

Effects of Parental Death on Labor Market Outcomes — source-faithful native replication

Bolero

Abstract

Registered research design and conditional interpretation policy; no outcomes observed.

Keywords: Pre-registration, research design

Effects of Parental Death on Labor Market Outcomes — source-faithful native replication

Replication pre-registration: the source paper's results are known; no new replication outcomes have been observed.

Research question

What does a source-faithful replication of Jensen and Zhang (2026) establish about parental death, adult labor outcomes, health and family support, within the source's actual identified scope?

Scope

Fixed journal question, original cohort/treatment and reported main/supplement analyses. Canonical two-stage replica conversion then Transcriber. Unresolved source choices are not implemented.

Graph overview

Dependency edges summarize the graph, not unconditional activation. Exact logical conditions and all local outcomes follow below; the separate graph atlas shows each Atom's branches.



Fig. 1 Registered Atom dependencies

Registered storyline

Population and descriptive evidence delimit context. Matched labor effects address the primary question; health, joint timing, childcare and surviving-parent support qualify mechanisms. Inheritance, formation, spouses and work choice supply distinct supporting/alternative-channel evidence. All components are interpreted through their actual estimates and uncertainty, including mixed findings and missing targets.

Source literature and access

Jensen, Mathias Fjællegaard, and Ning Zhang.2026. “Effects of Parental Death on Labor Market Outcomes.” American Economic Review116(5):1811–1836. DOI10.1257/aer.20240432. [Publisher article](#). Supplied26-page original journal PDF was read in full; its exact hash is in replication/source-provenance.yaml. The local article Markdown was read as a navigation aid and hash-checked; the original PDF governs scientific reading.

Jensen and Zhang, journal-linked Supplemental Appendix,31 pages, AEA material25093, [publisher material link](#), retrieved2026-09-05; SHA2567cebe88f0cdef5b8e8865d457f2abea64aad2691dbc8e9cf94f3faf5f1c51105. Full text and all source display pages were inspected. Contains Appendix A.A–G, FiguresB1–B23, TablesB1–B6. No original supplement is redistributed in this World.

Jensen and Zhang.2026. Data and Code for the article. ICPSR versionv1, DOI10.3886/E239029V1, [DataCite metadata](#). Publisher linking and metadata establish repository identity; archive access returned403, so remote byte correspondence and missing input files remain unverified. All36 available local author .do files were read, including sibling cleaning/analysis and FN programs; see replication/code-inspection-index.csv. Supplied analysis5/matching15 hashes match provenance. Programs are inspection-only and not redistributed.

The paper’s references provide its prior-work framing (bereavement, health shocks, child penalties and informal family care). This conversion makes no new novelty or comparative-literature claim and does not import scientific defaults from those works. In particular, the source FN robustness is extracted from this article/appendix and author programs, not redesigned from Fadlon–Nielsen papers. Earlier working-paper versions surfaced in live searches but are not the scientific specification for this journal conversion. Historical replication manuals/decks and earlier replica drafts are not sources.

Live retrieval queries, UTC dates, HTTP outcomes and access limitations are in replication/retrieval-log.md. OpenAlex exact-DOI lookup429; PMC exact-DOI search0 records; these are access/index outcomes, not evidence that a source or science is absent. The relevant publisher supplement was retrieved successfully. FSV metadata and official variable pages establish documentation only; see schema-evidence.md.

Shared data and result contract

This document is incorporated by reference into every Atom protocol. All 531 tables in the unchanged supplied boundary are allowed for planning; none is validated as mounted. Only documented fields and author source code were read. No row, micro-data summary, model estimate or empirical program was accessed/executed. The final package is a canonical pre-data replica registration, not a feasibility certificate. Each consuming Atom owns its linkage/coverage checks and their consequences. No separate diagnostic experiment is created.

D1. People, time, family and treatment

Unit of the main study is adult-child person-year; matched models duplicate a reused control separately for every source pair and cluster at $\text{person} \times \text{match}$, as specified. Event year D is calendar year of first parental death, $t = \text{year} - D$. Balanced 11-year means exactly one valid person-year at each $t = -5$ through 5, not merely 11 possibly duplicated rows. Replacing a count assertion by explicit key/year-set equality is meaning-preserving engineering; a mismatch stops that target, it does not license deduplication. Population source: FAIN 1981–85 references are assigned year-1 for 1980–84; BEF 1985 onward uses year-end snapshots, including December quarters in 2008+. Use source PNR, MOR_ID, FAR_ID, FOED_DAG, KOEN, ALDER, E_FAELE_ID, CIVST, KOM, REG. Historical date imputation in clean_6 uses Jan1(year-age) before precise birthdays; later available birthdays/gender/parent IDs are propagated as code. These are source definitions, not a permission to replace uncertain identities. Parental links must exist for both parents; parent birthdays and population presence in $t = -1$ are required. Source death records are DOD-SAARS2001 cumulative extract (PNR, C_LISTE_49, D_DODSDTO, C_SEX) and DODSAASG2019 cumulative extract (PNR, C_LISTE49, D_STATDATO, C_SEX). The catalog's current cumulative extract vintage differs; current data must not be presumed equal.

First death is the earlier linked mother/father death date. Source chooses mother if dates tie; keep A11 because the journal scientific treatment in simultaneous death is unspecified. Final sudden cause49-list codes are 24, 25, 26, 29, 46, 47, 49. The intermediate clean_8 list also includes 27/28, but final sample removes them in clean_15 lines 558–560 and descriptive consuming programs: this staged restriction is not a conflict. Final treated age 25–50 at event, first death 1985–2014; code initially allows 25–55 before final filtering. Original labor panel spans 1980–2019. Do not extend it to newer documented years. Second parental death within the main treated window is allowed in the main design; the source >5-year-gap robustness remains U03, not a universal exclusion. Control candidate first parental death is after D+5 or unknown, both parents remain in the population through the source pseudo window, and birth/age rules in matching-specification.md apply. Unknown future death is not invented survival beyond observed source follow-up.

D2. Education, geography and family fields

UDDA HFAUDD/HF_VFRA: choose latest education in year; source carries education forward, excludes HFAUDD1, maps DISCED level0/1/2 to below high school,3/5 to high-school/vocational,6/7/8 to college+. Years of schooling use author crosswalk months/12. Missing crosswalks are U02, not permission to infer categories from code digits. Cohabitation is nonempty source spouse/partner ID, not a newly restricted marriage variable. Married parent CIVST G or P is1, remaining/missing states0 in source; match-stage missing category2 applies where the variable is still missing. Child-parent records from MOR_ID/FAR_ID construct number of linked children born by each calendar year; source permits same-sex parental links. Youngest age is based on latest birth; integer age over20 is capped21 and no-child sentinel22 in matching. Source youngest-child tie identity remains A11/A06 where enrollment depends on it. Main analytical child count, descriptive cap5 and spouse cap5 are distinct consuming rules.

Geography uses source5-region REG, municipality KOM, and author crosswalk for11-region NUTS3. clean_6 maps old FAIN municipalities through BEF1986, explicitly patches municipalities793/719 to81 and627 to82; modern KOM/REG were revised historically according to FSV documentation. The exact vintage and crosswalk are required. Detailed municipality/11-region data begin1985. No distance in kilometers, updated boundary or geographic imputation is registered.

D3. Earnings, employment, jobs and wealth

Annual gross labor earnings including employment/self-employment are IND ERHVERVSINDK_GL1980-86, ERHVERVSINDK_13 from1987; clean_11 source converts to real2020DKK using the absent inflation factors and divides by1000 for raw earnings graphs. Source rank calculation among population ages20-55 with known birth/sex ranks ERHVERVSINDK_13 within calendar year with mean tied ranks, then $100 * (\text{rank} - 1) / (N - 1)$; no sex or birth-cohort rank denominator. Different annual extract definitions require provenance, not an automatic mapping. Source missing income joins are zero in the balanced sample only when absence is interpretable under a verified complete source extract. Genuine unspecified duplicate records stop rather than select a new row.

ATP employment=positive ATPXX (ATP_AAR alias in source vintage). Missing/negative ATP is0 as source; years1980/81 multiply1167/432; intensive margin capped at each year's99th percentile. Unconditional ATP includes zeros; conditional ATP retains positive observations. These are pension-contribution proxies, not observed work hours. Earnings p1/p99 winsorization applies in each specified analysis after its stated restrictions; index100 uses $t = -1$ mean over T/C by sex, except source group-specific denominators in childcare/geography. No universal preprocessing sample can replace the explicit module rules. The FN p5/p95 inconsistency is C10.

IDAN person-job records link PNR/year/LBNR to IDAS early ownership; owner07/08/09 is public;2007 owner99 carries prior08 as source. Early main job rule uses author IDANTYPE, dropsE/B, selects type3 if multiple and prioritizesH. Later

TYPE_2008 selects A/E/H/M/N/S; prioritizes H then E and wage where programmed. Author abbreviation TYPE may resolve TYPE_2008 under Stata varabbrev; this is not asserted to be a nonexistent-variable bug. Later sector classification excludes source8 prefixes and 32/42/52/62/99. The exact source sector tests must remain with their version (schema evidence). LBNR all-zero is missing; firm changes require adjacent valid different LBNRs. Public indicators after missing recode are conditioned on ATP employment where the source table requests it. AKM BESKST/BESKST13 identifies self-employment 01/02, and employee/self-employed DISCO codes are chosen as source; there is no new code harmonization at 2010.

Wealth: IND FORM 1984–97, FORMREST_NY05 from 1998, source maximum asset record within PNR/year, real 2020 DKK. Net wealth is measured excluding pension wealth per documentation, not an observed bequest. Table B4's 1000 DKK versus code 10000 DKK is C08. Missing author inflation factors blocks real amounts and source matching where real income is consumed; no constant source means or published values fill the gap.

D4. Health event aggregation

Hospital LPR_ADM source named LPRADM in paths: read 1980–2018, PNR, C_ADIAG, D_INDDTO, D_UDDTO. Source removes invalid/missing dates, deduplicates PNR/admission keeping latest discharge; unresolved tied records that alter diagnosis retain A11. Annual visits count admissions by admission year. Screening is diagnosis containing DZ00/DZ01 since 1994; final nonscreening is total-minus-screening in analysis_6. Days measure expands inclusive admission-to-discharge days, unions PNR/day to avoid overlap, counts by calendar year; outpatient visits are included. Do not reinterpret as inpatient bed-days. PSYK_ADM 2019 cumulative source counts 1995–2018 psychiatric admissions after same-date deduplication. No LPR3 substitution after 2018 is authorized.

SYSI 1990–2005 plus SSSY 2006–2020: aggregate YDLANT by PNR/year and six-digit SPECIALE. GP codes are 800100, 800101, 800102, 800105, 800124, 800125, 800126, 800200, 800201, 800202, 800203, (22 codes as written in clean_4). Private psychology/psychiatry uses prefix 63/24 followed by the missing author service mapping; prefixes alone do not replace that mapping. Negative count values set 0 and source pooled 99th-percentile cap where clean_3/4 applies. LMDB 1995–2021 source: row counts by PNR/year, ATC prefix mental N05/N06A/N06B/N06C, opioid N02A, alcohol N07BB. Derive any-use only for targets requesting it; B9 alcohol uses count. The current LMDB PNR12 requires an authenticated link map to source PNR. Dispensing/service use is not a clinical diagnosis or proof of ingestion.

D5. Childcare

DAGI 1995–2014 PNR/YDELSE: source young codes 151/152/154/200/201; older 153/224/225/226/227. BOERNFB 2015–2020 marks young enrollment by child-PNR/year; BOERNSB 2015–2020 older enrollment. Deduplicate only identical person-year enrollment events for any-use. Missing rows are source 0 within

complete available enrollment tables; absent years cannot be 0. Formal-care analyses restrict 2005–2019 as source and require full 11-year panels, not a new uptake study spanning all documented years. Current table coverage through 2021/2024 does not change original follow-up. Author variable abbreviation `formal_child_care` can resolve `formal_child_care_old5`; do not invent a bug from abbreviation alone. Youngest-child identity, age boundaries and actual lookup coverage remain gates.

D6. Source estimator and uncertainty translation

Eq1: sex-specific OLS, sum event indicators $k \neq -1$, $\text{age} \times \text{year}$ and $\text{education} \times \text{year}$ categorical controls; robust SE. Eq2: $Y = \sum [k = -5..5, k \neq -1] k T1(t=k) + \text{eventFE} + \text{person} \times \text{matchFE} + \text{yearFE} + \text{ageFE} + \text{error}$; source event range changes only as specified in Atom5. Eq3: $Y = T_{\text{post}} + \text{same FE}$. Eq4 replaces *Tpost* by *mother-firstTpost* and *father-firstTpost* jointly; same FE, split by adult sex. Post definition A03 and normalization A02 are retained explicitly. Cluster at $\text{person} \times \text{match}$ except spouse uses $\text{spouse} \times \text{match}$ and disputed FN source cluster C10. Source code's unweighted OLS is preserved; reused controls remain separate pair records. No bootstrap, randomization inference, IPW, alternative DiD, extra controls or multiplicity procedure is added.

Event plots explicitly compute $\text{coefficient} \pm 1.96 \times \text{SE}$ in author code; source pooled coefficient plots request 95% using `coefplot/reghdfe e(df_r)` semantics, so export author-corresponding CI endpoints and `df/method` rather than hard-code 1.96 for every interval. Tables with SE only retain SE; do not add new CIs or tests. Source printed F/p differences use the consuming joint-model covariance, never a synthetic difference from independent estimates. If `software/version/df` is unavailable, affected interval/contrast is unavailable even if a coefficient is known. Dynamic reference $t = -1$ is labelled omitted; no estimated zero or zero-width interval. Baselines, observation counts and reference units are target metadata. No p-value threshold, replication-success band, power calculation or minimum detectable effect is introduced.

D7. Pseudocode execution and failure boundaries

```
for target in the frozen inventory order:
  identify its source module, rows/panels, model and documented input years
  inspect only authorized runtime schema/metadata before requesting
    ↪ confidential computation
  if a scientific definition is unresolved: write no-estimate with exact
    ↪ gap ID; do not fit
  if required vintage, key, crosswalk or register-year is absent: write
    ↪ no-estimate with reason
  otherwise form precisely the source person/year/pair universe and
    ↪ transformations
  assert prescribed join cardinality, year set and source-identifiable
    ↪ treatment/covariates
  if the source model is unidentified or prescribed cells empty: retain
    ↪ unavailable target
  if software crashes, permissions fail or output is corrupt: record
    ↪ operational failure;
    no scientific verdict and no done() transition are implied
  else execute only the extracted, fully determined source specification
```

```

export every expected result and its actual uncertainty, including mixed
↪ findings
fill no missing result with a published number, fitted alternative, zero
↪ or interpolated point

```

Assertions verify implementation of the source unit and existing estimand. They are not empirical balance thresholds or new diagnostic studies. Runtime may not solve scientific ambiguities, regroup Atoms or choose a source branch. Resolution needs additional source-supported detail or separately authorized user extension, with a traceable amendment before dependent analysis. Such an extension is not currently authorized.

D8. Actual-result interface and local states

No actual-result file exists in this pre-data World. At runtime each Atom produces **results.csv**, **target-status.csv**, **interpretation.md** and the plot-ready outputs named in **figures.yaml**. CSV result key is **target_id,panel,series,x,statistic,term**; retain **outcome,estimand,sample_id,model_id,estimate,se,lower,upper,ci_method,df,unit,n,denominator,source_locator,statistic**. Numeric values null for unavailable rows. Source tables may have no interval or SE by design; use explicit **uncertainty=not_in_source**, not 0. Source-reported **F,p,baseline,N,pooledDID** and derived shares are separately keyed rows. **target-status.csv** records expected/available cells and exact non-estimate reason, never a success percentage used as evidence. **interpretation.md** cites actual row keys and distinguishes signed magnitude, uncertainty, subgroup contrasts and limitations. It must never infer a mediator from service utilization alone.

Local terminal states **evidence**, **partial**, **boundary** describe result coverage only: respectively all retained requested targets available; some available and some scientifically unavailable; no target estimable after a conclusive source/input boundary. The complete actual vector and source-qualified interpretation remain authoritative inside that state. **otherwise** is reserved for unresolved/inconclusive scope when a scientific disposition is possible; an operational failure is recorded separately and leaves the Atom pending, not done. Statistical imprecision is an actual estimate with its interval, not an unavailable target. No state denotes confirmed/refuted replication; mixed results stay mixed. All birth atom **_result** values remain prereg, with null strength/shape, empty evidence paths and no registered **_state**.

Figure input schemas are fixed output tables, not a new activation/control schema. Native activation is **TRUE** for every Atom; shared input construction entails no dependence on another Atom's scientific result. All target-specific gates are inside the consuming protocol. Molecule interpretation combines actual evidence vectors and restrictions using canonical grammar, never chooses execution from significance.

Identification and scope

The source question is fixed: the labor-market consequences of sudden first parental death and the health/family-support evidence discussed by Jensen and Zhang²⁰²⁶. The main design compares adults 25–50 whose first parent dies suddenly 1985–2014

with source-matched controls whose parents do not die during the comparison window. Matching, units, age/parent filters, fixed effects, clustering and actual event-time definitions are in data-contract.md and the Atom registrations. C01/A01/A02 prevent silently deciding disputed matching or normalization. No alternative design is added.

The source causal interpretation requires counterfactual labor trajectories of treated and matched controls to evolve similarly after the event, with sufficient overlap and no relevant anticipation of sudden loss. Predetermined matching variables are part of the source algorithm. Event-study pre-period estimates are presented as in the paper; there is no additional pretrend test or numerical acceptance threshold. Avoiding already-treated controls alone does not establish all DiD identification assumptions. Repeated controls retain the source person $\times$ match cluster; do not silently choose person-level clustering even where that seems attractive.

Source descriptive Eq1 estimates use no matched controls and remain descriptive. Population age-at-death distributions are conditional on the source event sample. Source service use and medication measures are proxies, not diagnoses or direct care hours. Childcare/proximity heterogeneity and post-event surviving-parent illness can reflect selection; they do not identify controlled direct effects or causal mediation. Joint onsets and source ratios are associational evidence. Parental wealth is not observed inheritance, and no null or small interaction proves absence of inheritance effects.

All source outcome modules and reported robustness are retained, but an unspecified model, matching vector, chronology, crosswalk or subgroup remains no-estimate. No new control, estimator, sensitivity battery, equivalence band, significance rule or power claim fills a source gap. The data boundary authorizes schema-only planning across 531 tables; it does not establish runtime access, overlap, follow-up or identifying variation. Routine key/window/model-identification checks live inside consuming Atoms and add no new experimental unit.

The registered target_strength for matched effects denotes the source intended causal target conditional on its assumptions. It is not a realized empirical strength. Molecule interpretation must read actual estimates and uncertainty and restrict conclusions to available, source-determined outputs. No outcome is currently observed.

Governance, ownership and handoff

Mode replica, preregistration mode canonical. Exactly two stages: source-faithful converter followed by Transcriber. No candidate hand, ideation, novelty contest or author-reviewer debate occurred. This stage assigns final Atom IDs and supplies native protocols, complete Molecule membership, source audit, manuscript and implemented figure sources. The self-check is not independent scientific acceptance.

Only schema/documentation, supplied public paper/code and publicly retrieved source materials were read. Original empirical programs were inspected but never executed. All 531 allowed FSV tables remain the planning boundary; validated_tables is empty. No confidential rows, computed outcome summaries or observed replication estimates were accessed. Published values are reference targets only, confined to explicitly labelled source comparison notes; current Atom results and manuscript replication slots remain unfilled.

Atom registration owns its protocol, activation and future local result states; Molecule membership owns ALL11 members and its registration owns story-line/interpretation. The DOT diagram is a derived input-flow aid. No new activation schema is introduced: TRUE is used for all Atoms. Input reuse is not evidence-dependent activation. Structural legacy ratings are neutral required compatibility fields, not scientific ratings from a tournament or empirical success scores. Molecule ratingC has the same administrative role at birth.

Runtime may implement only source-explicit or meaning-preserving translated specifications. It may not repair a paper/code conflict, fill a scientific default, add tests or regroup final modules. Any additional source material must be recorded with identity/date/locators and a traceable amendment before dependent analysis; separate scientific extensions require separate user authorization. An operational failure remains operational and pending, never a statistical null or automatic done(). Source/input boundaries and partial coverage must preserve every affected target.

The converter's synthetic interface tests run outside the World and contain no current results. Per-figure sources accept plot-ready reviewed output tables, never raw registers. Transcriber/host own full registration and paper LaTeX/PDF builds, copied class/style assets and ZIP packaging. Converter does not author those generated outputs or host-owned plot_style.py. Authored scientific sources must be preserved across transcription/build, including all unresolved decisions. No source PDF, screenshot, fulltext cache, original .do file, commit, push, release or external-system write belongs in this deliverable.

Registered interpretation rules

Restrictions take precedence over permissions; residual reporting does not invent a finding.

1. restriction

Condition: TRUE

Apply source-definition/vintage limits to every component. Even fully available results require outcome-specific signed estimates, source uncertainty, and the stated identification assumptions. No coverage state or paper comparison implies replication success.

```
kind: restriction
condition: 'TRUE'
says: Apply source-definition/vintage limits to every component. Even fully
      available results require outcome-specific signed estimates, source
      ↪ uncertainty,
      and the stated identification assumptions. No coverage state or paper
      ↪ comparison
      implies replication success.
```

2. primary

Condition: done(atom-000003)

Read labor results.csv and interpretation.md: report employment, conditional/unconditional ATP, raw/indexed earnings and parent/child-gender differences separately, together with source robustness, balance scope and missing targets. Base direction/magnitude claims on those actual estimates/intervals; mixed or imprecise findings stay mixed/imprecise.

```
kind: primary
condition: done(atom-000003)
says: 'Read labor results.csv and interpretation.md: report employment,
↪ conditional/unconditional
ATP, raw/indexed earnings and parent/child-gender differences separately,
together with source robustness, balance scope and missing targets. Base
direction/magnitude claims on those actual estimates/intervals; mixed or
imprecise findings stay mixed/imprecise.'
```

3. supplementary

Condition: done(atom-000004) OR done(atom-000005)

Use actual health-service, prescription and co-occurrence vectors to assess consistency with the source health explanation. Distinguish service use, onset timing, ratios and unavailable estimates. No causal mediation share follows from joint onset or completion.

```
kind: supplementary
condition: done(atom-000004) OR done(atom-000005)
says: Use actual health-service, prescription and co-occurrence vectors to
assess consistency with the source health explanation. Distinguish service
use, onset timing, ratios and unavailable estimates. No causal mediation
share follows from joint onset or completion.
```

4. supplementary

Condition: done(atom-000006) OR done(atom-000007)

Interpret childcare earnings contrasts, enrollment, surviving-parent illness/proximity and their source model boundaries jointly with actual labor evidence. Separate subgroup selection and post-event illness conditioning from causal claims about care.

```
kind: supplementary
condition: done(atom-000006) OR done(atom-000007)
says: Interpret childcare earnings contrasts, enrollment, surviving-parent
illness/proximity and their source model boundaries jointly with actual
labor evidence. Separate subgroup selection and post-event illness
↪ conditioning
from causal claims about care.
```

5. supplementary

Condition: done(atom-000008) OR done(atom-000009) OR done(atom-000010) OR done(atom-000011)

Report each actual inheritance, family-formation, spousal and job-choice result with its uncertainty and missing scope. Small/imprecise findings do not establish absence of a channel; no equivalence threshold or combined score is authorized.

```
kind: supplementary
condition: done(atom-000008) OR done(atom-000009) OR done(atom-000010) OR
done(atom-000011)
says: Report each actual inheritance, family-formation, spousal and job-choice
result with its uncertainty and missing scope. Small/imprecise findings
do not establish absence of a channel; no equivalence threshold or combined
score is authorized.
```

6. supplementary

Condition: done(atom-000001) OR done(atom-000002)

Use the actual population characteristics and descriptive sudden/non-sudden paths to delimit the source population and contextual dynamics. Do not turn conditional age distributions into population hazards or descriptive paths into matched causal evidence.

```
kind: supplementary
condition: done(atom-000001) OR done(atom-000002)
says: Use the actual population characteristics and descriptive
↪ sudden/non-sudden
paths to delimit the source population and contextual dynamics. Do not turn
conditional age distributions into population hazards or descriptive paths
into matched causal evidence.
```

7. residual

Condition: TRUE

Name every pending, partial, boundary, inconclusive or operationally failed component. No empirical replication has occurred at registration. If a component has no estimate, make no directional claim for its outcomes. Preserve unresolved source choices; compilation/conversion completion is not independent scientific acceptance.

```
kind: residual
condition: 'TRUE'
says: Name every pending, partial, boundary, inconclusive or operationally
failed component. No empirical replication has occurred at registration.
If a component has no estimate, make no directional claim for its outcomes.
Preserve unresolved source choices; compilation/conversion completion is
not independent scientific acceptance.
```

Preregistered epistemic graph

Each Atom is an interpretable question or diagnostic judgment. Local result states determine registered activation.

atom-000001

What population and pre-death characteristics underlie the source comparison of sudden and non-sudden first parental death?

Activation: TRUE

- **evidence:** All retained source targets have actual estimates and their source-required uncertainty/interpretation, with no unresolved target. — Report the complete actual outcome vector. Coverage alone gives no direction, causal conclusion or replication confirmation.
- **partial:** At least one source target has actual evidence and at least one has a conclusive source/input no-estimate disposition. — Report available signed estimates and uncertainty, including mixed findings, beside every missing target and its exact limitation. Do not generalize the available subset to the missing scope.
- **boundary:** No retained source target is estimable after a conclusive source-definition or input boundary; this is not an operational failure. — There is no scientific estimate for this source scope. The paper is neither confirmed nor refuted by inability to estimate.
- **otherwise:** otherwise — Retain unresolved or inconclusive scientific disposition with actual evidence if any. Operational failures remain pending and are not done() or a statistical null.

atom-000002

How do labor supply, earnings and mental-prescription use evolve around sudden and non-sudden first parental deaths in the source descriptive design?

Activation: TRUE

- **evidence:** All retained source targets have actual estimates and their source-required uncertainty/interpretation, with no unresolved target. — Report the complete actual outcome vector. Coverage alone gives no direction, causal conclusion or replication confirmation.
- **partial:** At least one source target has actual evidence and at least one has a conclusive source/input no-estimate disposition. — Report available signed estimates and uncertainty, including mixed findings, beside every missing target and its exact limitation. Do not generalize the available subset to the missing scope.
- **boundary:** No retained source target is estimable after a conclusive source-definition or input boundary; this is not an operational failure. — There is no scientific estimate for this source scope. The paper is neither confirmed nor refuted by inability to estimate.
- **otherwise:** otherwise — Retain unresolved or inconclusive scientific disposition with actual evidence if any. Operational failures remain pending and are not done() or a statistical null.

atom-000003

What are the source matched-DiD effects of sudden first parental death on labor outcomes, including its reported heterogeneity and robustness?

Activation: TRUE

- **evidence:** All retained source targets have actual estimates and their source-required uncertainty/interpretation, with no unresolved target. — Report the complete actual outcome vector. Coverage alone gives no direction, causal conclusion or replication confirmation.
- **partial:** At least one source target has actual evidence and at least one has a conclusive source/input no-estimate disposition. — Report available signed estimates and uncertainty, including mixed findings, beside every missing target and its exact limitation. Do not generalize the available subset to the missing scope.
- **boundary:** No retained source target is estimable after a conclusive source-definition or input boundary; this is not an operational failure. — There is no scientific estimate for this source scope. The paper is neither confirmed nor refuted by inability to estimate.
- **otherwise:** otherwise — Retain unresolved or inconclusive scientific disposition with actual evidence if any. Operational failures remain pending and are not done() or a statistical null.

atom-000004

How does parental loss affect the source health-service and prescription outcomes across parent and adult-child gender?

Activation: TRUE

- **evidence:** All retained source targets have actual estimates and their source-required uncertainty/interpretation, with no unresolved target. — Report the complete actual outcome vector. Coverage alone gives no direction, causal conclusion or replication confirmation.
- **partial:** At least one source target has actual evidence and at least one has a conclusive source/input no-estimate disposition. — Report available signed estimates and uncertainty, including mixed findings, beside every missing target and its exact limitation. Do not generalize the available subset to the missing scope.
- **boundary:** No retained source target is estimable after a conclusive source-definition or input boundary; this is not an operational failure. — There is no scientific estimate for this source scope. The paper is neither confirmed nor refuted by inability to estimate.
- **otherwise:** otherwise — Retain unresolved or inconclusive scientific disposition with actual evidence if any. Operational failures remain pending and are not done() or a statistical null.

atom-000005

What source joint timing links earnings declines to the start of health treatment after parental death?

Activation: TRUE

- **evidence:** All retained source targets have actual estimates and their source-required uncertainty/interpretation, with no unresolved target. — Report the complete actual outcome vector. Coverage alone gives no direction, causal conclusion or replication confirmation.
- **partial:** At least one source target has actual evidence and at least one has a conclusive source/input no-estimate disposition. — Report available signed estimates and uncertainty, including mixed findings, beside every missing target and its exact limitation. Do not generalize the available subset to the missing scope.
- **boundary:** No retained source target is estimable after a conclusive source-definition or input boundary; this is not an operational failure. — There is no scientific estimate for this source scope. The paper is neither confirmed nor refuted by inability to estimate.
- **otherwise:** otherwise — Retain unresolved or inconclusive scientific disposition with actual evidence if any. Operational failures remain pending and are not done() or a statistical null.

atom-000006

Does the source earnings response vary with young children, and does formal childcare use respond to parental loss?

Activation: TRUE

- **evidence:** All retained source targets have actual estimates and their source-required uncertainty/interpretation, with no unresolved target. — Report the complete actual outcome vector. Coverage alone gives no direction, causal conclusion or replication confirmation.
- **partial:** At least one source target has actual evidence and at least one has a conclusive source/input no-estimate disposition. — Report available signed estimates and uncertainty, including mixed findings, beside every missing target and its exact limitation. Do not generalize the available subset to the missing scope.
- **boundary:** No retained source target is estimable after a conclusive source-definition or input boundary; this is not an operational failure. — There is no scientific estimate for this source scope. The paper is neither confirmed nor refuted by inability to estimate.
- **otherwise:** otherwise — Retain unresolved or inconclusive scientific disposition with actual evidence if any. Operational failures remain pending and are not done() or a statistical null.

atom-000007

How do source earnings and proximity responses vary with the surviving parent's health?

Activation: TRUE

- **evidence:** All retained source targets have actual estimates and their source-required uncertainty/interpretation, with no unresolved target. — Report the complete actual outcome vector. Coverage alone gives no direction, causal conclusion or replication confirmation.
- **partial:** At least one source target has actual evidence and at least one has a conclusive source/input no-estimate disposition. — Report available signed estimates and uncertainty, including mixed findings, beside every missing target and its exact limitation. Do not generalize the available subset to the missing scope.
- **boundary:** No retained source target is estimable after a conclusive source-definition or input boundary; this is not an operational failure. — There is no scientific estimate for this source scope. The paper is neither confirmed nor refuted by inability to estimate.
- **otherwise:** otherwise — Retain unresolved or inconclusive scientific disposition with actual evidence if any. Operational failures remain pending and are not done() or a statistical null.

atom-000008

Does the source relationship between pre-death family wealth and earnings responses support its inheritance explanation?

Activation: TRUE

- **evidence:** All retained source targets have actual estimates and their source-required uncertainty/interpretation, with no unresolved target. — Report the complete actual outcome vector. Coverage alone gives no direction, causal conclusion or replication confirmation.
- **partial:** At least one source target has actual evidence and at least one has a conclusive source/input no-estimate disposition. — Report available signed estimates and uncertainty, including mixed findings, beside every missing target and its exact limitation. Do not generalize the available subset to the missing scope.
- **boundary:** No retained source target is estimable after a conclusive source-definition or input boundary; this is not an operational failure. — There is no scientific estimate for this source scope. The paper is neither confirmed nor refuted by inability to estimate.
- **otherwise:** otherwise — Retain unresolved or inconclusive scientific disposition with actual evidence if any. Operational failures remain pending and are not done() or a statistical null.

atom-000009

How does the source parental-death design affect cohabitation and number of children?

Activation: TRUE

- **evidence:** All retained source targets have actual estimates and their source-required uncertainty/interpretation, with no unresolved target. — Report the complete actual outcome vector. Coverage alone gives no direction, causal conclusion or replication confirmation.

- **partial:** At least one source target has actual evidence and at least one has a conclusive source/input no-estimate disposition. — Report available signed estimates and uncertainty, including mixed findings, beside every missing target and its exact limitation. Do not generalize the available subset to the missing scope.
- **boundary:** No retained source target is estimable after a conclusive source-definition or input boundary; this is not an operational failure. — There is no scientific estimate for this source scope. The paper is neither confirmed nor refuted by inability to estimate.
- **otherwise:** otherwise — Retain unresolved or inconclusive scientific disposition with actual evidence if any. Operational failures remain pending and are not done() or a statistical null.

atom-000010

What are the source earnings spillovers to adult partners with young children after a first parent-in-law death?

Activation: TRUE

- **evidence:** All retained source targets have actual estimates and their source-required uncertainty/interpretation, with no unresolved target. — Report the complete actual outcome vector. Coverage alone gives no direction, causal conclusion or replication confirmation.
- **partial:** At least one source target has actual evidence and at least one has a conclusive source/input no-estimate disposition. — Report available signed estimates and uncertainty, including mixed findings, beside every missing target and its exact limitation. Do not generalize the available subset to the missing scope.
- **boundary:** No retained source target is estimable after a conclusive source-definition or input boundary; this is not an operational failure. — There is no scientific estimate for this source scope. The paper is neither confirmed nor refuted by inability to estimate.
- **otherwise:** otherwise — Retain unresolved or inconclusive scientific disposition with actual evidence if any. Operational failures remain pending and are not done() or a statistical null.

atom-000011

What does the source job-transition analysis show about occupational, firm and sector responses to parental death?

Activation: TRUE

- **evidence:** All retained source targets have actual estimates and their source-required uncertainty/interpretation, with no unresolved target. — Report the complete actual outcome vector. Coverage alone gives no direction, causal conclusion or replication confirmation.
- **partial:** At least one source target has actual evidence and at least one has a conclusive source/input no-estimate disposition. — Report available signed estimates

and uncertainty, including mixed findings, beside every missing target and its exact limitation. Do not generalize the available subset to the missing scope.

- **boundary:** No retained source target is estimable after a conclusive source-definition or input boundary; this is not an operational failure. — There is no scientific estimate for this source scope. The paper is neither confirmed nor refuted by inability to estimate.
- **otherwise:** otherwise — Retain unresolved or inconclusive scientific disposition with actual evidence if any. Operational failures remain pending and are not done() or a statistical null.

Conversion handoff (no independent scientific review)

- Verdict: `ready_for_transcription`

Paper-to-protocol conversion, not an independent scientific acceptance. Published results are reference targets; no replication has been executed. Native manifest acceptance means the package was compiled and frozen for execution only.

Checks

- `artifact_contract`: pass

Issues

Independent checks

No independent scientific reviewer is part of this two-stage lane. The converter's self-audit is in `docs/replication/conversion-audit.md`.

Source-fidelity self-check and Stage 1 handoff

Status: complete source-to-protocol conversion with explicit unresolved targets. This is the converter's self-check, not independent scientific review or acceptance. No empirical replication has occurred. The package is a draft canonical native World, ready for the Transcriber stage to preserve its science and build generated reports.

Sources and boundaries

The original 26-page journal PDF, journal-linked 31-page supplement, supplied article Markdown navigation source and all 36 available author `.do` programs were inspected. The programs include the supplied main analysis/matching files, all sibling cleaning and analysis files, the spouse construction, future-treated construction/analysis and orchestration files. No empirical program was executed. Supplied source hashes match `source-provenance.yaml`; its World copy is byte-identical. The supplementary PDF hash, live retrieval dates and public-code identity/access limits are recorded separately. The AEA publisher link and DataCite metadata identify archive version v1, but archive HTTP403 prevents remote file-hash correspondence and retrieval of missing auxiliary mappings. That limitation remains explicit.

The full supplied boundary contains 531 unique allowed tables and no runtime-validated tables. Its copied bytes match the input. All 18 named planning registers fall inside it. Selected FSV field/coverage/history records and four live official variable pages were checked; no metadata fact was treated as proof of mounted confidential rows, linking success, overlap or identifying variation. No historical deck, replication manual or earlier replica draft was used as a scientific source.

Applicable supplied skills and the package's complete atom-granularity and paper-scaffold references were read before substantive extraction. Their native contract is applied with the user-directed replica exception: two stages only, final IDs assigned here, no candidate hand, ideation, novelty competition or reviewer debate.

Scientific coverage and granularity

The World has 11 scientific analysis-module Atoms and one complete-membership Molecule. Population characteristics, descriptive event paths, matched labor effects, health services, co-occurrence, childcare, surviving-parent care, inheritance, family formation, spouse earnings and job choices retain their distinct source questions. Related outcomes, panels, specifications and robustness stay inside those questions. Matching/balance and routine input checks are inside the consuming protocols; no technical linkage node or figure-per-Atom decomposition was added. `source-map.md` explains every separate Atom's purpose.

The inventory preserves all six main figures, 23 supplementary figures, the main table and six supplementary tables in physical source order. It additionally retains 14 prose/context target groups, including footnote robustness and unavailable wealth/health variants. Original panel/category/row scope, coefficients, source-required intervals or SE, printed comparisons, baselines and counts are specified by the protocols/interfaces. Abstract/introduction restatements bind to their underlying targets rather than being counted as new experiments. Code-only transforms, tests and supplementary-looking outputs absent from the paper/supplement are explicitly not rebuilt.

No unreported estimator, control, subgroup, empirical diagnostic, sensitivity battery, equivalence band, significance cutoff, scoring rule or power calculation was added. Native required scalar ratings are neutral compatibility metadata only. There is no replication-success threshold. Actual-result vectors retain direction, magnitude, uncertainty, mixed findings and partial scope; coverage states have no scientific direction. All current `atom_result` fields are `prereg`, `strengths/shapes` null, evidence paths empty, and `registered_state` absent.

Source disagreements and exact stop dispositions

The audit retained, rather than repaired, the journal/code matching-strata/history discrepancy, matching/version/tie ambiguities, age/year normalization, post-period wording, under-six summary row, parental-age histogram, descriptive prediction denominator, dynamic/model caption conflicts, private/combined psychiatry scope, childcare eligibility, surviving-parent illness/calendars, parental-wealth

units/ownership, region definitions/contrast models, age-split cutpoints, future-treated windows/clustering/winsorization, spouse construction/contrasts and occupation ordering. Missing auxiliary mappings and prose-only specifications remain unavailable. `deviations.md` names the exact decision, source locator, affected targets and no-estimate disposition for each.

The final check also retained the B1 printed40–45 versus coded40–44 category discrepancy without silently fixing its label or bin. Its cumulative denominator precedes source suppression; B3’s width2 and B4’s26-bin request are explicit source parameters, while exact origins and disputed variables remain unresolved. These are not permissions to invent missing bins.

Several apparent issues were resolved by tracing the consuming program, without scientific repair: broader intermediate age/cause pools have explicit final restrictions; nonscreening visits are explicitly overwritten as total-minus-screening; unique Stata abbreviations do not by themselves imply undefined fields; matching and balance recalculate father age rather than necessarily consume an erroneous intermediate gap; FN age is recomputed. A draft prose transcription of the rank numerator was corrected against `clean_11` line76 before final delivery: the source is $100 \cdot (\text{rank} - 1) / (N - 1)$. No alternative rank outcome was registered. The source’s intended effect of bereavement on joint health/earnings outcomes is retained conditionally on its DiD assumptions; this does not identify a causal health-to-earnings effect or a mediation share.

Native, manuscript and interface verification

The repository `native read_bundle(..., birth=True)` structural check passed for all 11 Atoms and all 11 ordered Molecule members. The exact manifest slug is `jensen-zhang-2026-parental-death-native`. Canonical activation is TRUE for every Atom; no scientific result chooses execution. Input reuse is described separately from evidence interpretation. All restrictions/primary/supplementary/residual conditions use the existing native grammar. No competing graph or new activation schema was introduced.

The complete prepared manuscript contains 63 distinct unfilled result slots; all have matching output/column/unit/reporting bindings. Every Atom has a narrative disposition. All 29 figure entries point to implemented render functions and registered Atom output interfaces. Thirty Python sources (29 figure wrappers plus a reusable render helper) parse successfully. They load no empirical files at import, fit no models, choose no source bins, compute no uncertainty and embed no synthetic or published estimates. Source-specific annotation families exclude unreported summaries. The helper supports the host’s `plot_style` interface while retaining a local synthetic-test fallback; no host-owned `plot_style.py` was authored.

All 29 renderers passed complete, partially unavailable and fully missing synthetic input checks: 87 renders. Five malformed-interface cases were rejected (reversed confidence interval, wrong target ID, duplicate result key, wrong units and numerical values attached to an unavailable result). The final B1/B3 source-detail changes received six additional targeted complete/partial/missing renders. These are software/figure tests only, not empirical tests. Synthetic inputs, test programs, reports, rendered PDFs and

inspection images are outside the World in `conversion/source-inspection/figure-tests` and adjacent inspection files.

An early layout check detected clipped long missing-result notes in an eight-panel figure and an edge category label; the reusable source layout was corrected and the suite passed. A full original event window is now shown even when endpoint results are missing. Partial-input Figure4 and all-missing FigureB10 were visually inspected. Long explanations remain below numbered callouts without clipping; missing estimates are not zeros, connected interpolations or substituted published values. PDF inspection confirms vector-only output for the synthetic renders. No full paper or registration compilation was attempted.

Figure-ID validation repair — 2026-09-05

The reported failure, `invalid or duplicate planned figure id: figure_b1`, was reproduced using the repository's complete `validate_package` check with `require_paper=True` and the full allowed-table set. The paper validator requires `figure[1-9][0-9]*`; the earlier `read_bundle` structural check alone did not check figure identifiers. All 23 supplementary display IDs had the invalid underscore form, although they were unique.

The roster now uses `figure1-figure29` in the same order. Supplement B.1–B.23 correspond to native `figure7-figure29`. Source target IDs, source captions, renderer code and configurations, CSV input filenames and keys, manuscript slots, Atom assignments and every other roster field are preserved. The display-to-source crosswalk and native PDF output naming are documented in `docs/paper/figure-interface.md` and linked from `source-map.md`. This is a meaning-preserving serialization repair; no scientific source choice was reopened or resolved.

After the repair, complete native package validation returned no errors for all 11 Atoms, 11 ordered Molecule members, 29 figures and the 531-table boundary. All 63 manuscript slots retain their matching result-map bindings. The existing synthetic harness was adapted outside the World to traverse the actual roster, load each bound renderer and input target identity, and export using the native figure ID. All 87 complete/partial/missing renders passed, including vector-only output and all 29 long partial-reason checks; the five malformed-input cases were rejected. The rendered partial Figure4 was visually inspected and its complete explanation remains readable.

Before/after hashes confirm that the only changed World files are `figures.yaml`, `figure-interface.md`, `source-map.md` and this audit. All Atom/Molecule registrations, source provenance, full data boundary, target inventory, deviations, manuscript/result slots and plotting sources remain byte-identical. Repair hashes, validation details, adapted test harness and synthetic output files are retained in `conversion/source-inspection/`, including `figure-id-repair-tests/`, outside the World. No empirical program or full LaTeX/PDF document build was run. This check is the converter's own interface check, not independent scientific acceptance or empirical replication.

Handoff

Final authored contents are confined to atoms/, molecules/ and prereg/. The World contains no source PDF, screenshot, fulltext extraction cache, original author program, confidential row file, actual-result CSV, generated paper/registration LaTeX/PDF, host-owned style/template files or ZIP. The only CSV is the source-program identity/inspection index. No repository resources/controller state, external systems or other revisions were modified; no commit, push or release occurred.

Transcriber/host own both full LaTeX/PDF builds and packaging. They must preserve Atom scientific protocols, complete membership, source audit, all no-estimate decisions, the authored manuscript, slot map and implemented figure sources. A successful build does not resolve scientific source ambiguity. Missing definitions may be supplied only through traceable source-supported clarification; no scientific extension or source-code repair has been authorized. The final deliverable records neither independent scientific acceptance nor observed replication results.

Source ambiguities, conflicts and unavailable targets

These are stop dispositions, not authorizations to repair or add science. E=source-explicit, T=meaning-preserving engineering, A=unresolved ambiguity, C=conflict, U=unavailable target/input. Only E/T authorize implementation. All rows below remain unresolved in this conversion. No paper/code preference is silently applied.

U01 — unavailable target input

Affected: All register analyses.

Source: Journal II; all cleaning programs; schema-evidence.md.

Missing decision: Historical extracts, common pseudonym namespace, field vintages and complete covered years have not been inspected. LMDb code uses PNR while current documentation exposes PNR12; early IDAN code uses IDANTYPE while current schema exposes TYPE.

Disposition: Obtain source-corresponding metadata and an authorized runtime field/link map. Do not equate differently named person keys or fabricate historical coverage. No affected estimate until verified.

U02 — unavailable target input

Affected: Education/matching; real earnings/wealth; psychology; 11-region geography.

Source: clean_3, clean_6, clean_10, clean_11; set_globals.

Missing decision: Missing psychology_psychiatry.dta, inflation_factor_2020.dta, n_audd_level_1114_k.dta, n_audd_pria_111x_k.dta and c_komlan_v4_k.dta, with their versions. The journal describes PRIS112 and 2020 prices but does not give the complete factor series.

Disposition: Stop each consuming derivation. Do not replace with guessed specialty, education, geography or inflation crosswalks. A checked equivalent supplied mapping is engineering only when it preserves every source category/value.

C01 — source-code conflict

Affected: All NN matched targets, including Table B.3 and spouse targets where shared construction is consumed.

Source: Journal III p1820 and fn14; `clean_15_matching.do` lines249–366.

Missing decision: Journal describes exact earnings and under-six groups and three-year earnings/employment matching. The code combines some earnings cells and young/childless/older-child cells into 34 groups; its distance list contains $t=-3$ earnings but no separate three annual employment indicators. The public sources do not identify which construction generated the journal results.

Disposition: Block NN sample-dependent estimates until the intended strata and history vector are established. Keep both descriptions in the audit; do not run two scientific variants or silently choose code.

A01 — unresolved source ambiguity

Affected: All matching-dependent targets.

Source: `clean_15` lines363–440; `clean_16`; `set_globals`; `do_all`.

Missing decision: Nearest-neighbor ties are reduced to `match_id1` following sorting and `seed123456`. Stata and `teffects/ado` versions are absent. Reused controls are relabelled with a finite duplicate-handling loop. Conflicting duplicate demographic/income rows have force/first-row rules without a stable full tie order.

Disposition: Require source software semantics and deterministic selected edges when a consequential tie/duplicate occurs. An isomorphic pair ID relabel is engineering; an altered match edge or tie selection is not. Stop affected targets if correspondence cannot be proved.

A02 — unresolved source ambiguity

Affected: Matched Eq2–4 estimates and source linear-controls robustness.

Source: Journal fn16 p1821; `analysis_5` and consuming `reghdfe` calls.

Missing decision: The text illustrates equal effects for two ages and two years without fixing which constraints apply to each restricted sample; code absorbs age/year/person effects without those explicit constraints. The separate linear-age/year robustness has no complete program or output specification.

Disposition: An algebraic demonstration that treatment coefficients and their variance are invariant under the source normalizations may establish engineering equivalence. Until then retain the exact missing normalization; never add a selected constraint or linear model. The prose robustness is no-estimate.

A03 — unresolved source ambiguity

Affected: Eq3/4 post-period summaries and co-occurrence immediate-year ratios.

Source: Journal equations3/4: after death/five years; `analysis_5` `post=time>=0`; Figure4 discussion; `analysis_14`.

Missing decision: Code includes event year0 through5 (six annual cells). Narrative uses after death/five years. Co-occurrence discussion says immediately following while code graphs a year0 spike.

Disposition: Event-time coordinates are explicit. Do not silently call six cells five years, or bind a prose ratio to year0 rather than1. Block ambiguous aggregate/ratio targets pending source clarification; dynamic targets are not blocked solely by this wording.

C02 — source-code conflict

Affected: Under-six rows in Table1, TableB1, TableB2, TableB3.

Source: Captions/row labels; analysis_2 and analysis_15 share_young construction.

Missing decision: Table labels say under6; code uses youngest age<=6.

Disposition: No estimate for this row until the age boundary is settled. Preserve other summary rows if their own dependencies are available.

C03 — source-code conflict

Affected: FigureB3 sudden-death age distribution.

Source: FigureB3 p9; analysis_2 histogram commands.

Missing decision: Non-sudden histogram uses parent age at death; sudden histogram uses first_death_age measured at t=-1.

Disposition: Block affected histogram comparison; do not shift one distribution by a year as a repair.

A04 — unresolved source ambiguity

Affected: FigureB1 and introductory91%/age shares; histogram endpoints.

Source: Journal p1811 and fn10; appendix B1/B3/B4; analysis_1/2.

Missing decision: B1 code forms cumulative age-at-first-death shares among deaths1985–2014, not an age-specific population risk. The prose91% at19+ is not determined by plotted five-year bins alone. Histogram bin26/default origins and suppression order lack a pinned Stata version.

Disposition: Use the source-defined conditional denominator for B1; explicitly describe it. Retain prose-only exact-age shares as unavailable if their exact denominator/year restriction cannot be determined. Do not read them from plotted bars or infer an age-specific death risk. Require source histogram algorithm/edges before reproducing bins.

C04 — source-code conflict

Affected: FiguresB5/B23 percentage-scaled descriptive paths.

Source: Journal Eq1 and figure captions; analysis_3/4 prediction and collapse blocks.

Missing decision: Regressions are sex-specific but predict without an estimation-sample restriction and then collapse both sexes before relabelling sex. This differs from the sex-specific counterfactual denominator described by the paper.

Disposition: Block plotted relative paths until prediction population and denominator are determined. Unscaled fitted coefficients are retained as protocol detail, not substituted into the published percentage display.

C05 — source-code conflict

Affected: Dynamic Figures4,B7,B12; Figure6 and B15 model labels.

Source: Figures4/B7/B12 captions cite Eq4 although panels and analysis_14/5 use Eq2 dynamics. Figure6/B15 captions cite parent-type Eq4 while analysis_12 uses illness-by-treatment Eq3-style terms..

Missing decision: The caption and program cannot jointly specify the reported model.

Disposition: Keep exact displayed estimands and candidate source formulations in protocols; block conflicting figure coefficients rather than correcting equation citations silently.

A05 — unresolved source ambiguity

Affected: Psychiatrist consultations prose target; TableB3 psychiatry row.

Source: Journal fn22 p1825; clean_2/3; analysis_6/15.

Missing decision: Text refers to private psychiatrist consultations; code creates a combined private-plus-hospital psychiatry measure. The private-service crosswalk is missing, and the intended combined/private target cannot be fully established.

Disposition: Retain a no-estimate disposition for the unresolved consultation target. Do not substitute all psychiatric encounters.

C06 — source-code conflict

Affected: Figure5; FigureB13 eligibility inherited from analysis8.

Source: Journal V.B: with/without young children; analysis_8 lines18–21.

Missing decision: Code drops childless adults before classifying young/older children. The stated without-young group does not explicitly exclude childless adults.

Disposition: Block the affected earnings subgroup comparison pending eligibility clarification; do not invent a childless stratum or silently exclude it.

A06 — unresolved source ambiguity

Affected: FigureB14 childcare uptake.

Source: FigureB14 under5/above5 labels; V.B age0–5; clean_9; analysis_8 final childcare sections.

Missing decision: Code uses youngest $\leq$ 5 and older6–10 at event; wording under5 is inconsistent at age5. Exact youngest-child identity for ties can affect uptake. Source empties age-ineligible values then codes missing uptake0.

Disposition: Preserve source6–10 and 0–5 as code evidence, but block disputed eligibility/ambiguous youngest-child identity. No guessed age5 rule, tie child or substitute childcare measure.

C07 — source-code conflict

Affected: Figure6, B15, B16 and eldercare interpretation.

Source: Journal V.C and Fig6; appendix B15/16; analysis_12 illness/calendar blocks.

Missing decision: Paper describes more than three months within five years; code flags any single annual hospital-days value ≥ 90 , includes $t=0$ in the post interval, and permits 2019 although days end 2018 then fills missing 0. B16 does not specify whether its displayed illness split is the pre or post definition; code produces both.

Disposition: Block illness-stratified outputs until accumulation window, strictness, calendar end and B16 definition are determined. Never fill an unavailable hospital year with zero or add a repaired definition.

U03 — unavailable target specification

Affected: Mental-prescription effects by surviving-parent illness; inpatient-health rematching; own-health rematching; second-death restriction; linear controls.

Source: Journal pp1820–21 fn14–16; p1831 eldercare paragraph.

Missing decision: These are reported in prose without complete corresponding programs/output definitions. Missing: inpatient-only count/codes and lookback, exact rematching vector/combinations and years, which outcomes/specifications were rerun, and the precise linear-control formulation. The second-death-gap > 5 restriction is stated, but rematching/reuse and the full output set are not.

Disposition: Inventory all claims, with no-estimate rows. Do not infer a sensitivity battery from one or more health variables. The supplied threshold > 5 is preserved; missing design decisions remain unresolved.

C08 — source-code conflict

Affected: TableB4 parental wealth interactions.

Source: Appendix A.A/TableB4 p30; analysis_9 lines 21–31, 75.

Missing decision: Table labels parental wealth in 1000 DKK; clean_12 retains DKK and analysis_9 divides by 10000. Pair-level max baseline wealth is not restricted to the treated member, potentially selecting control-side values.

Disposition: Block interaction coefficients and their interpretation pending unit and owner definition. No factor-ten coefficient repair or arbitrary selection of treated wealth.

U04 — unavailable target specification

Affected: Own assets, deceased-only assets, surviving-only assets interactions.

Source: Appendix A.A and fn29.

Missing decision: No consuming programs for three prose variants; timing/change-versus-level of own assets, scaling, interaction construction and estimators are not fully specified.

Disposition: Keep separate unavailable targets inside inheritance Atom. Do not invent own-wealth change or assign the main wealth specification to them.

C09 — source-code conflict

Affected: FigureB19 regional heterogeneity and displayed comparisons.

Source: Appendix B19 caption p25; analysis_10.

Missing decision: Caption says five regions, code uses 11-region proximity. Group-specific models interact fixed effects with proximity while pooled gender-contrast models use a different FE structure; one displayed men p-value is assigned an F statistic in code.

Disposition: Block conflicting coefficients/contrasts pending geography and contrast-model correspondence. Do not switch to 5 regions or repair the p-value assignment.

A07 — unresolved source ambiguity

Affected: TableB6 parent/child age heterogeneity.

Source: TableB6 p31; analysis_11 median splits.

Missing decision: Printed rows name70/38; code recalculates sample medians. The sources do not state whether printed cutpoints are fixed scientific bins or source-realized median labels.

Disposition: Retain cause and calendar strata as explicit. Block age-stratum targets pending fixed-versus-recomputed threshold clarification; no new cutpoints.

C10 — source-code conflict

Affected: FigureB21 future-treated comparison; FigureB22.

Source: Appendix A.F p5/B21/22; FN_clean; FN_analysis.

Missing decision: A.F example uses event window-3..5 while code uses -5..5. Captions cluster at individual level; code clusters person×stack. FN earnings code winsorizes p5/p95 while main98% rule is p1/p99 and text says same specifications.

Disposition: Block affected FN targets until window, cluster and earnings tail rules are determined. Employment has no winsor conflict but retains window/cluster ambiguity. No alternative-estimator experiment.

A08 — unresolved source-code execution

Affected: FN targets.

Source: FN_clean_1_treatment_controls.do fertility block.

Missing decision: The program drops age_of_youngest after merging fertility and subsequently references it. This is separate from the scientific window/cluster/winsor conflicts.

Disposition: Do not execute or silently fix author code. The translated independent earnings specification remains documented; do not infer a missing fertility construction.

A09 — unresolved source ambiguity

Affected: FigureB18 spouse sample and comparisons.

Source: Appendix A.C/B18; clean_16; analysis_16.

Missing decision: Spouse matching includes own-parent population presence rather than explicit survival, while prose says own parents alive; consecutive cohabitation duration lacks an explicit chronological sort; duplicate event partners are force-dropped. Separate spouse models absorb spouse age, whereas pooled gender tests use bereaved partner age and extra parental variables.

Disposition: Block source-dependent spouse estimates/contrasts until matching, tie/order and model consistency are determined. Do not equate presence with survival or repair pooled controls.

A10 — unresolved source ambiguity

Affected: Occupation switching columns in TableB5.

Source: Appendix A.D/B5; analysis_13.

Missing decision: Code compares four-digit occupation with up to eight earlier nonmissing records, but the within-person-match sort omits year. Source does not supply a stable ordering or a DISCO transition crosswalk.

Disposition: Block order-dependent occupation switching; do not invent a last-year rule or harmonization. Firm and sector targets remain separately retainable subject to shared matching/input gates.

A11 — unresolved source ambiguity

Affected: Record-dependent targets involving simultaneous deaths, duplicate IDs, missing outcomes or fractional ages.

Source: clean_6–8,11,13,15/16; source notes.

Missing decision: Code sometimes force-drops duplicates, assigns maternal first death on equal death dates and treats missing values as zero. Documentation cannot establish whether these cases occur or are genuine covered non-events.

Disposition: Retain each source rule and exact collision context. If a row has ambiguous identity or an unavailable register-year, stop the consuming target; do not impute or choose a new tie rule. Absence inside a verified complete source event register may follow the explicit source zero rule.

Resolved inspection questions, not source repairs

The final sudden-cause filter and final treated-age25–50 filter supersede broader intermediate construction by explicit author commands. Nonscreening hospital visits are explicitly overwritten by total-minus-screening in the consuming

program. Stata unique abbreviations can explain TYPE/TYPE_2008 and formal_child_care/formal_child_care_old5; they are not automatically undefined-variable errors. The mistaken father-child gap in an intermediate field does not establish erroneous matching father age: matching uses father's actual birthday, and balance code recalculates it. FN age is explicitly recomputed in its analysis. Rank is $100 \times (\text{rank}-1)/(\text{N}-1)$, verified against clean_11 line76. These checks prevent false conflict claims. No empirical execution was used.

Published rounding and point values are comparison targets only. Rebuilding every source target is currently underdetermined; a complete conversion preserves these boundaries rather than filling them. No scientific extension is authorized.

A04 additional exact display detail

analysis_1 labels the $\text{floor}(\text{age}/5)=8$ category40–45 although its coded ages are40–44. Preserve the printed label and record the disagreement; the affected bin target remains no-estimate until its intended boundary is established. No silent label repair is authorized. B1 cumulative numerator/denominator are calculated before the source<5 suppression. analysis_2 B3 explicitly sets width2; B4 sets26 bins. These specified parameters are retained, while source bin origins/edges and suppression/model correspondence remain gates.

Matching specification and unresolved correspondence

This is a transcription of source logic, not runnable author code and not permission to select code over journal. C01/A01/A02/A09 apply before scientific implementation. Main matching serves Atoms3–9/11; Atom10 constructs a separate spouse sample. Shared construction creates input dependencies, not a scientific-evidence dependency or an extra Atom.

Main source clean_15: candidate treated age25–55 initially, final25–50; final sudden codes24/25/26/29/46/47/49. Candidate controls receive pseudo D1985–2014, first parental death>D+5 or not observed, two known parent birthdays, both parents present all11 pseudo years, adult ages20–61 in source pool and pseudo-event age<=55, baseline $t=-1$ age>=24. Panel construction and final checks are in D1. For each D and sex concatenate treated/controls, retain source eligible baseline history and use three annual earnings categories at $t=-3,-2,-1$. Each category is0 unless both earnings>0 and ATP participation>0; otherwise pooled sex/year-specific positive-income tercile. Earnings category0 thus need not literally mean zero earnings; this is another exact-stratum correspondence issue within C01.

Let r =earnings category at-2, c =category at-1. Rows/columns below are0,1,2,3; entries are source group IDs. Young child at-1 means age<6; older6–21; childless sentinel22. These matrices document the literal code and are not authorized scientific alternatives to the journal exact-match statement.

Youngest category	Group matrix (rows r , columns c)
Young	[[1,2,3,3],[4,5,6,7],[8,8,9,10],[11,11,12,13]]

Youngest category	Group matrix (rows r, columns c)
Childless	[[14,15,3,3],[16,17,18,19],[20,20,21,22],[11,11,23,24]]
Older	[[25,26,3,3],[27,28,29,19],[30,30,31,32],[11,11,33,34]]

Within sex,D,group, source teffects nnmatch uses Mahalanobis distance over categorical indicators for t=-3 earnings category; floor(adult age at-1/5); child count at-1; cohabitation,5-region,public,education at-1; youngest-age category floor(age/5) with 21->5,22->6; floor(parent age/5). Source parent-age matching fields are constructed from D minus parent birth year and therefore fixed over the prehistory; do not silently shift to t=-1 in code transcription. Missing region->0,education->0,public->2 in match preparation; source arrays include extra variables/categories not consumed by the NN distance. Their presence is not permission to add them. SortPNR, source seed123456, retain first generated nearest-control index, handle reused controls with source matching/relabel logic. The matching call has a constant placeholder outcome0; it is a device to request match indices, not a scientific outcome. Preserve one treated/control pair and 11-year records only after source correspondence of ties/duplicates is established. There is no new caliper, propensity model, exact-match repair, weight or balance threshold.

Spouse source clean_16 additionally requires spouse baseline age >=24 (both treatment/control), main and spouse source histories, event-time partner and a full 11-row spouse window; it caps spouse children5. Age bins use floor(age/3) for spouse and biological adult. Main/spouse annual earnings use two positive-income categories conditional ATP>0; otherwise 0. Source group matrices below use spouse earnings r at-2,c at-1 in 0,1,2. Different year blocks are retained because the source actually uses distinct coarsening. Nearest-distance vector uses main earnings at-1,main ATP at-1,main age bin,main education at-1; spouse earnings/ATP at-3,spouse public/age bin/education at-1,spouse own-father/mother-in-pop at-1,youngest-age category,cohabitation tenure at-1. All are categorical source terms. Sort spouse ID; retain source first matching edge. A09 prevents substituting own-parent death indicators or a newly sorted tenure algorithm.

Spouse source years 1985,1990–2009,2011–2014

- Childless: [[1, 2, 2], [3, 4, 5], [6, 6, 7]]
- Young: [[8, 9, 9], [10, 11, 12], [13, 13, 14]]
- Older: [[15, 16, 16], [17, 18, 19], [20, 20, 21]]

Spouse source years 1986–1989

- Childless: [[1, 2, 2], [3, 4, 5], [6, 6, 7]]
- Young: [[8, 9, 9], [10, 11, 12], [13, 13, 14]]
- Older: [[15, 15, 15], [16, 16, 16], [17, 17, 17]]

Spouse source years 2010

- Childless: [[1, 1, 1], [2, 3, 4], [5, 5, 6]]

- Young: [[7, 8, 8], [9, 10, 11], [12, 12, 13]]
- Older: [[14, 14, 14], [15, 15, 16], [17, 17, 17]]

These are source-explicit candidate algorithms. They remain behind the documented source-consistency gates; runtime may not choose a matrix or matching construction based on observed balance or results.

Published reference values, not replication results

These checked anchors identify the source displays and interpretation targets. They are not current estimates, sample-size assumptions, runtime tuning criteria or manuscript result-slot values. Every source target remains in target-inventory.md even when its printed value is not transcribed here. A later comparison may show source and actual values with their original units and printed rounding; no numerical pass/fail band is registered.

Source locator	Reference quantity	Published value	Role
Journal Table 1, p1819	Sudden-death observations	206,286	Source descriptive population count only
Journal Table 1, p1819	Non-sudden-death observations	528,057	Source descriptive population count only
Journal Table 1, p1819	Total observations	734,343	Source descriptive population count only
Supplement Table B3, p13	Treated and matched-control baseline counts	206,282 each	Source matching table count; not automatically equal to Table 1
Journal Figure 1D, p1823	Men, indexed-earnings pooled DiD (SE)	−1.509 (0.148)	Source figure annotation; post aggregation remains A03
Journal Figure 1D, p1823	Women, indexed-earnings pooled DiD (SE)	−2.041 (0.171)	Source figure annotation; no assumed replication effect
Journal p1827	Joint onset / earnings-onset comparison, women / men	36% / 21%	Source prose ratios; event mapping remains A03

Source locator	Reference quantity	Published value	Role
Journal fn25, p1827	Joint onset / mental-treatment- onset comparison, women / men	17% / 30%	Source prose ratios, no new ratio inference
Supplement Table B6, p31	Printed parent-age and adult-age splits	70 and 38 years	Published labels; fixed-versus- recomputed cutpoints remain A07

Source point estimates and uncertainty elsewhere in the paper/supplement are inventory targets at their exact locators. This reference note deliberately contains no actual-result CSV, Atom evidence path or filled replication slot. It is not a substitute for reading the source figure/table when a source comparison is eventually prepared.

Live retrieval and access log

Date2026-09-05 UTC. Queries were restricted to source identity/version, missing supplement/code inputs, and documentation semantics; no novelty search or scientific extension.

- **aea:** <https://www.aeaweb.org/articles?id=10.1257/aer.20240432>; 2026-09-05T11:55:08.873719+00:00; HTTP 200; final <https://www.aeaweb.org/articles?id=10.1257/aer.20240432>; SHA256 f6a2bd250915d56b5c4a71e650c2d904dc08df5998672613c9c5502a4260bf9d.
- **appendix:** https://www.aeaweb.org/doi/suppl/10.1257/aer.20240432/suppl_file/aer.20240432_app.pdf; 2026-09-05T11:55:08.759823+00:00; HTTP 404; final https://www.aeaweb.org/doi/suppl/10.1257/aer.20240432/suppl_file/aer.20240432_app.pdf; SHA256 e55ff97d6642a9e0700c40fd5e398c463a3fafdee41d40ccaaa2600d7c1918a6.
- **appendix_pubs:** https://pubs.aeaweb.org/doi/suppl/10.1257/aer.20240432/suppl_file/aer.20240432_app.pdf; 2026-09-05T11:55:08.583244+00:00; HTTP 403; final https://pubs.aeaweb.org/doi/suppl/10.1257/aer.20240432/suppl_file/aer.20240432_app.pdf; SHA256 35877f7be8efa32fd4fadb9bce6336ca24d98b08dd100a511c787fe56c6d4269b.
- **icpsr:** <https://www.openicpsr.org/openicpsr/project/239029/version/V1/view>; 2026-09-05T11:55:08.800461+00:00; HTTP 403; final <https://www.openicpsr.org/openicpsr/project/239029/version/V1/view>; SHA256 934c6500b3a6a5b2c33f143b753d7f9b37f6874f08353e7321b01f1a20c9a40e.
- **openalex:** <https://api.openalex.org/works/https://doi.org/10.1257/aer.20240432>; 2026-09-05T11:55:08.594267+00:00; HTTP 429; final <https://api.openalex.org/works/https://doi.org/10.1257/aer.20240432>; SHA256 ea277d2335bab904012158e3e1510fd3ff41415a8f5fca5ee845bd61e5006362.
- **pmc:** <https://eutils.ncbi.nlm.nih.gov/entrez/eutils/esearch.fcgi?db=pmc&term=10.1257%2Faer.20240432&retmax=1>; 2026-09-05T11:55:08.785626+00:00; HTTP 200; final <https://eutils.ncbi.nlm.nih.gov/entrez/eutils/esearch.fcgi?db=pmc&term=10.1257%2Faer.20240432&retmax=1>; SHA256 c15df2d63636aa4e802fce4cbda9485c478675a298d140db238eec2626befa39.
- **supplement:** <https://www.aeaweb.org/articles/materials/25093>; 2026-09-05T11:55:24.643363+00:00; HTTP 200; final <https://assets.aeaweb.org/asset-server/files/24591.pdf>; SHA256 7cebe88f0cdef5b8e8865d457f2abae64aad2691dbc8e9cf94f3faf5f1c51105.

239029 Jensen Zhang; site.mathiasfjensen.com parental death replication. Results included AEA and older working-paper records; only AEA journal-linked materials were used scientifically. No accessible missing author-input bundle was found. Older decks/manuals/earlier replica drafts were not read or used.

Stata documentation lookup: site.stata.com "set varabbrev" "on"; official [pvarabbrev.pdf](#) supports variable-abbreviation semantics, only an engineering-language check. Official FSV variable URLs for BEF,LMDB,IDAN,DODSAASG were retrieved live200, inspected narrowly for key/vintage fields; supplementary local schema pages supply remaining field evidence. Documentation is not proof of rows or identifying variation.

Field and historical coverage evidence

Read2026-09-05; metadata only. Catalog layer FSV is planning evidence, not runtime availability. Selected field rows below are paraphrased from the relevant register pages, with exact field names and documented windows. Current table coverage is not the source extract vintage.

AKM

Repository locator: [resources/schema/data-registry-bolero/fsv/labour-income-and-wealth/AKM.md](#). [Official variable documentation](#). Status register, 1976–2024, update frequency: Year.

Field	Documented window	Meaning
BESKST	1976–2001	Employment status 1980 to 2001 (Income/AKM)
BESKST13	1991–2024	Code for the person's main source of income (INDKOMST/AKM)
DISCO08_LOEN_INDK	2010–2024	Trade code for the principal employee occupation during the year. (AKM)
DISCO08_SEL_INDK	2010–2024	Trade code for work in own business (AKM)
DISCOLOEN_INDK	1991–2009	Trade code for the principal employee occupation during the year. (AKM)
DISCOSEL_INDK	1991–2009	Trade code for work in own business (AKM)
PNR	1976–2024	Personal identification number

Runtime obligation: verify source-compatible vintage, key namespace and required year coverage; retain U01/U02 if unresolved.

BEF

Repository locator: `resources/schema/data-registry-bolero/fsv/population-and-elections/BEF.md`.
[Official variable documentation](#). Status register, 1985–2026, update frequency: Qtr.

Field	Documented window	Meaning
ALDER	198512–202606	Age at the reference date.
CIVST	198512–202606	Marital status
E_FAELE_ID	198512–202606	E_fælle
FAR_ID	198512–202606	Father’s person ID
FOED_DAG	198512–202606	Date of birth
KOEN	198512–202606	Sex
KOM	198512–202606	Municipality code
MOR_ID	198512–202606	The mother’s person ID
PNR	198512–202606	Personal identification number
REG	198512–202606	Region

Runtime obligation: verify source-compatible vintage, key namespace and required year coverage; retain U01/U02 if unresolved.

BOERNFB

Repository locator: `resources/schema/data-registry-bolero/fsv/living-conditions/BOERNFB.md`.
[Official variable documentation](#). Status period register, 2015–2021, update frequency: Closed.

Field	Documented window	Meaning
PNR	2015–2021	Personal identification number

Runtime obligation: verify source-compatible vintage, key namespace and required year coverage; retain U01/U02 if unresolved.

BOERNSB

Repository locator: `resources/schema/data-registry-bolero/fsv/living-conditions/BOERNSB.md`.
[Official variable documentation](#). Status period register, 2015–2024, update frequency: Year.

Field	Documented window	Meaning
KOM	2015–2024	Municipality code
PNR	2015–2024	Personal identification number

Runtime obligation: verify source-compatible vintage, key namespace and required year coverage; retain U01/U02 if unresolved.

DAGI

Repository locator: [resources/schema/data-registry-bolero/fsv/living-conditions/DAGI.md](#).
[Official variable documentation](#). Status register, 1995–2014, update frequency: Closed.

Field	Documented window	Meaning
PNR	1995–2014	Personal identification number
YDELSE	1995–2014	Benefit

Runtime obligation: verify source-compatible vintage, key namespace and required year coverage; retain U01/U02 if unresolved.

DODSAARS

Repository locator: [resources/schema/data-registry-bolero/fsv/health/DODSAARS.md](#).
[Official variable documentation](#). Event register, 1970–2001, update frequency: Closed.

Field	Documented window	Meaning
C_LISTE_49	2001	C_LISTE_49
D_DODSDTO	2001	Date of death
PNR	2001	Personal identification number

Runtime obligation: verify source-compatible vintage, key namespace and required year coverage; retain U01/U02 if unresolved.

DODSAASG

Repository locator: [resources/schema/data-registry-bolero/fsv/health/DODSAASG.md](#).
[Official variable documentation](#). Event register, 2002–2022, update frequency: Closed.

Field	Documented window	Meaning
C_LISTE49	2022	Diagnosis group, the 49-item list
D_STATDATO	2022	D_STATDATO
PNR	2022	Personal identification number

Runtime obligation: verify source-compatible vintage, key namespace and required year coverage; retain U01/U02 if unresolved.

FAIN

Repository locator: [resources/schema/data-registry-bolero/fsv/population-and-elections/FAIN.md](#).
[Official variable documentation](#). Status register, 1980–2007, update frequency: Closed.

Field	Documented window	Meaning
AEGTEFAELLE_ID	1980–2007	Spouse Person Id
ALDER	1980–2007	Age at the reference date.
CIVST	1980–2007	Marital status
FAR_ID	1980–2007	Father’s person ID
KOEN	1980–2007	Sex
KOM	1980–2007	Municipality code
MOR_ID	1980–2007	The mother’s person ID
PNR	1980–2007	Personal identification number

Runtime obligation: verify source-compatible vintage, key namespace and required year coverage; retain U01/U02 if unresolved.

IDAN

Repository locator: `resources/schema/data-registry-bolero/fsv/labour-income-and-wealth/IDAN.m`
[Official variable documentation](#). Status register, 1980–2024, update frequency: Year.

Field	Documented window	Meaning
ARB_SEKTORKODE	2008–2024	Sector code of the workplace
KOM	1980–2024	Municipality code
LBNR	1980–2024	Workplace serial number
PNR	1980–2024	Personal identification number
TYPE	1980–2007	Type of employment relationship
TYPE_2008	2008–2024	Type of job

Runtime obligation: verify source-compatible vintage, key namespace and required year coverage; retain U01/U02 if unresolved.

IDAP

Repository locator: `resources/schema/data-registry-bolero/fsv/labour-income-and-wealth/IDAP.m`
[Official variable documentation](#). Status register, 1980–2024, update frequency: Year.

Field	Documented window	Meaning
ATPXX	1980–2024	ATP converted to B rate
PNR	1980–2024	Personal identification number

Runtime obligation: verify source-compatible vintage, key namespace and required year coverage; retain U01/U02 if unresolved.

IDAS

Repository locator: [resources/schema/data-registry-bolero/fsv/labour-income-and-wealth/IDAS.md](#).
[Official variable documentation](#). Status register, 1980–2023, update frequency: Year.

Field	Documented window	Meaning
EJERKO	1980–2007	Owner code (recoded)
KOM	1980–2024	Workplace municipality code
LBNR	1980–2024	Workplace serial number

Runtime obligation: verify source-compatible vintage, key namespace and required year coverage; retain U01/U02 if unresolved.

IND

Repository locator: [resources/schema/data-registry-bolero/fsv/labour-income-and-wealth/IND.md](#).
[Official variable documentation](#). Status period register, 1980–2024, update frequency: Year.

Field	Documented window	Meaning
BESKST	1980–2001	Employment status 1980 to 2001 (Income/AKM)
BESKST13	1980–2024	Code for the person's most important source of income (INDKOMST/AKM)
ERHVERVSINDK_13	1980–2024	Business income, wages and net profit from self-employment incl. certain fees
ERHVERVSINDK_GL	1980–2013	Business income, wages and net profit from self-employment
FORM	1980–1997	Tax-assessed value of net wealth at the end of the year, excl. pension wealth
FORMREST_NY05	1997–2024	Net residual wealth at the end of the year, excl. pension wealth
PNR	1980–2024	Personal identification number

Runtime obligation: verify source-compatible vintage, key namespace and required year coverage; retain U01/U02 if unresolved.

LMDB

Repository locator: [resources/schema/data-registry-bolero/fsv/health/LMDB.md](#).
[Official variable documentation](#). Event register, 1995–2025, update frequency:
Half-year.

Field	Documented window	Meaning
ATC	199512–202512	ATC code, detailed (s)
EKSD	199512–202512	Processing date
KOM	199512–202512	Municipality code
PNR12	199512–202512	CPR number

Runtime obligation: verify source-compatible vintage, key namespace and required
year coverage; retain U01/U02 if unresolved.

LPR_ADM

Repository locator: [resources/schema/data-registry-bolero/fsv/health/LPR_ADM.md](#).
[Official variable documentation](#). Event register, 1977–2019, update frequency: Closed.

Field	Documented window	Meaning
C_ADIAG	1977–2019	Principal diagnosis
D_INDDTO	1977–2019	Date of admission
D_UDDTO	1977–2019	Date of discharge
PNR	1977–2019	Personal identification number

Runtime obligation: verify source-compatible vintage, key namespace and required
year coverage; retain U01/U02 if unresolved.

PSYK_ADM

Repository locator: [resources/schema/data-registry-bolero/fsv/health/PSYK_ADM.md](#).
[Official variable documentation](#). Spell/event-history (longitudinal) register,
1995–2019, update frequency: Closed.

Field	Documented window	Meaning
C_ADIAG	2019	Principal diagnosis
D_INDDTO	2019	Date of admission
D_UDDTO	2019	Date of discharge
PNR	2019	Personal identification number

Runtime obligation: verify source-compatible vintage, key namespace and required
year coverage; retain U01/U02 if unresolved.

SSSY

Repository locator: [resources/schema/data-registry-bolero/fsv/living-conditions/SSSY.md](#).
[Official variable documentation](#). Status period register, 2005–2025, update frequency: Year.

Field	Documented window	Meaning
PNR	2005–2025	Personal identification number
SPECIALE	2005–2025	6-digit speciality
YDLANT	2005–2025	Number of services within the specialty

Runtime obligation: verify source-compatible vintage, key namespace and required year coverage; retain U01/U02 if unresolved.

SYSI

Repository locator: [resources/schema/data-registry-bolero/fsv/living-conditions/SYSI.md](#).
[Official variable documentation](#). Status period register, 1990–2005, update frequency: Closed.

Field	Documented window	Meaning
PNR	1990–2005	Personal identification number
SPECIALE	1990–2005	6-digit speciality
YDLANT	1990–2005	Number of services within the specialty

Runtime obligation: verify source-compatible vintage, key namespace and required year coverage; retain U01/U02 if unresolved.

UDDA

Repository locator: [resources/schema/data-registry-bolero/fsv/education-and-knowledge/UDDA.md](#).
[Official variable documentation](#). Status register, 1980–2025, update frequency: Year.

Field	Documented window	Meaning
HFAUDD	198012–202509	Highest completed education
HF_VFRA	198012–202509	Time of completion of the highest completed education
PNR	198012–202509	Personal identification number

Runtime obligation: verify source-compatible vintage, key namespace and required year coverage; retain U01/U02 if unresolved.

Consequential historical differences

BEF history27 March2020 describes backdated municipality/region replacement; UDDA history19 April2021 describes rerun/relabelled September snapshots; IDAN history24 September2018 describes TYPE_2008 and researcher prioritization after non-November type3 disappears; AKM history7 February2017 describes corrected employment/occupation variables; LMDB currently lists PNR12, not source PNR. LPR_ADM closes March2019; source hospital analyses stop2018. These facts motivate exact source-vintage gates, not new harmonizations. DODSAARS and DODSAASG variable rows list cumulative-extract years2001/2022; that does not prove record coverage or source2019 parity. No historical pilot/grant table boundary constrains this run.

Source-to-Atom audit map

Scientific source: Jensen and Zhang2026, American Economic Review116(5),1811–1836, DOI10.1257/aer.20240432; journal-linked31-page supplement. Supplied original26-page PDF, entire supplement and all36 author .do files were inspected. Code index provides immutable hashes/line counts; the provenance copy records supplied source identities. Journal text is scientific authority; disagreements with supplement/code stay unresolved. There is no independent scientific acceptance.

E: source-explicit rules; T: meaning-preserving translation; A/C/U: unresolved ambiguity/conflict/unavailable target. Atom protocol prose distinguishes these; deviations.md gives the exact stop decision. Native card syntax, IDs, output keys, safe figure layout, key-cardinality assertions and CSV interfaces are T, never added estimators.

Native display identities

The main figures retain native IDs **figure1–figure6**; supplementary B.1–B.23 use **figure7–figure29** in their original order to satisfy the native figure-ID grammar. Their source target identities **figure_b1–figure_b23** remain unchanged in this audit, the inventory, Atom outputs, result slots and renderer input checks. The roster’s **code** and **inputs** fields preserve these bindings. The complete display-to-source crosswalk and PDF output convention are in [figure-interface.md](#). This serialization repair is T (meaning-preserving engineering), not a new scientific choice or a resolution of any source disagreement.

Analysis modules and split rationale

atom-000001 — Bereavement population

This is the source population/comparability question, not a standalone technical linkage experiment. Distributions and related baseline tables belong together.

Question: What population and pre-death characteristics underlie the source comparison of sudden and non-sudden first parental death?

Source targets: **figure_b1**, **figure_b2**, **figure_b3**, **figure_b4**, **table1**, **table_b1**, **table_b2**, **prose_01**, **prose_02**.

Tables: BEF, FAIN, DODSAARS, DODSAASG, IND, IDAP, UDDA.
Field/vintage evidence: schema-evidence.md. Exact unresolved choices: U01, U02, C02, C03, A04, A11.

atom-000002 — Descriptive event paths

Sudden and non-sudden paths answer the same uncontrolled event-time question; their external-validity comparison stays in one module.

Question: How do labor supply, earnings and mental-prescription use evolve around sudden and non-sudden first parental deaths in the source descriptive design?

Source targets: figure_b5, figure_b23, prose_13.

Tables: BEF, FAIN, DODSAARS, DODSAASG, IND, IDAP, UDDA, LMDB.
Field/vintage evidence: schema-evidence.md. Exact unresolved choices: U01, U02, C04, A11.

atom-000003 — Matched labor effects

Main outcomes, parent/child gender contrasts, labor robustness and balance evidence assess one labor-market question. Balance is embedded in this consuming protocol. Figures are not separate experiments.

Question: What are the source matched-DiD effects of sudden first parental death on labor outcomes, including its reported heterogeneity and robustness?

Source targets: figure1, figure2, figure_b6, figure_b7, figure_b8, figure_b19, figure_b20, figure_b21, figure_b22, table_b3, table_b6, prose_03, prose_04, prose_05, prose_06.

Tables: BEF, FAIN, DODSAARS, DODSAASG, IND, IDAP, IDAN, IDAS, UDDA, SYSI, SSSY, LPR_ADM, PSYK_ADM, LMDB. Field/vintage evidence: schema-evidence.md. Exact unresolved choices: U01, U02, C01, A01, A02, A03, C02, C05, A05, U03, C09, A07, C10, A08, A11.

atom-000004 — Health service effects

Health outcomes form the distinct health question, separate from earnings and from co-occurrence. Related services and prescription panels remain in one analysis.

Question: How does parental loss affect the source health-service and prescription outcomes across parent and adult-child gender?

Source targets: figure3, figure_b9, figure_b10, figure_b11, prose_07.

Tables: BEF, FAIN, DODSAARS, DODSAASG, IND, IDAP, IDAN, IDAS, UDDA, LPR_ADM, PSYK_ADM, SYSI, SSSY, LMDB. Field/vintage evidence: schema-evidence.md. Exact unresolved choices: U01, U02, C01, A01, A02, A03, A05, A11.

atom-000005 — Health earnings cooccurrence

Joint onset evidence asks a distinct timing/mechanism question; its related health panels and conditional-share claims stay together.

Question: What source joint timing links earnings declines to the start of health treatment after parental death?

Source targets: figure4, figure_b12, prose_08.

Tables: BEF, FAIN, DODSAARS, DODSAASG, IND, IDAP, IDAN, IDAS, UDDA, LPR_ADM, SYSL, SSSY, LMDB. Field/vintage evidence: schema-evidence.md. Exact unresolved choices: U01, U02, C01, A01, A02, A03, C05, A11.

atom-000006 — Childcare support

Earnings contrasts and formal-care uptake jointly address loss of informal childcare; separate measures/panels are not separate experimental nodes.

Question: Does the source earnings response vary with young children, and does formal childcare use respond to parental loss?

Source targets: figure5, figure_b13, figure_b14.

Tables: BEF, FAIN, DODSAARS, DODSAASG, IND, IDAP, IDAN, IDAS, UDDA, DAGI, BOERNFB, BOERNSB. Field/vintage evidence: schema-evidence.md. Exact unresolved choices: U01, U02, C01, A01, A02, A03, C06, A06, A11.

atom-000007 — Surviving parent care

Post/pre illness contrasts, residence response and source mental-health comparison address the surviving-parent support question and stay within one module.

Question: How do source earnings and proximity responses vary with the surviving parent's health?

Source targets: figure6, figure_b15, figure_b16, prose_09, prose_14.

Tables: BEF, FAIN, DODSAARS, DODSAASG, IND, IDAP, IDAN, IDAS, UDDA, LPR_ADM, LMDB. Field/vintage evidence: schema-evidence.md. Exact unresolved choices: U01, U02, C01, A01, A02, A03, C05, C07, U03, A11.

atom-000008 — Inheritance channel

All parental/own-wealth variants address the inheritance question; they are retained together even when unavailable.

Question: Does the source relationship between pre-death family wealth and earnings responses support its inheritance explanation?

Source targets: table_b4, prose_10, prose_11, prose_12.

Tables: BEF, FAIN, DODSAARS, DODSAASG, IND, IDAP, IDAN, IDAS, UDDA. Field/vintage evidence: schema-evidence.md. Exact unresolved choices: U01, U02, C01, A01, A02, A03, C08, U04, A11.

atom-000009 — Family formation

Cohabitation and fertility are the source family-formation analysis, independently interpretable from labor/health effects.

Question: How does the source parental-death design affect cohabitation and number of children?

Source targets: figure_b17.

Tables: BEF, FAIN, DODSAARS, DODSAASG, IND, IDAP, IDAN, IDAS, UDDA. Field/vintage evidence: schema-evidence.md. Exact unresolved choices: U01, U02, C01, A01, A02, A03, A11.

atom-000010 — Spousal earnings

The spouse outcome and its separately matched population are an independent spillover question; sharing family links does not merge it with adult-child earnings.

Question: What are the source earnings spillovers to adult partners with young children after a first parent-in-law death?

Source targets: figure_b18.

Tables: BEF, FAIN, DODSAARS, DODSAASG, IND, IDAP, IDAN, IDAS, UDDA. Field/vintage evidence: schema-evidence.md. Exact unresolved choices: U01, U02, A01, A02, A03, A09, A11.

atom-000011 — Work choice responses

The three job-choice outcomes form the original preference-related module; routine occupational coding is part of this analysis, not another Atom.

Question: What does the source job-transition analysis show about occupational, firm and sector responses to parental death?

Source targets: table_b5.

Tables: BEF, FAIN, DODSAARS, DODSAASG, IND, IDAP, IDAN, IDAS, UDDA, AKM. Field/vintage evidence: schema-evidence.md. Exact unresolved choices: U01, U02, C01, A01, A02, A03, A10, A11.

Input versus evidence relationships

Population, family linkage, matching and outcome construction are reused inputs. They do not imply that one scientific result licenses another analysis. Every activation is TRUE, with data/source gates inside its own protocol; no outcome-selected execution. Population and descriptive Atoms are retained because they answer independent source questions, not because they validate a data join. Labor evidence is primary; health, joint timing and family-support evidence qualify its explanation. Inheritance, family formation, spouses and work choices retain their independently interpretable questions. Molecule rules require actual component coefficients, intervals and source-based limitations; a complete/partial label alone carries no scientific conclusion.

Fidelity limits

No analysis is repaired to make it executable. Global NN construction ambiguity can block the entire matched family of targets; local panel precision or input gaps remain visible separately. The author programs are not redistributed or executed. No observed sample sizes, estimates, confirmation threshold, power, new controls, significance tests or independent-review claims enter this World.

Source target inventory

Every main and available supplementary display is below in physical source order. Within displays preserve source panel/category/outcome order and every plotted coefficient, printed interval, baseline, N and comparison. Prose-only claims follow in a

separate locator-ordered register. Module assignment is scientific; display order is presentational. All current results are unfilled. A target with unresolved decisions is retained with an explicit no-estimate disposition, not dropped.

Order / target	Source locator	Atom	Full target scope	Disposition
1. table1	Journal p1819, Table1	atom-000001	Sudden/non-sudden baseline characteristics. 18 rows including N; sudden/non-sudden means and source t-test p; Atom1 lists exact row order.	Source-specified scaffold; execute only E/T portions after source/input gates. See module gaps; no current values.
2. figure1	Journal p1823, Figure1	atom-000003	Labor outcomes by adult-child gender. A: Employment [Men, Women]; B: ATP conditional working [Men, Women]; C: Raw earnings [Men, Women]; D: Indexed earnings [Men, Women]	Source-specified scaffold; execute only E/T portions after source/input gates.
3. figure2	Journal p1824, Figure2	atom-000003	Earnings by parent and adult-child gender. A: Indexed earnings [Men, Women]	Source-specified scaffold; execute only E/T portions after source/input gates.

Order / target	Source locator	Atom	Full target scope	Disposition
4. figure3	Journal p1826, Figure3	atom-000004	Health-service and prescription effects. A: Psychologist visits [Men, Women]; B: Any mental prescriptions [Men, Women]; C: Any opioid prescriptions [Men, Women]; D: Hospital visits [Men, Women]	Source-specified scaffold; execute only E/T portions after source/input gates.
5. figure4	Journal p1828, Figure4	atom-000005	Earnings decline and mental-treatment onset. A: Earnings decline of at least 5% [Men, Women]; B: Start of earnings decline [Men, Women]; C: Start of mental prescriptions [Men, Women]; D: Joint mental and earnings onsets [Men, Women]	Source-specified scaffold; execute only E/T portions after source/input gates. Caption Eq4 versus dynamic code Eq2 remains C05.

Order / target	Source locator	Atom	Full target scope	Disposition
6. figure5	Journal p1829, Figure5	atom-000006	Earnings by young-child status. Mother: Mother death [Men, Women]; Father: Father death [Men, Women]	Source-specified scaffold; execute only E/T portions after source/input gates. Eligibility conflict C06 remains unresolved.
7. figure6	Journal p1831, Figure6	atom-000007	Earnings by surviving-parent health. A: Surviving-parent health after death [Men, Women]	Source-specified scaffold; execute only E/T portions after source/input gates. Illness definition, calendar coverage and model conflicts C05/C07.
8. figure_b1	Supplement p7, FigureB1	atom-000001	Cumulative age distribution of first parental death. A: Age at first parental death [All first deaths]	Source-specified scaffold; execute only E/T portions after source/input gates. Source denominator is deaths1985–2014, not an age-specific population risk; A04.

Order / target	Source locator	Atom	Full target scope	Disposition
9. figure_b2	Supplement p8, FigureB2	atom-000001	First-deceased parent gender. A: Parent gender by suddenness [Father, Mother]	Source-specified scaffold; execute only E/T portions after source/input gates.
10. figure_b3	Supplement p9, FigureB3	atom-000001	First-deceased parent age distribution. A: Parent age at first death [Non-sudden, Sudden]	Source-specified scaffold; execute only E/T portions after source/input gates. Source age variable mismatch C03; exact bins A04.
11. figure_b4	Supplement p9, FigureB4	atom-000001	Adult-child age distribution at first parental death. A: Adult age at first death [Non-sudden, Sudden]	Source-specified scaffold; execute only E/T portions after source/input gates. Exact source histogram edges required; A04.

Order / target	Source locator	Atom	Full target scope	Disposition
12. figure_b5	Supplement p10, FigureB5	atom-000002	Descriptive sudden-death paths. A: Employment [Men, Women]; B: ATP conditional working [Men, Women]; C: Raw earnings [Men, Women]; D: Any mental prescriptions [Men, Women]	Source-specified scaffold; execute only E/T portions after source/input gates. Prediction denominator C04.
13. table_b1	Supplement p11, TableB1	atom-000001	Characteristics by parent gender and suddenness. Same18 rows; mother sudden/non-sudden, father sudden/non-sudden.	Source-specified scaffold; execute only E/T portions after source/input gates. See module gaps; no current values.
14. table_b2	Supplement p12, TableB2	atom-000001	Sudden-death characteristics by both genders. Same18 rows; mother-daughters,father-daughters,mother-sons,father-sons.	Source-specified scaffold; execute only E/T portions after source/input gates. See module gaps; no current values.

Order / target	Source locator	Atom	Full target scope	Disposition
15. table_b3	Supplement p13, TableB3	atom-000003	Matched baseline characteristics. A targeted16 rows, B non-targeted9 rows, N; treatment/control. Exact rows in Atom3.	Source-specified scaffold; execute only E/T portions after source/input gates. See module gaps; no current values.
16. figure_b6	Supplement p14, FigureB6	atom-000003	Unconditional and conditional ATP margins. A: ATP including zeros [Men, Women]; B: ATP conditional working [Men, Women]	Source-specified scaffold; execute only E/T portions after source/input gates.

Order / target	Source locator	Atom	Full target scope	Disposition
17. figure_b7	Supplement p15, FigureB7	atom-000003	Dynamic labor effects by parent gender. A_Mother: A. Employment: mother [Men, Women]; A_Father: A. Employment: father [Men, Women]; B_Mother: B. ATP conditional working: mother [Men, Women]; B_Father: B. ATP conditional working: father [Men, Women]; C_Mother: C. Raw earnings: mother [Men, Women]; C_Father: C. Raw earnings: father [Men, Women]; D_Mother: D. Indexed earnings: mother [Men, Women]; D_Father: D. Indexed earnings: father [Men, Women]	Source- specified scaffold; execute only E/T portions after source/input gates. Caption Eq4 versus source Eq2 dynamic coefficients; C05.

Order / target	Source locator	Atom	Full target scope	Disposition
18. figure_b8	Supplement p16, FigureB8	atom-000003	Annual earnings percentile ranks. A: Earnings rank [Men, Women]	Source-specified scaffold; execute only E/T portions after source/input gates.
19. figure_b9	Supplement p16, FigureB9	atom-000004	Alcohol-treatment prescriptions and GP visits. A: Alcohol prescriptions [Men, Women]; B: GP visits [Men, Women]	Source-specified scaffold; execute only E/T portions after source/input gates.

Order / target	Source locator	Atom	Full target scope	Disposition
20. figure_b10	Supplement p17, FigureB10	atom-000004	Dynamic health effects by parent gender. A_Mother: A. Psychologist visits: mother [Men, Women]; A_Father: A. Psychologist visits: father [Men, Women]; B_Mother: B. Any mental prescriptions: mother [Men, Women]; B_Father: B. Any mental prescriptions: father [Men, Women]; C_Mother: C. Any opioid prescriptions: mother [Men, Women]; C_Father: C. Any opioid prescriptions: father [Men, Women]; D_Mother: D. Hospital visits: mother [Men, Women]; D_Father: D. Hospital visits: father [Men, Women]	Source-specified scaffold; execute only E/T portions after source/input gates.

Order / target	Source locator	Atom	Full target scope	Disposition
21. figure_b11	Supplement p18, FigureB11	atom-000004	Screening and nonscreening hospital visits. A: Screening [Men, Women]; B: Nonscreening [Men, Women]	Source-specified scaffold; execute only E/T portions after source/input gates.
22. figure_b12	Supplement p19, FigureB12	atom-000005	Additional health and earnings co-occurrence. A: Psychologist onset [Men, Women]; B: Joint psychologist/earnings onset [Men, Women]; C: Opioid onset [Men, Women]; D: Joint opioid/earnings onset [Men, Women]; E: Hospital visits [Men, Women]; F: Hospital visits $\times$ earnings onset [Men, Women]	Source-specified scaffold; execute only E/T portions after source/input gates. PanelF is hospital count times earnings-decline onset; caption/model conflict C05 remains.
23. figure_b13	Supplement p20, FigureB13	atom-000006	Earnings by youngest-child age. Mother: Mother death [Men, Women]; Father: Father death [Men, Women]	Source-specified scaffold; execute only E/T portions after source/input gates.

Order / target	Source locator	Atom	Full target scope	Disposition
24. figure_b14	Supplement p21, FigureB14	atom-000006	Formal childcare uptake. A: Young-child enrollment [Men, Women]; B: Older-child enrollment [Men, Women]	Source-specified scaffold; execute only E/T portions after source/input gates. Source labels under5/above5, code<=5 and6-10; A06.
25. figure_b15	Supplement p22, FigureB15	atom-000007	Earnings by surviving-parent prior health. A: Surviving-parent health before death [Men, Women]	Source-specified scaffold; execute only E/T portions after source/input gates. C05/C07.
26. figure_b16	Supplement p22, FigureB16	atom-000007	Residence near surviving parent. A: Same 11-region as surviving parent [Men, Women]	Source-specified scaffold; execute only E/T portions after source/input gates. Pre/post illness assignment unresolved C07.
27. figure_b17	Supplement p23, FigureB17	atom-000009	Cohabitation and number of children. A: Cohabitation [Men, Women]; B: Number of children [Men, Women]	Source-specified scaffold; execute only E/T portions after source/input gates.

Order / target	Source locator	Atom	Full target scope	Disposition
28. figure_b18	Supplement p24, FigureB18	atom-000010	Parent-in-law earnings effects. A: Spouse earnings with young children [Men, Women]	Source-specified scaffold; execute only E/T portions after source/input gates. Spouse matching and contrast ambiguities A09.
29. figure_b19	Supplement p25, FigureB19	atom-000003	Earnings by baseline parent proximity. A: Baseline parent/child region [Men, Women]	Source-specified scaffold; execute only E/T portions after source/input gates. Source caption5 regions versus code11; C09.
30. figure_b20	Supplement p26, FigureB20	atom-000003	Earnings by four proximity categories. Men: Men [Estimate]; Women: Women [Estimate]	Source-specified scaffold; execute only E/T portions after source/input gates.

Order / target	Source locator	Atom	Full target scope	Disposition
31. figure_b21	Supplement p27, FigureB21	atom-000003	Matched and future-treated control designs. A: Employment: men [NN, Future-treated]; B: Employment: women [NN, Future-treated]; C: Earnings: men [NN, Future-treated]; D: Earnings: women [NN, Future-treated]	Source-specified scaffold; execute only E/T portions after source/input gates. C10 blocks future-treated window, cluster and earnings tail choices.
32. figure_b22	Supplement p28, FigureB22	atom-000003	Future-treated earnings by parent gender. A: Indexed earnings [Men, Women]	Source-specified scaffold; execute only E/T portions after source/input gates. C10.

Order / target	Source locator	Atom	Full target scope	Disposition
33. figure_b23	Supplement p29, FigureB23	atom-000002	Descriptive non-sudden-death paths. A: Employment [Men, Women]; B: ATP conditional working [Men, Women]; C: Raw earnings [Men, Women]; D: Any mental prescriptions [Men, Women]	Source-specified scaffold; execute only E/T portions after source/input gates. Prediction denominator C04.
34. table_b4	Supplement p30, TableB4	atom-000008	Parental wealth interactions. 4 parent/child-gender columns; wealth interaction, treatment, SE, N, control means both sexes.	Source-specified scaffold; execute only E/T portions after source/input gates. See module gaps; no current values.
35. table_b5	Supplement p30, TableB5	atom-000011	Occupation, firm and sector. 6 columns: occupation men/women, firm men/women, public men/women, conditional employment men/women; coefficient/SE, N, p-value	Source-specified scaffold; execute only E/T portions after source/input gates. See module gaps; no current values.

Order / target	Source locator	Atom	Full target scope	Disposition
36. table_b6	Supplement p31, TableB6	atom-000003	Labor heterogeneity. 8 columns: cause, parent age, adult age, period, each by sex; all source categories, coefficient/SE, No	Source-specified scaffold; execute only E/T portions after source/input gates. See module gaps; current values.

Prose-only and contextual targets

These are retained separately rather than assigning numeric claims to an unrelated figure. Restatements in abstract/introduction/conclusion bind to the same display rows; they are not additional experiments.

Target	Locator	Atom	Claim scope	No-estimate/interpretation boundary
prose_01	Journal p1811, FigureB1 reference	atom-000001	Share first bereaved at19+	A04; exact source denominator and threshold extraction unresolved.
prose_02	Journal II.C pp1818–19	atom-000001	Sudden share and age25–50/<25/>50 population shares	Source prose and matching population denominator; do not assume published shares in runtime.
prose_03	Journal fn14 p1820	atom-000003	Rematching with parental inpatient hospital history	U03; missing inpatient definition/lookback/matching recipe/output set.

Target	Locator	Atom	Claim scope	No-estimate/interpretation boundary
prose_04	Journal fn14 p1820	atom-000003	Second parental death gap>5 and robustness	U03; published91% is reference only; reuse/rematching and full outcome set absent.
prose_05	Journal fn15 p1821	atom-000003	1995+ own/parent health rematching variants	U03; exact one-or-more variable combinations/output set absent.
prose_06	Journal fn16 p1821	atom-000003	Linear age/year robustness	U03/A02; precise source formulation absent.
prose_07	Journal fn22 p1825	atom-000004	Private psychiatrist consultations	A05/U02; combined/private target and crosswalk unresolved.
prose_08	Journal health co-occurrence pp1827–28	atom-000005	Four immediate-year conditional-share comparisons	A03; joint/earnings-onset and joint/mental-onset ratios by sex; year0/1 and denominator availability.
prose_09	Journal V.C p1831	atom-000007	Mental prescriptions by surviving-parent illness	U03/C07; no full source model/program; no-estimate.
prose_10	Supplement A.A p1	atom-000008	Own-asset interaction	U04; timing, scale, level/change and model unresolved.

Target	Locator	Atom	Claim scope	No-estimate/interpretation boundary
prose_11	Supplement A.A fn29 p1	atom-000008	Deceased-parent-only wealth	U04; exact specification absent.
prose_12	Supplement A.A fn29 p1	atom-000008	Surviving-parent-only wealth	U04; exact specification absent.
prose_13	Supplement A.G pp5–6	atom-000002	International generalization/bound argument	Not rebuilt: lower conceptual source discussion, no international analysis in the source; do not invent a test.
prose_14	Journal V.C pp1830–31	atom-000007	Surviving-parent illness share	C07;20%/80% are public reference values, not current subgroup sizes.

Source-code-only material not rebuilt

Source programs also calculate additional transformations (log/IHS earnings or ATP), six/ten income bins unused in final matching, part-time/full-time young-mother outcomes, extra hospital-onset/doubling and joint-any-decline series, nonsource psychiatric aggregates, FN parent-specific employment, and some unpublished joint significance tests. These are not main/supplement targets and do not authorize added analyses. They are retained here as explicit not-rebuilt code outputs, mapped to their consuming labor/health/co-occurrence/childcare modules for audit only. The source program also produces intermediate files and graphs that are not independent scientific targets. No historical deck is an authority for adding them.

Completeness rule

No target-count goal determines the 11-Atom grouping. Table 1 plus Tables B1–B6 = 7 tables; Figures 1–6 plus B1–B23 = 29 figures. All display targets and 14 prose/context target groups have registered membership/disposition. Display targets expand to all coefficients/rows through the Atom protocols and figure interfaces, not only one headline estimate. Printed numbers are never actual-result rows.

Execution supplement

Frozen parental-death reference implementation

This is a pre-data implementation of the existing **11 canonical Atoms**. It computes supported population summaries, shared constructions and numerical model kernels. It does **not** establish an executable, fully resolved matched replication: the frozen source audit explicitly forbids choosing between conflicting matching, scaling, post-period and model definitions. Those target cells stop with `STOP_UNRESOLVED_SOURCE`; independently supported components remain available. See [implementation report](#).

Entry point and environment

Use the controller Python; no installation or environment replacement is needed:

```
PD_PY=/home/shaoerzhuo/anaconda3/bin/python
PD_WORLD=/absolute/path/to/world
PD_ACTION=/absolute/path/to/writable/copied/execution
PD_INPUTS=/absolute/path/to/manifests
PD_OUTPUT=/absolute/path/to/new/output/root
"$PD_PY" "$PD_ACTION/run.py" --help
"$PD_PY" "$PD_ACTION/run.py" --atom atom-000001 --world "$PD_WORLD" --inputs
↪ "$PD_INPUTS/atom-000001.json" --output "$PD_OUTPUT/atom-000001"
↪ --dispatch-policy activation_respecting
```

Copy the complete execution directory, including `frozen-identity.json`, to the action directory. The World may be read-only. Every output directory must be empty/new, outside and not containing the World. No module writes to a canonical Atom, manuscript slot, original source or register input. Python bytecode output is disabled by both entry points.

For governed execution use the independent launcher in the trusted, read-only World:

```
"$PD_PY" "$PD_WORLD/docs/engineering/launch.py" --action "$PD_ACTION" --world
↪ "$PD_WORLD" --atom atom-000001 --inputs "$PD_INPUTS/atom-000001.json"
↪ --output "$PD_OUTPUT/atom-000001"
```

Substitute the Atom ID and its manifest/output in this launcher for every per-Atom command below. `--entry run_tests.py` runs the synthetic suite through the same verification. The World-side `engineering/execution-manifest.json` pins every execution file, including the adjacent scientific lock, CLI, tests and this README; it is separate from immutable scientific registration. The launcher checks the copied bytes **before importing any action code**, including `run.py`. Direct `run.py` also checks before analysis imports, but an arbitrarily replaced `run.py` cannot authenticate itself. The trust root is the protected World launcher/manifest supplied by the operator/controller, not a copied adjacent lock or a receipt's own checksum. The manifest is an author candidate for review, not an assertion of independent acceptance. Do not regenerate it at runtime. Python/library environment and exclusive

control of the action while importing remain operator obligations. Action bytecode caches and symlinked source files are rejected.

Environment: Python 3.12.4, NumPy 1.26.4, pandas 2.2.3, SciPy 1.15.2, PyYAML 6.0.2. Synthetic independent oracles use statsmodels 0.14.4; figure smoke tests use Matplotlib 3.10.0 and the unchanged source renderer. The revision suite exercises CSV, DTA and Parquet (PyArrow 19.0.0), including string codes and rejection of numeric code encodings. See the implementation report for actual run results. Stata is **unavailable**. Rscript exists but no R analysis is used or tested. `source_models.do` has not been executed or numerically validated. No source Stata/ado version is inferred from these Python versions.

Round 3 additionally rejects **all symlink entries**, including directories and dangling links, special files, `.pyc` files and `__pycache__` directories (even empty). A set `PYTHONPYCACHEPREFIX` or `sys.pycache_prefix` (including `-X pycache_prefix=...`) is rejected before action imports. Remove that setting in the governed launch environment; `-B` alone disables only cache writes and is insufficient. The World launcher is loaded directly from its protected source bytes. The code-identity walk enforces the same rules. Both CLI help paths avoid action imports and require no data. Missing/unreadable trusted launch material produces a sanitized `rc=2` operational failure, never a scientific result.

Run the complete synthetic suite with a new absolute scratch destination:

```
"$PD_PY" "$PD_ACTION/run_tests.py" --world "$PD_WORLD" --output  
↪ /absolute/scratch/new-parental-death-test
```

The test fixture seed is 271828. Other fixed seeds in tests specify numerical oracle fixtures only. The original matching seed 123456 is recorded in the source audit; no Python tie-breaking algorithm replaces it.

Explicit runtime input manifest

`--inputs` is one JSON file, not a data directory. Required top-level keys are `kind` (synthetic or governed), `registration_id` and `tables`. Optional keys are `dictionaries` and `evidence`. Unknown options, including scientific override/resolution flags, are rejected. The frozen registration identity is:

```
4116339700f6aeac0c7cb1961b8e90dee01325fcdabfe63599a2d930a2fc436a
```

Each table value is an entry or a list of source-vintage fragments. For example, this is a **format illustration**, not an assertion of available data:

```
{  
  "kind": "governed",  
  "registration_id":  
    ↪ "4116339700f6aeac0c7cb1961b8e90dee01325fcdabfe63599a2d930a2fc436a",  
  "tables": {  
    "BEF": {  
      "path": "/explicit/runtime/extract.csv",  
      "sha256": "<64 lowercase hexadecimal characters>",  
      "vintage": "<actual extract version>",  
    },  
  },  
}
```

```

    "source_locator": "<exact authoritative extract/schema locator>",
    "namespace": "<authenticated common person namespace>",
    "correspondence": "<documented source field/key and vintage
    → correspondence>",
    "fields": {"PNR": "physical_person", "MOR_ID": "physical_mother",
    → "FAR_ID": "physical_father", "FOED_DAG": "physical_birth", "KOEN":
    → "physical_sex"},
    "reference": {"year_field": "physical_year", "quarter_field":
    → "physical_reference_month"},
    "years": [1985],
    "complete_years": [1985],
    "date_formats": {"FOED_DAG": "%Y-%m-%d"}
  }
},
"ictionaries": {},
"evidence": []
}

```

Supply the actual required years; this one-year illustration cannot run the cohort. Annual references use `year_field` or a fixed `year`; population snapshots use `quarter_field` or fixed `quarter` with source month codes 3/6/9/12 (default 12 for annual source files). FAIN's reference year is shifted minus one in construction, never in a second adapter. Death, hospital and UDDA event years come from their explicit dates. `years` identifies extract/reference years; `complete_years` asserts covered event/observation years, which can differ for cumulative extracts. Missing complete coverage blocks the affected construction. Empty covered event cells can be zero; absent/unverified years cannot.

Physical field maps are injective, explicit and limited to the logical names below. CSV reads preserve identifier/code strings and leading zeros; empty strings become missing. Dates require explicit formats unless already typed by DTA/Parquet. Numeric fields are converted explicitly. Differently named person keys require an authenticated governed crosswalk before this adapter; renaming PNR12 to PNR does not establish correspondence. The code checks matching namespace declarations and content hashes, not the truth of an access authorization or metadata assertion.

Every mapped logical field outside `contracts.NUMERIC` and `contracts.DATES` must contain strings or missing values in all formats; the loader normalizes these to pandas string dtype. Numeric codes/IDs in DTA/Parquet are rejected, since leading zeros and original representations cannot be recovered safely. A source-backed lossless format adapter must supply strings explicitly and record that correspondence. No numeric-to-string or zero-padding guess is performed here. Dictionary code columns obey the same rule; documented numeric dictionary keys/values are explicitly converted. Blank strings normalize to missing across all three formats.

Dictionary entries need absolute `path`, `sha256`, `vintage` and exact `source_locator`. Their CSV column names are fixed:

Dictionary key	Required columns / source meaning
inflation	year, inflation_factor_2020; missing author factor series, clean_11/12

Dictionary key	Required columns / source meaning
education_level	HFAUDD, AUDD_LEVEL_L1L4_K; source DISCED level crosswalk, clean_10
education_months	HFAUDD, AUDD_PRIA_L1LX_K; source months crosswalk, clean_10; accepted for future source-established matching, not currently consumed
psychology	SPECIALE, psychology_dum, psychiatry_dum; exact author private-service mapping, clean_3
region11	KOM, KOMLAN_V4_K; source 11-region crosswalk; final geography models blocked
old_region	KOM, REG; checked source old-municipality mapping; final geography models blocked

The three dictionaries generated by `fixtures.make_inputs` contain **invented test-only values**, clearly labelled synthetic. They are never defaults for governed execution. No original scientific conflict is resolved by a manifest. Under frozen governance, establishing a new matching/prediction/model choice requires traceable source clarification before dependent analysis; this invocation does not authorize such an amendment.

Input field provenance

All names derive from [data contract D1–D5](#), [schema evidence](#), and the exact public cleaning programs indexed in [source map](#). These are the existing registered tables; the implementation imports no table whitelist from another study.

Tables	Source fields consumed or bound	Construction / source years
BEF, FAIN	PNR, MOR_ID, FAR_ID, KOEN, ALDER, E_FAELE_ID, CIVST, KOM, REG; BEF FOED_DAG	clean_6/7/13/14; FAIN1981–85 becomes1980–84; BEF1985–2020, annual panel ends2019. Actual FAIN spouse alias AEGTEFAELLE_ID requires explicit correspondence.

Tables	Source fields consumed or bound	Construction / source years
DODSAARS, DODSAASG	PNR; C_LISTE_49,D_DODSDT and C_LISTE49,D_STATDAT	clean_8; source cumulative2001/2019, death coverage1970– 2001/2001–2019; final sudden
IND	source C_SEX PNR, ERHVERVSINDK_GL, ERHVERVSINDK_13, FORM, FORMREST_NY05	codes24,25,26,29,46,47,49 clean_11/12; income GL through1986 then13; wealth FORM1984–97 then
IDAP	FORMREST_NY05 PNR, ATPXX, source ATP_AAR alias	FORMREST_NY05 clean_11;1980–2019; neg- ative/missing0,1980/81 correction1167/432, annualp99
IDAN, IDAS	PNR,TYPE,TYPE_2008,LBNR,ARBASE LBNR,EJERKO	LBNR,ARBASE,SEKTORKODE,SMALT_LOEN_BELOE job/owner1980–2007, late jobs2008–2019; source early type aliases explicitly mapped, no inferred late LBNR
AKM	PNR,BESKST,BESKST13,DISCOLOEN_INDK,DISCOSEL_INDK,DISCO08_LO	self-employment status vintage changes1991, DISCO changes2010; occupations start1991
UDDA	PNR,HFAUDD,HF_VFRA	clean_10; latest completion in year, carry forward; level mapping required
LPR_ADM, PSYK_ADM	PNR,C_ADIAG,D_INDDT	clean_11/2019 unions and admission- year counts;1980–2018 and1995–2018
SYSL, SSSY	PNR,SPECIALE,YDLANT	clean_3/4;1990–2005 and2006–2020; GP exact22 codes; private-service mapping required

Tables	Source fields consumed or bound	Construction / source years
LMDB	PNR,ATC plus explicit extract year	clean_5;1995–2021 source extraction, consuming windows narrower; current PNR12 correspondence unresolved until supplied
DAGI, BOERNFB, BOERNSB	PNR,YDELSE for DAGI; child PNR and year for latter two	clean_9;1995–2014 /2015–2020; final childcare window2005–2019 remains inside source gate

All tables supplied in a particular manifest must belong to that Atom’s exact **registration.tables**, even if another Atom permits them. Input tables can be omitted to obtain component-specific boundaries. Optional missing education, health or inflation never erases independent population/ATP results.

Common output and numerical interfaces

Every dispatched Atom writes **results.csv**, **target-status.csv**, **local-construction.csv**, **interpretation.md** and **receipt.json**. Each owned figure also writes **<source-target>-effects.csv** and **<source-target>-annotations.csv**, including header-only annotation tables where none are registered. No raw identifiers or register rows are exported.

The result key is **target_id,panel,series,x,statistic,term**; remaining columns are **outcome,estimand,sample_id,model_id,estimate,se,lower,upper,ci_method,df,unit,n**. Unavailable/pending rows have null numeric fields. Tables retain source mean/SE/p/N cells only; no source-absent intervals or tests are invented. Figure units, series, categories and event coordinates come from literal frozen renderer CONFIGs. Histogram geometry stays unresolved rather than guessing bins.

models.eq1 implements sex-specific categorical age×year and education×year OLS with HC1. **models.matched** constructs Eq2/3/4 with person×match, event, year and age effects and joint category interactions; Eq3/4 require explicit post times. Its default computable CR0 covariance is marked **CR0_not_source_inference** and cannot produce source intervals. Explicit source-established effective parameter count and residual df can be passed to the low-level kernel; test-supplied CR1 values validate algebra only. Frozen dispatch never bypasses C01/A01/A02/A03. **Fit.contrast** uses the same fitted joint covariance. **source_models.do** provides the actual reghdfe/lincom/test engine interface for later verified Stata execution; no numerical validation of that engine is claimed here. Dense categorical Python kernels are reference/oracle implementations, not a demonstrated national-register-scale engine.

Matched kernels check person-match-year uniqueness, nonmissing structural fields, exact event-year sets and common pair clocks before sex selection or source outcome listwise deletion. Conditional positive ATP may then lose outcome-missing observations; those deletions cannot hide duplicate/missing structural records. Hospital visit counts precede the invalid-interval day filter, exactly as clean_1:51/74. Missing discharge dates stop under A11 instead of selecting a different admission sibling.

`validation.validate_bundle(world, atom, output)` checks the full keyed result/unit/uncertainty roster and exact target-status/effects/annotations projections. `validation.render_checked(world, source_target, tables, pdf, width_mm=130.7, style_dir=None)` is the supported rendering entry: it validates exact annotation tuples and statistic-specific units, then calls the unchanged frozen template. Empty/partial annotations remain valid missing input. Calling a raw frozen renderer directly bypasses this added annotation guard. Known B20 label collisions, unavailable-histogram default axes and oversized boundary figures remain disclosed presentation limits; this adapter changes no layout, labels or bins.

Validation also recomputes data-independent source dispositions using `analyses.source_boundaries`: every module2-11 cell and Atom1's B3/B4/prose_01/02 must retain its exact status, gap IDs, reason and source locator. This includes cell-specific summary/occupation/age/balance additions; re-signing a boundary as "no effect" cannot remove a ledger gate. Pending outer activation is checked separately. Data-dependent Atom1 missingness and numeric truth still require the externally accepted receipt/input binding. A02 is an explicit matched-model target set: retained for linear-controls prose_06 and spouse B18, absent from B3 means, prose_08 ratios and prose_14 prevalence.

The complete suite includes `round2_tests.py` and `round3_tests.py`. Round-3 regressions inject real unchecked-hash bytecode into both `contracts` and `analyses`, exercise symlinked and relocated cache paths through both launch routes, compare both identity walks, check sanitized missing-launcher failure, assert structural numeric dtypes, verify the exact 526-observation partial-outcome sample (missing row0 and age-28 singleton row187), and check gap-set retention/removal over all50 targets. No such injected artifact is placed in the World.

Per-Atom entries

The exact commands below use the path variables above. All use the common outputs and column contracts. The full synthetic command above runs every entry under both dispatch policies and its supported construction/target path.

For Atoms3-11, "registered cohort/labor/education" means BEF,FAIN,DODSAARS,DODSAASG,IND,IDAP,IDAN,IDAS,UDDA; those are the shared registered matching inputs, even where unresolved matching prevents consumption of job fields. Additional outcome tables are named in each entry. The runtime checks the actual immutable per-Atom table list.

atom-000001 — Population and baseline descriptions

Command: `"$PD_PY" "$PD_ACTION/run.py" --atom atom-000001 --world "$PD_WORLD" --inputs "$PD_INPUTS/atom-000001.json" --output "$PD_OUTPUT/atom-000001".`

Implementation: analyses.population_results, summaries; construction.population, deaths, source_panel, family_counts, education, labour, atp; models.pooled_ttest. Required cohort tables: BEF,FAIN,DODSAARS,DODSAASG; optional row-specific IND,IDAP,UDDA and inflation/education dictionaries, with fields in the provenance table. B1 alone needs BEF1985–2014 and the death extracts; it does not require the balanced cohort inputs.

Targets: figure_b1/b2/b3/b4, table1/table_b1/table_b2, prose_01/02. Outputs use cumulative percent, parent-gender percent, density, source proportions/years/children/ATP/1000 real2020DKK/person counts and p probabilities. B1 denominator precedes count<5 suppression; bin8's conflicting40–45 label remains unavailable. B2 and supported baseline rows are estimated. C02 blocks only under-six rows; C03/A04 block histograms/exact prose shares. Tests: hand cumulative20/52/100 oracle, age25/50 and1985/2014 limits, exact death dates, second death, table t oracle, missing education/inflation, all missing inputs and renderer interfaces.

atom-000002 — Descriptive event paths

Command: "\$PD_PY" "\$PD_ACTION/run.py" --atom atom-000002 --world "\$PD_WORLD" --inputs "\$PD_INPUTS/atom-000002.json" --output "\$PD_OUTPUT/atom-000002".

Implementation: analyses.atom_000002, shared cohort/labor construction, construction.prescriptions, models.eq1, scaled_eq1. Inputs: BEF,FAIN,DODSAARS,DODSAASG,IND,IDAP,UDDA,LMDB; inflation/education dictionaries. Targets: figure_b5/b23 (percent of counterfactual mean), prose_13 (contextual nonempirical argument). C04 blocks the plotted scaling; unscaled Eq1 coefficients are not substituted. Tests: dense QR/statsmodels HC1 oracle, no missing-input replacement and both eight-panel interfaces. The CLI constructs inputs but does not fit the conflicting display.

atom-000003 — Matched labor effects, balance and robustness

Command: "\$PD_PY" "\$PD_ACTION/run.py" --atom atom-000003 --world "\$PD_WORLD" --inputs "\$PD_INPUTS/atom-000003.json" --output "\$PD_OUTPUT/atom-000003".

Implementation: analyses.atom_000003, labour_dynamics, missing_nonfigure; construction.candidate_controls, index_earnings, propagate_treated, proximity, labour, jobs; Eq2/3/4 kernels. Inputs: BEF,FAIN,DODSAARS,DODSAASG,IND,IDAP,IDAN,IDAS,UDDA,SYSI,SSSY,LPR_ADM,PSYK_ADM plus consuming dictionaries.

Targets: figure1/2, figure_b6/b7/b8/b19/b20/b21/b22, table_b3/b6, prose_03/04/05/06. Units: probability, ATP contributions,1000 real2020DKK,index points at baseline100,percentile ranks,source baseline characteristic units. Shared C01/A01/A02 blocks NN sample-dependent outputs. A03 affects pooled summaries; C05 B7; C09 B19; A07 age strata; C10/A08 future-treated; U03 prose variants. Cause/calendar categories remain explicit. TableB3's literal named/source rows total15 targeted although inventory says16: see report; no invented sixteenth row. Tests: tied rank0/50/50/100, singleton undefined rank, control death>D+5, full joint cluster oracle, both-sex labor paths, all source panels and absent inputs. Future-treated disputed choices are not implemented.

atom-000004 — Health services and prescriptions

Command: `"$PD_PY" "$PD_ACTION/run.py" --atom atom-000004
--world "$PD_WORLD" --inputs "$PD_INPUTS/atom-000004.json" --output
"$PD_OUTPUT/atom-000004".`

Implementation: `analyses.atom_000004, local_health, health_dynamics;
construction.hospital, prescriptions, services, covered_join;` joint-
parent Eq2 and Eq4 kernels. Inputs: registered cohort/labor/education tables plus
LPR_ADM,PSYK_ADM,SYSI,SSSY,LMDB and psychology mapping. Targets:
figure3,figure_b9/b10/b11,prose_07. Units: visits/year and probability; B9 alcohol is
prescription **count**. C01/A01/A02/A03 and private-versus-combined psychiatry A05
apply. Tests independently check inclusive overlap/cross-year union, latest-discharge
diagnosis tie, invalid dates, exact ATC counts, negative services after summing,
unavailable2019 versus covered absence, outcome-specific missingness and every
display.

atom-000005 — Earnings/health co-occurrence

Command: `"$PD_PY" "$PD_ACTION/run.py" --atom atom-000005
--world "$PD_WORLD" --inputs "$PD_INPUTS/atom-000005.json" --output
"$PD_OUTPUT/atom-000005".`

Implementation: `analyses.atom_000005; construction.onset_panel;` Eq2 ker-
nel with explicit event coordinates; `models.immediate_ratios`. Inputs: regis-
tered cohort/labor/education,LMDB,LPR_ADM,SYSI,SSSY and relevant dictio-
naries. Source matched history requires earnings `t=-6`, balance1995-2018 and
final `t=-4..5`. Targets: figure4,figure_b12,prose_08, probability/count and ratios.
C01/A01/A02/C05 block disputed figures; A03 blocks year0/1 prose ratios. No
mediation estimate is added. Tests: exact5% boundary, lag versus event clock,
zero denominator, persistent versus new declines, partial opioid data and hospital-
count×earnings-onset.

atom-000006 — Young children and formal childcare

Command: `"$PD_PY" "$PD_ACTION/run.py" --atom atom-000006
--world "$PD_WORLD" --inputs "$PD_INPUTS/atom-000006.json" --output
"$PD_OUTPUT/atom-000006".`

Implementation: `analyses.atom_000006; construction.birth_links,
family_counts, enrollment, index_earnings;` Eq4 interactions. Inputs: reg-
istered cohort/labor/education and DAGI,BOERNFB,BOERNSB. Targets:
figure5,figure_b13/b14; earnings index points/probability. C06 blocks childless eli-
gibility in F5/B13; A06 blocks age5/youngest-tie enrollment; shared matching/post
gates remain. Child enrollment indicators and counts are computed locally. Tests:
exact twin counts with ambiguous child identity, no-child sentinel22 versus age>20
sentinel21, and each missing/partial/full interface.

atom-000007 — Surviving-parent illness and proximity

Command: `"$PD_PY" "$PD_ACTION/run.py" --atom atom-000007
--world "$PD_WORLD" --inputs "$PD_INPUTS/atom-000007.json" --output
"$PD_OUTPUT/atom-000007".`

Implementation: `analyses.atom_000007;` inclusive `construction.hospital,`
`prescription/cohort construction` and `proximity` kernel. Inputs: registered

cohort/labor/education,LPR_ADM,LMDB; geography dictionaries for any future source-established proximity output. Targets: figure6,figure_b15/b16,prose_09/14; index points, probability and illness proportion. C07 prevents selecting annual $\geq$ 90 versus cumulative $\geq$ three-month illness, year0/2019 handling and B16 pre/post split; C05/U03 also apply. No illness stratum is guessed. Tests: union/calendar coverage, partial health inputs, geographic missingness, all three figure interfaces and both dispatch policies.

atom-000008 — Wealth interaction

Command: `"$PD_PY" "$PD_ACTION/run.py" --atom atom-000008 --world "$PD_WORLD" --inputs "$PD_INPUTS/atom-000008.json" --output "$PD_OUTPUT/atom-000008".`

Implementation: `analyses.atom_000008, missing_nonfigure; construction.wealth` and shared cohort. Inputs: registered cohort/labor/education and IND FORM/FORMREST_NY05, source inflation. Targets: table_b4 and prose_10/11/12. Annual wealth is real2020DKK locally; table interaction units remain explicitly unresolved (C08), not silently divided1000/10000. U04 blocks the three unspecified variants; matching/post gates remain. Tests: independent maximum $30 \times 2 = 60$ and next-vintage $40 \times 3 = 120$, mixed missing/nonmissing duplicate boundary, all table columns and absent-input dispatch.

atom-000009 — Family responses

Command: `"$PD_PY" "$PD_ACTION/run.py" --atom atom-000009 --world "$PD_WORLD" --inputs "$PD_INPUTS/atom-000009.json" --output "$PD_OUTPUT/atom-000009".`

Implementation: `analyses.atom_000009, family_models; family/cohabitation construction; joint-parent Eq4.` Inputs: BEF,FAIN,DODSAARS,DODSAASG,IND,IDAP,IDAN,IDAS,UDDA and source income/education dictionaries as registered. Target figure_b17: probability and children. No descriptive cap5 is applied to the family-response child count. C01/A01/A02/A03 prevent a matched estimate. Tests: family arithmetic, Eq4 kernel with explicitly test-supplied post times, missing post-time gate, both panels and comparison cells.

atom-000010 — Parent-in-law earnings

Command: `"$PD_PY" "$PD_ACTION/run.py" --atom atom-000010 --world "$PD_WORLD" --inputs "$PD_INPUTS/atom-000010.json" --output "$PD_OUTPUT/atom-000010".`

Implementation: `analyses.atom_000010; construction.spouse_histories` fixes event partner before spouse-year expansion; shared demographic/income functions and spouse-key-capable model kernel. Inputs: registered cohort/labor/education,IDAN,IDAS and dictionaries. Target figure_b18, index points. Separate spouse matching is not replaced by biological-child pairs. A09/A01/A02/A03 block presence-versus-survival, tenure/order, missing duplicate selection and pooled contrast correspondence. Tests: actual local spouse-history construction under both dispatches, cardinality validation and figure interface. No spouse matched-edge or source-engine inference validation is claimed.

atom-000011 — Occupation, firm and sector

Command: `"$PD_PY" "$PD_ACTION/run.py" --atom atom-000011
--world "$PD_WORLD" --inputs "$PD_INPUTS/atom-000011.json" --output
"$PD_OUTPUT/atom-000011".`

Implementation: `analyses.atom_000011, job_models; construction.occupation_codes,`
`jobs; Eq3.` Inputs: registered cohort/labor/education,AKM,IDAN,IDAS. Target
table `_b5` retains all 6 columns and coefficient/SE/N/premean cells,
probability/person-year counts. A10 stops occupation transitions whose eight-
prior-record order is unspecified; independently computable occupation codes and
firm/public construction survive. C01/A01/A02/A03 block matched estimates.
Tests: adjacent versus gapped employer years, public-code exclusions, main-job tie,
supported firm/public kernels versus blocked occupation and complete table interface.

Activation, evidence and portability

The native selected semantics are canonical, tri-valued Kleene logic; `activation.py`
is a meaning-preserving adaptation of `prereg-lane/native.py`. All eleven
frozen activations are TRUE, so no predecessor is required. Tests exercise
unknown/false/pending/no_result logic without editing the frozen activation. `direct`
bypasses only outer activation; it cannot resolve any target gate or retry an explicitly
terminal no_result Atom.

An optional evidence entry is `{"path":"/absolute/receipt.json","accepted":true,"input_id":"<externally
expected input identity>","implementation_id":"<trusted World manifest
implementation_id>","receipt_sha256":"<externally accepted receipt
digest>"}`. The accepting operator/controller supplies these expected values inde-
pendently of the received receipt; do not recalculate them from untrusted predecessor
files. The loader compares both expected hashes, frozen material and Atom protocol
identity, kind, input binding and every required hashed output. It then validates the
exact keyed vector, units, uncertainty and all derived CSV projections. Re-signing
a changed receipt cannot satisfy the externally pinned digest; re-signing a trun-
cated vector also fails structural validation. Pending/failure receipts cannot imply
`done()`. Local construction never supplies accepted predecessor evidence. Earlier exe-
cution versions are not automatically accepted: they require an explicitly reviewed
implementation binding; mutable realized Atom/Molecule text does not change the
existing valid binding.

Frozen identity hashes preregistration and membership separately from mutable
Atom result/front matter/body and Molecule interpretation. Frozen material content
is pinned; changing a protocol or source interface fails verification. Runtime locations
are excluded from scientific input identity, so copying a manifest's bound files and
changing only paths preserves identity. Evidence files and mutable result bodies are not
hashed as frozen registration. Each receipt records implementation hash, source/input
hash, dispatch policy, software and aggregate-output hashes.

The test runner writes synthetic inputs, receipts, numerical logs and all plot PDFs
only beneath its requested scratch output. It tests copied action code against a distinct
read-only World and changed realized result bodies, as well as protocol tampering,
corrupt evidence and wrong input bindings. Figures remain the existing templates;

the report records any rendering limitations. No synthetic preview is bound to a manuscript slot.

Engineering supplement

Implementation report — author round 3

Round 3 fixes all six supplied round-2 Opus findings without changing the frozen scientific package. The eleven exact executable entries, input provenance, output columns/units and source limits remain in [execution/README.md](#). No new scientific choice, Atom, observed result or independent acceptance is claimed. Earlier round responses below remain the record for every round-1 finding; this section supersedes their implementation-identity/cache assertions and loose sample-size assertion.

Responses to every round-2 reviewer finding

Finding	Author response	Regression evidence
OPUS-R02-01 — bytecode outside the hashed tree	<code>engineering/launch.py</code> and <code>contracts.code_identity</code> now use recursive <code>os.scandir</code> with no-follow stat checks. All symlinks (including directories/dangling links), special files, <code>.pyc</code> files and <code>__pycache__</code> directories are rejected. A set <code>PYTHONPYCACHEPREFIX</code> or <code>sys.pycache_prefix</code> is rejected before action imports in the trusted launcher and both direct entrypoints. The World launcher itself is compiled from protected source bytes, bypassing World-side cache loading.	<code>round3_tests</code> compiles actual unchecked-hash bytecode injecting a marker and the review's doubled B1 calculation, without altering the copied source. Tests cover both <code>contracts</code> and <code>analyses</code> , symlink caches, environment prefixes, <code>-X</code> <code>pycache_prefix</code> , both launch routes, directory/file/dangling symlinks, FIFO, empty cache and ordinary <code>pyc</code> . Both identity walks are checked independently; marker absence and no output directory prove rejection precedes action import.

Finding	Author response	Regression evidence
OPUS-R02-02 — boundary text/gaps/source locator can be relabelled a null	Extracted <code>analyses.source_boundaries</code> which recomputes <code>target_boundary</code> and all cell-specific additions without input access. <code>validation.validate_results</code> compares exact sta- tus/gap_ids/reason/source for every module2–11 cell and Atom1’s wholly source-blocked B3/B4/prose_01/02. Nonactivation has a separate expected disposition.	Per-field alterations rejects all ten blocked modules are rejected. A fully rewritten, consistently projected and re-signed synthetic truth. He asserting treatment/control equivalence is rejected even when its new checksum is externally rebound. Shuffled rows, legitimate optional-input loss, empty inputs and nonactivation remain valid. Data-dependent Atom1 explanations and numeric truth still require the externally accepted receipt/input binding.
OPUS-R02-03 — A02 over-attached to prose_08/prose_14	Added the explicit A02_TARGETS set. Removed A02 only from the two flagged ratio/share targets; retained it for matched effects, prose_06 linear-controls robustness and B18 spouse effects. Neither target becomes estimable: prose_08 retains C01/A01/A03; prose_14 retains C01/A01/C07.	Literal expected gap-set unions over all50 actual CLI targets test both removal and retention, plus cell-level A03 dynamic-versus-summary scope. Existing C02/A05/A07/A10 row gates remain.
OPUS-R02-04 — structural object/string column raises TypeError	<code>models.matched</code> checks <code>is_numeric_dtype</code> before calling <code>isfinite</code> or arithmetic and raises an A11 Boundary.	String-valued, object- valued and invalid-string year/time/age/treatment columns are rejected before missing-outcome/sex deletion. No coercion or source-silent value repair.

Finding	Author response	Regression evidence
OPUS-R02-05 — missing trusted launcher leaks traceback	Moved the entire bootstrap loader creation/read/compile/verification inside its sanitized failure handler. The callable trust check likewise converts unavailable trusted material to an IntegrityError.	A separate scratch World with <code>launch.py</code> removed returns rc=2 JSON, no traceback, no absolute path, and no output directory.
OPUS-R02-06 — partial-outcome sample assertion too loose	Replaced the loose positive/less-than checks with the exact source singleton sample: 528 minus row0's missing outcome and row187's unique age28 cell equals526.	The retained round2 regression now asserts the exact dropped indices {0,187} and sample loss2, alongside the independent singleton cascade and numerical model tests.

The trusted launcher/manifest remain a controller/operator-protected root. The copied action must not be concurrently modified during execution, and the Python/library environment must be trusted. These checks do not authenticate operator metadata declarations or reconstruct a statistical value without inputs. They prevent the concrete source-versus-bytecode and boundary-relabel counterexamples under the documented trust boundary. No manifest is regenerated at runtime.

Round-3 execution, coverage and limitations

The complete round-3 suite passed on 2026-09-06: **56 tests, zero failures/errors, 145 successful figure renders, no figure exception, process exit 0**. The numerical/construction/integration and reviewer-regression phase took 289.277 seconds; figure rendering followed separately. This was the only complete round-3 suite run. The source-supported estimates and scientific boundaries remain unchanged; the coverage and source-locator tables in the historical record below still apply.

Executed commands (use new output directories when reproducing):

```
PD_WORLD=/home/shaoerzhuo/Bolero-denmark/.prereg-runs/denmark-reference-code-
↪ 20260906/parental-death/v003/world
PD_SCRATCH=/home/shaoerzhuo/Bolero-denmark/.prereg-runs/denmark-reference-cod
↪ e-20260906/parental-death/scratch/r03/author
PD_SOURCE=/home/shaoerzhuo/Bolero-denmark/.prereg-runs/denmark-reference-code
↪ -20260906/parental-death/.controller/source
OPENBLAS_NUM_THREADS=1 /home/shaoerzhuo/anaconda3/bin/python -B
↪ "$PD_WORLD/prereg/engineering/launch.py" --world "$PD_WORLD" --action
↪ "$PD_WORLD/prereg/execution" --entry run_tests.py --output
↪ "$PD_SCRATCH/suite-01"
/home/shaoerzhuo/anaconda3/bin/python -B "$PD_WORLD/prereg/execution/run.py"
↪ --help
```

```

/home/shaoerzhuo/anaconda3/bin/python -B
↪ "$PD_WORLD/prereg/execution/run_tests.py" --help
OPENBLAS_NUM_THREADS=1 MPLCONFIGDIR="$PD_SCRATCH/inspection-matplotlib-cache"
↪ /home/shaoerzhuo/anaconda3/bin/python -B "$PD_SCRATCH/inspect_outputs.py"
↪ --world "$PD_WORLD" --source "$PD_SOURCE" --suite "$PD_SCRATCH/suite-01"
↪ --output "$PD_SCRATCH/inspection-02"

```

The suite's `test-log.txt` and `test-summary.json` retain the test evidence. It runs all eleven actual Atom commands under both dispatch policies, all eleven absent-input paths, supported local constructions and independent numerical oracles. It retains partial outcomes, native tri-valued activation, explicit `no_result` nonretry, read-only relocated Worlds with copied writable actions, realized-body mutation, protocol/input/evidence tampering and every new reviewer counterexample. Both help commands passed without data or World arguments.

Final author artifact inspection is recorded in `scratch/r03/author/inspection-02/audit.json`. It rechecks all eleven direct output bundles, their receipt/output hashes and exact projection schemas; confirms byte-identical result tables across the two dispatch policies; checks all 63 manuscript slot bindings and all 29 figure input bindings; and compares **all 122 original files byte-for-byte with zero changes, missing files or unexpected files outside the authorized execution/engineering trees**. The frozen Atom results and Molecule interpretation remain unmodified. Synthetic tables and previews are only in author scratch.

The current author-candidate implementation identity is `4443f312bb9ad04432a50618d126cba07e711a35b391b31a254d9d9f32b5369a`; every baseline direct receipt matches the final execution manifest. The scientific registration identity remains `4116339700f6aeac0c7cb1961b8e90dee01325fcdabfe63599a2d930a2fc436a`. No execution file changed after the successful suite.

Coverage inspection confirms 1,939 keyed synthetic result cells across 50 targets. Atom1 computes 140 of its 223 cells and remains partial. Atoms2–11 keep their source boundaries and produce no numerical target cells. Five targets contain computed components (B2 complete; B1 and Tables1/B1/B2 partial); 45 contain no numbers, including the contextual non-estimation target `prose_13`. These are engineering fixture counts, not Danish study outcomes.

All 145 PDFs are single-page, vector outputs of width 130.700 mm within 0.001 mm, using the tested DejaVu Sans fallback. They comprise 29 actual synthetic CLI projections and 116 deliberately fabricated interface-only fixtures for full, partial, missing and nonactivated dispositions; none fill a canonical result or manuscript slot. Contact sheets of all 29 actual outputs and enlarged B1/B20/B4/partial-F4 previews were visually inspected. The first inspection's restricted word check found two B20 overlaps; the expanded whole-label-span check confirms six collisions, three in each panel. This was a scratch inspection refinement, requiring no execution-code change or analysis rerun. B3/B4 still show default 0–1 axes without resolved age bins. Actual boundary F3/B7/B10/B12 heights remain 320.548/388.112/385.267/391.312 mm.

Versions confirmed in the controller environment: Python 3.12.4, NumPy 1.26.4, pandas 2.2.3, SciPy 1.15.2, PyYAML 6.0.2, statsmodels 0.14.4, Matplotlib 3.10.0 and PyArrow 19.0.0. No environment installation, commit, push, release, confidential-data read, remote execution or write outside the authorized trees was performed.

No additional Opus or Codex Sol xhigh agent was started. This invocation is author-only and explicitly prohibits spawning agents or writing reviewer outputs. The supplied reviews are addressed here; new independent checks remain for the controller's reviewer stage. There is no round-2 Codex review in the supplied directory.

Stata, `reghdfe` and `teffects` remain unavailable, with no environment installation attempted. The source Stata kernel has not been numerically executed. The matched Python kernels exercise only source-supported algebra with explicit test metadata; they are never promoted across unresolved matching, normalization or post-period gates. Public dictionaries, governed field/vintage/namespace truth, national-scale performance and governed font binding remain unvalidated.

The frozen Figure B20 label collision, B3/B4 default numeric axes without resolved bins, four over-height boundary figures and Table B3's 16-versus-15 targeted-row discrepancy remain disclosed. User scope forbids modifying their frozen templates/inventory. No numeric value, label or bin was fabricated to disguise them. Passing a synthetic suite is not empirical replication, causal validation, data availability or publication-layout approval.

Round-2 author record (historical)

This revision responds to both independent round-1 engineering reviews while preserving all frozen scientific material. The executable coverage remains the eleven entries documented in [execution/README.md](#): supported descriptive cells and source construction are implemented; 45 targets remain without numerical output because of source boundaries or registered non-estimation scope. There is no scientific amendment, new Atom, selected scientific default or newly observed result. The implementation/report below is an author response, not a second independent review.

Responses to every reviewer finding

Finding	Response and implementing source	Verification / remaining limit
R01-CODEX-01 — copied implementation and receipt identity	Added trusted World-side <code>engineering/launch.py</code> and <code>execution-manifest.json</code> the launcher compares every copied execution file before importing <code>run.py</code> or analysis code. <code>contracts.verify_world</code> requires this binding; the adjacent lock alone is insufficient. <code>run.load_evidence</code> requires externally accepted receipt digest and implementation ID matching the World manifest.	Tests alter <code>analyses.py</code> , <code>run.py</code> and <code>frozen-identity.json</code> independently and check rejection before an injected marker can execute. Tests reject missing expected receipt hashes, re-signed receipts and changed implementation IDs. The controller/operator must protect and accept the World launcher/manifest as its trust root. The author manifest is a review candidate, not a cryptographic claim of reviewer acceptance. An attacker allowed to replace both trust root and executable is outside this boundary. No change was made to the original manifest/CHECKSUMS.
R01-CODEX-02 — truncated predecessor vector	Added <code>validation.validate_results/unit</code> exact cell-key/unit roster; source uncertainty dispositions; recomputed status counts/reasons; exact effects/annotation projections; all projection files must be hash-bound. Validation runs on new exports and consumed evidence.	Tests delete a B1 row, <code>results/unit/uncertainty</code> , stale status/effects/annotations and omit a projection hash, starting from genuine synthetic CLI directories. Re-signing these fixtures does not authorize an invalid vector.

Finding	Response and implementing source	Verification / remaining limit
R01-CODEX-03 — outcome deletion hides duplicate structural keys	<code>models.matched</code> validates unique person-match-year, nonmissing structural fields, integral clocks, exact registered event-year sets, treatment constancy and common match event year before outcome/sex selection. It then uses the source regression's listwise missing-outcome rule (including conditional ATP).	Duplicate with one missing outcome, missing ID/match/year/time/age/treatment, absent year and inconsistent clock cases stop. A structurally complete panel with one missing outcome still fits the supported algebra. Existing clustered/dense-OLS oracles remain.
R01-CODEX-04 — unrestricted annotation terms/units	<code>validation.validate_annotations</code> binds exact target/panel/series/statistic/tuples and statistic-specific units before invoking the unchanged frozen renderer. Empty/partial annotation tables remain permitted missing input. All suite rendering now uses this guard.	<code>Annotations.render_checked</code> unregistered terms/term/panel/units and stale annotation projection. Mitigated at the authorized execution interface: the raw immutable <code>paper/figures/source_render.py</code> still lacks this check. Directly calling it bypasses the guard; the source owner must adopt this execution adapter. No template bytes were changed.
R01-CODEX-05 — A02 incorrectly assigned to B3 means	Split C01/A01 sample gates from A02 estimator normalization in <code>analyses.target_boundary</code> . Table B3 receives C01/A01, plus only its own under-six/psychiatry row gaps.	Exact gap-ID assertions verify B3 excludes A02 while matched effects obtain it. Availability is unchanged because B3 still needs an unresolved sample.

Finding	Response and implementing source	Verification / remaining limit
R01-CODEX-06 — B20 collision and empty histogram axes	Retained the exact frozen template; no replacement label, bin or axis was fabricated. The output interfaces are exercised and the layout defect is explicitly retained for the presentation owner.	Unresolved within author write scope. User explicitly forbids changing figure templates. B20 labels collide at 130.7 mm; B3/B4 missing-bin figures retain a default numeric age axis. Expanded boundary F3/B7/B10/B12 also exceed ordinary page height. Passing interface/PDF checks is not passing publication layout.
OPUS-R01-01 — invalid-interval admissions disappear from counts	<code>construction.hospital</code> now selects latest discharge and counts admission-year visits first, then removes reversed intervals only for day totals.	<code>clean_1_hospital_visits.do:20,51,74</code> is the precise order. Independent cases expect one visit/zero days for single and multiple reversed intervals; cross-year union oracle checks 2017/2018 totals and 2019 exclusion. D4's short summary about invalid dates is read in conjunction with its incorporated exact consuming source; no clinical definition changes.
OPUS-R01-02 — missing discharge silently selects another row	<code>construction.hospital</code> stops with A11 before dropping a missing-discharge record or selecting its sibling.	The two-record counterexample raises Hospital record without source discharge date ; both somatic and psychiatric paths tested. No missing-date imputation or replacement diagnosis.

Finding	Response and implementing source	Verification / remaining limit
OPUS-R01-03 — CSV/DTA/Parquet code dtype divergence	<code>contracts.source_strings</code> requires nonmissing logical codes/IDs to be strings, normalizes blank strings and pandas string dtype in every format. Numeric representations are rejected because lost leading zeros cannot be inferred. Dictionary code fields receive the same check; documented numeric dictionary values are explicit.	Valid string AKM, IDAN/IDAS and SYSI fixtures run in all three formats against hand-computable occupation/firm/public/GP oracles. Numeric occupation sentinel, service specialty, ownership and person/civil-status inputs stop for DTA and Parquet. No guessed field width or classification version.
OPUS-R01-04 — repeated STOP prefix	Removed STOP prefix from the shared NN reason; <code>target_boundary</code> adds it once.	String assertion checks a single prefix. No scientific gate changes.
OPUS-R01-05 — Table B3 row order	Emits employment, unconditional ATP, earnings, public in the named source order.	Independent literal 25-row list (15 targeted + 9 nontargeted + N), each with treatment/control, matches analysis 15:114 . Frozen inventory’s contradictory count of 16 targeted rows remains unresolved; no row is invented.

Round-2 validation and limitations

The reproducible entry is `/home/shaoerzhuo/anaconda3/bin/python <World>/prereg/engineering/launch.py --world <World> --action <copied-execution> --entry run_tests.py --output <absolute-author-scratch>`. Tests generate their own synthetic inputs/results outside the final World. Current round scratch is `parental-death/scratch/r02/author`. The final run passed **49 tests and 145 figure renders**, with zero failures/errors. The material audit found **122 original files unchanged** and no unexpected World files outside execution/engineering.

The candidate execution manifest is engineering identity only. Immutable registration/material identity remains independent of mutable Atom result bodies and Molecule current interpretation. Read-only relocation, accepted old receipts after

realized-body edits, protocol tampering and both dispatch policies remain covered. All actual activations are TRUE; unknown/false branches are native evaluator/inactive-export fixtures and do not imply changed scientific activation. No essential missing input is represented as an accepted predecessor or statistical null; no explicit `no_result` is retried.

The added guards introduce no statistical experiment. Source conflicts C01–C10/A01–A11/U01–U04 remain at the exact targets/locators documented below. Source Stata, ado versions, original reghdfe inference and matching/ties remain unavailable; Python algebra tests do not validate them. No confidential data, old empirical sessions, national-scale performance, actual runtime metadata truth or governed font binding were accessed/validated. No software was installed. No subagents or new independent reviewers were run because this invocation is author-only and prohibits spawning agents.

Final round-2 commands and evidence — 2026-09-06

The final suite was actually executed with:

```
PD_WORLD=/home/shaoerzhuo/Bolero-denmark/.prereg-runs/denmark-reference-code-
↳ 20260906/parental-death/v002/world
PD_SCRATCH=/home/shaoerzhuo/Bolero-denmark/.prereg-runs/denmark-reference-cod
↳ e-20260906/parental-death/scratch/r02/author
OPENBLAS_NUM_THREADS=1 /home/shaoerzhuo/anaconda3/bin/python -B
↳ "$PD_WORLD/prereg/engineering/launch.py" --world "$PD_WORLD" --action
↳ "$PD_WORLD/prereg/execution" --entry run_tests.py --output
↳ "$PD_SCRATCH/suite-02"
OPENBLAS_NUM_THREADS=1 MPLCONFIGDIR="$PD_SCRATCH/inspection-matplotlib-cache"
↳ /home/shaoerzhuo/anaconda3/bin/python -B "$PD_SCRATCH/inspect_outputs.py"
↳ --suite "$PD_SCRATCH/suite-02" --output "$PD_SCRATCH/final-inspection-02"
```

Use fresh output directories when reproducing these commands. The suite itself copies execution code to writable action directories and relocates read-only Worlds; the main suite command does not require the World to be writable. `run.py --help` also passed without inputs.

- First round-2 run, **suite-01**: 49 tests, one assertion failure, zero errors; figure rendering was correctly skipped after that failure. The new partial-outcome test incorrectly expected exactly one deleted row; source iterated singleton removal deletes an additional observation. Corrected that test expectation while retaining the independent singleton-rule and numerical covariance oracles. Numerical/integration runtime was 260.040 seconds.
- Final run, **suite-02/test-log.txt** and **test-summary.json**: 49 tests passed in 255.272 seconds, zero failures/errors; all 145 subsequent renders passed, with no figure exception. This includes all eleven Atoms under both policies, eleven absent-input runs, partial outcome paths, native tri-valued activation cases, explicit `no_result` nonretry, mutable realized bodies, protocol/input tampering, relocated read-only Worlds, copied-code mutations and all new reviewer counterexamples.
- Every frozen figure was rendered for five dispositions: full **interface-only invented** values, actual synthetic CLI exports, missing input, partial input and nonactivation.

Guard rejection of wrong units, missing required CIs and unregistered annotation keys/units also passed. The invented complete vectors are renderer tests, not estimates of blocked scientific targets.

- `final-inspection-02/audit.json` revalidates all eleven complete result bundles, records dimensions for all 145 PDFs, checks the original 122 files byte-for-byte and finds zero unexpected World files outside the allowed directories. The inspection helper’s first run produced the previews but failed to serialize a NumPy integer collision count; converting that scratch-only count to a Python integer fixed the helper. The successful second inspection required no execution-code change or statistical rerun.
- Current implementation identity is `83a78afd539e026c621e6ad47393dc9f2d2278cf924c41d4390bbf33901` it matches the separately supplied World engineering manifest and the final synthetic receipts. Registration identity remains `4116339700f6aeac0c7cb1961b8e90dee01325fcdabfe63599a2d930a2fc436a`. Neither identity hashes absolute runtime paths.
- Verified software: Python 3.12.4, NumPy 1.26.4, pandas 2.2.3, SciPy 1.15.2, PyYAML 6.0.2, statsmodels 0.14.4, Matplotlib 3.10.0 and PyArrow 19.0.0. No `stata`, `stata-se` or `stata-mp` executable was found. `/usr/bin/Rscript` exists but is not used/tested. Source Stata and ado inference remain unvalidated.

These are **synthetic engineering coverage counts**, not Danish results:

Atom	Expected keyed cells	Computed synthetic cells	Registered execution state
atom-000001	223	140	partial
atom-000002	161	0	boundary
atom-000003	669	0	boundary
atom-000004	337	0	boundary
atom-000005	274	0	boundary
atom-000006	116	0	boundary
atom-000007	56	0	boundary
atom-000008	25	0	boundary
atom-000009	36	0	boundary
atom-000010	18	0	boundary
atom-000011	24	0	boundary

Five of the 50 targets contain computable components: Figure B2 is complete in this fixture; Figure B1 and Tables 1/B1/B2 are partial. Other target cells remain explicitly unavailable, including the contextual, non-empirical prose_13. Local shared construction still executes where supported. These states do not establish data availability, matched estimation or empirical replication.

All 145 PDFs are 130.700 mm wide within 0.001 mm and contain no embedded raster images. Visual inspection covered both contact sheets of all 29 actual CLI outputs, plus enlarged B1, B20, B4 and partial F4. The successful inspection reproduced **three adjacent label collisions per B20 panel**; B3/B4 retain tick labels 0.0–1.0

under an age label despite unavailable bins. Actual boundary F3/B7/B10/B12 heights are respectively 320.548/388.112/385.267/391.312 mm. These are confirmed presentation limitations, not passed publication-layout checks. Contact sheets/previews exist under both inspection scratch directories; the successful measurements and audit are under `final-inspection-02/`.

The final World contains no synthetic result table, filled manuscript slot or canonical Atom result. No original scientific choice, frozen renderer or presentation file was altered. The author fixes and limitations are ready for controller-directed independent review; no new Opus/Codex review or acceptance is claimed here.

Prior-round implementation record (historical)

Everything below preserves the round-1 coverage, source-locator and test history. Its implementation digest, test counts and final validation appendix describe that earlier revision; the round-2 validation above supersedes that historical record where changed.

This bundle implements the frozen design’s determined construction and output contracts. It is **not a fully resolved executable empirical replication**. The frozen audit explicitly marks all matching-dependent output and the descriptive counterfactual scaling as unresolved. Supported Atom1 cells are computed; all affected siblings retain individually keyed no-estimate rows. Local construction and numerical model algebra run independently without being relabelled accepted predecessor evidence.

No confidential Danish/APTO data, old empirical session, observed outcome, empirical replication program or remote QGPU was accessed. Synthetic results remain in the author scratch tree. No original Atom, Molecule, membership, manifest, audit, manuscript, figure template or repository resource was edited. Comparison of all 122 immutable original files with the supplied World found zero changed or absent files. No commit/push/release, environment install, remote write or sub-agent invocation occurred.

Authority and implementation coverage

Read the repository lane AGENTS, full predata-engineering skill, preregistration-package skill and required contract/granularity/paper references, paper-to-replication skill, and the selected canonical semantics. Read the 11 member protocols/result regions, Molecule registration, shared data/identification/governance, source/matching/deviations/target audit and existing paper/result/figure interfaces. The package had no existing prereg/code implementation to import. Public original cleaning and analysis programs, including FN/spouse code, were inspected without executing them. Field meanings were checked against local FSV register documents listed in the frozen schema evidence.

The exact entry point, inputs, units, output columns, individual Atom functions and tests are in [execution README](#). No additional experiment registry or Atom was introduced.

Frozen scope	Executable supported work	Retained scientific boundary
Atom1 / B1	First linked death in1985–2014, five-year source categories, cumulative conditional denominator before count<5 suppression	A04 bin8 label40–45 versus coded40–44; exact19+ prose
Atom1 / B2, Table1/B1/B2	Age25–50 at event, first death1985–2014, parents known/present at-1, exact11 years; separate baseline outcomes/means; equal-variance Table1 t tests	C02 under-six row only; missing optional dictionaries/fields and empty groups remain local
Atom1 / B3/B4 and prose	Inputs/cohort are constructed; expected series and prose rows retained	C03 age clock; A04 source histogram origins/edges and prose denominators
Atom2 / B5/B23	Cohort/labor/Rx construction; independently executable Eq1 HC1 and scaling kernels	C04 prediction population/denominator; no unscaled substitution
Atom3 main labor, ATP/rank	Annual rank/ties, ATP scaling/caps, candidate-control timing, indexing, Eq2/3/4 model matrices and joint covariance kernels	C01/A01/A02 all NN sample-dependent targets; A03 pooled summaries
Atom3 balance/heterogeneity	Explicit source baseline rows; cause/period/age rows and source figure coordinates; geographic category kernel	C02/A05 special balance rows; A07 age threshold; C09 geography/contrast; TableB3 count inconsistency below
Atom3 FN/prose robustness	Source restrictions and each requested output preserved as audit/boundary; no extra alternatives	C10/A08 FN scientific/execution conflict; U03 missing rematch-ing/model/output specification

Frozen scope	Executable supported work	Retained scientific boundary
Atom4 health	Admission-year count, inclusive day union, screening subtraction, exact GP/ATC sets, private-service crosswalk join, coverage-aware missingness; joint-parent health Eq2 kernels	C01/A01/A02/A03; A05 private/combined target; U02 mapping
Atom5 co-occurrence	Extended earnings lag, source5% decline, onset versus continuation, joint mental/psych/opioid onset, hospital-count product, final-4..5 clock; ratio arithmetic only with supplied time	C01/A01/A02/C05 displayed model; A03 immediate year0/1
Atom6 children/care	Distinct linked births, same-sex parents, exact child counts, youngest category/tie, prescribed enrollment code sets	C06 childless comparison; A06 age5/tied youngest; matching/post gates
Atom7 surviving parent	Hospital-day/Rx input construction, shared parental linkage, geography arithmetic	C07 no invented illness stratum/calendar zero; C05/U03; matching/post gates
Atom8 wealth	Source annual vintage/max rule and real-DKK deflation	C08 unit/owner; U04 three absent variant models; matching/post gates
Atom9 family	Cohabitation and untruncated child count, joint-parent Eq4 kernels	C01/A01/A02/A03
Atom10 in-laws	Event spouse fixed before history expansion, key cardinality and exact-year balance; spouse subject supported in kernel	A09 distinct matching/presence/tenure/pooled model; A01/A02/A03

Frozen scope	Executable supported work	Retained scientific boundary
Atom11 jobs	Early/late job priorities, owner/public coding, adjacent firm changes, source four-digit occupations; firm/public Eq3 kernels	A10 occupation transition order; C01/A01/A02/A03

The CLI never obtains a match by accepting an arbitrary pre-matched panel. It also does not expose a choose-paper/choose-code flag. The standalone kernels accept mathematical panels so source-supported algebra can be tested; they are not evidence that matching, normalization, source uncertainty or a disputed model has been resolved.

Meaning-preserving adaptations

1. Historical extract paths are manifest inputs. Exact logical-to-physical aliases, source vintage, pseudonym correspondence, covered years and byte hashes are explicit. No local schema catalog is treated as evidence of mounted data.
2. Source yearly merges become checked many-to-one joins. The11-observation condition becomes unique person-year plus exact event-year-set equality. Neither operation silently deletes conflicting rows.
3. Hospital expansion into one row per person-day becomes an inclusive interval-union sweep split at calendar-year boundaries. Hand counterexamples independently verify overlap and cross-year totals. Outpatient days are not relabelled inpatient days.
4. Stata percentile order statistics are transcribed explicitly; pandas' default interpolated quantile is not substituted. Annual tied ranks use the source mean-rank formula and retain undefined singleton ranks.
5. Source OLS is computed by Frisch–Waugh–Lovell with the complete nuisance column space. Redundant nuisance coding can be absorbed without arbitrary treatment coefficients. Tests verify coefficient and HC1 covariance invariance against independent dense OLS/QR. This establishes a mathematical identity, **not** the unknown source reghdfe finite-sample convention in every restricted sample.
6. Cluster scores use person $\times$ match and the same residual/design matrix. A supplied effective parameter count/df has the usual CR1 algebra tested against statsmodels. Without source-established metadata, CR0 is explicitly unvalidated source inference and source CI export is blocked. Dynamic intervals use source1.96; pooled intervals require source residual df. Joint contrasts use off-diagonal covariance. No OLS substitute for an unspecified estimator, invented bootstrap or independent-SE contrast is used.
7. Python functions retain local frames in memory but export only aggregate results, dispositions and provenance. Real missing outcome definitions do not

- silently become zero. Missing inflation no longer reports completed real-earnings construction; nominal ranks remain independently possible.
8. Native activation is copied meaning-preservingly from the repository canonical implementation. All actual registered expressions are TRUE. Tests exercise false/unknown externally through the native evaluator and unactivated execution path; the frozen protocols are never changed to manufacture dependencies.
 9. The result exporter reads frozen literal figure CONFIGs for exact source target IDs, panels, categories, units and event coordinates. Native display IDs are used only by the original renderer output binding. No figure was redesigned or copied from a rendered source PDF.
 10. Scientific registration hashes exclude mutable result bodies and Molecule current interpretation. Source material identity, runtime input binding and implementation hash remain separate. Input paths do not enter input identity. Read-only relocated World tests use writable action code and output directories.

Exact unresolved target mapping

All IDs below are the existing [frozen deviations](#), not new scientific amendments.

- **U01:** Every consuming register component needs the authenticated historical extract/key/field/year correspondence. Source: JournalII; clean programs; schema-evidence. Cumulative DODSAARS2001 and DODSAASG2019 are not inferred from current2022 metadata; LMDb PNR12 correspondence is not inferred by renaming.
- **U02:** Only dependent education, real income/wealth, psychology and11-region constructions stop for absent author mappings. Source: clean_3,clean_6,clean_10,clean_11,set_globals. No guessed inflation or specialty dictionary.
- **C01/A01:** All NN sample-dependent target rows stop. JournalIII p1820/fn14 versus clean_15_matching.do249–366; match-edge/tie/relabel semantics363–440, clean_16 and unavailable engine versions. The34 source groups and source distance history are not silently preferred.
- **A02:** Matched Eq2–4/linear robustness; Journalfn16 and consuming reghdfe. Tested nuisance-space identity is not treated as a global resolution.
- **A03:** Every pooled Eq3/4 summary and prose_08; analysis_5 post $\geq$ 0 versus five-year after-death language; analysis_14 year0 versus prose immediately-following. It is not an independent gate on explicit dynamic coordinates.
- **C02:** Table1/B1/B2/B3 under-six row, analysis_2/15; other means remain.
- **C03/A04:** B3/B4, B1 disputed label and prose_01/02; analysis_1 lines53–71 and analysis_2 histograms. Source cumulative denominator is not population death risk.
- **C04:** B5/B23; analysis_3/4 regression, unrestricted predict and collapse blocks versus Eq1/captions.
- **C05:** F4/B7/B12 dynamic coefficients and F6/B15 model; captions versus analysis_14/5/12.
- **A05:** prose_07 and TableB3 psychiatry; Journalfn22, clean_2/3, analysis_6/15.
- **C06/A06:** F5/B13 childless eligibility (analysis_8 lines18–21), B14 age5/youngest identity (clean_9 and analysis_8).

- **C07:** F6/B15/B16/prose_14 and illness component of prose_09; JournalV.C versus analysis_12 annual ≥ 90 , event-year/window and 2019 zero; B16 pre/post split. No illness indicator is assigned.
- **U03:** prose_03/04/05/06/09; Journalfn14–16 and p1831. The known second-death gap >5 is not expanded into an unspecified rematching battery.
- **C08/U04:** TableB4 unit/owner (analysis_9 lines21–31,75; clean_12) and prose_10/11/12 absent asset variants (SupplementA.A/fn29).
- **C09:** B19 coefficient and comparison models, SupplementB19/analysis_10. Source F-as-p assignment is not repaired.
- **A07:** TableB6 age strata, analysis_11; source median versus printed70/38. Cause and period categories remain explicit.
- **C10/A08:** B21/B22 FN; SupplementA.F/B21/B22 versus FN_clean/FN_analysis. No choice of-3/-5 start, person/person $\times$ stack clustering, p1/99 versus p5/95 tails or removed-youngest-variable repair.
- **A09:** B18; clean_14/16 and analysis_16. Source spouse-present/survival, tenure/order and gender contrast controls remain unresolved.
- **A10:** TableB5 occupation transition; clean_11 and analysis_13 eight-prior-nonmissing-record scan lacks a deterministic within-person-year order.
- **A11:** Consequential simultaneous deaths, missing cause/identifier and duplicate link/row selection remain construction-level boundaries. Identical repeated events are deduplicated only where the source authorizes it.

Additional literal-interface inconsistency discovered in this author round: target-inventory says TableB3 has **16 targeted rows**, while Atom3's explicit row enumeration and analysis_15_balancing_test.do line114 contain **15** (age, male, three education groups, cohabitation, child count, youngest age, under6, two parent ages, employment, unconditional ATP, earnings, public), followed by 9 nontargeted rows and N. The implementation retains exactly those named rows, 50 mean cells, all currently source-blocked. It does not invent a sixteenth row, silently change the inventory or claim its count assertion is satisfied. The count/label discrepancy needs source clarification before claiming a complete table interface.

Tests and commands

All Python commands used `/home/shaoerzhuo/anaconda3/bin/python`. Full invocation:

```
/home/shaoerzhuo/anaconda3/bin/python <World>/prereg/execution/run_tests.py
↪ --output <author-scratch>/suite-03
```

The execution README gives a relocation-safe command. Actual scratch root: `/home/shaoerzhuo/Bolero-denmark/.prereg-runs/denmark-reference-code-20260906/parental-deat`

- `run.py --help`: passed without real inputs.
- Initial synthetic Atom1 CLI: partial; computed supported summary cells and kept source boundaries.

- **suite-01:**31 tests, zero assertion failures,2 errors; all numerical tests passed. Errors were a duplicate `E_FAELE_ID` column in spouse-history merge and copying a read-only Atom before test mutation. Fixed by distinct spouse-current-partner naming and chmod of scratch copies only.
- **suite-02:**36 tests, zero assertion failures,8 cascading errors from a duplicated uncertainty keyword introduced during the disposition fix. Fixed the exporter, then verified Atom1 CLI independently.
- **suite-03:**36 tests passed, zero failures/errors;116 figure renders passed. Numerical/integration runtime394.372seconds.
- **suite-04:**38 tests passed, zero failures/errors;145 figure renders passed, adding explicit nonactivation and occupation/enrollment checks. Numerical/integration runtime171.875seconds. The synthetic job fixture was narrowed to adult/spouse jobs to reduce cost; no cohort or construction rule changed.
- Final source audit identified omitted subgroup baseline rows in F5/B13/B19. Added the source sex-by-subgroup rows, independently checked counts8/8/4, and reran Atoms3 and6 on the synthetic manifests. Their scientific states remain boundary. This targeted follow-up tests the last output-only change; the unchanged numerical suite was not rerun unnecessarily. The current complete `run_tests.py` contains39 tests.

Independent numerical checks include Stata order-statistic hand values, $t=\sqrt{6}$ withdf4, HC1 versus statsmodels, full joint clustered covariance versus an independent statsmodels oracle, CIcritical-value distinction and nonidentified treatment rejection. Construction checks cover exact event dates and adult age/calendar endpoints, second death, missing/duplicate11-year panels, rank ties/singletons, inclusive hospital overlap, source service/ATC codes, incomplete years, twins, index denominators, employment transitions, wealth vintages, onset clocks and multiple-treatment control timing.

Integration checks execute all11 Atoms with both policies and all11 with absent inputs; optional missing education/inflation leaves independent outcomes; evidence checks cover wrong input/kind, explicit `no_result` nonretry, mutable Atom/Molecule realization, protocol tampering, corrupt output, copied input paths and relocated read-only Worlds.

Figure tests use unchanged source templates, explicitly invented interface-only vectors and actual synthetic CLI outputs. The final harness exercises full effects and available annotations, actual CLI rows, missing inputs, partial inputs with long reasons and nearby interval endpoints, and nonactivated rows:145 PDFs for29 figures $\times$ 5 dispositions. Missing required CIs and wrong units are rejected. Interface-only numeric vectors are not a replication of source-blocked analyses. Current World manuscript slots remain empty. No unchanged paper/PDF rebuild is claimed.

Limitations and handoff

Stata/reghdfe/teffects/ado inference is untested because Stata is unavailable. Source software versions remain A01; neither syntax nor Python numerical agreement establishes that engine's small-sample conventions. No national-scale memory/performance

validation occurred; dense numerical kernels are for reference checking. Governed data access, dictionaries, metadata truth and disclosure review remain unestablished.

No `runtime-figure-interface.md` or `runtime-execution-interface.md` was supplied/found in the World or current lane files. The frozen renderer's default local font is DejaVu Sans; governed Helvetica Neue files must be supplied through the existing `style_dir` interface. No such font set was located in the lane resources. Actual runtime font binding has not been validated and the templates were not altered to hide this limit.

There are no prior engineering reviewer findings in round1. This is the author's self-check, not an adversarial Opus or Codex Sol xhigh review. No independent review was run: the current invocation is author-only and expressly prohibits spawning agents. The controller can supply those findings to a subsequent author round. No review acceptance is asserted.

Synthetic tests establish implementation checks only. They are not empirical replication, causal validation or proof that the registered data are available.

Final validation appendix

Final author checks completed 2026-09-06.

- Full synthetic execution: `scratch/r01/author/suite-04/test-summary.json` records 38 passing numerical/construction/integration tests and 145 passing renders. This executes all 11 actual Atoms with both policies, all 11 absent-input paths, partial inputs and the relocation/evidence adversarial cases.
- Last output change: `scratch/r01/author/final-validation/targeted-test-log.txt` records the passing subgroup-baseline test. Updated Atoms 3 and 6 were actually rerun with direct dispatch under `baseline-final/`; both remain scientific boundary. The other 9 Atom implementations were unchanged.
- Final interface follow-up: `scratch/r01/author/final-validation/test-summary.json` records 145 successful renders, using those updated exports and the other 9 suite-04 exports. Full interface fixtures now include available source annotation cells, as well as effects. Missing required CI/wrong-unit rejection also passed. The reproducible follow-up script is `scratch/r01/author/final-validation.py`; the bundle's current `run_tests.py` incorporates the new baseline case in its complete 39-test suite.
- Final integrity audit: `scratch/r01/author/final-validation/integrity-audit.json` records 122 original files compared, zero changed originals and zero unexpected World files outside execution/engineering. All 11 Atom/Molecule registration/material checks pass. Implementation identity: `dd6cb47b82e60a36ffc26152075ae4937a3583d0e7c2f319a6e018d9b79ccc4f`.
- No synthetic estimated value was copied into this report, the World manuscript or a canonical Atom. The coverage states in synthetic runs are Atom1=partial and Atoms2-11=boundary; these are engineering dispositions, not evidence about Denmark.

Actual final 29 figure PDFs all have 130.700mm width (within 0.001mm of the requested width). Dimensions are recorded in `final-validation/dimensions.json`.

DejaVu Sans is embedded as CID TrueType, verified with pdffonts on B20. The available local font path was tested; unavailable governed Helvetica Neue binding remains unvalidated.

Visual inspection covered contact sheets of all 29 actual synthetic outputs, with enlarged checks of B1 partial cumulative bars, B7's eight panels and unavailable annotations, F1 partial vectors/nearby endpoints, B3 missing histograms, and B20 labels. Scratch previews are in visual-inspection/. The PNG contact sheets initially exposed a grayscale-preview decoding issue; RGB8 inspection copies showed the underlying PDFs correctly. This was an inspection artifact, not an altered result or template.

Presentation issues remain visible and are not counted as passing publication-layout checks. B20's four long category labels overlap at 130.7mm; reproduced and inspected in visual-inspection/figure26-detail.png. Boundary-heavy F3, B7, B10, B12 PDFs are 320.548, 388.112, 382.422, 391.312mm tall respectively, so ordinary manuscript-page placement has not been validated. The missing B3/B4 histogram template retains a default 0–1 horizontal range under an age-axis label despite explicitly saying no estimates/bins are available. No bins, result rows or shorter scientific labels were fabricated to conceal these issues. The user expressly prohibited changes to the fixed figure templates; this report identifies the exact issues for the authorized presentation owner rather than claiming the PDFs are publication-ready.

The tests distinguish file/schema compatibility from scientific fidelity, empirical feasibility and final typography. The bundle still needs source resolutions listed above and a verified source Stata/ado environment before the blocked analyses can produce faithful estimates. No independent reviewer acceptance has occurred in this author round.

Reference implementation review

Engineering status: REVISE — retained at handoff. Review rounds: 3. Revision slots used: 2 (initial draft excluded).

User-directed bounded handoff: no more than two author revisions. Proceeding to formatting is not reviewer acceptance. Failed/missing reviewer calls have no scientific verdict; original records remain available below.

This augmentation used schemas, documentation and synthetic fixtures only. Code review and package compilation do not establish scientific validity, empirical replication, data availability or independent acceptance of the original science. Original registrations, manuscript sources and scientific review remain unchanged. Unresolved source definitions remain unresolved; runnable components and blocked targets must be distinguished in prereg/execution/README.md.

The host checks prose Atom coverage, Python syntax and synthetic commands. These are engineering checks, not a scientific verdict.

Controller checks and review dispositions

```
{
  "round": 3,
  "accepted": false,
  "host": {
```

```

"status": "pass",
"findings": [],
"processes": [
  {
    "label": "r03-host-run",
    "command": [
      "/home/shaoerzhuo/anaconda3/bin/python",
      "/home/shaoerzhuo/Bolero-denmark/.prereg-runs/denmark-reference-cod
        ↪ e-20260906/parental-death/scratch/r03/host/attempt-01/world/pre
        ↪ reg/execution/run.py",
      "--help"
    ],
    "model": null,
    "cwd": "/home/shaoerzhuo/Bolero-denmark/.prereg-runs/denmark-referenc
        ↪ e-code-20260906/parental-death/scratch/r03/host/attempt-01",
    "log": "/home/shaoerzhuo/Bolero-denmark/.prereg-runs/denmark-referenc
        ↪ e-code-20260906/parental-death/logs/r03-host-run-a01.log",
    "created_at": "2026-09-06T08:29:03.294000+00:00",
    "status": "finished",
    "pid": 2343677,
    "start_ticks": "82141820",
    "state": "R",
    "started_at": "2026-09-06T08:29:03.301589+00:00",
    "returncode": 0,
    "completed_at": "2026-09-06T08:29:03.370350+00:00"
  },
  {
    "label": "r03-host-run_tests",
    "command": [
      "/home/shaoerzhuo/anaconda3/bin/python",
      "/home/shaoerzhuo/Bolero-denmark/.prereg-runs/denmark-reference-cod
        ↪ e-20260906/parental-death/scratch/r03/host/attempt-01/world/pre
        ↪ reg/execution/run_tests.py",
      "--output",
      "/home/shaoerzhuo/Bolero-denmark/.prereg-runs/denmark-reference-cod
        ↪ e-20260906/parental-death/scratch/r03/host/attempt-01/output"
    ],
    "model": null,
    "cwd": "/home/shaoerzhuo/Bolero-denmark/.prereg-runs/denmark-referenc
        ↪ e-code-20260906/parental-death/scratch/r03/host/attempt-01",
    "log": "/home/shaoerzhuo/Bolero-denmark/.prereg-runs/denmark-referenc
        ↪ e-code-20260906/parental-death/logs/r03-host-run_tests-a01.log",
    "created_at": "2026-09-06T08:29:03.377981+00:00",
    "status": "finished",
    "pid": 2343678,
    "start_ticks": "82141829",
    "state": "R",
    "started_at": "2026-09-06T08:29:03.386196+00:00",
    "returncode": 0,
    "completed_at": "2026-09-06T08:35:49.572236+00:00"
  }
],
"completed_at": "2026-09-06T08:35:49.585525+00:00",
"scope": "Structural coverage, syntax and synthetic execution; not
  ↪ scientific validation."
},
"reviews": {
  "opus": {
    "status": "complete",
    "verdict": "revise",

```

```

    "report": "/home/shaoerzhuo/Bolero-denmark/.prereg-runs/denmark-referen
    ↪ ce-code-20260906/parental-death/reviews/r03/opus.yaml"
  },
  "codex": {
    "status": "complete",
    "verdict": "revise",
    "report": "/home/shaoerzhuo/Bolero-denmark/.prereg-runs/denmark-referen
    ↪ ce-code-20260906/parental-death/reviews/r03/codex.yaml"
  }
},
"review_status": "revise",
"revision_slots_used": 2,
"continuation": "package_current_draft_with_unresolved_findings"
}

```

r01 / codex.md

Codex independent adversarial review — round 1

Verdict

Revise. The source-gated scientific choices are mostly implemented faithfully, but the execution/evidence trust chain and two exact interfaces accept falsified inputs. These are engineering defects, not requests for a new study or a different frozen protocol.

This review found no empirical Danish result. All reported numerical outputs were generated from fixtures labelled synthetic. A later acceptance would mean tested engineering/source fidelity within the explicit source limits—not scientific confirmation, replication success, or approval of the original study.

Scope and frozen identity

I read all 11 Atom front-matter result regions (`atom_result: prereg` at line 4) and protocols (starting at `atom-000001:58`, `000002:59`, `000003:66`, `000004:65`, `000005:64`, `000006:63`, `000007:61`, `000008:60`, `000009:59`, `000010:60` and `000011:61`), the Molecule membership and canonical interpretation (`molecules/molecule-000001.md:9-109`), shared D1-D8 (`prereg/data-contract.md:1-69`), matching transcription, deviations, target/source/schema audits, the 63 result-slot bindings, 29 figure bindings, original author programs and the execution code.

The World preserves the original package outside the new engineering implementation:

```

diff -rq --exclude=execution --exclude=engineering \
.../parental-death/.controller/source \
.../parental-death/v001/world
# exit 0; no output

```

The matching stop is not a fabricated data excuse. `deviations.md:25-33` documents that the journal requires exact earnings/young-child groups and three-year earnings/employment histories, while `clean/clean_15_matching.do:249-366` constructs 34 partly combined groups and a different distance vector. Lines 364-366

call the unavailable source Stata matching engine; lines 374-440 retain tie/reused-control handling whose version/order correspondence is unresolved. Blocking those matched-sample estimates is narrower and more faithful than fitting a substitute.

Advertised-suite result

I copied the action out of the read-only World and ran the entire advertised suite with the controller Python. It passed 39/39 tests with zero failures/errors and rendered all 145 variants. The tests include independent dense-OLS/statsmodels comparisons, clustered joint covariance, HC1, order statistics, cohort/event/calendar boundaries, complete-year absence semantics, all Atoms under both dispatch policies, missing optional inputs, Strong-Kleene activation, copied-layout execution, mutable realized Atom/Molecule bodies, protocol tampering and no-result retry rejection. The exact test names and outcome are in `scratch/r01/codex/advertised-suite/test-log.txt` and `test-summary.json`.

Direct dispatch did only bypass outer activation in the advertised tests: target-specific C/A/U gates remained numeric-null, and no blocked matched coefficient was promoted. `run.py:101-102` also rejects a prior `no_result` under both policies.

Per-Atom coverage

The counts below are actual reviewer-executed direct-dispatch synthetic outputs. “Available” means only a fixture result cell, never an empirical estimate.

Atom	Direct state and cells	Implemented/exercised path and assessment
atom-000001	partial; 140/223	<code>analyses.population_results</code> (<code>analyses.py:127-203</code>) computes B1 cumulative shares with pre-suppression denominator, B2 parent shares and unaffected Table 1/B1/B2 cells; C02/C03/A04 cells stay blocked. Local B1, balanced panel, family, ATP, earnings and education all constructed.

Atom	Direct state and cells	Implemented/exercised path and assessment
atom-000002	boundary; 0/161	atom_000002 (<code>analyses.py:368-375</code>) constructs the treated panel and mental prescriptions. Figure B5/B23 scaling stops at genuine C04; unscaled Eq1 is not substituted.
atom-000003	boundary; 0/669	atom_000003 (<code>analyses.py:378-387</code>) constructs cohort/family/labor/education, population ranks, jobs and all available health inputs. C01/A01 blocks the NN sample and target-specific gates remain. Table B3 count and A02 attribution issues are noted below.
atom-000004	boundary; 0/337	atom_000004 (<code>analyses.py:390-392</code>) exercises hospital, psychiatry, Rx, GP/private service and demographic construction; all effects require the unresolved NN sample, with A05/U02 retained for private psychiatry.

Atom	Direct state and cells	Implemented/exercised path and assessment
atom-000005	boundary; 0/274	atom_000005 (analyses.py:395-399) exercises health/demographic inputs. construction.onset_panel:502-529 contains the registered lag/onset arithmetic, but the required matched extended history cannot exist before C01; no fake panel is supplied. The PSYK fixture component correctly remained locally unavailable in this Atom's coverage window.
atom-000006	boundary; 0/116	atom_000006 (analyses.py:402-407) exercises enrollment plus fam- ily/demographic/labor inputs. C06/A06 and matching/post gates block the target estimates without discarding the constructors.
atom-000007	boundary; 0/56	atom_000007 (analyses.py:410-415) constructs parent hospital-day and mental-Rx inputs plus the common panel but assigns no guessed illness stratum under C07.

Atom	Direct state and cells	Implemented/exercised path and assessment
atom-000008	boundary; 0/25	atom_000008 (analyses.py:418-421) constructs annual maximum wealth with inflation and the common panel. C08/U04 and matching/model gates block numerical interactions.
atom-000009	boundary; 0/36	atom_000009 (analyses.py:424-425) constructs common demographic/family inputs. Family-formation Eq4 remains dependent on unresolved matching/post/model choices.
atom-000010	boundary; 0/18	atom_000010 (analyses.py:428-434) fixes the event partner before spouse history expansion. It does not substitute biological-child pairs for the distinct A09 spouse match.
atom-000011	boundary; 0/24	atom_000011 (analyses.py:437-442) exercises firm/public and four-digit occupation construction plus the common panel; A10 ordering and shared matching/post gates remain.

The shared constructors at **construction.py:48-595** retain FAIN year shifting, first-death time zero, age 25-50 and 1985-2014 limits, exact eleven-year balance, future control eligibility, source ATP and income units, coverage-aware zeroes, inclusive hospital-day unions, Rx prefixes, childcare codes, wealth vintages, job adjacency and spouse histories. The advertised boundary tests specifically exercise those risky clocks

and units. I found no source-supported target that was replaced solely by an always-missing wrapper; matched estimators are disconnected from target output because the sample itself is scientifically unresolved.

Findings requiring revision

1. High — implementation/evidence identity is not trusted

`contracts.verify_world` (`prereg/execution/contracts.py:46-67`) reads `frozen-identity.json` from beside the copied executable and verifies only the original frozen World materials. Neither that map nor `prereg/manifest.yaml:34+/CHECKSUMS.sha256` seals the added execution files. `contracts.code_identity (:260-263)` calculates a hash, but `run.py:110-124` only writes it into a self-hashed receipt. `load_evidence` (`run.py:29-88`) never compares it with a trusted expected value.

I copied the action and changed one line corresponding to `analyses.py:157` from `100*cumulative/total` to `200*cumulative/total`. Original and changed CLIs both exited 0 as `partial`. The four available B1 values changed from 20.8333/41.6667/62.5/96.875 to 41.6667/83.3333/125/193.75. Different implementation IDs were recorded, but the changed receipt was accepted by `load_evidence` as Atom 1 partial.

Narrow fix: bind a reviewed execution-file manifest/hash in trusted immutable World/controller metadata; compare the copied action before import; require evidence `implementation_id` and receipt digest to match externally bound expected values. An adjacent lock and recomputable receipt checksum cannot reject adversarial code. Regression: mutate `analyses.py`, `run.py` and the adjacent lock in separate cases and require rejection, while retaining the current legitimate realized Atom-body and Molecule-interpretation mutation behavior.

2. High — a truncated predecessor vector passes evidence validation

D8 (`data-contract.md:63-67`) requires the complete keyed vector. `load_evidence` checks column names, numeric nullness, statuses and only the set of `target_ids` (`run.py:64-81`); it neither validates the exact cell keys nor reconciles `target-status.csv` and figure projections.

My probe copied actual Atom 1 CLI output, deleted B1/A/All-first-deaths/0-4 (223 to 222 rows), updated the results hash and recomputed the receipt's unkeyed digest. The loader accepted `partial` and made a binding even though `target-status.csv` still claimed 13 B1 cells.

Narrow fix: validate the exact per-target key/unit/uncertainty roster; recompute target status; compare each plot table with the exact results projection; externally bind the receipt digest. Add deleted, extra, wrong-key/unit, stale-status and stale-figure tests based on a real CLI output layout.

3. Medium — outcome deletion precedes matched-panel cardinality validation

D1 (`data-contract.md:5-9`) requires one valid person-match-year and says mismatches stop rather than license deduplication. `models.matched` requires columns at line 160, drops rows missing the outcome at line 165, and only then calls `unique` at line 166. I appended a duplicate structural key with `y=missing` to the advertised algebra fixture. Two duplicate-key rows existed before outcome deletion, but Eq2 fit successfully and produced exactly the clean coefficients.

This does not alter today's outputs because all matched targets remain source-blocked. It does make the advertised future source-supported kernel unsafe. Check structural keys and event-year sets before outcome complete-case selection, then apply only an explicit source missing-outcome rule. Add the one-missing duplicate and missing-key boundary cases.

4. Medium — annotations accept unrelated schema substitutions

The shared renderer checks annotation panel, statistic family and duplicate keys (`source_render.py:38-46`), then only that an available value is finite (`:70-72`). It prints arbitrary series/term/unit at `:117-120`. Starting with actual Atom 3 Figure 1 tables, I changed one annotation to A / UNREGISTERED SERIES / F / UNREGISTERED TERM / 123 BANANAS. The renderer wrote a PDF.

This breaches the exact slot/interface binding in `results-map.md:11-12`; a permitted statistic family does not authorize arbitrary scope or units. Freeze and validate each target's annotation key tuples and statistic-specific units. Retain empty annotation tables for registered missing input, but reject unregistered extras.

5. Low — A02 is over-attributed to Table B3

`deviations.md:45-53` limits A02 to Eq2-Eq4 estimates and linear-controls robustness. Atom 3 (`atom-000003.md:89-95`) defines B3 as baseline means. Yet `analyses.target_boundary (:78-82)` adds a combined C01/A01/A02 boundary to every Atom 3 target; actual B3 rows therefore cite A02. C01/A01 already genuinely block the matched sample, so availability is unchanged. Separate sample-construction gates from estimator-normalization gates and regression-test exact target gap IDs.

6. Low — two actual boundary figures have misleading/colliding axes

At the registered 130.7 mm width, actual Figure B20 has overlapping four-category labels. The shared bar path (`source_render.py:165-174`) rotates labels only when category count exceeds six, even though the source labels in `figure_b20.py:2` are long. Actual no-bin Figure B4 shows matplotlib's numeric 0.0-1.0 x axis under "Age; source-determined bin edges"; the histogram path (`:158-164`) has no no-data x-axis behavior. The reasons are readable and no numerical estimate is plotted, but the display is not publication-safe.

Artifacts: `scratch/r01/codex/visual-review/figure26-actual.png` and `figure10-actual.png`. Fix label layout based on text extent/length and suppress

numeric histogram x ticks when bins are unavailable. Test bounding-box separation and no default numeric axis without changing labels or inventing bins.

Source limits and unavailable numerical results

- U01 remains real: historical extracts, namespace correspondence, vintages and complete years are unverified. U02 mappings/inflation are absent. No missing dataset alone is treated as a reference-code defect.
- Atom 2 B5/B23 stays no-estimate under C04. Atoms 3-9/11 remain no-estimate under C01/A01 and estimator-specific A02 plus their target gates. Atom 10 B18 remains no-estimate under A09/A01/A02/A03. C02-C10, A03-A11 and U03/U04 remain only where recorded; none was resolved.
- Atom 1 B3/B4 and exact prose denominators remain unavailable under C03/A04. Its synthetic partial cells demonstrate implementation paths only.
- Table B3 has a frozen literal inconsistency: `target-inventory.md:21` says 16 targeted rows, while Atom 3's named list and `analysis_15_balancing_test.do:114` contain 15. Keeping the 15 named rows and not inventing another is the conservative current behavior; a complete-interface claim remains unavailable.

Tests and limitations

Reviewer-authored probe source and output are `scratch/r01/codex/adversarial_tests.py` and `adversarial/probe-results.json`. The mutation copy and receipts are under `scratch/r01/codex/adversarial/`; nothing in the World was edited.

Stata was unavailable (`command -v stata-mp, stata-se, and stata` returned no path), so `source_models.do`, `teffects`, `reghdfe`, source residual degrees of freedom and engine-specific tie semantics were not numerically executed. I did not use real registers, test access/linkage truth, or rebuild the complete TeX manuscript. All 145 PDF variants were generated, but my independent visual inspection targeted actual Figure 1, B20 and B4 rather than every PDF at publication zoom.

The exact commands and outcomes, including failed invariants and limited checks, are recorded in the sibling YAML.

r01 / opus.md

Independent engineering review — Opus, round 1

Verdict: revise (three narrow code fixes with regression tests, two nits). Engineering and source fidelity only. Nothing here is scientific acceptance of the frozen protocol, an observed result, or a replication claim.

Reviewed World: `.prereg-runs/denmark-reference-code-20260906/parental-death/v001/world`. Everything I ran, wrote or rendered lives in `.prereg-runs/denmark-reference-code-20260906/parental-`. No prior engineering review or author response existed for this round, and I did not read the other reviewer.

1. What is actually in the bundle

`prereg/execution/` (12 files, ~2,700 lines of Python plus a 42-line Stata kernel) and `prereg/engineering/` (2 reports) are the only additions to the frozen package. I confirmed independently that the other 122 files are byte-identical to `.controller/source` and that nothing wrote into the World during any of my runs (`diff -rq`, then `sha256sum` of all 122 files, then a hash of the execution directory before and after every test — the copied action directory still hashes identically).

The package is a real implementation, not a wrapper. Atom 1 computes and exports 140 available cells. Atoms 2–11 return `boundary` with zero available cells — not because the code is missing, but because the frozen `deviations.md` blocks every one of their registered targets. My job was to decide whether that blocking is honest, whether the supported parts are faithful, and whether the plumbing can be abused.

2. Method

1. Mechanically cross-tabulated all eleven Atom protocols against the code: `Retained source targets:` vs `analyses.TARGETS`, `Unresolved target gates:` vs the gaps `target_boundary` actually emits, `registration.tables` vs what each module requests, and `figures.yaml`'s `atom:` binding vs target ownership. The 50 registered target IDs in `target-inventory.md` and in `analyses.TARGETS` agree exactly — no target added, none dropped, none silently regrouped.
2. Re-derived every one of Atom 1's *available* cells against the exact cited author programs (`analysis_1`, `analysis_2`, `clean_6`, `clean_7`, `clean_10`, `clean_11`, `clean_13`), because those are the only numbers this package produces.
3. Built independent statsmodels oracles for the model kernels rather than trusting the author's tests.
4. Wrote my own falsifying counterexamples for the risky constructions.
5. Ran the whole advertised suite, then re-ran all eleven CLIs myself, then attacked the dispatch/evidence/identity surface, then rendered all 29 figures from **my own** exports.

3. Per-Atom coverage

Atom	State in my run	What I verified	Gates I judged genuine
atom-000001	partial , 140/223 cells	Whole cohort chain against <code>clean_13:66-169</code> + <code>analysis_2:35-45</code> ; B1 against <code>analysis_1:9-71</code> ; all 18 Table 1 rows and the pooled t-test against <code>analysis_2:154-176</code> ; B2 against <code>analysis_2:97-112</code> ; the <code>attach_supported</code> component isolation by running with dictionaries/IND/IDAP/UDDA removed	C02 (under-six row only), C03 (B3), A04 (B4, bin 8, prose), A11 (simultaneous deaths, missing cause)
atom-000002	boundary	Cohort + <code>construction.prescriptions</code> run locally; <code>models.eq1</code> verified against dense HC1 OLS; C04 stops <i>before</i> estimation, and the unscaled coefficients are not substituted into the percentage display	C04

Atom	State in my run	What I verified	Gates I judged genuine
atom-000003	boundary	Rank formula vs clean_11:59-79; construction.jobs main-job selection proved equivalent to clean_11:229-277 by case enumeration; candidate_control vs clean_15:75-160,274 including the non-obvious drop if age2<24 & treatment==0; the table_b3 row roster (15 targeted + 9 non-targeted + N = 50 cells) vs analysis_15:114	C01/A01/A02, A03, C05 (B7), C09 (B19), C10/A08 (B21/B22), A07 (B6 age strata, row-level), A05 (psychiatry row only), C02 (sunder-six row only), U03 (prose)
atom-000004	boundary	construction.hosp (defect, §4.1/§4.2), services vs clean_4, prescriptions vs clean_5, covered_join zero-only-in- covered-years rule	C01/A01/A02/A03, A05, U02

Atom	State in my run	What I verified	Gates I judged genuine
atom-000005	boundary	<code>construction.onset</code> C01/A01/A02, line-by-line vs <code>analysis_14:246-256,365-379</code> (from <code>rose_bios</code>) including <code>drop if time==5</code> , the <code>fillna(0)</code> on the first lag, and start-of-decline semantics. All ten registered figure4/figure_b12 panel outcomes are constructed	C01, A01/A02, C05, A03
atom-000006	boundary	<code>enrollment</code> codes vs <code>clean_9:20,80</code> and the BOERNFB/BOERNSB split; <code>family_counts</code> twins/sentinel logic vs <code>clean_7:110-134</code> and <code>clean_13:148-155</code>	C06, A06
atom-000007	boundary	Hospital-day and prescription inputs constructed; no illness stratum is assigned anywhere — I grepped for it	C07, C05, U03
atom-000008	boundary	<code>construction.wealth</code> vs <code>clean_12</code> , including the correct refusal to take a numeric max when a (PNR, year) group mixes missing and non-missing net assets (Stata would pick the missing one)	C08, U04

Atom	State in my run	What I verified	Gates I judged genuine
atom-000009	boundary	Eq4 kernel with explicitly supplied post times; no descriptive cap-5 leaks into the family child count	C01/A01/A02/A03
atom-000010	boundary	<code>spouse_histories</code> fixes the event partner before expanding the spouse's calendar history; C01 correctly <i>not</i> applied (separate spouse matching), A09 applied	A09, A01/A02/A03
atom-000011	boundary	<code>occupation_codes</code> vs <code>clean_11:84-155</code> ; A10 correctly scoped to the occupation columns only (<code>analysis_13:15-23</code> really does sort <code>bysort PNR match_id:</code> with no sort variable), firm/public survive	A10 (row-level), C01/A01/A02/A03

Gate scoping is narrow, not blanket. C02 nulls only the under-six rows; A05 only `table_b3`'s psychiatry row; A07 only `table_b6`'s age strata; A10 only `table_b5`'s occupation columns. I confirmed this by reading the emitted `gap_ids`, not the prose.

I specifically probed the one case where over-blocking looked plausible — `figure_b20`, which atom-000003's protocol describes as “four mutually exclusive baseline categories”. Its frozen CONFIG settles it: unit `Index points`; `baseline 100`, `ci: True`, panels Men/Women. It is a treatment-effect display whose x-axis happens to be baseline proximity, so the C01 and A03 gates on it are correct.

4. Defects

4.1 `construction.hospital` drops invalid-interval admissions from the visit count (medium)

`clean_1_hospital_visits.do` collapses the annual visit/screening/nonscreening counts at line 51, *inside* a `preserve/restore` block, and only drops `D_INDDT0>D_UDDT0` at line 74, for the daily expansion. `construction.hospital` applies that restriction first, so it removes those admissions from the counts too.

`scratch/r01/opus/probe/p1_hospital.py`, hand oracle:

```
A visits rows: [] | source expects hospital=1 for p1/2000
B visits rows: [] | source expects hospital=1 (record with D_UDDT0
↪ 2000-01-08)
D visits: 2017 hospital=2, 2018 hospital=1
D days : 2017 12, 2018 27 <- matches my hand calculation exactly
```

Case D is the control: the interval-union sweep and the calendar-year split are correct, including the 2018/2019 clamp. Only the ordering is wrong. Neither case A (one record, no tie) nor case B (two records with distinct non-missing discharge dates, so `_n==_N` is deterministic) depends on any inference about Stata's sort order for missing values — that reasoning appears only in §4.2. And the rule is deliberate, not incidental: line 74 exists precisely because the authors contemplated `D_INDDT0>D_UDDT0` rows, and they put it *after* the visit collapse. A bad discharge date still counts as a hospital contact while contributing no days. That is source-explicit, not an A11 ambiguity, which is why the fix reproduces the ordering rather than adding a Boundary stop the way §4.2 does.

Fix. Do the (`PNR`, `D_INDDT0`) latest-discharge selection and the admission-year filter first, compute `visits` there, and apply `D_INDDT0 <= D_UDDT0` only to the day sweep. **Regression test.** A lone admission with `D_UDDT0 < D_INDDT0` must give `hospital=1` for its admission year and no `hospital_days` row.

4.2 `construction.hospital` silently drops missing discharge dates (low)

`records.dropna(subset=["PNR","D_INDDT0","D_UDDT0"])` discards them. In Stata missing sorts last, so `clean_1:20` would *select* exactly that record. Every sibling constructor in the same module (`deaths`, `prescriptions`, `services`) raises `Boundary(...,"A11")` on an equivalent missing value; this one does not, and A11's disposition is "stop the consuming target; do not impute or choose a new tie rule."

Probe case C: given two records for one admission date, one with a real discharge and one with a missing one, the code silently returns the *other* record's diagnosis and 3 days, with no gap row anywhere.

Fix. Raise `Boundary("Hospital record without source discharge date",["A11"])`. **Regression test.** Probe case C must return a Boundary.

4.3 CSV and DTA/Parquet inputs are not dtype-equivalent (medium)

`Inputs._read` reads CSV with `dtype=str`, `keep_default_na=False`, but hands back native dtypes for `.dta` and `.parquet`. `Inputs.table` only coerces the columns listed in `NUMERIC` and `DATES`. Everything else — the *code-valued* fields — keeps whatever dtype the file happened to have, and the downstream comparisons transcribed from the author’s Stata are string literals.

`scratch/r01/opus/probe/p2_format.py`, one AKM fragment written as both formats from the identical frame:

```
CSV dtypes DISCOLOEN_INDK object -> occupation ['<NA>', '2511']
DTA dtypes DISCOLOEN_INDK int32 -> occupation ['9999', '2511']
```

The 999999 “no occupation” sentinel (`clean_11:148–151`) is not masked on the DTA path, and `.astype("string").str[:4]` then manufactures a plausible four-digit occupation code. Nothing raises. The same class covers `jobs` (`TYPE`, `TYPE_2008`, `EJERKO`, `ARB_SEKTORKODE`), `services` (`SPECIALE` vs the 22 six-character GP codes — a float column becomes "800100.0"), and `population` (`CIVST`). `README.md` says DTA/Parquet “was not exercised”, which discloses the lack of testing but not the divergence.

Fix. In `Inputs.table`, coerce every mapped logical field that is neither `NUMERIC` nor `DATES` to pandas `"string"` dtype, or reject a non-string dtype with an `IntegrityError` (safer, since leading zeros cannot be recovered). **Regression test.** Build one fixture as `.csv` and `.dta` and assert `occupation_codes`, `jobs` and `services` return identical frames.

4.4–4.5 Nits

- Every NN-gated reason reads `STOP_UNRESOLVED_SOURCE: STOP_UNRESOLVED_SOURCE C01/A01/A02: ...` (`analyses.py:38` + `analyses.py:100`), visible in the rendered `figure_b20` footnote.
- `table_b3` emits its targeted rows as `... employment, earnings, atp, public` while both the protocol and `analysis_15:114` order them `... employment, ATP, earnings, public`. No numeric consequence — the interface is keyed, not ordered.

5. What checked out (the part that decides “revise, not reject”)

The risky **Atom 1** rules are all genuinely source-explicit. The one I most expected to be an invented default — `below_high_school = education.eq(1) | education.isna()` — is literally `replace hs_drop=1 if educ_level_str==""` (`analysis_2:164`, repeated at 198, 231, 269). Winsorization really does happen at `analysis_2:67–69`, before the `first_death_age_at_death<=100` and `years_from_first_death==–1` restrictions at lines 137–138, exactly as the code does it. Child count really is capped at 5 (`analysis_2:51`) while `age_of_youngest` is the *uncapped* variable (`clean_7:127`, distinct from `age_of_youngest_int`’s 21/22 sentinels at `clean_13:152–155`). The intensive margin really is ATP conditional on

positive (`analysis_2:155`). B1’s denominator really is computed before the `count<5` suppression (`analysis_1:62-67`), and the 40-45/40-44 label discrepancy is preserved as no-estimate rather than repaired.

The model kernels are independently correct (`probe/p3_models.py`, statsmodels 0.14.4): HC1 $\max|\text{diff}|$ 4.4e-16 and covariance 6.2e-18 against a dense non-redundant OLS with matching `df_resid`; supplied-metadata CR1 covariance 9.6e-18 against statsmodels cluster `use_correction=True`; matched Eq2 2.1e-14 and CR0 covariance 1.5e-15 against a fully dense LSDV fit with `person×match`, time, year and age dummies. The default clustered path is `CR0_not_source_inference`, refuses to emit intervals (`gaps A01/A02`) and returns `F=None`, `p=None`. A collinear treatment column stops rather than pseudo-inverting. `eq3` without `post_times` stops with A03. **A numerically runnable substitute model is not being passed off as the source estimator anywhere.**

Registration identity behaves exactly as challenge 6 requires. Editing an Atom’s activation → Frozen protocol or membership altered. Editing a figure CONFIG → Frozen material altered: `prereg/paper/figures/figure_b1.py`. Editing `figures.yaml` → rejected. Editing the *mutable* realized Atom front matter/body and the Molecule interpretation in place → still verifies and still runs. `frozen-identity.json` hashes 108 materials including every `figure*.py`; the 14 unhashed files are exactly the 11 Atom bodies, the Molecule body, `manifest.yaml` and `CHECKSUMS.sha256`, whose scientific content is covered by `registration_digests`.

Direct dispatch bypasses only the outer activation. A receipt whose `registered_state` contradicts its own rows is rejected under *both* policies. A synthetic receipt offered as governed evidence is rejected. A terminal `no_result` is not retried under either policy. An unknown manifest key (“scientific overrides are not authorized”), an out-of-scope table, a non-empty output directory and an output path inside the World all fail closed with non-leaking messages. Activation is native strong Kleene — NOT `done(x)` with no evidence is `None`, not `True` — and `activation.py` is byte-identical to the repository `native.py` evaluator apart from four removed unrelated constants.

Missing optional inputs degrade component-wise, not catastrophically. Removing both dictionaries → 108 available cells; also removing IND/IDAP/UDDA → 92; supplying only BEF and the two death extracts → B1 alone, 4 cells, with FAIN’s absence reported as FAIN1981–1985 needed for original 1980–1984 panel rather than substituted.

Figures. All 29 registered figures render from my own CLI exports, 0 failures, every page 130.70 mm wide. I independently reproduced the author’s disclosed presentation problems rather than taking them on trust: `figure_b12` 391.31 mm, `figure_b7` 388.11 mm, `figure_b10` 382.42 mm and `figure3` 320.55 mm tall (not placeable on an ordinary manuscript page); `figure_b20`’s four category labels overlap; `figure_b3` draws a default 0.0–1.0 axis under an age label. Every empty panel prints “No available estimates” with its exact reason and a footnoted gap. Nothing is fabricated to hide the boundary. The user prohibited template changes, and the author flagged these for the presentation owner instead of claiming publication readiness — the right call.

A note on the author’s self-report. Its two most checkable claims hold up independently: 39/39 tests pass in 171 s with 145 renders, and the 122 original files are unchanged. Its flagged `table_b3` count discrepancy is real — `analysis_15:114` lists 15 targeted variables, `target-inventory.md:21` says 16 — and the implementation emits the 15 named rows without inventing a sixteenth. That is the correct handling of a source disagreement.

6. A systematic behaviour worth stating

The bundle repeatedly converts a Stata “arbitrary but deterministic” rule into a hard stop: force-dropped income duplicates, missing IDAP PNR, unmapped education levels, main-job priority ties, same-admission diagnosis ties, simultaneous parental deaths, missing death causes. Each is authorized by A11/A01/U02, and stopping is the conservative direction (it never invents a value). But at governed runtime a *single* offending row can zero out an entire component. `attach_supported` keeps that component-scoped for Atom 1 — I verified that losing IND still leaves 92 cells — so this is a documented design choice, not a defect. It is worth the runtime operator knowing about. Findings 4.1 and 4.2 are the two places where the code departs from that otherwise consistent conservatism in the *unsafe* direction: it silently drops rows instead of either following the source or stopping.

7. Limitations of this review

- Stata is unavailable. `source_models.do` reads coherently against `models.py`’s stated conventions (`invttail(e(df_r),.025)` pooled, literal 1.96 dynamic, contrasts from the pooled interacted model’s joint covariance) but neither the author nor I executed it. No numerical validation of the source engine exists, and none is claimed.
- Parquet was not exercised at all. I exercised DTA only far enough to demonstrate 4.3.
- The evidence/predecessor machinery is unreachable in this World: all eleven activations are the literal TRUE, so `activation_respecting` and `direct` are behaviourally identical here. I exercised `load_evidence` by hand-constructing receipts.
- `proximity`, `scaled_eq1`, `immediate_ratios`, `labour_dynamics`, `health_dynamics`, `family_models`, `job_models` and `spouse_histories` have no reachable caller, because every consuming target is source-blocked. Their correctness rests on the synthetic tests plus my oracles, not on any integration path.
- I did not re-derive `clean_14/clean_16` line by line for Atom 10; A09 blocks every spouse target regardless.
- No real dictionary values exist, so every dictionary path ran with the clearly-labelled invented test dictionaries or with the dictionary absent.

8. Which numbers remain unavailable, and why

45 of the 50 registered source targets have no estimate. The cause is entirely the frozen `deviations.md`: C01/A01/A02 (matching strata, tie semantics, age/year normalization), A03 (post window), C04 (Eq1 counterfactual denominator), C05 (caption vs

program), C06/A06 (childcare eligibility), C07 (illness definition), C08/U04 (wealth units and absent variants), C09 (geography), C10/A08 (future-treated), A05, A07, A09, A10, U03. I checked each against its cited locator and found each genuine and target-scoped. On top of that, even once those resolve, **no source interval or p-value can be exported for any matched target** until a source-established effective parameter count and residual df are supplied — the code is right to block this, and it is right to label CR0 as **CR0_not_source_inference**.

The five targets that do produce values (figure_b1, figure_b2, table1, table_b1, table_b2) were computed from clearly-labelled synthetic fixtures only. They are engineering dispositions. They are not evidence about Denmark.

9. Requested action

Fix 4.1, 4.2 and 4.3 with the stated regression tests; optionally 4.4 and 4.5. No protocol change, no new experiment, no new research question, and no change to any frozen gate is requested or would be appropriate.

Proof artifacts (all outside the World)

scratch/r01/opus/action/	byte-identical copy of prereg/execution
→ used for every run	
scratch/r01/opus/suite-opus-01/	my run of the advertised 39-test suite +
→ 162 PDFs	
scratch/r01/opus/myrun/	my own CLI outputs for all 11 Atoms
scratch/r01/opus/adv/	adversarial
→ dispatch/evidence/identity/partial-input probes	
scratch/r01/opus/probe/p1_hospital.py	finding 4.1 / 4.2 counterexample
scratch/r01/opus/probe/p2_format.py	finding 4.3 counterexample
scratch/r01/opus/probe/p3_models.py	statsmodels oracles
scratch/r01/opus/renders/	29 figures rendered from my own exports +
→ PNG inspections	
scratch/r01/opus/xtab.py	protocol/code cross-tabulation

r02 / opus.md

Round-2 independent engineering / source-fidelity review — opus

Verdict: revise. One high defect, two low, three nits. Nine of the ten round-1 findings are genuinely fixed and I reproduced each against my own oracle rather than the author's.

This is engineering and source fidelity only. It is not scientific acceptance of the frozen registration, not a replication, and not evidence about Denmark. Every number below comes from clearly labelled synthetic fixtures.

World reviewed: v002/world. Exclusive scratch: scratch/r02/opus. Nothing was written into the World, the original package or any other reviewer's outputs. I did not read scratch/r02/codex or reviews/r02/codex.yaml.

1. What round 2 actually changed

`diff -rq v001/world v002/world` gives a narrow surface: `construction.py`, `contracts.py`, `models.py`, `analyses.py`, `run.py`, `run_tests.py`, `README.md`, plus three new files (`validation.py`, `round2_tests.py`, and the World-side engineering/`launch.py` + `execution-manifest.json`) and the two engineering reports. No frozen scientific material changed: the 122 original files are byte-identical to `.controller/source`, and no v001 test was deleted or renamed (39 → 49).

My review therefore concentrates on **the ten fixes and what they broke**, not on re-deriving the untouched constructors I already verified in round 1.

2. The headline defect — OPUS-R02-01

R01-CODEX-01 asked for a trusted launcher that compares the copied action against a World-side manifest before importing it. The author built exactly that, and where it looks it works: I mutated `construction.py`, `contracts.py`, `validation.py`, `analyses.py`, `run.py` and `frozen-identity.json`, added a file, deleted a file, planted a real `__pycache__` and symlinked a file — all rejected with `rc=2` and “*Copied execution differs from trusted World execution manifest*”. Stripping the bootstrap out of `run.py` is caught too, because `contracts.verify_world:47` calls `implementation_identity` independently. The digest plumbing is real, not vacuous: `contracts.digest` and `launch._digest` are identical implementations and `code_identity() == verify_action() ==` the manifest’s `83a78afd539e026c621e6ad47393dc9f2d2278cf924c41d4390bbf33901a562e`.

The manifest authenticates .py source. CPython executes bytecode. Two bytecode paths sit outside the tree `verify_action` scans:

Variant A — symlinked cache directory. `launch.py:33` enumerates with `Path.rglob("*")`, which on CPython 3.12 does not descend into a symlinked directory, and `is_file()` is `False` for the symlink itself. I printed what `verify_action` actually sees:

```
__pycache__  is_file=False is_symlink=True is_dir=True
```

That single entry fails the `is_file()` test at lines 34 and 36, so the explicit `__pycache__ / .pyc` guard *and* the hashing block are both skipped, and the `.pyc` inside is never enumerated. I compiled the copied `analyses.py` with codex’s exact round-1 mutation (`analyses.py:159`, `100 * → 200 * cumulative[i] / total`) using `py_compile.PycInvalidationMode.UNCHECKED_HASH` — which CPython loads without ever consulting the source — and symlinked `<action>/__pycache__` to it.

```
verify_action(W, action) ->
↳ 83a78afd539e026c621e6ad47393dc9f2d2278cf924c41d4390bbf33901a562e
python <action>/run.py --atom atom-000001 ... -> exit 0,
↳ registered_state=partial
figure_b1 available cells, trusted : [20.8333, 41.6667, 62.5, 96.875]
```

```
figure_b1 available cells, injected : [41.6667, 83.3333, 125.0, 193.75]
receipt implementation_id, BOTH runs : 83a78afd...
```

Two of those cumulative-distribution shares exceed 100 %. Running the same action through the **trusted World launcher** (`launch.py --world ... --action ... --entry run.py`) produces the identical doubled values. Feeding the resulting receipt to `load_evidence` from a separate, clean, manifest-verified action returns `{'atom-000001': 'partial'}` — it clears the external digest binding, the `implementation_id` binding and the whole new `validate_bundle` roster/status/uncertainty/projection validation, because only the cell **values** changed.

Variant B — relocated cache. `PYTHONPYCACHEPREFIX=<dir>` moves the byte-code cache out of the action directory entirely. With the same unchecked-hash `.pyc` pre-placed under the mirrored prefix path, an **unmodified** copied action run through `launch.py` also exits 0 and produces `[41.6667, 83.3333, 125.0, 193.75]`. No symlink, no write access to the action. `sys.dont_write_bytecode = True` (`launch.py:9`, `run.py:4`) stops *writing* a cache; it does not stop *reading* one.

Neither variant touches the launcher or the manifest, so this is inside — not outside — the author's stated trust boundary ("the operator must protect the World launcher/manifest"). The same blindness exists in `contracts.code_identity` (lines 289-292), which is why the receipt's self-reported ID also matches. A symlinked plain subdirectory is likewise invisible (`rc=0`); that one has no live exploit but is the same root cause.

Note that the R01-CODEX-02 guard is not a second line of defence here: `validation.expected_rows:16-19` imports `execute` from the *same* `analyses` module the run uses, so the roster gate would compare injected output against an injected oracle. Both round-2 guards share this single point of failure.

Narrow fix. In `verify_action`, replace `Path.rglob` with `os.walk(action, followlinks=False)` or recursive `os.scandir`, and raise for **any** entry that is a symlink or not a regular file — directories included — *before* the `__pycache__/.pyc` test. Additionally fail closed when `sys.pycache_prefix` or `PYTHONPYCACHEPREFIX` is set, since a relocated cache can never be covered by an in-directory manifest. **Placement matters:** that check must run in `run.py`'s bootstrap and in `launch.main` *before the first import of any action module*, not only inside `verify_action` — `contracts.implementation_identity` calls `verify_action` only after `contracts.py` has itself been imported, and a `verify_action` reached through an already-poisoned `contracts` is not a trust anchor. (Both my variants inject `analyses`, imported after the bootstrap, so the in-`verify_action` check would close them; stating the ordering closes the class.) Mirror the enumeration in `code_identity`. Regression tests: (a) symlinked `__pycache__` with an unchecked-hash `pyc` from a mutated `analyses.py`, (b) the same with `PYTHONPYCACHEPREFIX`, (c) a symlinked subdirectory — each rejected before any action module is imported, with a marker file proving no injected code ran.

3. Per-Atom coverage

All eleven Atoms executed under both dispatch policies from a `cp -a` copy of `prereg/execution`; the 22 `results.csv` files are byte-identical between policies, as expected when every registered activation is literally `TRUE`.

Atom	cells	available state	targets and what I checked this round
atom-000001	223	140 partial	figure_b1/b2/b3/b4, table1/b1/b2, prose_01/02. Roster gate under 4 partial-input manifests (108 / 92 / 4 / 92 available); B1 suppression and the A04 bin-8 row; the figure_b1/figure_b2 annotation_statistics == [] fact that makes the new validate_bundle roster check self-consistent.
atom-000002	161	0 boundary	figure_b5/b23 (C04), prose_13 (no gap — contextual, not empirical).

Atom	cells	available state	targets and what I checked this round
atom-000003	669	0 boundary	table_b3 row order fixed to analy- sis_15:114 (...employ- ment, atp, earnings, public) and gaps now C01;A01;A05;C02 with no A02. figure1 annotation guard probed with 7 mutations. Whole- bundle mutation probes run against this Atom's output.
atom-000004	337	0 boundary	local_health → the fixed construction.hospital (visits before the interval filter).

Atom	cells	available state	targets and what I checked this round
atom-000005	274	0 boundary	same hospital input; prose_08 still carries A02 though it is a ratio of onset counts (OPUS-R02- 03).
atom-000006	116	0 boundary	enrollment component; gaps C06/A06 unchanged.
atom-000007	56	0 boundary	hospital-day input; prose_14 still carries A02 though it is a share (OPUS-R02- 03).
atom-000008	25	0 boundary	wealth/inflation component; C08/U04 unchanged.
atom-000009	36	0 boundary	figure_b17; gaps unchanged.
atom-000010	18	0 boundary	figure_b18 correctly keeps A09;A01;A02;A03 via SPECIAL even though module 10 is excluded from the NN branch.

Atom	cells	available state	targets and what I checked this round
atom-000011	24	0 boundary	occupation_codes and jobs re-verified through all three input formats.

Gap IDs for all 50 targets were dumped from my own runs and checked line by line against `prereg/replication/deviations.md`. Every gate is genuine and target-scoped except the two A02 over-attachments in OPUS-R02-03.

4. The nine fixes, verified against my own oracles

OPUS-R01-01 / -02 (hospital). `construction.hospital` now does latest-discharge selection → admission-year filter → visit groupby → *then* the reversed-interval drop, which is `clean_1_hospital_visits.do` lines 20 → 26-27 → 50-51 → 69 → 74 → 78-86 exactly. My eight hand-computed cases:

case	source expectation	v002
A one reversed interval	hospital=1, no day row	(v001 returned an empty frame)
B two reversed records, same date	hospital=1 with the max-discharge diagnosis DZ001 → screening=1	
C missing discharge beside a valid sibling	stop (Stata sorts missing last, then <code>expand .</code> fails)	<code>Boundary(..., ["A11"])</code> ; v001 silently substituted the sibling
D valid cross-year stays	2017 v=2 d=12; 2018 v=1 d=27; 2019 clamped	unchanged
E tie on the maximum discharge	stop	A11
F psychiatric reversed interval	counts — <code>clean_2</code> has no interval filter at all	
G DZ001 in 1990	not screening (<code>clean_1:36</code> requires year 1994)	
H 2019 admission	excluded (<code>clean_1:26-27</code>)	

OPUS-R01-03 (formats). `contracts.source_strings` rejects a non-string code field rather than inventing zero padding — the conservative branch I asked

for. My own counterexample (the DISCO sentinel 999999 written as an integer) is now rejected on `.dta` and `.parquet`; zero-padded AKM/IDAN/IDAS fixtures give byte-equal `occupation_codes` and `jobs` frames across all three formats (`occupation` [`<NA>`, `'2511'`, `'6543'`], `public` [`1,0,0`], `firm_change` [`0,1,0`]), so the zero-padded LBNR join survives. Four code fields outside the author's suite — `LMDB.ATC`, `DAGI.YDELSE`, `IDAN.LBNR`, `BEF.E_FAELE_ID` — also fail closed. HFAUDD remaining in `contracts.NUMERIC` is **not** an inconsistency in that fix: `clean_10:25-26` compares `HFAUDD==.` and `HFAUDD==1` numerically, while `AUDD_LEVEL_L1L4_K=="0"` at lines 51-60 is a string comparison, matching `construction.education:218`.

R01-CODEX-03 (matched kernel). The structural gate now precedes outcome and sex deletion, which is precisely `data-contract.md D1`: *"Balanced 11-year means exactly one valid person-year at each $t=-5$ through 5 ... a mismatch stops that target, it does not license deduplication."* I reproduced codex's counterexample (a duplicated `PNR/match_id/year` with a missing outcome) — it stops under `sex=1`, without a sex split, and when both outcomes are present. Also stopping: an unbalanced person-match, a duplicated event time, a treatment flip inside a person-match, an inconsistent event clock, and `NaN PNR/match_id`. The estimator is untouched: `eq2` betas agree with a dense LSDV `statsmodels` fit on the identical post-singleton sample to **1.1e-13** ($n=262$).

The author's self-corrected test is right, and I re-derived it independently. Their note says the new partial-outcome test *"incorrectly expected exactly one deleted row; source iterated singleton removal deletes an additional observation."* In `fixtures.synthetic_matched()` the `age=28` cell has exactly one observation, so `reghdfe`-style iterated removal already takes `528 → 527` on the clean panel; blanking row 0's outcome gives 526, with dropped indices exactly `{0, 187}`. The code was right and the original expectation was wrong. (The replacement assertion is now too loose — `OPUS-R02-06`.)

R01-CODEX-05 / OPUS-R01-05 (table_b3). Row order matches `analysis_15:114`; `A02` is off `table_b3`; and `grep -c "STOP_UNRESOLVED_SOURCE: STOP_UNRESOLVED_SOURCE"` over all 22 of my exports matches nothing (`OPUS-R01-04` fixed). The remaining two-marker rows are two *distinct* clauses joined by `" | "`, not a doubled prefix.

R01-CODEX-02 (complete vector). `validate_bundle` rejects a deleted B1 row, an added row, a changed series, a changed unit, a changed uncertainty, a stale target-status, a stale effects projection and an omitted projection hash; it correctly accepts a row-order-only shuffle, since `D8` keys the CSV. The residual is `OPUS-R02-02`.

R01-CODEX-04 (annotations). `validation.render_checked` rejected all seven mutations I threw at the actual `atom-000003 figure1` table (unregistered series / term / statistic, wrong unit, wrong panel, wrong `target_id`, duplicate key), still permits an empty table as missing input, and refuses output inside the World. I independently confirmed the author's own disclosure that the frozen `source_render.py` is unchanged and still writes a PDF when `figure1.render()` is called directly with an unregistered annotation — the guard lives only at the execution interface, which is honest and explicitly scoped given the prohibition on editing frozen templates.

5. My other targeted tests

- **Advertised suite through the documented entry point:** 49 tests, 259.280 s, 0 failures, 0 errors, 145 renders, `stata_numerical_validation`: "not run; Stata unavailable".
- **Copied-layout CLI run twice:** both exit 0; no `__pycache__` or `.pyc` left behind.
- **Partial inputs:** dictionaries removed → 108; IND/IDAP/UDDA removed → 92; BEF+deaths only → 4; both removals → 92. All keep the full 223-row roster, so the new roster gate rejects no legitimate partial output.
- **Registration identity:** protocol activation, `figure_b1 CONFIG`, `figures.yaml`, the manifest's `implementation_id` and one manifest file hash are all rejected; appending to the mutable realized Atom body and to the Molecule interpretation still runs to `partial`.
- **Dispatch:** output inside the World, non-empty output dir, unknown manifest key, out-of-scope table, synthetic-as-governed receipt, `accepted=false`, and a terminal `no_result` under **both** policies all fail closed with a non-leaking `IntegrityError`.
- **Activation:** `activation.py` is AST-identical to the lane's `native.py` for all five shared definitions. `NOT done(...)` with no evidence → `None`; `unknown OR TRUE` → `True`; `unknown AND FALSE` → `False`; `import os, atom-1, __import__("os")`, `TRUE TRUE` and `Done(atom-000001)` all rejected by the grammar.
- **Figures:** 29/29 render from my own CLI exports, every page 130.700 mm wide.
- **Numerics:** HC1 2.1e-16 / cov 2.4e-17, supplied-metadata CR1 cov 2.9e-17 vs `statsmodels`; default clustered path is `CRO_not_source_inference` with `df=None`, refuses intervals/F/p; collinear treatment stops; `eq3` without post times, `immediate_ratios` without an immediate year and `scaled_eq1` with a zero denominator stop with A03/A03/C04.
- **World integrity:** 122 original files byte-identical; no unexpected World file; the sha256 of `prereg/execution` and `prereg/engineering` unchanged before and after every probe.

6. Remaining findings

id	sev	what	narrow fix
OPUS-R02-01	high	Bytecode outside the hashed tree defeats the new execution-identity anchor (symlinked <code>__pycache__</code> ; <code>PYTHONPYCACHEPREFIX</code>) and the fabricated receipt now carries the <i>trusted</i> implementation id.	Reject any symlink or non-regular entry in <code>verify_action</code> 's walk; fail closed on a set pycache fix; mirror in <code>code_identity</code> .
OPUS-R02-02	low	<code>validate_bundle</code> binds keys/units/status/ units/units/objects but not <code>gap_ids</code> , <code>reason</code> or <code>source_locator</code> ; a re-signed bundle whose every boundary reason reads “ <i>No effect detected: treatment and control equivalent</i> ” with empty gap IDs is accepted (state stays <code>boundary</code> , no number appears). A02 still reaches <code>prose_08</code> (a ratio of onset counts) and <code>prose_14</code> (a share), neither of which is a matched Eq2-4 estimate under <code>deviations.md</code> A02.	For modules 2-11 the gap set is uncertainly/objects — recompute <code>target_boundary</code> and require the exported <code>gap_ids</code> to match. That recomputation is load-bearing; a weaker “STOP marker implies 1 ledger ID” rule does <i>not</i> catch this, since the relabel removes marker and IDs together.
OPUS-R02-03	low	A02 still reaches <code>prose_08</code> (a ratio of onset counts) and <code>prose_14</code> (a share), neither of which is a matched Eq2-4 estimate under <code>deviations.md</code> A02.	Explicit A02 target set (must keep <code>prose_06</code> 's linear-controls robustness and <code>figure_b18</code>) + a per-target gap-ID regression over all 50 targets that asserts retention as well as removal.

id	sev	what	narrow fix
OPUS-R02-04	nit	<code>models.matched:105</code> raises a bare <code>TypeError</code> from <code>np.isfinite</code> on an object/string structural column instead of a <code>Boundary</code> .	Check <code>is_numeric_dtype</code> first and raise <code>Boundary(..., ["A11"])</code> .
OPUS-R02-05	nit	A World without <code>engineering/launch.py</code> makes <code>run.py</code> exit 1 with a full traceback and absolute paths, contradicting its own no- <code>traceback</code> policy and the documented <code>rc=2</code> JSON.	Move <code>launch.py</code> module inside the existing <code>try</code> .
OPUS-R02-06	nit	<code>assertLess(fit.n,assertEqual(len(clean)len(clean)) no longer pins the estimation sample (should be 526).</code>	<code>- fit.n, 2)</code> with the age-28 singleton named.

7. What remains unavailable, and why

45 of the 50 registered targets still produce **no number**, entirely because of the frozen deviations ledger and unavailable inputs — C01/A01 matching, A02 normalization, A03 post window, C04 Eq1 denominator, C05, C06/A06, C07, C08/U04, C09, C10/A08, A05, A07, A09, A10, C02, C03/A04, U03, and U01/U02 runtime inputs. Nothing in this review resolves any of them, and the code correctly declines to. The five targets that do produce values are Atom 1 descriptive cells computed from clearly labelled synthetic fixtures.

Also unresolved and explicitly *not* engineering blockers: Stata / `reghdfe` / `teffects` are absent, so `source_models.do` and the `reghdfe` small-sample convention remain unvalidated; input provenance (`kind`, `vintage`, `namespace`, `correspondence`) is an unverifiable operator assertion, and flipping `kind` to `governed` over synthetic fixtures runs to completion and removes the SYNTHETIC banner; receipt authenticity ultimately rests on the operator's external digest binding, since no check can validate a number without data; and the disclosed presentation limits reproduce under my own measurement — `figure_b20` has 6 adjacent word-box overlaps (3 per panel), `figure_b3/figure_b4` keep a default 0.0-1.0 axis under an age label, and `figure3 320.548 mm / figure_b10 385.269 mm / figure_b7 388.112 mm / figure_b12`

391.312 mm are not placeable on an ordinary page. The user forbids changing the frozen templates, so these stay disclosed rather than fixed. The frozen inventory's 16-vs-15 `table_b3` row-count contradiction is likewise still open; the code emits the 15 named rows and invents no sixteenth.

Proof artifacts are all under `scratch/r02/opus` (`action/`, `suite-01/`, `myrun/`, `probe/`). A self-contained reproduction of OPUS-R02-01, starting from a fresh `cp -a of prereg/execution`, is at `scratch/r02/opus/probe/REPRO-OPUS-R02-01.sh` `<world>` `<scratch>`.

r03 / codex.md

Codex round-3 independent engineering review

Verdict: **revise**. The candidate is substantially stronger than the prior rounds: it is a real eleven-Atom implementation, not a set of missing-input wrappers; it keeps unresolved science behind explicit boundaries; and its copied-layout, integrity, result-vector, activation, and rendering tests pass. Two narrow source-fidelity defects remain and both have independent, hand-constructed counterexamples.

This is an engineering judgment only. I did not observe Danish register data, accept the frozen scientific protocol, or infer any empirical effect.

Scope and evidence basis

I read the eleven complete Atom protocol/result documents, the full Molecule policy and membership, the shared D1-D8 contract, identification/governance/interpretation documents, all frozen deviation and matching records, the source inventory/provenance/schema/conversion records, all execution modules and tests, and the manuscript result/figure bindings. I also read the prior round findings and round-3 author response, without reading the other reviewer's round-3 output.

The exact local author program used below has SHA256 `397afe253637d66a2feb8eb85b2aa4353a4f43736956e5f748286264c0924ba5`, identical to `prereg/replication/code-inspection-index.csv:21`. No live source search was needed because the cited source snapshot and its registered hash were locally available.

The advertised suite was executed from a writable copy through the protected World launcher. It passed 56/56 tests and rendered 145/145 PDFs. I separately validated all 22 actual output bundles (11 Atoms under both dispatch policies), checked every registered checksum, confirmed no original-package differences outside `prereg/execution` and `prereg/engineering`, and exercised both help paths without data. The suite's actual activation-respecting result count was 1,939 cells: Atom 1 was partial with 140/223 available; Atoms 2-11 were boundaries with no numeric target cells. Passing those checks does not cover the two cases below.

Findings

CODEX-R03-01 — high: IDAP coverage metadata is ignored

The frozen contract requires no-estimate when a required register-year is absent (`prereg/data-contract.md:46-53`). The execution README is equally explicit: `complete_years` is the coverage assertion; missing complete coverage blocks the affected construction; absent/unverified years cannot become zero (`prereg/execution/README.md:75`). `Inputs.covered` exposes those years at `prereg/execution/contracts.py:251-255`.

The ATP branch does not consume them. `attach_supported` invokes `c.atp(inputs.table("IDAP"))` directly at `prereg/execution/analyses.py:123-125`. The immediately adjacent earnings branch does use `inputs.covered("IND")` at lines 126-133. `construction.atp` then turns source-missing ATP into zero and derives employment and conditional ATP at `prereg/execution/construction.py:271-286`. Every Atom reaches this shared branch through `local_demographics` (`analyses.py:374-379,382-456`), so its local-construction claim is broader than Atom 1 even though only Atom 1 currently exports these numeric summary cells.

My copied-layout test set both IDAP and IND `complete_years` to an empty list. The launcher returned `rc=0` and a valid `partial` receipt. Earnings correctly had zero numeric cells. Nevertheless, employment had eight numeric available cells and `atp_positive` had eight. The exact estimates/p-values are in `scratch/r03/codex/partial-coverage-output/results.csv:65-70,132-143,204-215`; the empty declarations are `partial-coverage-manifest.json:304,364`; and `local-construction.csv:5` still calls annual ATP `constructed_locally`. This is the forbidden transition from unverified coverage to evidence, not a missing-real-data complaint.

The narrow repair is to gate ATP on the exact IDAP years required by the consuming cohort/target before constructing or attaching it. A missing required year must yield a U01 boundary and null ATP/employment cells while leaving unrelated population/family outcomes intact. A regression should cover both `complete_years=[]` and one missing interior year, then assert null means and null pooled-test p-values.

CODEX-R03-02 — medium: post-2007 job fields cross the source boundary

The hash-matched author source reads IDAS ownership only from 1980 through 2007 (`clean/clean_11_employment_and_income.do:207-222`). Its post-2007 IDAN branches deliberately omit LBNR (:243-265). It merges early ownership at :297, calculates late public status from `ARB_SEKTORKODE` at :299-307, and calculates firm changes from the resulting identifier at :310-318.

The Python constructor does something different. `construction.jobs` merges every IDAS year and lets `public_early` override the sector rule (`prereg/execution/construction.py:482-496`). It also retains a supplied post-2007 LBNR and uses it in firm transitions (:497-503). That contradicts the

execution README’s own “early job/owner1980-2007” and “no inferred late LBNR” declaration (`prereg/execution/README.md:104`).

Two minimal hand cases in `scratch/r03/codex/adversarial_construction_tests.py` falsify the transcription:

- A 2008 primary job with sector 81 and a colliding 2008 IDAS `EJERK0=07` is private under the source’s sector rule (`public=0`), but Python returns 1.
- Adjacent 2007/2008 jobs with supplied `firm-1/firm-2` identifiers cannot be a source firm transition because the source does not load the latter identifier, so the source result is `[0,0]`; Python returns `[0,1]`.

The advertised job regression at `prereg/execution/run_tests.py:351-360` misses both branches: its IDAS data end in 2007 and its 2008 LBNR is missing. Restrict or reject IDAS after 2007 and force post-2007 LBNR missing, then add these two boundary fixtures. Current Table B.3/B.5 result rows remain unavailable under independent matching/model/A10 gates, so this is a latent constructor error—not a claim about observed table values.

Per-Atom coverage

Atom	Reviewer coverage and current synthetic disposition
atom-000001	Population, first-death clock, 1985/2014 and age-25/50 limits, balanced <code>-5..5</code> panel, B.1 denominator/suppression, B.2, Tables 1/B.1/B.2, optional outcomes and pooled tests reviewed. Actual fixture: 140/223 available, partial . CODEX-R03-01 directly corrupts ATP/employment availability when IDAP coverage is unverified. B.3/B.4/prose boundaries remain genuine.
atom-000002	Shared cohort/outcomes, prescriptions, Eq. 1/scaling kernel and B.5/B.23 bindings reviewed. Actual fixture: 0/161, boundary . C04 genuinely prevents substituting unscaled fitted paths for the requested percentage-of-counterfactual display; <code>prose_13</code> remains non-estimation. The shared ATP local status is subject to CODEX-R03-01.

Atom	Reviewer coverage and current synthetic disposition
atom-000003	Matching gates, ranks, jobs, health inputs, Eq. 2/3/4 algebra, clustered covariance, balance/robustness tables and all 10 figure/prose targets reviewed. Actual fixture: 0/669, boundary . The suite's independent dense OLS/statsmodels checks passed and CR0 was not promoted. C01/A01/A02/A03 and target-specific gates are genuine. Jobs are subject to CODEX-R03-02.
atom-000004	Hospital/prescription/service constructions, visit-before-interval filtering, discharge boundary, and Figure 3/B.9/B.10/B.11 interfaces reviewed. Actual fixture: 0/337, boundary ; local components execute. Matching/inference and A05/U02 boundaries correctly prevent numerical substitution.
atom-000005	Extended onset clock, lag construction, partial PSYK input, Figure 4/B.12 and prose ratios reviewed. Actual fixture: 0/274, boundary ; missing PSYK is localized while siblings construct. C01/A01/A03/C05 remain explicit, with no fake joint-onset results.
atom-000006	Child enrollment codes/years, youngest-child identity/tie construction, Figure 5/B.13/B.14 reviewed. Actual fixture: 0/116, boundary ; enrollment constructs. C06/A06 source conflicts remain narrow.
atom-000007	Surviving-parent linkage, hospital days, prescriptions, illness timing, Figure 6/B.15/B.16 and prose bindings reviewed. Actual fixture: 0/56, boundary ; no illness class is guessed across C07.

Atom	Reviewer coverage and current synthetic disposition
atom-000008	Wealth vintage/deflator construction, Table B.4 and three prose asset variants reviewed. Actual fixture: 0/25, boundary ; C08/U04 remain genuine unit/ownership/specification gaps rather than invented defaults.
atom-000009	Family/cohabitation path and Figure B.17 interface reviewed. Actual fixture: 0/36, boundary ; the shared sample/model gates remain explicit.
atom-000010	Event-partner/spouse calendar history, Figure B.18 and A09 restrictions reviewed. Actual fixture: 0/18, boundary ; direct dispatch does not manufacture spouse evidence.
atom-000011	Main-job firm/public and occupation constructors, Table B.5 roster/order and A10 reviewed. Actual fixture: 0/24, boundary . Local jobs execute but are wrong at the 2007/2008 boundary under CODEX-R03-02.

All eleven registered activation expressions are TRUE. The suite passed native unknown/false evaluator fixtures, direct dispatch, explicit `no_result` nonretry, missing-input runs, copied/relocated layouts, valid realized Atom/Molecule edits, and actual protocol/input/code tampering rejection. I found no fake predecessor evidence, silent retry, blanket branch forcing, or result-slot schema substitution in the reviewed paths.

Output and presentation checks

All 22 actual bundles passed `validation.validate_bundle`, including exact effects/annotation projections. The registration identity remained stable under legitimate realized-body and Molecule-interpretation changes while protocol/material tampering was rejected by the advertised regression. Registered checksums all passed and the immutable source versus reviewed World had no differences outside the authorized execution/engineering trees.

I rendered/opened selected actual interfaces: Figure 1, Figure B.4, Figure B.12, and the full-interface Figure B.20. Explicit unavailable reasons were visible and no fake omitted-reference estimates appeared. The known presentation limitations were reproduced: B.20 label collision, B.4's default 0-1 unavailable-bin axis, and very tall boundary pages (the selected Figure 1 and B.12 pages were about 272 mm and 391

mm tall). These are disclosed and not confused with scientific completion, but they remain publication-layout limits.

Tests and limitations

The exact commands and outcomes, including one reviewer path-assumption failure followed by the corrected 22-bundle validation, are recorded in `codex.yaml`. The independent proof artifacts are confined to `scratch/r03/codex`.

Stata, `reghdfe`, `teffects`, source ado/version behavior, and the original matching implementation were unavailable and not installed. No confidential data, real metadata truth, access authorization, crosswalk/dictionary truth, national-scale performance, or governed font binding was tested. Thus Atoms 2-11 still have no numerical scientific results; Atom 1's numbers are synthetic engineering fixtures only. The genuine C01-C10/A01-A11/U01-U04 source limits remain authoritative and do not imply null, equivalence, or failed replication.

r03 / opus.md

Round-3 independent engineering / source-fidelity review — Opus

Verdict: revise — one medium defect (unconditional), one low defect, two nits. Every fix requested is a narrow change to `prereg/execution/` or `prereg/engineering/` plus its regression test. Nothing here touches the frozen protocol, the frozen package, or the science.

World reviewed: `.prereg-runs/denmark-reference-code-20260906/parental-death/v003/world`
Scratch (all proof artifacts): `.../scratch/r03/opus/{action,suite-01,myrun,probe,renders}`
Nothing was written into the World. `diff -rq --exclude=execution --exclude=engineering` against `.controller/source` is clean (122 files byte-identical), the World holds 141 files with no `__pycache__/.pyc`/stray entry, and the `prereg/execution` and `prereg/engineering` directory digests are unchanged after every probe.

This is engineering and source fidelity only. It is **not** scientific acceptance of the frozen registration, **not** a replication, and **not** an observed empirical result. All numbers below come from clearly labelled synthetic fixtures.

1. What round 3 changed, and what I found

Round 3 answers the six round-2 Opus findings. **Five are completely fixed**, and I reproduced each against my own oracle rather than the author's tests. The sixth (OPUS-R02-02) is fixed for modules 2-11 but leaves a residual in Atom 1. Separately, the round-3 bytecode fix was applied to the action tree and to `launch.py` but **not** to the World-side figure modules, which is the medium finding.

Round-2 finding	Status	How I checked it
R02-01 bytecode outside the hashed tree	fixed for the action tree ; World-tree residual → OPUS-R03-01(b)	re-ran both exploit variants; 8 further mutations
R02-02 boundary relabelled as a null	fixed for modules 2–11 ; residual in Atom 1 → OPUS-R03-02	40 single-cell mutations + the whole-bundle counterexample
R02-03 A02 over-attached	fixed, narrowly	mechanical v002→v003 gap diff over all 50 targets
R02-04 structural dtype TypeError	fixed, no residual	dtype sweep incl. nullable Int64/Float64/boolean
R02-05 traceback leak on missing launcher	fixed	copied World with <code>launch.py</code> deleted
R02-06 loose sample assertion	fixed exactly as asked	read the new {0,187} / loss-2 assertions

Findings

OPUS-R03-01 — medium — `validation.py:150-177 (render_checked)`

(a) *The operative defect needs no tampering at all.* `render_checked` calls `validate_annotations(world, target, tables["annotations"])` at line 160 and passes `tables["effects"]` straight to the frozen template at line 171 without validating it. The annotation guard added in round 2 (for R01-CODEX-04) has no counterpart for effects.

`probe/p_effects_guard.py`, starting from my own `myrun/direct/atom-000003/figure1-effects.csv` (80 rows, **every** status `unavailable`, 0 numeric cells; the corresponding `results.csv` rows carry gaps C01;A01;A02 and C01;A01;A02;A03):

- I set 20 rows of one panel to `status="available"` with finite estimates and enclosing intervals and cleared their reason. **`render_checked` accepted** and wrote a 22,377-byte PDF of fabricated effects for a source-blocked target.
- **Control A** — the same fabricated numbers on rows left labelled `unavailable` are refused, by `source_renderer.py`'s own `ValueError: Unavailable targets cannot carry numerical estimates`. That check is in a **frozen** file and keys off the row's own status label, so flipping the label is enough to defeat the only guard on the path.
- **Control B** — the identical fabrication at the `results.csv` level *is* refused by `validate_results` (Source boundary disposition differs from frozen ledger).
- **Control C** — an unregistered annotation series *is* refused by `validate_annotations`.

So the author-side guard exists for annotations and for `results.csv`, and is simply absent for the effects table on the same entry point. `validate_bundle` does bind each figure projection to `results.csv`, but it is a separate call that `render_checked` does not make.

(b) *Defence-in-depth: the frozen guard (a) relies on is not itself authenticated at load time.* `render_checked` loads the World-side figure module with `spec.loader.exec_module` and puts the figures directory on `sys.path`, so it and `source_render/plot_style` resolve through CPython's cache-consulting import path. `verify_world` (`contracts.py:66-69`) hashes only the 108 files listed in `frozen-identity.json["materials"]` and never rejects an unexpected entry in a frozen tree, so a pre-placed `UNCHECKED_HASH .pyc` is invisible to it. On a **copied** World (`probe/p_worldcache{,2,3}.py`; the reviewed World was never written to) I compiled mutated copies with `py_compile.PycInvalidationMode.UNCHECKED_HASH`, asserted the World's `.py` bytes were still byte-identical to the reviewed World, and: `figure_b1`'s rendered title became `FABRICATED TITLE`; poisoning `source_render.py` — which the *frozen* `figure_b1.render()` loads through its own `exec_module` — removed the guard so that Control A above then rendered too; and `verify_world` returned normally in every case. The control that *does* work: editing `figure_b1.py` itself is rejected with `Frozen material altered: prereg/paper/figures/figure_b1.py`. `verify_world` also does not notice an added, unexpected `.py` file in the figures directory.

I checked the threat model rather than assuming it. `execution/README.md:19` says “The World may be read-only”, :21 “For governed execution use the independent launcher in the trusted, read-only World”, and :27 “The trust root is the protected World launcher/manifest supplied by the operator/controller”. The delivered World is `dr-xr-xr-x/r--r--` and my first attempt died with `PermissionError` until I `chmod`'d my own copy. Part (b) therefore requires violating the declared trust root — it is defence-in-depth, **not** a breach of the author's stated boundary, and I grade the finding on (a). Two further facts limit (b): `PYTHONPYCACHEPREFIX` is correctly rejected on this path too (because `render_checked` calls `verify_world` first), and the code does not write a cache into the World itself, because `contracts.implementation_identity` execs `launch.py`, whose line 9 sets `sys.dont_write_bytecode = True` as a side effect (confirmed with `PYTHONDONTWRITEBYTECODE` unset: `False` before `render_checked`, `True` after, no `.pyc`).

Fix: (a) guard the effects table the way `validate_annotations` already guards annotations — compare its (`target_id`, `panel`, `series`, `x`) keys and unit against `expected_rows(world, config["atom"])` for that target, and refuse any non-available row carrying a number plus any available row whose key is not registered; empty/partial tables must stay valid. (b) before `exec_module`, walk `world/prereg/paper/figures` with the same recursive `os.scandir/follow_symlinks=False` rule already used in `launch.verify_action` and `contracts.code_identity`, or (stronger, since it also covers the frozen module's own `exec_module(source_render)`) load the target with `exec(compile(path.read_bytes(), str(path), "exec"), module.__dict__)` as `run.py:49` already does for `launch.py`. Neither change touches a frozen file.

OPUS-R03-02 — low — analyses.py:484-497 (source_boundaries, Atom-1 branch)

The new ledger binding covers every module 2–11 cell and four of Atom 1’s nine targets. Twelve further Atom-1 cells carry frozen ledger codes on the input-dependent path and are unbound:

- `figure_b1 / A / All first deaths / x="40-45" / estimate` — A04
- the eleven under-six cells of `table1` (3), `table_b1` (4), `table_b2` (4) — C02

`probe/p_relabel_atom1.py` rewrites those 12 to `reason="No difference detected: source-equivalent categories, estimated null", gap_ids="", source_locator="Fabricated equivalence claim",` rebuilds every projection through `outputs.Results.write`, re-signs and externally rebinds the receipt. `validate_results` **accepts** and `load_evidence` returns `{'atom-000001': 'partial'}`. The control in the same script — the identical relabel on `prose_01`, which *is* bound — is correctly rejected with `Source boundary disposition` differs from frozen ledger.

These dispositions are not data-dependent in the ordinary sense: they are identical in my full-input run and in the suite’s `partial-education` and `relocated-inputs/partial-output` runs, and only in the `empty` run are they superseded by a U01 coverage boundary — presumably why they were left out. `execution/README.md` discloses the scope but calls the remainder “Data-dependent Atom1 missingness”, which these frozen A04/C02 conflicts are not. Impact is provenance only: no number is fabricated (`validate_results` still refuses numeric evidence on a non-available row) and no state changes.

Fix: bind these four groups conditionally — require `status == "unavailable"` and `gap_ids` nonempty containing either the ledger code (A04/C02) or U01 (the legitimate coverage supersession).

OPUS-R03-03 — nit — `analyses.py:49` vs `:90`. `figure_b18` is in A02_TARGETS but is module 10, which the enclosing guard excludes, so the entry is dead. Behaviour is correct because `SPECIAL["figure_b18"]` already supplies A02; my run confirms A09;A01;A02;A03.

OPUS-R03-04 — nit — `analyses.py:352-356` (`local_health`) called from `:412`. Atom 5 does not register `PSYK_ADM` (atoms 3 and 4 do) and its protocol never mentions psychiatric admissions, but the shared helper requests it unconditionally, so Atom 5’s `local-construction.csv` permanently carries `boundary | Missing runtime table PSYK_ADM | U01`. The row is unsatisfiable, not a fixture omission — supplying the table would be rejected by `contracts.Inputs.__init__:145` as out of the Atom’s registered scope. *Fix:* pass the table list from the caller.

2. Per-Atom coverage

All eleven Atoms execute end to end under **both** dispatch policies; `results.csv` is byte-identical between `activation_respecting` and `direct` for every Atom, as expected when every registered activation is literally `TRUE`. Totals: **1,939 keyed cells over 50 targets, zero duplicate** (`target_id,panel,series,x,statistic,term`) keys — so the new `result.rows.extend(source_boundaries(...).rows)` at

`analyses.py:480` introduces no overlap with the per-module rows. Counts are identical to my round-2 run.

Atom	State	Available / cells	Supported constructions (all <code>constructed_1</code> unless noted)	Round-3 ledger, I ran on it
000001	partial	140 / 223	6/6: B1 population, treated balanced-11 panel, family counts, ATP, real earnings, education	ledger-relabel counterexam- ple (defect), bound-target control, partial-input bundles, <code>figure_b1/figure_b2</code> renders, proto- col/CONFIG/figures.yaml/manifest tamper probes, both help paths, idempotent repeat run per-field ledger mutations rejected; <code>prose_13</code> confirmed non- estimation with empty gaps
000002	boundary	0 / 161	6/6 incl. mental prescriptions	

Atom	State	Available / cells	Supported constructions (all constructed_1 unless noted)	Round-3 checked; I ran on it
000003	boundary	0 / 669	13/13 incl. earnings ranks, main-job firm/public, LPR/PSYK visits, prescriptions, GP and psychology consultations	whole-bundle relabel coun- terexample now rejected ; table_b3 gaps C01;A01;A05;C02 (no A02); figure1 fabricated- effects render through render_checked (defect)
000004	boundary	0 / 337	11/11	per-field ledger mutations rejected; prose_07 A02 retention checked against deviations.md
000005	boundary	0 / 274	10/11 — PSYK_ADM row is the out-of-scope artefact (nit)	A02 removal from prose_08 ver- ified against target-inventory.md:57 ; dynamic-vs- summary A03 scope
000006	boundary	0 / 116	6/6 incl. child- level formal enrollment	per-field ledger mutations rejected

Atom	State	Available / cells	Supported constructions (all constructed_10 unless noted)	Round-3 checks I ran on it
000007	boundary	0 / 56	7/7 incl. surviving- parent hospital-day and mental- prescription inputs	A02 removal from prose_14 ver- ified against target-inventory.md:63
000008	boundary	0 / 25	6/6 incl. annual real wealth	per-field ledger mutations rejected
000009	boundary	0 / 36	5/5	per-field ledger mutations rejected
000010	boundary	0 / 18	6/6 incl. spouse calendar history	figure_b18 retains A02 through SPECIAL (nit: dead A02_TARGETS entry)
000011	boundary	0 / 24	7/7 incl. firm/publi- outcomes and four-digit occupation	per-field ledger mutations rejected; table_b5 gaps incl. A10

Every Atom therefore implements real supported construction; none is a wrapper that reports missing inputs.

3. My targeted tests

Distinguishing **my own executed** tests from unexecuted checks and unavailable software.

Executed by me

- **Advertised suite through the documented World-side launcher**, from a `cp -a` copy of `prereg/execution` into my scratch: `python -B <world>/prereg/engineering/launch.py --world <world> --action <scratch>/action --entry run_tests.py --output <scratch>/suite-01 → 56 tests, 0 failures, 0 errors, 285.363 s.`
- **All 11 Atoms × 2 policies** → 22/22 exit 0, byte-identical per policy, receipts all carry `implementation_id 4443f312...`, which equals my own recomputation of `_digest(files)` over the manifest's 15 entries and `verify_action()` over my copy.
- **probe/p_trust.sh** — pristine control plus 8 mutated actions, each through `run.py` and `launch.py`: symlinked `__pycache__` with `UNCHECKED_HASH` pyc (header flag asserted = 1), `PYTHONPYCACHEPREFIX` with a pre-placed pyc, symlinked subdirectory, `analyses.py` source mutation, real `__pycache__`, added file, deleted file, bootstrap-stripped `run.py`. → all `rc=2`, sanitized reason, no output directory; the doubled `figure_b1` values are never produced. **R02-01 is closed for the action tree.**
- **probe/p_effects_guard.py** — the unvalidated effects table on `render_checked`. **FOUND DEFECT** (OPUS-R03-01(a)); three controls isolate it.
- **probe/p_worldcache{,2,3}.py** — the World-tree bytecode residual, on a copied World. **FOUND DEFECT** (OPUS-R03-01(b), defence-in-depth).
- **probe/p_relabel_atom1.py** — Atom-1 ledger relabel. **FOUND DEFECT** (OPUS-R03-02).
- **probe/p_adversarial.py** — the round-2 whole-bundle relabel of atom-000003 now rejected at `load_evidence`; single-cell `gap_ids/reason/source_locator/status` mutations rejected on all of atoms 2–11; output inside the World, non-empty output directory, unknown manifest key, out-of-scope table, synthetic receipt as governed evidence, `accepted=false`, and `no_result` retry under **both** policies all fail closed with a non-leaking reason.
- **probe/p_identity.sh** — 5 copied Worlds. Protocol activation, figure `CONFIG`, figures `yaml` and manifest `implementation_id` tampering all rejected; appended realized Atom body **and** appended Molecule interpretation still run to `registered_state=partial`.
- **probe/p_dtype.py** — `str/object/category` raise `Boundary(..., ["A11"])`; nullable `Int64/Float64/boolean` fit identically (`n=262`), so the R02-04 fix did not over-reject; a genuine `pd.NA` in a nullable structural column is still caught.
- **probe/p_misc.py** — mechanical `v002→v003` gap-rule diff over all 50 targets: exactly two changes (`prose_08`, `prose_14` lose A02); `A02_TARGETS` contains no unregistered name; `figure_b18` is dead.
- **probe/p_models.py** — independent statsmodels 0.14.4 oracles on my own designs: `HC1 1.1e-15 / cov 1.2e-17` with matching rank 8 and df 192; supplied-metadata `CR1 cov 1.1e-17`; default clustered path is `CRO_not_source_inference` with `df=None`, refuses intervals with gaps `["A01", "A02"]` and returns `F=None` `p=None`; a collinear treatment column stops with `Source treatment coefficients not identified`; `matched eq2` agrees with a dense LSDV fit (time, year, age

and person×match dummies) on the identical 262-row post-singleton sample to $2.6e-14$.

- **probe/p_render.py** — all 29 registered figures rendered from **my own** CLI exports through `validation.render_checked`: 29/29, zero failures, every page 130.700 mm within 0.001 mm.
- **Copied-layout CLI** — `cp -a` of `prereg/execution` run twice with `PYTHONDONTWRITEBYTECODE` explicitly unset: both exit 0, byte-identical `results.csv`, zero `.pyc` left behind. `run.py --help` and `run_tests.py --help` both exit 0 with no World; the new `argument_parser` default roster matches `contracts.ROSTER`.
- **AST comparison** of `activation.py` against `tools/prereg-lane/src/prereg_lane/native.py` → all 5 shared top-level definitions identical, nothing extra vendored.
- **Integrity** — 122 original files byte-identical to `.controller/source`; `frozen-identity.json` byte-identical v002→v003 (registration id 4116339700f6... unchanged); no stray file in the World.

Not executed / unavailable

- **Stata, reghdfe, teffects** — command `-v stata / stata-se / stata-mp` finds nothing; I installed nothing. `source_models.do` has never been executed by the author or by me, so source-engine version, residual df, ties and exact numerical correspondence remain **unavailable**.
- **Parquet end-to-end** — exercised at the `contracts.Inputs/construction` level in round 2 only.
- **Governed pending/nonactivation bundle** — `validate_results'` pending branch (`validation.py:66-69`) is unreachable from the CLI here; exercised only by the author's `execute(..., active=False)` test.
- **Untouched constructors** — `source_panel`, `percentile`, `hospital`, `jobs`, `occupation_codes`, `onset_panel`, `candidate_controls`, `wealth`, `enrollment`, `education`, `prescriptions`, `services`, `B1/B2` and `spouse_histories` are byte-identical to v002 and were derived line-by-line against the author programs in my rounds 1 and 2; I did not derive them a third time.
- **Real data, linkage authorization, national-scale performance, governed font binding, TeX rebuild, exhaustive visual inspection** — not tested.

4. Which numerical results remain unavailable, and why

45 of the 50 registered targets produce no number, entirely because of the frozen deviations ledger and unavailable inputs — not because of anything in this code. I checked the exported gap set of every target against `deviations.md` and `target-inventory.md`: C01/A01 block all of Atoms 3–9 and 11; A02 adds estimator normalization on the matched Eq2–4 set plus `prose_06` and `figure_b18`; A03 the Eq3/4 post window; C04 the Eq1 counterfactual denominator; C05 caption-versus-program models; C06/A06 childcare eligibility; C07 the illness definition; C08/U04

wealth units; C09 geography; C10/A08 future-treated rules; A05 psychiatry; A07 age strata; A09 spouse matching; A10 occupation ordering; C02 the under-six boundary; C03/A04 histogram origins; U03 unspecified prose robustness.

The five targets that do produce values (`figure_b1`, `figure_b2`, `table1`, `table_b1`, `table_b2`) are **Atom 1 descriptive cells computed from clearly labelled synthetic fixtures**. They are not Danish results.

Even once those gates are resolved, **no source interval or p-value can be exported for any matched target**: the default clustered covariance is CR0, labelled `CR0_not_source_inference`, and `Fit.intervals/Fit.contrast` refuse intervals, F and p until a source-established effective parameter count and residual df are supplied. The `reghdfe` small-sample convention remains an unresolved prerequisite for every published number, and Stata is unavailable to settle it.

Input provenance and receipt authenticity remain unverifiable operator assertions inherent to a pre-data World. The disclosed frozen-template presentation limits (Figure B20 label collisions, B3/B4 default numeric axes, over-height F3/B7/B10/B12) are out of scope by user instruction and are not counted here as engineering blockers.

Public documentation and source-check log

Author round 3 — 2026-09-06

Read the complete eleven Atom registrations/protocols/result regions, complete Molecule registration/membership, D1–D8, identification/governance, source audit/deviations/matching/schema/inventory, all execution modules/tests and paper result-slot/figure interfaces. Applied the full lane AGENTS and predata-engineering, preregistration-package (all required references), paper-to-replication and canonical-preregistration skills. Read both round-1 reviews and the supplied round-2 Opus review. No round-2 Codex report exists in the supplied reviews directory; no review or reviewer acceptance is inferred from its absence.

No live search was used in this revision and no new scientific resolution metadata was accepted. All 36 local public author-code hashes match `replication/code-inspection-index.csv`. Source field meanings were checked with narrowly selected local FSV field/history rows for the same 18 registered tables. In particular: `BEF.md:40,44,48,54`, `FAIN.md:27,42,50`, `IDAP.md:38,166`, `IDAN.md:63,72,131`, `IDAS.md:84`, `AKM.md:33,41–45,133`, `IND.md:86–95`, `UDDA.md:34,37,109`, `DODSAARS.md:45,53`, `DODSAASG.md:53,64`, `LPR_ADM.md:28,61,63`, `PSYK_ADM.md:28,48,50`, `LMDB.md:62`, `DAGI.md:35`, `SYSI.md:40,44`, and `SSSY.md:42,47`; full repository directory locators remain in frozen `schema-evidence.md`. No modern extract, code classification or access availability is inferred from those metadata rows.

Round-3 source adjudication: `deviations.md A02` limits normalization to matched Eq2–4/linear-controls robustness; Atom5/prose_08 and Atom7/prose_14 identify ratios and prevalence, respectively. Their A03/C07 and sample gates stay intact. Rechecked local public `analysis_14_cooccurrence.do:245–424` (onsets/event clock),

`analysis_12_parental_health.do:1-100` (illness input and classification conflict), `analysis_5_main_analysis.do:1-100` (source model/interval conventions), complete `analysis_15_balancing_test.do` (15 named targeted rows at line114), and complete `clean_1_hospital_visits.do/clean_2_hospital_visits_psychiatry.do` (visit counting before interval filtering). These programs were read, never executed. Their hardcoded original data paths were never used.

The execution fixes concern authenticated source loading, exact boundary meta-data, dtype rejection and a synthetic sample oracle. They do not supply a source-silent estimator, normalization, clinical definition, policy date or table. The CPython cache counterexamples use locally compiled synthetic bytecode and controller Python only. No confidential data, old empirical session, remote QGPU, source dictionary values or measured outcomes were read. There is no `prereg/code/` directory or supplied additional runtime figure/execution interface in this World.

Author round 2 — 2026-09-06

Read both round-1 engineering reviews and all current execution modules, all eleven Atom protocols/result partitions, Molecule registration, shared D1–D8/identification/governance, replication audit/deviations/matching/schema/inventory and paper result/figure interfaces. The selected semantics are canonical. No `prereg/code` directory or additional runtime figure/execution interface was supplied. Applied the full lane AGENTS, predata-engineering, preregistration-package and required references, paper-to-replication and canonical-preregistration instructions. The original material is frozen and unchanged.

No live search was used in this revision; reviewer defects were adjudicated against local public source and existing schema documents. No new public source was treated as resolution metadata. Local exact rechecks:

Source	Locator	Revision use
Public author somatic hospital code	<code>resources/literature/papers/danishregister-recon/packs/data_check</code>	contact collapse; reversed intervals excluded only from day expansion. Full file read, not executed.
Public psychiatric hospital code	Same code root, <code>clean/clean_2_hospital_visits_psychiatry.do:12-26</code>	Same-admission selection counts; full file read, not executed.
Public labor construction	Same root, <code>clean/clean_11_employment_and_randomly_dla34-155,216-324</code>	Exact occupation job-code string semantics; full file read, not executed.

Source	Locator	Revision use
Public baseline table	Same root, analysis/analysis_15_balemc/ATP/estimates/00122, especially :114	Exact employ- row order, 15 named targeted rows.
Public spouse construction/orchestration	Same root, clean/clean_14_merge_spousing.do, do_all.do, set_globals.do	Existing source module and absence of source software pins; read, never executed, no original mounts accessed.
Local FSV	resources/schema/data- living-conditions/{SYS labour-income-and-wealth Admission/polecho/fsv/health/{LPR_ADM,PSYK_ADM}.md InCSSYsmdspecialty/code th/{ADM, DAN, DBAS}.md	vintage and TYPE history. Exact field rows queried with rg. These files are schema-only documentation, not runtime access evidence.

Dictionary values and historical extract correspondence remain U01/U02. Runtime code now rejects a numeric code representation when source string identity cannot be recovered; no leading-zero width, code mapping or historical classification is inferred from current metadata.

Prior round-1 retrieval record

Date: 2026-09-06. No confidential runtime was queried. Existing source-provenance.yaml, source-map.md, retrieval-log.md, code-inspection-index.csv, deviations.md and schema-evidence.md remain immutable.

Access	Exact query or local source	Exact URL / locator	Use and limit
Web search	site.stata.com summarize percentiles default definition P=N p integer average	Stata16 summarize manual , Methods and formulas	Default order statistic and adjacent averaging when Np is integer. This does not identify the original authors' Stata version or histogram algorithm.

Access	Exact query or local source	Exact URL / locator	Use and limit
Web search	<code>site.scorreia.com</code> <code>reghdfe nested fixed effects clusters degrees freedom</code>	reghdfe author FAQ , nested FE/cluster degrees of freedom	Confirms that FE nesting matters to finite-sample correction; no claim that full dense-design parameter count equals the unavailable original ado convention.
Local public code	<code>resources/literature/papers/danishPreregisterPublic</code>	<code>through16; analysis/analysis_1 through16; FN_analysis/*; set_globals.do; do_all.do</code>	<code>resources/public/con/packs/data_check_packs</code> source code only; no original program executed, no hardcoded original data mount used.
Local FSV documentation	<code>resources/schema/Data-registry</code>	<code>data-registry</code> page in frozen prereg/replication/schemas-evidence.md, including exact field rows and historical changes	See <code>fsv</code> field meanings/cumulative schemas-annual extract distinctions; catalog availability does not establish source-compatible data access.
Frozen package	all11 atoms, molecule-000001, prereg shared contracts, replication and paper interfaces	Relative World paths and frozen-identity.json material map	Scientific authority, exact source target ownership and ambiguity dispositions; no altered source choice.

No additional public evidence resolved C01–C10, A01–A11 or U01–U04. Public mathematical documentation was used for supported implementation, not as a scientific amendment. The original published outcome values were never inserted into runtime output or tests as expected replication coefficients. Test oracles are algebraic or invented synthetic constructions.

No live Danish/APTO schema or data access was attempted. Official URLs already recorded in frozen schema-evidence are provenance links; local FSV metadata was read rather than treating current official field listings as proof of historical extract parity.

Complete registered Molecule

```
id: molecule-000001
molecule_mode: prereg
molecule_shape: mechanism
topic: Parental death outcomes
rating: C
draft: false
literature: []
membership:
- atom: atom-000001
  role: supplementary
  description: This is the source population/comparability question, not a
    standalone technical linkage experiment. Distributions and related
    ↪ baseline
    tables belong together.
- atom: atom-000002
  role: supplementary
  description: Sudden and non-sudden paths answer the same uncontrolled
    ↪ event-time
    question; their external-validity comparison stays in one module.
- atom: atom-000003
  role: primary
  description: Main outcomes, parent/child gender contrasts, labor robustness
    and balance evidence assess one labor-market question. Balance is embedded
    in this consuming protocol. Figures are not separate experiments.
- atom: atom-000004
  role: secondary
  description: Health outcomes form the distinct health question, separate
    from earnings and from co-occurrence. Related services and prescription
    panels remain in one analysis.
- atom: atom-000005
  role: secondary
  description: Joint onset evidence asks a distinct timing/mechanism question;
    its related health panels and conditional-share claims stay together.
- atom: atom-000006
  role: secondary
  description: Earnings contrasts and formal-care uptake jointly address loss
    of informal childcare; separate measures/panels are not separate
    ↪ experimental
    nodes.
- atom: atom-000007
  role: secondary
  description: Post/pre illness contrasts, residence response and source
    ↪ mental-health
    comparison address the surviving-parent support question and stay within
    one module.
- atom: atom-000008
  role: supplementary
  description: All parental/own-wealth variants address the inheritance
    ↪ question;
    they are retained together even when unavailable.
- atom: atom-000009
  role: supplementary
  description: Cohabitation and fertility are the source family-formation
```

analysis, independently interpretable from labor/health effects.

- atom: atom-000010
 - role: supplementary
 - description: The spouse outcome and its separately matched population are an independent spillover question; sharing family links does not merge it with adult-child earnings.
- atom: atom-000011
 - role: supplementary
 - description: The three job-choice outcomes form the original
 - preference-related module; routine occupational coding is part of this analysis, not another Atom.

frontier:

- rule: Extract this journal study and preserve every target and source
 - boundary;
 - interpretation uses actual component evidence.
- avoid: No added science, outcome-selected activation, success threshold, independent-review claim or conflation of completion with confirmation.
- directions: []

preregistration:

- mode: canonical
- question: What does a source-faithful replication of Jensen and Zhang (2026) establish about parental death, adult labor outcomes, health and family support, within the source's actual identified scope?
- scope: Fixed journal question, original cohort/treatment and reported
 - main/supplement analyses. Canonical two-stage replica conversion then Transcriber.
 - Unresolved source choices are not implemented.

storyline: Population and descriptive evidence delimit context. Matched labor effects address the primary question; health, joint timing,

- childcare and surviving-parent support qualify mechanisms. Inheritance, formation, spouses and work choice supply distinct supporting/alternative-channel evidence. All components are interpreted through their actual estimates and uncertainty, including mixed findings and missing targets.

interpretation:

- kind: restriction
 - condition: 'TRUE'
 - says: Apply source-definition/vintage limits to every component. Even fully available results require outcome-specific signed estimates,
 - source uncertainty, and the stated identification assumptions. No coverage state or paper comparison implies replication success.
- kind: primary
 - condition: done(atom-000003)
 - says: 'Read labor results.csv and interpretation.md: report employment, conditional/unconditional ATP, raw/indexed earnings and
 - parent/child-gender differences separately, together with source robustness, balance scope and missing targets. Base direction/magnitude claims on those actual estimates/intervals; mixed or imprecise findings stay mixed/imprecise.'
- kind: supplementary
 - condition: done(atom-000004) OR done(atom-000005)
 - says: Use actual health-service, prescription and co-occurrence vectors to assess consistency with the source health explanation. Distinguish service use, onset timing, ratios and unavailable estimates. No causal mediation share follows from joint onset or completion.
- kind: supplementary
 - condition: done(atom-000006) OR done(atom-000007)

```

says: Interpret childcare earnings contrasts, enrollment, surviving-parent
      illness/proximity and their source model boundaries jointly with actual
      labor evidence. Separate subgroup selection and post-event illness
      ↪ conditioning
      from causal claims about care.
- kind: supplementary
condition: done(atom-000008) OR done(atom-000009) OR done(atom-000010)
      OR done(atom-000011)
says: Report each actual inheritance, family-formation, spousal and
      ↪ job-choice
      result with its uncertainty and missing scope. Small/imprecise findings
      do not establish absence of a channel; no equivalence threshold or
      ↪ combined
      score is authorized.
- kind: supplementary
condition: done(atom-000001) OR done(atom-000002)
says: Use the actual population characteristics and descriptive
      ↪ sudden/non-sudden
      paths to delimit the source population and contextual dynamics. Do not
      turn conditional age distributions into population hazards or
      ↪ descriptive
      paths into matched causal evidence.
- kind: residual
condition: 'TRUE'
says: Name every pending, partial, boundary, inconclusive or operationally
      failed component. No empirical replication has occurred at registration.
      If a component has no estimate, make no directional claim for its
      ↪ outcomes.
      Preserve unresolved source choices; compilation/conversion completion
      is not independent scientific acceptance.

```

Source-aligned storyline

Population and descriptive evidence delimit context. Matched labor effects address the primary question; health, joint timing, childcare and surviving-parent support qualify mechanisms. Inheritance, formation, spouses and work choice supply distinct supporting/alternative-channel evidence. All components are interpreted through their actual estimates and uncertainty, including mixed findings and missing targets.

Input reuse does not imply scientific-result dependence. All activations are TRUE, so no evidence sign, p-value or local state selects execution. Restriction, primary, supplementary and residual rules compose; none returns an automatic confirmation label. The membership above is the complete authority for all 11 Atoms. Atom registrations own protocols and local states.

All Atoms are pending preregistrations. No independent review or empirical replication has occurred.

atom-000001: Bereavement population

Bereavement population

Use shared definitions D1–D5 below. analysis_1 retains first deaths 1985–2014, forms five-year age-at-first-death bins with 60+ upper category, sums counts cumulatively over ordered bins, divides by all retained deaths, and suppresses source cells smaller

than5. B1 is conditional on this death-period sample; no age-specific hazard or population-risk denominator is created. Exact-age introductory shares are distinct prose targets; do not approximate19+ from five-year bars. For Table1, TablesB1/B2 and B2–B4, construct the treated-only balanced11-year panel of adults25–50 at first death1985–2014, two known parents present t=-1; retain both sudden and non-sudden groups. Use analysis_2 final cause set, not clean_8's intermediate broader set. Baseline is t=-1. Child count is capped5 in the source summaries; youngest age follows source definitions. Earnings are p1/p99 winsorized pooled across both death groups before baseline tabulation. First-death-parent-age<=100 is imposed where analysis_2 does so (B2/B3/tables); B4 precedes this restriction. Track this sample distinction, never harmonize it silently. Table1 rows in order: age; male; college+; high school/vocational; below high school; cohabitation; number of children; age youngest; under6; mother age; father age; mother married; father married; first-death parent age; employment; intensive margin conditional working; annual earnings; observations. Give sudden mean, non-sudden mean, source two-sample t-test p-value (pooled-variance Stata ttest as written) and counts. TableB1 repeats means by mother-sudden,mother-nonsudden,father-sudden,father-nonsudden; TableB2 by mother-daughter,father-daughter,mother-son,father-son for sudden deaths. Do not append t-tests absent from those tables. B2 displays father/mother percentages by suddenness. B3 displays parent age distributions; B4 adult age distributions. Histogram densities/bin edges must be source-determined; source suppression<5 is preserved and is not a new statistical threshold. PSEUDOCODE: authenticate definitions -> form source-specific sample -> resolve each target's gap list -> derive source fields -> tabulate named rows/categories in original order -> write estimate/denominator/units for every cell -> export figure interfaces. On C02 block only under-six rows; on C03 block B3; on A04 retain exact missing bin/denominator decisions. Do not execute blocked scientific alternatives.

Unresolved target gates: U01, U02, C02, C03, A04, A11.

Retained source targets: figure_b1, figure_b2, figure_b3, figure_b4, table1, table_b1, table_b2, prose_01, prose_02.

Incorporated specification and actual evidence

Read prereg/data-contract.md D1–D8, replication/matching-specification.md where applicable, schema-evidence.md, target-inventory.md and deviations.md as part of this protocol. Source target order, field definitions, sample restrictions, formulae, intervals, source-only tests and missing-data rules are fixed there. Conflicting decisions remain no-estimate; source-supported partial results do not fill another target's gap. No confidential input or result is present now.

```
INPUT = authenticated source-vintage tables named in registration.tables
TARGETS = this Atom's full source inventory, including unavailable prose
↪ claims
for target in TARGETS in source order:
  required = target's exact panel/outcome/sample/model/contrast
  ↪ specification
  blockers = unresolved source choices + missing required
  ↪ vintage/key/year/crosswalk
```

```

if blockers: emit target-status(no-estimate, exact reasons); emit
  ↳ numeric-null rows; continue
panel = derive required source panel, with prescribed source restrictions
  ↳ in prescribed order
verify unique source units, explicit observed year set and
  ↳ source-authorized join keys
if source estimator cannot identify required target: emit scientific
  ↳ boundary with reason; continue
result = evaluate only required source formula using prescribed inference
  ↳ semantics
if execution fails: log operational failure and leave pending; do not
  ↳ record a scientific null
else: emit all keyed estimates/SE/CI/baselines/N/source-reported
  ↳ contrasts, not a favorable subset
write results.csv; target-status.csv; interpretation.md; named figure output
  ↳ interfaces
interpret each outcome and contrast in its own units and uncertainty;
  ↳ preserve mixed results

```

Each figure output `<figure_id>-effects.csv` contains `target_id,panel,series,x,estimate,lower,upper,unit,status,reason,bin_left,bin_right`. For histogram source bins, edges are supplied only after source identification. `<figure_id>-annotations.csv` contains `target_id,panel,series,statistic,term,value,unit,status,reason` for source-reported DID,SE,difference,F,p,baseline,N. Tables/prose use `results.csv` key `target_id,panel,series,x,statistic,term` and D8 additional metadata. Numeric fields null when unavailable. Do not create actual CSV files during conversion. See `figures.yaml` and `results-map.md` for exact input names and manuscript bindings.

Scientific interpretation is outcome/estimand-specific: signed magnitude and source95% intervals (or source SE-only table), relevant source-reported subgroup contrasts, and source-compatible limitations. Statistical imprecision is retained as evidence, not an unavailable observation. There is no replication-success cutoff, invented power, equivalence band or significance-based execution. The native numeric ratings are required compatibility metadata (neutral3 and 0.5), not an elicited novelty ranking, empirical feasibility claim or success score. This replica lane has no tournament or reviewer.

Source precision at the display boundary: B1 is built directly from BEF year-end1985–2014 with both parent IDs known, retaining only the person’s first-death year. It does not inherit the25–50 restriction or11-year outcome balance. Compute cumulative numerator and total denominator BEFORE suppressing count<5 groups; export100*share for the percentage-labelled interface. The code’s40–45 printed label corresponds to $\text{floor}(\text{age}/5)=8$ (40–44); preserve this source label/calculation discrepancy under A04. The affected category is no-estimate until its intended boundary is source-resolved; do not silently relabel or rebin. B3 uses source histogram width2 after dropping age-at-death×suddenness cells<5; B4 uses bin(26) after its source age×suddenness suppression. Origins/edges and B3’s age-variable conflict remain unresolved. Table1/B1/B2 exclude nonpositive ATP for the intensive-margin means; TableB3’s intensive margin remains unconditional.

Complete registration fields

```
id: atom-000001
topic: Bereavement population
atom_result: prereg
atom_strength: null
atom_shape: null
evidence_paths: []
concepts: []
literature: []
bonds: []
preregistration:
  version: 1
  mode: canonical
  question: What population and pre-death characteristics underlie the source
    comparison of sudden and non-sudden first parental death?
  rationale: This is the source population/comparability question, not a
    ↪ standalone
    technical linkage experiment. Distributions and related baseline tables
    belong together.
  assumption: Source correspondence and input coverage must be established
    as specified. Interpretation stays within this source
    ↪ descriptive/associational
    scope; shared treatment does not establish causal mediation.
  target_strength: descriptive
  target_shape: comparison
  ratings:
    importance: 3
    novelty: 3
    feasibility: 3
    data_advantage: 3
    informativeness: 0.5
  tables:
    - BEF
    - FAIN
    - DODSAARS
    - DODSAASG
    - IND
    - IDAP
    - UDDA
  activation: 'TRUE'
  verdicts:
    - key: evidence
      region: All retained source targets have actual estimates and their
        ↪ source-required
        uncertainty/interpretation, with no unresolved target.
      says: Report the complete actual outcome vector. Coverage alone gives
        no direction, causal conclusion or replication confirmation.
    - key: partial
      region: At least one source target has actual evidence and at least one
        has a conclusive source/input no-estimate disposition.
      says: Report available signed estimates and uncertainty, including mixed
        findings, beside every missing target and its exact limitation. Do not
        generalize the available subset to the missing scope.
    - key: boundary
      region: No retained source target is estimable after a conclusive
        ↪ source-definition
        or input boundary; this is not an operational failure.
      says: There is no scientific estimate for this source scope. The paper
        is neither confirmed nor refuted by inability to estimate.
    - key: otherwise
```

```

region: otherwise
says: Retain unresolved or inconclusive scientific disposition with actual
      evidence if any. Operational failures remain pending and are not done()
      or a statistical null.
protocol: |
# Bereavement population

Use shared definitions D1-D5 below. analysis_1 retains first
↳ deaths1985-2014, forms five-year age-at-first-death bins with60+
↳ upper category, sums counts cumulatively over ordered bins, divides
↳ by all retained deaths, and suppresses source cells smaller than5. B1
↳ is conditional on this death-period sample; no age-specific hazard or
↳ population-risk denominator is created. Exact-age introductory shares
↳ are distinct prose targets; do not approximate19+ from five-year bars.
For Table1, TablesB1/B2 and B2-B4, construct the treated-only
↳ balanced11-year panel of adults25-50 at first death1985-2014, two
↳ known parents present t=-1; retain both sudden and non-sudden groups.
↳ Use analysis_2 final cause set, not clean_8's intermediate broader
↳ set. Baseline is t=-1. Child count is capped5 in the source
↳ summaries; youngest age follows source definitions. Earnings are
↳ p1/p99 winsorized pooled across both death groups before baseline
↳ tabulation. First-death-parent-age<=100 is imposed where analysis_2
↳ does so (B2/B3/tables); B4 precedes this restriction. Track this
↳ sample distinction, never harmonize it silently.
Table1 rows in order: age; male; college+; high school/vocational; below
↳ high school; cohabitation; number of children; age youngest; under6;
↳ mother age; father age; mother married; father married; first-death
↳ parent age; employment; intensive margin conditional working; annual
↳ earnings; observations. Give sudden mean, non-sudden mean, source
↳ two-sample t-test p-value (pooled-variance Stata ttest as written)
↳ and counts. TableB1 repeats means by
↳ mother-sudden,mother-nonsudden,father-sudden,father-nonsudden;
↳ TableB2 by mother-daughter,father-daughter,mother-son,father-son for
↳ sudden deaths. Do not append t-tests absent from those tables. B2
↳ displays father/mother percentages by suddenness. B3 displays parent
↳ age distributions; B4 adult age distributions. Histogram
↳ densities/bin edges must be source-determined; source suppression<5
↳ is preserved and is not a new statistical threshold.
PSEUDOCODE: authenticate definitions -> form source-specific sample ->
↳ resolve each target's gap list -> derive source fields -> tabulate
↳ named rows/categories in original order -> write
↳ estimate/denominator/units for every cell -> export figure
↳ interfaces. On C02 block only under-six rows; on C03 block B3; on A04
↳ retain exact missing bin/denominator decisions. Do not execute
↳ blocked scientific alternatives.

Unresolved target gates: U01, U02, C02, C03, A04, A11.

Retained source targets: figure_b1, figure_b2, figure_b3, figure_b4,
↳ table1, table_b1, table_b2, prose_01, prose_02.

## Incorporated specification and actual evidence

Read prereg/data-contract.md D1-D8, replication/matching-specification.md
↳ where applicable, schema-evidence.md, target-inventory.md and
↳ deviations.md as part of this protocol. Source target order, field
↳ definitions, sample restrictions, formulae, intervals, source-only
↳ tests and missing-data rules are fixed there. Conflicting decisions
↳ remain no-estimate; source-supported partial results do not fill
↳ another target's gap. No confidential input or result is present now.

```

```

...
INPUT = authenticated source-vintage tables named in registration.tables
TARGETS = this Atom's full source inventory, including unavailable prose
↳ claims
for target in TARGETS in source order:
    required = target's exact panel/outcome/sample/model/contrast
    ↳ specification
    blockers = unresolved source choices + missing required
    ↳ vintage/key/year/crosswalk
    if blockers: emit target-status(no-estimate, exact reasons); emit
    ↳ numeric-null rows; continue
    panel = derive required source panel, with prescribed source
    ↳ restrictions in prescribed order
    verify unique source units, explicit observed year set and
    ↳ source-authorized join keys
    if source estimator cannot identify required target: emit scientific
    ↳ boundary with reason; continue
    result = evaluate only required source formula using prescribed
    ↳ inference semantics
    if execution fails: log operational failure and leave pending; do not
    ↳ record a scientific null
    else: emit all keyed estimates/SE/CI/baselines/N/source-reported
    ↳ contrasts, not a favorable subset
write results.csv; target-status.csv; interpretation.md; named figure
↳ output interfaces
interpret each outcome and contrast in its own units and uncertainty;
↳ preserve mixed results
...

```

Each figure output `<figure\_id>-effects.csv` contains target_id,panel,series,x,estimate,lower,upper,unit,status,reason,bin_left,bin_right. For histogram source bins, edges are supplied only after source identification. `<figure\_id>-annotations.csv` contains target_id,panel,series,statistic,term,value,unit,status,reason for source-reported DID,SE,difference,F,p,baseline,N. Tables/prose use results.csv key target_id,panel,series,x,statistic,term and D8 additional metadata. Numeric fields null when unavailable. Do not create actual CSV files during conversion. See figures.yaml and results-map.md for exact input names and manuscript bindings.

Scientific interpretation is outcome/estimand-specific: signed magnitude and source95% intervals (or source SE-only table), relevant source-reported subgroup contrasts, and source-compatible limitations. Statistical imprecision is retained as evidence, not an unavailable observation. There is no replication-success cutoff, invented power, equivalence band or significance-based execution. The native numeric ratings are required compatibility metadata (neutral3 and0.5), not an elicited novelty ranking, empirical feasibility claim or success score. This replica lane has no tournament or reviewer.

Source precision at the display boundary: B1 is built directly from BEF

- ↪ year-end1985-2014 with both parent IDs known, retaining only the
- ↪ person's first-death year. It does not inherit the25-50 restriction
- ↪ ori1-year outcome balance. Compute cumulative numerator and total
- ↪ denominator BEFORE suppressing count<5 groups; export100*share for
- ↪ the percentage-labelled interface. The code's40-45 printed label
- ↪ corresponds to floor(age/5)=8 (40-44); preserve this source
- ↪ label/calculation discrepancy under A04. The affected category is
- ↪ no-estimate until its intended boundary is source-resolved; do not
- ↪ silently relabel or rebin. B3 uses source histogram width2 after
- ↪ dropping age-at-death*suddenness cells<5; B4 uses bin(26) after its
- ↪ source age*suddenness suppression. Origins/edges and B3's
- ↪ age-variable conflict remain unresolved. Table1/B1/B2 exclude
- ↪ nonpositive ATP for the intensive-margin means; TableB3's intensive
- ↪ margin remains unconditional.

Registered source module

What population and pre-death characteristics underlie the source comparison of sudden and non-sudden first parental death?

No current result, realized strength, evidence path or empirical replication is recorded. The preregistration protocol owns scientific specification and local states; manuscript/graph views are derived. Source locators, blockers and result interfaces are incorporated above. Scientific module purpose: This is the source population/comparability question, not a standalone technical linkage experiment. Distributions and related baseline tables belong together.

Reference implementation

Entry point and supported/unresolved components for atom-000001: [execution index](#). See [engineering review](#). This pointer does not amend the registered protocol.

atom-000002: Descriptive event paths

Descriptive event paths

Rebuild the treated-only balanced source panel separately for sudden (B5; analysis_3) and non-sudden death (B23; analysis_4). Outcomes are employment ATP>0, positive ATP intensive margin, real raw earnings, and any mental-health prescriptions (N05,N06A,N06B,N06C). No matched controls enter Eq1. Run each sex separately with event dummies k=-5..5 excluding-1, age-by-calendar-year and education-by-calendar-year categorical effects. Code uses heteroskedasticity-robust standard errors. Mental-prescription sample requires the complete11-year panel to begin1995 or later (therefore death2000-2014 given the labor upper end2019). Source earnings/ATP cleaning is D3; no new tail transformation is added. Source display conversion is100 times each event coefficient divided by its counterfactual fitted mean with the event coefficient removed; confidence endpoints are scaled likewise (code±1.96 robust SE before scaling). Prediction population is unresolved C04. Do not produce a percent path from a convenient denominator, nor present a different unscaled analysis as the

replication figure. $k=-1$ is an omitted reference, not an estimated zero with a fabricated interval. B5/B23 each preserve A employment, B conditional intensive margin, C raw earnings, D any mental prescriptions; two sexes per panel and all10 estimated event coefficients, source95% intervals and baseline denominator. PSEUDOCODE: build panel and outcome-specific availability -> generate categorical interactions and event design -> if denominator conflict unresolved, emit no-estimate for plotted targets with exact C04 reason; otherwise execute the source-established specification -> export full vector. Compare sudden/non-sudden trajectories only descriptively, preserving pre-event behavior; no new pretrend test, pooled effect or causal claim.

Unresolved target gates: U01, U02, C04, A11.

Retained source targets: figure_b5, figure_b23, prose_13.

Incorporated specification and actual evidence

Read prereg/data-contract.md D1–D8, replication/matching-specification.md where applicable, schema-evidence.md, target-inventory.md and deviations.md as part of this protocol. Source target order, field definitions, sample restrictions, formulae, intervals, source-only tests and missing-data rules are fixed there. Conflicting decisions remain no-estimate; source-supported partial results do not fill another target's gap. No confidential input or result is present now.

```
INPUT = authenticated source-vintage tables named in registration.tables
TARGETS = this Atom's full source inventory, including unavailable prose
↪ claims
for target in TARGETS in source order:
    required = target's exact panel/outcome/sample/model/contrast
    ↪ specification
    blockers = unresolved source choices + missing required
    ↪ vintage/key/year/crosswalk
    if blockers: emit target-status(no-estimate, exact reasons); emit
    ↪ numeric-null rows; continue
    panel = derive required source panel, with prescribed source restrictions
    ↪ in prescribed order
    verify unique source units, explicit observed year set and
    ↪ source-authorized join keys
    if source estimator cannot identify required target: emit scientific
    ↪ boundary with reason; continue
    result = evaluate only required source formula using prescribed inference
    ↪ semantics
    if execution fails: log operational failure and leave pending; do not
    ↪ record a scientific null
    else: emit all keyed estimates/SE/CI/baselines/N/source-reported
    ↪ contrasts, not a favorable subset
write results.csv; target-status.csv; interpretation.md; named figure output
↪ interfaces
interpret each outcome and contrast in its own units and uncertainty;
↪ preserve mixed results
```

Each figure output <figure_id>-effects.csv contains target_id,panel,series,x,estimate,lower,upper,unit,status,reason,bin_left,bin_right. For histogram source bins, edges are supplied only after source identification. <figure_id>-annotations.csv contains target_id,panel,series,statistic,term,value,unit,status,reason for source-reported

DID,SE,difference,F,p,baseline,N. Tables/prose use results.csv key target_id,panel,series,x,statistic,term and D8 additional metadata. Numeric fields null when unavailable. Do not create actual CSV files during conversion. See figures.yaml and results-map.md for exact input names and manuscript bindings.

Scientific interpretation is outcome/estimand-specific: signed magnitude and source95% intervals (or source SE-only table), relevant source-reported subgroup contrasts, and source-compatible limitations. Statistical imprecision is retained as evidence, not an unavailable observation. There is no replication-success cutoff, invented power, equivalence band or significance-based execution. The native numeric ratings are required compatibility metadata (neutral3 and0.5), not an elicited novelty ranking, empirical feasibility claim or success score. This replica lane has no tournament or reviewer.

Complete registration fields

```
id: atom-000002
topic: Descriptive event paths
atom_result: prereg
atom_strength: null
atom_shape: null
evidence_paths: []
concepts: []
literature: []
bonds: []
preregistration:
  version: 1
  mode: canonical
  question: How do labor supply, earnings and mental-prescription use evolve
    around sudden and non-sudden first parental deaths in the source
    ↪ descriptive
    design?
  rationale: Sudden and non-sudden paths answer the same uncontrolled
    ↪ event-time
    question; their external-validity comparison stays in one module.
  assumption: Source correspondence and input coverage must be established
    as specified. Interpretation stays within this source
    ↪ descriptive/associational
    scope; shared treatment does not establish causal mediation.
  target_strength: descriptive
  target_shape: temporal
  ratings:
    importance: 3
    novelty: 3
    feasibility: 3
    data_advantage: 3
    informativeness: 0.5
  tables:
    - BEF
    - FAIN
    - DODSAARS
    - DODSAASG
    - IND
    - IDAP
    - UDDA
    - LMDB
  activation: 'TRUE'
```

```

verdicts:
- key: evidence
  region: All retained source targets have actual estimates and their
    ↪ source-required
    uncertainty/interpretation, with no unresolved target.
  says: Report the complete actual outcome vector. Coverage alone gives
    no direction, causal conclusion or replication confirmation.
- key: partial
  region: At least one source target has actual evidence and at least one
    has a conclusive source/input no-estimate disposition.
  says: Report available signed estimates and uncertainty, including mixed
    findings, beside every missing target and its exact limitation. Do not
    generalize the available subset to the missing scope.
- key: boundary
  region: No retained source target is estimable after a conclusive
    ↪ source-definition
    or input boundary; this is not an operational failure.
  says: There is no scientific estimate for this source scope. The paper
    is neither confirmed nor refuted by inability to estimate.
- key: otherwise
  region: otherwise
  says: Retain unresolved or inconclusive scientific disposition with actual
    evidence if any. Operational failures remain pending and are not done()
    or a statistical null.
protocol: |
# Descriptive event paths

Rebuild the treated-only balanced source panel separately for sudden (B5;
↪ analysis_3) and non-sudden death (B23; analysis_4). Outcomes are
↪ employment ATP>0, positive ATP intensive margin, real raw earnings,
↪ and any mental-health prescriptions (N05,N06A,N06B,N06C). No matched
↪ controls enter Eq1. Run each sex separately with event dummies
↪ k=-5.5 excluding-1, age-by-calendar-year and
↪ education-by-calendar-year categorical effects. Code uses
↪ heteroskedasticity-robust standard errors. Mental-prescription sample
↪ requires the complete11-year panel to begin1995 or later (therefore
↪ death2000-2014 given the labor upper end2019). Source earnings/ATP
↪ cleaning is D3; no new tail transformation is added.
Source display conversion is100 times each event coefficient divided by
↪ its counterfactual fitted mean with the event coefficient removed;
↪ confidence endpoints are scaled likewise (code±1.96 robust SE before
↪ scaling). Prediction population is unresolved C04. Do not produce a
↪ percent path from a convenient denominator, nor present a different
↪ unscaled analysis as the replication figure. k=-1 is an omitted refer
↪ ence, not an estimated zero with a fabricated interval. B5/B23 each
↪ preserve A employment, B conditional intensive margin, C raw earnings
↪ , D any mental prescriptions; two sexes per panel and all10 estimated
↪ event coefficients, source95% intervals and baseline denominator.
PSEUDOCODE: build panel and outcome-specific availability -> generate
↪ categorical interactions and event design -> if denominator conflict
↪ unresolved, emit no-estimate for plotted targets with exact C04
↪ reason; otherwise execute the source-established specification ->
↪ export full vector. Compare sudden/non-sudden trajectories only
↪ descriptively, preserving pre-event behavior; no new pretrend test,
↪ pooled effect or causal claim.

Unresolved target gates: U01, U02, C04, A11.

Retained source targets: figure_b5, figure_b23, prose_13.

```

```

## Incorporated specification and actual evidence

Read prereg/data-contract.md D1-D8, replication/matching-specification.md
↳ where applicable, schema-evidence.md, target-inventory.md and
↳ deviations.md as part of this protocol. Source target order, field
↳ definitions, sample restrictions, formulae, intervals, source-only
↳ tests and missing-data rules are fixed there. Conflicting decisions
↳ remain no-estimate; source-supported partial results do not fill
↳ another target's gap. No confidential input or result is present now.

...

INPUT = authenticated source-vintage tables named in registration.tables
TARGETS = this Atom's full source inventory, including unavailable prose
↳ claims
for target in TARGETS in source order:
  required = target's exact panel/outcome/sample/model/contrast
  ↳ specification
  blockers = unresolved source choices + missing required
  ↳ vintage/key/year/crosswalk
  if blockers: emit target-status(no-estimate, exact reasons); emit
  ↳ numeric-null rows; continue
  panel = derive required source panel, with prescribed source
  ↳ restrictions in prescribed order
  verify unique source units, explicit observed year set and
  ↳ source-authorized join keys
  if source estimator cannot identify required target: emit scientific
  ↳ boundary with reason; continue
  result = evaluate only required source formula using prescribed
  ↳ inference semantics
  if execution fails: log operational failure and leave pending; do not
  ↳ record a scientific null
  else: emit all keyed estimates/SE/CI/baselines/N/source-reported
  ↳ contrasts, not a favorable subset
write results.csv; target-status.csv; interpretation.md; named figure
↳ output interfaces
interpret each outcome and contrast in its own units and uncertainty;
↳ preserve mixed results
...

Each figure output `<figure_id>-effects.csv` contains target_id,panel,ser
↳ ies,x,estimate,lower,upper,unit,status,reason,bin_left,bin_right. For
↳ histogram source bins, edges are supplied only after source
↳ identification. `<figure_id>-annotations.csv` contains
↳ target_id,panel,series,statistic,term,value,unit,status,reason for
↳ source-reported DID,SE,difference,F,p,baseline,N. Tables/prose use
↳ results.csv key target_id,panel,series,x,statistic,term and D8
↳ additional metadata. Numeric fields null when unavailable. Do not
↳ create actual CSV files during conversion. See figures.yaml and
↳ results-map.md for exact input names and manuscript bindings.

Scientific interpretation is outcome/estimand-specific: signed magnitude
↳ and source95% intervals (or source SE-only table), relevant
↳ source-reported subgroup contrasts, and source-compatible
↳ limitations. Statistical imprecision is retained as evidence, not an
↳ unavailable observation. There is no replication-success cutoff,
↳ invented power, equivalence band or significance-based execution. The
↳ native numeric ratings are required compatibility metadata (neutral3
↳ and0.5), not an elicited novelty ranking, empirical feasibility claim
↳ or success score. This replica lane has no tournament or reviewer.

```

Registered source module

How do labor supply, earnings and mental-prescription use evolve around sudden and non-sudden first parental deaths in the source descriptive design?

No current result, realized strength, evidence path or empirical replication is recorded. The preregistration protocol owns scientific specification and local states; manuscript/graph views are derived. Source locators, blockers and result interfaces are incorporated above. Scientific module purpose: Sudden and non-sudden paths answer the same uncontrolled event-time question; their external-validity comparison stays in one module.

Reference implementation

Entry point and supported/unresolved components for `atom-000002`: [execution index](#). See [engineering review](#). This pointer does not amend the registered protocol.

`atom-000003`: Matched labor effects

Matched labor effects

Use D1–D7 and exact matching transcription in `matching-specification.md`. Register source Eq2 dynamic effects for F1: A ATP>0; B ATP conditional positive; C earnings in1000 real2020DKK, pooled p1/p99 winsorized; D same earnings indexed to100 by sex-specific mean of all treated+control observations at $t=-1$. OLS absorbs $\text{person} \times \text{match}$, event time, calendar year, age; cluster $\text{person} \times \text{match}$. F1 includes source Eq3 summaries, standard errors, gender contrasts, F and p, sex baseline means. Eq3 aggregation window retains A03 rather than an invented choice. F2 uses Eq4 joint mother-death $\times$ post and father-death $\times$ post terms separately by sex; give four coefficients, source95% intervals, within-parent gender differences and within-sex mother-versus-father differences. Code tests differences with pooled interacted specifications; preserve the exact variance of the difference, not independent-SE arithmetic. B6 adds unconditional ATP including zeros beside conditional ATP; B7 shows F1's four outcomes dynamically for mother/father separately, with C05 equation-label conflict retained. B8 replaces earnings by annual cross-population rank among ages20–55 with known birth/sex: source egen rank, ties at average rank, $\text{percentile100} * (\text{rank} - 1) / (N - 1)$; no within-cohort ranking despite a code comment. Keep source ranks and denominators; if $N=1$, report undefined rather than inventing a percentile. B19 compares baseline child/deceased-parent proximity; caption5-region versus code11-region C09 blocks the comparison. B20 is four mutually exclusive baseline categories: same municipality; same11-region different municipality; same5-region different11-region; different5-region. Use $\text{event} \geq 1990$ as in `analysis_10` because detailed location input starts1985; subgroup-specific index/FE from the consuming block. Missing geography is not evidence of co-location. TableB6 consists of four separate source specifications by sex (8 columns), retaining all categories together: cause heart24/25, stroke26, respiratory29, accident46/47/49; parent-age split; adult-age split; death-period1985–94/1995–2004/2005–14. The age code uses sample medians while printed

labels70/38 are reference targets (A07). Effects are category×treatment×post with source fixed effects and cluster; retain SE and N, do not add cross-category tests. TableB3 is embedded evidence: at t=-1 report treatment/control means for targeted age,male,three education groups,cohabitation,child count,youngest age,under6,mother/father age,employment,unconditional intensive ATP,annual earnings,public sector; non-targeted mother/father any hospital,GP count,any hospital,any psychologist,any psychiatrist,any mental Rx,any alcohol Rx,any opioid; N. Follow source availability windows: GP/psych1990,hospital1994,Rx1995 with full-window eligibility. analysis_15 uses raw earnings despite calculating winsorized values and recalculates parent ages; do not use the erroneous late father-gap variable. Do not publish unreported balance tests or impose acceptance thresholds. Source FN robustness B21/B22 uses future sudden first deaths6 years later, five-year birth bins, same sex and parent gender, all eligible future-treated controls per stack, no1:1 NN. FN_clean loops index-death1985–2013, future controls up to2019, birth-bin index floor(birthyear/5)-385 in1..13, adult treated25–50 and future-control actual-event age>=30 within the source20–56 pool. It expands pseudo event-5..5, requires complete11 rows, recomputes age from birth. FN_analysis uses person×stack cluster, p5/p95 earnings winsorization and index100; A.F/B21–22 disagree (C10); do not implement these disputed choices. B21 panels A employment men,B women,C earnings men,D women overlay NN/FN; B22 earnings mother/father by sex. Author code-only parent-specific employment and extra transforms are excluded, not new robustness. Retain U03 footnote rematching, health matching, second-death and linear-controls claims with no-estimate. All explicit dynamic pre-periods are shown without adding a pretrend significance test. Source lower-bound/external-validity assertions remain contingent discussion, not new identification facts. PSEUDOCODE: resolve shared NN gates -> construct source-prescribed sample and output-specific transformations -> define source model columns -> fit only identified/authorized source specification -> export coefficients, covariance-based printed contrasts, intervals, baseline means, N and source notes -> label each target’s own availability. Any unresolved branch yields numeric-null rows; other authorized targets stay in inventory. Source targets are not selected by effect sign or p-value.

Unresolved target gates: U01, U02, C01, A01, A02, A03, C02, C05, A05, U03, C09, A07, C10, A08, A11.

Retained source targets: figure1, figure2, figure_b6, figure_b7, figure_b8, figure_b19, figure_b20, figure_b21, figure_b22, table_b3, table_b6, prose_03, prose_04, prose_05, prose_06.

Incorporated specification and actual evidence

Read prereg/data-contract.md D1–D8, replication/matching-specification.md where applicable, schema-evidence.md, target-inventory.md and deviations.md as part of this protocol. Source target order, field definitions, sample restrictions, formulae, intervals, source-only tests and missing-data rules are fixed there. Conflicting decisions remain no-estimate; source-supported partial results do not fill another target’s gap. No confidential input or result is present now.

```

INPUT = authenticated source-vintage tables named in registration.tables
TARGETS = this Atom's full source inventory, including unavailable prose
↪ claims
for target in TARGETS in source order:
    required = target's exact panel/outcome/sample/model/contrast
    ↪ specification
    blockers = unresolved source choices + missing required
    ↪ vintage/key/year/crosswalk
    if blockers: emit target-status(no-estimate, exact reasons); emit
    ↪ numeric-null rows; continue
    panel = derive required source panel, with prescribed source restrictions
    ↪ in prescribed order
    verify unique source units, explicit observed year set and
    ↪ source-authorized join keys
    if source estimator cannot identify required target: emit scientific
    ↪ boundary with reason; continue
    result = evaluate only required source formula using prescribed inference
    ↪ semantics
    if execution fails: log operational failure and leave pending; do not
    ↪ record a scientific null
    else: emit all keyed estimates/SE/CI/baselines/N/source-reported
    ↪ contrasts, not a favorable subset
write results.csv; target-status.csv; interpretation.md; named figure output
↪ interfaces
interpret each outcome and contrast in its own units and uncertainty;
↪ preserve mixed results

```

Each figure output `<figure_id>-effects.csv` contains `target_id,panel,series,x,estimate,lower,upper,unit,status,reason,bin_left,bin_right`. For histogram source bins, edges are supplied only after source identification. `<figure_id>-annotations.csv` contains `target_id,panel,series,statistic,term,value,unit,status,reason` for source-reported DID,SE,difference,F,p,baseline,N. Tables/prose use `results.csv` key `target_id,panel,series,x,statistic,term` and D8 additional metadata. Numeric fields null when unavailable. Do not create actual CSV files during conversion. See `figures.yaml` and `results-map.md` for exact input names and manuscript bindings.

Scientific interpretation is outcome/estimand-specific: signed magnitude and source95% intervals (or source SE-only table), relevant source-reported subgroup contrasts, and source-compatible limitations. Statistical imprecision is retained as evidence, not an unavailable observation. There is no replication-success cutoff, invented power, equivalence band or significance-based execution. The native numeric ratings are required compatibility metadata (neutral3 and 0.5), not an elicited novelty ranking, empirical feasibility claim or success score. This replica lane has no tournament or reviewer.

Complete registration fields

```

id: atom-000003
topic: Matched labor effects
atom_result: prereg
atom_strength: null
atom_shape: null
evidence_paths: []
concepts: []

```

```

literature: []
bonds: []
preregistration:
  version: 1
  mode: canonical
  question: What are the source matched-DiD effects of sudden first parental
    death on labor outcomes, including its reported heterogeneity and
    ↪ robustness?
  rationale: Main outcomes, parent/child gender contrasts, labor robustness
    and balance evidence assess one labor-market question. Balance is embedded
    in this consuming protocol. Figures are not separate experiments.
  assumption: Source correspondence and input coverage must be established
    as specified. Causal interpretation additionally requires the source
    ↪ parallel
    counterfactual trends/no relevant anticipation assumptions; the
    ↪ health/family
    patterns do not identify mediation.
  target_strength: causal
  target_shape: temporal
  ratings:
    importance: 3
    novelty: 3
    feasibility: 3
    data_advantage: 3
    informativeness: 0.5
  tables:
    - BEF
    - FAIN
    - DODSAARS
    - DODSAASG
    - IND
    - IDAP
    - IDAN
    - IDAS
    - UDDA
    - SYSI
    - SSSY
    - LPR_ADM
    - PSYK_ADM
    - LMDB
  activation: 'TRUE'
  verdicts:
    - key: evidence
      region: All retained source targets have actual estimates and their
        ↪ source-required
        uncertainty/interpretation, with no unresolved target.
      says: Report the complete actual outcome vector. Coverage alone gives
        no direction, causal conclusion or replication confirmation.
    - key: partial
      region: At least one source target has actual evidence and at least one
        has a conclusive source/input no-estimate disposition.
      says: Report available signed estimates and uncertainty, including mixed
        findings, beside every missing target and its exact limitation. Do not
        generalize the available subset to the missing scope.
    - key: boundary
      region: No retained source target is estimable after a conclusive
        ↪ source-definition
        or input boundary; this is not an operational failure.
      says: There is no scientific estimate for this source scope. The paper
        is neither confirmed nor refuted by inability to estimate.

```

- key: otherwise
region: otherwise
says: Retain unresolved or inconclusive scientific disposition with actual evidence if any. Operational failures remain pending and are not done() or a statistical null.
protocol: |
Matched labor effects

Use D1-D7 and exact matching transcription in matching-specification.md.
→ Register source Eq2 dynamic effects for F1: A ATP>0; B ATP conditional positive; C earnings in1000 real2020DKK, pooled p1/p99 winsorized; D same earnings indexed to100 by sex-specific mean of all treated+control observations at t=-1. OLS absorbs person*match, event time, calendar year, age; cluster person*match. F1 includes source Eq3 summaries, standard errors, gender contrasts, F and p, sex baseline means. Eq3 aggregation window retains A03 rather than an invented choice. F2 uses Eq4 joint mother-death*post and father-death*post terms separately by sex; give four coefficients, source95% intervals, within-parent gender differences and within-sex mother-versus-father differences. Code tests differences with pooled interacted specifications; preserve the exact variance of the difference, not independent-SE arithmetic.

B6 adds unconditional ATP including zeros beside conditional ATP; B7 shows F1's four outcomes dynamically for mother/father separately, with C05 equation-label conflict retained. B8 replaces earnings by annual cross-population rank among ages20-55 with known birth/sex: source egen rank, ties at average rank, percentile100*(rank-1)/(N-1); no within-cohort ranking despite a code comment. Keep source ranks and denominators; if N=1, report undefined rather than inventing a percentile. B19 compares baseline child/deceased-parent proximity; caption5-region versus code11-region C09 blocks the comparison. B20 is four mutually exclusive baseline categories: same municipality; same me11-region different municipality; same5-region different11-region; different5-region. Use event>=1990 as in analysis_10 because detailed location input starts1985; subgroup-specific index/FE from the consuming block. Missing geography is not evidence of co-location.

TableB6 consists of four separate source specifications by sex (8 columns), retaining all categories together: cause heart24/25, stroke26, respiratory29, accident46/47/49; parent-age split; adult-age split; death-period1985-94/1995-2004/2005-14. The age code uses sample medians while printed labels70/38 are reference targets (A07). Effects are category*treatment*post with source fixed effects and cluster; retain SE and N, do not add cross-category tests.

TableB3 is embedded evidence: at t=-1 report treatment/control means for targeted age,male,three education groups,cohabitation,child count,youngest age,under6,mother/father age,employment,unconditional intensive ATP,annual earnings,public sector; non-targeted mother/father any hospital,GP count,any hospital,any psychologist,any psychiatrist,any mental Rx,any alcohol Rx,any opioid; N. Follow source availability windows: GP/psych1990,hospital1994,Rx1995 with full-window eligibility. analysis_15 uses raw earnings despite calculating winsorized values and recalculates parent ages; do not use the erroneous late father-gap variable. Do not publish unreported balance tests or impose acceptance thresholds.

Source FN robustness B21/B22 uses future sudden first deaths6 years
 ↳ later, five-year birth bins, same sex and parent gender, all eligible
 ↳ future-treated controls per stack, no1:1 NN. FN_clean loops
 ↳ index-death1985-2013, future controls up to2019, birth-bin index
 ↳ floor(birthyear/5)-385 in1..13, adult treated25-50 and future-control
 ↳ actual-event age>=30 within the source20-56 pool. It expands pseudo
 ↳ event-5..5, requires complete11 rows, recomputes age from birth.
 ↳ FN_analysis uses person*stack cluster, p5/p95 earnings winsorization
 ↳ and index100; A.F/B21-22 disagree (C10); do not implement these
 ↳ disputed choices. B21 panels A employment men,B women,C earnings
 ↳ men,D women overlay NN/FN; B22 earnings mother/father by sex. Author
 ↳ code-only parent-specific employment and extra transforms are
 ↳ excluded, not new robustness.

Retain U03 footnote rematching, health matching, second-death and
 ↳ linear-controls claims with no-estimate. All explicit dynamic
 ↳ pre-periods are shown without adding a pretrend significance test.
 ↳ Source lower-bound/external-validity assertions remain contingent
 ↳ discussion, not new identification facts.

PSEUDOCODE: resolve shared NN gates -> construct source-prescribed sample
 ↳ and output-specific transformations -> define source model columns ->
 ↳ fit only identified/authorized source specification -> export
 ↳ coefficients, covariance-based printed contrasts, intervals, baseline
 ↳ means, N and source notes -> label each target's own availability.
 ↳ Any unresolved branch yields numeric-null rows; other authorized
 ↳ targets stay in inventory. Source targets are not selected by effect
 ↳ sign or p-value.

Unresolved target gates: U01, U02, C01, A01, A02, A03, C02, C05, A05,
 ↳ U03, C09, A07, C10, A08, A11.

Retained source targets: figure1, figure2, figure_b6, figure_b7,
 ↳ figure_b8, figure_b19, figure_b20, figure_b21, figure_b22, table_b3,
 ↳ table_b6, prose_03, prose_04, prose_05, prose_06.

Incorporated specification and actual evidence

Read prereg/data-contract.md D1-D8, replication/matching-specification.md
 ↳ where applicable, schema-evidence.md, target-inventory.md and
 ↳ deviations.md as part of this protocol. Source target order, field
 ↳ definitions, sample restrictions, formulae, intervals, source-only
 ↳ tests and missing-data rules are fixed there. Conflicting decisions
 ↳ remain no-estimate; source-supported partial results do not fill
 ↳ another target's gap. No confidential input or result is present now.

...

INPUT = authenticated source-vintage tables named in registration.tables
 TARGETS = this Atom's full source inventory, including unavailable prose
 ↳ claims

for target in TARGETS in source order:
 required = target's exact panel/outcome/sample/model/contrast
 ↳ specification
 blockers = unresolved source choices + missing required
 ↳ vintage/key/year/crosswalk
 if blockers: emit target-status(no-estimate, exact reasons); emit
 ↳ numeric-null rows; continue
 panel = derive required source panel, with prescribed source
 ↳ restrictions in prescribed order
 verify unique source units, explicit observed year set and
 ↳ source-authorized join keys

```

    if source estimator cannot identify required target: emit scientific
    ↪ boundary with reason; continue
    result = evaluate only required source formula using prescribed
    ↪ inference semantics
    if execution fails: log operational failure and leave pending; do not
    ↪ record a scientific null
    else: emit all keyed estimates/SE/CI/baselines/N/source-reported
    ↪ contrasts, not a favorable subset
write results.csv; target-status.csv; interpretation.md; named figure
↪ output interfaces
interpret each outcome and contrast in its own units and uncertainty;
↪ preserve mixed results
`..`

```

Each figure output ``<figure_id>-effects.csv`` contains `target_id,panel,series,x,estimate,lower,upper,unit,status,reason,bin_left,bin_right`. For histogram source bins, edges are supplied only after source identification. ``<figure_id>-annotations.csv`` contains `target_id,panel,series,statistic,term,value,unit,status,reason` for source-reported DID,SE,difference,F,p,baseline,N. Tables/prose use `results.csv` key `target_id,panel,series,x,statistic,term` and D8 additional metadata. Numeric fields null when unavailable. Do not create actual CSV files during conversion. See `figures.yaml` and `results-map.md` for exact input names and manuscript bindings.

Scientific interpretation is outcome/estimand-specific: signed magnitude ↪ and source95% intervals (or source SE-only table), relevant ↪ source-reported subgroup contrasts, and source-compatible ↪ limitations. Statistical imprecision is retained as evidence, not an ↪ unavailable observation. There is no replication-success cutoff, ↪ invented power, equivalence band or significance-based execution. The ↪ native numeric ratings are required compatibility metadata (neutral3 ↪ and0.5), not an elicited novelty ranking, empirical feasibility claim ↪ or success score. This replica lane has no tournament or reviewer.

Registered source module

What are the source matched-DiD effects of sudden first parental death on labor outcomes, including its reported heterogeneity and robustness?

No current result, realized strength, evidence path or empirical replication is recorded. The preregistration protocol owns scientific specification and local states; manuscript/graph views are derived. Source locators, blockers and result interfaces are incorporated above. Scientific module purpose: Main outcomes, parent/child gender contrasts, labor robustness and balance evidence assess one labor-market question. Balance is embedded in this consuming protocol. Figures are not separate experiments.

Reference implementation

Entry point and supported/unresolved components for `atom-000003`: [execution index](#). See [engineering review](#). This pointer does not amend the registered protocol.

atom-000004: Health service effects

Health service effects

Take the NN sample only after shared gates. D4 gives source encounter aggregation. F3 A annual psychologist consultations; B any mental-health prescription N05/N06A/N06B/N06C; C any opioid N02A; D annual non-psychiatric hospital visits including inpatient/outpatient. Estimate Eq4 mother/father $\times$ post separately by sex and source pooled contrasts. GP/psych services use years $\geq$ 1990, balanced11 rows (event $\geq$ 1995 through2014). Mental/opioid/alcohol prescriptions start1995, balanced11 (event2000–2014). Hospital outcomes restrict1994–2018 and rebalance11 (event1999–2013). Source psychiatry hospital data1995–2018 have their own complete-window sample. Different sample sizes are retained by outcome, never harmonized without source support. B9: A NUMBER of alcohol-addiction prescription records N07BB (not any-use); B annual GP consultations using source service-code list. B10: four dynamic outcomes in F3 order with mother/father subpanels and both sexes; source Eq2 dynamic interactions and95% intervals. B11: A screening hospital visits whose diagnosis containsDZ00 orDZ01; B nonscreening visits. Although clean_1 initially excludes allDZ in a provisional field, analysis_6 lines387–388 explicitly replaces nonscreening by total-minus-screening; this final consuming definition is source-explicit and is not an unresolved disagreement. Counts are annual encounters/services as code, never new diagnoses. TableB3's health means are supplied to the labor module with their exact sample windows, not a new health diagnostic. Private psychiatrist prose target remains A05/U02; no unreported combined psychiatric result replaces it. Preserve printed cross-sex and cross-parent contrasts and baseline means, including source significant and nonsignificant comparisons, without declaring a single health success label. PSEUDOCODE: authenticate complete source event register-year -> aggregate D4 -> join by authenticated PNR/year (absence0 only inside verified coverage) -> restrict each complete-window sample -> resolve target-specific gaps -> fit source Eq2/4 -> export outcome-specific coefficients/CI/contrast/N/baseline. If one service is unavailable, preserve the other eligible service targets; do not substitute a proxy. Health-utilization changes can be consistent with a health channel but cannot alone prove disease severity or causal mediation.

Unresolved target gates: U01, U02, C01, A01, A02, A03, A05, A11.

Retained source targets: figure3, figure_b9, figure_b10, figure_b11, prose_07.

Incorporated specification and actual evidence

Read prereg/data-contract.md D1–D8, replication/matching-specification.md where applicable, schema-evidence.md, target-inventory.md and deviations.md as part of this protocol. Source target order, field definitions, sample restrictions, formulae, intervals, source-only tests and missing-data rules are fixed there. Conflicting decisions remain no-estimate; source-supported partial results do not fill another target's gap. No confidential input or result is present now.

```

INPUT = authenticated source-vintage tables named in registration.tables
TARGETS = this Atom's full source inventory, including unavailable prose
↪ claims
for target in TARGETS in source order:
    required = target's exact panel/outcome/sample/model/contrast
    ↪ specification
    blockers = unresolved source choices + missing required
    ↪ vintage/key/year/crosswalk
    if blockers: emit target-status(no-estimate, exact reasons); emit
    ↪ numeric-null rows; continue
    panel = derive required source panel, with prescribed source restrictions
    ↪ in prescribed order
    verify unique source units, explicit observed year set and
    ↪ source-authorized join keys
    if source estimator cannot identify required target: emit scientific
    ↪ boundary with reason; continue
    result = evaluate only required source formula using prescribed inference
    ↪ semantics
    if execution fails: log operational failure and leave pending; do not
    ↪ record a scientific null
    else: emit all keyed estimates/SE/CI/baselines/N/source-reported
    ↪ contrasts, not a favorable subset
write results.csv; target-status.csv; interpretation.md; named figure output
↪ interfaces
interpret each outcome and contrast in its own units and uncertainty;
↪ preserve mixed results

```

Each figure output `<figure_id>-effects.csv` contains `target_id,panel,series,x,estimate,lower,upper,unit,status,reason,bin_left,bin_right`. For histogram source bins, edges are supplied only after source identification. `<figure_id>-annotations.csv` contains `target_id,panel,series,statistic,term,value,unit,status,reason` for source-reported DID,SE,difference,F,p,baseline,N. Tables/prose use `results.csv` key `target_id,panel,series,x,statistic,term` and D8 additional metadata. Numeric fields null when unavailable. Do not create actual CSV files during conversion. See `figures.yaml` and `results-map.md` for exact input names and manuscript bindings.

Scientific interpretation is outcome/estimand-specific: signed magnitude and source95% intervals (or source SE-only table), relevant source-reported subgroup contrasts, and source-compatible limitations. Statistical imprecision is retained as evidence, not an unavailable observation. There is no replication-success cutoff, invented power, equivalence band or significance-based execution. The native numeric ratings are required compatibility metadata (neutral3 and 0.5), not an elicited novelty ranking, empirical feasibility claim or success score. This replica lane has no tournament or reviewer.

Complete registration fields

```

id: atom-000004
topic: Health service effects
atom_result: prereg
atom_strength: null
atom_shape: null
evidence_paths: []
concepts: []

```

```

literature: []
bonds: []
preregistration:
  version: 1
  mode: canonical
  question: How does parental loss affect the source health-service and
    ↪ prescription
    outcomes across parent and adult-child gender?
  rationale: Health outcomes form the distinct health question, separate from
    earnings and from co-occurrence. Related services and prescription panels
    remain in one analysis.
  assumption: Source correspondence and input coverage must be established
    as specified. Causal interpretation additionally requires the source
    ↪ parallel
    counterfactual trends/no relevant anticipation assumptions; the
    ↪ health/family
    patterns do not identify mediation.
  target_strength: causal
  target_shape: comparison
  ratings:
    importance: 3
    novelty: 3
    feasibility: 3
    data_advantage: 3
    informativeness: 0.5
  tables:
    - BEF
    - FAIN
    - DODSAARS
    - DODSAASG
    - IND
    - IDAP
    - IDAN
    - IDAS
    - UDDA
    - LPR_ADM
    - PSYK_ADM
    - SYSI
    - SSSY
    - LMDB
  activation: 'TRUE'
  verdicts:
    - key: evidence
      region: All retained source targets have actual estimates and their
        ↪ source-required
        uncertainty/interpretation, with no unresolved target.
      says: Report the complete actual outcome vector. Coverage alone gives
        no direction, causal conclusion or replication confirmation.
    - key: partial
      region: At least one source target has actual evidence and at least one
        has a conclusive source/input no-estimate disposition.
      says: Report available signed estimates and uncertainty, including mixed
        findings, beside every missing target and its exact limitation. Do not
        generalize the available subset to the missing scope.
    - key: boundary
      region: No retained source target is estimable after a conclusive
        ↪ source-definition
        or input boundary; this is not an operational failure.
      says: There is no scientific estimate for this source scope. The paper
        is neither confirmed nor refuted by inability to estimate.

```

```

- key: otherwise
  region: otherwise
  says: Retain unresolved or inconclusive scientific disposition with actual
        evidence if any. Operational failures remain pending and are not done()
        or a statistical null.
protocol: |
  # Health service effects

  Take the NN sample only after shared gates. D4 gives source encounter agg
  ↪ regation. F3 A annual psychologist consultations; B any mental-health
  ↪ prescription N05/N06A/N06B/N06C; C any opioid N02A; D annual non-psyc
  ↪ hiatric hospital visits including inpatient/outpatient. Estimate Eq4
  ↪ mother/father*post separately by sex and source pooled contrasts.
  ↪ GP/psych services use years>=1990, balanced11 rows (event>=1995 throu
  ↪ gh2014). Mental/opioid/alcohol prescriptions start1995, balanced11
  ↪ (event2000-2014). Hospital outcomes restrict1994-2018 and rebalance11
  ↪ (event1999-2013). Source psychiatry hospital data1995-2018 have their
  ↪ own complete-window sample. Different sample sizes are retained by
  ↪ outcome, never harmonized without source support.
B9: A NUMBER of alcohol-addiction prescription records N07BB (not
  ↪ any-use); B annual GP consultations using source service-code list.
  ↪ B10: four dynamic outcomes in F3 order with mother/father subpanels
  ↪ and both sexes; source Eq2 dynamic interactions and95% intervals.
  ↪ B11: A screening hospital visits whose diagnosis containsDZ00 orDZ01;
  ↪ B nonscreening visits. Although clean_1 initially excludes allDZ in a
  ↪ provisional field, analysis_6 lines387-388 explicitly replaces
  ↪ nonscreening by total-minus-screening; this final consuming
  ↪ definition is source-explicit and is not an unresolved disagreement.
Counts are annual encounters/services as code, never new diagnoses. Table
  ↪ B3's health means are supplied to the labor module with their exact
  ↪ sample windows, not a new health diagnostic. Private psychiatrist pro
  ↪ se target remains A05/U02; no unreported combined psychiatric result
  ↪ replaces it. Preserve printed cross-sex and cross-parent contrasts
  ↪ and baseline means, including source significant and nonsignificant
  ↪ comparisons, without declaring a single health success label.
PSEUDOCODE: authenticate complete source event register-year -> aggregate
  ↪ D4 -> join by authenticated PNR/year (absence0 only inside verified
  ↪ coverage) -> restrict each complete-window sample -> resolve
  ↪ target-specific gaps -> fit source Eq2/4 -> export outcome-specific
  ↪ coefficients/CI/contrast/N/baseline. If one service is unavailable,
  ↪ preserve the other eligible service targets; do not substitute a
  ↪ proxy. Health-utilization changes can be consistent with a health
  ↪ channel but cannot alone prove disease severity or causal mediation.

Unresolved target gates: U01, U02, C01, A01, A02, A03, A05, A11.

Retained source targets: figure3, figure_b9, figure_b10, figure_b11,
  ↪ prose_07.

## Incorporated specification and actual evidence

Read prereg/data-contract.md D1-D8, replication/matching-specification.md
  ↪ where applicable, schema-evidence.md, target-inventory.md and
  ↪ deviations.md as part of this protocol. Source target order, field
  ↪ definitions, sample restrictions, formulae, intervals, source-only
  ↪ tests and missing-data rules are fixed there. Conflicting decisions
  ↪ remain no-estimate; source-supported partial results do not fill
  ↪ another target's gap. No confidential input or result is present now.
...

```

```

INPUT = authenticated source-vintage tables named in registration.tables
TARGETS = this Atom's full source inventory, including unavailable prose
↪ claims
for target in TARGETS in source order:
    required = target's exact panel/outcome/sample/model/contrast
    ↪ specification
    blockers = unresolved source choices + missing required
    ↪ vintage/key/year/crosswalk
    if blockers: emit target-status(no-estimate, exact reasons); emit
    ↪ numeric-null rows; continue
    panel = derive required source panel, with prescribed source
    ↪ restrictions in prescribed order
    verify unique source units, explicit observed year set and
    ↪ source-authorized join keys
    if source estimator cannot identify required target: emit scientific
    ↪ boundary with reason; continue
    result = evaluate only required source formula using prescribed
    ↪ inference semantics
    if execution fails: log operational failure and leave pending; do not
    ↪ record a scientific null
    else: emit all keyed estimates/SE/CI/baselines/N/source-reported
    ↪ contrasts, not a favorable subset
write results.csv; target-status.csv; interpretation.md; named figure
↪ output interfaces
interpret each outcome and contrast in its own units and uncertainty;
↪ preserve mixed results
...

```

Each figure output ``<figure_id>-effects.csv`` contains target_id,panel,series,x,estimate,lower,upper,unit,status,reason,bin_left,bin_right. For histogram source bins, edges are supplied only after source identification. ``<figure_id>-annotations.csv`` contains target_id,panel,series,statistic,term,value,unit,status,reason for source-reported DID,SE,difference,F,p,baseline,N. Tables/prose use results.csv key target_id,panel,series,x,statistic,term and D8 additional metadata. Numeric fields null when unavailable. Do not create actual CSV files during conversion. See figures.yaml and results-map.md for exact input names and manuscript bindings.

Scientific interpretation is outcome/estimand-specific: signed magnitude and source95% intervals (or source SE-only table), relevant source-reported subgroup contrasts, and source-compatible limitations. Statistical imprecision is retained as evidence, not an unavailable observation. There is no replication-success cutoff, invented power, equivalence band or significance-based execution. The native numeric ratings are required compatibility metadata (neutral3 and0.5), not an elicited novelty ranking, empirical feasibility claim or success score. This replica lane has no tournament or reviewer.

Registered source module

How does parental loss affect the source health-service and prescription outcomes across parent and adult-child gender?

No current result, realized strength, evidence path or empirical replication is recorded. The preregistration protocol owns scientific specification and local states; manuscript/graph views are derived. Source locators, blockers and result interfaces are incorporated above. Scientific module purpose: Health outcomes form the distinct

health question, separate from earnings and from co-occurrence. Related services and prescription panels remain in one analysis.

Reference implementation

Entry point and supported/unresolved components for `atom-000004`: [execution index](#). See [engineering review](#). This pointer does not amend the registered protocol.

`atom-000005`: Health earnings cooccurrence

Health earnings cooccurrence

Use `analysis_14`'s source extended panel to obtain earnings at $t=-6$ for lag construction, then retain matched eligible adults 25–50 and final sudden causes. Source winsorization `p1/p99` occurs before restricting calendar 1995–2018. Define year-on-year proportional change = $(\text{earnings}_t - \text{earnings}_{t-1}) / \text{earnings}_{t-1}$; decline is $\text{change} \leq -0.05$; undefined/missing ratio is coded 0 in this source program, not replaced with a log or denominator offset. Onset of decline = `decline_t` and not `decline_t-1`. Mental/psychologist/opioid onset = `any_use_t` and not `any_use_t-1`. Their joint outcome is product of both onsets. Missing lag handling follows the explicit code only inside complete covered years; unavailable registers cannot supply zero. After merging health, require full 11-year panel within 1995–2018 (death 2000–2013), then remove $t=-5$, yielding displayed -4..5 relative to omitted -1. F4 A indicator annual earnings fall $\geq 5\%$; B onset of that fall; C onset mental Rx; D joint mental/earnings onsets. B12 A psychologist onset, B joint psychologist/earnings onset, C opioid onset, D joint opioid/earnings onset, E hospital VISIT COUNT, F hospital count multiplied by earnings-decline ONSET. Despite loose caption language these last panels are counts in plot/program, not new binary hospital onsets. Do not include the other hospital-onset/doubling or any-decline variants produced only in author code. The displayed source program uses Eq2 event effects by sex, `person × match cluster`, `event/year/age/person × match FE`; preserve C05 because captions call this Eq4. Output complete dynamic vector and source Eq3 annotations where specified, with intervals and baselines. Prose ratios are joint-onset effect divided by earnings-onset effect and by mental-onset effect at the source immediate event year, separately by sex. Exact year 0 versus 1 is A03. No confidence intervals or ratio test are supplied in source, so none is invented. If a denominator is zero or unresolved, report undefined/no-estimate, not a regularized ratio or fabricated 0. PSEUDOCODE: verified lags -> source decline/onset arithmetic -> exact sample restrictions -> model/caption gate -> authorized dynamics -> ratio only if source event mapping determined -> output all panels and prose targets. Co-occurrence does not identify an indirect causal effect, fraction causally mediated, or an individual-level causal pathway.

Unresolved target gates: U01, U02, C01, A01, A02, A03, C05, A11.

Retained source targets: `figure4`, `figure_b12`, `prose_08`.

Incorporated specification and actual evidence

Read prereg/data-contract.md D1–D8, replication/matching-specification.md where applicable, schema-evidence.md, target-inventory.md and deviations.md as part of this protocol. Source target order, field definitions, sample restrictions, formulae, intervals, source-only tests and missing-data rules are fixed there. Conflicting decisions remain no-estimate; source-supported partial results do not fill another target’s gap. No confidential input or result is present now.

```
INPUT = authenticated source-vintage tables named in registration.tables
TARGETS = this Atom's full source inventory, including unavailable prose
→ claims
for target in TARGETS in source order:
  required = target's exact panel/outcome/sample/model/contrast
  → specification
  blockers = unresolved source choices + missing required
  → vintage/key/year/crosswalk
  if blockers: emit target-status(no-estimate, exact reasons); emit
  → numeric-null rows; continue
  panel = derive required source panel, with prescribed source restrictions
  → in prescribed order
  verify unique source units, explicit observed year set and
  → source-authorized join keys
  if source estimator cannot identify required target: emit scientific
  → boundary with reason; continue
  result = evaluate only required source formula using prescribed inference
  → semantics
  if execution fails: log operational failure and leave pending; do not
  → record a scientific null
  else: emit all keyed estimates/SE/CI/baselines/N/source-reported
  → contrasts, not a favorable subset
write results.csv; target-status.csv; interpretation.md; named figure output
→ interfaces
interpret each outcome and contrast in its own units and uncertainty;
→ preserve mixed results
```

Each figure output <figure_id>-effects.csv contains target_id,panel,series,x,estimate,lower,upper,unit,status,reason,bin_left,bin_right. For histogram source bins, edges are supplied only after source identification. <figure_id>-annotations.csv contains target_id,panel,series,statistic,term,value,unit,status,reason for source-reported DID,SE,difference,F,p,baseline,N. Tables/prose use results.csv key target_id,panel,series,x,statistic,term and D8 additional metadata. Numeric fields null when unavailable. Do not create actual CSV files during conversion. See figures.yaml and results-map.md for exact input names and manuscript bindings.

Scientific interpretation is outcome/estimand-specific: signed magnitude and source95% intervals (or source SE-only table), relevant source-reported subgroup contrasts, and source-compatible limitations. Statistical imprecision is retained as evidence, not an unavailable observation. There is no replication-success cutoff, invented power, equivalence band or significance-based execution. The native numeric ratings are required compatibility metadata (neutral3 and 0.5), not an elicited novelty ranking, empirical feasibility claim or success score. This replica lane has no tournament or reviewer.

The source describes a causal bereavement effect on co-occurrence under its DiD assumptions. That target is preserved; the relationship between health and earnings within the joint event remains associational and does not identify mediation.

Complete registration fields

```
id: atom-000005
topic: Health earnings cooccurrence
atom_result: prereg
atom_strength: null
atom_shape: null
evidence_paths: []
concepts: []
literature: []
bonds: []
preregistration:
  version: 1
  mode: canonical
  question: What source joint timing links earnings declines to the start
    of health treatment after parental death?
  rationale: Joint onset evidence asks a distinct timing/mechanism question;
    its related health panels and conditional-share claims stay together.
  assumption: Source matched-DiD counterfactual-trend and anticipation
    ↪ assumptions
    apply to the effect of parental death on joint outcomes. No causal effect
    of health on earnings, indirect effect or mediation share is identified
    by this co-occurrence design.
  target_strength: causal
  target_shape: temporal
  ratings:
    importance: 3
    novelty: 3
    feasibility: 3
    data_advantage: 3
    informativeness: 0.5
  tables:
    - BEF
    - FAIN
    - DODSAARS
    - DODSAASG
    - IND
    - IDAP
    - IDAN
    - IDAS
    - UDDA
    - LPR_ADM
    - SYSI
    - SSSY
    - LMDB
  activation: 'TRUE'
  verdicts:
    - key: evidence
      region: All retained source targets have actual estimates and their
        ↪ source-required
        uncertainty/interpretation, with no unresolved target.
      says: Report the complete actual outcome vector. Coverage alone gives
        no direction, causal conclusion or replication confirmation.
    - key: partial
      region: At least one source target has actual evidence and at least one
```

has a conclusive source/input no-estimate disposition.

says: Report available signed estimates and uncertainty, including mixed findings, beside every missing target and its exact limitation. Do not generalize the available subset to the missing scope.

- key: boundary
 - region: No retained source target is estimable after a conclusive
 - ↪ source-definition
 - or input boundary; this is not an operational failure.
 - says: There is no scientific estimate for this source scope. The paper is neither confirmed nor refuted by inability to estimate.
- key: otherwise
 - region: otherwise
 - says: Retain unresolved or inconclusive scientific disposition with actual evidence if any. Operational failures remain pending and are not done() or a statistical null.

protocol: |

Health earnings cooccurrence

Use analysis_14's source extended panel to obtain earnings at t=-6 for

- ↪ lag construction, then retain matched eligible adults25-50 and final sudden causes. Source winsorization p1/p99 occurs before restricting calendar1995-2018. Define year-on-year proportional change=(earnings_t-earnings_t-1)/earnings_t-1; decline is change<=-0.05;
- ↪ undefined/missing ratio is coded0 in this source program, not replaced with a log or denominator offset. Onset of decline=decline_t and not decline_t-1. Mental/psychologist/opioid onset=any use_t and not any use_t-1. Their joint outcome is product of both onsets.
- ↪ Missing lag handling follows the explicit code only inside complete covered years; unavailable registers cannot supply zero.

After merging health, require full11-year panel within1995-2018

- ↪ (death2000-2013), then remove t=-5, yielding displayed-4.5 relative to omitted-1. F4 A indicator annual earnings fall>=5%; B onset of that fall; C onset mental Rx; D joint mental/earnings onsets. B12 A psychologist onset,B joint psychologist/earnings onset,C opioid onset,D joint opioid/earnings onset,E hospital VISIT COUNT,F hospital count multiplied by earnings-decline ONSET. Despite loose caption language these last panels are counts in plot/program, not new binary hospital onsets. Do not include the other hospital-onset/doubling or any-decline variants produced only in author code.

The displayed source program uses Eq2 event effects by sex, person*match cluster, event/year/age/person*match FE; preserve C05 because captions call this Eq4. Output complete dynamic vector and source Eq3 annotations where specified, with intervals and baselines. Prose ratios are joint-onset effect divided by earnings-onset effect and by mental-onset effect at the source immediate event year, separately by sex. Exact year0 versus1 is A03. No confidence intervals or ratio test are supplied in source, so none is invented. If a denominator is zero or unresolved, report undefined/no-estimate, not a regularized ratio or fabricated0.

PSEUDOCODE: verified lags -> source decline/onset arithmetic -> exact

- ↪ sample restrictions -> model/caption gate -> authorized dynamics -> ratio only if source event mapping determined -> output all panels and prose targets. Co-occurrence does not identify an indirect causal effect, fraction causally mediated, or an individual-level causal pathway.

Unresolved target gates: U01, U02, C01, A01, A02, A03, C05, A11.

Retained source targets: figure4, figure_b12, prose_08.

```

## Incorporated specification and actual evidence

Read prereg/data-contract.md D1-D8, replication/matching-specification.md
↪ where applicable, schema-evidence.md, target-inventory.md and
↪ deviations.md as part of this protocol. Source target order, field
↪ definitions, sample restrictions, formulae, intervals, source-only
↪ tests and missing-data rules are fixed there. Conflicting decisions
↪ remain no-estimate; source-supported partial results do not fill
↪ another target's gap. No confidential input or result is present now.

...

INPUT = authenticated source-vintage tables named in registration.tables
TARGETS = this Atom's full source inventory, including unavailable prose
↪ claims
for target in TARGETS in source order:
  required = target's exact panel/outcome/sample/model/contrast
  ↪ specification
  blockers = unresolved source choices + missing required
  ↪ vintage/key/year/crosswalk
  if blockers: emit target-status(no-estimate, exact reasons); emit
  ↪ numeric-null rows; continue
  panel = derive required source panel, with prescribed source
  ↪ restrictions in prescribed order
  verify unique source units, explicit observed year set and
  ↪ source-authorized join keys
  if source estimator cannot identify required target: emit scientific
  ↪ boundary with reason; continue
  result = evaluate only required source formula using prescribed
  ↪ inference semantics
  if execution fails: log operational failure and leave pending; do not
  ↪ record a scientific null
  else: emit all keyed estimates/SE/CI/baselines/N/source-reported
  ↪ contrasts, not a favorable subset
write results.csv; target-status.csv; interpretation.md; named figure
↪ output interfaces
interpret each outcome and contrast in its own units and uncertainty;
↪ preserve mixed results
...

Each figure output `<figure_id>-effects.csv` contains target_id,panel,ser
↪ ies,x,estimate,lower,upper,unit,status,reason,bin_left,bin_right. For
↪ histogram source bins, edges are supplied only after source
↪ identification. `<figure_id>-annotations.csv` contains
↪ target_id,panel,series,statistic,term,value,unit,status,reason for
↪ source-reported DID,SE,difference,F,p,baseline,N. Tables/prose use
↪ results.csv key target_id,panel,series,x,statistic,term and D8
↪ additional metadata. Numeric fields null when unavailable. Do not
↪ create actual CSV files during conversion. See figures.yaml and
↪ results-map.md for exact input names and manuscript bindings.

Scientific interpretation is outcome/estimand-specific: signed magnitude
↪ and source95% intervals (or source SE-only table), relevant
↪ source-reported subgroup contrasts, and source-compatible
↪ limitations. Statistical imprecision is retained as evidence, not an
↪ unavailable observation. There is no replication-success cutoff,
↪ invented power, equivalence band or significance-based execution. The
↪ native numeric ratings are required compatibility metadata (neutral3
↪ and0.5), not an elicited novelty ranking, empirical feasibility claim
↪ or success score. This replica lane has no tournament or reviewer.

```

The source describes a causal bereavement effect on co-occurrence under
 ↳ its DiD assumptions. That target is preserved; the relationship
 ↳ between health and earnings within the joint event remains
 ↳ associational and does not identify mediation.

Registered source module

What source joint timing links earnings declines to the start of health treatment after parental death?

No current result, realized strength, evidence path or empirical replication is recorded. The preregistration protocol owns scientific specification and local states; manuscript/graph views are derived. Source locators, blockers and result interfaces are incorporated above. Scientific module purpose: Joint onset evidence asks a distinct timing/mechanism question; its related health panels and conditional-share claims stay together.

Reference implementation

Entry point and supported/unresolved components for `atom-000005`: [execution index](#). See [engineering review](#). This pointer does not amend the registered protocol.

`atom-000006`: Childcare support

Childcare support

Freeze treated adult's youngest-child group at event $t=0$ and propagate it to the matched pair; the source definition is `age0-5` versus `without-young`, with unresolved code removal of childless adults `C06`. `analysis__8` restricts first, winsorizes pooled earnings `p1/p99` after childless exclusion, indexes earnings to `baseline100` separately by `sex×young` group. `F5` panels mother/father each contain men/women with/without young children; estimate `Eq4` separately by child group, preserve cross-sex differences within parent/group and young-versus-no-young contrasts within sex. Do not invent additional childless estimates. `B13` refines to `youngest0-5` versus `6-14`; source keeps adults with `child≤14` at event, retaining the source winsorization stage then reindexes within these age groups. Same mother/father panels and contrasts. `B14` formal childcare: `D5` builds child-level enrollment; join youngest-child identifier to enrollment and then adult/pair. Restrict `calendar≥2005` and `complete11` observations through 2019, yielding event 2010-2014. PanelA code group `youngest≤5`; panelB6-10; wording disagreement `A06` remains. Source young-child values are blanked when `age>5` in follow-up then missing `set0`, and older code uses its corresponding enrolled-child series; record measurement implications without replacing definitions. Keep outcome as uptake probability, no invented hours/costs. PSEUDOCODE: resolve `C06/A06` eligibility before creating analytic subgroup -> source exact timing, youngest identity and pair propagation -> outcome-specific restriction and denominator -> `Eq4` and printed contrast models -> output all8 `F5/B13` coefficients plus source comparisons, `B14` mother/father×sex uptake with `CI/baselines/N`. Preserve uncertainty and

partial availability. Earnings heterogeneity and uptake are suggestive about family support; neither directly measures grandparent care hours nor establishes a mediation share.

Unresolved target gates: U01, U02, C01, A01, A02, A03, C06, A06, A11.

Retained source targets: figure5, figure_b13, figure_b14.

Incorporated specification and actual evidence

Read prereg/data-contract.md D1–D8, replication/matching-specification.md where applicable, schema-evidence.md, target-inventory.md and deviations.md as part of this protocol. Source target order, field definitions, sample restrictions, formulae, intervals, source-only tests and missing-data rules are fixed there. Conflicting decisions remain no-estimate; source-supported partial results do not fill another target’s gap. No confidential input or result is present now.

```
INPUT = authenticated source-vintage tables named in registration.tables
TARGETS = this Atom's full source inventory, including unavailable prose
↪ claims
for target in TARGETS in source order:
    required = target's exact panel/outcome/sample/model/contrast
    ↪ specification
    blockers = unresolved source choices + missing required
    ↪ vintage/key/year/crosswalk
    if blockers: emit target-status(no-estimate, exact reasons); emit
    ↪ numeric-null rows; continue
    panel = derive required source panel, with prescribed source restrictions
    ↪ in prescribed order
    verify unique source units, explicit observed year set and
    ↪ source-authorized join keys
    if source estimator cannot identify required target: emit scientific
    ↪ boundary with reason; continue
    result = evaluate only required source formula using prescribed inference
    ↪ semantics
    if execution fails: log operational failure and leave pending; do not
    ↪ record a scientific null
    else: emit all keyed estimates/SE/CI/baselines/N/source-reported
    ↪ contrasts, not a favorable subset
write results.csv; target-status.csv; interpretation.md; named figure output
↪ interfaces
interpret each outcome and contrast in its own units and uncertainty;
↪ preserve mixed results
```

Each figure output <figure_id>-effects.csv contains target_id,panel,series,x,estimate,lower,upper,unit,status,reason,bin_left,bin_right. For histogram source bins, edges are supplied only after source identification. <figure_id>-annotations.csv contains target_id,panel,series,statistic,term,value,unit,status,reason for source-reported DID,SE,difference,F,p,baseline,N. Tables/prose use results.csv key target_id,panel,series,x,statistic,term and D8 additional metadata. Numeric fields null when unavailable. Do not create actual CSV files during conversion. See figures.yaml and results-map.md for exact input names and manuscript bindings.

Scientific interpretation is outcome/estimand-specific: signed magnitude and source95% intervals (or source SE-only table), relevant source-reported subgroup

contrasts, and source-compatible limitations. Statistical imprecision is retained as evidence, not an unavailable observation. There is no replication-success cutoff, invented power, equivalence band or significance-based execution. The native numeric ratings are required compatibility metadata (neutral3 and 0.5), not an elicited novelty ranking, empirical feasibility claim or success score. This replica lane has no tournament or reviewer.

Complete registration fields

```
id: atom-000006
topic: Childcare support
atom_result: prereg
atom_strength: null
atom_shape: null
evidence_paths: []
concepts: []
literature: []
bonds: []
preregistration:
  version: 1
  mode: canonical
  question: Does the source earnings response vary with young children, and
    does formal childcare use respond to parental loss?
  rationale: Earnings contrasts and formal-care uptake jointly address loss
    of informal childcare; separate measures/panels are not separate
    ↪ experimental
    nodes.
  assumption: Source correspondence and input coverage must be established
    as specified. Causal interpretation additionally requires the source
    ↪ parallel
    counterfactual trends/no relevant anticipation assumptions; the
    ↪ health/family
    patterns do not identify mediation.
  target_strength: causal
  target_shape: comparison
ratings:
  importance: 3
  novelty: 3
  feasibility: 3
  data_advantage: 3
  informativeness: 0.5
tables:
- BEF
- FAIN
- DODSAARS
- DODSAASG
- IND
- IDAP
- IDAN
- IDAS
- UDDA
- DAGI
- BOERNFB
- BOERNSB
activation: 'TRUE'
verdicts:
- key: evidence
```

region: All retained source targets have actual estimates and their
↳ source-required
uncertainty/interpretation, with no unresolved target.
says: Report the complete actual outcome vector. Coverage alone gives
no direction, causal conclusion or replication confirmation.

- key: partial
region: At least one source target has actual evidence and at least one
has a conclusive source/input no-estimate disposition.
says: Report available signed estimates and uncertainty, including mixed
findings, beside every missing target and its exact limitation. Do not
generalize the available subset to the missing scope.

- key: boundary
region: No retained source target is estimable after a conclusive
↳ source-definition
or input boundary; this is not an operational failure.
says: There is no scientific estimate for this source scope. The paper
is neither confirmed nor refuted by inability to estimate.

- key: otherwise
region: otherwise
says: Retain unresolved or inconclusive scientific disposition with actual
evidence if any. Operational failures remain pending and are not done()
or a statistical null.

protocol: |
Childcare support

Freeze treated adult's youngest-child group at event t=0 and propagate it
↳ to the matched pair; the source definition is age0-5 versus
↳ without-young, with unresolved code removal of childless adults C06.
↳ analysis_8 restricts first, winsorizes pooled earnings p1/p99 after
↳ childless exclusion, indexes earnings to baseline100 separately by
↳ sex×young group. F5 panels mother/father each contain men/women
↳ with/without young children; estimate Eq4 separately by child group,
↳ preserve cross-sex differences within parent/group and
↳ young-versus-no-young contrasts within sex. Do not invent additional
↳ childless estimates.

B13 refines to youngest0-5 versus6-14; source keeps adults with child≤14
↳ at event, retaining the source winsorization stage then reindexes
↳ within these age groups. Same mother/father panels and contrasts. B14
↳ formal childcare: D5 builds child-level enrollment; join
↳ youngest-child identifier to enrollment and then adult/pair. Restrict
↳ calendar>=2005 and complete11 observations through2019, yielding
↳ event2010-2014. PanelA code group youngest≤5; panelB6-10; wording
↳ disagreement A06 remains. Source young-child values are blanked when
↳ age>5 in follow-up then missing set0, and older code uses its
↳ corresponding enrolled-child series; record measurement implications
↳ without replacing definitions. Keep outcome as uptake probability, no
↳ invented hours/costs.

PSEUDOCODE: resolve C06/A06 eligibility before creating analytic subgroup
↳ -> source exact timing, youngest identity and pair propagation ->
↳ outcome-specific restriction and denominator -> Eq4 and printed
↳ contrast models -> output all8 F5/B13 coefficients plus source
↳ comparisons, B14 mother/father×sex uptake with CI/baselines/N.
↳ Preserve uncertainty and partial availability. Earnings heterogeneity
↳ and uptake are suggestive about family support; neither directly
↳ measures grandparent care hours nor establishes a mediation share.

Unresolved target gates: U01, U02, C01, A01, A02, A03, C06, A06, A11.

Retained source targets: figure5, figure_b13, figure_b14.

```

## Incorporated specification and actual evidence

Read prereg/data-contract.md D1-D8, replication/matching-specification.md
↳ where applicable, schema-evidence.md, target-inventory.md and
↳ deviations.md as part of this protocol. Source target order, field
↳ definitions, sample restrictions, formulae, intervals, source-only
↳ tests and missing-data rules are fixed there. Conflicting decisions
↳ remain no-estimate; source-supported partial results do not fill
↳ another target's gap. No confidential input or result is present now.

...

INPUT = authenticated source-vintage tables named in registration.tables
TARGETS = this Atom's full source inventory, including unavailable prose
↳ claims
for target in TARGETS in source order:
  required = target's exact panel/outcome/sample/model/contrast
  ↳ specification
  blockers = unresolved source choices + missing required
  ↳ vintage/key/year/crosswalk
  if blockers: emit target-status(no-estimate, exact reasons); emit
  ↳ numeric-null rows; continue
  panel = derive required source panel, with prescribed source
  ↳ restrictions in prescribed order
  verify unique source units, explicit observed year set and
  ↳ source-authorized join keys
  if source estimator cannot identify required target: emit scientific
  ↳ boundary with reason; continue
  result = evaluate only required source formula using prescribed
  ↳ inference semantics
  if execution fails: log operational failure and leave pending; do not
  ↳ record a scientific null
  else: emit all keyed estimates/SE/CI/baselines/N/source-reported
  ↳ contrasts, not a favorable subset
write results.csv; target-status.csv; interpretation.md; named figure
↳ output interfaces
interpret each outcome and contrast in its own units and uncertainty;
↳ preserve mixed results
...

Each figure output `<figure_id>-effects.csv` contains target_id,panel,ser
↳ ies,x,estimate,lower,upper,unit,status,reason,bin_left,bin_right. For
↳ histogram source bins, edges are supplied only after source
↳ identification. `<figure_id>-annotations.csv` contains
↳ target_id,panel,series,statistic,term,value,unit,status,reason for
↳ source-reported DID,SE,difference,F,p,baseline,N. Tables/prose use
↳ results.csv key target_id,panel,series,x,statistic,term and D8
↳ additional metadata. Numeric fields null when unavailable. Do not
↳ create actual CSV files during conversion. See figures.yaml and
↳ results-map.md for exact input names and manuscript bindings.

Scientific interpretation is outcome/estimand-specific: signed magnitude
↳ and source95% intervals (or source SE-only table), relevant
↳ source-reported subgroup contrasts, and source-compatible
↳ limitations. Statistical imprecision is retained as evidence, not an
↳ unavailable observation. There is no replication-success cutoff,
↳ invented power, equivalence band or significance-based execution. The
↳ native numeric ratings are required compatibility metadata (neutral3
↳ and0.5), not an elicited novelty ranking, empirical feasibility claim
↳ or success score. This replica lane has no tournament or reviewer.

```

Registered source module

Does the source earnings response vary with young children, and does formal childcare use respond to parental loss?

No current result, realized strength, evidence path or empirical replication is recorded. The preregistration protocol owns scientific specification and local states; manuscript/graph views are derived. Source locators, blockers and result interfaces are incorporated above. Scientific module purpose: Earnings contrasts and formal-care uptake jointly address loss of informal childcare; separate measures/panels are not separate experimental nodes.

Reference implementation

Entry point and supported/unresolved components for `atom-000006`: [execution index](#). See [engineering review](#). This pointer does not amend the registered protocol.

`atom-000007`: Surviving parent care

Surviving parent care

Restrict the source NN panel to adults without young children (youngest<6 excluded at event; childless adults are retained in `analysis_12`). Surviving parent is the other parent relative to treated first death; propagate treated illness classification to the matched pair. D4 produces union hospital-days per calendar year including inpatient and outpatient encounters. Source text asks >three months over five years; code uses any annual total ≥ 90 in `t0..5` for F6 and `t-5..-1` for B15. Do not select one; C07 blocks affected classification and C05 the model label. Source earnings win-sORIZATION `p1/p99` follows the restriction; `index100` uses sex-specific `t=-1` means, not `sex $\times$ illness` means. The consuming model has illness-specific `treatment $\times$ post` terms with `event/year/age/person $\times$ match` FE, `clusterperson $\times$ match`. F6 and B15 show men/women by surviving-parent ill/not ill and source contrasts. Source calendar filter starts 1994 and rebalances 11 but lacks 2018 end, causing 2019 zero-fill against unavailable hospital-days; preserve this conflict. B16 outcome is same 11-NUTS3 region as surviving parent in each year, detailed region ≥ 1985 and hospital ≥ 1994 ; code drops pair records if surviving region unavailable; pre/post illness display assignment is unresolved. Retain prose claim comparing mental-prescription effects by illness as U03: missing exact sample, outcome time aggregation, model and output. Published 20% illness prevalence is reference only; do not use it as an assumed subgroup share or power input. No new pre/post illness alternative besides the source F6/B15 targets. PSEUDOCODE: verify years and parent identity -> unresolved accumulation/calendar/model/figure assignment gates -> only after source resolution derive illness variable -> source no-young restriction -> appropriate earnings or residence outcome -> source model/contrasts -> complete target vector or numeric-null reasons. Post-event illness is not randomized; conditioning on it does not identify a causal effect of eldercare. Interpret earnings, proximity and mental-health components separately, allowing mixed/partial results.

Unresolved target gates: U01, U02, C01, A01, A02, A03, C05, C07, U03, A11.
Retained source targets: figure6, figure_b15, figure_b16, prose_09, prose_14.

Incorporated specification and actual evidence

Read prereg/data-contract.md D1–D8, replication/matching-specification.md where applicable, schema-evidence.md, target-inventory.md and deviations.md as part of this protocol. Source target order, field definitions, sample restrictions, formulae, intervals, source-only tests and missing-data rules are fixed there. Conflicting decisions remain no-estimate; source-supported partial results do not fill another target’s gap. No confidential input or result is present now.

```
INPUT = authenticated source-vintage tables named in registration.tables
TARGETS = this Atom's full source inventory, including unavailable prose
  ↳ claims
for target in TARGETS in source order:
  required = target's exact panel/outcome/sample/model/contrast
  ↳ specification
  blockers = unresolved source choices + missing required
  ↳ vintage/key/year/crosswalk
  if blockers: emit target-status(no-estimate, exact reasons); emit
  ↳ numeric-null rows; continue
  panel = derive required source panel, with prescribed source restrictions
  ↳ in prescribed order
  verify unique source units, explicit observed year set and
  ↳ source-authorized join keys
  if source estimator cannot identify required target: emit scientific
  ↳ boundary with reason; continue
  result = evaluate only required source formula using prescribed inference
  ↳ semantics
  if execution fails: log operational failure and leave pending; do not
  ↳ record a scientific null
  else: emit all keyed estimates/SE/CI/baselines/N/source-reported
  ↳ contrasts, not a favorable subset
write results.csv; target-status.csv; interpretation.md; named figure output
  ↳ interfaces
interpret each outcome and contrast in its own units and uncertainty;
  ↳ preserve mixed results
```

Each figure output `<figure_id>-effects.csv` contains `target_id,panel,series,x,estimate,lower,upper,unit,status,reason,bin_left,bin_right`. For histogram source bins, edges are supplied only after source identification. `<figure_id>-annotations.csv` contains `target_id,panel,series,statistic,term,value,unit,status,reason` for source-reported DID,SE,difference,F,p,baseline,N. Tables/prose use results.csv key `target_id,panel,series,x,statistic,term` and D8 additional metadata. Numeric fields null when unavailable. Do not create actual CSV files during conversion. See figures.yaml and results-map.md for exact input names and manuscript bindings.

Scientific interpretation is outcome/estimand-specific: signed magnitude and source95% intervals (or source SE-only table), relevant source-reported subgroup contrasts, and source-compatible limitations. Statistical imprecision is retained as evidence, not an unavailable observation. There is no replication-success cutoff, invented power, equivalence band or significance-based execution. The native numeric ratings

are required compatibility metadata (neutral3 and 0.5), not an elicited novelty ranking, empirical feasibility claim or success score. This replica lane has no tournament or reviewer.

Complete registration fields

```
id: atom-000007
topic: Surviving parent care
atom_result: prereg
atom_strength: null
atom_shape: null
evidence_paths: []
concepts: []
literature: []
bonds: []
preregistration:
  version: 1
  mode: canonical
  question: How do source earnings and proximity responses vary with the
    ↪ surviving
      parent's health?
  rationale: Post/pre illness contrasts, residence response and source
    ↪ mental-health
      comparison address the surviving-parent support question and stay within
      one module.
  assumption: Source correspondence and input coverage must be established
    as specified. Interpretation stays within this source
    ↪ descriptive/associational
      scope; shared treatment does not establish causal mediation.
  target_strength: associational
  target_shape: comparison
  ratings:
    importance: 3
    novelty: 3
    feasibility: 3
    data_advantage: 3
    informativeness: 0.5
  tables:
    - BEF
    - FAIN
    - DODSAARS
    - DODSAASG
    - IND
    - IDAP
    - IDAN
    - IDAS
    - UDDA
    - LPR_ADM
    - LMDB
  activation: 'TRUE'
  verdicts:
    - key: evidence
      region: All retained source targets have actual estimates and their
        ↪ source-required
          uncertainty/interpretation, with no unresolved target.
      says: Report the complete actual outcome vector. Coverage alone gives
        no direction, causal conclusion or replication confirmation.
    - key: partial
      region: At least one source target has actual evidence and at least one
```

has a conclusive source/input no-estimate disposition.
says: Report available signed estimates and uncertainty, including mixed findings, beside every missing target and its exact limitation. Do not generalize the available subset to the missing scope.

- key: boundary
 - region: No retained source target is estimable after a conclusive
 - ↪ source-definition
 - or input boundary; this is not an operational failure.
 - says: There is no scientific estimate for this source scope. The paper is neither confirmed nor refuted by inability to estimate.
- key: otherwise
 - region: otherwise
 - says: Retain unresolved or inconclusive scientific disposition with actual evidence if any. Operational failures remain pending and are not done() or a statistical null.

protocol: |
Surviving parent care

Restrict the source NN panel to adults without young children (youngest<6
↪ excluded at event; childless adults are retained in analysis_12).
↪ Surviving parent is the other parent relative to treated first death;
↪ propagate treated illness classification to the matched pair. D4
↪ produces union hospital-days per calendar year including inpatient
↪ and outpatient encounters. Source text asks >three months over five
↪ years; code uses any annual total>=90 in t0..5 for F6 and t-5..-1 for
↪ B15. Do not select one; C07 blocks affected classification and C05
↪ the model label.

Source earnings winsorization p1/p99 follows the restriction; index100
↪ uses sex-specific t=-1 means, not sex*illness means. The consuming
↪ model has illness-specific treatment*post terms with
↪ event/year/age/person*match FE, clusterperson*match. F6 and B15 show
↪ men/women by surviving-parent ill/not ill and source contrasts.
↪ Source calendar filter starts1994 and rebalances11 but lacks2018 end,
↪ causing2019 zero-fill against unavailable hospital-days; preserve
↪ this conflict. B16 outcome is same11-NUTS3 region as surviving parent
↪ in each year, detailed region>=1985 and hospital>=1994; code drops
↪ pair records if surviving region unavailable; pre/post illness
↪ display assignment is unresolved.

Retain prose claim comparing mental-prescription effects by illness as
↪ U03: missing exact sample, outcome time aggregation, model and
↪ output. Published20% illness prevalence is reference only; do not use
↪ it as an assumed subgroup share or power input. No new pre/post
↪ illness alternative besides the source F6/B15 targets.

PSEUDOCODE: verify years and parent identity -> unresolved accumulation/c
↪ alendar/model/figure assignment gates -> only after source resolution
↪ derive illness variable -> source no-young restriction -> appropriate
↪ earnings or residence outcome -> source model/contrasts -> complete
↪ target vector or numeric-null reasons. Post-event illness is not
↪ randomized; conditioning on it does not identify a causal effect of
↪ eldercare. Interpret earnings, proximity and mental-health components
↪ separately, allowing mixed/partial results.

Unresolved target gates: U01, U02, C01, A01, A02, A03, C05, C07, U03, A11.

Retained source targets: figure6, figure_b15, figure_b16, prose_09,
↪ prose_14.

Incorporated specification and actual evidence

```

Read prereg/data-contract.md D1-D8, replication/matching-specification.md
↪ where applicable, schema-evidence.md, target-inventory.md and
↪ deviations.md as part of this protocol. Source target order, field
↪ definitions, sample restrictions, formulae, intervals, source-only
↪ tests and missing-data rules are fixed there. Conflicting decisions
↪ remain no-estimate; source-supported partial results do not fill
↪ another target's gap. No confidential input or result is present now.

```

```

...

```

```

INPUT = authenticated source-vintage tables named in registration.tables
TARGETS = this Atom's full source inventory, including unavailable prose
↪ claims
for target in TARGETS in source order:
  required = target's exact panel/outcome/sample/model/contrast
  ↪ specification
  blockers = unresolved source choices + missing required
  ↪ vintage/key/year/crosswalk
  if blockers: emit target-status(no-estimate, exact reasons); emit
  ↪ numeric-null rows; continue
  panel = derive required source panel, with prescribed source
  ↪ restrictions in prescribed order
  verify unique source units, explicit observed year set and
  ↪ source-authorized join keys
  if source estimator cannot identify required target: emit scientific
  ↪ boundary with reason; continue
  result = evaluate only required source formula using prescribed
  ↪ inference semantics
  if execution fails: log operational failure and leave pending; do not
  ↪ record a scientific null
  else: emit all keyed estimates/SE/CI/baselines/N/source-reported
  ↪ contrasts, not a favorable subset
write results.csv; target-status.csv; interpretation.md; named figure
↪ output interfaces
interpret each outcome and contrast in its own units and uncertainty;
↪ preserve mixed results
...

```

```

Each figure output `<figure_id>-effects.csv` contains target_id,panel,ser
↪ ies,x,estimate,lower,upper,unit,status,reason,bin_left,bin_right. For
↪ histogram source bins, edges are supplied only after source
↪ identification. `<figure_id>-annotations.csv` contains
↪ target_id,panel,series,statistic,term,value,unit,status,reason for
↪ source-reported DID,SE,difference,F,p,baseline,N. Tables/prose use
↪ results.csv key target_id,panel,series,x,statistic,term and D8
↪ additional metadata. Numeric fields null when unavailable. Do not
↪ create actual CSV files during conversion. See figures.yaml and
↪ results-map.md for exact input names and manuscript bindings.

```

```

Scientific interpretation is outcome/estimand-specific: signed magnitude
↪ and source95% intervals (or source SE-only table), relevant
↪ source-reported subgroup contrasts, and source-compatible
↪ limitations. Statistical imprecision is retained as evidence, not an
↪ unavailable observation. There is no replication-success cutoff,
↪ invented power, equivalence band or significance-based execution. The
↪ native numeric ratings are required compatibility metadata (neutral3
↪ and0.5), not an elicited novelty ranking, empirical feasibility claim
↪ or success score. This replica lane has no tournament or reviewer.

```

Registered source module

How do source earnings and proximity responses vary with the surviving parent's health?

No current result, realized strength, evidence path or empirical replication is recorded. The preregistration protocol owns scientific specification and local states; manuscript/graph views are derived. Source locators, blockers and result interfaces are incorporated above. Scientific module purpose: Post/pre illness contrasts, residence response and source mental-health comparison address the surviving-parent support question and stay within one module.

Reference implementation

Entry point and supported/unresolved components for `atom-000007`: [execution index](#). See [engineering review](#). This pointer does not amend the registered protocol.

atom-000008: Inheritance channel

Inheritance channel

Read D3 wealth: FORM1984-97, FORMREST_NY05 from1998, source2020 inflation factors, source maximum record if duplicate. analysis_9 links mother and father assets at $t=-1$, assigns first-deceased and surviving role, propagates each by pair maximum, sums them. Source missing values set0 only if coverage permits source non-event interpretation. TableB4 caption says1000DKK whereas program divides DKK sum by10000; C08 prevents choosing unit/owner. The target is interaction of treatment $\times$ post with parental baseline wealth plus treatment $\times$ post, no added main time-invariant wealth term required beyond individual FE. Earnings are source p1/p99 winsorized, sex-indexed100, with source event/year/age/person $\times$ match effects and person $\times$ match clustering. Four columns mother-men,mother-women,father-men,father-women; report interaction coefficient/SE, treatment coefficient/SE, observations, source control mean men and women. The own-asset interaction in A.A, deceased-only and surviving-only wealth in fn29 are separate target rows within this Atom but lack full specifications U04. Do not assert bequests are directly measured; parental net assets are a proxy. Do not construct an own-wealth change regression from the source prose. PSEUDOCODE: authenticate wealth vintages $\rightarrow$ gate C08 units/ownership $\rightarrow$ only resolved source unit and owner can form interaction $\rightarrow$ fit stated model $\rightarrow$ table rows and narrative $\rightarrow$ explicit no-estimate for U04. A small or imprecise interaction is not proof of no inheritance effect; no economically-negligible band or equivalence test is registered.

Unresolved target gates: U01, U02, C01, A01, A02, A03, C08, U04, A11.

Retained source targets: table_b4, prose_10, prose_11, prose_12.

Incorporated specification and actual evidence

Read prereg/data-contract.md D1-D8, replication/matching-specification.md where applicable, schema-evidence.md, target-inventory.md and deviations.md as part of this

protocol. Source target order, field definitions, sample restrictions, formulae, intervals, source-only tests and missing-data rules are fixed there. Conflicting decisions remain no-estimate; source-supported partial results do not fill another target's gap. No confidential input or result is present now.

```

INPUT = authenticated source-vintage tables named in registration.tables
TARGETS = this Atom's full source inventory, including unavailable prose
↳ claims
for target in TARGETS in source order:
  required = target's exact panel/outcome/sample/model/contrast
  ↳ specification
  blockers = unresolved source choices + missing required
  ↳ vintage/key/year/crosswalk
  if blockers: emit target-status(no-estimate, exact reasons); emit
  ↳ numeric-null rows; continue
  panel = derive required source panel, with prescribed source restrictions
  ↳ in prescribed order
  verify unique source units, explicit observed year set and
  ↳ source-authorized join keys
  if source estimator cannot identify required target: emit scientific
  ↳ boundary with reason; continue
  result = evaluate only required source formula using prescribed inference
  ↳ semantics
  if execution fails: log operational failure and leave pending; do not
  ↳ record a scientific null
  else: emit all keyed estimates/SE/CI/baselines/N/source-reported
  ↳ contrasts, not a favorable subset
write results.csv; target-status.csv; interpretation.md; named figure output
↳ interfaces
interpret each outcome and contrast in its own units and uncertainty;
↳ preserve mixed results

```

Each figure output `<figure_id>-effects.csv` contains `target_id,panel,series,x,estimate,lower,upper,unit,status,reason,bin_left,bin_right`. For histogram source bins, edges are supplied only after source identification. `<figure_id>-annotations.csv` contains `target_id,panel,series,statistic,term,value,unit,status,reason` for source-reported DID,SE,difference,F,p,baseline,N. Tables/prose use `results.csv` key `target_id,panel,series,x,statistic,term` and D8 additional metadata. Numeric fields null when unavailable. Do not create actual CSV files during conversion. See `figures.yaml` and `results-map.md` for exact input names and manuscript bindings.

Scientific interpretation is outcome/estimand-specific: signed magnitude and source95% intervals (or source SE-only table), relevant source-reported subgroup contrasts, and source-compatible limitations. Statistical imprecision is retained as evidence, not an unavailable observation. There is no replication-success cutoff, invented power, equivalence band or significance-based execution. The native numeric ratings are required compatibility metadata (neutral3 and 0.5), not an elicited novelty ranking, empirical feasibility claim or success score. This replica lane has no tournament or reviewer.

Complete registration fields

```
id: atom-000008
topic: Inheritance channel
atom_result: prereg
atom_strength: null
atom_shape: null
evidence_paths: []
concepts: []
literature: []
bonds: []
preregistration:
  version: 1
  mode: canonical
  question: Does the source relationship between pre-death family wealth and
    earnings responses support its inheritance explanation?
  rationale: All parental/own-wealth variants address the inheritance
    ↪ question;
    they are retained together even when unavailable.
  assumption: Source correspondence and input coverage must be established
    as specified. Interpretation stays within this source
    ↪ descriptive/associational
    scope; shared treatment does not establish causal mediation.
  target_strength: associational
  target_shape: comparison
  ratings:
    importance: 3
    novelty: 3
    feasibility: 3
    data_advantage: 3
    informativeness: 0.5
  tables:
    - BEF
    - FAIN
    - DODSAARS
    - DODSAASG
    - IND
    - IDAP
    - IDAN
    - IDAS
    - UDDA
  activation: 'TRUE'
  verdicts:
    - key: evidence
      region: All retained source targets have actual estimates and their
        ↪ source-required
        uncertainty/interpretation, with no unresolved target.
      says: Report the complete actual outcome vector. Coverage alone gives
        no direction, causal conclusion or replication confirmation.
    - key: partial
      region: At least one source target has actual evidence and at least one
        has a conclusive source/input no-estimate disposition.
      says: Report available signed estimates and uncertainty, including mixed
        findings, beside every missing target and its exact limitation. Do not
        generalize the available subset to the missing scope.
    - key: boundary
      region: No retained source target is estimable after a conclusive
        ↪ source-definition
        or input boundary; this is not an operational failure.
      says: There is no scientific estimate for this source scope. The paper
        is neither confirmed nor refuted by inability to estimate.
```

```

- key: otherwise
  region: otherwise
  says: Retain unresolved or inconclusive scientific disposition with actual
        evidence if any. Operational failures remain pending and are not done()
        or a statistical null.
protocol: |
  # Inheritance channel

Read D3 wealth: FORM1984-97, FORMREST_NY05 from1998, source2020 inflation
↳ factors, source maximum record if duplicate. analysis_9 links mother
↳ and father assets at t=-1, assigns first-deceased and surviving role,
↳ propagates each by pair maximum, sums them. Source missing values
↳ set0 only if coverage permits source non-event interpretation.
↳ TableB4 caption says1000DKK whereas program divides DKK sum by10000;
↳ C08 prevents choosing unit/owner. The target is interaction of
↳ treatment*post with parental baseline wealth plus treatment*post, no
↳ added main time-invariant wealth term required beyond individual FE.
↳ Earnings are source p1/p99 winsorized, sex-indexed100, with source
↳ event/year/age/person*match effects and person*match clustering. Four
↳ columns mother-men,mother-women,father-men,father-women; report
↳ interaction coefficient/SE, treatment coefficient/SE, observations,
↳ source control mean men and women.
The own-asset interaction in A.A, deceased-only and surviving-only wealth
↳ in fn29 are separate target rows within this Atom but lack full
↳ specifications U04. Do not assert bequests are directly measured;
↳ parental net assets are a proxy. Do not construct an own-wealth
↳ change regression from the source prose.
PSEUDOCODE: authenticate wealth vintages -> gate C08 units/ownership -> o
↳ nly resolved source unit and owner can form interaction -> fit stated
↳ model -> table rows and narrative -> explicit no-estimate for U04. A
↳ small or imprecise interaction is not proof of no inheritance effect;
↳ no economically-negligible band or equivalence test is registered.

Unresolved target gates: U01, U02, C01, A01, A02, A03, C08, U04, A11.

Retained source targets: table_b4, prose_10, prose_11, prose_12.

## Incorporated specification and actual evidence

Read prereg/data-contract.md D1-D8, replication/matching-specification.md
↳ where applicable, schema-evidence.md, target-inventory.md and
↳ deviations.md as part of this protocol. Source target order, field
↳ definitions, sample restrictions, formulae, intervals, source-only
↳ tests and missing-data rules are fixed there. Conflicting decisions
↳ remain no-estimate; source-supported partial results do not fill
↳ another target's gap. No confidential input or result is present now.
...

INPUT = authenticated source-vintage tables named in registration.tables
TARGETS = this Atom's full source inventory, including unavailable prose
↳ claims
for target in TARGETS in source order:
  required = target's exact panel/outcome/sample/model/contrast
  ↳ specification
  blockers = unresolved source choices + missing required
  ↳ vintage/key/year/crosswalk
  if blockers: emit target-status(no-estimate, exact reasons); emit
  ↳ numeric-null rows; continue
  panel = derive required source panel, with prescribed source
  ↳ restrictions in prescribed order

```

```

    verify unique source units, explicit observed year set and
    ↪ source-authorized join keys
    if source estimator cannot identify required target: emit scientific
    ↪ boundary with reason; continue
    result = evaluate only required source formula using prescribed
    ↪ inference semantics
    if execution fails: log operational failure and leave pending; do not
    ↪ record a scientific null
    else: emit all keyed estimates/SE/CI/baselines/N/source-reported
    ↪ contrasts, not a favorable subset
write results.csv; target-status.csv; interpretation.md; named figure
↪ output interfaces
interpret each outcome and contrast in its own units and uncertainty;
↪ preserve mixed results
`..

```

Each figure output ``<figure_id>-effects.csv`` contains `target_id,panel,series,x,estimate,lower,upper,unit,status,reason,bin_left,bin_right`. For histogram source bins, edges are supplied only after source identification. ``<figure_id>-annotations.csv`` contains `target_id,panel,series,statistic,term,value,unit,status,reason` for source-reported DID,SE,difference,F,p,baseline,N. Tables/prose use `results.csv` key `target_id,panel,series,x,statistic,term` and D8 additional metadata. Numeric fields null when unavailable. Do not create actual CSV files during conversion. See `figures.yaml` and `results-map.md` for exact input names and manuscript bindings.

Scientific interpretation is outcome/estimand-specific: signed magnitude
 ↪ and source95% intervals (or source SE-only table), relevant
 ↪ source-reported subgroup contrasts, and source-compatible
 ↪ limitations. Statistical imprecision is retained as evidence, not an
 ↪ unavailable observation. There is no replication-success cutoff,
 ↪ invented power, equivalence band or significance-based execution. The
 ↪ native numeric ratings are required compatibility metadata (neutral3
 ↪ and0.5), not an elicited novelty ranking, empirical feasibility claim
 ↪ or success score. This replica lane has no tournament or reviewer.

Registered source module

Does the source relationship between pre-death family wealth and earnings responses support its inheritance explanation?

No current result, realized strength, evidence path or empirical replication is recorded. The preregistration protocol owns scientific specification and local states; manuscript/graph views are derived. Source locators, blockers and result interfaces are incorporated above. Scientific module purpose: All parental/own-wealth variants address the inheritance question; they are retained together even when unavailable.

Reference implementation

Entry point and supported/unresolved components for `atom-000008`: [execution index](#). See [engineering review](#). This pointer does not amend the registered protocol.

atom-000009: Family formation

Family formation

Use the main NN sample and source D1/D5 demographics. B17 panelA cohabitation indicator E_FAELLE_ID nonempty, panelB number of linked children in the year. The fertility outcome is a stock count, not a new birth rate or conception measure. Use analysis_5's outcome loop for cohabitation and child, Eq4 mother/father×post separately by adult sex, event/year/age/person×match FE, clusterperson×match. Keep source child-count treatment from this consuming program (do not import the cap5 used only in descriptive tables). Mother/father×men/women effects, source95% intervals, sex differences per parent, parent differences per sex, baselines and N are the whole module. PSEUDOCODE: main-sample gate -> source yearly cohabitation/count fields -> source Eq4 and printed pooled contrasts -> emit B17 rows -> interpret signs/uncertainty alongside labor evidence. The source argues these responses inform family-behavior explanations, but a count response cannot establish a causal mediation effect or prove family choices irrelevant. No marriage-only outcome or new fertility horizon is added.

Unresolved target gates: U01, U02, C01, A01, A02, A03, A11.

Retained source targets: figure_b17.

Incorporated specification and actual evidence

Read prereg/data-contract.md D1–D8, replication/matching-specification.md where applicable, schema-evidence.md, target-inventory.md and deviations.md as part of this protocol. Source target order, field definitions, sample restrictions, formulae, intervals, source-only tests and missing-data rules are fixed there. Conflicting decisions remain no-estimate; source-supported partial results do not fill another target's gap. No confidential input or result is present now.

```
INPUT = authenticated source-vintage tables named in registration.tables
TARGETS = this Atom's full source inventory, including unavailable prose
↪ claims
for target in TARGETS in source order:
    required = target's exact panel/outcome/sample/model/contrast
    ↪ specification
    blockers = unresolved source choices + missing required
    ↪ vintage/key/year/crosswalk
    if blockers: emit target-status(no-estimate, exact reasons); emit
    ↪ numeric-null rows; continue
    panel = derive required source panel, with prescribed source restrictions
    ↪ in prescribed order
    verify unique source units, explicit observed year set and
    ↪ source-authorized join keys
    if source estimator cannot identify required target: emit scientific
    ↪ boundary with reason; continue
    result = evaluate only required source formula using prescribed inference
    ↪ semantics
    if execution fails: log operational failure and leave pending; do not
    ↪ record a scientific null
    else: emit all keyed estimates/SE/CI/baselines/N/source-reported
    ↪ contrasts, not a favorable subset
```

```

write results.csv; target-status.csv; interpretation.md; named figure output
↪ interfaces
interpret each outcome and contrast in its own units and uncertainty;
↪ preserve mixed results

```

Each figure output `<figure_id>-effects.csv` contains `target_id,panel,series,x,estimate,lower,upper,unit,status,reason,bin_left,bin_right`. For histogram source bins, edges are supplied only after source identification. `<figure_id>-annotations.csv` contains `target_id,panel,series,statistic,term,value,unit,status,reason` for source-reported DID,SE,difference,F,p,baseline,N. Tables/prose use results.csv key `target_id,panel,series,x,statistic,term` and D8 additional metadata. Numeric fields null when unavailable. Do not create actual CSV files during conversion. See figures.yaml and results.md for exact input names and manuscript bindings.

Scientific interpretation is outcome/estimand-specific: signed magnitude and source95% intervals (or source SE-only table), relevant source-reported subgroup contrasts, and source-compatible limitations. Statistical imprecision is retained as evidence, not an unavailable observation. There is no replication-success cutoff, invented power, equivalence band or significance-based execution. The native numeric ratings are required compatibility metadata (neutral3 and 0.5), not an elicited novelty ranking, empirical feasibility claim or success score. This replica lane has no tournament or reviewer.

Complete registration fields

```

id: atom-000009
topic: Family formation
atom_result: prereg
atom_strength: null
atom_shape: null
evidence_paths: []
concepts: []
literature: []
bonds: []
preregistration:
  version: 1
  mode: canonical
  question: How does the source parental-death design affect cohabitation
    and number of children?
  rationale: Cohabitation and fertility are the source family-formation
    ↪ analysis,
    independently interpretable from labor/health effects.
  assumption: Source correspondence and input coverage must be established
    as specified. Causal interpretation additionally requires the source
    ↪ parallel
    counterfactual trends/no relevant anticipation assumptions; the
    ↪ health/family
    patterns do not identify mediation.
  target_strength: causal
  target_shape: comparison
  ratings:
    importance: 3
    novelty: 3
    feasibility: 3

```

```

data_advantage: 3
informativeness: 0.5
tables:
- BEF
- FAIN
- DODSAARS
- DODSAASG
- IND
- IDAP
- IDAN
- IDAS
- UDDA
activation: 'TRUE'
verdicts:
- key: evidence
  region: All retained source targets have actual estimates and their
    ↪ source-required
    uncertainty/interpretation, with no unresolved target.
  says: Report the complete actual outcome vector. Coverage alone gives
    no direction, causal conclusion or replication confirmation.
- key: partial
  region: At least one source target has actual evidence and at least one
    has a conclusive source/input no-estimate disposition.
  says: Report available signed estimates and uncertainty, including mixed
    findings, beside every missing target and its exact limitation. Do not
    generalize the available subset to the missing scope.
- key: boundary
  region: No retained source target is estimable after a conclusive
    ↪ source-definition
    or input boundary; this is not an operational failure.
  says: There is no scientific estimate for this source scope. The paper
    is neither confirmed nor refuted by inability to estimate.
- key: otherwise
  region: otherwise
  says: Retain unresolved or inconclusive scientific disposition with actual
    evidence if any. Operational failures remain pending and are not done()
    or a statistical null.
protocol: |
  # Family formation

  Use the main NN sample and source D1/D5 demographics. B17 panelA cohabita
    ↪ tion indicator E_FAELLE_ID nonempty, panelB number of linked children
    ↪ in the year. The fertility outcome is a stock count, not a new birth
    ↪ rate or conception measure. Use analysis_5's outcome loop for
    ↪ cohabitation and child, Eq4 mother/father*post separately by adult
    ↪ sex, event/year/age/person*match FE, clusterperson*match. Keep source
    ↪ child-count treatment from this consuming program (do not import the
    ↪ cap5 used only in descriptive tables). Mother/father*men/women
    ↪ effects, source95% intervals, sex differences per parent, parent
    ↪ differences per sex, baselines and N are the whole module.
  PSEUDOCODE: main-sample gate -> source yearly cohabitation/count fields
    ↪ -> source Eq4 and printed pooled contrasts -> emit B17 rows ->
    ↪ interpret signs/uncertainty alongside labor evidence. The source
    ↪ argues these responses inform family-behavior explanations, but a
    ↪ count response cannot establish a causal mediation effect or prove
    ↪ family choices irrelevant. No marriage-only outcome or new fertility
    ↪ horizon is added.

  Unresolved target gates: U01, U02, C01, A01, A02, A03, A11.

```

Retained source targets: figure_b17.

Incorporated specification and actual evidence

```
Read prereg/data-contract.md D1-D8, replication/matching-specification.md
↳ where applicable, schema-evidence.md, target-inventory.md and
↳ deviations.md as part of this protocol. Source target order, field
↳ definitions, sample restrictions, formulae, intervals, source-only
↳ tests and missing-data rules are fixed there. Conflicting decisions
↳ remain no-estimate; source-supported partial results do not fill
↳ another target's gap. No confidential input or result is present now.

...

INPUT = authenticated source-vintage tables named in registration.tables
TARGETS = this Atom's full source inventory, including unavailable prose
↳ claims
for target in TARGETS in source order:
  required = target's exact panel/outcome/sample/model/contrast
  ↳ specification
  blockers = unresolved source choices + missing required
  ↳ vintage/key/year/crosswalk
  if blockers: emit target-status(no-estimate, exact reasons); emit
  ↳ numeric-null rows; continue
  panel = derive required source panel, with prescribed source
  ↳ restrictions in prescribed order
  verify unique source units, explicit observed year set and
  ↳ source-authorized join keys
  if source estimator cannot identify required target: emit scientific
  ↳ boundary with reason; continue
  result = evaluate only required source formula using prescribed
  ↳ inference semantics
  if execution fails: log operational failure and leave pending; do not
  ↳ record a scientific null
  else: emit all keyed estimates/SE/CI/baselines/N/source-reported
  ↳ contrasts, not a favorable subset
write results.csv; target-status.csv; interpretation.md; named figure
↳ output interfaces
interpret each outcome and contrast in its own units and uncertainty;
↳ preserve mixed results
...

Each figure output `<figure_id>-effects.csv` contains target_id,panel,ser
↳ ies,x,estimate,lower,upper,unit,status,reason,bin_left,bin_right. For
↳ histogram source bins, edges are supplied only after source
↳ identification. `<figure_id>-annotations.csv` contains
↳ target_id,panel,series,statistic,term,value,unit,status,reason for
↳ source-reported DID,SE,difference,F,p,baseline,N. Tables/prose use
↳ results.csv key target_id,panel,series,x,statistic,term and D8
↳ additional metadata. Numeric fields null when unavailable. Do not
↳ create actual CSV files during conversion. See figures.yaml and
↳ results-map.md for exact input names and manuscript bindings.
```

Scientific interpretation is outcome/estimand-specific: signed magnitude
 ↳ and source95% intervals (or source SE-only table), relevant
 ↳ source-reported subgroup contrasts, and source-compatible
 ↳ limitations. Statistical imprecision is retained as evidence, not an
 ↳ unavailable observation. There is no replication-success cutoff,
 ↳ invented power, equivalence band or significance-based execution. The
 ↳ native numeric ratings are required compatibility metadata (neutral3
 ↳ and0.5), not an elicited novelty ranking, empirical feasibility claim
 ↳ or success score. This replica lane has no tournament or reviewer.

Registered source module

How does the source parental-death design affect cohabitation and number of children?

No current result, realized strength, evidence path or empirical replication is recorded. The preregistration protocol owns scientific specification and local states; manuscript/graph views are derived. Source locators, blockers and result interfaces are incorporated above. Scientific module purpose: Cohabitation and fertility are the source family-formation analysis, independently interpretable from labor/health effects.

Reference implementation

Entry point and supported/unresolved components for `atom-000009`: [execution index](#). See [engineering review](#). This pointer does not amend the registered protocol.

atom-000010: Spousal earnings

Spousal earnings

Appendix A.C/B18 and `clean_16/analysis_16` define a separate matching procedure; do not reuse the main NN pair list as if it were the spouse design. Fix event-time cohabiting/married partner, build both biological adult and spouse histories, restrict biological bereaved adult25–50 and sudden first death; spouse youngest child<6 at event. Spouse panel requires11 observed calendar years1985–2019 (event>=1990), own parental links, matched counterpart without parent-in-law death within window. Parent-in-law mother/father role follows bereaved partner. Source matches biological adult sex; spouse sex becomes plotted sex. Do not assume same-sex couples absent. Detailed spouse matching source in `matching-specification.md` preserves age bins3 years, two positive-earnings quantiles, main/spouse histories, year-specific21-versus17 coarsened strata and the exact missing recipe boundary. The 17-group years1986–89/2010 and group table are source code, not invented new cohorts. A09 blocks population-presence versus survival, order/duplicate and contrast discrepancies. Source matching additionally uses cohabitation duration2, but the ordering to define duration is unresolved. After young-child restriction, source earnings p1/p99 winsorization and index100 by spouse sex and t=-1 mean across T/C. Eq4 in separate models uses spouse×match FE, spouse age, calendar year, event time; cluster spouse×match. Pooled gender contrast uses biological age and extra parental

variables, so cannot silently supply differences from a different model. Four plotted coefficients mother-in-law/father-in-law×men/women with source95% intervals and printed comparisons/baselines/N are retained. PSEUDOCODE: spouse-specific source matching gates -> event partner and young-child eligibility -> original complete-window restriction -> source normalization -> resolved Eq4 specification -> B18 estimates or numeric-null reasons. No new earnings response for older/childless partners, no inference that completion proves informal-care spillovers.

Unresolved target gates: U01, U02, A01, A02, A03, A09, A11.

Retained source targets: figure_b18.

Incorporated specification and actual evidence

Read prereg/data-contract.md D1–D8, replication/matching-specification.md where applicable, schema-evidence.md, target-inventory.md and deviations.md as part of this protocol. Source target order, field definitions, sample restrictions, formulae, intervals, source-only tests and missing-data rules are fixed there. Conflicting decisions remain no-estimate; source-supported partial results do not fill another target's gap. No confidential input or result is present now.

```
INPUT = authenticated source-vintage tables named in registration.tables
TARGETS = this Atom's full source inventory, including unavailable prose
↳ claims
for target in TARGETS in source order:
  required = target's exact panel/outcome/sample/model/contrast
  ↳ specification
  blockers = unresolved source choices + missing required
  ↳ vintage/key/year/crosswalk
  if blockers: emit target-status(no-estimate, exact reasons); emit
  ↳ numeric-null rows; continue
  panel = derive required source panel, with prescribed source restrictions
  ↳ in prescribed order
  verify unique source units, explicit observed year set and
  ↳ source-authorized join keys
  if source estimator cannot identify required target: emit scientific
  ↳ boundary with reason; continue
  result = evaluate only required source formula using prescribed inference
  ↳ semantics
  if execution fails: log operational failure and leave pending; do not
  ↳ record a scientific null
  else: emit all keyed estimates/SE/CI/baselines/N/source-reported
  ↳ contrasts, not a favorable subset
write results.csv; target-status.csv; interpretation.md; named figure output
↳ interfaces
interpret each outcome and contrast in its own units and uncertainty;
↳ preserve mixed results
```

Each figure output <figure_id>-effects.csv contains target_id,panel,series,x,estimate,lower,upper,unit,status,reason,bin_left,bin_right. For histogram source bins, edges are supplied only after source identification. <figure_id>-annotations.csv contains target_id,panel,series,statistic,term,value,unit,status,reason for source-reported DID,SE,difference,F,p,baseline,N. Tables/prose use results.csv key target_id,panel,series,x,statistic,term and D8 additional metadata. Numeric fields null

when unavailable. Do not create actual CSV files during conversion. See figures.yaml and results-map.md for exact input names and manuscript bindings.

Scientific interpretation is outcome/estimand-specific: signed magnitude and source95% intervals (or source SE-only table), relevant source-reported subgroup contrasts, and source-compatible limitations. Statistical imprecision is retained as evidence, not an unavailable observation. There is no replication-success cutoff, invented power, equivalence band or significance-based execution. The native numeric ratings are required compatibility metadata (neutral3 and0.5), not an elicited novelty ranking, empirical feasibility claim or success score. This replica lane has no tournament or reviewer.

Complete registration fields

```
id: atom-000010
topic: Spousal earnings
atom_result: prereg
atom_strength: null
atom_shape: null
evidence_paths: []
concepts: []
literature: []
bonds: []
preregistration:
  version: 1
  mode: canonical
  question: What are the source earnings spillovers to adult partners with
    young children after a first parent-in-law death?
  rationale: The spouse outcome and its separately matched population are
    an independent spillover question; sharing family links does not merge
    it with adult-child earnings.
  assumption: Source correspondence and input coverage must be established
    as specified. Causal interpretation additionally requires the source
    ↪ parallel
    counterfactual trends/no relevant anticipation assumptions; the
    ↪ health/family
    patterns do not identify mediation.
  target_strength: causal
  target_shape: comparison
  ratings:
    importance: 3
    novelty: 3
    feasibility: 3
    data_advantage: 3
    informativeness: 0.5
  tables:
    - BEF
    - FAIN
    - DODSAARS
    - DODSAASG
    - IND
    - IDAP
    - IDAN
    - IDAS
    - UDDA
  activation: 'TRUE'
  verdicts:
    - key: evidence
```

region: All retained source targets have actual estimates and their
↳ source-required
uncertainty/interpretation, with no unresolved target.
says: Report the complete actual outcome vector. Coverage alone gives
no direction, causal conclusion or replication confirmation.

- key: partial
region: At least one source target has actual evidence and at least one
has a conclusive source/input no-estimate disposition.
says: Report available signed estimates and uncertainty, including mixed
findings, beside every missing target and its exact limitation. Do not
generalize the available subset to the missing scope.

- key: boundary
region: No retained source target is estimable after a conclusive
↳ source-definition
or input boundary; this is not an operational failure.
says: There is no scientific estimate for this source scope. The paper
is neither confirmed nor refuted by inability to estimate.

- key: otherwise
region: otherwise
says: Retain unresolved or inconclusive scientific disposition with actual
evidence if any. Operational failures remain pending and are not done()
or a statistical null.

protocol: |
Spousal earnings

Appendix A.C/B18 and clean_16/analysis_16 define a separate matching
↳ procedure; do not reuse the main NN pair list as if it were the
↳ spouse design. Fix event-time cohabiting/married partner, build both
↳ biological adult and spouse histories, restrict biological bereaved
↳ adult25-50 and sudden first death; spouse youngest child<6 at event.
↳ Spouse panel requires11 observed calendar years1985-2019
↳ (event>=1990), own parental links, matched counterpart without
↳ parent-in-law death within window. Parent-in-law mother/father role
↳ follows bereaved partner. Source matches biological adult sex; spouse
↳ sex becomes plotted sex. Do not assume same-sex couples absent.

Detailed spouse matching source in matching-specification.md preserves
↳ age bins3 years, two positive-earnings quantiles, main/spouse
↳ histories, year-specific21-versus17 coarsened strata and the exact
↳ missing recipe boundary. The 17-group years1986-89/2010 and group
↳ table are source code, not invented new cohorts. A09 blocks
↳ population-presence versus survival, order/duplicate and contrast
↳ discrepancies. Source matching additionally uses cohabitation
↳ duration2, but the ordering to define duration is unresolved.

After young-child restriction, source earnings p1/p99 winsorization and
↳ index100 by spouse sex and t=-1 mean across T/C. Eq4 in separate
↳ models uses spouse*match FE, spouse age, calendar year, event time;
↳ cluster spouse*match. Pooled gender contrast uses biological age and
↳ extra parental variables, so cannot silently supply differences from
↳ a different model. Four plotted coefficients
↳ mother-in-law/father-in-law*men/women with source95% intervals and
↳ printed comparisons/baselines/N are retained.

PSEUDOCODE: spouse-specific source matching gates -> event partner and
↳ young-child eligibility -> original complete-window restriction ->
↳ source normalization -> resolved Eq4 specification -> B18 estimates
↳ or numeric-null reasons. No new earnings response for older/childless
↳ partners, no inference that completion proves informal-care
↳ spillovers.

Unresolved target gates: U01, U02, A01, A02, A03, A09, A11.

Retained source targets: figure_b18.

Incorporated specification and actual evidence

Read prereg/data-contract.md D1-D8, replication/matching-specification.md
→ where applicable, schema-evidence.md, target-inventory.md and
→ deviations.md as part of this protocol. Source target order, field
→ definitions, sample restrictions, formulae, intervals, source-only
→ tests and missing-data rules are fixed there. Conflicting decisions
→ remain no-estimate; source-supported partial results do not fill
→ another target's gap. No confidential input or result is present now.

...

INPUT = authenticated source-vintage tables named in registration.tables
TARGETS = this Atom's full source inventory, including unavailable prose
→ claims
for target in TARGETS in source order:
 required = target's exact panel/outcome/sample/model/contrast
 → specification
 blockers = unresolved source choices + missing required
 → vintage/key/year/crosswalk
 if blockers: emit target-status(no-estimate, exact reasons); emit
 → numeric-null rows; continue
 panel = derive required source panel, with prescribed source
 → restrictions in prescribed order
 verify unique source units, explicit observed year set and
 → source-authorized join keys
 if source estimator cannot identify required target: emit scientific
 → boundary with reason; continue
 result = evaluate only required source formula using prescribed
 → inference semantics
 if execution fails: log operational failure and leave pending; do not
 → record a scientific null
 else: emit all keyed estimates/SE/CI/baselines/N/source-reported
 → contrasts, not a favorable subset
write results.csv; target-status.csv; interpretation.md; named figure
→ output interfaces
interpret each outcome and contrast in its own units and uncertainty;
→ preserve mixed results
...

Each figure output `<figure\_id>-effects.csv` contains target_id,panel,series,x,estimate,lower,upper,unit,status,reason,bin_left,bin_right. For
→ histogram source bins, edges are supplied only after source
→ identification. `<figure\_id>-annotations.csv` contains
→ target_id,panel,series,statistic,term,value,unit,status,reason for
→ source-reported DID,SE,difference,F,p,baseline,N. Tables/prose use
→ results.csv key target_id,panel,series,x,statistic,term and D8
→ additional metadata. Numeric fields null when unavailable. Do not
→ create actual CSV files during conversion. See figures.yaml and
→ results-map.md for exact input names and manuscript bindings.

Scientific interpretation is outcome/estimand-specific: signed magnitude
 ↳ and source95% intervals (or source SE-only table), relevant
 ↳ source-reported subgroup contrasts, and source-compatible
 ↳ limitations. Statistical imprecision is retained as evidence, not an
 ↳ unavailable observation. There is no replication-success cutoff,
 ↳ invented power, equivalence band or significance-based execution. The
 ↳ native numeric ratings are required compatibility metadata (neutral3
 ↳ and0.5), not an elicited novelty ranking, empirical feasibility claim
 ↳ or success score. This replica lane has no tournament or reviewer.

Registered source module

What are the source earnings spillovers to adult partners with young children after a first parent-in-law death?

No current result, realized strength, evidence path or empirical replication is recorded. The preregistration protocol owns scientific specification and local states; manuscript/graph views are derived. Source locators, blockers and result interfaces are incorporated above. Scientific module purpose: The spouse outcome and its separately matched population are an independent spillover question; sharing family links does not merge it with adult-child earnings.

Reference implementation

Entry point and supported/unresolved components for atom-000010: [execution index](#). See [engineering review](#). This pointer does not amend the registered protocol.

atom-000011: Work choice responses

Work choice responses

TableB5 in A.D uses source main NN sample. Occupation is the first4 digits of the source main labor occupation (AKM employee code or valid self-employed code when self-employed); invalid999999/999900/000000 codes are missing. Source comparison searches up to8 earlier nonmissing records in person×match, with unspecified sort A10. Its source complete-panel minimum year≥1992 implies event≥1997. Do not introduce a DISCO harmonization or adjacent-year switch definition. Firm-to-firm change is1 when source consecutive annual main-job LBNRs are both nonmissing and differ,0 otherwise after source cleanup. Public sector is source public indicator conditional on employment ATP>0. The public classification uses early IDAS ownership07/08/09 and later IDAN sector codes; see D3. Estimate source Eq3 separately by sex for each outcome: columns occupation men,women; firm men,women; public-if-employed men,women. Preserve coefficients, clustered SE, observation N and mean over pre-event t<0. Table contains no extra group contrasts; do not create them or add job-quality/leisure outcomes. PSEUDOCODE: source main sample and field-vintage gate -> independently resolve occupation order versus firm/sector fields -> derive each authorized indicator without substituted missing definitions -> fit source Eq3

-> output six columns and premeans, or cell-specific reasons. Lack of a precise job-transition response is not an equivalence finding, nor direct evidence that preferences were unchanged.

Unresolved target gates: U01, U02, C01, A01, A02, A03, A10, A11.

Retained source targets: table_b5.

Incorporated specification and actual evidence

Read prereg/data-contract.md D1–D8, replication/matching-specification.md where applicable, schema-evidence.md, target-inventory.md and deviations.md as part of this protocol. Source target order, field definitions, sample restrictions, formulae, intervals, source-only tests and missing-data rules are fixed there. Conflicting decisions remain no-estimate; source-supported partial results do not fill another target's gap. No confidential input or result is present now.

```
INPUT = authenticated source-vintage tables named in registration.tables
TARGETS = this Atom's full source inventory, including unavailable prose
↳ claims
for target in TARGETS in source order:
  required = target's exact panel/outcome/sample/model/contrast
  ↳ specification
  blockers = unresolved source choices + missing required
  ↳ vintage/key/year/crosswalk
  if blockers: emit target-status(no-estimate, exact reasons); emit
  ↳ numeric-null rows; continue
  panel = derive required source panel, with prescribed source restrictions
  ↳ in prescribed order
  verify unique source units, explicit observed year set and
  ↳ source-authorized join keys
  if source estimator cannot identify required target: emit scientific
  ↳ boundary with reason; continue
  result = evaluate only required source formula using prescribed inference
  ↳ semantics
  if execution fails: log operational failure and leave pending; do not
  ↳ record a scientific null
  else: emit all keyed estimates/SE/CI/baselines/N/source-reported
  ↳ contrasts, not a favorable subset
write results.csv; target-status.csv; interpretation.md; named figure output
↳ interfaces
interpret each outcome and contrast in its own units and uncertainty;
↳ preserve mixed results
```

Each figure output <figure_id>-effects.csv contains target_id,panel,series,x,estimate,lower,upper,unit,status,reason,bin_left,bin_right. For histogram source bins, edges are supplied only after source identification. <figure_id>-annotations.csv contains target_id,panel,series,statistic,term,value,unit,status,reason for source-reported DID,SE,difference,F,p,baseline,N. Tables/prose use results.csv key target_id,panel,series,x,statistic,term and D8 additional metadata. Numeric fields null when unavailable. Do not create actual CSV files during conversion. See figures.yaml and results-map.md for exact input names and manuscript bindings.

Scientific interpretation is outcome/estimand-specific: signed magnitude and source95% intervals (or source SE-only table), relevant source-reported subgroup

contrasts, and source-compatible limitations. Statistical imprecision is retained as evidence, not an unavailable observation. There is no replication-success cutoff, invented power, equivalence band or significance-based execution. The native numeric ratings are required compatibility metadata (neutral3 and 0.5), not an elicited novelty ranking, empirical feasibility claim or success score. This replica lane has no tournament or reviewer.

Complete registration fields

```
id: atom-000011
topic: Work choice responses
atom_result: prereg
atom_strength: null
atom_shape: null
evidence_paths: []
concepts: []
literature: []
bonds: []
preregistration:
  version: 1
  mode: canonical
  question: What does the source job-transition analysis show about
    ↪ occupational,
    firm and sector responses to parental death?
  rationale: The three job-choice outcomes form the original
    ↪ preference-related
    module; routine occupational coding is part of this analysis, not another
    Atom.
  assumption: Source correspondence and input coverage must be established
    as specified. Causal interpretation additionally requires the source
    ↪ parallel
    counterfactual trends/no relevant anticipation assumptions; the
    ↪ health/family
    patterns do not identify mediation.
  target_strength: causal
  target_shape: comparison
  ratings:
    importance: 3
    novelty: 3
    feasibility: 3
    data_advantage: 3
    informativeness: 0.5
  tables:
    - BEF
    - FAIN
    - DODSAARS
    - DODSAASG
    - IND
    - IDAP
    - IDAN
    - IDAS
    - UDDA
    - AKM
  activation: 'TRUE'
  verdicts:
    - key: evidence
      region: All retained source targets have actual estimates and their
        ↪ source-required
```

uncertainty/interpretation, with no unresolved target.
 says: Report the complete actual outcome vector. Coverage alone gives no direction, causal conclusion or replication confirmation.

- key: partial
 - region: At least one source target has actual evidence and at least one has a conclusive source/input no-estimate disposition.
 - says: Report available signed estimates and uncertainty, including mixed findings, beside every missing target and its exact limitation. Do not generalize the available subset to the missing scope.
- key: boundary
 - region: No retained source target is estimable after a conclusive
 - ↪ source-definition
 - ↪ or input boundary; this is not an operational failure.
 - says: There is no scientific estimate for this source scope. The paper is neither confirmed nor refuted by inability to estimate.
- key: otherwise
 - region: otherwise
 - says: Retain unresolved or inconclusive scientific disposition with actual evidence if any. Operational failures remain pending and are not done() or a statistical null.

protocol: |
 # Work choice responses

TableB5 in A.D uses source main NN sample. Occupation is the first4 digit
 ↪ s of the source main labor occupation (AKM employee code or valid
 ↪ self-employed code when self-employed); invalid999999/999900/000000
 ↪ codes are missing. Source comparison searches up to8 earlier
 ↪ nonmissing records in person*match, with unspecified sort A10. Its
 ↪ source complete-panel minimum year>=1992 implies event>=1997. Do not
 ↪ introduce a DISCO harmonization or adjacent-year switch definition.
 Firm-to-firm change is1 when source consecutive annual main-job LBNRs are
 ↪ both nonmissing and differ,0 otherwise after source cleanup. Public
 ↪ sector is source public indicator conditional on employment ATP>0.
 ↪ The public classification uses early IDAS ownership07/08/09 and later
 ↪ IDAN sector codes; see D3. Estimate source Eq3 separately by sex for
 ↪ each outcome: columns occupation men,women; firm men,women; public-if
 ↪ -employed men,women. Preserve coefficients, clustered SE, observation
 ↪ N and mean over pre-event t<0. Table contains no extra group
 ↪ contrasts; do not create them or add job-quality/leisure outcomes.
 PSEUDOCODE: source main sample and field-vintage gate -> independently
 ↪ resolve occupation order versus firm/sector fields -> derive each
 ↪ authorized indicator without substituted missing definitions -> fit
 ↪ source Eq3 -> output six columns and premeans, or cell-specific reaso
 ↪ ns. Lack of a precise job-transition response is not an equivalence
 ↪ finding, nor direct evidence that preferences were unchanged.

Unresolved target gates: U01, U02, C01, A01, A02, A03, A10, A11.

Retained source targets: table_b5.

Incorporated specification and actual evidence

Read prereg/data-contract.md D1-D8, replication/matching-specification.md
 ↪ where applicable, schema-evidence.md, target-inventory.md and
 ↪ deviations.md as part of this protocol. Source target order, field
 ↪ definitions, sample restrictions, formulae, intervals, source-only
 ↪ tests and missing-data rules are fixed there. Conflicting decisions
 ↪ remain no-estimate; source-supported partial results do not fill
 ↪ another target's gap. No confidential input or result is present now.

```

...
INPUT = authenticated source-vintage tables named in registration.tables
TARGETS = this Atom's full source inventory, including unavailable prose
↳ claims
for target in TARGETS in source order:
    required = target's exact panel/outcome/sample/model/contrast
    ↳ specification
    blockers = unresolved source choices + missing required
    ↳ vintage/key/year/crosswalk
    if blockers: emit target-status(no-estimate, exact reasons); emit
    ↳ numeric-null rows; continue
    panel = derive required source panel, with prescribed source
    ↳ restrictions in prescribed order
    verify unique source units, explicit observed year set and
    ↳ source-authorized join keys
    if source estimator cannot identify required target: emit scientific
    ↳ boundary with reason; continue
    result = evaluate only required source formula using prescribed
    ↳ inference semantics
    if execution fails: log operational failure and leave pending; do not
    ↳ record a scientific null
    else: emit all keyed estimates/SE/CI/baselines/N/source-reported
    ↳ contrasts, not a favorable subset
write results.csv; target-status.csv; interpretation.md; named figure
↳ output interfaces
interpret each outcome and contrast in its own units and uncertainty;
↳ preserve mixed results
...

Each figure output `<figure_id>-effects.csv` contains target_id,panel,series,
↳ ies,x,estimate,lower,upper,unit,status,reason,bin_left,bin_right. For
↳ histogram source bins, edges are supplied only after source
↳ identification. `<figure_id>-annotations.csv` contains
↳ target_id,panel,series,statistic,term,value,unit,status,reason for
↳ source-reported DID,SE,difference,F,p,baseline,N. Tables/prose use
↳ results.csv key target_id,panel,series,x,statistic,term and D8
↳ additional metadata. Numeric fields null when unavailable. Do not
↳ create actual CSV files during conversion. See figures.yaml and
↳ results-map.md for exact input names and manuscript bindings.

Scientific interpretation is outcome/estimand-specific: signed magnitude
↳ and source95% intervals (or source SE-only table), relevant
↳ source-reported subgroup contrasts, and source-compatible
↳ limitations. Statistical imprecision is retained as evidence, not an
↳ unavailable observation. There is no replication-success cutoff,
↳ invented power, equivalence band or significance-based execution. The
↳ native numeric ratings are required compatibility metadata (neutral3
↳ and0.5), not an elicited novelty ranking, empirical feasibility claim
↳ or success score. This replica lane has no tournament or reviewer.

```

Registered source module

What does the source job-transition analysis show about occupational, firm and sector responses to parental death?

No current result, realized strength, evidence path or empirical replication is recorded. The preregistration protocol owns scientific specification and local states; manuscript/graph views are derived. Source locators, blockers and result interfaces are incorporated above. Scientific module purpose: The three job-choice outcomes

form the original preference-related module; routine occupational coding is part of this analysis, not another Atom.

Reference implementation

Entry point and supported/unresolved components for `atom-000011`: [execution index](#). See [engineering review](#). This pointer does not amend the registered protocol.