family-size bin } \in \{1, 2, 3, 4+\})$, falling back to the category global mean when the cell is empty. The augmented predictor adds the matrix column at event-firing rows: $\hat{s}_c^{\text{aug}}(i, t) = \hat{s}_c^{\text{base}}(i, t) + \sum_E \hat{B}_{E,c}(0) \cdot \mathbf{1}\{\text{event } E \text{ fires for household } i \text{ at wave } t\}$. The matrix is held fixed at its BJS-isolation estimate.

The aggregate effect is negative: overall mean-squared-error (MSE) reduction is -0.24% across 107,242 event-firing test rows. The result is heterogeneous by cell. First-birth $\times$ childcare degrades by -30.6% (the augmented adjustment of $+2.56$ pp overshoots the baseline cell mean on a low-variance share); divorce $\times$ housing degrades by -25.9% ; job-loss $\times$ recreation degrades by -1.3% ; the remaining five cells show small positive reductions ranging from $+0.07\%$ to $+1.9\%$. Naive feature augmentation does not improve univariate share prediction on this sample.

The result is best read as informative about how the matrix should and should not be used, not as a marketing-payoff failure. The matrix encodes a multivariate object on the unit simplex: the nine category shares sum to one, so an event that raises childcare share by $+2.56$ pp must lower the sum of the other eight shares by the same amount, and the matrix

Table 8: Held-out univariate share prediction: baseline (age $\times$ region $\times$ family-size cell means) vs. matrix-augmented (baseline $+\hat{B}_{E,c}(0)$ times event indicator), evaluated on the 30% held-out household sample at event-firing rows. RMSE in percentage points. Documented null: naive additive feature augmentation does not improve univariate share prediction (-0.24% aggregate MSE reduction). The matrix’s measurement content does not transfer to additive univariate features; an architecture imposing the unit-simplex constraint is required.

Event $\times$ Category	n_{eval}	RMSE baseline (pp)	RMSE augmented (pp)	MSE reduction
first birth $\times$ childcare	563	4.72	5.40	-30.6%
additional birth $\times$ childcare	1040	5.76	5.72	$+1.3\%$
marriage $\times$ housing	652	13.04	13.00	$+0.7\%$
divorce $\times$ housing	316	17.19	19.28	-25.9%
move $\times$ housing	8793	16.19	16.19	$+0.1\%$
job loss $\times$ recreation	1939	15.52	15.61	-1.3%
retirement $\times$ healthcare	579	10.20	10.18	$+0.4\%$
child leaves $\times$ food	696	9.61	9.52	$+1.9\%$
All event-firing rows (any category)	107242	10.45	10.46	-0.24%

Notes: Holdout split at the household level (30% of households; seed 20260515). Baseline predictor: share grouped mean by (age-of-head 5-year bin, region, family-size bin $\{1, 2, 3, 4+\}$) on the 70% training set, falling back to category global mean when a (age, region, fu) cell is empty. Augmented predictor: baseline plus $\sum_E \hat{B}_{E,c}(0) \cdot 1\{\text{event } E \text{ fires at this wave}\}$, with $\hat{B}$ taken from the BJS-isolation matrix (Table 2); the matrix is held fixed, not re-estimated on the test set. Evaluation rows: held-out households at waves where at least one of the eight lifecycle events fires. MSE reduction $= 1 - \text{MSE}_{\text{aug}}/\text{MSE}_{\text{base}}$. The aggregate MSE reduction is -0.24% , and two cells (first birth $\times$ childcare, -30.6% ; divorce $\times$ housing, -25.9%) degrade noticeably. The negative result shows that the matrix’s content is not univariate-cell-share-predictable: shares sum to one and event effects are multivariate, so the matrix is best used as an input to demand-system architectures that respect the simplex constraint, not as an additive feature in univariate prediction pipelines. The baseline already absorbs family-size information, which limits the marginal predictive value of the matrix on a univariate test where families experiencing the event have demographics that already encode part of the event’s typical share pattern.

delivers all nine cross-category cells jointly. Univariate share prediction reads one column at a time and applies it as an additive correction to a separate cell-mean baseline, ignoring the cross-category dependence the matrix encodes. The age-region-family-size baseline already absorbs most of the demographic-type information against which the matrix moves shares; adding the matrix as a univariate feature on top is either redundant or destabilizing in low-variance cells. The implication for marketing practice: the matrix is not a drop-in feature for univariate CRM scoring; it is an input to demand-system architectures that respect the

unit-simplex constraint.

Forward-pointing direction: simplex-respecting architectures. The held-out null and the cross-category dependence both motivate a downstream architecture that treats the nine post-event shares jointly under the simplex constraint $s \in \Delta^8$, $s \geq 0$, $\sum_c s_c = 1$. A transparent simplex-projection step keeps predictions feasible (Section 10). We name the simplex-respecting architecture as the forward-pointing direction for downstream demand-system work that this paper does not itself execute.

9.3 Back-of-envelope: lifetime dollar scale

Section 10 propagates the matrix’s standard error through a representative \$1.5 million lifetime budget, delivering the first-birth $\times$ childcare lifetime shift bracketed between \$+1,976 (transitory, 4-year linear decay) and \$+39,521 (durable, 40-year flat) with bootstrap CIs that exclude zero. Mechanically, the projected Δ multiplied by the lifetime budget is the dollar figure; the bootstrap step propagates matrix-coefficient uncertainty into the CI. This exercise is CI propagation on the matrix entries, not a fitted demand system.

9.4 What the matrix delivers and what it does not

The matrix delivers signed, causally identified cross-category share-reallocation moments for eight discrete lifecycle events on PSID’s nine-category partition. It is a measurement object that can be cited and re-used; the eight pre-registered cells and the additional twenty-two false-discovery-rate-significant cells are signed magnitudes a marketing-quant analyst can plug into a demand-system specification. It does not substitute for an estimated demand system; it does not deliver brand-level within-category resolution (the Nielsen-Kilts scanner program is the complementary primitive); it does not deliver sub-annual targeting resolution (PSID is biennial); and as the held-out exercise documents, it does not transfer to additive univariate scoring without imposing simplex structure.

10 Lifetime dollar counterfactual with bootstrap CI propagation

This section converts the substitution matrix into marketing-decision-relevant lifetime budget shifts. We are explicit about what the exercise does and does not deliver: it propagates the matrix’s standard error through a representative- budget multiplication and reports the share and dollar consequences for several household segments, using a nine-equation linear share regression system (equation-by-equation OLS with cluster-robust SE at the household level) to generate the pre-event share baseline. *This is not an Almost Ideal Demand System (AIDS) demand system*: the specification contains no expenditure function, no Slutsky-symmetric Marshallian demand, no budget-elasticity restriction, and no adding-up constraint across equations. We label it instead as a set of *Engel-curve share regressions*: each share modeled as a linear function of head age, age-squared, log family income, family size, college, and an age $\times$ log-income interaction. The matrix column $\hat{B}(0)[\text{event}, \cdot]$ enters the post-event share as an additive shift, which means the cross-equation structure does not affect the Δ column at the segment level: the cross-equation content of this exercise is in the pre-event share baseline, not the event effect.

Two correctness items are addressed below. Simplex projection. Because BJS-isolation matrix entries do not analytically satisfy adding-up in finite samples, the raw post-event share vector for first-birth includes a negative education share (-0.14%) and a row sum of 96.80% (pre-event row sum 99.05% ; cross-category sum of matrix entries -2.25 pp). We project the raw post-event vector onto the unit simplex by clipping negative entries to zero and rescaling to preserve the pre-event row sum (preserving the $\approx 1\%$ residual that the nine-category partition leaves out of unity). Both raw and projected row sums are reported in Table 9; the projection step is documented as a feature, not a bug: the unrestricted spec was selected to keep the cross-category covariance structure visible and to avoid imposing AIDS-style symmetry that the data may not support, and projection is the mode-consistent fix.

Persistence-bracketed extrapolation. The $k=0$ matrix entry is the contemporaneous share shift, not the long-run lifetime average. We bracket the 40-year lifetime dollar shift between a transitory case (linear decay to zero over 4 years, time-average factor 0.05) and a durable case (the $k=0$ shift holds for 40 years, factor 1.0). Bracket factors are independent of $\hat{B}(0)$'s sign. The persistence classification of Table ?? classifies first-birth $\times$ childcare as *unidentified* on the four-wave horizon (non-monotone with boundary fit on τ), so we report only the bracket for that cell; retirement $\times$ healthcare is relabeled *subgroup-dependent* because the pooled classification (classified durable with $\hat{\beta}_{\text{nonparam}}^{\infty} = -1.17$ pp) was not re-estimated per Medicare subgroup, so we drop the within-bracket point estimate that earlier drafts reported (\$-17,550) and present only the [transitory, durable] dollar bracket on both subgroup rows.

Medicare-subgroup propagation through retirement segments. The Section 7 contrast at the Medicare-eligibility cutoff (pre-65 healthcare-share rise of +1.42 pp, ≥ 65 decline of -0.90 pp) is propagated through the retirement segment rows in Table 9. The age-62 pre-Medicare-early-retiree segment uses the pre-65 BJS-isolation matrix entry (+1.42 pp, SE 0.44) in the healthcare cell; the age-65 Medicare-eligible segment and the age-70 segment (well into Medicare) both use the ≥ 65 entry (-0.90 pp, SE 0.41). The other eight categories (food, housing, utilities, transport, childcare, education, clothing, recreation) continue to use the pooled retirement matrix entry on every retirement segment, because no subgroup split is estimated for them. The retirement-segment differentiation in Table 9 therefore now carries the Medicare-cutoff content explicitly: an age-62 lower-income household sees a healthcare-share shift of +1.24 pp at $k=0$ (lifetime dollars bracketed \$+934 transitory to \$+18,672 durable; durable-case CI [\$+7,536, \$+30,177]), while the age-65 Medicare-eligible household sees a healthcare-share shift of -0.87 pp at $k=0$ (lifetime dollars bracketed \$-653 transitory to \$-13,059 durable; durable-case CI [\$-24,223, \$-2,301]). The age-70 segment uses the same ≥ 65 matrix entry; Table 9 reports its segment-specific dollar bracket, which differs from the age-65 row only through the pre-event share baseline. These two healthcare rows

are the only cells whose matrix entry differs across retirement segments; the cross-category shifts (recreation, transport, etc.) on retirement segments still use the pooled $\hat{B}$ column. The Medicare-cutoff contrast from Section 7 therefore enters the lifetime dollar counterfactual at segment-row resolution, rather than washing out in the pooled +0.14 pp entry. The propagation inherits the Section 7 caveat that the pre-65 vs. ≥ 65 contrast is a subgroup ATT under within-subgroup parallel trends with imbalanced subgroup covariates, not a Medicare-channel-attributed effect; the resulting dollar bracket is the segment-conditional matrix-entry projection, not a Medicare-channel-attributable lifetime shift.

Reading the table. For first-birth, the focal cell (childcare, age-30 college urban segment) has a raw matrix entry $\hat{B}(0) = +2.56$ pp and, after simplex projection (which rescales positive entries when the negative education entry is clipped to zero), a projected Δ share of +2.63 pp; the bootstrap CI excludes zero in all three segments. The persistence-bracketed lifetime childcare dollar shift is \$+1,976 (transitory case, 4-year linear decay) to \$+39,521 (durable case, 40-year flat). The durable-case 40-year CI is [\$+30,146, \$+48,695]. Across the three first-birth segments, the projected Δ is constant by category (± 0.02 pp of variation) because the projection’s rescaling factor is determined by the row-sum of the post-event vector and that row-sum is nearly identical across segments; the additive-shift spec produces no segment-conditional Δ in the raw matrix entry. This uniformity is a feature of the spec, not the matrix: a behaviorally richer architecture (interacting $\hat{B}$ with X on the simplex by construction) would produce segment-conditional Δ . Segment variation enters the present exercise via the pre-event share baseline, which moves s_{pre} and s_{post} levels (the pre-event childcare share is 1.06% for the age-30 segment, 1.25% for age-25 non-college rural, 0.92% for age-35 college suburban) but not the Δ . For retirement, the segment-row differentiation is genuine (not merely a baseline shift) for the healthcare cell because each row uses its own subgroup-conditional matrix entry; the remaining eight categories on each retirement row still use the pooled $\hat{B}$ entry and therefore vary only through the pre-event baseline.

Table 9: Lifetime dollar counterfactual with bootstrap CI propagation: first-birth and retirement events, multiple household segments. Pre-event baseline shares come from a nine-equation linear share regression system with no adding-up restriction (Engel-curve share regressions; cluster-robust OLS at the household level on the pre-event sample, waves 2005–2019). The event-effect column $\hat{B}(0)[\text{event}, \cdot]$ is plugged in as a fixed causal input (BJS-isolation specification, Table 2). **For retirement segments, the healthcare cell uses the Medicare-subgroup-conditional matrix entry: the age-62 (pre-Medicare) segment uses the pre-65 BJS-isolation entry (+1.42 pp, SE 0.44); the age-65 and age-70 segments use the ≥ 65 entry (−0.90 pp, SE 0.41). The other eight categories continue to use the pooled retirement entry on every segment.** Post-event shares are reported after projection onto the simplex (negative entries clipped to zero, then re-scaled to preserve the pre-event row sum); raw pre-projection row sums are also reported. The lifetime \$ bracket reports a persistence-bracketed extrapolation: lower bound assumes the share shift decays linearly to zero over 4 years post-event (transitory case); upper bound assumes the $k=0$ shift holds for 40 years (durable case). Confidence intervals are from a joint nonparametric cluster bootstrap ($B=1,000$) that resamples households for the share-regression step and independently perturbs each matrix entry from $N(\hat{\beta}, \widehat{SE}^2)$. CI columns report bounds on the 40-year (upper-bound, durable-case) lifetime \$ shift; multiply by 0.05 to obtain the transitory-case CI. Representative lifetime expenditure is \$1,500,000 (40 years $\times$ \$37,500/yr median household consumption).

Category	Pre-event share (%)	Post-event share (%)	Δ share (pp)	95% CI on Δ (pp)	Lifetime \$ bracket [transitory, durable]	95% CI on \$
First birth						
<i>age 30, college, urban, mid-income (\$85k)</i>						
Food	17.10	17.35	+0.25	[−0.93, +1.36]	[\$+189, \$+3,789]	[\$−, \$+]
Housing	41.04	42.18	+1.14	[−0.57, +2.71]	[\$+854, \$+17,078]	[\$−, \$+]
Utilities	4.96	5.56	+0.59*	[+0.12, +1.09]	[\$+445, \$+8,897]	[\$−, \$+]
Transport	18.78	18.11	−0.68	[−2.57, +0.96]	[\$−507, \$−10,143]	[\$−, \$+]
Healthcare	4.20	5.67	+1.48*	[+0.74, +2.24]	[\$+1,108, \$+22,164]	[\$−, \$+]
Childcare	1.06	3.70	+2.63*	[+2.01, +3.25]	[\$+1,976, \$+39,521]	[\$−, \$+]
Education	1.20	0.00	−1.20*	[−1.60, −0.54]	[\$−901, \$−18,030]	[\$−, \$+]
Clothing	2.75	2.63	−0.12	[−0.60, +0.38]	[\$−90, \$−1,807]	[\$−, \$+]
Recreation	7.95	3.85	−4.10*	[−5.94, −2.07]	[\$−3,073, \$−61,469]	[\$−, \$+]
Row sum (post, projected)	99.05	99.05	—	—	—	—
Row sum (post, raw, pre-projection)	—	96.80	—	—	—	—
<i>age 25, non-college, rural, lower-income (\$45k)</i>						
Food	18.00	18.30	+0.30	[−0.81, +1.41]	[\$+224, \$+4,482]	[\$−, \$+]
Housing	36.11	37.19	+1.09	[−0.44, +2.75]	[\$+814, \$+16,279]	[\$−, \$+]
Utilities	4.83	5.43	+0.60*	[+0.14, +1.10]	[\$+449, \$+8,972]	[\$−, \$+]
Transport	20.20	19.58	−0.62	[−2.38, +1.11]	[\$−463, \$−9,255]	[\$−, \$+]
Healthcare	3.62	5.09	+1.47*	[+0.70, +2.26]	[\$+1,104, \$+22,087]	[\$−, \$+]
Childcare	1.25	3.90	+2.64*	[+2.04, +3.24]	[\$+1,983, \$+39,667]	[\$−, \$+]
Education	2.77	1.45	−1.31*	[−1.93, −0.69]	[\$−985, \$−19,699]	[\$−, \$+]
Clothing	3.92	3.83	−0.09	[−0.60, +0.42]	[\$−67, \$−1,341]	[\$−, \$+]
Recreation	8.50	4.42	−4.08*	[−6.34, −2.02]	[\$−3,060, \$−61,193]	[\$−, \$+]
Row sum (post, projected)	99.19	99.19	—	—	—	—
Row sum (post, raw, pre-projection)	—	96.94	—	—	—	—
<i>age 35, college, suburban, higher-income (\$120k)</i>						
Food	16.52	16.75	+0.23	[−0.87, +1.37]	[\$+174, \$+3,471]	[\$−, \$+]
Housing	41.12	42.24	+1.12	[−0.53, +2.78]	[\$+839, \$+16,776]	[\$−, \$+]
Utilities	4.57	5.16	+0.58*	[+0.08, +1.05]	[\$+437, \$+8,730]	[\$−, \$+]
Transport	18.52	17.83	−0.69	[−2.54, +1.10]	[\$−518, \$−10,367]	[\$−, \$+]
Healthcare	5.09	6.58	+1.49*	[+0.68, +2.31]	[\$+1,120, \$+22,403]	[\$−, \$+]

Bootstrap. For each of $B=1,000$ replicates, we (i) cluster- resample households with replacement and re-fit the nine OLS share equations on the bootstrap sample, and (ii) independently perturb each matrix entry from $N(\hat{\beta}_{E,c}(0), \widehat{SE}_{E,c}^2)$. We then re-compute the share counterfactual for every segment (including the simplex projection) and the bracketed lifetime dollar shifts. For segments with a Medicare-subgroup override (age-62, age-65, age-70 retirement rows on the healthcare cell), the bootstrap perturbs the subgroup-specific $(\hat{\beta}, \widehat{SE})$, not the pooled values. The CI columns in Table 9 report the 2.5%/97.5% percentiles of the replicate distribution on the durable-case lifetime dollar shift; multiply by 0.05 for the transitory-case CI. Mechanically, the cluster- bootstrap on the share equations contributes only to the level of the pre-event share baseline, not to the Δ column: the Δ CI is driven by the matrix-entry SE propagation. We report the joint structure for transparency, not because the cluster bootstrap tightens the Δ CI.

What this exercise does not deliver. It does not estimate a demand-system in the AIDS / log-ratio / multinomial-logit sense. It does not produce segment-conditional Δ shares beyond what the additive-shift spec permits, except for the explicit Medicare-subgroup propagation on the retirement-healthcare cell described above. It does not impose simplex / adding- up in the estimation; the projection step is post-estimation and is documented as such. Each of these items is a real piece of structure that a follow-up paper could add to push toward a demand-system architecture that respects the unit simplex *by construction*, which Section 9 argues is where the matrix object’s marketing-decision payoff ultimately lives. The present section’s contribution is narrower: it transparently propagates the matrix’s SE through a representative-budget multiplication, brackets the extrapolation by persistence regime, propagates the Medicare-subgroup contrast explicitly through the retirement-healthcare row, and discloses the simplex failure and projection step honestly. The matrix is the input; the dollar counterfactual is one output a category manager can read.

11 Limitations

Biennial-panel resolution. PSID is biennial, so the first post-event wave $k = 0$ averages event timings of zero to two years post-event, and persistence half-life estimates below one wave fall at or below the sampling resolution. Several marketing-decision uses are blocked by the two-year sampling gap: Medigap and Medicare Advantage open-enrollment-window targeting (operating on month-scale windows), post-natal advertising sequences in the first six months after birth, and divorce-triggered housing campaigns whose value depends on within-six-month detection. The persistence-classification framework’s high “unidentified” share is itself an artifact of this resolution rather than an economic finding.⁴

Out-of-pocket healthcare measurement and Medicare channel attribution. PSID’s healthcare-expenditure variable captures out-of-pocket (OOP) spending, including premium payments by household members (private-market premia pre-65, Medicare Part B and Part D premia, Medigap and Medicare Advantage OOP premia). Employer premia deducted directly from paychecks are not in the measure. The pre-65 rise in healthcare share is therefore in part a measurement shift from paycheck-deducted premia to OOP-visible premia as workers transition from employer coverage to individual-market or Consolidated Omnibus Budget Reconciliation Act (COBRA) continuation. The ≥ 65 decline is constrained by Part B’s inclusion in OOP, so the channel-attribution discussion in Section 7 is intrinsically limited by what the measurement variable captures. The +2.32 pp contrast at the Medicare cutoff is consistent with multiple distinct sub-channels operating on the two sides; the paper does not point-identify which channels drive the pre-65 rise versus the ≥ 65 decline. The denominator-clean log-total-expenditure check rules out total-expenditure-denominator artifacts but cannot rule out unobserved premium-side measurement dynamics. The labor-

⁴HRS replication, with annual resolution and direct health-insurance variables on the age-50+ panel, would address several of these gaps for the retirement and Medicare-cutoff results. The 2021 and 2023 PSID demographic waves cover demographics but not expenditures and are not used in estimation, so the estimation window ends in 2019. The placebo diagnostic uses $B = 20$ replications, appropriate for a directional check but low for a formal p -value benchmark; the bootstrap inferential machinery uses $B = 1,000$.

force-attachment partial- R^2 benchmark is uninformative as constructed because the three proxies have near-zero partial- R^2 on both treatment and outcome; PSID lacks a direct health-insurance variable. An HRS replication would close several of these gaps and is the natural next step; the Consumer Expenditure Survey is a rotating panel and is unsuitable for the within-household event-study design adopted here.⁵

Targeting value depends on downstream model architecture. Naive univariate scoring pipelines do not benefit from the matrix and may degrade in low-variance cells (Section 9). Section 10 is matrix-entry CI propagation under a representative-budget shell, not a fitted simplex-respecting demand system; the representative-budget multiplication uses a median \$1.5 million lifetime budget that abstracts from hump-shaped lifecycle income profiles and from segment heterogeneity in durables and retirement-savings categories that the nine-category partition does not separately resolve.⁶

Inference machinery and category partition, honestly characterized. Section 4.6 documents four residual departures from canonical implementations: the $B = 1,000$ boot-

⁵Sample-selection caveats on adjacent events: the head-stable subset drops approximately 45% of divorces, so $\hat{\beta}_{\text{divorce, housing}}(0) = +4.58$ pp is the head-stable-divorce ATT, not the all-divorce ATT; the Internet Appendix reports the all-divorce-no-head-stability number as a sensitivity. The dCDH multi-treatment estimator support fails on the retirement and job-loss columns (absorbing treatments); identification on these two columns rests on BJS imputation with CS-DR co-primary triangulation rather than independent multi-estimator triangulation across identifying assumptions, and the Sun-Abraham cohort-saturated diagnostic in the Internet Appendix provides a separate cross-check that agrees directionally with the BJS headline. The retirement/job-loss identification stack is thinner than the rest of the matrix; HRS framing is the natural pressure-release valve. The custom Python BJS implementation was validated against the CS-DR co-primary triangulation as the independent-implementation check, not against a packaged reference implementation.

⁶Mechanism and persistence scope: the Section 5 discussion names three reduced-form channels (mechanical accounting, budget reallocation, durable preference), and the empirics deliver a persistence-classification framework that is a coarsening of the three-channel decomposition rather than a structural identification of its components. Posit A is narrowed to the cells where the channel firing is unambiguous: the persistence form is partially contradicted on two cells (retirement $\times$ healthcare, job-loss $\times$ recreation), supported on two cells (move $\times$ housing on both level and persistence, first-birth $\times$ childcare at the level prediction only), and unidentified on the parametric form for the remaining four headline cells (additional-birth $\times$ childcare, marriage $\times$ housing with a non-monotone trajectory, divorce $\times$ housing, child-leaves $\times$ food) on the available horizon (2 contradicted + 1 level-only + 1 level-and-persistence + 4 unidentified = 8 headline cells). We document per-cell persistence trajectories on the eight headline cells, and the classification (transitory / durable / subgroup-dependent / unidentified) is reported in Internet Appendix Section ??; only one cell (move $\times$ housing) is cleanly durable on the four-wave biennial horizon, and the high unidentified share is a sampling-resolution artifact.

strap is sub-sampling at rate 0.632 rather than canonical with-replacement; the HonestDiD $\bar{M}_{\text{approx}}$ value is a first-order indicator, not the canonical Rambachan-Roth fixed-length confidence interval (FLCI) bound committed to in the analysis plan (the R `HonestDiD` package was unavailable), and at $M = 2$ both retirement $\times$ healthcare and marriage $\times$ housing lose significance in the companion CI-inflation block — readers should treat the $M = 2$ reversal as the binding credibility ceiling and the sensitivity floor for these two cells; the Romano-Wolf joint-bootstrap stepdown is the headline multiplicity correction with Holm alongside as a conservative comparison; and the Medicare-eligibility regression-discontinuity check at age 65 is basic OLS, not Calonico-Cattaneo-Titiunik robust RD (`rdrobust` unavailable). On the category side, PSID’s nine-category partition does not separately resolve alcohol, tobacco, gifts, charity, household services, vacations, and durables; recreation pools entertainment, equipment, trips, and fees. The Nielsen-Kilts brand-level scanner program is the complementary primitive for within-category and brand-level resolution: the matrix delivers across-category budget allocation, and brand-within-category allocation requires complementary scanner data.

12 Robustness

12.1 Income-tercile heterogeneity on first birth (open puzzle from Section 5)

A pre-registered prediction from Section 5 stated that the first-birth $\times$ childcare share rise would be larger in higher-income terciles, on the reasoning that high-income households purchase paid daycare at higher rates while low-income households rely more on family-provided care. The estimate is the opposite. Bottom-tercile $\hat{\beta}_{\text{first birth, childcare}}(0) = +2.78$ pp; top-tercile = $+2.37$ pp; the T3 minus T1 difference is -0.41 pp. The predicted direction is contradicted on the income dimension. The liquid-wealth tercile split delivers a similar pattern: T1 = $+3.30$ pp, T3 = $+2.41$ pp, T3 minus T1 = -0.89 pp. See Appendix A for

the rolling event-cohort tercile assignment, which makes the tercile labels comparable within an event but not across the cross-section. We report this as a documented contradiction of the a-priori prediction. Section 5 discusses the open-puzzle interpretation and the candidate refined channel without claiming it is identified by this paper’s design.

The job-loss $\times$ recreation tercile split is consistent with the pre-registered budget-channel direction: T1 (low income) = -1.31 pp, T3 (high income) = -0.10 pp, T3 minus T1 = $+1.22$ pp. Low-income households cut recreation share more on job loss, as the budget-channel mechanism predicts.

12.2 Isolation-window sensitivity (± 1 vs. ± 2 waves)

The headline matrix uses the ± 1 -wave isolation sample. The ± 2 -wave sample drops approximately 40% of the headline samples but tightens the isolation guarantee. The headline coefficients are stable in sign across the two windows; magnitudes shift by less than one bootstrap SE for the eight headline cells. Full per-cell sensitivities are in the Internet Appendix.

12.3 FHS-anticipation adjustment on first birth

The Freyaldenhoven-Hansen-Shapiro anticipation control on first birth is implemented as a labor-supply-stability subset that excludes households whose head’s labor-force participation changes in the pre-event window (a plausible signal of anticipation behavior). The subset has $n \approx 25$ events; the FHS-adjusted estimate is lower than the unadjusted estimate. Power is limited at this n ; the FHS leg is reported as a sensitivity, not as the headline.

12.4 Romano-Wolf joint-bootstrap vs. Holm

The Romano-Wolf joint-bootstrap stepdown is the headline multiplicity correction; the conservative Holm approximation is reported alongside. The marriage $\times$ housing cell flips classification: $p_{\text{Holm}} = 0.062$ (Holm-marginal) vs. $p_{\text{joint-RW}} = 0.014$ (RW-significant). The headline

retains marriage $\times$ housing in the eight-cell set on the RW classification; the conservative Holm classification is reported in Table 3 for comparison.

12.5 Twenty-two BH-significant non-headline cells

Beyond the eight pre-registered headline cells, a further 22 cells of the 72-cell matrix $B(0)$ are significant under the Benjamini-Hochberg false-discovery-rate procedure at $q = 0.10$, of which 18 survive the tighter $q = 0.05$ threshold. The cells appear in the Internet Appendix with one-line annotations. The list shows dense substitution structure beyond the headline cells and surfaces several additional marketing-relevant patterns, e.g., divorce $\times$ recreation (-4.26 pp), child-leaves $\times$ housing (-2.49 pp), and additional-birth $\times$ recreation (near zero, in contrast to first-birth $\times$ recreation -4.18 pp). The pre-registration of the eight headline cells limits multiplicity exposure on the headline claims; the additional cells are exploratory and labeled as such.

12.6 Head-stability sensitivity on divorce $\times$ housing

The headline $\hat{\beta}_{\text{divorce, housing}}(0) = +4.58$ pp is the head-stable-divorce ATT. The all-divorce-no-head-stability sensitivity is reported in the Internet Appendix; the coefficient attenuates but retains the predicted positive sign. The qualitative Posit B asymmetry $\beta_{\text{divorce, housing}} > 0 > \beta_{\text{marriage, housing}}$ survives the sensitivity.

12.7 Doubly-robust CS triangulation on the eight headline cells

The doubly-robust Callaway-Sant’Anna estimator is reported for the eight headline cells. The coefficients agree directionally with the BJS-isolation headline on seven of eight cells; the marriage $\times$ housing cell shows the largest divergence (CS-plain- 2×2 -0.30 pp, CS doubly-robust -1.32 pp; BJS isolation -2.06 pp), which the marginal-Holm classification already flags.

13 Conclusion

This paper delivers two outputs and one extension. The primary marketing-decision-relevant output is matrix-entry CI propagation under a representative budget (Section 10): for a representative household, a first birth implies a lifetime reallocation toward childcare in the range \$2,000–\$39,500 and away from recreation in the range \$3,000–\$61,500, bracketed between a transitory case (4-year linear decay) and a durable case (40-year flat extrapolation), with bootstrap CIs propagated from the underlying causal-identification step. The primitive that makes the counterfactual possible is a 9×8 causal substitution matrix B for eight household lifecycle events on PSID 1999–2019, estimated using BJS imputation with CS-DR co-primary triangulation: 72 signed, causally identified cross-category share-point shifts, with eight cells pre-specified in our analysis plan and 22 further cells significant under Benjamini-Hochberg at $q = 0.10$ (18 at $q = 0.05$). We recommend that downstream marketing-quant work cite the matrix when the application requires causally identified across-category budget moments for discrete lifecycle events, and cite the lifetime-budget counterfactual when the application requires segment-level dollar-magnitude reallocation estimates with CI propagation. The simplex-respecting structural demand-system architecture that the matrix is intended to feed (Section 9) is not estimated here.

The framework also surfaces an age-at-retirement contrast in out-of-pocket healthcare share at the Medicare cutoff that the pooled retirement column conceals (Section 7). We present this as the highest-priority direction for HRS replication, with HRS follow-up (annual resolution, direct insurance-status variables) the natural identification infrastructure for the channel attribution.

A Variable construction and additional diagnostics

A.1 Expenditure-share construction

For each household-wave observation, the nine PSID expenditure category dollars are summed to form the total-expenditure denominator. Each category share is the category dollar amount divided by the denominator. Observations with zero total expenditure are dropped. Shares are winsorized at the 99.5th percentile within each category to guard against PSID coding outliers; winsorization affects fewer than 0.5% of observations per category. The binding caps after winsorization are 0.95 for housing share and food share, 0.90 for healthcare share, 0.85 for transport share, 0.80 for utilities share, 0.60 for education and recreation shares, 0.55 for childcare share, and 0.50 for clothing share. The fraction of observations at the cap is below 0.5% for every category (housing: approximately 0.42%; healthcare: approximately 0.31%; childcare: approximately 0.18%). Internet Appendix Section ?? reports sensitivity at 99% and 95% caps; the eight headline cells shift by less than one bootstrap SE under all three winsorization choices.

A.2 Event coding

Event indicators are constructed as follows (see the online replication archive for full code).

`event_first_birth`. First household-wave at which the family-composition history shows a child age zero present, conditional on no prior child in the household.

`event_additional_birth`. Subsequent household-waves at which the family-composition history shows a new child age zero present.

`event_marriage`. Wave at which marital status transitions from never-married, divorced, or widowed to married, with the head-stability flag satisfied (age-of-head wave-to-wave change in $[1, 4]$ years).

event_divorce. Wave at which marital status transitions from married to divorced or separated, with the head-stability flag satisfied.

event_move. Wave at which `moved_since` is 1. Restricted to cohort $g \geq 2005$ (variable available from 2003).

event_job_loss. Wave at which `employ_status1` transitions from 1 (working) to 2, 3, 4, 5 (laid off, unemployed, retired, disabled) and not retirement. Restricted to cohort $g \geq 2005$.

event_retirement. Wave at which `employ_status1` transitions from 1 (working) to 5 (retired). Restricted to cohort $g \geq 2005$.

event_child_leaves. Wave at which the youngest dependent child in the household at the prior wave is no longer present.

A.3 Diagnostic tables

Pre-trend joint χ^2 tests, HonestDiD $\bar{M}_{\text{approx}}$ values, placebo summaries, and Goodman-Bacon-style cohort-size-weighted decompositions are tabulated in the diagnostic table below. The CI-inflation block at $M \in \{0.5, 1, 2\}$, the canonical Romano-Wolf joint-bootstrap stepdown table, and the per-cell bootstrap-SE-vs-cluster-robust-SE comparison are in the Internet Appendix.

A.4 Tercile heterogeneity

A.5 Income-tercile and liquid-wealth-tercile assignment

Terciles are assigned on a rolling per-event-cohort basis using the pre-event three-wave mean of `labor_income_head` (for income terciles) or `wealth_wo_equity` (1999–2007) and `wealth_w_equity` (post-2007; for liquid-wealth terciles). The rolling assignment guarantees

Table 10: Tercile heterogeneity (BJS isolation, $k = 0$): first birth $\times$ childcare, job loss $\times$ recreation, by pre-event income and liquid-wealth tercile (rolling per event).

Event $\times$ Category	Tercile dim.	T1 (low)	T2	T3 (high)	T3-T1
first_birth $\times$ childcare	income	+2.78	+2.49	+2.37	-0.41
first_birth $\times$ childcare	liquid_wealth	+3.30	+1.92	+2.41	-0.89
job_loss $\times$ recreation	income	-1.31	-2.05	-0.10	+1.22
job_loss $\times$ recreation	liquid_wealth	-1.42	-0.69	-1.35	+0.07

Notes: Terciles assigned on rolling event-cohort basis using pre-event 3-wave mean of income (`labor_income_head`) or liquid wealth (`wealth_wo_equity` for 1999-2007, `wealth_w_equity` after). BJS imputation run separately per tercile subsample. Difference = T3 - T1. Holm-Bonferroni applied within Spec 6 family per §7.5.

that tercile membership is determined by the household’s pre-event position relative to its event-cohort peers, not by the full-sample distribution.

References

- Mark Aguiar and Erik Hurst. Consumption versus expenditure. *Journal of Political Economy*, 113(5):919–948, 2005. doi: 10.1086/491590.
- Mark Aguiar and Erik Hurst. Deconstructing life cycle expenditure. *Journal of Political Economy*, 121(3):437–492, 2013. doi: 10.1086/670740.
- James Banks, Richard Blundell, and Sarah Tanner. Is there a retirement-savings puzzle? *American Economic Review*, 88(4):769–788, 1998.
- B. Douglas Bernheim, Jonathan Skinner, and Steven Weinberg. What accounts for the variation in retirement wealth among u.s. households? *American Economic Review*, 91(4):832–857, 2001. doi: 10.1257/aer.91.4.832.
- Indrayudh Biswas, Pradeep K. Chintagunta, and Sanjay K. Dhar. How do U.S. households change their expenditure patterns in response to income or wealth shocks? Insights from NielsenIQ data. *Quantitative Marketing and Economics*, 23:419–445, 2025. doi: 10.1007/s11129-025-09297-4.

- Kirill Borusyak, Xavier Jaravel, and Jann Spiess. Revisiting event study designs: Robust and efficient estimation. *Review of Economic Studies*, 2024. doi: 10.1093/restud/rdae007.
- Bart J. Bronnenberg, Jean-Pierre H. Dubé, and Matthew Gentzkow. The evolution of brand preferences: Evidence from consumer migration. *American Economic Review*, 102(6): 2472–2508, 2012. doi: 10.1257/aer.102.6.2472.
- Brantly Callaway and Pedro H. C. Sant’Anna. Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230, 2021. doi: 10.1016/j.jeconom.2020.12.001.
- Sebastian Calonico, Matias D. Cattaneo, and Rocio Titiunik. Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6):2295–2326, 2014. doi: 10.3982/ECTA11757.
- Carlos Cinelli and Chad Hazlett. Making sense of sensitivity: Extending omitted variable bias. *Journal of the Royal Statistical Society: Series B*, 82(1):39–67, 2020. doi: 10.1111/rssb.12348.
- Clément de Chaisemartin and Xavier D’Haultfoeuille. Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–2996, 2020. doi: 10.1257/aer.20181169.
- Carlos Dobkin, Amy Finkelstein, Raymond Kluender, and Matthew J. Notowidigdo. The economic consequences of hospital admissions. *American Economic Review*, 108(2):308–352, 2018. doi: 10.1257/aer.20161038.
- Rex Y. Du and Wagner A. Kamakura. Where did all that money go? understanding how consumers allocate their consumption budget. *Journal of Marketing*, 72(6):109–131, 2008. doi: 10.1509/jmkg.72.6.109.

- Simon Freyaldenhoven, Christian Hansen, and Jesse M. Shapiro. Pre-event trends in the panel event-study design. *American Economic Review*, 109(9):3307–3338, 2019. doi: 10.1257/aer.20180609.
- Peter Ganong and Pascal Noel. Consumer spending during unemployment: Positive and normative implications. *American Economic Review*, 109(7):2383–2424, 2019. doi: 10.1257/aer.20170537.
- Andrew Goodman-Bacon. Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2):254–277, 2021. doi: 10.1016/j.jeconom.2021.03.014.
- Fatih Guvenen, Fatih Karahan, Serdar Ozkan, and Jae Song. What do data on millions of u.s. workers reveal about lifecycle earnings dynamics? *Econometrica*, 89(5):2303–2339, 2021. doi: 10.3982/ECTA14603.
- Michael D. Hurd and Susann Rohwedder. The retirement consumption puzzle: Actual spending change in panel data. *NBER Working Paper*, (13929), 2008. doi: 10.3386/w13929.
- Wagner A. Kamakura and Rex Y. Du. How economic contractions and expansions affect expenditure patterns. *Journal of Consumer Research*, 39(2):229–247, 2012. doi: 10.1086/662611.
- Eric Oblander and Daniel M. McCarthy. Frontiers: Estimating the long-term impact of major events on consumption patterns: Evidence from COVID-19. *Marketing Science*, 42(5):839–852, 2023. doi: 10.1287/mksc.2023.1443.
- Arna Olafsson and Michaela Pagel. Retirement puzzles: New evidence from personal finances. *Journal of Public Economics*, 234:105103, 2024. doi: 10.1016/j.jpubeco.2024.105103.

- Ashesh Rambachan and Jonathan Roth. A more credible approach to parallel trends. *Review of Economic Studies*, 90(5):2555–2591, 2023. doi: 10.1093/restud/rdad018.
- Jonathan Roth, Pedro H. C. Sant’Anna, Alyssa Bilinski, and John Poe. What’s trending in difference-in-differences? a synthesis of the recent econometrics literature. *Journal of Econometrics*, 235(2):2218–2244, 2023. doi: 10.1016/j.jeconom.2023.03.008.
- Pedro H. C. Sant’Anna and Jun Zhao. Doubly robust difference-in-differences estimators. *Journal of Econometrics*, 219(1):101–122, 2020. doi: 10.1016/j.jeconom.2020.06.003.
- Melvin Stephens Jr. The long-run consumption effects of earnings shocks. *Review of Economics and Statistics*, 83(1):28–36, 2001. doi: 10.1162/003465301750160018.
- Liyang Sun and Sarah Abraham. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2):175–199, 2021. doi: 10.1016/j.jeconom.2020.09.006.