

OPERATING CASE STUDY / OPERATIONAL EXCELLENCE

Executive Performance Review

A closed bridge can still tell the wrong story

An anonymized project case study of a recurring customer-defect and loss review: reconciling reported movement, distinguishing competing explanations and assigning evidence-backed action across two operating groups.

PROJECT CONTEXT / DISTRIBUTED SERVICE OPERATIONS

BLUF / BOTTOM LINE UP FRONT

Reconcile the movement. Then route the investigation.

RECALCULATED FINDING

The source attributes a rate increase to defect volume. Recalculation shows that count growth and exposure contraction contribute roughly equally. A separate loss breakdown leaves 5.33% of its reported total unallocated.

CONTRIBUTION IN THE SOURCE RECORD

Reporting design and analytical stewardship.
Registry-attested ownership; provisional hoop-level support.

EVIDENCE STATUS

Evidence-reconciliation and operating-design case. Field execution, causal customer impact and realized savings remain unestablished.

BLUF: analytical inference. Source contribution: historical record. Client and internal identities anonymized.

EXECUTIVE ABSTRACT

Use the bridge to allocate investigation before assigning cause or credit

A recurring executive review connected defect movement to named actions. Reconstructing its arithmetic shows why a reconciled report still needs a separate causal and financial warrant.

CONTEXT | A weekly explanation became an operating mechanism

[HR] The register describes paired reports for two operating groups, a fixed bucket-by-bucket layout, named action owners, source checks and reusable production instructions. The principal is recorded as the mechanism owner; operational teams retained responsibility for the countermeasures.

OBJECTIVE | Decide what the reported movement authorizes

[PD] Separate three questions: does the arithmetic reconcile, does the evidence discriminate among causes, and has the accountable owner approved an appropriate response? Require a distinct answer to each before assigning performance credit.

APPROACH | Reconstruct the accounting and challenge the explanation

[RC] Decompose a rate change, retain bucket offsets and residuals, audit competing revisions and compare cost with event volume. [AI] Test the source's common-cause and maturity interpretations. [PD] Specify conditional action routing and an independent decision evaluation.

FINDINGS | Arithmetic closure does not close the investigation

[RC] Count growth and exposure contraction contribute roughly equally to the headline rate rise. The supplied loss hierarchy leaves 5.33% unallocated, including 9.08% of the reported decline. A cost-to-count comparison implies a 5.15% increase in mean recorded cost per event, conditional on common scope.

DECISION VALUE | Preserve actuals and assign the unresolved work

[PD] The review owner should commission population, revision and residual reconciliation; the data owner should test recording changes against operational alternatives. Hold unsupported causal claims. Keep reported actuals visible until an authorized, traceable restatement replaces them.

IMPLICATIONS | Credit stewardship within its evidence boundary

[AI] Registry descriptions and ownership fields provide provisional hoop-level support for reporting stewardship. They do not independently authenticate authorship, audience use, field execution or savings. Evaluate the bridge's decision value separately from the customer effects of actions it recommends.

S1: selected project register, M01–M16 and linked evidence. HR historical record; RC recalculated; AI analytical inference; PD proposed design. Original files were not independently authenticated in this case.

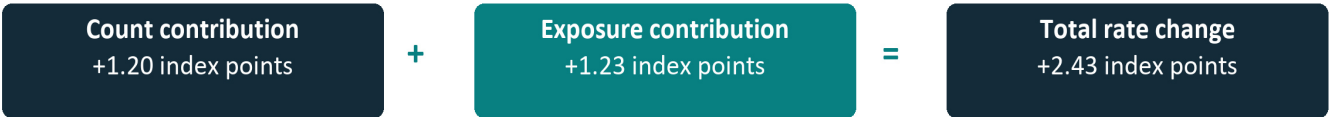
EXHIBIT 1 / RATE DECOMPOSITION

Count growth and exposure contraction contribute roughly equally to the rate rise

[RC] The reported change in defect volume is only one side of the rate identity. The source’s exposure decline contributes a similar amount.

THE RATE RISE HAS TWO SIMILAR ACCOUNTING COMPONENTS

Group A • prior rate = 100 • symmetric decomposition • rounded exposure input



Exposure contributes 50.61% of the decomposed rise.

Changing substitution order changes the allocation; averaging both orders is explicit.

This reconciles a rate identity. It does not identify an operational cause.

[HR] The source describes a higher defect count alongside lower shipped exposure, then characterizes the miss as a defect-volume issue. [RC] Reconstructing the rate from those inputs shows that both movements matter. The order-symmetric decomposition splits their interaction equally rather than assigning it to whichever factor is substituted first.

Accounting result	Decision consequence
[RC] Count and exposure contributions are similar.	[AI] The explanation cannot dismiss exposure while using a rate as its headline. Review the count and the denominator together.
[AI] The rate identity is descriptive.	[PD] Establish whether the recorded events belong to the same shipment cohort as the exposure before interpreting the rate as customer risk.

[AI] This finding changes the explanation, not the observed record. A denominator contraction can reflect calendar timing, scope or operating volume. It does not prove an improvement in customer experience, excuse a defect increase or identify what would have happened under a different operating decision.

EXECUTIVE ACTION

[PD] Require the analyst to show the numerator, exposure and comparison contract before recommending a cause-specific response. Keep the official actual distinct from any standardized sensitivity.

[RC] S1:M01/E04. Prior count, exposure and rate each indexed to 100. Symmetric decomposition uses count and inverse exposure; exposure change is rounded in the source. Das Gupta in references. No causal effect is estimated.

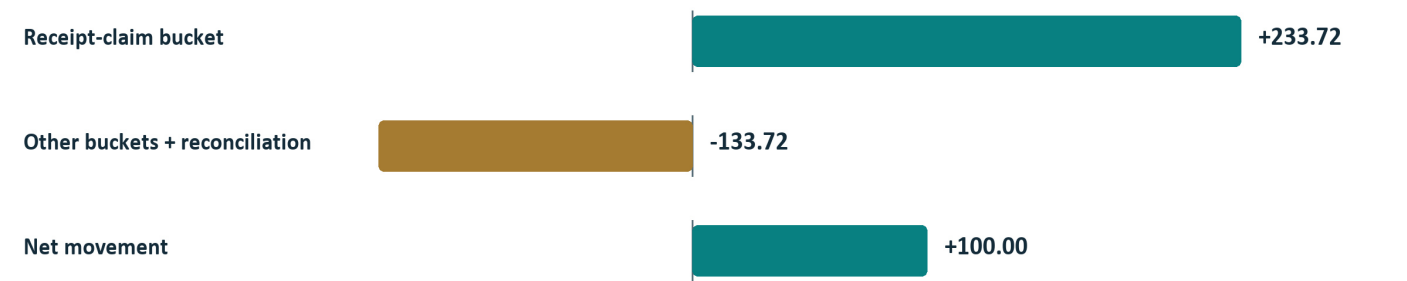
EXHIBIT 2 / OFFSETTING MOVEMENTS

Large opposing bucket changes sit behind the small net increase

[RC] The largest positive bucket contribution is 2.34 times the net headline increase because other buckets move in the opposite direction.

A LARGE DRIVER CAN EXCEED THE NET CHANGE

Group A • total weekly rate increase = 100 • accounting contribution scale



The source describes offsetting movements; the largest term is not the whole goal gap.

[AI] A bucket’s contribution can exceed the net movement because other buckets offset it. “The entire miss” confuses three different quantities: the weekly net increase, the gap to target and the causal contribution.

Bridge requirement	What it establishes
[PD] Common denominator and population	Bucket rate changes can be added only when their event sets, exposure, period and source vintage match the topline.
[PD] Exclusive partition; overlays separated	Each included event enters the additive hierarchy once. An overlapping label can be informative without becoming an additional bucket.
[RC] Residual remains visible	The displayed bucket changes do not close exactly to the topline. Preserve the remainder instead of adjusting a component to force agreement.

[AI] Rounded rates can generate a rate residual; they do not explain discrepancies between exact event counts. Define materiality from source precision and decision consequences. Approximate agreement remains distinct from exact closure.

MANAGEMENT IMPLICATION

[PD] Investigate the largest adverse bucket while retaining offsets and residuals. Neither the ranking nor the waterfall identifies a cause.

[RC] S1:M03/M05/M08; narrative Analysis_Performed. Figure uses one common protected rate scale. Contribution-to-net uses the reported net change, not the target gap. Exact residual arithmetic is retained privately.

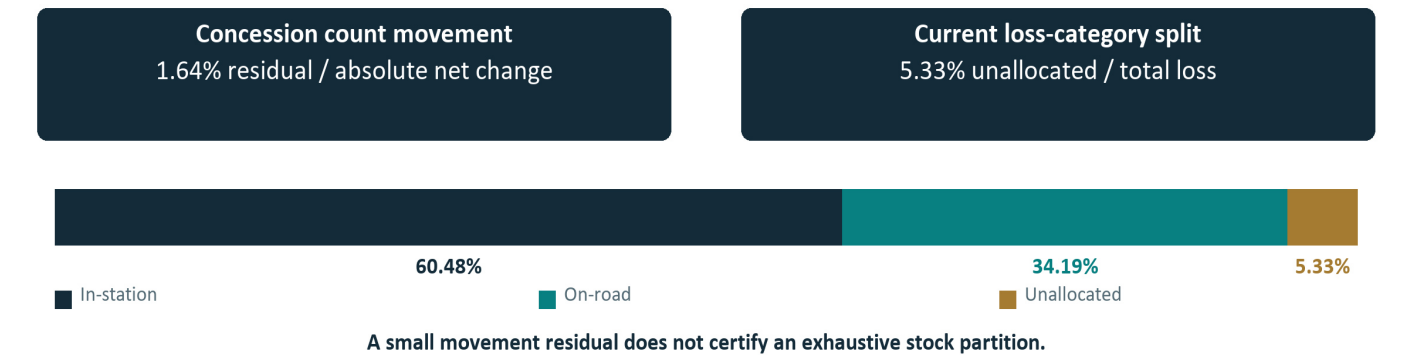
EXHIBIT 3 / PARTITION COMPLETENESS

The reported loss hierarchy leaves part of the total and its decline unassigned

[RC] The supplied in-station and on-road components leave 5.33% of the current loss total unallocated. That remainder contributes 9.08% of the reported decline.

CHECK THE MOVEMENT AND THE CURRENT TOTAL SEPARATELY

Different tests • different denominators • no pooling of these percentages



[AI] This does not identify an additional set of lost items beyond the reported total. It identifies an incomplete explanation of that total in the supplied extract. An omitted category, a population mismatch, a source revision or a transcription error could produce the remainder. The register does not resolve which explanation applies.

Two distinct checks	Required disposition
[RC] Level completeness	[PD] Reconcile each current-period component to the same total, with a complete category crosswalk and explicit residual.
[RC] Change completeness	[PD] Reconcile prior components on the same basis before allocating the total period change to operating owners.

[AI] A nearly reconciled concession bridge cannot certify this separate loss hierarchy. Likewise, a fall in reported loss alongside higher nonreceipt claims does not prove that the two event populations are disjoint. Overlap, recording lags and classification changes can produce different aggregate directions.

RELEASE RULE

[PD] Label the remainder “unallocated in the supplied extract” until the data owner resolves it. Do not rescale the named components to total 100 or assign their entire apparent improvement to a delivered action.

[RC] S1:M06/M08. Current residual share and residual contribution to the total decline use different denominators. Calculations assume common scope, clock and version; reconciliation remains open. Public figures expose no raw counts.

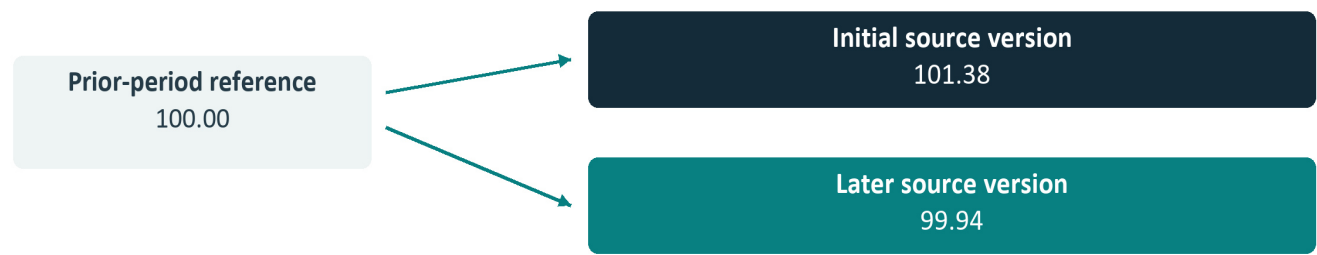
EXHIBIT 4 / REVISION LINEAGE

A downward revision requires more than an additions-only maturity explanation

[HR] Two versions of one reported period reverse its direction against the recorded prior. The later version is labeled matured; its revision mechanism is unverified.

ONE EVENT PERIOD CAN SUPPORT OPPOSITE WEEKLY SIGNS

Group A • prior-period rate = 100 • two source vintages of the same later week



A downward revision needs more than late additions at a fixed denominator.

Recover removals, reclassification, exposure versions and measurement clocks.

[AI] With a fixed eligible denominator and a cumulative numerator that only receives late additions, a later rate cannot fall. The observed downward revision therefore requires another process: removals, reversals, recoding, exposure revision, changed eligibility or a different time basis. The arithmetic rejects an additions-only explanation; it does not select the alternative.

Preserve both records	Answer different questions
[PD] As-issued version	What information was available when the executive decision was made? Retain the original and its extraction time.
[PD] Reconciled later version	What does the same cohort show after documented revisions? Retain revised event membership, exposure and reasons.

[AI] Two versions of one period are not two consecutive process observations. The direction of a rounded period comparison also does not establish a process shift. These versions demonstrate a revision-sensitive narrative, not a causal improvement or a validated maturation curve.

EXECUTIVE ACTION

[PD] Ask the data owner to reconcile the downward revision before using “matured” as a quality certificate. Publish the source fork and its decision consequence; do not silently overwrite the earlier decision record.

[HR/RC] S1:M09/D06/E11. Comparison prior =100 for the displayed versions. Source-labeled maturity is not independently verified. Croushore & Stark in references supports the distinction between decision-time and revised data.

EXHIBIT 5 / STANDARDIZATION AND SOURCE CONFLICT

A fixed-denominator sensitivity still needs a valid comparator

[RC] A calendar-distorted period’s rate looks less adverse under fixed exposure. Its reported growth also implies a prior value that differs from two unresolved source cuts.

A NORMALIZED READ STILL NEEDS THE RIGHT PRIOR

Holiday observation • source-linked alternatives • percentage change vs each prior

Prior implied by source Reported +46.90% Fixed-exposure read +7.14%	Alternate detail view Reported +45.34% Fixed-exposure read +6.00%	Alternate front view Reported +9.49% Fixed-exposure read -20.15%
---	---	--

The normalized direction changes under a different recorded comparator.

These views remain separate; no source choice is certified by the largest number.

[HR] The source chose to emphasize exposure compression after comparing the reported rate with a fixed-denominator calculation. [AI] That is useful accounting. Holding exposure fixed does not recreate a world without the calendar event, hold customer mix constant or establish the rate of mature harm in a matched shipment cohort.

Reconstruction branch	Interpretation
[RC] Source-implied prior	Reproduces the source’s reported growth and its normalized sensitivity within rounding. The implied comparator is not an independently recovered observation.
[HR/RC] Alternative recorded cuts	Changing the prior changes the normalized comparison and can change its sign. Their populations and filters remain unresolved; they cannot be treated as interchangeable baselines.

[AI] The right response is neither to average the cuts nor to choose the one producing the most reassuring story. Recover numerator membership, exposure, calendar window and extraction lineage, then identify the comparator that answers the declared question. Numerical sensitivity does not repair population incompatibility.

DECISION

[PD] Retain the reported actual. Use the normalized read as a labeled diagnostic sensitivity and hold its directional conclusion to a reconciled baseline. Do not claim a causal calendar effect or financial benefit from this calculation.

[HR/RC] S1:M10/D02–D03/E03/E07. Protected common scale in figure; alternative prior cuts are sensitivity branches, not equivalent observations. Source normalization is approximate. Das Gupta in references supports descriptive rate decomposition.

EXHIBIT 6 / FINANCIAL INTERPRETATION

Fewer recorded events do not deliver a proportional reduction in recorded cost

[RC] In Group B, the event count falls faster than gross concession cost. Their ratio implies a 5.15% increase in mean recorded cost per event, conditional on matching scope.

FEWER COUNTED EVENTS NEED NOT LOWER COST PROPORTIONALLY

Group B • each prior quantity = 100 • cost and event scopes must align

Gross cost 98.26	Conceded units 93.45	Cost per conceded unit 105.15
--	--	---

Mean recorded cost per conceded unit rises approximately 5.15%.

Composition, policy, timing and cost scope can change the mean.
No risk-adjusted severity, structural inefficiency or realized savings is established.

[AI] A count decline and a cost decline answer different questions. Average cost can change with severity, mix, reimbursement policy, timing or coverage. The ratio identifies neither the mechanism nor an effect of the bridge.

Finance validation	Evidence needed
[PD] Common accounting population	Link cost and events to the same eligibility, period, cohort, valuation policy and revision vintage.
[PD] Avoidable loss and realization	Identify the incremental mature events prevented by an action and the associated avoidable cost; distinguish expense, capacity and cash.
[PD] Incremental economics	Deduct delivery and validation cost, reconcile overlapping program credits and retain uncertainty in both effect and loss mix.

[AI] Recorded gross cost is an operating burden, not a reporting-mechanism saving. Do not annualize a single-period decline into an established benefit. An intervention’s economic endpoint must capture both frequency and severity.

MANAGEMENT IMPLICATION

[PD] Ask finance to validate the cost-event bridge before converting a lower defect count into a benefit claim. Evaluate the reporting mechanism’s decision value separately from the operational actions it informs.

[RC] S1:M02/E05. Separate prior count and cost indices =100; mean-cost index follows their ratio, not an average of indices. Approximately 5.15% is a same-scope accounting implication, not identified severity or savings.

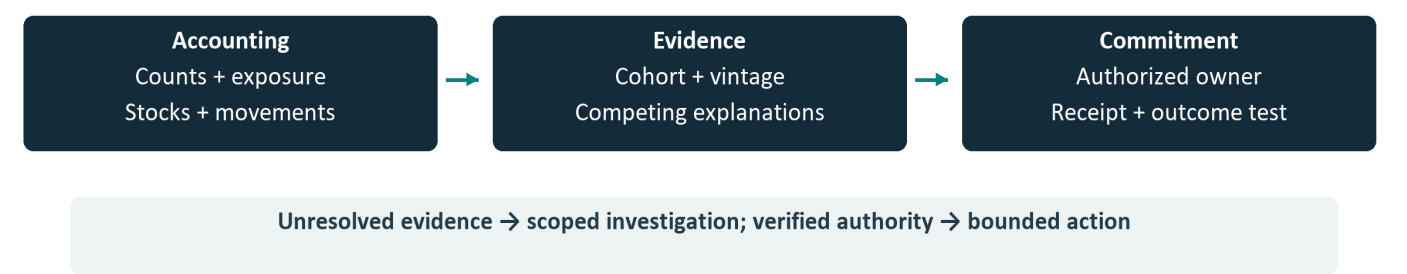
EXHIBIT 7 / CONDITIONAL ACTION ROUTING

A shared step should trigger a discriminating investigation and an authorized response

[AI] Synchronized changes across operating groups support a common-factor investigation. They do not distinguish a recording change from a shared operational shock.

THREE LEDGERS TURN THE BRIDGE INTO A DECISION RECORD

Proposed release protocol • each ledger has an owner and a completion test



A report can route a question before it can establish the answer.

[HR] Nonreceipt rises simultaneously across groups; the second group’s attended-delivery category also rises. [AI] Separately managed groups can share technology, policy, demand, weather and reporting changes. Weekly category changes do not establish daily negative-control behavior.

Three ledgers	Release authority and evidence
[PD] Accounting record	Data owner reconciles populations, versions, partitions and residuals; preserves actuals and traces authorized restatements.
[PD] Evidence record	Analytical owner tests competing explanations with linked events, configuration logs and prespecified contrasts. Retain unresolved status.
[PD] Commitment record	Operating owner accepts the action, deadline and verification requirement. Proposal, commitment, completion and effectiveness remain separate.

COO DECISION

[PD] Appoint the release owner and authorize bounded validation. Preserve actuals until a justified restatement is approved. Independently evaluate routing, loss and delay across all eligible opportunities, including holds, against reference evidence. Reviewer confidence is not ground truth. Test mature customer impact separately after verifying action delivery.

[HR/AI] S1:M03–M04/M13–M14/D07–D09; evidence E01–E06. [PD] Proposed routing and evaluation. Lipsitch et al. in references defines negative-control assumptions; no historical causal determination is established here.

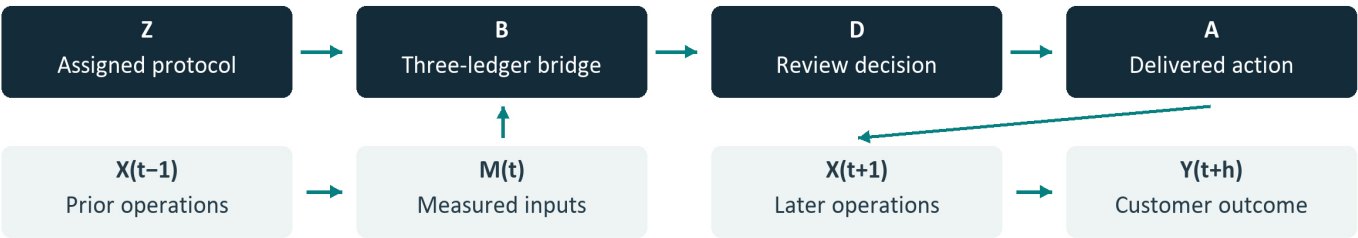
TECHNICAL FOUNDATION • THEORY OF CHANGE

The bridge changes a decision before it can change an outcome

[PD] The Three-ledger bridge separates accounting, evidence and commitment; each ledger controls a different failure in the decision pathway.

REPORTING POLICY AND CUSTOMER IMPACT FOLLOW DIFFERENT PATHS

Proposed time-indexed DAG • arrows are hypotheses, not estimated coefficients



D and A → L: 60-day decision-resource burden, the proposed primary endpoint.

Q quality, V vintage and E cohort/exposure → M; M → B.
C demand, staffing, mix and concurrent change → D, A, X and Y; U denotes unmeasured paths.
Later operations enter a future review cycle; no same-period feedback cycle is asserted.

[PD] Z assigns the governed bridge B; D is the resulting review decision, A the delivered operating action, X the operating state and Y the customer outcome. L is the primary decision-resource burden. Prior operations generate measurement M; quality Q, source vintage V and cohort/exposure E govern what M means. Shared causes C include demand, staffing, mix, calendar effects and concurrent programs; U represents unmeasured paths.

[AI] Accounting closure supports relationships among reported quantities. Configuration logs and event-level contrasts can test measurement explanations. An authorized action receipt supports delivery. None alone identifies the complete B→D→A→X→Y pathway. The historical record contains no design that isolates that pathway.

[PD] Index the pathway in time: the current review uses prior operating evidence; later operations enter a future review. Do not draw feedback as a same-period causal cycle. Estimate the policy effect without conditioning on post-assignment adoption, investigation or action completion. A residual is compatible with omitted paths; it does not name the path responsible.

Source S1: registry-attested reporting design, competing explanations and proposed actions. The DAG is a proposed causal model, not an identified historical effect.

TECHNICAL FOUNDATION • ESTIMANDS AND ELIGIBILITY

Keep every eligible review episode in the policy comparison

[PD] Freeze an episode before assignment: a source-defined issue, operating scope, review deadline and accountable decision role. Reports, claims and action lines are not interchangeable units.

Quantity	Exact contrast and boundary
Historical movement	[RC] Difference between reconciled period-level rates, counts or costs for the same construct, population, clock and vintage. This is descriptive; no intervention counterfactual is supplied.
Primary policy ITT	[PD] Expected 60-day avoidable decision-resource burden under assigned Three-ledger review versus the comparator, averaged across all upstream-frozen eligible episodes. Retain held, unresolved and noncompliant episodes.
Delivered-action effect	[PD] Effect of a specified, verified operating package on eligible operational opportunities. Review assignment and package receipt are different treatments; an as-treated contrast needs additional identification.
Customer effect	[PD] Assignment effect on a prespecified shipment-cohort outcome at a justified maturity horizon. Neither the resource endpoint nor the historical bridge identifies this separate effect.

[PD] The resource endpoint combines independently adjudicated avoidable work with preregistered unresolved/delay penalties. A justified investigation that finds no defect is not waste. Publish component outcomes—unsupported intervention, unnecessary investigation, justified hold, unresolved status, delay and customer-harm flags—alongside the aggregate.

[PD] Freeze loss weights, status precedence and observation rules before randomization. Incomplete ascertainment remains visible through missingness bounds and sensitivity; it does not become a zero-loss episode. Equal episode volumes and a constant policy effect underpin the illustrative power/economic alignment; heterogeneous effects require explicit target weights and recalculation.

Source S1 establishes reporting episodes and named actions, not observed policy effects. Prospective estimands and scoring rules are proposed.

TECHNICAL FOUNDATION • MEASUREMENT AND REVISION

A downward revision requires more than late additions

[AI] With a fixed eligible denominator and additions-only cumulative events, later ascertainment cannot reduce the rate. A downward source-labeled maturity revision falsifies that model.

Metric family	Validity contract
Concession rate	[PD] Bind unique eligible events to shipped exposure, business scope, event/claim clock, taxonomy and extraction vintage. A claim-flow ratio is not automatically shipment-cohort risk.
Lost and sub-buckets	[PD] Verify event overlap, an exhaustive partition and denominator alignment. Opposite aggregate movements do not prove disjoint underlying events.
Cost and exposure	[PD] Match cost and event scope; severity mix and refund policy can change cost per event.
Input indicators	[PD] Freeze eligibility, denominators, node crosswalk and goal vintage for dispatch exceptions, resolution, sorting and adoption. Resolve relative-percent versus percentage-point changes before comparison.
Action follow-through	[PD] Separate proposal, acceptance, authorization, completion and effect; a target is not a receipt.

[PD] Recover event, creation, update and extraction times; reversals; category changes; and exposure versions. Retain as-issued and harmonized series. Removals invalidate an additions-only completeness correction. Independent source/configuration evidence establishes reference facts; unresolved truth stays unknown.

[PD] Double-code 400 records, blinded and stratified by metric, node, vintage and dispute. Assumed design effect 1.5 gives worst-case agreement precision about ±6 percentage points at 95%. Pilot kappa's prevalence-dependent interval. Report agreement, confusion matrices, prevalence and kappa with clustered uncertainty; adjudicate after independent coding.

Source S1: source-labeled revisions, node proxies, taxonomy changes and action-stage ambiguity. Croushore–Stark supports vintage preservation. Audit size and precision assumptions are proposed.

TECHNICAL FOUNDATION · ACCOUNTING DECOMPOSITION

Separate event growth from exposure before naming the driver

[RC] For $R=N/E$, the rate ratio equals the event-count ratio divided by the exposure ratio. Both components can move the reported rate.

[RC] Let $n=N_i/N_o$ and $e=E_i/E_o$; set the prior rate index to 100. Symmetric event contribution= $50(n-1)(1+1/e)$. Symmetric exposure contribution= $50(n+1)(1/e-1)$. Their sum is $100(n/e-1)$. These expressions average the two substitution orders and split their interaction equally.

[AI] Fixing exposure at its prior level answers an accounting question. It does not identify the rate under an intervention that prevented the exposure change, nor a no-holiday outcome. The symmetric allocation is a declared convention, not causal responsibility. Rounded input changes require a rounding sensitivity.

Closure test	What must be established
Buckets to total	[PD] Mutually exclusive, exhaustive events; common clock, scope, denominator and vintage. Exclude an overlay only after establishing its overlap with the partition.
Residual tolerance	[PD] Preserve signed residuals. Derive rounding bounds from each term's precision. A count residual is not explained merely because displayed rates are rounded.
Mix adjustment	[PD] For common groups, w is exposure share and r is group rate; averages span the two periods. $\Delta R=\Sigma \text{average}(w)\times\Delta r+\Sigma \text{average}(r)\times\Delta w$. Recover group exposures; do not pool separately indexed groups.

INTERPRETATION

[AI] A closed bridge locates accounting movement. Source lineage determines comparability; a separate design identifies intervention effects.

Source S1 and protected recalculations: numerator/exposure changes and nonzero residuals. Das Gupta (1993), two-factor product and composition decomposition.

TECHNICAL FOUNDATION • VARIATION AND SIGNAL

A synchronized step deserves investigation before causal labeling

[AI] Shared timing supports a common-shock hypothesis. It does not distinguish policy, technology, operations, mix or measurement as the cause.

[RC] $MR_t = |x_t - x_{t-1}|$; tentative individual limits = $\text{mean}(x) \pm 2.66 \text{ mean}(MR)$, with moving-range upper limit = $3.268 \text{ mean}(MR)$. Every computable sequence is evaluated in the private ledger; each public series uses its own first observation = 100.

Actual sequence	Computed result and classification
Thirteen weekly pairs	[RC] Two observations each; no individual or moving-range crossings. Both observations necessarily fall inside self-estimated limits: non-diagnostic, not evidence of improvement or deterioration.
Two non-receipt daily prefixes	[RC] Four known observations each; no crossings. Tail values are ranges, not exact points. Limits span the suspected change: non-diagnostic; do not infer absence of a step.
Seven-day damage series	[RC] Index center 117.53; mean MR 9.37; individual limits 92.61–142.45; MR upper 30.61. No crossings: compatible with routine variation under tentative retrospective limits, not certified stability.

[AI] A comparison outcome is a diagnostic contrast until negative-control assumptions hold. Weekly attended-category growth is not a daily control series. Shared calendar, software or operational shocks can affect both networks.

[PD] Predeclare change windows and contrasts. Never create days from ranges, combine maturity versions as time points or treat gapped observations as consecutive weeks. Use p/Laney p' only with valid binary counts and actual denominators; recurrent events require a count-per-exposure model.

Source S1: short daily traces and revised weekly comparisons. Exact protected SPC results are supplied by the case calculation ledger. NIST, Laney and Lipsitch et al. govern interpretation.

TECHNICAL FOUNDATION • SELECTION AND HETEROGENEITY

An outlier list defines a selection problem, not an effect estimate

[AI] Chronic regional misses and acute site exceptions can justify different reviews; neither produces an untreated counterfactual by itself.

Moderator or selection	Mechanism and evaluation requirement
Worst-unit targeting	[PD] Freeze membership, ranking rule, source vintage and entry date before follow-up. Compare selected and nonselected units on pre-period levels, reliability and mature outcomes when available.
Node and operating scope	[PD] Verify business-line identity separately from node labels. Test whether a proposed action addresses the same process and authority constraints across scopes.
Attended versus unattended	[AI] Different disposition patterns can distinguish hypotheses, but mix, exposure and customer behavior can differ. A weekly count comparison is not a daily controlled contrast.
Calendar and maturity	[PD] Model exposure timing and cohort age explicitly. A holiday or a revision changes the observation process as well as the operational interpretation.
Action authority	[PD] A bridge should improve routing where a decision owner can act on verified evidence. It cannot overcome missing permissions, unavailable capacity or an untested package by changing prose.

[PD] Synthetic reversion sensitivity: for stationary equal-variance bivariate-normal scores with correlation ρ , a selected cohort's expected follow-up excess equals ρ times its initial excess. At assumed $\rho=.3/.6/.9$, expected excess falls 70%/40%/10% without intervention. No recovered panel estimates historical reliability.

[PD] Preregister mechanism-based interactions and report uncertainty. Do not average incompatible taxonomy definitions as effect heterogeneity. Selection sensitivity is not an estimate of intervention benefit.

Source S1: selected regional/site reviews, node proxies, disposition differences and authority constraints. Selection-aware comparisons and moderator tests are proposed.

TECHNICAL FOUNDATION • CONTRIBUTION ATTRIBUTION

Credit the reporting mechanism separately from operational results

[HR] The register attributes reporting design, analytical choices and codification to the case principal. It assigns data production, subject-matter review and field execution to other roles.

Contribution claim	Process-tracing grade and boundary
Designed the recurring bridge	[HR] Provisional hoop support: the register describes a dated format decision and ownership. Original decision records have not been independently authenticated in this case.
Made analytical method choices	[HR] Provisional hoop support: recorded normalization, reconciliation and source-quarantine decisions match the proposed role. The register is not independent corroboration of its underlying files.
Authored and codified production	[HR] Provisional hoop support: registered metadata and workflow descriptions support involvement. They do not establish sole authorship, sustained automation coverage or labor saved.
Changed review decisions	[AI] Unestablished as a measured effect: written recommendations and validation assignments support intent. A receipt and decision history are needed to confirm use.
Delivered customer or cash benefit	[AI] Unestablished: no isolated counterfactual connects this reporting mechanism to those outcomes. Reported concession costs belong to operations, not the bridge's benefit ledger.

[PD] A hoop test checks a necessary condition and cannot prove the claim alone. A smoking-gun test needs reliable evidence unlikely under competing explanations. Doubly decisive evidence must also exclude the relevant alternatives. Repeated summaries of one original are one evidence lineage, not multiple independent confirmations.

[AI] The defensible contribution is a method for producing a more auditable operating discussion. Higher attribution grades require original artifacts, release records, accepted commitments and verified downstream activity.

Source S1: registry-attested ownership, author metadata and responsibility boundaries. Grades apply to this case's available evidence, not to the principal's complete career record.

TECHNICAL FOUNDATION • PROSPECTIVE EVALUATION

Protect the full decision episode when teams cross into treatment

[PD] Randomize independent review teams to governed Three-ledger review or the comparator; evaluate the policy at the decision-resource endpoint it can directly change.

[PD] The proposed enrollment window spans twelve weekly periods. Six balanced sequences begin treatment in periods 3, 5, 7, 9 and 11, with the final sequence waitlisted until every enrolled episode's 60-day outcome closes. Eligibility is frozen upstream. Preserve each episode's assigned protocol and staff through closure; record spillovers and adherence.

Design element	Required implementation
Assignment and analysis	[PD] Stratify by operating scope and baseline decision burden; randomize sequence. Use team and calendar effects, retain ITT, and blind independent outcome adjudication to assignment where feasible.
Primary outcome	[PD] Sixty-day avoidable decision-resource burden with preregistered unresolved/delay penalties. Track separate harm, hold, latency and resource components; necessary negative investigations are not classified as waste.
Inference and power	[PD] The illustrative model assumes a common constant effect and known team-period covariance. Calibrate variance, serial dependence, volumes and finite-team inference in a pilot; evaluate heterogeneous or delayed effects explicitly.
Contamination rule	[PD] If protocol/staff separation through closure is infeasible, use a parallel cluster trial and recalculate power. A two-week switch must not silently redefine the control episode's 60-day exposure.

[PD] Inspect baseline/event-time patterns and effect dynamics. For a planned nonrandomized rollout, consider Callaway–Sant’Anna group-time effects with credible not-yet/never-treated controls, conditional parallel trends and anticipation/spillover checks. Use pretrend equivalence/sensitivity; nonsignificance proves nothing. Scale requires the lower benefit bound above the same economic break-even; already-avoidable burden receives no second avoidability haircut.

Prospective design; all sample sizes, covariance and financial inputs on the subsequent model pages are synthetic planning assumptions, not achieved power or historical savings.

TECHNICAL FOUNDATION • LITERATURE AND PROPOSITIONS

The Three-ledger bridge makes four operating claims falsifiable

[PD] The framework joins accounting reconciliation, evidence discrimination and authorized commitment. Its effectiveness remains a testable proposition.

Proposition	Mechanism, boundary and disconfirmation
P1 • Accounting	[PD] An explicit numerator/exposure and residual ledger reduces unsupported driver claims per eligible review. Falsified if blinded reanalysis finds no reduction, including withheld or unresolved episodes.
P2 • Evidence	[PD] Cohort/vintage lineage plus prespecified diagnostic contrasts improves operational-versus-measurement routing. Falsified if adjudicated routing fails to improve or delay harm offsets the gain.
P3 • Commitment	[PD] Accepted ownership and a verification receipt improve independently confirmed follow-through. Falsified by no follow-through gain despite comparable authority, capacity and eligibility.
P4 • Decision value	[PD] Combined ledgers reduce 60-day burden beyond incremental cost. Scale requires the lower benefit bound above break-even. An upper bound below break-even contradicts economic benefit; a crossing interval is inconclusive. A prespecified harm failure defeats adoption.

[AI] Das Gupta's standardization work separates rate and composition arithmetic; this case applies that distinction before prioritizing operational action. Croushore–Stark's real-time data work motivates separate as-issued and revised records. Lipsitch and colleagues require assumptions before a diagnostic contrast becomes a negative control.

[AI] NIST's process-chart logic distinguishes ordinary variation from a signal. The case extends these methods into an auditable decision contract; it does not establish a new statistical theorem. Its source-labeled downward revision contradicts an additions-only accrual explanation under a fixed denominator, not the broader possibility of legitimate data maturation.

Framework and propositions are proposed. Primary sources: Das Gupta (1993); Croushore–Stark (2000); Lipsitch et al. (2010); NIST/SEMATECH. No external validation or firm ranking is claimed.

TECHNICAL FOUNDATION · LIMITATIONS AND REPRODUCIBILITY

The evidence supports a reporting design and a bounded audit

[AI] Passing this case's review certifies consistency with the available record; it does not turn registry attestations into authenticated originals or descriptive changes into effects.

Limitation	Consequence and required evidence
Source authenticity	[HR] The project register is available; cited operational originals are not authenticated here. Recover immutable files and release logs before upgrading attribution or adoption claims.
Measurement	[AI] Source-vintage, scope, taxonomy and percentage-unit conflicts restrict comparisons. Preserve competing versions and explicit residuals rather than selecting the most favorable narrative.
Identification	[AI] Shared timing, reported actions and later metric changes do not isolate the bridge's effect. Obtain randomized assignment or a justified comparison with measured treatment delivery.
Economics and transport	[AI] No verified reporting ROI, released capacity or customer savings is observed. Synthetic assumptions need pilot calibration, local authority checks and an explicit financial-realization mechanism.
Privacy	[PD] Public indices and relative changes protect internal levels. Avoid complementary views or anchors that invert the transformation; retain raw records and identity mappings privately.

[PD] Reproduction starts with the unchanged register hash, frozen project extract and cell-level claim ledger. Run private evidence checks, reconcile each output to its source/version, then run the separately labeled synthetic decision model. Preserve formulas, rounding conventions, exclusions and all failed alternatives. Public tables must reconcile to the protected output, not a fresh uncontrolled pull.

[PD] Retain source-maturity alternatives as branches until comparable event/cohort records resolve them. Record code and data versions, parameter provenance and reviewer findings. Author-led, agent-assisted review checks arithmetic and claims; it is not an external academic or client audit.

Consolidated limitations across source access, measurement, causal attribution, valuation, transportability and confidentiality. Private annex contains raw source mappings and executable checks.

MODEL 1 / DECISION-RESOURCE ENDPOINT

Measure the cost of the decision pathway, including unresolved work

[PD] These four pages are synthetic planning models. They define a future test; they do not estimate historical customer or financial benefit.

[PD] Freeze all eligible weekly review episodes upstream of report production. Observe each episode for 60 days from its initial decision, then adjudicate using independent reference evidence and a codebook blind to assignment. Retain held, redirected and unresolved episodes; do not condition the denominator on actions completed.

Endpoint component	Measurement rule
Avoidable work	[PD] Record actual investigation, rework and remediation resources. Classify avoidability against the evidence available at the decision and the independently recovered facts. A justified investigation with a negative finding is not automatically waste.
Delay or unresolved status	[PD] Apply preregistered bounded penalties calibrated to the same resource-equivalent unit as observed work. Record time and reason separately. Do not count the same loss in both an observed cost and a penalty.
Required companion measures	[PD] Publish unweighted routing errors, resolution, hold frequency, cycle time, investigator workload and harmful premature actions. A lower average burden cannot excuse unacceptable harm or suppressed investigations.
Cost boundary	[PD] Exclude incremental bridge-production and validation effort from this endpoint; include it once as implementation cost K. Customer outcomes remain a separate endpoint.

[PD] The estimand is the intention-to-treat reduction in mean 60-day burden per all eligible episodes. Normalize to an independently measured control baseline of 100. Equal eligible volumes and baseline burdens across team-periods are planning assumptions; actual imbalance requires weights and new power.

[PD] For unresolved truth, retain a prespecified interval of plausible losses and test the investment decision at both bounds. Pilot classification reliability and loss weights. Independent adjudication means evidence-based review, not agreement among confident authors.

Proposed endpoint; no historical episode counts, error prevalence or resource burden are supplied.

MODEL 2 / ECONOMICS AND REALIZATION

A reduction in avoidable workload pays only when resources can be released

[PD] The same eligible episodes and 60-day outcome horizon define the burden reduction and the value pool.

[PD] Net value = $100 \times r \times f \times q - K$. Reduction r is the policy ITT on the defined burden; f is the share convertible to released resources or cash; q is the share not already credited elsewhere. The synthetic base uses $f = 0.80$, $q = 0.90$ and incremental cost $K = 4$.

Assumed burden reduction	f = 50%	f = 80%	f = 100%
4%	-2.20	-1.12	-0.40
8%	-0.40	+1.76	+3.20
12%	+1.40	+4.64	+6.80

[RC] Break-even fraction $r^* = K / (100 \times f \times q) = 0.0555556$, or 5.55556%. On a baseline of 100 this also equals 5.55556 index points. At 12%, base net value is +4.64 units. These are normalized resource-value units, not money or estimated savings.

[PD] The endpoint already excludes unavoidable work. Do not apply a second avoidability haircut. Financial realization is a separate mechanism: canceled external spend, released paid hours or deployable capacity with evidenced use. If the endpoint already measures exclusive realized cash, set f and q to 1.

[PD] Allocate setup and recurring review cost across the enrolled cohort using the same horizon; do not deduct setup per episode and again centrally. Keep delay penalties out of K when already in the endpoint. Jointly vary workload reduction, resource conversion, overlap and implementation cost; the grid is sensitivity, not a confidence interval.

DECISION RULE

[PD] Scale only if the benefit interval clears break-even, harm and resolution guardrails pass, and Finance verifies the resource-release mechanism. A reporting score or a gross concession charge cannot substitute for those tests.

Synthetic value model. Registered gross concession costs are operating charges, not benefits created by the reporting mechanism.

MODEL 3 / ROLLOUT AND ECONOMIC POWER

Protect the 60-day episode before selecting the rollout size

[PD] Randomize teams equally to six sequences over twelve enrollment weeks: switches in weeks 3, 5, 7, 9 and 11; the sixth stays on the comparator through final follow-up.

[PD] Freeze each episode’s assigned protocol and dedicated personnel through its 60-day outcome. A later team switch must not change an earlier episode’s treatment. If personnel, learning or shared actions defeat that separation, use a parallel cluster design and recalculate power. The waitlist does not switch automatically after week 12.

[PD] Fit team and enrollment-week effects with a constant immediate assignment effect. Positive reduction is minus the treatment coefficient. Covariance is $\sigma^2[\rho 11' + (1-\rho)AR1(\phi)]$: team-period mean SD $\sigma = 18$, persistent variance share $\rho = 0.25$, serial $\phi = 0.35$. Independent teams and Gaussian errors with known covariance are planning assumptions.

Independent review teams	Zero-null MDE	Economic power at 12%
36	8.21	59.4%
72	5.81	87.5%
108	4.74	96.8%
144	4.11	99.3%

[RC] MDE uses two-sided 5% size and 80% power against zero. Economic power is $\Pr(\text{lower 95\% bound} > 5.56\%)$ at an assumed true reduction of 12%. The minimum meeting BOTH $MDE \leq \text{break-even}$ and economic power $\geq 80\%$ is 84 teams, rising to 144 if $\sigma = 24$. The smaller 60/108 economic-only minima fail the MDE gate.

[PD] The two reported operating networks are not two adequate randomized teams. Map genuinely independent review units and resource ownership first. Pilot covariance, episode volume, long-tail costs, adjudication error, onboarding and spillovers; calibrate finite-cluster inference. These calculations do not power a customer-outcome trial.

Synthetic six-sequence team/calendar GLS. Profiled and full-design information calculations agree; no achieved trial power is asserted.

MODEL 4 / VALUE OF DIAGNOSTIC INFORMATION

Buy more evidence when it changes the best investigation

[PD] This decision model values evidence about an operational, measurement or mixed change. It does not assume that the bridge improves customer outcomes.

Action / state	Operational	Measurement	Mixed
Operational intervention	+18	-12	+6
Source repair	-6	+15	+4
Controlled joint diagnosis	+5	+5	+10

[PD] Prior state probabilities are 0.40, 0.35 and 0.25. Payoffs are invented net diagnostic-resource units including each action’s effort and delay. The best current action is controlled joint diagnosis. A preliminary audit yields the correct state signal with probability a and either wrong signal with probability (1-a)/2. Choose the action with the highest posterior expected payoff.

Signal accuracy a	Gross EVSI	Net after audit cost 2
45%	0.429	-1.571
70%	4.163	2.163
90%	7.188	5.188

[RC] EVSI equals expected optimal payoff with information minus optimal payoff now. The perfect-information ceiling is 8.700; all scenarios remain below it. Posterior probabilities, chosen actions and audit-cost sensitivities are retained in the package.

[PD] These ranges are hypothetical. Replace priors, audit likelihood and payoffs with pilot evidence before commissioning the study. A recorded validation hold establishes a decision, not its monetary value. Do not add EVSI to the operating-policy benefit when both concern the same prevented work.

Synthetic finite-state Bayesian decision model. Scalar joint-payoff enumeration and posterior optimization are compared in the audit.

Use the bridge to decide what must be verified next

[AI] The new finding is an interpretation correction: accounting closure alone cannot justify a cause, an intervention or a financial claim.

Claim class	Meaning in this case
[HR] Historical record	A statement or artifact described by the supplied project register. The underlying operational originals were not authenticated here.
[RC] Recalculated	Arithmetic derived from registered values or a declared synthetic model. The input class is stated beside the result.
[AI] Analytical inference	A reasoned interpretation or constraint. It does not imply an identified causal effect.
[PD] Proposed design	A future framework, protocol, experiment or planning scenario. It is not claimed as historically implemented.

[AI] Shewhart–Deming explains why acting on routine variation can worsen a process; the available short sequences do not certify stability. Campbell–Goodhart motivates scrutiny of scored indicators, without evidence of gaming here. Ittner–Larcker emphasizes the need to validate links between measurement and performance. Tucker–Edmondson distinguishes a workaround from preventing recurrence.

[AI] The Three-ledger bridge operationalizes those concerns through reconciliation, discriminating evidence and accepted commitments. Its weighted resource endpoint also needs the construct, weighting and sensitivity discipline of composite-indicator design. This case illustrates those principles and extends their application; it does not empirically validate the framework.

[PD] Reproduction chain: unchanged source-register hash → selected cells and claim ledger → raw arithmetic and source-version branches → protected exhibits → separately labeled synthetic model → PDF and review checks. The private package contains source locators, formulas, executable models, review findings and build files.

DISTRIBUTION

[PD] Share this anonymized PDF as the portfolio case. Keep the separate evidence annex and reproduction package private: they contain identities, internal source paths and raw values.

Source S1: supplied project register. Public historical quantities are indexed or relative; raw values and exact source dates remain private.

REFERENCES / PRIMARY METHODS AND SUBSTANTIVE LITERATURE

Reconciliation and measurement need explicit definitions

Sources support the methods and proposed design. They do not authenticate this project or establish its results.

R1. Das Gupta, P. (1993). Standardization and Decomposition of Rates: A User's Manual. U.S. Census Bureau, P23-186. Exact descriptive decomposition; the operative two-factor formula was checked against the official full text.

R2. Croushore, D., & Stark, T. (2000). A Real-Time Data Set for Macroeconomists: Does Data Vintage Matter for Forecasting? Federal Reserve Bank of Philadelphia, WP 00-6. Vintage preservation is a methodological analogy, not logistics evidence.

R3. Lipsitch, M., Tchetgen Tchetgen, E., & Cohen, T. (2010). Negative Controls: A Tool for Detecting Confounding and Bias in Observational Studies. *Epidemiology*, 21, 383–388. Exclusion and shared-bias assumptions for diagnostic contrasts.

R4. Wang, R. Y., & Strong, D. M. (1996). Beyond Accuracy: What Data Quality Means to Data Consumers. *Journal of Management Information Systems*, 12(4), 5–33. Context and fitness for use.

R5. W3C (2013). PROV Model Primer. Entities, activities, agents, versions and derivation. A provenance record supports assessment; it does not certify accuracy.

R6. OECD / European Commission JRC (2008). Handbook on Constructing Composite Indicators: Methodology and User Guide. Construct definition, normalization, weighting and sensitivity.

R7. Ittner, C. D., Larcker, D. F., & Randall, T. (2003). Performance implications of strategic performance measurement in financial services firms. *Accounting, Organizations and Society*. Measurement design and performance links.

R8. Campbell, D. T. (1976). Assessing the Impact of Planned Social Change. Original occasional-paper record, ERIC. Evaluative pressure and indicator validity.

Linked original research, institutional guidance and author manuscripts. Scope and access limits are recorded in the private research memo.

REFERENCES / PRIMARY METHODS AND SUBSTANTIVE LITERATURE

Variation, selection and learning require separate evidence

Sources support the methods and proposed design. They do not authenticate this project or establish its results.

R9. Goodhart, C. A. E. (1984). Problems of Monetary Management: The UK Experience. *Monetary Theory and Practice*, 91–121. Indicator relationships under control; publisher chapter record.

R10. Tucker, A. L., & Edmondson, A. C. (2003). Why Hospitals Don't Learn from Failures. *California Management Review*. Author working paper on workarounds and organizational learning.

R11. NIST/SEMATECH. e-Handbook, Individuals Control Charts. Successive moving ranges and individual limits. The case discloses the rounded 2.66 constant and short-series limitations.

R12. NIST/SEMATECH. e-Handbook, Attributes Control Charts. Proportions of defective units and counts of defects require different measurement contracts.

R13. Laney, D. B. (2002). Improved Control Charts for Attributes. *Quality Engineering*, 14(4), 531–537. Extra-binomial dispersion; abstract verified. No Laney chart is estimated without valid counts and denominators.

R14. The W. Edwards Deming Institute. The Funnel Experiment. Institutional explanation of tampering with a stable process; not a diagnosis that this project's process was stable.

R15. Barnett, A. G., van der Pols, J. C., & Dobson, A. J. (2005). Regression to the mean: what it is and how to deal with it. *International Journal of Epidemiology*, 34, 215–220. Extreme selection and repeated observation.

R16. Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*. Chance-corrected agreement; kappa is not reference truth.

Linked original research, institutional guidance and author manuscripts. Scope and access limits are recorded in the private research memo.

REFERENCES / PRIMARY METHODS AND SUBSTANTIVE LITERATURE

Causal attribution and investment need compatible tests

Sources support the methods and proposed design. They do not authenticate this project or establish its results.

R17. Collier, D. (2011). Understanding Process Tracing. PS: Political Science & Politics. Necessary and discriminating evidence for bounded contribution claims.

R18. Hernán, M. A., & Robins, J. M. Causal Inference: What If. Author-hosted text. Treatment definition, causal graphs, selection and policy estimands.

R19. Hussey, M. A., & Hughes, J. P. (2007). Design and analysis of stepped wedge cluster randomized trials. Contemporary Clinical Trials, 28, 182–191. Cluster rollout and power.

R20. Hemming, K., et al. (2015). The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting. BMJ, 350, h391. Calendar effects and implementation structure.

R21. Hemming, K., et al. (2018). Reporting of stepped wedge cluster randomised trials: extension of the CONSORT 2010 statement with explanation and elaboration. BMJ, 363, k1614. Sequences, periods and inference.

R22. Callaway, B., & Sant'Anna, P. H. C. (2021). Difference-in-Differences with Multiple Time Periods. Journal of Econometrics. Author manuscript: group-time comparisons for a justified nonrandom alternative.

R23. Roth, J. (2022). Pretest with Caution: Event-Study Estimates after Testing for Parallel Trends. American Economic Review: Insights. Passing a pre-trend test does not prove identification.

R24. Strong, M., Oakley, J. E., Brennan, A., & Breeze, P. (2015). Estimating the Expected Value of Sample Information Using the Probabilistic Sensitivity Analysis Sample. Medical Decision Making. Decision value from information.

Linked original research, institutional guidance and author manuscripts. Scope and access limits are recorded in the private research memo.

FINAL REFEREE REPORT / INTERNAL, AGENT-ASSISTED REVIEW

The revised case supports better investigation, with bounded attribution

[AI] The author and three parallel review agents checked source fidelity, methods and executive use. This is internal review, not external academic or client certification.

Major issue	Resolution in the revised case
Source and ownership	Register cells and source lineages are preserved; unrecovered originals remain an explicit boundary. Reporting stewardship receives provisional hoop support, not field-impact credit.
Reconciliation and cause	Count/exposure allocation, stock residual and conflicting vintages are exposed. Symmetry is an accounting convention. No shared timing, sign change or closure is promoted to a causal effect.
Evaluation and economics	Episode exposure stays fixed through 60 days. Both detectable break-even and economic-power gates apply; independent model implementations agree. Proposed value is kept separate from historical charges.
Confidentiality and reproduction	Public historical levels are indexed or relative; raw cells and identities remain private. Source, model, visual and artifact checks are retained for replication.

Minor revisions completed

[AI] The loss-partition exhibit identifies each segment. Two-point cost changes use descriptive language. Input indicators receive a measurement contract. An inconclusive interval is separated from evidence contradicting economic benefit. Source-version sensitivity is retained rather than silently choosing the favorable version.

Partner critique: would the executive act on page 1?

[AI] Yes: require numerator/exposure decomposition, preserve unresolved partitions and vintages, then assign a scoped investigation. The front page states that decision and the derived finding. It does not ask the executive to approve a savings claim unsupported by the evidence.

REVIEW DISPOSITION

[AI] Suitable for an anonymized operating-method portfolio case after the recorded checks. Customer impact, realized savings and broader implementation claims require additional evidence.

Internal author-led review; evidence and methods agents challenged the draft. Final artifact and visual checks are recorded in the private reproducibility package.