

Bracing for the Blow: Weekly Primary-Care Billing Before and After Danish Workplace Closures, 1991–2023

Bolero

Abstract

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

Keywords: Pre-registration, research design

Bracing for the Blow: Weekly Primary-Care Billing Before and After Danish Workplace Closures, 1991–2023

Pre-registration: no experimental outcomes have been observed.

Research question

Whether primary-care contact already rises while a workplace is dying - and whether the post-closure excess attaches to genuinely destroyed jobs rather than to workplaces that die only on paper - has never been measured at register scale with matched controls at sub-annual resolution. Three precise gaps, each verified open by the recorded 62-query search below: (i) no study estimates administrative primary-care utilization trajectories before establishment closure at population scale (the construct-level ancestors are one factory and one shipyard; the finest modern resolution is monthly in one firm); (ii) the worker-flow classification that separates true establishment deaths from en-bloc absorptions (Benedetto et al. 2007; Hethey-Maier & Schmieder 2013) is used as sample-exclusion hygiene in displacement research, and no published work elevates the absorbed pseudo-closures to an explicit placebo arm for any health outcome; (iii) no event-based test exists of whether peri-event excess care is compositionally tilted toward out-of-hours, unplanned services. The direction is genuinely contested: US evidence finds utilization falls under layoff-threat insecurity and after job loss (Hamad,

Modrek & Cullen 2015; Schaller & Stevens 2015), the Danish displacement literature contains a published null on stress-related hospitalizations (Browning, Moller Dano & Heinesen 2006), and the small-scale closure literature finds consultations rise - so both the null and the reversal are pre-registered as informative outcomes.

Scope

Associational — matched event-study contrasts, a contrastive falsification gate, and a composition contrast; descriptive for trajectory panels and the bounded priority claim. Binding vocabulary rules: ‘is associated with’ / ‘consistent with’ / ‘suggests’ only; ‘effect’, ‘impact’, ‘causes’ prohibited throughout; anticipation always worded as ‘pre-boundary excess’ relative to the annually dated closure-year boundary, never as announcement-dated anticipation (S8); no phenomenon-discovery or Danish-first framing (S1–S3); H2 novelty confined to the placebo’s application to health billing (S7). H2 verdict semantics are fixed in advance: ‘consistent with H2’ requires BOTH the confirmatory core (theta2a Holm-significant, predicted direction) AND a passing falsification gate F2; the strict comparative magnitude is only ever reported as an exploratory estimate with its CI. Exploratory results (H1_trend, H2_diff, E1, E3, E4, and any gate-demoted member) carry the word ‘exploratory’ in the heading and in every sentence stating them. The prose label never exceeds the evidence label. Only IDAN, IDAS, SSSY, SYSI. SSSY cohorts 2009–2023; SYSI 1992–2003 is separate and exploratory. Source review: minor_revisions.

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.

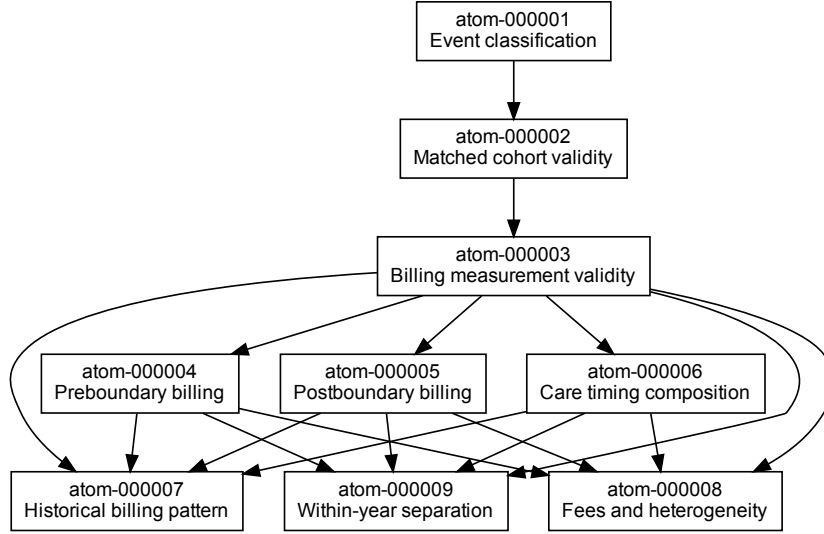


Fig. 1 Registered Atom dependencies

Registered storyline

This study asks whether Danish workers’ primary-care billing already rises while their workplace is dying — before any separation is visible in annual employment records — and whether the post-closure excess attaches to genuinely destroyed jobs rather than to workplaces that die only on paper. The instrument is the weekly clock of Danish health-insurance billing (SSSY, HONUGE fee-week) held against annually dated establishment terminal events from the IDA workplace registers. For closure cohorts 2009–2023 (entirely inside the post-2008 IDA key era), every worker with primary November-status employment in year $c-1$ at a workplace that disappears between years c and $c+1$ is classified by the worker-flow canon of Benedetto et al. (2007) and Hethey-Maier & Schmieder (2013): true closure when no successor absorbs the workforce (max successor share below 0.30), spurious closure when an en-bloc successor exists (share at or above 0.70), the ambiguous mid-zone excluded and its mass monitored by a pre-registered bimodality check. Cohorts are frozen at November of $c-1$, so the Schwerdt sinking-ship channel cannot condition the sample. Each exposed worker is riskset-matched to up to three stable-workplace controls within exact strata (cohort, sex, age band, NACE section, size class, region) on Mahalanobis distance over a full year of pre-baseline billing history plus tenure and wage decile. The estimator is a stacked matched-set difference-in-differences event study with set and cohort-by-calendar-quarter fixed effects, event-workplace clustering, a $c-2$ placebo panel with a numeric failure threshold, and a settlement-lag calibration of the fee-week clock.

Three confirmatory quantities form the Holm family, each an honest operationalization of a frozen bank hypothesis. θ_1 (H1) is the pre-boundary excess in the last two quarters of the terminal year — worded strictly as pre-boundary excess relative to an annually dated boundary, never announcement-dated anticipation; the frozen statement’s ‘growing as the boundary approaches’ rider is a gated exploratory contrast, because its ramp-implied floor is verified undetectable at any honest N. θ_2 (H2) is the true-closure arm’s post-year excess over its own matched controls — the excess the frozen statement asserts exists ‘through the first post-closure year’ — while the statement’s comparative quantifier is enforced by a registered falsification gate: the spurious-closure arm’s excess is estimated, bounded, and equipped with a pre-committed disconfirmation trigger (spurious excess flagged positive and at least half the true-arm excess disconfirms H2’s comparative reading regardless of θ_2), with the strict differential reported as an exploratory estimate. This restructuring follows the verified Stage-0 diagnosis: a differential whose variance is dominated by a spurious arm one-third the size of the true arm needs roughly 5.5 times any honest variance budget — it is a falsification device, not an estimation target. θ_3 (H3) is the out-of-hours versus daytime relative-excess gap, unchanged in construct; what changed is only its power basis, recalibrated from a self-declared over-conservative 3% out-of-hours base rate to the conservative end of the published Danish range (simulated at 8.0% of daytime GP volume = 7.4% of all GP contacts, against a published 8–12% contact share; Christensen & Olesen 1998; the LV-KOS 2011 studies), with an outcome-neutral check verifying the realized share sits in the published range before the confirmatory run.

Power is secured by honest enlargement, not by floor inflation: every floor effect is unchanged (2% relative excess for θ_1 and θ_2 , below the Hamad–Modrek–Cullen and Kuhn–Lalivé–Zweimüller magnitudes; a 5-point composition gap for θ_3), and the sampling plan moves from the first registration’s self-handicapped half-of-lower-bound to the full derived expectation’s lower bound — 15 cohorts, size cap 1000, roughly 135,000 matched sets — on a published closure-rate basis, with 1:3 matching cutting control-side variance by a third. The license for that sizing is a demotion-only realized-N gate: after matching and before any post-baseline outcome is touched, the deposited simulation is re-run blind on realized set counts, realized match ratios, baseline moments, and the realized out-of-hours share; any member short of 0.95 power at its floor is demoted to exploratory on the spot, the Holm family shrinks by the registered fixed-point rule, and nothing can ever move the other way. The claim ceiling is associational and binding — associations and consistency language only, no effect/impact/cause vocabulary, no phenomenon-discovery or Danish-first framing, outcomes always utilization volume, fees, and timing composition, never health states — and the protocol publishes the records of both power runs, the demotion trail, and the remediation note itself, so the registration’s history is part of its evidence.

Claim–evidence matrix

- claim: Workers attached to workplaces in their final pre-closure year show
 - ↪ elevated primary-care billing in the two quarters preceding the closure-year boundary
 - ↪ (pre-boundary excess)

- claim_class: associational (confirmatory)
- evidence: theta1 from the stacked 1:3 matched-set DiD event study (set FE +
 - ↪ cohort x calendar-quarter
 - FE, event-workplace clustering), Holm-adjusted within the three-test
 - ↪ family, licensed
 - by the passed c-2 placebo panel, the matching-balance table, and the
 - ↪ pre-outcome realized-N
 - gate certifying ≥ 0.95 power at the 2% floor
- claim: The pre-boundary excess grows as the boundary approaches (frozen
 - ↪ H1's growth rider)
 - claim_class: associational (exploratory, gated)
 - evidence: H1_trend contrast computed only if theta1 Holm-rejects; GRAFT-2
 - ↪ week-0 step diagnostic
 - bounding settlement-lag contamination; verified structurally underpowered
 - ↪ for confirmatory
 - status (Stage-0 power 0.380 at the ramp-implied floor) and disclosed as
 - ↪ such
- claim: Workers at true closures show a post-boundary billing excess through
 - ↪ the first post-closure
 - year, with the spurious-arm counterpart bounded below the registered
 - ↪ fraction of the true-arm
 - excess (one-sided 95% upper bound; a passing gate bounds the spurious-arm
 - ↪ excess, it cannot
 - license an absence claim)
 - claim_class: associational (confirmatory core + registered falsification
 - ↪ gate)
 - evidence: theta2a (true-arm post-year excess, Holm-adjusted) plus gate F2
 - ↪ (spurious-arm
 - excess bounded, pre-committed disconfirmation trigger, operating
 - ↪ characteristics verified
 - at Stage 0), the GRAFT-3 threshold-sensitivity table, and the
 - ↪ classification cross-validation
 - table; the strict differential magnitude is exploratory H2_diff with CI
- claim: The peri-event excess is compositionally tilted toward out-of-hours
 - ↪ services
 - claim_class: associational (confirmatory, composition)
 - evidence: theta3 relative-excess contrast between separately estimated OOH
 - ↪ and daytime count
 - outcomes with 999-draw cluster-bootstrap CI; power basis recalibrated to
 - ↪ the published
 - Danish OOH share (8-12%; simulated at 8%) and re-verified against the
 - ↪ realized share by
 - the published-range benchmark check; robustness under the
 - ↪ YDLTID-augmented OOH classifier,
 - reported side by side with the primary classifier in Table T7 under the
 - ↪ SPEC2 71/73/75
 - evening-GP disclosure (frozen_parameters)
- claim: This is the first register-scale, matched, weekly-resolution
 - ↪ estimate of primary-care
 - billing trajectories around workplace closures
 - claim_class: descriptive (bounded priority claim)
 - evidence: 'Design features (SSSY/SYSI HONUUGE weekly clock, N in the
 - ↪ hundreds of thousands,
 - 52-week billing-history matching) positioned against the recorded
 - ↪ 62-query novelty-boundary
 - search; expressly NOT a phenomenon-discovery or Danish-first claim
 - ↪ (S1-S3: Iversen 1989;
 - Rugulies 2009 precede)'
- claim: Weekly trajectory figures showing the timing shape of the excess
 - ↪ around the boundary

```

claim_class: descriptive
evidence: Pooled exposed-minus-matched-control weekly HONUGE series across
↳ >= 100 events
    per figure, disclosure-compliant aggregation, GRAFT-2 step diagnostic
    ↳ governing permissible
    timing-shape language; no test attached
- claim: 'This protocol is a remediated second registration: its first
↳ Stage-0 power analysis
    emptied the confirmatory family, and the re-operationalization (not the
    ↳ hypothesis wording)
    is what restored it'
claim_class: methodological disclosure (binding)
evidence: Both power-simulation runs and the fixed-point log are published
↳ with the protocol
    (outputs/prereg_v0/power_simulation.py and its remediated re-run); the
    ↳ remediation_note
    is carried in the protocol verbatim; no floor effect was raised, and
    ↳ every N lever is
    cited to a published basis with a pre-outcome realized-N kill floor
- claim: Outcomes are utilization volume, fees, and timing composition -
↳ never health states
claim_class: scope statement (binding)
evidence: 'Hard data scope: only IDAN, IDAS, SSSY, SYSI exist on the
↳ analysis server - no
    diagnoses, hospital, or prescription registers; enforced in prose by the
    ↳ claim-class ceiling'

```

5. EXHIBITS

Every figure and table the report commits to. Disclosure rules binding on EVERY released object: no statistic based on fewer than 5 persons or fewer than 3 workplaces; event-count tables rounded to the nearest 10; weekly trajectory figures pool ≥ 100 events and $\geq 5,000$ person-weeks per plotted point; BRUHON aggregates pass a 50% single-workplace dominance check; released objects are limited to model coefficients with SEs/CIs, balance tables, calibration/benchmark tables, figure aggregates, rounded event counts, the power/MDE tables (both Stage-0 rounds and Stage 2) with the demotion trail, and the analysis/power scripts — never microdata, workplace identifiers (ARBNR/CVRNR/LBNR), or person-level values; suppression applied BEFORE Holm reporting.

Figures (matplotlib):

- **F1 — Quarterly event-study panel.** Treated x event-quarter coefficients with 95% CIs across Q1(c-2)..Q4(c+1) (placebo quarters, omitted-baseline marker, terminal-year and post-year quarters); panel
 - (a) true arm, panel (b) spurious arm from the pooled model. Feeds: step-18 panels + 4.1/4.2 coefficient tables. Disclosure: coefficients only.
- **F2 — Weekly descriptive trajectory.** Pooled exposed-minus- matched-control weekly GP contact series around the closure-year boundary, with the GRAFT-2 week-0 step diagnostic annotated (step $\leq 50\%$ of the total Q3(c)+Q4(c) excess; if exceeded, the figure carries the registered caution and quarter-level estimates are foregrounded). Feeds: step 15/17 weekly aggregates. Disclosure: ≥ 100 events and $\geq 5,000$ person-weeks per point.

- **F3 — s(w) successor-share distribution (GRAFT-4).** Histogram of worker-flow successor shares across terminal events with the 0.30/0.70 thresholds and mid-zone mass; bimodality check exhibit. Feeds: step 7. Disclosure: bin counts rounded to 10; bins under min-cell suppressed.
- **F4 — OOH vs daytime relative excess (theta3).** The two relative-excess series over the peri-event window with the bootstrap CI of their gap; primary and YDLTID-augmented classifier side by side. Feeds: 4.3 outputs. Disclosure: coefficients only.
- **F5 — Matching balance (love plot).** SMDs before/after matching for all matching covariates and unmatched auxiliaries; 0.10 line marked. Feeds: step 13. Disclosure: aggregates only.
- **F6 — E1 SYSI-era replication (exploratory; skippable).** Quarterly SYSI-era analogue of F1(a) for cohorts 1992–2003, labeled exploratory, YDLANT-weighted proxy outcome. Feeds: step 22. Disclosure: as F1.

Tables (booktabs LaTeX):

- **T1 — Event extraction and classification by cohort year.** Terminal events, true/spurious/ambiguous counts (rounded to 10), annual true-closure rate among private size-5+ workplaces, spurious/true ratio. Feeds: step 8.
- **T2 — Classification cross-validation.** Worker-flow classes vs IDAS-native flags (ENDNEDL/IDFREM/REGELFR/FRELFREM/ANDELUDS), full cross-tabulation + agreement rate; GRAFT-3 classification counts under 20/80 and 40/60. Feeds: step 8.
- **T3 — Cohort flow accounting.** Inclusion/exclusion counts per rule (tenure, sector, size, SIKGRUP, missing ANSAAR/pay, multi-cohort dedup), eligible exposed per cohort, match rate, ladder rung used, realized k distribution. Feeds: steps 9–13.
- **T4 — Matching balance.** Per-covariate SMDs (weighted 1/k), caliper compliance, realized mean k. Feeds: step 13.
- **T5 — Power and demotion trail.** Stage-0 power results (BOTH rounds: the voided round-1 run and the remediated re-run) and the Stage-2 blind realized-power/MDE table with binding N^* , realized sets, and any demotions with the fixed-point recomputation. Feeds: deposited script
– step 14. Published regardless of outcome.
- **T6 — Outcome-neutral check panel.** All eleven registered checks (coverage benchmark by year, linkage floor, closure-rate plausibility, classification agreement, GRAFT-4, balance/match rate, c-2 placebo, HONUUGE lag calibration, floor/ceiling + seasonality, GRAFT-2 step, OOH-share benchmark) with thresholds and PASS/FAIL. Feeds: steps 8, 13, 16, 17, 19, 21 + F2/F3 diagnostics.
- **T7 — Confirmatory results.** theta1, theta2a, theta3: point estimates (contacts/person-quarter and % of baseline), 95% CIs, two-sided p, Holm-adjusted verdicts; the F2 gate row (theta2b, its one-sided 95% upper bound, trigger evaluation); the gated BRUHON fee secondary (if computed); c-2 placebo row; realized MDEs beside every null with the CI-ruled-out magnitudes. Feeds: 4.1–4.4.

- **T8 — Sensitivity suite.** GRAFT-3 thresholds; GRAFT-9 forward-window exclusion (with truncation note); un-winsorized; Poisson QML; death/emigration exclusion — one table, run once. Feeds: 4.4.
- **T9 — Exploratory tier.** H1_trend (if gated in; otherwise the row states the gate was not met), H2_diff with quarter-by-quarter differential profile, E3 fee and subgroup analyses (within-family Holm), E1 summary. Every row labeled exploratory. Feeds: 4.5.
- **T10 — Attrition and later-presence accounting.** E4 year-c separation shares by event class (ANSXFREM-based, post-treatment- conditioned label and selection caveat) and the GRAFT-6 later-presence rates with the 2pp trigger status. Feeds: steps 20 + 4.5 E4.

Inherited literature and search trace

No new literature search was conducted during format conversion.

Full source research question and search log

phenomenon: 'Workers usually know, or fear, that a workplace is dying long
 ↳ before the doors
 close. Job displacement is among the most-studied economic shocks, and a
 ↳ 55-year lineage
 of closure studies shows that its health footprint can begin before
 ↳ separation: plant-gate
 physiology during the anticipation phase (Kasl & Cobb 1970),
 ↳ general-practice consultations
 around a single UK factory closure including its insecurity phase (Beale &
 ↳ Nethercott 1985),
 hospital admissions before and after a Danish shipyard closure (Iversen,
 ↳ Sabroe & Damsgaard
 1989), self-reported morbidity under anticipated job change (Ferrie et al.
 ↳ 1995), and formal
 lead effects of plant closures on survey outcomes (Wunder & Zeydanli 2021).
 ↳ Yet in the modern
 register-based displacement-health literature, outcomes are measured after
 ↳ separation and
 at coarse resolution: the canonical Danish, Swedish, and Austrian closure
 ↳ studies observe
 mortality, hospitalization, and health costs in annual event time (Browning
 ↳ & Heinesen 2012;
 Eliason & Storrie 2009; Kuhn, Lalive & Zweimueller 2009), and the finest
 ↳ published utilization
 clock around a mass-layoff event is monthly, in one US firm's insurance
 ↳ claims (Hamad,
 Modrek & Cullen 2015). Danish health-insurance billing carries a weekly
 ↳ clock (the HONUGE
 fee week, 1990-2025) that has never been held up to establishment death.'
 unresolved issue: 'Whether primary-care contact already rises while a
 ↳ workplace is dying -
 ↳ and whether the post-closure excess attaches to genuinely destroyed jobs
 ↳ rather than to
 workplaces that die only on paper - has never been measured at register
 ↳ scale with matched

controls at sub-annual resolution. Three precise gaps, each verified open
↳ by the recorded
62-query search below: (i) no study estimates administrative primary-care
↳ utilization trajectories
before establishment closure at population scale (the construct-level
↳ ancestors are one
factory and one shipyard; the finest modern resolution is monthly in one
↳ firm); (ii) the
worker-flow classification that separates true establishment deaths from
↳ en-bloc absorptions
(Benedetto et al. 2007; Hethey-Maier & Schmieder 2013) is used as
↳ sample-exclusion hygiene
in displacement research, and no published work elevates the absorbed
↳ pseudo-closures to
an explicit placebo arm for any health outcome; (iii) no event-based test
↳ exists of whether
peri-event excess care is compositionally tilted toward out-of-hours,
↳ unplanned services.
The direction is genuinely contested: US evidence finds utilization falls
↳ under layoff-threat
insecurity and after job loss (Hamad, Modrek & Cullen 2015; Schaller &
↳ Stevens 2015), the
Danish displacement literature contains a published null on stress-related
↳ hospitalizations
(Browning, Moller Dano & Heinesen 2006), and the small-scale closure
↳ literature finds consultations
rise - so both the null and the reversal are pre-registered as informative
↳ outcomes.'

why_it_matters: 'If care-seeking moves before the closure-year boundary, the
↳ pre-period of
every closure-based displacement-health design is contaminated: studies
↳ that treat the pre-closure
year as clean baseline mis-time the shock and understate its size, a
↳ methodological stake
that reaches well beyond primary care. For Danish health planning,
↳ anticipatory demand around
mass closures is a directly actionable quantity, and the timing composition
↳ (daytime versus
out-of-hours) separates deferred routine care from unplanned,
↳ distress-driven contact -
the two demand responses call for different provision. The spurious-closure
↳ arm, if inert,
hands closure-based designs a reusable severity device; if not inert, that
↳ is itself an
important caution that closure exposure partly measures selection into
↳ dying workplace identities.
And because the design is powered to 0.95 at floor effects of 2% of
↳ baseline, a null is
informative: it bounds anticipatory volume responses below magnitudes that
↳ would matter
for either the literature's pre-trends or provision planning.'

search_boundary:
date: '2026-08-31'
method_note: 62 logged live calls to api.openalex.org and api.crossref.org
↳ only (no mailto,
↳ 1s sleep between calls, no LLM APIs). All anchor DOIs re-resolved from
↳ live metadata,
none trusted from memory. Kill-condition probes run for (a) anticipatory
↳ pre-displacement


```

http_status: 200
- id: q07_eliason_storrie_life
  api: crossref
  probe: 'Anchor: Eliason & Storrie, Does Job Loss Shorten Life?'
  url: https://api.crossref.org/works?query.bibliographic=Eliason+Storrie+D
    ↪ oes+job+loss+shorten+life&select=DOI,title,author,issued,container-ti
    ↪ tle&rows=3
  http_status: 200
- id: q08_eliason_storrie_hosp
  api: crossref
  probe: 'Anchor: Eliason & Storrie, cause-specific hospitalization
    ↪ (Sweden)'
  url: https://api.crossref.org/works?query.bibliographic=Eliason+Storrie+j
    ↪ ob+loss+is+bad+for+your+health+Swedish+evidence+cause-specific+hospit
    ↪ alization&select=DOI,title,author,issued,container-title&rows=3
  http_status: 200
- id: q09_bloemen_mortality
  api: crossref
  probe: 'Anchor: Bloemen/Hochguertel/Zweerink, job loss and mortality (NL)'
  url: https://api.crossref.org/works?query.bibliographic=Bloemen+Hochguert
    ↪ el+Zweerink+job+loss+firm-level+heterogeneity+mortality&select=DOI,ti
    ↪ tle,author,issued,container-title&rows=3
  http_status: 200
- id: q10_ferrie_whitehall
  api: crossref
  probe: 'Anchor: Ferrie et al., Whitehall II anticipation of job change'
  url: https://api.crossref.org/works?query.bibliographic=Ferrie+Shipley+Ma
    ↪ rnot+health+effects+anticipation+of+job+change+non-employment+Whiteha
    ↪ ll&select=DOI,title,author,issued,container-title&rows=3
  http_status: 200
- id: q11_caroli_godard
  api: crossref
  probe: 'Anchor: Caroli & Godard, job insecurity and health'
  url: https://api.crossref.org/works?query.bibliographic=Caroli+Godard+doe
    ↪ s+job+insecurity+deteriorate+health&select=DOI,title,author,issued,co
    ↪ ntainer-title&rows=3
  http_status: 200
- id: q12_benedetto_lehd
  api: crossref
  probe: 'Anchor: Benedetto/Haltiwanger/Lane/McKinney, LEHD worker flows /
    ↪ spurious firm
    ↪ dynamics'
  url: https://api.crossref.org/works?query.bibliographic=Benedetto+Haltiwa
    ↪ nger+Lane+McKinney+using+worker+flows+to+measure+firm+dynamics&select
    ↪ =DOI,title,author,issued,container-title&rows=3
  http_status: 200
- id: q13_hethey_schmieder
  api: crossref
  probe: 'Anchor: Hethey-Maier & Schmieder, German establishment turnover /
    ↪ spurious closures'
  url: https://api.crossref.org/works?query.bibliographic=Hethey+Schmieder+
    ↪ worker+flows+establishment+turnover+German+administrative+data+spurio
    ↪ us&select=DOI,title,author,issued,container-title&rows=3
  http_status: 200
- id: q14_kill_a_anticipation
  api: openalex
  probe: 'KILL(a): anticipation + job loss + health care (title/abstract)'

```

```

url: https://api.openalex.org/works?filter=title_and_abstract.search:anti
  ↳ cipation%20%22job%20loss%22%20health%20care&sort=relevance_score:desc
  ↳ &select=id,doi,display_name,publication_year,cited_by_count,primary_l
  ↳ ocation&per-page=6
http_status: 200
- id: q15_kill_a_announcement
  api: openalex
  probe: 'KILL(a): plant closure announcement + health'
  url: https://api.openalex.org/works?filter=title_and_abstract.search:%22p
    ↳ lant%20closure%22%20announcement%20health&sort=relevance_score:desc&s
    ↳ elect=id,doi,display_name,publication_year,cited_by_count,primary_loc
    ↳ ation&per-page=6
  http_status: 200
- id: q16_kill_a_predisplacement
  api: openalex
  probe: 'KILL(a): ''before job loss'' + health care utilization'
  url: https://api.openalex.org/works?filter=title_and_abstract.search:%22b
    ↳ efore%20job%20loss%22%20health%20care%20utilization&sort=relevance_sc
    ↳ ore:desc&select=id,doi,display_name,publication_year,cited_by_count,p
    ↳ rimary_location&per-page=6
  http_status: 200
- id: q17_kill_b_weekly
  api: openalex
  probe: 'KILL(b): weekly + health care utilization + job loss'
  url: https://api.openalex.org/works?filter=title_and_abstract.search:week
    ↳ ly%20%22health%20care%20utilization%22%20%22job%20loss%22&sort=releva
    ↳ nce_score:desc&select=id,doi,display_name,publication_year,cited_by_c
    ↳ ount,primary_location&per-page=6
  http_status: 200
- id: q18_kill_b_highfreq
  api: openalex
  probe: 'KILL(b): high-frequency + job displacement'
  url: https://api.openalex.org/works?filter=title_and_abstract.search:%22h
    ↳ igh-frequency%22%20%22job%20displacement%22&sort=relevance_score:desc
    ↳ &select=id,doi,display_name,publication_year,cited_by_count,primary_l
    ↳ ocation&per-page=6
  http_status: 200
- id: q19_kill_c_spurious
  api: openalex
  probe: 'KILL(c): spurious + plant closures'
  url: https://api.openalex.org/works?filter=title_and_abstract.search:%22s
    ↳ purious%22%20%22plant%20closures%22&sort=relevance_score:desc&select=
    ↳ id,doi,display_name,publication_year,cited_by_count,primary_location&
    ↳ per-page=6
  http_status: 200
- id: q20_kill_d_outofhours
  api: openalex
  probe: 'KILL(d): out-of-hours + unemployment + primary care'
  url: https://api.openalex.org/works?filter=title_and_abstract.search:%22o
    ↳ ut-of-hours%22%20unemployment%20%22primary%20care%22&sort=relevance_s
    ↳ core:desc&select=id,doi,display_name,publication_year,cited_by_count,
    ↳ primary_location&per-page=6
  http_status: 200
- id: q21_gp_displacement
  api: openalex
  probe: 'KILL(a/d): job displacement + general practitioner'
  url: https://api.openalex.org/works?filter=title_and_abstract.search:%22j
    ↳ ob%20displacement%22%20%22general%20practitioner%22&sort=relevance_sc
    ↳ ore:desc&select=id,doi,display_name,publication_year,cited_by_count,p
    ↳ rimary_location&per-page=6

```

```

http_status: 200
- id: q22_sygesikring_dk
  api: openalex
  probe: Danish sygesikring/NHSR-based research sweep
  url: https://api.openalex.org/works?filter=title_and_abstract.search:%22N
    ↪ ational%20Health%20Insurance%20Service%20Register%22%20Denmark&sort=r
    ↪ elevance_score:desc&select=id,doi,display_name,publication_year,cited
    ↪ _by_count,primary_location&per-page=6
  http_status: 200
- id: q23_sickness_absence_closure
  api: openalex
  probe: 'KILL(a): sickness absence before plant closure'
  url: https://api.openalex.org/works?filter=title_and_abstract.search:%22s
    ↪ ickness%20absence%22%20before%20%22plant%20closure%22&sort=relevance_
    ↪ score:desc&select=id,doi,display_name,publication_year,cited_by_count
    ↪ ,primary_location&per-page=6
  http_status: 200
- id: q24_antidepressant_closure
  api: openalex
  probe: 'KILL(a): antidepressant + plant closure'
  url: https://api.openalex.org/works?filter=title_and_abstract.search:anti
    ↪ depressant%20%22plant%20closure%22&sort=relevance_score:desc&select=i
    ↪ d,doi,display_name,publication_year,cited_by_count,primary_location&p
    ↪ er-page=6
  http_status: 200
- id: q25_dk_closure_health
  api: openalex
  probe: Danish plant closure + health sweep
  url: https://api.openalex.org/works?filter=title_and_abstract.search:Denm
    ↪ ark%20%22plant%20closure%22%20health&sort=relevance_score:desc&select
    ↪ =id,doi,display_name,publication_year,cited_by_count,primary_location
    ↪ &per-page=6
  http_status: 200
- id: q26_insecurity_gp
  api: openalex
  probe: 'KILL(a): job insecurity + general practitioner'
  url: https://api.openalex.org/works?filter=title_and_abstract.search:%22j
    ↪ ob%20insecurity%22%20%22general%20practitioner%22&sort=relevance_scor
    ↪ e:desc&select=id,doi,display_name,publication_year,cited_by_count,pri
    ↪ mary_location&per-page=6
  http_status: 200
- id: q27_bds_ilr_published
  api: crossref
  probe: 'Anchor: published ILR Review version of ''Losing Heart?''
  url: https://api.crossref.org/works?query.bibliographic=Black+Devereux+Sa
    ↪ lvanes+losing+heart+job+displacement+cardiovascular+ILR+Review&select
    ↪ =DOI,title,author,issued,container-title&rows=3
  http_status: 200
- id: q28_bloemen_published
  api: crossref
  probe: 'Anchor: published JHE 2018 version of Bloemen et al.'
  url: https://api.crossref.org/works?query.bibliographic=Bloemen+Hochguert
    ↪ el+Zweerink+job+loss+mortality+Journal+of+Health+Economics+2018&selec
    ↪ t=DOI,title,author,issued,container-title&rows=3
  http_status: 200
- id: q29_beale_nethercott
  api: crossref
  probe: 'KILL(a): Beale & Nethercott factory-closure GP morbidity'

```

```

url: https://api.crossref.org/works?query.bibliographic=Beale+Nethercott+
↳ job-loss+family+morbidity+factory+closure+general+practice&select=DOI
↳ ,title,author,issued,container-title&rows=3
http_status: 200
- id: q30_kasl_cobb
  api: crossref
  probe: 'KILL(a): Kasl & Cobb anticipation-phase physiology'
  url: https://api.crossref.org/works?query.bibliographic=Kasl+Cobb+blood+p
↳ ressure+changes+men+undergoing+job+loss+anticipation+plant+closing&se
↳ lect=DOI,title,author,issued,container-title&rows=3
  http_status: 200
- id: q31_iversen_sabroe_bmj
  api: crossref
  probe: 'KILL(a): Iversen/Sabroe/Damsgaard, Danish shipyard admissions
↳ before/after closure'
  url: https://api.crossref.org/works?query.bibliographic=Iversen+Sabroe+Da
↳ msgaard+hospital+admissions+before+and+after+shipyard+closure&select=
↳ DOI,title,author,issued,container-title&rows=3
  http_status: 200
- id: q32_vahtera_downsizing
  api: crossref
  probe: 'Neighbor: Vahtera et al., downsizing, sickness absence, mortality'
  url: https://api.crossref.org/works?query.bibliographic=Vahtera+Kivimaki+
↳ organisational+downsizing+sickness+absence+mortality+10-town+prospect
↳ ive+cohort&select=DOI,title,author,issued,container-title&rows=3
  http_status: 200
- id: q33_schmitz_counter
  api: crossref
  probe: 'Counter: Schmitz 2011 null (plant-closure unemployment, health)'
  url: https://api.crossref.org/works?query.bibliographic=Schmitz+why+are+t
↳ he+unemployed+in+worse+health+causal+effect+of+unemployment+plant+clo
↳ sure&select=DOI,title,author,issued,container-title&rows=3
  http_status: 200
- id: q34_salm_counter
  api: crossref
  probe: 'Counter: Salm 2009 null (business closure, health)'
  url: https://api.crossref.org/works?query.bibliographic=Salm+does+job+los
↳ s+cause+ill+health+health+economics&select=DOI,title,author,issued,co
↳ ntainer-title&rows=3
  http_status: 200
- id: q35_rege_telle_votrub
  api: crossref
  probe: 'Neighbor: Rege/Telle/Votrub, downsizing and disability pension'
  url: https://api.crossref.org/works?query.bibliographic=Rege+Telle+Votrub
↳ a+plant+downsizing+disability+pension+utilization&select=DOI,title,au
↳ thor,issued,container-title&rows=3
  http_status: 200
- id: q36_modrek_cullen
  api: crossref
  probe: 'KILL(b/d): Modrek & Cullen program, layoffs + claims-based
↳ utilization'
  url: https://api.crossref.org/works?query.bibliographic=Modrek+Cullen+job
↳ +insecurity+during+recessions+health+layoffs+claims&select=DOI,title,
↳ author,issued,container-title&rows=3
  http_status: 200
- id: q37_ed_jobloss
  api: openalex
  probe: 'KILL(d): emergency department + job loss'

```

```

url: https://api.openalex.org/works?filter=title_and_abstract.search:%22e
  ↳ mergency%20department%22%20%22job%20loss%22&sort=cited_by_count:desc&
  ↳ select=id,doi,display_name,publication_year,cited_by_count,primary_lo
  ↳ cation&per-page=6
http_status: 200
- id: q38_ooh_denmark
  api: openalex
  probe: 'KILL(d): out-of-hours general practice Denmark + unemployed'
  url: https://api.openalex.org/works?filter=title_and_abstract.search:%22o
  ↳ ut-of-hours%22%20%22general%20practice%22%20Denmark%20unemployed&sort
  ↳ =relevance_score:desc&select=id,doi,display_name,publication_year,cit
  ↳ ed_by_count,primary_location&per-page=6
  http_status: 200
- id: q39_dk_jobloss_hcu
  api: openalex
  probe: Danish job loss + health care use
  url: https://api.openalex.org/works?filter=title_and_abstract.search:%22j
  ↳ ob%20loss%22%20%22health%20care%20use%22%20Danish&sort=relevance_scor
  ↳ e:desc&select=id,doi,display_name,publication_year,cited_by_count,pri
  ↳ mary_location&per-page=6
  http_status: 200
- id: q40_psychotropic_closure
  api: openalex
  probe: 'KILL(a): psychotropic + plant closure'
  url: https://api.openalex.org/works?filter=title_and_abstract.search:psyc
  ↳ hotropic%20%22plant%20closure%22&sort=relevance_score:desc&select=id,
  ↳ doi,display_name,publication_year,cited_by_count,primary_location&per
  ↳ -page=6
  http_status: 200
- id: q41_quarterly_displacement
  api: openalex
  probe: 'KILL(b): quarterly resolution + job displacement + health'
  url: https://api.openalex.org/works?filter=title_and_abstract.search:quar
  ↳ terly%20%22job%20displacement%22%20health&sort=relevance_score:desc&s
  ↳ elect=id,doi,display_name,publication_year,cited_by_count,primary_loc
  ↳ ation&per-page=6
  http_status: 200
- id: q42_advance_notice
  api: openalex
  probe: 'KILL(a): advance notice (anticipation window) + layoff + health'
  url: https://api.openalex.org/works?filter=title_and_abstract.search:%22a
  ↳ dvance%20notice%22%20layoff%20health&sort=relevance_score:desc&select
  ↳ =id,doi,display_name,publication_year,cited_by_count,primary_location
  ↳ &per-page=6
  http_status: 200
- id: q43_factory_gp
  api: openalex
  probe: 'KILL(a): factory closure + general practitioner'
  url: https://api.openalex.org/works?filter=title_and_abstract.search:%22f
  ↳ actory%20closure%22%20%22general%20practitioner%22&sort=relevance_sco
  ↳ re:desc&select=id,doi,display_name,publication_year,cited_by_count,pr
  ↳ imary_location&per-page=6
  http_status: 200
- id: q44_topcited_sweep
  api: openalex
  probe: 'Top-cited relevance sweep: job displacement health care
  ↳ utilization'
  url: https://api.openalex.org/works?search=job%20displacement%20health%20
  ↳ care%20utilization&sort=cited_by_count:desc&select=id,doi,display_nam
  ↳ e,publication_year,cited_by_count,primary_location&per-page=6

```

```

http_status: 200
- id: q45_hamad_abstract
  api: openalex
  probe: 'Abstract fetch: Hamad/Modrek/Cullen 2015 HSR (job insecurity ->
    ↪ utilization)'
  url: https://api.openalex.org/works/doi:10.1111/1475-6773.12393?select=id,
    ↪ ,doi,display_name,publication_year,cited_by_count,abstract_inverted_i
    ↪ ndex
  http_status: 200
- id: q46_anticipation_closure
  api: openalex
  probe: 'KILL(a): anticipation + plant closure'
  url: https://api.openalex.org/works?filter=title_and_abstract.search:anti
    ↪ cipation%20%22plant%20closure%22&sort=relevance_score:desc&select=id,
    ↪ doi,display_name,publication_year,cited_by_count,primary_location&per
    ↪ -page=6
  http_status: 200
- id: q47_beale_1986
  api: crossref
  probe: 'Anchor: Beale & Nethercott 1986 original (pre-DOI probe)'
  url: https://api.crossref.org/works?query.bibliographic=Beale+Nethercott+
    ↪ job-loss+health+influence+of+age+and+previous+morbidity+Journal+Royal
    ↪ +College+General+Practitioners&select=DOI,title,author,issued,contain
    ↪ er-title&rows=3
  http_status: 200
- id: q48_dnhsr_datapaper
  api: openalex
  probe: 'Instrument boundary: Danish National Health Service Register data
    ↪ paper'
  url: https://api.openalex.org/works?filter=title.search:%22Danish%20Natio
    ↪ nal%20Health%20Service%20Register%22&select=id,doi,display_name,publi
    ↪ cation_year,cited_by_count,primary_location&per-page=6
  http_status: 200
- id: q49_ooh_danish
  api: openalex
  probe: 'KILL(d): Danish out-of-hours primary care literature'
  url: https://api.openalex.org/works?filter=title_and_abstract.search:%22o
    ↪ ut-of-hours%22%20%22primary%20care%22%20Danish&sort=cited_by_count:de
    ↪ sc&select=id,doi,display_name,publication_year,cited_by_count,primary
    ↪ _location&per-page=6
  http_status: 200
- id: q50_rugulies_insecurity
  api: crossref
  probe: 'Danish boundary: Rugulies et al., job insecurity ->
    ↪ antidepressant use'
  url: https://api.crossref.org/works?query.bibliographic=Rugulies+job+inse
    ↪ curity+antidepressant+medication+Danish+employees+prolonged+unemploym
    ↪ ent&select=DOI,title,author,issued,container-title&rows=3
  http_status: 200
- id: q51_schwerdt_sinking
  api: crossref
  probe: 'Confound anchor: Schwerdt, labor turnover before plant closure'
  url: https://api.crossref.org/works?query.bibliographic=Schwerdt+labor+tu
    ↪ rnover+before+plant+closure+leaving+the+sinking+ship&select=DOI,title
    ↪ ,author,issued,container-title&rows=3
  http_status: 200
- id: q52_utilisation_uk_sp
  api: openalex
  probe: Danish job-loss utilisation (British spelling)

```

```

url: https://api.openalex.org/works?filter=title_and_abstract.search:%22h
  ↪ ealthcare%20utilisation%22%20unemployment%20Denmark&sort=relevance_sc
  ↪ ore:desc&select=id,doi,display_name,publication_year,cited_by_count,p
  ↪ rimary_location&per-page=6
http_status: 200
- id: q53_gp_unemployed_dk
  api: openalex
  probe: Danish general practice + unemployed sweep
  url: https://api.openalex.org/works?filter=title_and_abstract.search:%22g
    ↪ eneral%20practice%22%20unemployed%20Danish&sort=cited_by_count:desc&s
    ↪ elect=id,doi,display_name,publication_year,cited_by_count,primary_loc
    ↪ ation&per-page=6
  http_status: 200
- id: q54_ida_datapaper
  api: crossref
  probe: IDA database documentation DOI probe (grey-literature check)
  url: https://api.crossref.org/works?query.bibliographic=Danish+integrated
    ↪ +database+for+labour+market+research+IDA&select=DOI,title,author,issu
    ↪ ed,container-title&rows=3
  http_status: 200
- id: q55_rx_displacement
  api: openalex
  probe: 'KILL(a): job displacement + antidepressants'
  url: https://api.openalex.org/works?filter=title_and_abstract.search:%22j
    ↪ ob%20displacement%22%20antidepressants&sort=cited_by_count:desc&selec
    ↪ t=id,doi,display_name,publication_year,cited_by_count,primary_locatio
    ↪ n&per-page=6
  http_status: 200
- id: q56_before_layoff
  api: openalex
  probe: 'KILL(a): ''before layoff'' + health'
  url: https://api.openalex.org/works?filter=title_and_abstract.search:%22b
    ↪ efore%20layoff%22%20health&sort=relevance_score:desc&select=id,doi,di
    ↪ splay_name,publication_year,cited_by_count,primary_location&per-page=6
  http_status: 200
- id: q57_jobloss_ooH
  api: openalex
  probe: 'KILL(d): job loss + out-of-hours'
  url: https://api.openalex.org/works?filter=title_and_abstract.search:%22j
    ↪ ob%20loss%22%20%22out-of-hours%22&sort=relevance_score:desc&select=id
    ↪ ,doi,display_name,publication_year,cited_by_count,primary_location&pe
    ↪ r-page=6
  http_status: 200
- id: q58_weekly_closure_es
  api: openalex
  probe: 'KILL(b): plant closures + weekly'
  url:
    ↪ https://api.openalex.org/works?filter=title_and_abstract.search:%22pl
    ↪ ant%20closures%22%20weekly&sort=relevance_score:desc&select=id,doi,di
    ↪ splay_name,publication_year,cited_by_count,primary_location&per-page=6
  http_status: 200
- id: q59_iversen_sabroe_psy
  api: crossref
  probe: 'Anchor: Iversen & Sabroe, well-being after company closedown'
  url: https://api.crossref.org/works?query.bibliographic=Iversen+Sabroe+ps
    ↪ ychological+well-being+unemployed+employed+company+closedown&select=D
    ↪ OI,title,author,issued,container-title&rows=3
  http_status: 200
- id: q60_masslayoff_hcu
  api: openalex

```

```

probe: 'KILL(b/d): mass layoffs + health care (top-cited)'
url: https://api.openalex.org/works?filter=title_and_abstract.search:%22m
  ↳ ass%20layoffs%22%20%22health%20care%22&sort=cited_by_count:desc&selec
  ↳ t=id,doi,display_name,publiation_year,cited_by_count,primary_locatio
  ↳ n&per-page=6
http_status: 200
- id: q61_jebo_lead_effects
  api: openalex
  probe: 'Abstract fetch: Wunder & Zeydanli 2021 JEBO lead effects of plant
    ↳ closures'
  url: https://api.openalex.org/works/doi:10.1016/j.jebo.2021.05.023?select
    ↳ =id,doi,display_name,publiation_year,cited_by_count,authorships,abst
    ↳ ract_inverted_index
  http_status: 200
- id: q62_ajph_mhcu
  api: openalex
  probe: 'Abstract fetch: Modrek/Hamad/Cullen 2014 AJPH mental health care
    ↳ utilization'
  url: https://api.openalex.org/works/doi:10.2105/ajph.2014.302219?select=i
    ↳ d,doi,display_name,publiation_year,cited_by_count,authorships,abstra
    ↳ ct_inverted_index
  http_status: 200
nearest_prior_work:
- ref: 'Beale N, Nethercott S (1985, J R Coll Gen Pract; DOI-verified 1997
  ↳ reprint), 'Job-loss
    and family morbidity: a study of a factory closure'''
  doi: 10.1007/978-1-349-25349-4_17
  verified_via: crossref q29/q47; original 1985/86 JRCPG articles are
    ↳ pre-DOI, reprint chapter
    is the resolvable record
  why_it_differs: 'THE construct-level ancestor of H1: GP consultation
    ↳ records around a
      single UK factory closure, with elevated consultations/referrals
        ↳ already during the
      pre-closure insecurity phase. Differs: one factory (~130 families), no
        ↳ population registers,
      no matched riskset controls, no weekly billing clock, descriptive
        ↳ rates. DS1.1 cannot
      claim the phenomenon is unobserved -- only that it has never been
        ↳ estimated at register
      scale with matched controls at weekly resolution.'
- ref: Kasl SV, Cobb S (1970, Psychosomatic Medicine), 'Blood Pressure
  ↳ Changes in Men Undergoing
    Job Loss'
  doi: 10.1097/00006842-197001000-00002
  verified_via: crossref q30
  why_it_differs: 'Anticipation-phase health measurement around plant
    ↳ closedown is 55 years
      old: the Michigan study measured physiology DURING the anticipation
        ↳ phase before shutdown.
      Differs: physiological outcomes not care-seeking, two plants,
        ↳ prospective panel not
      administrative data. Kills any 'first to look before the blow'
        ↳ rhetoric in general
      form.'
- ref: Iversen L, Sabroe S, Damsgaard MT (1989, BMJ), 'Hospital admissions
  ↳ before and after
    shipyard closure'
  doi: 10.1136/bmj.299.6707.1073

```

verified_via: crossref q31; sister papers Iversen & Sabroe 1988 J Social
 ↳ Issues 10.1111/j.1540-4560.1988.tb02096.x
 and JECH 10.1136/jech.42.4.396; 1987 Springer chapter
 ↳ 10.1007/978-3-642-83112-6_3 via
 openalex q25

why_it_differs: 'Danish, closure-based, and explicitly ''before and
 ↳ after'': hospital
 admissions around the Elsinore shipyard closure. Differs: hospital
 ↳ admissions not primary-care
 billing, one shipyard, 1980s, coarse periodization, no spurious-closure
 ↳ contrast. Removes
 the ''first Danish pre-closure utilization look'' claim; leaves the
 ↳ register-scale primary-care
 version open.'

- ref: 'Wunder C, Zeydanli T (2021, J Economic Behavior & Organization),
 ↳ ''The early costs
 of plant closures: Evidence on lead effects on workers'' subjective and
 ↳ objective outcomes'''
 doi: 10.1016/j.jebo.2021.05.023
 verified_via: openalex q58/q61 (abstract not deposited;
 ↳ title+venue+authors verified)

why_it_differs: 'Lead (anticipation) effects of plant closures are
 ↳ already a formalized
 estimand in economics, estimated on annual survey panels with
 ↳ subjective well-being
 and objective labor outcomes. Differs: no healthcare utilization
 ↳ outcomes, annual resolution,
 survey scale. DS1.1's novelty is the outcome domain (administrative
 ↳ billing) and the
 weekly clock, not the lead-effect concept.'

- ref: 'Hamad R, Modrek S, Cullen MR (2015, Health Services Research),
 ↳ ''The Effects of
 Job Insecurity on Health Care Utilization: Findings from a Panel of
 ↳ U.S. Workers'''
 doi: 10.1111/1475-6773.12393
 verified_via: crossref q36 + openalex abstract q45

why_it_differs: 'Closest high-frequency prior: monthly claims-based
 ↳ utilization around
 mass layoffs (segmented regressions, individual FE, high- vs low-layoff
 ↳ plants). Found
 outpatient utilization FELL at high-layoff plants (foregone care) while
 ↳ emergency use
 rose -- prior art for H3's unplanned-composition idea and directional
 ↳ counter-evidence
 for H1. Differs: US employer-insured single firm, survivors
 ↳ (insecurity) not closure-displaced
 workers, monthly not weekly, no closure register, no spurious-closure
 ↳ placebo.'

- ref: 'Modrek S, Hamad R, Cullen MR (2014, Am J Public Health),
 ↳ ''Psychological Well-Being
 During the Great Recession: Changes in Mental Health Care Utilization
 ↳ in an Occupational
 Cohort'''
 doi: 10.2105/ajph.2014.302219
 verified_via: openalex q60/q62 with abstract

why_it_differs: 'Mental-health service and medication utilization rose
 ↳ around the firm's
 2009 mass layoffs, more where layoffs were heavier. Same territory
 ↳ (labor shock -> utilization

trajectories in claims data) at yearly/piecewise resolution in one US
 ↳ firm; no closures,
 no primary-care billing system, no anticipation window keyed to
 ↳ establishment death.

- ref: Schaller J, Stevens AH (2015, J Health Economics), 'Short-run
 ↳ effects of job loss
 on health conditions, health insurance, and health care utilization'
 doi: 10.1016/j.jhealeco.2015.07.003
 verified_via: 'crossref q06 (note: correct DOI suffix is 2015.07.003, not
 ↳ the often-miscited
 05.003)'
 why_it_differs: 'The canonical job-loss-and-utilization paper: US survey
 ↳ data, post-displacement,
 found reduced insurance and REDUCED doctor visits with modest
 ↳ health-condition effects.
 Differs: no anticipation window, no administrative billing, US
 ↳ insurance channel dominates.
 Its utilization-drop result is publishable-null-adjacent
 ↳ counter-evidence for DS1.1's
 excess-contact hypotheses in the post period.'

- ref: Kuhn A, Lalive R, Zweimueller J (2009, J Health Economics), 'The
 ↳ public health costs
 of job loss'
 doi: 10.1016/j.jhealeco.2009.09.004
 verified_via: crossref q05; also surfaced by openalex kill-probes q24/q55
 ↳ as the only
 hit for displacement+antidepressants
 why_it_differs: 'Austrian plant-closure workers'' post-separation
 ↳ healthcare costs (hospitalizations,
 drug purchases): overall effects modest, concentrated in mental-health
 ↳ drugs for men.
 Differs: post-event only, multi-year aggregated costs, no weekly clock,
 ↳ no primary-care
 contact counts, no anticipation. Sets the expected-effect-size bar low.'

- ref: Browning M, Moller Dano A, Heinesen E (2006, Health Economics), 'Job
 ↳ displacement
 and stress-related health outcomes'
 doi: 10.1002/hec.1101
 verified_via: crossref q02
 why_it_differs: 'Danish displacement canon, and a NULL: no effect of
 ↳ displacement on stress-related
 hospitalizations. Differs: hospitalization not primary care, annual
 ↳ timing, post-event.
 Establishes that the Danish setting already contains a published null
 ↳ on displacement-health
 -- DS1.1's null branch is publishable and must be pre-registered as
 ↳ informative.'

- ref: Browning M, Heinesen E (2012, J Health Economics), 'Effect of job
 ↳ loss due to plant
 closure on mortality and hospitalization'
 doi: 10.1016/j.jhealeco.2012.03.001
 verified_via: crossref q03; also top hit of openalex Danish sweep q25
 why_it_differs: 'The closest Danish exposure construct:
 ↳ register-identified plant closures
 (IDA-based) with matched/duration designs. Differs: outcomes are
 ↳ mortality and cause-specific
 hospitalization, annual event time, post-closure only, no sygesikring
 ↳ billing, no spurious-closure
 placebo, no out-of-hours dimension. DS1.1 is the primary-care, weekly,
 ↳ pre-event complement

to this paper -- not a replacement.'

- ref: 'Ferrie JE, Shipley MJ, Marmot MG, Stansfeld S, Smith GD (1995, BMJ), 'Health effects of anticipation of job change and non-employment: longitudinal data from the Whitehall II study'''
doi: 10.1136/bmj.311.7015.1264
verified_via: crossref q10 (plus Ferrie 2002 JECH 10.1136/jech.56.6.450)
why_it_differs: 'Anticipation of job loss as a health exposure is a named, 30-year-old literature (self-reported morbidity, physiological measures, survey panels). Differs: no utilization records, no establishment events, no billing data. DS1.1 contributes measurement (administrative, weekly, matched), not the concept.'
- ref: Eliason M, Storrie D (2009, J Human Resources / Social Science & Medicine), 'Does Job Loss Shorten Life?' and 'Job loss is bad for your health'
doi: 10.1353/jhr.2009.0020
verified_via: 'crossref q07/q08 (SSM: 10.1016/j.socscimed.2009.01.021)'
why_it_differs: 'Swedish closure-based mortality and cause-specific hospitalization at annual resolution, post-event. No primary care, no anticipation, no weekly clock.'
- ref: 'Vahtera J, Kivimaki M, Pentti J, et al. (2004, BMJ), 'Organisational downsizing, sickness absence, and mortality: 10-town prospective cohort study'''
doi: 10.1136/bmj.37972.496262.0d
verified_via: crossref q32
why_it_differs: 'The downsizing-survivor health canon: sickness absence and mortality among stayers after downsizing. Adjacent to H1's insecurity mechanism but measures absence/mortality, not care-seeking volume or timing composition, and is keyed to downsizing episodes, not establishment deaths.'
- ref: Hethery-Maier T, Schmieder JF (2013, Schmollers Jahrbuch) and Benedetto G, Haltiwanger J, Lane J, McKinney K (2007, JBES), worker-flow classification of establishment births/deaths
doi: 10.3790/schm.133.4.477
verified_via: 'crossref q12/q13 (Benedetto: 10.1198/073500106000000620; NBER w19730)'
why_it_differs: 'The spurious-closure measurement canon: en-bloc worker flows identify ID reshuffling vs real establishment deaths. Used in displacement papers as sample-exclusion hygiene. No paper found that elevates the spurious group to an explicit placebo arm for HEALTH outcomes -- that H2 framing appears open. Danish IDA continuity conventions themselves are grey literature (q54 found no DOI-indexed documentation).'
- ref: Rugulies R, Thielen K, Nygaard E, Diderichsen F (2009, J Epidemiol Community Health), 'Job insecurity and the use of antidepressant medication among Danish employees'
doi: 10.1136/jech.2008.078493
verified_via: 'crossref q50 (note: 2009, not the remembered 2010)'

why_it_differs: 'Danish job insecurity -> medication utilization already
 ↳ exists (survey-measured
 insecurity, register antidepressant purchases). Differs: perceived
 ↳ insecurity not establishment
 events, prescriptions not primary-care contacts, annual follow-up.
 ↳ Squeezes the 'Danish
 insecurity-to-utilization' gap; DS1.1's remaining space is
 ↳ event-based exposure +
 billing outcomes + weekly timing.'

shrinkages:

- 'S1 (headline rhetoric): The abstract's 'nearly everything known about
 ↳ its health consequences
 is measured after separation' overstates. Anticipation-phase health
 ↳ around plant closure
 has a 55-year published lineage: Kasl & Cobb 1970 (physiology during
 ↳ anticipation), Beale
 & Nethercott 1985-88 (GP consultations around a factory closure incl.
 ↳ insecurity phase),
 Iversen/Sabroe/Damsgaard 1989 (Danish hospital admissions before and
 ↳ after shipyard closure),
 Ferrie 1995 Whitehall II (anticipation of job change), Wunder & Zeydanli
 ↳ 2021 (formal
 lead effects of plant closures). Rewrite: 'measured after separation IN
 ↳ THE MODERN REGISTER-BASED
 LITERATURE, and never at sub-annual resolution at population scale.''
- 'S2 (H1 not a phenomenon-first): Beale & Nethercott observed elevated GP
 ↳ consultation
 rates around a factory closure including the pre-closure insecurity
 ↳ period in the 1980s.
 H1's claim must be: first population-scale, register-based,
 ↳ riskset-matched, weekly-resolution
 estimate of pre-closure primary-care excess -- not first observation of
 ↳ anticipatory care-seeking.'
- 'S3 (Danish-first not available in general form):
 ↳ Iversen/Sabroe/Damsgaard 1989 BMJ is
 Danish, closure-based, and covers the pre-closure window (hospital
 ↳ admissions). Only 'first
 Danish register-scale PRIMARY-CARE billing look' survives.'
- 'S4 (estimand concept taken): 'Lead effects' of plant closures are an
 ↳ existing estimand
 (Wunder & Zeydanli 2021, JEB0). DS1.1's contribution is resolution
 ↳ (weekly billing) and
 outcome domain (utilization/fees/timing), not the idea of pre-event
 ↳ effects.'
- 'S5 (high-frequency increment, not category): Monthly claims-based
 ↳ utilization trajectories
 around mass-layoff events exist (Hamad/Modrek/Cullen 2015, segmented
 ↳ monthly regressions).
 Weekly is a 4x refinement plus population scale, and within this program
 ↳ GH1.2 already
 uses the HONUGE weekly clock -- instrument novelty is internal-relative,
 ↳ not absolute.'
- 'S6 (H3 has direct prior art): the 'unplanned care share rises under
 ↳ labor-shock stress'
 signature was already reported by Hamad et al. 2015 (outpatient down,
 ↳ emergency up under
 insecurity). H3 must cite it and position as the first
 ↳ within-primary-care OOH composition
 test around register-identified closures, and must acknowledge the
 ↳ possibility that TOTAL

contacts fall while OOH share rises.'

- 'S7 (H2 distinction is canon; only the placebo use is new):
 - ↪ true-vs-spurious establishment
 - death classification is established measurement methodology (Benedetto et al. 2007; Hethcote-Maier & Schmieder 2013) and standard exclusion hygiene in displacement papers (e.g., Kuhn et al. 2009). H2's novelty is confined to using absorbed pseudo-closures as an explicit placebo arm for health-billing outcomes -- claim exactly that, nothing more.'
- 'S8 (identification concession): with annual November-status closure
 - ↪ dating, DS1.1 measures pre-boundary excess, not announcement-locked anticipation; survey designs (Wunder & Zeydanli; Ferrie) and mandated-notice designs (e.g., the 2026 Brazil advance-notice working paper surfaced in q42) have sharper anticipation timing than DS1.1 ever will.
 - ↪ The protocol must not describe H1 as identifying announcement effects.'
- 'S9 (anticipatory attrition is a known confound, not a discovery):
 - ↪ Schwerdt 2011 (Labour Economics, 10.1016/j.labeco.2010.08.003) shows workers leave sinking ships early and leavers differ; November-status cohort anchoring inherits this selection and the protocol must present it as inherited literature, with the early-exit margin as a pre-registered robustness axis.'
- 'S10 (no instrument novelty in the register): the sygesikring/NHSR is a
 - ↪ mature, heavily-published research register (Andersen et al. 2011, 10.1177/1403494810394718; large descriptive OOH literature q49). Novelty lies solely in pairing it with IDAS establishment demography at weekly event time.'

kill_condition_probes:

- condition: (a) anticipatory healthcare utilization before
 - ↪ displacement/closure already published

queries:

- q14
- q15
- q16
- q23
- q24
- q26
- q29
- q30
- q31
- q40
- q42
- q43
- q46
- q55
- q56

outcome: NOT killed, but heavily shrunk. Construct-level priors exist at

- ↪ small scale (Beale & Nethercott single-factory GP records; Iversen 1989 single-shipyard admissions; Kasl-Cobb

physiology; Wunder-Zeydanli survey lead effects; Hamad 2015 survivor
 ↳ insecurity). No
 study found estimating population-scale administrative primary-care
 ↳ utilization trajectories
 before establishment closure with matched controls. H1 survives as a
 ↳ measurement-scale
 claim only (see S1-S4).
 - condition: (b) weekly/high-frequency utilization trajectories around
 ↳ displacement already
 published
 queries:
 - q17
 - q18
 - q36
 - q41
 - q44
 - q45
 - q58
 - q60
 - q62
 outcome: NOT killed. Finest found is monthly (Hamad/Modrek/Cullen 2015
 ↳ claims panel).
 Zero hits for weekly-resolution utilization around
 ↳ displacement/closure. Weekly at register
 scale appears open, but must be framed as a resolution increment over
 ↳ monthly, not a
 first look inside the year.
 - condition: (c) true-vs-spurious closure contrast in health outcomes
 ↳ already published
 queries:
 - q12
 - q13
 - q19
 - q51
 - q54
 outcome: NOT killed. The classification literature (Benedetto 2007;
 ↳ Hethey-Maier & Schmieder
 2013) is measurement-only; displacement-health papers use it to EXCLUDE
 ↳ spurious closures,
 and no work was found using absorbed pseudo-closures as an explicit
 ↳ placebo arm for
 any health outcome. H2's placebo framing appears genuinely open
 ↳ (subject to S7's narrow
 claim).
 - condition: (d) out-of-hours/unplanned composition shifts around job loss
 ↳ already published
 queries:
 - q20
 - q37
 - q38
 - q49
 - q57
 - q60
 outcome: NOT killed. Danish OOH literature is descriptive/cross-sectional
 ↳ (socioeconomic
 gradients in lægevagt use); no event-based OOH-composition study around
 ↳ job loss found.
 Nearest analog is Hamad 2015's ED-vs-outpatient shift under insecurity
 ↳ (S6). H3 survives
 with a citation duty to Hamad et al.

counter_evidence:

- 'Browning, Moller Dano & Heinesen 2006 (10.1002/hecl.1101): Danish
 - ↳ displacement had NO effect on stress-related hospitalizations -- a published Danish null in this exact institutional setting.'
- 'Schmitz 2011 (10.1016/j.labeco.2010.08.005): plant-closure unemployment
 - ↳ shows no causal health deterioration in German panel data; observed unemployed-sick correlations are selection.'
- 'Salm 2009 (10.1002/hecl.1537): exogenous job loss from business closures
 - ↳ (US HRS) -> no effect on physical or mental health.'
- 'Schaller & Stevens 2015 (10.1016/j.jhealeco.2015.07.003): job loss
 - ↳ REDUCES doctor visits and utilization in the US. The insurance channel is absent in Denmark, but time-cost/presenteeism channels could likewise push DS1.1's pre- and post-event contacts DOWN (job search, stigma, income-linked copay-free but time-priced visits); the protocol must pre-register the negative direction as interpretable, not just 'no anticipation'.'
- 'Hamad, Modrek & Cullen 2015 (10.1111/1475-6773.12393): under
 - ↳ layoff-threat insecurity, outpatient utilization FELL (foregone/deferred care) while emergency use rose -- directional counter-evidence for H1's excess-contact prediction, and simultaneously the strongest prior FOR H3's composition shift.'
- 'Kuhn, Lalive & Zweimueller 2009 (10.1016/j.jhealeco.2009.09.004):
 - ↳ overall post-job-loss health-cost effects were modest and concentrated in mental-health drugs among men -- expected GP-billing effect sizes are plausibly small; power analysis should assume modest effects.'
- 'Null-is-publishable note: given Browning et al. 2006 and Schmitz/Salm, a finding of no pre-boundary excess (or a utilization DROP) is an informative, publishable Registered Report outcome and must be written into the inference criteria.'

novelty_verdict: sharpen

verdict_rationale: 'PROCEED on the design, SHARPEN every first-ness claim.

- ↳ The precise empirical object -- population-scale weekly primary-care billing trajectories (volume, fees, out-of-hours composition) around register-identified establishment terminal events, with a true-vs-spurious closure placebo and 52-week billing-history matching -- has no published occupant: kill probes found nothing at weekly resolution, nothing using spurious closures as a health placebo, and nothing on OOH composition around job loss. But the anticipation territory as a CONCEPT is emphatically taken, with a 55-year lineage running
 - ↳ Kasl-Cobb 1970 -> Beale-Nethercott 1985 (GP consultations, including the pre-closure insecurity phase) -> Iversen-Sabroe-Damsgaard 1989 (Danish, shipyard, admissions before/after) -> Ferrie 1995 -> Rugulies 2009 (Danish

insecurity -> antidepressants) -> Hamad-Modrek-Cullen 2014/15 (monthly
 ↳ claims around mass
 layoffs) -> Wunder-Zeydanli 2021 (formal lead effects of plant closures).
 ↳ The abstract's
 'nearly everything known is measured after separation' must be
 ↳ rewritten, H1 must claim
 measurement scale and resolution (first register-scale, matched, weekly)
 ↳ rather than phenomenon
 discovery, H3 must cite Hamad et al. as direct prior art for the
 ↳ unplanned-care shift,
 and H2's novelty must be stated narrowly as the placebo elevation of a
 ↳ canonical measurement
 distinction. Direction of H1 is genuinely contested (Hamad and
 ↳ Schaller-Stevens found
 utilization falls), so the null and the negative direction must be
 ↳ pre-registered as informative
 outcomes.'

Data contract

Only IDAN, IDAS, SSSY, SYSI; actual data never accessed in conversion.

Current schema

registers:
 - register: IDA workplaces (IDAS)
 register_da: IDA arbejdssteder
 documentation: TIMES ida-databasen/ida-arbejdssteder (The IDA database /
 ↳ IDA workplaces)
 and HKV ida-arbejdssteder, valid from 01-01-1980
 years: 1980-2024 panel; event cohorts c in 2009-2023 (confirmatory);
 ↳ ENDNEDL ends 2023
 variables:
 - name: LBNR
 long_name: Workplace serial number
 role: key
 use: workplace identity; identity chains via IDFREM/IDTILB;
 ↳ identity-disappearance events
 between c and c+1
 - name: ARENR
 long_name: Arb number
 role: key
 use: workplace key, 2007+ vintage (confirmatory era)
 - name: CVRNR
 long_name: CVR number
 role: key
 use: firm key, 2007+ vintage (confirmatory era)
 - name: ENDNEDL
 long_name: Final closure of workplace
 role: exposure
 use: true-closure flag (with $s(w) < 0.30$); validity through 2023 caps the
 ↳ last cohort
 - name: IDFREM
 long_name: Identity of place of work forward in time
 role: exposure
 use: forward-continuation identity for spurious-closure classification;
 ↳ identity chain
 construction
 - name: IDTILB

- long_name: Identity of place of work backward in time
 - role: exposure
 - use: backward identity chain (control-stability and event-panel
↳ construction); HKV valid
from 01-01-1981
- name: REGELFR
 - long_name: Rule regarding retained workplace the following year
 - role: exposure
 - use: forward-identity rule for spurious-closure classification
- name: FRELFR
 - long_name: Firm-internal/external closure via absorption
 - role: exposure
 - use: en-bloc absorption indicator for spurious-closure classification
- name: ANDELOPS
 - long_name: Share of absorbed among employees
 - role: check
 - use: IDAS-native absorption share; classification cross-validation
↳ against worker-flow
s(w)
- name: ANDELUDS
 - long_name: Share of employees split off to another workplace among
↳ employees
 - role: check
 - use: IDAS-native split-off share; classification cross-validation
- name: AFGRATE
 - long_name: Exit rate in the subsequent annual period, in per cent
 - role: control
 - use: stable-control definition (AFGRATE < 30%); mass-separation exclusion
- name: ANTALXFR
 - long_name: Change in the number of employees in the following year
 - role: control
 - use: stable-control definition (|ANTALXFR| <= 30%)
- name: ANTNOV
 - long_name: Number of employees as at November
 - role: control
 - use: 'workplace size: eligibility 5-1000, exact-stratum size classes,
↳ event presence (ANTNOV
>= 5)'
- name: ARB_HOVED_BRA_DB07
 - long_name: Industry of the workplace
 - role: control
 - use: NACE section exact stratum (2008+ vintage)
- name: KOM
 - long_name: Municipality of the workplace
 - role: control
 - use: workplace municipality mapped to the 5 post-2007 regions (exact
↳ stratum)
 - verification: verified in the DST Forskningsservice four-register
↳ inventory for the analysis
server; absent from the public TIMES/HKV parses - disclosed
- name: ARBGNR
 - long_name: Employer number
 - role: key
 - use: legacy workplace key (pre-2008 vintage); E1 SYSI-era exploratory
↳ replication only
- name: ARBSTK
 - long_name: Workplace code
 - role: key
 - use: legacy workplace key (pre-2008 vintage); E1 exploratory only
- register: IDA employments (IDAN)

```

register_da: IDA ansaettelser
documentation: TIMES ida-databasen/ida-ansaettelser (The IDA database / IDA
↳ employments);
  HKV analogues Beskaeftigelsesoplysninger / Loenforhold / Ledighed der
  ↳ vedroerer IDA ansaettelser,
  valid from 01-01-1980
years: 1980-2024 panel; cohort anchoring at November of year c-1; worker
↳ flows into c+1
variables:
- name: PNR
  long_name: De-identified person number
  role: key
  use: person spine joining IDAN, SSSY, SYSI
  verification: verified in the DST Forskningssservice four-register
  ↳ inventory for the analysis
  ↳ server; absent from the public TIMES/HKV parses - disclosed
- name: LBNR
  long_name: Workplace serial number
  role: key
  use: employment-to-workplace link (with ARBNR/CVRNR, 2008+)
- name: ARBNR
  long_name: Arb number
  role: key
  use: workplace key on the employment relation, 2008+
- name: CVRNR
  long_name: CVR number
  role: key
  use: firm key on the employment relation, 2008+
- name: ANSAAR
  long_name: Year of employment at the workplace
  role: control
  use: tenure (c-1 minus ANSAAR); inclusion floor ANSAAR <= c-2;
  ↳ Mahalanobis component
- name: BREDT_LOEN_BELOEB
  long_name: Broad salary amount
  role: control
  use: primary-job rule (largest BREDT_LOEN_BELOEB in c-1); baseline wage
  ↳ decile (Mahalanobis component; E3 heterogeneity)
- name: LOENTIMER
  long_name: Paid hours (LOENTIMER)
  role: control
  use: deterministic primary-job tie-break (larger LOENTIMER, then lower
  ↳ ARBNR)
- name: ARB_SEKTORKODE
  long_name: Workplace sector code
  role: control
  use: private-sector filter via modal sector among year c-1 employees
  ↳ (2008+)
verification: 'Exception, disclosed (converse of the inventory-only
↳ variables): publicly
  documented (TIMES ida-ansaettelser) but NOT present in the recorded DST
  ↳ Forskningssservice
  four-register inventory extract for the analysis server, and the
  ↳ claimed 2008+ validity
  rests on the public parse alone. Server presence and coverage validated
  ↳ at Stage 1 before
  any classification work; frozen fallback registered in
  ↳ analysis_plan.contingent_decisions

```

```

    (private-sector filter from ARB_HOVED_BRA_DB07, NACE DB07 sections
    ↪ O/P/Q excluded).'
```

- name: ANSXFREM
 long_name: Change of employment up to the following year
 role: check
 use: E4 arm-comparative attrition accounting (forward-separation flag)
- name: ARBGNR
 long_name: Employer number
 role: key
 use: legacy key (pre-2008); E1 exploratory only
 verification: carried on legacy IDAN vintages per the Forskningservice
 ↪ inventory; the
 public parses document it on IDA workplaces
- name: ARBSTK
 long_name: Workplace code
 role: key
 use: legacy key (pre-2008); E1 exploratory only
 verification: carried on legacy IDAN vintages per the Forskningservice
 ↪ inventory; the
 public parses document it on IDA workplaces
- name: STILL
 long_name: Occupational position
 role: control
 use: legacy job characteristic for E1-era matching (SYSI lacks age/sex)
- name: TILKNYT
 long_name: Attachment to primary workplace
 role: control
 use: legacy primary-employment marker; E1 exploratory only
- name: ANSDAGE
 long_name: Number of days employed
 role: control
 use: legacy job characteristic for E1-era matching
- name: TIMELON
 long_name: Hourly wage in the employment
 role: control
 use: legacy wage measure for E1-era matching
- register: Health insurance statistics, current service series (SSSY)
 register_da: Sygesikringsstatistik (SSSY)
 documentation: TIMES sygesikringsstatistik (Health insurance statistics);
 ↪ HKV sygesikring
 - ydelser / honorarer / kontakter (SPEC2 valid from 01-01-2005)
 years: 2005-2025 service rows; extraction 2006-2025 (baselines c-3 onward,
 ↪ post-years through
 2024/2025)
 variables:
 - name: PNR
 long_name: De-identified person number
 role: key
 use: person spine
 verification: verified in the DST Forskningservice four-register
 ↪ inventory for the analysis
 server; absent from the public TIMES/HKV parses - disclosed
- name: HONUGE
 long_name: Fee week
 role: outcome
 use: 'the weekly clock: person x HONUGE week aggregation,
 ↪ ISO-week-to-quarter mapping'
- name: KONTAKT
 long_name: Contact
 role: outcome

```

use: row-level contact marker summed to quarterly GP contact counts
↳ (frozen fallback definition
    if in-environment validation fails)
- name: YDLANT
  long_name: Number of services within the specialty
  role: outcome
  use: service counts (descriptives; E1-era contact proxy weighting)
- name: BRUHON
  long_name: Gross fee to doctor etc.
  role: outcome
  use: 'gross fees: gated H2 fee secondary, E3 fee analogues, Mahalanobis
  ↳ log(1+BRUHON)
    component; nominal, compared only within cohort x calendar-quarter
    ↳ cells'
- name: SPEC2
  long_name: 2-digit specialisation
  role: outcome
  use: frozen GP (70-80) and OOH (81-84, 89) code lists; parsed labels mark
  ↳ 71/73/75 as
    "evening etc." GP services - kept daytime in the primary classifier by
    ↳ registered choice
    (see frozen_parameters.spec2_evening_gp_disclosure)
- name: YDLTID
  long_name: Time code for the benefit
  role: outcome
  use: robustness OOH classifier; missing/'not relevant' classified daytime
  ↳ (conservative
    against H3)
- name: AFRPER
  long_name: Settlement period
  role: check
  use: settlement-lag characterization
- name: BEHANDLINGSDATO
  long_name: The treatment date for the service
  role: check
  use: HONUGE settlement-lag calibration on 2021+ rows only (guard-band
  ↳ contingency); can
    calibrate but never replace HONUGE
- name: SIKGRUP
  long_name: Health insurance group
  role: check
  use: exclusion of modal group-2 insured (different billing regime)
- name: ALDERIMP
  long_name: Age at end of period, incl. imputed
  role: control
  use: age bands for exact strata (with explicit unknown stratum);
  ↳ eligibility 25-59
- name: KOENIMP
  long_name: Sex incl. imputed
  role: control
  use: sex stratum (with explicit unknown stratum); E3 heterogeneity
- name: YDERNR
  long_name: Provider number (ydernummer)
  role: key
  use: frozen fallback contact definition (distinct PNR x week x YDERNR x
  ↳ code-group cells)
- register: Health insurance statistics, legacy closed series (SYSI)
  register_da: Sygesikringsstatistik (SYSI)
  documentation: documented under TIMES sygesikringsstatistik and HKV
  ↳ sygesikring datasets

```

(valid from 01-01-1990); series closed 2005
years: 1990-2005; used ONLY in the exploratory SYSI-era replication E1
↪ (cohorts 1992-2003)

variables:

- name: PNR
long_name: De-identified person number
role: key
use: person spine
verification: verified in the DST Forskningsservice four-register
↪ inventory for the analysis
server; absent from the public TIMES/HKV parses - disclosed
- name: HONUGE
long_name: Fee week
role: outcome
use: weekly clock, legacy era
- name: SPECIALE
long_name: 6-digit speciality
role: outcome
use: era GP/00H classification (SPEC2 does not exist pre-2005); frozen
↪ era code list
- name: YDLANT
long_name: Number of services within the specialty
role: outcome
use: 'E1 contact proxy: service-row counts weighted by YDLANT (KONTAKT
↪ does not exist
pre-2005)'
- name: BRUHON
long_name: Gross fee to doctor etc.
role: outcome
use: era fee outcome
- name: YDLTID
long_name: Time code for the benefit
role: outcome
use: era time-of-day classification
- name: VAGTOMR
long_name: Out-of-hours (laegevagt) service area
role: outcome
use: era 00H marker, 1997-2005 only
verification: verified in the DST Forskningsservice four-register
↪ inventory for the analysis
server; absent from the public TIMES/HKV parses - disclosed
- name: AFRPER
long_name: Settlement period
role: check
use: era settlement-lag characterization
- name: SIKGRUP
long_name: Health insurance group
role: check
use: era exclusion of group-2 insured
- name: SIKREKOM
long_name: Municipality of the insured person
role: control
use: era patient geography (E1 only; SSSY has no patient geography)
- name: SIKREAMT
long_name: County of the insured person
role: control
use: era patient geography (E1 only)
verification: verified in the DST Forskningsservice four-register
↪ inventory for the analysis
server; absent from the public TIMES/HKV parses - disclosed

variable_verification_note: 'Every variable above was checked against the
↳ parsed public documentation
(data/mapping/times_variables_bilingual.csv;
↳ data/mapping/hkv_variables_bilingual.csv).
Exceptions carried with an explicit per-variable verification note: PNR,
↳ IDAS KOM, SYSI
VAGTOMR, and SYSI SIKREAMT are verified in the DST Forskningssservice
↳ four-register inventory
for the analysis server but are absent from the public TIMES/HKV parses;
↳ IDAN's legacy
workplace keys (ARBGNR/ARBSTK) are documented publicly on the workplace
↳ register and carried
on legacy IDAN vintages per the same inventory. Only these four registers
↳ exist on the analysis
server: there are no diagnosis, hospital, prescription, or population-spine
↳ registers in
scope, and outcomes are therefore utilization volume, fees, and timing
↳ composition - never
health states. One variable runs the other way and is disclosed as such:
↳ IDAN ARB_SEKTORKODE
- the private-sector eligibility filter - is publicly documented (TIMES
↳ ida-ansaettelser)
but absent from the recorded Forskningssservice inventory extract, its
↳ claimed 2008+ validity
resting on the public parse alone; its presence is validated at Stage 1
↳ before any classification
work, with a frozen ARB_HOVED_BRA_DB07-based fallback registered in
↳ analysis_plan.contingent_decisions.'

linkage:
keys:
- 'person spine: PNR (IDAN x SSSY x SYSI)'
- 'workplace: LBNR with ARBNR/CVRNR (IDAS x IDAN, 2008+ vintage - the
↳ entire confirmatory
tier); LBNR with ARBGNR/ARBSTK (pre-2008 vintage - E1 exploratory only);
↳ workplace identity
chains via IDFREM/IDTILB; never pooled across the 2007/08 IDA vintage
↳ break'
expected_linkage_rate_floor: '>= 80% of cohort workers (exposed + control
↳ pool) with >=
1 SSSY row of any kind in the 5-year window c-3..c+1 (outcome-neutral
↳ check; Danish annual
GP contact prevalence benchmark ~85%); below floor halts at Stage 3'

Population

definition: 'Registered cohorts (SSSY era): persons with primary
↳ November-status employment
in year c-1 (IDAN; primary job = largest BREDT_LOEN_BELOEB in c-1, ties
↳ broken by larger
LOENTIMER then lower ARBNR - deterministic) at (i) an event workplace whose
↳ IDAS identity
(LBNR, with ARBNR/CVRNR keys) is present in year c with ANTNOV >= 5 and
↳ absent in year c+1,
for closure years c in 2009-2023 (15 cohorts; c = 2023 is feasible because
↳ IDAS/IDAN extend
through 2024 for the c+1 identity-absence and worker-flow checks and SSSY
↳ through 2025 for
post-year outcomes; ENDNEDL ends 2023), or (ii) a stable control workplace
↳ (present with

unchanged identity c-2..c+1, no ENDNEDL, AFGRATE < 30% and |ANTALXFR| <= 30% in years c-1 and c). Attachment frozen at November of c-1 - before the terminal year - so sinking-ship attrition during year c cannot condition the sample (S9; Schwerdt 2011).
 Unit of analysis: person x event-quarter (aggregated from person x HONUFE fee-week), stacked by workplace-event cohort in 1:3 matched sets (one exposed worker + up to three riskset-matched controls, controls weighted 1/k within set).
 inclusion:
 - Age 25-59 at baseline, from ALDERIMP on the person's most recent SSSY row
 in c-3..c-1 (persons with no SSSY row in the lookup window enter an explicit 'unknown age/sex' matching stratum rather than being dropped)
 - 'Tenure >= 1 full year at the workplace: ANSAAR <= c-2 in the c-1 IDAN record'
 - 'Private-sector workplace: modal ARB_SEKTORKODE (IDAN, 2008+) among the workplace's year c-1 employees indicates private sector'
 - Workplace November size 5-1000 (ANTNOV in IDAS, year c-1) - cap raised from 500 to 1000 (Design B's cap), an honest eligibility enlargement admitting genuine large closures and their workers
 - 'Event workers: workplace classified true closure (ENDNEDL final closure AND worker-flow max-successor share s(w) < 0.30) or spurious closure (forward identity via IDFREME/REGELFR/FRELFREME AND s(w) >= 0.70)'
 exclusion:
 - Events with ambiguous successor share (0.30 <= s(w) < 0.70) - excluded from both classes, counts reported (GRAFT 4 monitors that this mid-zone holds < 35% of terminal events)
 - Workers whose modal SIKGRUP in baseline year c-1 SSSY rows is 2 (different billing regime; expected ~1-2%)
 - 'Workers appearing in more than one event cohort: assigned to the earliest event only'
 - Control-pool workers at workplaces experiencing any terminal event or mass separation (AFGRATE >= 30%) in c-2..c+1
 - Workers with missing ANSAAR or missing/zero annual pay in the c-1 IDAN record (share reported; expected < 2%)
 - 'Mass-separation events at surviving workplaces: excluded from the confirmatory tier entirely (scope discipline)'
 expected_n: 'Sampling plan sized, per the remediation directive's explicit authorization, at the FULL derived expectation's LOWER bound (not half of it): 15 cohorts x 10,000-20,000 eligible exposed workers/cohort-year = 150,000-300,000 exposed true-closure workers (the 500->1000 size-cap raise and the 15th cohort add further margin not counted in the lower

```

bound), yielding >= 135,000 matched sets at the registered >= 90%
↳ match-rate floor; spurious-closure
workers expected at 30-60% of the true count (45,000-90,000+ before
↳ matching). This sizing
is licensed only jointly with the PRE-OUTCOME REALIZED-N DEMOTION GATE
↳ (Stage 2): confirmatory
status requires realized matched sets >= N* per member (provisional: N*_H1
↳ = 100,000; N*_H2a
= 60,000 true + 15,000 spurious for gate F2; N*_H3 = 90,000; binding N*
↳ values are recomputed
by the re-run deposited simulation at Stage 0), with any shortfall
↳ resolving by demotion
to exploratory BEFORE any outcome data access - never by floor inflation.
↳ Project-viability
kill floors (500 true events / 10,000 sets; 150 spurious events / 3,000
↳ sets) are unchanged
and far below expectation.'
expected_n_basis: 'IDAS row scale from the Forskningservice inventory (~227k
↳ workplace rows/year,
~3.2M IDAN rows/year, ~88M SSSY rows/year); closure-rate basis now cited as
↳ required for
full-expectation sizing: Statistics Denmark business demography
↳ (Erhvervsdemografi/StatBank
DEMO series) establishment death rates of ~2-5% among employer workplaces,
↳ plus the Danish
displacement literature's demonstrated register cohort scale (Browning &
↳ Heinesen 2012,
J Health Econ, build Danish closure/mass-layoff displacement samples of
↳ comparable order
from the same IDA infrastructure); GP-contact base rates (~85% annual
↳ contact prevalence,
~4-5 contacts/person-year) from published DST health-insurance statistics;
↳ OOH share of
GP contact volume 8-12% from published Danish out-of-hours studies
↳ (Christensen & Olesen
1998 BMJ; Moth/Flarup/Christensen/Olesen/Vedsted LV-KOS 2011), simulated at
↳ the conservative
end (8%).'

```

Identification and claim ceiling

Source design

```

type: stacked matched difference-in-differences event study (matched
↳ event-study contrasts
  at quarterly confirmatory resolution with weekly descriptive trajectories;
  ↳ associational
  claim class)
design_name: 'Bracing for the Blow - remediated confirmatory design: stacked
↳ 1:3 riskset-matched
  difference-in-differences of quarterly GP billing around true vs spurious
  ↳ Danish workplace
  closures (SSSY era, closure cohorts 2009-2023), with the true-closure
  ↳ post-year excess as
  H2's confirmatory core, a spurious-arm falsification gate outside the Holm
  ↳ family, published-basis
  out-of-hours power calibration, and a pre-outcome realized-N demotion gate'
identifying_variation: 'Establishment terminal events in IDAS - workplace
↳ identities present

```

in year c (ANTNOV ≥ 5) and absent in year c+1 - across 15 closure cohorts
 ↪ 2009-2023, classified
 true vs spurious by IDAN worker-flow successor shares per the Benedetto et
 ↪ al. (2007) /
 Hethey-Maier & Schmieder (2013) canon, contrasted against riskset-matched
 ↪ workers at stable
 workplaces within exact strata (cohort x sex x age band x NACE section x
 ↪ size class x region)
 with 52 weeks of pre-baseline billing-history matching. Closure dating is
 ↪ annual November-status,
 so no sharp within-scope discontinuity exists: the claim ceiling is
 ↪ associational, all pre-event
 quantities are worded as pre-boundary excess (never announcement-dated
 ↪ anticipation), and
 the spurious-closure placebo plus the c-2 placebo panel carry the severity
 ↪ burden.'

alternatives_considered:
 - design: 'Round-1 conservative design (VOIDED): single-era stacked
 ↪ matched-PAIR (1:1) DiD,
 closure cohorts 2009-2022, sampling self-handicapped to half the lower
 ↪ bound of its own
 derived expectation (75k pairs), H2 as the true-minus-spurious
 ↪ differential, H3 simulated
 at a 3% OOH base rate'
 discrimination: identical contrast logic; differed only in matching ratio,
 ↪ sampling plan,
 H2's confirmatory construct, and H3's power basis
 why_not_selected: Its deposited Stage-0 power simulation emptied the
 ↪ confirmatory family
 via the registered fixed-point rule (H2 0.358, H1_trend 0.380, H3 0.614,
 ↪ H1 0.840) and
 the driver audit voided the resulting all-exploratory registration on
 ↪ 2026-08-31. The
 remediation (authorized honest levers only; no floor raised) produced the
 ↪ registered design;
 the full record is carried in remediation_record and the records of both
 ↪ power runs are
 published with the protocol.

- design: 'Design B, ambitious three-arm weekly event study 1991-2023:
 ↪ weekly-resolution confirmatory
 estimands, both eras (SYSI + SSSY) confirmatory, a mass-separation arm at
 ↪ surviving workplaces,
 ANTNOV cap 1000'
 discrimination: weekly confirmatory estimands would sharpen timing-shape
 ↪ claims; a second
 era and third arm would extend coverage and external validity
 why_not_selected: Weekly confirmatory estimands are hostage to the HONUKE
 ↪ settlement lag;
 the SYSI era cannot match on age or sex and sits across the 2007/08 IDA
 ↪ vintage break;
 the mass-separation arm breaks one-study scope discipline. Its honest
 ↪ elements - the ANTNOV
 1000 cap, pre-committed reversed-sign interpretations, and the absorption
 ↪ placebo - were
 absorbed into the registered design; the SYSI era survives as exploratory
 ↪ E1 and weekly
 resolution as descriptive trajectories under the GRAFT-2 step diagnostic.

threats:
 - threat: Anticipatory attrition - healthier/more mobile workers leave
 ↪ sinking workplaces

before closure (Schwerdt 2011), biasing late-window composition
addressed_by: Cohort frozen at November-status attachment in year c-1,
↳ before the terminal
year (S9); intention-to-follow regardless of subsequent separation; E4
↳ arm-comparative
attrition accounting panel quantifies the channel, labeled
↳ post-treatment-conditioned
descriptives

- threat: 'Billing responses to separations already realized within the
↳ terminal year - under
intention-to-follow, workers who separate from the closing workplace
↳ early in year c (the
wave E4 describes) contribute post-job-loss utilization to the
↳ pre-boundary Q3(c)/Q4(c)
window, so a Holm-significant θ_{t1} could be generated wholly by
↳ realized early job loss
with zero anticipation. Direction: inflates θ_{t1} relative to a
↳ pure-anticipation reading.
The frozen clause 'before any separation is visible in annual IDAN
↳ records'' holds at
the cohort-anchoring level (attachment frozen at November c-1) but not
↳ for individual
early leavers, whose separation is visible in the year-c IDAN record.'
addressed_by: The pre-committed H1 interpretation names both channels
↳ (anticipatory care-seeking
and/or billing responses to separations already realized early within the
↳ terminal year)
with the E4 year-c separation share bounding the latter; the E4
↳ arm-comparative separation
shares themselves; and the registered post-treatment-conditioned
↳ exploratory split of
 θ_{t1} by November-year-c attachment (E4), which localizes the channel
↳ without touching
the confirmatory estimand

- threat: Selection of sicker workers into dying workplaces (distressed-firm
↳ workforce composition)
masquerading as anticipation or as post-year excess
addressed_by: Exact + Mahalanobis matching on 52 weeks of pre-baseline
↳ billing (contacts,
fees, OOH, contact-weeks) in year c-1; c-2 placebo panel with numeric
↳ failure threshold;
falsification gate F2 - the spurious-closure arm separates identity-death
↳ from economic
job loss with a pre-committed disconfirmation trigger that can only take
↳ support away
from H2

- threat: HONUGE is the fee/settlement week, not the service week - billing
↳ lag blurs the
boundary and could relabel pre-boundary care as post-boundary
addressed_by: Quarter-level primary estimands (robust to ≤ 4 -week lags);
↳ BEHANDLINGSDATO
calibration (2021+) with the guard-band contingency; AFRPER used
↳ descriptively; GRAFT-2
week-0 step diagnostic bounds any timing-shape claim

- threat: Annual closure dating - the design cannot see announcement dates,
↳ so 'anticipation'
could be mislabeled
addressed_by: Estimand defined and worded exclusively as pre-boundary
↳ excess relative to

the closure-year boundary (S8); claim ceiling forbids announcement-locked
 ↳ language; gated
 H1_trend partial pattern pre-commits the flat-elevation reading when no
 ↳ within-year growth
 is seen

- threat: Local-economy co-shocks hitting exposed and control workers'
 ↳ regions simultaneously
 addressed_by: Exact matching on region and NACE section plus cohort x
 ↳ calendar-quarter fixed
 effects; residual local shocks affect controls too and attenuate the
 ↳ contrast (bias toward
 null, stated)
- threat: Control contamination by quietly dying or shrinking workplaces
 addressed_by: Stable-control definition (unchanged identity c-2..c+1, no
 ↳ ENDNEDL, AFGRAFE
 < 30%, |ANTALXFR| <= 30%); control-drop rule on data revision; GRAFT-9
 ↳ forward-window
 (c+2/c+3) sensitivity with registered direction note (truncated for the
 ↳ last two cohorts,
 disclosed)
- threat: Reverse causality at small firms - workforce or owner illness
 ↳ causes the closure
 addressed_by: Size floor ANTNOV >= 5; associational claim ceiling with the
 ↳ reverse channel
 named in the limitations; exploratory size heterogeneity (size classes
 ↳ now extend to 1000)
- threat: 2007/08 IDA vintage break contaminating any pooled estimate
 addressed_by: Confirmatory tier confined entirely to post-break cohorts (c
 ↳ >= 2009); SYSI-era
 analysis is exploratory replication (E1) and never pooled
- threat: Unobservable death/emigration (no population spine in scope)
 ↳ deflating post-event
 billing
 addressed_by: Named scope limitation; direction discussed (deflates the
 ↳ post window - conservative
 for theta1, attenuating for theta2a); pre-specified sensitivity excluding
 ↳ possible-death/emigration
 sets; GRAFT-6 2pp later-presence trigger forces best/worst-case bounds on
 ↳ the post-boundary
 quantities
- threat: 'Spurious-arm thinness - the structural cause of the first round''s
 ↳ collapse: any
 confirmatory quantity whose variance is dominated by the spurious arm is
 ↳ unpowerable at
 honest N'
 addressed_by: 'No confirmatory quantity now touches the spurious arm:
 ↳ theta2a is a true-arm-only
 contrast; the spurious arm serves as falsification gate F2 (whose job
 ↳ needs bounding,
 not detection, and whose operating characteristics are verified at Stage
 ↳ 0) and as the
 exploratory H2_diff; spurious-arm death caps the H2 verdict rather than
 ↳ silently weakening
 a test'
- threat: Residual underpowering if realized N, realized k, or the realized
 ↳ OOH share falls
 short of the sampling plan
 addressed_by: Pre-outcome Stage-2 realized-N gate with binding N*
 ↳ thresholds from the re-run

deposited simulation; blind re-simulation on realized
 ↳ Ns/k-distribution/baseline moments
 and realized OOH share; demotion-only resolution before any outcome
 ↳ access; realized-power/MDE
 table published regardless

falsification_tests:

- 'Falsification gate F2 (spurious-closure placebo arm): theta2b with the
 ↳ pre-committed disconfirmation
 trigger; outside the Holm family; verified operating characteristics
 ↳ registered (pass 0.886
 hypothesis-true / 0.978 inert; trigger 0.79 at floor-sized non-inert
 ↳ placebo, at 45,000
 spurious sets - the pre-matching worker count at the 30% band lower bound;
 ↳ the registered
 90% match-rate floor implies ~40,500 sets, so these characteristics are
 ↳ computed at a ~10%
 flattered lower-bound N); Stage-2 blind re-verification at realized N'
- 'c-2 pre-baseline placebo panel: exposed-minus-control contact difference
 ↳ across the four
 quarters of year c-2 with a numeric failure threshold - fails if and only
 ↳ if significant
 at $\alpha = 0.05$ AND $|\text{estimate}| > 1\%$ of control baseline mean, the single
 ↳ registered condition
 identical wherever this gate appears; one pre-specified re-match, then kill
 ↳ of the registered
 event-study tier'
- GRAFT 2 - week-0 boundary-step diagnostic on the weekly descriptive series
 ↳ (settlement-lag
 contamination bound on any timing-shape claim)
- GRAFT 3 - successor-share threshold sensitivities (20/80, 40/60 alongside
 ↳ primary 30/70)
 for theta2b/theta2_diff sign stability
- GRAFT 4 - s(w) bimodality check at Stage 1 (ambiguous mid-zone < 35% of
 ↳ terminal events)
- GRAFT 6 - later-presence differential trigger (2pp) forcing best/worst-case
 ↳ bounds on all
 post-boundary conclusions
- GRAFT 9 - control forward-window (c+2/c+3) contamination sensitivity with
 ↳ registered direction
 note (exclusion should weakly increase magnitudes; a decrease is flagged as
 ↳ diagnostic)
- 'Classification cross-validation: worker-flow classes vs IDAS-native flags
 ↳ (ENDNEDL/IDFREM/REGELFR/FRELFREM/ANDELUDS),
 >= 75% agreement kill floor, full cross-tabulation released'

narrative: 'This study asks whether Danish workers'' primary-care billing
 ↳ already rises while
 their workplace is dying - before any separation is visible in annual
 ↳ employment records
 - and whether the post-closure excess attaches to genuinely destroyed jobs
 ↳ rather than to
 workplaces that die only on paper. The instrument is the weekly clock of
 ↳ Danish health-insurance
 billing (SSSY, HONUUGE fee-week) held against annually dated establishment
 ↳ terminal events
 from the IDA workplace registers. For closure cohorts 2009-2023 (entirely
 ↳ inside the post-2008
 IDA key era), every worker with primary November-status employment in year
 ↳ c-1 at a workplace
 that disappears between years c and c+1 is classified by the worker-flow
 ↳ canon of Benedetto

et al. (2007) and Hethey-Maier & Schmieder (2013): true closure when no
 ↳ successor absorbs
 the workforce (max successor share below 0.30), spurious closure when an
 ↳ en-bloc successor
 exists (share at or above 0.70), the ambiguous mid-zone excluded and its
 ↳ mass monitored
 by a pre-registered bimodality check. Cohorts are frozen at November of
 ↳ c-1, so the Schwerdt
 sinking-ship channel cannot condition the sample. Each exposed worker is
 ↳ riskset-matched
 to up to three stable-workplace controls within exact strata (cohort, sex,
 ↳ age band, NACE
 section, size class, region) on Mahalanobis distance over a full year of
 ↳ pre-baseline billing
 history plus tenure and wage decile. The estimator is a stacked matched-set
 ↳ difference-in-differences
 event study with set and cohort-by-calendar-quarter fixed effects,
 ↳ event-workplace clustering,
 a c-2 placebo panel with a numeric failure threshold, and a settlement-lag
 ↳ calibration of
 the fee-week clock.\n\nThree confirmatory quantities form the Holm family,
 ↳ each an honest
 operationalization of a frozen bank hypothesis. θ_1 (H1) is the
 ↳ pre-boundary excess in
 the last two quarters of the terminal year - worded strictly as
 ↳ pre-boundary excess relative
 to an annually dated boundary, never announcement-dated anticipation; the
 ↳ frozen statement's
 'growing as the boundary approaches' rider is a gated exploratory
 ↳ contrast, because its
 ramp-implied floor is verified undetectable at any honest N. θ_{2a} (H2)
 ↳ is the true-closure
 arm's post-year excess over its own matched controls - the excess the
 ↳ frozen statement
 asserts exists 'through the first post-closure year' - while the
 ↳ statement's comparative
 quantifier is enforced by a registered falsification gate: the
 ↳ spurious-closure arm's excess
 is estimated, bounded, and equipped with a pre-committed disconfirmation
 ↳ trigger (spurious
 excess flagged positive and at least half the true-arm excess disconfirms
 ↳ H2's comparative
 reading regardless of θ_{2a}), with the strict differential reported as an
 ↳ exploratory
 estimate. This restructuring follows the verified Stage-0 diagnosis: a
 ↳ differential whose
 variance is dominated by a spurious arm one-third the size of the true arm
 ↳ needs roughly
 5.5 times any honest variance budget - it is a falsification device, not an
 ↳ estimation target.
 θ_3 (H3) is the out-of-hours versus daytime relative-excess gap,
 ↳ unchanged in construct;
 what changed is only its power basis, recalibrated from a self-declared
 ↳ over-conservative
 3% out-of-hours base rate to the conservative end of the published Danish
 ↳ range (simulated
 at 8.0% of daytime GP volume = 7.4% of all GP contacts, against a published
 ↳ 8-12% contact
 share; Christensen & Olesen 1998; the LV-KOS 2011 studies), with an
 ↳ outcome-neutral check

verifying the realized share sits in the published range before the
 → confirmatory run.\n\nPower
 is secured by honest enlargement, not by floor inflation: every floor
 → effect is unchanged
 (2% relative excess for theta1 and theta2a, below the Hamad-Modrek-Cullen
 → and Kuhn-Lalive-Zweimüller
 magnitudes; a 5-point composition gap for theta3), and the sampling plan
 → moves from the
 first registration's self-handicapped half-of-lower-bound to the full
 → derived expectation's
 lower bound - 15 cohorts, size cap 1000, roughly 135,000 matched sets - on
 → a published closure-rate
 basis, with 1:3 matching cutting control-side variance by a third. The
 → license for that
 sizing is a demotion-only realized-N gate: after matching and before any
 → post-baseline outcome
 is touched, the deposited simulation is re-run blind on realized set
 → counts, realized match
 ratios, baseline moments, and the realized out-of-hours share; any member
 → short of 0.95
 power at its floor is demoted to exploratory on the spot, the Holm family
 → shrinks by the
 registered fixed-point rule, and nothing can ever move the other way. The
 → claim ceiling
 is associational and binding - associations and consistency language only,
 → no effect/impact/cause
 vocabulary, no phenomenon-discovery or Danish-first framing, outcomes
 → always utilization
 volume, fees, and timing composition, never health states - and the
 → protocol publishes the
 records of both power runs, the demotion trail, and the remediation note
 → itself, so the
 registration's history is part of its evidence.'

Frozen scope and vocabulary

claim_class_ceiling: associational
 statement: 'Associational - matched event-study contrasts, a contrastive
 → falsification gate,
 and a composition contrast; descriptive for trajectory panels and the
 → bounded priority claim.
 Binding vocabulary rules: 'is associated with' / 'consistent with' /
 → 'suggests' only;
 'effect', 'impact', 'causes' prohibited throughout; anticipation
 → always worded as
 'pre-boundary excess' relative to the annually dated closure-year
 → boundary, never as announcement-dated
 anticipation (S8); no phenomenon-discovery or Danish-first framing (S1-S3);
 → H2 novelty confined
 to the placebo's application to health billing (S7). H2 verdict semantics
 → are fixed in
 advance: 'consistent with H2' requires BOTH the confirmatory core
 → (theta2a Holm-significant,
 predicted direction) AND a passing falsification gate F2; the strict
 → comparative magnitude
 is only ever reported as an exploratory estimate with its CI. Exploratory
 → results (H1_trend,
 H2_diff, E1, E3, E4, and any gate-demoted member) carry the word
 → 'exploratory' in the

heading and in every sentence stating them. The prose label never exceeds
↳ the evidence label.'

frozen_parameters:

- seed: 20260831
- gp_code_list: daytime GP = SPEC2 in {70..80}
- ooh_code_list: out-of-hours GP = SPEC2 in {81, 82, 83, 84, 89}
- ooh_robustness_classifier: adds YDLTID in {2, 3, 4, 5, 7, 8, 9}
- spec2_evening_gp_disclosure: 'Registered choice, disclosed: the parsed
↳ value labels mark
SPEC2 71 ("General practitioner, consultation, evening etc."), 73
↳ ("General practitioner,
telephone consultation, evening etc."), and 75 ("General practitioner,
↳ home visit, evening
etc.") as explicitly off-schedule GP services; they nonetheless remain on
↳ the daytime
side of the PRIMARY classifier because H3's out-of-hours construct and
↳ its published
power basis are keyed to the laegevagt out-of-hours service (SPEC2 81-84,
↳ 89), the quantity
the published Danish 8-12% share describes. Attenuation direction, both
↳ readings: counting
evening GP as daytime attenuates theta3 toward zero - conservative
↳ against the positive
H3 conclusion, anti-conservative for the null-bounding "compositionally
↳ unremarkable"
reading, whose Stage-2 statement therefore carries a mandatory caveat.
↳ In-environment
validation must verify that the YDLTID-augmented robustness classifier
↳ captures SPEC2
71/73/75 rows (their evening time codes 2-5); if YDLTID proves
↳ unpopulated on those rows,
the robustness classifier - and only the robustness classifier - is
↳ amended mechanically
and with disclosure to include SPEC2 71/73/75 directly. The SPEC2
↳ 71/73/75 volume share
is reported, and both classifiers appear side by side in Table T7.'
- successor_share_thresholds: 'true closure: ENDNEDL final closure AND
↳ worker-flow max-successor
share $s(w) < 0.30$; spurious closure: forward identity via
↳ IDFREEM/REGELFR/FRELFREM AND
 $s(w) \geq 0.70$; ambiguous mid-zone excluded; GRAFT-3 sensitivities
↳ 0.20/0.80 and 0.40/0.60'
- quarter_mapping: 'ISO week of HONUGE: 1-13 / 14-26 / 27-39 / 40-52; ISO
↳ week 53 to Q4'
- winsorization: person-quarter outcomes at the 99.9th percentile within era
↳ x calendar quarter
- matching: 1:3 nearest-neighbor Mahalanobis within exact strata, without
↳ replacement across
sets, target $k = 3$, minimum $k = 1$; calipers $|dif \text{ baseline contacts}| \leq 3$
↳ and $|dif \log(1+BRUHON)|$
 ≤ 0.5 ; exposed processed in seeded random order (seed 20260831)
- software: R ≥ 4.3 , fixest::feols (OLS) and fixest::fepois (Poisson QML
↳ robustness); versions
frozen in the registered script
- note: Code lists and thresholds frozen before any data access; code labels
↳ quoted verbatim
from parsed DST value labels and re-validated in-environment before
↳ estimation (frozen
fallback rules in contingent_decisions).

Governance and source status

Original audit remains minor_revisions; packaging is not acceptance.

Source fidelity and implementation order

This is a format conversion of DS1.1, not a new study or independent acceptance. Source audit: minor_revisions. No observed Danish outcomes are supplied. The source execution plan explicitly states that final prereg.yaml governs disagreements. Exact current YAML selections below therefore govern older execution-plan excerpts; these excerpts are retained as implementation detail, not a competing registration. See docs/provenance/source-discrepancies.md. The complete shared data/validity contracts are embedded in Atoms 1–3; read those native protocols when executing a downstream Atom, even without docs/. Do not substitute a scientific default for an unresolved source contradiction.

Preserve source stages 0 -> 1 -> 2 -> 3 -> 4 -> 5 inside the implementation. Source Stage-0 simulations are historical design evidence, not study outcomes; do not silently redo or improve them. Gate 14 must be evaluated from baseline data before post-baseline outcomes. All surviving Stage-4 quantities form ONE Holm family. The three question-level Atoms may share the same frozen family fit and tidy results; do not apply Holm separately per Atom or adjudicate a member before the complete surviving family's p-values exist. Cache/reuse identical governed intermediate results with provenance when available; a direct-dispatch call must obtain or compute its exact source-defined inputs, not pretend its predecessors succeeded. Stage-5 internal riders remain exploratory, after the Stage-4 family is frozen, even when housed in H1/H2.

Reference files under docs/execution/ are the ORIGINAL SKELETONS, not tested turnkey implementations. Adapt paths and complete engineering in a governed workspace; preserve source scientific rules and disclose unresolved choices. Write a result record with component estimates, uncertainty, family eligibility, source interpretation row, diagnostics and boundary/skip dispositions. Released outputs follow source disclosure rules; confidential rows stay governed. Local native states summarize the source judgment; neither a state name nor source power forecasts prove equivalence or causal identification. A bounded claim requires its actual interval and registered rule. Retain unresolved source wording conflicts; do not convert mere nonsignificance into scientific null.

Disclosure

```
environment: DST Forskningservice (no prior data access; gatekeeper
↳ certification per RR
    secondary-data rules). All analysis occurs inside the environment; only
    ↳ aggregate outputs
    leave. No microdata, workplace identifiers (ARBNR/CVRNR/LBNR), or
    ↳ person-level values are
    released.
minimum_cell_size: 5
minimum_cell_rules: no released statistic based on fewer than 5 persons or
↳ fewer than 3 workplaces;
```

```

event-count tables rounded to the nearest 10; weekly trajectory figures
↳ pool >= 100 events
and >= 5,000 person-weeks per plotted point; fee (BRUHON) aggregates
↳ subject to a 50% single-workplace
dominance check; min-cell suppression applied BEFORE Holm reporting so
↳ suppression can never
be outcome-selective
outputs_leaving_environment:
- model coefficients with SEs/CIs
- matching balance tables and realized-k distribution
- calibration and benchmark tables (coverage, linkage, HONUGE lag, OOH share)
- figure aggregates (weekly trajectory series under the pooling rules)
- event counts rounded to the nearest 10
- Stage-0 power/MDE tables - BOTH rounds (failed and remediated) - and the
↳ Stage-2 blind re-simulation
output with the demotion trail
- the c-2 placebo, gate F2 outputs, GRAFT-2/4 diagnostics, GRAFT-6
↳ later-presence rates, E4
attrition panel, and all outcome-neutral check outputs (the full audit
↳ trail exits under
the same rules)
- the analysis and power scripts
statement: 'All analysis occurs inside DST Forskningservice; only aggregate
↳ outputs leave.
Rules applied to every released object: (1) no statistic based on fewer
↳ than 5 persons or
fewer than 3 workplaces; (2) event-count tables rounded to the nearest 10;
↳ (3) weekly trajectory
figures pool >= 100 events and >= 5,000 person-weeks per plotted point; (4)
↳ fee (BRUHON)
statistics released only as cell means/coefficients over min-N cells, with
↳ a dominance check
that no single workplace contributes > 50% of any released fee aggregate;
↳ (5) released objects
are limited to: model coefficients with SEs/CIs, balance tables,
↳ calibration/benchmark tables,
figure aggregates, event counts (rounded), the Stage-0 (both rounds) and
↳ Stage-2 power/MDE
tables with the demotion trail, and the analysis/power scripts - never
↳ microdata, workplace
identifiers (ARBNR/CVRNR/LBNR), or person-level values; (6) min-cell
↳ suppression is applied
before Holm reporting so that suppression can never be outcome-selective;
↳ (7) the c-2 placebo,
gate F2 outputs, the GRAFT-2/4 diagnostics, the GRAFT-6 later-presence
↳ rates, the E4 attrition
panel, and all outcome-neutral check outputs are released under the same
↳ rules, ensuring
the full audit trail exits the environment.'
```

Remediation history

```

round1_void:
date: '2026-08-31'
voided_by: 'driver (power gate: NO_CONFIRMATORY_FAMILY - fixed-point
↳ iteration demoted all
tests; H2 0.358, H1_trend 0.380, H3 0.614, H1 0.840 alone)'
diagnosis: 'Conservative design underpowered at conservative low-end N: (a)
↳ H2''s spurious'
```

arm (~22.5k pairs) cannot detect the 2pp differential floor and needs
 ↳ ~5.5x the variance
 budget; (b) H3 was simulated at a deliberately conservative 3% OOH base
 ↳ rate versus the
 realistic 8-10% of GP volume - at 10% its z roughly doubles and clears;
 ↳ (c) H1 alone needs
 ~108k pairs (~136k under Holm), attainable at the FULL derived
 ↳ expectation of 150k-280k
 exposed workers rather than the conservative half.'

remediation_directive:

- 'AUTHORITY: re-operationalize the confirmatory structure; the frozen bank
 ↳ hypothesis WORDING
 is untouchable, but which statements carry confirmatory tests, and how
 ↳ each is operationalized,
 is open.'
- 'NO MANUFACTURED POWER: effect floors may NOT be inflated without a cited
 ↳ published basis;
 N assumptions may grow only through honest, defensible enlargements
 ↳ (cohort span, ANTNOV
 floor, sector/age-band inclusion, or sizing at the full derived
 ↳ expectation IF a published
 closure-rate basis supports it and a realized-N kill floor is registered
 ↳ at Stage 1).'
- 'H3 OOH base rate: recalibrate to a PUBLISHED Danish out-of-hours share
 ↳ of GP contact
 volume (cite the source in the power basis); 3% was self-declared
 ↳ over-conservative.'
- 'H2 restructure suggestion (designer''s choice): the spurious-closure
 ↳ placebo''s scientific
 role is falsification, not effect estimation - consider registering it as
 ↳ a bounded falsification
 gate (spurious-arm excess must be statistically bounded below a stated
 ↳ fraction of the
 true-closure excess, TOST-style) OUTSIDE the Holm family, with the H2
 ↳ statement''s confirmatory
 weight carried by the true-closure post-year excess it already contains,
 ↳ or demoting the
 differential to exploratory honestly.'
- If no confirmatory family honestly reaches 0.95, return registrable:false
 ↳ - the program
 substitutes runner-up DS1.3 rather than registering a vacuous or inflated
 ↳ protocol.
- 'HARD DATA SCOPE unchanged: only IDAN/IDAS/SSSY/SYSI (inventory in the
 ↳ entry''s HARD_DATA_SCOPE
 field); boundary directives S1-S10 in the entry''s BOUNDARY_DIRECTIVES
 ↳ field remain binding.'

round1_power_summary:

H1: 0.84
 H2_differential: 0.358
 H1_trend: 0.38
 H3: 0.614

verdict: NO_CONFIRMATORY_FAMILY - fixed-point iteration demoted all tests;
 ↳ registration
 voided by the driver audit 2026-08-31; full round-1 record (design,
 ↳ simulation, verification)
 retained in the program's stage-4 design file and published with the
 ↳ protocol

remediation_note: 'TRANSPARENT REMEDIATION RECORD (to be carried in the
 ↳ protocol verbatim).

WHAT FAILED STRUCTURALLY IN ROUND 1: the first Stage-0 simulation (seed
↳ 20260831, deposited
script, verified by independent re-run) emptied the confirmatory family via
↳ the registered
fixed-point rule - H2 0.358, H1_trend 0.380, H3 0.614, H1 0.840, all below
↳ the 0.95 NHB
target - and the audit voided the resulting all-exploratory registration.
↳ The diagnosis
identified four structural causes, none of them the hypotheses themselves:
↳ (1) the sampling
plan self-handicapped to HALF the lower bound of its own schema-derived
↳ expectation (75k
pairs vs 150k-280k expected) with 1:1 matching, although H1 needs ~108k
↳ pairs alone - attainable
at the full expectation; (2) H2's confirmatory quantity was the
↳ true-minus-spurious differential,
whose variance is dominated by the structurally thin spurious arm (~30-60%
↳ of the true arm)
and which needs ~5.5x any honest variance budget - infeasible at every N
↳ the registers can
supply, i.e., a construct-level failure, not a sample-size accident; (3)
↳ H3's power was
simulated at a 3% out-of-hours base rate that the design itself labeled
↳ deliberately over-conservative,
against a published Danish range of 8-12%; (4) H1's frozen growth rider
↳ was carried as
a near-confirmatory gated trend test at a ramp-implied 1.71% floor it can
↳ never detect.

WHAT WAS RE-OPERATIONALIZED AND WHY (frozen statement wording untouched
↳ throughout): (a)
H1's confirmatory estimand remains θ_1 , the statement's core assertion
↳ (elevated contact
in the two pre-boundary quarters); the strict 'growing as the boundary
↳ approaches' quantifier
is tiered exploratory as the gated H1_trend, disclosed as such, because its
↳ only confirmatory
operationalization is verified infeasible. (b) H2's confirmatory weight
↳ moves - exactly
as the audit directive suggests - onto the true-closure post-year excess
↳ the statement already
contains (θ_2 , the full four-quarter post-year window operationalizing
↳ 'through the
first post-closure year'), while the comparative quantifier ('larger and
↳ more persistent'
than spurious closures) is enforced by falsification gate F2 outside the
↳ Holm family: the
spurious-arm excess is estimated and bounded, with a pre-committed
↳ disconfirmation trigger
that caps H2's verdict regardless of θ_2 , and the strict differential
↳ is reported as
exploratory H2_diff with its CI. A falsification role needs bounding power,
↳ not detection
power, which the spurious arm can honestly supply; its operating
↳ characteristics are verified
in the Stage-0 re-run (0.886 pass under the hypothesis-true DGP, 0.978
↳ under a strictly
inert placebo; 0.79 trigger under the placebo-not-inert world at
↳ lower-bound spurious N)
and re-verified blind at realized N in Stage 2. No sign-ambiguous construct
↳ remains: under

the hypothesis-true world every registered confirmatory plim is strictly
 → positive, and every
 reversal has a pre-committed disconfirming interpretation. (c) H3's
 → construct is unchanged
 (it was already the repaired relative-excess contrast under which the Hamad
 → total-down/00H-up
 world is scoreable support); only its power basis is corrected to the
 → published 00H share,
 simulated at the conservative end (8%) and policed by a new outcome-neutral
 → published-range
 check plus realized-share re-simulation. WHY POWER IS NOW HONEST RATHER
 → THAN MANUFACTURED:
 no floor effect was raised - theta1 and theta2a keep the 2% floor (below
 → Hamad/Modrek/Cullen
 2015 and Kuhn/Lalive/Zweimüller 2009), theta3 keeps the 5-point floor; the
 → overdispersion
 split of the DGP is retained unchanged (not tuned upward); every N lever is
 → an audit-authorized
 honest enlargement (sizing at the full derived expectation's LOWER bound
 → on the published
 Erhvervsdemografi/business-demography and Browning-Heinesen basis; cohort
 → span extended
 to 2023, which the register windows support; ANTNOV cap raised to Design
 → B's 1000; 1:3
 riskset matching from a control pool orders of magnitude larger than the
 → exposed cohort);
 and the entire enlargement is licensed only jointly with a demotion-only
 → pre-outcome realized-N
 gate at Stage 2 (binding N* thresholds from the re-run deposited
 → simulation; provisional
 100k/60k+15k/90k sets) so that optimism about N can never buy a
 → confirmatory label that
 realized data do not support. By the verified diagnosis's own scaling
 → arithmetic, all three
 family members clear 0.95 at the registered plan (theta1: ~108k 1:1-pair
 → requirement falls
 ~25-33% under 1:3 matching against ~135k expected sets; theta2a: 4-quarter
 → averaging beats
 theta1 at equal N even with a registered serial-correlation discount;
 → theta3: z scales by
 ~sqrt(8/3) = 1.63 from the base-rate correction alone) - and the re-run of
 → the deposited
 script at Stage 0 is the binding verification, with a registered stop rule
 → that if the fixed-point
 iteration empties the family again, this candidate is declared
 → non-registrable and the program
 substitutes the runner-up rather than registering a vacuous or inflated
 → protocol. WHY THIS
 REMAINS AN HONEST TEST OF THE FROZEN HYPOTHESES: every frozen assertion is
 → either tested
 confirmatorily at unchanged floors (H1's pre-boundary excess; H2's
 → post-year excess with
 its comparative claim falsifiable by a gate that can only subtract support;
 → H3's composition
 tilt), or explicitly tiered exploratory with the infeasibility of its
 → confirmatory operationalization
 stated in the protocol (H1's growth rider; H2's strict differential
 → magnitude and persistence
 ordering) - nothing is quietly dropped, re-worded, or scored on a construct
 → whose plim under

the hypothesis-true world is ambiguous. The records of both power runs and
 ↳ this note are
 published with the registration.'

driver_reconciliation:
 date: '2026-08-31'
 basis: 'power_simulation verdict: registrable provided the F2
 ↳ characteristic targets are
 restated to the verified values and binding N*_H1 recorded as ~118,000
 ↳ sets rather than
 the provisional 100,000'

actions:
 - F2 operating characteristics restated from aspirational $\geq 0.90/\geq 0.95$ to
 ↳ VERIFIED 0.886
 pass / 0.79 trigger (0.978 inert reference) at lower-bound spurious N,
 ↳ with the Stage-2
 blind re-simulation at realized N registered as the binding
 ↳ re-verification; F2 explicitly
 outside the Holm family, no 0.95 requirement
 - N*_H1 binding realized-N gate recorded as 118,000 matched sets,
 ↳ superseding the provisional
 100,000
 - frozen bank hypothesis wording untouched; deterministic reconciliation
 ↳ per W4.7/F10.8
 precedent

Unchanged source timeline

anticipated_completion: 'Approximately 24 months from DST Forskningservice
 ↳ project approval:
 Stage 0 registration freeze and deposited-simulation re-run months 0-3 (no
 ↳ data access);
 Stage 1 event extraction and classification months 3-8 (IDAS + IDAN only);
 ↳ Stage 2 cohort
 construction, matching, and the pre-outcome realized-N gate months 8-12;
 ↳ Stage 3 outcome-neutral
 validation months 12-15; Stage 4 confirmatory estimation (frozen code,
 ↳ single run) months
 15-18; Stage 5 exploratory tier months 18-21; Stage-2 report and disclosure
 ↳ clearance months
 21-24. Support and access: DST Forskningservice project application
 ↳ prepared; the gatekeeper
 certifies no prior data access by any team member.'

stages:
 - stage: Stage 0 - Remediated registration freeze and re-run of the deposited
 ↳ power simulation
 (no data access; DST Forskningservice certifies no prior access)
 work: 'Freeze code lists, thresholds, seeds (20260831), and the full
 ↳ analysis script; re-run
 the deposited Stage-0 simulation (same DGP: baseline ~1.1
 ↳ contacts/person-quarter, 53.5%
 zero share, event ICC ~0.015, persistent/transient overdispersion split
 ↳ unchanged) under
 the REMEDIATED parameters: 1:3 matched sets, sampling plan at the full
 ↳ derived expectation's
 lower bound ($\geq 135,000$ true sets / $\geq 45,000$ spurious sets, published
 ↳ closure-rate basis),
 OOH base rate 8% of GP volume (published basis, conservative end),
 ↳ confirmatory family
 {theta1, theta2a, theta3} at unchanged floors (2%, 2%, 5-point), F2 gate
 ↳ operating characteristics

```

(verified Stage-0 values 0.886 pass / 0.79 trigger at lower-bound
↳ spurious N; re-verified
blind at realized N), and a registered serial-correlation discount on
↳ theta2a's 4-quarter
averaging. Output fixes the binding N* demotion thresholds (provisional:
↳ 100k / 60k+15k
/ 90k) and the realized-MDE machinery. Publish both the first-round
↳ (failed) and remediated
power runs with the protocol.'
stop_rule: STOP (do not register this candidate; the program substitutes
↳ the runner-up)
    if any confirmatory member fails 0.95 at its floor at the registered
    ↳ sampling plan and
    the registered fixed-point demotion would empty the family again. Floors
    ↳ may not be raised
    and N may not be asserted beyond the published-basis expectation to pass
    ↳ this gate.
- stage: Stage 1 - Event extraction and classification (IDAS + IDAN only; no
↳ outcome register
touched)
work: Build the workplace event panel for cohorts 2009-2023, compute
↳ worker-flow successor
shares s(w), classify true/spurious/ambiguous, validate ARB_SEKTORKODE
↳ presence and 2008+
coverage in the IDAN extract (frozen fallback in contingent_decisions),
↳ run the closure-rate
plausibility check against the published sizing basis, the classification
↳ cross-validation,
and the GRAFT-4 s(w) bimodality check.
stop_rule: STOP the project if true-closure events < 500. DROP the
↳ spurious-arm machinery
    (gate F2, theta2b, H2_diff - disclosed before any outcome access, H2
    ↳ verdict capped) if
    spurious events < 150, classification agreement < 75%, or bimodality
    ↳ fails unresolved.
    Report projected exposed N against the sampling plan; a projection below
    ↳ the smallest
    N* already forecasts (but does not yet execute) demotion.
- stage: Stage 2 - Cohort construction, 1:3 matching, and the pre-outcome
↳ realized-N gate
    (IDAN covariates + SSSY pre-boundary years c-3..c-1 only (no year-c or
    ↳ later rows))
work: 'Extract frozen November c-1 cohorts, build control pools, execute
↳ 1:3 matching with
    the coarsening ladder, produce balance tables and the realized k
    ↳ distribution; then run
    the BLIND realized-N gate: re-run the deposited simulation on realized
    ↳ set counts, realized
    k, baseline-year (c-1) moments, and the realized control OOH share - no
    ↳ post-baseline
    outcome touched. Publish the realized-power/MDE table.'
stop_rule: 'STOP if match rate < 90% after the full ladder or any
↳ matching-covariate SMD
    > 0.10 persists; matched sets < 10,000 stops the project; spurious sets <
    ↳ 3,000 drops
    the spurious-arm machinery. GATE (demotion-only, binding): any
    ↳ confirmatory member whose
    realized sets < N* or whose blind re-simulated power < 0.95 at its floor
    ↳ is demoted to

```

```

    exploratory NOW, before any outcome access; the Holm family shrinks by
    ↪ the registered
    fixed-point iteration; no promotion of anything, ever.'
- stage: Stage 3 - Outcome-neutral validation (no registered estimand
  ↪ computed)
work: Coverage benchmark vs published DST series, linkage floor,
  ↪ floor/ceiling checks, HONUGE
  lag calibration via BEHANDLINGSDATO, seasonality sanity, published-range
  ↪ OOH share benchmark,
  and the c-2 placebo panel.
stop_rule: STOP if the coverage benchmark misses  $\pm 10\%$  without documented
  ↪ reconciliation,
  the linkage floor (80%) fails, or the c-2 placebo fails after the single
  ↪ pre-specified
  re-match (which kills the registered event-study tier per the kill
  ↪ criteria, downgrading
  to descriptive trajectories only).
- stage: Stage 4 - Confirmatory estimation (frozen code, single run)
work: Estimate the surviving Holm family (theta1, theta2a, theta3 as gated
  ↪ at Stage 2) with
  Holm correction; evaluate gate F2 and its pre-committed trigger; compute
  ↪ the gated BRUHON
  fee secondary if theta2a rejects and the gated exploratory H1_trend if
  ↪ theta1 rejects;
  run the registered sensitivity suite and the GRAFT-6 later-presence
  ↪ trigger; produce the
  pre-committed boldface Summary verdict paragraph per hypothesis
  ↪ ('Consistent with' / 'Contrary
  to'); convert every null into CI-ruled-out magnitudes against realized
  ↪ MDEs.
stop_rule: 'No discretionary stop: all surviving registered tests are run
  ↪ and reported exactly
  once regardless of results (Stage-2 RR principle: decisions never based
  ↪ on outcomes);
  demoted members are reported under the exploratory heading with their
  ↪ demotion trail.'
- stage: Stage 5 - Exploratory tier (segregated by heading)
work: H2_diff differential and quarter-by-quarter persistence profile; E1
  ↪ SYSI-era replication
  (1992-2003, legacy keys, no age/sex matching, VAGTOMR/SIKREAMT disclosed
  ↪ as inventory-only
  variables); weekly-resolution trajectory figures under the GRAFT-2 step
  ↪ diagnostic; E3
  fees and heterogeneity (Holm within family as severity aid); E4 attrition
  ↪ accounting panel
  by event class.
stop_rule: Skip E1 entirely (reported as skipped, not failed) if SYSI-era
  ↪ event extraction
  yields < 300 true closures or the era's coverage benchmark fails; no
  ↪ exploratory result
  can revise any confirmatory verdict or Stage-4 language.

```

Registered interpretation rules

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

1. restriction

Condition: TRUE

Associational — matched event-study contrasts, a contrastive falsification gate, and a composition contrast; descriptive for trajectory panels and the bounded priority claim. Binding vocabulary rules: ‘is associated with’ / ‘consistent with’ / ‘suggests’ only; ‘effect’, ‘impact’, ‘causes’ prohibited throughout; anticipation always worded as ‘pre-boundary excess’ relative to the annually dated closure-year boundary, never as announcement-dated anticipation (S8); no phenomenon-discovery or Danish-first framing (S1–S3); H2 novelty confined to the placebo’s application to health billing (S7). H2 verdict semantics are fixed in advance: ‘consistent with H2’ requires BOTH the confirmatory core (theta2a Holm-significant, predicted direction) AND a passing falsification gate F2; the strict comparative magnitude is only ever reported as an exploratory estimate with its CI. Exploratory results (H1_trend, H2_diff, E1, E3, E4, and any gate-demoted member) carry the word ‘exploratory’ in the heading and in every sentence stating them. The prose label never exceeds the evidence label.

```
kind: restriction
condition: 'TRUE'
says: 'Associational - matched event-study contrasts, a contrastive
↳ falsification
gate, and a composition contrast; descriptive for trajectory panels and
the bounded priority claim. Binding vocabulary rules: ''is associated with''
/ ''consistent with'' / ''suggests'' only; ''effect'', ''impact'',
↳ ''causes''
prohibited throughout; anticipation always worded as ''pre-boundary excess''
relative to the annually dated closure-year boundary, never as
↳ announcement-dated
anticipation (S8); no phenomenon-discovery or Danish-first framing (S1-S3);
H2 novelty confined to the placebo's application to health billing (S7).
H2 verdict semantics are fixed in advance: ''consistent with H2'' requires
BOTH the confirmatory core (theta2a Holm-significant, predicted direction)
AND a passing falsification gate F2; the strict comparative magnitude is
only ever reported as an exploratory estimate with its CI. Exploratory
↳ results
(H1_trend, H2_diff, E1, E3, E4, and any gate-demoted member) carry the word
''exploratory'' in the heading and in every sentence stating them. The prose
label never exceeds the evidence label.'
```

2. restriction

Condition: state(atom-000003) == “descriptive_only”

The source placebo kill permits descriptive trajectories only; all event-study quantities remain unrun, not null.

```
kind: restriction
condition: state(atom-000003) == "descriptive_only"
says: The source placebo kill permits descriptive trajectories only; all
↳ event-study
quantities remain unrun, not null.
```

3. primary

Condition: done(atom-000004)

Use atom-000004 component results and its exact embedded source interpretation rows. Apply its boundary, eligibility, multiplicity and claim ceilings before quoting a scientific pattern; done does not imply support.

```
kind: primary
condition: done(atom-000004)
says: Use atom-000004 component results and its exact embedded source
↪ interpretation
  rows. Apply its boundary, eligibility, multiplicity and claim ceilings
  ↪ before
  quoting a scientific pattern; done does not imply support.
```

4. primary

Condition: done(atom-000005)

Use atom-000005 component results and its exact embedded source interpretation rows. Apply its boundary, eligibility, multiplicity and claim ceilings before quoting a scientific pattern; done does not imply support.

```
kind: primary
condition: done(atom-000005)
says: Use atom-000005 component results and its exact embedded source
↪ interpretation
  rows. Apply its boundary, eligibility, multiplicity and claim ceilings
  ↪ before
  quoting a scientific pattern; done does not imply support.
```

5. primary

Condition: done(atom-000006)

Use atom-000006 component results and its exact embedded source interpretation rows. Apply its boundary, eligibility, multiplicity and claim ceilings before quoting a scientific pattern; done does not imply support.

```
kind: primary
condition: done(atom-000006)
says: Use atom-000006 component results and its exact embedded source
↪ interpretation
  rows. Apply its boundary, eligibility, multiplicity and claim ceilings
  ↪ before
  quoting a scientific pattern; done does not imply support.
```

6. supplementary

Condition: done(atom-000007)

Use atom-000007 component results and its exact embedded source interpretation rows. Apply its boundary, eligibility, multiplicity and claim ceilings before quoting a scientific pattern; done does not imply support.

```

kind: supplementary
condition: done(atom-000007)
says: Use atom-000007 component results and its exact embedded source
↪ interpretation
  rows. Apply its boundary, eligibility, multiplicity and claim ceilings
    ↪ before
    quoting a scientific pattern; done does not imply support.

```

7. supplementary

Condition: done(atom-000008)

Use atom-000008 component results and its exact embedded source interpretation rows. Apply its boundary, eligibility, multiplicity and claim ceilings before quoting a scientific pattern; done does not imply support.

```

kind: supplementary
condition: done(atom-000008)
says: Use atom-000008 component results and its exact embedded source
↪ interpretation
  rows. Apply its boundary, eligibility, multiplicity and claim ceilings
    ↪ before
    quoting a scientific pattern; done does not imply support.

```

8. supplementary

Condition: done(atom-000009)

Use atom-000009 component results and its exact embedded source interpretation rows. Apply its boundary, eligibility, multiplicity and claim ceilings before quoting a scientific pattern; done does not imply support.

```

kind: supplementary
condition: done(atom-000009)
says: Use atom-000009 component results and its exact embedded source
↪ interpretation
  rows. Apply its boundary, eligibility, multiplicity and claim ceilings
    ↪ before
    quoting a scientific pattern; done does not imply support.

```

9. residual

Condition: TRUE

Retain every member, including source-killed, unactivated, boundary and operationally failed units. Report remaining source ambiguities and unobserved results explicitly.

```

kind: residual
condition: 'TRUE'
says: Retain every member, including source-killed, unactivated, boundary
      and operationally failed units. Report remaining source ambiguities and
      unobserved results explicitly.

```

Preregistered epistemic graph

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

atom-000001

Can true and administrative workplace closures be distinguished within the registered four-table scope?

Activation: TRUE

- **classifiable:** At least 500 true-closure events; source-required classification audits resolved. Preserve spurious-arm kill or availability as a component, without killing the true arm. — The true-closure event frame is usable within the audited classification scope; no health result exists.
- **project_stop:** Fewer than 500 true-closure events after the registered Stage-1 procedure. — Source project viability kill; do not estimate the registered outcome study.
- **boundary:** Required source-defined inputs or scientific definitions are demonstrably unavailable; retain available components and the exact unresolved scope. — An evidenced boundary, not a statistical null. A missing implementation or timeout alone is operational no_result, not this state.
- **otherwise:** otherwise — The registered judgment is unresolved; report all completed components without inventing support, equivalence or absent results.

atom-000002

Can the registered matched cohorts and pre-outcome inferential eligibility be established?

Activation: `state(atom-000001) == "classifiable"`

- **cohort_ready:** At least 10,000 true matched sets, registered balance/matching reconciled, and the baseline-only demotion trail archived. This may include zero surviving confirmatory members. — A matched cohort is available; the component-level demotion trail, not this state, determines confirmatory eligibility.
- **project_stop:** Fewer than 10,000 true matched sets after the source procedure. — Source project viability kill; no outcome study under this registration.
- **boundary:** Required source-defined inputs or scientific definitions are demonstrably unavailable; retain available components and the exact unresolved scope. — An evidenced boundary, not a statistical null. A missing implementation or timeout alone is operational no_result, not this state.
- **otherwise:** otherwise — The registered judgment is unresolved; report all completed components without inventing support, equivalence or absent results.

atom-000003

Do billing coverage, timing calibration and the registered placebo permit event-study interpretation?

Activation: `state(atom-000002) == "cohort_ready"`

- **descriptive_only**: The c-2 placebo fails the final YAML condition after the one registered re-match; other data needed for the source descriptive trajectory are usable. — The entire event-study tier is killed; only the source-permitted descriptive trajectories remain.
- **event_study_ready**: Registered coverage/linkage/timing checks resolved, and the final YAML c-2 placebo failure condition is not met after the permitted procedure. Carry H3 kill, F2 availability, warnings and demotions separately. — Event-study contrasts are eligible only at their registered component-specific inferential status; no causal identification is established.
- **boundary**: Required source-defined inputs or scientific definitions are demonstrably unavailable; retain available components and the exact unresolved scope. — An evidenced boundary, not a statistical null. A missing implementation or timeout alone is operational `no_result`, not this state.
- **otherwise**: otherwise — The registered judgment is unresolved; report all completed components without inventing support, equivalence or absent results.

atom-000004

Is primary-care billing elevated before the annual closure boundary, and is its registered timing pattern observed?

Activation: `state(atom-000003) == "event_study_ready" OR state(atom-000003) == "descriptive_only"`

- **descriptive_only**: Shared validity judgment is `descriptive_only`. — Produce only the already specified descriptive trajectories, not H1 or H1_trend estimates.
- **exploratory**: H1 was demoted by the pre-outcome gate and an interpretable source estimate is available. — Report exploratory estimates without confirmatory scoring; do not promote the internal trend rider.
- **preboundary_excess**: Eligible $\theta_1 > 0$ and its two-sided Holm test rejects. — Pre-boundary billing excess only. Apply the source H1_trend gate internally; distinguish supported growth from unresolved growth.
- **reversed**: Eligible $\theta_1 < 0$ and its two-sided Holm test rejects. — The pre-boundary decline contradicts H1; it is not support.
- **not_supported**: Eligible H1 is estimated but its Holm test does not reject. — Report the actual CI and realized MDE; no confirmation or automatic equivalence claim.
- **boundary**: Required source-defined inputs or scientific definitions are demonstrably unavailable; retain available components and the exact unresolved scope. — An evidenced boundary, not a statistical null. A missing implementation or timeout alone is operational `no_result`, not this state.
- **otherwise**: otherwise — The registered judgment is unresolved; report all completed components without inventing support, equivalence or absent results.

atom-000005

Do true closures show post-boundary billing excess, and how does the administrative-closure comparison constrain it?

Activation: `state(atom-000003) == "event_study_ready"`

- **exploratory**: H2 confirmatory core was demoted pre-outcome; source-defined components are estimable. — Report an exploratory core with all F2/arm restrictions; no confirmation.
- **comparative_disconfirmed**: Registered F2 trigger fires, regardless of the theta2a verdict. — The comparative H2 reading is disconfirmed; record the core and exploratory differential separately.
- **reversed**: Eligible $\text{theta2a} < 0$ and Holm rejects, with F2 not already determining the first-matching state. — Post-year billing declines; report attrition bounds and do not reframe as support.
- **comparative_unavailable**: Eligible $\text{theta2a} > 0$ and Holm rejects, but the registered comparative arm/gate is unavailable or bounds-only. — At most partially consistent; comparative quantifier untestable.
- **postboundary_excess**: Eligible $\text{theta2a} > 0$ and Holm rejects, F2 is evaluable and does not fire, and registered threshold signs are stable. — Quote the actual spurious-arm bound, never its absence. H2_diff remains exploratory.
- **not_supported**: Eligible core does not Holm-reject and no earlier region applies. — Report both arms with their actual bounds; retain the source non-inert-placebo row where applicable.
- **boundary**: Required source-defined inputs or scientific definitions are demonstrably unavailable; retain available components and the exact unresolved scope. — An evidenced boundary, not a statistical null. A missing implementation or timeout alone is operational no_result, not this state.
- **otherwise**: otherwise — The registered judgment is unresolved; report all completed components without inventing support, equivalence or absent results.

atom-000006

Is peri-event relative billing excess tilted toward out-of-hours rather than daytime primary care?

Activation: `state(atom-000003) == "event_study_ready"`

- **not_estimable**: The source H3-specific classifier or rare-OOH kill criterion is established. — Composition is not estimable under the registered construct; not a statistical null.
- **exploratory**: H3 was demoted pre-outcome and a valid source estimate is available. — Report exploratory composition with the primary/robustness classifier caveat.
- **ooh_tilt**: Eligible $\text{theta3} > 0$ and Holm rejects. — Composition claim only: retain the $\text{theta_day} \geq 0$ versus < 0 source interpretations; no diagnosis or health-state claim.
- **reversed**: Eligible $\text{theta3} < 0$ and Holm rejects. — A relative daytime tilt contradicts H3.

- **not_supported:** Eligible H3 is estimated but its Holm test does not reject. — Report the actual interval and classifier caveat; nonsignificance alone does not establish equivalence.
- **boundary:** Required source-defined inputs or scientific definitions are demonstrably unavailable; retain available components and the exact unresolved scope. — An evidenced boundary, not a statistical null. A missing implementation or timeout alone is operational `no_result`, not this state.
- **otherwise:** otherwise — The registered judgment is unresolved; report all completed components without inventing support, equivalence or absent results.

atom-000007

How does the source H1 analogue appear in the separate SYSI-era cohort?

Activation: `state(atom-000003) == "event_study_ready" AND done(atom-000004) AND done(atom-000005) AND done(atom-000006)`

- **exploratory_pattern:** The source-defined module yields interpretable estimates/descriptives after its own eligibility checks; retain unavailable components explicitly. — Report the complete source exploratory pattern without confirmatory or causal promotion.
- **source_skip:** The source-defined optional module/arm is unavailable or its registered skip criterion is met. — Record the source skip and its reason; never invent a negative or null finding.
- **boundary:** Required source-defined inputs or scientific definitions are demonstrably unavailable; retain available components and the exact unresolved scope. — An evidenced boundary, not a statistical null. A missing implementation or timeout alone is operational `no_result`, not this state.
- **otherwise:** otherwise — The registered judgment is unresolved; report all completed components without inventing support, equivalence or absent results.

atom-000008

How do source-defined fee and subgroup contrasts vary within the registered exploratory tier?

Activation: `state(atom-000003) == "event_study_ready" AND done(atom-000004) AND done(atom-000005) AND done(atom-000006)`

- **exploratory_pattern:** The source-defined module yields interpretable estimates/descriptives after its own eligibility checks; retain unavailable components explicitly. — Report the complete source exploratory pattern without confirmatory or causal promotion.
- **source_skip:** The source-defined optional module/arm is unavailable or its registered skip criterion is met. — Record the source skip and its reason; never invent a negative or null finding.
- **boundary:** Required source-defined inputs or scientific definitions are demonstrably unavailable; retain available components and the exact unresolved scope. — An

evidenced boundary, not a statistical null. A missing implementation or timeout alone is operational no_result, not this state.

- **otherwise:** otherwise — The registered judgment is unresolved; report all completed components without inventing support, equivalence or absent results.

atom-000009

How much within-year separation and later-presence selection limits the pre-boundary interpretation?

Activation: `state(atom-000003) == "event_study_ready" AND done(atom-000004) AND done(atom-000005) AND done(atom-000006)`

- **exploratory_pattern:** The source-defined module yields interpretable estimates/descriptives after its own eligibility checks; retain unavailable components explicitly. — Report the complete source exploratory pattern without confirmatory or causal promotion.
- **source_skip:** The source-defined optional module/arm is unavailable or its registered skip criterion is met. — Record the source skip and its reason; never invent a negative or null finding.
- **boundary:** Required source-defined inputs or scientific definitions are demonstrably unavailable; retain available components and the exact unresolved scope. — An evidenced boundary, not a statistical null. A missing implementation or timeout alone is operational no_result, not this state.
- **otherwise:** otherwise — The registered judgment is unresolved; report all completed components without inventing support, equivalence or absent results.

Audit Report — Stage-1 Registered Report Protocol

Candidate: DS1.1 FINAL VERDICT: minor_revisions Round-1 worst verdict: minor_revisions | Re-audit verdict: minor_revisions (improvement capped at one step) Unresolved fatal items: 0

Dimension scores (0-3)

Round-1 medians:

```
{
  "question_consequence": 2,
  "gap_validity": 3,
  "construct_validity": 2,
  "design_discrimination": 2,
  "claim_calibration": 2,
  "uncertainty": 3,
  "mechanism": 2,
  "generality": 3,
  "reproducibility": 3,
  "story_integrity": 3,
  "figure_table_integrity": 2,
  "implications": 3
}
```

Final (re-audit):

```
{
  "question_consequence": 2,
  "gap_validity": 3,
  "construct_validity": 3,
  "design_discrimination": 2,
  "claim_calibration": 3,
  "uncertainty": 3,
  "mechanism": 2,
  "generality": 3,
  "reproducibility": 3,
  "story_integrity": 3,
  "figure_table_integrity": 3,
  "implications": 3
}
```

NHB Stage-1 criteria failed by majority (round 1)

none

Fatal flags raised (round 1)

none

Findings (round 1, all auditors)

- [minor] (AU1) report.tex, Section 7 ‘Method and data-generating process’ (remediated element 4); cf. design table RQ3 row and power_simulation.py MU_OOH: Numeric inconsistency in the H3 power basis: the prose says the OOH base rate was recalibrated ‘to 8.8% of daytime GP volume’, but the deposited script sets MU_OOH = 0.088 contacts/person-quarter = 8.0% of the 1.1 daytime base (std-out prints ‘8.0% of daytime GP volume’, ‘7.4% of all GP contacts’), while the design table says ‘simulated at 8% of GP contact volume’ against a published range described as 8-12% of contact volume. — FIX: Correct ‘8.8%’ to ‘8.0% of daytime GP volume (7.4% of all GP contacts)’ and harmonize the design-table phrase so the simulated share is stated on the same denominator as the published 8-12% range.
- [minor] (AU1) report.tex design table, RQ2 interpretation column; prereg.yaml interpretations (H2 pass pattern): The pre-committed H2 pass conclusion — the excess ‘does not appear where workplace identity dies only on paper’ — can be printed under a registered pass state in which theta2b’s CI excludes zero positively but $\theta_{2b} < 50\%$ of θ_{2a} (gate passes by the 50% clause), and F2’s verified sensitivity to a floor-sized non-inert placebo is only 0.79; the frozen template can therefore assert an absence the gate has not established. — FIX: Reword the pass-state template to a bounded comparative — e.g., ‘appears at no more than half the true-arm magnitude (spurious-arm 95% upper bound: X)’ — using the theta2b one-sided upper bound the protocol already commits to reporting beside theta2a.
- [minor] (AU1) prereg.yaml power.dgp / expected_n; report.tex Section 7 (sampling remediation lever 1); power_simulation.py N_SP: The simulation’s 45,000 spurious sets are labeled ‘the 30% lower bound of the 30-60% band’, but the protocol’s

own accounting (spurious workers = 30% of 150,000 true workers = 45,000 workers before matching, then the 90% match-rate floor) implies ~40,500 sets; F2's registered operating characteristics (pass 0.886, trigger 0.79) are therefore computed at a ~10% flattered lower-bound N. — FIX: Either restate 45,000 as the pre-matching worker count and recompute/footnote F2 characteristics at 40,500 sets, or explicitly note the ~10% overstatement and lean on the already-registered Stage-2 blind re-verification at realized spurious N as the binding value.

- [minor] (AU1) report.tex Section 10 Timeline, Stage 2 heading; prereg.yaml timeline.stages (Stage 2) vs population.inclusion (ALDERIMP lookup) and contingent_decisions (placebo re-match): Stage 2's data scope is stated as 'IDAN covariates and SSSY baseline year c-1 only', but the age/sex inclusion rule reads ALDERIMP/KOENIMP from the most recent SSSY row in c-3..c-1, and the pre-specified placebo re-match adds year-c-2 baseline contacts to the Mahalanobis set — both require pre-baseline SSSY years beyond c-1. — FIX: Restate the Stage-2 SSSY scope as 'pre-boundary years c-3..c-1 only (no year-c or later rows)' — this preserves the pre-outcome blindness guarantee while matching the registered rules.
- [minor] (AU1) report.tex Introduction (Exploratory tier item, line ~103) vs Analysis plan 'Multiplicity' subsection; prereg.yaml analysis_plan.multiple_comparisons: The Introduction describes 'Five heading-segregated exploratory quantities, Holm-corrected within family as a severity aid', but the registered multiplicity rule applies within-family Holm only to (H2_diff; E1; E3; E4), with H1_trend gated and reported at unadjusted alpha=0.05 (as simulated in the joint-power computation). — FIX: Harmonize the Introduction sentence to exclude H1_trend from the within-family Holm claim (or state its gate-plus-unadjusted convention explicitly).
- [minor] (AU1) report.tex design table RQ1 interpretation column; prereg.yaml interpretations (H1 pass pattern): The H1 pass template names only 'anticipatory care-seeking under impending job loss'; because layoffs occur during terminal year c itself, a pre-boundary excess in Q3-Q4(c) can equally reflect realized early separations among the frozen c-1 cohort — an alternative the protocol quantifies only in the exploratory, post-treatment-conditioned E4 panel and does not name in the pre-committed conclusion. — FIX: Add one clause to the H1 pass template ('...or realized within-terminal-year separations, per the E4 accounting') so the frozen conclusion names the nearest alternative reading alongside anticipation.
- [major] (AU2) report.tex Section 6 check 7 and Section 5 contingent rule 6 / Section 7 H1 kill criterion; prereg.yaml outcome_neutral_checks item 7 vs kill_criteria H1: The load-bearing c-2 placebo gate has two contradictory registered conditions: the outcome-neutral check passes only if the placebo is 'not significant at alpha=0.05 AND point estimate <= 1% of control baseline mean', while the re-match/kill trigger fires only if it is 'significant at alpha=0.05 AND |estimate| > 1%'. These are not logical complements: a non-significant placebo at 1.5% of baseline (or a significant one at 0.5%) fails the check but does not meet the kill trigger, leaving the pipeline's action undefined — discretion in exactly the gate that kills the event-study tier. — FIX: Restate the check threshold, contingent rule 6, and the H1 placebo-kill criterion as one identical condition (e.g., fail iff significant at alpha=0.05

AND $|\text{estimate}| > 1\%$ of the control baseline mean), and state explicitly what happens in the two non-complementary cells.

- [major] (AU2) report.tex Introduction (frozen H1) and design-table H1 interpretation; prereg.yaml interpretations H1 and design threats item 1: theta1 is an intention-to-follow contrast on the November c-1 cohort, so workers separated in early year c (the Schwerdt wave that E4 only describes) contribute post-job-loss utilization to the ‘pre-boundary’ Q3/Q4(c) window. A Holm-significant theta1 can therefore be generated wholly by realized early job loss with zero anticipation — a mundane alternative the threat list does not name (it lists only compositional attrition bias). The pre-committed conclusion nonetheless foregrounds ‘consistent with anticipatory care-seeking under impending job loss’, and the frozen clause ‘before any separation is visible in annual IDAN records’ is false at the individual level for early leavers, whose separation is visible in the year-c IDAN record. — FIX: Add ‘billing responses to separations already realized within the terminal year’ as a named threat with direction; amend H1’s pre-committed interpretation sentence to read ‘consistent with anticipatory care-seeking and/or responses to early within-year separations, with the E4 year-c separation share bounding the latter channel’, and optionally register an exploratory theta1 split by November-year-c attachment labeled post-treatment-conditioned.
- [major] (AU2) report.tex design table RQ2 interpretation column; prereg.yaml interpretations H2 (full-success pattern): The pre-committed H2-confirmed sentence asserts an excess ‘that does not appear where workplace identity dies only on paper’. Gate F2 passes whenever theta2b’s CI excludes zero positively but $\text{theta2b} < 50\%$ of theta2a — i.e., with a spurious-arm excess that demonstrably DOES appear at up to 49% of the true-arm size — and its verified sensitivity to a floor-sized non-inert placebo is only 0.79 at plan N (~0.5 at the 15,000-set floor). A passing gate cannot license an absence claim; the pre-committed prose exceeds the evidence class of a bounding device. — FIX: Reword the pre-committed sentence to quote the registered quantity: ‘a post-year excess whose spurious-arm counterpart is bounded below X% of the true-arm excess (one-sided 95% upper bound)’, dropping ‘does not appear’.
- [minor] (AU2) report.tex Section 6 (preamble and check 10) vs Stage-3 description in Section 10; prereg.yaml outcome_neutral_checks item 10: The claim ‘Eleven checks, all orthogonal to the confirmatory hypotheses’ is overstated: the week-zero step diagnostic’s threshold is scaled by ‘the total Q3(c)+Q4(c) excess’ — the theta1 quantity itself — so running it requires computing a version of the confirmatory contrast. The protocol is also ambiguous about when it runs: Table T6 is fed by the ‘Stage-1 through Stage-3 check battery’ while Stage 3 promises ‘no registered estimand computed’. Impact is limited because Stage 4 runs all surviving tests regardless, but the orthogonality claim is factually wrong for check 10. — FIX: Schedule check 10 explicitly in Stage 4/5 (with the weekly descriptive series), and either drop check 10 from the ‘all orthogonal’ claim or rescale its threshold to an absolute contacts-per-week magnitude.
- [minor] (AU2) prereg.yaml interpretations H2 null row; report.tex design table RQ2: The interpretation grid is not exhaustive: the H2 null row is conditioned on ‘theta2b

also near zero’, so the realized pattern theta2a-null with a clearly positive theta2b (placebo arm alone elevated) has no pre-committed conclusion sentence — a discretion leak in a protocol that claims ‘nothing left to choose’. — FIX: Add one row: theta2a not Holm-significant AND theta2b positive/flagged -> pre-committed non-inert-placebo reading with both arms’ CI-ruled-out magnitudes.

- [minor] (AU2) power_simulation.py (SIG2_E=0.0375, ICC ~0.014) and report.tex Section 5 sensitivity subsection: A priori 0.95 power is certified at event-level ICC 0.0140 — the bottom of the design’s own stated plausibility band (0.01-0.05) — and the only sensitivity scenarios vary N and effect size, never the variance components. With H1 at 0.980 and a 13% N margin, power at the top of the band would plausibly fall below 0.95. The Stage-2 blind re-simulation on realized baseline moments mitigates this pre-outcome, but the Stage-1 certification is conditional on an unverified ICC and does not say so. — FIX: Add one registered scenario at ICC ~0.05 (and/or higher transient variance) to the sensitivity table, or state explicitly that the 0.95 certification is conditional on the band’s lower half pending the blind re-simulation.
- [minor] (AU2) report.tex Section 2 (Registers; clocks subsection) and Section 5; prereg.yaml data section: Three small internal inconsistencies: (a) SSSY extraction is 2007-2025 but the linkage floor and age/sex lookup use c-3..c-1/c+1 windows that reach 2006 for cohort 2009; (b) ‘ENDNEDL validity through 2023 caps the last cohort’ is asserted although the public HKV parse lists ENDNEDL as ‘Still valid’ (01-01-1980-), and unlike the four disclosed inventory-only variables the 2023 cap’s source is not flagged; (c) prose states the H3 power basis as ‘8% of GP contact volume’ while the script implements 8% of daytime volume = 7.4% of all contacts (conservative, but the wordings differ). — FIX: One sentence each: state the cohort-2009 window truncation rule, source the ENDNEDL 2023 cap to the analysis-server inventory explicitly, and unify the OOH base-rate wording to ‘8% of daytime GP volume (7.4% of all GP contacts)’.
- [major] (AU3) report.tex §5 ‘Preprocessing, in order’ item 1 and §2 SSSY bullet; prereg.yaml frozen_parameters (gp_code_list/oooh_code_list); stage4_design.json analysis_plan.preprocessing[0]: The frozen classifier ‘daytime GP = SPEC2 codes 70-80’ contradicts the parsed value labels the protocol claims to quote verbatim: data/schema_en_opus/value_labels.csv labels SPEC2 71 ‘General practitioner, consultation, evening etc.’, 73 ‘telephone consultation, evening etc.’, and 75 ‘home visit, evening etc.’. H3’s primary classifier therefore counts explicitly off-schedule GP care as daytime, attenuating theta3 (anti-conservative for the pre-committed null-bounding conclusion ‘compositionally unremarkable’), and drifts from the frozen H3 statement, which keys the construct to ‘out-of-hours and off-schedule services (YDLTID time-of-day codes)’ — YDLTID being demoted to a robustness classifier. — FIX: Before freeze, re-derive the code lists from the parsed labels: either move 71/73/75 to the off-schedule side (or a disclosed third ‘evening GP’ category), or keep them daytime with an explicit sentence stating the labels, the reason (lægevagt-keyed published power basis), and the attenuation direction for both the positive and the null reading; verify the YDLTID-augmented classifier catches 71/73/75 and report both classifiers side by side in Table T7.

- [major] (AU3) report.tex §2 IDAN bullet (‘ARB_SEKTORKODE (workplace sector, 2008+) applies the private-sector filter’) and §2 Population; prereg.yaml data.registers IDAN variables + variable_verification_note: ARB_SEKTORKODE — the private-sector eligibility filter for the entire confirmatory tier — is absent from the DS1.1 entry’s recorded Forskningsservice four-register inventory (stage4_design.json entry.registers, IDAN line) and appears in no analysis-server inventory extract anywhere in the harness outputs, unlike LOENTIMER [2008-2024], BEHANDLINGSDATO [2021-2025], and YDERNR [2005-2025], which sibling DS-wave inventories confirm. It is publicly documented (TIMES ida-ansaettelser) but the protocol’s verification note does not list it among the inventory exceptions, its claimed ‘2008+’ validity has no recorded source, and no contingent decision exists if it is absent in-environment — an unexecutable inclusion filter would stall Stage 2 with nothing registered to do about it. — FIX: Add ARB_SEKTORKODE to the per-variable verification-exception list (public-parse-verified, server presence unconfirmed) and register a frozen fallback: validate at Stage 1; if absent, define the private-sector filter from ARB_HOVED_BRA_DB07 sections (public-administration/health/education sections excluded), with the substitution and its count effect reported verbatim.
- [minor] (AU3) report.tex abstract, §1 final paragraph, §7 Verification (‘Both power runs ... are published with the protocol’); power_simulation.py header: ‘Both power runs are published with the protocol’ overstates the deposit: the deposited power_simulation.py is the remediated version that, per its own header, ‘overwrites the script per the remediation directive’; the voided round-1 run survives only as recorded stdout, parameters, and results inside stage4_design.json, so round 1 is documented but not re-runnable from the deposit bundle. — FIX: Deposit the round-1 script from git history alongside the remediated one (e.g., power_simulation_round1.py), or restate the sentence as ‘both power runs’ records’ throughout.
- [minor] (AU3) report.tex §2 SSSY bullet (‘extraction 2007-2025’) and §5 Preprocessing item 6, vs Population (ALDERIMP lookup c-3..c-1) and Linkage (80% floor over c-3..c+1): The registered SSSY extraction window 2007-2025 truncates the c-3 lookback for the first cohort (c=2009, c-3=2006, but SSSY rows exist from 2005): the ALDERIMP/KOENIMP lookup window and the 80% linkage-floor window are both shortened for that cohort without disclosure, mechanically inflating the unknown-age/sex stratum and slightly depressing the measured linkage rate in 2009. — FIX: Start extraction at 2006 (rows exist from 2005), or add one sentence disclosing the one-cohort truncation and excluding cohort 2009 from the linkage-floor denominator’s c-3 year.
- [minor] (AU3) report.tex \date field (‘git commit [COMMIT]’) and prereg.yaml git_commit: DRIVER_STAMPS_COMMIT: The commit-of-record placeholders are unstamped at audit time, so the ‘frozen at the same git commit’ claim binding the protocol, registration, and scripts together is not yet verifiable — acceptable per the stated driver workflow, but the freeze is not evidenced until stamped. — FIX: Stamp the commit into both files immediately after this audit, as the workflow already prescribes, and confirm the stamped commit contains the deposited scripts.

- [minor] (AU3) report.tex §10 Disclosure control and ethics; §11 Timeline (Stages 1-4 access scoping): The staged data-access discipline (outcome registers untouched until Stage 4; ‘blind’ Stage-2 gate on baseline-year SSSY only) is presented adjacent to the Forskningservice gatekeeper role, but DST certifies only no prior access and output disclosure control — all granted registers are simultaneously available on the server, so the intra-project staging is self-enforced, not gatekeeper-enforced; a reader could over-credit the blinding. — FIX: Add one sentence distinguishing gatekeeper-certified no-prior-access from self-enforced staging, and name the enforcement mechanism (e.g., logged extract scripts or separate project user authorizations per stage).

Revision round

Applied: - AU2 major (placebo gate contradiction): unified the c-2 placebo gate into one registered condition - FAILS iff significant at $\alpha=0.05$ AND $|\text{estimate}| > 1\%$ of the control baseline mean - stated identically in report.tex Section 3 falsification item 2, Section 5 contingent rule 7, Section 6 check 7, and the Section 7 H1 placebo kill, and in prereg.yaml falsification_tests, contingent_decisions, outcome_neutral_checks item 7, and both H1 kill-criterion entries; both non-complementary cells (significant at $\leq 1\%$; non-significant at $> 1\%$) are pre-committed as passes with mandatory disclosure (T7 placebo-row flag / caution paragraph against the theta1 floor and realized placebo MDE), no re-match, no kill, no discretion. - AU2 major (H1 early within-year separations): added the named threat ‘billing responses to separations already realized within the terminal year’ with direction (inflates theta1; frozen clause holds at cohort-anchoring level but not for individual early leavers) to report.tex Section 3 threats and prereg.yaml design.threats; amended the pre-committed H1-confirmed interpretation in the design table and prereg.yaml interpretations to ‘consistent with anticipatory care-seeking and/or billing responses to separations already realized early within the terminal year, with the E4 year-c separation share bounding the latter channel’; registered the optional post-treatment-conditioned exploratory theta1 split by November-year-c attachment inside E4 (intro exploratory tier, design table, Table T10, yaml E4 estimand/primary_quantity); Limitations/assumptions carry the corollary honestly. - AU2 major (H2 ‘does not appear’ overclaim): reworded the pre-committed H2 full-success conclusion in the design table and prereg.yaml interpretations to the registered bounding quantity - ‘a post-boundary billing excess ... whose spurious-arm counterpart is bounded below X% of the true-arm excess, where X is the realized one-sided 95% upper bound on theta2b as a share of theta2a-hat (a passing gate bounds the spurious-arm excess - it cannot license an absence claim)’ - and made the claim-evidence-matrix entries in both artifacts match; ‘does not appear’ dropped everywhere it stated the H2 claim. - AU3 major (SPEC2 71/73/75 evening GP): applied the audit’s keep-daytime-with-disclosure option, preserving the design’s frozen code lists: report.tex preprocessing item 1 and prereg.yaml frozen_parameters.spec2_evening_gp_disclosure now quote the three parsed ‘evening etc.’ labels verbatim, state the reason (H3 construct and published 8-12% power basis are keyed to the laegevagt service, SPEC2 81-84/89), state the attenuation direction for both the positive and the null reading, require

in-environment verification that the YDLTID-augmented classifier captures 71/73/75 (with a mechanical robustness-side-only amendment if YDLTID is unpopulated), report the 71/73/75 volume share, and put both classifiers side by side in Table T7; the H3-null pre-committed conclusion and Section 2 SSSY bullet and design-table RQ3 cell carry the matching caveat in both artifacts. - AU3 major (ARB_SEKTORKODE unverified in inventory): added the converse verification exception (publicly documented on TIMES ida-ansættelser, absent from the recorded Forskningservice inventory extract, 2008+ validity resting on the public parse alone) to report.tex Section 2 (intro note and IDAN bullet) and prereg.yaml (per-variable verification field and variable_verification_note); registered the frozen fallback as a new contingent rule in both artifacts (private-sector filter from ARB_HOVED_BRA_DB07, NACE DB07 sections O/P/Q excluded, substitution and count effect reported verbatim, no other filter change permitted); Stage-1 timeline work now includes the presence/coverage validation, and the failure-modes/Limitations passages name the fallback. Unfixable (design-level, recorded in Limitations): Recompiled: true

Re-audit findings

- [minor] report.tex, Section 5 ‘Method and data-generating process’ (line 287): The remediated OOH power basis is misstated as ‘recalibrated from the self-declared over-conservative 3% to 8.8% of daytime GP volume’. The registered value is $\mu_{OOH} = 0.088$ contacts/person-quarter on a 1.1 base = 8.0% of daytime volume, as stated in prereg.yaml power.dgp, power_simulation.py (MU_OOH comment), the deposited calibration stdout, and everywhere else in report.tex (Sections 3, 4, 5-verification); 8.8% would also contradict ‘the conservative end of the published Danish range’.
- [minor] prereg.yaml, hypotheses[H2].kill_criterion and kill_criteria[H2]: Both H2 kill-criterion strings still gloss the demotion threshold as ‘realized true matched sets $< N^*_{H2a}$ (provisional 60,000)’, while the binding value is 95,000 (yaml power.binding_n_star; report.tex Section 7 ‘binding value 95,000’). H1’s parallel string self-updates (‘provisional 100,000; binding value from the Stage-0 re-run’); H2’s does not. The blind re-simulated-power clause backstops any misreading, but a Stage-2 executor reading kill_criteria alone could apply 60,000.
- [minor] outputs/stage4_design.json (binding design) vs the revised registration: The revision round’s fixes were applied to report.tex and prereg.yaml only; stage4_design.json still carries the superseded content: the contradictory placebo condition (its outcome-neutral check requires ‘Not significant AND $\leq 1\%$ ’ while its contingent decision fails on ‘significant AND $> 1\%$ ’), the H2 ‘does not appear’ conclusion and claim-matrix line, the single-channel H1-confirmed interpretation, E4 without the theta1 attachment split, and no ARB_SEKTORKODE contingency. The yaml’s remediation_record documents the driver reconciliation but not these audit-round deltas, leaving a gap in the otherwise complete provenance chain.
- [minor] report.tex Section 2 (IDAS bullet, register-clocks subsection) and prereg.yaml data.registers[IDAS].years: ‘ENDNEDL ... validity through 2023 caps the last cohort’ is stated as fact but is not confirmable from the cited public parses: hkv_variables_bilingual.csv lists ENDNEDL valid from 1980 with valid_to ‘Still

valid’, and the cached TIMES page shows value labels valid to 31-12-3000 with no 2023 endpoint. Unlike the four inventory-only variables and ARB_SEKTORKODE, this claim carries no per-variable provenance note. Any error is bounded (a longer validity merely forgoes cohort 2024; a shorter one would be caught by the per-cohort closure-rate plausibility check).

- [minor] report.tex Population section and preprocessing item 6; prereg.yaml population.inclusion and data.registers[SSSY].years: SSSY extraction is registered as 2007-2025, but the age/sex lookup window (‘most recent SSSY row in c-3..c-1’) and the linkage-floor window (c-3..c+1) reference year 2006 for the first cohort (c = 2009), which lies outside the registered extraction even though SSSY rows exist from 2005. Practically immaterial (the ‘most recent row’ will come from 2007-2008 and the 80% floor is tolerant), but the registered windows and the registered extraction are formally inconsistent for one cohort.
- [minor] report.tex \NullOneText vs prereg.yaml hypotheses[H1].null_substantive: The H1 substantive null differs typographically between artifacts: tex has ‘\emph{decrease}’, ‘(deferred or forgone care’, ‘Hamad, Modrek and Cullen’; yaml has ‘DECREASE’, ‘(deferred/forgone care’, ‘Hamad, Modrek & Cullen’. The frozen H1/H2/H3 statements themselves are byte-identical (verified programmatically), and the H2/H3 nulls match; only H1’s null drifts.

Execution supplement

Source reference code — not turnkey

The three source code files are byte-identical to DS1.1. They contain useful functions and pseudocode, but missing runtime loaders and unfinished steps. No R execution or scientific/synthetic reanalysis was performed in conversion. Do not execute an unfilled skeleton as if it were a complete implementation. The native Atom protocols, not the code, own scientific definitions.

Atom	Source entry points / coverage	Required inputs and outputs
atom-000001	extraction_skeleton.R: build_identity_chains, assign_primary_jobs, compute_successor_shares, classify_events	IDAN/IDAS; interim/events_classified.parquet, source T1/T2/F3

Atom	Source entry points / coverage	Required inputs and outputs
atom-000002	extraction_skeleton.R: extract_exposed_cohort, build_control_pool, extract_ssy_prebaseline, build_matching_frame, run_matching, run_blind_gate; power_simulation.py	Events + baseline SSSY only; matched sets, realized_power_mde, demotion trail; T3–T5/F5
atom-000003	extraction_skeleton.R: extract_ssy_outcomes, aggre- gate_person_quarter, assemble_panels, calibrate_honuge_lag, run_c2_placebo_gate	Frozen roster and permitted SSSY; panels, coverage, lag, placebo, later-presence records; T6
atom-000004	analysis_skeleton.R: fit_event_study, compute_theta1, holm_adjudicate; H1_trend block	Complete surviving-family inference + panels; theta1/trend, F1/F2, T7–T9
atom-000005	analysis_skeleton.R: compute_theta2a, evaluate_gate_F2, holm_adjudicate; H2_diff/fee blocks	True/spurious panels and eligibility; theta2a/theta2b/F2/fee/differential, F1, T7–T9
atom-000006	analysis_skeleton.R: compute_theta3, holm_adjudicate	Source classifiers, full family and bootstrap; composition and robustness, F4/T7/T8
atom-000007	extraction_skeleton.R: e1_legacy_pipeline; analysis_skeleton.R E1 block	Legacy IDAN/IDAS/SYSI only; E1 panel, contrast, F6/T9 or source skip
atom-000008	analysis_skeleton.R exploratory E3 block	Frozen eligible panels; source fee/subgroup family, T9
atom-000009	analysis_skeleton.R exploratory E4 block	November attachment, later presence; source attrition/conditional theta1 split, T10

Reuse a shared governed family fit for Atoms 4–6; Holm is never recomputed on one Atom alone. When direct dispatch is selected, obtain/compute the exact registered prerequisites and preserve pre-outcome ordering. Missing code does not prohibit implementing an unambiguous protocol, but it does not authorize resolving source-silent scientific choices. Test completed adapters in their runtime workspace; no ready synthetic-test command is claimed here.

Complete registered Molecule

```
id: molecule-000001
molecule_mode: prereg
molecule_shape: chain
topic: Closures and primary care
rating: C
draft: false
membership:
- atom: atom-000001
  role: secondary
  description: Can true and administrative workplace closures be distinguished
    within the registered four-table scope?
- atom: atom-000002
  role: secondary
  description: Can the registered matched cohorts and pre-outcome inferential
    eligibility be established?
- atom: atom-000003
  role: secondary
  description: Do billing coverage, timing calibration and the registered
    placebo permit event-study interpretation?
- atom: atom-000004
  role: primary
  description: Is primary-care billing elevated before the annual closure
    boundary, and is its registered timing pattern observed?
- atom: atom-000005
  role: primary
  description: Do true closures show post-boundary billing excess, and how
    does the administrative-closure comparison constrain it?
- atom: atom-000006
  role: primary
  description: Is peri-event relative billing excess tilted toward
    ↪ out-of-hours
    rather than daytime primary care?
- atom: atom-000007
  role: supplementary
  description: How does the source H1 analogue appear in the separate SYSI-era
    cohort?
- atom: atom-000008
  role: supplementary
  description: How do source-defined fee and subgroup contrasts vary within
    the registered exploratory tier?
- atom: atom-000009
  role: supplementary
  description: How much within-year separation and later-presence selection
    limits the pre-boundary interpretation?
literature: []
frontier:
  rule: 'Whether primary-care contact already rises while a workplace is dying
    - and whether the post-closure excess attaches to genuinely destroyed
    jobs rather than to workplaces that die only on paper - has never been
```

measured at register scale with matched controls at sub-annual resolution. Three precise gaps, each verified open by the recorded 62-query search below: (i) no study estimates administrative primary-care utilization trajectories before establishment closure at population scale (the
→ construct-level
ancestors are one factory and one shipyard; the finest modern resolution is monthly in one firm); (ii) the worker-flow classification that
→ separates
true establishment deaths from en-bloc absorptions (Benedetto et al. 2007; Hethey-Maier & Schmieder 2013) is used as sample-exclusion hygiene in displacement research, and no published work elevates the absorbed
→ pseudo-closures
to an explicit placebo arm for any health outcome; (iii) no event-based test exists of whether peri-event excess care is compositionally tilted toward out-of-hours, unplanned services. The direction is genuinely
→ contested:
US evidence finds utilization falls under layoff-threat insecurity and after job loss (Hamad, Modrek & Cullen 2015; Schaller & Stevens 2015), the Danish displacement literature contains a published null on
→ stress-related
hospitalizations (Browning, Moller Dano & Heinesen 2006), and the
→ small-scale
closure literature finds consultations rise - so both the null and the reversal are pre-registered as informative outcomes.'

avoid: |-
Associational - matched event-study contrasts, a contrastive
→ falsification gate, and a composition contrast; descriptive for
→ trajectory panels and the bounded priority claim. Binding vocabulary
→ rules: 'is associated with' / 'consistent with' / 'suggests' only;
→ 'effect', 'impact', 'causes' prohibited throughout; anticipation
→ always worded as 'pre-boundary excess' relative to the annually dated
→ closure-year boundary, never as announcement-dated anticipation (S8);
→ no phenomenon-discovery or Danish-first framing (S1-S3); H2 novelty
→ confined to the placebo's application to health billing (S7). H2
→ verdict semantics are fixed in advance: 'consistent with H2' requires
→ BOTH the confirmatory core (theta2a Holm-significant, predicted
→ direction) AND a passing falsification gate F2; the strict
→ comparative magnitude is only ever reported as an exploratory
→ estimate with its CI. Exploratory results (H1_trend, H2_diff, E1, E3,
→ E4, and any gate-demoted member) carry the word 'exploratory' in the
→ heading and in every sentence stating them. The prose label never
→ exceeds the evidence label.
Only IDAN, IDAS, SSSY, SYSI. SSSY cohorts 2009-2023; SYSI 1992-2003 is
→ separate and exploratory. Source review: minor_revisions.

directions: []

preregistration:
mode: adaptive
question: 'Whether primary-care contact already rises while a workplace
is dying - and whether the post-closure excess attaches to genuinely
→ destroyed
jobs rather than to workplaces that die only on paper - has never been
measured at register scale with matched controls at sub-annual resolution. Three precise gaps, each verified open by the recorded 62-query search below: (i) no study estimates administrative primary-care utilization trajectories before establishment closure at population scale (the
→ construct-level
ancestors are one factory and one shipyard; the finest modern resolution is monthly in one firm); (ii) the worker-flow classification that
→ separates
true establishment deaths from en-bloc absorptions (Benedetto et al. 2007;

Hethey-Maier & Schmieder 2013) is used as sample-exclusion hygiene in displacement research, and no published work elevates the absorbed

- ↳ pseudo-closures

to an explicit placebo arm for any health outcome; (iii) no event-based test exists of whether peri-event excess care is compositionally tilted toward out-of-hours, unplanned services. The direction is genuinely

- ↳ contested:

US evidence finds utilization falls under layoff-threat insecurity and after job loss (Hamad, Modrek & Cullen 2015; Schaller & Stevens 2015), the Danish displacement literature contains a published null on

- ↳ stress-related

hospitalizations (Browning, Moller Dano & Heinesen 2006), and the

- ↳ small-scale

closure literature finds consultations rise - so both the null and the reversal are pre-registered as informative outcomes.'

scope: |-

Associational - matched event-study contrasts, a contrastive

- ↳ falsification gate, and a composition contrast; descriptive for
- ↳ trajectory panels and the bounded priority claim. Binding vocabulary
- ↳ rules: 'is associated with' / 'consistent with' / 'suggests' only;
- ↳ 'effect', 'impact', 'causes' prohibited throughout; anticipation
- ↳ always worded as 'pre-boundary excess' relative to the annually dated
- ↳ closure-year boundary, never as announcement-dated anticipation (S8);
- ↳ no phenomenon-discovery or Danish-first framing (S1-S3); H2 novelty
- ↳ confined to the placebo's application to health billing (S7). H2
- ↳ verdict semantics are fixed in advance: 'consistent with H2' requires
- ↳ BOTH the confirmatory core (theta2a Holm-significant, predicted
- ↳ direction) AND a passing falsification gate F2; the strict
- ↳ comparative magnitude is only ever reported as an exploratory
- ↳ estimate with its CI. Exploratory results (H1_trend, H2_diff, E1, E3,
- ↳ E4, and any gate-demoted member) carry the word 'exploratory' in the
- ↳ heading and in every sentence stating them. The prose label never
- ↳ exceeds the evidence label.

Only IDAN, IDAS, SSSY, SYSI. SSSY cohorts 2009-2023; SYSI 1992-2003 is

- ↳ separate and exploratory. Source review: minor_revisions.

storyline: |+

This study asks whether Danish workers' primary-care billing already
 → rises while their workplace is dying - before any separation is
 → visible in annual employment records - and whether the post-closure
 → excess attaches to genuinely destroyed jobs rather than to workplaces
 → that die only on paper. The instrument is the weekly clock of Danish
 → health-insurance billing (SSSY, HONUFE fee-week) held against
 → annually dated establishment terminal events from the IDA workplace
 → registers. For closure cohorts 2009-2023 (entirely inside the
 → post-2008 IDA key era), every worker with primary November-status
 → employment in year $c-1$ at a workplace that disappears between years c
 → and $c+1$ is classified by the worker-flow canon of Benedetto et al.
 → (2007) and Hethy-Maier & Schmieder (2013): true closure when no
 → successor absorbs the workforce (max successor share below 0.30),
 → spurious closure when an en-bloc successor exists (share at or above
 → 0.70), the ambiguous mid-zone excluded and its mass monitored by a
 → pre-registered bimodality check. Cohorts are frozen at November of
 → $c-1$, so the Schwerdt sinking-ship channel cannot condition the
 → sample. Each exposed worker is riskset-matched to up to three
 → stable-workplace controls within exact strata (cohort, sex, age band,
 → NACE section, size class, region) on Mahalanobis distance over a full
 → year of pre-baseline billing history plus tenure and wage decile. The
 → estimator is a stacked matched-set difference-in-differences event
 → study with set and cohort-by-calendar-quarter fixed effects, event-wo
 → rkplace clustering, a $c-2$ placebo panel with a numeric failure
 → threshold, and a settlement-lag calibration of the fee-week clock.

Three confirmatory quantities form the Holm family, each an honest
 → operationalization of a frozen bank hypothesis. θ_1 (H1) is the
 → pre-boundary excess in the last two quarters of the terminal year -
 → worded strictly as pre-boundary excess relative to an annually dated
 → boundary, never announcement-dated anticipation; the frozen statement
 → 's 'growing as the boundary approaches' rider is a gated exploratory
 → contrast, because its ramp-implied floor is verified undetectable at
 → any honest N . θ_2 (H2) is the true-closure arm's post-year excess
 → over its own matched controls - the excess the frozen statement
 → asserts exists 'through the first post-closure year' - while the
 → statement's comparative quantifier is enforced by a registered
 → falsification gate: the spurious-closure arm's excess is estimated,
 → bounded, and equipped with a pre-committed disconfirmation trigger (s
 → purious excess flagged positive and at least half the true-arm excess
 → disconfirms H2's comparative reading regardless of θ_2), with the
 → strict differential reported as an exploratory estimate. This
 → restructuring follows the verified Stage-0 diagnosis: a differential
 → whose variance is dominated by a spurious arm one-third the size of
 → the true arm needs roughly 5.5 times any honest variance budget - it
 → is a falsification device, not an estimation target. θ_3 (H3) is
 → the out-of-hours versus daytime relative-excess gap, unchanged in
 → construct; what changed is only its power basis, recalibrated from a
 → self-declared over-conservative 3% out-of-hours base rate to the
 → conservative end of the published Danish range (simulated at 8.0% of
 → daytime GP volume = 7.4% of all GP contacts, against a published
 → 8-12% contact share; Christensen & Olesen 1998; the LV-KOS 2011
 → studies), with an outcome-neutral check verifying the realized share
 → sits in the published range before the confirmatory run.

Power is secured by honest enlargement, not by floor inflation: every

- ↪ floor effect is unchanged (2% relative excess for θ_{11} and θ_{2a} ,
- ↪ below the Hamad-Modrek-Cullen and Kuhn-Lalive-Zweimüller magnitudes;
- ↪ a 5-point composition gap for θ_{33}), and the sampling plan moves
- ↪ from the first registration's self-handicapped half-of-lower-bound to
- ↪ the full derived expectation's lower bound - 15 cohorts, size cap
- ↪ 1000, roughly 135,000 matched sets - on a published closure-rate
- ↪ basis, with 1:3 matching cutting control-side variance by a third.
- ↪ The license for that sizing is a demotion-only realized-N gate: after
- ↪ matching and before any post-baseline outcome is touched, the deposit
- ↪ ed simulation is re-run blind on realized set counts, realized match
- ↪ ratios, baseline moments, and the realized out-of-hours share; any
- ↪ member short of 0.95 power at its floor is demoted to exploratory on
- ↪ the spot, the Holm family shrinks by the registered fixed-point rule,
- ↪ and nothing can ever move the other way. The claim ceiling is associa
- ↪ tional and binding - associations and consistency language only, no
- ↪ effect/impact/cause vocabulary, no phenomenon-discovery or Danish-fir
- ↪ st framing, outcomes always utilization volume, fees, and timing
- ↪ composition, never health states - and the protocol publishes the
- ↪ records of both power runs, the demotion trail, and the remediation
- ↪ note itself, so the registration's history is part of its evidence.

Claim-evidence matrix

```

```yaml
- claim: Workers attached to workplaces in their final pre-closure year
 ↪ show elevated primary-care
 ↪ billing in the two quarters preceding the closure-year boundary
 ↪ (pre-boundary excess)
 claim_class: associational (confirmatory)
 evidence: θ_{11} from the stacked 1:3 matched-set DiD event study (set
 ↪ FE + cohort x calendar-quarter
 ↪ FE, event-workplace clustering), Holm-adjusted within the three-test
 ↪ family, licensed
 ↪ by the passed c-2 placebo panel, the matching-balance table, and the
 ↪ pre-outcome realized-N
 ↪ gate certifying ≥ 0.95 power at the 2% floor
- claim: The pre-boundary excess grows as the boundary approaches (frozen
 ↪ H1's growth rider)
 claim_class: associational (exploratory, gated)
 evidence: H1_trend contrast computed only if θ_{11} Holm-rejects;
 ↪ GRAFT-2 week-0 step diagnostic
 ↪ bounding settlement-lag contamination; verified structurally
 ↪ underpowered for confirmatory
 ↪ status (Stage-0 power 0.380 at the ramp-implied floor) and disclosed
 ↪ as such
- claim: Workers at true closures show a post-boundary billing excess
 ↪ through the first post-closure
 ↪ year, with the spurious-arm counterpart bounded below the registered
 ↪ fraction of the true-arm
 ↪ excess (one-sided 95% upper bound; a passing gate bounds the
 ↪ spurious-arm excess, it cannot
 ↪ license an absence claim)
 claim_class: associational (confirmatory core + registered
 ↪ falsification gate)
 evidence: θ_{2a} (true-arm post-year excess, Holm-adjusted) plus gate
 ↪ F2 (spurious-arm
 ↪ excess bounded, pre-committed disconfirmation trigger, operating
 ↪ characteristics verified

```

- at Stage 0), the GRAFT-3 threshold-sensitivity table, and the
    - ↪ classification cross-validation
  - table; the strict differential magnitude is exploratory H2\_diff with
    - ↪ CI
- claim: The peri-event excess is compositionally tilted toward
  - ↪ out-of-hours services
  - claim\_class: associational (confirmatory, composition)
  - evidence: theta3 relative-excess contrast between separately estimated
    - ↪ OOH and daytime count
    - outcomes with 999-draw cluster-bootstrap CI; power basis recalibrated
      - ↪ to the published
    - Danish OOH share (8-12%; simulated at 8%) and re-verified against the
      - ↪ realized share by
    - the published-range benchmark check; robustness under the
      - ↪ YDLTID-augmented OOH classifier,
    - reported side by side with the primary classifier in Table T7 under
      - ↪ the SPEC2 71/73/75
    - evening-GP disclosure (frozen\_parameters)
- claim: This is the first register-scale, matched, weekly-resolution
  - ↪ estimate of primary-care
    - billing trajectories around workplace closures
  - claim\_class: descriptive (bounded priority claim)
  - evidence: 'Design features (SSSY/SYSI HONUUGE weekly clock, N in the
    - ↪ hundreds of thousands,
    - 52-week billing-history matching) positioned against the recorded
      - ↪ 62-query novelty-boundary
    - search; expressly NOT a phenomenon-discovery or Danish-first claim
      - ↪ (S1-S3: Iversen 1989;
      - Rugulies 2009 precede)'
- claim: Weekly trajectory figures showing the timing shape of the excess
  - ↪ around the boundary
  - claim\_class: descriptive
  - evidence: Pooled exposed-minus-matched-control weekly HONUUGE series
    - ↪ across >= 100 events
    - per figure, disclosure-compliant aggregation, GRAFT-2 step diagnostic
      - ↪ governing permissible
    - timing-shape language; no test attached
- claim: 'This protocol is a remediated second registration: its first
  - ↪ Stage-0 power analysis
    - emptied the confirmatory family, and the re-operationalization (not
      - ↪ the hypothesis wording)
    - is what restored it'
  - claim\_class: methodological disclosure (binding)
  - evidence: Both power-simulation runs and the fixed-point log are
    - ↪ published with the protocol
      - (outputs/prereg\_v0/power\_simulation.py and its remediated re-run);
    - ↪ the remediation\_note
    - is carried in the protocol verbatim; no floor effect was raised, and
      - ↪ every N lever is
    - cited to a published basis with a pre-outcome realized-N kill floor
- claim: Outcomes are utilization volume, fees, and timing composition -
  - ↪ never health states
  - claim\_class: scope statement (binding)
  - evidence: 'Hard data scope: only IDAN, IDAS, SSSY, SYSI exist on the
    - ↪ analysis server - no
    - diagnoses, hospital, or prescription registers; enforced in prose by
      - ↪ the claim-class ceiling'

## 5. EXHIBITS

Every figure and table the report commits to. Disclosure rules binding on EVERY released object: no statistic based on fewer than 5 persons or fewer than 3 workplaces; event-count tables rounded to the nearest 10; weekly trajectory figures pool  $\geq 100$  events and  $\geq 5,000$  person-weeks per plotted point; BRUHON aggregates pass a 50% single-workplace dominance check; released objects are limited to model coefficients with SEs/CIs, balance tables, calibration/benchmark tables, figure aggregates, rounded event counts, the power/MDE tables (both Stage-0 rounds and Stage 2) with the demotion trail, and the analysis/power scripts - never microdata, workplace identifiers (ARENR/CVRNR/LBNR), or person-level values; suppression applied BEFORE Holm reporting.

**\*\*Figures (matplotlib):\*\***

- **\*\*F1 - Quarterly event-study panel.\*\*** Treated x event-quarter coefficients with 95% CIs across  $Q1(c-2)..Q4(c+1)$  (placebo quarters, omitted-baseline marker, terminal-year and post-year quarters); panel (a) true arm, panel (b) spurious arm from the pooled model. Feeds: step-18 panels + 4.1/4.2 coefficient tables. Disclosure: coefficients only.
- **\*\*F2 - Weekly descriptive trajectory.\*\*** Pooled exposed-minus-matched-control weekly GP contact series around the closure-year boundary, with the GRAFT-2 week-0 step diagnostic annotated (step  $\leq 50\%$  of the total  $Q3(c)+Q4(c)$  excess; if exceeded, the figure carries the registered caution and quarter-level estimates are foregrounded). Feeds: step 15/17 weekly aggregates. Disclosure:  $\geq 100$  events and  $\geq 5,000$  person-weeks per point.
- **\*\*F3 - s(w) successor-share distribution (GRAFT-4).\*\*** Histogram of worker-flow successor shares across terminal events with the 0.30/0.70 thresholds and mid-zone mass; bimodality check exhibit. Feeds: step 7. Disclosure: bin counts rounded to 10; bins under min-cell suppressed.
- **\*\*F4 - OOH vs daytime relative excess (theta3).\*\*** The two relative-excess series over the peri-event window with the bootstrap CI of their gap; primary and YDLTID-augmented classifier side by side. Feeds: 4.3 outputs. Disclosure: coefficients only.
- **\*\*F5 - Matching balance (love plot).\*\*** SMDs before/after matching for all matching covariates and unmatched auxiliaries; 0.10 line marked. Feeds: step 13. Disclosure: aggregates only.
- **\*\*F6 - E1 SYSI-era replication (exploratory; skippable).\*\*** Quarterly SYSI-era analogue of F1(a) for cohorts 1992-2003, labeled exploratory, YDLANT-weighted proxy outcome. Feeds: step 22. Disclosure: as F1.

**\*\*Tables (booktabs LaTeX):\*\***

- **\*\*T1 - Event extraction and classification by cohort year.\*\*** Terminal events, true/spurious/ambiguous counts (rounded to 10), annual true-closure rate among private size-5+ workplaces, spurious/true ratio. Feeds: step 8.
- **\*\*T2 - Classification cross-validation.\*\*** Worker-flow classes vs IDAS-native flags (ENDNEDL/IDFREM/REGELFR/FRELFREM/ANDELUDS), full cross-tabulation + agreement rate; GRAFT-3 classification counts under 20/80 and 40/60. Feeds: step 8.
- **\*\*T3 - Cohort flow accounting.\*\*** Inclusion/exclusion counts per rule (tenure, sector, size, SIKGRUP, missing ANSAAR/pay, multi-cohort dedup), eligible exposed per cohort, match rate, ladder rung used, realized k distribution. Feeds: steps 9-13.
- **\*\*T4 - Matching balance.\*\*** Per-covariate SMDs (weighted  $1/k$ ), caliper compliance, realized mean k. Feeds: step 13.
- **\*\*T5 - Power and demotion trail.\*\*** Stage-0 power results (BOTH rounds:

the voided round-1 run and the remediated re-run) and the Stage-2 blind realized-power/MDE table with binding N\*, realized sets, and any demotions with the fixed-point recomputation. Feeds: deposited script + step 14. Published regardless of outcome.

- \*\*T6 - Outcome-neutral check panel.\*\* All eleven registered checks (coverage benchmark by year, linkage floor, closure-rate plausibility, classification agreement, GRAFT-4, balance/match rate, c-2 placebo, HONUGE lag calibration, floor/ceiling + seasonality, GRAFT-2 step, OOH-share benchmark) with thresholds and PASS/FAIL. Feeds: steps 8, 13, 16, 17, 19, 21 + F2/F3 diagnostics.
- \*\*T7 - Confirmatory results.\*\* theta1, theta2a, theta3: point estimates (contacts/person-quarter and % of baseline), 95% CIs, two-sided p, Holm-adjusted verdicts; the F2 gate row (theta2b, its one-sided 95% upper bound, trigger evaluation); the gated BRUHON fee secondary (if computed); c-2 placebo row; realized MDEs beside every null with the CI-ruled-out magnitudes. Feeds: 4.1-4.4.
- \*\*T8 - Sensitivity suite.\*\* GRAFT-3 thresholds; GRAFT-9 forward-window exclusion (with truncation note); un-winsorized; Poisson QML; death/emigration exclusion - one table, run once. Feeds: 4.4.
- \*\*T9 - Exploratory tier.\*\* H1\_trend (if gated in; otherwise the row states the gate was not met), H2\_diff with quarter-by-quarter differential profile, E3 fee and subgroup analyses (within-family Holm), E1 summary. Every row labeled exploratory. Feeds: 4.5.
- \*\*T10 - Attrition and later-presence accounting.\*\* E4 year-c separation shares by event class (ANSXFREM-based, post-treatment-conditioned label and selection caveat) and the GRAFT-6 later-presence rates with the 2pp trigger status. Feeds: steps 20 + 4.5 E4.

---

interpretation:

- kind: restriction
  - condition: 'TRUE'
  - says: 'Associational - matched event-study contrasts, a contrastive
    - ↪ falsification
      - gate, and a composition contrast; descriptive for trajectory panels and the bounded priority claim. Binding vocabulary rules: 'is
        - ↪ associated
        - with' / 'consistent with' / 'suggests' only; 'effect',
          - ↪ 'impact',
          - 'causes' prohibited throughout; anticipation always worded as
            - ↪ 'pre-boundary
            - excess' relative to the annually dated closure-year boundary, never as announcement-dated anticipation (S8); no phenomenon-discovery or Danish-first framing (S1-S3); H2 novelty confined to the placebo's application to health billing (S7). H2 verdict semantics are fixed in advance: 'consistent with H2' requires BOTH the confirmatory core (theta2a Holm-significant, predicted direction) AND a passing
              - ↪ falsification
        - gate F2; the strict comparative magnitude is only ever reported as an exploratory estimate with its CI. Exploratory results (H1\_trend,
          - ↪ H2\_diff,
- E1, E3, E4, and any gate-demoted member) carry the word 'exploratory' in the heading and in every sentence stating them. The prose label never exceeds the evidence label.'

- kind: restriction
- condition: state(atom-000003) == "descriptive\_only"
- says: The source placebo kill permits descriptive trajectories only; all event-study quantities remain unrun, not null.
- kind: primary

```

condition: done(atom-000004)
says: Use atom-000004 component results and its exact embedded source
 interpretation rows. Apply its boundary, eligibility, multiplicity and
 claim ceilings before quoting a scientific pattern; done does not imply
 support.
- kind: primary
condition: done(atom-000005)
says: Use atom-000005 component results and its exact embedded source
 interpretation rows. Apply its boundary, eligibility, multiplicity and
 claim ceilings before quoting a scientific pattern; done does not imply
 support.
- kind: primary
condition: done(atom-000006)
says: Use atom-000006 component results and its exact embedded source
 interpretation rows. Apply its boundary, eligibility, multiplicity and
 claim ceilings before quoting a scientific pattern; done does not imply
 support.
- kind: supplementary
condition: done(atom-000007)
says: Use atom-000007 component results and its exact embedded source
 interpretation rows. Apply its boundary, eligibility, multiplicity and
 claim ceilings before quoting a scientific pattern; done does not imply
 support.
- kind: supplementary
condition: done(atom-000008)
says: Use atom-000008 component results and its exact embedded source
 interpretation rows. Apply its boundary, eligibility, multiplicity and
 claim ceilings before quoting a scientific pattern; done does not imply
 support.
- kind: supplementary
condition: done(atom-000009)
says: Use atom-000009 component results and its exact embedded source
 interpretation rows. Apply its boundary, eligibility, multiplicity and
 claim ceilings before quoting a scientific pattern; done does not imply
 support.
- kind: residual
condition: 'TRUE'
says: Retain every member, including source-killed, unactivated, boundary
 and operationally failed units. Report remaining source ambiguities
 and unobserved results explicitly.

```

The native registrations contain the converted source science. Optional source files preserve provenance, not a competing editable authority.

## Current interpretation

All nine Atoms are pending. No empirical results or independent acceptance are claimed.

## atom-000001: Event classification

### Can true and administrative workplace closures be distinguished within the registered four-table scope?

#### Source fidelity and implementation order

This is a format conversion of DS1.1, not a new study or independent acceptance. Source audit: minor\_revisions. No observed Danish outcomes are supplied. The source execution plan explicitly states that final prereg.yaml governs disagreements. Exact current YAML selections below therefore govern older execution-plan excerpts; these excerpts are retained as implementation detail, not a competing registration. See docs/provenance/source-discrepancies.md. The complete shared data/validity contracts are embedded in Atoms 1–3; read those native protocols when executing a downstream Atom, even without docs/. Do not substitute a scientific default for an unresolved source contradiction.

Preserve source stages 0 -> 1 -> 2 -> 3 -> 4 -> 5 inside the implementation. Source Stage-0 simulations are historical design evidence, not study outcomes; do not silently redo or improve them. Gate 14 must be evaluated from baseline data before post-baseline outcomes. All surviving Stage-4 quantities form ONE Holm family. The three question-level Atoms may share the same frozen family fit and tidy results; do not apply Holm separately per Atom or adjudicate a member before the complete surviving family's p-values exist. Cache/reuse identical governed intermediate results with provenance when available; a direct-dispatch call must obtain or compute its exact source-defined inputs, not pretend its predecessors succeeded. Stage-5 internal riders remain exploratory, after the Stage-4 family is frozen, even when housed in H1/H2.

Reference files under docs/execution/ are the ORIGINAL SKELETONS, not tested turnkey implementations. Adapt paths and complete engineering in a governed workspace; preserve source scientific rules and disclose unresolved choices. Write a result record with component estimates, uncertainty, family eligibility, source interpretation row, diagnostics and boundary/skip dispositions. Released outputs follow source disclosure rules; confidential rows stay governed. Local native states summarize the source judgment; neither a state name nor source power forecasts prove equivalence or causal identification. A bounded claim requires its actual interval and registered rule. Retain unresolved source wording conflicts; do not convert mere nonsignificance into scientific null.

#### Frozen scope and vocabulary

```
claim_class_ceiling: associational
statement: 'Associational - matched event-study contrasts, a contrastive
↪ falsification gate,
 and a composition contrast; descriptive for trajectory panels and the
 ↪ bounded priority claim.
Binding vocabulary rules: 'is associated with' / 'consistent with' /
↪ 'suggests' only;
```

```

'effect', 'impact', 'causes' prohibited throughout; anticipation
↳ always worded as
'pre-boundary excess' relative to the annually dated closure-year
↳ boundary, never as announcement-dated
anticipation (S8); no phenomenon-discovery or Danish-first framing (S1-S3);
↳ H2 novelty confined
to the placebo's application to health billing (S7). H2 verdict semantics
↳ are fixed in
advance: 'consistent with H2' requires BOTH the confirmatory core
↳ (theta2a Holm-significant,
predicted direction) AND a passing falsification gate F2; the strict
↳ comparative magnitude
is only ever reported as an exploratory estimate with its CI. Exploratory
↳ results (H1_trend,
H2_diff, E1, E3, E4, and any gate-demoted member) carry the word
↳ 'exploratory' in the
heading and in every sentence stating them. The prose label never exceeds
↳ the evidence label.'
frozen_parameters:
seed: 20260831
gp_code_list: daytime GP = SPEC2 in {70..80}
ooh_code_list: out-of-hours GP = SPEC2 in {81, 82, 83, 84, 89}
ooh_robustness_classifier: adds YDLTID in {2, 3, 4, 5, 7, 8, 9}
spec2_evening_gp_disclosure: 'Registered choice, disclosed: the parsed
↳ value labels mark
SPEC2 71 ("General practitioner, consultation, evening etc."), 73
↳ ("General practitioner,
telephone consultation, evening etc."), and 75 ("General practitioner,
↳ home visit, evening
etc.") as explicitly off-schedule GP services; they nonetheless remain on
↳ the daytime
side of the PRIMARY classifier because H3's out-of-hours construct and
↳ its published
power basis are keyed to the laegevagt out-of-hours service (SPEC2 81-84,
↳ 89), the quantity
the published Danish 8-12% share describes. Attenuation direction, both
↳ readings: counting
evening GP as daytime attenuates theta3 toward zero - conservative
↳ against the positive
H3 conclusion, anti-conservative for the null-bounding "compositionally
↳ unremarkable"
reading, whose Stage-2 statement therefore carries a mandatory caveat.
↳ In-environment
validation must verify that the YDLTID-augmented robustness classifier
↳ captures SPEC2
71/73/75 rows (their evening time codes 2-5); if YDLTID proves
↳ unpopulated on those rows,
the robustness classifier - and only the robustness classifier - is
↳ amended mechanically
and with disclosure to include SPEC2 71/73/75 directly. The SPEC2
↳ 71/73/75 volume share
is reported, and both classifiers appear side by side in Table T7.'
successor_share_thresholds: 'true closure: ENDNEDL final closure AND
↳ worker-flow max-successor
share s(w) < 0.30; spurious closure: forward identity via
↳ IDFREEM/REGELFR/FRELFREM AND
s(w) >= 0.70; ambiguous mid-zone excluded; GRAFT-3 sensitivities
↳ 0.20/0.80 and 0.40/0.60'
quarter_mapping: 'ISO week of HONUGE: 1-13 / 14-26 / 27-39 / 40-52; ISO
↳ week 53 to Q4'

```

```

winsorization: person-quarter outcomes at the 99.9th percentile within era
↳ x calendar quarter
matching: 1:3 nearest-neighbor Mahalanobis within exact strata, without
↳ replacement across
 sets, target k = 3, minimum k = 1; calipers |dif baseline contacts| <= 3
 ↳ and |dif log(1+BRUHON)|
 <= 0.5; exposed processed in seeded random order (seed 20260831)
software: R >= 4.3, fixest::feols (OLS) and fixest::fepois (Poisson QML)
↳ robustness); versions
 frozen in the registered script
note: Code lists and thresholds frozen before any data access; code labels
↳ quoted verbatim
 from parsed DST value labels and re-validated in-environment before
 ↳ estimation (frozen
 fallback rules in contingent_decisions).

```

1. **IDAS ingest, 2007–2024.** Input: raw IDAS. Keep the D1 variable list; types cast; one row per (LBNR, year). Output: `interim/idas_panel.parquet`. Check: all 18 years present; row count per year within  $\pm 50\%$  of the  $\sim 227\text{k}/\text{year}$  inventory basis; no duplicate (LBNR, year); log per-year counts.
2. **Workplace identity chains.** Input: `idas_panel.parquet`. Build identity linkage via IDFREM (forward) and IDTILB (backward); resolve each workplace-year to a chained identity. Output: `interim/workplace_identity.parquet`. Check: forward/backward reciprocity rate reported; no cycles; share of years with broken chains reported.
3. **Terminal-event candidates, cohorts c in 2009–2023.** Input: steps 1–2. Flag identities present in year c with `ANTNOV`  $\geq 5$  and absent in year c+1 (IDAS extends through 2024, so c = 2023 is checkable). Attach `ENDNEDL`, `REGELFR`, `FRELFREM`, `ANDELOPS`, `ANDELUDS`, `AFGRATE`, `ANTALXFR`. Output: `interim/terminal_events.parquet`. Check: per-cohort event counts logged (rounded to 10 in any released object); validate the `ENDNEDL` flag-year convention against the identity panel — if it differs from “flagged in the workplace’s last active year”, apply the registered re-anchoring (c := last IDAS year with `ANTNOV`  $\geq 5$  for the identity) mechanically and report the recalibration verbatim.
4. **IDAN ingest, 2007–2024 (slim).** Input: raw IDAN. Keep the D2 confirmatory variables. Output: `interim/idan_slim.parquet`. Check:  $\sim 3.2\text{M}$  rows/year ( $\pm 50\%$ ); PNR and workplace keys non-null; per-year counts logged.
5. **Primary-job assignment, per person x year.** Input: `idan_slim.parquet`. Primary job = largest `BREDT_LOEN_BELOEB` in the year; ties broken by larger `LOENTIMER`, then lower `ARBNR` (deterministic). Output: `interim/primary_jobs.parquet`. Check: exactly one primary employment per person-year; re-run idempotence (identical output hash on repeat); multi-job share reported.
6. **Workplace sector assignment.** Input: steps 4–5. Modal `ARB_SEKTORKODE` among each workplace’s year c-1 primary employees; flag private-sector workplaces. Output: `interim/workplace_sector.parquet`. Check: private share of size-5+ workplaces reported; workplaces with no c-1 employees flagged, not dropped silently.

7. **Worker-flow successor shares  $s(w)$ .** Input: steps 3, 5. For each terminal event: among the event workplace’s year- $c$  November primary employees, the largest share whose year- $c+1$  primary workplace is a single successor identity. Computed on the year- $c$  workforce — deliberately independent of the frozen  $c-1$  cohort. Output: `interim/successor_shares.parquet`. Check: every terminal event carries  $s(w)$  or a documented reason (no year- $c$  primary employees); distribution summary logged.
8. **Event classification + cross-validation (Stage-1 gate).** Input: steps 3, 7. True closure = ENDNEDL final closure AND  $s(w) < 0.30$ ; spurious closure = forward identity via IDFREM/REGELFR/FRELFREM AND  $s(w) \geq 0.70$ ; else ambiguous-excluded. Also classify under the GRAFT-3 threshold pairs (0.20/0.80, 0.40/0.60). Cross-tabulate worker-flow classes against IDAS-native flags (ENDNEDL/IDFREM/REGELFR/FRELFREM/ANDELUDS). Output: `interim/events_classified.parquet` + released cross-tab. Checks (all Stage-1 stop rules): true-closure events  $\geq 500$  (project kill below); spurious events  $\geq 150$  (spurious-arm kill below); agreement  $\geq 75\%$  (75–90% proceeds with worker-flow classes as registered, full disagreement table published); GRAFT-4  $s(w)$  bimodality with ambiguous mid-zone  $< 35\%$  of terminal events; annual true-closure rate among private size-5+ workplaces in [1%, 8%] every cohort year; spurious/true ratio in [0.1, 1.5]. Any violation triggers the classification audit before any further stage. Report projected exposed  $N$  against the sampling plan. ## 1. DATA ACQUISITION

Request the following FOUR register datasets from DST — the study’s hard data scope; nothing else may be requested or used. Names are the documented TIMES/HKV identities. Every variable below is verified against the registration’s mapping bundle (`times_variables_bilingual.csv`, `hkv_variables_bilingual.csv`), with four disclosed exceptions verified only in the DST Forskningservice four-register inventory for the analysis server (PNR on all four registers; KOM on IDAS; VAGTOMR and SIKREAMT on SYSI) — these are flagged “inventory-verified” below and the disclosure must be carried into the report. DST delivers PNR as an anonymized project-specific person key; “PNR” below means that key.

## 1.1 Datasets and variables

**D1. IDA workplaces** (Danish: IDA arbejdssteder; TIMES `ida-databasen/ida-arbejdssteder`, HKV `ida-arbejdssteder`, valid from 01-01-1980). Request the 1980–2024 panel; the confirmatory tier consumes 2007–2024, the E1 exploratory tier 1990–2005. Unit: workplace x year.

Variable	Documented long name	Validity used	Role
LBNR	Workplace serial number	1980–2024	workplace identity key; identity chains via IDFREM/IDTILB; identity-disappearance events between $c$ and $c+1$
ARBNR	Arb number	2007–2024	workplace key, post-2008 vintage (entire confirmatory tier)
CVRNR	CVR number	2007–2024	firm key, post-2008 vintage
ENDNEDL	Final closure of workplace	through 2023	true-closure flag (with $s(w) < 0.30$ ); its 2023 end caps the last cohort
IDFREM	Identity of place of work forward in time	1980–2023	forward-continuation identity for spurious-closure classification; chain construction
IDTILB	Identity of place of work backward in time	1981–2024	backward identity chain (control stability, event-panel construction)
REGELFR	Rule regarding retained workplace the following year	1980–2023	forward-identity rule for spurious-closure classification
FRELFREM	Firm-internal/external closure via absorption	1980–2023	en-bloc absorption indicator for spurious-closure classification
ANDELOPS	Share of absorbed among employees	1981–2024	IDAS-native absorption share; classification cross-validation only

Variable	Documented long name	Validity used	Role
ANDELUDS	Share of employees split off to another workplace among employees	1980–2023	IDAS-native split-off share; cross-validation only
AFGRATE	Exit rate in the subsequent annual period, in per cent	1980–2023	stable-control definition (AFGRATE < 30%); mass-separation exclusion
ANTALXFR	Change in the number of employees in the following year	1980–2023	stable-control definition ( $ \text{ANTALXFR}  \leq 30\%$ )
ANTNOV	Number of employees as at November	1980–2024	size eligibility 5–1000; event presence (ANTNOV $\geq 5$ ); exact-stratum size classes
ARB_HOVED_BRAND	Industry of the workplace	2008–2024	NACE section exact stratum
KOM	Municipality of the workplace	1980–2024	workplace municipality mapped to the 5 post-2007 regions (exact stratum). INVENTORY-VERIFIED on IDAS (absent from the public TIMES/HKV parses for this register) — disclose legacy workplace key; E1 exploratory only
ARBGNR	Employer number	1980–2007	legacy workplace key; E1 exploratory only
ARBSTK	Workplace code	1980–2007	legacy workplace key; E1 exploratory only

**D2. IDA employments** (Danish: IDA ansaettelser; IDA HKV analogues and  
TIMES ida-databasen/ida-ansaettelser;  
beskaeftigelsesoplysninger-der-vedroerer-ida-ansaettelser)

loenforhold-der-vedroerer-ida-ansaettelser, valid from 01-01-1980). Request 1980–2024; confirmatory tier consumes 2007–2024 (cohort anchoring at November c-1, worker flows into c+1). Unit: person x employment relation x year.

Variable	Documented long name	Validity used	Role
PNR	De-identified person number	all years	person spine joining IDAN, SSSY, SYSL INVENTORY-VERIFIED (absent from public parses) — disclose
LBNR	Workplace serial number	all years	employment-to-workplace link (with ARBNR/CVRNR, 2008+)
ARBNR	Arb number	2008–2024	workplace key on the employment relation
CVRNR	CVR number	2008–2024	firm key on the employment relation
ANSAAR	Year of employment at the workplace	all years	tenure (c-1 minus ANSAAR); inclusion floor ANSAAR <= c-2; Mahalanobis component
BREDT_LOEN_BEDLOEB	2008 salary amount	2008–2024	primary-job rule (largest in c-1); baseline wage decile (Mahalanobis; E3 heterogeneity)
LOENTIMER	Paid hours (LOENTIMER)	2008–2024	deterministic primary-job tie-break (larger LOENTIMER, then lower ARBNR)
ARB_SEKTORKOD	Workplace sector code	2008–2024	private-sector filter via modal sector among year c-1 employees

Variable	Documented long name	Validity used	Role
ANSXFREM	Change of employment up to the following year	1980–2023	E4 attrition accounting (forward-separation flag)
ARBGNR	Employer number	1980–2007	legacy key; E1 only. Documented publicly on the workplace register; carried on legacy IDAN vintages per the Forskningservice inventory — disclose
ARBSTK	Workplace code	1980–2007	legacy key; E1 only (same disclosure)
STILL	Occupational position	1980–2007	legacy job characteristic for E1-era matching (SYSI lacks age/sex)
TILKNYT	Attachment to primary workplace	1980–2007	legacy primary-employment marker; E1 only
ANSDAGE	Number of days employed	1980–2007	legacy job characteristic for E1-era matching
TIMELON	Hourly wage in the employment	1980–2007	legacy wage measure for E1-era matching

**D3. Health insurance statistics, current service series (SSSY)**  
(Danish: Sygesikringsstatistik; TIMES sygesikringsstatistik; HKV sygesikring---ydelser / sygesikring---honorarer / sygesikring-kontakter datasets; SPEC2 valid from 01-01-2005). Request 2005–2025 service rows; the pipeline consumes 2007–2025 (baselines c-3 onward, post-years through 2024/2025). Unit: service row.

Variable	Documented long name	Validity used	Role
PNR	De-identified person number	all years	person spine. INVENTORY- VERIFIED — disclose
HONUGE	Fee week	all years	the weekly clock; person x HONUGE week aggregation; ISO-week-to- quarter mapping
KONTAKT	Contact	2005–2025	row-level contact marker summed to quarterly GP contact counts (frozen fallback if in-environment validation fails, Section 4.6)
YDLANT	Number of services within the specialty	all years	service counts (descriptives; E1-era contact proxy weighting)
BRUHON	Gross fee to doctor etc.	all years	gross fees: gated H2 fee secondary, E3 fee analogues, Mahalanobis $\log(1+\text{BRUHON})$ ; nominal, compared only within cohort x calendar-quarter cells
SPEC2	2-digit specialisation	2005–2025	frozen GP (70–80) and OOH (81–84, 89) code lists
YDLTID	Time code for the benefit	all years	robustness OOH classifier; missing/“not relevant” classified daytime (conservative against H3)

Variable	Documented long name	Validity used	Role
AFRPER	Settlement period	all years	settlement-lag characterization (descriptive)
BEHANDLINGSDAT	The treatment date for the service	2021–2025 only	HONUGE settlement-lag calibration; can calibrate but never replace HONUGE
SIKGRUP	Health insurance group	all years	exclusion of modal group-2 insured (different billing regime)
ALDERIMP	Age at end of period, incl. imputed	all years	age bands for exact strata (explicit unknown stratum); eligibility 25–59
KOENIMP	Sex incl. imputed	all years	sex stratum (explicit unknown stratum); E3 heterogeneity
YDERNR	Provider number (ydernummer)	all years	frozen fallback contact definition (distinct PNR x week x YDERNR x code-group cells); never exported

**D4. Health insurance statistics, legacy closed series (SYSI)** (Danish: Sygesikringsstatistik; documented under TIMES *sygesikringsstatistik* and the HKV sygesikring datasets, valid from 01-01-1990; series closed 2005). Request 1990–2005. Used ONLY in the exploratory SYSI-era replication E1 (cohorts 1992–2003); never pooled with the confirmatory era across the 2007/08 IDA vintage break.

Variable	Documented long name	Validity used	Role
PNR	De-identified person number	all years	person spine. INVENTORY-VERIFIED — disclose
HONUGE	Fee week	all years	weekly clock, legacy era

Variable	Documented long name	Validity used	Role
SPECIALE	6-digit speciality	all years	era GP/OOH classification (SPEC2 does not exist pre-2005); era code list frozen at Stage 0
YDLANT	Number of services within the specialty	all years	E1 contact proxy: service-row counts weighted by YDLANT (KONTAKT does not exist pre-2005)
BRUHON	Gross fee to doctor etc.	all years	era fee outcome
YDLTID	Time code for the benefit	all years	era time-of-day classification
VAGTOMR	Out-of-hours (laegevagt) service area	1997–2005	era OOH marker. INVENTORY-VERIFIED — disclose
AFRPER	Settlement period	all years	era settlement-lag characterization
SIKGRUP	Health insurance group	all years	era group-2 exclusion
SIKREKOM	Municipality of the insured person	all years	era patient geography (E1 only; SSSY has no patient geography)
SIKREAMT	County of the insured person	all years	era patient geography (E1 only). INVENTORY-VERIFIED — disclose

## 1.2 Join order, keys, and expected cardinalities

Person spine: PNR (IDAN x SSSY x SYSI). Workplace spine: LBNR with ARBNR/CVRNR (IDAS x IDAN, 2008+ vintage — the entire confirmatory tier); LBNR with ARBGNR/ARBSTK (pre-2008 — E1 only); workplace identity chains via IDFREM/IDTILB. Never join across the 2007/08 IDA vintage break.

Join order (confirmatory tier):

1. **IDAS(c) x IDAS(c+1) self-join** on workplace identity (LBNR chain): flags identity-disappearance events. 1:0/1 per identity. Expected scale: ~227k workplace rows/year (Forskningsservice inventory); terminal identity-disappearance events per cohort year in the low thousands before classification.
2. **IDAS <- IDAN** on (year, LBNR/ARBNR/CVRNR): workplace-year to employment rows, 1:many (mean ~14 employment rows per workplace-year; ~3.2M IDAN rows/year).
3. **IDAN(c) -> IDAN(c+1)** on PNR (primary jobs only): worker-flow successor shares. 1:0/1 per person.
4. **Cohort roster <- SSSY** on PNR: 1:many (person to service rows; ~88M SSSY rows/year population-wide; ~4–5 GP contacts/person-year, often multiple rows per contact — hence KONTAKT, not row counts).
5. **Matching:** exposed:controls = 1:up to 3, without replacement of controls across sets.
6. **Panel:** matched member x 16 event-quarters (Q1(c-2)..Q4(c+1)).

Expected cardinalities against the sampling plan (all realized counts go in the run manifest and exhibit T1): true-closure events ~500–1,700 per cohort year after classification (plan basis: 12,600 event workplaces over 15 cohorts); eligible exposed true-closure workers 10,000–20,000 per cohort, 150,000–300,000 total;  $\geq 135,000$  true 1:3 matched sets at the  $\geq 90\%$  match-rate floor; spurious-closure workers 30–60% of the true count (45,000–90,000 before matching;  $\geq 45,000$  sets at plan). Kill floors far below expectation: 500 true events / 10,000 sets (project); 150 spurious events / 3,000 spurious sets (spurious arm).

---

## Current schema and linkage

```

registers:
- register: IDA workplaces (IDAS)
 register_da: IDA arbejdssteder
 documentation: TIMES ida-databasen/ida-arbejdssteder (The IDA database /
 ↪ IDA workplaces)
 and HKV ida-arbejdssteder, valid from 01-01-1980
 years: 1980-2024 panel; event cohorts c in 2009-2023 (confirmatory);
 ↪ ENDNEDL ends 2023
 variables:
 - name: LBNR
 long_name: Workplace serial number
 role: key
 use: workplace identity; identity chains via IDFREM/IDTILB;
 ↪ identity-disappearance events
 between c and c+1
 - name: ARBNR
 long_name: Arb number
 role: key
 use: workplace key, 2007+ vintage (confirmatory era)
 - name: CVRNR
 long_name: CVR number
 role: key
 use: firm key, 2007+ vintage (confirmatory era)

```

- name: ENDNEDL  
long\_name: Final closure of workplace  
role: exposure  
use: true-closure flag (with  $s(w) < 0.30$ ); validity through 2023 caps the  
↳ last cohort
- name: IDFREM  
long\_name: Identity of place of work forward in time  
role: exposure  
use: forward-continuation identity for spurious-closure classification;  
↳ identity chain  
construction
- name: IDTILB  
long\_name: Identity of place of work backward in time  
role: exposure  
use: backward identity chain (control-stability and event-panel  
↳ construction); HKV valid  
from 01-01-1981
- name: REGELFR  
long\_name: Rule regarding retained workplace the following year  
role: exposure  
use: forward-identity rule for spurious-closure classification
- name: FRELFREM  
long\_name: Firm-internal/external closure via absorption  
role: exposure  
use: en-bloc absorption indicator for spurious-closure classification
- name: ANDELOPS  
long\_name: Share of absorbed among employees  
role: check  
use: IDAS-native absorption share; classification cross-validation  
↳ against worker-flow  
 $s(w)$
- name: ANDELUDS  
long\_name: Share of employees split off to another workplace among  
↳ employees  
role: check  
use: IDAS-native split-off share; classification cross-validation
- name: AFGRATE  
long\_name: Exit rate in the subsequent annual period, in per cent  
role: control  
use: stable-control definition ( $AFGRATE < 30\%$ ); mass-separation exclusion
- name: ANTALXFR  
long\_name: Change in the number of employees in the following year  
role: control  
use: stable-control definition ( $|ANTALXFR| \leq 30\%$ )
- name: ANTNOV  
long\_name: Number of employees as at November  
role: control  
use: 'workplace size: eligibility 5-1000, exact-stratum size classes,  
↳ event presence (ANTNOV  
 $\geq 5$ )'
- name: ARB\_HOVED\_BRA\_DB07  
long\_name: Industry of the workplace  
role: control  
use: NACE section exact stratum (2008+ vintage)
- name: KOM  
long\_name: Municipality of the workplace  
role: control  
use: workplace municipality mapped to the 5 post-2007 regions (exact  
↳ stratum)

```

verification: verified in the DST Forskningservice four-register
↳ inventory for the analysis
server; absent from the public TIMES/HKV parses - disclosed
- name: ARBGNR
 long_name: Employer number
 role: key
 use: legacy workplace key (pre-2008 vintage); E1 SYSI-era exploratory
 ↳ replication only
- name: ARBSTK
 long_name: Workplace code
 role: key
 use: legacy workplace key (pre-2008 vintage); E1 exploratory only
- register: IDA employments (IDAN)
 register_da: IDA ansaettelser
 documentation: TIMES ida-databasen/ida-ansaettelser (The IDA database / IDA
 ↳ employments);
 HKV analogues Beskaeftigelsesoplysninger / Loenforhold / Ledighed der
 ↳ vedroerer IDA ansaettelser,
 valid from 01-01-1980
years: 1980-2024 panel; cohort anchoring at November of year c-1; worker
↳ flows into c+1
variables:
- name: PNR
 long_name: De-identified person number
 role: key
 use: person spine joining IDAN, SSSY, SYSI
 verification: verified in the DST Forskningservice four-register
 ↳ inventory for the analysis
 server; absent from the public TIMES/HKV parses - disclosed
- name: LBNR
 long_name: Workplace serial number
 role: key
 use: employment-to-workplace link (with ARBNR/CVRNR, 2008+)
- name: ARBNR
 long_name: Arb number
 role: key
 use: workplace key on the employment relation, 2008+
- name: CVRNR
 long_name: CVR number
 role: key
 use: firm key on the employment relation, 2008+
- name: ANSAAR
 long_name: Year of employment at the workplace
 role: control
 use: tenure (c-1 minus ANSAAR); inclusion floor ANSAAR <= c-2;
 ↳ Mahalanobis component
- name: BREDT_LOEN_BELOEB
 long_name: Broad salary amount
 role: control
 use: primary-job rule (largest BREDT_LOEN_BELOEB in c-1); baseline wage
 ↳ decile (Mahalanobis
 component; E3 heterogeneity)
- name: LOENTIMER
 long_name: Paid hours (LOENTIMER)
 role: control
 use: deterministic primary-job tie-break (larger LOENTIMER, then lower
 ↳ ARBNR)
- name: ARB_SEKTORKODE
 long_name: Workplace sector code
 role: control

```

```

use: private-sector filter via modal sector among year c-1 employees
↳ (2008+)
verification: 'Exception, disclosed (converse of the inventory-only
↳ variables): publicly
documented (TIMES ida-ansaettelser) but NOT present in the recorded DST
↳ Forskningssservice
four-register inventory extract for the analysis server, and the
↳ claimed 2008+ validity
rests on the public parse alone. Server presence and coverage validated
↳ at Stage 1 before
any classification work; frozen fallback registered in
↳ analysis_plan.contingent_decisions
(private-sector filter from ARB_HOVED_BRA_DB07, NACE DB07 sections
↳ O/P/Q excluded).'
```

- name: ANSXFREM
 long\_name: Change of employment up to the following year
 role: check
 use: E4 arm-comparative attrition accounting (forward-separation flag)
- name: ARBGNR
 long\_name: Employer number
 role: key
 use: legacy key (pre-2008); E1 exploratory only
 verification: carried on legacy IDAN vintages per the Forskningssservice
 ↳ inventory; the
 public parses document it on IDA workplaces
- name: ARBSTK
 long\_name: Workplace code
 role: key
 use: legacy key (pre-2008); E1 exploratory only
 verification: carried on legacy IDAN vintages per the Forskningssservice
 ↳ inventory; the
 public parses document it on IDA workplaces
- name: STILL
 long\_name: Occupational position
 role: control
 use: legacy job characteristic for E1-era matching (SYSI lacks age/sex)
- name: TILKNYT
 long\_name: Attachment to primary workplace
 role: control
 use: legacy primary-employment marker; E1 exploratory only
- name: ANSDAGE
 long\_name: Number of days employed
 role: control
 use: legacy job characteristic for E1-era matching
- name: TIMELON
 long\_name: Hourly wage in the employment
 role: control
 use: legacy wage measure for E1-era matching
- register: Health insurance statistics, current service series (SSSY)
 register\_da: Sygesikringsstatistik (SSSY)
 documentation: TIMES sygesikringsstatistik (Health insurance statistics);
 ↳ HKV sygesikring
 - ydelser / honorarer / kontakter (SPEC2 valid from 01-01-2005)
 years: 2005-2025 service rows; extraction 2006-2025 (baselines c-3 onward,
 ↳ post-years through
 2024/2025)

```

variables:
- name: PNR
 long_name: De-identified person number
 role: key
```

```

use: person spine
verification: verified in the DST Forskningservice four-register
↳ inventory for the analysis
server; absent from the public TIMES/HKV parses - disclosed
- name: HONUKE
 long_name: Fee week
 role: outcome
 use: 'the weekly clock: person x HONUKE week aggregation,
 ↳ ISO-week-to-quarter mapping'
- name: KONTAKT
 long_name: Contact
 role: outcome
 use: row-level contact marker summed to quarterly GP contact counts
 ↳ (frozen fallback definition
 if in-environment validation fails)
- name: YDLANT
 long_name: Number of services within the specialty
 role: outcome
 use: service counts (descriptives; E1-era contact proxy weighting)
- name: BRUHON
 long_name: Gross fee to doctor etc.
 role: outcome
 use: 'gross fees: gated H2 fee secondary, E3 fee analogues, Mahalanobis
 ↳ log(1+BRUHON)
 component; nominal, compared only within cohort x calendar-quarter
 ↳ cells'
- name: SPEC2
 long_name: 2-digit specialisation
 role: outcome
 use: frozen GP (70-80) and OOH (81-84, 89) code lists; parsed labels mark
 ↳ 71/73/75 as
 "evening etc." GP services - kept daytime in the primary classifier by
 ↳ registered choice
 (see frozen_parameters.spec2_evening_gp_disclosure)
- name: YDLTID
 long_name: Time code for the benefit
 role: outcome
 use: robustness OOH classifier; missing/'not relevant' classified daytime
 ↳ (conservative
 against H3)
- name: AFRPER
 long_name: Settlement period
 role: check
 use: settlement-lag characterization
- name: BEHANDLINGSDATO
 long_name: The treatment date for the service
 role: check
 use: HONUKE settlement-lag calibration on 2021+ rows only (guard-band
 ↳ contingency); can
 calibrate but never replace HONUKE
- name: SIKGRUP
 long_name: Health insurance group
 role: check
 use: exclusion of modal group-2 insured (different billing regime)
- name: ALDERIMP
 long_name: Age at end of period, incl. imputed
 role: control
 use: age bands for exact strata (with explicit unknown stratum);
 ↳ eligibility 25-59
- name: KOENIMP

```

```

 long_name: Sex incl. imputed
 role: control
 use: sex stratum (with explicit unknown stratum); E3 heterogeneity
- name: YDERNR
 long_name: Provider number (ydernummer)
 role: key
 use: frozen fallback contact definition (distinct PNR x week x YDERNR x
 ↪ code-group cells)
- register: Health insurance statistics, legacy closed series (SYSI)
 register_da: Sygesikringsstatistik (SYSI)
 documentation: documented under TIMES sygesikringsstatistik and HKV
 ↪ sygesikring datasets
 (valid from 01-01-1990); series closed 2005
 years: 1990-2005; used ONLY in the exploratory SYSI-era replication E1
 ↪ (cohorts 1992-2003)
 variables:
- name: PNR
 long_name: De-identified person number
 role: key
 use: person spine
 verification: verified in the DST Forskningservice four-register
 ↪ inventory for the analysis
 server; absent from the public TIMES/HKV parses - disclosed
- name: HONUGE
 long_name: Fee week
 role: outcome
 use: weekly clock, legacy era
- name: SPECIALE
 long_name: 6-digit speciality
 role: outcome
 use: era GP/OOH classification (SPEC2 does not exist pre-2005); frozen
 ↪ era code list
- name: YDLANT
 long_name: Number of services within the specialty
 role: outcome
 use: 'E1 contact proxy: service-row counts weighted by YDLANT (KONTAKT
 ↪ does not exist
 pre-2005)'
- name: BRUHON
 long_name: Gross fee to doctor etc.
 role: outcome
 use: era fee outcome
- name: YDLTID
 long_name: Time code for the benefit
 role: outcome
 use: era time-of-day classification
- name: VAGTOMR
 long_name: Out-of-hours (laegevagt) service area
 role: outcome
 use: era OOH marker, 1997-2005 only
 verification: verified in the DST Forskningservice four-register
 ↪ inventory for the analysis
 server; absent from the public TIMES/HKV parses - disclosed
- name: AFRPER
 long_name: Settlement period
 role: check
 use: era settlement-lag characterization
- name: SIKGRUP
 long_name: Health insurance group
 role: check

```

```

 use: era exclusion of group-2 insured
- name: SIKREKOM
 long_name: Municipality of the insured person
 role: control
 use: era patient geography (E1 only; SSSY has no patient geography)
- name: SIKREAMT
 long_name: County of the insured person
 role: control
 use: era patient geography (E1 only)
 verification: verified in the DST Forskningservice four-register
 ↪ inventory for the analysis
 server; absent from the public TIMES/HKV parses - disclosed
variable_verification_note: 'Every variable above was checked against the
↪ parsed public documentation
(data/mapping/times_variables_bilingual.csv;
↪ data/mapping/hkv_variables_bilingual.csv).
Exceptions carried with an explicit per-variable verification note: PNR,
↪ IDAS KOM, SYSI
VAGTOMR, and SYSI SIKREAMT are verified in the DST Forskningservice
↪ four-register inventory
for the analysis server but are absent from the public TIMES/HKV parses;
↪ IDAN's legacy
workplace keys (ARBGNR/ARBSTK) are documented publicly on the workplace
↪ register and carried
on legacy IDAN vintages per the same inventory. Only these four registers
↪ exist on the analysis
server: there are no diagnosis, hospital, prescription, or population-spine
↪ registers in
scope, and outcomes are therefore utilization volume, fees, and timing
↪ composition - never
health states. One variable runs the other way and is disclosed as such:
↪ IDAN ARB_SEKTORKODE
- the private-sector eligibility filter - is publicly documented (TIMES
↪ ida-ansaettelser)
but absent from the recorded Forskningservice inventory extract, its
↪ claimed 2008+ validity
resting on the public parse alone; its presence is validated at Stage 1
↪ before any classification
work, with a frozen ARB_HOVED_BRA_DB07-based fallback registered in
↪ analysis_plan.contingent_decisions.'
linkage:
 keys:
 - 'person spine: PNR (IDAN x SSSY x SYSI)'
 - 'workplace: LBNR with ARBNR/CVRNR (IDAS x IDAN, 2008+ vintage - the
 ↪ entire confirmatory
 tier); LBNR with ARBGNR/ARBSTK (pre-2008 vintage - E1 exploratory only);
 ↪ workplace identity
 chains via IDFREM/IDTILB; never pooled across the 2007/08 IDA vintage
 ↪ break'
 expected_linkage_rate_floor: '>= 80% of cohort workers (exposed + control
 ↪ pool) with >=
 1 SSSY row of any kind in the 5-year window c-3..c+1 (outcome-neutral
 ↪ check; Danish annual
 GP contact prevalence benchmark ~85%); below floor halts at Stage 3'

```

## Current authoritative contingencies

```

- condition: ENDNEDL flag-year convention differs in-environment from the
↪ assumed 'flagged
 in the workplace's last active year'

```

```

rule: Re-anchor mechanically so that c always denotes the last IDAS year
↳ with ANTNOV >=
 5 for the identity; the estimand is unchanged; the recalibration is
 ↳ reported verbatim
- condition: ARB_SEKTORKODE is absent from the analysis-server IDAN extract
↳ (publicly documented
 on TIMES ida-ansaettelser but unconfirmed in the recorded
 ↳ Forskningservice inventory;
 presence and 2008+ coverage validated at Stage 1 before any
 ↳ classification work)
rule: 'Frozen fallback: define the private-sector filter from
↳ ARB_HOVED_BRA_DB07 instead
 - a workplace counts as private if and only if its NACE DB07 section is
 ↳ not 0 (public
 administration and defence), P (education), or Q (human health and social
 ↳ work); the substitution
 and its effect on eligible-workplace and worker counts are reported
 ↳ verbatim; no other
 filter change is permitted'
- condition: Match rate < 95% of eligible exposed workers at target k = 3
↳ under full exact
 strata
rule: 'Apply the pre-specified coarsening ladder in order, stopping at the
↳ first rung reaching
 95% of exposed matched with k >= 1: (1) NACE section -> 10 aggregate
 ↳ industry groups;
 (2) drop region stratum; (3) age bands 5-year -> 10-year. Never drop sex,
 ↳ cohort, size
 class, or the baseline-billing calipers; realized k distribution reported
 ↳ and fed to the
 Stage-2 blind re-simulation'
- condition: Median HONUGE settlement lag (BEHANDLINGSDATO calibration, 2021+
↳ rows) > 4 weeks
rule: 'Guard band: exclude Q1(c+1) from the theta2a/theta2b windows (means
↳ over Q2(c+1)..Q4(c+1))
 and report theta1 both as registered and with a 4-week-shifted boundary;
 ↳ the registered
 theta1 window remains primary'
- condition: KONTAKT does not behave as a row-level contact marker on
↳ in-environment validation
rule: 'Switch to the frozen fallback contact definition: count of distinct
↳ PNR x HONUGE-week
 x YDERNR x GP-code-group cells with >= 1 positive-fee row; switch and
 ↳ level effect reported;
 all estimands re-run under both definitions with the fallback as primary'
- condition: OLS and Poisson QML disagree in sign on any primary quantity
rule: Foreground the pre-committed OLS estimate, report both, flag the
↳ discordance in the
 Summary paragraph; no post hoc model selection
- condition: 'c-2 placebo panel FAILURE, defined as: placebo contrast
↳ significant at alpha
 = 0.05 AND |estimate| > 1% of control baseline mean (50% of the theta1
 ↳ floor effect) -
 the single registered failure condition, stated identically in
 ↳ outcome_neutral_checks
 and the H1 placebo kill'
rule: 'Execute the single pre-specified re-match adding year c-2 baseline
↳ contacts to the
 Mahalanobis set; if the placebo fails again under the same condition, the
 ↳ entire registered

```

event-study tier (confirmatory and exploratory event-study quantities) is  
 ↳ killed and the  
 study reports descriptive trajectories only. The two non-complementary  
 ↳ cells are pre-committed  
 passes with mandatory disclosure, never discretion: (a) significant at  
 ↳  $\alpha = 0.05$  with  
 $|\text{estimate}| \leq 1\%$  - no re-match, no kill; the Table T7 placebo row flags  
 ↳ the significance  
 and states that a sub-floor-scale imbalance was detected; (b) not  
 ↳ significant with  $|\text{estimate}|$   
 $> 1\%$  - no re-match, no kill; a mandatory caution paragraph reports the  
 ↳ placebo magnitude  
 against the  $\theta_1$  floor and the realized placebo MDE, and the H1/H2  
 ↳ Summary paragraphs  
 must cite it'

- condition: A matched control's own workplace acquires a terminal event  
 ↳ within  $c-2..c+1$  after  
 matching (data revision)  
 rule: Drop that control from the set; drop the whole set only if no control  
 ↳ remains; dropped  
 counts reported; no re-matching
- condition: YDLTID missing or code 0 ('not relevant') on a GP row  
 rule: Classify as daytime (conservative against the  $\theta_3$  prediction);  
 ↳ missingness share  
 reported; if  $> 20\%$  of peri-event GP rows, H3's kill criterion triggers
- condition: Exact tie in Mahalanobis distance for a nearest-neighbor match  
 rule: Deterministic tie-break by smallest absolute baseline-contact  
 ↳ difference, then seeded  
 RNG (seed 20260831)
- condition: IDAS-flag vs worker-flow classification agreement  $< 90\%$  but  $\geq$   
 ↳  $75\%$   
 rule: Proceed with the worker-flow classification as registered (the canon,  
 ↳ S7); publish  
 the full disagreement cross-tabulation; below  $75\%$ , H2's kill criterion  
 ↳ governs
- condition: 'GRAFT 3: sign of  $\theta_{2b}$  or  $\theta_{2\_diff}$  not stable across the  
 ↳ three successor-share  
 threshold pairs'  
 rule: Flag classification fragility in the pre-committed H2 Summary  
 ↳ paragraph; the 30/70  
 estimate remains primary; no threshold shopping
- condition: 'GRAFT 6: exposed-vs-control difference in later-presence rates  
 ↳ (any IDAN record  
 in year  $c+1$  or any SSSY row in  $c+1$  or later) exceeds 2 percentage points'  
 rule: A mandatory bounding paragraph accompanies all post-boundary  
 ↳ ( $\theta_{2a}/\theta_{2b}$ ) conclusions,  
 converting the differential into best/worst-case bounds; trigger and  
 ↳ threshold registered,  
 not discretionary
- condition: 'GRAFT 9: the control forward-window sensitivity yields SMALLER  
 ↳ magnitudes than  
 the primary estimates'  
 rule: Flag as diagnostic of something other than contamination (registered  
 ↳ direction note);  
 reported verbatim, no re-specification
- condition: 'STAGE-2 REALIZED-N GATE: realized matched sets (or spurious  
 ↳ sets for F2) fall  
 below a member's binding  $N^*$  from the Stage-0 re-run, or the Stage-2  
 ↳ blind re-simulation

```

 (realized Ns, realized k distribution, baseline-year moments only) shows
 ↪ power < 0.95
 for a member at its floor'
rule: That member is demoted to exploratory BEFORE any post-baseline
 ↪ outcome is accessed;
 the Holm family shrinks and alphas recompute by the registered
 ↪ fixed-point iteration among
 survivors; demotion-only - no member, threshold, or floor may move in the
 ↪ other direction;
 the realized-power/MDE table is published either way
- condition: Realized control-group OOH share of GP contacts < 5% (below the
 ↪ published 8-12%
 range)
rule: H3's power basis is re-verified in the Stage-2 blind re-simulation at
 ↪ the realized
 share; if < 0.95 at the 5-point floor, H3 demotes to exploratory per the
 ↪ realized-N gate;
 if realized share < 2%, H3's kill criterion governs instead
- condition: Spurious arm killed at Stage 1/2 (events < 150, sets < 3,000,
 ↪ classification
 agreement < 75%, or bimodality unresolved)
rule: Gate F2 and theta2_diff/theta2b are reported unavailable (drop
 ↪ disclosed before outcome
 access); theta2a remains in the Holm family if its own N* is met, but the
 ↪ H2 verdict is
 capped at 'partially consistent - comparative quantifier untestable' and
 ↪ the Discussion
 states that the excess cannot then be distinguished from selection into
 ↪ disappearing workplace
 identities

```

## Current shared design

```

type: stacked matched difference-in-differences event study (matched
 ↪ event-study contrasts
 at quarterly confirmatory resolution with weekly descriptive trajectories;
 ↪ associational
 claim class)
design_name: 'Bracing for the Blow - remediated confirmatory design: stacked
 ↪ 1:3 riskset-matched
 difference-in-differences of quarterly GP billing around true vs spurious
 ↪ Danish workplace
 closures (SSSY era, closure cohorts 2009-2023), with the true-closure
 ↪ post-year excess as
 H2's confirmatory core, a spurious-arm falsification gate outside the Holm
 ↪ family, published-basis
 out-of-hours power calibration, and a pre-outcome realized-N demotion gate'
identifying_variation: 'Establishment terminal events in IDAS - workplace
 ↪ identities present
 in year c (ANTNOV >= 5) and absent in year c+1 - across 15 closure cohorts
 ↪ 2009-2023, classified
 true vs spurious by IDAN worker-flow successor shares per the Benedetto et
 ↪ al. (2007) /
 Hethey-Maier & Schmieder (2013) canon, contrasted against riskset-matched
 ↪ workers at stable
 workplaces within exact strata (cohort x sex x age band x NACE section x
 ↪ size class x region)
 with 52 weeks of pre-baseline billing-history matching. Closure dating is
 ↪ annual November-status,

```

so no sharp within-scope discontinuity exists: the claim ceiling is  
↳ associational, all pre-event  
quantities are worded as pre-boundary excess (never announcement-dated  
↳ anticipation), and  
the spurious-closure placebo plus the c-2 placebo panel carry the severity  
↳ burden.'

alternatives\_considered:  
- design: 'Round-1 conservative design (VOIDED): single-era stacked  
↳ matched-PAIR (1:1) DiD,  
closure cohorts 2009-2022, sampling self-handicapped to half the lower  
↳ bound of its own  
derived expectation (75k pairs), H2 as the true-minus-spurious  
↳ differential, H3 simulated  
at a 3% OOH base rate'  
discrimination: identical contrast logic; differed only in matching ratio,  
↳ sampling plan,  
H2's confirmatory construct, and H3's power basis  
why\_not\_selected: Its deposited Stage-0 power simulation emptied the  
↳ confirmatory family  
via the registered fixed-point rule (H2 0.358, H1\_trend 0.380, H3 0.614,  
↳ H1 0.840) and  
the driver audit voided the resulting all-exploratory registration on  
↳ 2026-08-31. The  
remediation (authorized honest levers only; no floor raised) produced the  
↳ registered design;  
the full record is carried in remediation\_record and the records of both  
↳ power runs are  
published with the protocol.

- design: 'Design B, ambitious three-arm weekly event study 1991-2023:  
↳ weekly-resolution confirmatory  
estimands, both eras (SYSI + SSSY) confirmatory, a mass-separation arm at  
↳ surviving workplaces,  
ANTNOV cap 1000'  
discrimination: weekly confirmatory estimands would sharpen timing-shape  
↳ claims; a second  
era and third arm would extend coverage and external validity  
why\_not\_selected: Weekly confirmatory estimands are hostage to the HONUGE  
↳ settlement lag;  
the SYSI era cannot match on age or sex and sits across the 2007/08 IDA  
↳ vintage break;  
the mass-separation arm breaks one-study scope discipline. Its honest  
↳ elements - the ANTNOV  
1000 cap, pre-committed reversed-sign interpretations, and the absorption  
↳ placebo - were  
absorbed into the registered design; the SYSI era survives as exploratory  
↳ E1 and weekly  
resolution as descriptive trajectories under the GRAFT-2 step diagnostic.

threats:  
- threat: Anticipatory attrition - healthier/more mobile workers leave  
↳ sinking workplaces  
before closure (Schwerdt 2011), biasing late-window composition  
addressed\_by: Cohort frozen at November-status attachment in year c-1,  
↳ before the terminal  
year (S9); intention-to-follow regardless of subsequent separation; E4  
↳ arm-comparative  
attrition accounting panel quantifies the channel, labeled  
↳ post-treatment-conditioned  
descriptives  
- threat: 'Billing responses to separations already realized within the  
↳ terminal year - under

intention-to-follow, workers who separate from the closing workplace  
 ↪ early in year c (the wave E4 describes) contribute post-job-loss utilization to the  
 ↪ pre-boundary Q3(c)/Q4(c) window, so a Holm-significant  $\theta_{t1}$  could be generated wholly by  
 ↪ realized early job loss with zero anticipation. Direction: inflates  $\theta_{t1}$  relative to a  
 ↪ pure-anticipation reading.  
 The frozen clause ''before any separation is visible in annual IDAN  
 ↪ records'' holds at the cohort-anchoring level (attachment frozen at November c-1) but not  
 ↪ for individual early leavers, whose separation is visible in the year-c IDAN record.'  
 addressed\_by: The pre-committed H1 interpretation names both channels  
 ↪ (anticipatory care-seeking and/or billing responses to separations already realized early within the  
 ↪ terminal year) with the E4 year-c separation share bounding the latter; the E4  
 ↪ arm-comparative separation shares themselves; and the registered post-treatment-conditioned  
 ↪ exploratory split of  $\theta_{t1}$  by November-year-c attachment (E4), which localizes the channel  
 ↪ without touching the confirmatory estimand  
 - threat: Selection of sicker workers into dying workplaces (distressed-firm  
 ↪ workforce composition) masquerading as anticipation or as post-year excess  
 addressed\_by: Exact + Mahalanobis matching on 52 weeks of pre-baseline  
 ↪ billing (contacts, fees, OOH, contact-weeks) in year c-1; c-2 placebo panel with numeric  
 ↪ failure threshold; falsification gate F2 - the spurious-closure arm separates identity-death  
 ↪ from economic job loss with a pre-committed disconfirmation trigger that can only take  
 ↪ support away from H2  
 - threat: HONUGE is the fee/settlement week, not the service week - billing  
 ↪ lag blurs the boundary and could relabel pre-boundary care as post-boundary  
 addressed\_by: Quarter-level primary estimands (robust to  $\leq 4$ -week lags);  
 ↪ BEHANDLINGSDATO calibration (2021+) with the guard-band contingency; AFRPER used  
 ↪ descriptively; GRAFT-2 week-0 step diagnostic bounds any timing-shape claim  
 - threat: Annual closure dating - the design cannot see announcement dates,  
 ↪ so 'anticipation' could be mislabeled  
 addressed\_by: Estimand defined and worded exclusively as pre-boundary  
 ↪ excess relative to the closure-year boundary (S8); claim ceiling forbids announcement-locked  
 ↪ language; gated H1\_trend partial pattern pre-commits the flat-elevation reading when no  
 ↪ within-year growth is seen  
 - threat: Local-economy co-shocks hitting exposed and control workers'  
 ↪ regions simultaneously  
 addressed\_by: Exact matching on region and NACE section plus cohort x  
 ↪ calendar-quarter fixed effects; residual local shocks affect controls too and attenuate the  
 ↪ contrast (bias toward

```

 null, stated)
- threat: Control contamination by quietly dying or shrinking workplaces
 addressed_by: Stable-control definition (unchanged identity c-2..c+1, no
 ↳ ENDNEDL, AFGRAFE
 ↳ < 30%, |ANTALXFR| <= 30%); control-drop rule on data revision; GRAFT-9
 ↳ forward-window
 ↳ (c+2/c+3) sensitivity with registered direction note (truncated for the
 ↳ last two cohorts,
 disclosed)
- threat: Reverse causality at small firms - workforce or owner illness
 ↳ causes the closure
 addressed_by: Size floor ANTNOV >= 5; associational claim ceiling with the
 ↳ reverse channel
 ↳ named in the limitations; exploratory size heterogeneity (size classes
 ↳ now extend to 1000)
- threat: 2007/08 IDA vintage break contaminating any pooled estimate
 addressed_by: Confirmatory tier confined entirely to post-break cohorts (c
 ↳ >= 2009); SYSI-era
 ↳ analysis is exploratory replication (E1) and never pooled
- threat: Unobservable death/emigration (no population spine in scope)
 ↳ deflating post-event
 ↳ billing
 addressed_by: Named scope limitation; direction discussed (deflates the
 ↳ post window - conservative
 ↳ for theta1, attenuating for theta2a); pre-specified sensitivity excluding
 ↳ possible-death/emigration
 ↳ sets; GRAFT-6 2pp later-presence trigger forces best/worst-case bounds on
 ↳ the post-boundary
 quantities
- threat: 'Spurious-arm thinness - the structural cause of the first round's
 ↳ collapse: any
 ↳ confirmatory quantity whose variance is dominated by the spurious arm is
 ↳ unpowerable at
 honest N'
 addressed_by: 'No confirmatory quantity now touches the spurious arm:
 ↳ theta2a is a true-arm-only
 ↳ contrast; the spurious arm serves as falsification gate F2 (whose job
 ↳ needs bounding,
 ↳ not detection, and whose operating characteristics are verified at Stage
 ↳ 0) and as the
 ↳ exploratory H2_diff; spurious-arm death caps the H2 verdict rather than
 ↳ silently weakening
 a test'
- threat: Residual underpowering if realized N, realized k, or the realized
 ↳ OOH share falls
 ↳ short of the sampling plan
 addressed_by: Pre-outcome Stage-2 realized-N gate with binding N*
 ↳ thresholds from the re-run
 ↳ deposited simulation; blind re-simulation on realized
 ↳ Ns/k-distribution/baseline moments
 ↳ and realized OOH share; demotion-only resolution before any outcome
 ↳ access; realized-power/MDE
 table published regardless
falsification_tests:
- 'Falsification gate F2 (spurious-closure placebo arm): theta2b with the
 ↳ pre-committed disconfirmation
 ↳ trigger; outside the Holm family; verified operating characteristics
 ↳ registered (pass 0.886
 hypothesis-true / 0.978 inert; trigger 0.79 at floor-sized non-inert
 ↳ placebo, at 45,000

```

spurious sets - the pre-matching worker count at the 30% band lower bound;  
 → the registered  
 90% match-rate floor implies ~40,500 sets, so these characteristics are  
 → computed at a ~10%  
 flattered lower-bound N); Stage-2 blind re-verification at realized N'  
 - 'c-2 pre-baseline placebo panel: exposed-minus-control contact difference  
 → across the four  
 quarters of year c-2 with a numeric failure threshold - fails if and only  
 → if significant  
 at  $\alpha = 0.05$  AND  $|\text{estimate}| > 1\%$  of control baseline mean, the single  
 → registered condition  
 identical wherever this gate appears; one pre-specified re-match, then kill  
 → of the registered  
 event-study tier'  
 - GRAFT 2 - week-0 boundary-step diagnostic on the weekly descriptive series  
 → (settlement-lag  
 contamination bound on any timing-shape claim)  
 - GRAFT 3 - successor-share threshold sensitivities (20/80, 40/60 alongside  
 → primary 30/70)  
 for  $\theta_{2b}/\theta_{2d}$  sign stability  
 - GRAFT 4 - s(w) bimodality check at Stage 1 (ambiguous mid-zone < 35% of  
 → terminal events)  
 - GRAFT 6 - later-presence differential trigger (2pp) forcing best/worst-case  
 → bounds on all  
 post-boundary conclusions  
 - GRAFT 9 - control forward-window (c+2/c+3) contamination sensitivity with  
 → registered direction  
 note (exclusion should weakly increase magnitudes; a decrease is flagged as  
 → diagnostic)  
 - 'Classification cross-validation: worker-flow classes vs IDAS-native flags  
 → (ENDNEDL/IDFREM/REGELFR/FRELFREM/ANDELUDS),  
 >= 75% agreement kill floor, full cross-tabulation released'  
 narrative: 'This study asks whether Danish workers'' primary-care billing  
 → already rises while  
 their workplace is dying - before any separation is visible in annual  
 → employment records  
 - and whether the post-closure excess attaches to genuinely destroyed jobs  
 → rather than to  
 workplaces that die only on paper. The instrument is the weekly clock of  
 → Danish health-insurance  
 billing (SSSY, HONUUGE fee-week) held against annually dated establishment  
 → terminal events  
 from the IDA workplace registers. For closure cohorts 2009-2023 (entirely  
 → inside the post-2008  
 IDA key era), every worker with primary November-status employment in year  
 → c-1 at a workplace  
 that disappears between years c and c+1 is classified by the worker-flow  
 → canon of Benedetto  
 et al. (2007) and Hethey-Maier & Schmieder (2013): true closure when no  
 → successor absorbs  
 the workforce (max successor share below 0.30), spurious closure when an  
 → en-bloc successor  
 exists (share at or above 0.70), the ambiguous mid-zone excluded and its  
 → mass monitored  
 by a pre-registered bimodality check. Cohorts are frozen at November of  
 → c-1, so the Schwerdt  
 sinking-ship channel cannot condition the sample. Each exposed worker is  
 → riskset-matched  
 to up to three stable-workplace controls within exact strata (cohort, sex,  
 → age band, NACE

section, size class, region) on Mahalanobis distance over a full year of  
 ↪ pre-baseline billing  
 history plus tenure and wage decile. The estimator is a stacked matched-set  
 ↪ difference-in-differences  
 event study with set and cohort-by-calendar-quarter fixed effects,  
 ↪ event-workplace clustering,  
 a c-2 placebo panel with a numeric failure threshold, and a settlement-lag  
 ↪ calibration of  
 the fee-week clock.\n\nThree confirmatory quantities form the Holm family,  
 ↪ each an honest  
 operationalization of a frozen bank hypothesis.  $\theta_1$  (H1) is the  
 ↪ pre-boundary excess in  
 the last two quarters of the terminal year - worded strictly as  
 ↪ pre-boundary excess relative  
 to an annually dated boundary, never announcement-dated anticipation; the  
 ↪ frozen statement's  
 'growing as the boundary approaches' rider is a gated exploratory  
 ↪ contrast, because its  
 ramp-implied floor is verified undetectable at any honest N.  $\theta_{2a}$  (H2)  
 ↪ is the true-closure  
 arm's post-year excess over its own matched controls - the excess the  
 ↪ frozen statement  
 asserts exists 'through the first post-closure year' - while the  
 ↪ statement's comparative  
 quantifier is enforced by a registered falsification gate: the  
 ↪ spurious-closure arm's excess  
 is estimated, bounded, and equipped with a pre-committed disconfirmation  
 ↪ trigger (spurious  
 excess flagged positive and at least half the true-arm excess disconfirms  
 ↪ H2's comparative  
 reading regardless of  $\theta_{2a}$ ), with the strict differential reported as an  
 ↪ exploratory  
 estimate. This restructuring follows the verified Stage-0 diagnosis: a  
 ↪ differential whose  
 variance is dominated by a spurious arm one-third the size of the true arm  
 ↪ needs roughly  
 5.5 times any honest variance budget - it is a falsification device, not an  
 ↪ estimation target.  
 $\theta_3$  (H3) is the out-of-hours versus daytime relative-excess gap,  
 ↪ unchanged in construct;  
 what changed is only its power basis, recalibrated from a self-declared  
 ↪ over-conservative  
 3% out-of-hours base rate to the conservative end of the published Danish  
 ↪ range (simulated  
 at 8.0% of daytime GP volume = 7.4% of all GP contacts, against a published  
 ↪ 8-12% contact  
 share; Christensen & Olesen 1998; the LV-KOS 2011 studies), with an  
 ↪ outcome-neutral check  
 verifying the realized share sits in the published range before the  
 ↪ confirmatory run.\n\nPower  
 is secured by honest enlargement, not by floor inflation: every floor  
 ↪ effect is unchanged  
 (2% relative excess for  $\theta_1$  and  $\theta_{2a}$ , below the Hamad-Modrek-Cullen  
 ↪ and Kuhn-Lalive-Zweimüller  
 magnitudes; a 5-point composition gap for  $\theta_3$ ), and the sampling plan  
 ↪ moves from the  
 first registration's self-handicapped half-of-lower-bound to the full  
 ↪ derived expectation's  
 lower bound - 15 cohorts, size cap 1000, roughly 135,000 matched sets - on  
 ↪ a published closure-rate

basis, with 1:3 matching cutting control-side variance by a third. The  
 → license for that  
 sizing is a demotion-only realized-N gate: after matching and before any  
 → post-baseline outcome  
 is touched, the deposited simulation is re-run blind on realized set  
 → counts, realized match  
 ratios, baseline moments, and the realized out-of-hours share; any member  
 → short of 0.95  
 power at its floor is demoted to exploratory on the spot, the Holm family  
 → shrinks by the  
 registered fixed-point rule, and nothing can ever move the other way. The  
 → claim ceiling  
 is associational and binding - associations and consistency language only,  
 → no effect/impact/cause  
 vocabulary, no phenomenon-discovery or Danish-first framing, outcomes  
 → always utilization  
 volume, fees, and timing composition, never health states - and the  
 → protocol publishes the  
 records of both power runs, the demotion trail, and the remediation note  
 → itself, so the  
 registration's history is part of its evidence.'

## 2. ENVIRONMENT

Software stack (record exact versions in the run manifest at Stage 0 and freeze them):

- **python3 (>= 3.11) with polars** for ALL large-table preprocessing. Lazy scans only (`scan_parquet/scan_csv`) with projection and predicate pushdown; semi-joins against the cohort PNR roster before any collect; streaming collects for aggregations. Never materialize a person x week or person x month panel beyond need: the only dense person-time table permitted is matched members x 16 event-quarters (~540k true members + ~180k spurious members, x 16 = under 12M rows). Zero-filling happens at the quarter level for panel members only, never by cross-joining the population against a calendar.
- **R (>= 4.3) with fixest, invoked via Rscript**, for every fixed-effects estimation: `feols` (OLS, the pre-committed primary) and `fepois` (Poisson QML robustness). CR1 cluster-robust SEs on event workplace. No estimation in python; no data manipulation in R beyond reading the analysis parquet and writing tidy coefficient tables.
- **matplotlib** for all figures; **booktabs LaTeX** for all tables (no vertical rules). Every number in every exhibit is written by script from a parquet/CSV output — never hand-edited.
- **Memory discipline:** process cohort-by-cohort (`c = 2009..2023`) where possible; checkpoint every numbered pipeline step as parquet under `interim/` with a fixed name (given per step below); SSSY scans filter on (PNR in roster) AND (SPEC2 in frozen lists) at scan level; the full population is touched only by the aggregate coverage benchmark (step 16), which is computed as a streaming groupby per year, never collected at row level.
- **Seeded randomness:** a single root seed 20260831 (per the frozen registration) initializes every RNG: matching processing order, Mahalanobis tie-breaks, the

999-draw cluster bootstrap, and the Stage-2 blind re-simulation. No unseeded randomness anywhere.

- **Run manifest:** an append-only JSONL (`manifest/run_manifest.jsonl`), one record per pipeline step per run: step id, UTC timestamp, code version hash, input file(s) + row counts, output file + row count, parameters used, validation check results, PASS/FAIL. Every step below MUST log its row counts; a step whose check fails writes FAIL and halts downstream steps (Section 6.3). The manifest (aggregate counts only, event counts rounded per the disclosure rules) is part of the released audit trail.
- **Directory convention:** `raw/` (DST-delivered), `interim/` (parquet checkpoints), `analysis/` (panels), `out/tables/`, `out/figures/`, `manifest/`. All paths in code are relative to the project root inside the research environment; no absolute paths in any released script.
- **Frozen-code discipline:** the Stage-4 estimation code is frozen at the end of Stage 0 (hash recorded in the manifest); Stage 4 is a single run of that code. The sensitivity suite runs once and is reported together. Any code change after Stage 0 is a deviation and must be disclosed in running prose with reason and approver.

---

## Source disclosure rules

```
environment: DST Forskningservice (no prior data access; gatekeeper
↳ certification per RR
 secondary-data rules). All analysis occurs inside the environment; only
 ↳ aggregate outputs
 leave. No microdata, workplace identifiers (ARBNR/CVRNR/LBNR), or
 ↳ person-level values are
 released.
minimum_cell_size: 5
minimum_cell_rules: no released statistic based on fewer than 5 persons or
↳ fewer than 3 workplaces;
 event-count tables rounded to the nearest 10; weekly trajectory figures
 ↳ pool >= 100 events
 and >= 5,000 person-weeks per plotted point; fee (BRUHON) aggregates
 ↳ subject to a 50% single-workplace
 dominance check; min-cell suppression applied BEFORE Holm reporting so
 ↳ suppression can never
 be outcome-selective
outputs_leaving_environment:
- model coefficients with SEs/CIs
- matching balance tables and realized-k distribution
- calibration and benchmark tables (coverage, linkage, HONUGE lag, OOH share)
- figure aggregates (weekly trajectory series under the pooling rules)
- event counts rounded to the nearest 10
- Stage-0 power/MDE tables - BOTH rounds (failed and remediated) - and the
↳ Stage-2 blind re-simulation
 output with the demotion trail
- the c-2 placebo, gate F2 outputs, GRAFT-2/4 diagnostics, GRAFT-6
↳ later-presence rates, E4
 attrition panel, and all outcome-neutral check outputs (the full audit
 ↳ trail exits under
 the same rules)
- the analysis and power scripts
```

statement: 'All analysis occurs inside DST Forskningssservice; only aggregate  
 ↳ outputs leave.  
 Rules applied to every released object: (1) no statistic based on fewer  
 ↳ than 5 persons or  
 fewer than 3 workplaces; (2) event-count tables rounded to the nearest 10;  
 ↳ (3) weekly trajectory  
 figures pool  $\geq 100$  events and  $\geq 5,000$  person-weeks per plotted point; (4)  
 ↳ fee (BRUHON)  
 statistics released only as cell means/coefficients over min-N cells, with  
 ↳ a dominance check  
 that no single workplace contributes  $> 50\%$  of any released fee aggregate;  
 ↳ (5) released objects  
 are limited to: model coefficients with SEs/CIs, balance tables,  
 ↳ calibration/benchmark tables,  
 figure aggregates, event counts (rounded), the Stage-0 (both rounds) and  
 ↳ Stage-2 power/MDE  
 tables with the demotion trail, and the analysis/power scripts - never  
 ↳ microdata, workplace  
 identifiers (ARBNR/CVRNR/LBNR), or person-level values; (6) min-cell  
 ↳ suppression is applied  
 before Holm reporting so that suppression can never be outcome-selective;  
 ↳ (7) the c-2 placebo,  
 gate F2 outputs, the GRAFT-2/4 diagnostics, the GRAFT-6 later-presence  
 ↳ rates, the E4 attrition  
 panel, and all outcome-neutral check outputs are released under the same  
 ↳ rules, ensuring  
 the full audit trail exits the environment.'

## Complete registration fields

```
id: atom-000001
topic: Event classification
atom_result: prereg
atom_strength: null
atom_shape: null
evidence_paths: []
concepts: []
literature: []
bonds: []
preregistration:
 version: 1
 mode: adaptive
 question: Can true and administrative workplace closures be distinguished
 within the registered four-table scope?
 rationale: Source-existing shared validity judgment, required by several
 subsequent questions.
 assumption: |-
```

Associational - matched event-study contrasts, a contrastive  
 ↪ falsification gate, and a composition contrast; descriptive for  
 ↪ trajectory panels and the bounded priority claim. Binding vocabulary  
 ↪ rules: 'is associated with' / 'consistent with' / 'suggests' only;  
 ↪ 'effect', 'impact', 'causes' prohibited throughout; anticipation  
 ↪ always worded as 'pre-boundary excess' relative to the annually dated  
 ↪ closure-year boundary, never as announcement-dated anticipation (S8);  
 ↪ no phenomenon-discovery or Danish-first framing (S1-S3); H2 novelty  
 ↪ confined to the placebo's application to health billing (S7). H2  
 ↪ verdict semantics are fixed in advance: 'consistent with H2' requires  
 ↪ BOTH the confirmatory core (theta2a Holm-significant, predicted  
 ↪ direction) AND a passing falsification gate F2; the strict  
 ↪ comparative magnitude is only ever reported as an exploratory  
 ↪ estimate with its CI. Exploratory results (H1\_trend, H2\_diff, E1, E3,  
 ↪ E4, and any gate-demoted member) carry the word 'exploratory' in the  
 ↪ heading and in every sentence stating them. The prose label never  
 ↪ exceeds the evidence label.  
 Compatibility ratings below are neutral import metadata, not a new  
 ↪ scientific evaluation or empirical scores.  
 target\_strength: descriptive  
 target\_shape: comparison  
 ratings:  
   importance: 3  
   novelty: 3  
   feasibility: 3  
   data\_advantage: 3  
   informativeness: 0.5  
 tables:  
 - IDAN  
 - IDAS  
 activation: 'TRUE'  
 verdicts:  
 - key: classifiable  
   region: At least 500 true-closure events; source-required classification  
     audits resolved. Preserve spurious-arm kill or availability as a  
       ↪ component,  
       without killing the true arm.  
   says: The true-closure event frame is usable within the audited  
     ↪ classification  
       scope; no health result exists.  
 - key: project\_stop  
   region: Fewer than 500 true-closure events after the registered Stage-1  
     procedure.  
   says: Source project viability kill; do not estimate the registered  
     ↪ outcome  
       study.  
 - key: boundary  
   region: Required source-defined inputs or scientific definitions are  
     ↪ demonstrably  
       unavailable; retain available components and the exact unresolved scope.  
   says: An evidenced boundary, not a statistical null. A missing  
     ↪ implementation  
       or timeout alone is operational no\_result, not this state.  
 - key: otherwise  
   region: otherwise  
   says: The registered judgment is unresolved; report all completed  
     ↪ components  
       without inventing support, equivalence or absent results.  
 protocol: |

# Can true and administrative workplace closures be distinguished within  
↳ the registered four-table scope?

## Source fidelity and implementation order

This is a format conversion of DS1.1, not a new study or independent  
↳ acceptance.  
Source audit: minor\_revisions. No observed Danish outcomes are supplied.  
The source execution plan explicitly states that final prereg.yaml governs  
disagreements. Exact current YAML selections below therefore govern older  
execution-plan excerpts; these excerpts are retained as implementation  
↳ detail,  
not a competing registration. See docs/provenance/source-discrepancies.md.  
The complete shared data/validity contracts are embedded in Atoms 1--3;  
↳ read  
those native protocols when executing a downstream Atom, even without  
↳ docs/.  
Do not substitute a scientific default for an unresolved source  
↳ contradiction.

Preserve source stages 0 -> 1 -> 2 -> 3 -> 4 -> 5 inside the  
↳ implementation.  
Source Stage-0 simulations are historical design evidence, not study  
↳ outcomes;  
do not silently redo or improve them. Gate 14 must be evaluated from  
↳ baseline  
data before post-baseline outcomes. All surviving Stage-4 quantities form  
↳ ONE  
Holm family. The three question-level Atoms may share the same frozen  
↳ family  
fit and tidy results; do not apply Holm separately per Atom or adjudicate  
↳ a  
member before the complete surviving family's p-values exist. Cache/reuse  
identical governed intermediate results with provenance when available; a  
direct-dispatch call must obtain or compute its exact source-defined  
↳ inputs,  
not pretend its predecessors succeeded. Stage-5 internal riders remain  
exploratory, after the Stage-4 family is frozen, even when housed in  
↳ H1/H2.

Reference files under docs/execution/ are the ORIGINAL SKELETONS, not  
↳ tested  
turnkey implementations. Adapt paths and complete engineering in a  
↳ governed  
workspace; preserve source scientific rules and disclose unresolved  
↳ choices.  
Write a result record with component estimates, uncertainty, family  
↳ eligibility,  
source interpretation row, diagnostics and boundary/skip dispositions.  
↳ Released  
outputs follow source disclosure rules; confidential rows stay governed.  
Local native states summarize the source judgment; neither a state name  
↳ nor  
source power forecasts prove equivalence or causal identification. A  
↳ bounded  
claim requires its actual interval and registered rule. Retain unresolved  
↳ source  
wording conflicts; do not convert mere nonsignificance into scientific  
↳ null.

## Frozen scope and vocabulary

```
```yaml
claim_class_ceiling: associational
statement: 'Associational - matched event-study contrasts, a contrastive
↪ falsification gate,
and a composition contrast; descriptive for trajectory panels and the
↪ bounded priority claim.
Binding vocabulary rules: 'is associated with' / 'consistent with'
↪ / 'suggests' only;
'effect', 'impact', 'causes' prohibited throughout; anticipation
↪ always worded as
'pre-boundary excess' relative to the annually dated closure-year
↪ boundary, never as announcement-dated
anticipation (S8); no phenomenon-discovery or Danish-first framing
↪ (S1-S3); H2 novelty confined
to the placebo's application to health billing (S7). H2 verdict
↪ semantics are fixed in
advance: 'consistent with H2' requires BOTH the confirmatory core
↪ (theta2a Holm-significant,
predicted direction) AND a passing falsification gate F2; the strict
↪ comparative magnitude
is only ever reported as an exploratory estimate with its CI.
↪ Exploratory results (H1_trend,
H2_diff, E1, E3, E4, and any gate-demoted member) carry the word
↪ 'exploratory' in the
heading and in every sentence stating them. The prose label never
↪ exceeds the evidence label.'
frozen_parameters:
seed: 20260831
gp_code_list: daytime GP = SPEC2 in {70..80}
ooh_code_list: out-of-hours GP = SPEC2 in {81, 82, 83, 84, 89}
ooh_robustness_classifier: adds YDLTID in {2, 3, 4, 5, 7, 8, 9}
spec2_evening_gp_disclosure: 'Registered choice, disclosed: the parsed
↪ value labels mark
SPEC2 71 ("General practitioner, consultation, evening etc."), 73
↪ ("General practitioner,
telephone consultation, evening etc."), and 75 ("General
↪ practitioner, home visit, evening
etc.") as explicitly off-schedule GP services; they nonetheless
↪ remain on the daytime
side of the PRIMARY classifier because H3's out-of-hours construct
↪ and its published
power basis are keyed to the laegevagt out-of-hours service (SPEC2
↪ 81-84, 89), the quantity
the published Danish 8-12% share describes. Attenuation direction,
↪ both readings: counting
evening GP as daytime attenuates theta3 toward zero - conservative
↪ against the positive
H3 conclusion, anti-conservative for the null-bounding
↪ "compositionally unremarkable"
reading, whose Stage-2 statement therefore carries a mandatory
↪ caveat. In-environment
validation must verify that the YDLTID-augmented robustness
↪ classifier captures SPEC2
71/73/75 rows (their evening time codes 2-5); if YDLTID proves
↪ unpopulated on those rows,
the robustness classifier - and only the robustness classifier - is
↪ amended mechanically
```

```

and with disclosure to include SPEC2 71/73/75 directly. The SPEC2
  ↳ 71/73/75 volume share
is reported, and both classifiers appear side by side in Table T7.'
successor_share_thresholds: 'true closure: ENDNEDL final closure AND
  ↳ worker-flow max-successor
share s(w) < 0.30; spurious closure: forward identity via
  ↳ IDFREEM/REGELFR/FRELFREM AND
s(w) >= 0.70; ambiguous mid-zone excluded; GRAFT-3 sensitivities
  ↳ 0.20/0.80 and 0.40/0.60'
quarter_mapping: 'ISO week of HONUGE: 1-13 / 14-26 / 27-39 / 40-52; ISO
  ↳ week 53 to Q4'
winsorization: person-quarter outcomes at the 99.9th percentile within
  ↳ era x calendar quarter
matching: 1:3 nearest-neighbor Mahalanobis within exact strata, without
  ↳ replacement across
sets, target k = 3, minimum k = 1; calipers |dif baseline contacts|
  ↳ <= 3 and |dif log(1+BRUHON)|
  ↳ <= 0.5; exposed processed in seeded random order (seed 20260831)
software: R >= 4.3, fixest::feols (OLS) and fixest::fepois (Poisson QML
  ↳ robustness); versions
frozen in the registered script
note: Code lists and thresholds frozen before any data access; code
  ↳ labels quoted verbatim
from parsed DST value labels and re-validated in-environment before
  ↳ estimation (frozen
fallback rules in contingent_decisions).
...
1. **IDAS ingest, 2007-2024.** Input: raw IDAS. Keep the D1 variable
list; types cast; one row per (LBNR, year). Output:
`interim/idas_panel.parquet`. Check: all 18 years present; row count
per year within ±50% of the ~227k/year inventory basis; no duplicate
(LBNR, year); log per-year counts.
2. **Workplace identity chains.** Input: `idas_panel.parquet`. Build
identity linkage via IDFREEM (forward) and IDTILB (backward); resolve
each workplace-year to a chained identity. Output:
`interim/workplace_identity.parquet`. Check: forward/backward
reciprocity rate reported; no cycles; share of years with broken
chains reported.
3. **Terminal-event candidates, cohorts c in 2009-2023.** Input: steps
1-2. Flag identities present in year c with ANTNOV >= 5 and absent in
year c+1 (IDAS extends through 2024, so c = 2023 is checkable).
Attach ENDNEDL, REGELFR, FRELFREM, ANDELOPS, ANDELUDS, AFGRATE,
ANTALXFR. Output: `interim/terminal_events.parquet`. Check: per-cohort
event counts logged (rounded to 10 in any released object); validate
the ENDNEDL flag-year convention against the identity panel - if it
differs from "flagged in the workplace's last active year", apply the
registered re-anchoring (c := last IDAS year with ANTNOV >= 5 for the
identity) mechanically and report the recalibration verbatim.
4. **IDAN ingest, 2007-2024 (slim).** Input: raw IDAN. Keep the D2
confirmatory variables. Output: `interim/idan_slim.parquet`. Check:
~3.2M rows/year (±50%); PNR and workplace keys non-null; per-year
counts logged.
5. **Primary-job assignment, per person x year.** Input:
`idan_slim.parquet`. Primary job = largest BREDT_LOEN_BELOEB in the
year; ties broken by larger LOENTIMER, then lower ARBNR
(deterministic). Output: `interim/primary_jobs.parquet`. Check:
exactly one primary employment per person-year; re-run idempotence
(identical output hash on repeat); multi-job share reported.
6. **Workplace sector assignment.** Input: steps 4-5. Modal
ARB_SEKTORKODE among each workplace's year c-1 primary employees;

```

flag private-sector workplaces. Output:
`interim/workplace\_sector.parquet`. Check: private share of size-5+ workplaces reported; workplaces with no c-1 employees flagged, not dropped silently.

7. ****Worker-flow successor shares s(w).**** Input: steps 3, 5. For each terminal event: among the event workplace's year-c November primary employees, the largest share whose year-c+1 primary workplace is a single successor identity. Computed on the year-c workforce - deliberately independent of the frozen c-1 cohort. Output:
`interim/successor\_shares.parquet`. Check: every terminal event carries s(w) or a documented reason (no year-c primary employees); distribution summary logged.

8. ****Event classification + cross-validation (Stage-1 gate).**** Input: steps 3, 7. True closure = ENDNEDL final closure AND s(w) < 0.30; spurious closure = forward identity via IDFREM/REGELFR/FRELFREM AND s(w) >= 0.70; else ambiguous-excluded. Also classify under the GRAFT-3 threshold pairs (0.20/0.80, 0.40/0.60). Cross-tabulate worker-flow classes against IDAS-native flags (ENDNEDL/IDFREM/REGELFR/FRELFREM/ANDELUDS). Output:
`interim/events\_classified.parquet` + released cross-tab. Checks (all Stage-1 stop rules): true-closure events >= 500 (project kill below); spurious events >= 150 (spurious-arm kill below); agreement >= 75% (75-90% proceeds with worker-flow classes as registered, full disagreement table published); GRAFT-4 s(w) bimodality with ambiguous mid-zone < 35% of terminal events; annual true-closure rate among private size-5+ workplaces in [1%, 8%] every cohort year; spurious/true ratio in [0.1, 1.5]. Any violation triggers the classification audit before any further stage. Report projected exposed N against the sampling plan.

1. DATA ACQUISITION

Request the following FOUR register datasets from DST - the study's hard data scope; nothing else may be requested or used. Names are the documented TIMES/HKV identities. Every variable below is verified against the registration's mapping bundle (`times\_variables\_bilingual.csv`, `hkv\_variables\_bilingual.csv`), with four disclosed exceptions verified only in the DST Forskningservice four-register inventory for the analysis server (PNR on all four registers; KOM on IDAS; VAGTOMR and SIKREAMT on SYSI) - these are flagged "inventory-verified" below and the disclosure must be carried into the report. DST delivers PNR as an anonymized project-specific person key; "PNR" below means that key.

1.1 Datasets and variables

****D1. IDA workplaces**** (Danish: IDA arbejdssteder; TIMES
`ida-databasen/ida-arbejdssteder`, HKV `ida-arbejdssteder`, valid from 01-01-1980). Request the 1980-2024 panel; the confirmatory tier consumes 2007-2024, the E1 exploratory tier 1990-2005. Unit: workplace x year.

Variable	Documented long name	Validity used	Role
LBNR	Workplace serial number	1980-2024	workplace identity key;
↪	identity chains via IDFREM/IDTILB; identity-disappearance events		
↪	between c and c+1		
ARBNR	Arb number	2007-2024	workplace key, post-2008 vintage
↪	(entire confirmatory tier)		
CVRNR	CVR number	2007-2024	firm key, post-2008 vintage
ENDNEDL	Final closure of workplace	through 2023	true-closure flag
↪	(with s(w) < 0.30); its 2023 end caps the last cohort		

| IDFREEM | Identity of place of work forward in time | 1980-2023 |
 ↳ forward-continuation identity for spurious-closure classification;
 ↳ chain construction |
 | IDTILB | Identity of place of work backward in time | 1981-2024 |
 ↳ backward identity chain (control stability, event-panel construction)
 ↳ |
 | REGELFR | Rule regarding retained workplace the following year |
 ↳ 1980-2023 | forward-identity rule for spurious-closure classification
 ↳ |
 | FRELFR | Firm-internal/external closure via absorption | 1980-2023 |
 ↳ en-bloc absorption indicator for spurious-closure classification |
 | ANDELOPS | Share of absorbed among employees | 1981-2024 | IDAS-native
 ↳ absorption share; classification cross-validation only |
 | ANDELUDS | Share of employees split off to another workplace among
 ↳ employees | 1980-2023 | IDAS-native split-off share; cross-validation
 ↳ only |
 | AFGRATE | Exit rate in the subsequent annual period, in per cent |
 ↳ 1980-2023 | stable-control definition (AFGRATE < 30%);
 ↳ mass-separation exclusion |
 | ANTALXFR | Change in the number of employees in the following year |
 ↳ 1980-2023 | stable-control definition (\\|ANTALXFR\\| <= 30%) |
 | ANTNOV | Number of employees as at November | 1980-2024 | size
 ↳ eligibility 5-1000; event presence (ANTNOV >= 5); exact-stratum size
 ↳ classes |
 | ARB_HOVED_BRA_DB07 | Industry of the workplace | 2008-2024 | NACE
 ↳ section exact stratum |
 | KOM | Municipality of the workplace | 1980-2024 | workplace
 ↳ municipality mapped to the 5 post-2007 regions (exact stratum).
 ↳ INVENTORY-VERIFIED on IDAS (absent from the public TIMES/HKV parses
 ↳ for this register) - disclose |
 | ARBGNR | Employer number | 1980-2007 | legacy workplace key; E1
 ↳ exploratory only |
 | ARBSTK | Workplace code | 1980-2007 | legacy workplace key; E1
 ↳ exploratory only |

****D2. IDA employments**** (Danish: IDA ansaettelser; TIMES
 `ida-databasen/ida-ansaettelser`; HKV analogues
 `beskaeftigelsesoplysninger-der-vedroerer-ida-ansaettelser` and
 `loenforhold-der-vedroerer-ida-ansaettelser`, valid from 01-01-1980).
 Request 1980-2024; confirmatory tier consumes 2007-2024 (cohort anchoring
 at November c-1, worker flows into c+1). Unit: person x employment
 relation x year.

Variable	Documented long name	Validity used	Role
PNR	De-identified person number	all years	person spine joining
↳ IDAN, SSSY, SYSI.	INVENTORY-VERIFIED (absent from public parses) -		
↳ disclose			
LBNR	Workplace serial number	all years	employment-to-workplace
↳ link (with ARBNR/CVRNR, 2008+)			
ARBNR	Arb number	2008-2024	workplace key on the employment
↳ relation			
CVRNR	CVR number	2008-2024	firm key on the employment relation
ANSAAR	Year of employment at the workplace	all years	tenure (c-1
↳ minus ANSAAR)	inclusion floor ANSAAR <= c-2; Mahalanobis component		
BREDT_LOEN_BELOEB	Broad salary amount	2008-2024	primary-job rule
↳ (largest in c-1); baseline wage decile (Mahalanobis; E3			
↳ heterogeneity)			
LOENTIMER	Paid hours (LOENTIMER)	2008-2024	deterministic
↳ primary-job tie-break (larger LOENTIMER, then lower ARBNR)			

```

| ARB_SEKTORKODE | Workplace sector code | 2008-2024 | private-sector
↳ filter via modal sector among year c-1 employees |
| ANSXFREM | Change of employment up to the following year | 1980-2023 |
↳ E4 attrition accounting (forward-separation flag) |
| ARBGNR | Employer number | 1980-2007 | legacy key; E1 only. Documented
↳ publicly on the workplace register; carried on legacy IDAN vintages
↳ per the Forskningservice inventory - disclose |
| ARBSTK | Workplace code | 1980-2007 | legacy key; E1 only (same
↳ disclosure) |
| STILL | Occupational position | 1980-2007 | legacy job characteristic
↳ for E1-era matching (SYSI lacks age/sex) |
| TILKNYT | Attachment to primary workplace | 1980-2007 | legacy
↳ primary-employment marker; E1 only |
| ANSDAGE | Number of days employed | 1980-2007 | legacy job
↳ characteristic for E1-era matching |
| TIMELON | Hourly wage in the employment | 1980-2007 | legacy wage
↳ measure for E1-era matching |

```

****D3. Health insurance statistics, current service series (SSSY)****
(Danish: Sygesikringsstatistik; TIMES `sygesikringsstatistik`; HKV
`sygesikring---ydelser` / `sygesikring---honorarer` /
sygesikring-kontakter datasets; SPEC2 valid from 01-01-2005). Request
2005-2025 service rows; the pipeline consumes 2007-2025 (baselines c-3
onward, post-years through 2024/2025). Unit: service row.

```

| Variable | Documented long name | Validity used | Role |
|---|---|---|---|
| PNR | De-identified person number | all years | person spine.
↳ INVENTORY-VERIFIED - disclose |
| HONUGE | Fee week | all years | the weekly clock; person x HONUGE week
↳ aggregation; ISO-week-to-quarter mapping |
| KONTAKT | Contact | 2005-2025 | row-level contact marker summed to
↳ quarterly GP contact counts (frozen fallback if in-environment
↳ validation fails, Section 4.6) |
| YDLANT | Number of services within the specialty | all years | service
↳ counts (descriptives; E1-era contact proxy weighting) |
| BRUHON | Gross fee to doctor etc. | all years | gross fees: gated H2
↳ fee secondary, E3 fee analogues, Mahalanobis log(1+BRUHON); nominal,
↳ compared only within cohort x calendar-quarter cells |
| SPEC2 | 2-digit specialisation | 2005-2025 | frozen GP (70-80) and OOH
↳ (81-84, 89) code lists |
| YDLTID | Time code for the benefit | all years | robustness OOH
↳ classifier; missing/"not relevant" classified daytime (conservative
↳ against H3) |
| AFRPER | Settlement period | all years | settlement-lag
↳ characterization (descriptive) |
| BEHANDLINGSDATO | The treatment date for the service | 2021-2025 only |
↳ HONUGE settlement-lag calibration; can calibrate but never replace
↳ HONUGE |
| SIKGRUP | Health insurance group | all years | exclusion of modal
↳ group-2 insured (different billing regime) |
| ALDERIMP | Age at end of period, incl. imputed | all years | age bands
↳ for exact strata (explicit unknown stratum); eligibility 25-59 |
| KOENIMP | Sex incl. imputed | all years | sex stratum (explicit unknown
↳ stratum); E3 heterogeneity |
| YDERNR | Provider number (ydernummer) | all years | frozen fallback
↳ contact definition (distinct PNR x week x YDERNR x code-group cells);
↳ never exported |

```

****D4. Health insurance statistics, legacy closed series (SYSI)****

(Danish: Sygesikringsstatistik; documented under TIMES
`sygesikringsstatistik` and the HKV sygesikring datasets, valid from
01-01-1990; series closed 2005). Request 1990-2005. Used ONLY in the
exploratory SYSI-era replication E1 (cohorts 1992-2003); never pooled
with the confirmatory era across the 2007/08 IDA vintage break.

Variable	Documented long name	Validity used	Role
PNR	De-identified person number	all years	person spine.
INVENTORY-VERIFIED	- disclose		
HONUFE	Fee week	all years	weekly clock, legacy era
SPECIALE	6-digit speciality	all years	era GP/OOH classification
SPEC2	does not exist pre-2005		era code list frozen at Stage 0
YDLANT	Number of services within the specialty	all years	E1
contact proxy	service-row counts weighted by YDLANT		KONTAKT does not exist pre-2005
BRUHON	Gross fee to doctor etc.	all years	era fee outcome
YDLTID	Time code for the benefit	all years	era time-of-day classification
VAGTOMR	Out-of-hours (laegevagt) service area	1997-2005	era OOH marker. INVENTORY-VERIFIED - disclose
AFRPER	Settlement period	all years	era settlement-lag characterization
SIKGRUP	Health insurance group	all years	era group-2 exclusion
SIKREKOM	Municipality of the insured person	all years	era patient geography (E1 only; SSSY has no patient geography)
SIKREAMT	County of the insured person	all years	era patient geography (E1 only). INVENTORY-VERIFIED - disclose

1.2 Join order, keys, and expected cardinalities

Person spine: PNR (IDAN x SSSY x SYSI). Workplace spine: LBNR with ARBNR/CVRNR (IDAS x IDAN, 2008+ vintage - the entire confirmatory tier); LBNR with ARBGNR/ARBSTK (pre-2008 - E1 only); workplace identity chains via IDFREM/IDTILB. Never join across the 2007/08 IDA vintage break.

Join order (confirmatory tier):

1. **IDAS(c) x IDAS(c+1) self-join** on workplace identity (LBNR chain): flags identity-disappearance events. 1:0/1 per identity. Expected scale: ~227k workplace rows/year (Forskningsservice inventory); terminal identity-disappearance events per cohort year in the low thousands before classification.
2. **IDAS <- IDAN** on (year, LBNR/ARBNR/CVRNR): workplace-year to employment rows, 1:many (mean ~14 employment rows per workplace-year; ~3.2M IDAN rows/year).
3. **IDAN(c) -> IDAN(c+1)** on PNR (primary jobs only): worker-flow successor shares. 1:0/1 per person.
4. **Cohort roster <- SSSY** on PNR: 1:many (person to service rows; ~88M SSSY rows/year population-wide; ~4-5 GP contacts/person-year, often multiple rows per contact - hence KONTAKT, not row counts).
5. **Matching**: exposed:controls = 1:up to 3, without replacement of controls across sets.
6. **Panel**: matched member x 16 event-quarters (Q1(c-2)..Q4(c+1)).

Expected cardinalities against the sampling plan (all realized counts go in the run manifest and exhibit T1): true-closure events ~500-1,700 per cohort year after classification (plan basis: 12,600 event workplaces over 15 cohorts); eligible exposed true-closure workers 10,000-20,000 per cohort, 150,000-300,000 total; >= 135,000 true 1:3 matched sets at the

>= 90% match-rate floor; spurious-closure workers 30-60% of the true count (45,000-90,000 before matching; >= 45,000 sets at plan). Kill floors far below expectation: 500 true events / 10,000 sets (project); 150 spurious events / 3,000 spurious sets (spurious arm).

Current schema and linkage

```

```yaml
registers:
- register: IDA workplaces (IDAS)
 register_da: IDA arbejdssteder
 documentation: TIMES ida-databasen/ida-arbejdssteder (The IDA database
 ↳ / IDA workplaces)
 and HKV ida-arbejdssteder, valid from 01-01-1980
 years: 1980-2024 panel; event cohorts c in 2009-2023 (confirmatory);
 ↳ ENDNEDL ends 2023
 variables:
 - name: LBNR
 long_name: Workplace serial number
 role: key
 use: workplace identity; identity chains via IDFREM/IDTILB;
 ↳ identity-disappearance events
 between c and c+1
 - name: ARENR
 long_name: Arb number
 role: key
 use: workplace key, 2007+ vintage (confirmatory era)
 - name: CVRNR
 long_name: CVR number
 role: key
 use: firm key, 2007+ vintage (confirmatory era)
 - name: ENDNEDL
 long_name: Final closure of workplace
 role: exposure
 use: true-closure flag (with s(w) < 0.30); validity through 2023 caps
 ↳ the last cohort
 - name: IDFREM
 long_name: Identity of place of work forward in time
 role: exposure
 use: forward-continuation identity for spurious-closure
 ↳ classification; identity chain
 construction
 - name: IDTILB
 long_name: Identity of place of work backward in time
 role: exposure
 use: backward identity chain (control-stability and event-panel
 ↳ construction); HKV valid
 from 01-01-1981
 - name: REGELFR
 long_name: Rule regarding retained workplace the following year
 role: exposure
 use: forward-identity rule for spurious-closure classification
 - name: FRELFREM
 long_name: Firm-internal/external closure via absorption
 role: exposure
 use: en-bloc absorption indicator for spurious-closure classification
 - name: ANDELOPS

```

```

 long_name: Share of absorbed among employees
 role: check
 use: IDAS-native absorption share; classification cross-validation
 ↪ against worker-flow
 s(w)
- name: ANDELUDS
 long_name: Share of employees split off to another workplace among
 ↪ employees
 role: check
 use: IDAS-native split-off share; classification cross-validation
- name: AFGRATE
 long_name: Exit rate in the subsequent annual period, in per cent
 role: control
 use: stable-control definition (AFGRATE < 30%); mass-separation
 ↪ exclusion
- name: ANTALXFR
 long_name: Change in the number of employees in the following year
 role: control
 use: stable-control definition (|ANTALXFR| <= 30%)
- name: ANTNOV
 long_name: Number of employees as at November
 role: control
 use: 'workplace size: eligibility 5-1000, exact-stratum size classes,
 ↪ event presence (ANTNOV
 >= 5)'
- name: ARB_HOVED_BRA_DBO7
 long_name: Industry of the workplace
 role: control
 use: NACE section exact stratum (2008+ vintage)
- name: KOM
 long_name: Municipality of the workplace
 role: control
 use: workplace municipality mapped to the 5 post-2007 regions (exact
 ↪ stratum)
 verification: verified in the DST Forskningservice four-register
 ↪ inventory for the analysis
 server; absent from the public TIMES/HKV parses - disclosed
- name: ARBGNR
 long_name: Employer number
 role: key
 use: legacy workplace key (pre-2008 vintage); E1 SYSI-era exploratory
 ↪ replication only
- name: ARBSTK
 long_name: Workplace code
 role: key
 use: legacy workplace key (pre-2008 vintage); E1 exploratory only
- register: IDA employments (IDAN)
 register_da: IDA ansaettelser
 documentation: TIMES ida-databasen/ida-ansaettelser (The IDA database /
 ↪ IDA employments);
 HKV analogues Beskaeftigelsesoplysninger / Loenforhold / Ledighed der
 ↪ vedroerer IDA ansaettelser,
 valid from 01-01-1980
 years: 1980-2024 panel; cohort anchoring at November of year c-1;
 ↪ worker flows into c+1
 variables:
- name: PNR
 long_name: De-identified person number
 role: key
 use: person spine joining IDAN, SSSY, SYSI

```

verification: verified in the DST Forskningservice four-register  
 ↪ inventory for the analysis  
 server; absent from the public TIMES/HKV parses - disclosed

- name: LBNR
  - long\_name: Workplace serial number
  - role: key
  - use: employment-to-workplace link (with ARBNR/CVRNR, 2008+)
- name: ARBNR
  - long\_name: Arb number
  - role: key
  - use: workplace key on the employment relation, 2008+
- name: CVRNR
  - long\_name: CVR number
  - role: key
  - use: firm key on the employment relation, 2008+
- name: ANSAAR
  - long\_name: Year of employment at the workplace
  - role: control
  - use: tenure (c-1 minus ANSAAR); inclusion floor ANSAAR <= c-2;  
 ↪ Mahalanobis component
- name: BREDT\_LOEN\_BELOEB
  - long\_name: Broad salary amount
  - role: control
  - use: primary-job rule (largest BREDT\_LOEN\_BELOEB in c-1); baseline  
 ↪ wage decile (Mahalanobis  
 component; E3 heterogeneity)
- name: LOENTIMER
  - long\_name: Paid hours (LOENTIMER)
  - role: control
  - use: deterministic primary-job tie-break (larger LOENTIMER, then  
 ↪ lower ARBNR)
- name: ARB\_SEKTORKODE
  - long\_name: Workplace sector code
  - role: control
  - use: private-sector filter via modal sector among year c-1 employees  
 ↪ (2008+)
  - verification: 'Exception, disclosed (converse of the inventory-only  
 ↪ variables): publicly  
 documented (TIMES ida-ansaettelser) but NOT present in the recorded  
 ↪ DST Forskningservice  
 four-register inventory extract for the analysis server, and the  
 ↪ claimed 2008+ validity  
 rests on the public parse alone. Server presence and coverage  
 ↪ validated at Stage 1 before  
 any classification work; frozen fallback registered in  
 ↪ analysis\_plan.contingent\_decisions  
 (private-sector filter from ARB\_HOVED\_BRA\_DB07, NACE DB07 sections  
 ↪ O/P/Q excluded).'
- name: ANSXFREM
  - long\_name: Change of employment up to the following year
  - role: check
  - use: E4 arm-comparative attrition accounting (forward-separation flag)
- name: ARBGNR
  - long\_name: Employer number
  - role: key
  - use: legacy key (pre-2008); E1 exploratory only
  - verification: carried on legacy IDAN vintages per the  
 ↪ Forskningservice inventory; the  
 public parses document it on IDA workplaces
- name: ARESTK

```

 long_name: Workplace code
 role: key
 use: legacy key (pre-2008); E1 exploratory only
 verification: carried on legacy IDAN vintages per the
 ↪ Forskningservice inventory; the
 public parses document it on IDA workplaces
- name: STILL
 long_name: Occupational position
 role: control
 use: legacy job characteristic for E1-era matching (SYSI lacks
 ↪ age/sex)
- name: TILKNYT
 long_name: Attachment to primary workplace
 role: control
 use: legacy primary-employment marker; E1 exploratory only
- name: ANSDAGE
 long_name: Number of days employed
 role: control
 use: legacy job characteristic for E1-era matching
- name: TIMELON
 long_name: Hourly wage in the employment
 role: control
 use: legacy wage measure for E1-era matching
- register: Health insurance statistics, current service series (SSSY)
 register_da: Sygesikringsstatistik (SSSY)
 documentation: TIMES sygesikringsstatistik (Health insurance
 ↪ statistics); HKV sygesikring
 - ydelser / honorarer / kontakter (SPEC2 valid from 01-01-2005)
 years: 2005-2025 service rows; extraction 2006-2025 (baselines c-3
 ↪ onward, post-years through
 2024/2025)
variables:
- name: PNR
 long_name: De-identified person number
 role: key
 use: person spine
 verification: verified in the DST Forskningservice four-register
 ↪ inventory for the analysis
 server; absent from the public TIMES/HKV parses - disclosed
- name: HONUGE
 long_name: Fee week
 role: outcome
 use: 'the weekly clock: person x HONUGE week aggregation,
 ↪ ISO-week-to-quarter mapping'
- name: KONTAKT
 long_name: Contact
 role: outcome
 use: row-level contact marker summed to quarterly GP contact counts
 ↪ (frozen fallback definition
 if in-environment validation fails)
- name: YDLANT
 long_name: Number of services within the specialty
 role: outcome
 use: service counts (descriptives; E1-era contact proxy weighting)
- name: BRUHON
 long_name: Gross fee to doctor etc.
 role: outcome
 use: 'gross fees: gated H2 fee secondary, E3 fee analogues,
 ↪ Mahalanobis log(1+BRUHON)

```

```

 component; nominal, compared only within cohort x calendar-quarter
 ↪ cells'
- name: SPEC2
 long_name: 2-digit specialisation
 role: outcome
 use: frozen GP (70-80) and OOH (81-84, 89) code lists; parsed labels
 ↪ mark 71/73/75 as
 "evening etc." GP services - kept daytime in the primary classifier
 ↪ by registered choice
 (see frozen_parameters.spec2_evening_gp_disclosure)
- name: YDLTID
 long_name: Time code for the benefit
 role: outcome
 use: robustness OOH classifier; missing/'not relevant' classified
 ↪ daytime (conservative
 against H3)
- name: AFRPER
 long_name: Settlement period
 role: check
 use: settlement-lag characterization
- name: BEHANDLINGSDATO
 long_name: The treatment date for the service
 role: check
 use: HONUGE settlement-lag calibration on 2021+ rows only (guard-band
 ↪ contingency); can
 calibrate but never replace HONUGE
- name: SIKGRUP
 long_name: Health insurance group
 role: check
 use: exclusion of modal group-2 insured (different billing regime)
- name: ALDERIMP
 long_name: Age at end of period, incl. imputed
 role: control
 use: age bands for exact strata (with explicit unknown stratum);
 ↪ eligibility 25-59
- name: KOENIMP
 long_name: Sex incl. imputed
 role: control
 use: sex stratum (with explicit unknown stratum); E3 heterogeneity
- name: YDERNR
 long_name: Provider number (ydernummer)
 role: key
 use: frozen fallback contact definition (distinct PNR x week x YDERNR
 ↪ x code-group cells)
- register: Health insurance statistics, legacy closed series (SYSI)
 register_da: Sygesikringsstatistik (SYSI)
 documentation: documented under TIMES sygesikringsstatistik and HKV
 ↪ sygesikring datasets
 (valid from 01-01-1990); series closed 2005
 years: 1990-2005; used ONLY in the exploratory SYSI-era replication E1
 ↪ (cohorts 1992-2003)
 variables:
- name: PNR
 long_name: De-identified person number
 role: key
 use: person spine
 verification: verified in the DST Forskningservice four-register
 ↪ inventory for the analysis
 server; absent from the public TIMES/HKV parses - disclosed
- name: HONUGE

```

```

 long_name: Fee week
 role: outcome
 use: weekly clock, legacy era
- name: SPECIALE
 long_name: 6-digit speciality
 role: outcome
 use: era GP/OOH classification (SPEC2 does not exist pre-2005);
 ↪ frozen era code list
- name: YDLANT
 long_name: Number of services within the specialty
 role: outcome
 use: 'E1 contact proxy: service-row counts weighted by YDLANT
 ↪ (KONTAKT does not exist pre-2005)'
- name: BRUHON
 long_name: Gross fee to doctor etc.
 role: outcome
 use: era fee outcome
- name: YDLTID
 long_name: Time code for the benefit
 role: outcome
 use: era time-of-day classification
- name: VAGTOMR
 long_name: Out-of-hours (laegevagt) service area
 role: outcome
 use: era OOH marker, 1997-2005 only
 verification: verified in the DST Forskningservice four-register
 ↪ inventory for the analysis
 server; absent from the public TIMES/HKV parses - disclosed
- name: AFRPER
 long_name: Settlement period
 role: check
 use: era settlement-lag characterization
- name: SIKGRUP
 long_name: Health insurance group
 role: check
 use: era exclusion of group-2 insured
- name: SIKREKOM
 long_name: Municipality of the insured person
 role: control
 use: era patient geography (E1 only; SSSY has no patient geography)
- name: SIKREAMT
 long_name: County of the insured person
 role: control
 use: era patient geography (E1 only)
 verification: verified in the DST Forskningservice four-register
 ↪ inventory for the analysis
 server; absent from the public TIMES/HKV parses - disclosed
variable_verification_note: 'Every variable above was checked against the
↪ parsed public documentation
(data/mapping/times_variables_bilingual.csv;
 ↪ data/mapping/hkv_variables_bilingual.csv).
Exceptions carried with an explicit per-variable verification note:
 ↪ PNR, IDAS KOM, SYSI
VAGTOMR, and SYSI SIKREAMT are verified in the DST Forskningservice
 ↪ four-register inventory
for the analysis server but are absent from the public TIMES/HKV
 ↪ parses; IDAN's legacy
workplace keys (ARBGNR/ARBSTK) are documented publicly on the workplace
 ↪ register and carried

```

```

on legacy IDAN vintages per the same inventory. Only these four
↳ registers exist on the analysis
server: there are no diagnosis, hospital, prescription, or
↳ population-spine registers in
scope, and outcomes are therefore utilization volume, fees, and timing
↳ composition - never
health states. One variable runs the other way and is disclosed as
↳ such: IDAN ARB_SEKTORKODE
- the private-sector eligibility filter - is publicly documented (TIMES
↳ ida-ansaettelser)
but absent from the recorded Forskningsservice inventory extract, its
↳ claimed 2008+ validity
resting on the public parse alone; its presence is validated at Stage 1
↳ before any classification
work, with a frozen ARB_HOVED_BRA_DB07-based fallback registered in
↳ analysis_plan.contingent_decisions.'
linkage:
 keys:
 - 'person spine: PNR (IDAN x SSSY x SYSI)'
 - 'workplace: LBNR with ARBNR/CVRNR (IDAS x IDAN, 2008+ vintage - the
 ↳ entire confirmatory
 tier); LBNR with ARBGNR/ARBSTK (pre-2008 vintage - E1 exploratory
 ↳ only); workplace identity
 chains via IDFREEM/IDTILB; never pooled across the 2007/08 IDA vintage
 ↳ break'
 expected_linkage_rate_floor: '>= 80% of cohort workers (exposed +
 ↳ control pool) with >=
 1 SSSY row of any kind in the 5-year window c-3..c+1 (outcome-neutral
 ↳ check; Danish annual
 GP contact prevalence benchmark ~85%); below floor halts at Stage 3'
 ...

Current authoritative contingencies

```yaml
- condition: ENDNEDL flag-year convention differs in-environment from the
  ↳ assumed 'flagged
  in the workplace's last active year'
  rule: Re-anchor mechanically so that c always denotes the last IDAS
  ↳ year with ANTNOV >=
  5 for the identity; the estimand is unchanged; the recalibration is
  ↳ reported verbatim
- condition: ARB_SEKTORKODE is absent from the analysis-server IDAN
  ↳ extract (publicly documented
  on TIMES ida-ansaettelser but unconfirmed in the recorded
  ↳ Forskningsservice inventory;
  presence and 2008+ coverage validated at Stage 1 before any
  ↳ classification work)
  rule: 'Frozen fallback: define the private-sector filter from
  ↳ ARB_HOVED_BRA_DB07 instead
  - a workplace counts as private if and only if its NACE DB07 section
  ↳ is not 0 (public
  administration and defence), P (education), or Q (human health and
  ↳ social work); the substitution
  and its effect on eligible-workplace and worker counts are reported
  ↳ verbatim; no other
  filter change is permitted'
- condition: Match rate < 95% of eligible exposed workers at target k = 3
  ↳ under full exact
  strata

```

rule: 'Apply the pre-specified coarsening ladder in order, stopping at
↳ the first rung reaching
95% of exposed matched with $k \geq 1$: (1) NACE section -> 10 aggregate
↳ industry groups;
(2) drop region stratum; (3) age bands 5-year -> 10-year. Never drop
↳ sex, cohort, size
class, or the baseline-billing calipers; realized k distribution
↳ reported and fed to the
Stage-2 blind re-simulation'

- condition: Median HONUGE settlement lag (BEHANDLINGSDATO calibration,
↳ 2021+ rows) > 4 weeks

rule: 'Guard band: exclude $Q1(c+1)$ from the θ_{2a}/θ_{2b} windows
↳ (means over $Q2(c+1) \dots Q4(c+1)$)
and report θ_{2a} both as registered and with a 4-week-shifted
↳ boundary; the registered
 θ_{2a} window remains primary'

- condition: KONTAKT does not behave as a row-level contact marker on
↳ in-environment validation

rule: 'Switch to the frozen fallback contact definition: count of
↳ distinct PNR x HONUGE-week
x YDERNR x GP-code-group cells with ≥ 1 positive-fee row; switch and
↳ level effect reported;
all estimands re-run under both definitions with the fallback as
↳ primary'

- condition: OLS and Poisson QML disagree in sign on any primary quantity

rule: Foreground the pre-committed OLS estimate, report both, flag the
↳ discordance in the
Summary paragraph; no post hoc model selection

- condition: 'c-2 placebo panel FAILURE, defined as: placebo contrast
↳ significant at α
= 0.05 AND $|\text{estimate}| > 1\%$ of control baseline mean (50% of the
↳ θ_{2a} floor effect) -
the single registered failure condition, stated identically in
↳ outcome_neutral_checks
and the H1 placebo kill'

rule: 'Execute the single pre-specified re-match adding year c-2
↳ baseline contacts to the
Mahalanobis set; if the placebo fails again under the same condition,
↳ the entire registered
event-study tier (confirmatory and exploratory event-study
↳ quantities) is killed and the
study reports descriptive trajectories only. The two
↳ non-complementary cells are pre-committed
passes with mandatory disclosure, never discretion: (a) significant
↳ at $\alpha = 0.05$ with
 $|\text{estimate}| \leq 1\%$ - no re-match, no kill; the Table T7 placebo row
↳ flags the significance
and states that a sub-floor-scale imbalance was detected; (b) not
↳ significant with $|\text{estimate}|$
> 1% - no re-match, no kill; a mandatory caution paragraph reports
↳ the placebo magnitude
against the θ_{2a} floor and the realized placebo MDE, and the H1/H2
↳ Summary paragraphs
must cite it'

- condition: A matched control's own workplace acquires a terminal event
↳ within c-2..c+1 after
matching (data revision)

rule: Drop that control from the set; drop the whole set only if no
↳ control remains; dropped
counts reported; no re-matching

- condition: YDLTID missing or code 0 ('not relevant') on a GP row
 - rule: Classify as daytime (conservative against the theta3 prediction);
 - ↪ missingness share
 - reported; if > 20% of peri-event GP rows, H3's kill criterion triggers
- condition: Exact tie in Mahalanobis distance for a nearest-neighbor
 - ↪ match
 - rule: Deterministic tie-break by smallest absolute baseline-contact
 - ↪ difference, then seeded
 - RNG (seed 20260831)
- condition: IDAS-flag vs worker-flow classification agreement < 90% but
 - ↪ >= 75%
 - rule: Proceed with the worker-flow classification as registered (the
 - ↪ canon, S7); publish
 - the full disagreement cross-tabulation; below 75%, H2's kill
 - ↪ criterion governs
- condition: 'GRAFT 3: sign of theta2b or theta2_diff not stable across
 - ↪ the three successor-share
 - threshold pairs'
 - rule: Flag classification fragility in the pre-committed H2 Summary
 - ↪ paragraph; the 30/70
 - estimate remains primary; no threshold shopping
- condition: 'GRAFT 6: exposed-vs-control difference in later-presence
 - ↪ rates (any IDAN record
 - in year c+1 or any SSSY row in c+1 or later) exceeds 2 percentage
 - ↪ points'
 - rule: A mandatory bounding paragraph accompanies all post-boundary
 - ↪ (theta2a/theta2b) conclusions,
 - converting the differential into best/worst-case bounds; trigger and
 - ↪ threshold registered,
 - not discretionary
- condition: 'GRAFT 9: the control forward-window sensitivity yields
 - ↪ SMALLER magnitudes than
 - the primary estimates'
 - rule: Flag as diagnostic of something other than contamination
 - ↪ (registered direction note);
 - reported verbatim, no re-specification
- condition: 'STAGE-2 REALIZED-N GATE: realized matched sets (or spurious
 - ↪ sets for F2) fall
 - below a member's binding N* from the Stage-0 re-run, or the Stage-2
 - ↪ blind re-simulation
 - (realized Ns, realized k distribution, baseline-year moments only)
 - ↪ shows power < 0.95
 - for a member at its floor'
 - rule: That member is demoted to exploratory BEFORE any post-baseline
 - ↪ outcome is accessed;
 - the Holm family shrinks and alphas recompute by the registered
 - ↪ fixed-point iteration among
 - survivors; demotion-only - no member, threshold, or floor may move in
 - ↪ the other direction;
 - the realized-power/MDE table is published either way
- condition: Realized control-group OOH share of GP contacts < 5% (below
 - ↪ the published 8-12%
 - range)
 - rule: H3's power basis is re-verified in the Stage-2 blind
 - ↪ re-simulation at the realized
 - share; if < 0.95 at the 5-point floor, H3 demotes to exploratory per
 - ↪ the realized-N gate;
 - if realized share < 2%, H3's kill criterion governs instead
- condition: Spurious arm killed at Stage 1/2 (events < 150, sets <
 - ↪ 3,000, classification

```

    agreement < 75%, or bimodality unresolved)
rule: Gate F2 and theta2_diff/theta2b are reported unavailable (drop
↳ disclosed before outcome
    access); theta2a remains in the Holm family if its own N* is met, but
↳ the H2 verdict is
    capped at 'partially consistent - comparative quantifier untestable'
↳ and the Discussion
    states that the excess cannot then be distinguished from selection
↳ into disappearing workplace
    identities
...

## Current shared design

```yaml
type: stacked matched difference-in-differences event study (matched
↳ event-study contrasts
 at quarterly confirmatory resolution with weekly descriptive
↳ trajectories; associational
 claim class)
design_name: 'Bracing for the Blow - remediated confirmatory design:
↳ stacked 1:3 riskset-matched
 difference-in-differences of quarterly GP billing around true vs
↳ spurious Danish workplace
 closures (SSSY era, closure cohorts 2009-2023), with the true-closure
↳ post-year excess as
 H2's confirmatory core, a spurious-arm falsification gate outside the
↳ Holm family, published-basis
 out-of-hours power calibration, and a pre-outcome realized-N demotion
↳ gate'
identifying_variation: 'Establishment terminal events in IDAS - workplace
↳ identities present
 in year c (ANTNOV >= 5) and absent in year c+1 - across 15 closure
↳ cohorts 2009-2023, classified
 true vs spurious by IDAN worker-flow successor shares per the Benedetto
↳ et al. (2007) /
 Hethey-Maier & Schmieder (2013) canon, contrasted against
↳ riskset-matched workers at stable
 workplaces within exact strata (cohort x sex x age band x NACE section
↳ x size class x region)
 with 52 weeks of pre-baseline billing-history matching. Closure dating
↳ is annual November-status,
 so no sharp within-scope discontinuity exists: the claim ceiling is
↳ associational, all pre-event
 quantities are worded as pre-boundary excess (never announcement-dated
↳ anticipation), and
 the spurious-closure placebo plus the c-2 placebo panel carry the
↳ severity burden.'
alternatives_considered:
- design: 'Round-1 conservative design (VOIDED): single-era stacked
↳ matched-PAIR (1:1) DiD,
 closure cohorts 2009-2022, sampling self-handicapped to half the
↳ lower bound of its own
 derived expectation (75k pairs), H2 as the true-minus-spurious
↳ differential, H3 simulated
 at a 3% OOH base rate'
discrimination: identical contrast logic; differed only in matching
↳ ratio, sampling plan,
 H2's confirmatory construct, and H3's power basis

```

why\_not\_selected: Its deposited Stage-0 power simulation emptied the  
↳ confirmatory family  
via the registered fixed-point rule (H2 0.358, H1\_trend 0.380, H3  
↳ 0.614, H1 0.840) and  
the driver audit voided the resulting all-exploratory registration on  
↳ 2026-08-31. The  
remediation (authorized honest levers only; no floor raised) produced  
↳ the registered design;  
the full record is carried in remediation\_record and the records of  
↳ both power runs are  
published with the protocol.

- design: 'Design B, ambitious three-arm weekly event study 1991-2023:  
↳ weekly-resolution confirmatory  
estimands, both eras (SYSI + SSSY) confirmatory, a mass-separation  
↳ arm at surviving workplaces,  
ANTNOV cap 1000'  
discrimination: weekly confirmatory estimands would sharpen  
↳ timing-shape claims; a second  
era and third arm would extend coverage and external validity

why\_not\_selected: Weekly confirmatory estimands are hostage to the  
↳ HONUUGE settlement lag;  
the SYSI era cannot match on age or sex and sits across the 2007/08  
↳ IDA vintage break;  
the mass-separation arm breaks one-study scope discipline. Its honest  
↳ elements - the ANTNOV  
1000 cap, pre-committed reversed-sign interpretations, and the  
↳ absorption placebo - were  
absorbed into the registered design; the SYSI era survives as  
↳ exploratory E1 and weekly  
resolution as descriptive trajectories under the GRAFT-2 step  
↳ diagnostic.

threats:  
- threat: Anticipatory attrition - healthier/more mobile workers leave  
↳ sinking workplaces  
before closure (Schwerdt 2011), biasing late-window composition  
addressed\_by: Cohort frozen at November-status attachment in year c-1,  
↳ before the terminal  
year (S9); intention-to-follow regardless of subsequent separation;  
↳ E4 arm-comparative  
attrition accounting panel quantifies the channel, labeled  
↳ post-treatment-conditioned  
descriptives

- threat: 'Billing responses to separations already realized within the  
↳ terminal year - under  
intention-to-follow, workers who separate from the closing workplace  
↳ early in year c (the  
wave E4 describes) contribute post-job-loss utilization to the  
↳ pre-boundary Q3(c)/Q4(c)  
window, so a Holm-significant  $\theta_1$  could be generated wholly by  
↳ realized early job loss  
with zero anticipation. Direction: inflates  $\theta_1$  relative to a  
↳ pure-anticipation reading.  
The frozen clause ''before any separation is visible in annual IDAN  
↳ records'' holds at  
the cohort-anchoring level (attachment frozen at November c-1) but  
↳ not for individual  
early leavers, whose separation is visible in the year-c IDAN record.'  
addressed\_by: The pre-committed H1 interpretation names both channels  
↳ (anticipatory care-seeking

- and/or billing responses to separations already realized early within
    - ↪ the terminal year)
  - with the E4 year-c separation share bounding the latter; the E4
    - ↪ arm-comparative separation
  - shares themselves; and the registered post-treatment-conditioned
    - ↪ exploratory split of
  - theta1 by November-year-c attachment (E4), which localizes the
    - ↪ channel without touching
    - the confirmatory estimand
- threat: Selection of sicker workers into dying workplaces
  - ↪ (distressed-firm workforce composition)
  - masquerading as anticipation or as post-year excess
  - addressed\_by: Exact + Mahalanobis matching on 52 weeks of pre-baseline
    - ↪ billing (contacts,
    - fees, OOH, contact-weeks) in year c-1; c-2 placebo panel with numeric
      - ↪ failure threshold;
    - falsification gate F2 - the spurious-closure arm separates
      - ↪ identity-death from economic
    - job loss with a pre-committed disconfirmation trigger that can only
      - ↪ take support away
    - from H2
- threat: HONUGE is the fee/settlement week, not the service week -
  - ↪ billing lag blurs the
  - boundary and could relabel pre-boundary care as post-boundary
  - addressed\_by: Quarter-level primary estimands (robust to  $\leq 4$ -week
    - ↪ lags); BEHANDLINGSDATO
  - calibration (2021+) with the guard-band contingency; AFRPER used
    - ↪ descriptively; GRAFT-2
  - week-0 step diagnostic bounds any timing-shape claim
- threat: Annual closure dating - the design cannot see announcement
  - ↪ dates, so 'anticipation'
  - could be mislabeled
  - addressed\_by: Estimand defined and worded exclusively as pre-boundary
    - ↪ excess relative to
    - the closure-year boundary (S8); claim ceiling forbids
      - ↪ announcement-locked language; gated
    - H1\_trend partial pattern pre-commits the flat-elevation reading when
      - ↪ no within-year growth
    - is seen
- threat: Local-economy co-shocks hitting exposed and control workers'
  - ↪ regions simultaneously
  - addressed\_by: Exact matching on region and NACE section plus cohort x
    - ↪ calendar-quarter fixed
  - effects; residual local shocks affect controls too and attenuate the
    - ↪ contrast (bias toward
    - null, stated)
- threat: Control contamination by quietly dying or shrinking workplaces
  - addressed\_by: Stable-control definition (unchanged identity c-2..c+1,
    - ↪ no ENDNEDL, AFGRATE
  - < 30%, |ANTALXFR|  $\leq$  30%); control-drop rule on data revision;
    - ↪ GRAFT-9 forward-window
  - (c+2/c+3) sensitivity with registered direction note (truncated for
    - ↪ the last two cohorts,
    - disclosed)
- threat: Reverse causality at small firms - workforce or owner illness
  - ↪ causes the closure
  - addressed\_by: Size floor ANTNOV  $\geq 5$ ; associational claim ceiling with
    - ↪ the reverse channel
  - named in the limitations; exploratory size heterogeneity (size
    - ↪ classes now extend to 1000)

- threat: 2007/08 IDA vintage break contaminating any pooled estimate  
addressed\_by: Confirmatory tier confined entirely to post-break cohorts  
↳ (c >= 2009); SYSI-era  
analysis is exploratory replication (E1) and never pooled
- threat: Unobservable death/emigration (no population spine in scope)  
↳ deflating post-event  
billing  
addressed\_by: Named scope limitation; direction discussed (deflates the  
↳ post window - conservative  
for theta1, attenuating for theta2a); pre-specified sensitivity  
↳ excluding possible-death/emigration  
sets; GRAFT-6 2pp later-presence trigger forces best/worst-case  
↳ bounds on the post-boundary  
quantities
- threat: 'Spurious-arm thinness - the structural cause of the first  
round's collapse: any  
↳ confirmatory quantity whose variance is dominated by the spurious arm  
↳ is unpowerable at  
honest N'  
addressed\_by: 'No confirmatory quantity now touches the spurious arm:  
↳ theta2a is a true-arm-only  
contrast; the spurious arm serves as falsification gate F2 (whose job  
↳ needs bounding,  
not detection, and whose operating characteristics are verified at  
↳ Stage 0) and as the  
exploratory H2\_diff; spurious-arm death caps the H2 verdict rather  
↳ than silently weakening  
a test'
- threat: Residual underpowering if realized N, realized k, or the  
↳ realized OOH share falls  
short of the sampling plan  
addressed\_by: Pre-outcome Stage-2 realized-N gate with binding N\*  
↳ thresholds from the re-run  
deposited simulation; blind re-simulation on realized  
↳ Ns/k-distribution/baseline moments  
and realized OOH share; demotion-only resolution before any outcome  
↳ access; realized-power/MDE  
table published regardless
- falsification\_tests:
- 'Falsification gate F2 (spurious-closure placebo arm): theta2b with the  
↳ pre-committed disconfirmation  
trigger; outside the Holm family; verified operating characteristics  
↳ registered (pass 0.886  
hypothesis-true / 0.978 inert; trigger 0.79 at floor-sized non-inert  
↳ placebo, at 45,000  
spurious sets - the pre-matching worker count at the 30% band lower  
↳ bound; the registered  
90% match-rate floor implies ~40,500 sets, so these characteristics are  
↳ computed at a ~10%  
flattered lower-bound N); Stage-2 blind re-verification at realized N'
- 'c-2 pre-baseline placebo panel: exposed-minus-control contact  
↳ difference across the four  
quarters of year c-2 with a numeric failure threshold - fails if and  
↳ only if significant  
at alpha = 0.05 AND |estimate| > 1% of control baseline mean, the  
↳ single registered condition  
identical wherever this gate appears; one pre-specified re-match, then  
↳ kill of the registered  
event-study tier'

- GRAFT 2 - week-0 boundary-step diagnostic on the weekly descriptive
  - ↳ series (settlement-lag contamination bound on any timing-shape claim)
- GRAFT 3 - successor-share threshold sensitivities (20/80, 40/60)
  - ↳ alongside primary 30/70) for  $\theta_{2b}/\theta_{2\_diff}$  sign stability
- GRAFT 4 - s(w) bimodality check at Stage 1 (ambiguous mid-zone < 35% of terminal events)
  - ↳
- GRAFT 6 - later-presence differential trigger (2pp) forcing
  - ↳ best/worst-case bounds on all post-boundary conclusions
- GRAFT 9 - control forward-window (c+2/c+3) contamination sensitivity
  - ↳ with registered direction note (exclusion should weakly increase magnitudes; a decrease is flagged as diagnostic)
- 'Classification cross-validation: worker-flow classes vs IDAS-native flags (ENDNEDL/IDFREM/REGELFR/FRELFREM/ANDELUDS),  $\geq 75\%$  agreement kill floor, full cross-tabulation released'

narrative: 'This study asks whether Danish workers'' primary-care billing

- ↳ already rises while their workplace is dying - before any separation is visible in annual employment records
- and whether the post-closure excess attaches to genuinely destroyed jobs rather than to workplaces that die only on paper. The instrument is the weekly clock of Danish health-insurance billing (SSSY, HONUUGE fee-week) held against annually dated establishment terminal events from the IDA workplace registers. For closure cohorts 2009-2023 (entirely inside the post-2008 IDA key era), every worker with primary November-status employment in year c-1 at a workplace that disappears between years c and c+1 is classified by the worker-flow canon of Benedetto et al. (2007) and Hethey-Maier & Schmieder (2013): true closure when no successor absorbs the workforce (max successor share below 0.30), spurious closure when an en-bloc successor exists (share at or above 0.70), the ambiguous mid-zone excluded and its mass monitored by a pre-registered bimodality check. Cohorts are frozen at November of c-1, so the Schwerdt sinking-ship channel cannot condition the sample. Each exposed worker is riskset-matched to up to three stable-workplace controls within exact strata (cohort, sex, age band, NACE section, size class, region) on Mahalanobis distance over a full year of pre-baseline billing history plus tenure and wage decile. The estimator is a stacked matched-set difference-in-differences event study with set and cohort-by-calendar-quarter fixed effects, event-workplace clustering, a c-2 placebo panel with a numeric failure threshold, and a settlement-lag calibration of the fee-week clock.
- \n\nThree confirmatory quantities form the Holm family, each an honest operationalization of a frozen bank hypothesis.  $\theta_{11}$  (H1) is the pre-boundary excess in the last two quarters of the terminal year - worded strictly as pre-boundary excess relative

to an annually dated boundary, never announcement-dated anticipation;  
 → the frozen statement's  
 'growing as the boundary approaches' rider is a gated exploratory  
 → contrast, because its  
 ramp-implied floor is verified undetectable at any honest N. theta2a  
 → (H2) is the true-closure  
 arm's post-year excess over its own matched controls - the excess the  
 → frozen statement  
 asserts exists 'through the first post-closure year' - while the  
 → statement's comparative  
 quantifier is enforced by a registered falsification gate: the  
 → spurious-closure arm's excess  
 is estimated, bounded, and equipped with a pre-committed  
 → disconfirmation trigger (spurious  
 excess flagged positive and at least half the true-arm excess  
 → disconfirms H2's comparative  
 reading regardless of theta2a), with the strict differential reported  
 → as an exploratory  
 estimate. This restructuring follows the verified Stage-0 diagnosis: a  
 → differential whose  
 variance is dominated by a spurious arm one-third the size of the true  
 → arm needs roughly  
 5.5 times any honest variance budget - it is a falsification device,  
 → not an estimation target.  
 theta3 (H3) is the out-of-hours versus daytime relative-excess gap,  
 → unchanged in construct;  
 what changed is only its power basis, recalibrated from a self-declared  
 → over-conservative  
 3% out-of-hours base rate to the conservative end of the published  
 → Danish range (simulated  
 at 8.0% of daytime GP volume = 7.4% of all GP contacts, against a  
 → published 8-12% contact  
 share; Christensen & Olesen 1998; the LV-KOS 2011 studies), with an  
 → outcome-neutral check  
 verifying the realized share sits in the published range before the  
 → confirmatory run.  
 Power  
 is secured by honest enlargement, not by floor inflation: every floor  
 → effect is unchanged  
 (2% relative excess for theta1 and theta2a, below the  
 → Hamad-Modrek-Cullen and Kuhn-Lalive-Zweimüller  
 magnitudes; a 5-point composition gap for theta3), and the sampling  
 → plan moves from the  
 first registration's self-handicapped half-of-lower-bound to the full  
 → derived expectation's  
 lower bound - 15 cohorts, size cap 1000, roughly 135,000 matched sets -  
 → on a published closure-rate  
 basis, with 1:3 matching cutting control-side variance by a third. The  
 → license for that  
 sizing is a demotion-only realized-N gate: after matching and before  
 → any post-baseline outcome  
 is touched, the deposited simulation is re-run blind on realized set  
 → counts, realized match  
 ratios, baseline moments, and the realized out-of-hours share; any  
 → member short of 0.95  
 power at its floor is demoted to exploratory on the spot, the Holm  
 → family shrinks by the  
 registered fixed-point rule, and nothing can ever move the other way.  
 → The claim ceiling  
 is associational and binding - associations and consistency language  
 → only, no effect/impact/cause

```

vocabulary, no phenomenon-discovery or Danish-first framing, outcomes
→ always utilization
volume, fees, and timing composition, never health states - and the
→ protocol publishes the
records of both power runs, the demotion trail, and the remediation
→ note itself, so the
registration's history is part of its evidence.'
...

```

## ## 2. ENVIRONMENT

Software stack (record exact versions in the run manifest at Stage 0 and freeze them):

- **\*\*python3 (>= 3.11) with polars\*\*** for ALL large-table preprocessing. Lazy scans only (``scan_parquet`/`scan_csv``) with projection and predicate pushdown; semi-joins against the cohort PNR roster before any collect; streaming collects for aggregations. Never materialize a person x week or person x month panel beyond need: the only dense person-time table permitted is matched members x 16 event-quarters (~540k true members + ~180k spurious members, x 16 = under 12M rows). Zero-filling happens at the quarter level for panel members only, never by cross-joining the population against a calendar.
- **\*\*R (>= 4.3) with fixest, invoked via Rscript\*\***, for every fixed-effects estimation: ``feols`` (OLS, the pre-committed primary) and ``fepois`` (Poisson QML robustness). CR1 cluster-robust SEs on event workplace. No estimation in python; no data manipulation in R beyond reading the analysis parquet and writing tidy coefficient tables.
- **\*\*matplotlib\*\*** for all figures; **\*\*booktabs LaTeX\*\*** for all tables (no vertical rules). Every number in every exhibit is written by script from a parquet/CSV output - never hand-edited.
- **\*\*Memory discipline\*\*** process cohort-by-cohort (c = 2009..2023) where possible; checkpoint every numbered pipeline step as parquet under ``interim/`` with a fixed name (given per step below); SSSY scans filter on (PNR in roster) AND (SPEC2 in frozen lists) at scan level; the full population is touched only by the aggregate coverage benchmark (step 16), which is computed as a streaming groupby per year, never collected at row level.
- **\*\*Seeded randomness\*\*** a single root seed 20260831 (per the frozen registration) initializes every RNG: matching processing order, Mahalanobis tie-breaks, the 999-draw cluster bootstrap, and the Stage-2 blind re-simulation. No unseeded randomness anywhere.
- **\*\*Run manifest\*\*** an append-only JSONL (``manifest/run_manifest.jsonl``), one record per pipeline step per run: step id, UTC timestamp, code version hash, input file(s) + row counts, output file + row count, parameters used, validation check results, PASS/FAIL. Every step below MUST log its row counts; a step whose check fails writes FAIL and halts downstream steps (Section 6.3). The manifest (aggregate counts only, event counts rounded per the disclosure rules) is part of the released audit trail.
- **\*\*Directory convention\*\*** ``raw/`` (DST-delivered), ``interim/`` (parquet checkpoints), ``analysis/`` (panels), ``out/tables/``, ``out/figures/``, ``manifest/``. All paths in code are relative to the project root inside the research environment; no absolute paths in any released script.
- **\*\*Frozen-code discipline\*\*** the Stage-4 estimation code is frozen at the end of Stage 0 (hash recorded in the manifest); Stage 4 is a single run of that code. The sensitivity suite runs once and is reported together. Any code change after Stage 0 is a deviation and must be disclosed in running prose with reason and approver.

---

```

Source disclosure rules

```yaml
environment: DST Forskningservice (no prior data access; gatekeeper
  ↳ certification per RR
    secondary-data rules). All analysis occurs inside the environment; only
    ↳ aggregate outputs
  leave. No microdata, workplace identifiers (ARBNR/CVRNR/LBNR), or
    ↳ person-level values are
  released.
minimum_cell_size: 5
minimum_cell_rules: no released statistic based on fewer than 5 persons
  ↳ or fewer than 3 workplaces;
    event-count tables rounded to the nearest 10; weekly trajectory figures
    ↳ pool >= 100 events
    and >= 5,000 person-weeks per plotted point; fee (BRUHON) aggregates
    ↳ subject to a 50% single-workplace
    dominance check; min-cell suppression applied BEFORE Holm reporting so
    ↳ suppression can never
    be outcome-selective
outputs_leaving_environment:
  - model coefficients with SEs/CIs
  - matching balance tables and realized-k distribution
  - calibration and benchmark tables (coverage, linkage, HONUGE lag, OOH
    ↳ share)
  - figure aggregates (weekly trajectory series under the pooling rules)
  - event counts rounded to the nearest 10
  - Stage-0 power/MDE tables - BOTH rounds (failed and remediated) - and
    ↳ the Stage-2 blind re-simulation
    output with the demotion trail
  - the c-2 placebo, gate F2 outputs, GRAFT-2/4 diagnostics, GRAFT-6
    ↳ later-presence rates, E4
    attrition panel, and all outcome-neutral check outputs (the full audit
    ↳ trail exits under
    the same rules)
  - the analysis and power scripts
statement: 'All analysis occurs inside DST Forskningservice; only
  ↳ aggregate outputs leave.
    Rules applied to every released object: (1) no statistic based on fewer
    ↳ than 5 persons or
    fewer than 3 workplaces; (2) event-count tables rounded to the nearest
    ↳ 10; (3) weekly trajectory
    figures pool >= 100 events and >= 5,000 person-weeks per plotted point;
    ↳ (4) fee (BRUHON)
    statistics released only as cell means/coefficients over min-N cells,
    ↳ with a dominance check
    that no single workplace contributes > 50% of any released fee
    ↳ aggregate; (5) released objects
    are limited to: model coefficients with SEs/CIs, balance tables,
    ↳ calibration/benchmark tables,
    figure aggregates, event counts (rounded), the Stage-0 (both rounds)
    ↳ and Stage-2 power/MDE
    tables with the demotion trail, and the analysis/power scripts - never
    ↳ microdata, workplace
    identifiers (ARBNR/CVRNR/LBNR), or person-level values; (6) min-cell
    ↳ suppression is applied
    before Holm reporting so that suppression can never be
    ↳ outcome-selective; (7) the c-2 placebo,

```

```

gate F2 outputs, the GRAFT-2/4 diagnostics, the GRAFT-6 later-presence
→ rates, the E4 attrition
panel, and all outcome-neutral check outputs are released under the
→ same rules, ensuring
the full audit trail exits the environment.'
...

```

Pending registered experiment

No Danish study outcome has been observed for this bundle. See the complete registration above.

atom-000002: Matched cohort validity

Can the registered matched cohorts and pre-outcome inferential eligibility be established?

Source fidelity and implementation order

This is a format conversion of DS1.1, not a new study or independent acceptance. Source audit: minor_revisions. No observed Danish outcomes are supplied. The source execution plan explicitly states that final prereg.yaml governs disagreements. Exact current YAML selections below therefore govern older execution-plan excerpts; these excerpts are retained as implementation detail, not a competing registration. See docs/provenance/source-discrepancies.md. The complete shared data/validity contracts are embedded in Atoms 1–3; read those native protocols when executing a downstream Atom, even without docs/. Do not substitute a scientific default for an unresolved source contradiction.

Preserve source stages 0 -> 1 -> 2 -> 3 -> 4 -> 5 inside the implementation. Source Stage-0 simulations are historical design evidence, not study outcomes; do not silently redo or improve them. Gate 14 must be evaluated from baseline data before post-baseline outcomes. All surviving Stage-4 quantities form ONE Holm family. The three question-level Atoms may share the same frozen family fit and tidy results; do not apply Holm separately per Atom or adjudicate a member before the complete surviving family's p-values exist. Cache/reuse identical governed intermediate results with provenance when available; a direct-dispatch call must obtain or compute its exact source-defined inputs, not pretend its predecessors succeeded. Stage-5 internal riders remain exploratory, after the Stage-4 family is frozen, even when housed in H1/H2.

Reference files under docs/execution/ are the ORIGINAL SKELETONS, not tested turnkey implementations. Adapt paths and complete engineering in a governed workspace; preserve source scientific rules and disclose unresolved choices. Write a result record with component estimates, uncertainty, family eligibility, source interpretation row, diagnostics and boundary/skip dispositions. Released outputs follow source disclosure rules; confidential rows stay governed. Local native states summarize the source judgment; neither a state name nor source power forecasts prove equivalence or causal identification. A bounded claim requires its actual interval

and registered rule. Retain unresolved source wording conflicts; do not convert mere nonsignificance into scientific null.

Frozen scope and vocabulary

```
claim_class_ceiling: associational
statement: 'Associational - matched event-study contrasts, a contrastive
↳ falsification gate,
  and a composition contrast; descriptive for trajectory panels and the
  ↳ bounded priority claim.
Binding vocabulary rules: 'is associated with' / 'consistent with' /
↳ 'suggests' only;
'effect', 'impact', 'causes' prohibited throughout; anticipation
↳ always worded as
'pre-boundary excess' relative to the annually dated closure-year
↳ boundary, never as announcement-dated
anticipation (S8); no phenomenon-discovery or Danish-first framing (S1-S3);
↳ H2 novelty confined
to the placebo's application to health billing (S7). H2 verdict semantics
↳ are fixed in
advance: 'consistent with H2' requires BOTH the confirmatory core
↳ (theta2a Holm-significant,
predicted direction) AND a passing falsification gate F2; the strict
↳ comparative magnitude
is only ever reported as an exploratory estimate with its CI. Exploratory
↳ results (H1_trend,
H2_diff, E1, E3, E4, and any gate-demoted member) carry the word
↳ 'exploratory' in the
heading and in every sentence stating them. The prose label never exceeds
↳ the evidence label.'
```

```
frozen_parameters:
seed: 20260831
gp_code_list: daytime GP = SPEC2 in {70..80}
ooh_code_list: out-of-hours GP = SPEC2 in {81, 82, 83, 84, 89}
ooh_robustness_classifier: adds YDLTID in {2, 3, 4, 5, 7, 8, 9}
spec2_evening_gp_disclosure: 'Registered choice, disclosed: the parsed
↳ value labels mark
  SPEC2 71 ("General practitioner, consultation, evening etc."), 73
  ↳ ("General practitioner,
  telephone consultation, evening etc."), and 75 ("General practitioner,
  ↳ home visit, evening
  etc.") as explicitly off-schedule GP services; they nonetheless remain on
  ↳ the daytime
  side of the PRIMARY classifier because H3's out-of-hours construct and
  ↳ its published
  power basis are keyed to the laegevagt out-of-hours service (SPEC2 81-84,
  ↳ 89), the quantity
  the published Danish 8-12% share describes. Attenuation direction, both
  ↳ readings: counting
  evening GP as daytime attenuates theta3 toward zero - conservative
  ↳ against the positive
  H3 conclusion, anti-conservative for the null-bounding "compositionally
  ↳ unremarkable"
  reading, whose Stage-2 statement therefore carries a mandatory caveat.
  ↳ In-environment
  validation must verify that the YDLTID-augmented robustness classifier
  ↳ captures SPEC2
  71/73/75 rows (their evening time codes 2-5); if YDLTID proves
  ↳ unpopulated on those rows,
```

```

the robustness classifier - and only the robustness classifier - is
  ↳ amended mechanically
and with disclosure to include SPEC2 71/73/75 directly. The SPEC2
  ↳ 71/73/75 volume share
is reported, and both classifiers appear side by side in Table T7.'
successor_share_thresholds: 'true closure: ENDNEDL final closure AND
  ↳ worker-flow max-successor
  share s(w) < 0.30; spurious closure: forward identity via
  ↳ IDFREEM/REGELFR/FRELFREM AND
  s(w) >= 0.70; ambiguous mid-zone excluded; GRAFT-3 sensitivities
  ↳ 0.20/0.80 and 0.40/0.60'
quarter_mapping: 'ISO week of HONUGE: 1-13 / 14-26 / 27-39 / 40-52; ISO
  ↳ week 53 to Q4'
winsorization: person-quarter outcomes at the 99.9th percentile within era
  ↳ x calendar quarter
matching: 1:3 nearest-neighbor Mahalanobis within exact strata, without
  ↳ replacement across
  sets, target k = 3, minimum k = 1; calipers |dif baseline contacts| <= 3
  ↳ and |dif log(1+BRUHON)|
  <= 0.5; exposed processed in seeded random order (seed 20260831)
software: R >= 4.3, fixest::feols (OLS) and fixest::fepois (Poisson QML
  ↳ robustness); versions
  frozen in the registered script
note: Code lists and thresholds frozen before any data access; code labels
  ↳ quoted verbatim
  from parsed DST value labels and re-validated in-environment before
  ↳ estimation (frozen
  fallback rules in contingent_decisions).

```

9. **Exposed cohort extraction at November c-1.** Input: steps 3–8. Workers with primary November-status employment in year c-1 at an event workplace; apply inclusion rules (tenure ANSAAR <= c-2; private sector; ANTNOV 5–1000 in c-1) and the IDAN-side exclusions (missing ANSAAR or missing/zero annual pay, counts reported, expected < 2%; multi-cohort workers assigned to earliest event only). Age/sex and SIKGRUP screens complete in step 11. Output: `interim/exposed_cohort_raw.parquet`. Check: eligible exposed per cohort against the 10k–20k basis; exclusion accounting logged line-by-line (feeds exhibit T3).
10. **Stable-control pool.** Input: steps 1–6. Control workplaces: identity present and unchanged c-2..c+1, no ENDNEDL, AFGRATE < 30% and |ANTALXFR| <= 30% in years c-1 and c; exclude workplaces with any terminal event or mass separation (AFGRATE >= 30%) in c-2..c+1. Control workers: primary November c-1 employment at such workplaces, same inclusion rules as step 9. Output: `interim/control_pool.parquet`. Check: pool size per stratum orders of magnitude above exposed (log the ratio); zero overlap with any event workplace.
11. **SSSY pre-baseline extraction, years c-3..c-1 (Stage 2 begins; pre-baseline rows only).** Input: raw SSSY, roster = steps 9–10. Lazy scan filtered to roster PNRs; extract rows for the frozen GP/OOH code lists plus SIKGRUP/ALDERIMP/KOENIMP. Derive: age 25–59 at baseline from ALDERIMP on the person's most recent SSSY row in c-3..c-1 (no row -> explicit "unknown age/sex" stratum, not dropped); modal SIKGRUP in c-1 (modal group 2 -> excluded, share reported, expected ~1–2%); year c-1 matching covariates: total

GP contacts, OOH contacts, $\log(1+\text{BRUHON})$, weeks with ≥ 1 contact. Output: `interim/sssy_baseline.parquet`. Check: share of roster with ≥ 1 SSSY row in c-3..c-1 logged (precursor to the 80% linkage floor); unknown-stratum share reported; SIKGRUP exclusion share reported.

12. **Matching frame.** Input: steps 9–11 + IDAN covariates. Exact strata: cohort c x sex (KOENIMP, incl. unknown) x 5-year age band (ALDERIMP, incl. unknown) x NACE section (ARB_HOVED_BRA_DB07) x workplace size class (ANTNOV: 5–9/10–19/20–49/50–99/100–499/ 500–1000) x region (workplace KOM mapped to the 5 post-2007 regions). Mahalanobis components (all year c-1): total GP contacts, $\log(1+\text{BRUHON})$, OOH contacts, weeks with ≥ 1 contact, tenure (c-1 minus ANSAAR), wage decile (BREDT_LOEN_BELOEB, computed within cohort). Output: `interim/matching_frame.parquet`. Check: covariate completeness; stratum cardinalities and thin-stratum share logged.
13. **1:3 riskset matching.** Input: step 12. Nearest-neighbor Mahalanobis within exact strata, without replacement of controls across sets, target $k = 3$, minimum $k = 1$; calipers $|\text{dif baseline contacts}| \leq 3$ and $|\text{dif } \log(1+\text{BRUHON})| \leq 0.5$; exposed processed in seeded random order (seed 20260831); exact ties broken by smallest absolute baseline-contact difference, then seeded RNG. If match rate $< 95\%$ of eligible exposed at target $k = 3$ under full strata, apply the coarsening ladder IN ORDER, stopping at the first rung reaching 95% matched with $k \geq 1$: (1) NACE section $\rightarrow$ 10 aggregate industry groups; (2) drop region stratum; (3) age bands 5-year $\rightarrow$ 10-year. Never drop sex, cohort, size class, or the billing calipers. Run for the true arm and separately for the spurious arm (own controls). Outputs: `interim/matched_sets.parquet`, `interim/matched_sets_spurious.parquet`. Checks (Stage-2 stop rules): match rate $\geq 90\%$ after the full ladder; standardized mean difference ≤ 0.10 on every matching covariate and on unmatched auxiliaries (control aggregates weighted $1/k$ within set); realized k distribution and ladder rung logged; true sets $\geq 10,000$ (project kill below); spurious sets $\geq 3,000$ (spurious-arm kill below).
14. **STAGE-2 PRE-OUTCOME REALIZED-N GATE (blind; binding).** Input: realized set counts, realized k distribution, baseline-year (c-1) moments, realized control OOH share of baseline GP contacts — no post-baseline outcome touched. Re-run the deposited simulation (`power_simulation.py` machinery, seed 20260831) at these realized quantities; recompute F2 operating characteristics at realized spurious N. Decisions (demotion-only): any confirmatory member with realized sets $< N^*$ (H1 118,000; H2a 95,000 true + 15,000 spurious for F2; H3 90,000) or blind re-simulated power < 0.95 at its floor is demoted to exploratory NOW; the Holm family shrinks by the registered fixed-point iteration among survivors; nothing is ever promoted. If realized control OOH share $< 5\%$, H3's re-verification governs (demote if < 0.95 at the 5-point floor); $< 2\%$ invokes H3's kill criterion at Stage 3. Output: `interim/realized_power_mde.parquet` + the demotion trail (released regardless of outcome; feeds exhibit T5).

Population

definition: 'Registered cohorts (SSSY era): persons with primary
→ November-status employment
in year c-1 (IDAN; primary job = largest BREDT_LOEN_BELOEB in c-1, ties
→ broken by larger
LOENTIMER then lower ARBNR - deterministic) at (i) an event workplace whose
→ IDAS identity
(LBNR, with ARBNR/CVRNR keys) is present in year c with ANTNOV ≥ 5 and
→ absent in year c+1,
for closure years c in 2009-2023 (15 cohorts; c = 2023 is feasible because
→ IDAS/IDAN extend
through 2024 for the c+1 identity-absence and worker-flow checks and SSSY
→ through 2025 for
post-year outcomes; ENDNEDL ends 2023), or (ii) a stable control workplace
→ (present with
unchanged identity c-2..c+1, no ENDNEDL, AFGRATE $< 30\%$ and |ANTALXFR| $\leq$
→ 30% in years c-1
and c). Attachment frozen at November of c-1 - before the terminal year -
→ so sinking-ship
attrition during year c cannot condition the sample (S9; Schwerdt 2011).
→ Unit of analysis:
person x event-quarter (aggregated from person x HONUGE fee-week), stacked
→ by workplace-event
cohort in 1:3 matched sets (one exposed worker + up to three
→ riskset-matched controls, controls
weighted 1/k within set).'

inclusion:

- Age 25-59 at baseline, from ALDERIMP on the person's most recent SSSY row
→ in c-3..c-1 (persons
with no SSSY row in the lookup window enter an explicit 'unknown age/sex'
→ matching stratum
rather than being dropped)
- 'Tenure ≥ 1 full year at the workplace: ANSAAR $\leq c-2$ in the c-1 IDAN
→ record'
- 'Private-sector workplace: modal ARB_SEKTORKODE (IDAN, 2008+) among the
→ workplace's year
c-1 employees indicates private sector'
- Workplace November size 5-1000 (ANTNOV in IDAS, year c-1) - cap raised from
→ 500 to 1000
(Design B's cap), an honest eligibility enlargement admitting genuine large
→ closures and
their workers
- 'Event workers: workplace classified true closure (ENDNEDL final closure
→ AND worker-flow
max-successor share $s(w) < 0.30$) or spurious closure (forward identity via
→ IDFREM/REGELFR/FRELFREM
AND $s(w) \geq 0.70$)'

exclusion:

- Events with ambiguous successor share ($0.30 \leq s(w) < 0.70$) - excluded from
→ both classes,
counts reported (GRAFT 4 monitors that this mid-zone holds $< 35\%$ of
→ terminal events)
- Workers whose modal SIKGRUP in baseline year c-1 SSSY rows is 2 (different
→ billing regime;
expected ~1-2%)
- 'Workers appearing in more than one event cohort: assigned to the earliest
→ event only'
- Control-pool workers at workplaces experiencing any terminal event or mass
→ separation (AFGRATE
 $\geq 30\%$) in c-2..c+1

```

- Workers with missing ANSAAR or missing/zero annual pay in the c-1 IDAN
↳ record (share reported;
  expected < 2%)
- 'Mass-separation events at surviving workplaces: excluded from the
↳ confirmatory tier entirely
  (scope discipline)'
expected_n: 'Sampling plan sized, per the remediation directive''s explicit
↳ authorization,
  at the FULL derived expectation''s LOWER bound (not half of it): 15 cohorts
↳ x 10,000-20,000
  eligible exposed workers/cohort-year = 150,000-300,000 exposed true-closure
↳ workers (the
  500->1000 size-cap raise and the 15th cohort add further margin not counted
↳ in the lower
  bound), yielding >= 135,000 matched sets at the registered >= 90%
↳ match-rate floor; spurious-closure
  workers expected at 30-60% of the true count (45,000-90,000+ before
↳ matching). This sizing
  is licensed only jointly with the PRE-OUTCOME REALIZED-N DEMOTION GATE
↳ (Stage 2): confirmatory
  status requires realized matched sets >= N* per member (provisional: N*_H1
↳ = 100,000; N*_H2a
  = 60,000 true + 15,000 spurious for gate F2; N*_H3 = 90,000; binding N*
↳ values are recomputed
  by the re-run deposited simulation at Stage 0), with any shortfall
↳ resolving by demotion
  to exploratory BEFORE any outcome data access - never by floor inflation.
↳ Project-viability
  kill floors (500 true events / 10,000 sets; 150 spurious events / 3,000
↳ sets) are unchanged
  and far below expectation.'
expected_n_basis: 'IDAS row scale from the Forskningservice inventory (~227k
↳ workplace rows/year,
  ~3.2M IDAN rows/year, ~88M SSSY rows/year); closure-rate basis now cited as
↳ required for
  full-expectation sizing: Statistics Denmark business demography
↳ (Erhvervsdemografi/StatBank
  DEMO series) establishment death rates of ~2-5% among employer workplaces,
↳ plus the Danish
  displacement literature''s demonstrated register cohort scale (Browning &
↳ Heinesen 2012,
  J Health Econ, build Danish closure/mass-layoff displacement samples of
↳ comparable order
  from the same IDA infrastructure); GP-contact base rates (~85% annual
↳ contact prevalence,
  ~4-5 contacts/person-year) from published DST health-insurance statistics;
↳ OOH share of
  GP contact volume 8-12% from published Danish out-of-hours studies
↳ (Christensen & Olesen
  1998 BMJ; Moth/Flarup/Christensen/Olesen/Vedsted LV-KOS 2011), simulated at
↳ the conservative
  end (8%).'

```

Power and binding thresholds

```

method: simulation
tool: Python 3 (numpy) Monte Carlo simulation; root seed 20260831; nsim = 500
↳ primary scenario,
  250 per sensitivity scenario; exact set-level-difference equivalents of the
↳ registered fixest::feols

```

specification with CR1 event-workplace clustering
script: power_simulation.py
script_note: deposited beside this protocol (the REMEDIATED Stage-0 re-run of
↳ the deposited
↳ script; the round-1 record survives verbatim in the program's stage-4
↳ design file and is
summarized in remediation_record); harness re-run and independent driver
↳ re-run both reproduced
POWER_JSON exactly; driver stdout deposited as power_simulation_stdout.txt
alpha: 0.05
target_power: 0.95
dgp: 'REMEDIATED re-run of the deposited Stage-0 simulation (seed 20260831),
↳ same core DGP
calibration discipline as the failed round 1 (whose record survives in
↳ stage4_design.json):
person-quarter GP contact counts over a 12-event-quarter panel (4 baseline
↳ quarters in c-1;
Q1-Q4 of terminal year c; Q1-Q4 of c+1), conditionally Poisson with
↳ daytime-GP control mean
1.1 contacts/person-quarter (DST basis), persistent person gamma frailty
↳ var 1.25 (cancels
in within-set DiD) plus transient person-quarter lognormal shocks log-var
↳ 0.25 (does not
cancel) - overdispersion split retained UNCHANGED from round 1, jointly
↳ giving 53.5% zero
person-quarters (band 40-70%) and cross-sectional variance 3.49 - and
↳ event-workplace x
quarter lognormal shocks AR(1) rho 0.5, log-var 0.0375, realized
↳ event-level ICC ~0.014
(band 0.01-0.05). REMEDIATED elements, each an audit-authorized honest
↳ lever with no floor
raised: (1) sampling at the FULL derived expectation's LOWER bound (round
↳ 1 self-handicapped
to half): 135,000 true-closure 1:3 matched sets across 12,600 event
↳ workplaces (~10.7 sets/event,
lognormal size mix; 15 cohorts 2009-2023, Erhvervsdemografi closure-rate +
↳ Browning-Heinesen
register-scale basis) and 45,000 spurious sets across 4,200 events (30%
↳ lower bound of the
30-60% band); (2) 1:3 riskset matching - each set = 1 exposed + 3
↳ independent controls weighted
1/3, control counts simulated exactly as a single Poisson draw of the 3
↳ summed conditional
rates, controls inheriting the set's event cluster (CR1); (3) H2's
↳ confirmatory quantity
re-operationalized from the spurious-arm-dominated differential to theta2a,
↳ the true-arm-only
4-quarter post-year excess at the unchanged 2% floor (round 1's DGP used
↳ 2.5% true post-year;
lowering to the 2.0% floor is conservative), with the spurious arm serving
↳ only falsification
gate F2 (fire iff theta2b's two-sided 95% CI excludes zero positively AND
↳ theta2b >= 50%
of theta2a-hat) and exploratory H2_diff; (4) OOH base rate recalibrated
↳ from the self-declared
over-conservative 3% to mu_OOH = 0.088 = 8.0% of daytime GP volume = 7.4%
↳ of all GP contacts
(conservative end of the published Danish 8-12% range, Christensen & Olesen
↳ 1998; LV-KOS

2011); (5) the registered serial-correlation discount BUILT INTO the DGP
 ↪ rather than assumed
 away: person transient shocks are now AR(1) rho 0.5 (marginals, zero share,
 ↪ variance, ICC
 identical to round 1), which reduces the averaging gain of every
 ↪ multi-quarter window including
 theta2a's 4-quarter mean and weakly raises every SE vs round 1. Effects
 ↪ multiplicative
 on exposed means: linear ramp in year c with mean(Q3,Q4)=2% (theta1 floor;
 ↪ ramp-implied
 Q4-Q1 trend 1.714% for exploratory H1_trend); true post-year 2.0% (theta2a
 ↪ floor); spurious
 residual 0.5% in the hypothesis-true world (retained from round 1, NOT
 ↪ tuned to an idealized
 inert 0%, which is reported as reference); +5-point OOH-vs-daytime relative
 ↪ gap over Q3(c)..Q2(c+1)
 (theta3 floor). Estimators: exact set-level-difference equivalents of the
 ↪ registered feols
 spec (set FE + cohort x calendar-quarter FE reduce to set-level DiD with
 ↪ 1/k-weighted controls),
 CR1 cluster SEs on event workplace; H3 via the plug-in relative-excess
 ↪ contrast with cluster
 SE (delta approximation to the registered 999-draw cluster bootstrap);
 ↪ winsorization omitted
 (conservative). Holm-Bonferroni over the confirmatory family {H1=theta1,
 ↪ H2=theta2a, H3=theta3}
 at FWER 0.05, two-sided p, rejection counted only in the predicted
 ↪ direction; demotion-only
 fixed-point rule at the final family's Holm alphas. nsim=500 primary, 250
 ↪ per sensitivity
 scenario.'
 lowest_effect_basis: 'The registered floor effects, UNCHANGED from the failed
 ↪ round 1 (no
 floor was raised to buy power): H1 = +2% relative pre-boundary excess of
 ↪ the control baseline
 mean (0.022 contacts/person-quarter on a 1.1 base), set below Hamad, Modrek
 ↪ & Cullen (2015)
 layoff-insecurity utilization shifts of ~5-10%; H2 = +2% relative TRUE-ARM
 ↪ post-year excess
 (0.022 contacts/person-quarter; the re-operationalized theta2a - round 1's
 ↪ DGP gave the
 true arm 2.5%, so simulating at the 2.0% floor is a conservative change),
 ↪ set below Hamad/Modrek/Cullen
 (2015) and below Kuhn, Lalive & Zweimueller (2009) post-displacement
 ↪ utilization magnitudes;
 H3 = 5-point gap in relative peri-event excess between OOH and daytime
 ↪ contacts (0.05),
 below published composition shifts - what changed for H3 is only the power
 ↪ basis (OOH base
 rate 8% of daytime GP volume, conservative end of the published Danish
 ↪ 8-12% range), not
 the floor. Exploratory quantities: H1_trend at the ramp-implied 1.714% of
 ↪ baseline (0.019);
 H2_diff at the hypothesis-true +1.5pp differential. These are
 ↪ smallest-effects-of-interest
 per the NHB registered-report convention and the design's own commitments.'
 least_powered_test: 'H1 (theta1, pre-boundary excess): power 0.980 under
 ↪ full-family Holm
 - the least-powered CONFIRMATORY member. The fixed-point iteration
 ↪ CONVERGED AT STEP 1 with

```

zero demotions: final family = {H1, H2, H3} at its own Holm alphas
→ (H1=0.980, H2=0.998,
H3=1.000), all >= 0.95. H1 also carries the largest implied binding N*
→ (~118,000 sets vs
the provisional 100,000 and the 135,000-set plan), making it the member the
→ Stage-2 realized-N
demotion gate binds first.'
results:
- hypothesis_id: H1
test: 'CONFIRMATORY theta1: mean of Q3(c),Q4(c) set-level DiD coefficients
→ vs 4-quarter
c-1 baseline (1:3 sets, controls weighted 1/3); CR1 cluster z on event
→ workplace; Holm
within final family {H1,H2,H3} at FWER 0.05, directional'
assumed_effect: +2% of control baseline mean = 0.022
→ contacts/person-quarter (registered
floor, unchanged; linear ramp within year c)
assumed_n: 135,000 1:3 matched sets / 12,600 true-closure event workplaces
→ (full lower-bound
derived expectation, 15 cohorts 2009-2023)
power: 0.98
- hypothesis_id: H2
test: 'CONFIRMATORY theta2a (re-operationalized core): TRUE-ARM mean
→ post-year (Q1-Q4 of
c+1) DiD excess vs c-1 baseline, no spurious-arm variance; CR1 cluster z;
→ Holm within
final family, directional; serial-correlation discount active (AR(1) rho
→ 0.5 person shocks)'
assumed_effect: +2% relative post-year excess = 0.022
→ contacts/person-quarter (registered
floor, unchanged; round 1's 2.5% lowered to the floor - conservative)
assumed_n: 135,000 1:3 true matched sets / 12,600 event workplaces
power: 0.998
- hypothesis_id: H3
test: 'CONFIRMATORY theta3: 00H-vs-daytime relative peri-event excess gap
→ (Q3(c)-Q2(c+1)),
plug-in ratio contrast with event-workplace cluster SE (delta approx. to
→ registered 999-draw
cluster bootstrap); Holm within final family, directional'
assumed_effect: +5-point relative-excess gap (registered floor, unchanged);
→ 00H base recalibrated
to 0.088 = 8% of daytime GP volume (published Danish 8-12% range,
→ conservative end)
assumed_n: 135,000 1:3 true matched sets / 12,600 event workplaces
power: 1
- hypothesis_id: H1_trend
test: 'EXPLORATORY (gated, never confirmatory): beta_Q4(c) - beta_Q1(c)
→ within-terminal-year
contrast, computed only if theta1 Holm-rejects in predicted direction;
→ joint P(gate &
trend significant at alpha=0.05, directional)'
assumed_effect: 'ramp-implied +1.714% of baseline = 0.019
→ contacts/person-quarter (verified
structurally underpowered for confirmatory status in round 1: 0.380)'
assumed_n: 135,000 1:3 true matched sets / 12,600 event workplaces
power: 0.772
- hypothesis_id: H2_diff
test: 'EXPLORATORY: theta2a - theta2b true-minus-spurious post-year
→ differential (round

```

```

1''s infeasible confirmatory construct), unadjusted directional test at
  ↳ alpha=0.05, reported
  with CI under the exploratory flagging convention'
assumed_effect: +1.5pp differential in the hypothesis-true world (true
  ↳ +2.0% vs spurious
  residual +0.5%)
assumed_n: 135,000 true sets / 12,600 events vs 45,000 spurious sets /
  ↳ 4,200 events
power: 0.46
- hypothesis_id: F2_pass
test: 'F2 GATE OPERATING CHARACTERISTIC (severity device OUTSIDE the Holm
  ↳ family, consumes
  no alpha, not power-bearing): P(gate does NOT fire) under the
  ↳ hypothesis-true DGP; fire
  = theta2b two-sided 95% CI excludes zero positively AND theta2b >= 50% of
  ↳ theta2a-hat;
  registered target >= 0.90'
assumed_effect: spurious residual +0.5% of baseline (retained unchanged
  ↳ from round 1, not
  tuned to inert 0%; strictly-inert reference pass = 0.978)
assumed_n: 45,000 spurious 1:3 sets / 4,200 spurious-closure event
  ↳ workplaces
power: 0.886
- hypothesis_id: F2_trigger
test: 'F2 GATE OPERATING CHARACTERISTIC: P(gate fires) in the
  ↳ placebo-not-inert world; registered
  target >= 0.95 when theta2b = theta2a at floor'
assumed_effect: theta2b = +2.0% of baseline = theta2a at floor
assumed_n: 45,000 spurious 1:3 sets / 4,200 spurious-closure event
  ↳ workplaces
power: 0.79
binding_n_star:
rule: 'Binding pre-outcome demotion thresholds fixed by the deposited
  ↳ Stage-0 re-run (demotion-only:
  where the re-run implies a LOOSER threshold than the provisional value,
  ↳ the stricter provisional
  value stands)'
H1: 118,000 matched sets (re-run value; supersedes the provisional 100,000
  ↳ - tightened)
H2: 95,000 true matched sets (re-run value; supersedes the provisional
  ↳ 60,000 - tightened)
  plus >= 15,000 spurious sets for gate F2's registered operating
  ↳ characteristics
H3: 90,000 matched sets (provisional value retained as the stricter bound
  ↳ over the re-run's
  66,000)
note: 'The 135,000-set sampling plan covers all three, but H1's margin is
  ↳ ~13%: realized
  N below ~87% of the lower-bound expectation demotes H1 first. The Stage-2
  ↳ blind re-simulation
  (realized set counts, realized k distribution, baseline-year moments,
  ↳ realized OOH share)
  is the binding verification; any member short of 0.95 power at its floor
  ↳ demotes to exploratory
  before any post-baseline outcome is accessed.'
sensitivity_notes: 'FIXED-POINT FAMILY: converged at step 1 with ZERO
  ↳ demotions - final confirmatory
  family {H1, H2, H3} at its own Holm alphas: H1=0.980, H2=0.998, H3=1.000
  ↳ (nsim=500; MC SE

```

~0.006-0.014). Estimator recovery unbiased (mean theta-hats
 → 0.0220/0.0219/0.0501 vs targets
 0.022/0.022/0.050). The remediation's own scaling arithmetic is confirmed:
 → H1's mean $z=4.32$
 (design predicted 4.4-4.6 before the serial-correlation discount), H2a
 → $z=4.84$, H3 $z=5.79$,
 vs 4.039 needed for 0.95 at Holm worst-case $\alpha/3$. BINDING N^* (fixes the
 → Stage-2 demotion
 thresholds, z-squared scaling of mean SEs): $N^*_{H1} \sim 118,000$ sets - ABOVE
 → the provisional
 100,000, so the provisional gate must tighten to ~118k; $N^*_{H2a} \sim 95,000$
 → (above provisional
 60,000); $N^*_{H3} \sim 66,000$ (below provisional 90,000; the provisional value
 → may stand as the
 stricter bound). The 135,000-set plan covers all three, but H1's margin is
 → only ~13% -
 realized N below ~87% of the lower-bound expectation demotes H1 first.
 → SENSITIVITY N50 (67,500
 sets - the neighborhood of the provisional gates): full-family Holm powers
 → $H1=0.840$, $H2=0.904$,
 $H3=0.980$ - H1 and H2 fall below 0.95, directly validating that the
 → pre-outcome realized- N
 demotion gate is load-bearing, not decorative. SENSITIVITY EFF50 (floors
 → halved): $H1=0.512$,
 $H2=0.596$, $H3=0.800$ - the design still has essentially no confirmatory
 → resolution below its
 floors; nulls must be reported as CI-ruled-out magnitudes against realized
 → MDEs. CAVEATS
 DRIVING THE VERDICT (F2 gate, outside the Holm family, not power-bearing,
 → but with REGISTERED
 operating characteristics the design claims are Stage-0-verified): (1)
 → trigger probability
 when $\theta_{2b} = \theta_{2a}$ at floor = 0.790 at 45,000 spurious sets, clearly
 → BELOW the registered
 ≥ 0.95 target (MC SE ~0.018); z-scaling implies ~0.95 trigger sensitivity
 → needs ~75-85k
 spurious sets - within the 30-60% band's upper half (up to ~81k sets) but
 → above the lower-bound
 sizing, and far above the 15,000-set F2 viability gate (at 15k sets trigger
 → ~0.5, bounds-only
 reporting per the registered fallback). The gate remains VALID as a
 → severity device - it
 can only subtract support, never add it, and false-fire risk is low - but
 → its sensitivity
 to a floor-sized non-inert placebo is materially below the registered 0.95;
 → the registration
 should either restate the trigger target at its verified value (~0.79 at
 → lower-bound N)
 or add a Stage-2 recomputation of F2 characteristics at realized spurious
 → N . (2) Pass probability
 under the hypothesis-true DGP (spurious residual 0.5% retained from round
 → 1) = 0.886, marginally
 below the registered ≥ 0.90 (within ~1 MC SE, 0.014); under a strictly
 → inert placebo (0%)
 pass = 0.978, comfortably above - the miss reflects the honest 0.5%
 → residual convention,
 not gate malfunction; at N50 pass improves to 0.936 (lower spurious
 → precision fires less).
 (3) The serial-correlation discount (AR(1) ρ 0.5 person transient shocks,
 → marginals unchanged)

raised all SEs vs round 1's iid convention; results are therefore
 ↪ conservative relative
 to the round-1 variance budget - H1 at 0.980 already absorbs the discount.
 ↪ (4) Exploratory
 tiers confirm the remediation's tiering decisions: H1_trend joint gated
 ↪ power 0.772 and
 H2_diff 0.460 - both would fail as confirmatory members, exactly as
 ↪ diagnosed in round 1
 (0.380/0.358); their exploratory demotion is validated, not arbitrary.
 ↪ VERDICT RATIONALE:
 all three confirmatory members clear 0.95 with the family intact
 ↪ (adequately powered on
 the task's own criterion), but the F2 falsification gate misses its two
 ↪ registered operating-characteristic
 targets at lower-bound spurious N, so the design is powered_with_caveats:
 ↪ registrable, provided
 the F2 characteristic targets are restated to the verified values (or F2
 ↪ characteristics
 are re-verified at realized spurious N in the Stage-2 blind re-simulation)
 ↪ and the binding
 N*_H1 is recorded as ~118,000 sets rather than the provisional 100,000.'
 verdict: powered_with_caveats
 verification: 'Deposited script re-run verified by the harness (exit 0, seed
 ↪ 20260831, POWER_JSON
 matched exactly: H1 0.980, H2 0.998, H3 1.000, H1_trend 0.772, H2_diff
 ↪ 0.460, F2_pass 0.886,
 F2_trigger 0.790; calibration in band: zero-share 0.535, OOH 8.0% of
 ↪ daytime volume, ICC
 0.014). Independent driver reproduction on 2026-08-31: 3 differing stdout
 ↪ lines of 45, all
 wall-clock stamps; every numeric result identical. Driver reconciliation of
 ↪ the powered_with_caveats
 verdict, both actions carried in this protocol: (1) F2 operating
 ↪ characteristics restated
 from aspirational targets to the VERIFIED values with Stage-2
 ↪ re-verification at realized
 spurious N registered as binding; (2) N*_H1 recorded as 118,000 sets,
 ↪ superseding the provisional
 100,000. Frozen bank hypothesis wording untouched.'

Current authoritative contingencies

- condition: ENDNEDL flag-year convention differs in-environment from the
 ↪ assumed 'flagged
 in the workplace's last active year'
 rule: Re-anchor mechanically so that c always denotes the last IDAS year
 ↪ with ANTNOV >=
 5 for the identity; the estimand is unchanged; the recalibration is
 ↪ reported verbatim
 - condition: ARB_SEKTORKODE is absent from the analysis-server IDAN extract
 ↪ (publicly documented
 on TIMES ida-ansattelser but unconfirmed in the recorded
 ↪ Forskningservice inventory;
 presence and 2008+ coverage validated at Stage 1 before any
 ↪ classification work)
 rule: 'Frozen fallback: define the private-sector filter from
 ↪ ARB_HOVED_BRA_DB07 instead
 - a workplace counts as private if and only if its NACE DB07 section is
 ↪ not 0 (public

administration and defence), P (education), or Q (human health and social
↳ work); the substitution
and its effect on eligible-workplace and worker counts are reported
↳ verbatim; no other
filter change is permitted'

- condition: Match rate < 95% of eligible exposed workers at target k = 3
↳ under full exact
strata
rule: 'Apply the pre-specified coarsening ladder in order, stopping at the
↳ first rung reaching
95% of exposed matched with k >= 1: (1) NACE section -> 10 aggregate
↳ industry groups;
(2) drop region stratum; (3) age bands 5-year -> 10-year. Never drop sex,
↳ cohort, size
class, or the baseline-billing calipers; realized k distribution reported
↳ and fed to the
Stage-2 blind re-simulation'
- condition: Median HONUGE settlement lag (BEHANDLINGSDATO calibration, 2021+
↳ rows) > 4 weeks
rule: 'Guard band: exclude Q1(c+1) from the theta2a/theta2b windows (means
↳ over Q2(c+1)..Q4(c+1))
and report theta1 both as registered and with a 4-week-shifted boundary;
↳ the registered
theta1 window remains primary'
- condition: KONTAKT does not behave as a row-level contact marker on
↳ in-environment validation
rule: 'Switch to the frozen fallback contact definition: count of distinct
↳ PNR x HONUGE-week
x YDERNR x GP-code-group cells with >= 1 positive-fee row; switch and
↳ level effect reported;
all estimands re-run under both definitions with the fallback as primary'
- condition: OLS and Poisson QML disagree in sign on any primary quantity
rule: Foreground the pre-committed OLS estimate, report both, flag the
↳ discordance in the
Summary paragraph; no post hoc model selection
- condition: 'c-2 placebo panel FAILURE, defined as: placebo contrast
↳ significant at alpha
= 0.05 AND |estimate| > 1% of control baseline mean (50% of the theta1
↳ floor effect) -
the single registered failure condition, stated identically in
↳ outcome_neutral_checks
and the H1 placebo kill'
- rule: 'Execute the single pre-specified re-match adding year c-2 baseline
↳ contacts to the
Mahalanobis set; if the placebo fails again under the same condition, the
↳ entire registered
event-study tier (confirmatory and exploratory event-study quantities) is
↳ killed and the
study reports descriptive trajectories only. The two non-complementary
↳ cells are pre-committed
passes with mandatory disclosure, never discretion: (a) significant at
↳ alpha = 0.05 with
|estimate| <= 1% - no re-match, no kill; the Table T7 placebo row flags
↳ the significance
and states that a sub-floor-scale imbalance was detected; (b) not
↳ significant with |estimate|
> 1% - no re-match, no kill; a mandatory caution paragraph reports the
↳ placebo magnitude
against the theta1 floor and the realized placebo MDE, and the H1/H2
↳ Summary paragraphs

must cite it'

- condition: A matched control's own workplace acquires a terminal event
 - ↪ within c-2..c+1 after matching (data revision)
 - rule: Drop that control from the set; drop the whole set only if no control remains; dropped counts reported; no re-matching
- condition: YDLTID missing or code 0 ('not relevant') on a GP row
 - rule: Classify as daytime (conservative against the theta3 prediction);
 - ↪ missingness share reported; if > 20% of peri-event GP rows, H3's kill criterion triggers
- condition: Exact tie in Mahalanobis distance for a nearest-neighbor match
 - rule: Deterministic tie-break by smallest absolute baseline-contact difference, then seeded RNG (seed 20260831)
- condition: IDAS-flag vs worker-flow classification agreement < 90% but >= 75%
 - rule: Proceed with the worker-flow classification as registered (the canon, S7); publish
 - ↪ the full disagreement cross-tabulation; below 75%, H2's kill criterion governs
- condition: 'GRAFT 3: sign of theta2b or theta2_diff not stable across the three successor-share threshold pairs'
 - rule: Flag classification fragility in the pre-committed H2 Summary
 - ↪ paragraph; the 30/70 estimate remains primary; no threshold shopping
- condition: 'GRAFT 6: exposed-vs-control difference in later-presence rates (any IDAN record in year c+1 or any SSSY row in c+1 or later) exceeds 2 percentage points'
 - rule: A mandatory bounding paragraph accompanies all post-boundary
 - ↪ (theta2a/theta2b) conclusions, converting the differential into best/worst-case bounds; trigger and threshold registered, not discretionary
- condition: 'GRAFT 9: the control forward-window sensitivity yields SMALLER magnitudes than the primary estimates'
 - rule: Flag as diagnostic of something other than contamination (registered direction note);
 - ↪ reported verbatim, no re-specification
- condition: 'STAGE-2 REALIZED-N GATE: realized matched sets (or spurious sets for F2) fall below a member's binding N* from the Stage-0 re-run, or the Stage-2 blind re-simulation (realized Ns, realized k distribution, baseline-year moments only) shows power < 0.95 for a member at its floor'
 - rule: That member is demoted to exploratory BEFORE any post-baseline
 - ↪ outcome is accessed;
 - the Holm family shrinks and alphas recompute by the registered
 - ↪ fixed-point iteration among survivors; demotion-only - no member, threshold, or floor may move in the
 - ↪ other direction;
 - the realized-power/MDE table is published either way
- condition: Realized control-group OOH share of GP contacts < 5% (below the published 8-12% range)
 - rule: H3's power basis is re-verified in the Stage-2 blind re-simulation at
 - ↪ the realized

```

share; if < 0.95 at the 5-point floor, H3 demotes to exploratory per the
  ↪ realized-N gate;
  if realized share < 2%, H3's kill criterion governs instead
- condition: Spurious arm killed at Stage 1/2 (events < 150, sets < 3,000,
  ↪ classification
  agreement < 75%, or bimodality unresolved)
rule: Gate F2 and theta2_diff/theta2b are reported unavailable (drop
  ↪ disclosed before outcome
  access); theta2a remains in the Holm family if its own N* is met, but the
  ↪ H2 verdict is
  capped at 'partially consistent - comparative quantifier untestable' and
  ↪ the Discussion
  states that the excess cannot then be distinguished from selection into
  ↪ disappearing workplace
  identities

```

2. ENVIRONMENT

Software stack (record exact versions in the run manifest at Stage 0 and freeze them):

- **python3 (>= 3.11) with polars** for ALL large-table preprocessing. Lazy scans only (`scan_parquet/scan_csv`) with projection and predicate pushdown; semi-joins against the cohort PNR roster before any collect; streaming collects for aggregations. Never materialize a person x week or person x month panel beyond need: the only dense person-time table permitted is matched members x 16 event-quarters (~540k true members + ~180k spurious members, x 16 = under 12M rows). Zero-filling happens at the quarter level for panel members only, never by cross-joining the population against a calendar.
- **R (>= 4.3) with fixest, invoked via Rscript**, for every fixed-effects estimation: `feols` (OLS, the pre-committed primary) and `fepois` (Poisson QML robustness). CR1 cluster-robust SEs on event workplace. No estimation in python; no data manipulation in R beyond reading the analysis parquet and writing tidy coefficient tables.
- **matplotlib** for all figures; **booktabs LaTeX** for all tables (no vertical rules). Every number in every exhibit is written by script from a parquet/CSV output — never hand-edited.
- **Memory discipline**: process cohort-by-cohort ($c = 2009..2023$) where possible; checkpoint every numbered pipeline step as parquet under `interim/` with a fixed name (given per step below); SSSY scans filter on (PNR in roster) AND (SPEC2 in frozen lists) at scan level; the full population is touched only by the aggregate coverage benchmark (step 16), which is computed as a streaming groupby per year, never collected at row level.
- **Seeded randomness**: a single root seed 20260831 (per the frozen registration) initializes every RNG: matching processing order, Mahalanobis tie-breaks, the 999-draw cluster bootstrap, and the Stage-2 blind re-simulation. No unseeded randomness anywhere.
- **Run manifest**: an append-only JSONL (`manifest/run_manifest.jsonl`), one record per pipeline step per run: step id, UTC timestamp, code version hash, input file(s) + row counts, output file + row count, parameters used, validation check results, PASS/FAIL. Every step below MUST log its row counts; a step whose check

fails writes FAIL and halts downstream steps (Section 6.3). The manifest (aggregate counts only, event counts rounded per the disclosure rules) is part of the released audit trail.

- **Directory convention:** `raw/` (DST-delivered), `interim/` (parquet checkpoints), `analysis/` (panels), `out/tables/`, `out/figures/`, `manifest/`. All paths in code are relative to the project root inside the research environment; no absolute paths in any released script.
- **Frozen-code discipline:** the Stage-4 estimation code is frozen at the end of Stage 0 (hash recorded in the manifest); Stage 4 is a single run of that code. The sensitivity suite runs once and is reported together. Any code change after Stage 0 is a deviation and must be disclosed in running prose with reason and approver.

Source disclosure rules

```
environment: DST Forskningservice (no prior data access; gatekeeper
↳ certification per RR
secondary-data rules). All analysis occurs inside the environment; only
↳ aggregate outputs
leave. No microdata, workplace identifiers (ARBNR/CVRNR/LBNR), or
↳ person-level values are
released.
minimum_cell_size: 5
minimum_cell_rules: no released statistic based on fewer than 5 persons or
↳ fewer than 3 workplaces;
event-count tables rounded to the nearest 10; weekly trajectory figures
↳ pool >= 100 events
and >= 5,000 person-weeks per plotted point; fee (BRUHON) aggregates
↳ subject to a 50% single-workplace
dominance check; min-cell suppression applied BEFORE Holm reporting so
↳ suppression can never
be outcome-selective
outputs_leaving_environment:
- model coefficients with SEs/CIs
- matching balance tables and realized-k distribution
- calibration and benchmark tables (coverage, linkage, HONUGE lag, OOH share)
- figure aggregates (weekly trajectory series under the pooling rules)
- event counts rounded to the nearest 10
- Stage-0 power/MDE tables - BOTH rounds (failed and remediated) - and the
↳ Stage-2 blind re-simulation
output with the demotion trail
- the c-2 placebo, gate F2 outputs, GRAFT-2/4 diagnostics, GRAFT-6
↳ later-presence rates, E4
attrition panel, and all outcome-neutral check outputs (the full audit
↳ trail exits under
the same rules)
- the analysis and power scripts
statement: 'All analysis occurs inside DST Forskningservice; only aggregate
↳ outputs leave.
Rules applied to every released object: (1) no statistic based on fewer
↳ than 5 persons or
fewer than 3 workplaces; (2) event-count tables rounded to the nearest 10;
↳ (3) weekly trajectory
figures pool >= 100 events and >= 5,000 person-weeks per plotted point; (4)
↳ fee (BRUHON)
```

statistics released only as cell means/coefficients over min-N cells, with
 → a dominance check
 that no single workplace contributes > 50% of any released fee aggregate;
 → (5) released objects
 are limited to: model coefficients with SEs/CIs, balance tables,
 → calibration/benchmark tables,
 figure aggregates, event counts (rounded), the Stage-0 (both rounds) and
 → Stage-2 power/MDE
 tables with the demotion trail, and the analysis/power scripts - never
 → microdata, workplace
 identifiers (ARBNR/CVRNR/LBNR), or person-level values; (6) min-cell
 → suppression is applied
 before Holm reporting so that suppression can never be outcome-selective;
 → (7) the c-2 placebo,
 gate F2 outputs, the GRAFT-2/4 diagnostics, the GRAFT-6 later-presence
 → rates, the E4 attrition
 panel, and all outcome-neutral check outputs are released under the same
 → rules, ensuring
 the full audit trail exits the environment.'

Complete registration fields

```

id: atom-000002
topic: Matched cohort validity
atom_result: prereg
atom_strength: null
atom_shape: null
evidence_paths: []
concepts: []
literature: []
bonds: []
preregistration:
  version: 1
  mode: adaptive
  question: Can the registered matched cohorts and pre-outcome inferential
    eligibility be established?
  rationale: Source-existing shared validity judgment, required by several
    subsequent questions.
  assumption: |-
    Associational - matched event-study contrasts, a contrastive
    → falsification gate, and a composition contrast; descriptive for
    → trajectory panels and the bounded priority claim. Binding vocabulary
    → rules: 'is associated with' / 'consistent with' / 'suggests' only;
    → 'effect', 'impact', 'causes' prohibited throughout; anticipation
    → always worded as 'pre-boundary excess' relative to the annually dated
    → closure-year boundary, never as announcement-dated anticipation (S8);
    → no phenomenon-discovery or Danish-first framing (S1-S3); H2 novelty
    → confined to the placebo's application to health billing (S7). H2
    → verdict semantics are fixed in advance: 'consistent with H2' requires
    → BOTH the confirmatory core (theta2a Holm-significant, predicted
    → direction) AND a passing falsification gate F2; the strict
    → comparative magnitude is only ever reported as an exploratory
    → estimate with its CI. Exploratory results (H1_trend, H2_diff, E1, E3,
    → E4, and any gate-demoted member) carry the word 'exploratory' in the
    → heading and in every sentence stating them. The prose label never
    → exceeds the evidence label.
    Compatibility ratings below are neutral import metadata, not a new
    → scientific evaluation or empirical scores.
  target_strength: descriptive
  target_shape: comparison
  
```

```

ratings:
  importance: 3
  novelty: 3
  feasibility: 3
  data_advantage: 3
  informativeness: 0.5
tables:
- IDAN
- IDAS
- SSSY
activation: state(atom-000001) == "classifiable"
verdicts:
- key: cohort_ready
  region: At least 10,000 true matched sets, registered balance/matching
    reconciled, and the baseline-only demotion trail archived. This may
    include zero surviving confirmatory members.
  says: A matched cohort is available; the component-level demotion trail,
    not this state, determines confirmatory eligibility.
- key: project_stop
  region: Fewer than 10,000 true matched sets after the source procedure.
  says: Source project viability kill; no outcome study under this
    ↪ registration.
- key: boundary
  region: Required source-defined inputs or scientific definitions are
    ↪ demonstrably
    unavailable; retain available components and the exact unresolved scope.
  says: An evidenced boundary, not a statistical null. A missing
    ↪ implementation
    or timeout alone is operational no_result, not this state.
- key: otherwise
  region: otherwise
  says: The registered judgment is unresolved; report all completed
    ↪ components
    without inventing support, equivalence or absent results.
protocol: |
  # Can the registered matched cohorts and pre-outcome inferential
    ↪ eligibility be established?

## Source fidelity and implementation order

This is a format conversion of DS1.1, not a new study or independent
  ↪ acceptance.
Source audit: minor_revisions. No observed Danish outcomes are supplied.
The source execution plan explicitly states that final prereg.yaml governs
disagreements. Exact current YAML selections below therefore govern older
execution-plan excerpts; these excerpts are retained as implementation
  ↪ detail,
not a competing registration. See docs/provenance/source-discrepancies.md.
The complete shared data/validity contracts are embedded in Atoms 1--3;
  ↪ read
those native protocols when executing a downstream Atom, even without
  ↪ docs/.
Do not substitute a scientific default for an unresolved source
  ↪ contradiction.

Preserve source stages 0 -> 1 -> 2 -> 3 -> 4 -> 5 inside the
  ↪ implementation.
Source Stage-0 simulations are historical design evidence, not study
  ↪ outcomes;

```

do not silently redo or improve them. Gate 14 must be evaluated from
 ↳ baseline
 data before post-baseline outcomes. All surviving Stage-4 quantities form
 ↳ ONE
 Holm family. The three question-level Atoms may share the same frozen
 ↳ family
 fit and tidy results; do not apply Holm separately per Atom or adjudicate
 ↳ a
 member before the complete surviving family's p-values exist. Cache/reuse
 identical governed intermediate results with provenance when available; a
 direct-dispatch call must obtain or compute its exact source-defined
 ↳ inputs,
 not pretend its predecessors succeeded. Stage-5 internal riders remain
 exploratory, after the Stage-4 family is frozen, even when housed in
 ↳ H1/H2.

Reference files under docs/execution/ are the ORIGINAL SKELETONS, not
 ↳ tested
 turnkey implementations. Adapt paths and complete engineering in a
 ↳ governed
 workspace; preserve source scientific rules and disclose unresolved
 ↳ choices.
 Write a result record with component estimates, uncertainty, family
 ↳ eligibility,
 source interpretation row, diagnostics and boundary/skip dispositions.
 ↳ Released
 outputs follow source disclosure rules; confidential rows stay governed.
 Local native states summarize the source judgment; neither a state name
 ↳ nor
 source power forecasts prove equivalence or causal identification. A
 ↳ bounded
 claim requires its actual interval and registered rule. Retain unresolved
 ↳ source
 wording conflicts; do not convert mere nonsignificance into scientific
 ↳ null.

Frozen scope and vocabulary

```
```yaml
claim_class_ceiling: associational
statement: 'Associational - matched event-study contrasts, a contrastive
↳ falsification gate,
 and a composition contrast; descriptive for trajectory panels and the
 ↳ bounded priority claim.
Binding vocabulary rules: 'is associated with' / 'consistent with'
↳ / 'suggests' only;
'effect', 'impact', 'causes' prohibited throughout; anticipation
↳ always worded as
'pre-boundary excess' relative to the annually dated closure-year
↳ boundary, never as announcement-dated
anticipation (S8); no phenomenon-discovery or Danish-first framing
↳ (S1-S3); H2 novelty confined
to the placebo's application to health billing (S7). H2 verdict
↳ semantics are fixed in
advance: 'consistent with H2' requires BOTH the confirmatory core
↳ (theta2a Holm-significant,
predicted direction) AND a passing falsification gate F2; the strict
↳ comparative magnitude
is only ever reported as an exploratory estimate with its CI.
↳ Exploratory results (H1_trend,
```

H2\_diff, E1, E3, E4, and any gate-demoted member) carry the word  
 ↳ 'exploratory' in the  
 heading and in every sentence stating them. The prose label never  
 ↳ exceeds the evidence label.'

frozen\_parameters:  
 seed: 20260831  
 gp\_code\_list: daytime GP = SPEC2 in {70..80}  
 ooh\_code\_list: out-of-hours GP = SPEC2 in {81, 82, 83, 84, 89}  
 ooh\_robustness\_classifier: adds YDLTID in {2, 3, 4, 5, 7, 8, 9}  
 spec2\_evening\_gp\_disclosure: 'Registered choice, disclosed: the parsed  
 ↳ value labels mark  
 SPEC2 71 ("General practitioner, consultation, evening etc."), 73  
 ↳ ("General practitioner,  
 telephone consultation, evening etc."), and 75 ("General  
 ↳ practitioner, home visit, evening  
 etc.") as explicitly off-schedule GP services; they nonetheless  
 ↳ remain on the daytime  
 side of the PRIMARY classifier because H3's out-of-hours construct  
 ↳ and its published  
 power basis are keyed to the laegevagt out-of-hours service (SPEC2  
 ↳ 81-84, 89), the quantity  
 the published Danish 8-12% share describes. Attenuation direction,  
 ↳ both readings: counting  
 evening GP as daytime attenuates theta3 toward zero - conservative  
 ↳ against the positive  
 H3 conclusion, anti-conservative for the null-bounding  
 ↳ "compositionally unremarkable"  
 reading, whose Stage-2 statement therefore carries a mandatory  
 ↳ caveat. In-environment  
 validation must verify that the YDLTID-augmented robustness  
 ↳ classifier captures SPEC2  
 71/73/75 rows (their evening time codes 2-5); if YDLTID proves  
 ↳ unpopulated on those rows,  
 the robustness classifier - and only the robustness classifier - is  
 ↳ amended mechanically  
 and with disclosure to include SPEC2 71/73/75 directly. The SPEC2  
 ↳ 71/73/75 volume share  
 is reported, and both classifiers appear side by side in Table T7.'

successor\_share\_thresholds: 'true closure: ENDNEDL final closure AND  
 ↳ worker-flow max-successor  
 share  $s(w) < 0.30$ ; spurious closure: forward identity via  
 ↳ IDFREEM/REGELFR/FRELFREM AND  
 $s(w) \geq 0.70$ ; ambiguous mid-zone excluded; GRAFT-3 sensitivities  
 ↳ 0.20/0.80 and 0.40/0.60'

quarter\_mapping: 'ISO week of HONUGE: 1-13 / 14-26 / 27-39 / 40-52; ISO  
 ↳ week 53 to Q4'

winsorization: person-quarter outcomes at the 99.9th percentile within  
 ↳ era x calendar quarter

matching: 1:3 nearest-neighbor Mahalanobis within exact strata, without  
 ↳ replacement across  
 sets, target  $k = 3$ , minimum  $k = 1$ ; calipers  $|dif \text{ baseline contacts}|$   
 ↳  $\leq 3$  and  $|dif \log(1+BRUHON)|$   
 $\leq 0.5$ ; exposed processed in seeded random order (seed 20260831)

software: R  $\geq 4.3$ , fixest::feols (OLS) and fixest::fepois (Poisson QML  
 ↳ robustness); versions  
 frozen in the registered script

note: Code lists and thresholds frozen before any data access; code  
 ↳ labels quoted verbatim  
 from parsed DST value labels and re-validated in-environment before  
 ↳ estimation (frozen

```

... fallback rules in contingent_decisions).

9. **Exposed cohort extraction at November c-1.** Input: steps 3-8.
Workers with primary November-status employment in year c-1 at an
event workplace; apply inclusion rules (tenure ANSAAR <= c-2;
private sector; ANTNOV 5-1000 in c-1) and the IDAN-side exclusions
(missing ANSAAR or missing/zero annual pay, counts reported, expected
< 2%; multi-cohort workers assigned to earliest event only). Age/sex
and SIKGRUP screens complete in step 11. Output:
`interim/exposed_cohort_raw.parquet`. Check: eligible exposed per
cohort against the 10k-20k basis; exclusion accounting logged
line-by-line (feeds exhibit T3).

10. **Stable-control pool.** Input: steps 1-6. Control workplaces:
identity present and unchanged c-2..c+1, no ENDNEDL, AFGRATE < 30%
and |ANTALXFR| <= 30% in years c-1 and c; exclude workplaces with
any terminal event or mass separation (AFGRATE >= 30%) in c-2..c+1.
Control workers: primary November c-1 employment at such workplaces,
same inclusion rules as step 9. Output:
`interim/control_pool.parquet`. Check: pool size per stratum orders
of magnitude above exposed (log the ratio); zero overlap with any
event workplace.

11. **SSSY pre-baseline extraction, years c-3..c-1 (Stage 2 begins;
pre-baseline rows only).** Input: raw SSSY, roster = steps 9-10.
Lazy scan filtered to roster PNRs; extract rows for the frozen
GP/OOH code lists plus SIKGRUP/ALDERIMP/KOENIMP. Derive: age 25-59
at baseline from ALDERIMP on the person's most recent SSSY row in
c-3..c-1 (no row -> explicit "unknown age/sex" stratum, not
dropped); modal SIKGRUP in c-1 (modal group 2 -> excluded, share
reported, expected ~1-2%); year c-1 matching covariates: total GP
contacts, OOH contacts, log(1+BRUHON), weeks with >= 1 contact.
Output: `interim/sssy_baseline.parquet`. Check: share of roster with
>= 1 SSSY row in c-3..c-1 logged (precursor to the 80% linkage
floor); unknown-stratum share reported; SIKGRUP exclusion share
reported.

12. **Matching frame.** Input: steps 9-11 + IDAN covariates. Exact
strata: cohort c x sex (KOENIMP, incl. unknown) x 5-year age band
(ALDERIMP, incl. unknown) x NACE section (ARB_HOVED_BRA_DB07) x
workplace size class (ANTNOV: 5-9/10-19/20-49/50-99/100-499/
500-1000) x region (workplace KOM mapped to the 5 post-2007
regions). Mahalanobis components (all year c-1): total GP contacts,
log(1+BRUHON), OOH contacts, weeks with >= 1 contact, tenure (c-1
minus ANSAAR), wage decile (BREDT_LOEN_BELOEB, computed within
cohort). Output: `interim/matching_frame.parquet`. Check: covariate
completeness; stratum cardinalities and thin-stratum share logged.

13. **1:3 riskset matching.** Input: step 12. Nearest-neighbor
Mahalanobis within exact strata, without replacement of controls
across sets, target k = 3, minimum k = 1; calipers |dif baseline
contacts| <= 3 and |dif log(1+BRUHON)| <= 0.5; exposed processed in
seeded random order (seed 20260831); exact ties broken by smallest
absolute baseline-contact difference, then seeded RNG. If match rate
< 95% of eligible exposed at target k = 3 under full strata, apply
the coarsening ladder IN ORDER, stopping at the first rung reaching
95% matched with k >= 1: (1) NACE section -> 10 aggregate industry
groups; (2) drop region stratum; (3) age bands 5-year -> 10-year.
Never drop sex, cohort, size class, or the billing calipers. Run for
the true arm and separately for the spurious arm (own controls).
Outputs: `interim/matched_sets.parquet`,
`interim/matched_sets_spurious.parquet`. Checks (Stage-2 stop
rules): match rate >= 90% after the full ladder; standardized mean
difference <= 0.10 on every matching covariate and on unmatched

```

auxiliaries (control aggregates weighted 1/k within set); realized k distribution and ladder rung logged; true sets  $\geq 10,000$  (project kill below); spurious sets  $\geq 3,000$  (spurious-arm kill below).

14. **\*\*STAGE-2 PRE-OUTCOME REALIZED-N GATE (blind; binding).\*\*** Input: realized set counts, realized k distribution, baseline-year (c-1) moments, realized control OOH share of baseline GP contacts - no post-baseline outcome touched. Re-run the deposited simulation (``power_simulation.py`` machinery, seed 20260831) at these realized quantities; recompute F2 operating characteristics at realized spurious N. Decisions (demotion-only): any confirmatory member with realized sets  $< N^*$  (H1 118,000; H2a 95,000 true + 15,000 spurious for F2; H3 90,000) or blind re-simulated power  $< 0.95$  at its floor is demoted to exploratory NOW; the Holm family shrinks by the registered fixed-point iteration among survivors; nothing is ever promoted. If realized control OOH share  $< 5\%$ , H3's re-verification governs (demote if  $< 0.95$  at the 5-point floor);  $< 2\%$  invokes H3's kill criterion at Stage 3. Output: ``interim/realized_power_mde.parquet`` + the demotion trail (released regardless of outcome; feeds exhibit T5).

**## Population**

```

```yaml
definition: 'Registered cohorts (SSSY era): persons with primary
  ↳ November-status employment
    in year c-1 (IDAN; primary job = largest BREDT_LOEN_BELOEB in c-1, ties
      ↳ broken by larger
        LOENTIMER then lower ARBNR - deterministic) at (i) an event workplace
          ↳ whose IDAS identity
            (LBNR, with ARBNR/CVRNR keys) is present in year c with ANTNOV  $\geq 5$  and
              ↳ absent in year c+1,
                for closure years c in 2009-2023 (15 cohorts; c = 2023 is feasible
                  ↳ because IDAS/IDAN extend
                    through 2024 for the c+1 identity-absence and worker-flow checks and
                      ↳ SSSY through 2025 for
                        post-year outcomes; ENDNEDL ends 2023), or (ii) a stable control
                          ↳ workplace (present with
                            unchanged identity c-2..c+1, no ENDNEDL, AFGRATE  $< 30\%$  and |ANTALXFR|
                              ↳  $\leq 30\%$  in years c-1
                                and c). Attachment frozen at November of c-1 - before the terminal year
                                  ↳ - so sinking-ship
                                    attrition during year c cannot condition the sample (S9; Schwerdt
                                      ↳ 2011). Unit of analysis:
                                        person x event-quarter (aggregated from person x HONUFE fee-week),
                                          ↳ stacked by workplace-event
                                            cohort in 1:3 matched sets (one exposed worker + up to three
                                              ↳ riskset-matched controls, controls
                                                weighted 1/k within set).'
```

inclusion:

- Age 25-59 at baseline, from ALDERIMP on the person's most recent SSSY
 - ↳ row in c-3..c-1 (persons
 - with no SSSY row in the lookup window enter an explicit 'unknown
 ↳ age/sex' matching stratum
 rather than being dropped)
- 'Tenure ≥ 1 full year at the workplace: ANSAAR $\leq c-2$ in the c-1 IDAN
 - ↳ record'
- 'Private-sector workplace: modal ARB_SEKTORKODE (IDAN, 2008+) among the
 - ↳ workplace's year
 c-1 employees indicates private sector'

- Workplace November size 5-1000 (ANTNOV in IDAS, year c-1) - cap raised
 - ↳ from 500 to 1000
 - (Design B's cap), an honest eligibility enlargement admitting genuine
 - ↳ large closures and
 - their workers
- 'Event workers: workplace classified true closure (ENDNEDL final
 - ↳ closure AND worker-flow
 - max-successor share $s(w) < 0.30$) or spurious closure (forward identity
 - ↳ via IDFREM/REGELFR/FRELFREM
 - AND $s(w) \geq 0.70$)'

exclusion:

- Events with ambiguous successor share ($0.30 \leq s(w) < 0.70$) - excluded
 - ↳ from both classes,
 - counts reported (GRAFT 4 monitors that this mid-zone holds $< 35\%$ of
 - ↳ terminal events)
- Workers whose modal SIKGRUP in baseline year c-1 SSSY rows is 2
 - ↳ (different billing regime;
 - expected ~1-2%)
- 'Workers appearing in more than one event cohort: assigned to the
 - ↳ earliest event only'
- Control-pool workers at workplaces experiencing any terminal event or
 - ↳ mass separation (AFGRATE
 - $\geq 30\%$) in c-2..c+1
- Workers with missing ANSAAR or missing/zero annual pay in the c-1 IDAN
 - ↳ record (share reported;
 - expected $< 2\%$)
- 'Mass-separation events at surviving workplaces: excluded from the
 - ↳ confirmatory tier entirely
 - (scope discipline)'

expected_n: 'Sampling plan sized, per the remediation directive's

- ↳ explicit authorization,
- at the FULL derived expectation's LOWER bound (not half of it): 15
 - ↳ cohorts x 10,000-20,000
- eligible exposed workers/cohort-year = 150,000-300,000 exposed
 - ↳ true-closure workers (the
 - 500->1000 size-cap raise and the 15th cohort add further margin not
 - ↳ counted in the lower
 - bound), yielding $\geq 135,000$ matched sets at the registered $\geq 90\%$
 - ↳ match-rate floor; spurious-closure
 - workers expected at 30-60% of the true count (45,000-90,000+ before
 - ↳ matching). This sizing
 - is licensed only jointly with the PRE-OUTCOME REALIZED-N DEMOTION GATE
 - ↳ (Stage 2): confirmatory
 - status requires realized matched sets $\geq N^*$ per member (provisional:
 - ↳ $N^*_{H1} = 100,000$; N^*_{H2a}
 - $= 60,000$ true + 15,000 spurious for gate F2; $N^*_{H3} = 90,000$; binding N^*
 - ↳ values are recomputed
 - by the re-run deposited simulation at Stage 0), with any shortfall
 - ↳ resolving by demotion
 - to exploratory BEFORE any outcome data access - never by floor
 - ↳ inflation. Project-viability
 - kill floors (500 true events / 10,000 sets; 150 spurious events / 3,000
 - ↳ sets) are unchanged
 - and far below expectation.'

expected_n_basis: 'IDAS row scale from the Forskningservice inventory

- ↳ (~227k workplace rows/year,
- ~3.2M IDAN rows/year, ~88M SSSY rows/year); closure-rate basis now
 - ↳ cited as required for
- full-expectation sizing: Statistics Denmark business demography
 - ↳ (Erhvervsdemografi/StatBank

```

DEMO series) establishment death rates of ~2-5% among employer
→ workplaces, plus the Danish
displacement literature's demonstrated register cohort scale (Browning
→ & Heinesen 2012,
J Health Econ, build Danish closure/mass-layoff displacement samples of
→ comparable order
from the same IDA infrastructure); GP-contact base rates (~85% annual
→ contact prevalence,
~4-5 contacts/person-year) from published DST health-insurance
→ statistics; OOH share of
GP contact volume 8-12% from published Danish out-of-hours studies
→ (Christensen & Olesen
1998 BMJ; Moth/Flarup/Christensen/Olesen/Vedsted LV-KOS 2011),
→ simulated at the conservative
end (8%).'
'''

```

Power and binding thresholds

```

'''yaml
method: simulation
tool: Python 3 (numpy) Monte Carlo simulation; root seed 20260831; nsim =
→ 500 primary scenario,
250 per sensitivity scenario; exact set-level-difference equivalents of
→ the registered fixest::feols
specification with CR1 event-workplace clustering
script: power_simulation.py
script_note: deposited beside this protocol (the REMEDIATED Stage-0
→ re-run of the deposited
script; the round-1 record survives verbatim in the program's stage-4
→ design file and is
summarized in remediation_record); harness re-run and independent
→ driver re-run both reproduced
POWER_JSON exactly; driver stdout deposited as
→ power_simulation_stdout.txt
alpha: 0.05
target_power: 0.95
dgp: 'REMEDIATED re-run of the deposited Stage-0 simulation (seed
→ 20260831), same core DGP
calibration discipline as the failed round 1 (whose record survives in
→ stage4_design.json):
person-quarter GP contact counts over a 12-event-quarter panel (4
→ baseline quarters in c-1;
Q1-Q4 of terminal year c; Q1-Q4 of c+1), conditionally Poisson with
→ daytime-GP control mean
1.1 contacts/person-quarter (DST basis), persistent person gamma
→ frailty var 1.25 (cancels
in within-set DiD) plus transient person-quarter lognormal shocks
→ log-var 0.25 (does not
cancel) - overdispersion split retained UNCHANGED from round 1, jointly
→ giving 53.5% zero
person-quarters (band 40-70%) and cross-sectional variance 3.49 - and
→ event-workplace x
quarter lognormal shocks AR(1) rho 0.5, log-var 0.0375, realized
→ event-level ICC ~0.014
(band 0.01-0.05). REMEDIATED elements, each an audit-authorized honest
→ lever with no floor
raised: (1) sampling at the FULL derived expectation's LOWER bound
→ (round 1 self-handicapped

```

to half): 135,000 true-closure 1:3 matched sets across 12,600 event
 → workplaces (~10.7 sets/event,
 lognormal size mix; 15 cohorts 2009-2023, Erhvervsdemografi
 → closure-rate + Browning-Heinesen
 register-scale basis) and 45,000 spurious sets across 4,200 events (30%
 → lower bound of the
 30-60% band); (2) 1:3 riskset matching - each set = 1 exposed + 3
 → independent controls weighted
 1/3, control counts simulated exactly as a single Poisson draw of the 3
 → summed conditional
 rates, controls inheriting the set's event cluster (CR1); (3) H2's
 → confirmatory quantity
 re-operationalized from the spurious-arm-dominated differential to
 → theta2a, the true-arm-only
 4-quarter post-year excess at the unchanged 2% floor (round 1's DGP
 → used 2.5% true post-year;
 lowering to the 2.0% floor is conservative), with the spurious arm
 → serving only falsification
 gate F2 (fire iff theta2b's two-sided 95% CI excludes zero positively
 → AND theta2b >= 50%
 of theta2a-hat) and exploratory H2_diff; (4) OOH base rate recalibrated
 → from the self-declared
 over-conservative 3% to $\mu_{OOH} = 0.088 = 8.0\%$ of daytime GP volume =
 → 7.4% of all GP contacts
 (conservative end of the published Danish 8-12% range, Christensen &
 → Olesen 1998; LV-KOS
 2011); (5) the registered serial-correlation discount BUILT INTO the
 → DGP rather than assumed
 away: person transient shocks are now AR(1) rho 0.5 (marginals, zero
 → share, variance, ICC
 identical to round 1), which reduces the averaging gain of every
 → multi-quarter window including
 theta2a's 4-quarter mean and weakly raises every SE vs round 1.
 → Effects multiplicative
 on exposed means: linear ramp in year c with mean(Q3,Q4)=2% (theta1
 → floor; ramp-implied
 Q4-Q1 trend 1.714% for exploratory H1_trend); true post-year 2.0%
 → (theta2a floor); spurious
 residual 0.5% in the hypothesis-true world (retained from round 1, NOT
 → tuned to an idealized
 inert 0%, which is reported as reference); +5-point OOH-vs-daytime
 → relative gap over Q3(c)..Q2(c+1)
 (theta3 floor). Estimators: exact set-level-difference equivalents of
 → the registered feols
 spec (set FE + cohort x calendar-quarter FE reduce to set-level DiD
 → with 1/k-weighted controls),
 CR1 cluster SEs on event workplace; H3 via the plug-in relative-excess
 → contrast with cluster
 SE (delta approximation to the registered 999-draw cluster bootstrap);
 → winsorization omitted
 (conservative). Holm-Bonferroni over the confirmatory family
 → {H1=theta1, H2=theta2a, H3=theta3}
 at FWER 0.05, two-sided p, rejection counted only in the predicted
 → direction; demotion-only
 fixed-point rule at the final family's Holm alphas. nsim=500 primary,
 → 250 per sensitivity
 scenario.'
 lowest_effect_basis: 'The registered floor effects, UNCHANGED from the
 → failed round 1 (no

floor was raised to buy power): H1 = +2% relative pre-boundary excess
 ↪ of the control baseline
 mean (0.022 contacts/person-quarter on a 1.1 base), set below Hamad,
 ↪ Modrek & Cullen (2015)
 layoff-insecurity utilization shifts of ~5-10%; H2 = +2% relative
 ↪ TRUE-ARM post-year excess
 (0.022 contacts/person-quarter; the re-operationalized theta2a - round
 ↪ 1's DGP gave the
 true arm 2.5%, so simulating at the 2.0% floor is a conservative
 ↪ change), set below Hamad/Modrek/Cullen
 (2015) and below Kuhn, Lalive & Zweimueller (2009) post-displacement
 ↪ utilization magnitudes;
 H3 = 5-point gap in relative peri-event excess between OOH and daytime
 ↪ contacts (0.05),
 below published composition shifts - what changed for H3 is only the
 ↪ power basis (OOH base
 rate 8% of daytime GP volume, conservative end of the published Danish
 ↪ 8-12% range), not
 the floor. Exploratory quantities: H1_trend at the ramp-implied 1.714%
 ↪ of baseline (0.019);
 H2_diff at the hypothesis-true +1.5pp differential. These are
 ↪ smallest-effects-of-interest
 per the NHE registered-report convention and the design's own
 ↪ commitments.'

least_powered_test: 'H1 (theta1, pre-boundary excess): power 0.980 under
 ↪ full-family Holm
 - the least-powered CONFIRMATORY member. The fixed-point iteration
 ↪ CONVERGED AT STEP 1 with
 zero demotions: final family = {H1, H2, H3} at its own Holm alphas
 ↪ (H1=0.980, H2=0.998,
 H3=1.000), all >= 0.95. H1 also carries the largest implied binding N*
 ↪ (~118,000 sets vs
 the provisional 100,000 and the 135,000-set plan), making it the member
 ↪ the Stage-2 realized-N
 demotion gate binds first.'

results:

- hypothesis_id: H1
 - test: 'CONFIRMATORY theta1: mean of Q3(c),Q4(c) set-level DiD
 ↪ coefficients vs 4-quarter
 c-1 baseline (1:3 sets, controls weighted 1/3); CR1 cluster z on
 ↪ event workplace; Holm
 within final family {H1,H2,H3} at FWER 0.05, directional'
 assumed_effect: +2% of control baseline mean = 0.022
 ↪ contacts/person-quarter (registered
 floor, unchanged; linear ramp within year c)
 assumed_n: 135,000 1:3 matched sets / 12,600 true-closure event
 ↪ workplaces (full lower-bound
 derived expectation, 15 cohorts 2009-2023)
 power: 0.98
- hypothesis_id: H2
 - test: 'CONFIRMATORY theta2a (re-operationalized core): TRUE-ARM mean
 ↪ post-year (Q1-Q4 of
 c+1) DiD excess vs c-1 baseline, no spurious-arm variance; CR1
 ↪ cluster z; Holm within
 final family, directional; serial-correlation discount active (AR(1)
 ↪ rho 0.5 person shocks)'
 assumed_effect: +2% relative post-year excess = 0.022
 ↪ contacts/person-quarter (registered
 floor, unchanged; round 1's 2.5% lowered to the floor - conservative)
 assumed_n: 135,000 1:3 true matched sets / 12,600 event workplaces

```

power: 0.998
- hypothesis_id: H3
  test: 'CONFIRMATORY theta3: OOH-vs-daytime relative peri-event excess
    ↳ gap (Q3(c)-Q2(c+1)),
      plug-in ratio contrast with event-workplace cluster SE (delta approx.
        ↳ to registered 999-draw
          cluster bootstrap); Holm within final family, directional'
  assumed_effect: '+5-point relative-excess gap (registered floor,
    ↳ unchanged); OOH base recalibrated
      to 0.088 = 8% of daytime GP volume (published Danish 8-12% range,
        ↳ conservative end)
  assumed_n: 135,000 1:3 true matched sets / 12,600 event workplaces
  power: 1
- hypothesis_id: H1_trend
  test: 'EXPLORATORY (gated, never confirmatory): beta_Q4(c) - beta_Q1(c)
    ↳ within-terminal-year
      contrast, computed only if theta1 Holm-rejects in predicted
        ↳ direction; joint P(gate &
          trend significant at alpha=0.05, directional)'
  assumed_effect: 'ramp-implied +1.714% of baseline = 0.019
    ↳ contacts/person-quarter (verified
      structurally underpowered for confirmatory status in round 1: 0.380)'
  assumed_n: 135,000 1:3 true matched sets / 12,600 event workplaces
  power: 0.772
- hypothesis_id: H2_diff
  test: 'EXPLORATORY: theta2a - theta2b true-minus-spurious post-year
    ↳ differential (round
      1''s infeasible confirmatory construct), unadjusted directional test
        ↳ at alpha=0.05, reported
          with CI under the exploratory flagging convention'
  assumed_effect: '+1.5pp differential in the hypothesis-true world (true
    ↳ +2.0% vs spurious
      residual +0.5%)
  assumed_n: 135,000 true sets / 12,600 events vs 45,000 spurious sets /
    ↳ 4,200 events
  power: 0.46
- hypothesis_id: F2_pass
  test: 'F2 GATE OPERATING CHARACTERISTIC (severity device OUTSIDE the
    ↳ Holm family, consumes
      no alpha, not power-bearing): P(gate does NOT fire) under the
        ↳ hypothesis-true DGP; fire
          = theta2b two-sided 95% CI excludes zero positively AND theta2b >=
            ↳ 50% of theta2a-hat;
              registered target >= 0.90'
  assumed_effect: 'spurious residual +0.5% of baseline (retained unchanged
    ↳ from round 1, not
      tuned to inert 0%; strictly-inert reference pass = 0.978)
  assumed_n: 45,000 spurious 1:3 sets / 4,200 spurious-closure event
    ↳ workplaces
  power: 0.886
- hypothesis_id: F2_trigger
  test: 'F2 GATE OPERATING CHARACTERISTIC: P(gate fires) in the
    ↳ placebo-not-inert world; registered
      target >= 0.95 when theta2b = theta2a at floor'
  assumed_effect: 'theta2b = +2.0% of baseline = theta2a at floor
  assumed_n: 45,000 spurious 1:3 sets / 4,200 spurious-closure event
    ↳ workplaces
  power: 0.79
binding_n_star:

```

rule: 'Binding pre-outcome demotion thresholds fixed by the deposited
 ↳ Stage-0 re-run (demotion-only:
 where the re-run implies a LOOSER threshold than the provisional
 ↳ value, the stricter provisional
 value stands)'
 H1: 118,000 matched sets (re-run value; supersedes the provisional
 ↳ 100,000 - tightened)
 H2: 95,000 true matched sets (re-run value; supersedes the provisional
 ↳ 60,000 - tightened)
 plus $\geq 15,000$ spurious sets for gate F2's registered operating
 ↳ characteristics
 H3: 90,000 matched sets (provisional value retained as the stricter
 ↳ bound over the re-run's
 66,000)
 note: 'The 135,000-set sampling plan covers all three, but H1's margin
 ↳ is ~13%: realized
 N below ~87% of the lower-bound expectation demotes H1 first. The
 ↳ Stage-2 blind re-simulation
 (realized set counts, realized k distribution, baseline-year moments,
 ↳ realized OOH share)
 is the binding verification; any member short of 0.95 power at its
 ↳ floor demotes to exploratory
 before any post-baseline outcome is accessed.'
 sensitivity notes: 'FIXED-POINT FAMILY: converged at step 1 with ZERO
 ↳ demotions - final confirmatory
 family {H1, H2, H3} at its own Holm alphas: H1=0.980, H2=0.998,
 ↳ H3=1.000 (nsim=500; MC SE
 ~0.006-0.014). Estimator recovery unbiased (mean theta-hats
 ↳ 0.0220/0.0219/0.0501 vs targets
 0.022/0.022/0.050). The remediation's own scaling arithmetic is
 ↳ confirmed: H1's mean $z=4.32$
 (design predicted 4.4-4.6 before the serial-correlation discount), H2a
 ↳ $z=4.84$, H3 $z=5.79$,
 vs 4.039 needed for 0.95 at Holm worst-case $\alpha/3$. BINDING N^* (fixes
 ↳ the Stage-2 demotion
 thresholds, z -squared scaling of mean SEs): $N^*_{H1} \sim 118,000$ sets -
 ↳ ABOVE the provisional
 100,000, so the provisional gate must tighten to ~118k; $N^*_{H2a} \sim 95,000$
 ↳ (above provisional
 60,000); $N^*_{H3} \sim 66,000$ (below provisional 90,000; the provisional
 ↳ value may stand as the
 stricter bound). The 135,000-set plan covers all three, but H1's
 ↳ margin is only ~13% -
 realized N below ~87% of the lower-bound expectation demotes H1 first.
 ↳ SENSITIVITY N_{50} (67,500
 sets - the neighborhood of the provisional gates): full-family Holm
 ↳ powers H1=0.840, H2=0.904,
 H3=0.980 - H1 and H2 fall below 0.95, directly validating that the
 ↳ pre-outcome realized-N
 demotion gate is load-bearing, not decorative. SENSITIVITY EFF50
 ↳ (floors halved): H1=0.512,
 H2=0.596, H3=0.800 - the design still has essentially no confirmatory
 ↳ resolution below its
 floors; nulls must be reported as CI-ruled-out magnitudes against
 ↳ realized MDEs. CAVEATS
 DRIVING THE VERDICT (F2 gate, outside the Holm family, not
 ↳ power-bearing, but with REGISTERED
 operating characteristics the design claims are Stage-0-verified): (1)
 ↳ trigger probability

when theta2b = theta2a at floor = 0.790 at 45,000 spurious sets,
 → clearly BELOW the registered
 >= 0.95 target (MC SE ~0.018); z-scaling implies ~0.95 trigger
 → sensitivity needs ~75-85k
 spurious sets - within the 30-60% band's upper half (up to ~81k sets)
 → but above the lower-bound
 sizing, and far above the 15,000-set F2 viability gate (at 15k sets
 → trigger ~0.5, bounds-only
 reporting per the registered fallback). The gate remains VALID as a
 → severity device - it
 can only subtract support, never add it, and false-fire risk is low -
 → but its sensitivity
 to a floor-sized non-inert placebo is materially below the registered
 → 0.95; the registration
 should either restate the trigger target at its verified value (~0.79
 → at lower-bound N)
 or add a Stage-2 recomputation of F2 characteristics at realized
 → spurious N. (2) Pass probability
 under the hypothesis-true DGP (spurious residual 0.5% retained from
 → round 1) = 0.886, marginally
 below the registered >= 0.90 (within ~1 MC SE, 0.014); under a strictly
 → inert placebo (0%)
 pass = 0.978, comfortably above - the miss reflects the honest 0.5%
 → residual convention,
 not gate malfunction; at N50 pass improves to 0.936 (lower spurious
 → precision fires less).
 (3) The serial-correlation discount (AR(1) rho 0.5 person transient
 → shocks, marginals unchanged)
 raised all SEs vs round 1's iid convention; results are therefore
 → conservative relative
 to the round-1 variance budget - H1 at 0.980 already absorbs the
 → discount. (4) Exploratory
 tiers confirm the remediation's tiering decisions: H1_trend joint
 → gated power 0.772 and
 H2_diff 0.460 - both would fail as confirmatory members, exactly as
 → diagnosed in round 1
 (0.380/0.358); their exploratory demotion is validated, not arbitrary.
 → VERDICT RATIONALE:
 all three confirmatory members clear 0.95 with the family intact
 → (adequately powered on
 the task's own criterion), but the F2 falsification gate misses its
 → two registered operating-characteristic
 targets at lower-bound spurious N, so the design is
 → powered_with_caveats: registrable, provided
 the F2 characteristic targets are restated to the verified values (or
 → F2 characteristics
 are re-verified at realized spurious N in the Stage-2 blind
 → re-simulation) and the binding
 N*_H1 is recorded as ~118,000 sets rather than the provisional 100,000.'
 verdict: powered_with_caveats
 verification: 'Deposited script re-run verified by the harness (exit 0,
 → seed 20260831, POWER_JSON
 matched exactly: H1 0.980, H2 0.998, H3 1.000, H1_trend 0.772, H2_diff
 → 0.460, F2_pass 0.886,
 F2_trigger 0.790; calibration in band: zero-share 0.535, OOH 8.0% of
 → daytime volume, ICC
 0.014). Independent driver reproduction on 2026-08-31: 3 differing
 → stdout lines of 45, all
 wall-clock stamps; every numeric result identical. Driver
 → reconciliation of the powered_with_caveats

```

verdict, both actions carried in this protocol: (1) F2 operating
↳ characteristics restated
from aspirational targets to the VERIFIED values with Stage-2
↳ re-verification at realized
spurious N registered as binding; (2) N*_H1 recorded as 118,000 sets,
↳ superseding the provisional
100,000. Frozen bank hypothesis wording untouched.'
```

Current authoritative contingencies

```yaml
- condition: ENDNEDL flag-year convention differs in-environment from the
↳ assumed 'flagged
  in the workplace's last active year'
  rule: Re-anchor mechanically so that c always denotes the last IDAS
↳ year with ANTNOV >=
  5 for the identity; the estimand is unchanged; the recalibration is
↳ reported verbatim
- condition: ARB_SEKTORKODE is absent from the analysis-server IDAN
↳ extract (publicly documented
  on TIMES ida-ansaettelser but unconfirmed in the recorded
↳ Forskningservice inventory;
  presence and 2008+ coverage validated at Stage 1 before any
↳ classification work)
  rule: 'Frozen fallback: define the private-sector filter from
↳ ARB_HOVED_BRA_DB07 instead
  - a workplace counts as private if and only if its NACE DB07 section
↳ is not 0 (public
  administration and defence), P (education), or Q(human health and
↳ social work); the substitution
  and its effect on eligible-workplace and worker counts are reported
↳ verbatim; no other
  filter change is permitted'
- condition: Match rate < 95% of eligible exposed workers at target k = 3
↳ under full exact
  strata
  rule: 'Apply the pre-specified coarsening ladder in order, stopping at
↳ the first rung reaching
  95% of exposed matched with k >= 1: (1) NACE section -> 10 aggregate
↳ industry groups;
  (2) drop region stratum; (3) age bands 5-year -> 10-year. Never drop
↳ sex, cohort, size
  class, or the baseline-billing calipers; realized k distribution
↳ reported and fed to the
  Stage-2 blind re-simulation'
- condition: Median HONUGE settlement lag (BEHANDLINGSDATO calibration,
↳ 2021+ rows) > 4 weeks
  rule: 'Guard band: exclude Q1(c+1) from the theta2a/theta2b windows
↳ (means over Q2(c+1)..Q4(c+1))
  and report theta1 both as registered and with a 4-week-shifted
↳ boundary; the registered
  theta1 window remains primary'
- condition: KONTAKT does not behave as a row-level contact marker on
↳ in-environment validation
  rule: 'Switch to the frozen fallback contact definition: count of
↳ distinct PNR x HONUGE-week
  x YDERNR x GP-code-group cells with >= 1 positive-fee row; switch and
↳ level effect reported;

```

all estimands re-run under both definitions with the fallback as
↳ primary'

- condition: OLS and Poisson QML disagree in sign on any primary quantity
rule: Foreground the pre-committed OLS estimate, report both, flag the
↳ discordance in the
Summary paragraph; no post hoc model selection
- condition: 'c-2 placebo panel FAILURE, defined as: placebo contrast
↳ significant at alpha
= 0.05 AND |estimate| > 1% of control baseline mean (50% of the
↳ theta1 floor effect) -
the single registered failure condition, stated identically in
↳ outcome_neutral_checks
and the H1 placebo kill'
rule: 'Execute the single pre-specified re-match adding year c-2
↳ baseline contacts to the
Mahalanobis set; if the placebo fails again under the same condition,
↳ the entire registered
event-study tier (confirmatory and exploratory event-study
↳ quantities) is killed and the
study reports descriptive trajectories only. The two
↳ non-complementary cells are pre-committed
passes with mandatory disclosure, never discretion: (a) significant
↳ at alpha = 0.05 with
|estimate| <= 1% - no re-match, no kill; the Table T7 placebo row
↳ flags the significance
and states that a sub-floor-scale imbalance was detected; (b) not
↳ significant with |estimate|
> 1% - no re-match, no kill; a mandatory caution paragraph reports
↳ the placebo magnitude
against the theta1 floor and the realized placebo MDE, and the H1/H2
↳ Summary paragraphs
must cite it'
- condition: A matched control's own workplace acquires a terminal event
↳ within c-2..c+1 after
matching (data revision)
rule: Drop that control from the set; drop the whole set only if no
↳ control remains; dropped
counts reported; no re-matching
- condition: YDLTID missing or code 0 ('not relevant') on a GP row
rule: Classify as daytime (conservative against the theta3 prediction);
↳ missingness share
reported; if > 20% of peri-event GP rows, H3's kill criterion triggers
- condition: Exact tie in Mahalanobis distance for a nearest-neighbor
↳ match
rule: Deterministic tie-break by smallest absolute baseline-contact
↳ difference, then seeded
RNG (seed 20260831)
- condition: IDAS-flag vs worker-flow classification agreement < 90% but
↳ >= 75%
rule: Proceed with the worker-flow classification as registered (the
↳ canon, S7); publish
the full disagreement cross-tabulation; below 75%, H2's kill
↳ criterion governs
- condition: 'GRAFT 3: sign of theta2b or theta2_diff not stable across
↳ the three successor-share
threshold pairs'
rule: Flag classification fragility in the pre-committed H2 Summary
↳ paragraph; the 30/70
estimate remains primary; no threshold shopping

- condition: 'GRAFT 6: exposed-vs-control difference in later-presence
 - ↪ rates (any IDAN record in year c+1 or any SSSY row in c+1 or later) exceeds 2 percentage points'
 - rule: A mandatory bounding paragraph accompanies all post-boundary
 - ↪ (theta2a/theta2b) conclusions,
 - ↪ converting the differential into best/worst-case bounds; trigger and threshold registered,
 - ↪ not discretionary
- condition: 'GRAFT 9: the control forward-window sensitivity yields
 - ↪ SMALLER magnitudes than the primary estimates'
 - rule: Flag as diagnostic of something other than contamination
 - ↪ (registered direction note);
 - ↪ reported verbatim, no re-specification
- condition: 'STAGE-2 REALIZED-N GATE: realized matched sets (or spurious
 - ↪ sets for F2) fall below a member's binding N* from the Stage-0 re-run, or the Stage-2
 - ↪ blind re-simulation
 - ↪ (realized Ns, realized k distribution, baseline-year moments only)
 - ↪ shows power < 0.95
 - ↪ for a member at its floor'
 - rule: That member is demoted to exploratory BEFORE any post-baseline
 - ↪ outcome is accessed;
 - ↪ the Holm family shrinks and alphas recompute by the registered
 - ↪ fixed-point iteration among survivors; demotion-only - no member, threshold, or floor may move in
 - ↪ the other direction;
 - ↪ the realized-power/MDE table is published either way
- condition: Realized control-group OOH share of GP contacts < 5% (below
 - ↪ the published 8-12% range)
 - rule: H3's power basis is re-verified in the Stage-2 blind
 - ↪ re-simulation at the realized share; if < 0.95 at the 5-point floor, H3 demotes to exploratory per
 - ↪ the realized-N gate;
 - ↪ if realized share < 2%, H3's kill criterion governs instead
- condition: Spurious arm killed at Stage 1/2 (events < 150, sets <
 - ↪ 3,000, classification agreement < 75%, or bimodality unresolved)
 - rule: Gate F2 and theta2_diff/theta2b are reported unavailable (drop
 - ↪ disclosed before outcome access); theta2a remains in the Holm family if its own N* is met, but
 - ↪ the H2 verdict is capped at 'partially consistent - comparative quantifier untestable'
 - ↪ and the Discussion states that the excess cannot then be distinguished from selection
 - ↪ into disappearing workplace identities

2. ENVIRONMENT

Software stack (record exact versions in the run manifest at Stage 0 and freeze them):

- ****python3 (>= 3.11) with polars**** for ALL large-table preprocessing. Lazy scans only (`scan\_parquet`/`scan\_csv`) with projection and predicate pushdown; semi-joins against the cohort PNR roster before any collect; streaming collects for aggregations. Never materialize a person x week or person x month panel beyond need: the only dense

person-time table permitted is matched members x 16 event-quarters (~540k true members + ~180k spurious members, x 16 = under 12M rows). Zero-filling happens at the quarter level for panel members only, never by cross-joining the population against a calendar.

- ****R (>= 4.3) with fixest, invoked via Rscript**, for every fixed-effects estimation: `feols` (OLS, the pre-committed primary) and `fepois` (Poisson QML robustness). CR1 cluster-robust SEs on event workplace. No estimation in python; no data manipulation in R beyond reading the analysis parquet and writing tidy coefficient tables.**
- ****matplotlib** for all figures; **booktabs LaTeX** for all tables (no vertical rules). Every number in every exhibit is written by script from a parquet/CSV output - never hand-edited.**
- ****Memory discipline:** process cohort-by-cohort (c = 2009..2023) where possible; checkpoint every numbered pipeline step as parquet under `interim/` with a fixed name (given per step below); SSSY scans filter on (PNR in roster) AND (SPEC2 in frozen lists) at scan level; the full population is touched only by the aggregate coverage benchmark (step 16), which is computed as a streaming groupby per year, never collected at row level.**
- ****Seeded randomness:** a single root seed 20260831 (per the frozen registration) initializes every RNG: matching processing order, Mahalanobis tie-breaks, the 999-draw cluster bootstrap, and the Stage-2 blind re-simulation. No unseeded randomness anywhere.**
- ****Run manifest:** an append-only JSONL (`manifest/run\_manifest.jsonl`), one record per pipeline step per run: step id, UTC timestamp, code version hash, input file(s) + row counts, output file + row count, parameters used, validation check results, PASS/FAIL. Every step below MUST log its row counts; a step whose check fails writes FAIL and halts downstream steps (Section 6.3). The manifest (aggregate counts only, event counts rounded per the disclosure rules) is part of the released audit trail.**
- ****Directory convention:** `raw/` (DST-delivered), `interim/` (parquet checkpoints), `analysis/` (panels), `out/tables/`, `out/figures/`, `manifest/`. All paths in code are relative to the project root inside the research environment; no absolute paths in any released script.**
- ****Frozen-code discipline:** the Stage-4 estimation code is frozen at the end of Stage 0 (hash recorded in the manifest); Stage 4 is a single run of that code. The sensitivity suite runs once and is reported together. Any code change after Stage 0 is a deviation and must be disclosed in running prose with reason and approver.**

Source disclosure rules

```

```yaml
environment: DST Forskningservice (no prior data access; gatekeeper
 ↳ certification per RR
 secondary-data rules). All analysis occurs inside the environment; only
 ↳ aggregate outputs
 leave. No microdata, workplace identifiers (ARBNR/CVRNR/LBNR), or
 ↳ person-level values are
 released.
minimum_cell_size: 5
minimum_cell_rules: no released statistic based on fewer than 5 persons
 ↳ or fewer than 3 workplaces;
 event-count tables rounded to the nearest 10; weekly trajectory figures
 ↳ pool >= 100 events

```

```

and >= 5,000 person-weeks per plotted point; fee (BRUHON) aggregates
 ↳ subject to a 50% single-workplace
dominance check; min-cell suppression applied BEFORE Holm reporting so
 ↳ suppression can never
 be outcome-selective
outputs_leaving_environment:
- model coefficients with SEs/CIs
- matching balance tables and realized-k distribution
- calibration and benchmark tables (coverage, linkage, HONUGE lag, OOH
 ↳ share)
- figure aggregates (weekly trajectory series under the pooling rules)
- event counts rounded to the nearest 10
- Stage-0 power/MDE tables - BOTH rounds (failed and remediated) - and
 ↳ the Stage-2 blind re-simulation
 output with the demotion trail
- the c-2 placebo, gate F2 outputs, GRAFT-2/4 diagnostics, GRAFT-6
 ↳ later-presence rates, E4
 attrition panel, and all outcome-neutral check outputs (the full audit
 ↳ trail exits under
 the same rules)
- the analysis and power scripts
statement: 'All analysis occurs inside DST Forskningservice; only
 ↳ aggregate outputs leave.
 Rules applied to every released object: (1) no statistic based on fewer
 ↳ than 5 persons or
 fewer than 3 workplaces; (2) event-count tables rounded to the nearest
 ↳ 10; (3) weekly trajectory
 figures pool >= 100 events and >= 5,000 person-weeks per plotted point;
 ↳ (4) fee (BRUHON)
 statistics released only as cell means/coefficients over min-N cells,
 ↳ with a dominance check
 that no single workplace contributes > 50% of any released fee
 ↳ aggregate; (5) released objects
 are limited to: model coefficients with SEs/CIs, balance tables,
 ↳ calibration/benchmark tables,
 figure aggregates, event counts (rounded), the Stage-0 (both rounds)
 ↳ and Stage-2 power/MDE
 tables with the demotion trail, and the analysis/power scripts - never
 ↳ microdata, workplace
 identifiers (ARBNR/CVRNR/LBNR), or person-level values; (6) min-cell
 ↳ suppression is applied
 before Holm reporting so that suppression can never be
 ↳ outcome-selective; (7) the c-2 placebo,
 gate F2 outputs, the GRAFT-2/4 diagnostics, the GRAFT-6 later-presence
 ↳ rates, the E4 attrition
 panel, and all outcome-neutral check outputs are released under the
 ↳ same rules, ensuring
 the full audit trail exits the environment.'
...

```

## Pending registered experiment

No Danish study outcome has been observed for this bundle. See the complete registration above.

## atom-000003: Billing measurement validity

### Do billing coverage, timing calibration and the registered placebo permit event-study interpretation?

#### Source fidelity and implementation order

This is a format conversion of DS1.1, not a new study or independent acceptance. Source audit: minor\_revisions. No observed Danish outcomes are supplied. The source execution plan explicitly states that final prereg.yaml governs disagreements. Exact current YAML selections below therefore govern older execution-plan excerpts; these excerpts are retained as implementation detail, not a competing registration. See docs/provenance/source-discrepancies.md. The complete shared data/validity contracts are embedded in Atoms 1–3; read those native protocols when executing a downstream Atom, even without docs/. Do not substitute a scientific default for an unresolved source contradiction.

Preserve source stages 0 -> 1 -> 2 -> 3 -> 4 -> 5 inside the implementation. Source Stage-0 simulations are historical design evidence, not study outcomes; do not silently redo or improve them. Gate 14 must be evaluated from baseline data before post-baseline outcomes. All surviving Stage-4 quantities form ONE Holm family. The three question-level Atoms may share the same frozen family fit and tidy results; do not apply Holm separately per Atom or adjudicate a member before the complete surviving family's p-values exist. Cache/reuse identical governed intermediate results with provenance when available; a direct-dispatch call must obtain or compute its exact source-defined inputs, not pretend its predecessors succeeded. Stage-5 internal riders remain exploratory, after the Stage-4 family is frozen, even when housed in H1/H2.

Reference files under docs/execution/ are the ORIGINAL SKELETONS, not tested turnkey implementations. Adapt paths and complete engineering in a governed workspace; preserve source scientific rules and disclose unresolved choices. Write a result record with component estimates, uncertainty, family eligibility, source interpretation row, diagnostics and boundary/skip dispositions. Released outputs follow source disclosure rules; confidential rows stay governed. Local native states summarize the source judgment; neither a state name nor source power forecasts prove equivalence or causal identification. A bounded claim requires its actual interval and registered rule. Retain unresolved source wording conflicts; do not convert mere nonsignificance into scientific null.

#### Frozen scope and vocabulary

```
claim_class_ceiling: associational
statement: 'Associational - matched event-study contrasts, a contrastive
↪ falsification gate,
 and a composition contrast; descriptive for trajectory panels and the
 ↪ bounded priority claim.
Binding vocabulary rules: 'is associated with' / 'consistent with' /
↪ 'suggests' only;
```

```

 'effect', 'impact', 'causes' prohibited throughout; anticipation
 ↪ always worded as
 'pre-boundary excess' relative to the annually dated closure-year
 ↪ boundary, never as announcement-dated
 anticipation (S8); no phenomenon-discovery or Danish-first framing (S1-S3);
 ↪ H2 novelty confined
 to the placebo's application to health billing (S7). H2 verdict semantics
 ↪ are fixed in
 advance: 'consistent with H2' requires BOTH the confirmatory core
 ↪ (theta2a Holm-significant,
 predicted direction) AND a passing falsification gate F2; the strict
 ↪ comparative magnitude
 is only ever reported as an exploratory estimate with its CI. Exploratory
 ↪ results (H1_trend,
 H2_diff, E1, E3, E4, and any gate-demoted member) carry the word
 ↪ 'exploratory' in the
 heading and in every sentence stating them. The prose label never exceeds
 ↪ the evidence label.'
frozen_parameters:
 seed: 20260831
 gp_code_list: daytime GP = SPEC2 in {70..80}
 ooh_code_list: out-of-hours GP = SPEC2 in {81, 82, 83, 84, 89}
 ooh_robustness_classifier: adds YDLTID in {2, 3, 4, 5, 7, 8, 9}
 spec2_evening_gp_disclosure: 'Registered choice, disclosed: the parsed
 ↪ value labels mark
 SPEC2 71 ("General practitioner, consultation, evening etc."), 73
 ↪ ("General practitioner,
 telephone consultation, evening etc."), and 75 ("General practitioner,
 ↪ home visit, evening
 etc.") as explicitly off-schedule GP services; they nonetheless remain on
 ↪ the daytime
 side of the PRIMARY classifier because H3's out-of-hours construct and
 ↪ its published
 power basis are keyed to the laegevagt out-of-hours service (SPEC2 81-84,
 ↪ 89), the quantity
 the published Danish 8-12% share describes. Attenuation direction, both
 ↪ readings: counting
 evening GP as daytime attenuates theta3 toward zero - conservative
 ↪ against the positive
 H3 conclusion, anti-conservative for the null-bounding "compositionally
 ↪ unremarkable"
 reading, whose Stage-2 statement therefore carries a mandatory caveat.
 ↪ In-environment
 validation must verify that the YDLTID-augmented robustness classifier
 ↪ captures SPEC2
 71/73/75 rows (their evening time codes 2-5); if YDLTID proves
 ↪ unpopulated on those rows,
 the robustness classifier - and only the robustness classifier - is
 ↪ amended mechanically
 and with disclosure to include SPEC2 71/73/75 directly. The SPEC2
 ↪ 71/73/75 volume share
 is reported, and both classifiers appear side by side in Table T7.'
 successor_share_thresholds: 'true closure: ENDNEDL final closure AND
 ↪ worker-flow max-successor
 share s(w) < 0.30; spurious closure: forward identity via
 ↪ IDFREEM/REGELFR/FRELFREM AND
 s(w) >= 0.70; ambiguous mid-zone excluded; GRAFT-3 sensitivities
 ↪ 0.20/0.80 and 0.40/0.60'
 quarter_mapping: 'ISO week of HONUGE: 1-13 / 14-26 / 27-39 / 40-52; ISO
 ↪ week 53 to Q4'

```

```

winsorization: person-quarter outcomes at the 99.9th percentile within era
↳ x calendar quarter
matching: 1:3 nearest-neighbor Mahalanobis within exact strata, without
↳ replacement across
 sets, target k = 3, minimum k = 1; calipers |dif baseline contacts| <= 3
 ↳ and |dif log(1+BRUHON)|
 <= 0.5; exposed processed in seeded random order (seed 20260831)
software: R >= 4.3, fixest::feols (OLS) and fixest::fepois (Poisson QML)
↳ robustness); versions
 frozen in the registered script
note: Code lists and thresholds frozen before any data access; code labels
↳ quoted verbatim
 from parsed DST value labels and re-validated in-environment before
 ↳ estimation (frozen
 fallback rules in contingent_decisions).

```

15. **SSSY outcome extraction, 2007–2025, matched members only (Stage 3 begins).** Input: raw SSSY, roster = matched members. Lazy scan; person x HONUGE-week aggregation of KONTAKT (contacts), YDLANT (services), BRUHON (fees), separately for daytime GP (SPEC2 70–80) and OOH (SPEC2 81–84, 89), plus the YDLTID-augmented robustness OOH classification (YDLTID missing or 0 -> daytime, share logged). Validate KONTAKT as a row-level contact marker (distribution against YDLANT and row multiplicity); if validation fails, switch to the frozen fallback — count of distinct PNR x HONUGE-week x YDERNR x GP-code-group cells with  $\geq 1$  positive-fee row — report the switch and level effect, and re-run all estimands under both definitions with the fallback as primary. Output: `interim/sssy_weekly.parquet`. Check: value labels for SPEC2/YDLTID re-validated verbatim against in-environment documentation before anything downstream runs; YDLTID missing/“not relevant” share on peri-event GP rows logged ( $> 20\%$  with SPEC2 OOH codes absent triggers H3’s kill).
16. **Coverage benchmark (full extract, aggregate only).** Input: raw SSSY (streaming groupby per year; no row-level collect). Annual GP contacts per 1,000 population (SPEC2 70–80), 2008–2024, against the published DST StatBank sygesikring series. Output: `interim/coverage_benchmark.parquet`. Check: within  $\pm 10\%$  of the published series in every year; any year outside triggers documented reconciliation before estimation. Also compute the linkage floor: share of cohort workers (exposed + control pool) with  $\geq 1$  SSSY row of any kind in `c-3..c+1`;  $\geq 80\%$  required (halt at Stage 3 below).
17. **Person-quarter aggregation and winsorization.** Input: step 15. Map HONUGE ISO weeks to quarters (1–13/14–26/27–39/40–52; week 53 -> Q4); aggregate to person x event-quarter; compute 99.9th-percentile winsorization thresholds within era x calendar quarter, freeze them (log the thresholds), and winsorize contacts and BRUHON sums. Output: `interim/person_quarter.parquet`. Checks (floor/ceiling and parsing sanity): control share of zero person-quarters in [30%, 85%]; 99th percentile  $\leq 30$  contacts/quarter; control weekly series shows the summer trough (ISO weeks 28–31 mean  $\geq 10\%$  below the annual weekly mean); realized control OOH share of baseline GP contacts in [5%, 15%] (in [2%,

- 5%) the step-14 re-verification governs; < 2% H3 kill; > 15% code-list re-audit before estimation).
18. **Panel assembly.** Input: steps 13, 17. Sixteen event-quarters per member, Q1(c-2)..Q4(c+1), zero-filled where no billing rows exist (zero utilization by construction, never missingness); estimation panel = Q1(c-1)..Q4(c+1); placebo panel = Q1(c-2)..Q4(c-2) plus the c-1 baseline. Outputs: `analysis/analysis_panel.parquet`, `analysis/placebo_panel.parquet` (true arm); pooled panel with event-class indicators for the spurious arm. Check: rows = members x 16 exactly; weights present (exposed 1, controls 1/k\_s); event workplace cluster id attached to every member.
  19. **HONUGE settlement-lag calibration.** Input: SSSY 2021–2024 GP rows (BEHANDLINGSDATO vs HONUGE). Output: `interim/lag_calibration.parquet`
    - released table. Check:  $\geq 75\%$  of rows billed with HONUGE within 4 weeks of the treatment week. If median lag > 4 weeks: guard-band contingency — exclude Q1(c+1) from the theta2a/theta2b windows (means over Q2(c+1)..Q4(c+1)) and report theta1 both as registered and with a 4-week-shifted boundary (registered window remains primary). This table is reported BEFORE any event-study estimation.
  20. **GRAFT-6 later-presence accounting.** Input: IDAN c+1, SSSY c+1 onward, matched members. Exposed-vs-control rates of any IDAN record in c+1 or any SSSY row in c+1 or later. Output: `interim/later_presence.parquet`. Check: if the differential exceeds 2 percentage points, the mandatory bounding paragraph (best/worst case) accompanies every post-boundary conclusion — registered trigger, not discretionary.
  21. **c-2 placebo estimation (outcome-neutral; final Stage-3 gate).** Input: `analysis/placebo_panel.parquet`. The step-4.1 estimator run on year c-2 quarters against the c-1 baseline. Output: `interim/placebo_result.parquet`. Check: NOT significant at  $\alpha = 0.05$  AND  $|\text{point estimate}| \leq 1\%$  of the control baseline mean (50% of theta1's floor). On failure: execute the SINGLE pre-specified re-match adding year c-2 baseline contacts to the Mahalanobis set (back to step 13 once, flagged in the manifest); if the placebo still fails, the entire registered event-study tier (confirmatory and exploratory event-study quantities) is killed and the study reports descriptive trajectories only.

## Outcome-neutral checks

- check: 'Coverage benchmark: annual GP contacts per 1,000 population in the
  - ↪ `extract (SPEC2`
  - `70-80), 2008-2024, compared against published DST health-insurance`
  - ↪ `statistics (StatBank`
  - `sygesikring series)'`
 threshold: Within  $\pm 10\%$  of the published series in every year; any year
  - ↪ outside triggers
  - documented reconciliation before estimation
- check: 'Linkage/observability floor: share of cohort workers (exposed +
  - ↪ control pool) with
  - `>= 1 SSSY row of any kind in the 5-year window c-3..c+1'`

```

threshold: '>= 80% (Danish annual GP contact prevalence benchmark ~85%);
↳ below floor halts
 at Stage 3'
- check: 'Closure-rate plausibility: annual true-closure event rate among
↳ private size-5+
 workplaces, and spurious/true event ratio'
threshold: Rate in [1%, 8%] every cohort year (consistent with the
↳ published Erhvervsdemografi
 basis used for sizing); ratio in [0.1, 1.5]; violations trigger
 ↳ classification audit before
 any outcome data access
- check: 'Classification cross-validation: agreement between worker-flow
↳ classes (s(w) thresholds)
 and IDAS-native flags (ENDNEDL/IDFREM/REGELFR/FRELFREM/ANDELUDS)'
threshold: '>= 75% agreement over classified events (kill floor for the H2
↳ comparative machinery);
 full cross-tabulation released'
- check: 'GRAFT 4 - en-bloc successor-share bimodality (Stage 1):
↳ distribution of worker-flow
 successor shares s(w) across terminal events'
threshold: Visibly bimodal with the ambiguous mid-zone ($0.30 \leq s(w) < 0.70$) containing
↳ < 35% of terminal events, per the Hethy-Maier & Schmieder classification
 ↳ literature;
 violation triggers the classification audit and, if unresolved, the
 ↳ spurious-arm kill
- check: Matching balance and match rate
threshold: Standardized mean difference ≤ 0.10 on every matching covariate
↳ and on unmatched
 auxiliaries after matching (control aggregates weighted $1/k$ within set);
 ↳ match rate $\geq$
 90% of eligible exposed ($k \geq 1$) after the coarsening ladder; realized
 ↳ mean k reported
- check: 'Pre-baseline placebo (design validity, orthogonal to hypotheses):
↳ exposed-minus-control
 contact difference across the four quarters of year c-2'
threshold: 'FAILS if and only if the placebo contrast is significant at
↳ $\alpha = 0.05$ AND
 $|\text{estimate}| > 1\%$ of control baseline mean (= 50% of the θ_1 floor
 ↳ effect) - the single
 registered condition, stated identically in contingent_decisions and the
 ↳ H1 placebo kill;
 failure triggers the single pre-specified re-match, then kill of the
 ↳ registered event-study
 tier. The two non-complementary cells pass with mandatory disclosure and
 ↳ no discretion:
 significant with $|\text{estimate}| \leq 1\%$ passes, with the significance flagged
 ↳ in the Table T7
 placebo row; non-significant with $|\text{estimate}| > 1\%$ passes, with a
 ↳ mandatory caution paragraph
 quoting the magnitude against the θ_1 floor and the realized placebo
 ↳ MDE'
- check: HONUGE settlement-lag calibration on 2021-2024 GP rows via
↳ BEHANDLINGSDATO
 threshold: '>= 75% of rows billed with HONUGE within 4 weeks of the
 ↳ treatment week; median
 lag > 4 weeks triggers the guard-band contingency'
- check: 'No floor/ceiling in outcomes: distribution of control-group
↳ person-quarter GP contacts'

```

```

threshold: Share of zero person-quarters in [30%, 85%]; 99th percentile <=
↳ 30 contacts/quarter;
 control weekly series shows the expected summer trough (ISO weeks 28-31
 ↳ mean >= 10% below
 annual weekly mean - HONUKE parsing sanity)
- check: 'GRAFT 2 - week-0 boundary-step diagnostic on the weekly descriptive
↳ series: discontinuity
 in the exposed-minus-control series at the first HONUKE week of year c+1,
 ↳ relative to
 the total pre-boundary (Q3(c)+Q4(c)) excess'
threshold: 'Step <= 50% of the total pre-boundary excess; if exceeded: (a)
↳ quarter-level
 estimates foregrounded as the only interpretable resolution, (b) a
 ↳ caution paragraph converts
 the step size into an explicit bound on any timing-shape claim'
- check: 'Published-range OOH share benchmark (new, licenses H3''s remediated
↳ power basis):
 realized control-group OOH share of baseline-year GP contacts vs the
 ↳ published Danish
 8-12% range'
threshold: 'In [5%, 15%]: proceed as registered. In [2%, 5%]: H3''s
↳ realized-power re-verification
 governs (contingent decision). Below 2%: H3 kill criterion. Above 15%:
 ↳ code-list re-audit
 before estimation (possible OOH misclassification)'

```

## Current authoritative contingencies

```

- condition: ENDNEDL flag-year convention differs in-environment from the
↳ assumed 'flagged
 in the workplace's last active year'
rule: Re-anchor mechanically so that c always denotes the last IDAS year
↳ with ANTNOV >=
 5 for the identity; the estimand is unchanged; the recalibration is
 ↳ reported verbatim
- condition: ARB_SEKTORKODE is absent from the analysis-server IDAN extract
↳ (publicly documented
 on TIMES ida-ansaettelser but unconfirmed in the recorded
 ↳ Forskningservice inventory;
 presence and 2008+ coverage validated at Stage 1 before any
 ↳ classification work)
rule: 'Frozen fallback: define the private-sector filter from
↳ ARB_HOVED_BRA_DB07 instead
 - a workplace counts as private if and only if its NACE DB07 section is
 ↳ not 0 (public
 administration and defence), P (education), or Q (human health and social
 ↳ work); the substitution
 and its effect on eligible-workplace and worker counts are reported
 ↳ verbatim; no other
 filter change is permitted'
- condition: Match rate < 95% of eligible exposed workers at target k = 3
↳ under full exact
 strata
rule: 'Apply the pre-specified coarsening ladder in order, stopping at the
↳ first rung reaching
 95% of exposed matched with k >= 1: (1) NACE section -> 10 aggregate
 ↳ industry groups;
 (2) drop region stratum; (3) age bands 5-year -> 10-year. Never drop sex,
 ↳ cohort, size

```

```

class, or the baseline-billing calipers; realized k distribution reported
 ↳ and fed to the
 Stage-2 blind re-simulation'
- condition: Median HONUGE settlement lag (BEHANDLINGSDATO calibration, 2021+
 ↳ rows) > 4 weeks
 rule: 'Guard band: exclude Q1(c+1) from the theta2a/theta2b windows (means
 ↳ over Q2(c+1)..Q4(c+1))
 and report theta1 both as registered and with a 4-week-shifted boundary;
 ↳ the registered
 theta1 window remains primary'
- condition: KONTAKT does not behave as a row-level contact marker on
 ↳ in-environment validation
 rule: 'Switch to the frozen fallback contact definition: count of distinct
 ↳ PNR x HONUGE-week
 x YDERNR x GP-code-group cells with >= 1 positive-fee row; switch and
 ↳ level effect reported;
 all estimands re-run under both definitions with the fallback as primary'
- condition: OLS and Poisson QML disagree in sign on any primary quantity
 rule: Foreground the pre-committed OLS estimate, report both, flag the
 ↳ discordance in the
 Summary paragraph; no post hoc model selection
- condition: 'c-2 placebo panel FAILURE, defined as: placebo contrast
 ↳ significant at alpha
 = 0.05 AND |estimate| > 1% of control baseline mean (50% of the theta1
 ↳ floor effect) -
 the single registered failure condition, stated identically in
 ↳ outcome_neutral_checks
 and the H1 placebo kill'
 rule: 'Execute the single pre-specified re-match adding year c-2 baseline
 ↳ contacts to the
 Mahalanobis set; if the placebo fails again under the same condition, the
 ↳ entire registered
 event-study tier (confirmatory and exploratory event-study quantities) is
 ↳ killed and the
 study reports descriptive trajectories only. The two non-complementary
 ↳ cells are pre-committed
 passes with mandatory disclosure, never discretion: (a) significant at
 ↳ alpha = 0.05 with
 |estimate| <= 1% - no re-match, no kill; the Table T7 placebo row flags
 ↳ the significance
 and states that a sub-floor-scale imbalance was detected; (b) not
 ↳ significant with |estimate|
 > 1% - no re-match, no kill; a mandatory caution paragraph reports the
 ↳ placebo magnitude
 against the theta1 floor and the realized placebo MDE, and the H1/H2
 ↳ Summary paragraphs
 must cite it'
- condition: A matched control's own workplace acquires a terminal event
 ↳ within c-2..c+1 after
 matching (data revision)
 rule: Drop that control from the set; drop the whole set only if no control
 ↳ remains; dropped
 counts reported; no re-matching
- condition: YDLTID missing or code 0 ('not relevant') on a GP row
 rule: Classify as daytime (conservative against the theta3 prediction);
 ↳ missingness share
 reported; if > 20% of peri-event GP rows, H3's kill criterion triggers
- condition: Exact tie in Mahalanobis distance for a nearest-neighbor match
 rule: Deterministic tie-break by smallest absolute baseline-contact
 ↳ difference, then seeded

```

RNG (seed 20260831)

- condition: IDAS-flag vs worker-flow classification agreement < 90% but >=
  - ↳ 75%
    - rule: Proceed with the worker-flow classification as registered (the canon, S7); publish the full disagreement cross-tabulation; below 75%, H2's kill criterion governs
- condition: 'GRAFT 3: sign of theta2b or theta2\_diff not stable across the three successor-share threshold pairs'
  - ↳ rule: Flag classification fragility in the pre-committed H2 Summary paragraph; the 30/70 estimate remains primary; no threshold shopping
- condition: 'GRAFT 6: exposed-vs-control difference in later-presence rates (any IDAN record in year c+1 or any SSSY row in c+1 or later) exceeds 2 percentage points'
  - ↳ rule: A mandatory bounding paragraph accompanies all post-boundary (theta2a/theta2b) conclusions, converting the differential into best/worst-case bounds; trigger and threshold registered, not discretionary
- condition: 'GRAFT 9: the control forward-window sensitivity yields SMALLER magnitudes than the primary estimates'
  - ↳ rule: Flag as diagnostic of something other than contamination (registered direction note); reported verbatim, no re-specification
- condition: 'STAGE-2 REALIZED-N GATE: realized matched sets (or spurious sets for F2) fall below a member's binding N\* from the Stage-0 re-run, or the Stage-2 blind re-simulation (realized Ns, realized k distribution, baseline-year moments only) shows power < 0.95 for a member at its floor'
  - ↳ rule: That member is demoted to exploratory BEFORE any post-baseline outcome is accessed; the Holm family shrinks and alphas recompute by the registered fixed-point iteration among survivors; demotion-only - no member, threshold, or floor may move in the other direction; the realized-power/MDE table is published either way
- condition: Realized control-group OOH share of GP contacts < 5% (below the published 8-12% range)
  - ↳ rule: H3's power basis is re-verified in the Stage-2 blind re-simulation at the realized share; if < 0.95 at the 5-point floor, H3 demotes to exploratory per the realized-N gate; if realized share < 2%, H3's kill criterion governs instead
- condition: Spurious arm killed at Stage 1/2 (events < 150, sets < 3,000, classification agreement < 75%, or bimodality unresolved)
  - ↳ rule: Gate F2 and theta2\_diff/theta2b are reported unavailable (drop disclosed before outcome access); theta2a remains in the Holm family if its own N\* is met, but the H2 verdict is capped at 'partially consistent - comparative quantifier untestable' and the Discussion

states that the excess cannot then be distinguished from selection into  
→ disappearing workplace  
identities

## Source kill rules

- hypothesis\_id: H1  
criterion: 'PROJECT-VIABILITY KILL (unchanged): stop the project if Stage 1  
→ yields < 500  
true-closure events or Stage 2 yields < 10,000 matched sets. CONFIRMATORY  
→ DEMOTION GATE  
(new, pre-outcome): if realized matched sets < N\*\_H1 (provisional  
→ 100,000; binding value  
from the Stage-0 re-run) or the Stage-2 blind re-simulation shows power <  
→ 0.95 at the  
2% floor, theta1 demotes to exploratory before any post-baseline outcome  
→ is accessed.  
PLACEBO KILL: if the c-2 placebo fails after the single pre-specified  
→ re-match (significant  
at alpha = 0.05 AND |estimate| > 1% of control baseline mean - the single  
→ registered failure  
condition, identical in outcome\_neutral\_checks and contingent\_decisions),  
→ the entire registered  
event-study tier is killed and the study reports descriptive trajectories  
→ only.'
- hypothesis\_id: H2  
criterion: 'SPURIOUS-ARM KILL (comparative machinery): if < 150  
→ spurious-closure events  
or < 3,000 spurious matched sets survive Stage 2, or worker-flow vs  
→ IDAS-flag agreement  
< 75%, or the GRAFT-4 bimodality check fails unresolved - gate F2,  
→ theta2b, and H2\_diff  
are dropped (disclosed before outcome access) and the H2 verdict is  
→ capped at 'partially  
consistent - comparative quantifier untestable''. CONFIRMATORY DEMOTION  
→ GATE for theta2a:  
realized true matched sets < N\*\_H2a (provisional 60,000) or blind  
→ re-simulated power <  
0.95 at the 2% floor -> theta2a demotes to exploratory pre-outcome. Gate  
→ F2 additionally  
requires >= 15,000 spurious sets for its registered operating  
→ characteristics; below that,  
F2 reports bounds only and the verdict cap applies.'
- hypothesis\_id: H3  
criterion: 'KILL: if OOH contacts (SPEC2 81-84, 89) are < 2% of all  
→ control-group GP contacts  
in the baseline year (composition not estimable), or the OOH classifier  
→ is unresolvable  
(YDLTID missing/'not relevant' on > 20% of peri-event GP rows and SPEC2  
→ OOH codes absent).  
CONFIRMATORY DEMOTION GATE: realized control OOH share < 5% with blind  
→ re-simulated power  
< 0.95 at the 5-point floor, or realized matched sets < N\*\_H3  
→ (provisional 90,000) ->  
theta3 demotes to exploratory pre-outcome.'

## Complete current shared analysis plan

```

estimator: 'Stacked matched-set difference-in-differences event study,
↳ estimated by OLS (R
 >= 4.3, fixest::feols, versions frozen in the registered script): outcome
 ↳ y_iq (winsorized
 quarterly GP contact count) on treated x event-quarter indicators for the 8
 ↳ quarters Q1(c)..Q4(c+1),
 with matched-set fixed effects (set = 1 exposed + up to 3 controls,
 ↳ controls weighted 1/k_s
 within set) and cohort x calendar-quarter fixed effects; omitted baseline =
 ↳ the four quarters
 of year c-1; year c-2 quarters enter only the placebo panel. CR1
 ↳ cluster-robust SEs clustered
 on event workplace (each control inherits its set's event cluster).
 ↳ theta2a from the true-arm
 stacked model; theta2b (gate F2) and exploratory theta2_diff from the
 ↳ pooled model with
 event-class x event-quarter interactions; GRAFT-3 successor-share threshold
 ↳ sensitivities
 (20/80, 40/60 alongside primary 30/70) in one table. theta3: ratio contrast
 ↳ with percentile
 CIs from a cluster bootstrap over event workplaces (999 draws, RNG seed
 ↳ 20260831). GRAFT
 9: pre-specified sensitivity re-estimation of theta1/theta2a excluding sets
 ↳ whose control
 workplace acquires any terminal event in c+2 or c+3 (registered direction
 ↳ note: contamination
 attenuates toward null, so exclusion should weakly increase magnitudes; a
 ↳ decrease is flagged
 as diagnostic; applicable to cohorts c <= 2021 for the c+3 window and c <=
 ↳ 2022 for c+2,
 since IDAS ends 2024 - truncation disclosed). Matching estimator: 1:3
 ↳ nearest-neighbor Mahalanobis
 within exact strata, without replacement of controls across sets, target k
 ↳ = 3 with minimum
 k = 1 (realized k distribution reported and fed to the Stage-2 blind
 ↳ re-simulation), exposed
 processed in seeded random order (seed 20260831); exact strata = cohort c x
 ↳ sex (KOENIMP,
 incl. unknown) x 5-year age band (ALDERIMP, incl. unknown) x NACE section
 ↳ (ARB_HOVED_BRA_DB07)
 x workplace size class (ANTNOV: 5-9/10-19/20-49/50-99/100-499/500-1000) x
 ↳ region (workplace
 KOM mapped to the 5 post-2007 regions); Mahalanobis components (all year
 ↳ c-1): total GP
 contacts, log(1+BRUHON), OOH contacts, weeks with >= 1 contact, tenure (c-1
 ↳ - ANSAAR), wage
 decile (BREDT_LOEN_BELOEB); calipers |dif baseline contacts| <= 3 and |dif
 ↳ log(1+BRUHON)|
 <= 0.5. Poisson QML (fixest::fepois) pre-specified as the count-model
 ↳ robustness check.
 No causal machinery (no IV, no RD, no synthetic control): the claim ceiling
 ↳ is associational
 and the estimator is the simplest one that makes the contrasts exact.'
fixed_effects: matched-set fixed effects (set = 1 exposed + up to 3 controls
↳ weighted 1/k)
 + cohort x calendar-quarter fixed effects; omitted baseline = the four
 ↳ quarters of year
 c-1; year c-2 quarters enter only the placebo panel
controls:

```

- 'No post-matching regression covariates, by design: identification of the
  - ↪ descriptive contrast
  - rests entirely on exact strata + Mahalanobis matching + matched-set fixed
    - ↪ effects + cohort
    - x calendar-quarter fixed effects'
- 'Matching variables (the design's only covariates): KOENIMP, ALDERIMP,
  - ↪ ARB\_HOVED\_BRA\_DBO7,
  - ANTNOV size class, workplace KOM -> region, baseline-year GP contacts /
    - ↪  $\log(1+BRUHON)$  /
    - OOH contacts / contact-weeks (SSSY year c-1), tenure from ANSAAR, wage
      - ↪ decile from BREDT\_LOEN\_BELOEB
      - (IDAN year c-1)'

se\_clustering: CR1 cluster-robust SEs clustered on event workplace (each
 

- ↪ control inherits
- its set's event cluster); theta3 via cluster-bootstrap percentile CIs (999
  - ↪ draws, clusters
  - = event workplaces, seed 20260831)

preprocessing:

- 'Freeze code lists before data access: GP = SPEC2 in {70..80}; OOH = SPEC2
  - ↪ in {81,82,83,84,89};
  - robustness OOH classifier adds YDLTID in {2,3,4,5,7,8,9}; code labels
    - ↪ quoted verbatim from
    - parsed DST value labels and re-validated in-environment before estimation'
- 'IDAS event-panel construction per era key system (ARBNR/CVRNR/LBNR,
  - ↪ 2007+): workplace identity
  - = LBNR chain via IDFREM/IDTILB; identity-disappearance events between c and
    - ↪ c+1 flagged;
    - ENDNEDL, REGELFR, FRELFR, ANDELUPS, ANDELUDS attached'
- 'Worker-flow successor shares s(w) from IDAN: among the event workplace's
  - ↪ year-c November
  - primary employees, the largest share whose year-c+1 primary workplace is a
    - ↪ single successor
    - identity (computed on the year-c workforce, independent of the frozen c-1
      - ↪ cohort)'
- 'Event classification: true closure (ENDNEDL final AND  $s(w) < 0.30$ ),
  - ↪ spurious closure (forward
  - identity AND  $s(w) \geq 0.70$ ), else ambiguous-excluded; cross-tabulation
    - ↪ against IDAS flags
    - retained as an outcome-neutral check; GRAFT-4 bimodality check on s(w) at
      - ↪ Stage 1'
- Cohort extraction at November c-1; control pool from stable workplaces; 1:3
  - ↪ matching with
  - realized-k accounting
- 'SSSY extraction 2006-2025: person x HONUGE week aggregation of KONTAKT
  - ↪ (contacts), YDLANT
  - (services), BRUHON (fees) within frozen GP/OOH code lists; week-to-quarter
    - ↪ mapping by ISO
    - week of HONUGE (1-13/14-26/27-39/40-52; ISO week 53 to Q4)'
- Winsorization of person-quarter outcomes at the 99.9th percentile within
  - ↪ era x calendar
  - quarter
- 'Panel assembly: 16 event-quarters per set, Q1(c-2)..Q4(c+1); estimation
  - ↪ panel Q1(c-1)..Q4(c+1);
  - placebo panel Q1(c-2)..Q4(c-2)'
- BEHANDLINGSDATO (2021+) settlement-lag calibration table reported before
  - ↪ any event-study
  - estimation
- 'No deflation of BRUHON: nominal fees compared only within cohort x
  - ↪ calendar-quarter cells'

- 'Registered sensitivity suite (run once, reported together): (i) GRAFT-3
  - ↳ successor-share
  - thresholds for theta2b/theta2\_diff; (ii) GRAFT-9 control forward-window
    - ↳ exclusion; (iii)
  - un-winsorized outcomes; (iv) Poisson QML; (v) possible-death/emigration set
    - ↳ exclusion (see
    - missing\_data\_rule)'
- multiple\_comparisons: 'Holm-Bonferroni across the three confirmatory tests
  - ↳ (theta1, theta2a,
  - theta3) at family-wise alpha = 0.05, two-sided p-values with pre-registered
    - ↳ directions;
  - a Holm-significant estimate in the NON-predicted direction is
    - ↳ disconfirming, never supportive
  - (exemplar convention). Hierarchical gates: the BRUHON fee secondary is
    - ↳ tested only if theta2a
  - Holm-rejects; the exploratory H1\_trend contrast is computed only if theta1
    - ↳ Holm-rejects
  - in the predicted direction. Falsification gate F2 sits OUTSIDE the Holm
    - ↳ family as a registered
  - severity device with a pre-committed disconfirmation trigger (it consumes
    - ↳ no family alpha
  - and can only take support away, never add it). Exploratory families
    - ↳ (H2\_diff; E1; E3; E4)
  - receive Holm correction within family as a severity aid and are segregated
    - ↳ by heading. If
  - the Stage-2 realized-N gate demotes members, the family shrinks by the
    - ↳ registered fixed-point
  - iteration; the surviving family's alphas are those of the reduced family.
    - ↳ Null results
  - are bounded, not accepted: each null is converted into the effect
    - ↳ magnitudes the 95% CI
  - rules out, quoted against realized MDEs from the Stage-2 blind
    - ↳ re-simulation.'
- contingent\_decisions:
  - condition: ENDNEDL flag-year convention differs in-environment from the
    - ↳ assumed 'flagged
    - in the workplace's last active year'
  - rule: Re-anchor mechanically so that c always denotes the last IDAS year
    - ↳ with ANTNOV >=
    - 5 for the identity; the estimand is unchanged; the recalibration is
      - ↳ reported verbatim
  - condition: ARB\_SEKTORKODE is absent from the analysis-server IDAN extract
    - ↳ (publicly documented
    - on TIMES ida-ansattelser but unconfirmed in the recorded
      - ↳ Forskningsservice inventory;
    - presence and 2008+ coverage validated at Stage 1 before any
      - ↳ classification work)
    - rule: 'Frozen fallback: define the private-sector filter from
      - ↳ ARB\_HOVED\_BRA\_DB07 instead
    - a workplace counts as private if and only if its NACE DB07 section is
      - ↳ not 0 (public
      - administration and defence), P (education), or Q (human health and social
      - ↳ work); the substitution
    - and its effect on eligible-workplace and worker counts are reported
      - ↳ verbatim; no other
    - filter change is permitted'
  - condition: Match rate < 95% of eligible exposed workers at target k = 3
    - ↳ under full exact
    - strata

```

rule: 'Apply the pre-specified coarsening ladder in order, stopping at the
↳ first rung reaching
 95% of exposed matched with k >= 1: (1) NACE section -> 10 aggregate
 ↳ industry groups;
 (2) drop region stratum; (3) age bands 5-year -> 10-year. Never drop sex,
 ↳ cohort, size
 class, or the baseline-billing calipers; realized k distribution reported
 ↳ and fed to the
 Stage-2 blind re-simulation'
- condition: Median HONUGE settlement lag (BEHANDLINGSDATO calibration, 2021+
↳ rows) > 4 weeks
 rule: 'Guard band: exclude Q1(c+1) from the theta2a/theta2b windows (means
 ↳ over Q2(c+1)..Q4(c+1))
 and report theta1 both as registered and with a 4-week-shifted boundary;
 ↳ the registered
 theta1 window remains primary'
- condition: KONTAKT does not behave as a row-level contact marker on
↳ in-environment validation
 rule: 'Switch to the frozen fallback contact definition: count of distinct
 ↳ PNR x HONUGE-week
 x YDERNR x GP-code-group cells with >= 1 positive-fee row; switch and
 ↳ level effect reported;
 all estimands re-run under both definitions with the fallback as primary'
- condition: OLS and Poisson QML disagree in sign on any primary quantity
 rule: Foreground the pre-committed OLS estimate, report both, flag the
 ↳ discordance in the
 Summary paragraph; no post hoc model selection
- condition: 'c-2 placebo panel FAILURE, defined as: placebo contrast
↳ significant at alpha
 = 0.05 AND |estimate| > 1% of control baseline mean (50% of the theta1
 ↳ floor effect) -
 the single registered failure condition, stated identically in
 ↳ outcome_neutral_checks
 and the H1 placebo kill'
rule: 'Execute the single pre-specified re-match adding year c-2 baseline
↳ contacts to the
 Mahalanobis set; if the placebo fails again under the same condition, the
 ↳ entire registered
 event-study tier (confirmatory and exploratory event-study quantities) is
 ↳ killed and the
 study reports descriptive trajectories only. The two non-complementary
 ↳ cells are pre-committed
 passes with mandatory disclosure, never discretion: (a) significant at
 ↳ alpha = 0.05 with
 |estimate| <= 1% - no re-match, no kill; the Table T7 placebo row flags
 ↳ the significance
 and states that a sub-floor-scale imbalance was detected; (b) not
 ↳ significant with |estimate|
 > 1% - no re-match, no kill; a mandatory caution paragraph reports the
 ↳ placebo magnitude
 against the theta1 floor and the realized placebo MDE, and the H1/H2
 ↳ Summary paragraphs
 must cite it'
- condition: A matched control's own workplace acquires a terminal event
↳ within c-2..c+1 after
 matching (data revision)
 rule: Drop that control from the set; drop the whole set only if no control
 ↳ remains; dropped
 counts reported; no re-matching
- condition: YDLTID missing or code 0 ('not relevant') on a GP row

```

```

rule: Classify as daytime (conservative against the theta3 prediction);
 ↳ missingness share
 reported; if > 20% of peri-event GP rows, H3's kill criterion triggers
- condition: Exact tie in Mahalanobis distance for a nearest-neighbor match
rule: Deterministic tie-break by smallest absolute baseline-contact
 ↳ difference, then seeded
 RNG (seed 20260831)
- condition: IDAS-flag vs worker-flow classification agreement < 90% but >=
 ↳ 75%
rule: Proceed with the worker-flow classification as registered (the canon,
 ↳ S7); publish
 the full disagreement cross-tabulation; below 75%, H2's kill criterion
 ↳ governs
- condition: 'GRAFT 3: sign of theta2b or theta2_diff not stable across the
 ↳ three successor-share
 threshold pairs'
rule: Flag classification fragility in the pre-committed H2 Summary
 ↳ paragraph; the 30/70
 estimate remains primary; no threshold shopping
- condition: 'GRAFT 6: exposed-vs-control difference in later-presence rates
 ↳ (any IDAN record
 in year c+1 or any SSSY row in c+1 or later) exceeds 2 percentage points'
rule: A mandatory bounding paragraph accompanies all post-boundary
 ↳ (theta2a/theta2b) conclusions,
 converting the differential into best/worst-case bounds; trigger and
 ↳ threshold registered,
 not discretionary
- condition: 'GRAFT 9: the control forward-window sensitivity yields SMALLER
 ↳ magnitudes than
 the primary estimates'
rule: Flag as diagnostic of something other than contamination (registered
 ↳ direction note);
 reported verbatim, no re-specification
- condition: 'STAGE-2 REALIZED-N GATE: realized matched sets (or spurious
 ↳ sets for F2) fall
 below a member's binding N* from the Stage-0 re-run, or the Stage-2
 ↳ blind re-simulation
 (realized Ns, realized k distribution, baseline-year moments only) shows
 ↳ power < 0.95
 for a member at its floor'
rule: That member is demoted to exploratory BEFORE any post-baseline
 ↳ outcome is accessed;
 the Holm family shrinks and alphas recompute by the registered
 ↳ fixed-point iteration among
 survivors; demotion-only - no member, threshold, or floor may move in the
 ↳ other direction;
 the realized-power/MDE table is published either way
- condition: Realized control-group OOH share of GP contacts < 5% (below the
 ↳ published 8-12%
 range)
rule: H3's power basis is re-verified in the Stage-2 blind re-simulation at
 ↳ the realized
 share; if < 0.95 at the 5-point floor, H3 demotes to exploratory per the
 ↳ realized-N gate;
 if realized share < 2%, H3's kill criterion governs instead
- condition: Spurious arm killed at Stage 1/2 (events < 150, sets < 3,000,
 ↳ classification
 agreement < 75%, or bimodality unresolved)
rule: Gate F2 and theta2_diff/theta2b are reported unavailable (drop
 ↳ disclosed before outcome

```

```

access); theta2a remains in the Holm family if its own N* is met, but the
↪ H2 verdict is
capped at 'partially consistent - comparative quantifier untestable' and
↪ the Discussion
states that the excess cannot then be distinguished from selection into
↪ disappearing workplace
identities
outlier_rule: Person-quarter GP contact counts and BRUHON sums winsorized at
↪ the 99.9th percentile
within era x calendar quarter (thresholds computed pre-estimation and
↪ frozen); un-winsorized
estimates reported as sensitivity in the SI; no observation-level deletion
↪ for outlyingness.
missing_data_rule: Absence of SSSY rows is zero utilization by construction
↪ of the claims
universe, never missing data - person-quarters retained with y = 0. Persons
↪ with no SSSY
row in c-3..c-1 enter explicit 'unknown' age/sex matching strata. Workers
↪ with missing ANSAAR
or missing/zero annual pay in the c-1 IDAN record are excluded pre-matching
↪ with counts
reported. YDLTID missing -> daytime (rule above). Post-event disappearance
↪ from both IDAN
and SSSY (possible death/emigration - neither observable in scope) is
↪ retained under intention-to-follow;
a pre-specified sensitivity re-estimates excluding sets where the exposed
↪ member (or all
controls) has zero SSSY rows across c..c+1 AND no IDAN record in c+1. GRAFT
↪ 6 numeric trigger
as registered in contingent_decisions.

```

## 2. ENVIRONMENT

Software stack (record exact versions in the run manifest at Stage 0 and freeze them):

- **python3 (>= 3.11) with polars** for ALL large-table preprocessing. Lazy scans only (`scan_parquet/scan_csv`) with projection and predicate pushdown; semi-joins against the cohort PNR roster before any collect; streaming collects for aggregations. Never materialize a person x week or person x month panel beyond need: the only dense person-time table permitted is matched members x 16 event-quarters (~540k true members + ~180k spurious members, x 16 = under 12M rows). Zero-filling happens at the quarter level for panel members only, never by cross-joining the population against a calendar.
- **R (>= 4.3) with fixest, invoked via Rscript**, for every fixed-effects estimation: `feols` (OLS, the pre-committed primary) and `fepois` (Poisson QML robustness). CR1 cluster-robust SEs on event workplace. No estimation in python; no data manipulation in R beyond reading the analysis parquet and writing tidy coefficient tables.
- **matplotlib** for all figures; **booktabs LaTeX** for all tables (no vertical rules). Every number in every exhibit is written by script from a parquet/CSV output — never hand-edited.
- **Memory discipline:** process cohort-by-cohort ( $c = 2009..2023$ ) where possible; checkpoint every numbered pipeline step as parquet under `interim/` with a fixed name (given per step below); SSSY scans filter on (PNR in roster) AND (SPEC2

in frozen lists) at scan level; the full population is touched only by the aggregate coverage benchmark (step 16), which is computed as a streaming groupby per year, never collected at row level.

- **Seeded randomness:** a single root seed 20260831 (per the frozen registration) initializes every RNG: matching processing order, Mahalanobis tie-breaks, the 999-draw cluster bootstrap, and the Stage-2 blind re-simulation. No unseeded randomness anywhere.
- **Run manifest:** an append-only JSONL (`manifest/run_manifest.jsonl`), one record per pipeline step per run: step id, UTC timestamp, code version hash, input file(s) + row counts, output file + row count, parameters used, validation check results, PASS/FAIL. Every step below MUST log its row counts; a step whose check fails writes FAIL and halts downstream steps (Section 6.3). The manifest (aggregate counts only, event counts rounded per the disclosure rules) is part of the released audit trail.
- **Directory convention:** `raw/` (DST-delivered), `interim/` (parquet checkpoints), `analysis/` (panels), `out/tables/`, `out/figures/`, `manifest/`. All paths in code are relative to the project root inside the research environment; no absolute paths in any released script.
- **Frozen-code discipline:** the Stage-4 estimation code is frozen at the end of Stage 0 (hash recorded in the manifest); Stage 4 is a single run of that code. The sensitivity suite runs once and is reported together. Any code change after Stage 0 is a deviation and must be disclosed in running prose with reason and approver.

---

## Source disclosure rules

```
environment: DST Forskningservice (no prior data access; gatekeeper
↳ certification per RR
 secondary-data rules). All analysis occurs inside the environment; only
 ↳ aggregate outputs
 leave. No microdata, workplace identifiers (ARBNR/CVRNR/LBNR), or
 ↳ person-level values are
 released.
minimum_cell_size: 5
minimum_cell_rules: no released statistic based on fewer than 5 persons or
↳ fewer than 3 workplaces;
 event-count tables rounded to the nearest 10; weekly trajectory figures
 ↳ pool >= 100 events
 and >= 5,000 person-weeks per plotted point; fee (BRUHON) aggregates
 ↳ subject to a 50% single-workplace
 dominance check; min-cell suppression applied BEFORE Holm reporting so
 ↳ suppression can never
 be outcome-selective
outputs_leaving_environment:
- model coefficients with SEs/CIs
- matching balance tables and realized-k distribution
- calibration and benchmark tables (coverage, linkage, HONUGE lag, OOH share)
- figure aggregates (weekly trajectory series under the pooling rules)
- event counts rounded to the nearest 10
- Stage-0 power/MDE tables - BOTH rounds (failed and remediated) - and the
↳ Stage-2 blind re-simulation
```

output with the demotion trail

- the c-2 placebo, gate F2 outputs, GRAFT-2/4 diagnostics, GRAFT-6
  - ↪ later-presence rates, E4 attrition panel, and all outcome-neutral check outputs (the full audit trail exits under the same rules)
- the analysis and power scripts

statement: 'All analysis occurs inside DST Forskningservice; only aggregate outputs leave.

Rules applied to every released object: (1) no statistic based on fewer than 5 persons or fewer than 3 workplaces; (2) event-count tables rounded to the nearest 10; (3) weekly trajectory figures pool  $\geq 100$  events and  $\geq 5,000$  person-weeks per plotted point; (4) fee (BRUHON) statistics released only as cell means/coefficients over min-N cells, with a dominance check that no single workplace contributes  $> 50\%$  of any released fee aggregate; (5) released objects are limited to: model coefficients with SEs/CIs, balance tables, calibration/benchmark tables, figure aggregates, event counts (rounded), the Stage-0 (both rounds) and Stage-2 power/MDE tables with the demotion trail, and the analysis/power scripts - never microdata, workplace identifiers (ARBNR/CVRNR/LBNR), or person-level values; (6) min-cell suppression is applied before Holm reporting so that suppression can never be outcome-selective; (7) the c-2 placebo, gate F2 outputs, the GRAFT-2/4 diagnostics, the GRAFT-6 later-presence rates, the E4 attrition panel, and all outcome-neutral check outputs are released under the same rules, ensuring the full audit trail exits the environment.'

## Complete registration fields

```
id: atom-000003
topic: Billing measurement validity
atom_result: prereg
atom_strength: null
atom_shape: null
evidence_paths: []
concepts: []
literature: []
bonds: []
preregistration:
 version: 1
 mode: adaptive
 question: Do billing coverage, timing calibration and the registered placebo permit event-study interpretation?
 rationale: Source-existing shared validity judgment, required by several subsequent questions.
 assumption: |-
```

Associational - matched event-study contrasts, a contrastive  
 ↳ falsification gate, and a composition contrast; descriptive for  
 ↳ trajectory panels and the bounded priority claim. Binding vocabulary  
 ↳ rules: 'is associated with' / 'consistent with' / 'suggests' only;  
 ↳ 'effect', 'impact', 'causes' prohibited throughout; anticipation  
 ↳ always worded as 'pre-boundary excess' relative to the annually dated  
 ↳ closure-year boundary, never as announcement-dated anticipation (S8);  
 ↳ no phenomenon-discovery or Danish-first framing (S1-S3); H2 novelty  
 ↳ confined to the placebo's application to health billing (S7). H2  
 ↳ verdict semantics are fixed in advance: 'consistent with H2' requires  
 ↳ BOTH the confirmatory core (theta2a Holm-significant, predicted  
 ↳ direction) AND a passing falsification gate F2; the strict  
 ↳ comparative magnitude is only ever reported as an exploratory  
 ↳ estimate with its CI. Exploratory results (H1\_trend, H2\_diff, E1, E3,  
 ↳ E4, and any gate-demoted member) carry the word 'exploratory' in the  
 ↳ heading and in every sentence stating them. The prose label never  
 ↳ exceeds the evidence label.

Compatibility ratings below are neutral import metadata, not a new  
 ↳ scientific evaluation or empirical scores.

target\_strength: descriptive  
 target\_shape: comparison  
 ratings:  
 importance: 3  
 novelty: 3  
 feasibility: 3  
 data\_advantage: 3  
 informativeness: 0.5

tables:  
 - IDAN  
 - IDAS  
 - SSSY

activation: state(atom-000002) == "cohort\_ready"

verdicts:  
 - key: descriptive\_only  
 region: The c-2 placebo fails the final YAML condition after the one  
 ↳ registered  
 re-match; other data needed for the source descriptive trajectory are  
 usable.  
 says: The entire event-study tier is killed; only the source-permitted  
 descriptive trajectories remain.

- key: event\_study\_ready  
 region: Registered coverage/linkage/timing checks resolved, and the final  
 YAML c-2 placebo failure condition is not met after the permitted  
 ↳ procedure.  
 Carry H3 kill, F2 availability, warnings and demotions separately.  
 says: Event-study contrasts are eligible only at their registered  
 ↳ component-specific  
 inferential status; no causal identification is established.

- key: boundary  
 region: Required source-defined inputs or scientific definitions are  
 ↳ demonstrably  
 unavailable; retain available components and the exact unresolved scope.  
 says: An evidenced boundary, not a statistical null. A missing  
 ↳ implementation  
 or timeout alone is operational no\_result, not this state.

- key: otherwise  
 region: otherwise  
 says: The registered judgment is unresolved; report all completed  
 ↳ components  
 without inventing support, equivalence or absent results.

```

protocol: |
Do billing coverage, timing calibration and the registered placebo
↳ permit event-study interpretation?

Source fidelity and implementation order

This is a format conversion of DS1.1, not a new study or independent
↳ acceptance.
Source audit: minor_revisions. No observed Danish outcomes are supplied.
The source execution plan explicitly states that final prereg.yaml governs
disagreements. Exact current YAML selections below therefore govern older
execution-plan excerpts; these excerpts are retained as implementation
↳ detail,
not a competing registration. See docs/provenance/source-discrepancies.md.
The complete shared data/validity contracts are embedded in Atoms 1--3;
↳ read
those native protocols when executing a downstream Atom, even without
↳ docs/.
Do not substitute a scientific default for an unresolved source
↳ contradiction.

Preserve source stages 0 -> 1 -> 2 -> 3 -> 4 -> 5 inside the
↳ implementation.
Source Stage-0 simulations are historical design evidence, not study
↳ outcomes;
do not silently redo or improve them. Gate 14 must be evaluated from
↳ baseline
data before post-baseline outcomes. All surviving Stage-4 quantities form
↳ ONE
Holm family. The three question-level Atoms may share the same frozen
↳ family
fit and tidy results; do not apply Holm separately per Atom or adjudicate
↳ a
member before the complete surviving family's p-values exist. Cache/reuse
identical governed intermediate results with provenance when available; a
direct-dispatch call must obtain or compute its exact source-defined
↳ inputs,
not pretend its predecessors succeeded. Stage-5 internal riders remain
exploratory, after the Stage-4 family is frozen, even when housed in
↳ H1/H2.

Reference files under docs/execution/ are the ORIGINAL SKELETONS, not
↳ tested
turnkey implementations. Adapt paths and complete engineering in a
↳ governed
workspace; preserve source scientific rules and disclose unresolved
↳ choices.
Write a result record with component estimates, uncertainty, family
↳ eligibility,
source interpretation row, diagnostics and boundary/skip dispositions.
↳ Released
outputs follow source disclosure rules; confidential rows stay governed.
Local native states summarize the source judgment; neither a state name
↳ nor
source power forecasts prove equivalence or causal identification. A
↳ bounded
claim requires its actual interval and registered rule. Retain unresolved
↳ source
wording conflicts; do not convert mere nonsignificance into scientific
↳ null.

```

## Frozen scope and vocabulary

```
```yaml
claim_class_ceiling: associational
statement: 'Associational - matched event-study contrasts, a contrastive
↳ falsification gate,
and a composition contrast; descriptive for trajectory panels and the
↳ bounded priority claim.
Binding vocabulary rules: 'is associated with' / 'consistent with'
↳ / 'suggests' only;
'effect', 'impact', 'causes' prohibited throughout; anticipation
↳ always worded as
'pre-boundary excess' relative to the annually dated closure-year
↳ boundary, never as announcement-dated
anticipation (S8); no phenomenon-discovery or Danish-first framing
↳ (S1-S3); H2 novelty confined
to the placebo's application to health billing (S7). H2 verdict
↳ semantics are fixed in
advance: 'consistent with H2' requires BOTH the confirmatory core
↳ (theta2a Holm-significant,
predicted direction) AND a passing falsification gate F2; the strict
↳ comparative magnitude
is only ever reported as an exploratory estimate with its CI.
↳ Exploratory results (H1_trend,
H2_diff, E1, E3, E4, and any gate-demoted member) carry the word
↳ 'exploratory' in the
heading and in every sentence stating them. The prose label never
↳ exceeds the evidence label.'
frozen_parameters:
seed: 20260831
gp_code_list: daytime GP = SPEC2 in {70..80}
ooh_code_list: out-of-hours GP = SPEC2 in {81, 82, 83, 84, 89}
ooh_robustness_classifier: adds YDLTID in {2, 3, 4, 5, 7, 8, 9}
spec2_evening_gp_disclosure: 'Registered choice, disclosed: the parsed
↳ value labels mark
SPEC2 71 ("General practitioner, consultation, evening etc."), 73
↳ ("General practitioner,
telephone consultation, evening etc."), and 75 ("General
↳ practitioner, home visit, evening
etc.") as explicitly off-schedule GP services; they nonetheless
↳ remain on the daytime
side of the PRIMARY classifier because H3's out-of-hours construct
↳ and its published
power basis are keyed to the laegevagtt out-of-hours service (SPEC2
↳ 81-84, 89), the quantity
the published Danish 8-12% share describes. Attenuation direction,
↳ both readings: counting
evening GP as daytime attenuates theta3 toward zero - conservative
↳ against the positive
H3 conclusion, anti-conservative for the null-bounding
↳ "compositionally unremarkable"
reading, whose Stage-2 statement therefore carries a mandatory
↳ caveat. In-environment
validation must verify that the YDLTID-augmented robustness
↳ classifier captures SPEC2
71/73/75 rows (their evening time codes 2-5); if YDLTID proves
↳ unpopulated on those rows,
the robustness classifier - and only the robustness classifier - is
↳ amended mechanically
```

```

    and with disclosure to include SPEC2 71/73/75 directly. The SPEC2
    ↪ 71/73/75 volume share
    is reported, and both classifiers appear side by side in Table T7.'
    successor_share_thresholds: 'true closure: ENDNEDL final closure AND
    ↪ worker-flow max-successor
    share s(w) < 0.30; spurious closure: forward identity via
    ↪ IDFREEM/REGELFR/FRELFREM AND
    s(w) >= 0.70; ambiguous mid-zone excluded; GRAFT-3 sensitivities
    ↪ 0.20/0.80 and 0.40/0.60'
    quarter_mapping: 'ISO week of HONUGE: 1-13 / 14-26 / 27-39 / 40-52; ISO
    ↪ week 53 to Q4'
    winsorization: person-quarter outcomes at the 99.9th percentile within
    ↪ era x calendar quarter
    matching: 1:3 nearest-neighbor Mahalanobis within exact strata, without
    ↪ replacement across
    sets, target k = 3, minimum k = 1; calipers |dif baseline contacts|
    ↪ <= 3 and |dif log(1+BRUHON)|
    <= 0.5; exposed processed in seeded random order (seed 20260831)
    software: R >= 4.3, fixest::feols (OLS) and fixest::feopois (Poisson QML
    ↪ robustness); versions
    frozen in the registered script
    note: Code lists and thresholds frozen before any data access; code
    ↪ labels quoted verbatim
    from parsed DST value labels and re-validated in-environment before
    ↪ estimation (frozen
    fallback rules in contingent_decisions).
...
15. **SSSY outcome extraction, 2007-2025, matched members only (Stage 3
begins).** Input: raw SSSY, roster = matched members. Lazy scan;
person x HONUGE-week aggregation of KONTAKT (contacts), YDLANT
(services), BRUHON (fees), separately for daytime GP (SPEC2 70-80)
and OOH (SPEC2 81-84, 89), plus the YDLTID-augmented robustness OOH
classification (YDLTID missing or 0 -> daytime, share logged).
Validate KONTAKT as a row-level contact marker (distribution against
YDLANT and row multiplicity); if validation fails, switch to the
frozen fallback - count of distinct PNR x HONUGE-week x YDERNR x
GP-code-group cells with >= 1 positive-fee row - report the switch
and level effect, and re-run all estimands under both definitions
with the fallback as primary. Output:
`interim/sssy_weekly.parquet`. Check: value labels for SPEC2/YDLTID
re-validated verbatim against in-environment documentation before
anything downstream runs; YDLTID missing/"not relevant" share on
peri-event GP rows logged (> 20% with SPEC2 OOH codes absent
triggers H3's kill).
16. **Coverage benchmark (full extract, aggregate only).** Input: raw
SSSY (streaming groupby per year; no row-level collect). Annual GP
contacts per 1,000 population (SPEC2 70-80), 2008-2024, against the
published DST StatBank sygesikring series. Output:
`interim/coverage_benchmark.parquet`. Check: within ±10% of the
published series in every year; any year outside triggers documented
reconciliation before estimation. Also compute the linkage floor:
share of cohort workers (exposed + control pool) with >= 1 SSSY row
of any kind in c-3..c+1; >= 80% required (halt at Stage 3 below).
17. **Person-quarter aggregation and winsorization.** Input: step 15.
Map HONUGE ISO weeks to quarters (1-13/14-26/27-39/40-52; week 53 ->
Q4); aggregate to person x event-quarter; compute 99.9th-percentile
winsorization thresholds within era x calendar quarter, freeze them
(log the thresholds), and winsorize contacts and BRUHON sums.
Output: `interim/person_quarter.parquet`. Checks (floor/ceiling and
parsing sanity): control share of zero person-quarters in [30%,

```

- 85%]; 99th percentile ≤ 30 contacts/quarter; control weekly series shows the summer trough (ISO weeks 28-31 mean $\geq 10\%$ below the annual weekly mean); realized control OOH share of baseline GP contacts in [5%, 15%] (in [2%, 5%] the step-14 re-verification governs; $< 2\%$ H3 kill; $> 15\%$ code-list re-audit before estimation).
18. ****Panel assembly.**** Input: steps 13, 17. Sixteen event-quarters per member, Q1(c-2)..Q4(c+1), zero-filled where no billing rows exist (zero utilization by construction, never missingness); estimation panel = Q1(c-1)..Q4(c+1); placebo panel = Q1(c-2)..Q4(c-2) plus the c-1 baseline. Outputs: `analysis/analysis\_panel.parquet`, `analysis/placebo\_panel.parquet` (true arm); pooled panel with event-class indicators for the spurious arm. Check: rows = members x 16 exactly; weights present (exposed 1, controls 1/k_s); event workplace cluster id attached to every member.
 19. ****HONUGE settlement-lag calibration.**** Input: SSSY 2021-2024 GP rows (BEHANDLINGSDATO vs HONUGE). Output: `interim/lag\_calibration.parquet` + released table. Check: $\geq 75\%$ of rows billed with HONUGE within 4 weeks of the treatment week. If median lag > 4 weeks: guard-band contingency - exclude Q1(c+1) from the theta2a/theta2b windows (means over Q2(c+1)..Q4(c+1)) and report theta1 both as registered and with a 4-week-shifted boundary (registered window remains primary). This table is reported BEFORE any event-study estimation.
 20. ****GRAFT-6 later-presence accounting.**** Input: IDAN c+1, SSSY c+1 onward, matched members. Exposed-vs-control rates of any IDAN record in c+1 or any SSSY row in c+1 or later. Output: `interim/later\_presence.parquet`. Check: if the differential exceeds 2 percentage points, the mandatory bounding paragraph (best/worst case) accompanies every post-boundary conclusion - registered trigger, not discretionary.
 21. ****c-2 placebo estimation (outcome-neutral; final Stage-3 gate).**** Input: `analysis/placebo\_panel.parquet`. The step-4.1 estimator run on year c-2 quarters against the c-1 baseline. Output: `interim/placebo\_result.parquet`. Check: NOT significant at $\alpha = 0.05$ AND $|\text{point estimate}| \leq 1\%$ of the control baseline mean (50% of theta1's floor). On failure: execute the SINGLE pre-specified re-match adding year c-2 baseline contacts to the Mahalanobis set (back to step 13 once, flagged in the manifest); if the placebo still fails, the entire registered event-study tier (confirmatory and exploratory event-study quantities) is killed and the study reports descriptive trajectories only.

Outcome-neutral checks

```

```yaml
- check: 'Coverage benchmark: annual GP contacts per 1,000 population in
 ↳ the extract (SPEC2
 ↳ 70-80), 2008-2024, compared against published DST health-insurance
 ↳ statistics (StatBank
 sygesikring series)'
 threshold: Within $\pm 10\%$ of the published series in every year; any year
 ↳ outside triggers
 ↳ documented reconciliation before estimation
- check: 'Linkage/observability floor: share of cohort workers (exposed +
 ↳ control pool) with
 ↳ ≥ 1 SSSY row of any kind in the 5-year window c-3..c+1'
 threshold: ' $\geq 80\%$ (Danish annual GP contact prevalence benchmark
 ↳ $\sim 85\%$); below floor halts
 ↳ at Stage 3'
- check: 'Closure-rate plausibility: annual true-closure event rate among
 ↳ private size-5+
```

```

workplaces, and spurious/true event ratio'
threshold: Rate in [1%, 8%] every cohort year (consistent with the
↳ published Erhvervsdemografi
basis used for sizing); ratio in [0.1, 1.5]; violations trigger
↳ classification audit before
any outcome data access
- check: 'Classification cross-validation: agreement between worker-flow
↳ classes (s(w) thresholds)
and IDAS-native flags (ENDNEDL/IDFREM/REGELFR/FRELFREM/ANDELUDS)'
threshold: '>= 75% agreement over classified events (kill floor for the
↳ H2 comparative machinery);
full cross-tabulation released'
- check: 'GRAFT 4 - en-bloc successor-share bimodality (Stage 1):
↳ distribution of worker-flow
successor shares s(w) across terminal events'
threshold: 'Visibly bimodal with the ambiguous mid-zone ($0.30 \leq s(w) < 0.70$) containing
↳ < 35% of terminal events, per the Hethey-Maier & Schmieder
↳ classification literature;
violation triggers the classification audit and, if unresolved, the
↳ spurious-arm kill
- check: Matching balance and match rate
threshold: Standardized mean difference ≤ 0.10 on every matching
↳ covariate and on unmatched
auxiliaries after matching (control aggregates weighted $1/k$ within
↳ set); match rate $\geq$
90% of eligible exposed ($k \geq 1$) after the coarsening ladder;
↳ realized mean k reported
- check: 'Pre-baseline placebo (design validity, orthogonal to
↳ hypotheses): exposed-minus-control
contact difference across the four quarters of year c-2'
threshold: 'FAILS if and only if the placebo contrast is significant at
↳ $\alpha = 0.05$ AND
 $|\text{estimate}| > 1\%$ of control baseline mean (= 50% of the θ_1 floor
↳ effect) - the single
registered condition, stated identically in contingent_decisions and
↳ the H1 placebo kill;
failure triggers the single pre-specified re-match, then kill of the
↳ registered event-study
tier. The two non-complementary cells pass with mandatory disclosure
↳ and no discretion:
significant with $|\text{estimate}| \leq 1\%$ passes, with the significance
↳ flagged in the Table T7
placebo row; non-significant with $|\text{estimate}| > 1\%$ passes, with a
↳ mandatory caution paragraph
quoting the magnitude against the θ_1 floor and the realized
↳ placebo MDE'
- check: HONU GE settlement-lag calibration on 2021-2024 GP rows via
↳ BEHANDLINGSDATO
threshold: '>= 75% of rows billed with HONU GE within 4 weeks of the
↳ treatment week; median
lag > 4 weeks triggers the guard-band contingency'
- check: 'No floor/ceiling in outcomes: distribution of control-group
↳ person-quarter GP contacts'
threshold: Share of zero person-quarters in [30%, 85%]; 99th percentile
↳ ≤ 30 contacts/quarter;
control weekly series shows the expected summer trough (ISO weeks
↳ 28-31 mean $\geq 10\%$ below
annual weekly mean - HONU GE parsing sanity)

```

```

- check: 'GRAFT 2 - week-0 boundary-step diagnostic on the weekly
 ↳ descriptive series: discontinuity
 in the exposed-minus-control series at the first HONUGE week of year
 ↳ c+1, relative to
 the total pre-boundary (Q3(c)+Q4(c)) excess'
 threshold: 'Step <= 50% of the total pre-boundary excess; if exceeded:
 ↳ (a) quarter-level
 estimates foregrounded as the only interpretable resolution, (b) a
 ↳ caution paragraph converts
 the step size into an explicit bound on any timing-shape claim'
- check: 'Published-range OOH share benchmark (new, licenses H3''s
 ↳ remediated power basis):
 realized control-group OOH share of baseline-year GP contacts vs the
 ↳ published Danish
 8-12% range'
 threshold: 'In [5%, 15%]: proceed as registered. In [2%, 5%]: H3''s
 ↳ realized-power re-verification
 governs (contingent decision). Below 2%: H3 kill criterion. Above
 ↳ 15%: code-list re-audit
 before estimation (possible OOH misclassification)'
...

Current authoritative contingencies

```yaml
- condition: ENDNEDL flag-year convention differs in-environment from the
  ↳ assumed 'flagged
    in the workplace's last active year'
  rule: Re-anchor mechanically so that c always denotes the last IDAS
    ↳ year with ANTNOV >=
      5 for the identity; the estimand is unchanged; the recalibration is
      ↳ reported verbatim
- condition: ARB_SEKTORKODE is absent from the analysis-server IDAN
  ↳ extract (publicly documented
    on TIMES ida-ansættelser but unconfirmed in the recorded
    ↳ Forskningservice inventory;
      presence and 2008+ coverage validated at Stage 1 before any
      ↳ classification work)
  rule: 'Frozen fallback: define the private-sector filter from
    ↳ ARB_HOVED_BRA_DB07 instead
      - a workplace counts as private if and only if its NACE DB07 section
      ↳ is not 0 (public
        administration and defence), P (education), or Q (human health and
        ↳ social work); the substitution
          and its effect on eligible-workplace and worker counts are reported
          ↳ verbatim; no other
            filter change is permitted'
- condition: Match rate < 95% of eligible exposed workers at target k = 3
  ↳ under full exact
    strata
  rule: 'Apply the pre-specified coarsening ladder in order, stopping at
    ↳ the first rung reaching
      95% of exposed matched with k >= 1: (1) NACE section -> 10 aggregate
      ↳ industry groups;
        (2) drop region stratum; (3) age bands 5-year -> 10-year. Never drop
        ↳ sex, cohort, size
          class, or the baseline-billing calipers; realized k distribution
          ↳ reported and fed to the
            Stage-2 blind re-simulation'

```

- condition: Median HONUGE settlement lag (BEHANDLINGSDATO calibration,
↳ 2021+ rows) > 4 weeks
rule: 'Guard band: exclude Q1(c+1) from the theta2a/theta2b windows
↳ (means over Q2(c+1)..Q4(c+1))
and report theta1 both as registered and with a 4-week-shifted
↳ boundary; the registered
theta1 window remains primary'
- condition: KONTAKT does not behave as a row-level contact marker on
↳ in-environment validation
rule: 'Switch to the frozen fallback contact definition: count of
↳ distinct PNR x HONUGE-week
x YDERNR x GP-code-group cells with >= 1 positive-fee row; switch and
↳ level effect reported;
all estimands re-run under both definitions with the fallback as
↳ primary'
- condition: OLS and Poisson QML disagree in sign on any primary quantity
rule: Foreground the pre-committed OLS estimate, report both, flag the
↳ discordance in the
Summary paragraph; no post hoc model selection
- condition: 'c-2 placebo panel FAILURE, defined as: placebo contrast
↳ significant at alpha
= 0.05 AND |estimate| > 1% of control baseline mean (50% of the
↳ theta1 floor effect) -
the single registered failure condition, stated identically in
↳ outcome_neutral_checks
and the H1 placebo kill'
rule: 'Execute the single pre-specified re-match adding year c-2
↳ baseline contacts to the
Mahalanobis set; if the placebo fails again under the same condition,
↳ the entire registered
event-study tier (confirmatory and exploratory event-study
↳ quantities) is killed and the
study reports descriptive trajectories only. The two
↳ non-complementary cells are pre-committed
passes with mandatory disclosure, never discretion: (a) significant
↳ at alpha = 0.05 with
|estimate| <= 1% - no re-match, no kill; the Table T7 placebo row
↳ flags the significance
and states that a sub-floor-scale imbalance was detected; (b) not
↳ significant with |estimate|
> 1% - no re-match, no kill; a mandatory caution paragraph reports
↳ the placebo magnitude
against the theta1 floor and the realized placebo MDE, and the H1/H2
↳ Summary paragraphs
must cite it'
- condition: A matched control's own workplace acquires a terminal event
↳ within c-2..c+1 after
matching (data revision)
rule: Drop that control from the set; drop the whole set only if no
↳ control remains; dropped
counts reported; no re-matching
- condition: YDLTID missing or code 0 ('not relevant') on a GP row
rule: Classify as daytime (conservative against the theta3 prediction);
↳ missingness share
reported; if > 20% of peri-event GP rows, H3's kill criterion triggers
- condition: Exact tie in Mahalanobis distance for a nearest-neighbor
↳ match
rule: Deterministic tie-break by smallest absolute baseline-contact
↳ difference, then seeded
RNG (seed 20260831)

- condition: IDAS-flag vs worker-flow classification agreement < 90% but
 ↳ >= 75%
 rule: Proceed with the worker-flow classification as registered (the
 ↳ canon, S7); publish
 the full disagreement cross-tabulation; below 75%, H2's kill
 ↳ criterion governs
- condition: 'GRAFT 3: sign of theta2b or theta2_diff not stable across
 ↳ the three successor-share
 threshold pairs'
 rule: Flag classification fragility in the pre-committed H2 Summary
 ↳ paragraph; the 30/70
 estimate remains primary; no threshold shopping
- condition: 'GRAFT 6: exposed-vs-control difference in later-presence
 ↳ rates (any IDAN record
 in year c+1 or any SSSY row in c+1 or later) exceeds 2 percentage
 ↳ points'
 rule: A mandatory bounding paragraph accompanies all post-boundary
 ↳ (theta2a/theta2b) conclusions,
 converting the differential into best/worst-case bounds; trigger and
 ↳ threshold registered,
 not discretionary
- condition: 'GRAFT 9: the control forward-window sensitivity yields
 ↳ SMALLER magnitudes than
 the primary estimates'
 rule: Flag as diagnostic of something other than contamination
 ↳ (registered direction note);
 reported verbatim, no re-specification
- condition: 'STAGE-2 REALIZED-N GATE: realized matched sets (or spurious
 ↳ sets for F2) fall
 below a member's binding N* from the Stage-0 re-run, or the Stage-2
 ↳ blind re-simulation
 (realized Ns, realized k distribution, baseline-year moments only)
 ↳ shows power < 0.95
 for a member at its floor'
 rule: That member is demoted to exploratory BEFORE any post-baseline
 ↳ outcome is accessed;
 the Holm family shrinks and alphas recompute by the registered
 ↳ fixed-point iteration among
 survivors; demotion-only - no member, threshold, or floor may move in
 ↳ the other direction;
 the realized-power/MDE table is published either way
- condition: Realized control-group OOH share of GP contacts < 5% (below
 ↳ the published 8-12%
 range)
 rule: H3's power basis is re-verified in the Stage-2 blind
 ↳ re-simulation at the realized
 share; if < 0.95 at the 5-point floor, H3 demotes to exploratory per
 ↳ the realized-N gate;
 if realized share < 2%, H3's kill criterion governs instead
- condition: Spurious arm killed at Stage 1/2 (events < 150, sets <
 ↳ 3,000, classification
 agreement < 75%, or bimodality unresolved)
 rule: Gate F2 and theta2_diff/theta2b are reported unavailable (drop
 ↳ disclosed before outcome
 access); theta2a remains in the Holm family if its own N* is met, but
 ↳ the H2 verdict is
 capped at 'partially consistent - comparative quantifier untestable'
 ↳ and the Discussion
 states that the excess cannot then be distinguished from selection
 ↳ into disappearing workplace

```

... identities

## Source kill rules

```yaml
- hypothesis_id: H1
 criterion: 'PROJECT-VIABILITY KILL (unchanged): stop the project if
 ↳ Stage 1 yields < 500
 true-closure events or Stage 2 yields < 10,000 matched sets.
 ↳ CONFIRMATORY DEMOTION GATE
 (new, pre-outcome): if realized matched sets < N*_H1 (provisional
 ↳ 100,000; binding value
 from the Stage-0 re-run) or the Stage-2 blind re-simulation shows
 ↳ power < 0.95 at the
 2% floor, theta1 demotes to exploratory before any post-baseline
 ↳ outcome is accessed.
 PLACEBO KILL: if the c-2 placebo fails after the single pre-specified
 ↳ re-match (significant
 at alpha = 0.05 AND |estimate| > 1% of control baseline mean - the
 ↳ single registered failure
 condition, identical in outcome_neutral_checks and
 ↳ contingent_decisions), the entire registered
 event-study tier is killed and the study reports descriptive
 ↳ trajectories only.'
- hypothesis_id: H2
 criterion: 'SPURIOUS-ARM KILL (comparative machinery): if < 150
 ↳ spurious-closure events
 or < 3,000 spurious matched sets survive Stage 2, or worker-flow vs
 ↳ IDAS-flag agreement
 < 75%, or the GRAFT-4 bimodality check fails unresolved - gate F2,
 ↳ theta2b, and H2_diff
 are dropped (disclosed before outcome access) and the H2 verdict is
 ↳ capped at 'partially
 consistent - comparative quantifier untestable'. CONFIRMATORY
 ↳ DEMOTION GATE for theta2a:
 realized true matched sets < N*_H2a (provisional 60,000) or blind
 ↳ re-simulated power <
 0.95 at the 2% floor -> theta2a demotes to exploratory pre-outcome.
 ↳ Gate F2 additionally
 requires >= 15,000 spurious sets for its registered operating
 ↳ characteristics; below that,
 F2 reports bounds only and the verdict cap applies.'
- hypothesis_id: H3
 criterion: 'KILL: if OOH contacts (SPEC2 81-84, 89) are < 2% of all
 ↳ control-group GP contacts
 in the baseline year (composition not estimable), or the OOH
 ↳ classifier is unresolvable
 (YDLTID missing/'not relevant' on > 20% of peri-event GP rows and
 ↳ SPEC2 OOH codes absent).
 CONFIRMATORY DEMOTION GATE: realized control OOH share < 5% with
 ↳ blind re-simulated power
 < 0.95 at the 5-point floor, or realized matched sets < N*_H3
 ↳ (provisional 90,000) ->
 theta3 demotes to exploratory pre-outcome.'
...

Complete current shared analysis plan

```yaml

```

```

estimator: 'Stacked matched-set difference-in-differences event study,
↳ estimated by OLS (R
  >= 4.3, fixest::feols, versions frozen in the registered script):
  ↳ outcome y_iq (winsorized
  quarterly GP contact count) on treated x event-quarter indicators for
  ↳ the 8 quarters Q1(c)..Q4(c+1),
  with matched-set fixed effects (set = 1 exposed + up to 3 controls,
  ↳ controls weighted 1/k_s
  within set) and cohort x calendar-quarter fixed effects; omitted
  ↳ baseline = the four quarters
  of year c-1; year c-2 quarters enter only the placebo panel. CR1
  ↳ cluster-robust SEs clustered
  on event workplace (each control inherits its set's event cluster).
  ↳ theta2a from the true-arm
  stacked model; theta2b (gate F2) and exploratory theta2_diff from the
  ↳ pooled model with
  event-class x event-quarter interactions; GRAFT-3 successor-share
  ↳ threshold sensitivities
  (20/80, 40/60 alongside primary 30/70) in one table. theta3: ratio
  ↳ contrast with percentile
  CIs from a cluster bootstrap over event workplaces (999 draws, RNG seed
  ↳ 20260831). GRAFT
  9: pre-specified sensitivity re-estimation of theta1/theta2a excluding
  ↳ sets whose control
  workplace acquires any terminal event in c+2 or c+3 (registered
  ↳ direction note: contamination
  attenuates toward null, so exclusion should weakly increase magnitudes;
  ↳ a decrease is flagged
  as diagnostic; applicable to cohorts c <= 2021 for the c+3 window and c
  ↳ <= 2022 for c+2,
  since IDAS ends 2024 - truncation disclosed). Matching estimator: 1:3
  ↳ nearest-neighbor Mahalanobis
  within exact strata, without replacement of controls across sets,
  ↳ target k = 3 with minimum
  k = 1 (realized k distribution reported and fed to the Stage-2 blind
  ↳ re-simulation), exposed
  processed in seeded random order (seed 20260831); exact strata = cohort
  ↳ c x sex (KOENIMP,
  incl. unknown) x 5-year age band (ALDERIMP, incl. unknown) x NACE
  ↳ section (ARB_HOVED_BRA_DB07)
  x workplace size class (ANTNOV: 5-9/10-19/20-49/50-99/100-499/500-1000)
  ↳ x region (workplace
  KOM mapped to the 5 post-2007 regions); Mahalanobis components (all
  ↳ year c-1): total GP
  contacts, log(1+BRUHON), OOH contacts, weeks with >= 1 contact, tenure
  ↳ (c-1 - ANSAAR), wage
  decile (BREDT_LOEN_BELOEB); calipers |dif baseline contacts| <= 3 and
  ↳ |dif log(1+BRUHON)|
  <= 0.5. Poisson QML (fixest::fepois) pre-specified as the count-model
  ↳ robustness check.
  No causal machinery (no IV, no RD, no synthetic control): the claim
  ↳ ceiling is associational
  and the estimator is the simplest one that makes the contrasts exact.'
fixed_effects: matched-set fixed effects (set = 1 exposed + up to 3
↳ controls weighted 1/k)
  + cohort x calendar-quarter fixed effects; omitted baseline = the four
  ↳ quarters of year
  c-1; year c-2 quarters enter only the placebo panel
controls:

```

- 'No post-matching regression covariates, by design: identification of
 - ↳ the descriptive contrast
 - rests entirely on exact strata + Mahalanobis matching + matched-set
 - ↳ fixed effects + cohort
 - x calendar-quarter fixed effects'
- 'Matching variables (the design's only covariates): KOENIMP, ALDERIMP,
 - ↳ ARB_HOVED_BRA_DB07,
 - ANTNOV size class, workplace KOM -> region, baseline-year GP contacts /
 - ↳ $\log(1+BRUHON)$ /
 - OOH contacts / contact-weeks (SSSY year c-1), tenure from ANSAAR, wage
 - ↳ decile from BREDT_LOEN_BELOEB
 - (IDAN year c-1)'
- se_clustering: CR1 cluster-robust SEs clustered on event workplace (each
 - ↳ control inherits
 - its set's event cluster); theta3 via cluster-bootstrap percentile CIs
 - ↳ (999 draws, clusters
 - = event workplaces, seed 20260831)
- preprocessing:
 - 'Freeze code lists before data access: GP = SPEC2 in {70..80}; OOH =
 - ↳ SPEC2 in {81,82,83,84,89};
 - robustness OOH classifier adds YDLTID in {2,3,4,5,7,8,9}; code labels
 - ↳ quoted verbatim from
 - parsed DST value labels and re-validated in-environment before
 - ↳ estimation'
 - 'IDAS event-panel construction per era key system (ARBNR/CVRNR/LBNR,
 - ↳ 2007+): workplace identity
 - = LBNR chain via IDFREM/IDTILB; identity-disappearance events between c
 - ↳ and c+1 flagged;
 - ENDNEDL, REGELFR, FRELFR, ANDELOPS, ANDELUDS attached'
 - 'Worker-flow successor shares s(w) from IDAN:among the event
 - ↳ workplace's year-c November
 - primary employees, the largest share whose year-c+1 primary workplace
 - ↳ is a single successor
 - identity (computed on the year-c workforce, independent of the frozen
 - ↳ c-1 cohort)'
 - 'Event classification: true closure (ENDNEDL final AND $s(w) < 0.30$),
 - ↳ spurious closure (forward
 - identity AND $s(w) \geq 0.70$), else ambiguous-excluded; cross-tabulation
 - ↳ against IDAS flags
 - retained as an outcome-neutral check; GRAFT-4 bimodality check on s(w)
 - ↳ at Stage 1'
 - Cohort extraction at November c-1; control pool from stable workplaces;
 - ↳ 1:3 matching with
 - realized-k accounting
 - 'SSSY extraction 2006-2025: person x HONUGE week aggregation of KONTAKT
 - ↳ (contacts), YDLANT
 - (services), BRUHON (fees) within frozen GP/OOH code lists;
 - ↳ week-to-quarter mapping by ISO
 - week of HONUGE (1-13/14-26/27-39/40-52; ISO week 53 to Q4)'
 - Winsorization of person-quarter outcomes at the 99.9th percentile
 - ↳ within era x calendar
 - quarter
 - 'Panel assembly: 16 event-quarters per set, Q1(c-2)..Q4(c+1);
 - ↳ estimation panel Q1(c-1)..Q4(c+1);
 - placebo panel Q1(c-2)..Q4(c-2)'
 - BEHANDLINGSDATO (2021+) settlement-lag calibration table reported
 - ↳ before any event-study
 - estimation
 - 'No deflation of BRUHON: nominal fees compared only within cohort x
 - ↳ calendar-quarter cells'

- 'Registered sensitivity suite (run once, reported together): (i)
 - ↳ GRAFT-3 successor-share thresholds for $\theta_{2b}/\theta_{2_diff}$; (ii) GRAFT-9 control forward-window
 - ↳ exclusion; (iii) un-winsorized outcomes; (iv) Poisson QML; (v) possible-death/emigration
 - ↳ set exclusion (see `missing_data_rule`)'
- multiple_comparisons: 'Holm-Bonferroni across the three confirmatory
 - ↳ tests (θ_{1a} , θ_{2a} , θ_{3a}) at family-wise $\alpha = 0.05$, two-sided p-values with
 - ↳ pre-registered directions;
 - a Holm-significant estimate in the NON-predicted direction is
 - ↳ disconfirming, never supportive (exemplar convention). Hierarchical gates: the BRUHON fee secondary is
 - ↳ tested only if θ_{2a}
 - Holm-rejects; the exploratory H1_trend contrast is computed only if
 - ↳ θ_{1a} Holm-rejects
 - in the predicted direction. Falsification gate F2 sits OUTSIDE the Holm
 - ↳ family as a registered severity device with a pre-committed disconfirmation trigger (it
 - ↳ consumes no family α and can only take support away, never add it). Exploratory families
 - ↳ (H_{2_diff} ; E1; E3; E4)
 - receive Holm correction within family as a severity aid and are
 - ↳ segregated by heading. If the Stage-2 realized-N gate demotes members, the family shrinks by the
 - ↳ registered fixed-point iteration; the surviving family's α s are those of the reduced
 - ↳ family. Null results are bounded, not accepted: each null is converted into the effect
 - ↳ magnitudes the 95% CI rules out, quoted against realized MDEs from the Stage-2 blind
 - ↳ re-simulation.'
 - contingent_decisions:
 - condition: ENDNEDL flag-year convention differs in-environment from the
 - ↳ assumed 'flagged in the workplace's last active year'
 - rule: Re-anchor mechanically so that c always denotes the last IDAS
 - ↳ year with `ANTNOV` ≥ 5 for the identity; the estimand is unchanged; the recalibration is
 - ↳ reported verbatim
 - condition: ARB_SEKTORKODE is absent from the analysis-server IDAN
 - ↳ extract (publicly documented on TIMES ida-ansættelser but unconfirmed in the recorded
 - ↳ Forskningservice inventory; presence and 2008+ coverage validated at Stage 1 before any
 - ↳ classification work)
 - rule: 'Frozen fallback: define the private-sector filter from
 - ↳ ARB_HOVED_BRA_DB07 instead
 - a workplace counts as private if and only if its NACE DB07 section
 - ↳ is not 0 (public administration and defence), P (education), or Q (human health and
 - ↳ social work); the substitution and its effect on eligible-workplace and worker counts are reported
 - ↳ verbatim; no other filter change is permitted'
 - condition: Match rate $< 95\%$ of eligible exposed workers at target $k = 3$
 - ↳ under full exact strata

rule: 'Apply the pre-specified coarsening ladder in order, stopping at
↳ the first rung reaching
95% of exposed matched with $k \geq 1$: (1) NACE section -> 10 aggregate
↳ industry groups;
(2) drop region stratum; (3) age bands 5-year -> 10-year. Never drop
↳ sex, cohort, size
class, or the baseline-billing calipers; realized k distribution
↳ reported and fed to the
Stage-2 blind re-simulation'

- condition: Median HONUGE settlement lag (BEHANDLINGSDATO calibration,
↳ 2021+ rows) > 4 weeks

rule: 'Guard band: exclude $Q1(c+1)$ from the θ_{2a}/θ_{2b} windows
↳ (means over $Q2(c+1) \dots Q4(c+1)$)
and report θ_{2a} both as registered and with a 4-week-shifted
↳ boundary; the registered
 θ_{2a} window remains primary'

- condition: KONTAKT does not behave as a row-level contact marker on
↳ in-environment validation

rule: 'Switch to the frozen fallback contact definition: count of
↳ distinct PNR x HONUGE-week
x YDERNR x GP-code-group cells with ≥ 1 positive-fee row; switch and
↳ level effect reported;
all estimands re-run under both definitions with the fallback as
↳ primary'

- condition: OLS and Poisson QML disagree in sign on any primary quantity

rule: Foreground the pre-committed OLS estimate, report both, flag the
↳ discordance in the
Summary paragraph; no post hoc model selection

- condition: 'c-2 placebo panel FAILURE, defined as: placebo contrast
↳ significant at α
= 0.05 AND $|\text{estimate}| > 1\%$ of control baseline mean (50% of the
↳ θ_{2a} floor effect) -
the single registered failure condition, stated identically in
↳ outcome_neutral_checks
and the H1 placebo kill'

rule: 'Execute the single pre-specified re-match adding year c-2
↳ baseline contacts to the
Mahalanobis set; if the placebo fails again under the same condition,
↳ the entire registered
event-study tier (confirmatory and exploratory event-study
↳ quantities) is killed and the
study reports descriptive trajectories only. The two
↳ non-complementary cells are pre-committed
passes with mandatory disclosure, never discretion: (a) significant
↳ at $\alpha = 0.05$ with
 $|\text{estimate}| \leq 1\%$ - no re-match, no kill; the Table T7 placebo row
↳ flags the significance
and states that a sub-floor-scale imbalance was detected; (b) not
↳ significant with $|\text{estimate}|$
> 1% - no re-match, no kill; a mandatory caution paragraph reports
↳ the placebo magnitude
against the θ_{2a} floor and the realized placebo MDE, and the H1/H2
↳ Summary paragraphs
must cite it'

- condition: A matched control's own workplace acquires a terminal event
↳ within c-2..c+1 after
matching (data revision)

rule: Drop that control from the set; drop the whole set only if no
↳ control remains; dropped
counts reported; no re-matching

- condition: YDLTID missing or code 0 ('not relevant') on a GP row
 - rule: Classify as daytime (conservative against the theta3 prediction);
 - ↳ missingness share
 - reported; if > 20% of peri-event GP rows, H3's kill criterion triggers
- condition: Exact tie in Mahalanobis distance for a nearest-neighbor
 - ↳ match
 - rule: Deterministic tie-break by smallest absolute baseline-contact
 - ↳ difference, then seeded
 - RNG (seed 20260831)
- condition: IDAS-flag vs worker-flow classification agreement < 90% but
 - ↳ >= 75%
 - rule: Proceed with the worker-flow classification as registered (the
 - ↳ canon, S7); publish
 - the full disagreement cross-tabulation; below 75%, H2's kill
 - ↳ criterion governs
- condition: 'GRAFT 3: sign of theta2b or theta2_diff not stable across
 - ↳ the three successor-share
 - threshold pairs'
 - rule: Flag classification fragility in the pre-committed H2 Summary
 - ↳ paragraph; the 30/70
 - estimate remains primary; no threshold shopping
- condition: 'GRAFT 6: exposed-vs-control difference in later-presence
 - ↳ rates (any IDAN record
 - in year c+1 or any SSSY row in c+1 or later) exceeds 2 percentage
 - ↳ points'
 - rule: A mandatory bounding paragraph accompanies all post-boundary
 - ↳ (theta2a/theta2b) conclusions,
 - converting the differential into best/worst-case bounds; trigger and
 - ↳ threshold registered,
 - not discretionary
- condition: 'GRAFT 9: the control forward-window sensitivity yields
 - ↳ SMALLER magnitudes than
 - the primary estimates'
 - rule: Flag as diagnostic of something other than contamination
 - ↳ (registered direction note);
 - reported verbatim, no re-specification
- condition: 'STAGE-2 REALIZED-N GATE: realized matched sets (or spurious
 - ↳ sets for F2) fall
 - below a member's binding N* from the Stage-0 re-run, or the Stage-2
 - ↳ blind re-simulation
 - (realized Ns, realized k distribution, baseline-year moments only)
 - ↳ shows power < 0.95
 - for a member at its floor'
 - rule: That member is demoted to exploratory BEFORE any post-baseline
 - ↳ outcome is accessed;
 - the Holm family shrinks and alphas recompute by the registered
 - ↳ fixed-point iteration among
 - survivors; demotion-only - no member, threshold, or floor may move in
 - ↳ the other direction;
 - the realized-power/MDE table is published either way
- condition: Realized control-group OOH share of GP contacts < 5% (below
 - ↳ the published 8-12%
 - range)
 - rule: H3's power basis is re-verified in the Stage-2 blind
 - ↳ re-simulation at the realized
 - share; if < 0.95 at the 5-point floor, H3 demotes to exploratory per
 - ↳ the realized-N gate;
 - if realized share < 2%, H3's kill criterion governs instead
- condition: Spurious arm killed at Stage 1/2 (events < 150, sets <
 - ↳ 3,000, classification

```

    agreement < 75%, or bimodality unresolved)
rule: Gate F2 and theta2_diff/theta2b are reported unavailable (drop
↳ disclosed before outcome
    access); theta2a remains in the Holm family if its own N* is met, but
↳ the H2 verdict is
    capped at 'partially consistent - comparative quantifier untestable'
↳ and the Discussion
    states that the excess cannot then be distinguished from selection
↳ into disappearing workplace
    identities
outlier_rule: Person-quarter GP contact counts and BRUHON sums winsorized
↳ at the 99.9th percentile
    within era x calendar quarter (thresholds computed pre-estimation and
↳ frozen); un-winsorized
    estimates reported as sensitivity in the SI; no observation-level
↳ deletion for outlyingness.
missing_data_rule: Absence of SSSY rows is zero utilization by
↳ construction of the claims
    universe, never missing data - person-quarters retained with y = 0.
↳ Persons with no SSSY
    row in c-3..c-1 enter explicit 'unknown' age/sex matching strata.
↳ Workers with missing ANSAAR
    or missing/zero annual pay in the c-1 IDAN record are excluded
↳ pre-matching with counts
    reported. YDLTID missing -> daytime (rule above). Post-event
↳ disappearance from both IDAN
    and SSSY (possible death/emigration - neither observable in scope) is
↳ retained under intention-to-follow;
    a pre-specified sensitivity re-estimates excluding sets where the
↳ exposed member (or all
    controls) has zero SSSY rows across c..c+1 AND no IDAN record in c+1.
↳ GRAFT 6 numeric trigger
    as registered in contingent_decisions.
...

```

2. ENVIRONMENT

Software stack (record exact versions in the run manifest at Stage 0 and freeze them):

- **python3** (>= 3.11) with **polars** for ALL large-table preprocessing. Lazy scans only (``scan_parquet``/``scan_csv``) with projection and predicate pushdown; semi-joins against the cohort PNR roster before any collect; streaming collects for aggregations. Never materialize a person x week or person x month panel beyond need: the only dense person-time table permitted is matched members x 16 event-quarters (~540k true members + ~180k spurious members, x 16 = under 12M rows). Zero-filling happens at the quarter level for panel members only, never by cross-joining the population against a calendar.
- **R** (>= 4.3) with **fixest**, invoked via **Rscript**, for every fixed-effects estimation: ``feols`` (OLS, the pre-committed primary) and ``fepois`` (Poisson QML robustness). CR1 cluster-robust SEs on event workplace. No estimation in python; no data manipulation in R beyond reading the analysis parquet and writing tidy coefficient tables.
- **matplotlib** for all figures; **booktabs LaTeX** for all tables (no vertical rules). Every number in every exhibit is written by script from a parquet/CSV output - never hand-edited.
- **Memory discipline**: process cohort-by-cohort (c = 2009..2023) where possible; checkpoint every numbered pipeline step as parquet under ``interim/`` with a fixed name (given per step below); SSSY scans filter on (PNR in roster) AND (SPEC2 in frozen lists) at scan level; the full

- population is touched only by the aggregate coverage benchmark (step 16), which is computed as a streaming groupby per year, never collected at row level.
- ****Seeded randomness:**** a single root seed 20260831 (per the frozen registration) initializes every RNG: matching processing order, Mahalanobis tie-breaks, the 999-draw cluster bootstrap, and the Stage-2 blind re-simulation. No unseeded randomness anywhere.
- ****Run manifest:**** an append-only JSONL (``manifest/run_manifest.jsonl``), one record per pipeline step per run: step id, UTC timestamp, code version hash, input file(s) + row counts, output file + row count, parameters used, validation check results, PASS/FAIL. Every step below MUST log its row counts; a step whose check fails writes FAIL and halts downstream steps (Section 6.3). The manifest (aggregate counts only, event counts rounded per the disclosure rules) is part of the released audit trail.
- ****Directory convention:**** ``raw/`` (DST-delivered), ``interim/`` (parquet checkpoints), ``analysis/`` (panels), ``out/tables/``, ``out/figures/``, ``manifest/``. All paths in code are relative to the project root inside the research environment; no absolute paths in any released script.
- ****Frozen-code discipline:**** the Stage-4 estimation code is frozen at the end of Stage 0 (hash recorded in the manifest); Stage 4 is a single run of that code. The sensitivity suite runs once and is reported together. Any code change after Stage 0 is a deviation and must be disclosed in running prose with reason and approver.

Source disclosure rules

```

```yaml
environment: DST Forskningservice (no prior data access; gatekeeper
 ↳ certification per RR
 secondary-data rules). All analysis occurs inside the environment; only
 ↳ aggregate outputs
 ↳ leave. No microdata, workplace identifiers (ARBNR/CVRNR/LBNR), or
 ↳ person-level values are
 released.
minimum_cell_size: 5
minimum_cell_rules: no released statistic based on fewer than 5 persons
 ↳ or fewer than 3 workplaces;
 event-count tables rounded to the nearest 10; weekly trajectory figures
 ↳ pool >= 100 events
 and >= 5,000 person-weeks per plotted point; fee (BRUHON) aggregates
 ↳ subject to a 50% single-workplace
 dominance check; min-cell suppression applied BEFORE Holm reporting so
 ↳ suppression can never
 be outcome-selective
outputs_leaving_environment:
- model coefficients with SEs/CIs
- matching balance tables and realized-k distribution
- calibration and benchmark tables (coverage, linkage, HONUGE lag, OOH
 ↳ share)
- figure aggregates (weekly trajectory series under the pooling rules)
- event counts rounded to the nearest 10
- Stage-0 power/MDE tables - BOTH rounds (failed and remediated) - and
 ↳ the Stage-2 blind re-simulation
 ↳ output with the demotion trail
- the c-2 placebo, gate F2 outputs, GRAFT-2/4 diagnostics, GRAFT-6
 ↳ later-presence rates, E4

```

```

 attrition panel, and all outcome-neutral check outputs (the full audit
 ↪ trail exits under
 the same rules)
- the analysis and power scripts
statement: 'All analysis occurs inside DST Forskningservice; only
↪ aggregate outputs leave.
Rules applied to every released object: (1) no statistic based on fewer
 ↪ than 5 persons or
 fewer than 3 workplaces; (2) event-count tables rounded to the nearest
 ↪ 10; (3) weekly trajectory
 figures pool >= 100 events and >= 5,000 person-weeks per plotted point;
 ↪ (4) fee (BRUHON)
statistics released only as cell means/coefficients over min-N cells,
 ↪ with a dominance check
that no single workplace contributes > 50% of any released fee
 ↪ aggregate; (5) released objects
are limited to: model coefficients with SEs/CIs, balance tables,
 ↪ calibration/benchmark tables,
figure aggregates, event counts (rounded), the Stage-0 (both rounds)
 ↪ and Stage-2 power/MDE
tables with the demotion trail, and the analysis/power scripts - never
 ↪ microdata, workplace
 identifiers (ARBNR/CVRNR/LBNR), or person-level values; (6) min-cell
 ↪ suppression is applied
before Holm reporting so that suppression can never be
 ↪ outcome-selective; (7) the c-2 placebo,
gate F2 outputs, the GRAFT-2/4 diagnostics, the GRAFT-6 later-presence
 ↪ rates, the E4 attrition
panel, and all outcome-neutral check outputs are released under the
 ↪ same rules, ensuring
 the full audit trail exits the environment.'
...

```

## Pending registered experiment

No Danish study outcome has been observed for this bundle. See the complete registration above.

### atom-000004: Preboundary billing

**Is primary-care billing elevated before the annual closure boundary, and is its registered timing pattern observed?**

#### Source fidelity and implementation order

This is a format conversion of DS1.1, not a new study or independent acceptance. Source audit: minor\_revisions. No observed Danish outcomes are supplied. The source execution plan explicitly states that final prereg.yaml governs disagreements. Exact current YAML selections below therefore govern older execution-plan excerpts; these excerpts are retained as implementation detail, not a competing registration. See docs/provenance/source-discrepancies.md. The complete shared data/validity contracts are embedded in Atoms 1–3; read those native protocols when executing a

downstream Atom, even without docs/. Do not substitute a scientific default for an unresolved source contradiction.

Preserve source stages 0 -> 1 -> 2 -> 3 -> 4 -> 5 inside the implementation. Source Stage-0 simulations are historical design evidence, not study outcomes; do not silently redo or improve them. Gate 14 must be evaluated from baseline data before post-baseline outcomes. All surviving Stage-4 quantities form ONE Holm family. The three question-level Atoms may share the same frozen family fit and tidy results; do not apply Holm separately per Atom or adjudicate a member before the complete surviving family's p-values exist. Cache/reuse identical governed intermediate results with provenance when available; a direct-dispatch call must obtain or compute its exact source-defined inputs, not pretend its predecessors succeeded. Stage-5 internal riders remain exploratory, after the Stage-4 family is frozen, even when housed in H1/H2.

Reference files under docs/execution/ are the ORIGINAL SKELETONS, not tested turnkey implementations. Adapt paths and complete engineering in a governed workspace; preserve source scientific rules and disclose unresolved choices. Write a result record with component estimates, uncertainty, family eligibility, source interpretation row, diagnostics and boundary/skip dispositions. Released outputs follow source disclosure rules; confidential rows stay governed. Local native states summarize the source judgment; neither a state name nor source power forecasts prove equivalence or causal identification. A bounded claim requires its actual interval and registered rule. Retain unresolved source wording conflicts; do not convert mere nonsignificance into scientific null.

## Frozen scope and vocabulary

```
claim_class_ceiling: associational
statement: 'Associational - matched event-study contrasts, a contrastive
↳ falsification gate,
 and a composition contrast; descriptive for trajectory panels and the
 ↳ bounded priority claim.
Binding vocabulary rules: 'is associated with' / 'consistent with' /
↳ 'suggests' only;
'effect', 'impact', 'causes' prohibited throughout; anticipation
↳ always worded as
'pre-boundary excess' relative to the annually dated closure-year
↳ boundary, never as announcement-dated
anticipation (S8); no phenomenon-discovery or Danish-first framing (S1-S3);
↳ H2 novelty confined
to the placebo's application to health billing (S7). H2 verdict semantics
↳ are fixed in
advance: 'consistent with H2' requires BOTH the confirmatory core
↳ (theta2a Holm-significant,
predicted direction) AND a passing falsification gate F2; the strict
↳ comparative magnitude
is only ever reported as an exploratory estimate with its CI. Exploratory
↳ results (H1_trend,
H2_diff, E1, E3, E4, and any gate-demoted member) carry the word
↳ 'exploratory' in the
heading and in every sentence stating them. The prose label never exceeds
↳ the evidence label.'
frozen_parameters:
seed: 20260831
```

```

gp_code_list: daytime GP = SPEC2 in {70..80}
ooh_code_list: out-of-hours GP = SPEC2 in {81, 82, 83, 84, 89}
ooh_robustness_classifier: adds YDLTID in {2, 3, 4, 5, 7, 8, 9}
spec2_evening_gp_disclosure: 'Registered choice, disclosed: the parsed
↳ value labels mark
 SPEC2 71 ("General practitioner, consultation, evening etc."), 73
 ↳ ("General practitioner,
 telephone consultation, evening etc."), and 75 ("General practitioner,
 ↳ home visit, evening
 etc.") as explicitly off-schedule GP services; they nonetheless remain on
 ↳ the daytime
 side of the PRIMARY classifier because H3's out-of-hours construct and
 ↳ its published
 power basis are keyed to the laegevagt out-of-hours service (SPEC2 81-84,
 ↳ 89), the quantity
 the published Danish 8-12% share describes. Attenuation direction, both
 ↳ readings: counting
 evening GP as daytime attenuates theta3 toward zero - conservative
 ↳ against the positive
 H3 conclusion, anti-conservative for the null-bounding "compositionally
 ↳ unremarkable"
 reading, whose Stage-2 statement therefore carries a mandatory caveat.
 ↳ In-environment
 validation must verify that the YDLTID-augmented robustness classifier
 ↳ captures SPEC2
 71/73/75 rows (their evening time codes 2-5); if YDLTID proves
 ↳ unpopulated on those rows,
 the robustness classifier - and only the robustness classifier - is
 ↳ amended mechanically
 and with disclosure to include SPEC2 71/73/75 directly. The SPEC2
 ↳ 71/73/75 volume share
 is reported, and both classifiers appear side by side in Table T7.'
successor_share_thresholds: 'true closure: ENDNEDL final closure AND
↳ worker-flow max-successor
 share s(w) < 0.30; spurious closure: forward identity via
 ↳ IDFREEM/REGELFR/FRELFREM AND
 s(w) >= 0.70; ambiguous mid-zone excluded; GRAFT-3 sensitivities
 ↳ 0.20/0.80 and 0.40/0.60'
quarter_mapping: 'ISO week of HONUGE: 1-13 / 14-26 / 27-39 / 40-52; ISO
↳ week 53 to Q4'
winsorization: person-quarter outcomes at the 99.9th percentile within era
↳ x calendar quarter
matching: 1:3 nearest-neighbor Mahalanobis within exact strata, without
↳ replacement across
 sets, target k = 3, minimum k = 1; calipers |dif baseline contacts| <= 3
 ↳ and |dif log(1+BRUHON)|
 <= 0.5; exposed processed in seeded random order (seed 20260831)
software: R >= 4.3, fixest::feols (OLS) and fixest::fepois (Poisson QML
↳ robustness); versions
 frozen in the registered script
note: Code lists and thresholds frozen before any data access; code labels
↳ quoted verbatim
 from parsed DST value labels and re-validated in-environment before
 ↳ estimation (frozen
 fallback rules in contingent_decisions).

```

#### 4.1 theta1 — H1 pre-boundary excess (confirmatory)

- **Model:** `feols(y ~ i(event_quarter, treated, ref = baseline) | set_id + cohort^calquarter, weights = w, cluster = ~event_wp)` on the true-arm estimation panel.
- **Primary quantity:**  $\theta_1 = (\beta_{Q3c} + \beta_{Q4c})/2$ , by linear combination of coefficients; report in contacts per person-quarter AND as % of the matched-control baseline mean.
- **Test:** two-sided  $p$  from the CR1 covariance; Holm within the surviving family; support requires  $\theta_1 > 0$ . Floor: 2% of baseline (0.022 contacts/person-quarter).
- **Decision rules:** Holm-significant negative estimate = pre-registered informative reversal (deferred/forgone care reading), disconfirming, never reframed. Null = bounded: convert the 95% CI into the excess magnitudes ruled out, quoted against the realized MDE from step 14. Gate for H1\_trend: computed only if  $\theta_1$  Holm-rejects in the predicted direction.

### 4. ANALYSIS EXECUTION

Confirmatory tier first (Stage 4, frozen code, single run); exploratory tier after (Stage 5), segregated by heading. All estimation in R fixest via Rscript. Common estimator: outcome  $y_{iq}$  (winsorized quarterly GP contact count) on treated  $x$  event-quarter indicators for the 8 quarters  $Q1(c)..Q4(c+1)$ , matched-set fixed effects, cohort  $x$  calendar-quarter fixed effects, omitted baseline = the four quarters of year  $c-1$ ; observation weights: exposed 1, controls  $1/k_s$  within set; CR1 SEs clustered on event workplace (each control inherits its set's event cluster). Min-cell suppression is applied to all outputs BEFORE Holm reporting so suppression can never be outcome-selective.

#### 4.4 Family adjudication and reporting order

Holm-Bonferroni over the surviving confirmatory family (as gated at step 14) at family  $\alpha = 0.05$ . Report all surviving tests exactly once, regardless of results. Each confirmatory hypothesis closes with a boldface **Summary (H\_k)**. verdict paragraph opening “Consistent with” / “Contrary to our hypothesis”, per the pre-committed interpretation table in `prereg.yaml` (interpretations block — 16 rows covering positive, null, reversed, gate-fired, and partial patterns; apply them verbatim). Scoring attaches per the registered scoring\_rule: H1 confirmed only if  $\theta_1$  Holm-rejects positively AND the gated H1\_trend supports the growth rider; partially\_confirmed for flat elevation; reversed if negative. H2 confirmed only if  $\theta_{2a}$  Holm-rejects positively AND F2 does not fire AND the  $\theta_{2b}$  sign pattern is GRAFT-3-stable; reversed if  $\theta_{2a}$  is significantly negative OR F2 fires. H3 confirmed on a positive Holm rejection (including the total-down/OOH-up pattern); reversed if negative. Demoted or killed members attach no score; their trail is disclosed. Sensitivity suite (run once, one table): GRAFT-3; GRAFT-9 (re-estimate  $\theta_1/\theta_{2a}$  excluding sets whose control workplace acquires any terminal event in  $c+2$  or  $c+3$  — cohorts  $c \leq 2022$  for  $c+2$ ,  $c \leq 2021$  for  $c+3$ , truncation disclosed; registered direction note: exclusion should weakly INCREASE magnitudes, a decrease is flagged as diagnostic); un-winsorized outcomes; Poisson QML (`fepois`; if OLS and Poisson disagree in sign

on any primary quantity, foreground the pre-committed OLS estimate, report both, flag the discordance in the Summary paragraph); possible-death/emigration exclusion (drop sets where the exposed member, or all controls, has zero SSSY rows across  $c..c+1$  AND no IDAN record in  $c+1$ ).

## 5. EXHIBITS

Every figure and table the report commits to. Disclosure rules binding on EVERY released object: no statistic based on fewer than 5 persons or fewer than 3 workplaces; event-count tables rounded to the nearest 10; weekly trajectory figures pool  $\geq 100$  events and  $\geq 5,000$  person-weeks per plotted point; BRUHON aggregates pass a 50% single-workplace dominance check; released objects are limited to model coefficients with SEs/CIs, balance tables, calibration/benchmark tables, figure aggregates, rounded event counts, the power/MDE tables (both Stage-0 rounds and Stage 2) with the demotion trail, and the analysis/power scripts — never microdata, workplace identifiers (ARBNR/CVRNR/LBNR), or person-level values; suppression applied BEFORE Holm reporting.

### Figures (matplotlib):

- **F1 — Quarterly event-study panel.** Treated  $x$  event-quarter coefficients with 95% CIs across  $Q1(c-2)..Q4(c+1)$  (placebo quarters, omitted-baseline marker, terminal-year and post-year quarters); panel  
(a) true arm, panel (b) spurious arm from the pooled model. Feeds: step-18 panels + 4.1/4.2 coefficient tables. Disclosure: coefficients only.
- **F2 — Weekly descriptive trajectory.** Pooled exposed-minus- matched-control weekly GP contact series around the closure-year boundary, with the GRAFT-2 week-0 step diagnostic annotated (step  $\leq 50\%$  of the total  $Q3(c)+Q4(c)$  excess; if exceeded, the figure carries the registered caution and quarter-level estimates are foregrounded). Feeds: step 15/17 weekly aggregates. Disclosure:  $\geq 100$  events and  $\geq 5,000$  person-weeks per point.
- **F3 — s(w) successor-share distribution (GRAFT-4).** Histogram of worker-flow successor shares across terminal events with the 0.30/0.70 thresholds and mid-zone mass; bimodality check exhibit. Feeds: step 7. Disclosure: bin counts rounded to 10; bins under min-cell suppressed.
- **F4 — OOH vs daytime relative excess (theta3).** The two relative-excess series over the peri-event window with the bootstrap CI of their gap; primary and YDLTID-augmented classifier side by side. Feeds: 4.3 outputs. Disclosure: coefficients only.
- **F5 — Matching balance (love plot).** SMDs before/after matching for all matching covariates and unmatched auxiliaries; 0.10 line marked. Feeds: step 13. Disclosure: aggregates only.
- **F6 — E1 SYSI-era replication (exploratory; skippable).** Quarterly SYSI-era analogue of F1(a) for cohorts 1992–2003, labeled exploratory, YDLANT-weighted proxy outcome. Feeds: step 22. Disclosure: as F1.

### Tables (booktabs LaTeX):

- **T1 — Event extraction and classification by cohort year.** Terminal events, true/spurious/ambiguous counts (rounded to 10), annual true-closure rate among private size-5+ workplaces, spurious/true ratio. Feeds: step 8.
- **T2 — Classification cross-validation.** Worker-flow classes vs IDAS-native flags (ENDNEDL/IDFREM/REGELFR/FRELFREM/ANDELUDS), full cross-tabulation + agreement rate; GRAFT-3 classification counts under 20/80 and 40/60. Feeds: step 8.
- **T3 — Cohort flow accounting.** Inclusion/exclusion counts per rule (tenure, sector, size, SIKGRUP, missing ANSAAR/pay, multi-cohort dedup), eligible exposed per cohort, match rate, ladder rung used, realized k distribution. Feeds: steps 9–13.
- **T4 — Matching balance.** Per-covariate SMDs (weighted 1/k), caliper compliance, realized mean k. Feeds: step 13.
- **T5 — Power and demotion trail.** Stage-0 power results (BOTH rounds: the voided round-1 run and the remediated re-run) and the Stage-2 blind realized-power/MDE table with binding N\*, realized sets, and any demotions with the fixed-point recomputation. Feeds: deposited script  
– step 14. Published regardless of outcome.
- **T6 — Outcome-neutral check panel.** All eleven registered checks (coverage benchmark by year, linkage floor, closure-rate plausibility, classification agreement, GRAFT-4, balance/match rate, c-2 placebo, HONUUGE lag calibration, floor/ceiling + seasonality, GRAFT-2 step, OOH-share benchmark) with thresholds and PASS/FAIL. Feeds: steps 8, 13, 16, 17, 19, 21 + F2/F3 diagnostics.
- **T7 — Confirmatory results.** theta1, theta2a, theta3: point estimates (contacts/person-quarter and % of baseline), 95% CIs, two-sided p, Holm-adjusted verdicts; the F2 gate row (theta2b, its one-sided 95% upper bound, trigger evaluation); the gated BRUHON fee secondary (if computed); c-2 placebo row; realized MDEs beside every null with the CI-ruled-out magnitudes. Feeds: 4.1–4.4.
- **T8 — Sensitivity suite.** GRAFT-3 thresholds; GRAFT-9 forward-window exclusion (with truncation note); un-winsorized; Poisson QML; death/emigration exclusion — one table, run once. Feeds: 4.4.
- **T9 — Exploratory tier.** H1\_trend (if gated in; otherwise the row states the gate was not met), H2\_diff with quarter-by-quarter differential profile, E3 fee and subgroup analyses (within-family Holm), E1 summary. Every row labeled exploratory. Feeds: 4.5.
- **T10 — Attrition and later-presence accounting.** E4 year-c separation shares by event class (ANSXFREM-based, post-treatment- conditioned label and selection caveat) and the GRAFT-6 later-presence rates with the 2pp trigger status. Feeds: steps 20 + 4.5 E4.

---

## Current source hypotheses (all fields)

- id: H1  
tier: confirmatory

statement\_initial: 'H1 (anticipation): Compared to riskset-matched workers  
 ↳ at stable workplaces  
 ↳ with parallel 52-week baseline billing histories, workers employed at  
 ↳ workplaces in their  
 ↳ final year before permanent closure (IDAS ENDNEDL) will show elevated  
 ↳ weekly primary-care  
 ↳ contact rates (SYSI/SSSY contacts dated by HONUKE; YDLANT, KONTAKT)  
 ↳ already in the two  
 ↳ calendar quarters preceding the closure-year boundary, with the excess  
 ↳ growing as the  
 ↳ boundary approaches - before any separation is visible in annual IDAN  
 ↳ records.'

statement\_source: idea bank entry DS1.1 (data-scope wave DS1), hypothesis 1  
 ↳ of 3; wording  
 ↳ verbatim, frozen at bank creation. The statement's growth rider is tiered  
 ↳ exploratory  
 ↳ as H1\_trend (see remediation\_record); the wording itself is untouched.

estimand: 'Operationalization of the frozen H1 statement's feasible core  
 ↳ (the 'growing  
 ↳ as the boundary approaches' rider is tiered exploratory as H1\_trend -  
 ↳ see remediation\_note):  
 ↳ matched-set difference in quarterly GP contact counts (SSSY rows with  
 ↳ SPEC2 in {70..80},  
 ↳ contacts summed from KONTAKT over HONUKE weeks) between workers attached  
 ↳ (primary November  
 ↳ employment, year c-1) to workplaces in their final year c before true  
 ↳ permanent closure  
 ↳ and their 1:3 riskset-matched controls at stable workplaces (controls  
 ↳ weighted 1/k within  
 ↳ set), in the two calendar quarters immediately preceding the closure-year  
 ↳ boundary (Q3(c),  
 ↳ Q4(c)), relative to the matched-set difference across the four baseline  
 ↳ quarters of year  
 ↳ c-1. Closure cohorts c in 2009-2023. Pre-boundary excess ONLY - never  
 ↳ announcement-dated  
 ↳ anticipation (boundary directive S8).'

primary\_quantity: 'theta1 = (beta\_Q3c + beta\_Q4c)/2 from the stacked  
 ↳ matched DiD event-study  
 ↳ regression (matched-set FE + cohort x calendar-quarter FE, CR1 SEs  
 ↳ clustered on event  
 ↳ workplace); expressed in contacts per person-quarter and as % of  
 ↳ matched-control baseline  
 ↳ mean. Predicted direction theta1 > 0. Holm family member. Floor effect  
 ↳ UNCHANGED at 2%  
 ↳ of control baseline (0.022 contacts/person-quarter), set below Hamad,  
 ↳ Modrek & Cullen  
 ↳ (2015) layoff-insecurity utilization shifts (~5-10%). Power path (from  
 ↳ the verified Stage-0  
 ↳ diagnosis's own scaling): z ~ 3.0 at 75k 1:1 pairs implies z ~ 4.4-4.6  
 ↳ at >= 120k 1:3  
 ↳ sets (1:3 matching cuts required sets ~25-33%; full-lower-expectation  
 ↳ sizing supplies  
 ↳ ~135k sets); realized-N gate N\*\_H1 = 118,000 matched sets for power >=  
 ↳ 0.95 under Holm-3  
 ↳ (binding value from the deposited Stage-0 simulation; the earlier  
 ↳ provisional figure of  
 ↳ 100,000 is superseded).'

direction: 'positive (theta1 > 0: pre-boundary excess contact)'  
 comparator: 1:3 riskset-matched workers at stable workplaces (controls  
 ↳ weighted 1/k within

set), matched within exact strata on 52 weeks of year c-1 billing  
 ↪ history, tenure, and  
 wage decile  
 exposure\_variable: primary November-status employment (year c-1, IDAN) at a  
 ↪ workplace in  
 its final year c before true permanent closure (IDAS identity absent in  
 ↪ c+1; ENDNEDL final  
 closure; worker-flow max-successor share  $s(w) < 0.30$ )  
 outcome\_variable: quarterly GP contact counts (SSSY SPEC2 70-80; KONTAKT  
 ↪ summed over HONUGE  
 weeks), quarters Q3(c) and Q4(c) versus the four c-1 baseline quarters  
 forecast:  
 point: +1.5% of matched-control baseline mean (approx. +0.017 contacts  
 ↪ per person-quarter  
 on a 1.1 baseline)  
 interval\_90: '[-2.0%, +5.5%] of baseline (approx. [-0.022, +0.061]  
 ↪ contacts per person-quarter)'  
 basis: 'Agent forecast, recorded before any data access and kept separate  
 ↪ from any future  
 expert forecasts (collaboration roadmap rule). Upward anchors: the  
 ↪ construct-level ancestors  
 found elevated GP consultations around a factory closure including its  
 ↪ pre-closure insecurity  
 phase (Beale & Nethercott 1985), and Danish GP care is free at the  
 ↪ point of use with  
 no insurance-loss channel. Downward anchors: monthly US claims showed  
 ↪ outpatient utilization  
 FELL under layoff-threat insecurity (Hamad, Modrek & Cullen 2015), job  
 ↪ loss reduced  
 doctor visits in US survey data (Schaller & Stevens 2015), the Danish  
 ↪ displacement-health  
 literature contains a published null (Browning, Moller Dano & Heinesen  
 ↪ 2006), and annual  
 November-status closure dating dilutes any announcement-locked response  
 ↪ across the terminal  
 year. My central forecast sits BELOW the registered 2% floor: I  
 ↪ consider effects at  
 the floor slightly more likely than not to be absent, and the bounded  
 ↪ null is an outcome  
 I genuinely expect with ~40% probability.'  
 null\_substantive: 'Primary-care contact among workers at closing workplaces  
 ↪ does not move  
 before the closure-year boundary: any anticipatory response at or above  
 ↪ 2% of baseline  
 is ruled out at the design''s error rates, locating whatever billing  
 ↪ response exists at  
 or after separation. A Holm-significant DECREASE is not this null but a  
 ↪ pre-registered  
 informative reversal (deferred/forgone care under insecurity, per Hamad,  
 ↪ Modrek & Cullen  
 2015).'

kill\_criterion: 'PROJECT-VIABILITY KILL (unchanged): stop the project if  
 ↪ Stage 1 yields  
 < 500 true-closure events or Stage 2 yields < 10,000 matched sets.  
 ↪ CONFIRMATORY DEMOTION  
 GATE (new, pre-outcome): if realized matched sets <  $N \cdot H_1$  (provisional  
 ↪ 100,000; binding  
 value from the Stage-0 re-run) or the Stage-2 blind re-simulation shows  
 ↪ power < 0.95 at

```

the 2% floor, theta1 demotes to exploratory before any post-baseline
↳ outcome is accessed.
PLACEBO KILL: if the c-2 placebo fails after the single pre-specified
↳ re-match (significant
at alpha = 0.05 AND |estimate| > 1% of control baseline mean - the single
↳ registered failure
condition, identical in outcome_neutral_checks and contingent_decisions),
↳ the entire registered
event-study tier is killed and the study reports descriptive trajectories
↳ only.'
- id: H1_trend
tier: exploratory
statement_initial: "... with the excess growing as the boundary approaches
↳ ..."
statement_source: 'the growth rider of frozen bank hypothesis H1, excerpted
↳ verbatim; tiered
exploratory because its only confirmatory operationalization is verified
↳ structurally
underpowered (round-1 Stage-0 power 0.380 at the ramp-implied 1.71%
↳ floor; remediated
re-run gated joint power 0.772). Gated: computed only if theta1
↳ Holm-rejects in the predicted
direction.'
estimand: 'The frozen H1 statement''s strict growth quantifier (''with the
↳ excess growing
as the boundary approaches''), tiered EXPLORATORY because its
↳ ramp-implied floor (1.71%
of baseline for a single-quarter contrast) is verified structurally
↳ underpowered (Stage-0
power 0.380; ~4.5x the honest variance budget would be needed). Gated:
↳ the within-terminal-year
linear contrast beta_Q4c - beta_Q1c is computed and foregrounded ONLY if
↳ theta1's Holm-adjusted
test rejects in the predicted direction; otherwise the quarterly
↳ trajectory is shown descriptively
with no contrast statistic. No interpretation row references any
↳ unregistered quarter-ordering
claim.'
primary_quantity: beta_Q4c - beta_Q1c (contacts per person-quarter), 95%
↳ CI, CR1 event-workplace
clusters; weekly-resolution trajectory figures as descriptive support;
↳ exploratory flagging
convention only.
direction: positive (beta_Q4c - beta_Q1c > 0) under the frozen rider;
↳ exploratory flagging
convention only

```

## Current source interpretation rows

```

- hypothesis_id: H1
outcome_pattern: theta1 > 0, Holm-significant
conclusion: We will conclude that workers at closing workplaces show excess
↳ primary-care
contact in the final two pre-boundary quarters, consistent with
↳ anticipatory care-seeking
and/or billing responses to separations already realized early within the
↳ terminal year,
with the E4 year-c separation share bounding the latter channel - as a
↳ pre-boundary association

```

relative to an annually dated boundary, expressly not an  
 ↳ announcement-dated anticipation  
 effect (first register-scale, matched, weekly-resolution estimate in  
 ↳ primary care; not  
 a phenomenon discovery - Iversen 1989; Rugulies et al. 2009). The gated  
 ↳ H1\_trend contrast  
 (exploratory) then determines which timing-shape language is permitted.

- hypothesis\_id: H1  
 outcome\_pattern: theta1 not Holm-significant (NULL)  
 conclusion: 'We will conclude that we have no evidence that primary-care  
 ↳ contact is elevated  
   in the two pre-boundary quarters among workers at closing workplaces  
   ↳ relative to matched  
   stable-workplace controls, and we will report the excess magnitudes the  
   ↳ 95% CI rules out  
   ('the data are inconsistent with a pre-boundary excess larger than X% of  
   ↳ baseline'),  
   quoted against the realized MDE. This null is informative at the  
   ↳ registered  $\geq 0.95$  power:  
   anticipatory volume responses at or above 2% of baseline are ruled out at  
   ↳ the design's  
   error rates.'
- hypothesis\_id: H1  
 outcome\_pattern: theta1 < 0, Holm-significant (reversed)  
 conclusion: We will conclude - contrary to our hypothesis - that  
 ↳ pre-boundary primary-care  
   contact falls, a pattern consistent with forgone or deferred routine care  
   ↳ under job insecurity  
   as reported by Hamad, Modrek & Cullen (2015); this reversal is  
   ↳ pre-registered as an informative  
   outcome, is disconfirming for H1, and will not be reframed as support.
- hypothesis\_id: H1\_trend  
 outcome\_pattern: Gate met (theta1 Holm-rejects, predicted direction) AND  
 ↳ beta\_Q4c - beta\_Q1c  
   > 0 with 95% CI excluding zero (exploratory flag)  
 conclusion: We will report, as a gated exploratory estimate, that the  
 ↳ pre-boundary excess  
   grows within the terminal year toward the boundary - consistent with, but  
   ↳ not probative  
   of, intensifying anticipatory care-seeking; annual event dating still  
   ↳ forbids any announcement-locked  
   claim, and the growth quantifier of the frozen H1 statement is scored  
   ↳ exploratory-supported  
   only.
- hypothesis\_id: H1\_trend  
 outcome\_pattern: Gate met but trend CI includes zero (PARTIAL - significant  
 ↳ excess, no growth)  
 conclusion: 'We will conclude that the excess is a flat elevation across  
 ↳ the terminal year:  
   given annual event dating we cannot separate early-arriving anticipation  
   ↳ from time-invariant  
   residual selection; the flat profile weakens the anticipation reading,  
   ↳ and we will lean  
   on the F2 spurious-closure gate and the c-2 placebo panel to bound the  
   ↳ selection component.  
   H1's frozen growth rider is scored not supported at the resolution  
   ↳ available.'
- hypothesis\_id: H1\_trend  
 outcome\_pattern: Gate not met

conclusion: We will not compute or foreground the trend contrast; the  
 ↳ quarterly trajectory  
 is shown descriptively with no statistic attached, and no timing-shape  
 ↳ language appears  
 in the Summary.

## Source family scoring and limitations

scoring\_rule: 'Fixed in advance per the collaboration''s registered-report  
 ↳ conventions (trajectory  
 fields): at Stage 2 an outcome attaches to each confirmatory hypothesis''s  
 ↳ frozen statement\_initial  
 as confirmed | partially\_confirmed | reversed, adjudicated strictly against  
 ↳ the pre-committed  
 interpretation table in this protocol. H1 scores confirmed only if  
 ↳ theta1''s Holm-adjusted  
 two-sided test rejects with the point estimate in the predicted direction  
 ↳ AND the gated  
 exploratory H1\_trend contrast supports the frozen statement''s growth rider  
 ↳ (beta\_Q4c -  
 beta\_Q1c > 0 with 95% CI excluding zero); it scores partially\_confirmed  
 ↳ when theta1 Holm-rejects  
 in the predicted direction but the growth rider is unsupported (flat  
 ↳ elevation - the pre-committed  
 reading given annual event dating) or the trend gate is not computed;  
 ↳ reversed if theta1  
 is Holm-significant in the negative direction. H2 scores confirmed only if  
 ↳ theta2a Holm-rejects  
 in the predicted direction AND falsification gate F2 does not fire AND the  
 ↳ theta2b sign  
 pattern is stable across the GRAFT-3 successor-share thresholds;  
 ↳ partially\_confirmed when  
 theta2a Holm-rejects but the comparative quantifier is untestable  
 ↳ (spurious-arm kill - verdict  
 capped by pre-commitment at ''partially consistent - comparative quantifier  
 ↳ untestable'');  
 reversed if theta2a is Holm-significant in the negative direction OR gate  
 ↳ F2 fires (the  
 frozen statement''s comparative reading is disconfirmed by a non-inert  
 ↳ placebo regardless  
 of theta2a - registered as disconfirming, never reframed). H3 scores  
 ↳ confirmed only if theta3''s  
 Holm-adjusted test rejects in the predicted direction (including the  
 ↳ total-down/OOH-up pattern,  
 pre-registered as scoreable support for the composition claim); reversed if  
 ↳ Holm-significant  
 in the negative direction. A Holm-nonsignificant confirmatory result  
 ↳ attaches no confirmation:  
 it is reported as a bounded informative null - the 95% CI converted into  
 ↳ the excess magnitudes  
 the data rule out against realized MDEs - never ''accepted''. A member  
 ↳ demoted by the Stage-2  
 realized-N gate or killed by its registered criterion attaches no score;  
 ↳ the demotion/kill  
 trail is disclosed. Significant effects in the non-predicted direction are  
 ↳ disconfirming,  
 never supportive. Exploratory quantities (H1\_trend, H2\_diff, E1, E3, E4,  
 ↳ and any demoted  
 member) attach outcomes only under the exploratory heading and are never  
 ↳ scored as preregistered-confirmed.'

assumptions:

- SSSY (2005-2025) and SYSI (1990-2005) capture the universe of publicly
  - ↳ reimbursed primary-care
  - billing for insured residents; absence of rows means no billed care, not
  - missingness
- KONTAKT is a row-level contact marker whose sum yields contact counts
  - ↳ (in-environment validation
  - with a frozen fallback definition if not)
- ENDNEDL/IDFREM/IDTILB/REGELFR/FRELFREM semantics follow Statistics
  - ↳ Denmark's IDA workplace-continuity
  - documentation, cited directly (no DOI exists for these conventions, per
  - ↳ boundary directive);
  - the primary classification rests on the worker-flow measure  $s(w)$  from IDAN,
  - ↳ so misreading
  - an IDAS flag degrades only the cross-check
- November-status annual attachment measures workplace membership; intra-year
  - ↳ attachment changes
  - are unobservable ex ante and the frozen c-1 anchor is the design's answer
  - ↳ to that - with
  - the corollary, named as a design threat, that early year-c separations feed
  - ↳ realized post-job-loss
  - billing into the pre-boundary window under intention-to-follow, a channel
  - ↳ the E4 separation
  - shares and the registered theta1 attachment split bound but cannot remove
- ALDERIMP/KOENIMP (DST-imputed age/sex on SSSY rows) are accurate enough for
  - ↳ stratified matching;
  - persons absent from SSSY in c-3..c-1 are handled by explicit unknown strata
- HONUGE settlement lags are predominantly  $\leq 4$  weeks (calibrated on
  - ↳ BEHANDLINGSDATO 2021+;
  - guard-band contingency if not), and lag structure does not differ between
  - ↳ exposed and matched
  - controls within the same calendar quarters
- Death and emigration are unobservable in scope; their post-event incidence
  - ↳ is assumed small;
  - the deflating direction is conservative for the pre-boundary claim and
  - ↳ attenuating for theta2a,
  - with the GRAFT-6 trigger bounding the post-boundary side
- Nominal-fee comparisons within cohort x calendar-quarter cells require no
  - ↳ deflator
- Danish institutional stability of GP billing categories (SPEC2 70-84, 89)
  - ↳ across 2007-2024
  - at the 2-digit level
- 'PUBLISHED-BASIS SIZING ASSUMPTIONS (new, each with a registered failure
  - ↳ path): (i) Danish
  - establishment closure rates among private size-5+ workplaces are in the
  - ↳ published Erhvervsdemografi-consistent
  - 1.5-3% band, supporting 150,000-300,000 exposed workers over 15 cohorts -
  - ↳ failure path:
  - the pre-outcome realized-N gate demotes rather than stretches; (ii) the
  - ↳ Danish OOH share
  - of GP contact volume is in the published 8-12% band (Christensen & Olesen
  - ↳ 1998; LV-KOS 2011
  - studies), simulated at the conservative 8% - failure path: the
  - ↳ published-range benchmark
  - check plus realized-share re-simulation, then demotion or kill; (iii) the
  - ↳ control pool supports
  - 1:3 matching at  $\geq 90\%$  match rate - failure path: coarsening ladder,
  - ↳ realized-k accounting
  - in the blind re-simulation, then the gate'

- theta2a's 4-quarter averaging gain over theta1 survives serial correlation
  - ↳ in post-year shocks; a registered serial-correlation discount is built into the Stage-0 re-simulation rather than assumed away
- The persistent-vs-transient overdispersion split (persistent frailty var
  - ↳ 1.25, transient log-var 0.25) is retained unchanged from the first-round simulation -
    - ↳ deliberately NOT revised in the direction that would raise power, so the remediation cannot be accused of tuning the DGP
- failure\_modes:
  - 'Realized event or worker counts far below schema-derived expectations
    - ↳ (documentation-only corpus; no realized Ns exist) - the design now bears this risk openly: the Stage-2 realized-N gate demotes underpowered members loudly and pre-outcome instead of letting an underpowered confirmatory test run; a sparse-closure Denmark kills or demotes cleanly rather than quietly'
  - KONTAKT or ENDNEDL semantics differ from assumption - both carry frozen
    - ↳ fallback rules, but level shifts in outcomes or event dating would force reported recalibration and consume credibility
  - HONUGE settlement lag materially longer than 4 weeks - the quarter-level
    - ↳ estimands survive, but the weekly-resolution headline weakens to quarterly resolution in practice (GRAFT-2 makes this explicit and pre-committed)
  - ANDELOPS direction-of-absorption semantics may differ from the idea's gloss
    - ↳ - classification routes through IDAN worker flows, so this only weakens the IDAS cross-check, not the H2 machinery
  - Spurious-closure events too heterogeneous for a clean placebo - a non-inert
    - ↳ placebo now shows up as gate F2 firing (or exploratory H2\_diff reversal), which caps H2's verdict by pre-commitment instead of silently contaminating a differential test
  - Realized OOH share below the published range - caught by the
    - ↳ published-range benchmark check before estimation; H3 demotes or dies pre-outcome instead of running underpowered
  - Realized N lands below the full-expectation lower bound - the gate demotes
    - ↳ affected members to exploratory pre-outcome; the study survives with a smaller confirmatory family (or none, in which case it reports as registered exploratory estimation with the demotion trail published
      - the first-round outcome, now a contingency rather than the design)
  - Unobservable post-closure death/emigration deflates post-boundary billing -
    - ↳ direction stated, sensitivity pre-specified, GRAFT-6 2pp trigger forces best/worst-case bounds; no in-scope fix exists
  - Public/private sector misclassification via modal ARB\_SEKTORKODE for
    - ↳ multi-establishment

- firms - attenuates contrasts; share bounded in the balance tables; if
  - ↪ ARB\_SEKTORKODE itself
- proves absent in-environment (it is unconfirmed in the recorded
  - ↪ Forskningservice inventory),
- the frozen ARB\_HOVED\_BRA\_DB07 fallback in contingent\_decisions substitutes,
  - ↪ with counts
- reported
- Multi-job workers' primary-job assignment may misattribute exposure -
  - ↪ deterministic rule
  - biases toward null; share reported
- 1:3 matching without replacement may exhaust thin strata - realized  $k < 3$ 
  - ↪ accounted honestly
  - in the blind re-simulation rather than assumed; coarsening ladder
  - ↪ pre-committed
- The SYSI-era exploratory replication may be infeasible - bounded as
  - ↪ skipable Stage-5 work
  - that cannot contaminate the confirmatory tier
- Readers may over-read gate F2 as proof of causal job-loss effects - the
  - ↪ associational ceiling,
  - the 'consistent with' vocabulary cap, and the pre-committed F2-fired
  - ↪ interpretation guard
- the prose

## Source disclosure rules

environment: DST Forskningservice (no prior data access; gatekeeper  
 ↪ certification per RR  
 secondary-data rules). All analysis occurs inside the environment; only  
 ↪ aggregate outputs  
 leave. No microdata, workplace identifiers (ARBNR/CVRNR/LBNR), or  
 ↪ person-level values are  
 released.

minimum\_cell\_size: 5

minimum\_cell\_rules: no released statistic based on fewer than 5 persons or  
 ↪ fewer than 3 workplaces;  
 event-count tables rounded to the nearest 10; weekly trajectory figures  
 ↪ pool  $\geq 100$  events  
 and  $\geq 5,000$  person-weeks per plotted point; fee (BRUHON) aggregates  
 ↪ subject to a 50% single-workplace  
 dominance check; min-cell suppression applied BEFORE Holm reporting so  
 ↪ suppression can never  
 be outcome-selective

outputs\_leaving\_environment:

- model coefficients with SEs/CIs
- matching balance tables and realized-k distribution
- calibration and benchmark tables (coverage, linkage, HONUGE lag, OOH share)
- figure aggregates (weekly trajectory series under the pooling rules)
- event counts rounded to the nearest 10
- Stage-0 power/MDE tables - BOTH rounds (failed and remediated) - and the
  - ↪ Stage-2 blind re-simulation
  - output with the demotion trail
- the c-2 placebo, gate F2 outputs, GRAFT-2/4 diagnostics, GRAFT-6
  - ↪ later-presence rates, E4
  - attrition panel, and all outcome-neutral check outputs (the full audit
  - ↪ trail exits under
  - the same rules)
- the analysis and power scripts

statement: 'All analysis occurs inside DST Forskningservice; only aggregate  
 ↪ outputs leave.'

Rules applied to every released object: (1) no statistic based on fewer  
 → than 5 persons or  
 fewer than 3 workplaces; (2) event-count tables rounded to the nearest 10;  
 → (3) weekly trajectory  
 figures pool  $\geq 100$  events and  $\geq 5,000$  person-weeks per plotted point; (4)  
 → fee (BRUHON)  
 statistics released only as cell means/coefficients over min-N cells, with  
 → a dominance check  
 that no single workplace contributes  $> 50\%$  of any released fee aggregate;  
 → (5) released objects  
 are limited to: model coefficients with SEs/CIs, balance tables,  
 → calibration/benchmark tables,  
 figure aggregates, event counts (rounded), the Stage-0 (both rounds) and  
 → Stage-2 power/MDE  
 tables with the demotion trail, and the analysis/power scripts - never  
 → microdata, workplace  
 identifiers (ARBNR/CVRNR/LBNR), or person-level values; (6) min-cell  
 → suppression is applied  
 before Holm reporting so that suppression can never be outcome-selective;  
 → (7) the c-2 placebo,  
 gate F2 outputs, the GRAFT-2/4 diagnostics, the GRAFT-6 later-presence  
 → rates, the E4 attrition  
 panel, and all outcome-neutral check outputs are released under the same  
 → rules, ensuring  
 the full audit trail exits the environment.'

## Complete registration fields

```
id: atom-000004
topic: Preboundary billing
atom_result: prereg
atom_strength: null
atom_shape: null
evidence_paths: []
concepts: []
literature: []
bonds: []
preregistration:
 version: 1
 mode: adaptive
 question: Is primary-care billing elevated before the annual closure
 → boundary,
 and is its registered timing pattern observed?
 rationale: Source-existing coherent scientific question; its related
 → outcomes,
 riders and sensitivity specifications remain internal.
 assumption: |-
```

Associational - matched event-study contrasts, a contrastive  
 ↪ falsification gate, and a composition contrast; descriptive for  
 ↪ trajectory panels and the bounded priority claim. Binding vocabulary  
 ↪ rules: 'is associated with' / 'consistent with' / 'suggests' only;  
 ↪ 'effect', 'impact', 'causes' prohibited throughout; anticipation  
 ↪ always worded as 'pre-boundary excess' relative to the annually dated  
 ↪ closure-year boundary, never as announcement-dated anticipation (S8);  
 ↪ no phenomenon-discovery or Danish-first framing (S1-S3); H2 novelty  
 ↪ confined to the placebo's application to health billing (S7). H2  
 ↪ verdict semantics are fixed in advance: 'consistent with H2' requires  
 ↪ BOTH the confirmatory core (theta2a Holm-significant, predicted  
 ↪ direction) AND a passing falsification gate F2; the strict  
 ↪ comparative magnitude is only ever reported as an exploratory  
 ↪ estimate with its CI. Exploratory results (H1\_trend, H2\_diff, E1, E3,  
 ↪ E4, and any gate-demoted member) carry the word 'exploratory' in the  
 ↪ heading and in every sentence stating them. The prose label never  
 ↪ exceeds the evidence label.  
 Compatibility ratings below are neutral import metadata, not a new  
 ↪ scientific evaluation or empirical scores.  
 target\_strength: associational  
 target\_shape: comparison  
 ratings:  
   importance: 3  
   novelty: 3  
   feasibility: 3  
   data\_advantage: 3  
   informativeness: 0.5  
 tables:  
 - IDAN  
 - IDAS  
 - SSSY  
 activation: state(atom-000003) == "event\_study\_ready" OR state(atom-000003)  
   == "descriptive\_only"  
 verdicts:  
 - key: descriptive\_only  
   region: Shared validity judgment is descriptive\_only.  
   says: Produce only the already specified descriptive trajectories, not  
       H1 or H1\_trend estimates.  
 - key: exploratory  
   region: H1 was demoted by the pre-outcome gate and an interpretable source  
       estimate is available.  
   says: Report exploratory estimates without confirmatory scoring; do not  
       promote the internal trend rider.  
 - key: preboundary\_excess  
   region: Eligible theta1 > 0 and its two-sided Holm test rejects.  
   says: Pre-boundary billing excess only. Apply the source H1\_trend gate  
       internally; distinguish supported growth from unresolved growth.  
 - key: reversed  
   region: Eligible theta1 < 0 and its two-sided Holm test rejects.  
   says: The pre-boundary decline contradicts H1; it is not support.  
 - key: not\_supported  
   region: Eligible H1 is estimated but its Holm test does not reject.  
   says: Report the actual CI and realized MDE; no confirmation or automatic  
       equivalence claim.  
 - key: boundary  
   region: Required source-defined inputs or scientific definitions are  
       ↪ demonstrably  
       unavailable; retain available components and the exact unresolved scope.  
   says: An evidenced boundary, not a statistical null. A missing  
       ↪ implementation

```

 or timeout alone is operational no_result, not this state.
- key: otherwise
 region: otherwise
 says: The registered judgment is unresolved; report all completed
 ↪ components
 without inventing support, equivalence or absent results.
protocol: |
 # Is primary-care billing elevated before the annual closure boundary,
 ↪ and is its registered timing pattern observed?

Source fidelity and implementation order

This is a format conversion of DS1.1, not a new study or independent
↪ acceptance.
Source audit: minor_revisions. No observed Danish outcomes are supplied.
The source execution plan explicitly states that final prereg.yaml governs
disagreements. Exact current YAML selections below therefore govern older
execution-plan excerpts; these excerpts are retained as implementation
↪ detail,
not a competing registration. See docs/provenance/source-discrepancies.md.
The complete shared data/validity contracts are embedded in Atoms 1--3;
↪ read
those native protocols when executing a downstream Atom, even without
↪ docs/.
Do not substitute a scientific default for an unresolved source
↪ contradiction.

Preserve source stages 0 -> 1 -> 2 -> 3 -> 4 -> 5 inside the
↪ implementation.
Source Stage-0 simulations are historical design evidence, not study
↪ outcomes;
do not silently redo or improve them. Gate 14 must be evaluated from
↪ baseline
data before post-baseline outcomes. All surviving Stage-4 quantities form
↪ ONE
Holm family. The three question-level Atoms may share the same frozen
↪ family
fit and tidy results; do not apply Holm separately per Atom or adjudicate
↪ a
member before the complete surviving family's p-values exist. Cache/reuse
identical governed intermediate results with provenance when available; a
direct-dispatch call must obtain or compute its exact source-defined
↪ inputs,
not pretend its predecessors succeeded. Stage-5 internal riders remain
exploratory, after the Stage-4 family is frozen, even when housed in
↪ H1/H2.

Reference files under docs/execution/ are the ORIGINAL SKELETONS, not
↪ tested
turnkey implementations. Adapt paths and complete engineering in a
↪ governed
workspace; preserve source scientific rules and disclose unresolved
↪ choices.
Write a result record with component estimates, uncertainty, family
↪ eligibility,
source interpretation row, diagnostics and boundary/skip dispositions.
↪ Released
outputs follow source disclosure rules; confidential rows stay governed.
Local native states summarize the source judgment; neither a state name
↪ nor

```

```

source power forecasts prove equivalence or causal identification. A
→ bounded
claim requires its actual interval and registered rule. Retain unresolved
→ source
wording conflicts; do not convert mere nonsignificance into scientific
→ null.

Frozen scope and vocabulary

```yaml
claim_class_ceiling: associational
statement: 'Associational - matched event-study contrasts, a contrastive
→ falsification gate,
and a composition contrast; descriptive for trajectory panels and the
→ bounded priority claim.
Binding vocabulary rules: 'is associated with' / 'consistent with'
→ / 'suggests' only;
'effect', 'impact', 'causes' prohibited throughout; anticipation
→ always worded as
'pre-boundary excess' relative to the annually dated closure-year
→ boundary, never as announcement-dated
anticipation (S8); no phenomenon-discovery or Danish-first framing
→ (S1-S3); H2 novelty confined
to the placebo's application to health billing (S7). H2 verdict
→ semantics are fixed in
advance: 'consistent with H2' requires BOTH the confirmatory core
→ (theta2a Holm-significant,
predicted direction) AND a passing falsification gate F2; the strict
→ comparative magnitude
is only ever reported as an exploratory estimate with its CI.
→ Exploratory results (H1_trend,
H2_diff, E1, E3, E4, and any gate-demoted member) carry the word
→ 'exploratory' in the
heading and in every sentence stating them. The prose label never
→ exceeds the evidence label.'
frozen_parameters:
seed: 20260831
gp_code_list: daytime GP = SPEC2 in {70..80}
ooh_code_list: out-of-hours GP = SPEC2 in {81, 82, 83, 84, 89}
ooh_robustness_classifier: adds YDLTID in {2, 3, 4, 5, 7, 8, 9}
spec2_evening_gp_disclosure: 'Registered choice, disclosed: the parsed
→ value labels mark
SPEC2 71 ("General practitioner, consultation, evening etc."), 73
→ ("General practitioner,
telephone consultation, evening etc."), and 75 ("General
→ practitioner, home visit, evening
etc.") as explicitly off-schedule GP services; they nonetheless
→ remain on the daytime
side of the PRIMARY classifier because H3's out-of-hours construct
→ and its published
power basis are keyed to the lægevagts out-of-hours service (SPEC2
→ 81-84, 89), the quantity
the published Danish 8-12% share describes. Attenuation direction,
→ both readings: counting
evening GP as daytime attenuates theta3 toward zero - conservative
→ against the positive
H3 conclusion, anti-conservative for the null-bounding
→ "compositionally unremarkable"
reading, whose Stage-2 statement therefore carries a mandatory
→ caveat. In-environment

```

```

validation must verify that the YDLTID-augmented robustness
  ↳ classifier captures SPEC2
71/73/75 rows (their evening time codes 2-5); if YDLTID proves
  ↳ unpopulated on those rows,
the robustness classifier - and only the robustness classifier - is
  ↳ amended mechanically
and with disclosure to include SPEC2 71/73/75 directly. The SPEC2
  ↳ 71/73/75 volume share
is reported, and both classifiers appear side by side in Table T7.'
successor_share_thresholds: 'true closure: ENDNEDL final closure AND
  ↳ worker-flow max-successor
share s(w) < 0.30; spurious closure: forward identity via
  ↳ IDFREEM/REGELFR/FRELFREM AND
s(w) >= 0.70; ambiguous mid-zone excluded; GRAFT-3 sensitivities
  ↳ 0.20/0.80 and 0.40/0.60'
quarter_mapping: 'ISO week of HONUGE: 1-13 / 14-26 / 27-39 / 40-52; ISO
  ↳ week 53 to Q4'
winsorization: person-quarter outcomes at the 99.9th percentile within
  ↳ era x calendar quarter
matching: 1:3 nearest-neighbor Mahalanobis within exact strata, without
  ↳ replacement across
sets, target k = 3, minimum k = 1; calipers |dif baseline contacts|
  ↳ <= 3 and |dif log(1+BRUHON)|
<= 0.5; exposed processed in seeded random order (seed 20260831)
software: R >= 4.3, fixest::feols (OLS) and fixest::fepois (Poisson QML
  ↳ robustness); versions
frozen in the registered script
note: Code lists and thresholds frozen before any data access; code
  ↳ labels quoted verbatim
from parsed DST value labels and re-validated in-environment before
  ↳ estimation (frozen
  ↳ fallback rules in contingent_decisions).
...

```

4.1 theta1 - H1 pre-boundary excess (confirmatory)

```

- **Model:** `feols(y ~ i(event_quarter, treated, ref = baseline) |
  set_id + cohort^calquarter, weights = w, cluster = ~event_wp)` on the
  true-arm estimation panel.
- **Primary quantity:** theta1 = (beta_Q3c + beta_Q4c)/2, by linear
  combination of coefficients; report in contacts per person-quarter AND
  as % of the matched-control baseline mean.
- **Test:** two-sided p from the CR1 covariance; Holm within the
  surviving family; support requires theta1 > 0. Floor: 2% of baseline
  (0.022 contacts/person-quarter).
- **Decision rules:** Holm-significant negative estimate = pre-registered
  informative reversal (deferred/forgone care reading), disconfirming,
  never reframed. Null = bounded: convert the 95% CI into the excess
  magnitudes ruled out, quoted against the realized MDE from step 14.
  Gate for H1_trend: computed only if theta1 Holm-rejects in the
  predicted direction.

```

4. ANALYSIS EXECUTION

Confirmatory tier first (Stage 4, frozen code, single run); exploratory tier after (Stage 5), segregated by heading. All estimation in R fixest via Rscript. Common estimator: outcome y_iq (winsorized quarterly GP contact count) on treated x event-quarter indicators for the 8 quarters Q1(c)..Q4(c+1), matched-set fixed effects, cohort x calendar-quarter fixed effects, omitted baseline = the four quarters of year c-1; observation weights: exposed 1, controls 1/k_s within set; CR1 SEs

clustered on event workplace (each control inherits its set's event cluster). Min-cell suppression is applied to all outputs BEFORE Holm reporting so suppression can never be outcome-selective.

4.4 Family adjudication and reporting order

Holm-Bonferroni over the surviving confirmatory family (as gated at step 14) at family alpha = 0.05. Report all surviving tests exactly once, regardless of results. Each confirmatory hypothesis closes with a boldface ****Summary (H_k).**** verdict paragraph opening "Consistent with" / "Contrary to our hypothesis", per the pre-committed interpretation table in `prereg.yaml` (interpretations block - 16 rows covering positive, null, reversed, gate-fired, and partial patterns; apply them verbatim). Scoring attaches per the registered scoring_rule: H1 confirmed only if theta1 Holm-rejects positively AND the gated H1_trend supports the growth rider; partially-confirmed for flat elevation; reversed if negative. H2 confirmed only if theta2a Holm-rejects positively AND F2 does not fire AND the theta2b sign pattern is GRAFT-3-stable; reversed if theta2a is significantly negative OR F2 fires. H3 confirmed on a positive Holm rejection (including the total-down/OOH-up pattern); reversed if negative. Demoted or killed members attach no score; their trail is disclosed. Sensitivity suite (run once, one table): GRAFT-3; GRAFT-9 (re-estimate theta1/theta2a excluding sets whose control workplace acquires any terminal event in c+2 or c+3 - cohorts c <= 2022 for c+2, c <= 2021 for c+3, truncation disclosed; registered direction note: exclusion should weakly INCREASE magnitudes, a decrease is flagged as diagnostic); un-winsorized outcomes; Poisson QML (`fepois`; if OLS and Poisson disagree in sign on any primary quantity, foreground the pre-committed OLS estimate, report both, flag the discordance in the Summary paragraph); possible-death/emigration exclusion (drop sets where the exposed member, or all controls, has zero SSSY rows across c..c+1 AND no IDAN record in c+1).

5. EXHIBITS

Every figure and table the report commits to. Disclosure rules binding on EVERY released object: no statistic based on fewer than 5 persons or fewer than 3 workplaces; event-count tables rounded to the nearest 10; weekly trajectory figures pool >= 100 events and >= 5,000 person-weeks per plotted point; BRUHON aggregates pass a 50% single-workplace dominance check; released objects are limited to model coefficients with SEs/CIs, balance tables, calibration/benchmark tables, figure aggregates, rounded event counts, the power/MDE tables (both Stage-0 rounds and Stage 2) with the demotion trail, and the analysis/power scripts - never microdata, workplace identifiers (ARBNR/CVRNR/LBNR), or person-level values; suppression applied BEFORE Holm reporting.

****Figures (matplotlib):****

- ****F1 - Quarterly event-study panel.**** Treated x event-quarter coefficients with 95% CIs across Q1(c-2)..Q4(c+1) (placebo quarters, omitted-baseline marker, terminal-year and post-year quarters); panel (a) true arm, panel (b) spurious arm from the pooled model. Feeds: step-18 panels + 4.1/4.2 coefficient tables. Disclosure: coefficients only.
- ****F2 - Weekly descriptive trajectory.**** Pooled exposed-minus-matched-control weekly GP contact series around the closure-year boundary, with the GRAFT-2 week-0 step diagnostic annotated (step <= 50% of the total Q3(c)+Q4(c) excess; if exceeded, the figure carries the registered caution and quarter-level estimates are foregrounded).

Feeds: step 15/17 weekly aggregates. Disclosure: ≥ 100 events and $\geq 5,000$ person-weeks per point.

- **F3 - s(w) successor-share distribution (GRAFT-4).** Histogram of worker-flow successor shares across terminal events with the 0.30/0.70 thresholds and mid-zone mass; bimodality check exhibit. Feeds: step 7. Disclosure: bin counts rounded to 10; bins under min-cell suppressed.
- **F4 - OOH vs daytime relative excess (theta3).** The two relative-excess series over the peri-event window with the bootstrap CI of their gap; primary and YDLTID-augmented classifier side by side. Feeds: 4.3 outputs. Disclosure: coefficients only.
- **F5 - Matching balance (love plot).** SMDs before/after matching for all matching covariates and unmatched auxiliaries; 0.10 line marked. Feeds: step 13. Disclosure: aggregates only.
- **F6 - E1 SYSI-era replication (exploratory; skippable).** Quarterly SYSI-era analogue of F1(a) for cohorts 1992-2003, labeled exploratory, YDLANT-weighted proxy outcome. Feeds: step 22. Disclosure: as F1.

****Tables (booktabs LaTeX):****

- **T1 - Event extraction and classification by cohort year.** Terminal events, true/spurious/ambiguous counts (rounded to 10), annual true-closure rate among private size-5+ workplaces, spurious/true ratio. Feeds: step 8.
- **T2 - Classification cross-validation.** Worker-flow classes vs IDAS-native flags (ENDNEDL/IDFREM/REGELFR/FRELFREM/ANDELUDS), full cross-tabulation + agreement rate; GRAFT-3 classification counts under 20/80 and 40/60. Feeds: step 8.
- **T3 - Cohort flow accounting.** Inclusion/exclusion counts per rule (tenure, sector, size, SIKGRUP, missing ANSAAR/pay, multi-cohort dedup), eligible exposed per cohort, match rate, ladder rung used, realized k distribution. Feeds: steps 9-13.
- **T4 - Matching balance.** Per-covariate SMDs (weighted 1/k), caliper compliance, realized mean k. Feeds: step 13.
- **T5 - Power and demotion trail.** Stage-0 power results (BOTH rounds: the voided round-1 run and the remediated re-run) and the Stage-2 blind realized-power/MDE table with binding N*, realized sets, and any demotions with the fixed-point recomputation. Feeds: deposited script + step 14. Published regardless of outcome.
- **T6 - Outcome-neutral check panel.** All eleven registered checks (coverage benchmark by year, linkage floor, closure-rate plausibility, classification agreement, GRAFT-4, balance/match rate, c-2 placebo, HONUGE lag calibration, floor/ceiling + seasonality, GRAFT-2 step, OOH-share benchmark) with thresholds and PASS/FAIL. Feeds: steps 8, 13, 16, 17, 19, 21 + F2/F3 diagnostics.
- **T7 - Confirmatory results.** theta1, theta2a, theta3: point estimates (contacts/person-quarter and % of baseline), 95% CIs, two-sided p, Holm-adjusted verdicts; the F2 gate row (theta2b, its one-sided 95% upper bound, trigger evaluation); the gated BRUHON fee secondary (if computed); c-2 placebo row; realized MDEs beside every null with the CI-ruled-out magnitudes. Feeds: 4.1-4.4.
- **T8 - Sensitivity suite.** GRAFT-3 thresholds; GRAFT-9 forward-window exclusion (with truncation note); un-winsorized; Poisson QML; death/emigration exclusion - one table, run once. Feeds: 4.4.
- **T9 - Exploratory tier.** H1_trend (if gated in; otherwise the row states the gate was not met), H2_diff with quarter-by-quarter differential profile, E3 fee and subgroup analyses (within-family Holm), E1 summary. Every row labeled exploratory. Feeds: 4.5.
- **T10 - Attrition and later-presence accounting.** E4 year-c separation shares by event class (ANSXFREM-based, post-treatment-conditioned label and selection caveat) and the GRAFT-6 later-presence

rates with the 2pp trigger status. Feeds: steps 20 + 4.5 E4.

Current source hypotheses (all fields)

```
```yaml
- id: H1
 tier: confirmatory
 statement_initial: 'H1 (anticipation): Compared to riskset-matched
 ↳ workers at stable workplaces
 ↳ with parallel 52-week baseline billing histories, workers employed at
 ↳ workplaces in their
 ↳ final year before permanent closure (IDAS ENDNEDL) will show elevated
 ↳ weekly primary-care
 ↳ contact rates (SYSI/SSSY contacts dated by HONUKE; YDLANT, KONTAKT)
 ↳ already in the two
 ↳ calendar quarters preceding the closure-year boundary, with the
 ↳ excess growing as the
 ↳ boundary approaches - before any separation is visible in annual IDAN
 ↳ records.'
 statement_source: idea bank entry DS1.1 (data-scope wave DS1),
 ↳ hypothesis 1 of 3; wording
 ↳ verbatim, frozen at bank creation. The statement's growth rider is
 ↳ tiered exploratory
 ↳ as H1_trend (see remediation_record); the wording itself is untouched.
 estimand: 'Operationalization of the frozen H1 statement's feasible
 ↳ core (the 'growing
 ↳ as the boundary approaches' rider is tiered exploratory as H1_trend
 ↳ - see remediation_note):
 ↳ matched-set difference in quarterly GP contact counts (SSSY rows with
 ↳ SPEC2 in {70..80},
 ↳ contacts summed from KONTAKT over HONUKE weeks) between workers
 ↳ attached (primary November
 ↳ employment, year c-1) to workplaces in their final year c before true
 ↳ permanent closure
 ↳ and their 1:3 riskset-matched controls at stable workplaces (controls
 ↳ weighted 1/k within
 ↳ set), in the two calendar quarters immediately preceding the
 ↳ closure-year boundary (Q3(c),
 ↳ Q4(c)), relative to the matched-set difference across the four
 ↳ baseline quarters of year
 ↳ c-1. Closure cohorts c in 2009-2023. Pre-boundary excess ONLY - never
 ↳ announcement-dated
 ↳ anticipation (boundary directive S8).'
 primary_quantity: 'theta1 = (beta_Q3c + beta_Q4c)/2 from the stacked
 ↳ matched DiD event-study
 ↳ regression (matched-set FE + cohort x calendar-quarter FE, CR1 SEs
 ↳ clustered on event
 ↳ workplace); expressed in contacts per person-quarter and as % of
 ↳ matched-control baseline
 ↳ mean. Predicted direction theta1 > 0. Holm family member. Floor
 ↳ effect UNCHANGED at 2%
 ↳ of control baseline (0.022 contacts/person-quarter), set below Hamad,
 ↳ Modrek & Cullen
 ↳ (2015) layoff-insecurity utilization shifts (~5-10%). Power path
 ↳ (from the verified Stage-0
 ↳ diagnosis's own scaling): z ~ 3.0 at 75k 1:1 pairs implies z ~
 ↳ 4.4-4.6 at >= 120k 1:3
```

sets (1:3 matching cuts required sets ~25-33%; full-lower-expectation  
↳ sizing supplies  
~135k sets); realized-N gate  $N*_H1 = 118,000$  matched sets for power  
↳  $\geq 0.95$  under Holm-3  
(binding value from the deposited Stage-0 simulation; the earlier  
↳ provisional figure of  
100,000 is superseded).'

direction: 'positive ( $\theta_1 > 0$ : pre-boundary excess contact)'  
comparator: 1:3 riskset-matched workers at stable workplaces (controls  
↳ weighted  $1/k$  within  
set), matched within exact strata on 52 weeks of year c-1 billing  
↳ history, tenure, and  
wage decile

exposure\_variable: primary November-status employment (year c-1, IDAN)  
↳ at a workplace in  
its final year c before true permanent closure (IDAS identity absent  
↳ in c+1; ENDNEDL final  
closure; worker-flow max-successor share  $s(w) < 0.30$ )

outcome\_variable: quarterly GP contact counts (SSSY SPEC2 70-80;  
↳ KONTAKT summed over HONUGE  
weeks), quarters Q3(c) and Q4(c) versus the four c-1 baseline quarters

forecast:  
point: +1.5% of matched-control baseline mean (approx. +0.017  
↳ contacts per person-quarter  
on a 1.1 baseline)  
interval\_90: '[-2.0%, +5.5%] of baseline (approx. [-0.022, +0.061]  
↳ contacts per person-quarter)'  
basis: 'Agent forecast, recorded before any data access and kept  
↳ separate from any future  
expert forecasts (collaboration roadmap rule). Upward anchors: the  
↳ construct-level ancestors  
found elevated GP consultations around a factory closure including  
↳ its pre-closure insecurity  
phase (Beale & Nethercott 1985), and Danish GP care is free at the  
↳ point of use with  
no insurance-loss channel. Downward anchors: monthly US claims  
↳ showed outpatient utilization  
FELL under layoff-threat insecurity (Hamad, Modrek & Cullen 2015),  
↳ job loss reduced  
doctor visits in US survey data (Schaller & Stevens 2015), the  
↳ Danish displacement-health  
literature contains a published null (Browning, Moller Dano &  
↳ Heinesen 2006), and annual  
November-status closure dating dilutes any announcement-locked  
↳ response across the terminal  
year. My central forecast sits BELOW the registered 2% floor: I  
↳ consider effects at  
the floor slightly more likely than not to be absent, and the  
↳ bounded null is an outcome  
I genuinely expect with ~40% probability.'

null\_substantive: 'Primary-care contact among workers at closing  
↳ workplaces does not move  
before the closure-year boundary: any anticipatory response at or  
↳ above 2% of baseline  
is ruled out at the design's error rates, locating whatever billing  
↳ response exists at  
or after separation. A Holm-significant DECREASE is not this null but  
↳ a pre-registered  
informative reversal (deferred/forgone care under insecurity, per  
↳ Hamad, Modrek & Cullen

```

2015).'
kill_criterion: 'PROJECT-VIABILITY KILL (unchanged): stop the project
↳ if Stage 1 yields
 < 500 true-closure events or Stage 2 yields < 10,000 matched sets.
 ↳ CONFIRMATORY DEMOTION
GATE (new, pre-outcome): if realized matched sets < N*_H1
↳ (provisional 100,000; binding
value from the Stage-0 re-run) or the Stage-2 blind re-simulation
↳ shows power < 0.95 at
the 2% floor, theta1 demotes to exploratory before any post-baseline
↳ outcome is accessed.
PLACEBO KILL: if the c-2 placebo fails after the single pre-specified
↳ re-match (significant
at alpha = 0.05 AND |estimate| > 1% of control baseline mean - the
↳ single registered failure
condition, identical in outcome_neutral_checks and
↳ contingent_decisions), the entire registered
event-study tier is killed and the study reports descriptive
↳ trajectories only.'
- id: H1_trend
tier: exploratory
statement_initial: '"... with the excess growing as the boundary
↳ approaches ..."'
statement_source: 'the growth rider of frozen bank hypothesis H1,
↳ excerpted verbatim; tiered
exploratory because its only confirmatory operationalization is
↳ verified structurally
underpowered (round-1 Stage-0 power 0.380 at the ramp-implied 1.71%
↳ floor; remediated
re-run gated joint power 0.772). Gated: computed only if theta1
↳ Holm-rejects in the predicted
direction.'
estimand: 'The frozen H1 statement''s strict growth quantifier (''with
↳ the excess growing
as the boundary approaches''), tiered EXPLORATORY because its
↳ ramp-implied floor (1.71%
of baseline for a single-quarter contrast) is verified structurally
↳ underpowered (Stage-0
power 0.380; ~4.5x the honest variance budget would be needed).
↳ Gated: the within-terminal-year
linear contrast beta_Q4c - beta_Q1c is computed and foregrounded ONLY
↳ if theta1''s Holm-adjusted
test rejects in the predicted direction; otherwise the quarterly
↳ trajectory is shown descriptively
with no contrast statistic. No interpretation row references any
↳ unregistered quarter-ordering
claim.'
primary_quantity: beta_Q4c - beta_Q1c (contacts per person-quarter),
↳ 95% CI, CR1 event-workplace
clusters; weekly-resolution trajectory figures as descriptive
↳ support; exploratory flagging
convention only.
direction: positive (beta_Q4c - beta_Q1c > 0) under the frozen rider;
↳ exploratory flagging
convention only
...

Current source interpretation rows

```yaml

```

- hypothesis_id: H1
outcome_pattern: $\theta_1 > 0$, Holm-significant
conclusion: We will conclude that workers at closing workplaces show
↳ excess primary-care
contact in the final two pre-boundary quarters, consistent with
↳ anticipatory care-seeking
and/or billing responses to separations already realized early within
↳ the terminal year,
with the E4 year-c separation share bounding the latter channel - as
↳ a pre-boundary association
relative to an annually dated boundary, expressly not an
↳ announcement-dated anticipation
effect (first register-scale, matched, weekly-resolution estimate in
↳ primary care; not
a phenomenon discovery - Iversen 1989; Rugulies et al. 2009). The
↳ gated H1_trend contrast
(exploratory) then determines which timing-shape language is
↳ permitted.
- hypothesis_id: H1
outcome_pattern: θ_1 not Holm-significant (NULL)
conclusion: 'We will conclude that we have no evidence that
↳ primary-care contact is elevated
in the two pre-boundary quarters among workers at closing workplaces
↳ relative to matched
stable-workplace controls, and we will report the excess magnitudes
↳ the 95% CI rules out
('the data are inconsistent with a pre-boundary excess larger than
↳ X% of baseline'),
quoted against the realized MDE. This null is informative at the
↳ registered ≥ 0.95 power:
anticipatory volume responses at or above 2% of baseline are ruled
↳ out at the design's
error rates.'
- hypothesis_id: H1
outcome_pattern: $\theta_1 < 0$, Holm-significant (reversed)
conclusion: We will conclude - contrary to our hypothesis - that
↳ pre-boundary primary-care
contact falls, a pattern consistent with forgone or deferred routine
↳ care under job insecurity
as reported by Hamad, Modrek & Cullen (2015); this reversal is
↳ pre-registered as an informative
outcome, is disconfirming for H1, and will not be reframed as support.
- hypothesis_id: H1_trend
outcome_pattern: Gate met (θ_1 Holm-rejects, predicted direction)
↳ AND $\beta_{Q4c} - \beta_{Q1c}$
 > 0 with 95% CI excluding zero (exploratory flag)
conclusion: We will report, as a gated exploratory estimate, that the
↳ pre-boundary excess
grows within the terminal year toward the boundary - consistent with,
↳ but not probative
of, intensifying anticipatory care-seeking; annual event dating still
↳ forbids any announcement-locked
claim, and the growth quantifier of the frozen H1 statement is scored
↳ exploratory-supported
only.
- hypothesis_id: H1_trend
outcome_pattern: Gate met but trend CI includes zero (PARTIAL -
↳ significant excess, no growth)
conclusion: 'We will conclude that the excess is a flat elevation
↳ across the terminal year:

```

    given annual event dating we cannot separate early-arriving
    ↪ anticipation from time-invariant
    residual selection; the flat profile weakens the anticipation
    ↪ reading, and we will lean
    on the F2 spurious-closure gate and the c-2 placebo panel to bound
    ↪ the selection component.
    H1's frozen growth rider is scored not supported at the resolution
    ↪ available.'
- hypothesis_id: H1_trend
  outcome_pattern: Gate not met
  conclusion: We will not compute or foreground the trend contrast; the
    ↪ quarterly trajectory
    is shown descriptively with no statistic attached, and no
    ↪ timing-shape language appears
    in the Summary.
...

## Source family scoring and limitations

```yaml
scoring_rule: 'Fixed in advance per the collaboration's
 ↪ registered-report conventions (trajectory
 fields): at Stage 2 an outcome attaches to each confirmatory
 ↪ hypothesis's frozen statement_initial
 as confirmed | partially_confirmed | reversed, adjudicated strictly
 ↪ against the pre-committed
 interpretation table in this protocol. H1 scores confirmed only if
 ↪ theta1's Holm-adjusted
 two-sided test rejects with the point estimate in the predicted
 ↪ direction AND the gated
 exploratory H1_trend contrast supports the frozen statement's growth
 ↪ rider (beta_Q4c -
 beta_Q1c > 0 with 95% CI excluding zero); it scores partially_confirmed
 ↪ when theta1 Holm-rejects
 in the predicted direction but the growth rider is unsupported (flat
 ↪ elevation - the pre-committed
 reading given annual event dating) or the trend gate is not computed;
 ↪ reversed if theta1
 is Holm-significant in the negative direction. H2 scores confirmed only
 ↪ if theta2a Holm-rejects
 in the predicted direction AND falsification gate F2 does not fire AND
 ↪ the theta2b sign
 pattern is stable across the GRAFT-3 successor-share thresholds;
 ↪ partially_confirmed when
 theta2a Holm-rejects but the comparative quantifier is untestable
 ↪ (spurious-arm kill - verdict
 capped by pre-commitment at 'partially consistent - comparative
 ↪ quantifier untestable');
 reversed if theta2a is Holm-significant in the negative direction OR
 ↪ gate F2 fires (the
 frozen statement's comparative reading is disconfirmed by a non-inert
 ↪ placebo regardless
 of theta2a - registered as disconfirming, never reframed). H3 scores
 ↪ confirmed only if theta3's
 Holm-adjusted test rejects in the predicted direction (including the
 ↪ total-down/00H-up pattern,
 pre-registered as scoreable support for the composition claim);
 ↪ reversed if Holm-significant
 in the negative direction. A Holm-nonsignificant confirmatory result
 ↪ attaches no confirmation:

```

it is reported as a bounded informative null - the 95% CI converted  
 ↳ into the excess magnitudes  
 the data rule out against realized MDEs - never 'accepted'. A member  
 ↳ demoted by the Stage-2  
 realized-N gate or killed by its registered criterion attaches no  
 ↳ score; the demotion/kill  
 trail is disclosed. Significant effects in the non-predicted direction  
 ↳ are disconfirming,  
 never supportive. Exploratory quantities (H1\_trend, H2\_diff, E1, E3,  
 ↳ E4, and any demoted  
 member) attach outcomes only under the exploratory heading and are  
 ↳ never scored as preregistered-confirmed.'

assumptions:

- SSSY (2005-2025) and SYSI (1990-2005) capture the universe of publicly  
 ↳ reimbursed primary-care  
 billing for insured residents; absence of rows means no billed care,  
 ↳ not missingness
- KONTAKT is a row-level contact marker whose sum yields contact counts  
 ↳ (in-environment validation  
 with a frozen fallback definition if not)
- ENDNEDL/IDFREM/IDTILB/REGELFR/FRELFREM semantics follow Statistics  
 ↳ Denmark's IDA workplace-continuity  
 documentation, cited directly (no DOI exists for these conventions, per  
 ↳ boundary directive);  
 the primary classification rests on the worker-flow measure s(w) from  
 ↳ IDAN, so misreading  
 an IDAS flag degrades only the cross-check
- November-status annual attachment measures workplace membership;  
 ↳ intra-year attachment changes  
 are unobservable ex ante and the frozen c-1 anchor is the design's  
 ↳ answer to that - with  
 the corollary, named as a design threat, that early year-c separations  
 ↳ feed realized post-job-loss  
 billing into the pre-boundary window under intention-to-follow, a  
 ↳ channel the E4 separation  
 shares and the registered theta1 attachment split bound but cannot  
 ↳ remove
- ALDERIMP/KOENIMP (DST-imputed age/sex on SSSY rows) are accurate enough  
 ↳ for stratified matching;  
 persons absent from SSSY in c-3..c-1 are handled by explicit unknown  
 ↳ strata
- HONUUGE settlement lags are predominantly <= 4 weeks (calibrated on  
 ↳ BEHANDLINGSDATO 2021+;  
 guard-band contingency if not), and lag structure does not differ  
 ↳ between exposed and matched  
 controls within the same calendar quarters
- Death and emigration are unobservable in scope; their post-event  
 ↳ incidence is assumed small;  
 the deflating direction is conservative for the pre-boundary claim and  
 ↳ attenuating for theta2a,  
 with the GRAFT-6 trigger bounding the post-boundary side
- Nominal-fee comparisons within cohort x calendar-quarter cells require  
 ↳ no deflator
- Danish institutional stability of GP billing categories (SPEC2 70-84,  
 ↳ 89) across 2007-2024  
 at the 2-digit level
- 'PUBLISHED-BASIS SIZING ASSUMPTIONS (new, each with a registered  
 ↳ failure path): (i) Danish  
 establishment closure rates among private size-5+ workplaces are in the  
 ↳ published Erhvervsdemografi-consistent

1.5-3% band, supporting 150,000-300,000 exposed workers over 15 cohorts  
 ↳ - failure path:  
 the pre-outcome realized-N gate demotes rather than stretches; (ii) the  
 ↳ Danish OOH share  
 of GP contact volume is in the published 8-12% band (Christensen &  
 ↳ Olesen 1998; LV-KOS 2011  
 studies), simulated at the conservative 8% - failure path: the  
 ↳ published-range benchmark  
 check plus realized-share re-simulation, then demotion or kill; (iii)  
 ↳ the control pool supports  
 1:3 matching at  $\geq 90\%$  match rate - failure path: coarsening ladder,  
 ↳ realized-k accounting  
 in the blind re-simulation, then the gate'

- theta2a's 4-quarter averaging gain over theta1 survives serial  
 ↳ correlation in post-year  
 shocks; a registered serial-correlation discount is built into the  
 ↳ Stage-0 re-simulation  
 rather than assumed away
- The persistent-vs-transient overdispersion split (persistent frailty  
 ↳ var 1.25, transient  
 log-var 0.25) is retained unchanged from the first-round simulation -  
 ↳ deliberately NOT revised  
 in the direction that would raise power, so the remediation cannot be  
 ↳ accused of tuning  
 the DGP

failure\_modes:

- 'Realized event or worker counts far below schema-derived expectations  
 ↳ (documentation-only  
 corpus; no realized Ns exist) - the design now bears this risk openly:  
 ↳ the Stage-2 realized-N  
 gate demotes underpowered members loudly and pre-outcome instead of  
 ↳ letting an underpowered  
 confirmatory test run; a sparse-closure Denmark kills or demotes  
 ↳ cleanly rather than quietly'
- KONTAKT or ENDNEDL semantics differ from assumption - both carry frozen  
 ↳ fallback rules,  
 but level shifts in outcomes or event dating would force reported  
 ↳ recalibration and consume  
 credibility
- HONUKE settlement lag materially longer than 4 weeks - the  
 ↳ quarter-level estimands survive,  
 but the weekly-resolution headline weakens to quarterly resolution in  
 ↳ practice (GRAFT-2  
 makes this explicit and pre-committed)
- ANDELOPS direction-of-absorption semantics may differ from the idea's  
 ↳ gloss - classification  
 routes through IDAN worker flows, so this only weakens the IDAS  
 ↳ cross-check, not the H2  
 machinery
- Spurious-closure events too heterogeneous for a clean placebo - a  
 ↳ non-inert placebo now  
 shows up as gate F2 firing (or exploratory H2\_diff reversal), which  
 ↳ caps H2's verdict by  
 pre-commitment instead of silently contaminating a differential test
- Realized OOH share below the published range - caught by the  
 ↳ published-range benchmark check  
 before estimation; H3 demotes or dies pre-outcome instead of running  
 ↳ underpowered
- Realized N lands below the full-expectation lower bound - the gate  
 ↳ demotes affected members

- to exploratory pre-outcome; the study survives with a smaller
  - ↳ confirmatory family (or none,
  - in which case it reports as registered exploratory estimation with the
  - ↳ demotion trail published
  - the first-round outcome, now a contingency rather than the design)
- Unobservable post-closure death/emigration deflates post-boundary
  - ↳ billing - direction stated,
  - sensitivity pre-specified, GRAFT-6 2pp trigger forces best/worst-case
  - ↳ bounds; no in-scope
  - fix exists
- Public/private sector misclassification via modal ARB\_SEKTORKODE for
  - ↳ multi-establishment
  - firms - attenuates contrasts; share bounded in the balance tables; if
  - ↳ ARB\_SEKTORKODE itself
  - proves absent in-environment (it is unconfirmed in the recorded
  - ↳ Forskningservice inventory),
  - the frozen ARB\_HOVED\_BRA\_DB07 fallback in contingent\_decisions
  - ↳ substitutes, with counts
  - reported
- Multi-job workers' primary-job assignment may misattribute exposure -
  - ↳ deterministic rule
  - biases toward null; share reported
- 1:3 matching without replacement may exhaust thin strata - realized  $k <$ 
  - ↳ 3 accounted honestly
  - in the blind re-simulation rather than assumed; coarsening ladder
  - ↳ pre-committed
- The SYSI-era exploratory replication may be infeasible - bounded as
  - ↳ skipable Stage-5 work
  - that cannot contaminate the confirmatory tier
- Readers may over-read gate F2 as proof of causal job-loss effects - the
  - ↳ associational ceiling,
  - the 'consistent with' vocabulary cap, and the pre-committed F2-fired
  - ↳ interpretation guard
  - the prose

## ## Source disclosure rules

```

```yaml
environment: DST Forskningservice (no prior data access; gatekeeper
↳ certification per RR
secondary-data rules). All analysis occurs inside the environment; only
↳ aggregate outputs
leave. No microdata, workplace identifiers (ARBNR/CVRNR/LBNR), or
↳ person-level values are
released.
minimum_cell_size: 5
minimum_cell_rules: no released statistic based on fewer than 5 persons
↳ or fewer than 3 workplaces;
event-count tables rounded to the nearest 10; weekly trajectory figures
↳ pool  $\geq$  100 events
and  $\geq$  5,000 person-weeks per plotted point; fee (BRUHON) aggregates
↳ subject to a 50% single-workplace
dominance check; min-cell suppression applied BEFORE Holm reporting so
↳ suppression can never
be outcome-selective
outputs_leaving_environment:
- model coefficients with SEs/CIs
- matching balance tables and realized-k distribution

```

- calibration and benchmark tables (coverage, linkage, HONUGE lag, OOH
 ↳ share)
- figure aggregates (weekly trajectory series under the pooling rules)
- event counts rounded to the nearest 10
- Stage-0 power/MDE tables - BOTH rounds (failed and remediated) - and
 ↳ the Stage-2 blind re-simulation output with the demotion trail
- the c-2 placebo, gate F2 outputs, GRAFT-2/4 diagnostics, GRAFT-6
 ↳ later-presence rates, E4 attrition panel, and all outcome-neutral check outputs (the full audit
 ↳ trail exits under the same rules)
- the analysis and power scripts

statement: 'All analysis occurs inside DST Forskningservice; only
 ↳ aggregate outputs leave.
 Rules applied to every released object: (1) no statistic based on fewer
 ↳ than 5 persons or fewer than 3 workplaces; (2) event-count tables rounded to the nearest
 ↳ 10; (3) weekly trajectory figures pool ≥ 100 events and $\geq 5,000$ person-weeks per plotted point;
 ↳ (4) fee (BRUHON) statistics released only as cell means/coefficients over min-N cells,
 ↳ with a dominance check that no single workplace contributes $> 50\%$ of any released fee
 ↳ aggregate; (5) released objects are limited to: model coefficients with SEs/CIs, balance tables,
 ↳ calibration/benchmark tables, figure aggregates, event counts (rounded), the Stage-0 (both rounds)
 ↳ and Stage-2 power/MDE tables with the demotion trail, and the analysis/power scripts - never
 ↳ microdata, workplace identifiers (ARBNR/CVRNR/LBNR), or person-level values; (6) min-cell
 ↳ suppression is applied before Holm reporting so that suppression can never be
 ↳ outcome-selective; (7) the c-2 placebo, gate F2 outputs, the GRAFT-2/4 diagnostics, the GRAFT-6 later-presence
 ↳ rates, the E4 attrition panel, and all outcome-neutral check outputs are released under the
 ↳ same rules, ensuring the full audit trail exits the environment.'

...

Pending registered experiment

No Danish study outcome has been observed for this bundle. See the complete registration above.

atom-000005: Postboundary billing

Do true closures show post-boundary billing excess, and how does the administrative-closure comparison constrain it?

Source fidelity and implementation order

This is a format conversion of DS1.1, not a new study or independent acceptance. Source audit: minor_revisions. No observed Danish outcomes are supplied. The source execution plan explicitly states that final prereg.yaml governs disagreements. Exact current YAML selections below therefore govern older execution-plan excerpts; these excerpts are retained as implementation detail, not a competing registration. See docs/provenance/source-discrepancies.md. The complete shared data/validity contracts are embedded in Atoms 1–3; read those native protocols when executing a downstream Atom, even without docs/. Do not substitute a scientific default for an unresolved source contradiction.

Preserve source stages 0 -> 1 -> 2 -> 3 -> 4 -> 5 inside the implementation. Source Stage-0 simulations are historical design evidence, not study outcomes; do not silently redo or improve them. Gate 14 must be evaluated from baseline data before post-baseline outcomes. All surviving Stage-4 quantities form ONE Holm family. The three question-level Atoms may share the same frozen family fit and tidy results; do not apply Holm separately per Atom or adjudicate a member before the complete surviving family's p-values exist. Cache/reuse identical governed intermediate results with provenance when available; a direct-dispatch call must obtain or compute its exact source-defined inputs, not pretend its predecessors succeeded. Stage-5 internal riders remain exploratory, after the Stage-4 family is frozen, even when housed in H1/H2.

Reference files under docs/execution/ are the ORIGINAL SKELETONS, not tested turnkey implementations. Adapt paths and complete engineering in a governed workspace; preserve source scientific rules and disclose unresolved choices. Write a result record with component estimates, uncertainty, family eligibility, source interpretation row, diagnostics and boundary/skip dispositions. Released outputs follow source disclosure rules; confidential rows stay governed. Local native states summarize the source judgment; neither a state name nor source power forecasts prove equivalence or causal identification. A bounded claim requires its actual interval and registered rule. Retain unresolved source wording conflicts; do not convert mere nonsignificance into scientific null.

Frozen scope and vocabulary

```
claim_class_ceiling: associational
statement: 'Associational - matched event-study contrasts, a contrastive
↪ falsification gate,
  and a composition contrast; descriptive for trajectory panels and the
↪ bounded priority claim.'
```

Binding vocabulary rules: 'is associated with' / 'consistent with' /
 ↳ 'suggests' only;
 'effect', 'impact', 'causes' prohibited throughout; anticipation
 ↳ always worded as
 'pre-boundary excess' relative to the annually dated closure-year
 ↳ boundary, never as announcement-dated
 anticipation (S8); no phenomenon-discovery or Danish-first framing (S1-S3);
 ↳ H2 novelty confined
 to the placebo's application to health billing (S7). H2 verdict semantics
 ↳ are fixed in
 advance: 'consistent with H2' requires BOTH the confirmatory core
 ↳ (theta2a Holm-significant,
 predicted direction) AND a passing falsification gate F2; the strict
 ↳ comparative magnitude
 is only ever reported as an exploratory estimate with its CI. Exploratory
 ↳ results (H1_trend,
 H2_diff, E1, E3, E4, and any gate-demoted member) carry the word
 ↳ 'exploratory' in the
 heading and in every sentence stating them. The prose label never exceeds
 ↳ the evidence label.'

frozen_parameters:
 seed: 20260831
 gp_code_list: daytime GP = SPEC2 in {70..80}
 ooh_code_list: out-of-hours GP = SPEC2 in {81, 82, 83, 84, 89}
 ooh_robustness_classifier: adds YDLTID in {2, 3, 4, 5, 7, 8, 9}
 spec2_evening_gp_disclosure: 'Registered choice, disclosed: the parsed
 ↳ value labels mark
 SPEC2 71 ("General practitioner, consultation, evening etc."), 73
 ↳ ("General practitioner,
 telephone consultation, evening etc."), and 75 ("General practitioner,
 ↳ home visit, evening
 etc.") as explicitly off-schedule GP services; they nonetheless remain on
 ↳ the daytime
 side of the PRIMARY classifier because H3's out-of-hours construct and
 ↳ its published
 power basis are keyed to the laegevagt out-of-hours service (SPEC2 81-84,
 ↳ 89), the quantity
 the published Danish 8-12% share describes. Attenuation direction, both
 ↳ readings: counting
 evening GP as daytime attenuates theta3 toward zero - conservative
 ↳ against the positive
 H3 conclusion, anti-conservative for the null-bounding "compositionally
 ↳ unremarkable"
 reading, whose Stage-2 statement therefore carries a mandatory caveat.
 ↳ In-environment
 validation must verify that the YDLTID-augmented robustness classifier
 ↳ captures SPEC2
 71/73/75 rows (their evening time codes 2-5); if YDLTID proves
 ↳ unpopulated on those rows,
 the robustness classifier - and only the robustness classifier - is
 ↳ amended mechanically
 and with disclosure to include SPEC2 71/73/75 directly. The SPEC2
 ↳ 71/73/75 volume share
 is reported, and both classifiers appear side by side in Table T7.'

successor_share_thresholds: 'true closure: ENDNEDL final closure AND
 ↳ worker-flow max-successor
 share $s(w) < 0.30$; spurious closure: forward identity via
 ↳ IDFREEM/REGELFR/FRELFREEM AND
 $s(w) \geq 0.70$; ambiguous mid-zone excluded; GRAFT-3 sensitivities
 ↳ 0.20/0.80 and 0.40/0.60'

```

quarter_mapping: 'ISO week of HONUGE: 1-13 / 14-26 / 27-39 / 40-52; ISO
↳ week 53 to Q4'
winsorization: person-quarter outcomes at the 99.9th percentile within era
↳ x calendar quarter
matching: 1:3 nearest-neighbor Mahalanobis within exact strata, without
↳ replacement across
  sets, target k = 3, minimum k = 1; calipers |dif baseline contacts| <= 3
  ↳ and |dif log(1+BRUHON)|
  <= 0.5; exposed processed in seeded random order (seed 20260831)
software: R >= 4.3, fixest::feols (OLS) and fixest::fepois (Poisson QML
↳ robustness); versions
  frozen in the registered script
note: Code lists and thresholds frozen before any data access; code labels
↳ quoted verbatim
  from parsed DST value labels and re-validated in-environment before
  ↳ estimation (frozen
  fallback rules in contingent_decisions).

```

4.2 theta2a — H2 confirmatory core, and gate F2

- **theta2a:** from the true-arm model of 4.1: mean of the four post-year coefficients, k in $\{Q1(c+1)..Q4(c+1)\}$ (guard band: Q2–Q4 only if step 19 triggered). Two-sided p , Holm member, support requires $\theta_{2a} > 0$; floor 2% of baseline.
- **Gate F2 (OUTSIDE the Holm family; consumes no alpha):** estimate the pooled model with event-class x event-quarter interactions (true + spurious arms, each with own matched controls); θ_{2b} = spurious-arm post-year mean excess. Pre-committed trigger: F2 FIRES iff θ_{2b} 's two-sided 95% CI excludes zero in the positive direction AND $\theta_{2b} \geq 50\%$ of θ_{2a} 's point estimate. If F2 fires, H2's comparative reading is disconfirmed regardless of θ_{2a} . ALWAYS report θ_{2b} 's one-sided 95% upper bound beside θ_{2a} . F2's registered operating characteristics are the Stage-0-verified values (pass 0.886 / inert-reference 0.978 / trigger 0.79 at 45,000 spurious sets), re-verified blind at realized N in step 14; with $< 15,000$ spurious sets F2 reports bounds only and the H2 verdict is capped at "partially consistent — comparative quantifier untestable".
- **Gated fee secondary:** BRUHON analogue of θ_{2a} , tested ONLY if θ_{2a} Holm-rejects; nominal fees, contrasts only within cohort x calendar-quarter cells (no deflation).
- **GRAFT-3:** re-estimate θ_{2b} and (exploratory) θ_{2_diff} under classification thresholds 0.20/0.80 and 0.40/0.60 alongside 0.30/0.70; one table; sign instability is flagged in the pre-committed H2 Summary paragraph; the 30/70 estimate remains primary; no threshold shopping.

4. ANALYSIS EXECUTION

Confirmatory tier first (Stage 4, frozen code, single run); exploratory tier after (Stage 5), segregated by heading. All estimation in R `fixest` via `Rscript`. Common estimator: outcome `y_iq` (winsorized quarterly GP contact count) on treated x event-quarter indicators for the 8 quarters $Q1(c)..Q4(c+1)$, matched-set fixed effects, cohort x calendar-quarter fixed effects, omitted baseline = the four quarters of year $c-1$; observation weights: exposed 1, controls $1/k_s$ within set; CR1 SEs clustered on event

workplace (each control inherits its set’s event cluster). Min-cell suppression is applied to all outputs BEFORE Holm reporting so suppression can never be outcome-selective.

4.4 Family adjudication and reporting order

Holm–Bonferroni over the surviving confirmatory family (as gated at step 14) at family $\alpha = 0.05$. Report all surviving tests exactly once, regardless of results. Each confirmatory hypothesis closes with a boldface **Summary (H_k)**. verdict paragraph opening “Consistent with” / “Contrary to our hypothesis”, per the pre-committed interpretation table in `prereg.yaml` (interpretations block — 16 rows covering positive, null, reversed, gate-fired, and partial patterns; apply them verbatim). Scoring attaches per the registered `scoring_rule`: H1 confirmed only if `theta1` Holm-rejects positively AND the gated `H1_trend` supports the growth rider; partially_confirmed for flat elevation; reversed if negative. H2 confirmed only if `theta2a` Holm-rejects positively AND F2 does not fire AND the `theta2b` sign pattern is GRAFT-3-stable; reversed if `theta2a` is significantly negative OR F2 fires. H3 confirmed on a positive Holm rejection (including the total-down/OOH-up pattern); reversed if negative. Demoted or killed members attach no score; their trail is disclosed. Sensitivity suite (run once, one table): GRAFT-3; GRAFT-9 (re-estimate `theta1/theta2a` excluding sets whose control workplace acquires any terminal event in `c+2` or `c+3` — cohorts `c <= 2022` for `c+2`, `c <= 2021` for `c+3`, truncation disclosed; registered direction note: exclusion should weakly INCREASE magnitudes, a decrease is flagged as diagnostic); un-winsorized outcomes; Poisson QML (`fepois`; if OLS and Poisson disagree in sign on any primary quantity, foreground the pre-committed OLS estimate, report both, flag the discordance in the Summary paragraph); possible-death/emigration exclusion (drop sets where the exposed member, or all controls, has zero SSSY rows across `c..c+1` AND no IDAN record in `c+1`).

4.5 Exploratory tier (Stage 5; heading-segregated; Holm within family as severity aid)

- **H1_trend (gated)**: `beta_Q4c – beta_Q1c` from the true-arm model, 95% CI, CR1 clusters — computed ONLY if `theta1` Holm-rejected positively; otherwise the quarterly trajectory is shown descriptively with no contrast statistic and no timing-shape language in the Summary.
- **H2_diff**: `theta2_diff = mean_k[beta_true,k – beta_spurious,k]` over the post-year quarters from the pooled model, 95% CI, plus the quarter-by-quarter differential profile (the “more persistent” rider). A flagged estimate with `spurious > true` triggers the pre-committed non-inert-placebo interpretation and reinforces whatever F2 showed.
- **E1**: SYSI-era analogue of `theta1` from step 22, 95% CI, descriptive replication only; never pooled across the vintage break.
- **E3**: BRUHON fee analogues of `theta2a/theta2_diff` beyond the gated secondary; heterogeneity of `theta1/theta2a` by sex (KOENIMP), age band (ALDERIMP), and baseline-year wage decile (BREDT_LOEN_BELOEB); Holm within this family.

- **E4:** arm-comparative attrition panel — share of each frozen November c-1 cohort no longer attached to the event workplace by November of year c (IDAN year-c records; ANSXFREM as the forward-separation flag), by event class; explicitly labeled post-treatment-conditioned descriptives with a selection caveat; no test.
- No exploratory result can revise any confirmatory verdict or Stage-4 language. Every sentence stating an exploratory result contains the word “exploratory”.

5. EXHIBITS

Every figure and table the report commits to. Disclosure rules binding on EVERY released object: no statistic based on fewer than 5 persons or fewer than 3 workplaces; event-count tables rounded to the nearest 10; weekly trajectory figures pool ≥ 100 events and $\geq 5,000$ person-weeks per plotted point; BRUHON aggregates pass a 50% single-workplace dominance check; released objects are limited to model coefficients with SEs/CIs, balance tables, calibration/benchmark tables, figure aggregates, rounded event counts, the power/MDE tables (both Stage-0 rounds and Stage 2) with the demotion trail, and the analysis/power scripts — never microdata, workplace identifiers (ARBNR/CVRNR/LBNR), or person-level values; suppression applied BEFORE Holm reporting.

Figures (matplotlib):

- **F1 — Quarterly event-study panel.** Treated x event-quarter coefficients with 95% CIs across Q1(c-2)..Q4(c+1) (placebo quarters, omitted-baseline marker, terminal-year and post-year quarters); panel
(a) true arm, panel (b) spurious arm from the pooled model. Feeds: step-18 panels + 4.1/4.2 coefficient tables. Disclosure: coefficients only.
- **F2 — Weekly descriptive trajectory.** Pooled exposed-minus- matched-control weekly GP contact series around the closure-year boundary, with the GRAFT-2 week-0 step diagnostic annotated (step $\leq 50\%$ of the total Q3(c)+Q4(c) excess; if exceeded, the figure carries the registered caution and quarter-level estimates are foregrounded). Feeds: step 15/17 weekly aggregates. Disclosure: ≥ 100 events and $\geq 5,000$ person-weeks per point.
- **F3 — s(w) successor-share distribution (GRAFT-4).** Histogram of worker-flow successor shares across terminal events with the 0.30/0.70 thresholds and mid-zone mass; bimodality check exhibit. Feeds: step 7. Disclosure: bin counts rounded to 10; bins under min-cell suppressed.
- **F4 — OOH vs daytime relative excess (theta3).** The two relative-excess series over the peri-event window with the bootstrap CI of their gap; primary and YDLTID-augmented classifier side by side. Feeds: 4.3 outputs. Disclosure: coefficients only.
- **F5 — Matching balance (love plot).** SMDs before/after matching for all matching covariates and unmatched auxiliaries; 0.10 line marked. Feeds: step 13. Disclosure: aggregates only.

- **F6 — E1 SYSI-era replication (exploratory; skippable).** Quarterly SYSI-era analogue of F1(a) for cohorts 1992–2003, labeled exploratory, YDLANT-weighted proxy outcome. Feeds: step 22. Disclosure: as F1.

Tables (booktabs LaTeX):

- **T1 — Event extraction and classification by cohort year.** Terminal events, true/spurious/ambiguous counts (rounded to 10), annual true-closure rate among private size-5+ workplaces, spurious/true ratio. Feeds: step 8.
- **T2 — Classification cross-validation.** Worker-flow classes vs IDAS-native flags (ENDNEDL/IDFREM/REGELFR/FRELFREM/ANDELUDS), full cross-tabulation + agreement rate; GRAFT-3 classification counts under 20/80 and 40/60. Feeds: step 8.
- **T3 — Cohort flow accounting.** Inclusion/exclusion counts per rule (tenure, sector, size, SIKGRUP, missing ANSAAR/pay, multi-cohort dedup), eligible exposed per cohort, match rate, ladder rung used, realized k distribution. Feeds: steps 9–13.
- **T4 — Matching balance.** Per-covariate SMDs (weighted 1/k), caliper compliance, realized mean k. Feeds: step 13.
- **T5 — Power and demotion trail.** Stage-0 power results (BOTH rounds: the voided round-1 run and the remediated re-run) and the Stage-2 blind realized-power/MDE table with binding N^* , realized sets, and any demotions with the fixed-point recomputation. Feeds: deposited script
– step 14. Published regardless of outcome.
- **T6 — Outcome-neutral check panel.** All eleven registered checks (coverage benchmark by year, linkage floor, closure-rate plausibility, classification agreement, GRAFT-4, balance/match rate, c-2 placebo, HONUUGE lag calibration, floor/ceiling + seasonality, GRAFT-2 step, OOH-share benchmark) with thresholds and PASS/FAIL. Feeds: steps 8, 13, 16, 17, 19, 21 + F2/F3 diagnostics.
- **T7 — Confirmatory results.** theta1, theta2a, theta3: point estimates (contacts/person-quarter and % of baseline), 95% CIs, two-sided p, Holm-adjusted verdicts; the F2 gate row (theta2b, its one-sided 95% upper bound, trigger evaluation); the gated BRUHON fee secondary (if computed); c-2 placebo row; realized MDEs beside every null with the CI-ruled-out magnitudes. Feeds: 4.1–4.4.
- **T8 — Sensitivity suite.** GRAFT-3 thresholds; GRAFT-9 forward-window exclusion (with truncation note); un-winsorized; Poisson QML; death/emigration exclusion — one table, run once. Feeds: 4.4.
- **T9 — Exploratory tier.** H1_trend (if gated in; otherwise the row states the gate was not met), H2_diff with quarter-by-quarter differential profile, E3 fee and subgroup analyses (within-family Holm), E1 summary. Every row labeled exploratory. Feeds: 4.5.
- **T10 — Attrition and later-presence accounting.** E4 year-c separation shares by event class (ANSXFREM-based, post-treatment- conditioned label and selection caveat) and the GRAFT-6 later-presence rates with the 2pp trigger status. Feeds: steps 20 + 4.5 E4.

Current source hypotheses (all fields)

```
- id: H2
tier: confirmatory
statement_initial: 'H2 (true vs spurious closure): Compared to workers
↳ whose workplace disappears
  only administratively - absorbed into a successor unit with a high
  ↳ absorbed-worker share
  (IDAS ANDELOPS) and preserved forward identity (IDFREM, REGELFR), so
  ↳ employment continues
  - workers at true permanent closures (ENDNEDL with low ANDELOPS) will
  ↳ show a larger and
  more persistent post-event excess in weekly contacts and gross fees
  ↳ (BRUHON) through the
  first post-closure year.'
```

statement_source: idea bank entry DS1.1, hypothesis 2 of 3; wording

```
↳ verbatim, frozen at
  bank creation. Confirmatory weight carried by theta2a (the true-arm
  ↳ post-year excess the
  statement asserts); the comparative quantifier is enforced by
  ↳ falsification gate F2 outside
  the Holm family and reported as exploratory H2_diff (see
  ↳ remediation_record); the wording
  itself is untouched.
```

estimand: 'Re-operationalization of the frozen H2 statement per the

```
↳ remediation directive:
  the statement's asserted excess is tested confirmatorily and its
  ↳ comparative quantifier
  is carried by a registered falsification gate. (a) CONFIRMATORY CORE
  ↳ theta2a: the post-boundary
  excess - mean over Q1(c+1)..Q4(c+1) of exposed-minus-matched-control
  ↳ quarterly GP contacts,
  relative to year c-1 baseline - for workers at true closures (identity
  ↳ disappears, ENDNEDL
  final-closure flag, worker-flow max-successor share  $s(w) < 0.30$ ,
  ↳ Hethy-Maier & Schmieder
  2013 / Benedetto et al. 2007 canon) vs their own 1:3 riskset-matched
  ↳ stable-workplace
  controls; 'through the first post-closure year' operationalized as the
  ↳ full four-quarter
  post-year window. (b) FALSIFICATION GATE F2 (severity device OUTSIDE the
  ↳ Holm family;
  not a power-bearing test): theta2b, the identical quantity for the
  ↳ spurious-closure arm
  (identity disappears but IDFREM/REGELFR/FRELFREM indicate en-bloc
  ↳ continuation and  $s(w)$ 
   $\geq 0.70$ , own matched controls), with the pre-committed disconfirmation
  ↳ trigger: if theta2b's
  95% CI excludes zero in the positive direction AND theta2b  $\geq 50\%$  of
  ↳ theta2a's point
  estimate, H2's comparative reading is disconfirmed regardless of
  ↳ theta2a's verdict;
  theta2b's one-sided 95% upper bound is always reported beside theta2a.
  ↳ GRAFT-3 successor-share
  threshold sensitivities (20/80, 40/60 alongside primary 30/70) retained.
  ↳ The strict comparative
  magnitude quantifier ('larger and more persistent post-event excess'
  ↳ than the spurious
  arm) is tiered exploratory as H2_diff. Novelty confined to the placebo's
  ↳ use for health
  billing (S7).'
```

primary_quantity: 'theta2a = mean_k[beta_true,k], k in {Q1(c+1)..Q4(c+1)}',
 ↳ from the true-arm
 stacked model; contacts per person-quarter and % of matched-control
 ↳ baseline; 95% CI,
 CR1 event-workplace clusters. Predicted direction theta2a > 0. Holm
 ↳ family member. Floor
 effect 2% relative post-year excess (0.022 contacts/person-quarter), set
 ↳ below Hamad/Modrek/Cullen
 (2015) and below post-displacement utilization/expenditure increases in
 ↳ Kuhn, Lalive &
 Zweimüller (2009, J Health Econ). Gross fees (BRUHON): same contrast,
 ↳ hierarchically gated
 secondary tested only if theta2a Holm-rejects. Power path: theta2a
 ↳ averages 4 post-year
 quarters (vs theta1's 2), so even with a registered serial-correlation
 ↳ discount its z
 exceeds theta1's at equal N; provisional gates N*_H2a = 60,000 true
 ↳ matched sets, and
 >= 15,000 spurious sets for F2's operating characteristics (VERIFIED
 ↳ Stage-0 values at
 lower-bound spurious N of 45,000 sets: gate pass probability 0.886 under
 ↳ the hypothesis-true
 DGP with spurious residual +0.5% of baseline, 0.978 under a strictly
 ↳ inert placebo; trigger
 probability 0.79 when theta2b = theta2a at floor. These verified values
 ↳ are the registered
 operating characteristics; the Stage-2 blind re-simulation at realized
 ↳ spurious N is the
 binding re-verification. F2 is a severity device outside the Holm family
 ↳ and carries no
 0.95 power requirement).'
 direction: 'positive (theta2a > 0: true-arm post-year excess), with gate F2
 ↳ required NOT
 to fire for the comparative reading'
 comparator: 'for theta2a: the true-closure workers'' own 1:3
 ↳ riskset-matched stable-workplace
 controls; the frozen statement's spurious-closure comparison is carried
 ↳ by gate F2 (theta2b,
 spurious arm vs its own matched controls) and exploratory H2_diff'
 exposure_variable: true permanent closure (ENDNEDL final closure AND s(w) <
 ↳ 0.30) versus
 spurious closure (forward identity via IDFREEM/REGELFR/FRELFREM AND s(w)
 ↳ >= 0.70) for the
 gate arm; ambiguous mid-zone excluded
 outcome_variable: quarterly GP contact counts, Q1(c+1)..Q4(c+1) versus c-1
 ↳ baseline; gross
 fees (BRUHON) as the hierarchically gated secondary
 forecast:
 point: +3.0% of matched-control baseline mean (approx. +0.033 contacts
 ↳ per person-quarter)
 interval_90: '[-1.0%, +8.0%] of baseline (approx. [-0.011, +0.088]
 ↳ contacts per person-quarter)'
 basis: 'Agent forecast, recorded before any data access and kept separate
 ↳ from any future
 expert forecasts (collaboration roadmap rule). Upward anchors:
 ↳ displaced workers face
 lower time costs of daytime care, and post-displacement increases in
 ↳ mental-health service
 and medication use are documented (Kuhn, Lalive & Zweimueller 2009 -
 ↳ modest overall,

concentrated in mental-health drugs; Modrek, Hamad & Cullen 2014).

- Downward anchors:
- the published Danish null on stress-related hospitalizations (Browning,
- Moller Dano
- & Heinesen 2006), US post-loss utilization drops (Schaller & Stevens
- 2015; largely an
- insurance channel absent in Denmark), and unobservable death/emigration
- deflating the
- post window. I judge the post-year excess more likely than the
- pre-boundary one. I also
- forecast gate F2 does NOT fire with probability ~0.75: reorganization
- stress makes a
- small positive spurious-arm excess plausible, but I expect it below
- half of theta2a.'

null_substantive: Workers at true closures show no post-boundary excess in

- primary-care
- billing through the first post-closure year relative to matched controls
- (post-year excess
- at or above 2% of baseline ruled out); together with theta2b near zero
- this locates any
- billing response, if theta1 found one, strictly before the boundary.

null_with_positive_placebo: 'PRE-COMMITTED (audit-added row, closes the

- non-exhaustive grid):
- if theta2a is not Holm-significant AND theta2b is flagged positive (95%
- CI excludes zero),
- we will conclude the placebo arm is not inert - the billing excess
- appears where employment
- continued under a new workplace identity - and read the pattern as
- evidence against closure-specific
- job-loss exposure driving billing, reporting both arms'' CI-ruled-out
- magnitudes; H2''s
- verdict is then "not supported, non-inert placebo", never partially
- supported.'

kill_criterion: 'SPURIOUS-ARM KILL (comparative machinery): if < 150

- spurious-closure events
- or < 3,000 spurious matched sets survive Stage 2, or worker-flow vs
- IDAS-flag agreement
- < 75%, or the GRAFT-4 bimodality check fails unresolved - gate F2,
- theta2b, and H2_diff
- are dropped (disclosed before outcome access) and the H2 verdict is
- capped at 'partially
- consistent - comparative quantifier untestable''. CONFIRMATORY DEMOTION
- GATE for theta2a:
- realized true matched sets < N*_H2a (provisional 60,000) or blind
- re-simulated power <
- 0.95 at the 2% floor -> theta2a demotes to exploratory pre-outcome. Gate
- F2 additionally
- requires >= 15,000 spurious sets for its registered operating
- characteristics; below that,
- F2 reports bounds only and the verdict cap applies.'

falsification_gate_F2: 'Registered severity device OUTSIDE the Holm family

- (consumes no
- alpha, can only subtract support): theta2b, the spurious-arm post-year
- excess, fires the
- pre-committed disconfirmation trigger iff its two-sided 95% CI excludes
- zero positively
- AND theta2b >= 50% of theta2a''s point estimate - disconfirming H2''s
- comparative reading
- regardless of theta2a. Verified operating characteristics at lower-bound
- spurious N (45,000

```

sets): pass 0.886 under the hypothesis-true DGP (spurious residual
↳ +0.5%), 0.978 under
a strictly inert placebo; trigger 0.79 when theta2b = theta2a at floor.
↳ These verified
values ARE the registered operating characteristics (aspirational >= 0.90
↳ / >= 0.95 targets
restated per the driver reconciliation); the Stage-2 blind re-simulation
↳ at realized spurious
N is the binding re-verification. F2 requires >= 15,000 spurious sets;
↳ below that it reports
bounds only and the H2 verdict cap applies.'
- id: H2_diff
tier: exploratory
statement_initial: '"... a larger and more persistent post-event excess in
↳ weekly contacts
and gross fees (BRUHON) through the first post-closure year ..."'
statement_source: the strict comparative-magnitude quantifier of frozen
↳ bank hypothesis
H2, excerpted verbatim; tiered exploratory because the
↳ true-minus-spurious differential
is verified structurally infeasible as a confirmatory test (variance
↳ dominated by the
spurious arm; ~5.5x any honest variance budget; round-1 power 0.358,
↳ remediated re-run
0.460).
estimand: 'The frozen H2 statement''s strict differential quantifier,
↳ tiered EXPLORATORY
because it is verified structurally infeasible as a confirmatory test:
↳ theta2_diff = true-minus-spurious
post-year differential, whose variance is dominated by the spurious arm
↳ and would need
~5.5x the variance budget of any honest spurious-arm N (Stage-0
↳ diagnosis; power 0.358
at the 2pp floor). Reported with 95% CI under the exploratory flagging
↳ convention, plus
the quarter-by-quarter differential profile across Q1(c+1)..Q4(c+1)
↳ addressing the statement''s
'more persistent'' rider.'
primary_quantity: theta2_diff = mean_k[beta_true,k - beta_spurious,k], k in
↳ {Q1(c+1)..Q4(c+1)},
from the pooled stacked model with event-class x event-quarter
↳ interactions; 95% CI; exploratory;
a flagged estimate in the NON-predicted direction (spurious > true)
↳ triggers the pre-committed
non-inert-placebo interpretation.
direction: positive (true > spurious) under the frozen rider; a flagged
↳ estimate in the
NON-predicted direction triggers the pre-committed non-inert-placebo
↳ interpretation

```

Current source interpretation rows

```

- hypothesis_id: H2
outcome_pattern: theta2a > 0, Holm-significant, AND gate F2 passes
↳ (theta2b's disconfirmation
trigger does not fire), sign stable across GRAFT-3 thresholds
conclusion: 'We will conclude that workers at true closures show a
↳ post-boundary billing
excess through the first post-closure year whose spurious-arm counterpart
↳ is bounded below

```

X% of the true-arm excess, where X is the realized one-sided 95% upper
 ↳ bound on θ_{2b}
 as a share of θ_{2a} -hat (the registered gate quantity; a passing gate
 ↳ bounds the spurious-arm
 excess - it cannot license an absence claim) - consistent with job-loss
 ↳ exposure rather
 than dying-firm selection driving the association (classification is
 ↳ measurement canon:
 Benedetto et al. 2007; Hethey-Maier & Schmieder 2013; our claim is only
 ↳ its placebo use
 for health billing). The strict comparative magnitude ('larger and more
 ↳ persistent')
 is reported from the exploratory H2_diff with its CI, never as a
 ↳ confirmed quantity.'

- hypothesis_id: H2
 outcome_pattern: $\theta_{2a} > 0$, Holm-significant, BUT gate F2 fires (θ_{2b}
 ↳ CI excludes zero
 positively AND $\theta_{2b} \geq 50\%$ of θ_{2a})
 conclusion: 'We will conclude that although true-closure workers show a
 ↳ post-year excess,
 the comparative reading of H2 is disconfirmed: the placebo is not inert,
 ↳ and the shared
 excess is consistent with selection into dying workplace identities or
 ↳ reorganization
 stress rather than job loss per se - an explicit caution for
 ↳ closure-based designs in
 the displacement literature. The θ_{1} interpretation is correspondingly
 ↳ weakened in
 the Discussion. This pattern is registered as disconfirming for H2's
 ↳ comparative claim
 regardless of θ_{2a} 's significance.'

- hypothesis_id: H2
 outcome_pattern: θ_{2a} not Holm-significant (NULL), θ_{2b} also near zero
 conclusion: We will conclude that we have no evidence of post-boundary
 ↳ excess primary-care
 billing after workplace terminal events of either kind, report the
 ↳ CI-ruled-out magnitudes
 for both arms, and note that this locates any billing response (per the
 ↳ θ_{1} result)
 before the boundary.

- hypothesis_id: H2
 outcome_pattern: $\theta_{2a} < 0$, Holm-significant (reversed)
 conclusion: We will conclude - contrary to our hypothesis - that
 ↳ true-closure workers' post-year
 contact falls relative to matched controls, consistent with
 ↳ out-of-population attrition
 (death/emigration invisible in scope; GRAFT-6 bounds mandatory) or
 ↳ post-displacement care
 deferral; disconfirming for H2 and reported with the bounding machinery,
 ↳ never reframed.

- hypothesis_id: H2_diff
 outcome_pattern: Exploratory flag with $\theta_{2_diff} < 0$ (spurious exceeds
 ↳ true)
 conclusion: We will report - contrary to the frozen comparative expectation
 ↳ - that spurious-closure
 workers show the larger excess, indicating the placebo is not inert
 ↳ (reorganizations may

carry their own insecurity exposure); this exploratory pattern reinforces
 ↳ whatever gate
 F2 showed and caps the interpretive weight the true/spurious contrast can
 ↳ bear.

Source family scoring and limitations

scoring_rule: 'Fixed in advance per the collaboration''s registered-report
 ↳ conventions (trajectory
 fields): at Stage 2 an outcome attaches to each confirmatory hypothesis''s
 ↳ frozen statement_initial
 as confirmed | partially_confirmed | reversed, adjudicated strictly against
 ↳ the pre-committed
 interpretation table in this protocol. H1 scores confirmed only if
 ↳ theta1''s Holm-adjusted
 two-sided test rejects with the point estimate in the predicted direction
 ↳ AND the gated
 exploratory H1_trend contrast supports the frozen statement''s growth rider
 ↳ (beta_Q4c -
 beta_Q1c > 0 with 95% CI excluding zero); it scores partially_confirmed
 ↳ when theta1 Holm-rejects
 in the predicted direction but the growth rider is unsupported (flat
 ↳ elevation - the pre-committed
 reading given annual event dating) or the trend gate is not computed;
 ↳ reversed if theta1
 is Holm-significant in the negative direction. H2 scores confirmed only if
 ↳ theta2a Holm-rejects
 in the predicted direction AND falsification gate F2 does not fire AND the
 ↳ theta2b sign
 pattern is stable across the GRAFT-3 successor-share thresholds;
 ↳ partially_confirmed when
 theta2a Holm-rejects but the comparative quantifier is untestable
 ↳ (spurious-arm kill - verdict
 capped by pre-commitment at ''partially consistent - comparative quantifier
 ↳ untestable'');
 reversed if theta2a is Holm-significant in the negative direction OR gate
 ↳ F2 fires (the
 frozen statement''s comparative reading is disconfirmed by a non-inert
 ↳ placebo regardless
 of theta2a - registered as disconfirming, never reframed). H3 scores
 ↳ confirmed only if theta3''s
 Holm-adjusted test rejects in the predicted direction (including the
 ↳ total-down/OOH-up pattern,
 pre-registered as scoreable support for the composition claim); reversed if
 ↳ Holm-significant
 in the negative direction. A Holm-nonsignificant confirmatory result
 ↳ attaches no confirmation:
 it is reported as a bounded informative null - the 95% CI converted into
 ↳ the excess magnitudes
 the data rule out against realized MDEs - never ''accepted''. A member
 ↳ demoted by the Stage-2
 realized-N gate or killed by its registered criterion attaches no score;
 ↳ the demotion/kill
 trail is disclosed. Significant effects in the non-predicted direction are
 ↳ disconfirming,
 never supportive. Exploratory quantities (H1_trend, H2_diff, E1, E3, E4,
 ↳ and any demoted
 member) attach outcomes only under the exploratory heading and are never
 ↳ scored as preregistered-confirmed.'

assumptions:

- SSSY (2005-2025) and SYSI (1990-2005) capture the universe of publicly
 - ↳ reimbursed primary-care billing for insured residents; absence of rows means no billed care, not missingness
- KONTAKT is a row-level contact marker whose sum yields contact counts
 - ↳ (in-environment validation with a frozen fallback definition if not)
- ENDNEDL/IDFREM/IDTILB/REGELFR/FRELFREM semantics follow Statistics
 - ↳ Denmark's IDA workplace-continuity documentation, cited directly (no DOI exists for these conventions, per boundary directive);
 - the primary classification rests on the worker-flow measure $s(w)$ from IDAN,
 - ↳ so misreading an IDAS flag degrades only the cross-check
- November-status annual attachment measures workplace membership; intra-year
 - ↳ attachment changes are unobservable ex ante and the frozen c-1 anchor is the design's answer
 - ↳ to that - with the corollary, named as a design threat, that early year-c separations feed realized post-job-loss billing into the pre-boundary window under intention-to-follow, a channel
 - ↳ the E4 separation shares and the registered theta1 attachment split bound but cannot remove
- ALDERIMP/KOENIMP (DST-imputed age/sex on SSSY rows) are accurate enough for
 - ↳ stratified matching; persons absent from SSSY in c-3..c-1 are handled by explicit unknown strata
- HONUGE settlement lags are predominantly ≤ 4 weeks (calibrated on
 - ↳ BEHANDLINGSDATO 2021+; guard-band contingency if not), and lag structure does not differ between
 - ↳ exposed and matched controls within the same calendar quarters
- Death and emigration are unobservable in scope; their post-event incidence
 - ↳ is assumed small; the deflating direction is conservative for the pre-boundary claim and
 - ↳ attenuating for theta2a, with the GRAFT-6 trigger bounding the post-boundary side
- Nominal-fee comparisons within cohort x calendar-quarter cells require no
 - ↳ deflator
- Danish institutional stability of GP billing categories (SPEC2 70-84, 89)
 - ↳ across 2007-2024 at the 2-digit level
- 'PUBLISHED-BASIS SIZING ASSUMPTIONS (new, each with a registered failure
 - ↳ path): (i) Danish establishment closure rates among private size-5+ workplaces are in the
 - ↳ published Erhvervsdemografi-consistent 1.5-3% band, supporting 150,000-300,000 exposed workers over 15 cohorts -
 - ↳ failure path: the pre-outcome realized-N gate demotes rather than stretches; (ii) the
 - ↳ Danish OOH share of GP contact volume is in the published 8-12% band (Christensen & Olesen
 - ↳ 1998; LV-KOS 2011 studies), simulated at the conservative 8% - failure path: the
 - ↳ published-range benchmark check plus realized-share re-simulation, then demotion or kill; (iii) the
 - ↳ control pool supports 1:3 matching at $\geq 90\%$ match rate - failure path: coarsening ladder,
 - ↳ realized-k accounting in the blind re-simulation, then the gate'
- theta2a's 4-quarter averaging gain over theta1 survives serial correlation
 - ↳ in post-year

shocks; a registered serial-correlation discount is built into the Stage-0
 ↳ re-simulation
 rather than assumed away

- The persistent-vs-transient overdispersion split (persistent frailty var
 ↳ 1.25, transient
 log-var 0.25) is retained unchanged from the first-round simulation -
 ↳ deliberately NOT revised
 in the direction that would raise power, so the remediation cannot be
 ↳ accused of tuning
 the DGP

failure_modes:

- 'Realized event or worker counts far below schema-derived expectations
 ↳ (documentation-only
 corpus; no realized Ns exist) - the design now bears this risk openly: the
 ↳ Stage-2 realized-N
 gate demotes underpowered members loudly and pre-outcome instead of letting
 ↳ an underpowered
 confirmatory test run; a sparse-closure Denmark kills or demotes cleanly
 ↳ rather than quietly'
- KONTAKT or ENDNEDL semantics differ from assumption - both carry frozen
 ↳ fallback rules,
 but level shifts in outcomes or event dating would force reported
 ↳ recalibration and consume
 credibility
- HONUUGE settlement lag materially longer than 4 weeks - the quarter-level
 ↳ estimands survive,
 but the weekly-resolution headline weakens to quarterly resolution in
 ↳ practice (GRAFT-2
 makes this explicit and pre-committed)
- ANDELOPS direction-of-absorption semantics may differ from the idea's gloss
 ↳ - classification
 routes through IDAN worker flows, so this only weakens the IDAS
 ↳ cross-check, not the H2
 machinery
- Spurious-closure events too heterogeneous for a clean placebo - a non-inert
 ↳ placebo now
 shows up as gate F2 firing (or exploratory H2_diff reversal), which caps
 ↳ H2's verdict by
 pre-commitment instead of silently contaminating a differential test
- Realized OOH share below the published range - caught by the
 ↳ published-range benchmark check
 before estimation; H3 demotes or dies pre-outcome instead of running
 ↳ underpowered
- Realized N lands below the full-expectation lower bound - the gate demotes
 ↳ affected members
 to exploratory pre-outcome; the study survives with a smaller confirmatory
 ↳ family (or none,
 in which case it reports as registered exploratory estimation with the
 ↳ demotion trail published
 - the first-round outcome, now a contingency rather than the design)
- Unobservable post-closure death/emigration deflates post-boundary billing -
 ↳ direction stated,
 sensitivity pre-specified, GRAFT-6 2pp trigger forces best/worst-case
 ↳ bounds; no in-scope
 fix exists
- Public/private sector misclassification via modal ARB_SEKTORKODE for
 ↳ multi-establishment
 firms - attenuates contrasts; share bounded in the balance tables; if
 ↳ ARB_SEKTORKODE itself

- proves absent in-environment (it is unconfirmed in the recorded
 - ↪ Forskningservice inventory),
- the frozen ARB_HOVED_BRA_DB07 fallback in contingent_decisions substitutes,
 - ↪ with counts reported
- Multi-job workers' primary-job assignment may misattribute exposure -
 - ↪ deterministic rule
 - biases toward null; share reported
- 1:3 matching without replacement may exhaust thin strata - realized $k < 3$
 - ↪ accounted honestly
 - in the blind re-simulation rather than assumed; coarsening ladder
 - ↪ pre-committed
- The SYSI-era exploratory replication may be infeasible - bounded as
 - ↪ skipable Stage-5 work
 - that cannot contaminate the confirmatory tier
- Readers may over-read gate F2 as proof of causal job-loss effects - the
 - ↪ associational ceiling,
 - the 'consistent with' vocabulary cap, and the pre-committed F2-fired
 - ↪ interpretation guard
- the prose

Source disclosure rules

environment: DST Forskningservice (no prior data access; gatekeeper

- ↪ certification per RR
- secondary-data rules). All analysis occurs inside the environment; only
 - ↪ aggregate outputs
- leave. No microdata, workplace identifiers (ARBNR/CVRNR/LBNR), or
 - ↪ person-level values are released.

minimum_cell_size: 5

minimum_cell_rules: no released statistic based on fewer than 5 persons or

- ↪ fewer than 3 workplaces;
 - event-count tables rounded to the nearest 10; weekly trajectory figures
 - ↪ pool ≥ 100 events
 - and $\geq 5,000$ person-weeks per plotted point; fee (BRUHON) aggregates
 - ↪ subject to a 50% single-workplace
- dominance check; min-cell suppression applied BEFORE Holm reporting so
 - ↪ suppression can never be outcome-selective

outputs_leaving_environment:

- model coefficients with SEs/CIs
- matching balance tables and realized-k distribution
- calibration and benchmark tables (coverage, linkage, HONUGE lag, OOH share)
- figure aggregates (weekly trajectory series under the pooling rules)
- event counts rounded to the nearest 10
- Stage-0 power/MDE tables - BOTH rounds (failed and remediated) - and the
 - ↪ Stage-2 blind re-simulation
- output with the demotion trail
- the c-2 placebo, gate F2 outputs, GRAFT-2/4 diagnostics, GRAFT-6
 - ↪ later-presence rates, E4
- attrition panel, and all outcome-neutral check outputs (the full audit
 - ↪ trail exits under the same rules)
- the analysis and power scripts

statement: 'All analysis occurs inside DST Forskningservice; only aggregate

- ↪ outputs leave.

Rules applied to every released object: (1) no statistic based on fewer

- ↪ than 5 persons or

fewer than 3 workplaces; (2) event-count tables rounded to the nearest 10;
 ↳ (3) weekly trajectory
 figures pool ≥ 100 events and $\geq 5,000$ person-weeks per plotted point; (4)
 ↳ fee (BRUHON)
 statistics released only as cell means/coefficients over min-N cells, with
 ↳ a dominance check
 that no single workplace contributes $> 50\%$ of any released fee aggregate;
 ↳ (5) released objects
 are limited to: model coefficients with SEs/CIs, balance tables,
 ↳ calibration/benchmark tables,
 figure aggregates, event counts (rounded), the Stage-0 (both rounds) and
 ↳ Stage-2 power/MDE
 tables with the demotion trail, and the analysis/power scripts - never
 ↳ microdata, workplace
 identifiers (ARBNR/CVRNR/LBNR), or person-level values; (6) min-cell
 ↳ suppression is applied
 before Holm reporting so that suppression can never be outcome-selective;
 ↳ (7) the c-2 placebo,
 gate F2 outputs, the GRAFT-2/4 diagnostics, the GRAFT-6 later-presence
 ↳ rates, the E4 attrition
 panel, and all outcome-neutral check outputs are released under the same
 ↳ rules, ensuring
 the full audit trail exits the environment.'

Complete registration fields

```
id: atom-000005
topic: Postboundary billing
atom_result: prereg
atom_strength: null
atom_shape: null
evidence_paths: []
concepts: []
literature: []
bonds: []
preregistration:
  version: 1
  mode: adaptive
  question: Do true closures show post-boundary billing excess, and how does
    the administrative-closure comparison constrain it?
  rationale: Source-existing coherent scientific question; its related
    ↳ outcomes,
    riders and sensitivity specifications remain internal.
  assumption: |-
```

Associational - matched event-study contrasts, a contrastive
 ↪ falsification gate, and a composition contrast; descriptive for
 ↪ trajectory panels and the bounded priority claim. Binding vocabulary
 ↪ rules: 'is associated with' / 'consistent with' / 'suggests' only;
 ↪ 'effect', 'impact', 'causes' prohibited throughout; anticipation
 ↪ always worded as 'pre-boundary excess' relative to the annually dated
 ↪ closure-year boundary, never as announcement-dated anticipation (S8);
 ↪ no phenomenon-discovery or Danish-first framing (S1-S3); H2 novelty
 ↪ confined to the placebo's application to health billing (S7). H2
 ↪ verdict semantics are fixed in advance: 'consistent with H2' requires
 ↪ BOTH the confirmatory core (theta2a Holm-significant, predicted
 ↪ direction) AND a passing falsification gate F2; the strict
 ↪ comparative magnitude is only ever reported as an exploratory
 ↪ estimate with its CI. Exploratory results (H1_trend, H2_diff, E1, E3,
 ↪ E4, and any gate-demoted member) carry the word 'exploratory' in the
 ↪ heading and in every sentence stating them. The prose label never
 ↪ exceeds the evidence label.
 Compatibility ratings below are neutral import metadata, not a new
 ↪ scientific evaluation or empirical scores.
 target_strength: associational
 target_shape: comparison
 ratings:
 importance: 3
 novelty: 3
 feasibility: 3
 data_advantage: 3
 informativeness: 0.5
 tables:
 - IDAN
 - IDAS
 - SSSY
 activation: state(atom-000003) == "event_study_ready"
 verdicts:
 - key: exploratory
 region: H2 confirmatory core was demoted pre-outcome; source-defined
 ↪ components
 are estimable.
 says: Report an exploratory core with all F2/arm restrictions; no
 ↪ confirmation.
 - key: comparative_disconfirmed
 region: Registered F2 trigger fires, regardless of the theta2a verdict.
 says: The comparative H2 reading is disconfirmed; record the core and
 exploratory differential separately.
 - key: reversed
 region: Eligible theta2a < 0 and Holm rejects, with F2 not already
 ↪ determining
 the first-matching state.
 says: Post-year billing declines; report attrition bounds and do not
 ↪ reframe
 as support.
 - key: comparative_unavailable
 region: Eligible theta2a > 0 and Holm rejects, but the registered
 ↪ comparative
 arm/gate is unavailable or bounds-only.
 says: At most partially consistent; comparative quantifier untestable.
 - key: postboundary_excess
 region: Eligible theta2a > 0 and Holm rejects, F2 is evaluable and does
 not fire, and registered threshold signs are stable.
 says: Quote the actual spurious-arm bound, never its absence. H2_diff
 remains exploratory.

- key: not_supported
 - region: Eligible core does not Holm-reject and no earlier region applies.
 - says: Report both arms with their actual bounds; retain the source
 - ↪ non-inert-placebo
 - row where applicable.
- key: boundary
 - region: Required source-defined inputs or scientific definitions are
 - ↪ demonstrably
 - unavailable; retain available components and the exact unresolved scope.
 - says: An evidenced boundary, not a statistical null. A missing
 - ↪ implementation
 - or timeout alone is operational no_result, not this state.
- key: otherwise
 - region: otherwise
 - says: The registered judgment is unresolved; report all completed
 - ↪ components
 - without inventing support, equivalence or absent results.

protocol: |

Do true closures show post-boundary billing excess, and how does the

↪ administrative-closure comparison constrain it?

Source fidelity and implementation order

This is a format conversion of DS1.1, not a new study or independent

↪ acceptance.

Source audit: minor_revisions. No observed Danish outcomes are supplied.

The source execution plan explicitly states that final prereg.yaml governs

disagreements. Exact current YAML selections below therefore govern older

execution-plan excerpts; these excerpts are retained as implementation

↪ detail,

not a competing registration. See docs/provenance/source-discrepancies.md.

The complete shared data/validity contracts are embedded in Atoms 1--3;

↪ read

those native protocols when executing a downstream Atom, even without

↪ docs/.

Do not substitute a scientific default for an unresolved source

↪ contradiction.

Preserve source stages 0 -> 1 -> 2 -> 3 -> 4 -> 5 inside the

↪ implementation.

Source Stage-0 simulations are historical design evidence, not study

↪ outcomes;

do not silently redo or improve them. Gate 14 must be evaluated from

↪ baseline

data before post-baseline outcomes. All surviving Stage-4 quantities form

↪ ONE

Holm family. The three question-level Atoms may share the same frozen

↪ family

fit and tidy results; do not apply Holm separately per Atom or adjudicate

↪ a

member before the complete surviving family's p-values exist. Cache/reuse

identical governed intermediate results with provenance when available; a

direct-dispatch call must obtain or compute its exact source-defined

↪ inputs,

not pretend its predecessors succeeded. Stage-5 internal riders remain

exploratory, after the Stage-4 family is frozen, even when housed in

↪ H1/H2.

Reference files under docs/execution/ are the ORIGINAL SKELETONS, not

↪ tested

turnkey implementations. Adapt paths and complete engineering in a
 ↳ governed
 workspace; preserve source scientific rules and disclose unresolved
 ↳ choices.
 Write a result record with component estimates, uncertainty, family
 ↳ eligibility,
 source interpretation row, diagnostics and boundary/skip dispositions.
 ↳ Released
 outputs follow source disclosure rules; confidential rows stay governed.
 Local native states summarize the source judgment; neither a state name
 ↳ nor
 source power forecasts prove equivalence or causal identification. A
 ↳ bounded
 claim requires its actual interval and registered rule. Retain unresolved
 ↳ source
 wording conflicts; do not convert mere nonsignificance into scientific
 ↳ null.

Frozen scope and vocabulary

```
```yaml
claim_class_ceiling: associational
statement: 'Associational - matched event-study contrasts, a contrastive
↳ falsification gate,
 and a composition contrast; descriptive for trajectory panels and the
 ↳ bounded priority claim.
Binding vocabulary rules: 'is associated with' / 'consistent with'
↳ / 'suggests' only;
'effect', 'impact', 'causes' prohibited throughout; anticipation
↳ always worded as
'pre-boundary excess' relative to the annually dated closure-year
↳ boundary, never as announcement-dated
anticipation (S8); no phenomenon-discovery or Danish-first framing
↳ (S1-S3); H2 novelty confined
to the placebo's application to health billing (S7). H2 verdict
↳ semantics are fixed in
advance: 'consistent with H2' requires BOTH the confirmatory core
↳ (theta2a Holm-significant,
predicted direction) AND a passing falsification gate F2; the strict
↳ comparative magnitude
is only ever reported as an exploratory estimate with its CI.
↳ Exploratory results (H1_trend,
H2_diff, E1, E3, E4, and any gate-demoted member) carry the word
↳ 'exploratory' in the
heading and in every sentence stating them. The prose label never
↳ exceeds the evidence label.'
```

frozen\_parameters:

```
seed: 20260831
gp_code_list: daytime GP = SPEC2 in {70..80}
ooh_code_list: out-of-hours GP = SPEC2 in {81, 82, 83, 84, 89}
ooh_robustness_classifier: adds YDLTID in {2, 3, 4, 5, 7, 8, 9}
spec2_evening_gp_disclosure: 'Registered choice, disclosed: the parsed
↳ value labels mark
 SPEC2 71 ("General practitioner, consultation, evening etc."), 73
 ↳ ("General practitioner,
 telephone consultation, evening etc."), and 75 ("General
 ↳ practitioner, home visit, evening
 etc.") as explicitly off-schedule GP services; they nonetheless
 ↳ remain on the daytime
```

side of the PRIMARY classifier because H3's out-of-hours construct  
 ↳ and its published  
 power basis are keyed to the laegevagt out-of-hours service (SPEC2  
 ↳ 81-84, 89), the quantity  
 the published Danish 8-12% share describes. Attenuation direction,  
 ↳ both readings: counting  
 evening GP as daytime attenuates theta3 toward zero - conservative  
 ↳ against the positive  
 H3 conclusion, anti-conservative for the null-bounding  
 ↳ "compositionally unremarkable"  
 reading, whose Stage-2 statement therefore carries a mandatory  
 ↳ caveat. In-environment  
 validation must verify that the YDLTID-augmented robustness  
 ↳ classifier captures SPEC2  
 71/73/75 rows (their evening time codes 2-5); if YDLTID proves  
 ↳ unpopulated on those rows,  
 the robustness classifier - and only the robustness classifier - is  
 ↳ amended mechanically  
 and with disclosure to include SPEC2 71/73/75 directly. The SPEC2  
 ↳ 71/73/75 volume share  
 is reported, and both classifiers appear side by side in Table T7.'  
 successor\_share\_thresholds: 'true closure: ENDNEDL final closure AND  
 ↳ worker-flow max-successor  
 share  $s(w) < 0.30$ ; spurious closure: forward identity via  
 ↳ IDFREEM/REGELFR/FRELFREM AND  
 $s(w) \geq 0.70$ ; ambiguous mid-zone excluded; GRAFT-3 sensitivities  
 ↳ 0.20/0.80 and 0.40/0.60'  
 quarter\_mapping: 'ISO week of HONUGE: 1-13 / 14-26 / 27-39 / 40-52; ISO  
 ↳ week 53 to Q4'  
 winsorization: person-quarter outcomes at the 99.9th percentile within  
 ↳ era x calendar quarter  
 matching: 1:3 nearest-neighbor Mahalanobis within exact strata, without  
 ↳ replacement across  
 sets, target  $k = 3$ , minimum  $k = 1$ ; calipers  $|\text{dif baseline contacts}|$   
 ↳  $\leq 3$  and  $|\text{dif log}(1+\text{BRUHON})|$   
 $\leq 0.5$ ; exposed processed in seeded random order (seed 20260831)  
 software: R  $\geq 4.3$ , fixest::feols (OLS) and fixest::fepois (Poisson QML  
 ↳ robustness); versions  
 frozen in the registered script  
 note: Code lists and thresholds frozen before any data access; code  
 ↳ labels quoted verbatim  
 from parsed DST value labels and re-validated in-environment before  
 ↳ estimation (frozen  
 fallback rules in contingent\_decisions).  
 ...

### 4.2 theta2a - H2 confirmatory core, and gate F2

- \*\*theta2a:\*\* from the true-arm model of 4.1: mean of the four post-year coefficients,  $k$  in  $\{Q1(c+1)..Q4(c+1)\}$  (guard band: Q2-Q4 only if step 19 triggered). Two-sided  $p$ , Holm member, support requires  $\text{theta2a} > 0$ ; floor 2% of baseline.
- \*\*Gate F2 (OUTSIDE the Holm family; consumes no alpha):\*\* estimate the pooled model with event-class x event-quarter interactions (true + spurious arms, each with own matched controls);  $\text{theta2b}$  = spurious-arm post-year mean excess. Pre-committed trigger: F2 FIRES iff  $\text{theta2b}$ 's two-sided 95% CI excludes zero in the positive direction AND  $\text{theta2b} \geq 50\%$  of  $\text{theta2a}$ 's point estimate. If F2 fires, H2's comparative reading is disconfirmed regardless of  $\text{theta2a}$ . ALWAYS report  $\text{theta2b}$ 's one-sided 95% upper bound beside  $\text{theta2a}$ . F2's registered operating characteristics are the Stage-0-verified values (pass 0.886 /

- inert-reference 0.978 / trigger 0.79 at 45,000 spurious sets), re-verified blind at realized N in step 14; with < 15,000 spurious sets F2 reports bounds only and the H2 verdict is capped at "partially consistent - comparative quantifier untestable".
- **\*\*Gated fee secondary:\*\*** BRUHON analogue of theta2a, tested ONLY if theta2a Holm-rejects; nominal fees, contrasts only within cohort x calendar-quarter cells (no deflation).
- **\*\*GRAFT-3:\*\*** re-estimate theta2b and (exploratory) theta2\_diff under classification thresholds 0.20/0.80 and 0.40/0.60 alongside 0.30/0.70; one table; sign instability is flagged in the pre-committed H2 Summary paragraph; the 30/70 estimate remains primary; no threshold shopping.

#### ## 4. ANALYSIS EXECUTION

Confirmatory tier first (Stage 4, frozen code, single run); exploratory tier after (Stage 5), segregated by heading. All estimation in R fixest via Rscript. Common estimator: outcome y\_iq (winsorized quarterly GP contact count) on treated x event-quarter indicators for the 8 quarters Q1(c)..Q4(c+1), matched-set fixed effects, cohort x calendar-quarter fixed effects, omitted baseline = the four quarters of year c-1; observation weights: exposed 1, controls 1/k\_s within set; CR1 SEs clustered on event workplace (each control inherits its set's event cluster). Min-cell suppression is applied to all outputs BEFORE Holm reporting so suppression can never be outcome-selective.

##### ### 4.4 Family adjudication and reporting order

Holm-Bonferroni over the surviving confirmatory family (as gated at step 14) at family alpha = 0.05. Report all surviving tests exactly once, regardless of results. Each confirmatory hypothesis closes with a boldface **\*\*Summary (H\_k):\*\*** verdict paragraph opening "Consistent with" / "Contrary to our hypothesis", per the pre-committed interpretation table in `prereg.yaml` (interpretations block - 16 rows covering positive, null, reversed, gate-fired, and partial patterns; apply them verbatim). Scoring attaches per the registered scoring\_rule: H1 confirmed only if theta1 Holm-rejects positively AND the gated H1\_trend supports the growth rider; partially\_confirmed for flat elevation; reversed if negative. H2 confirmed only if theta2a Holm-rejects positively AND F2 does not fire AND the theta2b sign pattern is GRAFT-3-stable; reversed if theta2a is significantly negative OR F2 fires. H3 confirmed on a positive Holm rejection (including the total-down/OOH-up pattern); reversed if negative. Demoted or killed members attach no score; their trail is disclosed. Sensitivity suite (run once, one table): GRAFT-3; GRAFT-9 (re-estimate theta1/theta2a excluding sets whose control workplace acquires any terminal event in c+2 or c+3 - cohorts c <= 2022 for c+2, c <= 2021 for c+3, truncation disclosed; registered direction note: exclusion should weakly INCREASE magnitudes, a decrease is flagged as diagnostic); un-winsorized outcomes; Poisson QML (`fepois`); if OLS and Poisson disagree in sign on any primary quantity, foreground the pre-committed OLS estimate, report both, flag the discordance in the Summary paragraph; possible-death/emigration exclusion (drop sets where the exposed member, or all controls, has zero SSSY rows across c..c+1 AND no IDAN record in c+1).

##### ### 4.5 Exploratory tier (Stage 5; heading-segregated; Holm within family ↳ as severity aid)

- **\*\*H1\_trend (gated):\*\*** beta\_Q4c - beta\_Q1c from the true-arm model, 95% CI, CR1 clusters - computed ONLY if theta1 Holm-rejected positively; otherwise the quarterly trajectory is shown descriptively with no

- contrast statistic and no timing-shape language in the Summary.
- **\*\*H2\_diff:\*\*** theta2\_diff = mean\_k[beta\_true,k - beta\_spurious,k] over the post-year quarters from the pooled model, 95% CI, plus the quarter-by-quarter differential profile (the "more persistent" rider). A flagged estimate with spurious > true triggers the pre-committed non-inert-placebo interpretation and reinforces whatever F2 showed.
- **\*\*E1:\*\*** SYSI-era analogue of theta1 from step 22, 95% CI, descriptive replication only; never pooled across the vintage break.
- **\*\*E3:\*\*** BRUHON fee analogues of theta2a/theta2\_diff beyond the gated secondary; heterogeneity of theta1/theta2a by sex (KOENIMP), age band (ALDERIMP), and baseline-year wage decile (BREDT\_LOEN\_BELOEB); Holm within this family.
- **\*\*E4:\*\*** arm-comparative attrition panel - share of each frozen November c-1 cohort no longer attached to the event workplace by November of year c (IDAN year-c records; ANSXFREM as the forward-separation flag), by event class; explicitly labeled post-treatment-conditioned descriptives with a selection caveat; no test.
- No exploratory result can revise any confirmatory verdict or Stage-4 language. Every sentence stating an exploratory result contains the word "exploratory".

## ## 5. EXHIBITS

Every figure and table the report commits to. Disclosure rules binding on EVERY released object: no statistic based on fewer than 5 persons or fewer than 3 workplaces; event-count tables rounded to the nearest 10; weekly trajectory figures pool  $\geq 100$  events and  $\geq 5,000$  person-weeks per plotted point; BRUHON aggregates pass a 50% single-workplace dominance check; released objects are limited to model coefficients with SEs/CIs, balance tables, calibration/benchmark tables, figure aggregates, rounded event counts, the power/MDE tables (both Stage-0 rounds and Stage 2) with the demotion trail, and the analysis/power scripts - never microdata, workplace identifiers (ARENR/CVRNR/LBNR), or person-level values; suppression applied BEFORE Holm reporting.

### **\*\*Figures (matplotlib):\*\***

- **\*\*F1 - Quarterly event-study panel.\*\*** Treated x event-quarter coefficients with 95% CIs across Q1(c-2)..Q4(c+1) (placebo quarters, omitted-baseline marker, terminal-year and post-year quarters); panel (a) true arm, panel (b) spurious arm from the pooled model. Feeds: step-18 panels + 4.1/4.2 coefficient tables. Disclosure: coefficients only.
- **\*\*F2 - Weekly descriptive trajectory.\*\*** Pooled exposed-minus-matched-control weekly GP contact series around the closure-year boundary, with the GRAFT-2 week-0 step diagnostic annotated (step  $\leq 50\%$  of the total Q3(c)+Q4(c) excess; if exceeded, the figure carries the registered caution and quarter-level estimates are foregrounded). Feeds: step 15/17 weekly aggregates. Disclosure:  $\geq 100$  events and  $\geq 5,000$  person-weeks per point.
- **\*\*F3 - s(w) successor-share distribution (GRAFT-4).\*\*** Histogram of worker-flow successor shares across terminal events with the 0.30/0.70 thresholds and mid-zone mass; bimodality check exhibit. Feeds: step 7. Disclosure: bin counts rounded to 10; bins under min-cell suppressed.
- **\*\*F4 - OOH vs daytime relative excess (theta3).\*\*** The two relative-excess series over the peri-event window with the bootstrap CI of their gap; primary and YDLTID-augmented classifier side by side. Feeds: 4.3 outputs. Disclosure: coefficients only.
- **\*\*F5 - Matching balance (love plot).\*\*** SMDs before/after matching for

all matching covariates and unmatched auxiliaries; 0.10 line marked.  
Feeds: step 13. Disclosure: aggregates only.

- \*\*F6 - E1 SYSI-era replication (exploratory; skippable).\*\* Quarterly SYSI-era analogue of F1(a) for cohorts 1992-2003, labeled exploratory, YDLANT-weighted proxy outcome. Feeds: step 22. Disclosure: as F1.

**\*\*Tables (booktabs LaTeX):\*\***

- \*\*T1 - Event extraction and classification by cohort year.\*\* Terminal events, true/spurious/ambiguous counts (rounded to 10), annual true-closure rate among private size-5+ workplaces, spurious/true ratio. Feeds: step 8.
- \*\*T2 - Classification cross-validation.\*\* Worker-flow classes vs IDAS-native flags (ENDNEDL/IDFREM/REGELFR/FRELFREM/ANDELUDS), full cross-tabulation + agreement rate; GRAFT-3 classification counts under 20/80 and 40/60. Feeds: step 8.
- \*\*T3 - Cohort flow accounting.\*\* Inclusion/exclusion counts per rule (tenure, sector, size, SIKGRUP, missing ANSAAR/pay, multi-cohort dedup), eligible exposed per cohort, match rate, ladder rung used, realized k distribution. Feeds: steps 9-13.
- \*\*T4 - Matching balance.\*\* Per-covariate SMDs (weighted 1/k), caliper compliance, realized mean k. Feeds: step 13.
- \*\*T5 - Power and demotion trail.\*\* Stage-0 power results (BOTH rounds: the voided round-1 run and the remediated re-run) and the Stage-2 blind realized-power/MDE table with binding N\*, realized sets, and any demotions with the fixed-point recomputation. Feeds: deposited script + step 14. Published regardless of outcome.
- \*\*T6 - Outcome-neutral check panel.\*\* All eleven registered checks (coverage benchmark by year, linkage floor, closure-rate plausibility, classification agreement, GRAFT-4, balance/match rate, c-2 placebo, HONUGE lag calibration, floor/ceiling + seasonality, GRAFT-2 step, OOH-share benchmark) with thresholds and PASS/FAIL. Feeds: steps 8, 13, 16, 17, 19, 21 + F2/F3 diagnostics.
- \*\*T7 - Confirmatory results.\*\* theta1, theta2a, theta3: point estimates (contacts/person-quarter and % of baseline), 95% CIs, two-sided p, Holm-adjusted verdicts; the F2 gate row (theta2b, its one-sided 95% upper bound, trigger evaluation); the gated BRUHON fee secondary (if computed); c-2 placebo row; realized MDEs beside every null with the CI-ruled-out magnitudes. Feeds: 4.1-4.4.
- \*\*T8 - Sensitivity suite.\*\* GRAFT-3 thresholds; GRAFT-9 forward-window exclusion (with truncation note); un-winsorized; Poisson QML; death/emigration exclusion - one table, run once. Feeds: 4.4.
- \*\*T9 - Exploratory tier.\*\* H1\_trend (if gated in; otherwise the row states the gate was not met), H2\_diff with quarter-by-quarter differential profile, E3 fee and subgroup analyses (within-family Holm), E1 summary. Every row labeled exploratory. Feeds: 4.5.
- \*\*T10 - Attrition and later-presence accounting.\*\* E4 year-c separation shares by event class (ANSXFREM-based, post-treatment-conditioned label and selection caveat) and the GRAFT-6 later-presence rates with the 2pp trigger status. Feeds: steps 20 + 4.5 E4.

---

**## Current source hypotheses (all fields)**

```

```yaml
- id: H2
  tier: confirmatory

```

```

statement_initial: 'H2 (true vs spurious closure): Compared to workers
→ whose workplace disappears
  only administratively - absorbed into a successor unit with a high
  → absorbed-worker share
  (IDAS ANDELOPS) and preserved forward identity (IDFREM, REGELFR), so
  → employment continues
- workers at true permanent closures (ENDNEDL with low ANDELOPS) will
  → show a larger and
  more persistent post-event excess in weekly contacts and gross fees
  → (BRUHON) through the
  first post-closure year.'
statement_source: idea bank entry DS1.1, hypothesis 2 of 3; wording
→ verbatim, frozen at
  bank creation. Confirmatory weight carried by theta2a (the true-arm
  → post-year excess the
  statement asserts); the comparative quantifier is enforced by
  → falsification gate F2 outside
  the Holm family and reported as exploratory H2_diff (see
  → remediation_record); the wording
  itself is untouched.
estimand: 'Re-operationalization of the frozen H2 statement per the
→ remediation directive:
  the statement's asserted excess is tested confirmatorily and its
  → comparative quantifier
  is carried by a registered falsification gate. (a) CONFIRMATORY CORE
  → theta2a: the post-boundary
  excess - mean over Q1(c+1)..Q4(c+1) of exposed-minus-matched-control
  → quarterly GP contacts,
  relative to year c-1 baseline - for workers at true closures
  → (identity disappears, ENDNEDL
  final-closure flag, worker-flow max-successor share  $s(w) < 0.30$ ,
  → Hethy-Maier & Schmieder
  2013 / Benedetto et al. 2007 canon) vs their own 1:3 riskset-matched
  → stable-workplace
  controls; 'through the first post-closure year' operationalized as
  → the full four-quarter
  post-year window. (b) FALSIFICATION GATE F2 (severity device OUTSIDE
  → the Holm family;
  not a power-bearing test): theta2b, the identical quantity for the
  → spurious-closure arm
  (identity disappears but IDFREM/REGELFR/FRELFREM indicate en-bloc
  → continuation and  $s(w)$ 
   $\geq 0.70$ , own matched controls), with the pre-committed
  → disconfirmation trigger: if theta2b's
  95% CI excludes zero in the positive direction AND theta2b  $\geq 50\%$  of
  → theta2a's point
  estimate, H2's comparative reading is disconfirmed regardless of
  → theta2a's verdict;
  theta2b's one-sided 95% upper bound is always reported beside
  → theta2a. GRAFT-3 successor-share
  threshold sensitivities (20/80, 40/60 alongside primary 30/70)
  → retained. The strict comparative
  magnitude quantifier ('larger and more persistent post-event
  → excess' than the spurious
  arm) is tiered exploratory as H2_diff. Novelty confined to the
  → placebo's use for health
  billing (S7).'
primary_quantity: 'theta2a = mean_k[beta_true,k], k in
→ {Q1(c+1)..Q4(c+1)}, from the true-arm

```

stacked model; contacts per person-quarter and % of matched-control
 ↳ baseline; 95% CI,
 CR1 event-workplace clusters. Predicted direction $\theta_{2a} > 0$. Holm
 ↳ family member. Floor
 effect 2% relative post-year excess (0.022 contacts/person-quarter),
 ↳ set below Hamad/Modrek/Cullen
 (2015) and below post-displacement utilization/expenditure increases
 ↳ in Kuhn, Lalive &
 Zweimüller (2009, J Health Econ). Gross fees (BRUHON): same contrast,
 ↳ hierarchically gated
 secondary tested only if θ_{2a} Holm-rejects. Power path: θ_{2a}
 ↳ averages 4 post-year
 quarters (vs θ_{1a} 's 2), so even with a registered
 ↳ serial-correlation discount its z
 exceeds θ_{1a} 's at equal N; provisional gates $N \cdot H_{2a} = 60,000$ true
 ↳ matched sets, and
 $\geq 15,000$ spurious sets for F2's operating characteristics (VERIFIED
 ↳ Stage-0 values at
 lower-bound spurious N of 45,000 sets: gate pass probability 0.886
 ↳ under the hypothesis-true
 DGP with spurious residual +0.5% of baseline, 0.978 under a strictly
 ↳ inert placebo; trigger
 probability 0.79 when $\theta_{2b} = \theta_{2a}$ at floor. These verified
 ↳ values are the registered
 operating characteristics; the Stage-2 blind re-simulation at
 ↳ realized spurious N is the
 binding re-verification. F2 is a severity device outside the Holm
 ↳ family and carries no
 0.95 power requirement).'
 direction: 'positive ($\theta_{2a} > 0$: true-arm post-year excess), with
 ↳ gate F2 required NOT
 to fire for the comparative reading'
 comparator: 'for θ_{2a} : the true-closure workers'' own 1:3
 ↳ riskset-matched stable-workplace
 controls; the frozen statement's spurious-closure comparison is
 ↳ carried by gate F2 (θ_{2b} ,
 spurious arm vs its own matched controls) and exploratory H_{2_diff} '
 exposure_variable: true permanent closure (ENDNEDL final closure AND
 ↳ $s(w) < 0.30$) versus
 spurious closure (forward identity via IDFRE/REGELFR/FRELFRE AND
 ↳ $s(w) \geq 0.70$) for the
 gate arm; ambiguous mid-zone excluded
 outcome_variable: quarterly GP contact counts, $Q_1(c+1) \dots Q_4(c+1)$ versus
 ↳ c-1 baseline; gross
 fees (BRUHON) as the hierarchically gated secondary
 forecast:
 point: +3.0% of matched-control baseline mean (approx. +0.033
 ↳ contacts per person-quarter)
 interval_90: '[-1.0%, +8.0%] of baseline (approx. [-0.011, +0.088]
 ↳ contacts per person-quarter)'
 basis: 'Agent forecast, recorded before any data access and kept
 ↳ separate from any future
 expert forecasts (collaboration roadmap rule). Upward anchors:
 ↳ displaced workers face
 lower time costs of daytime care, and post-displacement increases
 ↳ in mental-health service
 and medication use are documented (Kuhn, Lalive & Zweimueller 2009
 ↳ - modest overall,
 concentrated in mental-health drugs; Modrek, Hamad & Cullen 2014).
 ↳ Downward anchors:

the published Danish null on stress-related hospitalizations
 ↪ (Browning, Moller Dano
 & Heinesen 2006), US post-loss utilization drops (Schaller &
 ↪ Stevens 2015; largely an
 insurance channel absent in Denmark), and unobservable
 ↪ death/emigration deflating the
 post window. I judge the post-year excess more likely than the
 ↪ pre-boundary one. I also
 forecast gate F2 does NOT fire with probability ~0.75:
 ↪ reorganization stress makes a
 small positive spurious-arm excess plausible, but I expect it below
 ↪ half of theta2a.'
 null_substantive: Workers at true closures show no post-boundary excess
 ↪ in primary-care
 billing through the first post-closure year relative to matched
 ↪ controls (post-year excess
 at or above 2% of baseline ruled out); together with theta2b near
 ↪ zero this locates any
 billing response, if theta1 found one, strictly before the boundary.
 null_with_positive_placebo: 'PRE-COMMITTED (audit-added row, closes the
 ↪ non-exhaustive grid):
 if theta2a is not Holm-significant AND theta2b is flagged positive
 ↪ (95% CI excludes zero),
 we will conclude the placebo arm is not inert - the billing excess
 ↪ appears where employment
 continued under a new workplace identity - and read the pattern as
 ↪ evidence against closure-specific
 job-loss exposure driving billing, reporting both arms'' CI-ruled-out
 ↪ magnitudes; H2''s
 verdict is then "not supported, non-inert placebo", never partially
 ↪ supported.'
 kill_criterion: 'SPURIOUS-ARM KILL (comparative machinery): if < 150
 ↪ spurious-closure events
 or < 3,000 spurious matched sets survive Stage 2, or worker-flow vs
 ↪ IDAS-flag agreement
 < 75%, or the GRAFT-4 bimodality check fails unresolved - gate F2,
 ↪ theta2b, and H2_diff
 are dropped (disclosed before outcome access) and the H2 verdict is
 ↪ capped at 'partially
 consistent - comparative quantifier untestable''. CONFIRMATORY
 ↪ DEMOTION GATE for theta2a:
 realized true matched sets < N*_H2a (provisional 60,000) or blind
 ↪ re-simulated power <
 0.95 at the 2% floor -> theta2a demotes to exploratory pre-outcome.
 ↪ Gate F2 additionally
 requires >= 15,000 spurious sets for its registered operating
 ↪ characteristics; below that,
 F2 reports bounds only and the verdict cap applies.'
 falsification_gate_F2: 'Registered severity device OUTSIDE the Holm
 ↪ family (consumes no
 alpha, can only subtract support): theta2b, the spurious-arm
 ↪ post-year excess, fires the
 pre-committed disconfirmation trigger iff its two-sided 95% CI
 ↪ excludes zero positively
 AND theta2b >= 50% of theta2a''s point estimate - disconfirming H2''s
 ↪ comparative reading
 regardless of theta2a. Verified operating characteristics at
 ↪ lower-bound spurious N (45,000
 sets): pass 0.886 under the hypothesis-true DGP (spurious residual
 ↪ +0.5%), 0.978 under

```

a strictly inert placebo; trigger 0.79 when theta2b = theta2a at
↳ floor. These verified
values ARE the registered operating characteristics (aspirational >=
↳ 0.90 / >= 0.95 targets
restated per the driver reconciliation); the Stage-2 blind
↳ re-simulation at realized spurious
N is the binding re-verification. F2 requires >= 15,000 spurious
↳ sets; below that it reports
bounds only and the H2 verdict cap applies.'
- id: H2_diff
tier: exploratory
statement_initial: '"... a larger and more persistent post-event excess
↳ in weekly contacts
and gross fees (BRUHON) through the first post-closure year ..."'
statement_source: the strict comparative-magnitude quantifier of frozen
↳ bank hypothesis
H2, excerpted verbatim; tiered exploratory because the
↳ true-minus-spurious differential
is verified structurally infeasible as a confirmatory test (variance
↳ dominated by the
spurious arm; ~5.5x any honest variance budget; round-1 power 0.358,
↳ remediated re-run
0.460).
estimand: 'The frozen H2 statement''s strict differential quantifier,
↳ tiered EXPLORATORY
because it is verified structurally infeasible as a confirmatory
↳ test: theta2_diff = true-minus-spurious
post-year differential, whose variance is dominated by the spurious
↳ arm and would need
~5.5x the variance budget of any honest spurious-arm N (Stage-0
↳ diagnosis; power 0.358
at the 2pp floor). Reported with 95% CI under the exploratory
↳ flagging convention, plus
the quarter-by-quarter differential profile across Q1(c+1)..Q4(c+1)
↳ addressing the statement''s
'more persistent'' rider.'
primary_quantity: theta2_diff = mean_k[beta_true,k - beta_spurious,k],
↳ k in {Q1(c+1)..Q4(c+1)},
from the pooled stacked model with event-class x event-quarter
↳ interactions; 95% CI; exploratory;
a flagged estimate in the NON-predicted direction (spurious > true)
↳ triggers the pre-committed
non-inert-placebo interpretation.
direction: positive (true > spurious) under the frozen rider; a flagged
↳ estimate in the
NON-predicted direction triggers the pre-committed non-inert-placebo
↳ interpretation
...

## Current source interpretation rows

```yaml
- hypothesis_id: H2
outcome_pattern: theta2a > 0, Holm-significant, AND gate F2 passes
↳ (theta2b's disconfirmation
trigger does not fire), sign stable across GRAFT-3 thresholds
conclusion: 'We will conclude that workers at true closures show a
↳ post-boundary billing
excess through the first post-closure year whose spurious-arm
↳ counterpart is bounded below

```

X% of the true-arm excess, where X is the realized one-sided 95%  
 ↳ upper bound on theta2b  
 as a share of theta2a-hat (the registered gate quantity; a passing  
 ↳ gate bounds the spurious-arm  
 excess - it cannot license an absence claim) - consistent with  
 ↳ job-loss exposure rather  
 than dying-firm selection driving the association (classification is  
 ↳ measurement canon:  
 Benedetto et al. 2007; Hethey-Maier & Schmieder 2013; our claim is  
 ↳ only its placebo use  
 for health billing). The strict comparative magnitude ('larger and  
 ↳ more persistent')  
 is reported from the exploratory H2\_diff with its CI, never as a  
 ↳ confirmed quantity.'

- hypothesis\_id: H2  
 outcome\_pattern: theta2a > 0, Holm-significant, BUT gate F2 fires  
 ↳ (theta2b CI excludes zero  
 positively AND theta2b >= 50% of theta2a)  
 conclusion: 'We will conclude that although true-closure workers show a  
 ↳ post-year excess,  
 the comparative reading of H2 is disconfirmed: the placebo is not  
 ↳ inert, and the shared  
 excess is consistent with selection into dying workplace identities  
 ↳ or reorganization  
 stress rather than job loss per se - an explicit caution for  
 ↳ closure-based designs in  
 the displacement literature. The theta1 interpretation is  
 ↳ correspondingly weakened in  
 the Discussion. This pattern is registered as disconfirming for H2's  
 ↳ comparative claim  
 regardless of theta2a's significance.'

- hypothesis\_id: H2  
 outcome\_pattern: theta2a not Holm-significant (NULL), theta2b also near  
 ↳ zero  
 conclusion: We will conclude that we have no evidence of post-boundary  
 ↳ excess primary-care  
 billing after workplace terminal events of either kind, report the  
 ↳ CI-ruled-out magnitudes  
 for both arms, and note that this locates any billing response (per  
 ↳ the theta1 result)  
 before the boundary.

- hypothesis\_id: H2  
 outcome\_pattern: theta2a < 0, Holm-significant (reversed)  
 conclusion: We will conclude - contrary to our hypothesis - that  
 ↳ true-closure workers' post-year  
 contact falls relative to matched controls, consistent with  
 ↳ out-of-population attrition  
 (death/emigration invisible in scope; GRAFT-6 bounds mandatory) or  
 ↳ post-displacement care  
 deferral; disconfirming for H2 and reported with the bounding  
 ↳ machinery, never reframed.

- hypothesis\_id: H2\_diff  
 outcome\_pattern: Exploratory flag with theta2\_diff < 0 (spurious  
 ↳ exceeds true)  
 conclusion: We will report - contrary to the frozen comparative  
 ↳ expectation - that spurious-closure  
 workers show the larger excess, indicating the placebo is not inert  
 ↳ (reorganizations may  
 carry their own insecurity exposure); this exploratory pattern  
 ↳ reinforces whatever gate

F2 showed and caps the interpretive weight the true/spurious contrast  
 ↳ can bear.

## Source family scoring and limitations

```

```yaml
scoring_rule: 'Fixed in advance per the collaboration''s
↳ registered-report conventions (trajectory
  fields): at Stage 2 an outcome attaches to each confirmatory
  ↳ hypothesis''s frozen statement_initial
  as confirmed | partially_confirmed | reversed, adjudicated strictly
  ↳ against the pre-committed
  interpretation table in this protocol. H1 scores confirmed only if
  ↳ theta1''s Holm-adjusted
  two-sided test rejects with the point estimate in the predicted
  ↳ direction AND the gated
  exploratory H1_trend contrast supports the frozen statement''s growth
  ↳ rider (beta_Q4c -
  beta_Q1c > 0 with 95% CI excluding zero); it scores partially_confirmed
  ↳ when theta1 Holm-rejects
  in the predicted direction but the growth rider is unsupported (flat
  ↳ elevation - the pre-committed
  reading given annual event dating) or the trend gate is not computed;
  ↳ reversed if theta1
  is Holm-significant in the negative direction. H2 scores confirmed only
  ↳ if theta2a Holm-rejects
  in the predicted direction AND falsification gate F2 does not fire AND
  ↳ the theta2b sign
  pattern is stable across the GRAFT-3 successor-share thresholds;
  ↳ partially_confirmed when
  theta2a Holm-rejects but the comparative quantifier is untestable
  ↳ (spurious-arm kill - verdict
  capped by pre-commitment at ''partially consistent - comparative
  ↳ quantifier untestable'');
  reversed if theta2a is Holm-significant in the negative direction OR
  ↳ gate F2 fires (the
  frozen statement''s comparative reading is disconfirmed by a non-inert
  ↳ placebo regardless
  of theta2a - registered as disconfirming, never reframed). H3 scores
  ↳ confirmed only if theta3''s
  Holm-adjusted test rejects in the predicted direction (including the
  ↳ total-down/00H-up pattern,
  pre-registered as scoreable support for the composition claim);
  ↳ reversed if Holm-significant
  in the negative direction. A Holm-nonsignificant confirmatory result
  ↳ attaches no confirmation:
  it is reported as a bounded informative null - the 95% CI converted
  ↳ into the excess magnitudes
  the data rule out against realized MDEs - never ''accepted''. A member
  ↳ demoted by the Stage-2
  realized-N gate or killed by its registered criterion attaches no
  ↳ score; the demotion/kill
  trail is disclosed. Significant effects in the non-predicted direction
  ↳ are disconfirming,
  never supportive. Exploratory quantities (H1_trend, H2_diff, E1, E3,
  ↳ E4, and any demoted
  member) attach outcomes only under the exploratory heading and are
  ↳ never scored as preregistered-confirmed.'
assumptions:

```

- SSSY (2005-2025) and SYSI (1990-2005) capture the universe of publicly
 - ↪ reimbursed primary-care billing for insured residents; absence of rows means no billed care,
 - ↪ not missingness
- KONTAKT is a row-level contact marker whose sum yields contact counts
 - ↪ (in-environment validation with a frozen fallback definition if not)
- ENDNEDL/IDFREM/IDTILB/REGELFR/FRELFREM semantics follow Statistics
 - ↪ Denmark's IDA workplace-continuity documentation, cited directly (no DOI exists for these conventions, per
 - ↪ boundary directive);
 - the primary classification rests on the worker-flow measure $s(w)$ from
 - ↪ IDAN, so misreading an IDAS flag degrades only the cross-check
- November-status annual attachment measures workplace membership;
 - ↪ intra-year attachment changes are unobservable ex ante and the frozen c-1 anchor is the design's
 - ↪ answer to that - with the corollary, named as a design threat, that early year-c separations
 - ↪ feed realized post-job-loss billing into the pre-boundary window under intention-to-follow, a
 - ↪ channel the E4 separation shares and the registered theta1 attachment split bound but cannot
 - ↪ remove
- ALDERIMP/KOENIMP (DST-imputed age/sex on SSSY rows) are accurate enough
 - ↪ for stratified matching; persons absent from SSSY in c-3..c-1 are handled by explicit unknown
 - ↪ strata
- HONUGE settlement lags are predominantly ≤ 4 weeks (calibrated on
 - ↪ BEHANDLINGSDATO 2021+; guard-band contingency if not), and lag structure does not differ
 - ↪ between exposed and matched controls within the same calendar quarters
- Death and emigration are unobservable in scope; their post-event
 - ↪ incidence is assumed small; the deflating direction is conservative for the pre-boundary claim and
 - ↪ attenuating for theta2a, with the GRAFT-6 trigger bounding the post-boundary side
- Nominal-fee comparisons within cohort x calendar-quarter cells require
 - ↪ no deflator
- Danish institutional stability of GP billing categories (SPEC2 70-84,
 - ↪ 89) across 2007-2024 at the 2-digit level
- 'PUBLISHED-BASIS SIZING ASSUMPTIONS (new, each with a registered
 - ↪ failure path): (i) Danish establishment closure rates among private size-5+ workplaces are in the
 - ↪ published Erhvervsdemografi-consistent 1.5-3% band, supporting 150,000-300,000 exposed workers over 15 cohorts
 - ↪ - failure path: the pre-outcome realized-N gate demotes rather than stretches; (ii) the
 - ↪ Danish OOH share of GP contact volume is in the published 8-12% band (Christensen &
 - ↪ Olesen 1998; LV-KOS 2011 studies), simulated at the conservative 8% - failure path: the
 - ↪ published-range benchmark check plus realized-share re-simulation, then demotion or kill; (iii)
 - ↪ the control pool supports 1:3 matching at $\geq 90\%$ match rate - failure path: coarsening ladder,
 - ↪ realized-k accounting in the blind re-simulation, then the gate'

- theta2a's 4-quarter averaging gain over theta1 survives serial
 - ↳ correlation in post-year shocks; a registered serial-correlation discount is built into the
 - ↳ Stage-0 re-simulation rather than assumed away
 - The persistent-vs-transient overdispersion split (persistent frailty
 - ↳ var 1.25, transient log-var 0.25) is retained unchanged from the first-round simulation -
 - ↳ deliberately NOT revised in the direction that would raise power, so the remediation cannot be
 - ↳ accused of tuning the DGP
- failure_modes:
- 'Realized event or worker counts far below schema-derived expectations
 - ↳ (documentation-only corpus; no realized Ns exist) - the design now bears this risk openly:
 - ↳ the Stage-2 realized-N gate demotes underpowered members loudly and pre-outcome instead of
 - ↳ letting an underpowered confirmatory test run; a sparse-closure Denmark kills or demotes
 - ↳ cleanly rather than quietly'
 - KONTAKT or ENDNEDL semantics differ from assumption - both carry frozen
 - ↳ fallback rules, but level shifts in outcomes or event dating would force reported
 - ↳ recalibration and consume credibility
 - HONUGE settlement lag materially longer than 4 weeks - the
 - ↳ quarter-level estimands survive, but the weekly-resolution headline weakens to quarterly resolution in
 - ↳ practice (GRAFT-2 makes this explicit and pre-committed)
 - ANDELOPS direction-of-absorption semantics may differ from the idea's
 - ↳ gloss - classification routes through IDAN worker flows, so this only weakens the IDAS
 - ↳ cross-check, not the H2 machinery
 - Spurious-closure events too heterogeneous for a clean placebo - a
 - ↳ non-inert placebo now shows up as gate F2 firing (or exploratory H2_diff reversal), which
 - ↳ caps H2's verdict by pre-commitment instead of silently contaminating a differential test
 - Realized OOH share below the published range - caught by the
 - ↳ published-range benchmark check before estimation; H3 demotes or dies pre-outcome instead of running
 - ↳ underpowered
 - Realized N lands below the full-expectation lower bound - the gate
 - ↳ demotes affected members to exploratory pre-outcome; the study survives with a smaller
 - ↳ confirmatory family (or none, in which case it reports as registered exploratory estimation with the
 - ↳ demotion trail published
 - the first-round outcome, now a contingency rather than the design)
 - Unobservable post-closure death/emigration deflates post-boundary
 - ↳ billing - direction stated, sensitivity pre-specified, GRAFT-6 2pp trigger forces best/worst-case
 - ↳ bounds; no in-scope fix exists
 - Public/private sector misclassification via modal ARB_SEKTORKODE for
 - ↳ multi-establishment

```

    firms - attenuates contrasts; share bounded in the balance tables; if
    ↪ ARB_SEKTORKODE itself
    proves absent in-environment (it is unconfirmed in the recorded
    ↪ Forskningsservice inventory),
    the frozen ARB_HOVED_BRA_DB07 fallback in contingent_decisions
    ↪ substitutes, with counts
    reported
- Multi-job workers' primary-job assignment may misattribute exposure -
↪ deterministic rule
  biases toward null; share reported
- 1:3 matching without replacement may exhaust thin strata - realized k <
↪ 3 accounted honestly
  in the blind re-simulation rather than assumed; coarsening ladder
  ↪ pre-committed
- The SYSI-era exploratory replication may be infeasible - bounded as
↪ skipable Stage-5 work
  that cannot contaminate the confirmatory tier
- Readers may over-read gate F2 as proof of causal job-loss effects - the
↪ associational ceiling,
  the 'consistent with' vocabulary cap, and the pre-committed F2-fired
  ↪ interpretation guard
  the prose
...

```

Source disclosure rules

```

...yaml
environment: DST Forskningsservice (no prior data access; gatekeeper
↪ certification per RR
  secondary-data rules). All analysis occurs inside the environment; only
  ↪ aggregate outputs
  leave. No microdata, workplace identifiers (ARBNR/CVRNR/LBNR), or
  ↪ person-level values are
  released.
minimum_cell_size: 5
minimum_cell_rules: no released statistic based on fewer than 5 persons
↪ or fewer than 3 workplaces;
  event-count tables rounded to the nearest 10; weekly trajectory figures
  ↪ pool >= 100 events
  and >= 5,000 person-weeks per plotted point; fee (BRUHON) aggregates
  ↪ subject to a 50% single-workplace
  dominance check; min-cell suppression applied BEFORE Holm reporting so
  ↪ suppression can never
  be outcome-selective
outputs_leaving_environment:
- model coefficients with SEs/CIs
- matching balance tables and realized-k distribution
- calibration and benchmark tables (coverage, linkage, HONUGE lag, OOH
↪ share)
- figure aggregates (weekly trajectory series under the pooling rules)
- event counts rounded to the nearest 10
- Stage-0 power/MDE tables - BOTH rounds (failed and remediated) - and
↪ the Stage-2 blind re-simulation
  output with the demotion trail
- the c-2 placebo, gate F2 outputs, GRAFT-2/4 diagnostics, GRAFT-6
↪ later-presence rates, E4
  attrition panel, and all outcome-neutral check outputs (the full audit
  ↪ trail exits under
  the same rules)
- the analysis and power scripts

```

```

statement: 'All analysis occurs inside DST Forskningservice; only
↳ aggregate outputs leave.
Rules applied to every released object: (1) no statistic based on fewer
↳ than 5 persons or
fewer than 3 workplaces; (2) event-count tables rounded to the nearest
↳ 10; (3) weekly trajectory
figures pool >= 100 events and >= 5,000 person-weeks per plotted point;
↳ (4) fee (BRUHON)
statistics released only as cell means/coefficients over min-N cells,
↳ with a dominance check
that no single workplace contributes > 50% of any released fee
↳ aggregate; (5) released objects
are limited to: model coefficients with SEs/CIs, balance tables,
↳ calibration/benchmark tables,
figure aggregates, event counts (rounded), the Stage-0 (both rounds)
↳ and Stage-2 power/MDE
tables with the demotion trail, and the analysis/power scripts - never
↳ microdata, workplace
identifiers (ARBNR/CVRNR/LBNR), or person-level values; (6) min-cell
↳ suppression is applied
before Holm reporting so that suppression can never be
↳ outcome-selective; (7) the c-2 placebo,
gate F2 outputs, the GRAFT-2/4 diagnostics, the GRAFT-6 later-presence
↳ rates, the E4 attrition
panel, and all outcome-neutral check outputs are released under the
↳ same rules, ensuring
the full audit trail exits the environment.'
...

```

Pending registered experiment

No Danish study outcome has been observed for this bundle. See the complete registration above.

atom-000006: Care timing composition

Is peri-event relative billing excess tilted toward out-of-hours rather than daytime primary care?

Source fidelity and implementation order

This is a format conversion of DS1.1, not a new study or independent acceptance. Source audit: minor_revisions. No observed Danish outcomes are supplied. The source execution plan explicitly states that final prereg.yaml governs disagreements. Exact current YAML selections below therefore govern older execution-plan excerpts; these excerpts are retained as implementation detail, not a competing registration. See docs/provenance/source-discrepancies.md. The complete shared data/validity contracts are embedded in Atoms 1–3; read those native protocols when executing a downstream Atom, even without docs/. Do not substitute a scientific default for an unresolved source contradiction.

Preserve source stages 0 -> 1 -> 2 -> 3 -> 4 -> 5 inside the implementation. Source Stage-0 simulations are historical design evidence, not study outcomes; do not

silently redo or improve them. Gate 14 must be evaluated from baseline data before post-baseline outcomes. All surviving Stage-4 quantities form ONE Holm family. The three question-level Atoms may share the same frozen family fit and tidy results; do not apply Holm separately per Atom or adjudicate a member before the complete surviving family's p-values exist. Cache/reuse identical governed intermediate results with provenance when available; a direct-dispatch call must obtain or compute its exact source-defined inputs, not pretend its predecessors succeeded. Stage-5 internal riders remain exploratory, after the Stage-4 family is frozen, even when housed in H1/H2.

Reference files under docs/execution/ are the ORIGINAL SKELETONS, not tested turnkey implementations. Adapt paths and complete engineering in a governed workspace; preserve source scientific rules and disclose unresolved choices. Write a result record with component estimates, uncertainty, family eligibility, source interpretation row, diagnostics and boundary/skip dispositions. Released outputs follow source disclosure rules; confidential rows stay governed. Local native states summarize the source judgment; neither a state name nor source power forecasts prove equivalence or causal identification. A bounded claim requires its actual interval and registered rule. Retain unresolved source wording conflicts; do not convert mere nonsignificance into scientific null.

Frozen scope and vocabulary

```
claim_class_ceiling: associational
statement: 'Associational - matched event-study contrasts, a contrastive
↳ falsification gate,
  and a composition contrast; descriptive for trajectory panels and the
  ↳ bounded priority claim.
Binding vocabulary rules: 'is associated with' / 'consistent with' /
↳ 'suggests' only;
'effect', 'impact', 'causes' prohibited throughout; anticipation
↳ always worded as
'pre-boundary excess' relative to the annually dated closure-year
↳ boundary, never as announcement-dated
anticipation (S8); no phenomenon-discovery or Danish-first framing (S1-S3);
↳ H2 novelty confined
to the placebo's application to health billing (S7). H2 verdict semantics
↳ are fixed in
advance: 'consistent with H2' requires BOTH the confirmatory core
↳ (theta2a Holm-significant,
predicted direction) AND a passing falsification gate F2; the strict
↳ comparative magnitude
is only ever reported as an exploratory estimate with its CI. Exploratory
↳ results (H1_trend,
H2_diff, E1, E3, E4, and any gate-demoted member) carry the word
↳ 'exploratory' in the
heading and in every sentence stating them. The prose label never exceeds
↳ the evidence label.'
frozen_parameters:
seed: 20260831
gp_code_list: daytime GP = SPEC2 in {70..80}
ooh_code_list: out-of-hours GP = SPEC2 in {81, 82, 83, 84, 89}
ooh_robustness_classifier: adds YDLTID in {2, 3, 4, 5, 7, 8, 9}
spec2_evening_gp_disclosure: 'Registered choice, disclosed: the parsed
↳ value labels mark
```

SPEC2 71 ("General practitioner, consultation, evening etc."), 73
 ↳ ("General practitioner,
 telephone consultation, evening etc."), and 75 ("General practitioner,
 ↳ home visit, evening
 etc.") as explicitly off-schedule GP services; they nonetheless remain on
 ↳ the daytime
 side of the PRIMARY classifier because H3's out-of-hours construct and
 ↳ its published
 power basis are keyed to the laegevagt out-of-hours service (SPEC2 81-84,
 ↳ 89), the quantity
 the published Danish 8-12% share describes. Attenuation direction, both
 ↳ readings: counting
 evening GP as daytime attenuates theta3 toward zero - conservative
 ↳ against the positive
 H3 conclusion, anti-conservative for the null-bounding "compositionally
 ↳ unremarkable"
 reading, whose Stage-2 statement therefore carries a mandatory caveat.
 ↳ In-environment
 validation must verify that the YDLTID-augmented robustness classifier
 ↳ captures SPEC2
 71/73/75 rows (their evening time codes 2-5); if YDLTID proves
 ↳ unpopulated on those rows,
 the robustness classifier - and only the robustness classifier - is
 ↳ amended mechanically
 and with disclosure to include SPEC2 71/73/75 directly. The SPEC2
 ↳ 71/73/75 volume share
 is reported, and both classifiers appear side by side in Table T7.'
 successor_share_thresholds: 'true closure: ENDNEDL final closure AND
 ↳ worker-flow max-successor
 share s(w) < 0.30; spurious closure: forward identity via
 ↳ IDFREEM/REGELFR/FRELFREM AND
 s(w) >= 0.70; ambiguous mid-zone excluded; GRAFT-3 sensitivities
 ↳ 0.20/0.80 and 0.40/0.60'
 quarter_mapping: 'ISO week of HONUGE: 1-13 / 14-26 / 27-39 / 40-52; ISO
 ↳ week 53 to Q4'
 winsorization: person-quarter outcomes at the 99.9th percentile within era
 ↳ x calendar quarter
 matching: 1:3 nearest-neighbor Mahalanobis within exact strata, without
 ↳ replacement across
 sets, target k = 3, minimum k = 1; calipers |dif baseline contacts| <= 3
 ↳ and |dif log(1+BRUHON)|
 <= 0.5; exposed processed in seeded random order (seed 20260831)
 software: R >= 4.3, fixest::feols (OLS) and fixest::fepois (Poisson QML
 ↳ robustness); versions
 frozen in the registered script
 note: Code lists and thresholds frozen before any data access; code labels
 ↳ quoted verbatim
 from parsed DST value labels and re-validated in-environment before
 ↳ estimation (frozen
 fallback rules in contingent_decisions).

4.3 theta3 — H3 out-of-hours composition (confirmatory)

- **Model:** the 4.1 specification run separately with $y = \text{OOH contacts (SPEC2 81-84, 89)}$ and $y = \text{daytime GP contacts (SPEC2 70-80)}$.
- **Primary quantity:** $\text{theta3} = (\text{theta_OOH} / \mu_{\text{OOH}, \text{ctrl}}) - (\text{theta_day} / \mu_{\text{day}, \text{ctrl}})$, where theta_x = mean event-study coefficient over the peri-event

window $Q3(c)..Q2(c+1)$ for outcome x , and $\mu_{x,ctrl}$ = the matched-control mean of x over the same window.

- **Inference:** cluster-bootstrap percentile CI, 999 draws, resampling event workplaces, seed 20260831; Holm member; support requires $\theta_3 > 0$; floor = 5-point relative-excess gap.
- **Robustness:** repeat under the YDLTID-augmented OOH classifier. Pre-committed interpretations distinguish $\theta_{day} \geq 0$ (disproportionately OOH excess) from $\theta_{day} < 0$ (total-down/OOH-up composition shift — scoreable support for H3 that simultaneously disconfirms a volume-increase reading of θ_1).

4. ANALYSIS EXECUTION

Confirmatory tier first (Stage 4, frozen code, single run); exploratory tier after (Stage 5), segregated by heading. All estimation in R fixest via Rscript. Common estimator: outcome y_{iq} (winsorized quarterly GP contact count) on treated x event-quarter indicators for the 8 quarters $Q1(c)..Q4(c+1)$, matched-set fixed effects, cohort x calendar-quarter fixed effects, omitted baseline = the four quarters of year $c-1$; observation weights: exposed 1, controls $1/k_s$ within set; CR1 SEs clustered on event workplace (each control inherits its set's event cluster). Min-cell suppression is applied to all outputs BEFORE Holm reporting so suppression can never be outcome-selective.

4.4 Family adjudication and reporting order

Holm-Bonferroni over the surviving confirmatory family (as gated at step 14) at family $\alpha = 0.05$. Report all surviving tests exactly once, regardless of results. Each confirmatory hypothesis closes with a boldface **Summary (H_k)**. verdict paragraph opening “Consistent with” / “Contrary to our hypothesis”, per the pre-committed interpretation table in `prereg.yaml` (interpretations block — 16 rows covering positive, null, reversed, gate-fired, and partial patterns; apply them verbatim). Scoring attaches per the registered `scoring_rule`: H1 confirmed only if θ_1 Holm-rejects positively AND the gated $H1_trend$ supports the growth rider; partially-confirmed for flat elevation; reversed if negative. H2 confirmed only if θ_{2a} Holm-rejects positively AND F2 does not fire AND the θ_{2b} sign pattern is GRAFT-3-stable; reversed if θ_{2a} is significantly negative OR F2 fires. H3 confirmed on a positive Holm rejection (including the total-down/OOH-up pattern); reversed if negative. Demoted or killed members attach no score; their trail is disclosed. Sensitivity suite (run once, one table): GRAFT-3; GRAFT-9 (re-estimate θ_1/θ_{2a} excluding sets whose control workplace acquires any terminal event in $c+2$ or $c+3$ — cohorts $c \leq 2022$ for $c+2$, $c \leq 2021$ for $c+3$, truncation disclosed; registered direction note: exclusion should weakly INCREASE magnitudes, a decrease is flagged as diagnostic); un-winsorized outcomes; Poisson QML (`fepois`; if OLS and Poisson disagree in sign on any primary quantity, foreground the pre-committed OLS estimate, report both, flag the discordance in the Summary paragraph); possible-death/emigration exclusion (drop sets where the exposed member, or all controls, has zero SSSY rows across $c..c+1$ AND no IDAN record in $c+1$).

5. EXHIBITS

Every figure and table the report commits to. Disclosure rules binding on EVERY released object: no statistic based on fewer than 5 persons or fewer than 3 workplaces; event-count tables rounded to the nearest 10; weekly trajectory figures pool ≥ 100 events and $\geq 5,000$ person-weeks per plotted point; BRUHON aggregates pass a 50% single-workplace dominance check; released objects are limited to model coefficients with SEs/CIs, balance tables, calibration/benchmark tables, figure aggregates, rounded event counts, the power/MDE tables (both Stage-0 rounds and Stage 2) with the demotion trail, and the analysis/power scripts — never microdata, workplace identifiers (ARBNR/CVRNR/LBNR), or person-level values; suppression applied BEFORE Holm reporting.

Figures (matplotlib):

- **F1 — Quarterly event-study panel.** Treated $\times$ event-quarter coefficients with 95% CIs across Q1(c-2)..Q4(c+1) (placebo quarters, omitted-baseline marker, terminal-year and post-year quarters); panel
(a) true arm, panel (b) spurious arm from the pooled model. Feeds: step-18 panels + 4.1/4.2 coefficient tables. Disclosure: coefficients only.
- **F2 — Weekly descriptive trajectory.** Pooled exposed-minus- matched-control weekly GP contact series around the closure-year boundary, with the GRAFT-2 week-0 step diagnostic annotated (step $\leq 50\%$ of the total Q3(c)+Q4(c) excess; if exceeded, the figure carries the registered caution and quarter-level estimates are foregrounded). Feeds: step 15/17 weekly aggregates. Disclosure: ≥ 100 events and $\geq 5,000$ person-weeks per point.
- **F3 — s(w) successor-share distribution (GRAFT-4).** Histogram of worker-flow successor shares across terminal events with the 0.30/0.70 thresholds and mid-zone mass; bimodality check exhibit. Feeds: step 7. Disclosure: bin counts rounded to 10; bins under min-cell suppressed.
- **F4 — OOH vs daytime relative excess (theta3).** The two relative-excess series over the peri-event window with the bootstrap CI of their gap; primary and YDLTID-augmented classifier side by side. Feeds: 4.3 outputs. Disclosure: coefficients only.
- **F5 — Matching balance (love plot).** SMDs before/after matching for all matching covariates and unmatched auxiliaries; 0.10 line marked. Feeds: step 13. Disclosure: aggregates only.
- **F6 — E1 SYSI-era replication (exploratory; skippable).** Quarterly SYSI-era analogue of F1(a) for cohorts 1992–2003, labeled exploratory, YDLANT-weighted proxy outcome. Feeds: step 22. Disclosure: as F1.

Tables (booktabs LaTeX):

- **T1 — Event extraction and classification by cohort year.** Terminal events, true/spurious/ambiguous counts (rounded to 10), annual true-closure rate among private size-5+ workplaces, spurious/true ratio. Feeds: step 8.

- **T2 — Classification cross-validation.** Worker-flow classes vs IDAS-native flags (ENDNEDL/IDFREM/REGELFR/FRELFREM/ANDELUDS), full cross-tabulation + agreement rate; GRAFT-3 classification counts under 20/80 and 40/60. Feeds: step 8.
- **T3 — Cohort flow accounting.** Inclusion/exclusion counts per rule (tenure, sector, size, SIKGRUP, missing ANSAAR/pay, multi-cohort dedup), eligible exposed per cohort, match rate, ladder rung used, realized k distribution. Feeds: steps 9–13.
- **T4 — Matching balance.** Per-covariate SMDs (weighted 1/k), caliper compliance, realized mean k. Feeds: step 13.
- **T5 — Power and demotion trail.** Stage-0 power results (BOTH rounds: the voided round-1 run and the remediated re-run) and the Stage-2 blind realized-power/MDE table with binding N*, realized sets, and any demotions with the fixed-point recomputation. Feeds: deposited script
– step 14. Published regardless of outcome.
- **T6 — Outcome-neutral check panel.** All eleven registered checks (coverage benchmark by year, linkage floor, closure-rate plausibility, classification agreement, GRAFT-4, balance/match rate, c-2 placebo, HONUUGE lag calibration, floor/ceiling + seasonality, GRAFT-2 step, OOH-share benchmark) with thresholds and PASS/FAIL. Feeds: steps 8, 13, 16, 17, 19, 21 + F2/F3 diagnostics.
- **T7 — Confirmatory results.** theta1, theta2a, theta3: point estimates (contacts/person-quarter and % of baseline), 95% CIs, two-sided p, Holm-adjusted verdicts; the F2 gate row (theta2b, its one-sided 95% upper bound, trigger evaluation); the gated BRUHON fee secondary (if computed); c-2 placebo row; realized MDEs beside every null with the CI-ruled-out magnitudes. Feeds: 4.1–4.4.
- **T8 — Sensitivity suite.** GRAFT-3 thresholds; GRAFT-9 forward-window exclusion (with truncation note); un-winsorized; Poisson QML; death/emigration exclusion — one table, run once. Feeds: 4.4.
- **T9 — Exploratory tier.** H1_trend (if gated in; otherwise the row states the gate was not met), H2_diff with quarter-by-quarter differential profile, E3 fee and subgroup analyses (within-family Holm), E1 summary. Every row labeled exploratory. Feeds: 4.5.
- **T10 — Attrition and later-presence accounting.** E4 year-c separation shares by event class (ANSXFREM-based, post-treatment- conditioned label and selection caveat) and the GRAFT-6 later-presence rates with the 2pp trigger status. Feeds: steps 20 + 4.5 E4.

Current source hypotheses (all fields)

```
- id: H3
  tier: confirmatory
  statement_initial: 'H3 (distress composition): Compared to the out-of-hours
    ↪ share of contemporaneous
      contacts among their matched controls, closure-exposed workers''
    ↪ peri-event excess contacts
```

will contain a disproportionately higher share of out-of-hours and
 ↳ off-schedule services
 (YDLTID time-of-day codes; VAGTOMR in the SYSI era), consistent with
 ↳ unplanned, distress-driven
 care-seeking rather than deferred routine care.'

statement_source: idea bank entry DS1.1, hypothesis 3 of 3; wording
 ↳ verbatim, frozen at
 bank creation. Construct unchanged from the bank statement
 ↳ (relative-excess contrast,
 sign-unambiguous under the Hamad total-down/OOH-up world); only the power
 ↳ basis was recalibrated
 to the published Danish OOH share (see remediation_record).

estimand: 'Unchanged construct (already repaired against
 ↳ share-among-utilizers pathologies;
 sign-unambiguous under the Hamad total-down/OOH-up world, S6): within
 ↳ true-closure stacks,
 the difference in RELATIVE peri-event excess (event-study coefficient
 ↳ divided by matched-control
 same-window mean) between out-of-hours GP contacts (SPEC2 in
 ↳ {81,82,83,84,89}) and daytime
 GP contacts (SPEC2 in {70..80}), peri-event window Q3(c)..Q2(c+1). What
 ↳ changed is ONLY
 the power basis: the Stage-0 simulation's OOH base rate is recalibrated
 ↳ from the self-declared
 over-conservative 3% to 8.0% of daytime GP volume (7.4% of all GP
 ↳ contacts) - the conservative
 end of the published Danish out-of-hours share (~8-12%: Christensen &
 ↳ Olesen 1998, BMJ
 316:1502; Moth, Flarup, Christensen, Olesen & Vedsted, LV-KOS 2011
 ↳ out-of-hours contact
 studies; the audit-verified diagnosis records that at ~10% the test's z
 ↳ roughly doubles
 and clears).'

primary_quantity: 'theta3 = (theta_OOH / mu_OOH,ctrl) - (theta_day /
 ↳ mu_day,ctrl), where
 theta_x = mean event-study coefficient over the peri-event window for
 ↳ outcome x and mu_x,ctrl
 = matched-control mean of x over the same window; cluster-bootstrap
 ↳ percentile CI (999
 draws, seed 20260831, clusters = event workplaces). Predicted direction
 ↳ theta3 > 0. Holm
 family member. Floor effect UNCHANGED at a 5-point relative-excess gap,
 ↳ below the Hamad/Modrek/Cullen
 (2015) ER-up/outpatient-down composition shifts. Power path: z scales ~
 ↳ sqrt(OOH base
 rate), x1.63 moving 3% -> 8%, plus the N and 1:3-matching levers;
 ↳ provisional gate N*_H3
 = 90,000 true matched sets; an outcome-neutral check verifies the
 ↳ realized control OOH
 share sits in the published range before the confirmatory run.'

direction: 'positive (theta3 > 0: OOH relative excess exceeds daytime
 ↳ relative excess)'

comparator: matched controls' contemporaneous out-of-hours and daytime GP
 ↳ contact means
 over the same peri-event window (each arm's excess scaled by its own
 ↳ control mean before
 differencing)

exposure_variable: true-closure exposure over the peri-event window
 ↳ Q3(c)..Q2(c+1)

outcome_variable: out-of-hours GP contacts (SPEC2 81-84, 89; robustness
 ↳ classifier adds
 YDLTID 2-5, 7-9) versus daytime GP contacts (SPEC2 70-80), as relative
 ↳ peri-event excesses

forecast:
 point: +2.5-point gap in relative peri-event excess (OOH minus daytime)
 interval_90: '[-3.0, +9.0] points'
 basis: 'Agent forecast, recorded before any data access and kept separate
 ↳ from any future
 expert forecasts (collaboration roadmap rule). Upward anchors: the
 ↳ strongest prior art
 found by the search - emergency use rose while outpatient use fell
 ↳ under layoff-threat
 insecurity (Hamad, Modrek & Cullen 2015) - and the low access cost of
 ↳ the Danish laegevag.
 Downward anchors: OOH contact may track genuine acute need rather than
 ↳ distress, Danish
 OOH gatekeeping triages non-urgent contacts back to daytime care, and
 ↳ if the peri-event
 excess is dominated by deferred routine care (using remaining
 ↳ employment time for scheduled
 visits) the gap reverses - the reversal is pre-registered as
 ↳ interpretable. My central
 forecast sits below the 5-point floor; the composition tilt is the
 ↳ hypothesis I hold
 with the widest uncertainty.'

null_substantive: 'The peri-event excess is compositionally unremarkable:
 ↳ the out-of-hours
 share of excess contacts does not differ from the daytime share by 5
 ↳ points or more (bootstrap
 CI converted to ruled-out gaps against the realized MDE), consistent with
 ↳ scheduled as
 much as distress-driven care.'

kill_criterion: 'KILL: if OOH contacts (SPEC2 81-84, 89) are < 2% of all
 ↳ control-group GP
 contacts in the baseline year (composition not estimable), or the OOH
 ↳ classifier is unresolvable
 (YDLTID missing/'not relevant' on > 20% of peri-event GP rows and SPEC2
 ↳ OOH codes absent).
 CONFIRMATORY DEMOTION GATE: realized control OOH share < 5% with blind
 ↳ re-simulated power
 < 0.95 at the 5-point floor, or realized matched sets < N*_H3
 ↳ (provisional 90,000) ->
 theta3 demotes to exploratory pre-outcome.'

Current source interpretation rows

- hypothesis_id: H3
 outcome_pattern: theta3 > 0, Holm-significant, with theta_day >= 0
 conclusion: We will conclude that the peri-event excess is
 ↳ disproportionately out-of-hours,
 consistent with unplanned, distress-driven care-seeking rather than
 ↳ deferred routine care
 (composition claim only; no diagnosis data exist in scope, so no
 ↳ health-state claim is
 made).
- hypothesis_id: H3
 outcome_pattern: theta3 > 0, Holm-significant, with theta_day < 0 (total
 ↳ down, OOH up)

```

conclusion: We will conclude that closure exposure is associated with a
↳ compositional shift
- daytime contact falls while out-of-hours contact rises - matching the
↳ forgone-routine-care-plus-acute-distress
pattern in the prior insecurity literature (Hamad, Modrek & Cullen 2015);
↳ this pattern
supports H3's composition claim while disconfirming any volume-increase
↳ reading of theta1
(S6).
- hypothesis_id: H3
outcome_pattern: theta3 not Holm-significant (NULL)
conclusion: We will conclude that we have no evidence that the peri-event
↳ excess is compositionally
tilted toward out-of-hours care, and report the relative-excess
↳ differentials the 95%
bootstrap CI rules out against the realized MDE; any theta1 excess is
↳ then compositionally
unremarkable - consistent with scheduled as much as distress-driven care
↳ - quoted with
the registered caveat that the primary classifier counts SPEC2 71/73/75
↳ evening-GP services
as daytime, which attenuates theta3 and correspondingly weakens the null
↳ bound (frozen_parameters.spec2_evening_gp_disclosure).
- hypothesis_id: H3
outcome_pattern: theta3 < 0, Holm-significant (reversed)
conclusion: We will conclude - contrary to our hypothesis - that the
↳ peri-event excess is
disproportionately daytime/scheduled, consistent with planned or deferred
↳ routine care
(e.g., using remaining employment time or sick-leave access for scheduled
↳ visits) rather
than acute distress; H3 is disconfirmed.

```

Source family scoring and limitations

```

scoring_rule: 'Fixed in advance per the collaboration''s registered-report
↳ conventions (trajectory
fields): at Stage 2 an outcome attaches to each confirmatory hypothesis''s
↳ frozen statement_initial
as confirmed | partially_confirmed | reversed, adjudicated strictly against
↳ the pre-committed
interpretation table in this protocol. H1 scores confirmed only if
↳ theta1''s Holm-adjusted
two-sided test rejects with the point estimate in the predicted direction
↳ AND the gated
exploratory H1_trend contrast supports the frozen statement''s growth rider
↳ (beta_Q4c -
beta_Q1c > 0 with 95% CI excluding zero); it scores partially_confirmed
↳ when theta1 Holm-rejects
in the predicted direction but the growth rider is unsupported (flat
↳ elevation - the pre-committed
reading given annual event dating) or the trend gate is not computed;
↳ reversed if theta1
is Holm-significant in the negative direction. H2 scores confirmed only if
↳ theta2a Holm-rejects
in the predicted direction AND falsification gate F2 does not fire AND the
↳ theta2b sign
pattern is stable across the GRAFT-3 successor-share thresholds;
↳ partially_confirmed when

```

theta2a Holm-rejects but the comparative quantifier is untestable
 ↪ (spurious-arm kill - verdict
 capped by pre-commitment at 'partially consistent - comparative quantifier
 ↪ untestable');
 reversed if theta2a is Holm-significant in the negative direction OR gate
 ↪ F2 fires (the
 frozen statement's comparative reading is disconfirmed by a non-inert
 ↪ placebo regardless
 of theta2a - registered as disconfirming, never reframed). H3 scores
 ↪ confirmed only if theta3's
 Holm-adjusted test rejects in the predicted direction (including the
 ↪ total-down/OOH-up pattern,
 pre-registered as scoreable support for the composition claim); reversed if
 ↪ Holm-significant
 in the negative direction. A Holm-nonsignificant confirmatory result
 ↪ attaches no confirmation:
 it is reported as a bounded informative null - the 95% CI converted into
 ↪ the excess magnitudes
 the data rule out against realized MDEs - never 'accepted'. A member
 ↪ demoted by the Stage-2
 realized-N gate or killed by its registered criterion attaches no score;
 ↪ the demotion/kill
 trail is disclosed. Significant effects in the non-predicted direction are
 ↪ disconfirming,
 never supportive. Exploratory quantities (H1_trend, H2_diff, E1, E3, E4,
 ↪ and any demoted
 member) attach outcomes only under the exploratory heading and are never
 ↪ scored as preregistered-confirmed.'

assumptions:

- SSSY (2005-2025) and SYSI (1990-2005) capture the universe of publicly
 ↪ reimbursed primary-care
 billing for insured residents; absence of rows means no billed care, not
 ↪ missingness
- KONTAKT is a row-level contact marker whose sum yields contact counts
 ↪ (in-environment validation
 with a frozen fallback definition if not)
- ENDNEDL/IDFREM/IDTILB/REGELFR/FRELFREM semantics follow Statistics
 ↪ Denmark's IDA workplace-continuity
 documentation, cited directly (no DOI exists for these conventions, per
 ↪ boundary directive);
 the primary classification rests on the worker-flow measure s(w) from IDAN,
 ↪ so misreading
 an IDAS flag degrades only the cross-check
- November-status annual attachment measures workplace membership; intra-year
 ↪ attachment changes
 are unobservable ex ante and the frozen c-1 anchor is the design's answer
 ↪ to that - with
 the corollary, named as a design threat, that early year-c separations feed
 ↪ realized post-job-loss
 billing into the pre-boundary window under intention-to-follow, a channel
 ↪ the E4 separation
 shares and the registered theta1 attachment split bound but cannot remove
- ALDERIMP/KOENIMP (DST-imputed age/sex on SSSY rows) are accurate enough for
 ↪ stratified matching;
 persons absent from SSSY in c-3..c-1 are handled by explicit unknown strata
- HONUGE settlement lags are predominantly <= 4 weeks (calibrated on
 ↪ BEHANDLINGSDATO 2021+;
 guard-band contingency if not), and lag structure does not differ between
 ↪ exposed and matched
 controls within the same calendar quarters

- Death and emigration are unobservable in scope; their post-event incidence
 - ↪ is assumed small;
 - the deflating direction is conservative for the pre-boundary claim and
 - ↪ attenuating for theta2a,
 - with the GRAFT-6 trigger bounding the post-boundary side
- Nominal-fee comparisons within cohort x calendar-quarter cells require no
 - ↪ deflator
- Danish institutional stability of GP billing categories (SPEC2 70-84, 89)
 - ↪ across 2007-2024
 - at the 2-digit level
- 'PUBLISHED-BASIS SIZING ASSUMPTIONS (new, each with a registered failure
 - ↪ path): (i) Danish establishment closure rates among private size-5+ workplaces are in the
 - ↪ published Erhvervsdemografi-consistent
 - 1.5-3% band, supporting 150,000-300,000 exposed workers over 15 cohorts -
 - ↪ failure path:
 - the pre-outcome realized-N gate demotes rather than stretches; (ii) the
 - ↪ Danish OOH share
 - of GP contact volume is in the published 8-12% band (Christensen & Olesen
 - ↪ 1998; LV-KOS 2011
 - studies), simulated at the conservative 8% - failure path: the
 - ↪ published-range benchmark
 - check plus realized-share re-simulation, then demotion or kill; (iii) the
 - ↪ control pool supports
 - 1:3 matching at >= 90% match rate - failure path: coarsening ladder,
 - ↪ realized-k accounting
 - in the blind re-simulation, then the gate'
 - theta2a's 4-quarter averaging gain over theta1 survives serial correlation
 - ↪ in post-year
 - shocks; a registered serial-correlation discount is built into the Stage-0
 - ↪ re-simulation
 - rather than assumed away
- The persistent-vs-transient overdispersion split (persistent frailty var
 - ↪ 1.25, transient
 - log-var 0.25) is retained unchanged from the first-round simulation -
 - ↪ deliberately NOT revised
 - in the direction that would raise power, so the remediation cannot be
 - ↪ accused of tuning
 - the DGP
- failure_modes:
 - 'Realized event or worker counts far below schema-derived expectations
 - ↪ (documentation-only
 - corpus; no realized Ns exist) - the design now bears this risk openly: the
 - ↪ Stage-2 realized-N
 - gate demotes underpowered members loudly and pre-outcome instead of letting
 - ↪ an underpowered
 - confirmatory test run; a sparse-closure Denmark kills or demotes cleanly
 - ↪ rather than quietly'
 - KONTAKT or ENDNEDL semantics differ from assumption - both carry frozen
 - ↪ fallback rules,
 - but level shifts in outcomes or event dating would force reported
 - ↪ recalibration and consume
 - credibility
 - HONUGE settlement lag materially longer than 4 weeks - the quarter-level
 - ↪ estimands survive,
 - but the weekly-resolution headline weakens to quarterly resolution in
 - ↪ practice (GRAFT-2
 - makes this explicit and pre-committed)
 - ANDELOPS direction-of-absorption semantics may differ from the idea's gloss
 - ↪ - classification

- routes through IDAN worker flows, so this only weakens the IDAS
 - ↪ cross-check, not the H2 machinery
- Spurious-closure events too heterogeneous for a clean placebo - a non-inert
 - ↪ placebo now shows up as gate F2 firing (or exploratory H2_diff reversal), which caps
 - ↪ H2's verdict by pre-commitment instead of silently contaminating a differential test
- Realized OOH share below the published range - caught by the
 - ↪ published-range benchmark check before estimation; H3 demotes or dies pre-outcome instead of running
 - ↪ underpowered
- Realized N lands below the full-expectation lower bound - the gate demotes
 - ↪ affected members to exploratory pre-outcome; the study survives with a smaller confirmatory
 - ↪ family (or none, in which case it reports as registered exploratory estimation with the
 - ↪ demotion trail published
 - the first-round outcome, now a contingency rather than the design)
- Unobservable post-closure death/emigration deflates post-boundary billing -
 - ↪ direction stated, sensitivity pre-specified, GRAFT-6 2pp trigger forces best/worst-case
 - ↪ bounds; no in-scope fix exists
- Public/private sector misclassification via modal ARB_SEKTORKODE for
 - ↪ multi-establishment firms - attenuates contrasts; share bounded in the balance tables; if
 - ↪ ARB_SEKTORKODE itself proves absent in-environment (it is unconfirmed in the recorded
 - ↪ Forskningservice inventory), the frozen ARB_HOVED_BRA_DB07 fallback in contingent_decisions substitutes,
 - ↪ with counts reported
- Multi-job workers' primary-job assignment may misattribute exposure -
 - ↪ deterministic rule biases toward null; share reported
- 1:3 matching without replacement may exhaust thin strata - realized $k < 3$
 - ↪ accounted honestly in the blind re-simulation rather than assumed; coarsening ladder
 - ↪ pre-committed
- The SYSI-era exploratory replication may be infeasible - bounded as
 - ↪ skippable Stage-5 work that cannot contaminate the confirmatory tier
- Readers may over-read gate F2 as proof of causal job-loss effects - the
 - ↪ associational ceiling, the 'consistent with' vocabulary cap, and the pre-committed F2-fired
 - ↪ interpretation guard the prose

Source disclosure rules

environment: DST Forskningservice (no prior data access; gatekeeper
 ↪ certification per RR secondary-data rules). All analysis occurs inside the environment; only
 ↪ aggregate outputs leave. No microdata, workplace identifiers (ARBNR/CVRNR/LBNR), or
 ↪ person-level values are released.
 minimum_cell_size: 5

```

minimum_cell_rules: no released statistic based on fewer than 5 persons or
↳ fewer than 3 workplaces;
  event-count tables rounded to the nearest 10; weekly trajectory figures
  ↳ pool >= 100 events
  and >= 5,000 person-weeks per plotted point; fee (BRUHON) aggregates
  ↳ subject to a 50% single-workplace
  dominance check; min-cell suppression applied BEFORE Holm reporting so
  ↳ suppression can never
  be outcome-selective
outputs_leaving_environment:
- model coefficients with SEs/CIs
- matching balance tables and realized-k distribution
- calibration and benchmark tables (coverage, linkage, HONUGE lag, OOH share)
- figure aggregates (weekly trajectory series under the pooling rules)
- event counts rounded to the nearest 10
- Stage-0 power/MDE tables - BOTH rounds (failed and remediated) - and the
↳ Stage-2 blind re-simulation
  output with the demotion trail
- the c-2 placebo, gate F2 outputs, GRAFT-2/4 diagnostics, GRAFT-6
↳ later-presence rates, E4
  attrition panel, and all outcome-neutral check outputs (the full audit
  ↳ trail exits under
  the same rules)
- the analysis and power scripts
statement: 'All analysis occurs inside DST Forskningservice; only aggregate
↳ outputs leave.
  Rules applied to every released object: (1) no statistic based on fewer
  ↳ than 5 persons or
  fewer than 3 workplaces; (2) event-count tables rounded to the nearest 10;
  ↳ (3) weekly trajectory
  figures pool >= 100 events and >= 5,000 person-weeks per plotted point; (4)
  ↳ fee (BRUHON)
  statistics released only as cell means/coefficients over min-N cells, with
  ↳ a dominance check
  that no single workplace contributes > 50% of any released fee aggregate;
  ↳ (5) released objects
  are limited to: model coefficients with SEs/CIs, balance tables,
  ↳ calibration/benchmark tables,
  figure aggregates, event counts (rounded), the Stage-0 (both rounds) and
  ↳ Stage-2 power/MDE
  tables with the demotion trail, and the analysis/power scripts - never
  ↳ microdata, workplace
  identifiers (ARBNR/CVRNR/LBNR), or person-level values; (6) min-cell
  ↳ suppression is applied
  before Holm reporting so that suppression can never be outcome-selective;
  ↳ (7) the c-2 placebo,
  gate F2 outputs, the GRAFT-2/4 diagnostics, the GRAFT-6 later-presence
  ↳ rates, the E4 attrition
  panel, and all outcome-neutral check outputs are released under the same
  ↳ rules, ensuring
  the full audit trail exits the environment.'
```

Complete registration fields

```

id: atom-000006
topic: Care timing composition
atom_result: prereg
atom_strength: null
atom_shape: null
evidence_paths: []
```

```

concepts: []
literature: []
bonds: []
preregistration:
  version: 1
  mode: adaptive
  question: Is peri-event relative billing excess tilted toward out-of-hours
    rather than daytime primary care?
  rationale: Source-existing coherent scientific question; its related
    ↳ outcomes,
    ↳ riders and sensitivity specifications remain internal.
  assumption: |-
    Associational - matched event-study contrasts, a contrastive
    ↳ falsification gate, and a composition contrast; descriptive for
    ↳ trajectory panels and the bounded priority claim. Binding vocabulary
    ↳ rules: 'is associated with' / 'consistent with' / 'suggests' only;
    ↳ 'effect', 'impact', 'causes' prohibited throughout; anticipation
    ↳ always worded as 'pre-boundary excess' relative to the annually dated
    ↳ closure-year boundary, never as announcement-dated anticipation (S8);
    ↳ no phenomenon-discovery or Danish-first framing (S1-S3); H2 novelty
    ↳ confined to the placebo's application to health billing (S7). H2
    ↳ verdict semantics are fixed in advance: 'consistent with H2' requires
    ↳ BOTH the confirmatory core (theta2a Holm-significant, predicted
    ↳ direction) AND a passing falsification gate F2; the strict
    ↳ comparative magnitude is only ever reported as an exploratory
    ↳ estimate with its CI. Exploratory results (H1_trend, H2_diff, E1, E3,
    ↳ E4, and any gate-demoted member) carry the word 'exploratory' in the
    ↳ heading and in every sentence stating them. The prose label never
    ↳ exceeds the evidence label.
    Compatibility ratings below are neutral import metadata, not a new
    ↳ scientific evaluation or empirical scores.
  target_strength: associational
  target_shape: comparison
  ratings:
    importance: 3
    novelty: 3
    feasibility: 3
    data_advantage: 3
    informativeness: 0.5
  tables:
    - IDAN
    - IDAS
    - SSSY
  activation: state(atom-000003) == "event_study_ready"
  verdicts:
    - key: not_estimable
      region: The source H3-specific classifier or rare-OOH kill criterion is
        established.
      says: Composition is not estimable under the registered construct; not
        a statistical null.
    - key: exploratory
      region: H3 was demoted pre-outcome and a valid source estimate is
        ↳ available.
      says: Report exploratory composition with the primary/robustness
        ↳ classifier
        ↳ caveat.
    - key: ooh_tilt
      region: Eligible theta3 > 0 and Holm rejects.
      says: 'Composition claim only: retain the theta_day >= 0 versus < 0 source
        interpretations; no diagnosis or health-state claim.'
```

- key: reversed
region: Eligible $\theta_3 < 0$ and Holm rejects.
says: A relative daytime tilt contradicts H3.
- key: not_supported
region: Eligible H3 is estimated but its Holm test does not reject.
says: Report the actual interval and classifier caveat; nonsignificance alone does not establish equivalence.
- key: boundary
region: Required source-defined inputs or scientific definitions are
 ↳ demonstrably
 unavailable; retain available components and the exact unresolved scope.
says: An evidenced boundary, not a statistical null. A missing
 ↳ implementation
 or timeout alone is operational no_result, not this state.
- key: otherwise
region: otherwise
says: The registered judgment is unresolved; report all completed
 ↳ components
 without inventing support, equivalence or absent results.

protocol: |
Is peri-event relative billing excess tilted toward out-of-hours rather
 ↳ than daytime primary care?

Source fidelity and implementation order

This is a format conversion of DS1.1, not a new study or independent
 ↳ acceptance.
Source audit: minor_revisions. No observed Danish outcomes are supplied.
The source execution plan explicitly states that final prereg.yaml governs
disagreements. Exact current YAML selections below therefore govern older
execution-plan excerpts; these excerpts are retained as implementation
 ↳ detail,
not a competing registration. See docs/provenance/source-discrepancies.md.
The complete shared data/validity contracts are embedded in Atoms 1--3;
 ↳ read
those native protocols when executing a downstream Atom, even without
 ↳ docs/.
Do not substitute a scientific default for an unresolved source
 ↳ contradiction.

Preserve source stages 0 -> 1 -> 2 -> 3 -> 4 -> 5 inside the
 ↳ implementation.
Source Stage-0 simulations are historical design evidence, not study
 ↳ outcomes;
do not silently redo or improve them. Gate 14 must be evaluated from
 ↳ baseline
data before post-baseline outcomes. All surviving Stage-4 quantities form
 ↳ ONE
Holm family. The three question-level Atoms may share the same frozen
 ↳ family
fit and tidy results; do not apply Holm separately per Atom or adjudicate
 ↳ a
member before the complete surviving family's p-values exist. Cache/reuse
identical governed intermediate results with provenance when available; a
direct-dispatch call must obtain or compute its exact source-defined
 ↳ inputs,
not pretend its predecessors succeeded. Stage-5 internal riders remain
exploratory, after the Stage-4 family is frozen, even when housed in
 ↳ H1/H2.

Reference files under docs/execution/ are the ORIGINAL SKELETONS, not
↳ tested
turnkey implementations. Adapt paths and complete engineering in a
↳ governed
workspace; preserve source scientific rules and disclose unresolved
↳ choices.
Write a result record with component estimates, uncertainty, family
↳ eligibility,
source interpretation row, diagnostics and boundary/skip dispositions.
↳ Released
outputs follow source disclosure rules; confidential rows stay governed.
Local native states summarize the source judgment; neither a state name
↳ nor
source power forecasts prove equivalence or causal identification. A
↳ bounded
claim requires its actual interval and registered rule. Retain unresolved
↳ source
wording conflicts; do not convert mere nonsignificance into scientific
↳ null.

Frozen scope and vocabulary

```
```yaml
claim_class_ceiling: associational
statement: 'Associational - matched event-study contrasts, a contrastive
↳ falsification gate,
 and a composition contrast; descriptive for trajectory panels and the
↳ bounded priority claim.
Binding vocabulary rules: 'is associated with' / 'consistent with'
↳ / 'suggests' only;
'effect', 'impact', 'causes' prohibited throughout; anticipation
↳ always worded as
'pre-boundary excess' relative to the annually dated closure-year
↳ boundary, never as announcement-dated
anticipation (S8); no phenomenon-discovery or Danish-first framing
↳ (S1-S3); H2 novelty confined
to the placebo's application to health billing (S7). H2 verdict
↳ semantics are fixed in
advance: 'consistent with H2' requires BOTH the confirmatory core
↳ (theta2a Holm-significant,
predicted direction) AND a passing falsification gate F2; the strict
↳ comparative magnitude
is only ever reported as an exploratory estimate with its CI.
↳ Exploratory results (H1_trend,
H2_diff, E1, E3, E4, and any gate-demoted member) carry the word
↳ 'exploratory' in the
heading and in every sentence stating them. The prose label never
↳ exceeds the evidence label.'
```

frozen\_parameters:

```
seed: 20260831
gp_code_list: daytime GP = SPEC2 in {70..80}
ooh_code_list: out-of-hours GP = SPEC2 in {81, 82, 83, 84, 89}
ooh_robustness_classifier: adds YDLTID in {2, 3, 4, 5, 7, 8, 9}
spec2_evening_gp_disclosure: 'Registered choice, disclosed: the parsed
↳ value labels mark
 SPEC2 71 ("General practitioner, consultation, evening etc."), 73
↳ ("General practitioner,
 telephone consultation, evening etc."), and 75 ("General
↳ practitioner, home visit, evening
```

```

etc.") as explicitly off-schedule GP services; they nonetheless
 ↳ remain on the daytime
side of the PRIMARY classifier because H3's out-of-hours construct
 ↳ and its published
power basis are keyed to the laegevagt out-of-hours service (SPEC2
 ↳ 81-84, 89), the quantity
the published Danish 8-12% share describes. Attenuation direction,
 ↳ both readings: counting
evening GP as daytime attenuates theta3 toward zero - conservative
 ↳ against the positive
H3 conclusion, anti-conservative for the null-bounding
 ↳ "compositionally unremarkable"
reading, whose Stage-2 statement therefore carries a mandatory
 ↳ caveat. In-environment
validation must verify that the YDLTID-augmented robustness
 ↳ classifier captures SPEC2
71/73/75 rows (their evening time codes 2-5); if YDLTID proves
 ↳ unpopulated on those rows,
the robustness classifier - and only the robustness classifier - is
 ↳ amended mechanically
and with disclosure to include SPEC2 71/73/75 directly. The SPEC2
 ↳ 71/73/75 volume share
is reported, and both classifiers appear side by side in Table T7.'
successor_share_thresholds: 'true closure: ENDNEDL final closure AND
 ↳ worker-flow max-successor
share s(w) < 0.30; spurious closure: forward identity via
 ↳ IDFREEM/REGELFR/FRELFREM AND
s(w) >= 0.70; ambiguous mid-zone excluded; GRAFT-3 sensitivities
 ↳ 0.20/0.80 and 0.40/0.60'
quarter_mapping: 'ISO week of HONUGE: 1-13 / 14-26 / 27-39 / 40-52; ISO
 ↳ week 53 to Q4'
winsorization: person-quarter outcomes at the 99.9th percentile within
 ↳ era x calendar quarter
matching: 1:3 nearest-neighbor Mahalanobis within exact strata, without
 ↳ replacement across
sets, target k = 3, minimum k = 1; calipers |dif baseline contacts|
 ↳ <= 3 and |dif log(1+BRUHON)|
<= 0.5; exposed processed in seeded random order (seed 20260831)
software: R >= 4.3, fixest::feols (OLS) and fixest::fepois (Poisson QML
 ↳ robustness); versions
frozen in the registered script
note: Code lists and thresholds frozen before any data access; code
 ↳ labels quoted verbatim
from parsed DST value labels and re-validated in-environment before
 ↳ estimation (frozen
fallback rules in contingent_decisions).
...

4.3 theta3 - H3 out-of-hours composition (confirmatory)

- **Model:** the 4.1 specification run separately with y = OOH contacts
(SPEC2 81-84, 89) and y = daytime GP contacts (SPEC2 70-80).
- **Primary quantity:** theta3 = (theta_OOH / mu_OOH,ctrl) - (theta_day /
mu_day,ctrl), where theta_x = mean event-study coefficient over the
peri-event window Q3(c)..Q2(c+1) for outcome x, and mu_x,ctrl = the
matched-control mean of x over the same window.
- **Inference:** cluster-bootstrap percentile CI, 999 draws, resampling
event workplaces, seed 20260831; Holm member; support requires theta3
> 0; floor = 5-point relative-excess gap.
- **Robustness:** repeat under the YDLTID-augmented OOH classifier.
Pre-committed interpretations distinguish theta_day >= 0

```

(disproportionately 00H excess) from  $\theta_{day} < 0$  (total-down/00H-up composition shift - scoreable support for H3 that simultaneously disconfirms a volume-increase reading of  $\theta_{day}$ ).

#### ## 4. ANALYSIS EXECUTION

Confirmatory tier first (Stage 4, frozen code, single run); exploratory tier after (Stage 5), segregated by heading. All estimation in R fixest via Rscript. Common estimator: outcome  $y_{iq}$  (winsorized quarterly GP contact count) on treated  $x$  event-quarter indicators for the 8 quarters  $Q1(c) \dots Q4(c+1)$ , matched-set fixed effects, cohort  $x$  calendar-quarter fixed effects, omitted baseline = the four quarters of year  $c-1$ ; observation weights: exposed 1, controls  $1/k_s$  within set; CR1 SEs clustered on event workplace (each control inherits its set's event cluster). Min-cell suppression is applied to all outputs BEFORE Holm reporting so suppression can never be outcome-selective.

##### ### 4.4 Family adjudication and reporting order

Holm-Bonferroni over the surviving confirmatory family (as gated at step 14) at family  $\alpha = 0.05$ . Report all surviving tests exactly once, regardless of results. Each confirmatory hypothesis closes with a boldface **\*\*Summary ( $H_k$ ).** verdict paragraph opening "Consistent with" / "Contrary to our hypothesis", per the pre-committed interpretation table in ``prereg.yaml`` (interpretations block - 16 rows covering positive, null, reversed, gate-fired, and partial patterns; apply them verbatim). Scoring attaches per the registered scoring\_rule: H1 confirmed only if  $\theta_{day}$  Holm-rejects positively AND the gated H1\_trend supports the growth rider; partially\_confirmed for flat elevation; reversed if negative. H2 confirmed only if  $\theta_{day2a}$  Holm-rejects positively AND F2 does not fire AND the  $\theta_{day2b}$  sign pattern is GRAFT-3-stable; reversed if  $\theta_{day2a}$  is significantly negative OR F2 fires. H3 confirmed on a positive Holm rejection (including the total-down/00H-up pattern); reversed if negative. Demoted or killed members attach no score; their trail is disclosed. Sensitivity suite (run once, one table): GRAFT-3; GRAFT-9 (re-estimate  $\theta_{day1}/\theta_{day2a}$  excluding sets whose control workplace acquires any terminal event in  $c+2$  or  $c+3$  - cohorts  $c \leq 2022$  for  $c+2$ ,  $c \leq 2021$  for  $c+3$ , truncation disclosed; registered direction note: exclusion should weakly INCREASE magnitudes, a decrease is flagged as diagnostic); un-winsorized outcomes; Poisson QML (``fepois``; if OLS and Poisson disagree in sign on any primary quantity, foreground the pre-committed OLS estimate, report both, flag the discordance in the Summary paragraph); possible-death/emigration exclusion (drop sets where the exposed member, or all controls, has zero SSSY rows across  $c \dots c+1$  AND no IDAN record in  $c+1$ ).

#### ## 5. EXHIBITS

Every figure and table the report commits to. Disclosure rules binding on EVERY released object: no statistic based on fewer than 5 persons or fewer than 3 workplaces; event-count tables rounded to the nearest 10; weekly trajectory figures pool  $\geq 100$  events and  $\geq 5,000$  person-weeks per plotted point; BRUHON aggregates pass a 50% single-workplace dominance check; released objects are limited to model coefficients with SEs/CIs, balance tables, calibration/benchmark tables, figure aggregates, rounded event counts, the power/MDE tables (both Stage-0 rounds and Stage 2) with the demotion trail, and the analysis/power scripts - never microdata, workplace identifiers (ARENR/CVRNR/LBNR), or person-level values; suppression applied BEFORE Holm reporting.

**\*\*Figures (matplotlib):\*\***

- **\*\*F1 - Quarterly event-study panel.\*\*** Treated x event-quarter coefficients with 95% CIs across Q1(c-2)..Q4(c+1) (placebo quarters, omitted-baseline marker, terminal-year and post-year quarters); panel (a) true arm, panel (b) spurious arm from the pooled model. Feeds: step-18 panels + 4.1/4.2 coefficient tables. Disclosure: coefficients only.
- **\*\*F2 - Weekly descriptive trajectory.\*\*** Pooled exposed-minus-matched-control weekly GP contact series around the closure-year boundary, with the GRAFT-2 week-0 step diagnostic annotated (step <= 50% of the total Q3(c)+Q4(c) excess; if exceeded, the figure carries the registered caution and quarter-level estimates are foregrounded). Feeds: step 15/17 weekly aggregates. Disclosure: >= 100 events and >= 5,000 person-weeks per point.
- **\*\*F3 - s(w) successor-share distribution (GRAFT-4).\*\*** Histogram of worker-flow successor shares across terminal events with the 0.30/0.70 thresholds and mid-zone mass; bimodality check exhibit. Feeds: step 7. Disclosure: bin counts rounded to 10; bins under min-cell suppressed.
- **\*\*F4 - OOH vs daytime relative excess (theta3).\*\*** The two relative-excess series over the peri-event window with the bootstrap CI of their gap; primary and YDLTID-augmented classifier side by side. Feeds: 4.3 outputs. Disclosure: coefficients only.
- **\*\*F5 - Matching balance (love plot).\*\*** SMDs before/after matching for all matching covariates and unmatched auxiliaries; 0.10 line marked. Feeds: step 13. Disclosure: aggregates only.
- **\*\*F6 - E1 SYSI-era replication (exploratory; skipable).\*\*** Quarterly SYSI-era analogue of F1(a) for cohorts 1992-2003, labeled exploratory, YDLANT-weighted proxy outcome. Feeds: step 22. Disclosure: as F1.

**\*\*Tables (booktabs LaTeX):\*\***

- **\*\*T1 - Event extraction and classification by cohort year.\*\*** Terminal events, true/spurious/ambiguous counts (rounded to 10), annual true-closure rate among private size-5+ workplaces, spurious/true ratio. Feeds: step 8.
- **\*\*T2 - Classification cross-validation.\*\*** Worker-flow classes vs IDAS-native flags (ENDNEDL/IDFREM/REGELFR/FRELFREM/ANDELUDS), full cross-tabulation + agreement rate; GRAFT-3 classification counts under 20/80 and 40/60. Feeds: step 8.
- **\*\*T3 - Cohort flow accounting.\*\*** Inclusion/exclusion counts per rule (tenure, sector, size, SIKGRUP, missing ANSAAR/pay, multi-cohort dedup), eligible exposed per cohort, match rate, ladder rung used, realized k distribution. Feeds: steps 9-13.
- **\*\*T4 - Matching balance.\*\*** Per-covariate SMDs (weighted 1/k), caliper compliance, realized mean k. Feeds: step 13.
- **\*\*T5 - Power and demotion trail.\*\*** Stage-0 power results (BOTH rounds: the voided round-1 run and the remediated re-run) and the Stage-2 blind realized-power/MDE table with binding N\*, realized sets, and any demotions with the fixed-point recomputation. Feeds: deposited script + step 14. Published regardless of outcome.
- **\*\*T6 - Outcome-neutral check panel.\*\*** All eleven registered checks (coverage benchmark by year, linkage floor, closure-rate plausibility, classification agreement, GRAFT-4, balance/match rate, c-2 placebo, HONUGE lag calibration, floor/ceiling + seasonality, GRAFT-2 step, OOH-share benchmark) with thresholds and PASS/FAIL. Feeds: steps 8, 13, 16, 17, 19, 21 + F2/F3 diagnostics.
- **\*\*T7 - Confirmatory results.\*\*** theta1, theta2a, theta3: point estimates (contacts/person-quarter and % of baseline), 95% CIs, two-sided p, Holm-adjusted verdicts; the F2 gate row (theta2b, its

- one-sided 95% upper bound, trigger evaluation); the gated BRUHON fee secondary (if computed); c-2 placebo row; realized MDEs beside every null with the CI-ruled-out magnitudes. Feeds: 4.1-4.4.
- \*\*T8 - Sensitivity suite.\*\* GRAFT-3 thresholds; GRAFT-9 forward-window exclusion (with truncation note); un-winsorized; Poisson QML; death/emigration exclusion - one table, run once. Feeds: 4.4.
  - \*\*T9 - Exploratory tier.\*\* H1\_trend (if gated in; otherwise the row states the gate was not met), H2\_diff with quarter-by-quarter differential profile, E3 fee and subgroup analyses (within-family Holm), E1 summary. Every row labeled exploratory. Feeds: 4.5.
  - \*\*T10 - Attrition and later-presence accounting.\*\* E4 year-c separation shares by event class (ANSXFREM-based, post-treatment-conditioned label and selection caveat) and the GRAFT-6 later-presence rates with the 2pp trigger status. Feeds: steps 20 + 4.5 E4.

---

## Current source hypotheses (all fields)

```

```yaml
- id: H3
  tier: confirmatory
  statement_initial: 'H3 (distress composition): Compared to the
    ↳ out-of-hours share of contemporaneous
      contacts among their matched controls, closure-exposed workers'
    ↳ peri-event excess contacts
      will contain a disproportionately higher share of out-of-hours and
    ↳ off-schedule services
      (YDLTID time-of-day codes; VAGTOMR in the SYSI era), consistent with
    ↳ unplanned, distress-driven
      care-seeking rather than deferred routine care.'
  statement_source: idea bank entry DS1.1, hypothesis 3 of 3; wording
    ↳ verbatim, frozen at
      bank creation. Construct unchanged from the bank statement
    ↳ (relative-excess contrast,
      sign-unambiguous under the Hamad total-down/OOH-up world); only the
    ↳ power basis was recalibrated
      to the published Danish OOH share (see remediation_record).
  estimand: 'Unchanged construct (already repaired against
    ↳ share-among-utilizers pathologies;
      sign-unambiguous under the Hamad total-down/OOH-up world, S6): within
    ↳ true-closure stacks,
      the difference in RELATIVE peri-event excess (event-study coefficient
    ↳ divided by matched-control
      same-window mean) between out-of-hours GP contacts (SPEC2 in
    ↳ {81,82,83,84,89}) and daytime
      GP contacts (SPEC2 in {70..80}), peri-event window Q3(c)..Q2(c+1).
    ↳ What changed is ONLY
      the power basis: the Stage-0 simulation's OOH base rate is
    ↳ recalibrated from the self-declared
      over-conservative 3% to 8.0% of daytime GP volume (7.4% of all GP
    ↳ contacts) - the conservative
      end of the published Danish out-of-hours share (~8-12%: Christensen &
    ↳ Olesen 1998, BMJ
      316:1502; Moth, Flarup, Christensen, Olesen & Vedsted, LV-KOS 2011
    ↳ out-of-hours contact
      studies; the audit-verified diagnosis records that at ~10% the
    ↳ test's z roughly doubles
      and clears).'
```

primary_quantity: 'theta3 = (theta_OOH / mu_OOH,ctrl) - (theta_day /
↳ mu_day,ctrl), where
theta_x = mean event-study coefficient over the peri-event window for
↳ outcome x and mu_x,ctrl
= matched-control mean of x over the same window; cluster-bootstrap
↳ percentile CI (999
draws, seed 20260831, clusters = event workplaces). Predicted
↳ direction theta3 > 0. Holm
family member. Floor effect UNCHANGED at a 5-point relative-excess
↳ gap, below the Hamad/Modrek/Cullen
(2015) ER-up/outpatient-down composition shifts. Power path: z scales
↳ ~ sqrt(OOH base
rate), 11.63 moving 3% -> 8%, plus the N and 1:3-matching levers;
↳ provisional gate N*_H3
= 90,000 true matched sets; an outcome-neutral check verifies the
↳ realized control OOH
share sits in the published range before the confirmatory run.'
direction: 'positive (theta3 > 0: OOH relative excess exceeds daytime
↳ relative excess)'
comparator: matched controls' contemporaneous out-of-hours and daytime
↳ GP contact means
over the same peri-event window (each arm's excess scaled by its own
↳ control mean before
differencing)
exposure_variable: true-closure exposure over the peri-event window
↳ Q3(c)..Q2(c+1)
outcome_variable: out-of-hours GP contacts (SPEC2 81-84, 89; robustness
↳ classifier adds
YDLTID 2-5, 7-9) versus daytime GP contacts (SPEC2 70-80), as
↳ relative peri-event excesses
forecast:
point: +2.5-point gap in relative peri-event excess (OOH minus
↳ daytime)
interval_90: '[-3.0, +9.0] points'
basis: 'Agent forecast, recorded before any data access and kept
↳ separate from any future
expert forecasts (collaboration roadmap rule). Upward anchors: the
↳ strongest prior art
found by the search - emergency use rose while outpatient use fell
↳ under layoff-threat
insecurity (Hamad, Modrek & Cullen 2015) - and the low access cost
↳ of the Danish laegevagts.
Downward anchors: OOH contact may track genuine acute need rather
↳ than distress, Danish
OOH gatekeeping triages non-urgent contacts back to daytime care,
↳ and if the peri-event
excess is dominated by deferred routine care (using remaining
↳ employment time for scheduled
visits) the gap reverses - the reversal is pre-registered as
↳ interpretable. My central
forecast sits below the 5-point floor; the composition tilt is the
↳ hypothesis I hold
with the widest uncertainty.'
null_substantive: 'The peri-event excess is compositionally
↳ unremarkable: the out-of-hours
share of excess contacts does not differ from the daytime share by 5
↳ points or more (bootstrap
CI converted to ruled-out gaps against the realized MDE), consistent
↳ with scheduled as
much as distress-driven care.'

```

kill_criterion: 'KILL: if OOH contacts (SPEC2 81-84, 89) are < 2% of
→ all control-group GP
  contacts in the baseline year (composition not estimable), or the OOH
  → classifier is unresolvable
  (YDLTID missing/'not relevant' on > 20% of peri-event GP rows and
  → SPEC2 OOH codes absent).
CONFIRMATORY DEMOTION GATE: realized control OOH share < 5% with
→ blind re-simulated power
< 0.95 at the 5-point floor, or realized matched sets < N*_H3
→ (provisional 90,000) ->
  theta3 demotes to exploratory pre-outcome.'
...

```

Current source interpretation rows

```

```yaml
- hypothesis_id: H3
 outcome_pattern: theta3 > 0, Holm-significant, with theta_day >= 0
 conclusion: We will conclude that the peri-event excess is
 → disproportionately out-of-hours,
 consistent with unplanned, distress-driven care-seeking rather than
 → deferred routine care
 (composition claim only; no diagnosis data exist in scope, so no
 → health-state claim is
 made).
- hypothesis_id: H3
 outcome_pattern: theta3 > 0, Holm-significant, with theta_day < 0
 → (total down, OOH up)
 conclusion: We will conclude that closure exposure is associated with a
 → compositional shift
 - daytime contact falls while out-of-hours contact rises - matching
 → the forgone-routine-care-plus-acute-distress
 pattern in the prior insecurity literature (Hamad, Modrek & Cullen
 → 2015); this pattern
 supports H3's composition claim while disconfirming any
 → volume-increase reading of theta1
 (S6).
- hypothesis_id: H3
 outcome_pattern: theta3 not Holm-significant (NULL)
 conclusion: We will conclude that we have no evidence that the
 → peri-event excess is compositionally
 tilted toward out-of-hours care, and report the relative-excess
 → differentials the 95%
 bootstrap CI rules out against the realized MDE; any theta1 excess is
 → then compositionally
 unremarkable - consistent with scheduled as much as distress-driven
 → care - quoted with
 the registered caveat that the primary classifier counts SPEC2
 → 71/73/75 evening-GP services
 as daytime, which attenuates theta3 and correspondingly weakens the
 → null bound (frozen_parameters.spec2_evening_gp_disclosure).
- hypothesis_id: H3
 outcome_pattern: theta3 < 0, Holm-significant (reversed)
 conclusion: We will conclude - contrary to our hypothesis - that the
 → peri-event excess is
 disproportionately daytime/scheduled, consistent with planned or
 → deferred routine care
 (e.g., using remaining employment time or sick-leave access for
 → scheduled visits) rather
 than acute distress; H3 is disconfirmed.

```

```

...

Source family scoring and limitations

```yaml
scoring_rule: 'Fixed in advance per the collaboration''s
  ↳ registered-report conventions (trajectory
    fields): at Stage 2 an outcome attaches to each confirmatory
      ↳ hypothesis''s frozen statement_initial
    as confirmed | partially_confirmed | reversed, adjudicated strictly
      ↳ against the pre-committed
    interpretation table in this protocol. H1 scores confirmed only if
      ↳ theta1''s Holm-adjusted
    two-sided test rejects with the point estimate in the predicted
      ↳ direction AND the gated
    exploratory H1_trend contrast supports the frozen statement''s growth
      ↳ rider (beta_Q4c -
    beta_Q1c > 0 with 95% CI excluding zero); it scores partially_confirmed
      ↳ when theta1 Holm-rejects
    in the predicted direction but the growth rider is unsupported (flat
      ↳ elevation - the pre-committed
    reading given annual event dating) or the trend gate is not computed;
      ↳ reversed if theta1
    is Holm-significant in the negative direction. H2 scores confirmed only
      ↳ if theta2a Holm-rejects
    in the predicted direction AND falsification gate F2 does not fire AND
      ↳ the theta2b sign
    pattern is stable across the GRAFT-3 successor-share thresholds;
      ↳ partially_confirmed when
    theta2a Holm-rejects but the comparative quantifier is untestable
      ↳ (spurious-arm kill - verdict
    capped by pre-commitment at ''partially consistent - comparative
      ↳ quantifier untestable'');
    reversed if theta2a is Holm-significant in the negative direction OR
      ↳ gate F2 fires (the
    frozen statement''s comparative reading is disconfirmed by a non-inert
      ↳ placebo regardless
    of theta2a - registered as disconfirming, never reframed). H3 scores
      ↳ confirmed only if theta3''s
    Holm-adjusted test rejects in the predicted direction (including the
      ↳ total-down/00H-up pattern,
    pre-registered as scoreable support for the composition claim);
      ↳ reversed if Holm-significant
    in the negative direction. A Holm-nonsignificant confirmatory result
      ↳ attaches no confirmation:
    it is reported as a bounded informative null - the 95% CI converted
      ↳ into the excess magnitudes
    the data rule out against realized MDEs - never ''accepted''. A member
      ↳ demoted by the Stage-2
    realized-N gate or killed by its registered criterion attaches no
      ↳ score; the demotion/kill
    trail is disclosed. Significant effects in the non-predicted direction
      ↳ are disconfirming,
    never supportive. Exploratory quantities (H1_trend, H2_diff, E1, E3,
      ↳ E4, and any demoted
    member) attach outcomes only under the exploratory heading and are
      ↳ never scored as preregistered-confirmed.'
assumptions:
- SSSY (2005-2025) and SYSI (1990-2005) capture the universe of publicly
  ↳ reimbursed primary-care

```

- billing for insured residents; absence of rows means no billed care,
 - not missingness
- KONTAKT is a row-level contact marker whose sum yields contact counts
 - (in-environment validation
 - with a frozen fallback definition if not)
- ENDNEDL/IDFREM/IDTILB/REGELFR/FRELFREM semantics follow Statistics
 - Denmark's IDA workplace-continuity
 - documentation, cited directly (no DOI exists for these conventions, per
 - boundary directive);
 - the primary classification rests on the worker-flow measure $s(w)$ from
 - IDAN, so misreading
 - an IDAS flag degrades only the cross-check
- November-status annual attachment measures workplace membership;
 - intra-year attachment changes
 - are unobservable ex ante and the frozen c-1 anchor is the design's
 - answer to that - with
 - the corollary, named as a design threat, that early year-c separations
 - feed realized post-job-loss
 - billing into the pre-boundary window under intention-to-follow, a
 - channel the E4 separation
 - shares and the registered theta1 attachment split bound but cannot
 - remove
- ALDERIMP/KOENIMP (DST-imputed age/sex on SSSY rows) are accurate enough
 - for stratified matching;
 - persons absent from SSSY in c-3..c-1 are handled by explicit unknown
 - strata
- HONUGE settlement lags are predominantly ≤ 4 weeks (calibrated on
 - BEHANDLINGSDATO 2021+;
 - guard-band contingency if not), and lag structure does not differ
 - between exposed and matched
 - controls within the same calendar quarters
- Death and emigration are unobservable in scope; their post-event
 - incidence is assumed small;
 - the deflating direction is conservative for the pre-boundary claim and
 - attenuating for theta2a,
 - with the GRAFT-6 trigger bounding the post-boundary side
- Nominal-fee comparisons within cohort x calendar-quarter cells require
 - no deflator
- Danish institutional stability of GP billing categories (SPEC2 70-84,
 - 89) across 2007-2024
 - at the 2-digit level
- 'PUBLISHED-BASIS SIZING ASSUMPTIONS (new, each with a registered
 - failure path): (i) Danish
 - establishment closure rates among private size-5+ workplaces are in the
 - published Erhvervsdemografi-consistent
 - 1.5-3% band, supporting 150,000-300,000 exposed workers over 15 cohorts
 - - failure path:
 - the pre-outcome realized-N gate demotes rather than stretches; (ii) the
 - Danish OOH share
 - of GP contact volume is in the published 8-12% band (Christensen &
 - Olesen 1998; LV-KOS 2011
 - studies), simulated at the conservative 8% - failure path: the
 - published-range benchmark
 - check plus realized-share re-simulation, then demotion or kill; (iii)
 - the control pool supports
 - 1:3 matching at $\geq 90\%$ match rate - failure path: coarsening ladder,
 - realized-k accounting
 - in the blind re-simulation, then the gate'
- theta2a's 4-quarter averaging gain over theta1 survives serial
 - correlation in post-year

- shocks; a registered serial-correlation discount is built into the
 - ↳ Stage-0 re-simulation
 rather than assumed away
- The persistent-vs-transient overdispersion split (persistent frailty
 - ↳ var 1.25, transient
 log-var 0.25) is retained unchanged from the first-round simulation -
 - ↳ deliberately NOT revised
 in the direction that would raise power, so the remediation cannot be
 - ↳ accused of tuning
 the DGP
- failure_modes:
 - 'Realized event or worker counts far below schema-derived expectations
 - ↳ (documentation-only
 corpus; no realized Ns exist) - the design now bears this risk openly:
 - ↳ the Stage-2 realized-N
 gate demotes underpowered members loudly and pre-outcome instead of
 - ↳ letting an underpowered
 confirmatory test run; a sparse-closure Denmark kills or demotes
 - ↳ cleanly rather than quietly'
 - KONTAKT or ENDNEDL semantics differ from assumption - both carry frozen
 - ↳ fallback rules,
 but level shifts in outcomes or event dating would force reported
 - ↳ recalibration and consume
 credibility
 - HONUGE settlement lag materially longer than 4 weeks - the
 - ↳ quarter-level estimands survive,
 but the weekly-resolution headline weakens to quarterly resolution in
 - ↳ practice (GRAFT-2
 makes this explicit and pre-committed)
 - ANDELOPS direction-of-absorption semantics may differ from the idea's
 - ↳ gloss - classification
 routes through IDAN worker flows, so this only weakens the IDAS
 - ↳ cross-check, not the H2
 machinery
 - Spurious-closure events too heterogeneous for a clean placebo - a
 - ↳ non-inert placebo now
 shows up as gate F2 firing (or exploratory H2_diff reversal), which
 - ↳ caps H2's verdict by
 pre-commitment instead of silently contaminating a differential test
 - Realized OOH share below the published range - caught by the
 - ↳ published-range benchmark check
 before estimation; H3 demotes or dies pre-outcome instead of running
 - ↳ underpowered
 - Realized N lands below the full-expectation lower bound - the gate
 - ↳ demotes affected members
 to exploratory pre-outcome; the study survives with a smaller
 - ↳ confirmatory family (or none,
 in which case it reports as registered exploratory estimation with the
 - ↳ demotion trail published
 - the first-round outcome, now a contingency rather than the design)
 - Unobservable post-closure death/emigration deflates post-boundary
 - ↳ billing - direction stated,
 sensitivity pre-specified, GRAFT-6 2pp trigger forces best/worst-case
 - ↳ bounds; no in-scope
 fix exists
 - Public/private sector misclassification via modal ARB_SEKTORKODE for
 - ↳ multi-establishment
 firms - attenuates contrasts; share bounded in the balance tables; if
 - ↳ ARB_SEKTORKODE itself

```

proves absent in-environment (it is unconfirmed in the recorded
  ↳ Forskningservice inventory),
the frozen ARB_HOVED_BRA_DB07 fallback in contingent_decisions
  ↳ substitutes, with counts
reported
- Multi-job workers' primary-job assignment may misattribute exposure -
  ↳ deterministic rule
  ↳ biases toward null; share reported
- 1:3 matching without replacement may exhaust thin strata - realized k <
  ↳ 3 accounted honestly
  ↳ in the blind re-simulation rather than assumed; coarsening ladder
  ↳ pre-committed
- The SYSI-era exploratory replication may be infeasible - bounded as
  ↳ skipable Stage-5 work
  ↳ that cannot contaminate the confirmatory tier
- Readers may over-read gate F2 as proof of causal job-loss effects - the
  ↳ associational ceiling,
  ↳ the 'consistent with' vocabulary cap, and the pre-committed F2-fired
  ↳ interpretation guard
the prose
...

## Source disclosure rules

```yaml
environment: DST Forskningservice (no prior data access; gatekeeper
 ↳ certification per RR
 ↳ secondary-data rules). All analysis occurs inside the environment; only
 ↳ aggregate outputs
 ↳ leave. No microdata, workplace identifiers (ARBNR/CVRNR/LBNR), or
 ↳ person-level values are
 ↳ released.
minimum_cell_size: 5
minimum_cell_rules: no released statistic based on fewer than 5 persons
 ↳ or fewer than 3 workplaces;
 ↳ event-count tables rounded to the nearest 10; weekly trajectory figures
 ↳ pool >= 100 events
 ↳ and >= 5,000 person-weeks per plotted point; fee (BRUHON) aggregates
 ↳ subject to a 50% single-workplace
 ↳ dominance check; min-cell suppression applied BEFORE Holm reporting so
 ↳ suppression can never
 ↳ be outcome-selective
outputs_leaving_environment:
- model coefficients with SEs/CIs
- matching balance tables and realized-k distribution
- calibration and benchmark tables (coverage, linkage, HONUGE lag, OOH
 ↳ share)
- figure aggregates (weekly trajectory series under the pooling rules)
- event counts rounded to the nearest 10
- Stage-0 power/MDE tables - BOTH rounds (failed and remediated) - and
 ↳ the Stage-2 blind re-simulation
 ↳ output with the demotion trail
- the c-2 placebo, gate F2 outputs, GRAFT-2/4 diagnostics, GRAFT-6
 ↳ later-presence rates, E4
 ↳ attrition panel, and all outcome-neutral check outputs (the full audit
 ↳ trail exits under
 ↳ the same rules)
- the analysis and power scripts
statement: 'All analysis occurs inside DST Forskningservice; only
 ↳ aggregate outputs leave.

```

```

Rules applied to every released object: (1) no statistic based on fewer
→ than 5 persons or
fewer than 3 workplaces; (2) event-count tables rounded to the nearest
→ 10; (3) weekly trajectory
figures pool >= 100 events and >= 5,000 person-weeks per plotted point;
→ (4) fee (BRUHON)
statistics released only as cell means/coefficients over min-N cells,
→ with a dominance check
that no single workplace contributes > 50% of any released fee
→ aggregate; (5) released objects
are limited to: model coefficients with SEs/CIs, balance tables,
→ calibration/benchmark tables,
figure aggregates, event counts (rounded), the Stage-0 (both rounds)
→ and Stage-2 power/MDE
tables with the demotion trail, and the analysis/power scripts - never
→ microdata, workplace
identifiers (ARBNR/CVRNR/LBNR), or person-level values; (6) min-cell
→ suppression is applied
before Holm reporting so that suppression can never be
→ outcome-selective; (7) the c-2 placebo,
gate F2 outputs, the GRAFT-2/4 diagnostics, the GRAFT-6 later-presence
→ rates, the E4 attrition
panel, and all outcome-neutral check outputs are released under the
→ same rules, ensuring
the full audit trail exits the environment.'
...

```

## Pending registered experiment

No Danish study outcome has been observed for this bundle. See the complete registration above.

## atom-000007: Historical billing pattern

### How does the source H1 analogue appear in the separate SYSI-era cohort?

#### Source fidelity and implementation order

This is a format conversion of DS1.1, not a new study or independent acceptance. Source audit: minor\_revisions. No observed Danish outcomes are supplied. The source execution plan explicitly states that final prereg.yaml governs disagreements. Exact current YAML selections below therefore govern older execution-plan excerpts; these excerpts are retained as implementation detail, not a competing registration. See docs/provenance/source-discrepancies.md. The complete shared data/validity contracts are embedded in Atoms 1-3; read those native protocols when executing a downstream Atom, even without docs/. Do not substitute a scientific default for an unresolved source contradiction.

Preserve source stages 0 -> 1 -> 2 -> 3 -> 4 -> 5 inside the implementation. Source Stage-0 simulations are historical design evidence, not study outcomes; do not silently redo or improve them. Gate 14 must be evaluated from baseline data before post-baseline outcomes. All surviving Stage-4 quantities form ONE Holm family. The

three question-level Atoms may share the same frozen family fit and tidy results; do not apply Holm separately per Atom or adjudicate a member before the complete surviving family's p-values exist. Cache/reuse identical governed intermediate results with provenance when available; a direct-dispatch call must obtain or compute its exact source-defined inputs, not pretend its predecessors succeeded. Stage-5 internal riders remain exploratory, after the Stage-4 family is frozen, even when housed in H1/H2.

Reference files under docs/execution/ are the ORIGINAL SKELETONS, not tested turnkey implementations. Adapt paths and complete engineering in a governed workspace; preserve source scientific rules and disclose unresolved choices. Write a result record with component estimates, uncertainty, family eligibility, source interpretation row, diagnostics and boundary/skip dispositions. Released outputs follow source disclosure rules; confidential rows stay governed. Local native states summarize the source judgment; neither a state name nor source power forecasts prove equivalence or causal identification. A bounded claim requires its actual interval and registered rule. Retain unresolved source wording conflicts; do not convert mere nonsignificance into scientific null.

## Frozen scope and vocabulary

```
claim_class_ceiling: associational
statement: 'Associational - matched event-study contrasts, a contrastive
↳ falsification gate,
 and a composition contrast; descriptive for trajectory panels and the
 ↳ bounded priority claim.
Binding vocabulary rules: 'is associated with' / 'consistent with' /
↳ 'suggests' only;
'effect', 'impact', 'causes' prohibited throughout; anticipation
↳ always worded as
'pre-boundary excess' relative to the annually dated closure-year
↳ boundary, never as announcement-dated
anticipation (S8); no phenomenon-discovery or Danish-first framing (S1-S3);
↳ H2 novelty confined
to the placebo's application to health billing (S7). H2 verdict semantics
↳ are fixed in
advance: 'consistent with H2' requires BOTH the confirmatory core
↳ (theta2a Holm-significant,
predicted direction) AND a passing falsification gate F2; the strict
↳ comparative magnitude
is only ever reported as an exploratory estimate with its CI. Exploratory
↳ results (H1_trend,
H2_diff, E1, E3, E4, and any gate-demoted member) carry the word
↳ 'exploratory' in the
heading and in every sentence stating them. The prose label never exceeds
↳ the evidence label.'
frozen_parameters:
seed: 20260831
gp_code_list: daytime GP = SPEC2 in {70..80}
ooh_code_list: out-of-hours GP = SPEC2 in {81, 82, 83, 84, 89}
ooh_robustness_classifier: adds YDLTID in {2, 3, 4, 5, 7, 8, 9}
spec2_evening_gp_disclosure: 'Registered choice, disclosed: the parsed
↳ value labels mark
 SPEC2 71 ("General practitioner, consultation, evening etc."), 73
 ↳ ("General practitioner,
```

telephone consultation, evening etc."), and 75 ("General practitioner,  
 ↪ home visit, evening  
 etc.") as explicitly off-schedule GP services; they nonetheless remain on  
 ↪ the daytime  
 side of the PRIMARY classifier because H3's out-of-hours construct and  
 ↪ its published  
 power basis are keyed to the laegevagt out-of-hours service (SPEC2 81-84,  
 ↪ 89), the quantity  
 the published Danish 8-12% share describes. Attenuation direction, both  
 ↪ readings: counting  
 evening GP as daytime attenuates theta3 toward zero - conservative  
 ↪ against the positive  
 H3 conclusion, anti-conservative for the null-bounding "compositionally  
 ↪ unremarkable"  
 reading, whose Stage-2 statement therefore carries a mandatory caveat.  
 ↪ In-environment  
 validation must verify that the YDLTID-augmented robustness classifier  
 ↪ captures SPEC2  
 71/73/75 rows (their evening time codes 2-5); if YDLTID proves  
 ↪ unpopulated on those rows,  
 the robustness classifier - and only the robustness classifier - is  
 ↪ amended mechanically  
 and with disclosure to include SPEC2 71/73/75 directly. The SPEC2  
 ↪ 71/73/75 volume share  
 is reported, and both classifiers appear side by side in Table T7.'  
 successor\_share\_thresholds: 'true closure: ENDNEDL final closure AND  
 ↪ worker-flow max-successor  
 share s(w) < 0.30; spurious closure: forward identity via  
 ↪ IDFREME/REGELFR/FRELFREME AND  
 s(w) >= 0.70; ambiguous mid-zone excluded; GRAFT-3 sensitivities  
 ↪ 0.20/0.80 and 0.40/0.60'  
 quarter\_mapping: 'ISO week of HONUGE: 1-13 / 14-26 / 27-39 / 40-52; ISO  
 ↪ week 53 to Q4'  
 winsorization: person-quarter outcomes at the 99.9th percentile within era  
 ↪ x calendar quarter  
 matching: 1:3 nearest-neighbor Mahalanobis within exact strata, without  
 ↪ replacement across  
 sets, target k = 3, minimum k = 1; calipers |dif baseline contacts| <= 3  
 ↪ and |dif log(1+BRUHON)|  
 <= 0.5; exposed processed in seeded random order (seed 20260831)  
 software: R >= 4.3, fixest::feols (OLS) and fixest::fepois (Poisson QML  
 ↪ robustness); versions  
 frozen in the registered script  
 note: Code lists and thresholds frozen before any data access; code labels  
 ↪ quoted verbatim  
 from parsed DST value labels and re-validated in-environment before  
 ↪ estimation (frozen  
 fallback rules in contingent\_decisions).

22. **E1 legacy pipeline (Stage 5; skippable).** Input: IDAS/IDAN 1990-2005  
 (legacy keys ARBGNR/ARBSTK/LBNR), SYSI 1990-2005. SYSI-era replication  
 of the H1 contrast for cohorts c in 1992-2003 (baseline 1991+, post-year <= 2004):  
 era GP/OOH classification from SPECIALE (era code list frozen at Stage 0) with  
 VAGTOMR (1997-2005) as the era OOH marker; contacts proxied by service-row  
 counts weighted by YDLANT; matching WITHOUT age/sex (SYSI lacks them),  
 using legacy job characteristics (STILL, ANSDAGE, TIMELON, TILKNYT) and  
 SIKREKOM/ SIKREAMT patient geography (both used only here; inventory-only

disclosure carried). Output: `analysis/e1_panel.parquet`. Check: skip E1 entirely — reported as skipped, not failed — if SYSI-era extraction yields  $< 300$  true closures or the era’s coverage benchmark fails. Never pooled with the confirmatory era.

---

#### 4.5 Exploratory tier (Stage 5; heading-segregated; Holm within family as severity aid)

- **H1\_trend (gated):**  $\beta_{Q4c} - \beta_{Q1c}$  from the true-arm model, 95% CI, CR1 clusters — computed ONLY if  $\theta_1$  Holm-rejected positively; otherwise the quarterly trajectory is shown descriptively with no contrast statistic and no timing-shape language in the Summary.
- **H2\_diff:**  $\theta_2\_diff = \text{mean}_k[\beta_{true,k} - \beta_{spurious,k}]$  over the post-year quarters from the pooled model, 95% CI, plus the quarter-by-quarter differential profile (the “more persistent” rider). A flagged estimate with  $\text{spurious} > \text{true}$  triggers the pre-committed non-inert-placebo interpretation and reinforces whatever F2 showed.
- **E1:** SYSI-era analogue of  $\theta_1$  from step 22, 95% CI, descriptive replication only; never pooled across the vintage break.
- **E3:** BRUHON fee analogues of  $\theta_{2a}/\theta_2\_diff$  beyond the gated secondary; heterogeneity of  $\theta_1/\theta_{2a}$  by sex (KOENIMP), age band (ALDERIMP), and baseline-year wage decile (BREDDT\_LOEN\_BELOEB); Holm within this family.
- **E4:** arm-comparative attrition panel — share of each frozen November c-1 cohort no longer attached to the event workplace by November of year c (IDAN year-c records; ANSXFREM as the forward-separation flag), by event class; explicitly labeled post-treatment-conditioned descriptives with a selection caveat; no test.
- No exploratory result can revise any confirmatory verdict or Stage-4 language. Every sentence stating an exploratory result contains the word “exploratory”.

### 4. ANALYSIS EXECUTION

Confirmatory tier first (Stage 4, frozen code, single run); exploratory tier after (Stage 5), segregated by heading. All estimation in R fixest via Rscript. Common estimator: outcome  $y_{iq}$  (winsorized quarterly GP contact count) on treated  $x$  event-quarter indicators for the 8 quarters  $Q1(c)..Q4(c+1)$ , matched-set fixed effects, cohort  $x$  calendar-quarter fixed effects, omitted baseline = the four quarters of year c-1; observation weights: exposed 1, controls  $1/k_s$  within set; CR1 SEs clustered on event workplace (each control inherits its set’s event cluster). Min-cell suppression is applied to all outputs BEFORE Holm reporting so suppression can never be outcome-selective.

### 5. EXHIBITS

Every figure and table the report commits to. Disclosure rules binding on EVERY released object: no statistic based on fewer than 5 persons or fewer than 3 workplaces; event-count tables rounded to the nearest 10; weekly trajectory figures pool

$\geq 100$  events and  $\geq 5,000$  person-weeks per plotted point; BRUHON aggregates pass a 50% single-workplace dominance check; released objects are limited to model coefficients with SEs/CIs, balance tables, calibration/benchmark tables, figure aggregates, rounded event counts, the power/MDE tables (both Stage-0 rounds and Stage 2) with the demotion trail, and the analysis/power scripts — never microdata, workplace identifiers (ARBNR/CVRNR/LBNR), or person-level values; suppression applied BEFORE Holm reporting.

#### Figures (matplotlib):

- **F1 — Quarterly event-study panel.** Treated  $x$  event-quarter coefficients with 95% CIs across  $Q1(c-2)..Q4(c+1)$  (placebo quarters, omitted-baseline marker, terminal-year and post-year quarters); panel  
(a) true arm, panel (b) spurious arm from the pooled model. Feeds: step-18 panels + 4.1/4.2 coefficient tables. Disclosure: coefficients only.
- **F2 — Weekly descriptive trajectory.** Pooled exposed-minus- matched-control weekly GP contact series around the closure-year boundary, with the GRAFT-2 week-0 step diagnostic annotated (step  $\leq 50\%$  of the total  $Q3(c)+Q4(c)$  excess; if exceeded, the figure carries the registered caution and quarter-level estimates are foregrounded). Feeds: step 15/17 weekly aggregates. Disclosure:  $\geq 100$  events and  $\geq 5,000$  person-weeks per point.
- **F3 —  $s(w)$  successor-share distribution (GRAFT-4).** Histogram of worker-flow successor shares across terminal events with the 0.30/0.70 thresholds and mid-zone mass; bimodality check exhibit. Feeds: step 7. Disclosure: bin counts rounded to 10; bins under min-cell suppressed.
- **F4 — OOH vs daytime relative excess (theta3).** The two relative-excess series over the peri-event window with the bootstrap CI of their gap; primary and YDLTID-augmented classifier side by side. Feeds: 4.3 outputs. Disclosure: coefficients only.
- **F5 — Matching balance (love plot).** SMDs before/after matching for all matching covariates and unmatched auxiliaries; 0.10 line marked. Feeds: step 13. Disclosure: aggregates only.
- **F6 — E1 SYSI-era replication (exploratory; skippable).** Quarterly SYSI-era analogue of F1(a) for cohorts 1992–2003, labeled exploratory, YDLANT-weighted proxy outcome. Feeds: step 22. Disclosure: as F1.

#### Tables (booktabs LaTeX):

- **T1 — Event extraction and classification by cohort year.** Terminal events, true/spurious/ambiguous counts (rounded to 10), annual true-closure rate among private size-5+ workplaces, spurious/true ratio. Feeds: step 8.
- **T2 — Classification cross-validation.** Worker-flow classes vs IDAS-native flags (ENDNEDL/IDFREM/REGELFR/FRELFREM/ANDELUDS), full cross-tabulation + agreement rate; GRAFT-3 classification counts under 20/80 and 40/60. Feeds: step 8.

- **T3 — Cohort flow accounting.** Inclusion/exclusion counts per rule (tenure, sector, size, SIKGRUP, missing ANSAAR/pay, multi-cohort dedup), eligible exposed per cohort, match rate, ladder rung used, realized k distribution. Feeds: steps 9–13.
- **T4 — Matching balance.** Per-covariate SMDs (weighted 1/k), caliper compliance, realized mean k. Feeds: step 13.
- **T5 — Power and demotion trail.** Stage-0 power results (BOTH rounds: the voided round-1 run and the remediated re-run) and the Stage-2 blind realized-power/MDE table with binding  $N^*$ , realized sets, and any demotions with the fixed-point recomputation. Feeds: deposited script  
– step 14. Published regardless of outcome.
- **T6 — Outcome-neutral check panel.** All eleven registered checks (coverage benchmark by year, linkage floor, closure-rate plausibility, classification agreement, GRAFT-4, balance/match rate, c-2 placebo, HONUUGE lag calibration, floor/ceiling + seasonality, GRAFT-2 step, OOH-share benchmark) with thresholds and PASS/FAIL. Feeds: steps 8, 13, 16, 17, 19, 21 + F2/F3 diagnostics.
- **T7 — Confirmatory results.** theta1, theta2a, theta3: point estimates (contacts/person-quarter and % of baseline), 95% CIs, two-sided p, Holm-adjusted verdicts; the F2 gate row (theta2b, its one-sided 95% upper bound, trigger evaluation); the gated BRUHON fee secondary (if computed); c-2 placebo row; realized MDEs beside every null with the CI-ruled-out magnitudes. Feeds: 4.1–4.4.
- **T8 — Sensitivity suite.** GRAFT-3 thresholds; GRAFT-9 forward-window exclusion (with truncation note); un-winsorized; Poisson QML; death/emigration exclusion — one table, run once. Feeds: 4.4.
- **T9 — Exploratory tier.** H1\_trend (if gated in; otherwise the row states the gate was not met), H2\_diff with quarter-by-quarter differential profile, E3 fee and subgroup analyses (within-family Holm), E1 summary. Every row labeled exploratory. Feeds: 4.5.
- **T10 — Attrition and later-presence accounting.** E4 year-c separation shares by event class (ANSXFREM-based, post-treatment- conditioned label and selection caveat) and the GRAFT-6 later-presence rates with the 2pp trigger status. Feeds: steps 20 + 4.5 E4.

---

## Current source hypotheses (all fields)

```
- id: E1
 tier: exploratory
 statement_initial: null
 statement_source: no frozen bank statement - protocol-defined exploratory
 ↪ module (SYSI-era
 descriptive replication of the H1 contrast)
 estimand: SYSI-era replication of the H1 contrast for closure cohorts c in
 ↪ 1992-2003 (baseline
 1991+, post-year <= 2004; legacy keys ARBGNR/ARBSTK/LBNR; matching
 ↪ without age/sex, which
 SYSI lacks; SIKREKOM/SIKREAMT and VAGTOMR used only here and disclosed as
 ↪ verified in
```

the Forskningservice inventory but absent from public TIMES/HKV  
 ↳ metadata). Never pooled  
 with the 2009-2023 era across the 2007/08 IDA vintage break.  
 primary\_quantity: SYSI-era analogue of theta1 (contacts proxied by  
 ↳ service-row counts weighted  
 by YDLANT; KONTAKT does not exist pre-2005); 95% CI; descriptive  
 ↳ replication only.

## Current source interpretation rows

□

## Source family scoring and limitations

scoring\_rule: 'Fixed in advance per the collaboration's registered-report  
 ↳ conventions (trajectory  
 fields): at Stage 2 an outcome attaches to each confirmatory hypothesis's  
 ↳ frozen statement\_initial  
 as confirmed | partially\_confirmed | reversed, adjudicated strictly against  
 ↳ the pre-committed  
 interpretation table in this protocol. H1 scores confirmed only if  
 ↳ theta1's Holm-adjusted  
 two-sided test rejects with the point estimate in the predicted direction  
 ↳ AND the gated  
 exploratory H1\_trend contrast supports the frozen statement's growth rider  
 ↳ (beta\_Q4c -  
 beta\_Q1c > 0 with 95% CI excluding zero); it scores partially\_confirmed  
 ↳ when theta1 Holm-rejects  
 in the predicted direction but the growth rider is unsupported (flat  
 ↳ elevation - the pre-committed  
 reading given annual event dating) or the trend gate is not computed;  
 ↳ reversed if theta1  
 is Holm-significant in the negative direction. H2 scores confirmed only if  
 ↳ theta2a Holm-rejects  
 in the predicted direction AND falsification gate F2 does not fire AND the  
 ↳ theta2b sign  
 pattern is stable across the GRAFT-3 successor-share thresholds;  
 ↳ partially\_confirmed when  
 theta2a Holm-rejects but the comparative quantifier is untestable  
 ↳ (spurious-arm kill - verdict  
 capped by pre-commitment at 'partially consistent - comparative quantifier  
 ↳ untestable');  
 reversed if theta2a is Holm-significant in the negative direction OR gate  
 ↳ F2 fires (the  
 frozen statement's comparative reading is disconfirmed by a non-inert  
 ↳ placebo regardless  
 of theta2a - registered as disconfirming, never reframed). H3 scores  
 ↳ confirmed only if theta3's  
 Holm-adjusted test rejects in the predicted direction (including the  
 ↳ total-down/OOH-up pattern,  
 pre-registered as scoreable support for the composition claim); reversed if  
 ↳ Holm-significant  
 in the negative direction. A Holm-nonsignificant confirmatory result  
 ↳ attaches no confirmation:  
 it is reported as a bounded informative null - the 95% CI converted into  
 ↳ the excess magnitudes  
 the data rule out against realized MDEs - never 'accepted'. A member  
 ↳ demoted by the Stage-2  
 realized-N gate or killed by its registered criterion attaches no score;  
 ↳ the demotion/kill

trail is disclosed. Significant effects in the non-predicted direction are  
↳ disconfirming,  
never supportive. Exploratory quantities (H1\_trend, H2\_diff, E1, E3, E4,  
↳ and any demoted  
member) attach outcomes only under the exploratory heading and are never  
↳ scored as preregistered-confirmed.'

assumptions:

- SSSY (2005-2025) and SYSI (1990-2005) capture the universe of publicly  
↳ reimbursed primary-care  
billing for insured residents; absence of rows means no billed care, not  
↳ missingness
- KONTAKT is a row-level contact marker whose sum yields contact counts  
↳ (in-environment validation  
with a frozen fallback definition if not)
- ENDNEDL/IDFREM/IDTILB/REGELFR/FRELFREM semantics follow Statistics  
↳ Denmark's IDA workplace-continuity  
documentation, cited directly (no DOI exists for these conventions, per  
↳ boundary directive);  
the primary classification rests on the worker-flow measure s(w) from IDAN,  
↳ so misreading  
an IDAS flag degrades only the cross-check
- November-status annual attachment measures workplace membership; intra-year  
↳ attachment changes  
are unobservable ex ante and the frozen c-1 anchor is the design's answer  
↳ to that - with  
the corollary, named as a design threat, that early year-c separations feed  
↳ realized post-job-loss  
billing into the pre-boundary window under intention-to-follow, a channel  
↳ the E4 separation  
shares and the registered theta1 attachment split bound but cannot remove
- ALDERIMP/KOENIMP (DST-imputed age/sex on SSSY rows) are accurate enough for  
↳ stratified matching;  
persons absent from SSSY in c-3..c-1 are handled by explicit unknown strata
- HONUGE settlement lags are predominantly <= 4 weeks (calibrated on  
↳ BEHANDLINGSDATO 2021+;  
guard-band contingency if not), and lag structure does not differ between  
↳ exposed and matched  
controls within the same calendar quarters
- Death and emigration are unobservable in scope; their post-event incidence  
↳ is assumed small;  
the deflating direction is conservative for the pre-boundary claim and  
↳ attenuating for theta2a,  
with the GRAFT-6 trigger bounding the post-boundary side
- Nominal-fee comparisons within cohort x calendar-quarter cells require no  
↳ deflator
- Danish institutional stability of GP billing categories (SPEC2 70-84, 89)  
↳ across 2007-2024  
at the 2-digit level
- 'PUBLISHED-BASIS SIZING ASSUMPTIONS (new, each with a registered failure  
↳ path): (i) Danish  
establishment closure rates among private size-5+ workplaces are in the  
↳ published Erhvervsdemografi-consistent  
1.5-3% band, supporting 150,000-300,000 exposed workers over 15 cohorts -  
↳ failure path:  
the pre-outcome realized-N gate demotes rather than stretches; (ii) the  
↳ Danish OOH share  
of GP contact volume is in the published 8-12% band (Christensen & Olesen  
↳ 1998; LV-KOS 2011  
studies), simulated at the conservative 8% - failure path: the  
↳ published-range benchmark

- check plus realized-share re-simulation, then demotion or kill; (iii) the
  - ↳ control pool supports
  - 1:3 matching at  $\geq 90\%$  match rate - failure path: coarsening ladder,
    - ↳ realized-k accounting
    - in the blind re-simulation, then the gate'
- theta2a's 4-quarter averaging gain over theta1 survives serial correlation
  - ↳ in post-year
  - shocks; a registered serial-correlation discount is built into the Stage-0
    - ↳ re-simulation
    - rather than assumed away
- The persistent-vs-transient overdispersion split (persistent frailty var
  - ↳ 1.25, transient
  - log-var 0.25) is retained unchanged from the first-round simulation -
    - ↳ deliberately NOT revised
    - in the direction that would raise power, so the remediation cannot be
      - ↳ accused of tuning
      - the DGP
- failure\_modes:
  - 'Realized event or worker counts far below schema-derived expectations
    - ↳ (documentation-only
    - corpus; no realized Ns exist) - the design now bears this risk openly: the
      - ↳ Stage-2 realized-N
      - gate demotes underpowered members loudly and pre-outcome instead of letting
        - ↳ an underpowered
        - confirmatory test run; a sparse-closure Denmark kills or demotes cleanly
          - ↳ rather than quietly'
    - KONTAKT or ENDNEDL semantics differ from assumption - both carry frozen
      - ↳ fallback rules,
      - but level shifts in outcomes or event dating would force reported
        - ↳ recalibration and consume
        - credibility
    - HONUGE settlement lag materially longer than 4 weeks - the quarter-level
      - ↳ estimands survive,
      - but the weekly-resolution headline weakens to quarterly resolution in
        - ↳ practice (GRAFT-2
        - makes this explicit and pre-committed)
    - ANDELOPS direction-of-absorption semantics may differ from the idea's gloss
      - ↳ - classification
      - routes through IDAN worker flows, so this only weakens the IDAS
        - ↳ cross-check, not the H2
        - machinery
    - Spurious-closure events too heterogeneous for a clean placebo - a non-inert
      - ↳ placebo now
      - shows up as gate F2 firing (or exploratory H2\_diff reversal), which caps
        - ↳ H2's verdict by
        - pre-commitment instead of silently contaminating a differential test
    - Realized OOH share below the published range - caught by the
      - ↳ published-range benchmark check
      - before estimation; H3 demotes or dies pre-outcome instead of running
        - ↳ underpowered
    - Realized N lands below the full-expectation lower bound - the gate demotes
      - ↳ affected members
      - to exploratory pre-outcome; the study survives with a smaller confirmatory
        - ↳ family (or none,
        - in which case it reports as registered exploratory estimation with the
          - ↳ demotion trail published
          - the first-round outcome, now a contingency rather than the design)
    - Unobservable post-closure death/emigration deflates post-boundary billing -
      - ↳ direction stated,

- sensitivity pre-specified, GRAFT-6 2pp trigger forces best/worst-case
  - ↪ bounds; no in-scope
- fix exists
- Public/private sector misclassification via modal ARB\_SEKTORKODE for
  - ↪ multi-establishment
- firms - attenuates contrasts; share bounded in the balance tables; if
  - ↪ ARB\_SEKTORKODE itself
- proves absent in-environment (it is unconfirmed in the recorded
  - ↪ Forskningservice inventory),
- the frozen ARB\_HOVED\_BRA\_DB07 fallback in contingent\_decisions substitutes,
  - ↪ with counts
- reported
- Multi-job workers' primary-job assignment may misattribute exposure -
  - ↪ deterministic rule
- biases toward null; share reported
- 1:3 matching without replacement may exhaust thin strata - realized  $k < 3$ 
  - ↪ accounted honestly
- in the blind re-simulation rather than assumed; coarsening ladder
  - ↪ pre-committed
- The SYSI-era exploratory replication may be infeasible - bounded as
  - ↪ skippable Stage-5 work
- that cannot contaminate the confirmatory tier
- Readers may over-read gate F2 as proof of causal job-loss effects - the
  - ↪ associational ceiling,
- the 'consistent with' vocabulary cap, and the pre-committed F2-fired
  - ↪ interpretation guard
- the prose

## Source disclosure rules

environment: DST Forskningservice (no prior data access; gatekeeper

- ↪ certification per RR

secondary-data rules). All analysis occurs inside the environment; only

- ↪ aggregate outputs

leave. No microdata, workplace identifiers (ARBNR/CVRNR/LBNR), or

- ↪ person-level values are

released.

minimum\_cell\_size: 5

minimum\_cell\_rules: no released statistic based on fewer than 5 persons or

- ↪ fewer than 3 workplaces;

event-count tables rounded to the nearest 10; weekly trajectory figures

- ↪ pool  $\geq 100$  events

and  $\geq 5,000$  person-weeks per plotted point; fee (BRUHON) aggregates

- ↪ subject to a 50% single-workplace

dominance check; min-cell suppression applied BEFORE Holm reporting so

- ↪ suppression can never

be outcome-selective

outputs\_leaving\_environment:

- model coefficients with SEs/CIs
- matching balance tables and realized-k distribution
- calibration and benchmark tables (coverage, linkage, HONUGE lag, OOH share)
- figure aggregates (weekly trajectory series under the pooling rules)
- event counts rounded to the nearest 10
- Stage-0 power/MDE tables - BOTH rounds (failed and remediated) - and the
  - ↪ Stage-2 blind re-simulation
- output with the demotion trail
- the c-2 placebo, gate F2 outputs, GRAFT-2/4 diagnostics, GRAFT-6
  - ↪ later-presence rates, E4
- attrition panel, and all outcome-neutral check outputs (the full audit
  - ↪ trail exits under

the same rules)

- the analysis and power scripts

statement: 'All analysis occurs inside DST Forskningservice; only aggregate  
 ↳ outputs leave.

Rules applied to every released object: (1) no statistic based on fewer  
 ↳ than 5 persons or  
 fewer than 3 workplaces; (2) event-count tables rounded to the nearest 10;  
 ↳ (3) weekly trajectory  
 figures pool  $\geq 100$  events and  $\geq 5,000$  person-weeks per plotted point; (4)  
 ↳ fee (BRUHON)  
 statistics released only as cell means/coefficients over min-N cells, with  
 ↳ a dominance check  
 that no single workplace contributes  $> 50\%$  of any released fee aggregate;  
 ↳ (5) released objects  
 are limited to: model coefficients with SEs/CIs, balance tables,  
 ↳ calibration/benchmark tables,  
 figure aggregates, event counts (rounded), the Stage-0 (both rounds) and  
 ↳ Stage-2 power/MDE  
 tables with the demotion trail, and the analysis/power scripts - never  
 ↳ microdata, workplace  
 identifiers (ARBNR/CVRNR/LBNR), or person-level values; (6) min-cell  
 ↳ suppression is applied  
 before Holm reporting so that suppression can never be outcome-selective;  
 ↳ (7) the c-2 placebo,  
 gate F2 outputs, the GRAFT-2/4 diagnostics, the GRAFT-6 later-presence  
 ↳ rates, the E4 attrition  
 panel, and all outcome-neutral check outputs are released under the same  
 ↳ rules, ensuring  
 the full audit trail exits the environment.'

## Complete registration fields

```
id: atom-000007
topic: Historical billing pattern
atom_result: prereg
atom_strength: null
atom_shape: null
evidence_paths: []
concepts: []
literature: []
bonds: []
preregistration:
 version: 1
 mode: adaptive
 question: How does the source H1 analogue appear in the separate SYSI-era cohort?
 rationale: Source-existing coherent scientific question; its related
 ↳ outcomes,
 riders and sensitivity specifications remain internal.
 assumption: |-
```

Associational - matched event-study contrasts, a contrastive  
 ↪ falsification gate, and a composition contrast; descriptive for  
 ↪ trajectory panels and the bounded priority claim. Binding vocabulary  
 ↪ rules: 'is associated with' / 'consistent with' / 'suggests' only;  
 ↪ 'effect', 'impact', 'causes' prohibited throughout; anticipation  
 ↪ always worded as 'pre-boundary excess' relative to the annually dated  
 ↪ closure-year boundary, never as announcement-dated anticipation (S8);  
 ↪ no phenomenon-discovery or Danish-first framing (S1-S3); H2 novelty  
 ↪ confined to the placebo's application to health billing (S7). H2  
 ↪ verdict semantics are fixed in advance: 'consistent with H2' requires  
 ↪ BOTH the confirmatory core (theta2a Holm-significant, predicted  
 ↪ direction) AND a passing falsification gate F2; the strict  
 ↪ comparative magnitude is only ever reported as an exploratory  
 ↪ estimate with its CI. Exploratory results (H1\_trend, H2\_diff, E1, E3,  
 ↪ E4, and any gate-demoted member) carry the word 'exploratory' in the  
 ↪ heading and in every sentence stating them. The prose label never  
 ↪ exceeds the evidence label.  
 Compatibility ratings below are neutral import metadata, not a new  
 ↪ scientific evaluation or empirical scores.  
 target\_strength: descriptive  
 target\_shape: comparison  
 ratings:  
   importance: 3  
   novelty: 3  
   feasibility: 3  
   data\_advantage: 3  
   informativeness: 0.5  
 tables:  
 - IDAN  
 - IDAS  
 - SYSI  
 activation: state(atom-000003) == "event\_study\_ready" AND done(atom-000004)  
   AND done(atom-000005) AND done(atom-000006)  
 verdicts:  
 - key: exploratory\_pattern  
   region: The source-defined module yields interpretable  
     ↪ estimates/descriptives  
       after its own eligibility checks; retain unavailable components  
       ↪ explicitly.  
   says: Report the complete source exploratory pattern without confirmatory  
     or causal promotion.  
 - key: source\_skip  
   region: The source-defined optional module/arm is unavailable or its  
     ↪ registered  
       skip criterion is met.  
   says: Record the source skip and its reason; never invent a negative or  
     null finding.  
 - key: boundary  
   region: Required source-defined inputs or scientific definitions are  
     ↪ demonstrably  
       unavailable; retain available components and the exact unresolved scope.  
   says: An evidenced boundary, not a statistical null. A missing  
     ↪ implementation  
       or timeout alone is operational no\_result, not this state.  
 - key: otherwise  
   region: otherwise  
   says: The registered judgment is unresolved; report all completed  
     ↪ components  
       without inventing support, equivalence or absent results.  
 protocol: |

# How does the source H1 analogue appear in the separate SYSI-era cohort?

## Source fidelity and implementation order

This is a format conversion of DS1.1, not a new study or independent  
↳ acceptance.  
Source audit: minor\_revisions. No observed Danish outcomes are supplied.  
The source execution plan explicitly states that final prereg.yaml governs  
disagreements. Exact current YAML selections below therefore govern older  
execution-plan excerpts; these excerpts are retained as implementation  
↳ detail,  
not a competing registration. See docs/provenance/source-discrepancies.md.  
The complete shared data/validity contracts are embedded in Atoms 1--3;  
↳ read  
those native protocols when executing a downstream Atom, even without  
↳ docs/.  
Do not substitute a scientific default for an unresolved source  
↳ contradiction.

Preserve source stages 0 -> 1 -> 2 -> 3 -> 4 -> 5 inside the  
↳ implementation.  
Source Stage-0 simulations are historical design evidence, not study  
↳ outcomes;  
do not silently redo or improve them. Gate 14 must be evaluated from  
↳ baseline  
data before post-baseline outcomes. All surviving Stage-4 quantities form  
↳ ONE  
Holm family. The three question-level Atoms may share the same frozen  
↳ family  
fit and tidy results; do not apply Holm separately per Atom or adjudicate  
↳ a  
member before the complete surviving family's p-values exist. Cache/reuse  
identical governed intermediate results with provenance when available; a  
direct-dispatch call must obtain or compute its exact source-defined  
↳ inputs,  
not pretend its predecessors succeeded. Stage-5 internal riders remain  
exploratory, after the Stage-4 family is frozen, even when housed in  
↳ H1/H2.

Reference files under docs/execution/ are the ORIGINAL SKELETONS, not  
↳ tested  
turnkey implementations. Adapt paths and complete engineering in a  
↳ governed  
workspace; preserve source scientific rules and disclose unresolved  
↳ choices.  
Write a result record with component estimates, uncertainty, family  
↳ eligibility,  
source interpretation row, diagnostics and boundary/skip dispositions.  
↳ Released  
outputs follow source disclosure rules; confidential rows stay governed.  
Local native states summarize the source judgment; neither a state name  
↳ nor  
source power forecasts prove equivalence or causal identification. A  
↳ bounded  
claim requires its actual interval and registered rule. Retain unresolved  
↳ source  
wording conflicts; do not convert mere nonsignificance into scientific  
↳ null.

## Frozen scope and vocabulary

```

```yaml
claim_class_ceiling: associational
statement: 'Associational - matched event-study contrasts, a contrastive
↳ falsification gate,
and a composition contrast; descriptive for trajectory panels and the
↳ bounded priority claim.
Binding vocabulary rules: 'is associated with' / 'consistent with'
↳ / 'suggests' only;
'effect', 'impact', 'causes' prohibited throughout; anticipation
↳ always worded as
'pre-boundary excess' relative to the annually dated closure-year
↳ boundary, never as announcement-dated
anticipation (S8); no phenomenon-discovery or Danish-first framing
↳ (S1-S3); H2 novelty confined
to the placebo's application to health billing (S7). H2 verdict
↳ semantics are fixed in
advance: 'consistent with H2' requires BOTH the confirmatory core
↳ (theta2a Holm-significant,
predicted direction) AND a passing falsification gate F2; the strict
↳ comparative magnitude
is only ever reported as an exploratory estimate with its CI.
↳ Exploratory results (H1_trend,
H2_diff, E1, E3, E4, and any gate-demoted member) carry the word
↳ 'exploratory' in the
heading and in every sentence stating them. The prose label never
↳ exceeds the evidence label.'
frozen_parameters:
seed: 20260831
gp_code_list: daytime GP = SPEC2 in {70..80}
ooh_code_list: out-of-hours GP = SPEC2 in {81, 82, 83, 84, 89}
ooh_robustness_classifier: adds YDLTID in {2, 3, 4, 5, 7, 8, 9}
spec2_evening_gp_disclosure: 'Registered choice, disclosed: the parsed
↳ value labels mark
SPEC2 71 ("General practitioner, consultation, evening etc."), 73
↳ ("General practitioner,
telephone consultation, evening etc."), and 75 ("General
↳ practitioner, home visit, evening
etc.") as explicitly off-schedule GP services; they nonetheless
↳ remain on the daytime
side of the PRIMARY classifier because H3's out-of-hours construct
↳ and its published
power basis are keyed to the laegevagt out-of-hours service (SPEC2
↳ 81-84, 89), the quantity
the published Danish 8-12% share describes. Attenuation direction,
↳ both readings: counting
evening GP as daytime attenuates theta3 toward zero - conservative
↳ against the positive
H3 conclusion, anti-conservative for the null-bounding
↳ "compositionally unremarkable"
reading, whose Stage-2 statement therefore carries a mandatory
↳ caveat. In-environment
validation must verify that the YDLTID-augmented robustness
↳ classifier captures SPEC2
71/73/75 rows (their evening time codes 2-5); if YDLTID proves
↳ unpopulated on those rows,
the robustness classifier - and only the robustness classifier - is
↳ amended mechanically
and with disclosure to include SPEC2 71/73/75 directly. The SPEC2
↳ 71/73/75 volume share

```

```

is reported, and both classifiers appear side by side in Table T7.'
successor_share_thresholds: 'true closure: ENDNEDL final closure AND
↳ worker-flow max-successor
share s(w) < 0.30; spurious closure: forward identity via
↳ IDFREEM/REGELFR/FRELFREM AND
s(w) >= 0.70; ambiguous mid-zone excluded; GRAFT-3 sensitivities
↳ 0.20/0.80 and 0.40/0.60'
quarter_mapping: 'ISO week of HONUGE: 1-13 / 14-26 / 27-39 / 40-52; ISO
↳ week 53 to Q4'
winsorization: person-quarter outcomes at the 99.9th percentile within
↳ era x calendar quarter
matching: 1:3 nearest-neighbor Mahalanobis within exact strata, without
↳ replacement across
sets, target k = 3, minimum k = 1; calipers |dif baseline contacts|
↳ <= 3 and |dif log(1+BRUHON)|
<= 0.5; exposed processed in seeded random order (seed 20260831)
software: R >= 4.3, fixest::feols (OLS) and fixest::fepois (Poisson QML
↳ robustness); versions
frozen in the registered script
note: Code lists and thresholds frozen before any data access; code
↳ labels quoted verbatim
from parsed DST value labels and re-validated in-environment before
↳ estimation (frozen
fallback rules in contingent_decisions).
...

22. **E1 legacy pipeline (Stage 5; skippable).** Input: IDAS/IDAN
1990-2005 (legacy keys ARBGNR/ARBSTK/LBNR), SYSI 1990-2005. SYSI-era
replication of the H1 contrast for cohorts c in 1992-2003 (baseline
1991+, post-year <= 2004): era GP/OOH classification from SPECIALE
(era code list frozen at Stage 0) with VAGTOMR (1997-2005) as the
era OOH marker; contacts proxied by service-row counts weighted by
YDLANT; matching WITHOUT age/sex (SYSI lacks them), using legacy job
characteristics (STILL, ANSDAGE, TIMELON, TILKNYT) and SIKREKOM/
SIKREAMT patient geography (both used only here; inventory-only
disclosure carried). Output: `analysis/e1_panel.parquet`. Check:
skip E1 entirely - reported as skipped, not failed - if SYSI-era
extraction yields < 300 true closures or the era's coverage
benchmark fails. Never pooled with the confirmatory era.

---

### 4.5 Exploratory tier (Stage 5; heading-segregated; Holm within family
↳ as severity aid)

- **H1_trend (gated):** beta_Q4c - beta_Q1c from the true-arm model, 95%
CI, CR1 clusters - computed ONLY if theta1 Holm-rejected positively;
otherwise the quarterly trajectory is shown descriptively with no
contrast statistic and no timing-shape language in the Summary.
- **H2_diff:** theta2_diff = mean_k[beta_true,k - beta_spurious,k] over
the post-year quarters from the pooled model, 95% CI, plus the
quarter-by-quarter differential profile (the "more persistent" rider).
A flagged estimate with spurious > true triggers the pre-committed
non-inert-placebo interpretation and reinforces whatever F2 showed.
- **E1:** SYSI-era analogue of theta1 from step 22, 95% CI, descriptive
replication only; never pooled across the vintage break.
- **E3:** BRUHON fee analogues of theta2a/theta2_diff beyond the gated
secondary; heterogeneity of theta1/theta2a by sex (KOENIMP), age band
(ALDERIMP), and baseline-year wage decile (BREDT_LOEN_BELOEB); Holm
within this family.
- **E4:** arm-comparative attrition panel - share of each frozen

```

- November c-1 cohort no longer attached to the event workplace by November of year c (IDAN year-c records; ANSXFREM as the forward-separation flag), by event class; explicitly labeled post-treatment-conditioned descriptives with a selection caveat; no test.
- No exploratory result can revise any confirmatory verdict or Stage-4 language. Every sentence stating an exploratory result contains the word "exploratory".

4. ANALYSIS EXECUTION

Confirmatory tier first (Stage 4, frozen code, single run); exploratory tier after (Stage 5), segregated by heading. All estimation in R fixest via Rscript. Common estimator: outcome y_iq (winsorized quarterly GP contact count) on treated x event-quarter indicators for the 8 quarters Q1(c)..Q4(c+1), matched-set fixed effects, cohort x calendar-quarter fixed effects, omitted baseline = the four quarters of year c-1; observation weights: exposed 1, controls 1/k_s within set; CR1 SEs clustered on event workplace (each control inherits its set's event cluster). Min-cell suppression is applied to all outputs BEFORE Holm reporting so suppression can never be outcome-selective.

5. EXHIBITS

Every figure and table the report commits to. Disclosure rules binding on EVERY released object: no statistic based on fewer than 5 persons or fewer than 3 workplaces; event-count tables rounded to the nearest 10; weekly trajectory figures pool ≥ 100 events and $\geq 5,000$ person-weeks per plotted point; BRUHON aggregates pass a 50% single-workplace dominance check; released objects are limited to model coefficients with SEs/CIs, balance tables, calibration/benchmark tables, figure aggregates, rounded event counts, the power/MDE tables (both Stage-0 rounds and Stage 2) with the demotion trail, and the analysis/power scripts - never microdata, workplace identifiers (ARENR/CVRNR/LBNR), or person-level values; suppression applied BEFORE Holm reporting.

****Figures (matplotlib):****

- ****F1 - Quarterly event-study panel.**** Treated x event-quarter coefficients with 95% CIs across Q1(c-2)..Q4(c+1) (placebo quarters, omitted-baseline marker, terminal-year and post-year quarters); panel (a) true arm, panel (b) spurious arm from the pooled model. Feeds: step-18 panels + 4.1/4.2 coefficient tables. Disclosure: coefficients only.
- ****F2 - Weekly descriptive trajectory.**** Pooled exposed-minus-matched-control weekly GP contact series around the closure-year boundary, with the GRAFT-2 week-0 step diagnostic annotated (step $\leq 50\%$ of the total Q3(c)+Q4(c) excess; if exceeded, the figure carries the registered caution and quarter-level estimates are foregrounded). Feeds: step 15/17 weekly aggregates. Disclosure: ≥ 100 events and $\geq 5,000$ person-weeks per point.
- ****F3 - s(w) successor-share distribution (GRAFT-4).**** Histogram of worker-flow successor shares across terminal events with the 0.30/0.70 thresholds and mid-zone mass; bimodality check exhibit. Feeds: step 7. Disclosure: bin counts rounded to 10; bins under min-cell suppressed.
- ****F4 - OOH vs daytime relative excess (theta3).**** The two relative-excess series over the peri-event window with the bootstrap CI of their gap; primary and YDLTID-augmented classifier side by side. Feeds: 4.3 outputs. Disclosure: coefficients only.
- ****F5 - Matching balance (love plot).**** SMDs before/after matching for

all matching covariates and unmatched auxiliaries; 0.10 line marked.
Feeds: step 13. Disclosure: aggregates only.

- **F6 - E1 SYSI-era replication (exploratory; skippable).** Quarterly SYSI-era analogue of F1(a) for cohorts 1992-2003, labeled exploratory, YDLANT-weighted proxy outcome. Feeds: step 22. Disclosure: as F1.

****Tables (booktabs LaTeX):****

- **T1 - Event extraction and classification by cohort year.** Terminal events, true/spurious/ambiguous counts (rounded to 10), annual true-closure rate among private size-5+ workplaces, spurious/true ratio. Feeds: step 8.
- **T2 - Classification cross-validation.** Worker-flow classes vs IDAS-native flags (ENDNEDL/IDFREM/REGELFR/FRELFREM/ANDELUDS), full cross-tabulation + agreement rate; GRAFT-3 classification counts under 20/80 and 40/60. Feeds: step 8.
- **T3 - Cohort flow accounting.** Inclusion/exclusion counts per rule (tenure, sector, size, SIKGRUP, missing ANSAAR/pay, multi-cohort dedup), eligible exposed per cohort, match rate, ladder rung used, realized k distribution. Feeds: steps 9-13.
- **T4 - Matching balance.** Per-covariate SMDs (weighted 1/k), caliper compliance, realized mean k. Feeds: step 13.
- **T5 - Power and demotion trail.** Stage-0 power results (BOTH rounds: the voided round-1 run and the remediated re-run) and the Stage-2 blind realized-power/MDE table with binding N*, realized sets, and any demotions with the fixed-point recomputation. Feeds: deposited script + step 14. Published regardless of outcome.
- **T6 - Outcome-neutral check panel.** All eleven registered checks (coverage benchmark by year, linkage floor, closure-rate plausibility, classification agreement, GRAFT-4, balance/match rate, c-2 placebo, HONUGE lag calibration, floor/ceiling + seasonality, GRAFT-2 step, OOH-share benchmark) with thresholds and PASS/FAIL. Feeds: steps 8, 13, 16, 17, 19, 21 + F2/F3 diagnostics.
- **T7 - Confirmatory results.** theta1, theta2a, theta3: point estimates (contacts/person-quarter and % of baseline), 95% CIs, two-sided p, Holm-adjusted verdicts; the F2 gate row (theta2b, its one-sided 95% upper bound, trigger evaluation); the gated BRUHON fee secondary (if computed); c-2 placebo row; realized MDEs beside every null with the CI-ruled-out magnitudes. Feeds: 4.1-4.4.
- **T8 - Sensitivity suite.** GRAFT-3 thresholds; GRAFT-9 forward-window exclusion (with truncation note); un-winsorized; Poisson QML; death/emigration exclusion - one table, run once. Feeds: 4.4.
- **T9 - Exploratory tier.** H1_trend (if gated in; otherwise the row states the gate was not met), H2_diff with quarter-by-quarter differential profile, E3 fee and subgroup analyses (within-family Holm), E1 summary. Every row labeled exploratory. Feeds: 4.5.
- **T10 - Attrition and later-presence accounting.** E4 year-c separation shares by event class (ANSXFREM-based, post-treatment-conditioned label and selection caveat) and the GRAFT-6 later-presence rates with the 2pp trigger status. Feeds: steps 20 + 4.5 E4.

Current source hypotheses (all fields)

```

```yaml
- id: E1
 tier: exploratory
 statement_initial: null

```

```

statement_source: no frozen bank statement - protocol-defined
 ↳ exploratory module (SYSI-era
 descriptive replication of the H1 contrast)
estimand: SYSI-era replication of the H1 contrast for closure cohorts c
 ↳ in 1992-2003 (baseline
 1991+, post-year <= 2004; legacy keys ARBGNR/ARBSTK/LBNR; matching
 ↳ without age/sex, which
 SYSI lacks; SIKREKOM/SIKREAMT and VAGTOMR used only here and
 ↳ disclosed as verified in
 the Forskningservice inventory but absent from public TIMES/HKV
 ↳ metadata). Never pooled
 with the 2009-2023 era across the 2007/08 IDA vintage break.
primary_quantity: SYSI-era analogue of theta1 (contacts proxied by
 ↳ service-row counts weighted
 by YDLANT; KONTAKT does not exist pre-2005); 95% CI; descriptive
 ↳ replication only.
...

```

## Current source interpretation rows

```

```yaml
[]
```

```

## Source family scoring and limitations

```

```yaml
scoring_rule: 'Fixed in advance per the collaboration''s
  ↳ registered-report conventions (trajectory
    fields): at Stage 2 an outcome attaches to each confirmatory
    ↳ hypothesis''s frozen statement_initial
    as confirmed | partially_confirmed | reversed, adjudicated strictly
    ↳ against the pre-committed
    interpretation table in this protocol. H1 scores confirmed only if
    ↳ theta1''s Holm-adjusted
    two-sided test rejects with the point estimate in the predicted
    ↳ direction AND the gated
    exploratory H1_trend contrast supports the frozen statement''s growth
    ↳ rider (beta_Q4c -
    beta_Q1c > 0 with 95% CI excluding zero); it scores partially_confirmed
    ↳ when theta1 Holm-rejects
    in the predicted direction but the growth rider is unsupported (flat
    ↳ elevation - the pre-committed
    reading given annual event dating) or the trend gate is not computed;
    ↳ reversed if theta1
    is Holm-significant in the negative direction. H2 scores confirmed only
    ↳ if theta2a Holm-rejects
    in the predicted direction AND falsification gate F2 does not fire AND
    ↳ the theta2b sign
    pattern is stable across the GRAFT-3 successor-share thresholds;
    ↳ partially_confirmed when
    theta2a Holm-rejects but the comparative quantifier is untestable
    ↳ (spurious-arm kill - verdict
    capped by pre-commitment at ''partially consistent - comparative
    ↳ quantifier untestable'');
    reversed if theta2a is Holm-significant in the negative direction OR
    ↳ gate F2 fires (the
    frozen statement''s comparative reading is disconfirmed by a non-inert
    ↳ placebo regardless

```

of theta2a - registered as disconfirming, never reframed). H3 scores
 ↳ confirmed only if theta3's
 Holm-adjusted test rejects in the predicted direction (including the
 ↳ total-down/OOH-up pattern,
 pre-registered as scoreable support for the composition claim);
 ↳ reversed if Holm-significant
 in the negative direction. A Holm-nonsignificant confirmatory result
 ↳ attaches no confirmation:
 it is reported as a bounded informative null - the 95% CI converted
 ↳ into the excess magnitudes
 the data rule out against realized MDEs - never 'accepted'. A member
 ↳ demoted by the Stage-2
 realized-N gate or killed by its registered criterion attaches no
 ↳ score; the demotion/kill
 trail is disclosed. Significant effects in the non-predicted direction
 ↳ are disconfirming,
 never supportive. Exploratory quantities (H1_trend, H2_diff, E1, E3,
 ↳ E4, and any demoted
 member) attach outcomes only under the exploratory heading and are
 ↳ never scored as preregistered-confirmed.'

assumptions:

- SSSY (2005-2025) and SYSI (1990-2005) capture the universe of publicly
 ↳ reimbursed primary-care
 billing for insured residents; absence of rows means no billed care,
 ↳ not missingness
- KONTAKT is a row-level contact marker whose sum yields contact counts
 ↳ (in-environment validation
 with a frozen fallback definition if not)
- ENDNEDL/IDFREM/IDTILE/REGELFR/FRELFREM semantics follow Statistics
 ↳ Denmark's IDA workplace-continuity
 documentation, cited directly (no DOI exists for these conventions, per
 ↳ boundary directive);
 the primary classification rests on the worker-flow measure s(w) from
 ↳ IDAN, so misreading
 an IDAS flag degrades only the cross-check
- November-status annual attachment measures workplace membership;
 ↳ intra-year attachment changes
 are unobservable ex ante and the frozen c-1 anchor is the design's
 ↳ answer to that - with
 the corollary, named as a design threat, that early year-c separations
 ↳ feed realized post-job-loss
 billing into the pre-boundary window under intention-to-follow, a
 ↳ channel the E4 separation
 shares and the registered theta1 attachment split bound but cannot
 ↳ remove
- ALDERIMP/KOENIMP (DST-imputed age/sex on SSSY rows) are accurate enough
 ↳ for stratified matching;
 persons absent from SSSY in c-3..c-1 are handled by explicit unknown
 ↳ strata
- HONUIGE settlement lags are predominantly <= 4 weeks (calibrated on
 ↳ BEHANDLINGSDATO 2021+;
 guard-band contingency if not), and lag structure does not differ
 ↳ between exposed and matched
 controls within the same calendar quarters
- Death and emigration are unobservable in scope; their post-event
 ↳ incidence is assumed small;
 the deflating direction is conservative for the pre-boundary claim and
 ↳ attenuating for theta2a,
 with the GRAFT-6 trigger bounding the post-boundary side

- Nominal-fee comparisons within cohort x calendar-quarter cells require
 - ↳ no deflator
- Danish institutional stability of GP billing categories (SPEC2 70-84, 89) across 2007-2024
 - ↳ at the 2-digit level
- 'PUBLISHED-BASIS SIZING ASSUMPTIONS (new, each with a registered failure path): (i) Danish establishment closure rates among private size-5+ workplaces are in the
 - ↳ published Erhvervsdemografi-consistent 1.5-3% band, supporting 150,000-300,000 exposed workers over 15 cohorts
 - ↳ - failure path: the pre-outcome realized-N gate demotes rather than stretches; (ii) the Danish OOH share of GP contact volume is in the published 8-12% band (Christensen & Olesen 1998; LV-KOS 2011 studies), simulated at the conservative 8% - failure path: the published-range benchmark check plus realized-share re-simulation, then demotion or kill; (iii) the control pool supports 1:3 matching at $\geq 90\%$ match rate - failure path: coarsening ladder, realized-k accounting in the blind re-simulation, then the gate'
- theta2a's 4-quarter averaging gain over theta1 survives serial correlation in post-year shocks; a registered serial-correlation discount is built into the
 - ↳ Stage-0 re-simulation rather than assumed away
- The persistent-vs-transient overdispersion split (persistent frailty var 1.25, transient log-var 0.25) is retained unchanged from the first-round simulation -
 - ↳ deliberately NOT revised in the direction that would raise power, so the remediation cannot be accused of tuning the DGP
- failure_modes:
 - 'Realized event or worker counts far below schema-derived expectations
 - ↳ (documentation-only corpus; no realized Ns exist) - the design now bears this risk openly:
 - ↳ the Stage-2 realized-N gate demotes underpowered members loudly and pre-outcome instead of letting an underpowered confirmatory test run; a sparse-closure Denmark kills or demotes cleanly rather than quietly'
 - KONTAKT or ENDNEDL semantics differ from assumption - both carry frozen
 - ↳ fallback rules, but level shifts in outcomes or event dating would force reported
 - ↳ recalibration and consume credibility
 - HONUGE settlement lag materially longer than 4 weeks - the
 - ↳ quarter-level estimands survive, but the weekly-resolution headline weakens to quarterly resolution in practice (GRAFT-2 makes this explicit and pre-committed)
 - ANDELOPS direction-of-absorption semantics may differ from the idea's
 - ↳ gloss - classification routes through IDAN worker flows, so this only weakens the IDAS
 - ↳ cross-check, not the H2 machinery
 - Spurious-closure events too heterogeneous for a clean placebo - a
 - ↳ non-inert placebo now

- shows up as gate F2 firing (or exploratory H2_diff reversal), which
 - ↳ caps H2's verdict by pre-commitment instead of silently contaminating a differential test
- Realized OOH share below the published range - caught by the
 - ↳ published-range benchmark check before estimation; H3 demotes or dies pre-outcome instead of running
 - ↳ underpowered
- Realized N lands below the full-expectation lower bound - the gate
 - ↳ demotes affected members to exploratory pre-outcome; the study survives with a smaller
 - ↳ confirmatory family (or none, in which case it reports as registered exploratory estimation with the
 - ↳ demotion trail published
 - the first-round outcome, now a contingency rather than the design)
- Unobservable post-closure death/emigration deflates post-boundary
 - ↳ billing - direction stated, sensitivity pre-specified, GRAFT-6 2pp trigger forces best/worst-case
 - ↳ bounds; no in-scope fix exists
- Public/private sector misclassification via modal ARB_SEKTORKODE for
 - ↳ multi-establishment firms - attenuates contrasts; share bounded in the balance tables; if
 - ↳ ARB_SEKTORKODE itself proves absent in-environment (it is unconfirmed in the recorded
 - ↳ Forskningsservice inventory), the frozen ARB_HOVED_BRA_DB07 fallback in contingent_decisions
 - ↳ substitutes, with counts reported
- Multi-job workers' primary-job assignment may misattribute exposure -
 - ↳ deterministic rule biases toward null; share reported
- 1:3 matching without replacement may exhaust thin strata - realized $k <$
 - ↳ 3 accounted honestly in the blind re-simulation rather than assumed; coarsening ladder
 - ↳ pre-committed
- The SYSI-era exploratory replication may be infeasible - bounded as
 - ↳ skippable Stage-5 work that cannot contaminate the confirmatory tier
- Readers may over-read gate F2 as proof of causal job-loss effects - the
 - ↳ associational ceiling, the 'consistent with' vocabulary cap, and the pre-committed F2-fired
 - ↳ interpretation guard
 - the prose

Source disclosure rules

```

```yaml
environment: DST Forskningsservice (no prior data access; gatekeeper
↳ certification per RR
secondary-data rules). All analysis occurs inside the environment; only
↳ aggregate outputs
leave. No microdata, workplace identifiers (ARBNR/CVRNR/LBNR), or
↳ person-level values are
released.
minimum_cell_size: 5
minimum_cell_rules: no released statistic based on fewer than 5 persons
↳ or fewer than 3 workplaces;
event-count tables rounded to the nearest 10; weekly trajectory figures
↳ pool >= 100 events

```

```

and >= 5,000 person-weeks per plotted point; fee (BRUHON) aggregates
 ↳ subject to a 50% single-workplace
dominance check; min-cell suppression applied BEFORE Holm reporting so
 ↳ suppression can never
 be outcome-selective
outputs_leaving_environment:
- model coefficients with SEs/CIs
- matching balance tables and realized-k distribution
- calibration and benchmark tables (coverage, linkage, HONUGE lag, OOH
 ↳ share)
- figure aggregates (weekly trajectory series under the pooling rules)
- event counts rounded to the nearest 10
- Stage-0 power/MDE tables - BOTH rounds (failed and remediated) - and
 ↳ the Stage-2 blind re-simulation
 output with the demotion trail
- the c-2 placebo, gate F2 outputs, GRAFT-2/4 diagnostics, GRAFT-6
 ↳ later-presence rates, E4
 attrition panel, and all outcome-neutral check outputs (the full audit
 ↳ trail exits under
 the same rules)
- the analysis and power scripts
statement: 'All analysis occurs inside DST Forskningservice; only
 ↳ aggregate outputs leave.
 Rules applied to every released object: (1) no statistic based on fewer
 ↳ than 5 persons or
 fewer than 3 workplaces; (2) event-count tables rounded to the nearest
 ↳ 10; (3) weekly trajectory
 figures pool >= 100 events and >= 5,000 person-weeks per plotted point;
 ↳ (4) fee (BRUHON)
 statistics released only as cell means/coefficients over min-N cells,
 ↳ with a dominance check
 that no single workplace contributes > 50% of any released fee
 ↳ aggregate; (5) released objects
 are limited to: model coefficients with SEs/CIs, balance tables,
 ↳ calibration/benchmark tables,
 figure aggregates, event counts (rounded), the Stage-0 (both rounds)
 ↳ and Stage-2 power/MDE
 tables with the demotion trail, and the analysis/power scripts - never
 ↳ microdata, workplace
 identifiers (ARBNR/CVRNR/LBNR), or person-level values; (6) min-cell
 ↳ suppression is applied
 before Holm reporting so that suppression can never be
 ↳ outcome-selective; (7) the c-2 placebo,
 gate F2 outputs, the GRAFT-2/4 diagnostics, the GRAFT-6 later-presence
 ↳ rates, the E4 attrition
 panel, and all outcome-neutral check outputs are released under the
 ↳ same rules, ensuring
 the full audit trail exits the environment.'
...

```

## Pending registered experiment

No Danish study outcome has been observed for this bundle. See the complete registration above.

## atom-000008: Fees and heterogeneity

### How do source-defined fee and subgroup contrasts vary within the registered exploratory tier?

#### Source fidelity and implementation order

This is a format conversion of DS1.1, not a new study or independent acceptance. Source audit: minor\_revisions. No observed Danish outcomes are supplied. The source execution plan explicitly states that final prereg.yaml governs disagreements. Exact current YAML selections below therefore govern older execution-plan excerpts; these excerpts are retained as implementation detail, not a competing registration. See docs/provenance/source-discrepancies.md. The complete shared data/validity contracts are embedded in Atoms 1–3; read those native protocols when executing a downstream Atom, even without docs/. Do not substitute a scientific default for an unresolved source contradiction.

Preserve source stages 0 -> 1 -> 2 -> 3 -> 4 -> 5 inside the implementation. Source Stage-0 simulations are historical design evidence, not study outcomes; do not silently redo or improve them. Gate 14 must be evaluated from baseline data before post-baseline outcomes. All surviving Stage-4 quantities form ONE Holm family. The three question-level Atoms may share the same frozen family fit and tidy results; do not apply Holm separately per Atom or adjudicate a member before the complete surviving family's p-values exist. Cache/reuse identical governed intermediate results with provenance when available; a direct-dispatch call must obtain or compute its exact source-defined inputs, not pretend its predecessors succeeded. Stage-5 internal riders remain exploratory, after the Stage-4 family is frozen, even when housed in H1/H2.

Reference files under docs/execution/ are the ORIGINAL SKELETONS, not tested turnkey implementations. Adapt paths and complete engineering in a governed workspace; preserve source scientific rules and disclose unresolved choices. Write a result record with component estimates, uncertainty, family eligibility, source interpretation row, diagnostics and boundary/skip dispositions. Released outputs follow source disclosure rules; confidential rows stay governed. Local native states summarize the source judgment; neither a state name nor source power forecasts prove equivalence or causal identification. A bounded claim requires its actual interval and registered rule. Retain unresolved source wording conflicts; do not convert mere nonsignificance into scientific null.

#### Frozen scope and vocabulary

```
claim_class_ceiling: associational
statement: 'Associational - matched event-study contrasts, a contrastive
↪ falsification gate,
 and a composition contrast; descriptive for trajectory panels and the
 ↪ bounded priority claim.
Binding vocabulary rules: 'is associated with' / 'consistent with' /
↪ 'suggests' only;
```

```

'effect', 'impact', 'causes' prohibited throughout; anticipation
↳ always worded as
'pre-boundary excess' relative to the annually dated closure-year
↳ boundary, never as announcement-dated
anticipation (S8); no phenomenon-discovery or Danish-first framing (S1-S3);
↳ H2 novelty confined
to the placebo's application to health billing (S7). H2 verdict semantics
↳ are fixed in
advance: 'consistent with H2' requires BOTH the confirmatory core
↳ (theta2a Holm-significant,
predicted direction) AND a passing falsification gate F2; the strict
↳ comparative magnitude
is only ever reported as an exploratory estimate with its CI. Exploratory
↳ results (H1_trend,
H2_diff, E1, E3, E4, and any gate-demoted member) carry the word
↳ 'exploratory' in the
heading and in every sentence stating them. The prose label never exceeds
↳ the evidence label.'
frozen_parameters:
seed: 20260831
gp_code_list: daytime GP = SPEC2 in {70..80}
ooh_code_list: out-of-hours GP = SPEC2 in {81, 82, 83, 84, 89}
ooh_robustness_classifier: adds YDLTID in {2, 3, 4, 5, 7, 8, 9}
spec2_evening_gp_disclosure: 'Registered choice, disclosed: the parsed
↳ value labels mark
SPEC2 71 ("General practitioner, consultation, evening etc."), 73
↳ ("General practitioner,
telephone consultation, evening etc."), and 75 ("General practitioner,
↳ home visit, evening
etc.") as explicitly off-schedule GP services; they nonetheless remain on
↳ the daytime
side of the PRIMARY classifier because H3's out-of-hours construct and
↳ its published
power basis are keyed to the laegevagt out-of-hours service (SPEC2 81-84,
↳ 89), the quantity
the published Danish 8-12% share describes. Attenuation direction, both
↳ readings: counting
evening GP as daytime attenuates theta3 toward zero - conservative
↳ against the positive
H3 conclusion, anti-conservative for the null-bounding "compositionally
↳ unremarkable"
reading, whose Stage-2 statement therefore carries a mandatory caveat.
↳ In-environment
validation must verify that the YDLTID-augmented robustness classifier
↳ captures SPEC2
71/73/75 rows (their evening time codes 2-5); if YDLTID proves
↳ unpopulated on those rows,
the robustness classifier - and only the robustness classifier - is
↳ amended mechanically
and with disclosure to include SPEC2 71/73/75 directly. The SPEC2
↳ 71/73/75 volume share
is reported, and both classifiers appear side by side in Table T7.'
successor_share_thresholds: 'true closure: ENDNEDL final closure AND
↳ worker-flow max-successor
share s(w) < 0.30; spurious closure: forward identity via
↳ IDFREEM/REGELFR/FRELFREM AND
s(w) >= 0.70; ambiguous mid-zone excluded; GRAFT-3 sensitivities
↳ 0.20/0.80 and 0.40/0.60'
quarter_mapping: 'ISO week of HONUGE: 1-13 / 14-26 / 27-39 / 40-52; ISO
↳ week 53 to Q4'

```

```

winsorization: person-quarter outcomes at the 99.9th percentile within era
↳ x calendar quarter
matching: 1:3 nearest-neighbor Mahalanobis within exact strata, without
↳ replacement across
 sets, target k = 3, minimum k = 1; calipers |dif baseline contacts| <= 3
 ↳ and |dif log(1+BRUHON)|
 <= 0.5; exposed processed in seeded random order (seed 20260831)
software: R >= 4.3, fixest::feols (OLS) and fixest::fepois (Poisson QML
↳ robustness); versions
 frozen in the registered script
note: Code lists and thresholds frozen before any data access; code labels
↳ quoted verbatim
 from parsed DST value labels and re-validated in-environment before
 ↳ estimation (frozen
 fallback rules in contingent_decisions).

```

#### 4.5 Exploratory tier (Stage 5; heading-segregated; Holm within family as severity aid)

- **H1\_trend (gated):**  $\beta_{Q4c} - \beta_{Q1c}$  from the true-arm model, 95% CI, CR1 clusters — computed ONLY if  $\theta_1$  Holm-rejected positively; otherwise the quarterly trajectory is shown descriptively with no contrast statistic and no timing-shape language in the Summary.
- **H2\_diff:**  $\theta_{2\_diff} = \text{mean}_k[\beta_{\text{true},k} - \beta_{\text{spurious},k}]$  over the post-year quarters from the pooled model, 95% CI, plus the quarter-by-quarter differential profile (the “more persistent” rider). A flagged estimate with  $\text{spurious} > \text{true}$  triggers the pre-committed non-inert-placebo interpretation and reinforces whatever F2 showed.
- **E1:** SYSI-era analogue of  $\theta_1$  from step 22, 95% CI, descriptive replication only; never pooled across the vintage break.
- **E3:** BRUHON fee analogues of  $\theta_{2a}/\theta_{2\_diff}$  beyond the gated secondary; heterogeneity of  $\theta_1/\theta_{2a}$  by sex (KOENIMP), age band (ALDERIMP), and baseline-year wage decile (BREDDT\_LOEN\_BELOEB); Holm within this family.
- **E4:** arm-comparative attrition panel — share of each frozen November c-1 cohort no longer attached to the event workplace by November of year c (IDAN year-c records; ANSXFREM as the forward-separation flag), by event class; explicitly labeled post-treatment-conditioned descriptives with a selection caveat; no test.
- No exploratory result can revise any confirmatory verdict or Stage-4 language. Every sentence stating an exploratory result contains the word “exploratory”.

### 4. ANALYSIS EXECUTION

Confirmatory tier first (Stage 4, frozen code, single run); exploratory tier after (Stage 5), segregated by heading. All estimation in R `fixest` via `Rscript`. Common estimator: outcome `y_iq` (winsorized quarterly GP contact count) on treated `x` event-quarter indicators for the 8 quarters  $Q1(c) \dots Q4(c+1)$ , matched-set fixed effects, cohort `x` calendar-quarter fixed effects, omitted baseline = the four quarters of year c-1; observation weights: exposed 1, controls  $1/k_s$  within set; CR1 SEs clustered on event workplace (each control inherits its set’s event cluster). Min-cell suppression is applied to all outputs BEFORE Holm reporting so suppression can never be outcome-selective.

## 5. EXHIBITS

Every figure and table the report commits to. Disclosure rules binding on EVERY released object: no statistic based on fewer than 5 persons or fewer than 3 workplaces; event-count tables rounded to the nearest 10; weekly trajectory figures pool  $\geq 100$  events and  $\geq 5,000$  person-weeks per plotted point; BRUHON aggregates pass a 50% single-workplace dominance check; released objects are limited to model coefficients with SEs/CIs, balance tables, calibration/benchmark tables, figure aggregates, rounded event counts, the power/MDE tables (both Stage-0 rounds and Stage 2) with the demotion trail, and the analysis/power scripts — never microdata, workplace identifiers (ARBNR/CVRNR/LBNR), or person-level values; suppression applied BEFORE Holm reporting.

### Figures (matplotlib):

- **F1 — Quarterly event-study panel.** Treated  $x$  event-quarter coefficients with 95% CIs across Q1(c-2)..Q4(c+1) (placebo quarters, omitted-baseline marker, terminal-year and post-year quarters); panel  
(a) true arm, panel (b) spurious arm from the pooled model. Feeds: step-18 panels + 4.1/4.2 coefficient tables. Disclosure: coefficients only.
- **F2 — Weekly descriptive trajectory.** Pooled exposed-minus- matched-control weekly GP contact series around the closure-year boundary, with the GRAFT-2 week-0 step diagnostic annotated (step  $\leq 50\%$  of the total Q3(c)+Q4(c) excess; if exceeded, the figure carries the registered caution and quarter-level estimates are foregrounded). Feeds: step 15/17 weekly aggregates. Disclosure:  $\geq 100$  events and  $\geq 5,000$  person-weeks per point.
- **F3 — s(w) successor-share distribution (GRAFT-4).** Histogram of worker-flow successor shares across terminal events with the 0.30/0.70 thresholds and mid-zone mass; bimodality check exhibit. Feeds: step 7. Disclosure: bin counts rounded to 10; bins under min-cell suppressed.
- **F4 — OOH vs daytime relative excess (theta3).** The two relative-excess series over the peri-event window with the bootstrap CI of their gap; primary and YDLTID-augmented classifier side by side. Feeds: 4.3 outputs. Disclosure: coefficients only.
- **F5 — Matching balance (love plot).** SMDs before/after matching for all matching covariates and unmatched auxiliaries; 0.10 line marked. Feeds: step 13. Disclosure: aggregates only.
- **F6 — E1 SYSI-era replication (exploratory; skippable).** Quarterly SYSI-era analogue of F1(a) for cohorts 1992–2003, labeled exploratory, YDLANT-weighted proxy outcome. Feeds: step 22. Disclosure: as F1.

### Tables (booktabs LaTeX):

- **T1 — Event extraction and classification by cohort year.** Terminal events, true/spurious/ambiguous counts (rounded to 10), annual true-closure rate among private size-5+ workplaces, spurious/true ratio. Feeds: step 8.

- **T2 — Classification cross-validation.** Worker-flow classes vs IDAS-native flags (ENDNEDL/IDFREM/REGELFR/FRELFREM/ANDELUDS), full cross-tabulation + agreement rate; GRAFT-3 classification counts under 20/80 and 40/60. Feeds: step 8.
- **T3 — Cohort flow accounting.** Inclusion/exclusion counts per rule (tenure, sector, size, SIKGRUP, missing ANSAAR/pay, multi-cohort dedup), eligible exposed per cohort, match rate, ladder rung used, realized k distribution. Feeds: steps 9–13.
- **T4 — Matching balance.** Per-covariate SMDs (weighted 1/k), caliper compliance, realized mean k. Feeds: step 13.
- **T5 — Power and demotion trail.** Stage-0 power results (BOTH rounds: the voided round-1 run and the remediated re-run) and the Stage-2 blind realized-power/MDE table with binding N\*, realized sets, and any demotions with the fixed-point recomputation. Feeds: deposited script  
– step 14. Published regardless of outcome.
- **T6 — Outcome-neutral check panel.** All eleven registered checks (coverage benchmark by year, linkage floor, closure-rate plausibility, classification agreement, GRAFT-4, balance/match rate, c-2 placebo, HONUUGE lag calibration, floor/ceiling + seasonality, GRAFT-2 step, OOH-share benchmark) with thresholds and PASS/FAIL. Feeds: steps 8, 13, 16, 17, 19, 21 + F2/F3 diagnostics.
- **T7 — Confirmatory results.** theta1, theta2a, theta3: point estimates (contacts/person-quarter and % of baseline), 95% CIs, two-sided p, Holm-adjusted verdicts; the F2 gate row (theta2b, its one-sided 95% upper bound, trigger evaluation); the gated BRUHON fee secondary (if computed); c-2 placebo row; realized MDEs beside every null with the CI-ruled-out magnitudes. Feeds: 4.1–4.4.
- **T8 — Sensitivity suite.** GRAFT-3 thresholds; GRAFT-9 forward-window exclusion (with truncation note); un-winsorized; Poisson QML; death/emigration exclusion — one table, run once. Feeds: 4.4.
- **T9 — Exploratory tier.** H1\_trend (if gated in; otherwise the row states the gate was not met), H2\_diff with quarter-by-quarter differential profile, E3 fee and subgroup analyses (within-family Holm), E1 summary. Every row labeled exploratory. Feeds: 4.5.
- **T10 — Attrition and later-presence accounting.** E4 year-c separation shares by event class (ANSXFREM-based, post-treatment- conditioned label and selection caveat) and the GRAFT-6 later-presence rates with the 2pp trigger status. Feeds: steps 20 + 4.5 E4.

---

## Current source hypotheses (all fields)

```

- id: E3
 tier: exploratory
 statement_initial: null
 statement_source: no frozen bank statement - protocol-defined exploratory
 ↪ module (fee extensions
 and heterogeneity)

```

```

estimand: 'Heterogeneity and fee extensions: theta2a and theta2_diff
↳ analogues for gross
 fees (BRUHON, nominal DKK; calendar price levels difference out within
 ↳ cohort x calendar-quarter
 cells) beyond the gated H2 fee secondary, and heterogeneity of
 ↳ theta1/theta2a by sex (KOENIMP),
 age band (ALDERIMP), and baseline-year wage decile (BREDT_LOEN_BELOEB).'
primary_quantity: Fee analogues and subgroup theta1/theta2a with Holm
↳ correction within
 this exploratory family (severity aid only).

```

## Current source interpretation rows

[]

## Source family scoring and limitations

```

scoring_rule: 'Fixed in advance per the collaboration's registered-report
↳ conventions (trajectory
 fields): at Stage 2 an outcome attaches to each confirmatory hypothesis's
 ↳ frozen statement_initial
 as confirmed | partially_confirmed | reversed, adjudicated strictly against
 ↳ the pre-committed
 interpretation table in this protocol. H1 scores confirmed only if
 ↳ theta1's Holm-adjusted
 two-sided test rejects with the point estimate in the predicted direction
 ↳ AND the gated
 exploratory H1 trend contrast supports the frozen statement's growth rider
 ↳ (beta_Q4c -
 beta_Q1c > 0 with 95% CI excluding zero); it scores partially_confirmed
 ↳ when theta1 Holm-rejects
 in the predicted direction but the growth rider is unsupported (flat
 ↳ elevation - the pre-committed
 reading given annual event dating) or the trend gate is not computed;
 ↳ reversed if theta1
 is Holm-significant in the negative direction. H2 scores confirmed only if
 ↳ theta2a Holm-rejects
 in the predicted direction AND falsification gate F2 does not fire AND the
 ↳ theta2b sign
 pattern is stable across the GRAFT-3 successor-share thresholds;
 ↳ partially_confirmed when
 theta2a Holm-rejects but the comparative quantifier is untestable
 ↳ (spurious-arm kill - verdict
 capped by pre-commitment at 'partially consistent - comparative quantifier
 ↳ untestable');
 reversed if theta2a is Holm-significant in the negative direction OR gate
 ↳ F2 fires (the
 frozen statement's comparative reading is disconfirmed by a non-inert
 ↳ placebo regardless
 of theta2a - registered as disconfirming, never reframed). H3 scores
 ↳ confirmed only if theta3's
 Holm-adjusted test rejects in the predicted direction (including the
 ↳ total-down/OOH-up pattern,
 pre-registered as scoreable support for the composition claim); reversed if
 ↳ Holm-significant
 in the negative direction. A Holm-nonsignificant confirmatory result
 ↳ attaches no confirmation:
 it is reported as a bounded informative null - the 95% CI converted into
 ↳ the excess magnitudes

```

the data rule out against realized MDEs - never 'accepted'. A member  
 ↳ demoted by the Stage-2  
 realized-N gate or killed by its registered criterion attaches no score;  
 ↳ the demotion/kill  
 trail is disclosed. Significant effects in the non-predicted direction are  
 ↳ disconfirming,  
 never supportive. Exploratory quantities (H1\_trend, H2\_diff, E1, E3, E4,  
 ↳ and any demoted  
 member) attach outcomes only under the exploratory heading and are never  
 ↳ scored as preregistered-confirmed.'

assumptions:

- SSSY (2005-2025) and SYSI (1990-2005) capture the universe of publicly  
 ↳ reimbursed primary-care  
 billing for insured residents; absence of rows means no billed care, not  
 ↳ missingness
- KONTAKT is a row-level contact marker whose sum yields contact counts  
 ↳ (in-environment validation  
 with a frozen fallback definition if not)
- ENDNEDL/IDFREM/IDTILB/REGELFR/FRELFREM semantics follow Statistics  
 ↳ Denmark's IDA workplace-continuity  
 documentation, cited directly (no DOI exists for these conventions, per  
 ↳ boundary directive);  
 the primary classification rests on the worker-flow measure s(w) from IDAN,  
 ↳ so misreading  
 an IDAS flag degrades only the cross-check
- November-status annual attachment measures workplace membership; intra-year  
 ↳ attachment changes  
 are unobservable ex ante and the frozen c-1 anchor is the design's answer  
 ↳ to that - with  
 the corollary, named as a design threat, that early year-c separations feed  
 ↳ realized post-job-loss  
 billing into the pre-boundary window under intention-to-follow, a channel  
 ↳ the E4 separation  
 shares and the registered theta1 attachment split bound but cannot remove
- ALDERIMP/KOENIMP (DST-imputed age/sex on SSSY rows) are accurate enough for  
 ↳ stratified matching;  
 persons absent from SSSY in c-3..c-1 are handled by explicit unknown strata
- HONUGE settlement lags are predominantly <= 4 weeks (calibrated on  
 ↳ BEHANDLINGSDATO 2021+;  
 guard-band contingency if not), and lag structure does not differ between  
 ↳ exposed and matched  
 controls within the same calendar quarters
- Death and emigration are unobservable in scope; their post-event incidence  
 ↳ is assumed small;  
 the deflating direction is conservative for the pre-boundary claim and  
 ↳ attenuating for theta2a,  
 with the GRAFT-6 trigger bounding the post-boundary side
- Nominal-fee comparisons within cohort x calendar-quarter cells require no  
 ↳ deflator
- Danish institutional stability of GP billing categories (SPEC2 70-84, 89)  
 ↳ across 2007-2024  
 at the 2-digit level
- 'PUBLISHED-BASIS SIZING ASSUMPTIONS (new, each with a registered failure  
 ↳ path): (i) Danish  
 establishment closure rates among private size-5+ workplaces are in the  
 ↳ published Erhvervsdemografi-consistent  
 1.5-3% band, supporting 150,000-300,000 exposed workers over 15 cohorts -  
 ↳ failure path:  
 the pre-outcome realized-N gate demotes rather than stretches; (ii) the  
 ↳ Danish OOH share

of GP contact volume is in the published 8-12% band (Christensen & Olesen  
 ↳ 1998; LV-KOS 2011  
 studies), simulated at the conservative 8% - failure path: the  
 ↳ published-range benchmark  
 check plus realized-share re-simulation, then demotion or kill; (iii) the  
 ↳ control pool supports  
 1:3 matching at  $\geq 90\%$  match rate - failure path: coarsening ladder,  
 ↳ realized-k accounting  
 in the blind re-simulation, then the gate'

- theta2a's 4-quarter averaging gain over theta1 survives serial correlation  
 ↳ in post-year  
 shocks; a registered serial-correlation discount is built into the Stage-0  
 ↳ re-simulation  
 rather than assumed away
- The persistent-vs-transient overdispersion split (persistent frailty var  
 ↳ 1.25, transient  
 log-var 0.25) is retained unchanged from the first-round simulation -  
 ↳ deliberately NOT revised  
 in the direction that would raise power, so the remediation cannot be  
 ↳ accused of tuning  
 the DGP

failure\_modes:

- 'Realized event or worker counts far below schema-derived expectations  
 ↳ (documentation-only  
 corpus; no realized Ns exist) - the design now bears this risk openly: the  
 ↳ Stage-2 realized-N  
 gate demotes underpowered members loudly and pre-outcome instead of letting  
 ↳ an underpowered  
 confirmatory test run; a sparse-closure Denmark kills or demotes cleanly  
 ↳ rather than quietly'
- KONTAKT or ENDNEDL semantics differ from assumption - both carry frozen  
 ↳ fallback rules,  
 but level shifts in outcomes or event dating would force reported  
 ↳ recalibration and consume  
 credibility
- HONUUGE settlement lag materially longer than 4 weeks - the quarter-level  
 ↳ estimands survive,  
 but the weekly-resolution headline weakens to quarterly resolution in  
 ↳ practice (GRAFT-2  
 makes this explicit and pre-committed)
- ANDELOPS direction-of-absorption semantics may differ from the idea's gloss  
 ↳ - classification  
 routes through IDAN worker flows, so this only weakens the IDAS  
 ↳ cross-check, not the H2  
 machinery
- Spurious-closure events too heterogeneous for a clean placebo - a non-inert  
 ↳ placebo now  
 shows up as gate F2 firing (or exploratory H2\_diff reversal), which caps  
 ↳ H2's verdict by  
 pre-commitment instead of silently contaminating a differential test
- Realized OOH share below the published range - caught by the  
 ↳ published-range benchmark check  
 before estimation; H3 demotes or dies pre-outcome instead of running  
 ↳ underpowered
- Realized N lands below the full-expectation lower bound - the gate demotes  
 ↳ affected members  
 to exploratory pre-outcome; the study survives with a smaller confirmatory  
 ↳ family (or none,  
 in which case it reports as registered exploratory estimation with the  
 ↳ demotion trail published

- the first-round outcome, now a contingency rather than the design)
- Unobservable post-closure death/emigration deflates post-boundary billing -
  - ↳ direction stated,
  - sensitivity pre-specified, GRAFT-6 2pp trigger forces best/worst-case
  - ↳ bounds; no in-scope
  - fix exists
- Public/private sector misclassification via modal ARB\_SEKTORKODE for
  - ↳ multi-establishment
  - firms - attenuates contrasts; share bounded in the balance tables; if
  - ↳ ARB\_SEKTORKODE itself
  - proves absent in-environment (it is unconfirmed in the recorded
  - ↳ Forskningsservice inventory),
  - the frozen ARB\_HOVED\_BRA\_DB07 fallback in contingent\_decisions substitutes,
  - ↳ with counts
  - reported
- Multi-job workers' primary-job assignment may misattribute exposure -
  - ↳ deterministic rule
  - biases toward null; share reported
- 1:3 matching without replacement may exhaust thin strata - realized  $k < 3$ 
  - ↳ accounted honestly
  - in the blind re-simulation rather than assumed; coarsening ladder
  - ↳ pre-committed
- The SYSI-era exploratory replication may be infeasible - bounded as
  - ↳ skipable Stage-5 work
  - that cannot contaminate the confirmatory tier
- Readers may over-read gate F2 as proof of causal job-loss effects - the
  - ↳ associational ceiling,
  - the 'consistent with' vocabulary cap, and the pre-committed F2-fired
  - ↳ interpretation guard
  - the prose

## Source disclosure rules

environment: DST Forskningsservice (no prior data access; gatekeeper  
 ↳ certification per RR  
 secondary-data rules). All analysis occurs inside the environment; only  
 ↳ aggregate outputs  
 leave. No microdata, workplace identifiers (ARBNR/CVRNR/LBNR), or  
 ↳ person-level values are  
 released.

minimum\_cell\_size: 5

minimum\_cell\_rules: no released statistic based on fewer than 5 persons or  
 ↳ fewer than 3 workplaces;  
 event-count tables rounded to the nearest 10; weekly trajectory figures  
 ↳ pool  $\geq 100$  events  
 and  $\geq 5,000$  person-weeks per plotted point; fee (BRUHON) aggregates  
 ↳ subject to a 50% single-workplace  
 dominance check; min-cell suppression applied BEFORE Holm reporting so  
 ↳ suppression can never  
 be outcome-selective

outputs\_leaving\_environment:

- model coefficients with SEs/CIs
- matching balance tables and realized-k distribution
- calibration and benchmark tables (coverage, linkage, HONUGE lag, OOH share)
- figure aggregates (weekly trajectory series under the pooling rules)
- event counts rounded to the nearest 10
- Stage-0 power/MDE tables - BOTH rounds (failed and remediated) - and the
- ↳ Stage-2 blind re-simulation
- output with the demotion trail

- the c-2 placebo, gate F2 outputs, GRAFT-2/4 diagnostics, GRAFT-6 later-presence rates, E4 attrition panel, and all outcome-neutral check outputs (the full audit trail exits under the same rules)
- the analysis and power scripts

statement: 'All analysis occurs inside DST Forskningservice; only aggregate outputs leave.

Rules applied to every released object: (1) no statistic based on fewer than 5 persons or fewer than 3 workplaces; (2) event-count tables rounded to the nearest 10; (3) weekly trajectory figures pool  $\geq 100$  events and  $\geq 5,000$  person-weeks per plotted point; (4) fee (BRUHON) statistics released only as cell means/coefficients over min-N cells, with a dominance check that no single workplace contributes  $> 50\%$  of any released fee aggregate; (5) released objects are limited to: model coefficients with SEs/CIs, balance tables, calibration/benchmark tables, figure aggregates, event counts (rounded), the Stage-0 (both rounds) and Stage-2 power/MDE tables with the demotion trail, and the analysis/power scripts - never microdata, workplace identifiers (ARBNR/CVRNR/LBNR), or person-level values; (6) min-cell suppression is applied before Holm reporting so that suppression can never be outcome-selective; (7) the c-2 placebo, gate F2 outputs, the GRAFT-2/4 diagnostics, the GRAFT-6 later-presence rates, the E4 attrition panel, and all outcome-neutral check outputs are released under the same rules, ensuring the full audit trail exits the environment.'

## Complete registration fields

```
id: atom-000008
topic: Fees and heterogeneity
atom_result: prereg
atom_strength: null
atom_shape: null
evidence_paths: []
concepts: []
literature: []
bonds: []
preregistration:
 version: 1
 mode: adaptive
 question: How do source-defined fee and subgroup contrasts vary within the
 registered exploratory tier?
 rationale: Source-existing coherent scientific question; its related
 outcomes,
 riders and sensitivity specifications remain internal.
 assumption: |-
```

Associational - matched event-study contrasts, a contrastive  
 ↪ falsification gate, and a composition contrast; descriptive for  
 ↪ trajectory panels and the bounded priority claim. Binding vocabulary  
 ↪ rules: 'is associated with' / 'consistent with' / 'suggests' only;  
 ↪ 'effect', 'impact', 'causes' prohibited throughout; anticipation  
 ↪ always worded as 'pre-boundary excess' relative to the annually dated  
 ↪ closure-year boundary, never as announcement-dated anticipation (S8);  
 ↪ no phenomenon-discovery or Danish-first framing (S1-S3); H2 novelty  
 ↪ confined to the placebo's application to health billing (S7). H2  
 ↪ verdict semantics are fixed in advance: 'consistent with H2' requires  
 ↪ BOTH the confirmatory core (theta2a Holm-significant, predicted  
 ↪ direction) AND a passing falsification gate F2; the strict  
 ↪ comparative magnitude is only ever reported as an exploratory  
 ↪ estimate with its CI. Exploratory results (H1\_trend, H2\_diff, E1, E3,  
 ↪ E4, and any gate-demoted member) carry the word 'exploratory' in the  
 ↪ heading and in every sentence stating them. The prose label never  
 ↪ exceeds the evidence label.  
 Compatibility ratings below are neutral import metadata, not a new  
 ↪ scientific evaluation or empirical scores.  
 target\_strength: associational  
 target\_shape: comparison  
 ratings:  
   importance: 3  
   novelty: 3  
   feasibility: 3  
   data\_advantage: 3  
   informativeness: 0.5  
 tables:  
 - IDAN  
 - IDAS  
 - SSSY  
 activation: state(atom-000003) == "event\_study\_ready" AND done(atom-000004)  
   AND done(atom-000005) AND done(atom-000006)  
 verdicts:  
 - key: exploratory\_pattern  
   region: The source-defined module yields interpretable  
     ↪ estimates/descriptives  
       after its own eligibility checks; retain unavailable components  
       ↪ explicitly.  
   says: Report the complete source exploratory pattern without confirmatory  
     or causal promotion.  
 - key: source\_skip  
   region: The source-defined optional module/arm is unavailable or its  
     ↪ registered  
       skip criterion is met.  
   says: Record the source skip and its reason; never invent a negative or  
     null finding.  
 - key: boundary  
   region: Required source-defined inputs or scientific definitions are  
     ↪ demonstrably  
       unavailable; retain available components and the exact unresolved scope.  
   says: An evidenced boundary, not a statistical null. A missing  
     ↪ implementation  
       or timeout alone is operational no\_result, not this state.  
 - key: otherwise  
   region: otherwise  
   says: The registered judgment is unresolved; report all completed  
     ↪ components  
       without inventing support, equivalence or absent results.  
 protocol: |

# How do source-defined fee and subgroup contrasts vary within the  
↳ registered exploratory tier?

## Source fidelity and implementation order

This is a format conversion of DS1.1, not a new study or independent  
↳ acceptance.  
Source audit: minor\_revisions. No observed Danish outcomes are supplied.  
The source execution plan explicitly states that final prereg.yaml governs  
disagreements. Exact current YAML selections below therefore govern older  
execution-plan excerpts; these excerpts are retained as implementation  
↳ detail,  
not a competing registration. See docs/provenance/source-discrepancies.md.  
The complete shared data/validity contracts are embedded in Atoms 1--3;  
↳ read  
those native protocols when executing a downstream Atom, even without  
↳ docs/.  
Do not substitute a scientific default for an unresolved source  
↳ contradiction.

Preserve source stages 0 -> 1 -> 2 -> 3 -> 4 -> 5 inside the  
↳ implementation.  
Source Stage-0 simulations are historical design evidence, not study  
↳ outcomes;  
do not silently redo or improve them. Gate 14 must be evaluated from  
↳ baseline  
data before post-baseline outcomes. All surviving Stage-4 quantities form  
↳ ONE  
Holm family. The three question-level Atoms may share the same frozen  
↳ family  
fit and tidy results; do not apply Holm separately per Atom or adjudicate  
↳ a  
member before the complete surviving family's p-values exist. Cache/reuse  
identical governed intermediate results with provenance when available; a  
direct-dispatch call must obtain or compute its exact source-defined  
↳ inputs,  
not pretend its predecessors succeeded. Stage-5 internal riders remain  
exploratory, after the Stage-4 family is frozen, even when housed in  
↳ H1/H2.

Reference files under docs/execution/ are the ORIGINAL SKELETONS, not  
↳ tested  
turnkey implementations. Adapt paths and complete engineering in a  
↳ governed  
workspace; preserve source scientific rules and disclose unresolved  
↳ choices.  
Write a result record with component estimates, uncertainty, family  
↳ eligibility,  
source interpretation row, diagnostics and boundary/skip dispositions.  
↳ Released  
outputs follow source disclosure rules; confidential rows stay governed.  
Local native states summarize the source judgment; neither a state name  
↳ nor  
source power forecasts prove equivalence or causal identification. A  
↳ bounded  
claim requires its actual interval and registered rule. Retain unresolved  
↳ source  
wording conflicts; do not convert mere nonsignificance into scientific  
↳ null.

## Frozen scope and vocabulary

```
```yaml
claim_class_ceiling: associational
statement: 'Associational - matched event-study contrasts, a contrastive
↪ falsification gate,
and a composition contrast; descriptive for trajectory panels and the
↪ bounded priority claim.
Binding vocabulary rules: 'is associated with' / 'consistent with'
↪ / 'suggests' only;
'effect', 'impact', 'causes' prohibited throughout; anticipation
↪ always worded as
'pre-boundary excess' relative to the annually dated closure-year
↪ boundary, never as announcement-dated
anticipation (S8); no phenomenon-discovery or Danish-first framing
↪ (S1-S3); H2 novelty confined
to the placebo's application to health billing (S7). H2 verdict
↪ semantics are fixed in
advance: 'consistent with H2' requires BOTH the confirmatory core
↪ (theta2a Holm-significant,
predicted direction) AND a passing falsification gate F2; the strict
↪ comparative magnitude
is only ever reported as an exploratory estimate with its CI.
↪ Exploratory results (H1_trend,
H2_diff, E1, E3, E4, and any gate-demoted member) carry the word
↪ 'exploratory' in the
heading and in every sentence stating them. The prose label never
↪ exceeds the evidence label.'
frozen_parameters:
seed: 20260831
gp_code_list: daytime GP = SPEC2 in {70..80}
ooh_code_list: out-of-hours GP = SPEC2 in {81, 82, 83, 84, 89}
ooh_robustness_classifier: adds YDLTID in {2, 3, 4, 5, 7, 8, 9}
spec2_evening_gp_disclosure: 'Registered choice, disclosed: the parsed
↪ value labels mark
SPEC2 71 ("General practitioner, consultation, evening etc."), 73
↪ ("General practitioner,
telephone consultation, evening etc."), and 75 ("General
↪ practitioner, home visit, evening
etc.") as explicitly off-schedule GP services; they nonetheless
↪ remain on the daytime
side of the PRIMARY classifier because H3's out-of-hours construct
↪ and its published
power basis are keyed to the laegevagt out-of-hours service (SPEC2
↪ 81-84, 89), the quantity
the published Danish 8-12% share describes. Attenuation direction,
↪ both readings: counting
evening GP as daytime attenuates theta3 toward zero - conservative
↪ against the positive
H3 conclusion, anti-conservative for the null-bounding
↪ "compositionally unremarkable"
reading, whose Stage-2 statement therefore carries a mandatory
↪ caveat. In-environment
validation must verify that the YDLTID-augmented robustness
↪ classifier captures SPEC2
71/73/75 rows (their evening time codes 2-5); if YDLTID proves
↪ unpopulated on those rows,
the robustness classifier - and only the robustness classifier - is
↪ amended mechanically
```

```

    and with disclosure to include SPEC2 71/73/75 directly. The SPEC2
    ↪ 71/73/75 volume share
    is reported, and both classifiers appear side by side in Table T7.'
    successor_share_thresholds: 'true closure: ENDNEDL final closure AND
    ↪ worker-flow max-successor
    share s(w) < 0.30; spurious closure: forward identity via
    ↪ IDFREEM/REGELFR/FRELFREM AND
    s(w) >= 0.70; ambiguous mid-zone excluded; GRAFT-3 sensitivities
    ↪ 0.20/0.80 and 0.40/0.60'
    quarter_mapping: 'ISO week of HONUGE: 1-13 / 14-26 / 27-39 / 40-52; ISO
    ↪ week 53 to Q4'
    winsorization: person-quarter outcomes at the 99.9th percentile within
    ↪ era x calendar quarter
    matching: 1:3 nearest-neighbor Mahalanobis within exact strata, without
    ↪ replacement across
    sets, target k = 3, minimum k = 1; calipers |dif baseline contacts|
    ↪ <= 3 and |dif log(1+BRUHON)|
    <= 0.5; exposed processed in seeded random order (seed 20260831)
    software: R >= 4.3, fixest::feols (OLS) and fixest::fepois (Poisson QML
    ↪ robustness); versions
    frozen in the registered script
    note: Code lists and thresholds frozen before any data access; code
    ↪ labels quoted verbatim
    from parsed DST value labels and re-validated in-environment before
    ↪ estimation (frozen
    fallback rules in contingent_decisions).
...
### 4.5 Exploratory tier (Stage 5; heading-segregated; Holm within family
    ↪ as severity aid)

- **H1_trend (gated):** beta_Q4c - beta_Q1c from the true-arm model, 95%
  CI, CR1 clusters - computed ONLY if theta1 Holm-rejected positively;
  otherwise the quarterly trajectory is shown descriptively with no
  contrast statistic and no timing-shape language in the Summary.
- **H2_diff:** theta2_diff = mean_k[beta_true,k - beta_spurious,k] over
  the post-year quarters from the pooled model, 95% CI, plus the
  quarter-by-quarter differential profile (the "more persistent" rider).
  A flagged estimate with spurious > true triggers the pre-committed
  non-inert-placebo interpretation and reinforces whatever F2 showed.
- **E1:** SYSI-era analogue of theta1 from step 22, 95% CI, descriptive
  replication only; never pooled across the vintage break.
- **E3:** BRUHON fee analogues of theta2a/theta2_diff beyond the gated
  secondary; heterogeneity of theta1/theta2a by sex (KOENIMP), age band
  (ALDERIMP), and baseline-year wage decile (BREDT_LOEN_BELOEB); Holm
  within this family.
- **E4:** arm-comparative attrition panel - share of each frozen
  November c-1 cohort no longer attached to the event workplace by
  November of year c (IDAN year-c records; ANSXFREM as the
  forward-separation flag), by event class; explicitly labeled
  post-treatment-conditioned descriptives with a selection caveat; no
  test.
- No exploratory result can revise any confirmatory verdict or Stage-4
  language. Every sentence stating an exploratory result contains the
  word "exploratory".

```

4. ANALYSIS EXECUTION

Confirmatory tier first (Stage 4, frozen code, single run); exploratory tier after (Stage 5), segregated by heading. All estimation in R fixest via Rscript. Common estimator: outcome y_iq (winsorized quarterly GP

contact count) on treated x event-quarter indicators for the 8 quarters Q1(c)..Q4(c+1), matched-set fixed effects, cohort x calendar-quarter fixed effects, omitted baseline = the four quarters of year c-1; observation weights: exposed 1, controls 1/k_s within set; CR1 SEs clustered on event workplace (each control inherits its set's event cluster). Min-cell suppression is applied to all outputs BEFORE Holm reporting so suppression can never be outcome-selective.

5. EXHIBITS

Every figure and table the report commits to. Disclosure rules binding on EVERY released object: no statistic based on fewer than 5 persons or fewer than 3 workplaces; event-count tables rounded to the nearest 10; weekly trajectory figures pool ≥ 100 events and $\geq 5,000$ person-weeks per plotted point; BRUHON aggregates pass a 50% single-workplace dominance check; released objects are limited to model coefficients with SEs/CIs, balance tables, calibration/benchmark tables, figure aggregates, rounded event counts, the power/MDE tables (both Stage-0 rounds and Stage 2) with the demotion trail, and the analysis/power scripts - never microdata, workplace identifiers (ARBNR/CVRNR/LBNR), or person-level values; suppression applied BEFORE Holm reporting.

Figures (matplotlib):

- **F1 - Quarterly event-study panel.** Treated x event-quarter coefficients with 95% CIs across Q1(c-2)..Q4(c+1) (placebo quarters, omitted-baseline marker, terminal-year and post-year quarters); panel (a) true arm, panel (b) spurious arm from the pooled model. Feeds: step-18 panels + 4.1/4.2 coefficient tables. Disclosure: coefficients only.
- **F2 - Weekly descriptive trajectory.** Pooled exposed-minus-matched-control weekly GP contact series around the closure-year boundary, with the GRAFT-2 week-0 step diagnostic annotated (step $\leq 50\%$ of the total Q3(c)+Q4(c) excess; if exceeded, the figure carries the registered caution and quarter-level estimates are foregrounded). Feeds: step 15/17 weekly aggregates. Disclosure: ≥ 100 events and $\geq 5,000$ person-weeks per point.
- **F3 - s(w) successor-share distribution (GRAFT-4).** Histogram of worker-flow successor shares across terminal events with the 0.30/0.70 thresholds and mid-zone mass; bimodality check exhibit. Feeds: step 7. Disclosure: bin counts rounded to 10; bins under min-cell suppressed.
- **F4 - OOH vs daytime relative excess (theta3).** The two relative-excess series over the peri-event window with the bootstrap CI of their gap; primary and YDLTID-augmented classifier side by side. Feeds: 4.3 outputs. Disclosure: coefficients only.
- **F5 - Matching balance (love plot).** SMDs before/after matching for all matching covariates and unmatched auxiliaries; 0.10 line marked. Feeds: step 13. Disclosure: aggregates only.
- **F6 - E1 SYSI-era replication (exploratory; skippable).** Quarterly SYSI-era analogue of F1(a) for cohorts 1992-2003, labeled exploratory, YDLANT-weighted proxy outcome. Feeds: step 22. Disclosure: as F1.

Tables (booktabs LaTeX):

- **T1 - Event extraction and classification by cohort year.** Terminal events, true/spurious/ambiguous counts (rounded to 10), annual true-closure rate among private size-5+ workplaces, spurious/true ratio. Feeds: step 8.
- **T2 - Classification cross-validation.** Worker-flow classes vs IDAS-native flags (ENDNEDL/IDFREM/REGELFR/FRELFREM/ANDELUDS), full

- cross-tabulation + agreement rate; GRAFT-3 classification counts under 20/80 and 40/60. Feeds: step 8.
- **T3 - Cohort flow accounting.** Inclusion/exclusion counts per rule (tenure, sector, size, SIKGRUP, missing ANSAAR/pay, multi-cohort dedup), eligible exposed per cohort, match rate, ladder rung used, realized k distribution. Feeds: steps 9-13.
 - **T4 - Matching balance.** Per-covariate SMDs (weighted 1/k), caliper compliance, realized mean k. Feeds: step 13.
 - **T5 - Power and demotion trail.** Stage-0 power results (BOTH rounds: the voided round-1 run and the remediated re-run) and the Stage-2 blind realized-power/MDE table with binding N*, realized sets, and any demotions with the fixed-point recomputation. Feeds: deposited script + step 14. Published regardless of outcome.
 - **T6 - Outcome-neutral check panel.** All eleven registered checks (coverage benchmark by year, linkage floor, closure-rate plausibility, classification agreement, GRAFT-4, balance/match rate, c-2 placebo, HONUGE lag calibration, floor/ceiling + seasonality, GRAFT-2 step, OOH-share benchmark) with thresholds and PASS/FAIL. Feeds: steps 8, 13, 16, 17, 19, 21 + F2/F3 diagnostics.
 - **T7 - Confirmatory results.** theta1, theta2a, theta3: point estimates (contacts/person-quarter and % of baseline), 95% CIs, two-sided p, Holm-adjusted verdicts; the F2 gate row (theta2b, its one-sided 95% upper bound, trigger evaluation); the gated BRUHON fee secondary (if computed); c-2 placebo row; realized MDEs beside every null with the CI-ruled-out magnitudes. Feeds: 4.1-4.4.
 - **T8 - Sensitivity suite.** GRAFT-3 thresholds; GRAFT-9 forward-window exclusion (with truncation note); un-winsorized; Poisson QML; death/emigration exclusion - one table, run once. Feeds: 4.4.
 - **T9 - Exploratory tier.** H1_trend (if gated in; otherwise the row states the gate was not met), H2_diff with quarter-by-quarter differential profile, E3 fee and subgroup analyses (within-family Holm), E1 summary. Every row labeled exploratory. Feeds: 4.5.
 - **T10 - Attrition and later-presence accounting.** E4 year-c separation shares by event class (ANSXFREM-based, post-treatment-conditioned label and selection caveat) and the GRAFT-6 later-presence rates with the 2pp trigger status. Feeds: steps 20 + 4.5 E4.

Current source hypotheses (all fields)

```

```yaml
- id: E3
 tier: exploratory
 statement_initial: null
 statement_source: no frozen bank statement - protocol-defined
 ↳ exploratory module (fee extensions
 and heterogeneity)
 estimand: 'Heterogeneity and fee extensions: theta2a and theta2_diff
 ↳ analogues for gross
 fees (BRUHON, nominal DKK; calendar price levels difference out
 ↳ within cohort x calendar-quarter
 cells) beyond the gated H2 fee secondary, and heterogeneity of
 ↳ theta1/theta2a by sex (KOENIMP),
 age band (ALDERIMP), and baseline-year wage decile
 ↳ (BREDT_LOEN_BELOEB).'
 primary_quantity: Fee analogues and subgroup theta1/theta2a with Holm
 ↳ correction within
 this exploratory family (severity aid only).
```

```

...

Current source interpretation rows

```yaml
[]
```

Source family scoring and limitations

```yaml
scoring_rule: 'Fixed in advance per the collaboration''s
  ↳ registered-report conventions (trajectory
    fields): at Stage 2 an outcome attaches to each confirmatory
    ↳ hypothesis''s frozen statement_initial
    as confirmed | partially_confirmed | reversed, adjudicated strictly
    ↳ against the pre-committed
    interpretation table in this protocol. H1 scores confirmed only if
    ↳ theta1''s Holm-adjusted
    two-sided test rejects with the point estimate in the predicted
    ↳ direction AND the gated
    exploratory H1_trend contrast supports the frozen statement''s growth
    ↳ rider (beta_Q4c -
    beta_Q1c > 0 with 95% CI excluding zero); it scores partially_confirmed
    ↳ when theta1 Holm-rejects
    in the predicted direction but the growth rider is unsupported (flat
    ↳ elevation - the pre-committed
    reading given annual event dating) or the trend gate is not computed;
    ↳ reversed if theta1
    is Holm-significant in the negative direction. H2 scores confirmed only
    ↳ if theta2a Holm-rejects
    in the predicted direction AND falsification gate F2 does not fire AND
    ↳ the theta2b sign
    pattern is stable across the GRAFT-3 successor-share thresholds;
    ↳ partially_confirmed when
    theta2a Holm-rejects but the comparative quantifier is untestable
    ↳ (spurious-arm kill - verdict
    capped by pre-commitment at ''partially consistent - comparative
    ↳ quantifier untestable'');
    reversed if theta2a is Holm-significant in the negative direction OR
    ↳ gate F2 fires (the
    frozen statement''s comparative reading is disconfirmed by a non-inert
    ↳ placebo regardless
    of theta2a - registered as disconfirming, never reframed). H3 scores
    ↳ confirmed only if theta3''s
    Holm-adjusted test rejects in the predicted direction (including the
    ↳ total-down/OOH-up pattern,
    pre-registered as scoreable support for the composition claim);
    ↳ reversed if Holm-significant
    in the negative direction. A Holm-nonsignificant confirmatory result
    ↳ attaches no confirmation:
    it is reported as a bounded informative null - the 95% CI converted
    ↳ into the excess magnitudes
    the data rule out against realized MDEs - never ''accepted''. A member
    ↳ demoted by the Stage-2
    realized-N gate or killed by its registered criterion attaches no
    ↳ score; the demotion/kill
    trail is disclosed. Significant effects in the non-predicted direction
    ↳ are disconfirming,

```

never supportive. Exploratory quantities (H1_trend, H2_diff, E1, E3, E4, and any demoted member) attach outcomes only under the exploratory heading and are never scored as preregistered-confirmed.'

assumptions:

- SSSY (2005-2025) and SYSI (1990-2005) capture the universe of publicly reimbursed primary-care billing for insured residents; absence of rows means no billed care, not missingness
- KONTAKT is a row-level contact marker whose sum yields contact counts (in-environment validation with a frozen fallback definition if not)
- ENDNEDL/IDFREM/IDTILB/REGELFR/FRELFREM semantics follow Statistics Denmark's IDA workplace-continuity documentation, cited directly (no DOI exists for these conventions, per boundary directive); the primary classification rests on the worker-flow measure s(w) from IDAN, so misreading an IDAS flag degrades only the cross-check
- November-status annual attachment measures workplace membership; intra-year attachment changes are unobservable ex ante and the frozen c-1 anchor is the design's answer to that - with the corollary, named as a design threat, that early year-c separations feed realized post-job-loss billing into the pre-boundary window under intention-to-follow, a channel the E4 separation shares and the registered theta1 attachment split bound but cannot remove
- ALDERIMP/KOENIMP (DST-imputed age/sex on SSSY rows) are accurate enough for stratified matching; persons absent from SSSY in c-3..c-1 are handled by explicit unknown strata
- HONUGE settlement lags are predominantly ≤ 4 weeks (calibrated on BEHANDLINGSDATO 2021+; guard-band contingency if not), and lag structure does not differ between exposed and matched controls within the same calendar quarters
- Death and emigration are unobservable in scope; their post-event incidence is assumed small; the deflating direction is conservative for the pre-boundary claim and attenuating for theta2a, with the GRAFT-6 trigger bounding the post-boundary side
- Nominal-fee comparisons within cohort x calendar-quarter cells require no deflator
- Danish institutional stability of GP billing categories (SPEC2 70-84, 89) across 2007-2024 at the 2-digit level
- 'PUBLISHED-BASIS SIZING ASSUMPTIONS (new, each with a registered failure path): (i) Danish establishment closure rates among private size-5+ workplaces are in the published Erhvervsdemografi-consistent 1.5-3% band, supporting 150,000-300,000 exposed workers over 15 cohorts - failure path: the pre-outcome realized-N gate demotes rather than stretches; (ii) the Danish OOH share of GP contact volume is in the published 8-12% band (Christensen & Olesen 1998; LV-KOS 2011 studies), simulated at the conservative 8% - failure path: the published-range benchmark

- check plus realized-share re-simulation, then demotion or kill; (iii)
 - ↪ the control pool supports
- 1:3 matching at $\geq 90\%$ match rate - failure path: coarsening ladder,
 - ↪ realized-k accounting
- in the blind re-simulation, then the gate'
- theta2a's 4-quarter averaging gain over theta1 survives serial
 - ↪ correlation in post-year
- shocks; a registered serial-correlation discount is built into the
 - ↪ Stage-0 re-simulation
- rather than assumed away
- The persistent-vs-transient overdispersion split (persistent frailty
 - ↪ var 1.25, transient
- log-var 0.25) is retained unchanged from the first-round simulation -
 - ↪ deliberately NOT revised
- in the direction that would raise power, so the remediation cannot be
 - ↪ accused of tuning
- the DGP
- failure_modes:
 - 'Realized event or worker counts far below schema-derived expectations
 - ↪ (documentation-only
 - corpus; no realized Ns exist) - the design now bears this risk openly:
 - ↪ the Stage-2 realized-N
 - gate demotes underpowered members loudly and pre-outcome instead of
 - ↪ letting an underpowered
 - confirmatory test run; a sparse-closure Denmark kills or demotes
 - ↪ cleanly rather than quietly'
 - KONTAKT or ENDNEDL semantics differ from assumption - both carry frozen
 - ↪ fallback rules,
 - but level shifts in outcomes or event dating would force reported
 - ↪ recalibration and consume
 - credibility
 - HONUGE settlement lag materially longer than 4 weeks - the
 - ↪ quarter-level estimands survive,
 - but the weekly-resolution headline weakens to quarterly resolution in
 - ↪ practice (GRAFT-2
 - makes this explicit and pre-committed)
 - ANDELOPS direction-of-absorption semantics may differ from the idea's
 - ↪ gloss - classification
 - routes through IDAN worker flows, so this only weakens the IDAS
 - ↪ cross-check, not the H2
 - machinery
 - Spurious-closure events too heterogeneous for a clean placebo - a
 - ↪ non-inert placebo now
 - shows up as gate F2 firing (or exploratory H2_diff reversal), which
 - ↪ caps H2's verdict by
 - pre-commitment instead of silently contaminating a differential test
 - Realized OOH share below the published range - caught by the
 - ↪ published-range benchmark check
 - before estimation; H3 demotes or dies pre-outcome instead of running
 - ↪ underpowered
 - Realized N lands below the full-expectation lower bound - the gate
 - ↪ demotes affected members
 - to exploratory pre-outcome; the study survives with a smaller
 - ↪ confirmatory family (or none,
 - in which case it reports as registered exploratory estimation with the
 - ↪ demotion trail published
 - the first-round outcome, now a contingency rather than the design)
 - Unobservable post-closure death/emigration deflates post-boundary
 - ↪ billing - direction stated,

```

sensitivity pre-specified, GRAFT-6 2pp trigger forces best/worst-case
↳ bounds; no in-scope
fix exists
- Public/private sector misclassification via modal ARB_SEKTORKODE for
↳ multi-establishment
firms - attenuates contrasts; share bounded in the balance tables; if
↳ ARB_SEKTORKODE itself
proves absent in-environment (it is unconfirmed in the recorded
↳ Forskningsservice inventory),
the frozen ARB_HOVED_BRA_DB07 fallback in contingent_decisions
↳ substitutes, with counts
reported
- Multi-job workers' primary-job assignment may misattribute exposure -
↳ deterministic rule
biases toward null; share reported
- 1:3 matching without replacement may exhaust thin strata - realized k <
↳ 3 accounted honestly
in the blind re-simulation rather than assumed; coarsening ladder
↳ pre-committed
- The SYSI-era exploratory replication may be infeasible - bounded as
↳ skippable Stage-5 work
that cannot contaminate the confirmatory tier
- Readers may over-read gate F2 as proof of causal job-loss effects - the
↳ associational ceiling,
the 'consistent with' vocabulary cap, and the pre-committed F2-fired
↳ interpretation guard
the prose
...

```

Source disclosure rules

```

```yaml
environment: DST Forskningsservice (no prior data access; gatekeeper
↳ certification per RR
secondary-data rules). All analysis occurs inside the environment; only
↳ aggregate outputs
leave. No microdata, workplace identifiers (ARBNR/CVRNR/LBNR), or
↳ person-level values are
released.
minimum_cell_size: 5
minimum_cell_rules: no released statistic based on fewer than 5 persons
↳ or fewer than 3 workplaces;
event-count tables rounded to the nearest 10; weekly trajectory figures
↳ pool >= 100 events
and >= 5,000 person-weeks per plotted point; fee (BRUHON) aggregates
↳ subject to a 50% single-workplace
dominance check; min-cell suppression applied BEFORE Holm reporting so
↳ suppression can never
be outcome-selective
outputs_leaving_environment:
- model coefficients with SEs/CIs
- matching balance tables and realized-k distribution
- calibration and benchmark tables (coverage, linkage, HONUGE lag, OOH
↳ share)
- figure aggregates (weekly trajectory series under the pooling rules)
- event counts rounded to the nearest 10
- Stage-0 power/MDE tables - BOTH rounds (failed and remediated) - and
↳ the Stage-2 blind re-simulation
output with the demotion trail

```

```

- the c-2 placebo, gate F2 outputs, GRAFT-2/4 diagnostics, GRAFT-6
 ↳ later-presence rates, E4
 attrition panel, and all outcome-neutral check outputs (the full audit
 ↳ trail exits under
 the same rules)
- the analysis and power scripts
statement: 'All analysis occurs inside DST Forskningservice; only
 ↳ aggregate outputs leave.
 Rules applied to every released object: (1) no statistic based on fewer
 ↳ than 5 persons or
 fewer than 3 workplaces; (2) event-count tables rounded to the nearest
 ↳ 10; (3) weekly trajectory
 figures pool >= 100 events and >= 5,000 person-weeks per plotted point;
 ↳ (4) fee (BRUHON)
 statistics released only as cell means/coefficients over min-N cells,
 ↳ with a dominance check
 that no single workplace contributes > 50% of any released fee
 ↳ aggregate; (5) released objects
 are limited to: model coefficients with SEs/CIs, balance tables,
 ↳ calibration/benchmark tables,
 figure aggregates, event counts (rounded), the Stage-0 (both rounds)
 ↳ and Stage-2 power/MDE
 tables with the demotion trail, and the analysis/power scripts - never
 ↳ microdata, workplace
 identifiers (ARBNR/CVRNR/LBNR), or person-level values; (6) min-cell
 ↳ suppression is applied
 before Holm reporting so that suppression can never be
 ↳ outcome-selective; (7) the c-2 placebo,
 gate F2 outputs, the GRAFT-2/4 diagnostics, the GRAFT-6 later-presence
 ↳ rates, the E4 attrition
 panel, and all outcome-neutral check outputs are released under the
 ↳ same rules, ensuring
 the full audit trail exits the environment.'
...

```

## Pending registered experiment

No Danish study outcome has been observed for this bundle. See the complete registration above.

### atom-000009: Within-year separation

## How much within-year separation and later-presence selection limits the pre-boundary interpretation?

### Source fidelity and implementation order

This is a format conversion of DS1.1, not a new study or independent acceptance. Source audit: minor\_revisions. No observed Danish outcomes are supplied. The source execution plan explicitly states that final prereg.yaml governs disagreements. Exact current YAML selections below therefore govern older execution-plan excerpts; these excerpts are retained as implementation detail, not a competing registration. See docs/provenance/source-discrepancies.md. The complete shared data/validity contracts are embedded in Atoms 1–3; read those native protocols when executing a

downstream Atom, even without docs/. Do not substitute a scientific default for an unresolved source contradiction.

Preserve source stages 0 -> 1 -> 2 -> 3 -> 4 -> 5 inside the implementation. Source Stage-0 simulations are historical design evidence, not study outcomes; do not silently redo or improve them. Gate 14 must be evaluated from baseline data before post-baseline outcomes. All surviving Stage-4 quantities form ONE Holm family. The three question-level Atoms may share the same frozen family fit and tidy results; do not apply Holm separately per Atom or adjudicate a member before the complete surviving family's p-values exist. Cache/reuse identical governed intermediate results with provenance when available; a direct-dispatch call must obtain or compute its exact source-defined inputs, not pretend its predecessors succeeded. Stage-5 internal riders remain exploratory, after the Stage-4 family is frozen, even when housed in H1/H2.

Reference files under docs/execution/ are the ORIGINAL SKELETONS, not tested turnkey implementations. Adapt paths and complete engineering in a governed workspace; preserve source scientific rules and disclose unresolved choices. Write a result record with component estimates, uncertainty, family eligibility, source interpretation row, diagnostics and boundary/skip dispositions. Released outputs follow source disclosure rules; confidential rows stay governed. Local native states summarize the source judgment; neither a state name nor source power forecasts prove equivalence or causal identification. A bounded claim requires its actual interval and registered rule. Retain unresolved source wording conflicts; do not convert mere nonsignificance into scientific null.

## Frozen scope and vocabulary

```
claim_class_ceiling: associational
statement: 'Associational - matched event-study contrasts, a contrastive
↳ falsification gate,
 and a composition contrast; descriptive for trajectory panels and the
↳ bounded priority claim.
Binding vocabulary rules: 'is associated with' / 'consistent with' /
↳ 'suggests' only;
'effect', 'impact', 'causes' prohibited throughout; anticipation
↳ always worded as
'pre-boundary excess' relative to the annually dated closure-year
↳ boundary, never as announcement-dated
anticipation (S8); no phenomenon-discovery or Danish-first framing (S1-S3);
↳ H2 novelty confined
to the placebo's application to health billing (S7). H2 verdict semantics
↳ are fixed in
advance: 'consistent with H2' requires BOTH the confirmatory core
↳ (theta2a Holm-significant,
predicted direction) AND a passing falsification gate F2; the strict
↳ comparative magnitude
is only ever reported as an exploratory estimate with its CI. Exploratory
↳ results (H1_trend,
H2_diff, E1, E3, E4, and any gate-demoted member) carry the word
↳ 'exploratory' in the
heading and in every sentence stating them. The prose label never exceeds
↳ the evidence label.'
frozen_parameters:
seed: 20260831
```

```

gp_code_list: daytime GP = SPEC2 in {70..80}
ooh_code_list: out-of-hours GP = SPEC2 in {81, 82, 83, 84, 89}
ooh_robustness_classifier: adds YDLTID in {2, 3, 4, 5, 7, 8, 9}
spec2_evening_gp_disclosure: 'Registered choice, disclosed: the parsed
 ↳ value labels mark
 SPEC2 71 ("General practitioner, consultation, evening etc."), 73
 ↳ ("General practitioner,
 telephone consultation, evening etc."), and 75 ("General practitioner,
 ↳ home visit, evening
 etc.") as explicitly off-schedule GP services; they nonetheless remain on
 ↳ the daytime
 side of the PRIMARY classifier because H3's out-of-hours construct and
 ↳ its published
 power basis are keyed to the laegevagt out-of-hours service (SPEC2 81-84,
 ↳ 89), the quantity
 the published Danish 8-12% share describes. Attenuation direction, both
 ↳ readings: counting
 evening GP as daytime attenuates theta3 toward zero - conservative
 ↳ against the positive
 H3 conclusion, anti-conservative for the null-bounding "compositionally
 ↳ unremarkable"
 reading, whose Stage-2 statement therefore carries a mandatory caveat.
 ↳ In-environment
 validation must verify that the YDLTID-augmented robustness classifier
 ↳ captures SPEC2
 71/73/75 rows (their evening time codes 2-5); if YDLTID proves
 ↳ unpopulated on those rows,
 the robustness classifier - and only the robustness classifier - is
 ↳ amended mechanically
 and with disclosure to include SPEC2 71/73/75 directly. The SPEC2
 ↳ 71/73/75 volume share
 is reported, and both classifiers appear side by side in Table T7.'
successor_share_thresholds: 'true closure: ENDNEDL final closure AND
 ↳ worker-flow max-successor
 share s(w) < 0.30; spurious closure: forward identity via
 ↳ IDFREEM/REGELFR/FRELFREM AND
 s(w) >= 0.70; ambiguous mid-zone excluded; GRAFT-3 sensitivities
 ↳ 0.20/0.80 and 0.40/0.60'
quarter_mapping: 'ISO week of HONUGE: 1-13 / 14-26 / 27-39 / 40-52; ISO
 ↳ week 53 to Q4'
winsorization: person-quarter outcomes at the 99.9th percentile within era
 ↳ x calendar quarter
matching: 1:3 nearest-neighbor Mahalanobis within exact strata, without
 ↳ replacement across
 sets, target k = 3, minimum k = 1; calipers |dif baseline contacts| <= 3
 ↳ and |dif log(1+BRUHON)|
 <= 0.5; exposed processed in seeded random order (seed 20260831)
software: R >= 4.3, fixest::feols (OLS) and fixest::fepois (Poisson QML
 ↳ robustness); versions
 frozen in the registered script
note: Code lists and thresholds frozen before any data access; code labels
 ↳ quoted verbatim
 from parsed DST value labels and re-validated in-environment before
 ↳ estimation (frozen
 fallback rules in contingent_decisions).

```

#### 4.5 Exploratory tier (Stage 5; heading-segregated; Holm within family as severity aid)

- **H1\_trend (gated):**  $\text{beta\_Q4c} - \text{beta\_Q1c}$  from the true-arm model, 95% CI, CR1 clusters — computed ONLY if theta1 Holm-rejected positively; otherwise the quarterly trajectory is shown descriptively with no contrast statistic and no timing-shape language in the Summary.
- **H2\_diff:**  $\text{theta2\_diff} = \text{mean\_k}[\text{beta\_true,k} - \text{beta\_spurious,k}]$  over the post-year quarters from the pooled model, 95% CI, plus the quarter-by-quarter differential profile (the “more persistent” rider). A flagged estimate with  $\text{spurious} > \text{true}$  triggers the pre-committed non-inert-placebo interpretation and reinforces whatever F2 showed.
- **E1:** SYSL-era analogue of theta1 from step 22, 95% CI, descriptive replication only; never pooled across the vintage break.
- **E3:** BRUHON fee analogues of theta2a/theta2\_diff beyond the gated secondary; heterogeneity of theta1/theta2a by sex (KOENIMP), age band (ALDERIMP), and baseline-year wage decile (BREDT\_LOEN\_BELOEB); Holm within this family.
- **E4:** arm-comparative attrition panel — share of each frozen November c-1 cohort no longer attached to the event workplace by November of year c (IDAN year-c records; ANSXFREM as the forward-separation flag), by event class; explicitly labeled post-treatment-conditioned descriptives with a selection caveat; no test.
- No exploratory result can revise any confirmatory verdict or Stage-4 language. Every sentence stating an exploratory result contains the word “exploratory”.

### 4. ANALYSIS EXECUTION

Confirmatory tier first (Stage 4, frozen code, single run); exploratory tier after (Stage 5), segregated by heading. All estimation in R fixest via Rscript. Common estimator: outcome  $y_{iq}$  (winsorized quarterly GP contact count) on treated  $x$  event-quarter indicators for the 8 quarters  $Q1(c)..Q4(c+1)$ , matched-set fixed effects, cohort  $x$  calendar-quarter fixed effects, omitted baseline = the four quarters of year  $c-1$ ; observation weights: exposed 1, controls  $1/k_s$  within set; CR1 SEs clustered on event workplace (each control inherits its set’s event cluster). Min-cell suppression is applied to all outputs BEFORE Holm reporting so suppression can never be outcome-selective.

### 5. EXHIBITS

Every figure and table the report commits to. Disclosure rules binding on EVERY released object: no statistic based on fewer than 5 persons or fewer than 3 workplaces; event-count tables rounded to the nearest 10; weekly trajectory figures pool  $\geq 100$  events and  $\geq 5,000$  person-weeks per plotted point; BRUHON aggregates pass a 50% single-workplace dominance check; released objects are limited to model coefficients with SEs/CIs, balance tables, calibration/benchmark tables, figure aggregates, rounded event counts, the power/MDE tables (both Stage-0 rounds and Stage 2) with the demotion trail, and the analysis/power scripts — never microdata, workplace identifiers (ARBNR/CVRNR/LBNR), or person-level values; suppression applied BEFORE Holm reporting.

### Figures (matplotlib):

- **F1 — Quarterly event-study panel.** Treated x event-quarter coefficients with 95% CIs across Q1(c-2)..Q4(c+1) (placebo quarters, omitted-baseline marker, terminal-year and post-year quarters); panel  
(a) true arm, panel (b) spurious arm from the pooled model. Feeds: step-18 panels + 4.1/4.2 coefficient tables. Disclosure: coefficients only.
- **F2 — Weekly descriptive trajectory.** Pooled exposed-minus- matched-control weekly GP contact series around the closure-year boundary, with the GRAFT-2 week-0 step diagnostic annotated (step  $\leq 50\%$  of the total Q3(c)+Q4(c) excess; if exceeded, the figure carries the registered caution and quarter-level estimates are foregrounded). Feeds: step 15/17 weekly aggregates. Disclosure:  $\geq 100$  events and  $\geq 5,000$  person-weeks per point.
- **F3 — s(w) successor-share distribution (GRAFT-4).** Histogram of worker-flow successor shares across terminal events with the 0.30/0.70 thresholds and mid-zone mass; bimodality check exhibit. Feeds: step 7. Disclosure: bin counts rounded to 10; bins under min-cell suppressed.
- **F4 — OOH vs daytime relative excess (theta3).** The two relative-excess series over the peri-event window with the bootstrap CI of their gap; primary and YDLTID-augmented classifier side by side. Feeds: 4.3 outputs. Disclosure: coefficients only.
- **F5 — Matching balance (love plot).** SMDs before/after matching for all matching covariates and unmatched auxiliaries; 0.10 line marked. Feeds: step 13. Disclosure: aggregates only.
- **F6 — E1 SYSI-era replication (exploratory; skippable).** Quarterly SYSI-era analogue of F1(a) for cohorts 1992–2003, labeled exploratory, YDLANT-weighted proxy outcome. Feeds: step 22. Disclosure: as F1.

### Tables (booktabs LaTeX):

- **T1 — Event extraction and classification by cohort year.** Terminal events, true/spurious/ambiguous counts (rounded to 10), annual true-closure rate among private size-5+ workplaces, spurious/true ratio. Feeds: step 8.
- **T2 — Classification cross-validation.** Worker-flow classes vs IDAS-native flags (ENDNEDL/IDFREM/REGELFR/FRELFREM/ANDELUDS), full cross-tabulation + agreement rate; GRAFT-3 classification counts under 20/80 and 40/60. Feeds: step 8.
- **T3 — Cohort flow accounting.** Inclusion/exclusion counts per rule (tenure, sector, size, SIKGRUP, missing ANSAAR/pay, multi-cohort dedup), eligible exposed per cohort, match rate, ladder rung used, realized k distribution. Feeds: steps 9–13.
- **T4 — Matching balance.** Per-covariate SMDs (weighted  $1/k$ ), caliper compliance, realized mean k. Feeds: step 13.
- **T5 — Power and demotion trail.** Stage-0 power results (BOTH rounds: the voided round-1 run and the remediated re-run) and the Stage-2 blind realized-power/MDE table with binding  $N^*$ , realized sets, and any demotions with the fixed-point recomputation. Feeds: deposited script

- step 14. Published regardless of outcome.
- **T6 — Outcome-neutral check panel.** All eleven registered checks (coverage benchmark by year, linkage floor, closure-rate plausibility, classification agreement, GRAFT-4, balance/match rate, c-2 placebo, HONUGE lag calibration, floor/ceiling + seasonality, GRAFT-2 step, OOH-share benchmark) with thresholds and PASS/FAIL. Feeds: steps 8, 13, 16, 17, 19, 21 + F2/F3 diagnostics.
- **T7 — Confirmatory results.** theta1, theta2a, theta3: point estimates (contacts/person-quarter and % of baseline), 95% CIs, two-sided p, Holm-adjusted verdicts; the F2 gate row (theta2b, its one-sided 95% upper bound, trigger evaluation); the gated BRUHON fee secondary (if computed); c-2 placebo row; realized MDEs beside every null with the CI-ruled-out magnitudes. Feeds: 4.1–4.4.
- **T8 — Sensitivity suite.** GRAFT-3 thresholds; GRAFT-9 forward-window exclusion (with truncation note); un-winsorized; Poisson QML; death/emigration exclusion — one table, run once. Feeds: 4.4.
- **T9 — Exploratory tier.** H1\_trend (if gated in; otherwise the row states the gate was not met), H2\_diff with quarter-by-quarter differential profile, E3 fee and subgroup analyses (within-family Holm), E1 summary. Every row labeled exploratory. Feeds: 4.5.
- **T10 — Attrition and later-presence accounting.** E4 year-c separation shares by event class (ANSXFREM-based, post-treatment- conditioned label and selection caveat) and the GRAFT-6 later-presence rates with the 2pp trigger status. Feeds: steps 20 + 4.5 E4.

---

## Current source hypotheses (all fields)

```

- id: E4
 tier: exploratory
 statement_initial: null
 statement_source: no frozen bank statement - protocol-defined exploratory
 ↪ module (arm-comparative
 attrition accounting; Schwerdt sinking-ship channel)
 estimand: 'Arm-comparative attrition accounting panel (GRAFT 8): share of
 ↪ each frozen November
 c-1 cohort no longer attached to the event workplace by November of year
 ↪ c (IDAN year-c
 records; ANSXFREM as the forward-separation flag), reported by event
 ↪ class (true closure,
 spurious closure, stable control), explicitly labeled
 ↪ post-treatment-conditioned descriptives
 with a selection caveat. Quantifies the Schwerdt (2011) sinking-ship
 ↪ channel the frozen-cohort
 design guards against. Includes a registered post-treatment-conditioned
 ↪ exploratory split
 of the theta1 contrast by whether the exposed worker retains
 ↪ November-year-c attachment
 to the event workplace, bounding the realized-early-separation channel
 ↪ named in design.threats.'
 primary_quantity: Class-specific year-c separation shares with
 ↪ disclosure-compliant cell

```

sizes; descriptive only, no test; plus the theta1 split by  
 ↳ November-year-c attachment  
 (post-treatment-conditioned, selection caveat mandatory, exploratory  
 ↳ flagging convention  
 only).

## Current source interpretation rows

□

## Source family scoring and limitations

scoring\_rule: 'Fixed in advance per the collaboration's registered-report  
 ↳ conventions (trajectory  
 fields): at Stage 2 an outcome attaches to each confirmatory hypothesis's  
 ↳ frozen statement\_initial  
 as confirmed | partially\_confirmed | reversed, adjudicated strictly against  
 ↳ the pre-committed  
 interpretation table in this protocol. H1 scores confirmed only if  
 ↳ theta1's Holm-adjusted  
 two-sided test rejects with the point estimate in the predicted direction  
 ↳ AND the gated  
 exploratory H1\_trend contrast supports the frozen statement's growth rider  
 ↳ (beta\_Q4c -  
 beta\_Q1c > 0 with 95% CI excluding zero); it scores partially\_confirmed  
 ↳ when theta1 Holm-rejects  
 in the predicted direction but the growth rider is unsupported (flat  
 ↳ elevation - the pre-committed  
 reading given annual event dating) or the trend gate is not computed;  
 ↳ reversed if theta1  
 is Holm-significant in the negative direction. H2 scores confirmed only if  
 ↳ theta2a Holm-rejects  
 in the predicted direction AND falsification gate F2 does not fire AND the  
 ↳ theta2b sign  
 pattern is stable across the GRAFT-3 successor-share thresholds;  
 ↳ partially\_confirmed when  
 theta2a Holm-rejects but the comparative quantifier is untestable  
 ↳ (spurious-arm kill - verdict  
 capped by pre-commitment at 'partially consistent - comparative quantifier  
 ↳ untestable');  
 reversed if theta2a is Holm-significant in the negative direction OR gate  
 ↳ F2 fires (the  
 frozen statement's comparative reading is disconfirmed by a non-inert  
 ↳ placebo regardless  
 of theta2a - registered as disconfirming, never reframed). H3 scores  
 ↳ confirmed only if theta3's  
 Holm-adjusted test rejects in the predicted direction (including the  
 ↳ total-down/OOH-up pattern,  
 pre-registered as scoreable support for the composition claim); reversed if  
 ↳ Holm-significant  
 in the negative direction. A Holm-nonsignificant confirmatory result  
 ↳ attaches no confirmation:  
 it is reported as a bounded informative null - the 95% CI converted into  
 ↳ the excess magnitudes  
 the data rule out against realized MDEs - never 'accepted'. A member  
 ↳ demoted by the Stage-2  
 realized-N gate or killed by its registered criterion attaches no score;  
 ↳ the demotion/kill

trail is disclosed. Significant effects in the non-predicted direction are  
↳ disconfirming,  
never supportive. Exploratory quantities (H1\_trend, H2\_diff, E1, E3, E4,  
↳ and any demoted  
member) attach outcomes only under the exploratory heading and are never  
↳ scored as preregistered-confirmed.'

assumptions:

- SSSY (2005-2025) and SYSI (1990-2005) capture the universe of publicly  
↳ reimbursed primary-care  
billing for insured residents; absence of rows means no billed care, not  
↳ missingness
- KONTAKT is a row-level contact marker whose sum yields contact counts  
↳ (in-environment validation  
with a frozen fallback definition if not)
- ENDNEDL/IDFREM/IDTILB/REGELFR/FRELFREM semantics follow Statistics  
↳ Denmark's IDA workplace-continuity  
documentation, cited directly (no DOI exists for these conventions, per  
↳ boundary directive);  
the primary classification rests on the worker-flow measure s(w) from IDAN,  
↳ so misreading  
an IDAS flag degrades only the cross-check
- November-status annual attachment measures workplace membership; intra-year  
↳ attachment changes  
are unobservable ex ante and the frozen c-1 anchor is the design's answer  
↳ to that - with  
the corollary, named as a design threat, that early year-c separations feed  
↳ realized post-job-loss  
billing into the pre-boundary window under intention-to-follow, a channel  
↳ the E4 separation  
shares and the registered theta1 attachment split bound but cannot remove
- ALDERIMP/KOENIMP (DST-imputed age/sex on SSSY rows) are accurate enough for  
↳ stratified matching;  
persons absent from SSSY in c-3..c-1 are handled by explicit unknown strata
- HONUGE settlement lags are predominantly <= 4 weeks (calibrated on  
↳ BEHANDLINGSDATO 2021+;  
guard-band contingency if not), and lag structure does not differ between  
↳ exposed and matched  
controls within the same calendar quarters
- Death and emigration are unobservable in scope; their post-event incidence  
↳ is assumed small;  
the deflating direction is conservative for the pre-boundary claim and  
↳ attenuating for theta2a,  
with the GRAFT-6 trigger bounding the post-boundary side
- Nominal-fee comparisons within cohort x calendar-quarter cells require no  
↳ deflator
- Danish institutional stability of GP billing categories (SPEC2 70-84, 89)  
↳ across 2007-2024  
at the 2-digit level
- 'PUBLISHED-BASIS SIZING ASSUMPTIONS (new, each with a registered failure  
↳ path): (i) Danish  
establishment closure rates among private size-5+ workplaces are in the  
↳ published Erhvervsdemografi-consistent  
1.5-3% band, supporting 150,000-300,000 exposed workers over 15 cohorts -  
↳ failure path:  
the pre-outcome realized-N gate demotes rather than stretches; (ii) the  
↳ Danish OOH share  
of GP contact volume is in the published 8-12% band (Christensen & Olesen  
↳ 1998; LV-KOS 2011  
studies), simulated at the conservative 8% - failure path: the  
↳ published-range benchmark

- check plus realized-share re-simulation, then demotion or kill; (iii) the
  - ↪ control pool supports
  - 1:3 matching at  $\geq 90\%$  match rate - failure path: coarsening ladder,
    - ↪ realized-k accounting
    - in the blind re-simulation, then the gate'
- theta2a's 4-quarter averaging gain over theta1 survives serial correlation
  - ↪ in post-year
  - shocks; a registered serial-correlation discount is built into the Stage-0
    - ↪ re-simulation
    - rather than assumed away
- The persistent-vs-transient overdispersion split (persistent frailty var
  - ↪ 1.25, transient
  - log-var 0.25) is retained unchanged from the first-round simulation -
    - ↪ deliberately NOT revised
    - in the direction that would raise power, so the remediation cannot be
      - ↪ accused of tuning
      - the DGP
- failure\_modes:
  - 'Realized event or worker counts far below schema-derived expectations
    - ↪ (documentation-only
    - corpus; no realized Ns exist) - the design now bears this risk openly: the
      - ↪ Stage-2 realized-N
      - gate demotes underpowered members loudly and pre-outcome instead of letting
        - ↪ an underpowered
        - confirmatory test run; a sparse-closure Denmark kills or demotes cleanly
          - ↪ rather than quietly'
    - KONTAKT or ENDNEDL semantics differ from assumption - both carry frozen
      - ↪ fallback rules,
      - but level shifts in outcomes or event dating would force reported
        - ↪ recalibration and consume
        - credibility
    - HONUGE settlement lag materially longer than 4 weeks - the quarter-level
      - ↪ estimands survive,
      - but the weekly-resolution headline weakens to quarterly resolution in
        - ↪ practice (GRAFT-2
        - makes this explicit and pre-committed)
    - ANDELOPS direction-of-absorption semantics may differ from the idea's gloss
      - ↪ - classification
      - routes through IDAN worker flows, so this only weakens the IDAS
        - ↪ cross-check, not the H2
        - machinery
    - Spurious-closure events too heterogeneous for a clean placebo - a non-inert
      - ↪ placebo now
      - shows up as gate F2 firing (or exploratory H2\_diff reversal), which caps
        - ↪ H2's verdict by
        - pre-commitment instead of silently contaminating a differential test
    - Realized OOH share below the published range - caught by the
      - ↪ published-range benchmark check
      - before estimation; H3 demotes or dies pre-outcome instead of running
        - ↪ underpowered
    - Realized N lands below the full-expectation lower bound - the gate demotes
      - ↪ affected members
      - to exploratory pre-outcome; the study survives with a smaller confirmatory
        - ↪ family (or none,
        - in which case it reports as registered exploratory estimation with the
          - ↪ demotion trail published
          - the first-round outcome, now a contingency rather than the design)
    - Unobservable post-closure death/emigration deflates post-boundary billing -
      - ↪ direction stated,

- sensitivity pre-specified, GRAFT-6 2pp trigger forces best/worst-case
  - ↪ bounds; no in-scope
- fix exists
- Public/private sector misclassification via modal ARB\_SEKTORKODE for
  - ↪ multi-establishment
- firms - attenuates contrasts; share bounded in the balance tables; if
  - ↪ ARB\_SEKTORKODE itself
- proves absent in-environment (it is unconfirmed in the recorded
  - ↪ Forskningservice inventory),
- the frozen ARB\_HOVED\_BRA\_DB07 fallback in contingent\_decisions substitutes,
  - ↪ with counts
- reported
- Multi-job workers' primary-job assignment may misattribute exposure -
  - ↪ deterministic rule
- biases toward null; share reported
- 1:3 matching without replacement may exhaust thin strata - realized  $k < 3$ 
  - ↪ accounted honestly
- in the blind re-simulation rather than assumed; coarsening ladder
  - ↪ pre-committed
- The SYSI-era exploratory replication may be infeasible - bounded as
  - ↪ skippable Stage-5 work
- that cannot contaminate the confirmatory tier
- Readers may over-read gate F2 as proof of causal job-loss effects - the
  - ↪ associational ceiling,
- the 'consistent with' vocabulary cap, and the pre-committed F2-fired
  - ↪ interpretation guard
- the prose

## Source disclosure rules

environment: DST Forskningservice (no prior data access; gatekeeper

- ↪ certification per RR

secondary-data rules). All analysis occurs inside the environment; only

- ↪ aggregate outputs

leave. No microdata, workplace identifiers (ARBNR/CVRNR/LBNR), or

- ↪ person-level values are

released.

minimum\_cell\_size: 5

minimum\_cell\_rules: no released statistic based on fewer than 5 persons or

- ↪ fewer than 3 workplaces;

event-count tables rounded to the nearest 10; weekly trajectory figures

- ↪ pool  $\geq 100$  events

and  $\geq 5,000$  person-weeks per plotted point; fee (BRUHON) aggregates

- ↪ subject to a 50% single-workplace

dominance check; min-cell suppression applied BEFORE Holm reporting so

- ↪ suppression can never

be outcome-selective

outputs\_leaving\_environment:

- model coefficients with SEs/CIs
- matching balance tables and realized-k distribution
- calibration and benchmark tables (coverage, linkage, HONUGE lag, OOH share)
- figure aggregates (weekly trajectory series under the pooling rules)
- event counts rounded to the nearest 10
- Stage-0 power/MDE tables - BOTH rounds (failed and remediated) - and the
  - ↪ Stage-2 blind re-simulation
- output with the demotion trail
- the c-2 placebo, gate F2 outputs, GRAFT-2/4 diagnostics, GRAFT-6
  - ↪ later-presence rates, E4
- attrition panel, and all outcome-neutral check outputs (the full audit
  - ↪ trail exits under

the same rules)

- the analysis and power scripts

statement: 'All analysis occurs inside DST Forskningssservice; only aggregate  
 ↳ outputs leave.

Rules applied to every released object: (1) no statistic based on fewer  
 ↳ than 5 persons or  
 fewer than 3 workplaces; (2) event-count tables rounded to the nearest 10;  
 ↳ (3) weekly trajectory  
 figures pool  $\geq 100$  events and  $\geq 5,000$  person-weeks per plotted point; (4)  
 ↳ fee (BRUHON)  
 statistics released only as cell means/coefficients over min-N cells, with  
 ↳ a dominance check  
 that no single workplace contributes  $> 50\%$  of any released fee aggregate;  
 ↳ (5) released objects  
 are limited to: model coefficients with SEs/CIs, balance tables,  
 ↳ calibration/benchmark tables,  
 figure aggregates, event counts (rounded), the Stage-0 (both rounds) and  
 ↳ Stage-2 power/MDE  
 tables with the demotion trail, and the analysis/power scripts - never  
 ↳ microdata, workplace  
 identifiers (ARBNR/CVRNR/LBNR), or person-level values; (6) min-cell  
 ↳ suppression is applied  
 before Holm reporting so that suppression can never be outcome-selective;  
 ↳ (7) the c-2 placebo,  
 gate F2 outputs, the GRAFT-2/4 diagnostics, the GRAFT-6 later-presence  
 ↳ rates, the E4 attrition  
 panel, and all outcome-neutral check outputs are released under the same  
 ↳ rules, ensuring  
 the full audit trail exits the environment.'

## Complete registration fields

```
id: atom-000009
topic: Within-year separation
atom_result: prereg
atom_strength: null
atom_shape: null
evidence_paths: []
concepts: []
literature: []
bonds: []
preregistration:
 version: 1
 mode: adaptive
 question: How much within-year separation and later-presence selection
 ↳ limits
 the pre-boundary interpretation?
 rationale: Source-existing coherent scientific question; its related
 ↳ outcomes,
 riders and sensitivity specifications remain internal.
 assumption: |-
```

Associational - matched event-study contrasts, a contrastive  
 ↪ falsification gate, and a composition contrast; descriptive for  
 ↪ trajectory panels and the bounded priority claim. Binding vocabulary  
 ↪ rules: 'is associated with' / 'consistent with' / 'suggests' only;  
 ↪ 'effect', 'impact', 'causes' prohibited throughout; anticipation  
 ↪ always worded as 'pre-boundary excess' relative to the annually dated  
 ↪ closure-year boundary, never as announcement-dated anticipation (S8);  
 ↪ no phenomenon-discovery or Danish-first framing (S1-S3); H2 novelty  
 ↪ confined to the placebo's application to health billing (S7). H2  
 ↪ verdict semantics are fixed in advance: 'consistent with H2' requires  
 ↪ BOTH the confirmatory core (theta2a Holm-significant, predicted  
 ↪ direction) AND a passing falsification gate F2; the strict  
 ↪ comparative magnitude is only ever reported as an exploratory  
 ↪ estimate with its CI. Exploratory results (H1\_trend, H2\_diff, E1, E3,  
 ↪ E4, and any gate-demoted member) carry the word 'exploratory' in the  
 ↪ heading and in every sentence stating them. The prose label never  
 ↪ exceeds the evidence label.  
 Compatibility ratings below are neutral import metadata, not a new  
 ↪ scientific evaluation or empirical scores.  
 target\_strength: descriptive  
 target\_shape: comparison  
 ratings:  
   importance: 3  
   novelty: 3  
   feasibility: 3  
   data\_advantage: 3  
   informativeness: 0.5  
 tables:  
 - IDAN  
 - IDAS  
 - SSSY  
 activation: state(atom-000003) == "event\_study\_ready" AND done(atom-000004)  
   AND done(atom-000005) AND done(atom-000006)  
 verdicts:  
 - key: exploratory\_pattern  
   region: The source-defined module yields interpretable  
     ↪ estimates/descriptives  
       after its own eligibility checks; retain unavailable components  
       ↪ explicitly.  
   says: Report the complete source exploratory pattern without confirmatory  
     or causal promotion.  
 - key: source\_skip  
   region: The source-defined optional module/arm is unavailable or its  
     ↪ registered  
       skip criterion is met.  
   says: Record the source skip and its reason; never invent a negative or  
     null finding.  
 - key: boundary  
   region: Required source-defined inputs or scientific definitions are  
     ↪ demonstrably  
       unavailable; retain available components and the exact unresolved scope.  
   says: An evidenced boundary, not a statistical null. A missing  
     ↪ implementation  
       or timeout alone is operational no\_result, not this state.  
 - key: otherwise  
   region: otherwise  
   says: The registered judgment is unresolved; report all completed  
     ↪ components  
       without inventing support, equivalence or absent results.  
 protocol: |

```

How much within-year separation and later-presence selection limits the
↳ pre-boundary interpretation?

Source fidelity and implementation order

This is a format conversion of DS1.1, not a new study or independent
↳ acceptance.
Source audit: minor_revisions. No observed Danish outcomes are supplied.
The source execution plan explicitly states that final prereg.yaml governs
disagreements. Exact current YAML selections below therefore govern older
execution-plan excerpts; these excerpts are retained as implementation
↳ detail,
not a competing registration. See docs/provenance/source-discrepancies.md.
The complete shared data/validity contracts are embedded in Atoms 1--3;
↳ read
those native protocols when executing a downstream Atom, even without
↳ docs/.
Do not substitute a scientific default for an unresolved source
↳ contradiction.

Preserve source stages 0 -> 1 -> 2 -> 3 -> 4 -> 5 inside the
↳ implementation.
Source Stage-0 simulations are historical design evidence, not study
↳ outcomes;
do not silently redo or improve them. Gate 14 must be evaluated from
↳ baseline
data before post-baseline outcomes. All surviving Stage-4 quantities form
↳ ONE
Holm family. The three question-level Atoms may share the same frozen
↳ family
fit and tidy results; do not apply Holm separately per Atom or adjudicate
↳ a
member before the complete surviving family's p-values exist. Cache/reuse
identical governed intermediate results with provenance when available; a
direct-dispatch call must obtain or compute its exact source-defined
↳ inputs,
not pretend its predecessors succeeded. Stage-5 internal riders remain
exploratory, after the Stage-4 family is frozen, even when housed in
↳ H1/H2.

Reference files under docs/execution/ are the ORIGINAL SKELETONS, not
↳ tested
turnkey implementations. Adapt paths and complete engineering in a
↳ governed
workspace; preserve source scientific rules and disclose unresolved
↳ choices.
Write a result record with component estimates, uncertainty, family
↳ eligibility,
source interpretation row, diagnostics and boundary/skip dispositions.
↳ Released
outputs follow source disclosure rules; confidential rows stay governed.
Local native states summarize the source judgment; neither a state name
↳ nor
source power forecasts prove equivalence or causal identification. A
↳ bounded
claim requires its actual interval and registered rule. Retain unresolved
↳ source
wording conflicts; do not convert mere nonsignificance into scientific
↳ null.

```

## Frozen scope and vocabulary

```
```yaml
claim_class_ceiling: associational
statement: 'Associational - matched event-study contrasts, a contrastive
↪ falsification gate,
  and a composition contrast; descriptive for trajectory panels and the
  ↪ bounded priority claim.
Binding vocabulary rules: 'is associated with' / 'consistent with'
↪ / 'suggests' only;
'effect', 'impact', 'causes' prohibited throughout; anticipation
↪ always worded as
'pre-boundary excess' relative to the annually dated closure-year
↪ boundary, never as announcement-dated
anticipation (S8); no phenomenon-discovery or Danish-first framing
↪ (S1-S3); H2 novelty confined
to the placebo's application to health billing (S7). H2 verdict
↪ semantics are fixed in
advance: 'consistent with H2' requires BOTH the confirmatory core
↪ (theta2a Holm-significant,
predicted direction) AND a passing falsification gate F2; the strict
↪ comparative magnitude
is only ever reported as an exploratory estimate with its CI.
↪ Exploratory results (H1_trend,
H2_diff, E1, E3, E4, and any gate-demoted member) carry the word
↪ 'exploratory' in the
heading and in every sentence stating them. The prose label never
↪ exceeds the evidence label.'
frozen_parameters:
seed: 20260831
gp_code_list: daytime GP = SPEC2 in {70..80}
ooh_code_list: out-of-hours GP = SPEC2 in {81, 82, 83, 84, 89}
ooh_robustness_classifier: adds YDLTID in {2, 3, 4, 5, 7, 8, 9}
spec2_evening_gp_disclosure: 'Registered choice, disclosed: the parsed
↪ value labels mark
  SPEC2 71 ("General practitioner, consultation, evening etc."), 73
  ↪ ("General practitioner,
  telephone consultation, evening etc."), and 75 ("General
  ↪ practitioner, home visit, evening
  etc.") as explicitly off-schedule GP services; they nonetheless
  ↪ remain on the daytime
  side of the PRIMARY classifier because H3's out-of-hours construct
  ↪ and its published
  power basis are keyed to the laegevagtt out-of-hours service (SPEC2
  ↪ 81-84, 89), the quantity
  the published Danish 8-12% share describes. Attenuation direction,
  ↪ both readings: counting
  evening GP as daytime attenuates theta3 toward zero - conservative
  ↪ against the positive
  H3 conclusion, anti-conservative for the null-bounding
  ↪ "compositionally unremarkable"
  reading, whose Stage-2 statement therefore carries a mandatory
  ↪ caveat. In-environment
  validation must verify that the YDLTID-augmented robustness
  ↪ classifier captures SPEC2
  71/73/75 rows (their evening time codes 2-5); if YDLTID proves
  ↪ unpopulated on those rows,
  the robustness classifier - and only the robustness classifier - is
  ↪ amended mechanically
```

```

    and with disclosure to include SPEC2 71/73/75 directly. The SPEC2
    ↪ 71/73/75 volume share
    is reported, and both classifiers appear side by side in Table T7.'
    successor_share_thresholds: 'true closure: ENDNEDL final closure AND
    ↪ worker-flow max-successor
    share s(w) < 0.30; spurious closure: forward identity via
    ↪ IDFREEM/REGELFR/FRELFREM AND
    s(w) >= 0.70; ambiguous mid-zone excluded; GRAFT-3 sensitivities
    ↪ 0.20/0.80 and 0.40/0.60'
    quarter_mapping: 'ISO week of HONUGE: 1-13 / 14-26 / 27-39 / 40-52; ISO
    ↪ week 53 to Q4'
    winsorization: person-quarter outcomes at the 99.9th percentile within
    ↪ era x calendar quarter
    matching: 1:3 nearest-neighbor Mahalanobis within exact strata, without
    ↪ replacement across
    sets, target k = 3, minimum k = 1; calipers |dif baseline contacts|
    ↪ <= 3 and |dif log(1+BRUHON)|
    <= 0.5; exposed processed in seeded random order (seed 20260831)
    software: R >= 4.3, fixest::feols (OLS) and fixest::feols (Poisson QML
    ↪ robustness); versions
    frozen in the registered script
    note: Code lists and thresholds frozen before any data access; code
    ↪ labels quoted verbatim
    from parsed DST value labels and re-validated in-environment before
    ↪ estimation (frozen
    ↪ fallback rules in contingent_decisions).
...
### 4.5 Exploratory tier (Stage 5; heading-segregated; Holm within family
↪ as severity aid)

- **H1_trend (gated):** beta_Q4c - beta_Q1c from the true-arm model, 95%
  CI, CR1 clusters - computed ONLY if theta1 Holm-rejected positively;
  otherwise the quarterly trajectory is shown descriptively with no
  contrast statistic and no timing-shape language in the Summary.
- **H2_diff:** theta2_diff = mean_k[beta_true,k - beta_spurious,k] over
  the post-year quarters from the pooled model, 95% CI, plus the
  quarter-by-quarter differential profile (the "more persistent" rider).
  A flagged estimate with spurious > true triggers the pre-committed
  non-inert-placebo interpretation and reinforces whatever F2 showed.
- **E1:** SYSI-era analogue of theta1 from step 22, 95% CI, descriptive
  replication only; never pooled across the vintage break.
- **E3:** BRUHON fee analogues of theta2a/theta2_diff beyond the gated
  secondary; heterogeneity of theta1/theta2a by sex (KOENIMP), age band
  (ALDERIMP), and baseline-year wage decile (BREDT_LOEN_BELOEB); Holm
  within this family.
- **E4:** arm-comparative attrition panel - share of each frozen
  November c-1 cohort no longer attached to the event workplace by
  November of year c (IDAN year-c records; ANSXFREM as the
  forward-separation flag), by event class; explicitly labeled
  post-treatment-conditioned descriptives with a selection caveat; no
  test.
- No exploratory result can revise any confirmatory verdict or Stage-4
  language. Every sentence stating an exploratory result contains the
  word "exploratory".

```

4. ANALYSIS EXECUTION

Confirmatory tier first (Stage 4, frozen code, single run); exploratory tier after (Stage 5), segregated by heading. All estimation in R fixest via Rscript. Common estimator: outcome y_iq (winsorized quarterly GP

contact count) on treated x event-quarter indicators for the 8 quarters Q1(c)..Q4(c+1), matched-set fixed effects, cohort x calendar-quarter fixed effects, omitted baseline = the four quarters of year c-1; observation weights: exposed 1, controls 1/k_s within set; CR1 SEs clustered on event workplace (each control inherits its set's event cluster). Min-cell suppression is applied to all outputs BEFORE Holm reporting so suppression can never be outcome-selective.

5. EXHIBITS

Every figure and table the report commits to. Disclosure rules binding on EVERY released object: no statistic based on fewer than 5 persons or fewer than 3 workplaces; event-count tables rounded to the nearest 10; weekly trajectory figures pool ≥ 100 events and $\geq 5,000$ person-weeks per plotted point; BRUHON aggregates pass a 50% single-workplace dominance check; released objects are limited to model coefficients with SEs/CIs, balance tables, calibration/benchmark tables, figure aggregates, rounded event counts, the power/MDE tables (both Stage-0 rounds and Stage 2) with the demotion trail, and the analysis/power scripts - never microdata, workplace identifiers (ARBNR/CVRNR/LBNR), or person-level values; suppression applied BEFORE Holm reporting.

Figures (matplotlib):

- **F1 - Quarterly event-study panel.** Treated x event-quarter coefficients with 95% CIs across Q1(c-2)..Q4(c+1) (placebo quarters, omitted-baseline marker, terminal-year and post-year quarters); panel (a) true arm, panel (b) spurious arm from the pooled model. Feeds: step-18 panels + 4.1/4.2 coefficient tables. Disclosure: coefficients only.
- **F2 - Weekly descriptive trajectory.** Pooled exposed-minus-matched-control weekly GP contact series around the closure-year boundary, with the GRAFT-2 week-0 step diagnostic annotated (step $\leq 50\%$ of the total Q3(c)+Q4(c) excess; if exceeded, the figure carries the registered caution and quarter-level estimates are foregrounded). Feeds: step 15/17 weekly aggregates. Disclosure: ≥ 100 events and $\geq 5,000$ person-weeks per point.
- **F3 - s(w) successor-share distribution (GRAFT-4).** Histogram of worker-flow successor shares across terminal events with the 0.30/0.70 thresholds and mid-zone mass; bimodality check exhibit. Feeds: step 7. Disclosure: bin counts rounded to 10; bins under min-cell suppressed.
- **F4 - OOH vs daytime relative excess (theta3).** The two relative-excess series over the peri-event window with the bootstrap CI of their gap; primary and YDLTID-augmented classifier side by side. Feeds: 4.3 outputs. Disclosure: coefficients only.
- **F5 - Matching balance (love plot).** SMDs before/after matching for all matching covariates and unmatched auxiliaries; 0.10 line marked. Feeds: step 13. Disclosure: aggregates only.
- **F6 - E1 SYSI-era replication (exploratory; skippable).** Quarterly SYSI-era analogue of F1(a) for cohorts 1992-2003, labeled exploratory, YDLANT-weighted proxy outcome. Feeds: step 22. Disclosure: as F1.

Tables (booktabs LaTeX):

- **T1 - Event extraction and classification by cohort year.** Terminal events, true/spurious/ambiguous counts (rounded to 10), annual true-closure rate among private size-5+ workplaces, spurious/true ratio. Feeds: step 8.
- **T2 - Classification cross-validation.** Worker-flow classes vs IDAS-native flags (ENDNEDL/IDFREM/REGELFR/FRELFREM/ANDELUDS), full

- cross-tabulation + agreement rate; GRAFT-3 classification counts under 20/80 and 40/60. Feeds: step 8.
- **T3 - Cohort flow accounting.** Inclusion/exclusion counts per rule (tenure, sector, size, SIKGRUP, missing ANSAAR/pay, multi-cohort dedup), eligible exposed per cohort, match rate, ladder rung used, realized k distribution. Feeds: steps 9-13.
 - **T4 - Matching balance.** Per-covariate SMDs (weighted 1/k), caliper compliance, realized mean k. Feeds: step 13.
 - **T5 - Power and demotion trail.** Stage-0 power results (BOTH rounds: the voided round-1 run and the remediated re-run) and the Stage-2 blind realized-power/MDE table with binding N*, realized sets, and any demotions with the fixed-point recomputation. Feeds: deposited script + step 14. Published regardless of outcome.
 - **T6 - Outcome-neutral check panel.** All eleven registered checks (coverage benchmark by year, linkage floor, closure-rate plausibility, classification agreement, GRAFT-4, balance/match rate, c-2 placebo, HONUGE lag calibration, floor/ceiling + seasonality, GRAFT-2 step, OOH-share benchmark) with thresholds and PASS/FAIL. Feeds: steps 8, 13, 16, 17, 19, 21 + F2/F3 diagnostics.
 - **T7 - Confirmatory results.** theta1, theta2a, theta3: point estimates (contacts/person-quarter and % of baseline), 95% CIs, two-sided p, Holm-adjusted verdicts; the F2 gate row (theta2b, its one-sided 95% upper bound, trigger evaluation); the gated BRUHON fee secondary (if computed); c-2 placebo row; realized MDEs beside every null with the CI-ruled-out magnitudes. Feeds: 4.1-4.4.
 - **T8 - Sensitivity suite.** GRAFT-3 thresholds; GRAFT-9 forward-window exclusion (with truncation note); un-winsorized; Poisson QML; death/emigration exclusion - one table, run once. Feeds: 4.4.
 - **T9 - Exploratory tier.** H1_trend (if gated in; otherwise the row states the gate was not met), H2_diff with quarter-by-quarter differential profile, E3 fee and subgroup analyses (within-family Holm), E1 summary. Every row labeled exploratory. Feeds: 4.5.
 - **T10 - Attrition and later-presence accounting.** E4 year-c separation shares by event class (ANSXFREM-based, post-treatment-conditioned label and selection caveat) and the GRAFT-6 later-presence rates with the 2pp trigger status. Feeds: steps 20 + 4.5 E4.

Current source hypotheses (all fields)

```

```yaml
- id: E4
 tier: exploratory
 statement_initial: null
 statement_source: no frozen bank statement - protocol-defined
 ↳ exploratory module (arm-comparative
 attrition accounting; Schwerdt sinking-ship channel)
 estimand: 'Arm-comparative attrition accounting panel (GRAFT 8): share
 ↳ of each frozen November
 c-1 cohort no longer attached to the event workplace by November of
 ↳ year c (IDAN year-c
 records; ANSXFREM as the forward-separation flag), reported by event
 ↳ class (true closure,
 spurious closure, stable control), explicitly labeled
 ↳ post-treatment-conditioned descriptives
 with a selection caveat. Quantifies the Schwerdt (2011) sinking-ship
 ↳ channel the frozen-cohort

```

```

 design guards against. Includes a registered
 ↪ post-treatment-conditioned exploratory split
 of the theta1 contrast by whether the exposed worker retains
 ↪ November-year-c attachment
 to the event workplace, bounding the realized-early-separation
 ↪ channel named in design.threats.'
primary_quantity: Class-specific year-c separation shares with
 ↪ disclosure-compliant cell
 sizes; descriptive only, no test; plus the theta1 split by
 ↪ November-year-c attachment
 (post-treatment-conditioned, selection caveat mandatory, exploratory
 ↪ flagging convention
 only).
...

Current source interpretation rows

```yaml
[]
```

Source family scoring and limitations

```yaml
scoring_rule: 'Fixed in advance per the collaboration''s
    ↪ registered-report conventions (trajectory
    fields): at Stage 2 an outcome attaches to each confirmatory
    ↪ hypothesis''s frozen statement_initial
    as confirmed | partially_confirmed | reversed, adjudicated strictly
    ↪ against the pre-committed
    interpretation table in this protocol. H1 scores confirmed only if
    ↪ theta1''s Holm-adjusted
    two-sided test rejects with the point estimate in the predicted
    ↪ direction AND the gated
    exploratory H1_trend contrast supports the frozen statement''s growth
    ↪ rider (beta_Q4c -
    beta_Q1c > 0 with 95% CI excluding zero); it scores partially_confirmed
    ↪ when theta1 Holm-rejects
    in the predicted direction but the growth rider is unsupported (flat
    ↪ elevation - the pre-committed
    reading given annual event dating) or the trend gate is not computed;
    ↪ reversed if theta1
    is Holm-significant in the negative direction. H2 scores confirmed only
    ↪ if theta2a Holm-rejects
    in the predicted direction AND falsification gate F2 does not fire AND
    ↪ the theta2b sign
    pattern is stable across the GRAFT-3 successor-share thresholds;
    ↪ partially_confirmed when
    theta2a Holm-rejects but the comparative quantifier is untestable
    ↪ (spurious-arm kill - verdict
    capped by pre-commitment at ''partially consistent - comparative
    ↪ quantifier untestable'');
    reversed if theta2a is Holm-significant in the negative direction OR
    ↪ gate F2 fires (the
    frozen statement''s comparative reading is disconfirmed by a non-inert
    ↪ placebo regardless
    of theta2a - registered as disconfirming, never reframed). H3 scores
    ↪ confirmed only if theta3''s
    Holm-adjusted test rejects in the predicted direction (including the
    ↪ total-down/00H-up pattern,

```

pre-registered as scoreable support for the composition claim);
 ↳ reversed if Holm-significant
 in the negative direction. A Holm-nonsignificant confirmatory result
 ↳ attaches no confirmation:
 it is reported as a bounded informative null - the 95% CI converted
 ↳ into the excess magnitudes
 the data rule out against realized MDEs - never 'accepted'. A member
 ↳ demoted by the Stage-2
 realized-N gate or killed by its registered criterion attaches no
 ↳ score; the demotion/kill
 trail is disclosed. Significant effects in the non-predicted direction
 ↳ are disconfirming,
 never supportive. Exploratory quantities (H1_trend, H2_diff, E1, E3,
 ↳ E4, and any demoted
 member) attach outcomes only under the exploratory heading and are
 ↳ never scored as preregistered-confirmed.'

assumptions:

- SSSY (2005-2025) and SYSI (1990-2005) capture the universe of publicly
 ↳ reimbursed primary-care
 billing for insured residents; absence of rows means no billed care,
 ↳ not missingness
- KONTAKT is a row-level contact marker whose sum yields contact counts
 ↳ (in-environment validation
 with a frozen fallback definition if not)
- ENDNEDL/IDFREM/IDTILE/REGELFR/FRELFREM semantics follow Statistics
 ↳ Denmark's IDA workplace-continuity
 documentation, cited directly (no DOI exists for these conventions, per
 ↳ boundary directive);
 the primary classification rests on the worker-flow measure s(w) from
 ↳ IDAN, so misreading
 an IDAS flag degrades only the cross-check
- November-status annual attachment measures workplace membership;
 ↳ intra-year attachment changes
 are unobservable ex ante and the frozen c-1 anchor is the design's
 ↳ answer to that - with
 the corollary, named as a design threat, that early year-c separations
 ↳ feed realized post-job-loss
 billing into the pre-boundary window under intention-to-follow, a
 ↳ channel the E4 separation
 shares and the registered theta1 attachment split bound but cannot
 ↳ remove
- ALDERIMP/KOENIMP (DST-imputed age/sex on SSSY rows) are accurate enough
 ↳ for stratified matching;
 persons absent from SSSY in c-3..c-1 are handled by explicit unknown
 ↳ strata
- HONUUGE settlement lags are predominantly <= 4 weeks (calibrated on
 ↳ BEHANDLINGSDATO 2021+;
 guard-band contingency if not), and lag structure does not differ
 ↳ between exposed and matched
 controls within the same calendar quarters
- Death and emigration are unobservable in scope; their post-event
 ↳ incidence is assumed small;
 the deflating direction is conservative for the pre-boundary claim and
 ↳ attenuating for theta2a,
 with the GRAFT-6 trigger bounding the post-boundary side
- Nominal-fee comparisons within cohort x calendar-quarter cells require
 ↳ no deflator
- Danish institutional stability of GP billing categories (SPEC2 70-84,
 ↳ 89) across 2007-2024
 at the 2-digit level

- 'PUBLISHED-BASIS SIZING ASSUMPTIONS (new, each with a registered failure path): (i) Danish establishment closure rates among private size-5+ workplaces are in the published Erhvervsdemografi-consistent 1.5-3% band, supporting 150,000-300,000 exposed workers over 15 cohorts
 - - failure path: the pre-outcome realized-N gate demotes rather than stretches; (ii) the Danish OOH share of GP contact volume is in the published 8-12% band (Christensen & Olesen 1998; LV-KOS 2011 studies), simulated at the conservative 8% - failure path: the published-range benchmark check plus realized-share re-simulation, then demotion or kill; (iii) the control pool supports 1:3 matching at $\geq 90\%$ match rate - failure path: coarsening ladder, realized-k accounting in the blind re-simulation, then the gate'
 - theta2a's 4-quarter averaging gain over theta1 survives serial correlation in post-year shocks; a registered serial-correlation discount is built into the Stage-0 re-simulation rather than assumed away
 - The persistent-vs-transient overdispersion split (persistent frailty var 1.25, transient log-var 0.25) is retained unchanged from the first-round simulation - deliberately NOT revised in the direction that would raise power, so the remediation cannot be accused of tuning the DGP
- failure_modes:
- 'Realized event or worker counts far below schema-derived expectations (documentation-only corpus; no realized Ns exist) - the design now bears this risk openly:
 - the Stage-2 realized-N gate demotes underpowered members loudly and pre-outcome instead of letting an underpowered confirmatory test run; a sparse-closure Denmark kills or demotes cleanly rather than quietly'
 - KONTAKT or ENDNEDL semantics differ from assumption - both carry frozen fallback rules, but level shifts in outcomes or event dating would force reported recalibration and consume credibility
 - HONUGE settlement lag materially longer than 4 weeks - the quarter-level estimands survive, but the weekly-resolution headline weakens to quarterly resolution in practice (GRAFT-2 makes this explicit and pre-committed)
 - ANDELOPS direction-of-absorption semantics may differ from the idea's gloss - classification routes through IDAN worker flows, so this only weakens the IDAS cross-check, not the H2 machinery
 - Spurious-closure events too heterogeneous for a clean placebo - a non-inert placebo now shows up as gate F2 firing (or exploratory H2_diff reversal), which caps H2's verdict by pre-commitment instead of silently contaminating a differential test
 - Realized OOH share below the published range - caught by the published-range benchmark check

- before estimation; H3 demotes or dies pre-outcome instead of running
 - ↪ underpowered
- Realized N lands below the full-expectation lower bound - the gate
 - ↪ demotes affected members
 - to exploratory pre-outcome; the study survives with a smaller
 - ↪ confirmatory family (or none,
 - in which case it reports as registered exploratory estimation with the
 - ↪ demotion trail published
 - the first-round outcome, now a contingency rather than the design)
- Unobservable post-closure death/emigration deflates post-boundary
 - ↪ billing - direction stated,
 - sensitivity pre-specified, GRAFT-6 2pp trigger forces best/worst-case
 - ↪ bounds; no in-scope
 - fix exists
- Public/private sector misclassification via modal ARB_SEKTORKODE for
 - ↪ multi-establishment
 - firms - attenuates contrasts; share bounded in the balance tables; if
 - ↪ ARB_SEKTORKODE itself
 - proves absent in-environment (it is unconfirmed in the recorded
 - ↪ Forskningservice inventory),
 - the frozen ARB_HOVED_BRA_DB07 fallback in contingent_decisions
 - ↪ substitutes, with counts
 - reported
- Multi-job workers' primary-job assignment may misattribute exposure -
 - ↪ deterministic rule
 - biases toward null; share reported
- 1:3 matching without replacement may exhaust thin strata - realized $k <$
 - ↪ 3 accounted honestly
 - in the blind re-simulation rather than assumed; coarsening ladder
 - ↪ pre-committed
- The SYSI-era exploratory replication may be infeasible - bounded as
 - ↪ skipable Stage-5 work
 - that cannot contaminate the confirmatory tier
- Readers may over-read gate F2 as proof of causal job-loss effects - the
 - ↪ associational ceiling,
 - the 'consistent with' vocabulary cap, and the pre-committed F2-fired
 - ↪ interpretation guard
 - the prose

Source disclosure rules

```
```yaml
environment: DST Forskningservice (no prior data access; gatekeeper
 ↪ certification per RR
 secondary-data rules). All analysis occurs inside the environment; only
 ↪ aggregate outputs
 leave. No microdata, workplace identifiers (ARBNR/CVRNR/LBNR), or
 ↪ person-level values are
 released.
minimum_cell_size: 5
minimum_cell_rules: no released statistic based on fewer than 5 persons
 ↪ or fewer than 3 workplaces;
 event-count tables rounded to the nearest 10; weekly trajectory figures
 ↪ pool $\geq$ 100 events
 and $\geq$ 5,000 person-weeks per plotted point; fee (BRUHON) aggregates
 ↪ subject to a 50% single-workplace
 dominance check; min-cell suppression applied BEFORE Holm reporting so
 ↪ suppression can never
 be outcome-selective
```

```

outputs_leaving_environment:
- model coefficients with SEs/CIs
- matching balance tables and realized-k distribution
- calibration and benchmark tables (coverage, linkage, HONUGE lag, OOH
 ↳ share)
- figure aggregates (weekly trajectory series under the pooling rules)
- event counts rounded to the nearest 10
- Stage-0 power/MDE tables - BOTH rounds (failed and remediated) - and
 ↳ the Stage-2 blind re-simulation
 output with the demotion trail
- the c-2 placebo, gate F2 outputs, GRAFT-2/4 diagnostics, GRAFT-6
 ↳ later-presence rates, E4
 attrition panel, and all outcome-neutral check outputs (the full audit
 ↳ trail exits under
 the same rules)
- the analysis and power scripts
statement: 'All analysis occurs inside DST Forskningssservice; only
 ↳ aggregate outputs leave.
 Rules applied to every released object: (1) no statistic based on fewer
 ↳ than 5 persons or
 fewer than 3 workplaces; (2) event-count tables rounded to the nearest
 ↳ 10; (3) weekly trajectory
 figures pool >= 100 events and >= 5,000 person-weeks per plotted point;
 ↳ (4) fee (BRUHON)
 statistics released only as cell means/coefficients over min-N cells,
 ↳ with a dominance check
 that no single workplace contributes > 50% of any released fee
 ↳ aggregate; (5) released objects
 are limited to: model coefficients with SEs/CIs, balance tables,
 ↳ calibration/benchmark tables,
 figure aggregates, event counts (rounded), the Stage-0 (both rounds)
 ↳ and Stage-2 power/MDE
 tables with the demotion trail, and the analysis/power scripts - never
 ↳ microdata, workplace
 identifiers (ARBNR/CVRNR/LBNR), or person-level values; (6) min-cell
 ↳ suppression is applied
 before Holm reporting so that suppression can never be
 ↳ outcome-selective; (7) the c-2 placebo,
 gate F2 outputs, the GRAFT-2/4 diagnostics, the GRAFT-6 later-presence
 ↳ rates, the E4 attrition
 panel, and all outcome-neutral check outputs are released under the
 ↳ same rules, ensuring
 the full audit trail exits the environment.'
...

```

## Pending registered experiment

No Danish study outcome has been observed for this bundle. See the complete registration above.