

OPERATING CASE STUDY / OPERATIONAL EXCELLENCE

Recovery Roadmap & Outcome Alignment

A detailed roadmap can still lack a sufficient path to the outcome

An executive recovery-plan case that separates line specificity, calendar feasibility, input-response evidence and approval. It tests when a quantified roadmap supports a commitment.

RECOVERY ROADMAP & OUTCOME SUFFICIENCY

BLUF / BOTTOM LINE UP FRONT

Make every action auditable. Test whether the whole plan can meet the goal.

ANALYTICAL FINDING

Specific owners and monthly contributions expose gaps; they do not close them. A complete input trajectory is not a causal output equation, and summed monthly rates are not a weighted horizon result.

CONTRIBUTION IN THE SOURCE RECORD

Provisional registry-level hoop for the review artifact, line contract and analytical framing. Programme execution and outcomes remain separate.

EVIDENCE STATUS

Exact review, roadmap and sign-off originals remain unrecovered. A related planning workbook cannot authenticate release, acceptance or impact.

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

STRUCTURED ABSTRACT

Require a sufficient output bridge before accepting the recovery commitment

[AI] Specific actions make a plan auditable. They do not establish that the portfolio has enough capacity, starts early enough or causes the promised customer outcome.

Element	Decision-relevant conclusion
Context [HR]	A planning lead reportedly rebuilt a challenged monthly review into a structured recovery roadmap with action, owner, mechanism, cadence and monthly contribution fields. The artifact was described as complete but awaiting sign-off; sending and acceptance remain unconfirmed.
Objective [AI]	Determine whether the listed initiatives can close the required gap under a common output definition, exposure calendar, feasible resource envelope and accepted decision rights.
Approach [RC/AI]	Reconcile the commitment arithmetic and source versions privately. Separate a line-level format test from a portfolio sufficiency test. Construct independent examples for weighted timing and causal input-output ambiguity, then evaluate a finite shadow commitment rule.
Findings [AI]	Under the stated additive basis, the listed future stack falls short even at full delivery. Summing monthly rates does not produce an annual weighted rate. A perfectly predictive input equation can still fail under intervention.
Decision value [PD]	The sponsor approves the goal and commitment; programme owners accept scope and dates; the resource owner verifies capacity; measurement and Finance owners validate original outcomes and incremental cost. A qualified roadmap remains feasible when sufficiency is unresolved.
Implications [PD]	Use the Input-to-Outcome Sufficiency Contract (IOSC). The hypothetical paid rules improve on a naive commitment but lose to a free qualified plan at the main prices. A perfect-information ceiling rules out the higher review fees within that law.

[PD] Preserve customer promises and remedies. Validate the response and timing assumptions before a confirmatory decision; a generated roadmap or successful arithmetic check is not implementation evidence.

[HR] historical record; [RC] recalculated; [AI] analytical inference; [PD] proposed design. Historical quantities, identities and source coordinates remain private. Public constructions are independent hypotheticals.

EXHIBIT 1 / THE LINE CONTRACT

Make missing commitments visible, then test the portfolio they form

[AI] A required field can reveal a gap without closing it. A filled field can make a claim inspectable without making it true.

Roadmap field	Additional acceptance evidence
Action and mechanism	Executable change, eligible population, expected pathway and contrary evidence.
Named owner	Accepted authority, accountability and available delivery capacity.
Reporting cadence	A reliable measurement definition, decision threshold and escalation response.
Period contribution	Activation date, exposure weight, overlap and output-response basis.
Bridge to goal	Complete feasible portfolio, common units, uncertainty and explicit residual gap.

[HR/AI] The source replaced broad focus areas with structured programme cards and visibly marked missing ownership or cadence. That is an auditability improvement. Calling the document complete does not establish that every initiative has an accepted owner or that the largest unresolved line can be delivered.

[AI] A portfolio can reconcile against its own table while missing required work, resource conflicts or an unlisted prerequisite. Conversely, a visible gap is not evidence that no useful work exists. Keep the original matrix, programme boundaries and acceptance records distinct from the rendered presentation.

ACCEPT THE COMMITMENT, NOT ONLY ITS FORMAT

[PD] Test line completeness, authorization, calendar feasibility and output sufficiency separately. The sponsor can accept a qualified plan with named unresolved conditions. The reporting author cannot manufacture ownership, capacity or expected benefit to make the table appear complete.

[HR/AI] This case concerns a single review and recovery-roadmap deliverable. It is distinct from the parent recurring monthly mechanism and from the programmes listed in the plan.

EXHIBIT 2 / THE CALENDAR BRIDGE

Weight the contribution when the original exposure actually occurs

[AI] Rate reductions add within a common period and population only under a valid decomposition. Across periods, the horizon result requires exposure weights.

Independent invented plan	Early-heavy calendar	Late-heavy calendar
Three period reductions	2, 4, 6	2, 4, 6
Original exposure shares	0.60, 0.30, 0.10	0.10, 0.30, 0.60
Sum of period reductions	12 rate-period units	12 rate-period units
Exposure-weighted reduction	3	5
Horizon target	4: not met	4: met

[PD] The same ramp and the same sum yield opposite target conclusions. The weighted result is the sum of each period reduction times its original exposure share. A later activation shifts the contribution into a different exposure period; calling it an annual benefit does not recover the missed earlier opportunities.

[RC/AI] In the private source reconstruction, the displayed future stack is below the remaining requirement even if every listed line delivers fully, conditional on a common additive basis. This is a sufficiency gap in the shown bridge, not proof that the complete operational portfolio is impossible: the original full matrix is not authenticated, and missing or noncomparable lines may require reconciliation.

KEEP THREE TESTS SEPARATE

[PD] Check the period arithmetic, the exposure-weighted horizon and the activation/resource schedule. A prior delivery ratio is a sensitivity input, not a causal productivity estimate or an unavoidable ceiling on future execution.

[PD] The three-period numerical example is invented. [RC/AI] Source sums retain their monthly-rate-period units privately; no historical annual rate, currency value or impact is inferred.

EXHIBIT 3 / INPUT-TO-OUTPUT SUFFICIENCY

Even perfect prediction does not prove that forcing inputs will improve output

[PD] An exact observational counterexample isolates the difference between a checkable equation and a causal commitment.

Three invented worlds	Same observed fit	After setting all inputs to goal
Inputs cause improvement	$Y = 120 - 10(X_1 + X_2 + X_3)$	Mean failures 90
Inputs only reveal context	Same equation in every observed cell	Mean failures 105
Forcing inputs adds harm	Same equation in every observed cell	Mean failures 120
Common observational baseline	All eight input patterns equally likely	Mean failures 105

[PD] Each cell contains 1,000 original eligible units. Three independent fair contextual variables U determine observed inputs $X = U$. The worlds have structural outcomes $120 - 10 \cdot \text{sum}(X)$, $120 - 10 \cdot \text{sum}(U)$, and $120 + 10 \cdot \text{sum}(X) - 20 \cdot \text{sum}(U)$. All produce the same observed data and a unique perfect multivariate regression. Setting $X = (1,1,1)$ while preserving U separates benefit, no effect and harm.

[AI] A weekly glidepath calculated from selected inputs is therefore a conditional forecast. Without a validated intervention-response equation, its arithmetic does not prove that hitting the input targets will hit the customer outcome. This construction does not show that the historical inputs are harmful or useless; it shows what their association cannot identify.

TEST THE ACTION THAT MOVES THE INPUT

[PD] Specify how the input changes, which other work is displaced and which original-customer outcome is measured. A credible intervention comparison can identify total policy effect; achieving a mediator target alone cannot.

[PD] Source-independent structural construction, independently enumerated. No source regression coefficients, programme effects or field sample sizes calibrate it.

EXHIBIT 4 / RESOURCE AND CUSTOMER COUNTER-METRICS

Protecting one workstream requires accounting for the work it displaces

[AI] A staffing standard can be a legitimate operating choice while still carrying capacity, service and cost consequences that the roadmap has not measured.

Proposed lever	Outcome and burden to retain
Protect problem-solving labour	Incremental recovery, delayed ordinary work, overtime and forgone alternatives.
Train or coach a selected cohort	Original eligible cohort, actual participation, selection/reversion and comparison outcomes.
Change a defect status or coding instruction	All-cause original physical outcome; cross-bucket transfers and recording changes.
Inspect weak regions and outlier sites	Exposure, overlap, escalation capacity and effects beyond the selected tail.

[HR/AI] The source treats protected labour as an execution standard and cites a single site as a proof point. That observation can motivate testing. It does not establish a transportable causal effect, costless reallocation or a guarantee that every site has spare capacity. A stated relative risk for failed problem-solving does not itself identify the effect of staffing more hours.

[AI] A parent missing-status category must not be confused with its nested same-driver-return subset. Returning a parcel does not by itself reveal original custody, fault or preventability. Moving a record from one defect label to another can change programme credit while leaving total customer failure unchanged.

RETAIN THE ORIGINAL CUSTOMER FRAME

[PD] Measure physical receipt against unchanged promises, later recovery and unresolved outcomes separately. Count diverted, recoded and delayed units. Price every added effort and adverse effect; preserve entitled remedies and do not use the roadmap to authorize individual sanctions.

[HR/AI] No historical labour trade-off, avoided cost or causal programme benefit is quantified in the selected record. These are required evaluation dimensions, not achieved findings.

EXHIBIT 5 / A CONTROLLED REVISION

A late correction must update the evidence bridge, not only the headline

[AI] A change propagated consistently through an application can still lack a valid source. Consistency of copies is different from correctness of the underlying value.

Revision step	Evidence retained
Record the requested correction	Source, event window, population, maturity, metric definition and requesting authority.
Compare with the frozen version	Numerator, denominator, actual/plan role and effect on every dependent calculation.
Rebuild the affected output	Traceable transformations and tests, with unresolved differences visibly retained.
Authorize publication	Named release owner, accepted exceptions and immutable as-of decision record.
Preserve the prior release	Version bridge and corrected interpretation, without erasing the original information set.

[HR/AI] The derivative record describes a late directed headline edit after a locked-basis rule and a preserved pre-edit version. It also reports that the replacement matched an earlier-year value. Equality raises a discriminating question; without the original change record it does not prove intent or conclusively establish the error's cause.

[AI] Multiple checks by agents using the same locked packet test consistency with that packet. They are not independent evidence that the packet is correct. A generated Q&A entry, a clean layout or a verification label cannot authenticate source accuracy, executive acceptance or operating delivery.

APPLY THE SAME CHANGE CONTROL TO EVERY ROLE

[PD] A senior request can be a valid correction. It still needs a defined basis and approval for its dependencies. Preserve both versions until the metric owner resolves the fork; do not silently relabel the old value as current.

[HR/AI] Exact application, source packs, verification ledger and sign-off artifacts remain unrecovered. Public wording does not authenticate derivative timestamps or imply an intentional misstatement.

EXHIBIT 6 / OWNERSHIP AND OUTCOME TIMING

Credit the review design without importing programme ownership or target failure

[HR/AI] The record describes a completed artifact and a plan held for approval. Those are different completion states, and neither establishes a mature operating result.

Claim	Supported boundary
Author of the review and line contract	Provisional registry-level hoop for framing, construction and measurement decisions.
Named goal owner	Shared accountability; not sole target authority or ownership of every programme.
Listed programme owner or regional lead	Separate delivery responsibility and acceptance evidence.
Later network performance update	Different scope and an observation before the named target date do not prove that target failed.
Related planning workbook	Context with unresolved version; cannot authenticate this review's release or sign-off.

[AI] A reported continuation of a mechanism is not the same as measured effectiveness. A booking line can persist while its causal value remains unidentified. A comparison against the wrong population, target basis or maturity window cannot serve as the missing target-date post-mortem.

[PD] Recover the actual submission and acceptance trail, the original roadmap matrix and its output equation. Define the follow-up cohort and target horizon before scoring delivery. Separate plan, forecast, activation, accepted ownership, realized outcome and personal contribution in the evidence ledger.

THE CASE IS A DESIGN AND RECONCILIATION CASE

[AI] It supports an auditable-plan lesson and exposes sufficiency gaps. No verified sign-off, historical cash benefit, mature causal improvement or authenticated target-date failure is claimed.

[HR/AI] Exact review and decision originals were not recovered in bounded searches. The source carries no historical dollar benefit; the model uses independently invented decision-resource units.

EXHIBIT 7 / THE COMMITMENT DECISION

Buy evidence only when it can change an affordable, authorized decision

[PD] IOSC joins action rights, resource availability, activation, output response and original exposure. It does not let a complete roadmap substitute for any of them.

Commitment gate	Responsible evidence owner
Goal and original horizon	Sponsor and measurement owner fix the target, population and exposure calendar.
Executable action and prerequisites	Programme and resource owners accept scope, dates and capacity.
Input-to-output response	Independent evaluation separates predicted association from intervention effect.
Economic and service choice	Finance and service owners price displaced work, false commitment and conditional release.
Publication and follow-up	Release owner retains uncertainty; measurement owner scores the original horizon.

[PD] The invented shadow model compares current-plan commitment, a costed authorized acceleration and a qualified roadmap. The paid rules sharply reduce false commitments relative to a naive monthly-sum rule, yet every offered package loses to free qualification at the main prices. Strong same-calendar and same-signal policies remain in the comparison.

[PD] Perfect knowledge of the constructed timing and response states is worth at most 0.70 per packet. Review fees above that ceiling cannot be justified by any classifier using the same decision menu. That bound is conditional on the prices and opportunities; it is not a universal argument against analysis.

NEXT ACTION: CLOSE THE RESPONSE AND CALENDAR GAPS

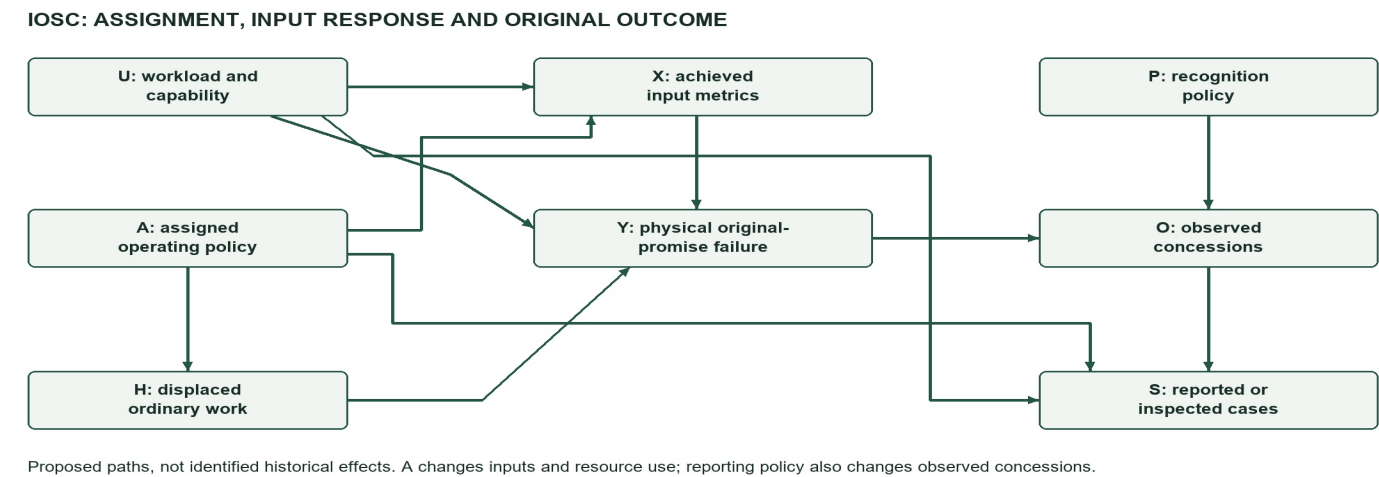
[PD] Validate the causal bridge, available capacity and exposure forecast, then register a test against the strongest feasible comparator. A larger sample can improve precision; it cannot repair a negative purchase value or create authority to make a promise.

[PD] All model probabilities, prices and allocations are hypothetical. Qualification is a conditional roadmap for one review cycle, not indefinite delay or altered customer entitlement.

TECHNICAL ANALYSIS

Make the causal bridge an explicit commitment

[PD] The Input-to-Outcome Sufficiency Contract (IOSC) joins a specific authorized action, its activation and resource requirements, the input response, the original-cohort output and a calibrated commitment rule.



[PD] A is the assigned operating policy; X the achieved inputs; U contextual workload and capability; H displaced work; Y physical original-promise failure; P recognition policy; O observed concessions. U affects X and Y. A may affect Y through X and through H. Y and P affect O. S selects reported or inspected cases. Randomizing A can identify its total effect; conditioning on achieved X or selected S generally requires extra assumptions.

[HR/AI] This child review translated an executive request into action, owner, mechanism, cadence and quantified contribution fields. That is a useful audit of plan completeness. A displayed gap remains unresolved; a filled field does not prove accepted ownership, available capacity or causal output sufficiency. Exact June originals were not recovered in bounded searches; the related planning artifact cannot authenticate this review, its final headline, sending or acceptance.

[AI/PD] Shmueli distinguishes prediction from causal explanation. Here the regression is cited, but coefficients, fitted population, residual uncertainty and intervention validation are not supplied. Composite weights are not causal slopes. VanderWeele shows why a strongly associated surrogate can still mislead about intervention direction. A coherent target table must therefore be accompanied by evidence that the authorized action changes the desired physical outcome, including other service pathways.

[HR/AI] Contribution is a provisional registry-level hoop for framing, synthesis and artifact direction. It does not establish sole programme ownership, authorship of the cited regression, approved execution or realized causal benefit.

Basis: historical record [HR], analytical interpretation [AI], proposed design [PD]. Shmueli (2010); VanderWeele (2013). Graph is a proposed causal model, not an identified historical result.

TECHNICAL ANALYSIS

Perfect prediction can imply three opposite decisions

[AI] Independent construction: unobserved U_1 , U_2 and U_3 are independent fair binary contextual factors. Each cell contains 1,000 original eligible units. Observed inputs X_i equal U_i , and failures Y equal 120 minus 10 times their sum.

[AI] All eight input combinations occur. The unique full-rank observational regression has intercept 120 and three slopes of -10, with zero residual error. These are exact population cells, not a small-sample fitted coincidence. The same complete observed distribution arises from each structural law below.

Structural world	Failure law	Mean after all inputs set to 1
Beneficial response	$Y = 120 - 10 \sum(X)$	90
Context-only proxy	$Y = 120 - 10 \sum(U)$	105
Adverse response	$Y = 120 - 20 \sum(U) + 10 \sum(X)$	120

[AI] In the observational regime $X=U$, every law reduces to $Y=120-10 \sum(U)$; mean failure count is 105. Intervene by setting all inputs to one while preserving U , original eligibility and each structural equation. The respective gains are +15, zero and -15 failures per 1,000 original units. All counterfactual counts remain valid nonnegative integers within the original cohort.

[AI] Full observed support for X does not establish causal exchangeability. U is unobserved; within each U stratum, natural X has no variation. A predictive coefficient is not an intervention coefficient. The proof does not establish that a historical input is useless or harmful; it establishes that the supplied kind of predictive evidence cannot choose among these explanations. More observations from the same unchanged observational regime do not remove this ambiguity.

[PD] Specify the actual action that can change the input, its direct and displaced-work pathways, and its feasible population. Validate total outcome effects first. Imai, Keele and Tingley explain why randomized treatment alone does not identify mediation through a nonrandom achieved input. An input goal can remain a diagnostic milestone while its output guarantee is held pending causal evidence.

Basis: wholly hypothetical structural proof [AI], independently enumerated across all eight contexts and eight input interventions. No historical numerical coordinates or fitted historical coefficients.

TECHNICAL ANALYSIS

Reconcile rates, recognition and the full resource consequence

[PD] Lock the population, metric, denominator, policy, time window and extraction vintage for every baseline, input, output and target. Preserve contradictory versions until the responsible owner resolves them.

[AI/PD] If period t has original exposure N_t and incremental reduction dt per million, implied prevented units are $\text{sum}(N_t * dt) / 1,000,000$. The corresponding horizon rate reduction is $\text{sum}(N_t * dt) / \text{sum}(N_t)$. Summing dt alone yields a rate-period quantity. Endpoint attainment, exposure-weighted horizon attainment and annual attainment are different claims. Activation delay and shared capacity must be integrated before a monthly contribution becomes an executable bridge.

[HR/RC] The supplied H2 component totals leave a residual even under their own additive rate-period reading; alternative operations totals create another basis fork. A booked-to-plan ratio does not isolate adherence from exposure, efficacy, timing or recognition. Earlier network/YTD performance cannot demonstrate a later business-unit weekly deadline miss. It can motivate concern without settling that endpoint.

[AI/PD] Policy subtraction defines an accounting scenario unless a no-policy counterfactual is identified. Recoding an event into another defect bucket does not establish total prevention. A later return cannot rescue a missed original promise. Track physical service, recognized concessions and access to remedies separately, with original identities retained through reclassification.

[PD] Validate input labels and physical outcomes using a stratified sample of successes, failures, unresolved and recoded records. Two coders blinded to assignment assess independent evidence; retain sampling weights, confusion matrices, prevalence and cluster-bootstrap kappa. Four hundred independent binary assessments give worst-case simple-agreement precision near plus/minus 4.9 percentage points, not kappa power or clustered precision.

[AI/PD] Holmstrom and Milgrom motivate measuring attention shifted across tasks. Price protected labor, backfill, late work, verification and ordinary-flow harm. Resource savings and released cash require separate evidence. This record supplies no historical dollar benefit or quantified displacement cost; a no-headcount request does not imply a costless intervention.

Basis: source reconciliation [HR/RC], dimensional identities [AI], measurement and costing proposals [PD]. Holmstrom/Milgrom (1991); Cohen (1960).

TECHNICAL ANALYSIS

Evaluate a roadmap decision and an operating policy separately

[PD] First test shadow commitment rules on original planning packets with common calendar and authority information. A later field trial asks whether an authorized operating change improves service.

[PD] The compact benchmark uses a fixed original exposure calendar, uncertain activation timing and uncertain incremental response. Compare immediate commitment, paid earlier-start commitment and a qualified release, with all fixed actions, the best common-calendar rule, the best rule using the same own-packet signal and fuller verification. Qualification is a useful bounded decision for this cycle, not an indefinite delay or a claim of zero value.

[PD] The shadow estimand is $E[L(\text{usual}) - L(\text{rule})]$ per original packet before subtracting review fee K in the same units. L prices false commitments, qualification and selected acceleration. The hidden constructed response is an assumed benchmark truth, not an empirically observed field counterfactual. Fees apply to every assigned packet; action costs apply whenever selected, including mistaken commitments. Lock the recommendation at intake plus 72 hours; adjudicate the fixed 90-day horizon 30 days afterward, with reference hidden from tuning.

[PD] For a subsequent operating trial, randomize independent labor-sharing station domains to the approved package or usual practice, stratifying on preassignment workload, stage mix, backlog and staffing capacity. Freeze original bookings and promises. Estimate $E[Y(\text{usual}) - Y(\text{package})]$ and resource burden per original booking, including ordinary work; keep customer remedies unchanged. Use preassignment expected activation timing and baseline mix as prespecified moderators. Achieved input values are mediators, not baseline strata.

[PD] Exact finite calculations audit the declared shadow law; normal planning remains conditional. Compare detection at economic break-even, MDE and full-response clearance separately. Recrutable domains, shared-domain variance, reference error and delay must be measured before choosing a confirmatory sample. Add equivalence or noninferiority limits for material ordinary-service harm only with approved margins. More sample size cannot repair an inferior comparator or unidentified causal response.

[PD] Hernan and colleagues motivate retaining all assigned originals instead of evaluating only returned parcels, achieved goals or reported failures. If later rollout is observational, require credible time-matched controls, pretrend and equivalence diagnostics, policy/maturity alignment and confounding sensitivity; a nonsignificant pretrend test does not establish comparability.

Basis: prospective design [PD], not a retrospective trial. Imai/Keele/Tingley (2010); Hernan et al. (2004); Hudgens/Halloran (2008). Exact model clocks are specified with its law.

TECHNICAL ANALYSIS

State what could refute the bridge and what remains unknown

[PD] A precise plan is testable. It becomes a defensible commitment only when its outcome, calendar, authority and resource claims survive the relevant tests.

Proposition	Prespecified rejection or hold rule
P1: predictive fit establishes input sufficiency	Reject the general inference using observationally equivalent opposite-effect worlds; hold the field causal claim without intervention evidence.
P2: monthly plan closure establishes horizon attainment	Reject when correctly exposure-weighted, activation-adjusted outcomes miss the stated horizon target; unresolved units or dates hold the bridge.
P3: protected staffing is a costless lever	Reject a zero-resource or no-displacement claim when measured incremental work or ordinary-service harm is material.
P4: selected-site improvement establishes effect	Require a credible same-eligibility comparison. Imprecision holds attribution; an upper bound below the claimed gain refutes its magnitude.
P5: paid review deserves adoption	Require full-cost decision gain against the strongest feasible cheaper rule plus an agreed absolute false-commitment tolerance.

[HR/RC] The private inventory contains 92 vectors: 22 source/sensitivity, 20 versions, 32 designs and 18 profiles. No source/sensitivity X or moving-range crossing establishes an operating signal; no stable phase is authenticated. Eleven X and six MR crossings occur in engineered ramps or an artifact-time profile. Tentative XmR uses 2.66 and 3.267. One mechanical two-network Laney calculation is ineligible; no eligible Laney panel or missing observations is invented.

[PD] Under an assumed linear conditional-mean model, expected untreated follow-up after selection at x is $m + \rho \times (x - m)$, where m is population mean and rho baseline persistence. This illustrates regression to the mean, not a measured correction for the selected site.

Limitations and identifying assumptions

[AI/PD] Exact original artifacts, reconciled headline, fitted regression, accepted sign-off and a matched endpoint postmortem are unavailable. The synthetic law requires calibrated activation and response likelihoods, stable authority, original exposure and priced losses. Physical reference finality, spillovers, competing work and transport across networks remain open. Neither published plan arithmetic nor a shadow decision improvement supplies the missing causal field evidence.

Basis: source/SPC scope [HR/RC], new proof [AI], falsification and limits [PD]. Barnett et al. (2005).

Test the commitment against its calendar and response assumptions

[PD] This entirely invented shadow benchmark prices a roadmap decision. It neither authorizes a programme nor estimates historical customer benefit.

[PD] Each original packet has three periods and observed exposure weights: early-heavy (0.6, 0.3, 0.1) or late-heavy (0.1, 0.3, 0.6), equally likely. The weighted target is 4. Hidden response/timing states are full/timely, full/late, half/timely and half/late, with probabilities 0.40/0.20/0.25/0.15. Timely reductions are 2/4/6; late reductions are 0/2/4, multiplied by response. Approved acceleration advances one period, capped at 6.

Evidence version	Timing / response accuracy by grade	Capture by grade	Fee
Base	0.85/0.65 for both flags	0.90/0.70	0.40
Transport stress	Timing 0.85/0.65; response 0.65/0.55	0.90/0.70	0.40
Corroborated	0.95/0.85 for both flags	0.95/0.85	0.90
Fuller review	0.98/0.95 for both flags	0.98/0.95	1.30

[PD] Three packets share an unobserved grade, with probabilities 0.70/0.30, within each independent domain. Capture and the two conditionally independent flag errors follow that grade. Decisions see only their own calendar and flags; grade, future truth and other packets' signals are hidden.

[PD] Usual commits the current plan because its monthly sum equals three times the target. The tested rule requires a positive response flag. With positive timing it accelerates early-heavy calendars and commits current late-heavy plans; with negative timing it accelerates only late-heavy plans. Otherwise it qualifies. All rules have the same approved acceleration rights; missing authority requires qualification outside this admissible fixture.

PRICE EVERY ORIGINAL DECISION

[PD] False full commitment costs 6; qualification costs 2; correct commitment costs 0. Acceleration costs 1 whenever selected, including mistakes. Review fees apply to every packet, including no capture. Qualification releases a conditional roadmap for this cycle; customer promises and remedies stay unchanged.

[PD] Invented resource units per packet. Freeze at intake; shadow lock +72 hours; constructed horizon 90 days, hidden reference 30 days later. Reference cannot reveal field counterfactuals without an identifying design. Common intake and research adjudication are separately budgeted.

PROPOSED EVALUATION / 2

Precision against a weak comparator is not a purchase case

[PD] Randomize independent domains, retaining every original packet. The declared shadow contrast compares the paid rule with usual commitment; it does not establish superiority to qualification.

Context	Domains / arm	80% MDE	Break-even fee	Zero-test power at break-even	Joint floor at full law
Base	16	1.406	0.40	12.0%	>99.9%
Base	128	0.499	0.40	61.3%	>99.9%
Transport stress	128	0.496	0.40	61.7%	>99.9%
Corroborated	256	0.354	0.90	>99.9%	>99.9%

[PD] The first three allocations fail 80% break-even detection. Corroborated passes conditionally but requires 512 domains and 1,536 packets; recruitment is unvalidated. Full-law joint power and break-even detection concern different effects.

[PD] For sensitivity, use the tested rule with probability λ per original packet, usual otherwise, while paying the full fee. Exact shared-grade moments give variance ceilings over λ from 0 to 1. The directional radius is $1.96 \cdot \sqrt{V_{\max} / n}$, with n domains per arm. Gates require loss reduction above the fee and wrong-commitment reduction above zero. MDE is beneficial zero-test detection, not an economic clearance threshold.

Independent-arm finite calculation	Base, 128 / arm	Corroborated, 256 / arm
Zero-null loss / risk tails	2.116% / 1.352%	2.223% / 1.387%
At break-even: positive zero-test detection	61.112%	>99.9%
At break-even: economic false clearance	2.415%	2.463%

[PD] Eighteen exact marginal tails check zero, economic-boundary and full laws. Joint floors use a union bound. Selected null tails are conservative, without uniform coverage. No minimum-sample search or Monte Carlo is used.

DO NOT CONFUSE RELATIVE WITH ABSOLUTE SAFETY

[PD] Relative improvement does not establish acceptable absolute risk. Free qualification has zero false full commitments and lower priced loss in every main context. More observations cannot reverse that mean comparison.

[PD] Normal planning is approximate; exact tails are conditional on the stated finite law. Prototype randomization, empirical likelihood calibration, authority and an independently agreed absolute-risk limit are prerequisites to confirmatory use.

Keep the cheaper decision and every adverse cost visible

[PD] Entries are total loss including review fees / unconditional false full-commitment risk. Lower is better.

Policy	Base	Transport stress	Corroborated
Usual current plan	4.8000 / 80.00%	4.8000 / 80.00%	4.8000 / 80.00%
Always accelerate	4.0000 / 50.00%	4.0000 / 50.00%	4.0000 / 50.00%
Free qualification / common-calendar optimum	2.0000 / 0.00%	2.0000 / 0.00%	2.0000 / 0.00%
Tested paid rule	2.3530 / 7.82%	2.6344 / 11.94%	2.5016 / 3.57%
Own-calendar + signal optimum	2.3027 / 4.64%	2.4000 / 0.00%	2.5016 / 3.57%
Fuller-review optimum	2.7194 / 1.44%	2.7194 / 1.44%	2.7194 / 1.44%

[PD] Enumerate 9 common-calendar and 59,049 own-calendar/signal mappings, including no capture; no cross-packet pooling. Base favorable reductions are 4.60911; added qualification/false-acceleration/reservation burdens are 1.76207. Net versus usual = $a * h * 4.60911 - 1.76207 - 0.40$ fee. Only favorable reductions receive realization a and overlap/transfer retention h . All adverse costs remain fully charged.

$a \setminus h$	0.60	0.80	1.00
0.50	-0.779	-0.318	0.142
0.75	-0.088	0.603	1.295
1.00	0.603	1.525	2.447

[PD] Prices are judgments. With qualification 3 and reservation 1, base costs 2.97717 versus common 3: the ordering narrowly reverses. With qualification 1, free qualification costs 1 versus tested 1.72876. Set prices prospectively; numerical checks cannot validate them.

REFERENCE ERROR CAN ERASE A NARROW ADVANTAGE

[PD] At main prices, variable loss lies in 0 to 7. Reassigning declared truth on epsilon of original packets in each arm permits up to $14 * \text{epsilon}$ loss-contrast error and $2 * \text{epsilon}$ risk error, with verified fees fixed.

[PD] The bound confines truth to declared response/timing states. New calendars, authority failures, costs, outcome support or interference require new bounds. These are decision-resource sensitivities, not cash savings or statistical confidence intervals.

PROPOSED INFORMATION VALUE / 4

Apply an information ceiling before paying for a better review

[PD] Within the locked action menu and prices, even perfect timing and response information cannot justify the higher review fees.

[PD] Perfect state knowledge chooses current only where it truly attains the target, acceleration where it adds sufficient exposure-weighted response, and qualification otherwise. Its expected terminal loss is 1.30, including 0.30 of reservation cost, versus 2.00 for free qualification. Therefore any information about these states is worth at most 0.70 before fees. This is a constructed-information bound, not an observable field oracle.

Offered evidence	Best gross information value	Fee	Net versus common rule
Base	0.097335	0.400000	-0.302665
Transport stress	0.000000	0.400000	-0.400000
Corroborated	0.398440	0.900000	-0.501560

[PD] Corroborated and fuller-review fees of 0.90 and 1.30 exceed the 0.70 ceiling. No better classifier alone can rescue those purchases under this fixture. Changing delay prices, authority, feasible actions or the response law changes the ceiling and requires a new decision analysis.

[PD] Separately, choose which validation study to run: response validation, calendar validation or defer. States are response-useful, calendar-useful and neither, with prior 0.45/0.35/0.20. Payoffs are (7, -2, -1), (-2, 6, -1) and (0, 0, 0). Without a signal the best expected utility is 2.25; perfect information adds at most 3.00.

Research scenario	Correct-state signal q	EVSI	Study cost	Net EVSI
Low	0.50	0.38750	0.50000	-0.11250
Base	0.75	1.59375	1.25000	0.34375
High	0.90	2.37500	2.25000	0.12500

RESEARCH VALUE IS A SEPARATE DECISION

[PD] Each signal assigns probability q to the true state and (1 - q) / 2 to each other state; all 27 contingent research policies are enumerated. Low/base/high are coherent alternative experiments, not a confidence interval. Do not add these utilities to operational savings or the roadmap information ceiling.

[PD] Proposed next step: calibrate calendar/response evidence, loss prices and reference validity; register a comparison against qualification before recruiting. A subsequent causal operating-policy trial would need original customer outcomes, full resource costs and its own power design.

REPRODUCIBILITY / READER PROTOCOL

Reproduce the bridge before approving the commitment

[AI] The Input-to-Outcome Sufficiency Contract makes action specificity a testable plan input. It requires separate evidence for execution feasibility, physical outcome response and the value of additional verification.

Contract layer	Required record and independent check
Measurement and version	Freeze original eligible identities, promises, denominators, outcome windows and recognition rules. Recalculate competing actuals against their own plans and periods. Preserve disagreements until the authorized owner resolves them.
Action and authority	Specify the executable action, accepted owner, decision rights, dependency, staffing and activation clock. A completed template or reported production event does not establish acceptance or implementation.
Input and causal response	Retain the fitted population, coefficients, residuals and validation data for any predictive relationship. Test the total effect of the feasible action, including pathways outside the targeted input and harm to ordinary work.
Calendar and exposure	Convert each period contribution into original outcome units using its exposure. Integrate activation lag, overlap and resource constraints. Keep endpoint targets, weighted horizon rates and sums of monthly rates distinct.
Commitment and information	Compare a conditional release, authorized acceleration and further validation under the same information rights. Charge attempted reviews and selected actions. Retain stronger cheaper comparators even when they overturn a favorable result against usual practice.

[PD] Rebuild the shadow law from its fixed calendar, hidden response states, signal likelihoods and loss table. Enumerate all states and policy observations; replicate means, shared-domain covariance, finite marginal tails, comparator choices and price sensitivities. Never substitute the constructed hidden truth for a historical counterfactual.

[AI] GAO schedule guidance motivates linked dependencies and resource-feasible timing. The proposed contract adds a separate causal response requirement and an information-purchase test. Neither a technically coherent schedule nor a perfect observational regression proves the full commitment.

[AI] The private archive retains source cells, raw alternatives, recovery limits, calculation code and internal reviews. Public examples are independently constructed; historical numerical coordinates and identity mappings remain private.

Protected source register; proposed IOSC reproduction protocol; GAO-16-89G. Each claim retains its historical, recalculated, analytical or proposed-design classification.

REFERENCES / 1

Primary sources and their role in this case

[AI] These sources inform prediction, causal pathways, measurement, schedule feasibility and decision value. They do not authenticate internal releases, acceptance or historical impact.

NIST/SEMATECH. e-Handbook of Statistical Methods, section 6.3.2.2: Individuals Control Charts. Official method. Distinguish sparse source observations from engineered target trajectories and valid operating phases.

Laney, D. B. (2002). Improved Control Charts for Attributes. *Quality Engineering*, 14(4), 531–537. Primary publisher abstract. Require aligned unique binary outcomes and repeated original exposures before a Laney chart.

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20, 37–46. Primary publisher record. Validate input classification and original-promise outcomes with blinded repeated coding.

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. A selected poor-performing station can improve without a policy effect.

Hudgens, M. G., & Halloran, M. E. (2008). Toward Causal Inference with Interference. *Journal of the American Statistical Association*, 103, 832–842. Original manuscript. Randomize shared labor domains and retain ordinary-work spillovers in the target outcome.

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 rollout requires credible group-time contrasts with stable outcome definitions.

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. Value readiness and response evidence against strong common-calendar alternatives.

Original author, journal, institutional and official methods sources. Access scope and application retained privately. Source support is distinct from internal historical evidence.

REFERENCES / 2

Primary sources and their role in this case

[AI] These sources inform prediction, causal pathways, measurement, schedule feasibility and decision value. They do not authenticate internal releases, acceptance or historical impact.

Collier, D. (2011). Understanding Process Tracing. *PS: Political Science & Politics*, 44, 823–830. Primary publication. Separate artifact direction, sign-off, implementation and realized effect in the contribution chain.

Tucker, A. L., & Edmondson, A. C. (2003). Why Hospitals Don't Learn from Failures. *California Management Review*, 45(2), 55-72. Author manuscript. Treat repeated execution problems as possible capacity and process constraints.

The W. Edwards Deming Institute. The Funnel Experiment. Official exposition. A planned downward ramp is not a control limit or evidence of a stable process.

Campbell, D. T. (1976). Assessing the Impact of Planned Social Change. Original institutional record. Track recoding and recognition changes alongside physical outcomes under metric pressure.

Goodhart, C. A. E. (1992). The Objectives for, and Conduct of, Monetary Policy in the 1990s. Reserve Bank of Australia Annual Conference. Original author paper, section 1(b)(i). Revalidate the link from a managed input target to the desired output.

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. Evaluate the causal mechanism and resource trade-offs behind a performance scorecard.

OECD / European Commission JRC (2008). Handbook on Constructing Composite Indicators. Official handbook. Composite weights are design choices, not causal input-response coefficients.

Original author, journal, institutional and official methods sources. Access scope and application retained privately. Source support is distinct from internal historical evidence.

REFERENCES / 3

Primary sources and their role in this case

[AI] These sources inform prediction, causal pathways, measurement, schedule feasibility and decision value. They do not authenticate internal releases, acceptance or historical impact.

Shmueli, G. (2010). To Explain or to Predict? *Statistical Science* 25(3), 289-310. doi:10.1214/10-STS330. Separate predictive fit from the causal effect of changing an input.

VanderWeele, T. J. (2013). Surrogate Measures and Consistent Surrogates. *Biometrics* 69(3), 561-565. doi:10.1111/biom.12071. Require assumptions linking input changes to final outcomes; strong association does not exclude a surrogate paradox.

Imai, K., Keele, L., and Tingley, D. (2010). A General Approach to Causal Mediation Analysis. *Psychological Methods* 15(4), 309-334. doi:10.1037/a0020761. Estimate total effects first; input mediation requires additional identification and sensitivity analysis.

Hernan, M. A., Hernandez-Diaz, S., and Robins, J. M. (2004). A Structural Approach to Selection Bias. *Epidemiology* 15(5), 615-625. doi:10.1097/01.ede.0000135174.63482.43. Retain the original assigned cohort rather than conditioning effects on attained goals or reported problem cases.

Holmstrom, B., and Milgrom, P. (1991). Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design. *Journal of Law, Economics, and Organization* 7(special issue), 24-52. doi:10.1093/jleo/7.special_issue.24. Measure work displaced by protecting a favored metric; resource reallocation can change other service dimensions.

U.S. Government Accountability Office (2015). *Schedule Assessment Guide: Best Practices for Project Schedules*, GAO-16-89G. Best Practices 3 and 8. A credible schedule connects dependencies, resources, uncertainty and completion confidence. A prescribed finish date does not establish feasibility. Applied as schedule-governance guidance, not validation of this case or its causal assumptions.

Original author, journal, institutional and official methods sources. Access scope and application retained privately. Source support is distinct from internal historical evidence.

Approve the next test, not an unsupported guarantee

[AI] Author-led, agent-assisted review; not external peer review. The case establishes a reproducible boundary between a specific action plan and a defensible outcome commitment.

Major issue	Resolution and remaining boundary
Detailed action fields treated as sufficiency	Separate template completeness, accepted accountability, feasible activation and causal response. The historical record supports framing and analytical direction at a provisional contribution grade.
Regression fit used as an intervention guarantee	The full-support structural proof preserves the entire observational distribution while intervention effects change sign. State which new design could distinguish those worlds.
Monthly values summed into an annual outcome	Retain rate-period arithmetic privately. Require exposure weights, compatible scope, timing and overlap before asserting a horizon effect. A listed gap is not a universal capability ceiling.
Metric improvement or policy subtraction sold as impact	Distinguish routine variation, recognition changes, population changes and physical outcomes. Earlier network performance cannot settle a later business-unit endpoint.
A paid review compared only with a weak baseline	Retain free qualification, all common-calendar choices, the best same-signal rule, fuller review and the information ceiling. Preserve adverse purchase results and price reversals.
Power or internal review treated as validation	Conditional precision concerns the stated law. It neither repairs an inferior comparator nor authenticates the source. Physical service effects require their own prospective evaluation.

Partner critique: can a COO act on page one?

[AI] Yes. Name the accountable decision owner, reconcile the outcome and calendar basis, verify authorized capacity, and qualify any unresolved output promise. Purchase a validation step only when its expected decision gain exceeds its full cost against the strongest feasible cheaper rule.

Remaining limits and minor issues

[AI] The technical layer consolidates source recovery, final approval, confounding, ascertainment, selection, displacement and transport limits. Version and unit corrections have been retained, not replaced with an invented final number. Internal review cannot certify adoption, field efficacy, realized savings or a professional ranking.

REVIEW DISPOSITION / PASS WITHIN STATED SCOPE

[AI] Publish the operating-design reconciliation, falsifiable sufficiency contract and conditional shadow evaluation. Preserve unresolved historical outcome impact and require a field test before promising the causal gain.

Final source, methods, independent numerical, partner and actual-page reviews. Internal publication readiness is distinct from external assurance or outcome validation.