

OPERATING CASE STUDY / CUSTOMER EXPERIENCE & LOGISTICS

Arrival Reliability & Customer Exposure

Turn a timing correlation into a decision that can withstand a causal test

An anonymized operating case on sizing concession exposure around inbound arrival cutoffs, reconciling package and unit measures, and testing a proposed upstream explanation. It separates reported exposure, causal effect and preventable customer harm.

PROJECT CONTEXT / INBOUND TIMING, DATA INTEGRITY & DECISION DESIGN

BLUF / BOTTOM LINE UP FRONT

Freeze the clocks and original cohort.
Test prevention before assigning value.

ANALYTICAL FINDING

Agreement between two share calculations can be an algebraic identity. A lower unadjusted concession rate can also coexist with higher risk in every operating stratum. Neither result, alone, validates the causal story.

CONTRIBUTION IN THE SOURCE RECORD

Scoped the analysis, commissioned partner-built data, reconciled measures and challenged the framing. Provisional registry hoop support.

EVIDENCE STATUS

Originals unrecovered. The reply is recorded as a draft; complete linkage, causal effects and realized financial benefit remain unestablished.

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

EXECUTIVE ABSTRACT / DECISION AND OWNERSHIP

Validate the exposure before investing in a timing intervention

[AI/PD] The proposed Timing-to-Outcome Contract (TOC) fixes the clocks, original cohort, data grain and mature outcome before interpreting exposure or testing an intervention. A reported timing association is an input to investigation, not a measured preventable effect.

Abstract element	Decision-relevant synthesis
Context [HR]	An executive challenge asked how much customer-concession exposure came from inbound arrival timing rather than downstream execution. The author scoped an inducted-volume analysis, commissioned partner-built data and reconciled competing measures.
Objective [PD]	Estimate a clearly defined exposure and decide whether an approved pre-dispatch timing package can improve mature customer outcomes at acceptable incremental cost. Keep accounting attribution separate.
Approach [RC/AI]	Reconcile arrival definitions, package-versus-unit grain, denominator choice, category membership, join multiplicity and maturity. Test the inference from crude group rates and from an under-mature allocation snapshot.
Findings [AI]	Two share calculations agree mechanically when their common denominator cancels. Unadjusted risk can reverse the within-stratum relationship. Truncated row sums and unidentified join multiplicity do not establish a reliable financial floor.
Decision value [PD]	Approve a scoped reconstruction of the original eligible population and a bounded pre-dispatch evaluation. Require valid joins, frozen clocks, independent outcomes and a detectable economic threshold before allocating intervention funds.
Implications [PD]	Data owners certify lineage and completeness; Operations specifies the physical treatment; an independent evaluator measures original-cohort outcomes; Finance validates avoidable costs. A changed reporting label earns no prevention credit.

EXECUTIVE RELEASE CONDITION

[PD] Accept the exposure only after category, footprint, unit and time alignment. Test the operating package on all original eligible shipments, including units not inducted or recovered. Preserve customer promises and existing entitlements.

Basis: project register S1. [HR] historical record; [RC] recalculated; [AI] analytical inference; [PD] proposed design. This is an analytical design case; historical customer effects and realized savings are unestablished.

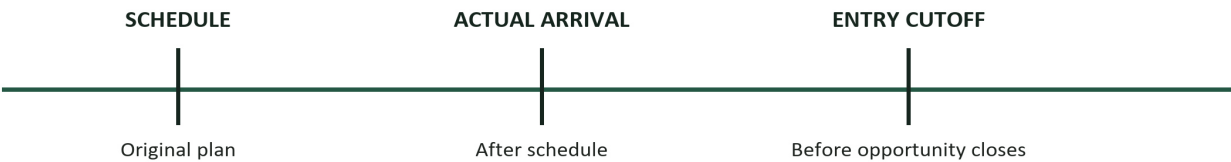
EXHIBIT 1 / FROZEN CLOCKS

Arrival after schedule and arrival after cutoff are different exposures

[HR] One analysis used a bounded breach of the station’s entry cutoff; a later join compared actual with scheduled inbound arrival. They also used different weeks, populations and denominators. [AI] These quantities cannot be treated as interchangeable estimates of one exposure.

ONE ARRIVAL CAN MISS ITS SCHEDULE AND STILL PRECEDE THE CUTOFF

Hypothetical ordering only / the original schedule, entry cutoff and promise remain frozen



Customer promise and observation horizon are separate clocks; neither changes through relabeling.

Basis: original TOC clock distinction. Real schedule–cutoff ordering and processing capacity require verification.

Clock or boundary [PD]	Required interpretation
Original schedule	The planned arrival known before treatment. A later schedule edit can change the late flag without changing physical arrival or customer outcome. Retain both versions.
Entry cutoff	The last arrival time relevant to a defined processing opportunity. Calendar, cycle, timezone and capacity rules determine its relation to the schedule; no universal ordering is assumed.
Customer promise	The original commitment used to evaluate on-time receipt. Operational recovery can bridge a missed inbound schedule without proving that the original delay was harmless.
Observation cutoff	The date at which delivery, nonreceipt and concessions are read. Event time and posting time differ; a later administrative record is not necessarily a later physical event.

FREEZE THE QUESTION BEFORE QUERYING

[PD] Specify which clock is breached, which units are eligible and which opportunity is lost. A pre-dispatch intervention must change actual operating opportunity; editing a timestamp or attribution field is not the treatment.

Basis: S1 bounded cutoff-breach population versus actual-after-scheduled-arrival join. Diagram is hypothetical; no internal clock, calendar or breach-window value is disclosed.

EXHIBIT 2 / ALGEBRA AND MEMBERSHIP

Two agreeing share calculations can repeat the same information

[HR] The author reconciled an exposed-concession unit share with the ratio of two rates. The source calls their agreement independent convergence. [RC/AI] When both rates use the same shipped denominator, that denominator cancels. Agreement checks arithmetic; it does not independently validate category membership or footprint.

THE SHARED EXPOSURE DENOMINATOR CANCELS

Hypothetical: exposed concessions E = 12; target category L = 80; shipped exposure S = 10,000

UNITS-SHARE PATH
 $12 / 80 = 15\%$

RATE-SHARE PATH
 $(12 / 10,000) / (80 / 10,000) = 15\%$

Same arithmetic, not independent corroboration. Verify E actually belongs to L.

Basis: algebraic identity for $S > 0$ and $L > 0$. No historical population, count or proportion is disclosed.

[RC/AI] Let E be exposed-concession units, L the target concession-category units and S shipped exposure. Then $(E/S)/(L/S)=E/L$ for $S>0$ and $L>0$. In the hypothetical example, $12/80=15\%$ and $(12/10,000)/(80/10,000)=15\%$. These are the same calculation expressed two ways.

Validation still required [PD]	Decision consequence
Category membership	Verify that E belongs to L under the same category definition. The available brief names exposed-and-conceded units but does not reproduce the target-category filter. A quoted “of which” does not supply that predicate.
Footprint and weighting	Match geography, business line and eligibility. A small package share in a peripheral segment does not bound its unit-weighted contribution without the segment’s units and concession mix.
Independent reconciliation	Rebuild from unique package and concession identifiers, actual units and source controls. Distinguish shared arithmetic from independently recovered inputs or separately validated extraction paths.

NAME THE DENOMINATOR AND THE CLAIM

[AI/PD] A share of all concessions answers a different question from a share of the late-delivery category. Neither share is a causal attributable fraction. Reconcile the numerator’s membership before presenting a part-of-whole percentage.

Basis: S1 exposure tie-out, denominator substitution and unresolved numerator/denominator alignment. All numbers in the figure and equation example are hypothetical.

EXHIBIT 3 / OPERATING MIX

A lower aggregate rate can conceal higher risk in every comparable group

[HR] The later analysis compared unadjusted concession risks on late and on-time inbound packages and retired a broad upstream-driver narrative. No covariate adjustment, interval or formal test is recorded. [AI] The comparison challenges the headline story but does not identify a causal null.

STRATUM RISK CAN RISE WHILE THE CRUDE RATE FALLS

Wholly hypothetical Simpson construction / identical stratum definitions, different group mix

Group	Low-risk share / risk	High-risk share / risk	Crude risk
On-time arrival	50% / 5%	50% / 15%	10%
Late arrival	80% / 7%	20% / 17%	9%

Within both strata: +2 percentage points. Crude comparison: -1 point.

Basis: original counterexample, not a fitted historical model. It establishes possibility, not the actual confounder.

[RC/AI] In the hypothetical construction, on-time risk is 5% in the low-risk stratum and 15% in the high-risk stratum. Late-arrival risk is 7% and 17%, respectively: two points higher in each stratum. Yet a different mix produces aggregate risks of 10% on time and 9% late. Standardizing both groups to the same mix restores the positive within-stratum difference.

Threat [AI]	Required design response [PD]
Different operating mix	Compare pre-existing route, cycle, node, shipment and demand factors with common support. Predeclare adjustment and report standardized risks with uncertainty.
Recovery and selection	Recovery effort can mask part of a delay’s burden. Inducted, successfully joined or conceded-only cohorts can select on post-arrival variables; retain the original eligible manifest.
Exposure is not treatment	The effect of an arrival delay differs from the effect of a feasible warning and handoff package. Specify the intervention and its costs rather than treating a descriptive group difference as its expected benefit.

REJECT AN OVERSTATEMENT WITHOUT INVENTING ITS OPPOSITE

[AI] The source did not establish that lateness caused the observed concessions. It also did not establish that lateness was harmless. A counterexample demonstrates non-identification; it does not prove that this particular mix reversal occurred historically.

Basis: S1 crude comparison and cycle-mix caveat; original hypothetical Simpson construction. Risks are probabilities, not odds. No historical risk or complementary rate is disclosed.

EXHIBIT 4 / JOIN AND EXPORT INTEGRITY

Quoting only a ratio does not repair an unidentified join

[HR] A joined dataset produced totals much larger than the reference dashboard, while another file is recorded at a tool’s row limit. The source retained scope warnings and rejected the capped extract as a complete answer. [AI] Multiplicity and truncation affect different quantities; neither is repaired by presenting percentages alone.

A RATIO IS NOT IMMUNE TO DIFFERENTIAL DUPLICATION

Hypothetical join audit / true unique population has 10 exposed and 90 unexposed cases

UNIQUE CASES 10 / (10 + 90) 10%	FOUR COPIES OF EVERY CASE 40 / (40 + 360) 10%	FOUR COPIES OF EXPOSED ONLY 40 / (40 + 90) 30.77%
---------------------------------------	---	---

A larger joined denominator does not establish whether duplication is balanced.

Basis: synthetic multiplicity counterexample. Recover keys, weights and coverage before interpreting the historical ratio.

[RC/AI] The hypothetical true population has 10 exposed and 90 unexposed cases. Four copies of every case preserve the 10% share. Four copies of exposed cases only change it to 40/130=30.77%. A ratio survives duplication only under sufficient balancing conditions, not because numerator and denominator came from the same join.

Release check [PD]	Failure it prevents
Entity and transaction grain	Test join cardinality by exposure and outcome; deduplicate on valid entity/transaction keys. SUM(DISTINCT amount) removes repeated values, not duplicate transactions.
Complete scoped extraction	Verify expected counts, pagination, filters, dates and typed timestamps. Pre-aggregate at the source using the intended grain, then reconcile back to distinct entities.
Value and sign discipline	A capped row sum can overcount duplicated charges or omit reversals. A lower-bound claim needs unique valid records, a nested target population and an appropriate nonnegative measure.
Period-specific economics	Use the actual eligible transaction values after reconciliation. Do not multiply a cross-category blended unit cost into a new population or annualize one incomplete day.

KEEP THE REJECTED FINANCIAL NUMBER REJECTED

[HR/AI] The author withheld a blended annualization from the draft response. That restraint is supported; it does not turn a different unreconciled row sum into measured savings or a guaranteed exposure floor.

Basis: S1 inflated join base, capped extract, distinct-value query issue and explicitly rejected blended annualization. Figure is hypothetical; actual counts and financial values remain private.

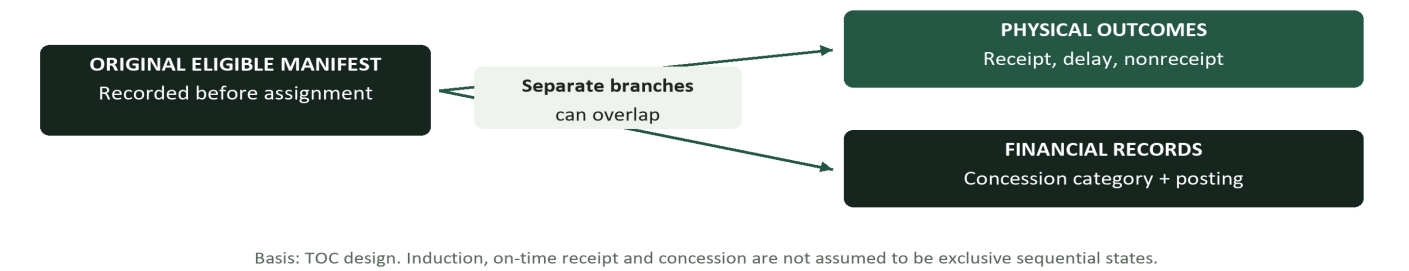
EXHIBIT 5 / ELIGIBILITY AND OUTCOME BRANCHES

Keep the original manifest when measuring recovery and customer harm

[HR] The source changed a recovery-oriented feed to include inducted packages whether or not they later arrived on time. That broadened the analysis, but the induct scan still excluded part of the original arrival population and could not isolate the requested dispatched-versus-stranded condition.

FREEZE ELIGIBILITY BEFORE THE OPERATING PACKAGE CAN CHANGE IT

Proposed original-manifest evaluation / retain noninducted, diverted and unresolved units



Observation [AI]	What must remain visible [PD]
Induction	A recorded scan can support a processing event. It does not alone establish continuous custody, a usable recovery window or dispatch; retain noninducted and missing-scan units.
On-time delivery	Measure against the original customer promise on the original eligible cohort. Record incremental recovery resources; a recovered outcome may hide a costly underlying delay.
Concession	This is a financial/customer-service event with its own category and posting clock. It can overlap other process states; it is not automatically the final stage of an exclusive physical funnel.
Dispatch or stranded status	Recover first relevant in-transit and induction events, align them to the cutoff and test the requested boundary directly. A missing timestamp cannot be supplied by a different event's presence.

[PD] Start evaluation before dispatch on a frozen manifest and retain diversions, nonreceipt, late reporting and unresolved observations. Link later physical states and concession transactions as separate branches. Estimate package and unit-weighted quantities separately, using actual contents rather than a transportable average units-per-package factor.

RECOVERY IS A MEDIATOR, NOT AN ELIGIBILITY FILTER

[AI/PD] An operating package can change who is inducted, recovered or observed. Restricting analysis to those successful states can change the population being compared. The total offer effect belongs to all original eligible shipments.

Basis: S1 inducted-volume scope and unavailable dispatch boundary; proposed TOC original-cohort design. Diagram states an evaluation contract, not a historical deployment.

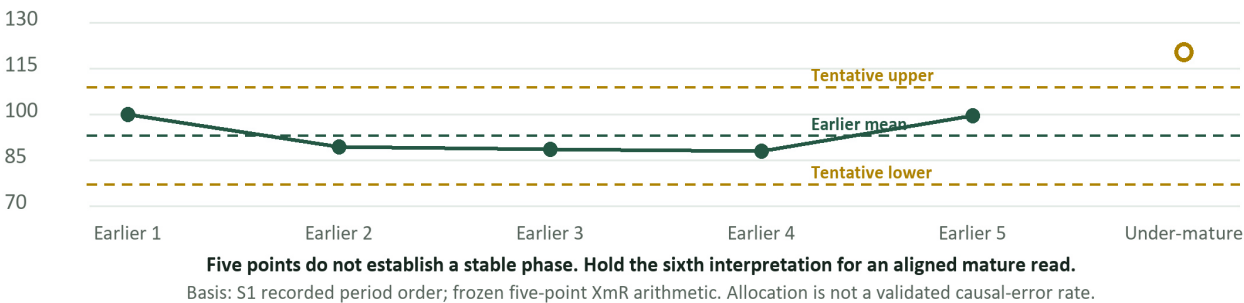
EXHIBIT 6 / MATURITY AND PROCESS INTERPRETATION

Treat an under-mature point as unresolved even when it crosses a fitted limit

[HR] The record supplies five earlier owner-allocation shares and a sixth, under-mature share. It interprets the earlier band as chronic and the last point as a maturation artifact. [RC/AI] The last point exceeds tentative limits fitted to the first five, but its different maturity prevents a like-for-like process interpretation.

THE LAST POINT IS OUTSIDE FROZEN LIMITS BUT HAS DIFFERENT MATURITY

Historical recorded allocation share / first period = 100; one numeric view only



[RC/AI] The chart reports one historical series indexed to its first period=100. Tentative individuals limits use the earlier five observations: $\text{mean}(x) \pm 2.66 \times \text{mean}(|x_t - x_{t-1}|)$. Freeze that reference before plotting the hollow under-mature point. Five observations do not establish a stable process; refitting the limits with the disputed point would answer a different question.

Observed fact	Interpretation and next check
Earlier observations [RC/AI]	No earlier point crosses its fitted arithmetic bounds. This does not prove chronic causal misattribution, station responsibility or process stability.
Under-mature point [HR/AI]	Its bound crossing is arithmetic. Both an incomplete reporting window and a genuine process change remain possible explanations until an aligned mature read is available.
Ownership share [AI/PD]	A station allocation among late arrivals is not itself an error rate. Validate the policy, supported cause and response owner independently before naming it misattribution.

REOPEN ON A FIXED MATURITY READ

[PD] Compare every original cohort at the same outcome horizon and preserve later revisions. Do not claim either deterioration or an explained artifact from an under-mature point; do not label within-limit changes as improvement.

Basis: S1 recorded allocation-share series, indexed from the first period. Raw proportions, complementary allocation shares, counts and transformation key remain private. This is a tentative process screen, not causal evaluation.

EXHIBIT 7 / ANALYTICAL LEADERSHIP AND ACCEPTED DELIVERY

Answer the requested decision in its required unit before adding the wider story

[HR] The author scoped the analysis, wrote reconciliation logic and rejected unsupported framing and annualization. The record also says the reply remained a draft and a later executive deliverable did not answer the requested question. The rebuilt monthly trend and impact-ranked analysis were not completed in the supplied record.

ANALYSIS, ACCEPTANCE AND OPERATING IMPACT ARE DIFFERENT MILESTONES

Historical contribution and proposed delivery contract / no inherited ownership of partner engineering



A drafted reply is not a sent answer. Internal review is not recipient acceptance.
An accepted analysis is not an implemented intervention or a realized saving.

Basis: S1 analytical role, partner-built pipeline, draft response and rejected/uncompleted executive deliverable.

Contribution or decision	Evidence boundary and accountable role
Scope and reconciliation [HR]	Provisional registry hoop support for the author’s brief, commissioned analysis and tie-out. Engineering, extraction and validated query construction are credited to partners; original files are unrecovered.
Framing discipline [HR/AI]	The author changed the narrative and withheld unsupported financial language. A rejected hypothesis does not make the replacement attribution story causally identified.
Acceptance [HR/PD]	A drafted reply and internal reviewer approval do not prove recipient acceptance. Record the literal question, required period, measure, maturity and acceptance evidence before marking an answer delivered.
Impact-ranked comparison [AI/PD]	A sibling outcome category, a subset of the target outcome and a possible input are different objects. Build a consistent decomposition or use separate views; do not rank their overlapping totals as additive causes.

THE COO’S NEXT MOVE

[PD] Commission the narrow evidence repair and a pre-dispatch treatment specification. Approve evaluation only after the original cohort, clocks, join grain and maturity are verified. Keep reporting repair, accepted delivery and measured operating benefit as distinct milestones.

Basis: S1 role boundaries, draft status, executive rejection and uncompleted rebuild. Personal-contribution grades are provisional; no sole engineering ownership, client acceptance or realized saving is asserted.

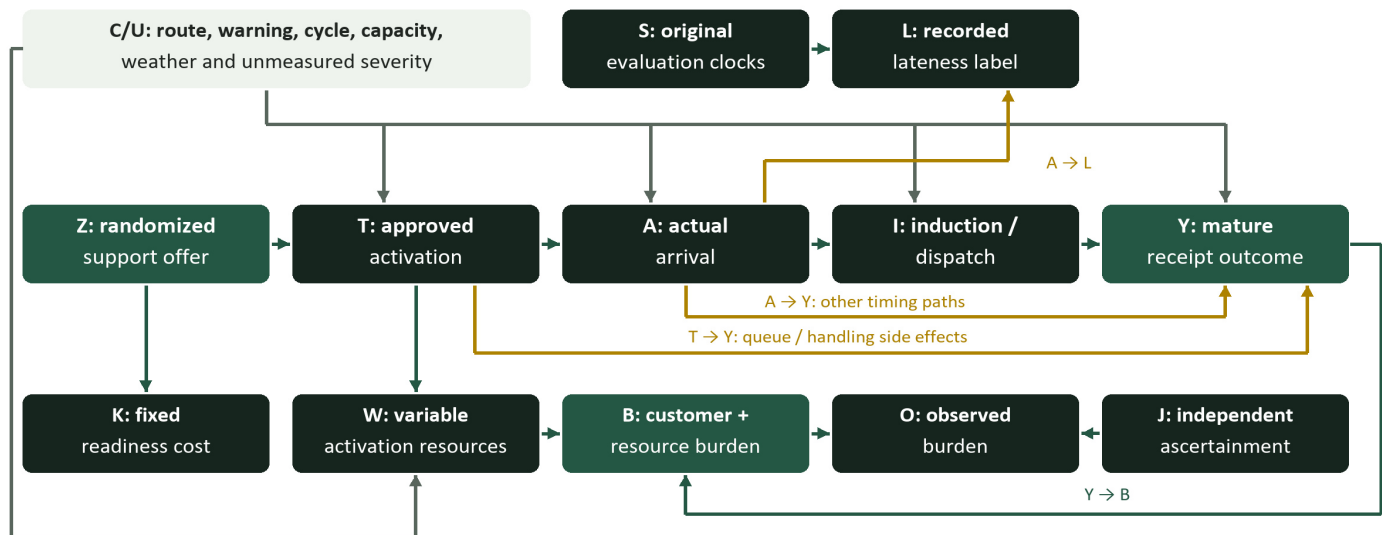
TECHNICAL ANALYSIS

Arrival, induction and customer outcome are different variables

[AI/PD] The proposed Timing-to-Outcome Contract (TOC) fixes the original cohort and clocks, validates the joins, and tests an operating intervention. A recorded late-arrival label is neither a causal effect nor an ownership ruling.

TEST AN OPERATING OFFER, NOT A CHANGE TO THE CLOCK LABEL

Proposed TOC graph / assignment precedes the arrival and induction it may change



Primary ITT retains every original manifest unit, including noninducted and diverted units.

$C/U \rightarrow T, A, I, Y, W$. Conditioning on I can open $A \rightarrow I \leftarrow \text{capacity} \rightarrow Y$.

K is charged outside B once. S is the saved evaluation clock; L is not a customer outcome.

Crossing lines without junction dots do not add edges. Safety and spillover need separate tests.

[PD] Z assigns approved pre-dispatch support; T is activation, A actual arrival, I induction/dispatch and Y mature receipt disposition. T can affect arrival, variable resources W and queue or handling harm. B combines customer and resource costs; O observes B under independent ascertainment J . Fixed readiness K is charged outside B once.

[AI/PD] C/U includes route, baseline warning, cycle, weather, available capacity and unmeasured severity. These can affect activation, arrival, induction, outcomes and work. S is the frozen evaluation schedule/cutoff; A and S define label L . Changing S alone can change L without changing A or Y .

[AI] Selecting $I=1$ may condition on a common effect of arrival and capacity while removing a pathway through failed induction. The total offer effect therefore retains every original manifest unit. The graph proposes mechanisms to test; the register does not identify their coefficients or demonstrate an intervention effect.

Basis: registered timing and induction definitions [HR]; proposed causal structure [AI/PD]. Hernán et al.2004; Hernán–VanderWeele2011. Original customer-outcome effects are unestimated.

TECHNICAL ANALYSIS

Define the clocks before estimating exposure

[HR/AI] The initial restricted cutoff-tier analysis and the later arrival-versus-schedule comparison change the exposure, period and population. Their estimates cannot be interpreted as successive measurements of one fixed causal quantity.

Quantity [AI/PD]	Definition and denominator
Schedule lateness	Actual arrival exceeds the original scheduled arrival, among all eligible original units.
Cutoff lateness	Actual arrival exceeds the original station cutoff; separately retain delay severity and excluded tiers.
Concession share	Eligible late-exposed concession units divided by all eligible units in the specified concession bucket.
Intervention effect	Usual-minus-offer mature burden over all original assigned manifest units, including noninducted and diverted units.

[AI] Without a verified ordering of schedule and cutoff, neither nesting direction is established. Arrival can be after schedule but before cutoff, or after cutoff but before schedule. A bounded tier additionally omits later arrivals. The clock definitions require time zone, daylight-saving handling, date rollover, original version and actual event provenance; a display fragment is not a timestamp.

[HR/AI] The first query changed an on-time-delivery numerator to any inducted package. That broadens the outcome condition but does not reveal whether the package was stranded before dispatch. Induction, dispatch, timely receipt and later concession are not mutually exclusive funnel states: an on-time delivery may still generate a concession. Missing scans must remain unknown unless independently resolved.

[PD] Keep event time, extraction time and customer-claim observation time separate. A concession observed by a fixed horizon is a horizon-specific outcome, not ultimate lifetime risk. Freeze the promise before assignment so a later promise extension cannot improve the endpoint. Audit whether selected service options disappear on the concession side of the join.

[AI] A valid timing exposure still does not determine causal or financial ownership. Policy allocation requires an accepted rule, while preventability requires an intervention contrast. The source left that allocation boundary unanswered.

Basis: incompatible registered exposure definitions, induction scope and unresolved attribution boundary [HR]. Compound-treatment and target-trial methods define the proposed estimands; no historical causal null follows.

TECHNICAL ANALYSIS

Two matching calculations can share the same error

[AI/RC] Re-expressing one numerator as a rate does not create an independent validation path. Algebra can verify internal consistency while leaving entity scope and source truth unresolved.

[AI/RC] Let A be the exposed eligible concession units, B the eligible units in the target bucket, and N a common shipped exposure. For $B > 0$ and $N > 0$, $(A/N)/(B/N) = A/B$. The common exposure cancels. This equality remains true if N is the wrong exposure or if A and B share a duplicated join. Rounded displayed rates add rounding differences but no independent evidence.

Acceptance test [PD]	What it can establish	What remains open
Rate-versus-count ratio	Arithmetic consistency under a common base	Correct grain, eligibility and source accuracy
Independent entity reconciliation	Membership and cardinality across source ledgers	Physical meaning and causal effect
Versioned bucket cross-tab	Whether exposed records belong to the stated target bucket	Whether their harm was preventable by this intervention

[HR/AI] The registered initial numerator describes inducted packages later conceded, while the headline denominator is a specific late-delivery bucket. The recovered register does not reproduce the required bucket filter. Before calling the ratio a subset share, verify that every numerator unit belongs to that bucket on the same entity, time and population basis. A bucket mix from a different extract cannot repair this missing membership test.

[PD] Reconcile source counts before and after each join, list unmatched entities, and inspect one-to-many cardinalities. Preserve separate physical-package, conceded-unit and financial-event grains. A conversion factor estimated from one mixture cannot automatically convert another mixture. Cross-check a blinded sample against evidence outside the joined extract.

[AI] A percentage presentation does not cure an unexplained join base that differs from the operating topline. Nor does a common proportional rescaling repair a population mismatch. Treat the conflict as a release condition with an accountable owner, rather than selecting the denominator that produces the preferred narrative.

Basis: exact shared-denominator identity [AI/RC]; registered footprint, bucket and join-base ambiguities [HR]. The identity is reproduced privately with explicitly hypothetical values.

TECHNICAL ANALYSIS

A lower crude risk can coexist with harm in every stratum

[AI/RC] This independent hypothetical example disproves the claim that a lower unadjusted late-arrival risk must refute causal harm. It does not estimate the historical bias or imply that the example occurred.

Baseline stratum	On-time risk	Late risk	Mix: on-time / late
Lower baseline risk	5%	7%	50% / 80%
Higher baseline risk	15%	17%	50% / 20%
Crude mixture	10%	9%	Different populations
Same 50/50 mixture	10%	12%	Common target population

[AI/RC] Define $Y(a)=1[U<b(C)+0.02a]$, with U uniform on $[0,1]$, $a=0$ for on-time arrival, $a=1$ for late arrival and $b(C)$ the table's on-time risk in stratum C . Exposure is independent of U within C . Late arrival raises risk by two percentage points in each stratum. Different observed mixtures nevertheless produce a one-point lower crude risk; a common mixture restores the two-point increase. These explicit hypothetical assumptions complete the counterexample.

[HR/AI] The register calls reported probabilities odds and treats their crude contrast as a causal refutation. They are risks; an odds contrast would transform each risk by $p/(1-p)$. Neither scale supplies exchangeability, precision or valid observation. The absence of a reported interval is not evidence of nonsignificance, equivalence or no operational effect.

[AI/PD] Baseline route, service window, cycle, weather and capacity may confound arrival and harm. Induction or successful linkage can introduce additional selection. Predefine baseline adjustment and common support; do not adjust away post-arrival induction when estimating the total effect. Sensitivity analyses must include unmeasured confounding and differential outcome capture.

[PD] The defensible historical decision is to withhold an unsupported driver claim and an unsupported harmlessness claim. Test the specific timing intervention on the original cohort. This retains the value of challenging a preferred explanation without replacing it with opposite causal certainty.

Basis: independently derived exact counterexample [AI/RC]; Simpson1951 and selection-bias methods. All percentages in this table are invented for the proof and are unrelated to client measurements.

TECHNICAL ANALYSIS

Truncation, duplication and successful execution are separate tests

[HR/AI] The register describes a capped single-day extract, broader scope than requested and damaged timestamp typing. These defects limit its use; they do not all bias a total in the same direction.

[AI/RC] A sum over a verified unique eligible subset is guaranteed to be a lower bound on the full same-population total when all target event amounts are nonnegative. A truncated row stream need not satisfy those conditions. Duplicated events or out-of-scope records can overstate the target despite missing rows. Conversely, SUM(DISTINCT amount) can collapse two genuine events with equal values. Deduplicate by the authenticated event key and version, not by the numeric amount.

Measurement object [PD]	Independent acceptance evidence
Physical unit	Stable entity linkage, no repeated-count inflation, original manifest retained.
Financial event	Valid event ID, units/currency, reversals, nonoverlap and eligible bucket membership.
Timestamp	Typed event/date/zone with provenance; missing and impossible sequences adjudicated.
Delivered dataset	Snapshot/hash, load completeness, consumer readback and reproducible source-to-output lineage.

[PD] Audit a random core of 200 future unit-context records and a separate exceptional-record sample. Two blinded coders apply a frozen codebook for clock validity, entity linkage, bucket eligibility and maturity. An independent-binomial worst-case agreement screen has approximately ± 6.9 percentage-point precision at 95%; repeated lanes, source families or shared failures require cluster simulation and larger samples. This is not promised precision for kappa.

[PD] Retain first-pass disagreement matrices, raw agreement, category prevalence and nominal kappa intervals. Adjudicate using independent manifests and event records, allowing unresolved status. Agreement between two users of the same faulty feed is not accuracy. Weighted exceptional samples need their design weights before population extrapolation.

[AI] Query execution, a pipeline listing and a dashboard response establish different software facts. Repeated source-to-consumer checks are needed to distinguish a live asset from a complete, current and decision-fit daily dataset. Neither achievement alone demonstrates a customer effect.

Basis: registered extract and query limitations [HR]; PostgreSQL aggregate semantics; Akidau et al.2015; Cohen1960. Sample size is prospective, hypothetical and conditional on dependence assumptions.

TECHNICAL ANALYSIS

A maturity exception is a reason to investigate, not an explanation

[HR/AI] The source excludes the latest allocation ratio because its observation window is incomplete, while describing earlier ratios as structural and stable. A stated maturity problem makes the last point incomparable; it does not prove why the point moved.

[PD] Preserve each original cohort at repeated as-of cuts, including downward revisions, reversals and disappearing joins. Plot the revision curve separately from the sequence of different event cohorts. A processing timestamp or apparently flat arrival-rate series cannot authenticate complete loading, invariant scope or claim maturity.

[AI/RC] For a suitable individual-value sequence, the mechanical XmR screen uses $MR_t = |x_t - x_{t-1}|$, center $\text{mean}(x)$, individual limits $\text{mean}(x) \pm 2.66 \text{ mean}(MR)$, and moving-range upper limit $3.267 \text{ mean}(MR)$. Freeze any candidate baseline before inspecting later observations. Do not pool an incomplete point into the baseline to explain away its departure; do not call a departure causal.

[RC/AI] The private audit contains 27 screens and three held-out endpoint checks: recorded series, earlier-period references, sparse comparisons, source versions and derived rates. All refitted screens lie within their tentative bands. Only the final allocation endpoint exceeds both frozen earlier-reference individual and moving-range limits, including reported-rounding sensitivity. This is an arithmetic alert, not a validated process signal.

[HR/AI] The available earlier allocation sequence is too short, and its measurement basis too weakly authenticated, to establish routine variation or a stable chronic defect. A later point above a frozen arithmetic band remains a conditional measurement alert. It could reflect maturation, changed mix, loading or a real change; those alternatives require evidence.

[PD] Rate dispersion analysis requires repeated valid numerators and exposures at a common grain. Rounded percentages, version alternatives and extracted totals do not supply a Laney panel. Inventory every provided ordered sequence, but keep sparse comparisons and artifact profiles distinct from an operating phase.

[AI/PD] No fixed worst-N treatment cohort or empirical regression-to-the-mean fraction is established. For a future selected-site study, freeze eligibility before assignment and use a concurrent comparison. Under an explicitly assumed linear reliability model $E[\text{next} - \mu | \text{baseline}] = p(\text{baseline} - \mu)$, a selected excess d has expected natural shrinkage $(1-p)d$. Sweep plausible p ; do not report that formula as an observed correction.

[PD] Prespecify route, business line, node type, service window and cutoff slack as moderators. Different recovery capacity may change the action effect even when aggregate arrival shares remain similar. Test interactions with uncertainty rather than comparing separate significance labels.

Basis: registered ordered ratios and incomplete observation [HR]; mechanical process-screen formulas [RC]; prospective selection and moderation safeguards [PD]. Final screen inventory is retained in the private audit.

TECHNICAL ANALYSIS

Credit the analytical discipline without certifying its conclusions

[HR/AI] The register supports a bounded contribution: specifying an inducted-volume query, commissioning a package-level join, reconciling bases, rejecting blended monetization and preserving superseded work. These are analytical and governance contributions, not demonstrated operational prevention.

[HR] Data engineering and the dashboard belonged to partners. A source-listed asset persisted, but sustained valid refresh was not independently reproduced. The outward reply remains a draft in the record; receipt, an accepted ownership boundary and execution of the proposed attribution remedy are not established. A later reviewer rejected a deliverable that did not answer the requested view.

[AI] In process-tracing terms, derivative descriptions provide provisional hoop evidence for direction and output production. Original scripts, immutable run logs, reviewer receipt, accepted rules and downstream readback would strengthen the sequence. None alone would identify customer impact. Repeated summaries sharing the same query are not independent causal replications.

[AI] Simpson and structural selection methods explain why a crude risk contrast can mislead. Compound-treatment theory explains why changed clocks and intervention versions cannot share an effect by name. The Dataflow model adds a distinct implementation lesson: event-time completeness and processing-time availability require explicit revision semantics. These are complementary constraints, not a composite score that averages away missing evidence.

[AI] Deming cautions against reacting to an unstable measurement stream. Campbell/Goodhart makes moving the schedule a concrete optimization risk: a better label can coexist with unchanged arrival and customer harm. Tucker–Edmondson motivates learning from the gap between an analytically interesting output and the actual decision request. Ittner–Larcker–Randall does not imply that adding measurements automatically creates value.

[AI] Consolidated limits are unrecovered originals, changing exposure and footprint, disputed bucket nesting, unresolved join multiplicity, incomplete maturity, observational selection, unknown action authority and no post-intervention outcome panel. These limits do not prove the historical claims false or the intervention useless. They define which inferences remain unestablished and which evidence would change the decision.

Basis: derivative contribution and implementation record [HR]; bounded process tracing [AI]. Source authentication remains incomplete. No attribution, adoption, customer or monetary outcome is borrowed from adjacent projects.

TECHNICAL ANALYSIS

Randomize the offer before the arrival it intends to change

[PD] The prospective test compares approved early-departure and handoff support with usual approved practice across independent route/station capacity domains. It evaluates an operating offer, not reassignment of an accounting label.

[PD] Freeze the manifest, original schedule, station cutoff and customer promise before random assignment. A pre-dispatch warning triggers support; future actual arrival is unavailable to the rule. The synthetic model uses latent would-be arrival windows to generate outcomes, not as information granted to operators. False-positive warnings, unsuccessful attempts and early-window no-ops still incur activation work.

[PD] Enroll six fixed weekly cohorts with equal externally scheduled exposure per domain in the planning design. Retain noninducted, diverted, refused and nonreceived units. Independent disposition adjudication occurs at original promise plus 35 days, with customer-promise miss and resource-weighted burden reported over the same original units. Unknown truth is bounded; a missing join is not an on-time outcome.

[PD] The primary contrast is mean usual-minus-offer total variable burden; fixed readiness is charged once to the same service pool. Physical receipt categories and their costs remain visible. The planning law allows actual earlier arrival and charges activation, including failed or unnecessary attempts. A schedule-only relabeling scenario has no physical benefit.

[PD] Randomize domains large enough to contain shared dispatch, labor and receiving capacity. Record cross-domain queues and diversion; parallel assignment does not remove spillover automatically. Rushed handling, worker safety, downstream congestion and displacement of unselected freight need separate prespecified harm gates. Promise-miss power does not size these additional safeguards.

[PD] The model pages compare detection at economic break-even with joint economic and customer gates, then stress heterogeneity and queue harm. These are conditional planning calculations, not achieved power or available recruitment. Pilot the reference process, warning calibration, covariance and feasible independent domains; redesign and re-power if assumptions fail. More correlated packages cannot replace missing independent domains.

Basis: new prospective TOC design [PD]; target-trial principles. The historical study did not randomize timing or measure this intervention. All planning probabilities, costs and sample sizes are hypothetical.

TECHNICAL ANALYSIS

Four propositions make the contract falsifiable

[AI/PD] TOC is useful only if its claims can fail on a frozen population, against accepted evidence and thresholds. A data repair and a customer intervention have different acceptance tests.

Proposition [PD]	Test that can reject it
P1: Exposure identity survives replay	Freeze manifest and original clocks. Recompute both timing labels; unresolved identity, version or bucket membership blocks a causal release.
P2: The daily view is decision-fit	Independently reconcile entities, financial events, unmatched rows, typed times and later revisions. A successful job without valid consumer output fails.
P3: The offer improves mature service	Randomized all-manifest customer contrast must clear its prespecified threshold; faster arrival labels alone do not satisfy it.
P4: Net value survives full costs	Economic lower bound must exceed readiness cost while customer and separate harm gates pass; discount favorable avoidable benefits, not added work or adverse outcomes.

[PD] Preregister a meaningful effect, error rates, maturity horizon, analysis weights and stop rules before inspecting outcomes. A confidence interval spanning material harm and benefit is inconclusive; nonsignificance does not falsify a useful mechanism. Test equivalence only against justified margins, with adequate precision.

[PD] If later rollout is nonrandomized and staggered, define group-time effects among treated domains against comparable not-yet-treated domains. Require no anticipation, overlap, defensible parallel trends and explicit spillover limits. Event studies and joint pretrend/equivalence diagnostics can challenge assumptions but cannot prove them. Mature-outcome and covariance pilots must precede field power claims.

[AI] Keep reporting allocation outside the prevention total unless an implemented response mechanism changes actual outcomes. Do not monetize a blended concession rate as savings or annualize one selected extract. Value of information may justify a bounded measurement study, but cannot override missing authority or outcome-truth requirements.

[AI] The immediate release is an auditable exposure-and-data contract with a specified identifying test. The larger causal and economic claims remain conditional until the evidence supports them.

Basis: proposed four-part TOC framework [PD]; causal interpretation and limits [AI]. Any future staggered field valuation uses its stated treated-population effect, not the randomized planning effect by substitution.

PROSPECTIVE MODEL / FROZEN ARRIVAL CONTRACT

Test earlier physical arrival while preserving original promises

[PD] Randomize independent route/station capacity domains to an approved pre-dispatch warning, protected departure and handoff offer or usual practice. This wholly hypothetical field-planning model retains every original manifest package. No schedule relabeling, refund hold or intentional customer delay is permitted.

[PD] Six cohorts contain 40 packages each initially, or 160 each after redesign. Equal frozen counts align domain and package means. Independent disposition is measured at original promise + 35 days, including noninduction, diversion and nonreceipt. Calibrate this maturity window before use.

Potential arrival window	Warning probability	On time / late / nonreceived
Before scheduled arrival	0.15	0.94 / 0.05 / 0.01
After schedule; before cutoff	0.40	0.88 / 0.10 / 0.02
After station cutoff	0.80	0.50 / 0.42 / 0.08

[PD] The middle window has share 0.24. After-cutoff share is 0.08 or 0.16 across equally likely domains. A flagged package moves one window earlier with probability 0.50; an early-window flag cannot improve its window. Every activation costs 0.15 hypothetical resource units, including ineffective and unnecessary attempts.

A WARNING CANNOT USE THE FUTURE ARRIVAL

[PD] Potential untreated windows are latent variables in the invented law. Operational warnings must use only pre-departure information. Source lateness labels do not establish prediction quality, preventability or treatment eligibility.

[PD] Terminal outcomes are mutually exclusive. Conditional packages are independent within a shared domain timing grade; domains are independent. The trial tests the whole offer, not an execution-conditioned effect.

PROSPECTIVE MODEL / JOINT DECISION POWER

A larger package sample can recover power lost to queue spillover

[PD] Exact finite episode moments support conditional Gaussian planning. The variance envelope covers the full effectiveness path; actual boundary variance is used for power. This is an approximation under known hypothetical parameters, not achieved field confidence.

Design	Domains/arm / packages/domain	MDE (index)	Break-even detection	Joint power floor
Initial	64 / 240	> full path	20.9%	26.0%
Base candidate	360 / 240	5.36	81.3%	91.9%
Same / stress	360 / 240	6.11	69.8%	61.4%
Revised / stress	360 / 960	4.63	91.5%	86.4%

[PD] Usual variable burden is 100; fixed readiness cost is 5.44 index points. Require 80% correct-direction detection at the positive 2.5% tail of a two-sided 5% zero-effect test, plus an 80% joint floor for economic and promise-miss superiority. Each gate uses a one-sided 97.5% lower bound.

[PD] Effectiveness scales the 0.50 move probability. Stress widens after-cutoff shares to 0.06 / 0.18 and diverts 0.5% of otherwise on-time early-window packages into late receipt, including unflagged packages. Sampling 960 instead of 240 already scheduled eligible packages restores conditional power; do not create operating volume. Recalibrate if capacity or mix changes.

CALIBRATION AND RECRUITMENT COME BEFORE A TRIAL

[PD] Minimum domains per arm are 349 base, 542 under stress and 307 after the sampling redesign. A 360-per-arm candidate requires 720 independent capacity domains; availability is unestablished. Damage, rushing, safety, external spillover and missing-reference guardrails require separate margins and power.

[PD] Joint floor = max(0, economic power + promise power – 1), without endpoint independence. Seven 30,000-trial finite-process checks assess approximation at zero, break-even and full-effect settings.

PROSPECTIVE MODEL / ARRIVAL ECONOMICS

Charge unnecessary activations and displaced queues in full

[PD] Hypothetical resource units per original package value on-time, late-received and nonreceived outcomes at 0 / 5 / 25; common effort is 0.05. Usual variable burden is 1.102, indexed to 100.

[PD] Full effectiveness avoids 0.1728 physical burden but adds 0.0432 activation effort. Stress adds 0.01504 queue harm. Fixed readiness cost $K = 0.06$ allocates forecasting 0.020, capacity 0.025 and coordination 0.015; research-only audit cost is separate.

Avoidable share a	h = 0.80	h = 0.90	h = 1.00
0.60	-3.20	-2.26	-1.32
0.80	-0.69	0.56	1.81
1.00	1.81	3.38	4.95

[PD] Cells are net index points; h is the non-overlap share. $Net_raw = a \times h \times 0.1728 - 0.0432 - 0.01504 - K$; $Net_index = 100 \times Net_raw / 1.102$. Discount only favorable avoided harm. New activation, queue harm and readiness costs stay fully charged. At $a = h = 1$, changing K from 0.04 to 0.08 changes net from 6.77 to 3.14 index points.

[PD] If half of each arm’s day-35 nonreceipts later arrive late, net falls from 4.95 to 1.90 index points. This different maturity estimand requires new power; original promises remain missed.

KNOWN CLOCKS DO NOT RESOLVE UNKNOWN OUTCOMES

[PD] Keep unverified dispositions in bounded outcome totals. Field value needs original eligible volume $\times$ mature randomized outcome-vector ITT benefit, less all added harm and costs. A later nonrandomized rollout instead needs an identified treated-population ATT. Neither field effect is observed here.

[PD] Shares must match category-specific realizability; reprice and re-power if they differ. Readiness allocations use the same original-package service pool across designs; larger research samples do not create benefits.

PROSPECTIVE MODEL / DECISION-INFORMATION VALUE

Buy evidence that distinguishes a timing intervention from a clock repair

[PD] This separate utility model values the next investment decision. First identify insufficient evidence or unsafe capacity; otherwise distinguish a supported physical timing opportunity from a measurement issue. All priors and utilities are hypothetical.

Next investment action	Physical opportunity	Clock / population issue	Insufficient evidence
Bounded timing trial	13	-7	-11
Reconcile clocks and population	2	9	-3
Hold investment; obtain evidence	1	1	5

[PD] Prior probabilities are 0.40 / 0.35 / 0.25. The best prior action, reconciliation, has utility 3.20; perfect information can add at most 6.40. Signal accuracy is q, with either wrong signal probability (1 - q)/2. Evaluate all 27 signal-to-action policies.

Signal quality q	EVSI	Evidence cost	Net EVSI
0.55	1.09	0.50	0.59
0.75	3.45	1.25	2.20
0.90	5.22	2.75	2.47

INFORMATION VALUE DOES NOT REQUIRE POSITIVE OPERATING IMPACT

[PD] Evidence can redirect funding toward reconciliation or an investment hold. Acquisition cost is separate from subsequent action costs already included in utility. Do not add EVSI to the operating-benefit calculation. Customer protection and authority gates remain mandatory regardless of expected utility.

[PD] This finite Bayesian exercise values a changed decision. It does not estimate historical avoided expenditure, establish a source hypothesis or authorize a live timing change.

Reproduce the exposure before evaluating the intervention

[AI] This is an exposure-measurement and operating-evaluation case. Its historical contribution is analytical direction and reconciliation. Its new contribution is a contract connecting frozen clocks, original populations, mature outcomes and a falsifiable action test.

Claim class	Meaning and replication boundary
[HR] Historical record	A statement or action attributed by the supplied register. Original correspondence, queries, extracts and run logs remain unrecovered. Inherited confidence labels and repeated summaries are not independent authentication.
[RC] Recalculated	Denominator identities, join counterexamples, tentative process limits and finite prospective calculations. Arithmetic closure does not establish category nesting, representative selection or causal validity.
[AI] Analytical inference	Interpretations bounded by their assumptions: equal ratios can share a dependency; a crude risk comparison can reverse after stratification; incomplete maturity can change an allocation without proving a physical change.
[PD] Proposed design	The Timing-to-Outcome Contract, clock and cohort freeze, independent outcome reference, randomized operating offer, economic stress tests and information-value study. These are new designs, not historical implementations.

Replication sequence

[PD] Freeze source versions and reproduce every selected cell. Retain competing timing definitions and rejected numbers. Reconstruct entity keys, join multiplicities, target-category predicates, typed timestamps and the original cohort before interpreting any exposure share.

[PD] Replay each cohort at fixed maturity; retain unresolved outcomes. Recompute process screens on comparable phases. Independently enumerate the proposed treatment law, then verify costs, covariance, break-even and recruitment before commissioning a field experiment.

THE DECISION THE CASE CHANGES

[AI/PD] Replace confidence in a reconciled percentage with an auditable test of actual timing and mature customer outcomes. A valid query can describe an exposure; prevention requires a defined action and a credible comparison.

Historical public metrics use a single baseline-100 normalization. Internal counts, raw rates, clock names, source identities, personnel and version-specific calculations remain in the private annex.

RESEARCH REFERENCES / TIMING, SELECTION AND DATA CONTRACTS

Make the timing claim survive a change in population and clock

Primary research and official methods inform the interpretation and proposed evaluation. They do not authenticate unavailable project originals, identify the historical timing effect or supply a benchmark improvement.

Simpson, E. H. (1951). The Interpretation of Interaction in Contingency Tables. *Journal of the Royal Statistical Society B*, 13, 238–241. DOI 10.1111/j.2517-6161.1951.tb00088.x. Full original paper. Aggregate associations can conceal different conditional relationships. The case derives a transparent numerical reversal; it does not claim the historical comparison has that exact mechanism.

Hernán, M. A., Hernández-Díaz, S., & Robins, J. M. (2004). A Structural Approach to Selection Bias. *Epidemiology*, 15, 615–625. DOI 10.1097/01.ede.0000135174.63482.43. Primary indexed abstract; author PDF direct access unavailable. Conditioning on a shared consequence can induce association. Restricting to inducted shipments therefore needs an explicit causal justification.

Hernán, M. A., & VanderWeele, T. J. (2011). Compound Treatments and Transportability of Causal Inference. *Epidemiology*, 22, 368–377. DOI 10.1097/EDE.0b013e3182109296. Primary indexed text; direct PMC access challenge. A named exposure can contain materially different versions. Earlier arrival, a changed schedule and a changed accounting label are not one treatment.

Akidau, T., et al. (2015). The Dataflow Model: A Practical Approach to Balancing Correctness, Latency, and Cost in Massive-Scale, Unbounded, Out-of-Order Data Processing. *PVLDB*, 8, 1792–1803. DOI 10.14778/2824032.2824076. Full primary paper. Event time and processing time require distinct semantics. The case extends that discipline to cohort maturity and revision disclosure.

Hernán, M. A., & Robins, J. M. (2016). Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *American Journal of Epidemiology*, 183, 758–764. DOI 10.1093/aje/kwv254. Full primary paper. Explicit eligibility, assignment, follow-up and estimands expose design errors. The proposed study freezes original manifests before an operating offer can affect selection.

PostgreSQL Global Development Group. Official documentation, Value Expressions, section on Aggregate Expressions. DISTINCT applies to input values. Summing distinct quantity values cannot substitute for deduplicating physical entities or financial-event identifiers. This establishes a SQL semantic distinction, not the correctness of an unrecovered historical query.

Access scope and citation verification are recorded privately. Methodological analogies do not import external effect sizes, clinical standards or certification into this operating case.

RESEARCH REFERENCES / MEASUREMENT AND EVALUATION

Match uncertainty and comparison to the operating decision

Primary research and official methods inform the interpretation and proposed evaluation. They do not authenticate unavailable project originals, identify the historical timing effect or supply a benchmark improvement.

NIST/SEMATECH. e-Handbook of Statistical Methods, section 6.3.2.2: Individuals Control Charts. Official method. Moving-range limits require ordered comparable observations. This case separately reports tentative refitted screens and a held-out endpoint against an earlier frozen band; neither converts incomplete maturity into a causal signal.

Laney, D. B. (2002). Improved Control Charts for Attributes. *Quality Engineering*, 14(4), 531–537. Primary publisher abstract. Dispersion correction needs eligible repeated counts and denominators. Large extracts and rounded rates do not provide a compatible panel. The case reports the missing prerequisites instead of inventing a correction.

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20, 37–46. Primary publisher record. Agreement should be assessed using independent coding, confusion cells and uncertainty. Reliability of an ownership taxonomy remains distinct from physical timing validity and causal attribution.

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 adverse units changes interpretation of subsequent movement. The proposed study fixes eligibility and comparisons before assignment; no empirical recovery correction is inferred from aggregate exposures.

Hudgens, M. G., & Halloran, M. E. (2008). Toward Causal Inference With Interference. *Journal of the American Statistical Association*, 103, 832–842. Primary manuscript. Shared dispatch and receiving capacity can transmit effects between units. Domain boundaries and queue-displacement monitoring are therefore substantive assumptions, not merely standard-error choices.

Callaway, B., & Sant'Anna, P. H. C. (2021). Difference-in-Differences with Multiple Time Periods. *Journal of Econometrics*, 225, 200–230. Author manuscript. A later nonrandom staggered rollout needs explicit treatment timing, eligible comparisons and group-time effects. That design requires assumptions and power beyond the randomized hypothetical planning model.

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. Evidence earns value through changed choices. The separate finite-state calculation bounds information value and deducts evidence cost without claiming a historical saving.

Access scope and citation verification are recorded privately. Methodological analogies do not import external effect sizes, clinical standards or certification into this operating case.

RESEARCH REFERENCES / MEASUREMENT, LEARNING AND CONTRIBUTION

Link the corrected measure to an accepted action and outcome

Primary research and official methods inform the interpretation and proposed evaluation. They do not authenticate unavailable project originals, identify the historical timing effect or supply a benchmark improvement.

Collier, D. (2011). Understanding Process Tracing. *PS: Political Science & Politics*, 44, 823–830. Primary publication. Contribution claims need discriminating evidence. Scoping, commissioning, reconciling, drafting, delivering and obtaining acceptance are separate events. Derivative records support provisional roles without substituting for original artifacts.

Tucker, A. L., & Edmondson, A. C. (2003). Why Hospitals Don't Learn from Failures. *California Management Review*, 45(2), 55–72. Author manuscript. Local fixes may leave system causes intact. Here an analytically interesting view that fails the requested decision motivates explicit consumer acceptance and operational readback, without importing healthcare effect estimates.

The W. Edwards Deming Institute. The Funnel Experiment. Official exposition. Reacting to routine variation can destabilize a process. This case treats a maturity-incompatible endpoint as a measurement investigation and avoids calling within-limit movement improvement or deterioration. A frozen arithmetic crossing alone does not authorize an operating correction.

Campbell, D. T. (1976). Assessing the Impact of Planned Social Change. Original institutional record. Consequential indicators can alter behavior and recording. Freezing schedules and customer promises prevents retrospective relabeling from satisfying the proposed physical-outcome test. No historical manipulation intent is inferred.

Goodhart, C. A. E. Problems of Monetary Management: The U.K. Experience. 1975 conference; revised volume listed as 1976 in the RBA archive. A targeted measure may cease to represent its original construct. The case distinguishes an improved lateness label from earlier arrival, reliable handoff and a mature customer benefit.

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. Measurement must connect to strategic and operating use. More data, a daily refresh and numerical agreement do not themselves establish an accepted decision or value realization.

OECD / European Commission JRC (2008). Handbook on Constructing Composite Indicators. Official handbook. Normalization, weights and aggregation affect interpretation. The case retains separate physical outcomes and added work, tests avoidability and overlap, and publishes one historical normalization to reduce inversion risk across portfolio documents.

Access scope and citation verification are recorded privately. Methodological analogies do not import external effect sizes, clinical standards or certification into this operating case.

FINAL REFEREE REPORT / INTERNAL AGENT-ASSISTED REVIEW

Accept the diagnostic advance; require a separate prevention test

[AI] This is author-led, agent-assisted review, not external peer review, client approval or a consulting-firm ranking. Its verdict concerns the bounded historical interpretation and proposed validation design; no historical customer benefit is certified.

Major issue	Disposition in the revised case
Agreement and category nesting	Exposed the shared-denominator identity. Retained the missing target-bucket predicate and incompatible footprints; identical arithmetic cannot independently authenticate a numerator.
Crude association and selection	Derived a reversal counterexample. Defined an original-manifest estimand and retained noninducted, diverted and unresolved units. No causal null is inferred from the crude comparison.
Clocks and outcome maturity	Separated schedule and entry-cutoff exposure. Kept original promises fixed. Reported the held-out crossing as a conditional measurement alert, not established deterioration or a proven artifact.
Economics and evaluation	Charged false warnings, failed attempts and queue harm. Compared break-even with domain-level design requirements. All costs, probabilities, covariance and recruitment remain hypothetical.
Contribution and delivery	Credited direction and reconciliation provisionally; preserved partner engineering, a draft reply and the rejected executive view. No accepted rollout, prevention or saving is invented.

Partner critique: can a COO act on page one?

[AI] Yes: commission the frozen-clock and entity-grain reconciliation, require the requested consumer view, and fund a calibrated test only after the measurement contract passes. The sponsor can recognize analytical leadership without converting exposure into an investment return.

Minor revisions and remaining limits

[AI] Preserve windows, financial-event keys, revisions and role boundaries. Recover originals, calibrate warnings and outcome ascertainment, and verify domain recruitment before committing to the proposed test.

REVIEW DISPOSITION / PASS WITHIN STATED SCOPE

[AI] The case supports diagnostic leadership and a falsifiable timing-to-outcome contract. It establishes no historical timing effect, verified misallocation rate, implemented prevention benefit or realized financial saving.

Scope: internal source, scientific, numerical and executive review. Acceptance remains bounded by the selected record and explicit hypothetical analysis.