

OPERATING CASE STUDY / CUSTOMER EXPERIENCE & LOGISTICS

Recovery Process & Customer Risk

Turn a contested performance signal into a defensible operating decision

An anonymized analytical case examining whether stronger exception recovery creates downstream damage. It reconciles measurement errors, tests the logic of the historical verdict and designs a prospective customer-safety evaluation.

PROJECT CONTEXT / RECOVERY QUALITY & CAUSAL DECISION DISCIPLINE

BLUF / BOTTOM LINE UP FRONT

Keep the operating question open.
Test customer harm before expanding recovery.

ANALYTICAL FINDING

A positive delayed causal effect can coexist with a weakening, sign-reversing correlation. The historical pattern does not establish no harm; it supports a controlled test of a defined recovery package.

CONTRIBUTION IN THE SOURCE RECORD

Authored the diagnostic, corrected metric comparisons and proposed a forward test. Provisional registry hoop support; delivery unconfirmed.

EVIDENCE STATUS

Originals unrecovered. Historical causal effects and benefit are not identified. The reconciliation and evaluation design use explicit hypothetical test inputs.

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

EXECUTIVE ABSTRACT / DECISION AND OWNERSHIP

Preserve approved practice while resolving the customer-harm question

[AI/PD] A recovery score describes process performance. It does not identify the customer effect of a specific recovery action. The proposed operating decision separates the evidence needed to change practice from the evidence needed to investigate a risk.

Abstract element	Decision-relevant synthesis
Context [HR]	A senior reviewer questioned whether stronger exception recovery was increasing damage. The author compared weekly movements, process-specific measures, two networks and daily lag patterns, then consolidated revised findings.
Objective [PD]	Decide whether to retain, constrain or expand a defined recovery-assistance package using mature customer outcomes and net resource value. Avoid an unsupported shutdown or an unsupported declaration of safety.
Approach [RC/AI]	Reconcile metric bases and observation windows; distinguish detection from physical origin; test the logic of the causal verdict; specify a clustered prospective comparison and decision-relevant power.
Findings [AI]	Lagged correlation decay does not refute harm. Aligning denominators corrects a rate comparison but does not make process detections a subset of customer claims. Stable early extracts do not establish ultimate outcome maturity.
Decision value [PD]	Keep the harm hypothesis testable. Replace an uncalibrated correlation trigger with mature-outcome safety and value gates, plus immediate review of credible severe incidents.
Implications [PD]	The operating sponsor owns the action version and scope. Quality owns the harm margin and incident response. Analytics owns cohort linkage, independent ascertainment and the prespecified evaluation.

EXECUTIVE ACTION AND THRESHOLD

[PD] Preserve current approved safeguards. Fund calibration and a bounded randomized rollout only after eligibility, custody boundaries and outcome capture are validated. Expand only when the upper harm bound is below the agreed margin and the lower net-value bound exceeds zero.

Basis: S1 selected project register. [HR] historical record; [RC] recalculated; [AI] analytical inference; [PD] proposed design. No identified historical effect or realized saving.

EXHIBIT 1 / HISTORICAL QUESTION AND DEFENSIBLE ANSWER

Separate an unsupported accusation from a demonstrated absence of harm

[HR] The recorded investigation challenged a plausible mechanism: recovered packages could sustain or carry damage into later stages. Its strongest operating contribution was to expose competing explanations and improve the measurement basis. [AI] Its stated causal rejection exceeded the available identification.

A DIAGNOSTIC VERDICT IS NOT AN IDENTIFIED CUSTOMER EFFECT

Historical question + analytical reassessment / unsupported harm is not demonstrated safety



Diagnostic lens	What it can establish
Weekly direction [HR/AI]	Recovery and damage moved in opposite directions in the principal network. This weakens a simple contemporaneous story; changing mix, reporting and other operations can still mask a harmful effect.
Process fingerprint [HR/AI]	A genuinely pre-action damage event cannot be a required consequence of later handling. Verify event order and coding before using a reporting bucket to reject a mechanism; detected share does not bound preventable harm.
Second network [HR/AI]	The comparison provides a directional cross-check. Different populations and treatment histories prevent it from serving as an exchangeable untreated control.
Lag pattern [HR/AI]	The reported association weakens and changes sign across lags. Without an identified model of common drivers, dynamics and observation timing, that pattern cannot distinguish harm from no harm.

DECISION THAT SURVIVES THE REVIEW

[AI/PD] The record does not justify throttling approved recovery solely because two metrics correlate. It also does not justify closing the harm question. Retain safeguards and commission a prospective comparison of an explicit action package.

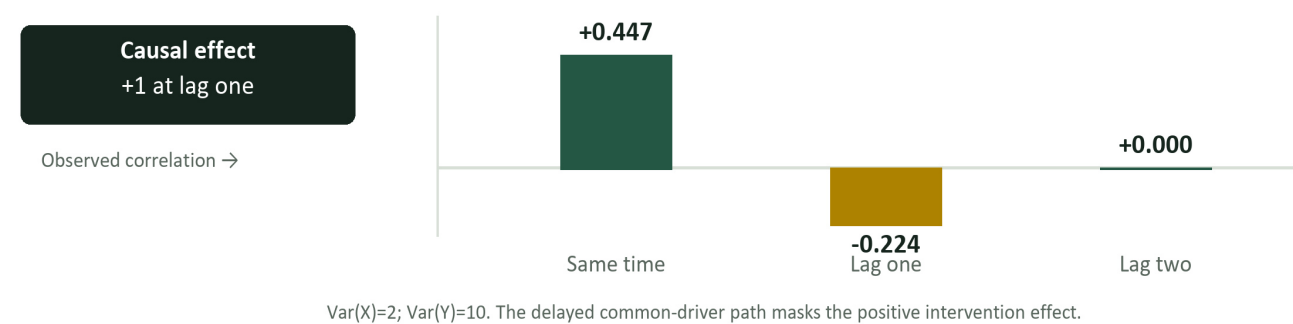
Basis: S1 master, M01–M12, M23 and D01–D05. The historical verdict is attributed; its causal interpretation is reassessed in this case.

EXHIBIT 2 / A FALSIFIABLE COUNTEREXAMPLE

A real delayed harmful effect can hide behind a sign-reversing correlation

[AI] A causal claim about all lag patterns can be disproved by one valid counterexample. The synthetic system below contains a positive delayed effect of X on Y, yet its observed correlation is positive at the same time and negative one period later. It is a logical demonstration, not a fit to the project data.

POSITIVE CAUSAL EFFECT, NEGATIVE DELAYED ASSOCIATION
Synthetic structural counterexample / all quantities are hypothetical, not fitted project values



[AI] Let independent, centered, unit-variance U , e and v generate $X_t = U_t + e_t$ and $Y_t = 2U_t + X_{t-1} - 3U_{t-1} + v_t$. An intervention that raises X_t by one unit raises Y_{t+1} by one unit, holding the other causes fixed. The direct delayed causal coefficient is positive.

Quantity	Derived implication
Observed association [RC]	Var(X) = 2; Var(Y) = 10. The same-time correlation is $2/\sqrt{20} \approx +0.447$. At lag one it is $-1/\sqrt{20} \approx -0.224$; at lag two it is zero.
Unmeasured path [AI]	The common driver raises X now and separately lowers Y later. Its delayed path outweighs the positive $X \rightarrow Y$ path in the observational covariance.
Operating consequence [PD]	Use the lag pattern to specify candidate mechanisms and sampling windows. Require an identifying comparison before making a no-harm or harm claim.

WHAT CHANGES

[AI] Replace “a real mechanism must strengthen with lag” with a model-specific prediction. A failed correlation prediction can reject a restricted model; it cannot eliminate every harmful recovery mechanism.

Basis: original synthetic structural counterexample, independently reproduced in the private annex. All numbers on this exhibit are hypothetical analytical quantities.

EXHIBIT 3 / MEASUREMENT EQUIVALENCE

Correct the denominator and then test whether the events are comparable

[HR] The author corrected a comparison that divided a process-detection rate by a customer-claim rate calculated on a different exposure base. [AI] This repairs one arithmetic error. It does not prove that the detected process events belong to the same customer-claim cohort or represent damage created by the process.

COMPARABILITY REQUIRES FOUR BRIDGES

Analytical validity chain / one repaired denominator does not establish event membership



Process detections / customer claims is a ratio of different constructs.
It becomes a share only after subset membership is established.

Basis: S1 M10/M11/M13/M20. Common-basis arithmetic is necessary but insufficient.

Validation layer	Required bridge
Exposure basis [RC/AI]	Express both rates on a documented common denominator for a like-for-like rate comparison. Preserve the original denominator and rounding interval privately; never substitute an inferred count for an observed exposure.
Event identity [AI/PD]	Join package identifiers and event definitions. Resolve repeat detections, multiple scans and claims before treating a process count as part of a customer-outcome total.
Cohort and maturity [AI/PD]	Align origin cohort, eligible population, inclusion rules and follow-up. The same calendar window can contain different package cohorts and differently matured outcomes.
Mechanism and actionability [AI/PD]	A detection location is not a verified damage-creation location. A booked category is not a count of preventable events. Estimate preventability through a defined action and comparison.

USE THE CORRECTION WITHOUT OVEREXTENDING IT

[AI] The corrected comparison is numerically smaller on the common exposure basis. The claim that recovery touches only that proportion of the damage problem remains unestablished because event overlap and causal origin are missing.

Basis: S1 M10/M11/M13/M20, D02/D05. Arithmetic correction is reported and recalculated; event-membership and preventability conclusions are analytical extensions.

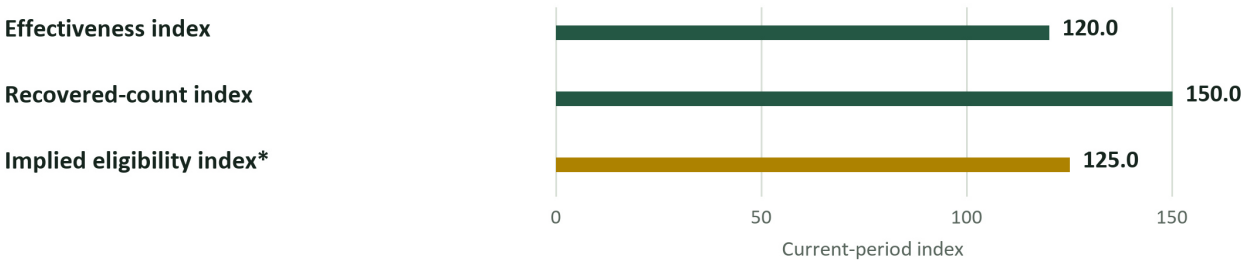
EXHIBIT 4 / SCORE AND WORKLOAD

Separate a changing recovery score from the population entering recovery

[HR] Recovered-unit count rose substantially more than the reported effectiveness rate. The source flags possible coverage or definition changes. [AI] A higher recovered count therefore cannot be read as pure process improvement or as a clean increase in treatment dose.

RECOVERED COUNT COMBINES EFFECTIVENESS AND ELIGIBLE VOLUME

Hypothetical normalized illustration / each prior-period quantity = 100; not project values



*150 = 120 × 125 / 100. The identity requires matching populations and definitions.

[AI] If recovered count N and effectiveness p refer to the same eligible population, then $N = pE$, where E is eligibility volume. Therefore $N_1/N_0 = (p_1/p_0)(E_1/E_0)$. The implied change in E is a conditional reconciliation, not an observed denominator.

Question before attribution	Evidence to retain
Did the process change? [PD]	Record the precise assistance, handling, routing or inspection action offered to each eligible episode. A measured score is an outcome of those actions and the case mix.
Did eligibility change? [PD]	Freeze an upstream eligibility definition before assignment. Audit reclassification, intake coverage, exception severity and excluded episodes; preserve unknown eligibility as a separate state.
Did outcomes improve? [PD]	Follow every assigned eligible episode through intact receipt, damaged receipt or unresolved/nonreceipt. Do not condition the primary outcome on successful recovery or delivery.

NO NUMERIC SUBSTITUTION FOR A MISSING FIELD

[AI] The conditional implied pool helps locate a reconciliation problem. It cannot supply an actual count denominator for a rate-control chart, repair a missing daily panel or identify which recovery actions changed.

Basis: S1 M01/M02 and definition-change warning. Figure uses hypothetical indices, each reference = 100; no actual score ratio or inferred eligibility level is disclosed.

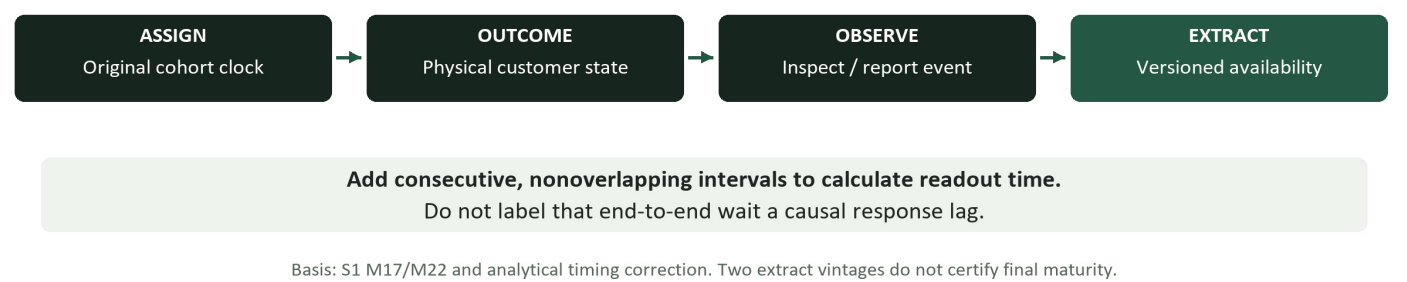
EXHIBIT 5 / EVENT TIME AND INFORMATION TIME

Keep the operating clock separate from the reporting clock

[HR] The investigation separated process timing from extract backfill and compared early and later readings of the same reporting weeks. [AI] That comparison describes a limited revision window. It does not measure all later customer claims or validate a universal input-to-outcome lag.

EVENT LATENCY AND INFORMATION DELAY CAN COMPOSE

Proposed timing contract / schematic sequence; spacing is not measured elapsed time



Clock	Decision discipline
Operating event time [PD]	Link original exception, assistance assignment, handling, delivery and physical condition. Pre-specify the exposure window; record actual events rather than assigning every unit an aggregate average lag.
Outcome observation time [PD]	Record customer reporting and independent inspection separately. Preserve late events and nonreceipt. Use a fixed original-promise follow-up clock so treatment cannot shift eligibility by delaying completion.
Extraction vintage [HR/PD]	Retain each extract date, cohort key, definition and revision. A stable value across two vintages is stable only within that observed interval.
Decision time [PD]	Use early reads for operational investigation. Release high-stakes benefit and safety conclusions only after the prespecified horizon and ascertainment checks are complete.

RETIRE THE AUTOMATIC CREDIT RULE

[AI/PD] Pairing a score today with a later damage rate does not identify the score’s causal contribution. Estimate the assigned package effect on original cohorts; use event-time diagnostics to explain the mechanism and reporting vintage to audit completeness.

Basis: S1 M17/M22 and timing narrative. Displayed maturity-pair arithmetic and quoted changes are reconciled privately; future customer-tail completeness remains unmeasured.

EXHIBIT 6 / SAFETY, VALUE AND STOPPING RULES

Require both customer protection and net value before expanding the package

[PD] The prospective decision concerns a versioned recovery-assistance package offered to upstream eligible episodes. Current approved practice is the comparator. A package can increase intact receipt and damaged receipt simultaneously, making a recovery score alone an inadequate decision criterion.

TWO DECISION GATES MUST PASS ON THE SAME ASSIGNED COHORT

Proposed rollout / randomized offer, mature customer states, fixed original eligibility

CUSTOMER HARM
Upper bound below agreed margin

AND

NET RESOURCE VALUE
Lower bound above zero

Maintain separate severe-incident and data-integrity safeguards.
Inconclusive evidence does not establish an acceptable harm bound.

Basis: proposed clustered evaluation and hypothetical decision model. No achieved field assurance.

Decision gate	Prespecified requirement
Customer harm [PD]	The upper confidence bound for additional mature damaged receipt must fall below a justified harm margin. Choose the margin before sizing the study; do not enlarge it to obtain a passing sample requirement.
Net resource value [PD]	The lower bound must exceed zero after valuing avoided unresolved/nonreceipt, additional damage and incremental work. Count every assigned eligible episode and prevent benefit overlap.
Other safeguards [PD]	Maintain severe-incident review, service-delay, safety and data-integrity controls. A powered damage/value model does not confer assurance on endpoints that have not been sized.
Inconclusive result [PD]	Retain a bounded status and obtain the missing information. Failure to establish excess harm is different from establishing that the package stays within the accepted margin.

REPLACE THE CONJUNCTIVE CORRELATION WATCH
[AI/PD] A trigger requiring several aggregate metrics to rise together can miss offsetting harm or a changing detection process. Allow credible incident evidence and ascertainment failures to trigger review independently of the aggregate score.

Basis: proposed evaluation and model, pages 19–22. S1 D07 supplies an uncalibrated historical watch rule; its execution is not established.

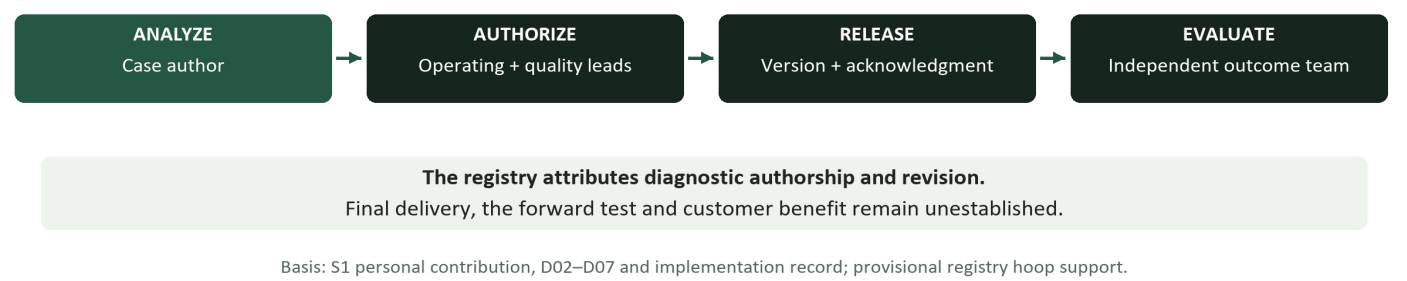
EXHIBIT 7 / OWNERSHIP AND DURABLE LEARNING

Preserve the author’s analytical contribution and verify the decision endpoint

[HR] The register attributes the analysis, response to reviewer objections, denominator correction and consolidation to the case author. It does not establish delivery of the final version, operation of the next-period test or ownership of the underlying recovery program.

AUTHORSHIP, RELEASE AND OPERATING EFFECT ARE DISTINCT

Historical contribution boundary + proposed ownership controls



Contribution or handoff	Evidence grade and next control
Diagnostic authorship [HR/AI]	Provisional registry hoop support. The record describes authorship and revisions; authenticate the original artifacts and metadata before strengthening the personal-contribution grade.
Metric and narrative revision [HR/AI]	Provisional registry hoop support. The record preserves a reviewer challenge and a corrected comparison. The unrecovered originals prevent independent confirmation of the revision sequence.
Program delivery and impact [HR/AI]	Unestablished. Authorship of the investigation does not establish control of the operating program, implementation of a later action plan or a causal customer benefit.
Release and follow-through [PD]	Freeze one approved artifact with its evidence version; record recipient acknowledgment, superseded versions, named test owner and scheduled decision date. A rendered report alone does not close the operating loop.

THE REUSABLE CAPABILITY

[AI] The case demonstrates a documented approach to challenging an executive hypothesis and correcting its own evidence. This portfolio review extends that approach by testing the causal verdict itself and replacing it with an identifiable decision design.

Basis: S1 master contribution, D02–D07 and implementation record. Grades are internal process-tracing assessments, not external certification or demonstrated outcome attribution.

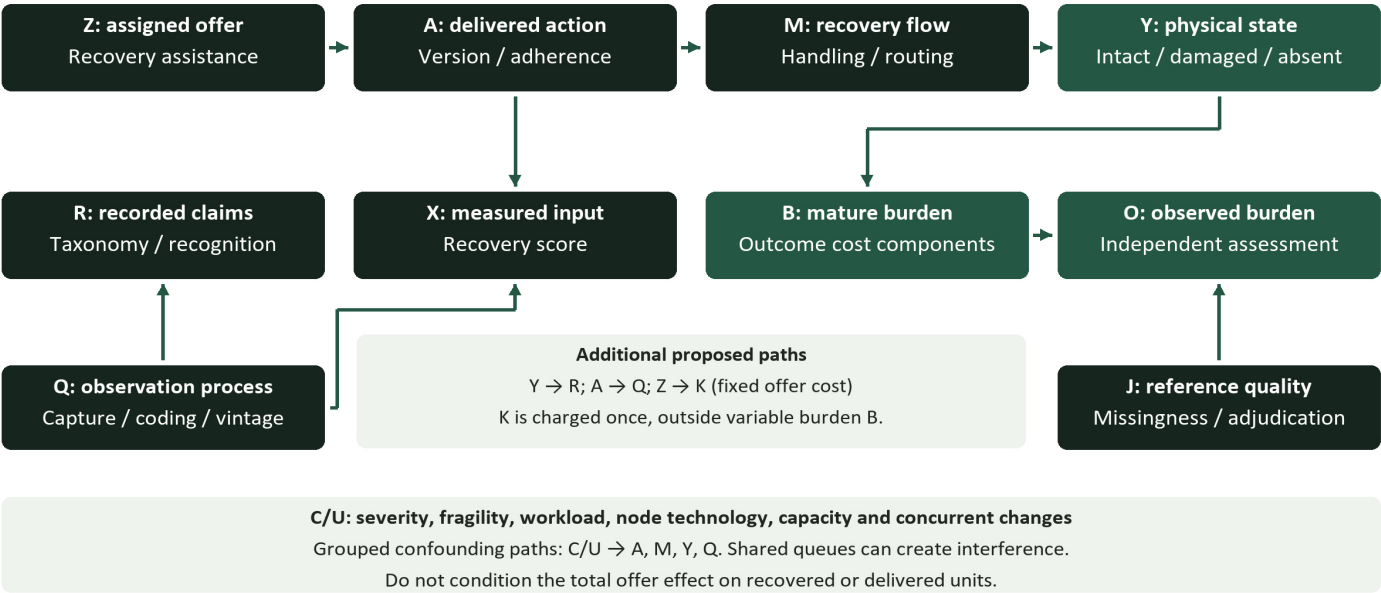
TECHNICAL ANALYSIS

Assign a defined action before interpreting a recovery score

[AI/PD] The proposed Temporal Evidence Contract (TEC) links an assigned recovery offer to physical customer states and independently measured burden. The historical score and claim series are observations of that system; they are not interchangeable with treatment and effect.

TEST THE ACTION; DO NOT TREAT THE SCORE AS THE ACTION

Proposed Temporal Evidence Contract / no historical causal edge is identified



Y denotes the underlying customer state; J affects its ascertainment, not physical damage.
X is a process measurement. Recorded claims R and independent outcome assessment O are distinct.

[PD] Z is the assigned offer; A is delivered assistance; M is recovery and handling flow. Physical state Y distinguishes intact receipt, damaged receipt and unresolved/nonreceipt. Y determines variable outcome burden B, independently observed as O through reference quality J. Fixed offer cost K is committed by Z and charged once outside B.

[AI/PD] A and observation process Q can change measured input X. Physical state and Q jointly affect recorded claims R. C/U names severity, fragility, workload, node technology, capacity and concurrent changes. Reporting shifts can therefore change a claim rate without an equivalent change in physical harm.

[PD] Estimate the total offer effect across the original eligible cohort. Conditioning on recovered or delivered units selects a post-assignment population. Shared queues and staffing can contaminate a comparison; randomize the domain containing those dependencies. These arrows are hypotheses, not historical effect estimates.

Basis: registered recovery/damage question [HR]; proposed TEC causal model [PD]. The separate synthetic counterexample disproves a universal inference rule, not this operating mechanism.

TECHNICAL ANALYSIS

A failed lag prediction rejects only the model that made it

[AI] The assertion that a real delayed mechanism must strengthen its observed lag correlation is false without restrictive assumptions. A positive action effect can coexist with negative delayed association.

Synthetic structural system [AI/RC]	Implication
$X_t = U_t + e_t$	U, e and v are mutually independent over time, centered and unit variance.
$Y_t = 2U_t + X_{(t-1)} - 3U_{(t-1)} + v_t$	Holding the disturbances fixed, increasing X_t by one increases $Y_{(t+1)}$ by one.
$Var(X)=2; Var(Y)=10$	Correlations at lags zero, one and two are +0.447, -0.224 and zero.

[AI/RC] The same-time covariance is 2. The next-period covariance is $Var(X_t)-3Cov(X_t,U_t)=-1$. The delayed common-driver path masks the positive causal coefficient. No sampling error is needed. This constructed process is not a probability model of damaged parcels and is not fitted to the source.

[AI] More generally, $Y_t = \sum \beta_j X_{(t-j)} + g(C_t,C_{(t-1)},...) + \text{error}$ implies a cross-covariance that combines lag coefficients, exposure autocovariance and common-cause covariance. Its largest value need not occur at the physical response delay; its sign need not match a particular β_j .

[AI] Granger’s original framework concerns incremental predictive information conditional on a specified information set. A short bivariate correlation scan is not that test. Even a predictive finding requires further assumptions to support an intervention claim. Distributed-lag methods describe cumulative and delayed associations; they do not remove confounding by themselves.

[HR/AI] The source retained an inconvenient same-time association and a later sign change. That disclosure is useful. The inference “not identified by these diagnostics” is defensible; “no causal harm” is not. Several correlated diagnostic lenses do not create independent replications or a calibrated probability of safety.

[PD] Specify the intervention, population, causal alternatives, lag restrictions and detectable harm before testing. A restricted mechanism can be contradicted by precise incompatible evidence. A wide interval, low power or an unadjusted sign change leaves competing explanations open.

Basis: exact synthetic counterexample [AI/RC]; Granger1969; Gasparrini et al.2010; Greenland et al.2016. No historical coefficients are re-estimated without the paired panel.

TECHNICAL ANALYSIS

Separate the clocks, then compose only nonoverlapping intervals

[AI] Operating response, observation and extract revision answer different questions. They can occur sequentially, but their sum is a readout delay rather than a physical causal lag.

Clock or contrast	Required definition [PD]
Action to physical outcome	Original eligibility, assigned action, event timestamp and physical condition on the same unit.
Outcome to observation	Independent inspection, customer report or scan; preserve never-observed and unresolved cases.
Observation to extraction	Recorded event key, processing time, extract version and subsequent corrections.
Decision readout	Original-promise horizon plus the prespecified observation and availability requirements.

[AI] If the physical outcome occurs two periods after action and its observation becomes available two periods later, the decision can read it four periods after action. That does not make the causal response four periods. Conversely, do not add a posting delay twice when an observed operational interval already includes it.

[HR/AI] The source compares early and later extract vintages of a few reporting weeks. Those pairs measure revision over that interval only. Small backfill does not establish ultimate customer-claim completeness, physical truth or a common maturity curve across origin categories.

[AI] A service policy describes intended timing; it does not measure the distribution of actual latency. Median availability is insufficient when the decision concerns a late tail. Reclassification can also move cases between buckets without adding a new physical event. Preserve additions, removals and category migrations separately.

[PD] Build an event-by-report-delay triangle from matched units and frozen cohort definitions. Distinguish eventual reporting from never-recorded damage; a nowcast of the former does not identify the latter. Validate an independent late follow-up sample before asserting completeness.

[PD] Release early readings as investigation signals. Release causal credit or safety claims only after the declared original-cohort horizon and reference checks. The prospective model’s horizon is a design assumption, not a maturity fact imported from two historical extracts.

Basis: registered two-vintage probe [HR]; temporal accounting [AI]; Bastos et al.2019. The source’s automatic timing-credit rule is not adopted as a causal estimator.

TECHNICAL ANALYSIS

Validate event membership after correcting the rate denominator

[HR/AI] The registered denominator correction improves arithmetic comparability. It does not make process detections a subset of customer claims or identify where physical damage was created.

Construct	Measurement contract [PD]
Recovery effectiveness	Resolved eligible episodes / all eligible episodes under one versioned definition; audit exclusions and reclassification.
Process damage detection	Unique linked physical unit, detection time and reason; collapse repeated scans only under the correct entity rule.
Recorded damage claim	Claim category, recognition time and customer-unit link; keep other claim buckets and unresolved mappings.
Physical customer state	Independent condition and receipt evidence on every original eligible unit, with missing-truth bounds.

[AI] A ratio between counts becomes a share only after the numerator is verified as a subset of the denominator on the same unit, population and horizon. A detection total may include duplicates, diverted units or events never conceded. A small early-booked fraction is not a bound on all actionable damage.

[AI] Calling early recognition a lower bound on preventable opportunity additionally requires true-positive classification and a feasible action that would change the outcome. False positives, different cohorts or unpreventable events break that implication. Neither an assumed fraction of late claims nor a reporting-node label repairs the missing evidence.

[AI/PD] If an event truly precedes recovery, it cannot be damage created by later rehandling. It can describe baseline condition. Calling it a negative control requires verified timing and no effect of treatment on its capture or classification; those conditions are not established by its name.

[PD] Audit a random core of 240 linked exception histories plus prespecified weighted rare cases. Two blinded coders independently assess eligibility, physical condition, event order and claim linkage. Report confusion matrices, prevalence, raw agreement and kappa intervals. The iid raw-agreement screen is about ±6.3 percentage points; cluster dependence and kappa precision require pilot simulation.

[PD] Test measurement invariance across node and business line. Reference evidence must not merely repeat the same scan used by the process score. Unknown outcomes remain explicit; an empty field is not a verified negative.

Basis: registered detection/claim mismatch [HR]; construct and nesting analysis [AI]; proposed blinded audit [PD]; Cohen1960. No count denominator is fabricated from a rounded rate.

TECHNICAL ANALYSIS

Define harm and value on every originally eligible exception

[AI/PD] The outcome of a recovery offer can include both more intact receipt and more damaged receipt. The decision needs the joint customer distribution, not a score conditional on successful recovery.

Estimand	Population, sign and boundary [PD]
Descriptive lag association	Correlation of explicitly matched observed variables at fixed lags and one frozen window; not an intervention effect.
Damage ITT increase	Risk under assigned offer minus risk under usual approved practice, per original eligible exception at the fixed horizon.
Resource-benefit ITT	Usual-minus-offer mature cost-weighted burden over the same original units, before separately charged fixed offer cost.
Delivered-action effect	A different causal quantity requiring identification of actual execution; conditioning on completed recovery does not estimate it.

[AI] For a matched definition, recovered count N equals effectiveness p times eligible volume E. Changes in N combine changes in p and E. The source-relative reconciliation can suggest an eligibility or coverage change; it cannot supply the missing observed denominator or a clean treatment dose.

[PD] Freeze eligibility upstream of action and retain refusal, hold, reroute, nonexecution, nonreceipt and late observation. Define disjoint final states and prioritize adjudication rules before labels are seen. Do not classify unknown physical condition as either harmless or damaged solely to complete a table.

[PD] The working model distinguishes intact receipt, damaged receipt and unresolved/nonreceipt. It makes a positive damage increment visible rather than hiding it inside a larger recovery gain. A noninferiority margin represents tolerated excess risk; it requires independent Quality approval and is not evidence of zero harm.

[PD] Price each outcome component once on the same horizon. Preserve the negative value of additional damage when valuing avoided nonreceipt. Discount only genuinely avoidable, nonoverlapping benefit; charge additional harm and work fully unless a separately justified accounting contract says otherwise.

[PD] Report standardized business-line and node-type contrasts with original exposure weights. The safety margin and service obligation may differ by fragility or customer consequence. A pooled favorable mean cannot erase an adverse subgroup or a severe incident.

Basis: descriptive metric identity [AI]; proposed original-cohort estimands [PD]; surrogate-outcome literature. No historical ITT, delivered-action effect or cash benefit is identified.

TECHNICAL ANALYSIS

Treat sparse movements and selected lag results as diagnostics

[HR/AI] The register contains short, changing observation windows and multiple related diagnostics. Its reported directions can guide investigation; they cannot establish a stable process or an identified mechanism.

Variation and multiplicity

[PD] For compatible ordered observations, compute moving ranges $MR_t = |x_t - x_{(t-1)}|$, individual limits $\text{mean}(x) \pm 2.66 \times \text{mean}(MR)$, and MR upper limit $3.267 \times \text{mean}(MR)$. Mixed vintage, incompatible populations, serial dependence and few points can make the computed bands non-diagnostic. A zero moving range is not proof of stability.

[HR/RC] The private audit contains 39 XmR screens: 19 ordered context series, one exact-rate derivation, three cohort snapshots, four vintage accumulations, two sparse endpoints, three model profiles, two denominator comparisons and five rounded-prior reconstructions. No point or moving range crosses its provisional band; no stable operating phase is established. A separate two-period Laney calculation assumes unique binary outcomes; one moving range and unauthenticated process definitions leave zero eligible Laney panels.

[AI] Lagging short panels drops different boundary observations. The shorter and longer reported windows therefore need separate frozen analyses. Controlling for volume changes residual degrees of freedom and still leaves other common causes. Predeclare candidate lags, outcomes and cumulative contrasts; report the full family rather than the most favorable result.

Selection and heterogeneity

[AI/PD] The second network is a directional cross-check, not an exchangeable untreated control. Severity, fragility, automation, shared staffing and observation quality can modify both action delivery and outcomes. Estimate subgroup interactions with uncertainty; do not infer a mechanism from opposite aggregate trends alone.

[PD] No measured worst-N treatment cohort or repeat-ranking reliability is supplied here. If future work selects a worst-decile normal score, its expected initial excess is 1.755 SD. Assumed repeat reliabilities 0.25/0.50/0.75 imply untreated reversion of 1.32/0.88/0.44 SD. This sensitivity is not an estimated historical correction.

[PD] Keep selected and nonselected eligible units, pre-action measures and repeated untreated ranks. Register phase changes and exposure weights before seeing results. Neither within-band movement nor a significance failure should be narrated as demonstrated improvement, deterioration or absence of harm.

Basis: source window and selection limits [HR/AI]; proposed XmR, selection sensitivity and moderation [PD]; NIST; Laney; Barnett. Final source inventory and exact raw limits remain private.

TECHNICAL ANALYSIS

Credit the analytical revisions without certifying the final causal verdict

[HR/AI] The register attributes question framing, analysis, reviewer response, denominator correction and consolidation to the case author. It also records unresolved delivery and follow-through. Those are distinct contribution claims.

Claim	Process-tracing grade and limit [HR/AI]
Diagnostic authorship	Provisional registry hoop support for framing, analysis and comparison. Authenticate originals before strengthening the attribution.
Response to challenge	Provisional hoop support for revising the metric basis and retaining contrary signals. Revision does not certify causal validity.
Consolidated release	A rendered and designated master artifact is recorded; receipt and accepted decision remain unestablished.
Operating or customer effect	Unestablished. Program ownership, downstream action delivery and mature causal outcomes are not implied by report authorship.

[AI] Several agents or diagnostic lenses reading overlapping source material do not provide independent empirical replications. The source’s confidence score is a workflow judgment, not a calibrated posterior or frequentist safety bound. Repeated summaries cannot upgrade provisional provenance to a smoking-gun test.

Consolidated limits and reproducibility

[HR/AI] Original-source authentication, paired daily observations, contemporaneous action exposure and mature unit-level outcomes remain incomplete. Measurement definitions, window choices, category mapping and source vintages differ. No identified comparison separates recovery effects from case mix, calendar, workload or reporting.

[AI] The source’s causal rejection, automatic credit rule and opportunity-floor language exceed what these diagnostics establish. Correcting those interpretations preserves the useful diagnostic contribution; it does not identify the opposite effect. Missing follow-through establishes neither completion nor failure outside the supplied corpus.

[PD] Freeze source/version hashes, a claim ledger, raw/indexed calculations, entity and timing rules, source conflicts, independent truth audits and analysis code. Preserve the original report and its corrections. Public artifacts retain anonymous conclusions; raw values, identities, original paths and transformation keys remain private.

[PD] Distinguish trial design from achieved assurance. Recruitment, cluster isolation, maturity, variance and costs require pilot evidence. A simulation demonstrates behavior under its stated law; it does not authenticate the field or validate an unsized safety endpoint.

Basis: registered authorship/revisions and delivery gaps [HR]; internal process-tracing assessment [AI]. Supported case: evidence reconciliation and proposed evaluation, not demonstrated impact.

TECHNICAL ANALYSIS

Power the customer-harm gate as well as the financial contrast

[PD] Propose matched independent custody domains randomized to a fixed recovery-assistance offer or unchanged approved practice. Randomize the domain containing shared staff, capacity and routing; preserve original eligibility and the same outcome reference process.

[PD] The illustrative design has 200 total eligible exception units per domain, divided into eight cohorts of 25. Follow each at its original promise plus 35 days; validate that horizon before the field test. Retain all original units and use predefined bounds for unresolved physical truth. Equal cohort sizes align domain and unit weights; unequal volumes require new weights and power.

[PD] Prespecify two gates: an upper confidence bound for added damaged receipt below a Quality-approved margin, and a lower bound for net resource benefit above zero. The model's illustrative margin is not established safe tolerance. Do not enlarge it to meet a recruitment budget. A positive assumed damage increase must remain visible.

[PD] Derive paired covariance from disjoint outcome categories and a declared domain-risk mixture. Match on pretreatment severity; model arm dependence from matching separately from within-domain clustering. Pilot those quantities. Use the actual effect-boundary outcome law when testing power, rather than only a large favorable alternative.

[PD] Compare the zero-effect MDE with economic break-even on the same original-unit pool. Also calculate joint probability of satisfying damage and value criteria; passing economic detection alone is insufficient. A conservative working-distribution envelope and direct finite simulations support conditional planning, not exact field power.

[PD] Account for avoided nonreceipt, added damaged receipt and fixed incremental handling/support expense separately. Operational verification belongs in incremental cost; research-only adjudication is separate. Independent information-value analysis purchases better evidence about the next decision and must not be added to the same avoided customer loss.

[PD] Maintain separately calibrated severe-incident, delay, safety and reference-integrity safeguards. A correlation threshold cannot veto credible incident evidence. If cluster independence, recruitment or outcome capture fail, redesign and re-power rather than substituting more correlated episodes.

[PD] For a later staggered rollout, estimate group-time ATT against not-yet-treated domains with no anticipation, common support and interference controls. Baseline/event-time diagnostics and equivalence checks inform assumptions; nonsignificant pretrends do not prove them. This separate design has no inherited power guarantee.

Basis: proposed paired cluster trial and synthetic model, pages19–22 [PD]; Schuirmann; Hudgens–Halloran; Callaway–Sant’Anna. No trial, safe margin or field funding result is reported.

TECHNICAL ANALYSIS

Turn timing discipline into four testable operating propositions

[AI/PD] TEC extends the source’s useful metric and clock corrections by testing the inferential rules themselves. It separates a diagnostic pattern from an intervention decision and makes release conditions observable.

Proposition [PD]	Evidence that challenges it [PD]
P1: Versioned event and availability clocks reduce decision errors.	A blinded replay showing no prespecified useful reduction in direction, eligibility or maturity errors challenges the benefit of the contract.
P2: The defined recovery offer creates net value within the agreed harm limit.	An upper net-benefit bound below zero or a lower harm bound above tolerance contradicts the claim; wide intervals remain inconclusive.
P3: Linked mechanism measures predict effect location better than booking labels.	Prespecified out-of-sample comparisons showing no useful calibration or discrimination gain challenge this proposed diagnostic advantage.
P4: Governed release and retesting improve correction uptake and later decisions.	A controlled second-cycle comparison showing no useful gain over document completion challenges the learning mechanism.

[AI] Granger makes the information set and observation frequency part of the inference problem. Distributed-lag research accommodates multiple response times and cumulative effects. The synthetic counterexample extends this case by demonstrating why a sign-reversing correlation cannot serve as a universal causal falsifier.

[AI] Surrogate research explains why a favorable process score and its association with an outcome do not determine the treatment’s customer effect. Equivalence and noninferiority require explicit bounds; failure to reject a zero-effect test cannot supply them. Reporting-delay methods also distinguish eventual observation from physical events that never enter the record.

[AI] Deming’s variation discipline cautions against reacting to unqualified movement. Campbell/Goodhart motivate preserving the original cohort so score pressure cannot redefine success. Ittner and composite-indicator guidance require explicit component weights and separate harm reporting rather than letting a favorable recovery total conceal damage.

[AI/PD] Tucker and Edmondson distinguish immediate workarounds from learning that changes recurrent failures. A corrected report becomes an operating capability only when the approved version reaches its owner and a subsequent test informs action. The register does not establish that final loop.

[PD] Preregister contrasts, harm margins and revision rules. Publish unsupported mechanisms as unresolved or contradicted only within the tested assumptions. Do not use an attractive narrative, an internal confidence score or an economic payoff to override missing authority or physical truth.

Basis: proposed TEC propositions [PD]; Granger; Gasparrini; VanderWeele; Bastos; Schuirmann; Deming; Campbell; Ittner; OECD; Tucker–Edmondson. None is historically confirmed.

PROSPECTIVE MODEL / RECOVERY AND HARM

Test a recovery package against a damage margin, not against a correlation

[PD] Every number on these four pages is hypothetical. Randomize a fixed recovery-assistance package within matched independent custody-domain pairs; preserve current approved practice in the comparator. A process score is neither assignment nor a treatment version.

Contract	Prespecified planning choice
Cohort and clock	200 original eligible exception units per domain: eight cohorts of 25. Independent outcome adjudication at original promise + 35 days; calibrate that horizon before a field trial.
Whole-policy ITT	Retain sideline, failed recovery, withdrawal and nonreceipt. Equal frozen pools align domain and unit means; changed weights require new power.
Reference truth	Intact receipt, damaged receipt and unresolved nonreceipt are exclusive terminal states. Unknown truth remains bounded, not a clean receipt. Recording labels do not establish physical damage.

[PD] Low/high-risk usual nonreceipt probabilities are 0.18/0.26; damaged-receipt probabilities are 0.06/0.10. The package adds 0.08 receipt probability; 2.5% of those extra receipts are damaged. Thus damage rises by 0.20 percentage points in the assumed beneficial scenario.

A TOLERATED BOUND IS NOT PROOF OF NO HARM

[PD] A candidate +1.00-point absolute damage margin requires Quality approval before enrollment. It is not an established tolerance. More recovery and positive net value cannot substitute for this safety decision.

[PD] Domain-pair grades have correlation 0.8, with 0.2 stress. Conditional parcel outcomes are multinomial; pair types induce clustering. Delay, unsafe dispatch and missing-truth safeguards require separate sizing.

PROSPECTIVE MODEL / JOINT ACCEPTANCE AND POWER

Size the damage bound and the economic decision together

[PD] Exact paired categorical moments feed a known-law Gaussian planning approximation. It is not an exact finite-sample guarantee. Both gates must clear: upper 97.5% damage bound below the approved margin and lower 97.5% variable-benefit bound above cost.

Design	Pairs	MDE (index)	Joint-power floor	Result
Initial	20	12.35	8.2%	Fails
Base candidate	140	4.68	81.5%	Conditional pass
Same size / stress	140	6.88	65.1%	Fails
Revised / stress	240	5.25	88.0%	Conditional pass

[PD] The initial study detects economic break-even with 88.4% power, yet its joint floor is only 8.25%. Economic precision alone cannot answer the harm question.

[PD] Cost is 13.89 index points. Also require 80% two-sided 5% detection power at that break-even. Base minimum: 136 pairs; weaker matching: 195. The reported joint floor is one minus the sum of gate-failure probabilities, not two separate 80% claims.

[PD] Economic variance is bounded over the declared recovery-effect path. Damage variance is bounded over all damaged fractions of the extra recoveries. Simulations regenerate categorical outcomes at economic and safety boundaries using their own covariance. Pilot-estimated uncertainty and studentized or randomization inference need new calibration.

RECRUITMENT AND APPROVAL REMAIN OPEN

[PD] The revised candidate requires 480 independent domains and 96,000 original exception episodes. Neither availability nor independent capacity is established. Start with calibration; do not enlarge the harm margin to rescue an infeasible study.

[PD/RC] MDE and cost use index points; usual harm-resource burden=100. Power is conditional, not achieved. A smaller Quality margin can require far more domains. The two modeled gates do not size every safety safeguard.

PROSPECTIVE MODEL / FULL CONSEQUENCE ECONOMICS

Keep the cost of added damaged receipts in the business case

[PD] Nonreceipt and damaged receipt cost 80 and 50 hypothetical resource units. Usual burden is 21.6 per original episode, indexed to 100. Added recovery avoids 6.4 units of nonreceipt cost but adds 0.1 of damage cost; neither amount is historical.

[PD] Fixed incremental cost $K_{raw}=3$ per original episode allocates handling 1.2, capacity/support 0.8 and production verification 1.0 over the same prototype service pool. Research-only adjudication is separate. This is not an observed expense or a scaled operating forecast.

Avoidable / non-overlap	$K_{index}=9.26$	$K_{index}=13.89$	$K_{index}=18.52$
0.50 / 0.80	+2.13	-2.50	-7.13
0.75 / 0.90	+10.28	+5.65	+1.02
1.00 / 1.00	+19.91	+15.28	+10.65

[PD] $Net_{raw} = a \cdot h \cdot 80 \cdot \text{nonreceipt-ITT reduction} - 50 \cdot \text{damage-ITT increase} - K_{raw}$. $Net_{index}=100 \cdot Net_{raw}/21.6$. Discount avoided nonreceipt only; retain added damage and new operating expense in full. If severity or outcome prices change, reprice the full vector and recompute power.

POSITIVE VALUE CAN COEXIST WITH HARM

[PD] The base scenario has positive value and a real damage increase below the proposed margin. A favorable economic test does not prove zero harm. Field valuation requires eligible original volume, mature signed outcome contrasts and local avoidable costs; those inputs are not established.

[PD/RC] Grid shows net index points, not money. Economics and noninferiority share the original eligible population and maturity window. Delay, diversion and unsafe-content costs must be added before implementation if present.

PROSPECTIVE MODEL / EVIDENCE BEFORE ESCALATION

Buy evidence that distinguishes a harmful package from an unresolved measurement

[PD] A separate three-state decision model compares a controlled recovery trial, safeguard redesign and a hold while repairing the decision basis. It values choosing the right next action even when the operating package is harmful or its effect is unidentified.

Action / utility	Acceptable + valuable	Harm exceeds margin	Other / unresolved
Controlled trial	14	-18	-7
Redesign safeguards	5	12	-5
Hold / repair basis	-3	-4	11

[PD] Prior probabilities are 0.45/0.25/0.30. Excess harm takes priority; acceptable and valuable requires valid evidence. All other states include nonpositive value or an unresolved basis. Signals identify the correct state with probability q and either alternative with (1-q)/2.

Signal quality q	EVSI	Evidence cost	Net EVSI
0.50	0.5000	0.40	0.1000
0.75	4.6750	1.30	3.3750
0.90	7.1800	2.40	4.7800

INFORMATION CANNOT GRANT OPERATING AUTHORITY

[PD] Exhaustive enumeration of 27 signal-action policies gives prior value 3.75 and perfect-information value 8.85 above that baseline. Utilities already include subsequent action expense. Subtract acquisition cost once; do not add this utility to operating savings or override a safety hold.

[PD/RC] All priors, likelihoods, utilities and costs are hypothetical. EVSI is bounded by EVPI within this decision model; no historical decision saving or delivered customer benefit is asserted.

Reproduce the interpretation before accepting the operating verdict

[AI] This case combines a documented diagnostic, a new critique of its causal logic and a prospective decision design. Its contribution is stronger measurement and decision discipline; it does not establish that the historical recovery program caused, prevented or avoided harm.

Claim class	Meaning and replication boundary
[HR] Historical record	A statement, action or value attributed to the selected register. Transcription checks verify the register extraction; they do not authenticate an unavailable underlying artifact.
[RC] Recalculated	Source-relative arithmetic, source-conflict reconciliation, tentative process limits and explicit hypothetical model calculations. A correct calculation can still compare different constructs.
[AI] Analytical inference	A conclusion tied to stated premises: a failed universal lag rule, incomplete event equivalence, conditional eligible-pool expansion or a bounded interpretation of extract maturity.
[PD] Proposed design	The Temporal Evidence Contract, randomized recovery-assistance test, harm margin, measurement audit, economic model and information-value analysis. Their parameters require local validation.

Reproduction sequence

[PD] Check the source hash and row/cell crosswalk. Reconcile count, rate, denominator, population, window and vintage. Preserve contradictory versions and missing fields; do not manufacture a daily panel from reported correlation coefficients.

[PD] Expand the synthetic structural equations and recover their covariances independently. Rebuild the prospective cluster law, mature outcome probabilities and economic thresholds. Recompute the harm and value gates, sample-size search, sensitivity cases and separate information-value model.

THE CORE CHANGE IN INTERPRETATION

[AI] An executive hypothesis can remain unsupported without being disproved. The decision becomes more defensible when the operating action, mature customer outcome, tolerated harm and required evidence are specified separately.

Public source metrics are relative or indexed. Raw values, identity mappings, source paths, source-recovery status and all executable calculations are retained in the separate private package.

RESEARCH REFERENCES / CAUSAL AND TEMPORAL INTERPRETATION

Distinguish observed temporal patterns from intervention effects

Primary research and official methods support the interpretation and proposed evaluation. They neither authenticate missing project originals nor supply historical customer effects.

Granger, C. W. J. (1969). Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica*, 37(3), 424–438. DOI 10.2307/1912791. Full original reproduction. Predictive direction depends on the information set and observation frequency. The project's pairwise correlation scan is not an identified intervention effect or a completed Granger test.

Gasparrini, A., Armstrong, B., & Kenward, M. G. (2010). Distributed Lag Non-linear Models. *Statistics in Medicine*, 29, 2224–2234. DOI 10.1002/sim.3940. Full primary manuscript. Lag-response shape and cumulative effects require explicit modeling. This case does not fit a distributed-lag model to the absent daily panel.

VanderWeele, T. J. (2013). Surrogate Measures and Consistent Surrogates. *Biometrics*, 69, 561–569. DOI 10.1111/biom.12071. Indexed primary manuscript; direct opening restricted. A favorable intermediate measure need not imply a favorable final outcome. The proposed test measures recovery, damaged receipt and nonreceipt separately.

Bastos, L. S., et al. (2019). A Modelling Approach for Correcting Reporting Delays in Disease Surveillance Data. *Statistics in Medicine*, 38, 4363–4377. DOI 10.1002/sim.8303. Full author manuscript. Reporting-delay methods concern events eventually observed; they do not recover never-reported physical truth without additional assumptions. The case applies that distinction to claim maturity.

Greenland, S., et al. (2016). Statistical Tests, P Values, Confidence Intervals, and Power: A Guide to Misinterpretations. *European Journal of Epidemiology*, 31, 337–350. DOI 10.1007/s10654-016-0149-3. Primary essay and indexed manuscript. Non-significance is not proof of a null effect; a confidence interval is not a probability assigned to the hypothesis.

Schirmann, D. J. (1987). A Comparison of the Two One-Sided Tests Procedure and the Power Approach for Assessing the Equivalence of Average Bioavailability. *Journal of Pharmacokinetics and Biopharmaceutics*, 15, 657–680. DOI 10.1007/BF01068419. Primary abstract. The case borrows the need for a prespecified margin; it uses a one-sided harm gate, not a clinical equivalence claim or an externally validated operating tolerance.

Access scope and citation metadata are recorded privately. Continuing references support established methods; no external effect size is imported.

RESEARCH REFERENCES / IDENTIFICATION AND UNCERTAINTY

Make variation, comparison and information value explicit

Primary research and official methods support the interpretation and proposed evaluation. They neither authenticate missing project originals nor supply historical customer effects.

NIST/SEMATECH. e-Handbook of Statistical Methods, section 6.3.2.2: Individuals Control Charts. Official method. Tentative moving-range limits are computed where the source provides usable sequences. A two-point calculation, extract revision or lag ladder does not establish a stable operating process.

Laney, D. B. (2002). Improved Control Charts for Attributes. *Quality Engineering*, 14(4), 531–537. Publisher abstract. Extra-binomial dispersion requires actual aligned counts and exposures. Rounded rates and implied eligibility do not provide a valid repeated-rate panel or authorize invented denominators.

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20, 37–46. Primary publisher record. The proposed blinded audit reports agreement, a confusion matrix and kappa with uncertainty. Agreement between coders does not by itself establish correct physical condition.

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. Primary abstract. Selecting extreme stations or windows can produce apparent recovery without an intervention. This case uses explicit sensitivity assumptions where untreated repeat observations are absent.

Hudgens, M. G., & Halloran, M. E. (2008). Toward Causal Inference With Interference. *Journal of the American Statistical Association*, 103(482), 832–842. Primary manuscript. Shared custody, labor and routing can expose comparison units to the same action. The proposed assignment boundary must contain or explicitly model these spillovers.

Callaway, B., & Sant'Anna, P. H. C. (2021). Difference-in-Differences with Multiple Time Periods. *Journal of Econometrics*, 225, 200–230. Author manuscript. Group-time effects inform a staggered observational alternative. That alternative requires its own parallel-trends assessment, overlap, treatment timing and power analysis.

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*, 35, 570–583. Author repository. This case uses a separate finite-state information model and does not add its decision utility to operating benefit.

Access scope and citation metadata are recorded privately. Continuing references support established methods; no external effect size is imported.

RESEARCH REFERENCES / OPERATING MEASUREMENT AND LEARNING

Connect performance indicators to accountable operating action

Primary research and official methods support the interpretation and proposed evaluation. They neither authenticate missing project originals nor supply historical customer effects.

Collier, D. (2011). Understanding Process Tracing. *PS: Political Science & Politics*, 44, 823–830. Primary publication. Authorship, revision, delivery and causal contribution are distinct claims. Registry references to metadata remain provisional until the underlying artifacts are authenticated.

Tucker, A. L., & Edmondson, A. C. (2003). Why Hospitals Don't Learn from Failures. *California Management Review*, 45(2), 55–72. Author manuscript. Local recovery and system learning are different mechanisms. A higher recovery score does not establish that upstream failure causes have been removed.

The W. Edwards Deming Institute. The Funnel Experiment. Official exposition. Reacting to routine variation can worsen a process. Here the short and changing observation bases prevent a stable-process conclusion; the case does not estimate a historical tampering effect.

Campbell, D. T. (1976). Assessing the Impact of Planned Social Change. Original institutional record. Consequential indicators can distort behavior. Fixing eligibility upstream and retaining unresolved cases reduces the opportunity to obtain a favorable score by changing the measured population. No intentional manipulation is alleged.

Goodhart, C. A. E. Problems of Monetary Management: The U.K. Experience. 1975 conference; revised volume listed as 1976 in the RBA archive. The measurement critique motivates separate process, service and customer-condition measures. The case does not treat a target score as a validated causal surrogate.

Ittner, C. D., Larcker, D. F., & Randall, T. (2003). Performance Implications of Strategic Performance Measurement in Financial Services Firms. *Accounting, Organizations and Society*, 28, 715–741. Author manuscript. The value of a measurement system depends on its relationship to decisions and outcomes. More metrics or a formal review alone do not identify an operating effect.

OECD / European Commission JRC (2008). Handbook on Constructing Composite Indicators. Official handbook. Weighting and compensation require explicit choices. The case preserves a separate customer-harm gate so a favorable monetary aggregate cannot compensate for an unacceptable damage increase.

Access scope and citation metadata are recorded privately. Continuing references support established methods; no external effect size is imported.

FINAL REFEREE REPORT / INTERNAL AGENT-ASSISTED REVIEW

Accept the case within its evidence and identification boundaries

[AI] This is author-led, agent-assisted review. It is not external peer review, client approval or a certification of consulting-firm rank. The supported contribution is an evidence-reconciliation and prospective evaluation case.

Major issue	Disposition in the revised case
Causal rejection	Replaced the universal lag-strengthening rule with an explicit synthetic counterexample. The historical diagnosis remains documented; no-harm and harmful-effect claims remain unidentified.
Measurement equivalence	Separated rate-denominator correction from event membership, cohort alignment and physical origin. Conditional inferred eligibility is not treated as an observed count.
Timing and maturity	Separated operating events, outcome observation and extraction vintage. Consecutive intervals can compose into readout time without becoming a causal response lag. Limited backfill does not establish finality.
Evaluation and economics	Defined an offered action package, original eligible cohorts, clustered assignment, mature customer states and separate harm/value gates. Power is conditional on hypothetical assumptions; severe-event and data-quality safeguards remain separate.
Contribution and evidence	Bounded authorship and revision claims with provisional registry hoop support. Delivery, the forward test, program ownership and causal outcomes remain unestablished. Independent calculations and actual page review are retained privately.

Partner critique: can a COO act on page one?

[AI] Yes: preserve approved safeguards, retain the harm question, and commission a bounded test of a specified package. Quality sets the acceptable margin before sizing; the sponsor owns scope, and the evaluator owns mature outcome verification.

Minor revisions and remaining limits

[AI] Source confidence scores are not calibrated probabilities. Model values, recruitment, covariance and ascertainment require calibration. The case retains an unresolved-effect conclusion rather than converting review completion into proof of safety.

REVIEW DISPOSITION / PASS WITHIN STATED SCOPE

[AI] The document supports source reconciliation, a new causal-logic finding and a reproducible prospective decision design. It establishes no realized saving, historical no-harm effect or implemented expansion.

Scope: internal scientific, evidence, numerical and executive review. Critical limitations constrain the verdict; passing a design case does not certify a historical impact.