BEYOND COST SENSITIVITY: OUTCOME-AWARE OFF-POLICY EVALUATION VIA LOCAL DISCLOSURE

Anonymous authors
Paper under double-blind review

ABSTRACT

We study off-policy evaluation under local disclosure when policies attach fixed fees or subsidies to strategic actions. Ratio-only disclosure groups applicants with the same cost sensitivity $\alpha = \Gamma/V$, although their absolute stakes and default risks can differ. We characterize when this grouping retains enough information for evaluation. For zero-price targets, covered outcome averages by action and α suffice. For a priced target, two worlds can agree under every zero-transfer experiment over a fixed library yet have default risks zero and one. We use randomized cardinal anchors to label outcome-relevant (V, Γ) cells. A common shift of the displayed transfers preserves choices and matches the focal transfer to the target. Horvitz-Thompson moments then consistently estimate approval, risk, and value. One frozen log also supports uniform evaluation of a finite predeclared policy class after data-dependent selection. In the coarse independent experiment, direct labels improve the minimax mean squared error from $\Theta((Npg^2)^{-1})$ to $\Theta((Npg)^{-1})$. Our budget analysis accounts for recruitment, follow-up, and gross subsidies. A complete priced-menu simulation evaluates eight policies from three logging menus and verifies transfer matching. Further experiments recover the rare-cell rates and test scalar-type insufficiency with public credit covariates.

1 INTRODUCTION

Banks and public agencies use off-policy evaluation, or OPE, to assess policies before deployment. Applicants may pay old debt or move savings after learning how approval works. Their submitted covariates then depend on the policy (Hardt et al., 2016; Perdomo et al., 2020).

With local information disclosure, or LID, personalized truthful recommendations reveal an interval for hidden cost sensitivity α. Vo et al. (2026) use these intervals to learn its distribution and construct a one-shot, strategy-aware OPE estimator. Individual response costs remain unobserved.

We study a fixed fee or subsidy t_a attached to an action with truthful approval p_a and burden d_a. Type (V, Γ) has utility

$$U_a = Vp_a - \Gamma d_a - t_a. \tag{1}$$

At zero price, choices depend only on $\alpha = \Gamma/V$. Consider b with $(p, d, t) = (.4, 0, 0)$ and a with $(p, d, t) = (.8, .2, 1)$. Types $(1, 1)$ and $(10, 10)$ share $\alpha = 1$, for which $U_a - U_b = .2V - 1$. They choose different actions. Absolute scale affects target behavior and can affect default. Here V measures willingness to pay, distinct from poverty or welfare.

For zero-price targets, covered outcome averages by action and α suffice even when risk varies within an α fiber. For a priced target, two worlds can agree under every zero-transfer experiment over a fixed library yet have default risks zero and one. Vo et al. (2026) already allow general scalar outcomes. We study the cardinal heterogeneity that monetary targets make relevant to these outcomes.

When does ratio-only LID suffice, and how can one budgeted log evaluate and select among priced policies? Calibration randomizes the monetary terms in equation 1 to expose a type cell through choice. A common shift matches the focal target transfer without changing choices. Under selected-action transfer exclusion, randomized approval identifies the default mass of the labelled (V, Γ) cell. These moments are estimated once and reused for policy selection. Figure 1 illustrates the scale alias and calibrated cell.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

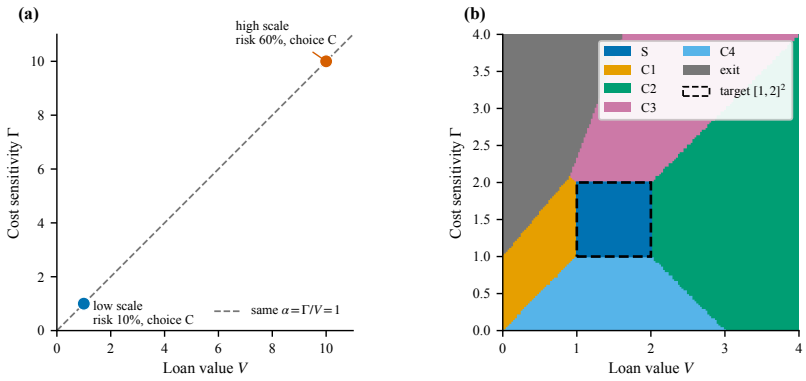


Figure 1: Scale alias and direct cell disclosure. Left, types $(1, 1)$ and $(10, 10)$ share $\alpha = 1$. Their choices agree without a fee and separate with one. Right, a five-offer menu labels $[1, 2]^2$. A common shift matches any focal target transfer.

We make the following contributions.

- We characterize the boundary between zero-price sufficiency and priced-target non-identification in Propositions 1 and 2.
- We introduce focal transfer matching in Lemma 1. Theorem 1 characterizes when a truthful action library can label a bounded outcome-relevant cell.
- We establish consistent risk and value estimation and a uniform finite-policy guarantee from one frozen log in Theorems 2 and 3. In the declared coarse-log model, Theorem 4 improves the rare-cell minimax rate from $\Theta((Npg^2)^{-1})$ to $\Theta((Npg)^{-1})$ through direct labels. Corollary 1 gives the corresponding budget criterion.

2 Outcome-Aware Strategic OPE under Local Disclosure

We follow the one-shot lending interaction of Vo et al. (2026). Payments before reapplication modify the application, for example by paying old debt. They do not repay the new loan whose default we observe.

2.1 Problem formulation

An applicant has base covariates $X^b \in \mathcal{X}$, with finite $\mathcal{X}$ for exposition. Policy π draws the initial approval $T^b \sim \text{Bernoulli}(\pi(X^b))$.

Local feedback. After rejection, $T^b = 0$, explanation e supplies a finite set of truthful offers $\mathcal{A}_e^+(X^b)$ and a displayed no-loan action $\varnothing$. Each offer $a \in \mathcal{A}_e^+$ specifies approval $p_{\pi,a}$, primitive burden d_a, and possibly a policy transfer $t_{\pi,e,a}$, such as old-debt matching or a fixed application fee. Calibration assigns a transfer vector k, positive for fees and negative for subsidies. The no-loan action has $p_\varnothing = d_\varnothing = 0$. Its transfer moves with the other alternatives in this closed-menu experiment.

Strategic response. The applicant has an unobserved monetary-equivalent type $\theta = (V, \Gamma) \in \mathcal{T} \subset (0, \infty) \times \mathbb{R}_+$. The type determines the choice

$$S(\theta; e, k) \in \arg\max_{a \in \mathcal{A}_e} \{u_a(\theta; k) := V p_{\pi,a} - \Gamma d_a - k_a\}. \tag{2}$$

Here V values the loan opportunity in money units, and Γ measures sensitivity to the disclosed burden in the same units. Targets use $k = t_{\pi,e}$; calibration uses the assigned vector. With equal transfers, dividing by V gives $p_{\pi,a} - \alpha d_a$ with $\alpha = \Gamma/V$, recovering the ratio-only LID response. Unequal fixed-money transfers make absolute scale relevant to choice. Ties have zero probability or follow a fixed rule.

Loan outcome. If $S = a \in \mathcal{A}_e^+$, final approval T is drawn with probability $p_{\pi,a}$. Choosing $\varnothing$ ends the interaction without a loan. Let $D_a(k_a) \in \{0, 1\}$ denote potential default on the new loan after offer a is selected and approved at realized transfer k_a. Only TD is observed.

Our primary estimand is default among target-policy approvals. A bank may also assign known gain G_a to repayment and loss L_a to default, yielding payoff $(1 - D)G_a - DL_a$ for an approved loan. The same quantities identify this value. We write D^{base} for default after initial approval, before disclosure.

Observed data and objective. A post-rejection calibration record is

$$O_i = (X_i^b, T_i^b, M_i, S_i, T_i, T_i D_i), \tag{3}$$

where M_i is the assigned menu. All menu, initial-approval, and final-approval propensities are known. Initially approved observations also record $T_i^b D_i^{\text{base}}$. From one frozen log $\mathcal{D}_0 = \{O_i\}_{i=1}^N$, the decision maker estimates approval mass A_π, approved-default mass B_π, risk $R(\pi) = B_\pi / A_\pi$, and linear payoff $J(\pi)$ for policies declared before calibration.

We impose the following design and outcome assumptions.

Assumption 1 (One reusable pre-selection log)**.** *A finite candidate class $\Pi = \{\pi_1, \ldots, \pi_K\}$ is declared before calibration. Before seeing any calibration outcome, the decision maker commits to one distribution over a finite menu library. Each iid subject receives one menu and contributes at most one downstream outcome. The log is then frozen and an arbitrary, possibly data-dependent $\widehat{\pi} \in \Pi$ is selected from it; no outcome collected after selection is used to evaluate or select $\widehat{\pi}$. Target-policy menus are not prohibited from the library, but each included menu consumes assignment probability before the target is known.*

For $K = 1$ with a target fixed before calibration, direct rollout is the natural benchmark. We evaluate candidate classes from one shared log before selection.

Assumption 2 (Closed-menu response)**.** *Every enrolled applicant chooses from exactly the displayed alternatives, including $\varnothing$, with a fixed tie rule. There is no additional outside option whose utility remains unshifted when the experimental menu is shifted.*

Assumption 3 (Randomized calibration and outcome observation)**.** *Applicants and their potential variables are iid. Within each stratum, initial and final approval draws are independent across applicants. Calibration menus follow a predeclared independent categorical design or complete randomization with positive arm fractions. Allocation is independent of types and potential defaults. Conditional on $(X^b, \theta, S = a, e)$, final approval is independent of $D_a(k_a)$, with known probability $p_a \geq p_{\min} > 0$.*

Assumption 4 (Selected-action transfer exclusion)**.** *For an enrolled applicant, the potential default distribution given $(X^b, \theta, S = a)$ depends on the menu only through the transfer actually attached to a. Unchosen offers have no direct effect. Thus a focal action with $k_a = t$ reveals $q_a^t(\theta, x) := \mathbb{P}(D_a(t) = 1 \mid \theta, X^b = x)$.*

Assumption 4 is weaker than outcome neutrality, but rules out effects of unchosen offers on default. Such offers may change financial behavior. Appendix Proposition 5 gives a randomized falsification test. Non-rejection does not validate transport.

2.2 Policy value decomposition and target outcome moments

Evaluation requires each choice probability and its default mass. For policy π, let $\tau_\pi(e \mid x)$ be the explanation distribution and $C_{\pi,x,e,a}$ the type region choosing a under transfers $t_{\pi,e}$. Define

$$g_{\pi,x,e,a} = F_x(C_{\pi,x,e,a}), \qquad m_{\pi,x,e,a} = \int_{C_{\pi,x,e,a}} q_a^{t_{\pi,e,a}}(\theta, x)\, dF_x(\theta), \tag{4}$$

where F_x is the conditional type distribution. For the initially approved branch, let $m_x^{\text{base}} = \mathbb{P}(D^{\text{base}} = 1 \mid X^b = x)$. Summing over the initial decision, explanation, and selected offer gives

$$A_\pi = \sum_x P_X(x)\Big[\pi(x) + (1-\pi(x)) \sum_e \tau_\pi(e \mid x) \sum_{a \in \mathcal{A}_e^+} p_{\pi,a} g_{\pi,x,e,a}\Big], \tag{5}$$

$$B_\pi = \sum_x P_X(x)\Big[\pi(x) m_x^{\text{base}} + (1-\pi(x)) \sum_e \tau_\pi(e \mid x) \sum_{a \in \mathcal{A}_e^+} p_{\pi,a} m_{\pi,x,e,a}\Big]. \tag{6}$$

The target default rate is

$$R(\pi) = B_\pi / A_\pi, \tag{7}$$

provided $A_\pi > 0$. Equations 5 and 6 separate the observed base distribution and known policy, explanation, and approval probabilities from hidden-region masses g and default masses m. These moments depend on type region, action, and target transfer, replacing the observed-state outcome regression in Vo et al. (2026). Applicants with the same observed action and α can have different hidden scales. Ratio-only LID can therefore recover strategic action probabilities without recovering the default mass needed by a priced target.

We first isolate the missing information, then characterize when heterogeneity at fixed α preserves identification.

Proposition 1 (Ratio-only equivalence with missing outcome coverage)**.** *There exist two data-generating processes with the same law under every zero-transfer menu from a fixed action library, but different default rates under a priced target policy. This remains true when $\alpha = \Gamma/V$ is known exactly.*

Two indistinguishable worlds. Take one stratum, zero initial approval, and equally likely types $(1,1)$ and $(10,10)$. The library contains free exit, $b : (p,d) = (.4,0)$, and $a : (.8,.2)$. World A assigns risks $(q_a, q_b) = (0,1)$ to the low type and $(1,0)$ to the high type; world B reverses them. Risks are transfer-neutral. Every zero-transfer submenu selects the same action at both scales and reveals the same equal mixture of risks. At $t_a = 1, t_b = 0$, $u_a - u_b = .2V - 1$. The low type chooses b, the high type chooses a, and both prefer their choice to exit. Approval mass is .6 in both worlds, yet every approved type defaults in A and none in B. Appendix Table 3 gives the complete observed law.

Proposition 2 (When the scalar response type is sufficient)**.** *Suppose every target transfer is zero, so choice is a known function of (x, e, α), and the log identifies F_x^α and, for every target-used action, $\bar{q}_a(x,\alpha) = \mathbb{E}[q_a^0(V,\Gamma,x) \mid X^b = x, \alpha]$ under a target-stable conditional type law. Suppose further that standard logging overlap and outcome stability identify P_X and the initially approved moment m_x^{base} in every target-used stratum, and that $A_\pi > 0$. Then $(A_\pi, B_\pi, R(\pi))$ are identified without recovering (V,Γ) jointly.*

Covered fiber-average outcomes suffice for zero-price targets without joint-scale recovery. Fixed-money incentives can separate types with the same α; calibrated cells measure their action outcomes at the target transfer.

2.3 Learning outcome-relevant hidden cells

An action-specific monetary anchor makes absolute scale visible in the response. Write $z_a = (p_a, -d_a)$, so $u_a(\theta; k) = \theta^\top z_a - k_a$. Observing $S = s$ places the type in the polyhedral censoring region

$$C_s(k) = \bigcap_{j \neq s} \{\theta : (z_s - z_j)^\top \theta \geq k_s - k_j\}. \tag{8}$$

The region is the two-dimensional analogue of the interval construction in Vo et al. (2026).

Randomized two-dimensional thresholds identify the joint type distribution under a rank and support condition. Appendix Proposition 3 gives the result and a bivariate censored-choice likelihood analogous to Vo et al. (2026). Our estimator uses only the finite type-cell moments needed by candidate policies.

Default can depend on transfers. A common shift matches focal coordinate k_s to target transfer t without changing any choice.

Lemma 1 (Focal transfer matching)**.** *For any menu k, focal action s, and desired target transfer t, define $k^{\langle s,t\rangle} = k + (t - k_s)\mathbf{1}$. Every type makes the same choice under $k^{\langle s,t\rangle}$ and k, while $k_s^{\langle s,t\rangle} = t$.*

The shift adds the same constant to every utility. The menu therefore preserves the type label, and under Assumption 4, approved focal choosers reveal $D_s(t)$. Randomized approval identifies the corresponding default mass at the target transfer. Zero-transfer centering is the special case $t = 0$.

Whether the focal choice can identify a bounded cell depends on the truthful action library.

Theorem 1 (Bounded focal cell)**.** *Over the full plane $\mathbb{R}^2$, there exists a transfer vector for which the nonempty focal region $C_s(k)$ is bounded if and only if*

$$z_s \in \operatorname{int}\operatorname{conv}\{z_j : j \neq s\}. \tag{9}$$

Any such menu has a choice-equivalent version matched to any focal transfer t. With a relative-transfer cap $|k_j - k_s| \leq H$, the desired cell is feasible only when its defining offsets obey that cap.

The equivalence follows from known power-cell geometry (Allevi et al., 2024) and does not require knowledge of the type support. On $\mathbb{R}^2_+$ or known compact support, the condition remains sufficient but need not be necessary. Appendix A.1 gives a five-offer construction. Section 3 uses these target-matched cells to estimate response and outcome moments without recovering individual types.

3 Outcome-Aware LID Estimator

The menus in Section 2 reveal hidden cell membership. We estimate the resulting response and default masses once, then reuse them across the predeclared target class Π.

Definition 1 (Reusable outcome basis)**.** *A collection $\mathcal{H} = \{(a_\ell, t_\ell, B_\ell, M_\ell)\}_{\ell=1}^L$ is Π-covering if each focal-matched M_ℓ has transfer t_ℓ and satisfies $\mathbf{1}\{S(M_\ell) = a_\ell\} = \mathbf{1}\{\theta \in B_\ell\}$ F almost surely, and if every $(A_\pi, B_\pi, J(\pi))$ is a known affine function of $g_\ell = F(B_\ell)$, $m_\ell = \int_{B_\ell} q_{a_\ell}^{t_\ell}(\theta)dF(\theta)$, and shared base-stratum moments. Target transfers lie on the finite calibrated grid. Continuous transfers require approximation or outcome smoothness.*

3.1 Cell mass and target-matched default mass

Fix base stratum x, action a, target transfer t, and type cell B. A direct menu $M_{a,t,B}$ labels the cell at the focal transfer:

$$\mathbf{1}\{S(M_{a,t,B}) = a\} = \mathbf{1}\{\theta \in B\}, \qquad k_a(M_{a,t,B}) = t. \tag{10}$$

Broad menus $M_1, \ldots, M_L$ can also recover B when their focal-choice indicators span its indicator pointwise:

$$\mathbf{1}\{\theta \in B\} = \sum_{\ell=1}^{L} c_\ell \mathbf{1}\{S(M_\ell) = a\}, \qquad k_a(M_\ell) = t. \tag{11}$$

Rectangle inclusion and exclusion gives such a span. Direct labelling has $L = 1$ and $c_1 = 1$. Appendix C.2 certifies rank, residual, and conditioning by atomizing a finite library. Appendix C.7 gives signed-span constants.

Let $N_x = \sum_i \mathbf{1}\{X_i^b = x\}$ count logged applicants in stratum x and $n_x^- = \sum_i \mathbf{1}\{X_i^b = x, T_i^b = 0\}$ those reaching calibration. The logging initial-approval probability is $\pi_0(x)$. Rejected applicants receive menu M_ℓ with known probability $\rho_\ell(x) > 0$. For $N_x > 0$, Assumptions 3 and 4 give

$$\widehat{g}_{a,B,x} = \frac{1}{N_x} \sum_{i:X_i^b = x, T_i^b = 0} \sum_{\ell=1}^{L} \frac{c_\ell \mathbf{1}\{M_i = M_\ell, S_i = a\}}{[1 - \pi_0(x)]\rho_\ell(x)}, \tag{12}$$

$$\widehat{m}_{a,B,x} = \frac{1}{N_x} \sum_{i:X_i^b = x, T_i^b = 0} \sum_{\ell=1}^{L} \frac{c_\ell \mathbf{1}\{M_i = M_\ell, S_i = a\} T_i D_i}{[1 - \pi_0(x)]\rho_\ell(x) p_{\ell,a}(x)}. \tag{13}$$

The weights correct for initial rejection, menu assignment, and selective default observation after final approval. Thus

$$\mathbb{E}[\widehat{g}_{a,B,x}] = F_x(B), \qquad \mathbb{E}[\widehat{m}_{a,B,x}] = \int_B q_a^t(\theta, x) dF_x(\theta). \tag{14}$$

When $F_x(B) > 0$, the ratio consistently estimates average cell risk. The linear masses remain defined for an empty cell. Appendix Proposition 5 tests exclusion using randomized duplicate menus: unequal default masses for the same certified cell and transfer reject the restriction. Non-rejection does not establish target transport.

The initially approved branch uses ordinary inverse propensity weighting, given in Appendix C. As in ordinary OPE, it requires outcome stability. Conditional on base stratum and hidden type, logging and target approval reveal the same potential default on the new loan.

3.2 Target default rate and bank value

Suppose a finite calibration partition refines every target choice region $C_{\pi,x,e,a}$. We obtain $\widehat{A}_\pi$ and $\widehat{B}_\pi$ by substituting equation 12, equation 13, equation 26, and $\widehat{P}_X(x)$ into equation 5 and equation 6. The outcome-aware LID estimator is their ratio:

$$\widehat{R}_{\text{OA-LID}}(\pi) = \widehat{B}_\pi / \widehat{A}_\pi. \tag{15}$$

The masses are unbiased; their ratio is consistent but not exactly unbiased. Known repayment payoff G_a and default payoff $-L_a$ give bank value

$$J(\pi) = \sum_x P_X(x) \Big\{ \pi(x) \big[G_{\text{base}} - (G_{\text{base}} + L_{\text{base}}) m_x^{\text{base}} \big] + (1 - \pi(x)) \sum_e \tau_\pi(e \mid x) \sum_{a \in \mathcal{A}_e^+} p_{\pi,a} \big[G_a g_{\pi,x,e,a} - (G_a + L_a) m_{\pi,x,e,a} \big] \Big\}. \tag{16}$$

Replacing (P_X, g, m) by $(\widehat{P}_X, \widehat{g}, \widehat{m})$ gives $\widehat{J}_{\text{OA-LID}}(\pi)$, separating approval volume, default risk, and payoff. We now establish identification and consistency.

Theorem 2 (Identification and consistency)**.** *Suppose iid applicants satisfy $P_X(x) > 0$ on finite $\mathcal{X}$; target explanation sets, action sets, and calibration spans are finite; all required menu, initial-decision, and approval propensities are known and bounded away from zero; Assumptions 3 and 4 and target outcome stability hold; and $A_\pi \geq a_0 > 0$. If equation 11 represents every target region indexed by type and action, then $(A_\pi, B_\pi, R(\pi), J(\pi))$ are identified. Moreover,*

$$(\widehat{A}_\pi, \widehat{B}_\pi, \widehat{R}_{\text{OA-LID}}(\pi), \widehat{J}_{\text{OA-LID}}(\pi)) \xrightarrow{p} (A_\pi, B_\pi, R(\pi), J(\pi)). \tag{17}$$

Theorem 3 (Uniform evaluation from one reusable log)**.** *Suppose $\mathcal{X}$ and the number of blocks indexed by action and stratum are fixed, $P_X(x) \geq x_0 > 0$, and logging initial approval and rejection probabilities are at least $r_0 > 0$ in every target-used stratum. Suppose also that a direct Π-covering basis has L independently and equally assigned arms, final approval at least p_0, mutually exclusive cells within each block, aggregation weights in $[0, 1]$, and $\inf_{\pi \in \Pi} A_\pi \geq a_0$. Under Assumptions 1 through 4 and target outcome stability, with probability at least $1 - \delta$,*

$$\max_{\pi \in \Pi} \max\{|\widehat{A}_\pi - A_\pi|, |\widehat{B}_\pi - B_\pi|\} \leq C \left[\sqrt{\frac{L \log(4K/\delta)}{N p_0}} + \frac{L \log(4K/\delta)}{N p_0} \right] =: \varepsilon_N, \tag{18}$$

where C depends only on (x_0, r_0) and the fixed blocks. If $\varepsilon_N < a_0/2$, then $\max_{\pi \in \Pi} |\widehat{R}_\pi - R(\pi)| \leq 4\varepsilon_N / a_0$, including for a data-dependent selected policy.

For a direct nonnegative basis, the leading square-root term gives a $\sqrt{K/L}$ gain over independent K-policy rollout. Appendix C.7 gives coefficient-dependent constants for signed spans and the optimized benchmark for well-conditioned target spans.

For continuous boundaries, Appendix C.6 bounds the additional risk error by $2\epsilon_\pi / a_0$, where ϵ_π is the target-weighted mass of the boundary approximation error.

3.3 THE RARE-CELL INFORMATION GAP

How much information does a direct label provide? We compare the focal choice in Table 2, which reveals hidden cell membership, with independent broad-menu observations under the following coarse protocol.

Definition 2 (Coarse independent experiments)**.** *Fix a known design $\mathfrak{d}$ containing F, a cell B of mass $0 < g \le g_0$, focal transfer t, approval $p \ge p_0$, and menu quantities. Each iid subject sees one menu, so $O = (M, S, T, TD)$ has law $P^{\otimes N}$ and contains neither θ, cell membership, a refining covariate, nor repeated responses. The direct experiment has one focal menu matched to t with $\mathbf{1}\{S = s\} = \mathbf{1}\{\theta \in B\}$. The broad experiment has fixed L, arm probabilities at least ρ_0, coefficients satisfying $\max_\ell |c_\ell| \le C_c < \infty$ and giving equation 11 F almost surely, and background mass at least $2\beta_0 > 0$ in every arm touching B. Only the focal risk q_s^t varies nonparametrically, with cell mean in $[1/4, 3/4]$.*

For $E \in \{D, W\}$, let $\mathcal{P}_E(\mathfrak{d})$ be the induced coarse-log laws and define $\mathfrak{R}_E(N; \mathfrak{d}) = \inf_{\widehat{q}} \sup_{P \in \mathcal{P}_E(\mathfrak{d})} \mathbb{E}_{P^{\otimes N}} (\widehat{q} - q_B)^2$.

Theorem 4 (Direct versus broad rare-cell rates)**.** *For every admissible fixed design with $g_0 < 1 - 2\beta_0$ and $Npg^2 \ge 1$, constants depending only on the fixed bounds give*

$$\mathfrak{R}_D(N; \mathfrak{d}) \asymp \frac{1}{Npg}, \tag{19}$$

$$\mathfrak{R}_W(N; \mathfrak{d}) \asymp \frac{1}{Npg^2}. \tag{20}$$

A direct cell default label occurs with probability gp, so the sample mean gives the upper bound. A change Δ in q_B gives KL divergence $\Theta(pg\Delta^2)$. A broad arm dilutes this change to $g\Delta$ in a background mixture with mass bounded away from zero, giving KL divergence $\Theta(pg^2\Delta^2)$. Signed inversion using Horvitz-Thompson weights attains the matching upper bound. Appendix C gives the two-point lower bounds and finite-submodel constants.

The separation holds for the coarse independent protocol. Observed membership, coupled menus, shrinking broad regions, and outcome smoothness give different information regimes. To see the dilution, a 1% cell changes a 50% mixture by only $0.01/0.50$ times its own risk change. A direct label preserves the full contrast and saves order $1/g$ in sample size.

3.4 BUDGET-AWARE DISCLOSURE DESIGN

Direct labels can reduce enrolment while subsidizing nonfocal choices. Budget efficiency depends on recruitment, outcome follow-up, and menu payments.

Cost of a matched menu. Matching focal action a from k_a to target t sends action b to $k_b + t - k_a$. Its gross subsidy $(k_a - k_b - t)_+$ has expectation

$$Q_{a,t} = \sum_b \varsigma_b (k_a - k_b - t)_+, \tag{21}$$

where ς_b is the choice share. Fees do not offset subsidies. The caps $|k_b - k_a| \le H$ and $|t| \le T$ give $Q_{a,t} \le H + T$ and, at $t = 0$, $Q_{a,0} \le H(1 - \varsigma_a)$. For design $m \in \{D, W\}$, the per-applicant cost is

$$C_m = c_m + \lambda Q_m, \tag{22}$$

where c_m covers recruitment, logging, and outcome follow-up. Expected budget M supports $N_m \simeq M/C_m$ subjects. Deterministic cost caps in Appendix D give a conservative hard-budget sample size.

Corollary 1 (Information per expected budget)**.** *For $E \in \{D, W\}$, let p_E be its final approval probability and let $0 < C_E < \infty$. Set $N_E(M) = \lfloor M/C_E \rfloor$ and define $\mathfrak{R}_E^{\mathrm{bud}}(M; \mathfrak{d}) = \mathfrak{R}_E(N_E(M); \mathfrak{d})$. If $M \ge 2 \max_E C_E$ and $N_E(M) p_E g^2 \ge 1$, Theorem 4 gives*

$$\mathfrak{R}_D^{\mathrm{bud}}(M; \mathfrak{d}) \asymp \frac{C_D}{M p_D g}, \qquad \mathfrak{R}_W^{\mathrm{bud}}(M; \mathfrak{d}) \asymp \frac{C_W}{M p_W g^2}, \quad \frac{\mathfrak{R}_D^{\mathrm{bud}}(M; \mathfrak{d})}{\mathfrak{R}_W^{\mathrm{bud}}(M; \mathfrak{d})} \asymp \frac{p_W}{p_D} g \frac{C_D}{C_W}. \tag{23}$$

With uniform theorem constants and bounded approval ratios, $C_D/C_W = o(1/g)$ favors direct labels. Enrolment is fixed under expected cost, not realized payments; Appendix D.3 gives the constant sandwich and sufficient inequalities.

Broad inversion requires order $1/g$ more subjects in this comparison. Only localization costs of the same order erase the budget gain. Appendix E.7 gives a normalized crossing. Reuse favors the basis when $LC_{\text{basis}} < KC_{\text{rollout}}$, with an optimized benchmark from a well-conditioned target span.

Pilot choice shares estimate Q. Appendix D gives certification, explicit constants, hard-budget bounds, and cost-weighted allocation. Related work studies adaptive instrument assignment (Chandak et al., 2024; Oprescu et al., 2025) and budget-constrained experimental design (Yamin et al., 2025).

4 Experiments

Controlled experiments test identification, priced-menu evaluation, reuse, and rare-cell rates. The semi-synthetic study uses public credit covariates. All strategic types and outcomes are simulated. Appendix E gives the protocols.

4.1 Controlled synthetic data

Priced menus. Uniform types on $[.5, 2] \times [.25, 2]$ choose among three priced offers and free exit. We simulate initial approval, utility choice, final approval, and default at the selected transfer. Three logging menus cover eight policies through five action-specific moments. One region uses signed coefficients $(1, 1, -1)$. Appendix E.4 gives the menus and coverage certificate; Appendix E.5 gives the direct hidden-atom check.

Baselines and evaluation. Comparators use independent eight-policy rollout or pool those same records across matched cells. Optimized three-target collection is observation-equivalent to the reusable design. At six sizes with 1,000 repetitions each, we measure maximum risk error and regret for $J = A - 3B$. Gross subsidies follow realized choices and are not offset by fees.

With fixed cost $c = .02$, expected budget 1374.48 admits 48,000 reusable-design applicants and 48,912 rollout applicants. Mean maximum risk errors are .00670, .01286, and .00740 for reusable, independent, and pooled estimation. Mean regrets are .001120, .004470, and .001535. Figure 2 shows both criteria; Appendix Table 6 reports Monte Carlo uncertainty and equal-enrollment results.

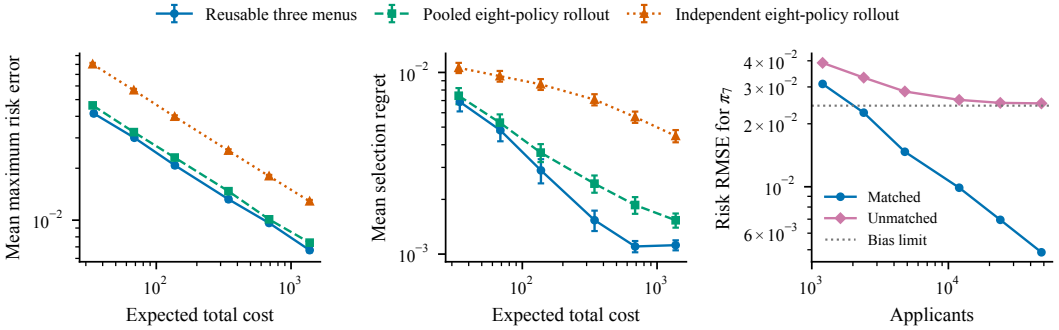


Figure 2: Complete priced-menu simulation with 1,000 repetitions per size. Left and center: risk error and selection regret at equal expected cost, including gross subsidies; bars show 1.96 Monte Carlo standard errors. Right: shifting one logging menu preserves choices but mismatches target transfers. The risk error for π_7 approaches a nonzero bias.

Transfer matching. Subtracting .1 from every E3 transfer, including exit, preserves choices but lowers each approved action's default probability by .03. Reusing these outcomes at the original target prices contaminates pooled duplicate cells and gives maximum asymptotic risk bias .02441. At $N = 48{,}000$, mean maximum error is .02466, versus .00670 with matched transfers. The mismatched population objective selects π_6 instead of the true optimum π_4.

Scalar type insufficiency. Exact enumeration verifies that identical zero-transfer laws can give priced-target risks 1 and 0. A common $.4$ shift preserves all $25{,}921$ checked choices and matches the focal transfer.

Rare-cell and budget laws. With $10{,}000$ runs per configuration, the fitted MSE exponents in g are -1.011 for direct labels and -1.927 for broad measurements. They match the predicted exponents -1 and -2. The normalized budget check crosses at $c = .002$ for $g = .01$, $Q_D = .198$, and $Q_W = 0$.

At $N = 200{,}000$, duplicate menus yield selection and exclusion-null p-values $.580$ and $.583$. An injected $.16$ context effect gives contrast $.00936$ against analytic $.01$, with $p < 8 \times 10^{-33}$.

4.2 Semi-synthetic credit covariates

We use eight selected covariates from UCI South German Credit (UCI, 2019) and simulate types, menus, approvals, and defaults. Approval and inverse weighting use the selected offer's probability. The direct cell estimator is compared with a favorable oracle-aided α-only baseline at six sample sizes, with 400 repetitions each. Target risk is $.26524$, while scale averaging gives the baseline limit $.39386$. At $N = 20{,}000$, the respective RMSEs are $.00640$ and $.12900$, as shown in Appendix Figure 6.

5 Related Work

Strategic evaluation. Classical and uniform OPE evaluate changes in action assignment under a fixed environment (Dudík et al., 2011; Uehara et al., 2020; Yin & Wang, 2021). In strategic learning, applicants can also change their submitted covariates in response to the policy (Hardt et al., 2016; Perdomo et al., 2020). Vo et al. (2026) use LID to learn scalar sensitivity for general outcomes. Our design exposes absolute scale and identifies target-matched action-outcome moments under prices. Strategic decision and treatment methods instead optimize policies or response costs (Tsirtsis & Gomez-Rodriguez, 2020; Rosenfeld & Rosenfeld, 2024; Tsirtsis et al., 2024; Munro, 2025). They do not identify these moments from one reusable pre-policy LID log.

Disclosure and recourse. Common shifts and affine-menu geometry complement tools from screening, strategic data procurement, and cost-aware collection (Rochet & Choné, 1998; Frongillo & Kash, 2021; Allevi et al., 2024; Cai et al., 2015; Chandak et al., 2024; Harding et al., 2026; Yamin et al., 2025). Our design connects these tools to outcome-aware LID through a truthful recourse library. Each calibration menu's relative transfers expose an outcome cell and determine validity and subsidy. Closed-menu participation, transfer exclusion, suitable geometry, and affordability govern this construction. Fees can instrument demand (Ai et al., 2024). Deployed assessment rules can serve as instruments for causal recovery from strategic responses (Harris et al., 2022). Recourse supplies feasible actions (Tsirtsis & Gomez-Rodriguez, 2020; Vo et al., 2026). Here transfers expose absolute scale and the library labels the target-matched cell.

6 Conclusion and Discussion

Zero-price targets may need only covered outcomes indexed by action and α. Priced targets can require absolute scale. Cardinal anchors and focal matching produce target-valid cells whose Horvitz-Thompson moments uniformly estimate approval, risk, and value for a predeclared class. Designed disclosure creates outcome information before policy choice. Randomized duplicate menus make the outcome restriction refutable.

The guarantees apply to truthful finite libraries, closed shifted menus, transfer exclusion, and stable outcomes. The rare-cell separation concerns the coarse independent model, where direct labels remove one rare-cell factor under the budget criterion. Direct rollout is the natural benchmark for one predetermined target. Well-conditioned spans can attain the direct-basis reuse rate. Extensions include continuous-price grids, legally compliant menus, welfare objectives, and dynamic responses.

AI USE STATEMENT

Large language models were used for language polishing.

ETHICS STATEMENT

This work studies pre-deployment evaluation rather than prescribing a lending policy. Any application involving fees or subsidies must comply with applicable law and review participation, liquidity, welfare, consent, fairness, and outcome impacts. Experiments use simulated outcomes, with public covariates entering only the labelled semi-synthetic study.

REPRODUCIBILITY STATEMENT

Appendices state assumptions and proofs. Result tables accompany the experimental panels and distinguish simulated outcomes from observed public covariates.

REFERENCES

Rui Ai, David Simchi-Levi, and Feng Zhu. Dynamic service fee pricing under strategic behavior: Actions as instruments and phase transition. In *Advances in Neural Information Processing Systems*, volume 37, pp. 53525–53570, 2024. doi: 10.52202/079017-1695. URL https://proceedings.neurips.cc/paper_files/paper/2024/hash/6008facedac886858acdb76c09da2097-Abstract-Conference.html.

Elisabetta Allevi, Juan Enrique Martínez-Legaz, and Rossana Riccardi. On the basic properties and the structure of power cells. *Journal of Optimization Theory and Applications*, 203(2):1246–1262, 2024. doi: 10.1007/s10957-024-02435-0. URL https://link.springer.com/article/10.1007/s10957-024-02435-0.

Yang Cai, Constantinos Daskalakis, and Christos Papadimitriou. Optimum statistical estimation with strategic data sources. In *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pp. 280–296. PMLR, 2015. URL https://proceedings.mlr.press/v40/Cai15.html.

Yash Chandak, Shiv Shankar, Vasilis Syrgkanis, and Emma Brunskill. Adaptive instrument design for indirect experiments. In *International Conference on Learning Representations*, pp. 46639–46688, 2024. URL https://proceedings.iclr.cc/paper_files/paper/2024/hash/cbca4c1890be5a49b440b258c18ab39e-Abstract-Conference.html.

Miroslav Dudík, John Langford, and Lihong Li. Doubly robust policy evaluation and learning. In *Proceedings of the 28th International Conference on Machine Learning*, pp. 1097–1104, 2011. URL https://icml.cc/2011/papers/554_icmlpaper.pdf.

Rafael M. Frongillo and Ian A. Kash. General truthfulness characterizations via convex analysis. *Games and Economic Behavior*, 130:636–662, 2021. doi: 10.1016/j.geb.2021.09.010. URL https://www.sciencedirect.com/science/article/pii/S0899825621001330.

Michael O. Harding, Vikas Singh, and Kirthevasan Kandasamy. Learning from biased and costly data sources: Minimax-optimal data collection under a budget (extended abstract). In *Proceedings of the 39th Conference on Learning Theory*, volume 336 of *Proceedings of Machine Learning Research*, pp. 3184–3184. PMLR, 2026. URL https://proceedings.mlr.press/v336/harding26a.html.

Moritz Hardt, Nimrod Megiddo, Christos H. Papadimitriou, and Mary Wootters. Strategic classification. In *Proceedings of the 2016 ACM Conference on Innovations in Theoretical Computer Science*, pp. 111–122. ACM, 2016. doi: 10.1145/2840728.2840730. URL https://doi.org/10.1145/2840728.2840730.

Keegan Harris, Dung Daniel T Ngo, Logan Stapleton, Hoda Heidari, and Steven Wu. Strategic instrumental variable regression: Recovering causal relationships from strategic responses. In *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pp. 8502–8522. PMLR, 2022. URL https://proceedings.mlr.press/v162/harris22a.html.

Evan Munro. Treatment allocation with strategic agents. *Management Science*, 71(1):123–145, 2025. doi: 10.1287/mnsc.2022.01629. URL https://pubsonline.informs.org/doi/10.1287/mnsc.2022.01629.

Miruna Oprescu, Brian Cho, and Nathan Kallus. Efficient adaptive experimentation with noncompliance. In *Advances in Neural Information Processing Systems*, volume 38, pp. 42462–42498. Curran Associates, Inc., 2025. doi: 10.52202/085713-1414. URL https://proceedings.neurips.cc/paper_files/paper/2025/hash/3ca380c5fe9f174a71a230478741169f-Abstract-Conference.html.

Juan Perdomo, Tijana Zrnic, Celestine Mendler-Dünner, and Moritz Hardt. Performative prediction. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pp. 7599–7609. PMLR, 2020. URL https://proceedings.mlr.press/v119/perdomo20a.html.

Jean-Charles Rochet and Philippe Choné. Ironing, sweeping, and multidimensional screening. *Econometrica*, 66(4):783–826, 1998. doi: 10.2307/2999574. URL https://www.jstor.org/stable/2999574.

Elan Rosenfeld and Nir Rosenfeld. One-shot strategic classification under unknown costs. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 42719–42741. PMLR, 2024. URL https://proceedings.mlr.press/v235/rosenfeld24a.html.

Stratis Tsirtsis and Manuel Gomez-Rodriguez. Decisions, counterfactual explanations and strategic behavior. In *Advances in Neural Information Processing Systems*, volume 33, pp. 16749–16760. Curran Associates, Inc., 2020. URL https://proceedings.neurips.cc/paper/2020/hash/c2ba1bc54b239208cb37b901c0d3b363-Abstract.html.

Stratis Tsirtsis, Behzad Tabibian, Moein Khajehnejad, Adish Singla, Bernhard Schölkopf, and Manuel Gomez-Rodriguez. Optimal decision making under strategic behavior. *Management Science*, 70(12):8506–8519, 2024. doi: 10.1287/mnsc.2021.02567. URL https://pubsonline.informs.org/doi/10.1287/mnsc.2021.02567.

UCI. South german credit. UCI Machine Learning Repository, 2019. URL https://archive.ics.uci.edu/dataset/522/south+german+credit. Dataset.

Masatoshi Uehara, Masahiro Kato, and Shota Yasui. Off-policy evaluation and learning for external validity under a covariate shift. In *Advances in Neural Information Processing Systems*, volume 33, pp. 49–61. Curran Associates, Inc., 2020. URL https://proceedings.neurips.cc/paper/2020/hash/0084ae4bc24c0795d1e6a4f58444d39b-Abstract.html.

Kiet Q. H. Vo, Abbavaram Gowtham Reddy, Julian Rodemann, Siu Lun Chau, and Krikamol Muandet. Off-policy evaluation with strategic agents via local disclosure. In *International Conference on Machine Learning*, 2026. doi: 10.48550/arXiv.2606.07308. URL https://arxiv.org/abs/2606.07308. Accepted at ICML 2026; proceedings volume and page numbers were not yet available on 2026-09-01.

Khurram Yamin, Edward Kennedy, and Bryan Wilder. Dependent randomized rounding for budget constrained experimental design. In *Proceedings of the Forty-first Conference on Uncertainty in Artificial Intelligence*, volume 286 of *Proceedings of Machine Learning Research*, pp. 4681–4700. PMLR, 2025. URL https://proceedings.mlr.press/v286/yamin25a.html.

Ming Yin and Yu-Xiang Wang. Optimal uniform OPE and model-based offline reinforcement learning in time-homogeneous, reward-free and task-agnostic settings. In *Advances in Neural Information Processing Systems*, volume 34, pp. 12890–12903, 2021. URL https://proceedings.neurips.cc/paper/2021/hash/6b3c49bdba5be0d322334e30c459f8bd-Abstract.html.

A Notation and Additional Modelling Discussion

Table 1: Core notation.

symbol	meaning
$X^b = x$	pre-strategic applicant covariates observed before local disclosure
$\varnothing; a, b, s$	the displayed no-loan action; genuine offers, with s focal
p_a, d_a, t_a, k_a	approval, primitive old-debt modification burden, target transfer, and calibration transfer for action a
V, Γ	monetary-equivalent loan-opportunity value and burden sensitivity
$\alpha = \Gamma/V$	relative cost sensitivity visible to an unpriced choice model
S, T, D	selected action, final approval, and new-loan default indicator
$q_a^t(\theta, x)$	default probability after action a with realized transfer t
$C_a(k)$	hidden-type region selecting action a under transfer menu k
$g_B, m_{a,t,B}$	cell mass and its action-a, transfer-t default mass
$A_\pi, B_\pi, R(\pi)$	target approval mass, approved-and-default mass, and their ratio
g, p, N	target cell mass, outcome-observation probability, and number of subjects
Q, c, C, M, H	expected gross subsidy, nontransfer per-person cost, total per-person cost, expected total budget, and relative-transfer cap

Scope of cardinal interpretation. Suppose primitive psychological utility is $vp_a - \gamma d_a - \mu k_a$, where μ is the individual's positive marginal utility of money. The model then identifies $V = v/\mu$ and $\Gamma = \gamma/\mu$. The common psychological scale (v, γ, μ) is observationally irrelevant. Thus V measures willingness to pay in the experiment's monetary unit. It is distinct from poverty, medical urgency, social welfare, and moral deservingness.

Old debt versus the prospective loan. d_a represents a modification undertaken before reapplication, such as paying an old balance, increasing documented savings, or completing verification. Default D concerns the subsequently issued new loan. Focal matching reproduces target transfer t_a; it does not erase d_a or assume the transfer has no effect on default.

Closed menus and participation. Focal matching shifts every displayed alternative, including $\varnothing$. An external option whose utility does not shift can change menu participation. The closed-menu analysis therefore corresponds to voluntary enrollment in a calibration study. Every displayed alternative receives the displayed credit or charge, including "take no new loan". Enrollment design and ethical review remain application-level requirements.

Target coverage. The indicator-span condition in equation 11 is an experimental analogue of overlap. The convex-hull condition in Theorem 1 guarantees that some bounded focal region can exist. Exact finite evaluation applies to target regions represented by the indicator span; Equation 31 covers finite approximations.

Relation to the outcome model of Vo et al. (2026). Default is an instance of the general scalar outcome Y allowed by Vo et al. (2026). Their transport step models the conditional outcome using the observed final action and covariates together with exogenous noise. Our motivating case allows outcome-relevant heterogeneity within an observed action and scalar response type. It focuses on calibration data that expose the corresponding outcome mass indexed by type and action.

Proposition 3 (Joint type identification)**.** *Suppose each identification menu contains exactly three displayed alternatives (j_0, s, j_1) and let $A = [(z_s - z_{j_0})^\top; (z_s - z_{j_1})^\top]$. If A is nonsingular and an independent two-dimensional threshold assignment has positive density on a rectangle covering the support of $A\theta$, then one-shot choices identify $F_x(V, \Gamma)$.*

Optional structural type model. An analogous structural implementation can posit $(\log V, \log \Gamma) \mid X^b = x \sim \mathcal{N}_2(\mu_\eta(x), \Sigma_\eta)$ and maximize

$$\widehat{\eta} \in \arg\max_{\eta} \sum_i \log \int_{C_{S_i}(k_i)} f_\eta(\theta \mid X_i^b)\, d\theta. \tag{24}$$

The structural model is useful when the action library does not directly realize every cell. Standard sufficient conditions require a compact parameter space containing the truth and almost-everywhere continuity in η. Realized censoring regions must have positive probability near the truth. The log likelihood must also have an integrable envelope and a unique expected maximizer. Proposition 3 supplies design-level identification when threshold support covers type support. The consistency argument assumes correct log-normal specification.

A.1 Five-offer hand-checkable closed menu

The five-offer menu below imposes exactly $V \geq 1$, $V \leq 2$, $\Gamma \leq 2$, and $\Gamma \geq 1$. Its exit does not bind, so

$$S = s \quad \Longleftrightarrow \quad (V, \Gamma) \in [1, 2] \times [1, 2], \tag{25}$$

under a focal tie rule, or almost surely when boundary faces have zero mass. A common shift can charge the focal group any target transfer while preserving this label.

Table 2: A five-offer closed menu that directly labels a hidden type cell. Negative k denotes a subsidy. The displayed exit shifts with every offer.

action	approval p	old-debt burden d	transfer k
focal s	0.80	0.50	0.00
c_1	0.60	0.50	-0.20
c_2	1.00	0.50	0.40
c_3	0.80	0.30	0.40
c_4	0.80	0.70	-0.20
exit $\varnothing$	0.00	0.00	0.30

B Proofs for Identification and Disclosure Geometry

This appendix conditions on a fixed base stratum x whenever doing so simplifies notation.

B.1 Proof of Proposition 1

Let X^b be constant, initial approval be zero, and the type be $(1, 1)$ or $(10, 10)$ with probability $1/2$ each. Both have $\alpha = 1$. The closed library contains $\varnothing : (p, d) = (0, 0)$ and two genuine offers $b : (0.4, 0)$ and $a : (0.8, 0.2)$. Under any nonempty zero-transfer submenu of $\{a, b\}$, utility differences equal V times the same normalized difference, so both scales select the same offer. The chosen offer also has positive utility and beats $\varnothing$.

In world A, set $(q_a, q_b) = (0, 1)$ for the low type and $(1, 0)$ for the high type. In world B, reverse these four values and let them be invariant to transfer. Any genuine offer selected by a zero-transfer submenu has the same approval probability across types. Its approved default law is Bernoulli$(1/2)$ in both worlds. Thus menu assignment, choice, approval, and observed default have the same one-subject law under every zero-transfer experiment over this fixed library. The iid logs have the same product law.

Table 3: Observed law under every nonempty zero-transfer submenu. The selected action is deterministic for a given submenu. Each row holds in both worlds.

selected action	$\mathbb{P}(T = 0, TD = 0)$	$\mathbb{P}(T = 1, TD = 0)$	$\mathbb{P}(T = 1, TD = 1)$
b	$3/5$	$1/5$	$1/5$
a	$1/5$	$2/5$	$2/5$

For the priced target set $(t_b, t_a) = (0, 1)$. Then

$$u_a - u_b = (0.8 - 0.4)V - 0.2\Gamma - 1 = 0.2V - 1.$$

The low type chooses b and the high type chooses a. Both offers beat $\varnothing$. The common approval mass is $\frac{1}{2}(0.4) + \frac{1}{2}(0.8) = 0.6$. Both selected types default in world A and neither defaults in world B.

Target default among approvals is therefore 1 and 0, proving non-identification from any ratio-only log. □

B.2 Proof of Proposition 2

For each (x, e), let $a_\pi(x, e, \alpha)$ denote the target-selected action. Since its choice regions are measurable in α, the post-rejection default contribution equals

$$\int p_{\pi, a_\pi(x,e,\alpha)} \bar{q}_{a_\pi(x,e,\alpha)}(x, \alpha)\, dF_x^\alpha(\alpha).$$

The analogous expression without $\bar{q}$ gives approval mass. Every integrand and F_x^α is identified by assumption. The stated logging overlap and outcome stability identify P_X and m_x^{base}, so summing over (x, e) identifies A_π and B_π, hence $R(\pi)$ when $A_\pi > 0$. No joint scale distribution is needed. □

B.3 Proof of Proposition 3

For the transfer vector $(k_{j_0}, k_s, k_{j_1}) = (0, c_0, c_0 - c_1)$,

$$u_s - u_{j_0} = (z_s - z_{j_0})^\top \theta - c_0,$$
$$u_s - u_{j_1} = (z_s - z_{j_1})^\top \theta - c_1.$$

Because the menu contains exactly these three actions and ties have probability zero,

$$S = s \iff W_0 \geq c_0,\ W_1 \geq c_1, \qquad W = A\theta.$$

Random thresholds are independent of type, hence

$$\mathbb{P}(S = s \mid C_0 = c_0, C_1 = c_1, X^b = x) = \mathbb{P}(W_0 \geq c_0, W_1 \geq c_1 \mid x).$$

Threshold support covering the support of W identifies its joint survival function, which uniquely determines its probability law. If A is nonsingular, the measurable bijection $\theta = A^{-1}W$ identifies F_x.

For necessity in an unrestricted two-dimensional type model, suppose $\operatorname{rank}(A) < 2$. There is nonzero h with $Ah = 0$. Every pairwise action contrast lies in the row span of A, since $z_{j_0} - z_{j_1} = (z_s - z_{j_1}) - (z_s - z_{j_0})$. Thus sufficiently small perturbations θ and $\theta + \varepsilon h$ in the positive orthant produce identical utilities up to action-independent constants under every transfer vector, yet define different type distributions. □

Structural-likelihood consistency. Under the regularity conditions listed in Appendix A, the per-observation log likelihood in equation 24 is continuous and dominated. A uniform law of large numbers gives uniform convergence of the empirical criterion to its expectation. Identification and the Kullback-Leibler inequality make the true parameter the unique expected maximizer. The argmax theorem therefore gives $\widehat{\eta} \xrightarrow{p} \eta_0$. This is the standard censored-choice M-estimation argument. The finite-cell estimator does not rely on it.

B.4 Proof of Lemma 1

For every action j,

$$u_j(\theta; k + (t - k_s)\mathbf{1}) = u_j(\theta; k) - (t - k_s).$$

Adding the same constant to all utilities leaves their argmax set unchanged. The focal component is $k_s + (t - k_s) = t$. Relative differences also remain unchanged, as does a relative-transfer cap. □

B.5 Proof of Theorem 1

Write $v_j = z_s - z_j$ and $c_j = k_s - k_j$. The focal region is the polyhedron

$$P_s = \{\theta : v_j^\top \theta \geq c_j\ \forall j \neq s\}.$$

Its recession cone is

$$\operatorname{rec}(P_s) = \{h : v_j^\top h \geq 0\ \forall j \neq s\}.$$

A nonempty polyhedron is bounded if and only if its recession cone is $\{0\}$. By finite-dimensional separation,

$$\{h : v_j^\top h \geq 0 \; \forall j\} = \{0\} \quad \iff \quad \operatorname{cone}\{v_j : j \neq s\} = \mathbb{R}^2 \quad \iff \quad 0 \in \operatorname{int}\operatorname{conv}\{v_j : j \neq s\}.$$

Since $\operatorname{conv}\{z_s - z_j\} = z_s - \operatorname{conv}\{z_j\}$, the last condition is equivalent to $z_s \in \operatorname{int}\operatorname{conv}\{z_j : j \neq s\}$.

If the condition holds, choose any interior point θ_0 and offsets $c_j = v_j^\top \theta_0 - \varepsilon$ for small $\varepsilon > 0$. Then θ_0 is a strict feasible point and the recession condition makes the region bounded. Set $k_s = 0, k_j = -c_j$, then apply Lemma 1 to match any desired focal transfer. Conversely, any nonempty bounded focal polyhedron has trivial recession cone and satisfies the convex-hull condition. □

B.6 Verification of the five-offer closed-menu cell

Using Table 2, the four focal comparisons are

$$u_s - u_{c_1} = 0.2V - 0.2, \qquad u_s - u_{c_2} = -0.2V + 0.4,$$
$$u_s - u_{c_3} = -0.2\Gamma + 0.4, \qquad u_s - u_{c_4} = 0.2\Gamma - 0.2.$$

They are all nonnegative exactly when $1 \leq V \leq 2$ and $1 \leq \Gamma \leq 2$. On this rectangle the focal offer also beats exit because $u_s - u_\varnothing = 0.8V - 0.5\Gamma + 0.3 \geq 0.1$. Outside it, a listed competitor already beats s, so exit cannot enlarge the focal region. The displayed representative has focal transfer zero. A common shift matches any target transfer. The approval values are a common 0.3 upward shift of the simplest genuine-offer construction. This adds $0.3V$ to every genuine-offer utility without changing their pairwise choices.

C Proofs for Estimation and Statistical Rates

C.1 Identification of the cell moments

Condition on $X^b = x$ and on reaching the LID calibration stage. For any arm ℓ,

$$\mathbb{E}\left[\frac{\mathbf{1}\{M = M_\ell, S = a\}}{\rho_\ell(x)}\middle| x\right] = \mathbb{P}(S(M_\ell) = a \mid x).$$

Multiplying by c_ℓ, summing, and applying the pointwise span equation 11 gives $F_x(B)$, proving the first identity in equation 14.

For the outcome moment, iterated expectation, randomized approval, and selected-action transfer exclusion give

$$\mathbb{E}\left[\frac{\mathbf{1}\{M = M_\ell, S = a\}TD}{\rho_\ell(x)p_{\ell,a}(x)}\middle| x, \theta\right]$$
$$= \mathbf{1}\{S(M_\ell;\theta) = a\}\mathbb{E}[D_a(t) \mid x, \theta] = \mathbf{1}\{S(M_\ell;\theta) = a\}q_a^t(\theta, x).$$

The focal transfer is t in every arm. Summing the signed indicators and integrating yields $\int_B q_a^t dF_x$, proving the second identity.

Multiplying the main-text estimator by $\widehat{P}_X(x) = N_x/N$ gives exactly the all-log contribution

$$\frac{\mathbf{1}\{X_i = x, T_i^b = 0, M_i = M_\ell, S_i = a\}}{[1 - \pi_0(x)]\rho_\ell(x)}$$

for mass and additionally $T_i D_i / p_{\ell,a}$ for default mass. Since initial randomization depends only on observed x, their averages are unbiased for $P_X(x)$ times the conditional moments above. The analogous all-log base contribution is $\mathbf{1}\{X_i = x, T_i^b = 1\}D_i^{\text{base}}/\pi_0(x)$. The uniform theorem uses these exact estimators.

For the initially approved branch, ordinary inverse propensity weighting gives the default mass when $\pi_0(x) > 0$:

$$\widehat{m}_x^{\text{base}} = \frac{1}{N_x}\sum_{i:X_i^b = x}\frac{T_i^b D_i^b}{\pi_0(x)}. \tag{26}$$

C.2 FINITE COVERING-BASIS CERTIFICATE

Fix a block $b = (x, a, t)$. Let $C_{b1}, \dots, C_{bJ_b}$ be the distinct target regions used by Π and let

$$G_{br} = \{\theta : S(M_{br}; \theta) = a\}, \qquad r = 1, \dots, R_b,$$

be candidate focal regions with $k_a(M_{br}) = t$. Their membership patterns induce finitely many atoms $\mathcal{A}_{b1}, \dots, \mathcal{A}_{bQ_b}$. The fixed tie rule assigns boundary points. A support-wide certificate keeps every geometrically nonempty atom. Rows known to have zero F_x mass may instead be removed.

Define the zero-one matrices

$$(H_b)_{qj} = \mathbf{1}\{\mathcal{A}_{bq} \subseteq C_{bj}\}, \qquad (\Phi_b)_{qr} = \mathbf{1}\{\mathcal{A}_{bq} \subseteq G_{br}\}.$$

Proposition 4 (Finite covering certificate)**.** *For a menu subset $\mathcal{B}_b \subseteq \{1, \dots, R_b\}$, every target indicator lies in the focal span on the retained atoms if and only if*

$$\operatorname{rank}[\Phi_{b,\mathcal{B}_b} \;\; H_b] = \operatorname{rank}(\Phi_{b,\mathcal{B}_b}). \tag{27}$$

If $\Phi_{b,\mathcal{B}_b}$ has full column rank, the coefficient matrix is unique and equals

$$C_b = \Phi^{\dagger}_{b,\mathcal{B}_b} H_b, \qquad H_b = \Phi_{b,\mathcal{B}_b} C_b. \tag{28}$$

The same coefficients identify target mass. Under Assumption 4, they identify target-matched calibration default mass because every retained menu has focal transfer t. Target outcome stability transports this moment to deployment.

Proof. Each target and focal indicator is constant on every atom. The pointwise indicator identities are therefore the matrix equation $H_b = \Phi_{b,\mathcal{B}_b} C_b$. Such a matrix exists exactly when every column of H_b belongs to the column space of $\Phi_{b,\mathcal{B}_b}$. This is equivalent to equation 27. Full column rank gives the unique least-squares solution in equation 28. Transfer matching permits the same multiplication after every focal indicator is weighted by q^t_a. □

For a finite candidate library, the minimum basis size is

$$L^{\star}_b = \min_{\mathcal{B} \subseteq \{1, \dots, R_b\}} \{|\mathcal{B}| : \operatorname{rank}[\Phi_{b,\mathcal{B}} \;\; H_b] = \operatorname{rank}(\Phi_{b,\mathcal{B}})\}.$$

Exhaustive subset search certifies this minimum. Column-pivoted QR can screen subsets, but acceptance requires exact row reduction and $\|H_b - \Phi_{b,\mathcal{B}_b} C_b\|_{\max} = 0$. Backward deletion certifies inclusion minimality, not minimum cardinality.

The certificate reports

$$\chi_{b,1} = \max_j \|C_{b,:,j}\|_1, \qquad \chi_{b,2} = \max_j \|C_{b,:,j}\|_2, \qquad \chi_{b,\infty} = \max_j \|C_{b,:,j}\|_\infty.$$

For a positive diagonal audit weight matrix D_b, it also reports the smallest singular value and condition number of $D_b^{1/2} \Phi_{b,\mathcal{B}_b}$. Exact span establishes identification. Coefficient norms and singular values diagnose statistical amplification and numerical stability.

For approximate coefficients $\widetilde{c}_{bj}$, define the atom-weighted residual

$$\delta_{bj} = \sum_q F_x(\mathcal{A}_{bq}) \, |(H_b)_{qj} - (\Phi_{b,\mathcal{B}_b} \widetilde{c}_{bj})_q|. \tag{29}$$

The cell-mass error is at most δ_{bj}. Since $0 \le q^t_a \le 1$, Assumption 4 gives the same bound for target-matched calibration default mass. Target outcome stability transports this bound to deployment. Target-weighted residuals can therefore be added to the statistical error without imposing smoothness on outcomes.

Proposition 5 (Duplicate-menu falsification moment)**.** *Fix (x, a, t, B) with $P_X(x) > 0$ and $1 - \pi_0(x) > 0$. Let $M^{(0)}$ and $M^{(1)}$ have common final approval probability $p_a(x) > 0$ and satisfy*

$$k_a(M^{(0)}) = k_a(M^{(1)}) = t, \qquad \mathbf{1}\{S(M^{(0)}; \theta) = a\} = \mathbf{1}\{S(M^{(1)}; \theta) = a\} = \mathbf{1}\{\theta \in B\}$$

F_x almost surely. Among initially rejected applicants, assign the menus with known first-order probabilities $\rho_0(x), \rho_1(x) > 0$ and define

$$\widehat{\mu}_{r,x} = \frac{1}{N_x} \sum_{i:X_i^b = x} \frac{\mathbf{1}\{T_i^b = 0, M_i = M^{(r)}, S_i = a\} T_i D_i}{[1 - \pi_0(x)]\rho_r(x) p_a(x)}, \qquad r \in \{0, 1\}. \tag{30}$$

Under Assumptions 3 and 4,

$$\mathbb{E}[\widehat{\mu}_{0,x} \mid N_x > 0] = \mathbb{E}[\widehat{\mu}_{1,x} \mid N_x > 0] = \int_B q_a^t(\theta, x) dF_x(\theta).$$

Under independent menu assignment and positive contrast variance, the studentized contrast is asymptotically standard normal. Alternatively, under the iid superpopulation model and complete randomization, fixed-arm Neyman studentization has the same limit when both arm fractions converge to positive constants and the limiting contrast variance is positive. The contrast without $T_i D_i / p_a(x)$ tests equality of selection probabilities, which is necessary for focal-cell equality.

C.3 Proof of Proposition 5

Define the corresponding selection moment by

$$\widehat{\nu}_{r,x} = \frac{1}{N_x} \sum_{i:X_i^b = x} \frac{\mathbf{1}\{T_i^b = 0, M_i = M^{(r)}, S_i = a\}}{[1 - \pi_0(x)]\rho_r(x)}.$$

For $r \in \{0, 1\}$, condition on $(X^b, \theta) = (x, \theta)$. Initial rejection and menu assignment cancel their inverse probabilities. The duplicate-cell identity then gives

$$\mathbb{E}[\widehat{\nu}_{r,x} \mid N_x > 0] = F_x(B).$$

Randomized final approval, the common focal transfer, and Assumption 4 give

$$\mathbb{E}[T_i D_i \mid x, \theta, M_i = M^{(r)}, S_i = a] = p_a(x) q_a^t(\theta, x).$$

Substitution into equation 30 and integration over B yield

$$\mathbb{E}[\widehat{\mu}_{r,x} \mid N_x > 0] = \int_B q_a^t(\theta, x) dF_x(\theta).$$

Because $P_X(x) > 0$, N_x diverges in probability. Under independent menu assignment, the inverse probabilities are bounded and the summands are iid and bounded. The central limit theorem and consistency of the sample variance prove the first claim. Under the iid superpopulation model and complete randomization, the difference-in-means central limit theorem and consistent within-arm variances prove the second.

A nonzero population selection contrast implies unequal selection probabilities. It therefore rules out equal focal cells under the common type law. The sample contrast tests this implication. A zero population contrast is only necessary because distinct cells can have equal mass.

When the sample selection contrast is compatible with zero, rejection of a zero default-mass contrast rejects a common outcome law for the two calibration contexts. It falsifies the maintained model used for transport but does not identify which contextual channel failed. Non-rejection neither proves pointwise exclusion inside B nor verifies stability in an unobserved deployment environment. □

C.4 Proof of Theorem 2

The expressions for A_π and B_π use known policy and design probabilities, empirical base-stratum frequencies, and the unbiased finite moments established above. Under exact indicator spans, the joint law of the calibration data identifies every component. It therefore identifies their finite sums A_π and B_π. Because $A_\pi > 0$, it also identifies $R(\pi) = B_\pi / A_\pi$. The bank value is a known finite linear combination of the same mass and default-mass components.

The inverse probabilities are uniformly bounded by assumption, and every outcome is bounded. The weak law of large numbers therefore gives joint convergence of the finite vector of estimated moments.

Finite sums preserve convergence, so $(\widehat{A}_\pi, \widehat{B}_\pi, \widehat{J}_\pi) \to_p (A_\pi, B_\pi, J_\pi)$. Since the denominator is bounded away from zero, the continuous mapping theorem gives $\widehat{R}_{\text{OA-LID}}(\pi) \to_p R(\pi)$. A multivariate central limit theorem and the delta method additionally give asymptotic normality whenever the limiting covariance is nonsingular. The ratio need not be unbiased at finite N. □

For the uniform statement, write the finite vector of all calibrated mass and default-mass errors as Δ_N. Uniformly bounded ℓ_1 aggregation coefficients imply that every policy's approval, default, and value errors are at most $C\|\Delta_N\|_\infty$. Since the vector is finite, $\|\Delta_N\|_\infty = o_p(1)$. The ratio map is uniformly Lipschitz on denominators at least a_0, yielding the claim.

C.5 Proof of Theorem 3

Use the all-log inverse-propensity form above. Consider one post-rejection action by stratum block and a fixed policy π. For approved-default mass, write its direct-basis contribution as $N^{-1}\sum_i Z_{i,\pi}$. If $w_{\pi\ell} \in [0,1]$ is the aggregation weight and $p_\ell \geq p_0$ is arm-ℓ approval, then

$$\mathbb{E}Z_{i,\pi}^2 \leq \frac{P_X(x)L}{1-\pi_0(x)} \sum_{\ell=1}^{L} \frac{w_{\pi\ell}^2 F_x(B_\ell)}{p_\ell} \leq \frac{L}{r_0 p_0}.$$

The first bound retains each actual p_ℓ. The last inequality uses $p_\ell \geq p_0$ and disjointness of the cells. Also $|Z_{i,\pi}| \leq L/(r_0 p_0)$. Approval-mass summands omit p_ℓ and satisfy the same coarse bounds. Initial-branch, stratum-frequency, and fixed-block terms are bounded by constants depending only on (x_0, r_0) and the block count. Bernstein's inequality therefore bounds either target mass error by

$$C\left\{\sqrt{\frac{L\log(4K/\delta)}{Np_0}} + \frac{L\log(4K/\delta)}{Np_0}\right\}$$

simultaneously over K policies and two masses, by a union bound. This proves equation 18. If the bound is $\varepsilon_N < a_0/2$, then $\widehat{A}_\pi \geq a_0/2$ and, using $B_\pi \leq A_\pi$,

$$\left|\frac{\widehat{B}_\pi}{\widehat{A}_\pi} - \frac{B_\pi}{A_\pi}\right| \leq \frac{\varepsilon_N}{a_0/2} + \frac{B_\pi \varepsilon_N}{A_\pi(a_0/2)} \leq \frac{4\varepsilon_N}{a_0}.$$

The event is simultaneous over Π, so it also covers any data-dependent $\widehat{\pi}$. For the policy-wise benchmark, equal allocation across K target menus repeats the same calculation with per-menu probability $1/K$, replacing L by K. □

C.6 Boundary approximation

For continuous boundaries, let $\widetilde{C}_{\pi,x,e,a}$ be the calibrated approximation and let ϵ_π be the target-weighted mass of its symmetric difference from $C_{\pi,x,e,a}$. Bounded approval and default indicators give both mass errors at most ϵ_π. If both approval denominators are at least a_0, then

$$|R(\pi) - \widetilde{R}(\pi)| \leq 2\epsilon_\pi / a_0. \tag{31}$$

A finite grid therefore carries a separate boundary-approximation error. It does not exactly recover an arbitrary continuous risk surface.

C.7 Signed spans and frozen-log design audit

Fix one block with J_b target cells. For target j, let c_{jr} be its coefficient on focal region G_r. Write $r_x = 1 - \pi_0(x)$ and $g_r = F_x(G_r)$. Assume $n_x \geq 1$, $\delta \in (0,1)$, $r_x > 0$, $\rho_r > 0$, and $p_r > 0$, with categorical arm assignments independent across applicants conditional on x. For mass and default mass, define

$$v_j^g = \sum_r \frac{c_{jr}^2 g_r}{r_x \rho_r}, \qquad u_j^g = 1 + \max_r \frac{|c_{jr}|}{r_x \rho_r},$$
$$v_j^m = \sum_r \frac{c_{jr}^2 g_r}{r_x \rho_r p_r}, \qquad u_j^m = 1 + \max_r \frac{|c_{jr}|}{r_x \rho_r p_r}. \tag{32}$$

Conditional on $N_x = n_x$, Bernstein's inequality gives with probability at least $1 - \delta$

$$|\widehat{z}_j - z_j| \le \sqrt{\frac{2v_j^z \log(4J_b/\delta)}{n_x}} + \frac{u_j^z \log(4J_b/\delta)}{3n_x}, \qquad j \le J_b, \quad z \in \{g, m\}, \tag{33}$$

simultaneously for the J_b target cells in this block. Indeed, only one arm can be assigned to an applicant. For cell j, the score second moments are bounded by v_j^g and v_j^m, while the centered score magnitudes are bounded by u_j^g and u_j^m. Bernstein's inequality and a union bound prove equation 33. The direct uniform theorem is the well-conditioned special case with nonnegative coefficients, mutually exclusive cells, and equal arm assignment.

Before outcomes are observed, the frozen-log certificate records the candidate policy class, menu library, assignment probabilities, and block index (x, a, t). Every arm contributing to an outcome moment uses focal action a and focal transfer t. Indicator identities hold pointwise on the certified support. The same coefficients aggregate mass and default mass, and initial approval, menu assignment, and final approval probabilities are known and positive.

The direct uniform guarantee additionally records cell exclusivity, aggregation weights in $[0, 1]$, and the constants x_0, r_0, p_0, a_0. Cell mass omits the final-approval inverse probability, while default mass includes it. Empty cells remain linear masses. The initially approved branch is estimated separately. Data-dependent policy selection is permitted only within the class fixed before outcomes.

Corollary 2 (Bounded context dependence)**.** *Under the reusable-log, randomization, direct-basis, positivity, and exclusivity conditions used in Theorem 3, replace exact calibration-to-target outcome transport by the bounds below. Assume $\varepsilon_N < a_0/2$ and $\inf_{\pi\in\Pi} A_\pi \ge a_0 > 0$, and write the target approval and default masses as*

$$A_\pi = \sum_h w_{\pi h}, \qquad B_\pi^{\mathrm{tar}} = \sum_h w_{\pi h} q_{\pi h}^{\mathrm{tar}},$$

where $w_{\pi h} \ge 0$ are the approved-mass weights of the base and calibrated cell components. Let $B_\pi^{\mathrm{cal}} = \sum_h w_{\pi h} q_{\pi h}^{\mathrm{cal}}$, define $R_\pi^{\mathrm{cal}} = B_\pi^{\mathrm{cal}}/A_\pi$ and $R_\pi^{\mathrm{tar}} = B_\pi^{\mathrm{tar}}/A_\pi$, and suppose $|q_{\pi h}^{\mathrm{cal}} - q_{\pi h}^{\mathrm{tar}}| \le \eta_{\pi h}$. Define $\beta_\pi = \sum_h w_{\pi h}\eta_{\pi h}$. Then

$$|B_\pi^{\mathrm{cal}} - B_\pi^{\mathrm{tar}}| \le \beta_\pi, \qquad |R_\pi^{\mathrm{cal}} - R_\pi^{\mathrm{tar}}| \le \frac{\beta_\pi}{A_\pi}. \tag{34}$$

The concentration argument gives $\max_{\pi\in\Pi} |\widehat{R}_\pi - R_\pi^{\mathrm{cal}}| \le 4\varepsilon_N/a_0$ on an event of probability at least $1 - \delta$. Consequently,

$$\max_{\pi\in\Pi} |\widehat{R}_\pi - R_\pi^{\mathrm{tar}}| \le \frac{4\varepsilon_N}{a_0} + \max_{\pi\in\Pi} \frac{\beta_\pi}{A_\pi}. \tag{35}$$

If every $\eta_{\pi h} \le \eta$, the second term is at most η. Setting $\eta = 0$ recovers the original uniform conclusion exactly.

Proof. The default-mass difference is a weighted sum of component differences, so the triangle inequality gives the first bound. Approval mass does not depend on the outcome restriction. Division by A_π gives the risk bound. The bounded-score concentration proof of Theorem 3 applies unchanged around the calibration moments, and a second triangle inequality gives equation 35. Since $\sum_h w_{\pi h} = A_\pi$, a common radius contributes at most η. □

C.8 Finite-cell approximation

Let (A, B) and $(\widetilde{A}, \widetilde{B})$ denote the true and approximating target approval masses and approved-default masses. A disagreement of the true and approximating action kernels on probability mass ϵ_π can change an approval indicator or an approved-default indicator by at most one. Thus

$$|A - \widetilde{A}| \le \epsilon_\pi, \qquad |B - \widetilde{B}| \le \epsilon_\pi.$$

If $A, \widetilde{A} \ge a_0$ and $0 \le \widetilde{B} \le \widetilde{A}$, then

$$\begin{aligned}\left|\frac{B}{A} - \frac{\widetilde{B}}{\widetilde{A}}\right| &\le \frac{|B - \widetilde{B}|}{A} + \frac{\widetilde{B}|A - \widetilde{A}|}{A\widetilde{A}} \\ &\le \frac{\epsilon_\pi}{a_0} + \frac{\epsilon_\pi}{a_0} = \frac{2\epsilon_\pi}{a_0}.\end{aligned}$$

This proves equation 31. A margin or boundary-mass assumption converts cell diameter into a rate for ϵ_π. The present theorem leaves ϵ_π explicit.

C.9 Proof of Theorem 4

Fix an admissible design $\mathfrak{d}$ from Definition 2. All design quantities are known: F, B, g, p, t, allocation probabilities ρ_ℓ, regions R_ℓ, and coefficients c_ℓ. Across subjects the experiment is iid, so $O_{1:N} \sim P^{\otimes N}$. The supremum ranges over measurable focal risks $q_s^t : \mathbb{R}^2 \to [0,1]$ whose cell average lies in $[1/4, 3/4]$. Other-action outcomes are nuisance laws. Constants may depend on $(L, \rho_0, p_0, \beta_0, C_c, g_0)$ but not (N, p, g).

Direct upper bound. Let I_B be the observed indicator for both focal choice and cell membership, with $\mathbb{P}(I_B = 1) = g$. Conditional on $I_B = 1$, approval $T \sim \mathrm{Bernoulli}(p)$ is independent. Since g is part of the known design, the unbiased estimator

$$\widehat{q}_D = \frac{1}{Npg} \sum_{i=1}^{N} I_{B,i} T_i D_i$$

has

$$\mathrm{Var}(\widehat{q}_D) = \frac{pgq_B(1 - pgq_B)}{Np^2g^2} \le \frac{1}{Npg}.$$

Direct lower bound. Take $q_\pm = 1/2 \pm \Delta$. Membership and approval laws are identical. Only a default observed with probability gp changes. The one-subject KL divergence satisfies

$$\mathrm{KL}(P_+ \| P_-) = gp\,\mathrm{KL}(\mathrm{Bern}(q_+) \| \mathrm{Bern}(q_-)) \le Cgp\Delta^2$$

for sufficiently small fixed Δ. Choosing $\Delta = c/\sqrt{Npg}$ makes the N-sample total variation bounded away from one. Le Cam's two-point lemma yields risk at least $c'/(Npg)$. If Npg is bounded, use a fixed Δ to obtain a constant lower bound.

Broad upper bound. For arm ℓ, let

$$\widehat{m}_\ell = \frac{1}{N} \sum_{i=1}^{N} \frac{\mathbf{1}\{M_i = \ell, S_i = s\} T_i D_i}{\rho_\ell p}.$$

It is unbiased for $m_\ell = \int_{R_\ell} q_s^t(\theta) F(d\theta)$ and has variance at most $1/(N\rho_0 p)$. The pointwise contrast gives $gq_B = \sum_\ell c_\ell m_\ell$, so $g^{-1} \sum_\ell c_\ell \widehat{m}_\ell$ has variance at most $C/(Npg^2)$.

Broad lower bound. Use the same pair with $q_s^t = 1/2 \pm \Delta$ on B and $1/2$ off B, while keeping the nuisance laws identical. In any arm intersecting B, changing the sign perturbs each default or nondefault category among selected and approved observations by at most $2\rho_\ell pg\Delta$. Its background probability is at least $\rho_\ell p\beta_0$ because $F(R_\ell \setminus B) \ge 2\beta_0$. The categorical KL bound gives one-subject KL at most $Cpg^2\Delta^2$. With $\Delta = c/\sqrt{Npg^2} \le 1/4$, Le Cam yields $\Omega((Npg^2)^{-1})$, completing the proof. □

Alternative information regimes. If I_B is observed in the broad arm, its conditional default labels reduce the problem to the direct experiment. If broad regions shrink with g rather than retaining background mass, their categorical baseline probabilities also shrink and the KL denominator changes. Repeated or coupled queries and smooth outcome restrictions similarly define separate statistical experiments with different information rates.

D Budget Derivations

D.1 Gross subsidy identity and caps

After matching focal action a to target transfer t, action b has transfer $\widetilde{k}_b = k_b + t - k_a$. A negative transfer is paid by the experiment, hence its gross subsidy is $(-\widetilde{k}_b)_+ = (k_a - k_b - t)_+$. Averaging

over choice shares gives

$$Q_{a,t} = \sum_b \varsigma_b (k_a - k_b - t)_+.$$

If $|k_b - k_a| \le H$ and $|t| \le T$, each subsidy is at most $H + T$. Together with deterministic nontransfer cap $\bar{c}_m$, this gives

$$N_m^{\text{safe}} = \left\lfloor \frac{M}{\bar{c}_m + \lambda(H_m + T_m)} \right\rfloor. \tag{36}$$

At $t = 0$, the focal term vanishes and the sharper bound $Q_{a,0} \le H(1 - \varsigma_a)$ follows.

D.2 Proof of Corollary 1

For $0 < C_E < \infty$, Theorem 4 gives, up to its design-specific constants,

$$\mathfrak{R}_D(N_D; \mathfrak{d}) \asymp \frac{1}{N_D p_D g}, \qquad \mathfrak{R}_W(N_W; \mathfrak{d}) \asymp \frac{1}{N_W p_W g^2}$$

in the stated non-saturated regime. Enrolling the fixed number $N_E(M) = \lfloor M/C_E \rfloor$ costs at most M in expectation. When $M/C_E \ge 2$, one has $M/(2C_E) \le N_E(M) \le M/C_E$. Substitution proves the first two relations in equation 23, and division proves the third. Over a sequence with uniform theorem constants and bounded approval ratios, $C_D/C_W = o(1/g)$ sends the risk ratio to zero. Enrollment is fixed before outcomes and realized subsidies. Adaptive stopping requires a separate analysis. The explicit-constant bounds below give exact crossing points. □

D.3 Explicit constants and decision regions

Consider a declared design envelope on which the rare-cell theorem constants admit positive uniform lower bounds and finite uniform upper bounds:

$$\begin{aligned} \frac{\underline{c}_D}{N p_D g} \le \mathfrak{R}_D(N; \mathfrak{d}) \le \frac{\overline{c}_D}{N p_D g}, \\ \frac{\underline{c}_W}{N p_W g^2} \le \mathfrak{R}_W(N; \mathfrak{d}) \le \frac{\overline{c}_W}{N p_W g^2}. \end{aligned} \tag{37}$$

The constants may depend uniformly on the fixed envelope $(L, \rho_0, p_0, \beta_0, C_c, g_0)$, but not on N, p_E, or g. Since $M/(2C_E) \le N_E(M) \le M/C_E$ when $M \ge 2C_E$,

$$\begin{aligned} \frac{\underline{c}_D C_D}{M p_D g} \le \mathfrak{R}_D^{\text{bud}} \le \frac{2\overline{c}_D C_D}{M p_D g}, \\ \frac{\underline{c}_W C_W}{M p_W g^2} \le \mathfrak{R}_W^{\text{bud}} \le \frac{2\overline{c}_W C_W}{M p_W g^2}. \end{aligned} \tag{38}$$

Consequently,

$$\frac{\underline{c}_D}{2\overline{c}_W} \frac{p_W}{p_D} g \frac{C_D}{C_W} \le \frac{\mathfrak{R}_D^{\text{bud}}}{\mathfrak{R}_W^{\text{bud}}} \le \frac{2\overline{c}_D}{\underline{c}_W} \frac{p_W}{p_D} g \frac{C_D}{C_W}. \tag{39}$$

Direct labelling is strictly better whenever

$$\frac{C_D}{C_W} < \frac{\underline{c}_W}{2\overline{c}_D} \frac{p_D}{p_W} \frac{1}{g}. \tag{40}$$

Broad inversion is strictly better whenever

$$\frac{C_D}{C_W} > \frac{2\overline{c}_W}{\underline{c}_D} \frac{p_D}{p_W} \frac{1}{g}.$$

The interval between these sufficient conditions depends on the minimax constants. It does not alter the uniform conclusion that $C_D/C_W = o(1/g)$ favors direct labelling. The executable condition

$$M p_E g^2 \ge 2C_E$$

implies $N_E(M) p_E g^2 \ge 1$ and therefore places both designs in the stated regime.

The expected-cost certificate and the hard-budget certificate answer different questions. A pilot estimates C_E and supports fixed expected enrollment. A deterministic cap uses equation 36. Replacing C_E by the capped per-person cost in equation 38 yields the conservative hard-budget version.

D.4 The normalized hand calculation

For $g = 0.01$, $Q_D = 0.198$, $Q_W = 0$, common base cost c, equal approval probabilities, and equal leading constants, equal proxy risk requires $N_W = N_D/g$. At $N_D = 10{,}000$, equating total costs gives

$$\begin{aligned} 10{,}000(c + 0.198) &= 1{,}000{,}000c, \\ 0.01c + 0.00198 &= c, \\ c &= 0.00198/0.99 = 0.002. \end{aligned}$$

At $c = 1$, the two costs are $10{,}000(1.198) = 11{,}980$ and 10^6. At $c = 0.001$, they are $10{,}000(0.199) = 1{,}990$ and $1{,}000$.

D.5 Pilot certificate

The pilot subsidies s_{S_i} are independent and lie in $[0, H]$. Hoeffding's inequality gives

$$\mathbb{P}(|\widehat{Q} - Q| > t) \le 2\exp\left(-\frac{2n_0 t^2}{H^2}\right).$$

Setting the right-hand side equal to δ gives, with probability at least $1 - \delta$,

$$|\widehat{Q} - Q| \le H\sqrt{\frac{\log(2/\delta)}{2n_0}}. \tag{41}$$

For $(n_0, H, \delta) = (1000, 0.2, 0.05)$,

$$0.2\sqrt{\frac{\log 40}{2000}} = 0.008589\ldots.$$

The certificate assumes that the pilot and main-study choice distributions coincide.

D.6 Multi-arm allocation

For $M > 0$, $v_j > 0$, $C_j > 0$, and positive N_j, minimize $\sum_j v_j/N_j$ subject to $\sum_j C_j N_j \le M$. Zero-variance arms may be removed. The Lagrangian is

$$\mathcal{L} = \sum_j \frac{v_j}{N_j} + \eta\left(\sum_j C_j N_j - M\right).$$

The first-order condition gives $N_j = \sqrt{v_j/(\eta C_j)}$. Enforcing the active budget constraint yields

$$\sqrt{\eta} = \frac{\sum_\ell \sqrt{v_\ell C_\ell}}{M},$$

and therefore

$$N_j^* = \frac{M\sqrt{v_j/C_j}}{\sum_\ell \sqrt{v_\ell C_\ell}}, \qquad \sum_j \frac{v_j}{N_j^*} = \frac{(\sum_j \sqrt{v_j C_j})^2}{M}. \tag{42}$$

Convexity in each $N_j > 0$ proves global optimality of the continuous allocation. The Cauchy-Schwarz inequality gives the same result. The result applies to continuous allocation and does not cover integer rounding, unknown v_j, or random stopping under realized payments.

E Experimental Details

E.1 Reproduction and artifacts

The controlled suite uses seed 20260901 and an independent seed, 20261808, for the duplicate-menu diagnostic. Run the following commands from the project root to regenerate the results:

```
/opt/anaconda3/bin/python paper/experiments/run_synthetic_suite.py
/opt/anaconda3/bin/python paper/figures/gen_fig_alias_cell.py
/opt/anaconda3/bin/python paper/figures/gen_fig_mse_scaling.py
/opt/anaconda3/bin/python paper/figures/gen_fig_budget_tradeoff.py
/opt/anaconda3/bin/python \
  paper/figures/gen_fig_priced_target_reuse.py
```

The recorded runtime of the numerical suite is under two seconds. Raw outputs are stored as tidy CSV files with a machine-readable JSON summary. The summary records the exact two-world log law, closed-menu transfer matching, duplicate-menu geometry, and sampled menu scoring. It also records hidden-atom policy reconstruction, target-span conditioning, menu geometry, Monte Carlo slopes, and the normalized budget crossing.

E.2 Priced-target alias and focal-cell experiment

For the alias witness, actions S, C have $(p, d, k) = (0.8, 0.5, 0)$ and $(0.7, 0.3, 0)$. Types $(1, 1)$ and $(10, 10)$ therefore have utilities $(0.3, 0.4)$ and $(3, 4)$. A fee $k_C = 0.15$ changes their choices to S and C, respectively. Target behavior therefore depends on scale. The script also enumerates 81 rational zero-fee actions and verifies complete ranking equivalence.

The stronger check for Proposition 1 treats a and b as genuine approval offers and retains a separate explicit exit in every closed menu. It enumerates the three nonempty zero-transfer offer submenus $\{b\}, \{a\}, \{a, b\}$. In both worlds, their exact laws of (S, T, TD) coincide. When b is selected, the nonzero masses are $(T, TD) = (0, 0), (1, 0), (1, 1)$ with probabilities $3/5, 1/5, 1/5$. When a is selected, the probabilities are $1/5, 2/5, 2/5$. Under $(t_b, t_a) = (0, 1)$, low scale selects b and high scale selects a. Both worlds have approval mass $3/5$, while their default masses are $3/5$ and 0, yielding risks 1 and 0.

The focal-cell experiment uses the actions and transfers in Table 2. The displayed exit has base transfer 0.3. Exact arithmetic on a 161×161 rational grid checks $S \iff (V, \Gamma) \in [1, 2]^2$ under the focal tie rule. Appendix A.1 proves the identity analytically. To check transfer matching, we add 0.4 to all six displayed transfers. All 25,921 grid points have zero choice mismatches and zero pairwise-transfer-difference error. The focal transfer becomes 0.4, as in Lemma 1.

The separate stochastic $t = 0$ recovery draws 200,000 independent uniform types. The true cell mass is $1/16$. Potential default is Bernoulli with

$$q_s^0(V, \Gamma) = 0.05 + 0.08V + 0.10\Gamma,$$

which lies in $[0.05, 0.77]$ on the support, and approval is independent Bernoulli with probability 0.8. The analytic cell mean is $0.05 + 0.08(1.5) + 0.10(1.5) = 0.32$. The run selected 12,262 subjects, approved 9,820, and observed 3,125 defaults, producing estimate 0.318565.

E.3 Duplicate-menu falsification protocol

Both arms retain the same shifted six-action menu and focal transfer $t = 0.4$. Each adds one context offer that is strictly dominated by the displayed exit throughout $[0, 4]^2$. Arm A adds $(p, d, k) = (0.1, 1, 5.4)$ and arm B adds $(0.9, 0.1, 5.4)$. Exact rational arithmetic on a 161×161 grid for each arm gives zero choice mismatches, zero focal-cell mismatches, and zero context-offer selections. Both menus therefore label $[1, 2]^2$ with focal action S at the same transfer.

The experiment assigns $N = 200{,}000$ subjects in equal numbers by complete randomization. Scoring each arm from its displayed menu gives zero sampled choice and focal-cell mismatches. Final approval is independent with probability 0.8. Fixed-arm Neyman standard errors yield normal approximations. The HT selection-mass contrast is -0.0006, with $z = -0.5532$ and $p = 0.5801$. Under the exclusion-null generator, the arm B minus arm A HT default-mass contrast is -0.0003875. Its standard error is 0.0007061, with $z = -0.5488$ and $p = 0.5831$.

The second generator retains the same types, assignments, approvals, and default uniforms. It adds 0.16 to default probability for focal choosers in arm B. The analytic default-mass contrast is 0.01. The estimate is 0.0093625, with standard error 0.0007845, $z = 11.9336$, and $p = 7.91 \times 10^{-33}$. The contrast detects the injected menu-context dependence. It does not establish stability in a deployment environment.

Table 4: Fixed offers in the complete priced-menu experiment. A positive transfer is a fee; a negative transfer is a subsidy.

Action	Approval p_a	Burden d_a	Transfer t_a
s	.8	.5	.1
a_1	.6	.5	$-.1$
a_2	.8	.3	.4
$\varnothing$	0	0	0

Table 5: Population quantities obtained by integrating the utility-defined regions and the stated default law. Subsidy is per enrolled applicant, before any fee revenue.

Policy	A_π	B_π	$R(\pi)$	$J(\pi)$	Gross subsidy
π_1	0.676952	0.195902	0.289388	0.089247	0.012952
π_2	0.698286	0.208325	0.298337	0.073311	0.000000
π_3	0.684571	0.204829	0.299207	0.070085	0.012952
π_4	0.690667	0.199398	0.288703	0.092473	0.000000
π_5	0.687619	0.202113	0.293932	0.081279	0.006476
π_6	0.763429	0.224247	0.293736	0.090689	0.009714
π_7	0.605714	0.185411	0.306103	0.049481	0.016190
π_8	0.687619	0.202113	0.293932	0.081279	0.006476

E.4 Complete priced-menu experiment

Each applicant draws a type and chooses the displayed alternative with the highest utility. We then draw approval for the selected offer and observe default only after approval. There is one base stratum, with (V, Γ) uniform on $[.5, 2] \times [.25, 2]$. The support has area $21/8$. Table 4 specifies the action library. The four menus are

$$E_1 = \{s, a_1, \varnothing\}, \quad E_2 = \{s, a_2, \varnothing\}, \quad E_3 = \{s, a_1, a_2, \varnothing\}, \quad E_4 = \{s, \varnothing\}.$$

All four target menus allow free exit. Utilities are $Vp_a - \Gamma d_a - t_a$ with a fixed action-order tie rule. Boundary ties have zero population mass. A lower-approval offer remains competitive because of its subsidy.

Population outcomes and candidates. For each genuine offer, the simulated default probability is

$$q_a^t(V, \Gamma) = .05 + .08V + .10\Gamma + \delta_a + .3t, \qquad (\delta_s, \delta_{a_1}, \delta_{a_2}) = (0, .05, -.03).$$

These probabilities lie between .135 and .5 over the support at the target transfers. The initially approved branch has default probability $.07 + .08V + .10\Gamma$, with population mean .2825. Approval and default uniforms are independent conditional on type and selected action. No outcome labels from a credit data set enter this experiment.

Each policy is specified by its initial approval probability and its conditional menu distribution:

$$\pi_1 = (.2, E_1), \qquad \pi_2 = (.2, E_2), \quad \pi_3 = (.2, E_3), \quad \pi_4 = (.2, E_4),$$
$$\pi_5 = (.2, \tfrac{1}{2}E_1 + \tfrac{1}{2}E_2), \quad \pi_6 = (.4, E_3), \quad \pi_7 = (0, E_3), \quad \pi_8 = (.2, \tfrac{1}{4}\textstyle\sum_{e=1}^{4} E_e).$$

The mixtures randomize the displayed menu after initial rejection. The objective $J(\pi) = A_\pi - 3B_\pi$ is fixed before simulation. Polygon integration gives all population quantities in Table 5. The optimal policy is π_4, with $J = .09247333$. The gap to π_6 is .00178476. Policies π_5 and π_8 have identical population quantities. Including the base-outcome moment, the eight policy rows have coefficient rank four.

Coverage from actual choices. Write $C_{e,a}$ for the region selecting a in E_e. Utility comparison gives $s \succeq a_1$ if $V \geq 1$, $s \succeq a_2$ if $\Gamma \leq 1.5$, and $s \succeq \varnothing$ if $\Gamma \leq 1.6V - .2$. On the support,

$$C_{1,a_1} = C_{3,a_1}, \qquad C_{2,a_2} = C_{3,a_2}, \qquad \mathbf{1}_{C_{4,s}} = \mathbf{1}_{C_{1,s}} + \mathbf{1}_{C_{2,s}} - \mathbf{1}_{C_{3,s}}.$$

The region excluded by both comparisons, $V < 1$ and $\Gamma > 1.5$, cannot prefer s to exit. The three nonzero s indicator patterns across E_1, E_2, E_3 are $(1,1,1)$, $(1,0,0)$, and $(0,1,0)$, so the coverage matrix has rank three. The E4 coefficient vector is $(1,1,-1)$ and its absolute coefficient sum is three. The areas are

$$\big(|C_{1,s}|, |C_{2,s}|, |C_{3,s}|, |C_{4,s}|\big) = \left(\frac{131}{80}, \frac{519}{320}, \frac{399}{320}, \frac{161}{80}\right), \quad |C_{1,a_1}| = \frac{17}{40}, \quad |C_{2,a_2}| = \frac{27}{64}.$$

An independent rational-arithmetic verifier enumerates all six positive-area joint-choice atoms and integrates each affine outcome law. It agrees with the simulation's polygon integrator to numerical precision.

Frozen-log estimators. The reusable design sets initial approval to .2 and assigns equal enrollment to E_1, E_2, E_3. Fixed index groups receive the three menus independently of their type draws. This has the same distribution as complete randomization of iid applicants. A menu is displayed only after initial rejection. Let r_e denote the unconditional probability of initial rejection and menu E_e. Then $r_e = .8/3$ for the three logging menus. For $a = s$, the estimates are

$$\widehat{g}_{e,s} = \frac{1}{Nr_e} \sum_i \mathbf{1}\{T_i^b = 0, M_i = E_e, S_i = s\},$$

$$\widehat{m}_{e,s} = \frac{1}{Nr_e p_s} \sum_i \mathbf{1}\{T_i^b = 0, M_i = E_e, S_i = s\} T_i D_i.$$

For a_1, we pool the E1 and E3 numerators. The denominator is $N(r_1 + r_3)$ for g and includes the additional factor p_{a_1} for m. We similarly pool E2 and E3 for a_2. The initial outcome moment is $\sum_i T_i^b D_i^{\text{base}}/(.2N)$. In total, we estimate five distinct action-specific moments from three physical menus, plus the initial outcome moment.

The cells overlap across menus, and the E4 reconstruction is signed. Appendix C.7 gives the relevant coefficient-dependent guarantee. The direct nonnegative-basis constant in Theorem 3 is a separate case.

Independent rollout allocates $N/8$ applicants to each full target policy and directly estimates its approval and approved-default frequencies. Pooled rollout uses the same records. Its initial approval probability averages to .2, and its unconditional rejected-menu probabilities are $(.175, .175, .325, .125)$. It pools the repeated a_1 and a_2 cells using these known exposures. All four s regions, including E4, are estimated directly. The comparator thus uses both sharing and target-menu observations.

Equal enrollment across the first three target policies produces the same observable experiment as the reusable design. The corresponding estimates agree within 10^{-14} in every simulated batch.

All estimators substitute their moments into the policy decomposition and select the largest estimated J from the frozen log. To define risk numerically, we divide by $\max\{\widehat{A}, .05\}$. This floor is never active in the recorded runs, and $\min_\pi A_\pi = .605714$. Each applicant contributes at most one observed downstream outcome. The recorded joint estimates preserve correlations from shared observations.

Enrollment and actual menu cost. Only a_1 receives a subsidy in the matched experiment. Its selection probability per rejected applicant is $17/105$ in E1 and E3. Thus the reusable design's expected gross subsidy per enrolled applicant is

$$Q_{\text{reuse}} = .8\frac{2}{3}(.1)\frac{17}{105} = \frac{68}{7875} = .00863492.$$

The eight-policy design has $Q_{\text{rollout}} = 17/2100 = .00809524$, shared by its independent and pooled estimators. Fees are recorded separately and never offset these payments. With $c = .02$ and unit subsidy weight, an expected ceiling $M = N_{\text{reuse}}(c + Q_{\text{reuse}})$ admits

$$N_{\text{rollout}} = 24 \left\lfloor \frac{M}{24(c + Q_{\text{rollout}})} \right\rfloor.$$

Enrollment is fixed in advance. Realized subsidies need not equal their expectation. We also record the accounting sensitivity at $c = 1$. The reusable design has a higher per-person subsidy in this example. The matched-budget comparison includes that difference.

Table 6: Complete-menu results at the largest enrollment or expected-budget ceiling. Entries after $\pm$ are Monte Carlo standard errors over $1{,}000$ independent repetitions. Maximum risk error and regret are computed over the same eight predeclared policies.

Design	N	Expected cost	Max. risk error	Selection regret
Equal enrollment				
Reusable three menus	48,000	1374.48	0.00670 ± 0.00010	0.001120 ± 0.000036
Pooled eight-policy rollout	48,000	1348.57	0.00741 ± 0.00009	0.001658 ± 0.000072
Independent eight-policy rollout	48,000	1348.57	0.01273 ± 0.00012	0.004662 ± 0.000183
Equal expected budget				
Reusable three menus	48,000	1374.48	0.00670 ± 0.00010	0.001120 ± 0.000036
Pooled eight-policy rollout	48,912	1374.19	0.00740 ± 0.00009	0.001535 ± 0.000070
Independent eight-policy rollout	48,912	1374.19	0.01286 ± 0.00012	0.004470 ± 0.000177

Transfer-matching ablation. We subtract .1 from every E3 transfer, including its exit, and keep the original target policies. Each E3 utility increases by .1, leaving choices unchanged. Each genuine action's default probability falls by .03. Matched and shifted runs use the same types, assignments, approvals, and default uniforms.

The wrong-price estimator pools these outcomes as if their selected transfers had not changed. Its E3 s moment is shifted by $-.03g_{3,s}$. The pooled a_1 and a_2 moments shift by $-.015g_{1,a_1}$ and $-.015g_{2,a_2}$. Substituting all three changes into the policy decomposition gives the limiting risk biases

$$(-.001722, -.002209, -.017279, .013205, -.001970, -.011621, -.024410, -.001970).$$

The limiting estimated objective selects π_6, whose true regret is .00178476. This ablation changes selected transfers, whereas the duplicate-menu diagnostic above changes only unchosen context at a fixed selected transfer. The shifted E3 exit receives a .1 subsidy. This diagnostic is excluded from the matched-budget design comparison.

Execution and uncertainty. We use seed 20260907 and six sample sizes: $1200, 2400, 4800, 12000, 24000, 48000$. Each size has $1{,}000$ repetitions. Every replication draws continuous types and evaluates the actual utility maximizer rather than sampling an abstract atom index. Independent and pooled rollout share their raw records. The same reusable run appears under both comparison axes. The protocol and objective were fixed before the first run.

The main simulation took 28.54 seconds with Python 3.12.7 and NumPy 1.26.4 on a local CPU. The result tables are accompanied by raw policy estimates, selection decisions, realized subsidies, fees, and standard errors. The accompanying records also contain the subject-level execution witness and source hash. Run the following commands from the project root to reproduce the experiment:

```
/opt/anaconda3/bin/python \
  paper/experiments/priced_menu/test_priced_menu.py
/opt/anaconda3/bin/python \
  paper/experiments/priced_menu/verify_geometry.py
/opt/anaconda3/bin/python \
  paper/experiments/priced_menu/run_priced_menu.py
MPLCONFIGDIR=/private/tmp/oa-lid-mpl /opt/anaconda3/bin/python \
  paper/figures/gen_fig_priced_menu.py
```

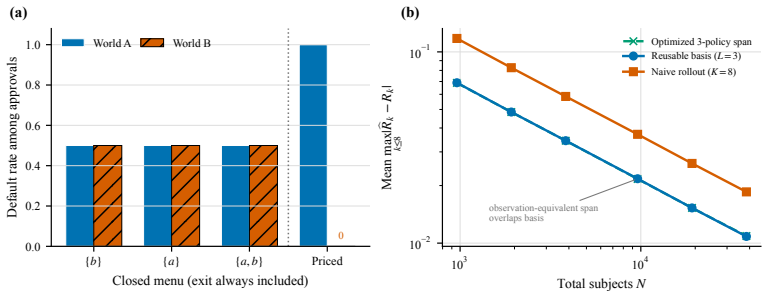


Figure 3: Identification and the direct hidden-atom reuse check. Left: identical zero-transfer laws yield priced risks 1 and 0. Right: a three-arm basis improves independent eight-policy rollout and matches an observation-equivalent span over 20,000 repetitions per size.

E.5 Direct hidden-atom reuse protocol

The hidden atom $Z \in \{1, 2, 3\}$ has mass $g = (0.2, 0.3, 0.5)$, matched target transfer $t = (0, 0.4, 0.8)$, and default risk $q = (0.1, 0.4, 0.8)$. The eight target policies have atom-specific approval matrix

$$W = \begin{bmatrix} .8 & 0 & 0 \\ 0 & .8 & 0 \\ 0 & 0 & .8 \\ .8 & .8 & 0 \\ .8 & 0 & .8 \\ 0 & .8 & .8 \\ .8 & .8 & .8 \\ .2 & .5 & .8 \end{bmatrix}.$$

Thus the analytic target quantities are $A_k = \sum_\ell W_{k\ell} g_\ell$, $B_k = \sum_\ell W_{k\ell} g_\ell q_\ell$, and $R_k = B_k / A_k$; $\min_k A_k = 0.16$.

The reusable design allocates $N/3$ subjects to each focal-matched arm. In arm ℓ, the choice label reveals $I\{Z = \ell\}$. A selected subject is independently approved with probability $p = 0.8$. Default is observed only after approval. With $n_\ell = N/3$, the direct moments are

$$\widehat{g}_\ell = \frac{1}{n_\ell} \sum_i I\{Z_i = \ell\}, \qquad \widehat{m}_\ell = \frac{1}{n_\ell p} \sum_i I\{Z_i = \ell, T_i = 1, D_i = 1\}.$$

Substitution into $W\widehat{g}$ and $W\widehat{m}$ evaluates all eight policies from one frozen log. The naive comparator instead allocates $N/8$ independent subjects to each complete target rollout and estimates its (A_k, B_k, R_k) directly. Every design uses one menu and at most one downstream outcome per subject. A unit per-subject cost makes N a normalized enrollment budget. This normalization excludes unequal monetary arm costs.

We use $N \in \{960, 1920, 3840, 9600, 19200, 38400\}$ and 20,000 repetitions at each point. At the largest N, mean maximum error over the eight risks is 0.01084 for the basis and 0.01853 for naive rollout, a ratio of 1.710; their slopes from regressions of logarithmic error on logarithmic N are -0.502 and -0.501. For the predeclared value $J_k = A_k - 2.4B_k$, P4 is uniquely optimal. Selection from the same frozen log gives mean regret 0.000361 for the basis and 0.001979 for naive rollout at the largest N.

The first three policies satisfy $W_{1:3} = 0.8I_3$, with rank 3, smallest singular value 0.8, condition number 1, and zero reconstruction residual. An independently simulated rollout of this optimized target span has mean maximum error 0.01089 and slope -0.500. This optimized target-span baseline matches the basis within Monte Carlo error.

E.6 MSE protocol and exact checks

The direct estimator grants knowledge of g and uses

$$\widehat{q}_D = \frac{X_D}{Npg}, \qquad X_D \sim \text{Binomial}(N, pgq),$$

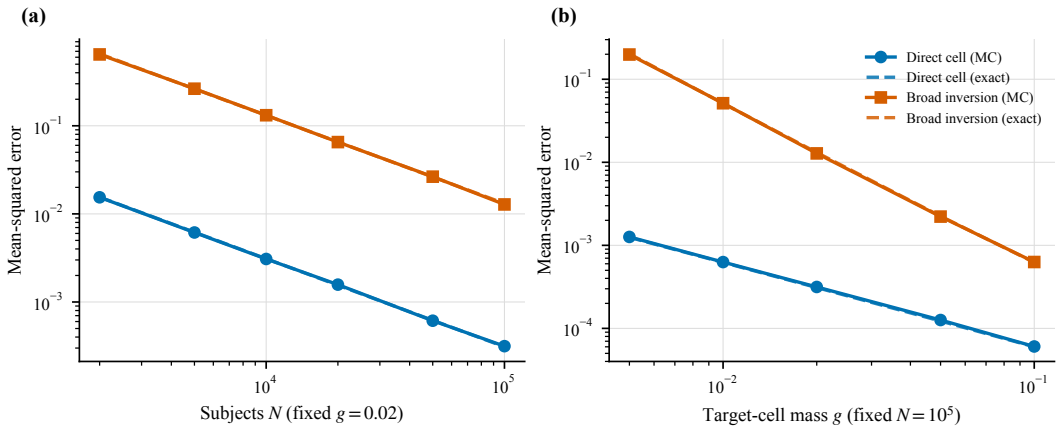


Figure 4: Direct-cell versus broad-inversion MSE in the restricted coarse-log model. Solid markers are Monte Carlo results and dashed curves are exact variances. Fitted exponents in N are -0.996 for direct labelling and -1.001 for broad inversion. The corresponding exponents in g are -1.011 and -1.927.

so its exact MSE is

$$\frac{pgq(1-pgq)}{N(pg)^2}.$$

For broad inversion, each subject is assigned to the plus or minus arm with probability $1/2$. With background mass $h = 0.5$, background risk $r_0 = 0.2$, target risk $q = 0.5$, and $p = 0.8$, the two default-event probabilities are

$$\eta_+ = \tfrac{1}{2}p(hr_0 + gq), \qquad \eta_- = \tfrac{1}{2}phr_0.$$

The joint event counts are drawn from the exact multinomial distribution, and

$$\widehat{q}_W = \frac{X_+ - X_-}{N(1/2)pg}.$$

The exact variance uses

$$\mathrm{Var}(X_+ - X_-) = N\{\eta_+ + \eta_- - (\eta_+ - \eta_-)^2\}.$$

We use $N \in \{2000, 5000, 10^4, 2 \cdot 10^4, 5 \cdot 10^4, 10^5\}$, $g \in \{0.005, 0.01, 0.02, 0.05, 0.1\}$, and 10,000 replicates at every point. Across the resulting 60 combinations of method and configuration, empirical MSE divided by exact MSE lies between 0.975 and 1.029.

E.7 Budget sweep

At $g = .01$, $Q_D = .198$, $Q_W = 0$, and equal approval probabilities and leading constants, the proxy crossing is

$$10^4(c + .198) = 10^6 c \quad \Longleftrightarrow \quad c = .002. \tag{43}$$

Equation 43 is a normalized accounting comparison.

The sweep fixes $g = 0.01$, $N_D = 10^4$, $N_W = 10^6$, $Q_W = 0$, and varies c logarithmically from 10^{-4} to 10 for $Q_D \in \{0, 0.05, 0.1, 0.198, 0.3\}$. It plots both total costs and the exact ratio

$$\frac{N_D(c + Q_D)}{N_W c} = g\frac{c + Q_D}{c}.$$

This experiment fixes statistical constants and approval probabilities to isolate the accounting corollary.

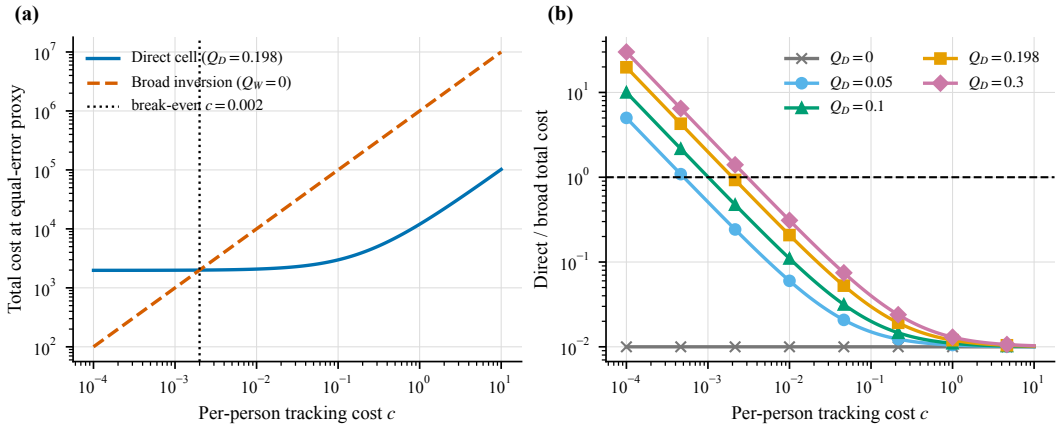


Figure 5: Normalized budget proxy for $g = 0.01$ and $N_D = 10^4$. Direct and broad costs cross at $c = 0.002$ when $Q_D = 0.198$ and $Q_W = 0$. The costs are normalized simulation quantities.

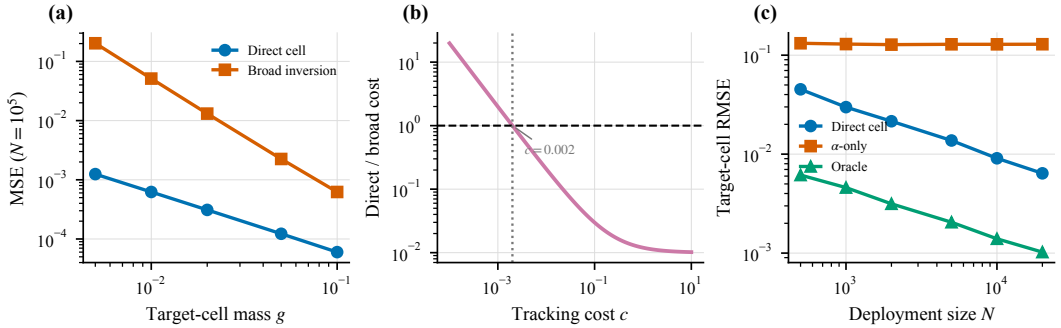


Figure 6: Additional theorem-directed checks, from left to right: exact rare-cell MSE at $N = 10^5$; the normalized ratio of direct to broad cost for $g = 0.01$, $Q_D = 0.198$, and $Q_W = 0$; and a semi-synthetic South German Credit stress test in which the direct-cell estimator converges while an α-only comparator retains structural bias. All hidden types and outcomes are simulated.

E.8 Semi-synthetic credit protocol

We use 1,000 profiles from the UCI South German Credit data set under CC BY 4.0 (UCI, 2019). Its UCI ID is 522 and its DOI is 10.24432/C5X89F. The real inputs are eight financial covariates: duration, transformed amount, checking status, savings, employment duration, instalment-rate category, credit history, and number of existing credits. We exclude age, the combined personal status and sex attribute, and foreign-worker status. The historical credit outcome is never used.

Real covariates determine probabilities over $V \in \{0.8, 1.4, 2.6\}$ and $\alpha \in \{0.75, 1, 1.25\}$. We set $\Gamma = V\alpha$, so each ratio occurs at three absolute scales. The focal menu selects $1 \leq V \leq 2$ and $1 \leq \Gamma \leq 2$. This experiment uses the $t = 0$ member of the finite transfer grid. New-loan default is a simulated logistic function of V, Γ, and a real-covariate financial-risk score. It is independent of the selected logging action. Removing this action mismatch favors the α-only comparator.

Approval uses the same coefficient that enters choice utility. A subject selecting offer S_i from Table 2 receives approval with probability p_{S_i}. The focal propensity is 0.8; other selected offers have propensities 0.6, 0.8, or 1. Ties follow the displayed order. All nine supported type cells select a genuine offer, so every simulated type has positive approval probability. The exit has zero approval probability and contributes no observed default.

Let $Z_i = \frac{T_i D_i}{p_{S_i}}$ when $p_{S_i} > 0$ and $Z_i = 0$ otherwise. Let F_i indicate focal choice. The direct-cell estimate is $\frac{\sum_i F_i Z_i}{\sum_i F_i}$. The favorable α-only estimate averages Z_i within each exact ratio group and

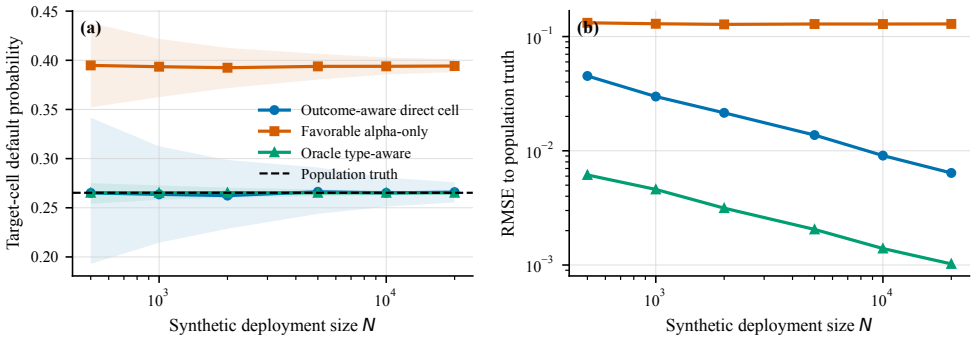


Figure 7: Semi-synthetic credit experiment with 400 independent repetitions per sample size. The left panel shows mean estimates and the fifth to ninety-fifth percentiles across repetitions. The right panel shows RMSE. Eight real UCI financial covariates determine the simulated type distribution and outcome probabilities; choices and approvals follow the displayed menu.

combines those averages with the oracle target-cell ratio weights. It uses all approved outcomes and their selected-action propensities but pools absolute scales within each ratio. The third estimator is an oracle that observes each target subject's simulated default probability.

Subjects are sampled with replacement from the 1,000 profiles. For $N \in \{500, 1000, 2000, 5000, 10^4, 2 \cdot 10^4\}$, we run 400 repetitions with seed 20260901.

The exact target-cell mass is 0.30, its risk is 0.26524143, and the α-only population limit is 0.39385701. At $N = 20{,}000$, direct, α-only, and oracle RMSEs are 0.00639794, 0.12900236, and 0.00102284, respectively. Their Monte Carlo standard errors are 0.00024151, 0.00019837, and 0.00003632. Logarithmic RMSE regressions on logarithmic sample size give slopes -0.5247, -0.0045, and -0.4937 over the six values of N. Within the exact $\alpha = 1$ cells, simulated risks at low, middle, and high scale are 0.11488, 0.26112, and 0.70725.

Run the following commands to reproduce the experiment:

```
cd paper
/opt/anaconda3/bin/python \
 experiments/semisynthetic_credit/run_semisynthetic_credit.py
/opt/anaconda3/bin/python \
 experiments/semisynthetic_credit/test_semisynthetic_credit.py
/opt/anaconda3/bin/python \
 figures/semisynthetic_credit/gen_fig_semisynthetic_credit.py
```

The run takes about two seconds with Python 3.12.7 and NumPy 1.26.4. CSV files store every repetition and population cell. JSON records the complete menu, selected propensities, data checksum, and source checksum. The six checks cover rectangle choice geometry, positivity on the type support, zero-propensity handling, exact inverse-propensity identities, a seeded approval-frequency check, and agreement between saved replicates and summary statistics.

E.9 Empirical evidence represented by the experiments

The controlled experiments instantiate selected theorem predictions under the stated simulated configurations. The semi-synthetic study combines public covariates with simulated strategic types, policies, and new-loan outcomes. The empirical claims concern theorem recovery and finite-sample behavior; counterfactual lending effects and legal validation remain deployment questions.