# PARTIAL IDENTIFICATION OF NETWORK CAUSAL EFFECTS UNDER STRUCTURAL SUPPORT DEFICIENCY

**Anonymous authors**
Paper under double-blind review

## ABSTRACT

Repeated network experiments can leave policy contrasts unidentified even as sampling error vanishes: operational constraints assign zero probability to relevant exposures. We characterize this limitation through a compact completion polytope combining outcome bounds, scientific linear restrictions, and bidirectional transformation distances respecting one-way operational boundaries. Under the finite-population mean-response model, its target image is a sharp identified set: diameter $\omega$ implies worst-case squared-error risk at least $\omega^2/4$ for every point estimator under any number of unchanged-design repetitions. Optimal dual faces characterize when these bounds coincide and their local sensitivity to supported means and restrictions. To infer this set, we project a design-based confidence polytope for supported means, obtaining finite-sample full-set coverage; independent stable waves permit affine refinement through exact common-dual vertex certification. Matching upper and block-experiment lower bounds give minimax expected interval length of order $D_0 + M\sqrt{(1+\Delta)/(Np)}$ for the specified fixed-dimensional classes with uniformly bounded endpoint sensitivity, identification-diameter bound $D_0$, supported-cell probabilities $\geq p$, and dependency degree $\leq \Delta$. In 600-repetition experiments, widths approach the $0.20$ identification plateau, while projection attains $0.9917$–$0.9967$ full-set coverage versus $\leq 0.0067$ for raw optimization. The predicted dependence scaling appears in a variance log–log slope of $0.9976$ against $1 + \Delta$; across 4,000 end-to-end runs, dependence-calibrated full-set coverage is $1.000$, versus $0.910$–$0.918$ with a unit-iid radius. In a six-cell insurance-network audit, descriptive projection spans all held-out empirical means at average width $0.922$, versus unrestricted width $1.000$. These results connect the limits of causal learning to experimental design and specify when greater precision requires new identifying information.

**Keywords:** partial identification; network interference; exposure mapping; structural positivity; randomization inference; linear programming duality; dependency graph.

## 1 INTRODUCTION

Network experiments operate under joint constraints. A platform may change regional subsidies while capping neighboring congestion; a service system may randomize outreach while limiting simultaneous demand. These guardrails remove entire treatment configurations. Exposure mappings translate the remaining assignments into local environments (Hudgens & Halloran, 2008; Manski, 2013; Aronow & Samii, 2017), but some policy-relevant exposures receive exactly zero probability. Repeating the experiment estimates observed exposures more precisely while leaving these missing counterfactuals unresolved.

Three kinds of absence have different implications. A *sample zero* has positive design probability but is unobserved in one sample. A *current-design zero* is a meaningful exposure excluded by the deployed randomization. A *globally infeasible label* lies outside the scientific state space and is removed from the target. We study current-design zeros. A prediction model crosses these holes through its extrapolation assumptions; our objective is to characterize the full range supported by explicit scientific restrictions.



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.

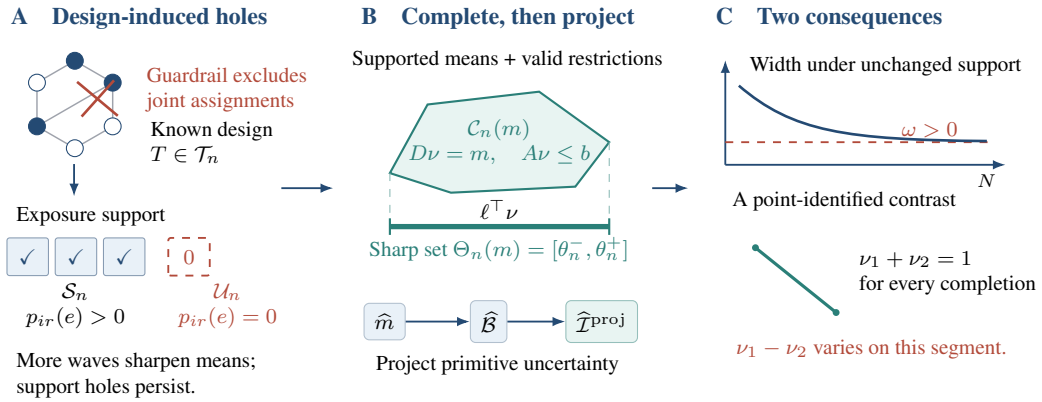


Figure 1: **Support constraints, completion, and target identification.** (A) Operational guardrails create zero-probability exposure cells. (B) Supported means and mechanism-valid restrictions define a completion polytope; its target projection gives a sharp identified interval, and projecting a confidence region accounts for sampling uncertainty. (C) More observations can leave a positive identification diameter, whereas a contrast constant along all feasible missing directions is point identified. The curve and segment are schematic consequences of Theorems 3.1–3.2, Corollary 3.1, and Theorem 5.2.

We encode admissible comparisons by a directed graph of mechanism-preserving transformations. Two-way path costs constrain exposure-mean differences while respecting one-way operational boundaries. Outcome bounds and linear scientific restrictions then define a completion polytope. Its image under the chosen contrast is sharp, and its optimal dual faces identify which supported means and restrictions determine each endpoint. Figure 1 shows the resulting separation between learning an observed mean and identifying a counterfactual target.

The remaining obstacle is statistical. Even unbiased feasible estimates can produce inward-biased endpoints after optimization, causing the plug-in interval to miss part of the population identified set. We instead project a simultaneous design-based confidence region through the completion constraints. Independent waves additionally permit an affine refinement: a pilot region certifies one dual representation at every vertex, and a separate wave evaluates the certified endpoints. An explicit block experiment matches the attainable interval scale with a lower bound.

The results separate identification, inference, and experimental design. Sharp completion characterizes both persistent nonidentification and contrasts that remain point identified under incomplete support. Projection covers the complete identified set in finite samples, while an exact certificate permits affine endpoint evaluation. Dual sensitivity quantifies the value of supported means and structural restrictions. Matching bounds establish the minimax interval scale $D_0 + M\sqrt{(1+\Delta)/(Np)}$ on the stated fixed-dimensional classes. Simulations evaluate full-set coverage, and observed insurance-network data test extrapolation across missing exposure cells.

## 2 Related Work

**Network interference and support.** Exposure mappings organize causal effects under partial and general interference (Hudgens & Halloran, 2008; Tchetgen Tchetgen & VanderWeele, 2012; Aronow & Samii, 2017; Basse & Airoldi, 2018). Subsequent work develops randomization tests, observational identification, cluster designs, approximate interference, bipartite experiments, and marketplace randomization (Athey et al., 2018; Forastiere et al., 2021; Jagadeesan et al., 2020; Leung, 2022a;b; Zigler & Papadogeorgou, 2021; Harshaw et al., 2023; Masoero et al., 2026). Rooted-network comparisons transfer information across comparable neighborhoods (Auerbach et al., 2026). Our transformation graph specifies admissible comparisons between exposure means after the assignment law removes target cells.

**Partial identification and optimized inference.** Weak overlap affects precision and regular estimation (Crump et al., 2009; Khan & Tamer, 2010; D'Amour et al., 2021); exact support failure changes identification. Sharp Lipschitz LP bounds cover off-policy values (Khan et al., 2024) and network policy evaluation (Wang et al., 2026). We study exposure contrasts under constrained finite-population randomization and dependent design-based scores. The distinction between identification and inference is classical (Manski, 2003; Imbens & Manski, 2004; Chernozhukov et al., 2007; Tamer, 2010; Kaido et al., 2019; Molinari, 2020); recent methods address perturbed LP values, conditional LPs, and feasibility testing (Cho & Russell, 2024; Voronin, 2025; Ben-Michael, 2025; Goff & Mbakop, 2026). Known assignment probabilities support our finite-sample primitive region and full-set projection.

**Scientific restrictions and uncertainty.** Misspecified mappings and networks affect causal interpretation (Sävje et al., 2021; Sävje, 2024; Weinstein & Nevo, 2026; Wikle & Zigler, 2024); recent work studies sensitivity, learned mappings, and mapping-agnostic inference (Schröder et al., 2026; Huber et al., 2026; He & Tchetgen Tchetgen, 2026). We condition on a declared mapping and propagate bounded mapping error in Appendix K. Metric restrictions formalize information transfer (Villani, 2009; Blanchet & Murthy, 2019); polyhedral sensitivity (Rockafellar, 1997; Rockafellar & Wets, 1998; Bonnans & Shapiro, 2013) characterizes the local sensitivity of our interval endpoints.

## 3 Framework and Identification

The identification argument separates the information fixed by the design from the variation left to scientific restrictions. We first quantify the estimation risk that persists under unchanged support, then show that the completion interval exhausts the admissible target values (Theorems 3.1 and 3.2). Dual sensitivity identifies the constraints carrying this uncertainty, and target constancy characterizes when it disappears (Proposition 3.1 and Corollary 3.1).

**Model and support.** Fix a population $V_n$, network $G_n$, and pretreatment information $\mathcal{F}_0$. In wave $r$, a known design draws $T^{(r)} \in \mathcal{T}_n$, inducing exposure $E_i^{(r)} = d_i(T^{(r)}, G_n, X_{1:n}) \in \mathcal{E}$. The maintained mean-response model is

$$\mathbb{E}\{Y_i^{(r),\mathrm{obs}} \mid T^{(r)}, \mathcal{F}_0\} = m_i(E_i^{(r)}), \qquad 0 \le Y_i^{(r),\mathrm{obs}}, m_i(e) \le M.$$

Write $\mu_{0e} = n^{-1}\sum_i m_i(e)$ and $\theta_0 = \ell^\top \mu_0$ for a pre-specified contrast. With known $p_{ir}(e) = \mathbb{P}(E_i^{(r)} = e \mid \mathcal{F}_0)$, assume the global partition

$$\mathcal{S}_n = \{e : \min_{i,r} p_{ir}(e) \ge p_{\min} > 0\}, \qquad \mathcal{U}_n = \{e : p_{ir}(e) = 0 \,\forall i, r\}, \quad \mathcal{E} = \mathcal{S}_n \dot\cup \mathcal{U}_n.$$

Globally infeasible labels are excluded from $\mathcal{E}$. Let $D$ select the supported coordinates, so the design identifies $m_0 = D\mu_0$. Appendix J extends the target to eligible populations when support varies by unit.

**Scientific completion.** Let $\mathcal{V} = (\mathcal{E}, \mathcal{R})$ be a directed graph of admissible transformations. Each edge preserves the mechanism being compared and has a nonnegative scientific cost. Paths compose these local comparisons; a missing path has cost $+\infty$. If $c_\to(e, e')$ is the shortest directed transformation cost, use $d(e, e') = \max\{c_\to(e, e'), c_\to(e', e)\}$. The restriction $|\nu_e - \nu_{e'}| \le L d(e, e')$ applies to mutually reachable pairs. One-way knowledge instead gives directed rows, such as $\nu_{e'} - \nu_e \le L c_\to(e, e')$. This distinction preserves operational boundaries even when exposure labels are numerically close. Collect these comparisons, outcome bounds, and auxiliary restrictions in $A\nu \le b$, where $A$ has $r_A$ rows. Monotonicity and benchmark moments are examples; all rows are maintained valid at $\mu_0$. The graph, costs, and slope bound $L$ form part of the scientific specification. For $K = |\mathcal{E}|$, define

$$\mathcal{C}_n(m) = \{\nu \in \mathbb{R}^K : D\nu = m,\ A\nu \le b\}, \quad \mathcal{M}_n = \{m : \mathcal{C}_n(m) \ne \varnothing\}, \quad \Theta_n(m) = \ell^\top \mathcal{C}_n(m).$$

For $m \in \mathcal{M}_n$, the compact convex image is an interval $\Theta_n(m) = [\theta_n^-(m), \theta_n^+(m)]$, with diameter $\omega_n(m)$. These conditions are recorded as Assumptions A.1–A.3.

**Theorem 3.1** (Persistent nonidentification)**.** *Under the maintained exposure, support, and structural conditions, suppose $\omega_n(m_0) > 0$. For every number of unchanged-design waves $R$, there exist admissible $P^-, P^+$ with the same observed-data law and targets $\theta_n^-(m_0), \theta_n^+(m_0)$. Every observed-data estimator therefore satisfies*

$$\max_{\sigma\in\{-,+\}} \mathbb{E}_{P^\sigma}\{\widehat{\theta}_R - \theta(P^\sigma)\}^2 \geq \omega_n(m_0)^2/4.$$

**Theorem 3.2** (Sharp completion)**.** *Under the same conditions, $\Theta_n(m_0)$ is exactly the set of target values generated by bounded finite-population mean-response schedules satisfying the stated restrictions and supported means $m_0$. Every value in the interval is attainable. Any two completions can be embedded in admissible submodels with identical observed-data laws under every repetition of the current design.*

**Why the set is sharp.** Given a completion $\nu$, assign every unit the constant response mean $\widetilde{m}_i(e) = \nu_e$. These schedules satisfy the bounds and mean restrictions. All completions share supported coordinates, and the design never reveals the remaining coordinates, so the schedules can produce identical observations. Convex combinations of endpoint completions attain every intermediate target. The same indistinguishability proves Theorem 3.1: under the common observed law, the larger squared error at two targets separated by $\omega_n(m_0)$ is at least $\omega_n(m_0)^2/4$.

For objective $c \in \mathbb{R}^K$, set

$$\phi_{n,c}(m) = \min_{\nu\in\mathcal{C}_n(m)} c^\top \nu, \qquad \mathcal{D}_c = \{(\rho,\lambda) : D^\top\rho - A^\top\lambda = c,\ \lambda \geq 0\}.$$

Write $\theta_n^-(m) = \phi_{n,\ell}(m)$ and $\theta_n^+(m) = -\phi_{n,-\ell}(m)$; omit the argument at $m = m_0$. The equality multiplier $\rho$ prices supported means, while $\lambda$ prices structural restrictions.

**Proposition 3.1** (Dual sensitivity)**.** *For $m \in \mathcal{M}_n$, primal and dual optima are attained and $\phi_{n,c}(m) = \max_{(\rho,\lambda)\in\mathcal{D}_c}(\rho^\top m - \lambda^\top b)$. Let $\mathcal{D}_c^*(m)$ be the optimal face. Along a feasible joint direction $(h,k)$,*

$$\phi'_{n,c}\{(m,b);(h,k)\} = \max_{(\rho,\lambda)\in\mathcal{D}_c^*(m)}(\rho^\top h - \lambda^\top k).$$

*Relaxing row $j$ has right derivative $-\min_{\mathcal{D}_c^*(m)}\lambda_j$. A positive minimum identifies a first-order bottleneck across all optimal representations. For $m \in \mathrm{ri}(\mathcal{M}_n)$, set $\mathsf{T}_n = \mathrm{lin}(\mathcal{M}_n - \mathcal{M}_n)$. The relative subdifferential is the projected optimal face:*

$$\partial_{\mathsf{T}_n}\phi_{n,c}(m) = \{\Pi_{\mathsf{T}_n}\rho : \exists\lambda,\ (\rho,\lambda) \in \mathcal{D}_c^*(m)\}.$$

**Corollary 3.1** (Target-specific point identification)**.** *For $m \in \mathcal{M}_n$, the following are equivalent: $\Theta_n(m)$ is a singleton; $\phi_{n,\ell}(m) + \phi_{n,-\ell}(m) = 0$; and $\ell^\top(\nu - \nu') = 0$ for all $\nu, \nu' \in \mathcal{C}_n(m)$.*

Thus incomplete support can leave the full mean vector undetermined while identifying one contrast. In Figure 1C, all completions satisfy $\nu_1+\nu_2 = 1$, although $\nu_1-\nu_2$ varies. Precision and identification require different experimental changes.

## 4 Estimation and Certified Inference

We now propagate uncertainty in supported means through the identification geometry. The primitive variance bound calibrates a simultaneous region, whose projection supplies the default interval (Theorem 4.1). A common dual representation converts both endpoints to affine functions precisely when a finite vertex test succeeds; this certificate defines the optional branch in Figure 2 (Theorem 4.2).

For independent wave blocks $\mathcal{R}_w$, let $N_w = n|\mathcal{R}_w|$, $s = |\mathcal{S}_n|$, and $\Delta_w$ be the maximum degree of a dependency graph for the unit–wave score vectors, conditional on $\mathcal{F}_0$. Estimate supported means by

$$\widehat{\mu}_e^{(w)} = \frac{1}{N_w}\sum_{r\in\mathcal{R}_w}\sum_{i=1}^{n}\frac{\mathbf{1}\{E_i^{(r)} = e\}Y_i^{(r),\mathrm{obs}}}{p_{ir}(e)}.$$

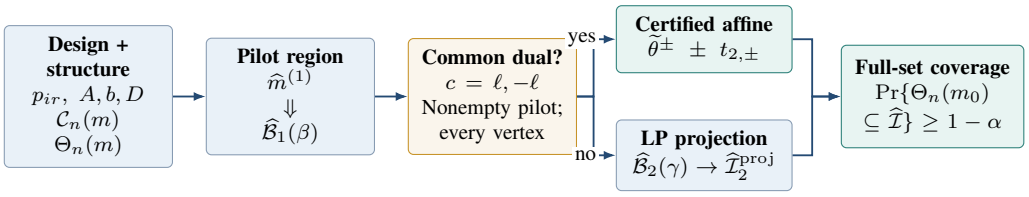


Figure 2: **From identification to inference.** Sharp completion gives the population interval (Theorem 3.2); unbiased primitives calibrate the pilot region (Theorem 4.1); the common-dual vertex test selects affine evaluation or projection (Theorem 4.2). Independent evaluation gives full-set coverage (Theorem 5.1), and the width and lower-bound results separate identification from sampling (Theorems 5.2 and 5.3).

**Theorem 4.1** (Unbiased primitives and network variance)**.** *Under the exposure and support conditions and stable independent waves (Assumption A.4), for $e \in \mathcal{S}_n$ and $a \in \mathbb{R}^s$,*

$$\mathbb{E}(\widehat{\mu}_e^{(w)} \mid \mathcal{F}_0) = \mu_{0e}, \qquad \mathrm{Var}\{a^\top \widehat{m}^{(w)} \mid \mathcal{F}_0\} \leq \frac{M^2(1+\Delta_w)}{N_w p_{\min,w}} \|a\|_2^2.$$

*Here $p_{\min,w} = \min_{i,r\in\mathcal{R}_w,e\in\mathcal{S}_n} p_{ir}(e)$.*

The effective exposure count is $N_w p_{\min,w}/(1+\Delta_w)$. Endpoint optimization nevertheless changes the direction of error: for an unbiased feasible $\widehat{m}$, convexity and concavity give $\mathbb{E}\phi_{n,\ell}(\widehat{m}) \geq \theta_n^-$ and $\mathbb{E}\{-\phi_{n,-\ell}(\widehat{m})\} \leq \theta_n^+$ (Proposition C.1). Projection accounts for this optimization step:

$$\widehat{\mathcal{B}}_w(\eta) = \{m \in \mathcal{M}_n : \|m - \widehat{m}^{(w)}\|_\infty \leq r_w(\eta)\}, \qquad r_w(\eta) = M\sqrt{\frac{(1+\Delta_w)s}{N_w p_{\min,w}\eta}},$$

$$\widehat{\mathcal{I}}_w^{\mathrm{proj}}(\eta) = \left[\min_{m\in\widehat{\mathcal{B}}_w(\eta)} \phi_{n,\ell}(m),\ \max_{m\in\widehat{\mathcal{B}}_w(\eta)} -\phi_{n,-\ell}(m)\right].$$

The variance bound and a union argument imply $\mathbb{P}\{m_0 \in \widehat{\mathcal{B}}_w(\eta) \mid \mathcal{F}_0\} \geq 1 - \eta$. If the region is empty, the procedure reports incompatibility and returns the structural-only range over $\{\nu : A\nu \leq b\}$.

**Theorem 4.2** (Exact common-dual certificate)**.** *For a nonempty compact polytope $B \subseteq \mathcal{M}_n$ and $(\bar{\rho}, \bar{\lambda}) \in \mathcal{D}_c$, dual optimality at every vertex of $B$ is equivalent to dual optimality throughout $B$. A common pair therefore exists exactly when*

$$D^\top \rho - A^\top \lambda = c, \quad \lambda \geq 0, \qquad \rho^\top v - \lambda^\top b = \phi_{n,c}(v) \quad \textit{for all } v \in \mathrm{vert}(B)$$

*is feasible. The certificate accommodates multiple optimal duals.*

**Why vertices suffice.** Every point in $B$ is a convex combination of vertices. If one feasible dual pair attains the value at all vertices, its affine objective equals the corresponding convex combination of vertex values. Convexity bounds the primal value above by this combination, while weak duality bounds it below by the same affine objective. Equality follows throughout $B$. A check at the center provides only one equality and leaves room for a change of active face elsewhere in the pilot region.

**Certified evaluation.** Choose $\beta + \gamma_- + \gamma_+ \leq \alpha$, with positive budgets, and set $\gamma = \gamma_- + \gamma_+$. Test both objectives $c = \ell, -\ell$ on $\widehat{\mathcal{B}}_1(\beta)$. If the pilot region is nonempty and both tests pass, retain common pairs $(\widehat{\rho}_\pm, \widehat{\lambda}_\pm)$ and use the independent evaluation block to compute

$$\widetilde{\theta}^- = \widehat{\rho}_-^\top \widehat{m}^{(2)} - \widehat{\lambda}_-^\top b, \qquad \widetilde{\theta}^+ = -(\widehat{\rho}_+^\top \widehat{m}^{(2)} - \widehat{\lambda}_+^\top b),$$

$$t_{2,\pm} = M\|\widehat{\rho}_\pm\|_2 \sqrt{\frac{1+\Delta_2}{N_2 p_{\min,2}} \frac{1-\gamma_\pm}{\gamma_\pm}}.$$

Return the hull of $\widetilde{\theta}^{-} - t_{2,-}$ and $\widetilde{\theta}^{+} + t_{2,+}$. If either certificate fails, return $\widehat{\mathcal{I}}_2^{\mathrm{proj}}(\gamma)$, with the empty-region rule above. The pilot fixes the branch before evaluation data are used. Algorithm 1 gives full pseudocode, and Appendix M gives an exact vertex-separation implementation.

**Algorithm 1** Certified hybrid inference under structural support deficiency

**Require:** $A, b, D, \ell, M$; independent pilot/evaluation wave blocks; $\beta+\gamma_-+\gamma_+ \le \alpha$, $\gamma = \gamma_- + \gamma_+$.

1: Compute $\widehat{m}^{(1)}$, $r_1(\beta)$, and $\widehat{\mathcal{B}}_1(\beta)$.
2: **if** $\widehat{\mathcal{B}}_1(\beta) = \varnothing$ **then**
3: Mark both endpoint objectives uncertified.
4: **else**
5: **for** $c \in \{\ell, -\ell\}$ **do**
6: Use Theorem 4.2 to test for a common dual pair on $\widehat{\mathcal{B}}_1(\beta)$, and retain one common pair if it exists.
7: **if** both endpoint objectives admit a common pair **then**
8: Set

$$\widetilde{\theta}^{-} = \widehat{\rho}_-^{\top}\widehat{m}^{(2)} - \widehat{\lambda}_-^{\top} b, \qquad \widetilde{\theta}^{+} = -\{\widehat{\rho}_+^{\top}\widehat{m}^{(2)} - \widehat{\lambda}_+^{\top} b\}.$$

9: Return the interval hull of $\widetilde{\theta}^{-} - t_{2,-}$ and $\widetilde{\theta}^{+} + t_{2,+}$.
10: **else**
11: Construct

$$\widehat{\mathcal{C}}_2(\gamma) := \{\nu : A\nu \le b,\ \left\| D\nu - \widehat{m}^{(2)} \right\|_\infty \le r_2(\gamma)\}.$$

12: Return the minimum and maximum of $\ell^{\top}\nu$ over $\widehat{\mathcal{C}}_2(\gamma)$. If the set is empty, flag incompatibility and return the structural-only range over $\{\nu : A\nu \le b\}$.
13: Report the selected branch, confidence radii, certificate status, multiplier ranges over each optimal dual face, and numerical optimality gaps.

## 5 Statistical Guarantees

The statistical argument proceeds from coverage to width, optimality, and structural sensitivity. Pilot certification and independent evaluation give finite-sample coverage (Theorem 5.1); polyhedral sensitivity bounds the sampling enlargement (Theorem 5.2); block experiments establish the minimax scale (Theorem 5.3 and Corollary 5.1). Row-wise relaxation then prices uncertainty in the scientific restrictions (Theorem 5.4).

**Theorem 5.1** (Finite-sample full-set coverage)**.** *Under the maintained model, valid structure, known support, and stable independent waves, the certified procedure with $\beta + \gamma_- + \gamma_+ \le \alpha$ satisfies*

$$\mathbb{P}\{\Theta_n(m_0) \subseteq \widehat{\mathcal{I}}^{\mathrm{hyb}}_{n,1-\alpha} \mid \mathcal{F}_0\} \ge 1 - \alpha.$$

*The one-wave projection with error budget $\eta$ likewise covers $\Theta_n(m_0)$ with probability at least $1-\eta$.*

**Why selection preserves coverage.** On the event that the pilot region contains $m_0$, a successful certificate makes the selected affine formulas exact at both population endpoints. Conditional on the pilot, independence leaves the evaluation means unbiased, and the two one-sided variance bounds spend $\gamma_- + \gamma_+$. On the projection branch, the evaluation region contains $m_0$ with the same error budget. Adding the pilot error $\beta$ proves the guarantee. The error bound holds conditional on either pilot-selected branch, so averaging over the pilot preserves coverage.

For a compact convex $K_0 \subset \mathrm{ri}(\mathcal{M}_n)$, let $\Gamma_{n,c}(K_0) = \sup_{m\in K_0,\ g\in\partial_{\mathsf{T}_n}\phi_{n,c}(m)} \|g\|_1$, where $\mathsf{T}_n = \mathrm{lin}(\mathcal{M}_n - \mathcal{M}_n)$. These relative slope moduli are finite; set $\Gamma_n(K_0) = \max_{c=\ell,-\ell}\Gamma_{n,c}(K_0)$.

**Theorem 5.2** (Identification–sampling decomposition)**.** *Let $B = \{m \in \mathcal{M}_n : \|m - \widehat{m}\|_\infty \le r\}$ be nonempty and contained in $K_0$. Then*

$$\mathrm{diam}(\widehat{\mathcal{I}}^{\mathrm{proj}}) \le \sup_{m\in B} \omega_n(m) + 2r\Gamma_n(K_0).$$

*If $m_0 \in B$, then*

$$\mathrm{diam}(\widehat{\mathcal{I}}^{\mathrm{proj}}) \le \omega_n(m_0) + 2r\{\Gamma_{n,\ell}(K_0) + \Gamma_{n,-\ell}(K_0)\}.$$

With bounded slope moduli and fixed $s$, the sampling term has order $M\sqrt{(1+\Delta)/(Np_{\min})}$ at fixed confidence. The identification term depends on support and scientific restrictions. Under a positive common-dual margin, certification succeeds with probability tending to one; dependency-graph normal limits then refine the affine branch (Condition D.1, Proposition D.1, and Theorem D.3).

**Theorem 5.3** (Identification plus statistical lower bound). *Fix $\alpha \in (0, 1/2)$, $p \in (0, 1]$, $q = 1 + \Delta$ dividing $n$, $N = nR$, and $D_0 \in [0, M]$. Choose $c > 0$ with $\kappa_c = 1 - 2\alpha - \sqrt{8}c > 0$ and $\epsilon_N = c\sqrt{q/(Np)} \le 1/4$. There exist admissible block experiments $P_\pm$, with score degree $\Delta$, exposure probability $p$, and*

$$\Theta(P_\pm) = [M(1/2 \pm \epsilon_N),\ M(1/2 \pm \epsilon_N) + D_0], \qquad \mathrm{KL}(P_+, P_-) \le 16c^2,$$

*such that every interval covering $\Theta(P)$ with probability at least $1 - \alpha$ for both experiments obeys*

$$\sup_{P \in \{P_+, P_-\}} \mathbb{E}_P \operatorname{diam}(\widehat{I}_N) \ge (1-\alpha)D_0 + 2\kappa_c c M \sqrt{\frac{1+\Delta}{Np}}.$$

**Why dependence and support multiply.** The lower-bound construction groups $q$ units around one shared response, leaving $N/q$ independent block–wave observations. Only a fraction $p$ reveals the supported exposure, so perturbations of size $\sqrt{q/(Np)}$ remain statistically difficult to distinguish. A separate unsupported coordinate contributes diameter $D_0$ under either law. The testing argument therefore adds a statistical separation cost to the cost of covering the unidentified segment. This construction explains the same two terms that appear in the projection width bound.

**Corollary 5.1** (Minimax interval scale). *Fix $\alpha \in (0, 1/2)$ and $K, s, M$. Consider a class $\mathcal{P}_N$ with $p_{\min} \ge p_\star \in (0, 1/2]$, score degree at most $\bar{\Delta}$, $\sup_{m \in \mathcal{M}_n} \omega_n(m) \le D_0$ uniformly, and both endpoint maps globally $\bar{\Gamma}$-Lipschitz in $\|\cdot\|_\infty$, for fixed $1 \le \bar{\Gamma} < \infty$. Suppose the class contains the pair in Theorem 5.3 with $p = p_\star$, $\Delta = \bar{\Delta}$, and $c = (1-2\alpha)/(2\sqrt{8})$. Along $1 + \bar{\Delta} \mid n$, for sufficiently large $N$,*

$$\inf_{\widehat{I}:\ \inf_{P \in \mathcal{P}_N} \mathbb{P}_P\{\Theta(P) \subseteq \widehat{I}\} \ge 1-\alpha} \ \sup_{P \in \mathcal{P}_N} \mathbb{E}_P \operatorname{diam}(\widehat{I}) \asymp_{\alpha, K, s, \bar{\Gamma}} D_0 + M\sqrt{\frac{1+\bar{\Delta}}{Np_\star}}.$$

**Sensitivity to scientific restrictions.** The preceding guarantees use a valid structural system. Relaxation makes the consequence of a declared violation explicit and connects robustness to the same dual prices that govern precision (Theorem 5.4).

**Theorem 5.4** (Slackened structural validity). *Let $\varepsilon \in \mathbb{R}^{r_A}_+$, replace $b$ by $b + \varepsilon$, and write $\phi^\varepsilon_{n,c}(m)$ for the relaxed value. If both programs are feasible, every $(\rho^*, \lambda^*) \in \mathcal{D}^*_c(m)$ satisfies*

$$0 \le \phi_{n,c}(m) - \phi^\varepsilon_{n,c}(m) \le (\lambda^*)^\top \varepsilon.$$

*If $A\mu_0 \le b + \varepsilon$, the relaxed completion set is valid for the true mean vector.*

Taking the smallest penalty over the optimal face gives the strongest bound of this form. Distance uncertainty contributes row slack proportional to $L$, while certified primal–dual gaps add explicit endpoint enlargements (Corollary D.3 and Proposition D.2). Thus scientific calibration, sampling uncertainty, and numerical error each enter a stated part of the reported interval. Proofs and expanded statements appear in the appendix.

## 6 Synthetic and Field-Data Experiments

The experiments examine identification width, plug-in bias, projection coverage, vertex certification, and the effect of network dependence. The release uses master seed 20260916, 600 repetitions per component cell, and 1,000 repetitions at each of four end-to-end network sizes. Scripts archive numerical tables, vector figures, solver diagnostics, and file hashes. Appendix E gives the full factor grid and software versions; the Reproducibility statement records the independent replay.

**Identification and projection.** With unchanged support, the population diameter is 0.20: mean projected width falls from 0.520 after one wave to 0.220 after 256 waves, with full-set coverage 0.995–0.998. At $N = 30{,}000$, the positivity experiment gives width 0.311 at $p = 0$, against diameter 0.30, and 0.0229 at $p = 0.30$. On a common polytope, $\mu_{u_2} + \mu_{u_3} = 1$ is point identified while $\mu_{u_2} - \mu_{u_3} \in [-0.5, 0.1]$. Figure 3 displays this geometry.

At 250 observations per supported cell, raw lower and upper endpoint biases are $+0.0187$ and $-0.0169$. Across sample sizes 50–2500, raw full-set coverage is at most 0.0067, whereas projection achieves 0.9917–0.9967. The certificate experiment checks 18,000 pilot boxes: exact vertex certification produces zero false passes and activates with probability one at normalized distance from the dual kink at least four. A center-only rule falsely certifies 56.7% of boxes. Figures 5 and 10 report the diagnostics.

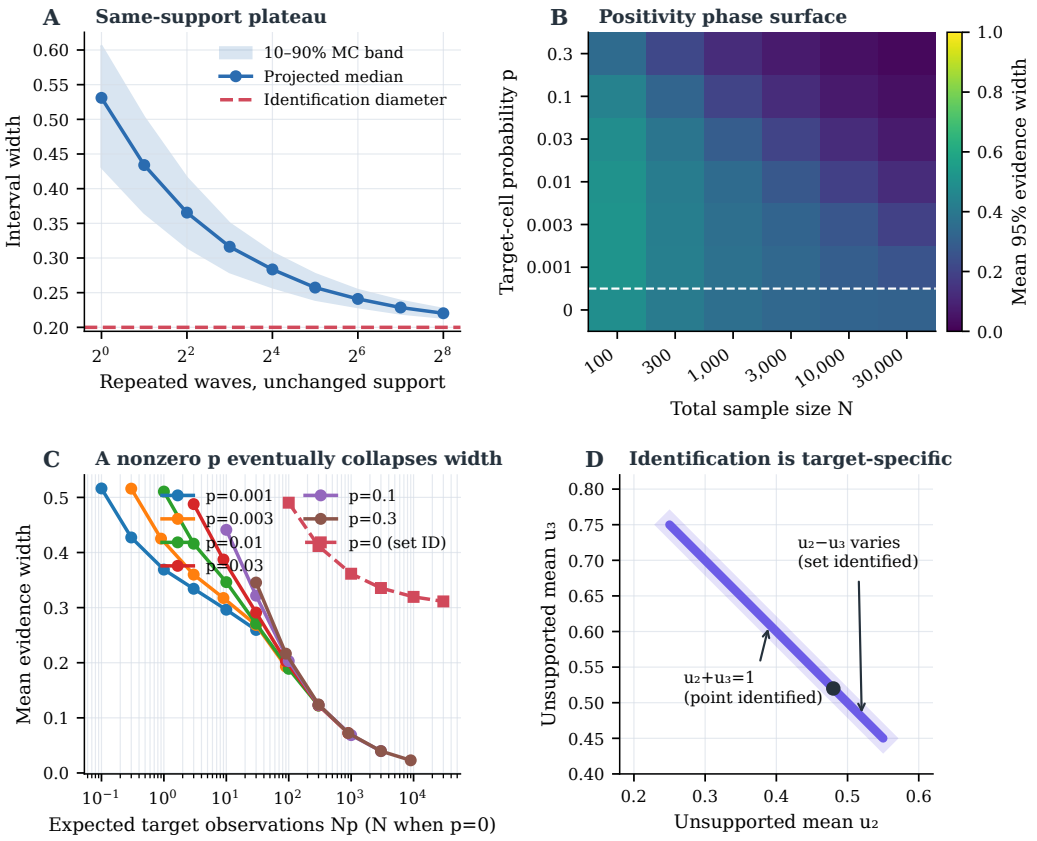


Figure 3: **Identification phase atlas.** Repetition reduces interval width toward the fixed-support diameter (A); positive target-cell probability changes the limiting width (B–C); a common completion polytope point identifies one contrast while leaving another set identified (D). Bands in A are 10th–90th Monte Carlo quantiles across 600 repetitions.

**Network dependence and structural calibration.** A union of six-node rings executes assignment, exposure mapping, Horvitz–Thompson estimation, a graph-Hoeffding primitive region, and LP projection. Its sharp interval is $[0.47, 0.62]$. Across 4,000 runs, dependence-calibrated full-set coverage is 1.000, with mean width falling from 0.739 to 0.359; a unit-iid radius covers only 0.910–0.918. The block experiment yields variance slope 0.9976 against $1 + \Delta$. Mechanism-valid distance reduces projected width from 0.605 to 0.225 at $n = 2500$, retaining coverage one. A true restriction violation of 0.10 requires slack 0.10 to recover coverage. Figures 6–8 separate these mechanisms.

**Field-data stress test.** The public insurance-network replication (Leung, 2020; Cai et al., 2015; 2019) contains 142 households and 343 non-self dependency links. Own information assignment

crossed with treated-friend counts $0, 1, \geq 2$ forms six cells of sizes $41, 21, 9, 43, 19, 9$. Each leave-cell-out fit calibrates structure on the other five cells. Raw structural and Euclidean intervals span four of six held-out means at average widths $0.600$ and $0.570$; every interval intersects its unadjusted Wilson reference interval. The largest gap occurs in the intensive/high-peer cell: adoption $0.778$, structural bounds $[0.125, 0.524]$, and Euclidean bounds $[0.182, 0.524]$. Logistic and linear extrapolation have mean absolute errors $0.293$ and $0.299$.

Wilson-based projection spans all six held-out means at mean width $0.922$, compared with unrestricted width $1.000$. The high-peer endpoint is determined by the control/high-peer and intensive/one-peer cells (Figure 4). Appendix E.1 extends this audit to paired-cell holdouts. These observed-mean comparisons use training-calibrated restrictions and descriptive Wilson intervals. The simulation experiments evaluate the design-based coverage guarantee.

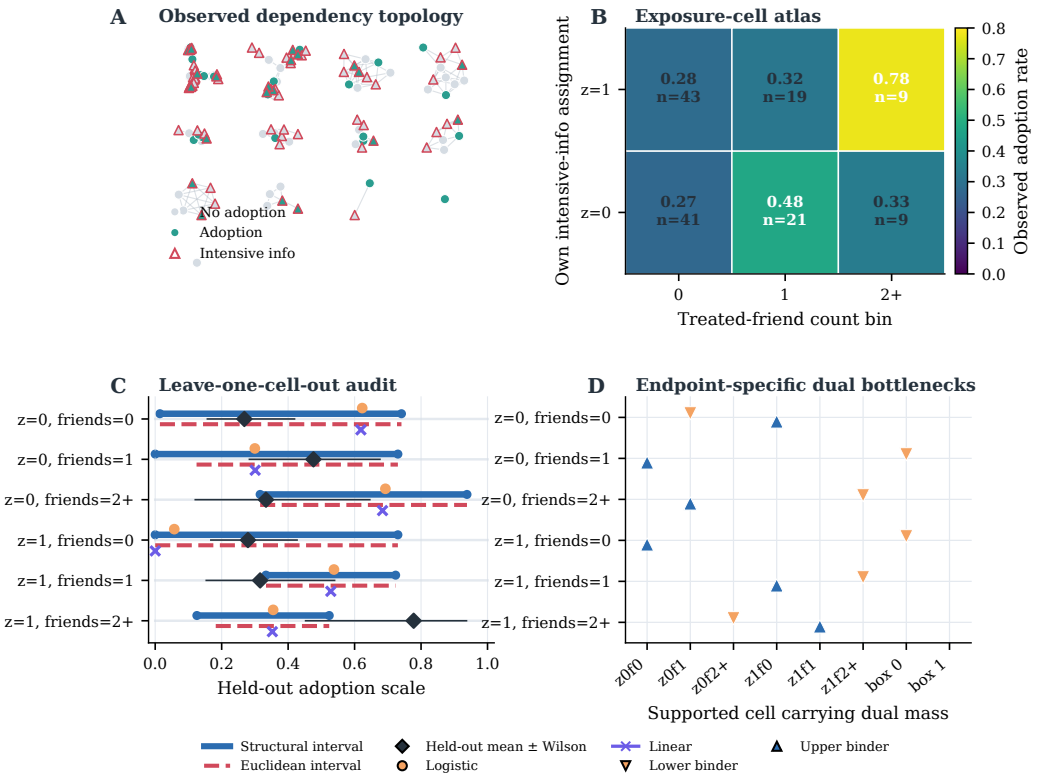


Figure 4: **Insurance-network stress test.** The bundled dependency graph (A) and exposure cells (B) define six leave-cell-out fits (C). Diamonds and thin bars are held-out means and unadjusted Wilson references; thick blue and dashed red lines are structural and Euclidean intervals; circles and crosses are logistic and linear predictions. Endpoint binders (D) identify which retained cells constrain each missing target.

## 7 Discussion

The interval separates incomplete identification from sampling uncertainty. The completion polytope gives the population range, and projection accounts for estimating supported means. Repeated experiments reduce sampling uncertainty; additional target-relevant support or stronger valid restrictions reduce the identification diameter.

Dual multipliers distinguish where more observations improve endpoint precision from where relaxing structural assumptions changes identification. Equality multipliers govern the former; minimum structural multipliers over the optimal face govern the latter. These derivatives quantify local changes under the maintained model.

### AI USE STATEMENT

We used generative AI tools to support literature discovery, develop the theoretical framework and mathematical claims, draft and check proof arguments and proof text, design experiments and interpret their results, revise the manuscript, and develop experiment code, figures, and Lean formalizations. We manually checked the AI-assisted mathematical arguments and proof steps. We take full responsibility for the accuracy, originality, and integrity of this work, including its mathematical arguments, citations, code, and reported results.

### ETHICS STATEMENT

Our field-data analysis reuses publicly released, processed insurance-network data (Cai et al., 2015; Leung, 2020). The supplementary artifact records the data provenance and preserves the source attribution and upstream license. Policy applications should assess which populations and policy options the experimental design leaves unsupported and justify the restrictions used to extrapolate beyond that support.

### REPRODUCIBILITY STATEMENT

The supplementary artifact contains the Python suite, fixed seeds, numerical outputs, vector figures, file hashes, and a replay checker. An independent full replay reproduced all 39 numerical outputs of the four original modules within absolute tolerance $10^{-8}$ and relative tolerance $10^{-10}$. The comparison is recorded in `experiment_release/REPLAY_STATUS.md`. The paired-cell extension includes its fixed protocol, input hashes, held-out-outcome perturbation checks, and independent LP verification (Appendix E.1). The `LeanProofs` project kernel-checks 13 deterministic propositions (Appendix F); the mathematical appendix gives the remaining proofs.

## REFERENCES


Peter M Aronow and Cyrus Samii. Estimating average causal effects under general interference, with application to a social network experiment. *The Annals of Applied Statistics*, 11(4):1912–1947, 2017.

Susan Athey, Dean Eckles, and Guido W Imbens. Exact p-values for network interference. *Journal of the American Statistical Association*, 113(521):230–240, 2018.

Eric Auerbach, Hongchang Guo, and Max Tabord-Meehan. The local approach to causal inference under network interference. *Quantitative Economics*, 17(1):173–199, 2026. doi: 10.3982/QE2484.

Guillaume W Basse and Edoardo M Airoldi. Limitations of design-based causal inference and a/b testing under arbitrary and network interference. *Sociological Methodology*, 48(1):136–151, 2018.

Eli Ben-Michael. Partial identification via conditional linear programs: Estimation and policy learning. *arXiv preprint arXiv:2506.12215*, 2025.

Jose Blanchet and Karthyek Murthy. Quantifying distributional model risk via optimal transport. *Mathematics of Operations Research*, 44(2):565–600, 2019.

J Frédéric Bonnans and Alexander Shapiro. *Perturbation analysis of optimization problems*. Springer Science & Business Media, 2013.

Jing Cai, Alain De Janvry, and Elisabeth Sadoulet. Social networks and the decision to insure. *American Economic Journal: Applied Economics*, 7(2):81–108, 2015.

Jing Cai, Alain De Janvry, and Elisabeth Sadoulet. Replication data for: Social networks and the decision to insure. *Nashville, TN: American Economic Association, Ann Arbor, MI: Inter-university Consortium for Political and Social Research*, pp. 10–12, 2019.

Victor Chernozhukov, Han Hong, and Elie Tamer. Estimation and confidence regions for parameter sets in econometric models. *Econometrica*, 75(5):1243–1284, 2007.

JoonHwan Cho and Thomas M. Russell. Simple inference on functionals of set-identified parameters defined by linear moments. *Journal of Business & Economic Statistics*, 42(2):563–578, 2024.

Richard K Crump, V Joseph Hotz, Guido W Imbens, and Oscar A Mitnik. Dealing with limited overlap in estimation of average treatment effects. *Biometrika*, 96(1):187–199, 2009.

Alexander D'Amour, Peng Ding, Avi Feller, Lihua Lei, and Jasjeet Sekhon. Overlap in observational studies with high-dimensional covariates. *Journal of Econometrics*, 221(2):644–654, 2021.

Laura Forastiere, Edoardo M Airoldi, and Fabrizia Mealli. Identification and estimation of treatment and interference effects in observational studies on networks. *Journal of the American Statistical Association*, 116(534):901–918, 2021.

Leonard Goff and Eric Mbakop. Testing the solvability of systems of linear inequalities. *arXiv preprint arXiv:2506.06776*, 2026.

Christopher Harshaw, Fredrik Sävje, David Eisenstat, Vahab Mirrokni, and Jean Pouget-Abadie. Design and analysis of bipartite experiments under a linear exposure-response model. *Electronic Journal of Statistics*, 17(1):464–518, 2023.

Yihui He and Eric J. Tchetgen Tchetgen. A general exposure-mapping-agnostic framework for causal inference under interference. *arXiv preprint arXiv:2607.04644*, 2026.

Martin Huber, Jannis Kueck, and Mara Mattes. Learning and testing exposure mappings of interference using graph convolutional autoencoders. In *Proceedings of the Fifth Conference on Causal Learning and Reasoning*, volume 323 of *Proceedings of Machine Learning Research*, pp. 656–675, 2026.

Michael G Hudgens and M Elizabeth Halloran. Toward causal inference with interference. *Journal of the American Statistical Association*, 103(482):832–842, 2008.

Guido W Imbens and Charles F Manski. Confidence intervals for partially identified parameters. *Econometrica*, 72(6):1845–1857, 2004.

Ravi Jagadeesan, Natesh S Pillai, and Alexander Volfovsky. Designs for estimating the treatment effect in networks with interference. *The Annals of Statistics*, 48(2):679–712, 2020.

Svante Janson. Normal convergence by higher semiinvariants with applications to sums of dependent random variables and random graphs. *The Annals of Probability*, 16(1):305–312, 1988.

Hiroaki Kaido, Francesca Molinari, and Jörg Stoye. Confidence intervals for projections of partially identified parameters. *Econometrica*, 87(4):1397–1432, 2019.

Samir Khan, Martin Saveski, and Johan Ugander. Off-policy evaluation beyond overlap: Sharp partial identification under smoothness. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 23734–23757, 2024.

Shakeeb Khan and Elie Tamer. Irregular identification, support conditions, and inverse weight estimation. *Econometrica*, 78(6):2021–2042, 2010.

Michael P. Leung. Treatment and spillover effects under network interference. *The Review of Economics and Statistics*, 102(2):368–380, 2020. doi: 10.1162/rest_a_00818.

Michael P Leung. Causal inference under approximate neighborhood interference. *Econometrica*, 90(1):267–293, 2022a.

Michael P Leung. Rate-optimal cluster-randomized designs for spatial interference. *The Annals of Statistics*, 50(5):3064–3087, 2022b.

C. F. Manski. Identification of treatment response with social interactions. *The Econometrics Journal*, 16(1):S1–S23, 2013.

Charles F Manski. *Partial identification of probability distributions*. Springer, 2003.

Lorenzo Masoero, Suhas Vijaykumar, Thomas S Richardson, James McQueen, Ido Rosen, Brian Burdick, Pat Bajari, and Guido Imbens. Multiple randomization designs: Estimation and inference with interference. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 88(3):958–977, 2026.

Francesca Molinari. Microeconometrics with partial identification. *Handbook of Econometrics*, 7 (Part A):355–486, 2020.

R Tyrrell Rockafellar. *Convex analysis*, volume 28. Princeton University Press, 1997.

R Tyrrell Rockafellar and Roger JB Wets. *Variational analysis*. Springer, 1998.

Fredrik Sävje. Causal inference with misspecified exposure mappings: separating definitions and assumptions. *Biometrika*, 111(1):1–15, 2024.

Fredrik Sävje, Peter Aronow, and Michael Hudgens. Average treatment effects in the presence of unknown interference. *Annals of Statistics*, 49(2):673–701, 2021.

Maresa Schröder, Miruna Oprescu, Stefan Feuerriegel, and Nathan Kallus. Causal inference on networks under misspecified exposure mappings: A partial identification framework. *arXiv preprint arXiv:2602.03459*, 2026.

Elie Tamer. Partial identification in econometrics. *Annu. Rev. Econ.*, 2(1):167–195, 2010.

Eric J. Tchetgen Tchetgen and Tyler J. VanderWeele. On causal inference in the presence of interference. *Statistical Methods in Medical Research*, 21(1):55–75, 2012.

Cédric Villani. *Optimal transport: old and new*, volume 338. Springer, 2009.

Andrei Voronin. Linear programming approach to partially identified econometric models. *arXiv preprint arXiv:2503.14940*, 2025.

Ziyan Wang, Yiran Liu, and Zhiheng Zhang. Partial identification of policy values under network interference. In *Forty-third International Conference on Machine Learning*, 2026.

Bar Weinstein and Daniel Nevo. Causal inference with misspecified network interference structure. *Biometrics*, 82(1):ujag023, 2026.

Nathan B Wikle and Corwin M Zigler. Causal health impacts of power plant emission controls under modeled and uncertain physical process interference. *The Annals of Applied Statistics*, 18 (4):2753–2774, 2024.

Corwin M Zigler and Georgia Papadogeorgou. Bipartite causal inference with interference. *Statistical Science: a review journal of the Institute of Mathematical Statistics*, 36(1):109–123, 2021.

## APPENDIX CONTENTS


**A Problem Setup** **14**
A.1 Finite-population mean-response model with repeated randomization . . . . . 14
A.2 Structural support . . . . . 14
A.3 Mechanism-consistent structural restrictions . . . . . 14
A.4 Assumptions, diagnostics, and fallbacks . . . . . 15

**B Expanded Identification Results** **15**
B.1 A fundamental negative result . . . . . 16
B.2 Sharpness and endpoint geometry . . . . . 16
B.3 Primal–dual representation and bottleneck sensitivity . . . . . 17

**C Estimation and Certified Hybrid Inference** **18**
C.1 Replicated Horvitz–Thompson primitives . . . . . 18
C.2 Why direct plug-in endpoints are biased inward . . . . . 18
C.3 Polyhedral confidence regions . . . . . 19
C.4 An exact common-dual certificate . . . . . 19
C.5 Certified hybrid algorithm . . . . . 19

**D Expanded Statistical Results** **20**
D.1 Finite-sample coverage of the complete identified set . . . . . 20
D.2 Width control and polyhedral calmness . . . . . 20
D.3 Stable-face asymptotic refinement . . . . . 21
D.4 Structural and computational robustness . . . . . 21
D.5 An explicit minimax lower bound . . . . . 22

**E Experimental Design and Additional Evidence** **23**
E.1 Paired exposure-cell holdouts on observed insurance data . . . . . 23

**F Machine-Checked Deterministic Core** **25**

**G Notation and Implementation Objects** **26**

**H Auxiliary Lemmas** **27**

**I Proofs of Main Results** **31**
I.1 Proof of Theorem 3.1 . . . . . 31
I.2 Proof of Theorem 3.2 . . . . . 31
I.3 Proof of Proposition 3.1 . . . . . 32
I.4 Proof of Corollaries 3.1 and B.2 . . . . . 32
I.5 Proof of Theorem 4.1 . . . . . 32
I.6 Proof of Proposition C.1 . . . . . 32
I.7 Proof of Theorem 4.2 . . . . . 33
I.8 Proof of Theorem 5.1 . . . . . 33
I.9 Proof of Theorem 5.2 and Corollary D.1 . . . . . 34
I.10 Proof of Proposition D.1 . . . . . 34
I.11 Proof of Theorem D.3 and Corollary D.2 . . . . . 34
I.12 Proofs of Theorem 5.4, Corollary D.3, and Proposition D.2 . . . . . 35
I.13 Proof of Theorem 5.3 and Corollary 5.1 . . . . . 35

**J Heterogeneous and Partially Supported Exposure Cells** **36**

**K Exposure-Mapping Approximation** **37**

**L Variance Estimation and Zero Joint-Inclusion Probabilities** **37**

**M Common-Dual Cutting-Plane Implementation** **39**

# A Problem Setup

## A.1 Finite-population mean-response model with repeated randomization

Fix a finite population $V_n = \{1, \dots, n\}$, a graph $G_n = (V_n, E_n)$, and pretreatment covariates $X_{1:n}$. All are conditioned on throughout. A feasible global intervention is a vector $T \in \mathcal{T}_n$, where $\mathcal{T}_n$ need not be a Cartesian product. Budgets, safety rules, fairness constraints, and operational conflicts may exclude joint assignments even when each coordinate action is individually feasible.

The experiment can be repeated across stable, nonoverlapping waves $r = 1, \dots, R$. In wave $r$, an assignment $T^{(r)}$ is drawn from a known design $\mathcal{P}_{n,r}$. A fixed exposure mapping produces

$$E_i^{(r)} = d_i\Big(T^{(r)}, G_n, X_{1:n}\Big) \in \mathcal{E} = \{1, \dots, K\}. \tag{1}$$

For each unit–exposure pair there is a bounded mean response $m_i(e) \in [0, M]$. The observed outcome satisfies the mean-exposure restriction

$$\mathbb{E}\Big(Y_i^{(r),\mathrm{obs}} \mid T^{(r)}, \mathcal{F}_0\Big) = m_i\Big(E_i^{(r)}\Big), \qquad 0 \le Y_i^{(r),\mathrm{obs}} \le M, \tag{2}$$

where $\mathcal{F}_0$ contains the fixed population, graph, covariates, and mean-response schedule. Deterministic finite-population potential outcomes are included as the degenerate special case. The exposure means and target contrast are

$$\mu_{0e} := \frac{1}{n}\sum_{i=1}^{n} m_i(e), \qquad \mu_0 = (\mu_{01}, \dots, \mu_{0K})^\top, \qquad \theta_0 := \ell^\top \mu_0, \tag{3}$$

with $\ell \in \mathbb{R}^K$ fixed before outcomes are inspected.

## A.2 Structural support

Let

$$p_{ir}(e) := \mathbb{P}_{\mathcal{P}_{n,r}}\{E_i^{(r)} = e\}.$$

The main text adopts a transparent global-cell convention: each exposure label is either uniformly supported or absent under every current design wave. Specifically, for a known $p_{\min} > 0$,

$$\begin{aligned} \mathcal{S}_n &:= \{e \in \mathcal{E} : \min_{i,r} p_{ir}(e) \ge p_{\min}\}, \\ \mathcal{U}_n &:= \{e \in \mathcal{E} : p_{ir}(e) = 0 \text{ for every } i, r\}, \qquad \mathcal{S}_n \dot{\cup} \mathcal{U}_n = \mathcal{E}. \end{aligned} \tag{4}$$

This convention directly represents an exposure template that the operational policy never allows. Unit-specific eligibility and partially supported cells can be handled by expanding the primitive coordinate system to eligible-population means; Appendix J gives the exact modification and states the additional transport interpretation required when eligible populations differ.

Let $s = |\mathcal{S}_n|$, let $D \in \mathbb{R}^{s \times K}$ select the supported coordinates, and write

$$m_0 := D\mu_0 = \mu_{0,\mathcal{S}_n}. \tag{5}$$

The current design identifies $m_0$, but it contains no direct observations from $e \in \mathcal{U}_n$.

## A.3 Mechanism-consistent structural restrictions

Let $\mathcal{V} = (\mathcal{E}, \mathcal{R})$ be a directed graph of scientifically admissible local transformations. An edge $(e, e') \in \mathcal{R}$ means that exposure $e$ can be changed to $e'$ without crossing an unmodeled mechanism or operational rule boundary; its cost is $w(e, e') \in [0, \infty)$. Let $c_{\to}(e, e')$ be the directed shortest-path cost, with value $+\infty$ if no finite path exists. Because the main restriction below is symmetric, define the two-way extended distance

$$d(e, e') := \begin{cases} \max\{c_{\to}(e, e'), c_{\to}(e', e)\}, & c_{\to}(e, e'), c_{\to}(e', e) < \infty, \\ +\infty, & \text{otherwise.} \end{cases} \tag{6}$$

Symmetric restrictions use mutually reachable pairs. Directional scientific knowledge enters through one-sided rows such as $\nu_{e'} - \nu_e \le L c_{\to}(e, e')$.

For a supported mean vector $m$, define the structural-completion set

$$\begin{aligned}\mathcal{C}_n(m) := \{\nu \in \mathbb{R}^K : 0 \le \nu_e \le M \quad (e \in \mathcal{E}),\\ D\nu = m,\\ |\nu_e - \nu_{e'}| \le L d(e,e') \quad \text{when } d(e,e') < \infty,\\ H\nu \le h\},\end{aligned} \tag{7}$$

where $H\nu \le h$ contains optional scientifically justified linear restrictions, including monotonicity, signed spillover restrictions, benchmark moments, and capacity-calibration inequalities. No row is imposed when $d(e,e') = \infty$. Collect all box, distance, and auxiliary rows in $A \in \mathbb{R}^{r_A \times K}$ and $b \in \mathbb{R}^{r_A}$. Then

$$\mathcal{C}_n(m) = \{\nu \in \mathbb{R}^K : D\nu = m,\ A\nu \le b\}, \qquad \mathcal{M}_n := \{m : \mathcal{C}_n(m) \ne \varnothing\}. \tag{8}$$

**Definition A.1** (Population partial-identification region)**.** *For $m_0 \in \mathcal{M}_n$, define*

$$\Theta_n(m_0) := \{\ell^\top \nu : \nu \in \mathcal{C}_n(m_0)\} = [\theta_n^-, \theta_n^+], \tag{9}$$

*where*

$$\theta_n^- := \min_{\nu \in \mathcal{C}_n(m_0)} \ell^\top \nu, \qquad \theta_n^+ := \max_{\nu \in \mathcal{C}_n(m_0)} \ell^\top \nu, \qquad \omega_n(m_0) := \theta_n^+ - \theta_n^-. \tag{10}$$

*The object $\Theta_n(m_0)$ is a population identified set. A confidence interval must additionally account for estimation of $m_0$.*

**Example A.1** (A stylized subsidy–throttling contrast)**.** *Let $e_{ab}$ index own-subsidy level $a \in \{0,1,2\}$ and neighboring-throttle indicator $b \in \{0,1\}$. Suppose $e_{21}$ is structurally absent, while*

$$\mu_{00} = 0.40, \quad \mu_{10} = 0.46, \quad \mu_{20} = 0.52, \quad \mu_{01} = 0.36, \quad \mu_{11} = 0.42$$

*are supported. If $M = 1$, $L = 0.12$, and the transformation graph gives unit two-way distance from $e_{21}$ to both $e_{20}$ and $e_{11}$, then*

$$\mu_{21} \in [\mu_{20} - 0.12, \mu_{20} + 0.12] \cap [\mu_{11} - 0.12, \mu_{11} + 0.12] = [0.40, 0.54].$$

*The target $\mu_{21} - \mu_{00}$ is therefore sharply restricted to $[0, 0.14]$, compared with the unrestricted box interval $[-0.40, 0.60]$. The lower endpoint is controlled by the link to $e_{20}$, whereas the upper endpoint is controlled by the link to $e_{11}$; the dual multipliers recover exactly this asymmetry.*

### A.4 Assumptions, diagnostics, and fallbacks

**Assumption A.1** (Mean-exposure validity)**.** *The mapping $d_i$ is fixed before outcome analysis, $K < \infty$, and (2) holds with a common finite bound $M$.*

**Assumption A.2** (Known design and global-cell support)**.** *The assignment probabilities $p_{ir}(e)$ are known for $e \in \mathcal{S}_n$, and the partition (4) holds.*

**Assumption A.3** (Structural validity)**.** *The population mean vector satisfies $A\mu_0 \le b$; equivalently, $\mu_0 \in \mathcal{C}_n(m_0)$.*

**Assumption A.4** (Stable independent randomization waves)**.** *Pilot and evaluation waves target the same fixed $m_0$. Conditional on $\mathcal{F}_0$, distinct waves are independent, the graph and exposure mapping are stable, and there is no carryover. Within a wave, the score vectors admit a dependency graph with maximum degree $\Delta_w$.*

## B Expanded Identification Results

For a generic objective $c \in \mathbb{R}^K$, define the endpoint value

$$\phi_{n,c}(m) := \min\{c^\top \nu : D\nu = m,\ A\nu \le b\}. \tag{11}$$

Then $\theta_n^- = \phi_{n,\ell}(m_0)$ and $\theta_n^+ = -\phi_{n,-\ell}(m_0)$.

Table 1: Role, diagnostic content, and fallback for each maintained assumption.

| Assumption | Why it enters | Observable diagnostic | Response to violation |
|---|---|---|---|
| Mean–exposure validity | Defines the finite exposure state and links observed outcomes to its cell means | Pre-specification audit; residual contrasts within nominal exposure cells; alternative mappings | Use the mapping-error enlargement in Appendix K, refine the mapping, or report mapping-indexed bounds |
| Known design and support | Makes zero probability an identified property of the randomization law and calibrates Horvitz–Thompson scores | Enumerate feasible assignments; compare analytic and simulated exposure frequencies | Redefine eligible-population targets as in Appendix J or enlarge probability uncertainty |
| Structural validity | Supplies the scientific content that narrows unsupported means | Hold out supported cells; inspect rule-specific slack paths and dual prices | Add explicit row slack via Theorem 5.4, weaken the rule graph, or return to box bounds |
| Stable independent waves | Separates pilot selection from evaluation and controls score dependence | Check randomization logs, carryover windows, graph stability, and component sizes | Use the one-wave projected procedure; enlarge the dependency bound; abandon the affine refinement |

### B.1 A FUNDAMENTAL NEGATIVE RESULT

The first theorem separates support deficiency from ordinary imprecision: whenever the target varies over the completion set, no amount of data from the unchanged design can select one point uniformly over the maintained model.

**Theorem B.1** (Persistent nonidentification under an unchanged design)**.** *Suppose Assumptions A.1–A.3 hold and $\omega_n(m_0) > 0$. For every number of repeated waves $R$ and every estimator $\widehat{\theta}_R$ measurable with respect to their observed data, the model class contains two admissible data-generating processes $P^-$ and $P^+$ that induce the same observed-data law under the current design and satisfy*

$$\theta(P^-) = \theta_n^-, \qquad \theta(P^+) = \theta_n^+.$$

*Consequently,*

$$\max_{\sigma\in\{-,+\}} \mathbb{E}_{P^\sigma}\left(\widehat{\theta}_R - \theta(P^\sigma)\right)^2 \geq \frac{\omega_n(m_0)^2}{4}. \tag{12}$$

*In particular, point estimation cannot be uniformly consistent unless the identified-set diameter vanishes.*

**Interpretation.** The obstruction is exact: lower and upper completions agree on every exposure generated by the design yet disagree on the target, so the corresponding worlds are observationally equivalent. Their separation and the resulting risk persist with arbitrarily many repetitions. Stronger scientific restrictions or a design that creates target-relevant support resolve the ambiguity.

### B.2 SHARPNESS AND ENDPOINT GEOMETRY

**Theorem B.2** (Sharp structural completion)**.** *Under Assumptions A.1–A.3, $\Theta_n(m_0)$ is exactly the set of target values generated by finite-population mean-response schedules whose exposure-mean vector satisfies the stated restrictions. More precisely:*

*(i) for every $\vartheta \in \Theta_n(m_0)$, there exists a bounded mean-response schedule $\{\widetilde{m}_i(e)\}$ satisfying the exposure model and all structural restrictions, with supported mean vector $m_0$ and target $\vartheta$;*

*(ii) every admissible schedule with supported mean vector $m_0$ has a target in $\Theta_n(m_0)$;*

*(iii) any two completions $\nu, \nu' \in \mathcal{C}_n(m_0)$ can be embedded in two admissible submodels that have the same observed-data distribution under every repetition of the current design and targets $\ell^\top \nu$ and $\ell^\top \nu'$, respectively.*

*The schedules in (i)–(iii) may vary across unsupported exposures while preserving every supported population mean.*

Theorem 3.2 identifies the completion interval with the target values compatible with the supported means and structural restrictions.

### B.3 Primal–dual representation and bottleneck sensitivity

The dual-feasible set for objective $c$ is

$$\mathcal{D}_c := \{(\rho, \lambda) : D^\top \rho - A^\top \lambda = c,\ \lambda \geq 0\}. \tag{13}$$

For $m \in \mathcal{M}_n$, linear-programming duality gives

$$\phi_{n,c}(m) = \max_{(\rho,\lambda)\in\mathcal{D}_c} \{\rho^\top m - \lambda^\top b\}. \tag{14}$$

The equality multiplier $\rho$ prices supported-cell information; the nonnegative multiplier $\lambda$ prices the structural right-hand sides.

**Proposition B.1** (Duality and set-valued sensitivity)**.** *For every $m \in \mathcal{M}_n$, the primal and dual optima in (11)–(14) are attained and have equal value. Let*

$$\mathcal{D}_c^*(m) := \arg\max_{(\rho,\lambda)\in\mathcal{D}_c} \{\rho^\top m - \lambda^\top b\}.$$

*For $m \in \mathrm{ri}(\mathcal{M}_n)$, let $\mathsf{T}_n := \mathrm{lin}(\mathcal{M}_n - \mathcal{M}_n)$ and let $\Pi_{\mathsf{T}_n}$ be orthogonal projection onto $\mathsf{T}_n$. Then*

$$\partial_{\mathsf{T}_n} \phi_{n,c}(m) = \{\Pi_{\mathsf{T}_n} \rho : \exists \lambda, (\rho, \lambda) \in \mathcal{D}_c^*(m)\}, \tag{15}$$

*where the right-hand side is understood as the convex projected optimal face. For every feasible direction $h \in \mathsf{T}_n$,*

$$\phi'_{n,c}(m; h) = \max_{(\rho,\lambda)\in\mathcal{D}_c^*(m)} \rho^\top h. \tag{16}$$

*Jointly as a function of $(m, b)$, every optimal pair supplies a subgradient $(\rho, -\lambda)$, and for every feasible joint direction $(h, k)$,*

$$\phi'_{n,c}\{(m, b); (h, k)\} = \max_{(\rho,\lambda)\in\mathcal{D}_c^*(m)} \{\rho^\top h - \lambda^\top k\}. \tag{17}$$

*In particular, the right derivative from relaxing row $j$ is*

$$\frac{d}{dt_+} \phi_{n,c}(m; b + te_j)\bigg|_{t=0} = -\min_{(\rho,\lambda)\in\mathcal{D}_c^*(m)} \lambda_j. \tag{18}$$

*Hence row $j$ is an indispensable first-order bottleneck precisely when its minimum multiplier over the complete optimal dual face is positive.*

**Dual-face audit.** For each supported cell and structural row, we report its attainable multiplier range over the optimal dual face, allowing an unbounded upper endpoint. A strictly positive minimum certifies that every dual representation prices that row; a range containing zero records degeneracy or redundancy. The ranges quantify local sensitivity under the specified structural system. The validity of each row is a separate scientific assumption.

**Corollary B.1** (Target-specific point identification)**.** *For any $m \in \mathcal{M}_n$, the following are equivalent:*

*(i) $\Theta_n(m)$ is a singleton;*

*(ii) $\phi_{n,\ell}(m) + \phi_{n,-\ell}(m) = 0$;*

*(iii)* $\ell^\top(\nu - \nu') = 0$ *for every* $\nu, \nu' \in \mathcal{C}_n(m)$.

*Therefore partial exposure support can coexist with point identification of a particular contrast, even though the full mean vector is not point identified.*

**Corollary B.2** (Full-support and unrestricted-box limits)**.** *If* $\mathcal{S}_n = \mathcal{E}$*, then* $\mathcal{C}_n(m_0) = \{\mu_0\}$ *and* $\Theta_n(m_0) = \{\ell^\top \mu_0\}$*. If unsupported coordinates are constrained only by* $0 \le \nu_e \le M$*, then*

$$\theta_n^- = \sum_{e \in \mathcal{S}_n} \ell_e \mu_{0e} + M \sum_{e \in \mathcal{U}_n : \ell_e < 0} \ell_e, \tag{19}$$

$$\theta_n^+ = \sum_{e \in \mathcal{S}_n} \ell_e \mu_{0e} + M \sum_{e \in \mathcal{U}_n : \ell_e > 0} \ell_e. \tag{20}$$

# C Estimation and Certified Hybrid Inference

## C.1 Replicated Horvitz–Thompson primitives

Partition the randomization waves into a pilot block $\mathcal{R}_1$ and an evaluation block $\mathcal{R}_2$, with $R_w = |\mathcal{R}_w|$ and effective unit–wave count $N_w = nR_w$. For $e \in \mathcal{S}_n$, define

$$\widehat{\mu}_e^{(w)} := \frac{1}{nR_w} \sum_{r \in \mathcal{R}_w} \sum_{i=1}^{n} \frac{\mathbf{1}\{E_i^{(r)} = e\} Y_i^{(r),\mathrm{obs}}}{p_{ir}(e)}, \qquad \widehat{m}^{(w)} = (\widehat{\mu}_e^{(w)} : e \in \mathcal{S}_n). \tag{21}$$

Let $\Delta_w$ be the maximum degree of a dependency graph for the $N_w$ unit–wave score vectors. Independence across waves makes this graph the disjoint union of its within-wave components.

**Theorem C.1** (Primitive unbiasedness and network-dependent variance)**.** *Under Assumptions A.1, A.2, and A.4,*

$$\mathbb{E}\Big(\widehat{\mu}_e^{(w)} \mid \mathcal{F}_0\Big) = \mu_{0e} \qquad (e \in \mathcal{S}_n). \tag{22}$$

*For every* $a \in \mathbb{R}^s$*,*

$$\mathrm{Var}\Big\{a^\top(\widehat{m}^{(w)} - m_0) \mid \mathcal{F}_0\Big\} \le \frac{M^2(1 + \Delta_w)}{N_w p_{\min,w}} \|a\|_2^2, \tag{23}$$

*where* $p_{\min,w} := \min_{i,r \in \mathcal{R}_w, e \in \mathcal{S}_n} p_{ir}(e)$*. More sharply, the factor* $p_{\min,w}^{-1} \|a\|_2^2$ *may be replaced by*

$$\max_{i,r} \sum_{e \in \mathcal{S}_n} \frac{a_e^2}{p_{ir}(e)}.$$

The bound reveals the statistical effective sample size

$$N_{\mathrm{eff},w} \asymp \frac{N_w p_{\min,w}}{1 + \Delta_w}. \tag{24}$$

It is a sampling quantity. It has no effect on the population diameter $\omega_n(m_0)$.

## C.2 Why direct plug-in endpoints are biased inward

The lower endpoint is a convex function of the supported means and the upper endpoint is concave. Unbiased primitive estimation therefore does not imply unbiased endpoint estimation.

**Proposition C.1** (One-sided optimization bias)**.** *Let* $\widehat{m}$ *be integrable, satisfy* $\mathbb{E}\widehat{m} = m$*, and belong to* $\mathcal{M}_n$ *almost surely. Then*

$$\mathbb{E}\phi_{n,\ell}(\widehat{m}) \ge \phi_{n,\ell}(m), \qquad \mathbb{E}\{-\phi_{n,-\ell}(\widehat{m})\} \le -\phi_{n,-\ell}(m). \tag{25}$$

*Thus the expected plug-in endpoints lie inside the population endpoints. Infeasible estimates are handled by the projected procedure defined below.*

### C.3 POLYHEDRAL CONFIDENCE REGIONS

For $\eta \in (0,1)$, define the coordinate radius

$$r_w(\eta) := M\sqrt{\frac{s(1+\Delta_w)}{N_w p_{\min,w}\eta}} \tag{26}$$

and the feasible confidence polytope

$$\widehat{\mathcal{B}}_w(\eta) := \{m \in \mathcal{M}_n : \left\|m - \widehat{m}^{(w)}\right\|_\infty \le r_w(\eta)\}. \tag{27}$$

A coordinatewise Chebyshev bound and a union bound give

$$\mathbb{P}\{m_0 \in \widehat{\mathcal{B}}_w(\eta) \mid \mathcal{F}_0\} \ge 1-\eta. \tag{28}$$

The polyhedral form gives a finite vertex certificate and represents interval projection as a linear program.

### C.4 AN EXACT COMMON-DUAL CERTIFICATE

For a nonempty polytope $B \subseteq \mathcal{M}_n$, define the common-optimal dual set

$$\mathcal{D}_c^*(B) := \bigcap_{v \in \mathrm{vert}(B)} \mathcal{D}_c^*(v). \tag{29}$$

A pair in $\mathcal{D}_c^*(B)$ represents the endpoint value by the same affine function at every $m \in B$.

**Theorem C.2** (Finite-vertex certificate for a common dual face)**.** *Let $B \subseteq \mathcal{M}_n$ be a nonempty compact polytope and let $(\bar\rho, \bar\lambda) \in \mathcal{D}_c$. The following are equivalent:*

*(i) $(\bar\rho, \bar\lambda)$ is dual-optimal for every $m \in B$;*

*(ii) it is dual-optimal at every vertex $v \in \mathrm{vert}(B)$.*

*Consequently, $\mathcal{D}_c^*(B) \neq \varnothing$ can be checked by solving the endpoint LP at the vertices and then the linear feasibility system*

$$\begin{cases} D^\top \rho - A^\top \lambda = c, \quad \lambda \ge 0, \\ \rho^\top v - \lambda^\top b = \phi_{n,c}(v), \quad v \in \mathrm{vert}(B). \end{cases} \tag{30}$$

*For large vertex sets, Appendix M gives a finite constraint-generation implementation based on an exact vertex-separation oracle; worst-case complexity follows the combinatorics of the pilot polytope.*

Theorem 4.2 checks optimality over the entire pilot polytope. Multiple optimal dual solutions are allowed whenever their common intersection is nonempty.

### C.5 CERTIFIED HYBRID ALGORITHM

Choose deterministic error budgets

$$\beta, \gamma_-, \gamma_+ \in (0,1), \qquad \beta + \gamma_- + \gamma_+ \le \alpha, \qquad \gamma := \gamma_- + \gamma_+. \tag{31}$$

For a pilot-certified equality multiplier $\widehat\rho_\sigma$, $\sigma \in \{-,+\}$, define

$$V_{2,\sigma} := \frac{M^2(1+\Delta_2)}{N_2 p_{\min,2}} \left\|\widehat\rho_\sigma\right\|_2^2, \qquad t_{2,\sigma}(\gamma_\sigma) := \sqrt{V_{2,\sigma}\frac{1-\gamma_\sigma}{\gamma_\sigma}}. \tag{32}$$

Algorithm 1 in the main text gives the complete procedure.

The pilot fixes the branch, and the evaluation block supplies the selected interval. This separation preserves the conditional error budget in Theorem 5.1.

# D Expanded Statistical Results

## D.1 Finite-sample coverage of the complete identified set

**Theorem D.1** (Certified hybrid finite-sample coverage)**.** *Suppose Assumptions A.1–A.4 hold and Algorithm 1 is run with (31). Let $\widehat{\mathcal{I}}^{\mathrm{hyb}}_{n,1-\alpha}$ denote the interval returned by its pre-specified branch. Then, conditional on the fixed population,*

$$\mathbb{P}\Big\{\Theta_n(m_0) \subseteq \widehat{\mathcal{I}}^{\mathrm{hyb}}_{n,1-\alpha} \mid \mathcal{F}_0\Big\} \geq 1-\alpha. \tag{33}$$

*The guarantee follows from the finite-sample variance bound and one-sided tail inequalities.*

**Certification mechanism.** The pilot event proves that the selected affine dual functions are correct at every supported-mean vector inside the pilot confidence polytope. Pilot coverage therefore transfers the representation to the unknown $m_0$. A failed certificate routes the evaluation confidence box through the complete structural program.

**Remark D.1** (One-wave use)**.** *With one wave, the projected branch uses the full data and a single confidence polytope. Independent stable waves supply the separation required for pilot-selected affine evaluation.*

## D.2 Width control and polyhedral calmness

The next result quantifies informativeness by decomposing projected width into a population identification term and a statistical movement term.

For a compact convex set $K_0 \subset \mathrm{ri}(\mathcal{M}_n)$, define the relative slope moduli

$$\begin{aligned} \Gamma_{n,c}(K_0) &:= \sup\{\|g\|_1 : m \in K_0,\ g \in \partial_{\mathsf{T}_n}\phi_{n,c}(m)\}, \\ \Gamma_n(K_0) &:= \max\{\Gamma_{n,\ell}(K_0), \Gamma_{n,-\ell}(K_0)\}. \end{aligned} \tag{34}$$

Because the endpoint values are finite polyhedral functions, these quantities are finite on compact subsets of the relative interior. They are directly related to the projected equality multipliers in Proposition 3.1.

Let

$$\widehat{\mathcal{I}}^{\mathrm{proj}}_n(r) := \left[\min_{\substack{A\nu\leq b \\ \|D\nu-\widehat{m}\|_\infty\leq r}} \ell^\top\nu,\ \max_{\substack{A\nu\leq b \\ \|D\nu-\widehat{m}\|_\infty\leq r}} \ell^\top\nu\right]. \tag{35}$$

**Theorem D.2** (Identification–sampling width decomposition)**.** *Let $B := \{m \in \mathcal{M}_n : \|m-\widehat{m}\|_\infty \leq r\}$ be nonempty and contained in a compact convex $K_0 \subset \mathrm{ri}(\mathcal{M}_n)$. Then*

$$\mathrm{diam}\{\widehat{\mathcal{I}}^{\mathrm{proj}}_n(r)\} \leq \sup_{m\in B}\omega_n(m) + 2r\Gamma_n(K_0). \tag{36}$$

*If additionally $m_0 \in B$, then*

$$\mathrm{diam}\{\widehat{\mathcal{I}}^{\mathrm{proj}}_n(r)\} \leq \omega_n(m_0) + 2r\{\Gamma_{n,\ell}(K_0) + \Gamma_{n,-\ell}(K_0)\}. \tag{37}$$

*Thus the first term is removable only by changing the identifying information, whereas the second decreases with the effective sample size.*

**Corollary D.1** (Projected upper rate)**.** *Take $r = r_2(\eta)$ with fixed $\eta \in (0,1)$, and suppose $s$, $\Gamma_n(K_0)$, and $\Gamma_{n,\ell}(K_0) + \Gamma_{n,-\ell}(K_0)$ are uniformly bounded. Assume that whenever $B = \{m \in \mathcal{M}_n : \|m-\widehat{m}\|_\infty \leq r\}$ is nonempty, it is contained in a fixed compact convex $K_0 \subset \mathrm{ri}(\mathcal{M}_n)$. Then, pathwise on $B \neq \varnothing$,*

$$\left[\mathrm{diam}\{\widehat{\mathcal{I}}^{\mathrm{proj}}_n(r)\} - \sup_{m\in B}\omega_n(m)\right]_+ \leq 2r\Gamma_n(K_0) = O\left(M\sqrt{\frac{1+\Delta_2}{N_2 p_{\min,2}}}\right). \tag{38}$$

*Moreover, on $m_0 \in B$, an event of probability at least $1-\eta$,*

$$\left[\mathrm{diam}\{\widehat{\mathcal{I}}^{\mathrm{proj}}_n(r)\} - \omega_n(m_0)\right]_+ \leq 2r\{\Gamma_{n,\ell}(K_0) + \Gamma_{n,-\ell}(K_0)\} = O\left(M\sqrt{\frac{1+\Delta_2}{N_2 p_{\min,2}}}\right). \tag{39}$$

*The rate is invariant to how many unsupported cells exist except through the structural sensitivity modulus: only target-relevant completion directions matter.*

### D.3 STABLE-FACE ASYMPTOTIC REFINEMENT

All asymptotic statements use a triangular array indexed by $t$, with $N_{2,t} = n_t R_{2,t} \to \infty$. The exposure dimensions $K$ and $s$ are fixed, $M$ is uniformly bounded, and the graph, response schedule, randomization law, exposure probabilities, dependency degree, dual slopes, and score variances may vary with $t$. We suppress the index. The displayed margin, variance, and dependency conditions hold uniformly along the array; either $n_t$, $R_{2,t}$, or both may diverge.

For a deterministic direction $a \in \mathbb{R}^s$, define the centered unit–wave score

$$\xi_{ir}(a) := \sum_{e \in \mathcal{S}_n} a_e \left\{ \frac{\mathbf{1}\{E_i^{(r)} = e\} Y_i^{(r),\mathrm{obs}}}{p_{ir}(e)} - m_i(e) \right\}, \tag{40}$$

for $r \in \mathcal{R}_2$, and let

$$\sigma_{N_2}^2(a) := \mathrm{Var}\left( N_2^{-1/2} \sum_{r \in \mathcal{R}_2} \sum_{i=1}^{n} \xi_{ir}(a) \mid \mathcal{F}_0 \right). \tag{41}$$

**Condition D.1** (Positive dual-face margin and full-dimensional supported moments)**.** *For each objective $c \in \{\ell, -\ell\}$, there exist a dual-feasible pair $(\rho_c, \lambda_c) \in \mathcal{D}_c$ and a constant $\delta_c > 0$ such that the pair is optimal for every*

$$m \in \mathcal{M}_n, \qquad \|m - m_0\|_\infty \le \delta_c.$$

*For the asymptotic refinement, the supported-moment domain is locally full dimensional, so $\mathsf{T}_n = \mathbb{R}^s$. Along the asymptotic sequence, the pilot error budget satisfies $\beta_{N_1} \to 0$ and $r_1(\beta_{N_1}) \to 0$.*

**Proposition D.1** (Consistency of the polyhedral certificate)**.** *Under Condition D.1, with probability tending to one, both certificates pass and every retained equality multiplier equals the corresponding population affine slope near $m_0$.*

**Diagnostic meaning.** Condition D.1 places the supported mean a positive distance inside a dual critical region; local full dimensionality identifies the equality-multiplier slope. The empirical certificate audits that geometry: a pilot polytope contained in one region activates affine evaluation, while a boundary crossing selects projection. The finite-sample theorem covers both branches. The Gaussian refinement below uses the stated local full-dimensionality condition to identify a unique slope.

**Assumption D.1** (Dependency-graph CLT)**.** *For the population affine directions $a \in \{\rho_-, \rho_+\}$ supplied by Condition D.1, $\sigma_{N_2}^2(a) \to \sigma^2(a) \in (0, \infty)$. For some fixed integer $k \ge 3$,*

$$N_2(1 + \Delta_2)^{k-1} \left\{ \frac{2M \|a\|_1}{p_{\min,2} \sqrt{N_2}\, \sigma_{N_2}(a)} \right\}^k \longrightarrow 0. \tag{42}$$

**Theorem D.3** (Endpoint asymptotic normality)**.** *Under Condition D.1, Assumption D.1, and the finite-sample assumptions, the affine endpoint estimators in Algorithm 1, extended by any fixed values on certificate-failure events, satisfy*

$$\frac{\sqrt{N_2}(\widetilde{\theta}^- - \theta_n^-)}{\sigma_{N_2}(\rho_-)} \rightsquigarrow N(0,1), \qquad \frac{\sqrt{N_2}(\widetilde{\theta}^+ - \theta_n^+)}{\sigma_{N_2}(\rho_+)} \rightsquigarrow N(0,1). \tag{43}$$

**Corollary D.2** (Gaussian confidence enlargement)**.** *With consistent standard-error estimators, the interval*

$$\left[ \widetilde{\theta}^- - z_{1-\alpha/2} \widehat{\sigma}_- / \sqrt{N_2},\ \widetilde{\theta}^+ + z_{1-\alpha/2} \widehat{\sigma}_+ / \sqrt{N_2} \right] \tag{44}$$

*contains $\Theta_n(m_0)$ with probability at least $1 - \alpha + o(1)$. This is an efficiency refinement; Theorem 5.1 is the finite-sample guarantee.*

### D.4 STRUCTURAL AND COMPUTATIONAL ROBUSTNESS

**Theorem D.4** (Slackened structural validity)**.** *Let $\varepsilon \in \mathbb{R}_+^{r_A}$, replace $b$ by $b + \varepsilon$, and write $\phi_{n,c}^{\varepsilon}(m)$ for the relaxed value. If the original and relaxed programs are feasible and $(\rho^*, \lambda^*) \in \mathcal{D}_c^*(m)$, then*

$$0 \le \phi_{n,c}(m) - \phi_{n,c}^{\varepsilon}(m) \le (\lambda^*)^\top \varepsilon. \tag{45}$$

*If $A\mu_0 \le b + \varepsilon$, the relaxed constraints define a valid completion set for the true mean vector.*

**Corollary D.3** (Distance and rule-graph uncertainty)**.** *Suppose the true structural distance $d_0$ and the analyst's distance $\widetilde{d}$ satisfy $d_0(e,e') \le \widetilde{d}(e,e') + \delta_{ee'}$ on every imposed finite link, and the mean response obeys $|\mu_{0e} - \mu_{0e'}| \le L d_0(e,e')$. Then the analyst's distance rows are valid after adding slack $L\delta_{ee'}$ to their right-hand sides. The endpoint change is therefore bounded by the corresponding dual-weighted distance error through Theorem 5.4.*

**Proposition D.2** (Certified numerical optimization error)**.** *Suppose an implementation returns an interval $[\widehat{L}, \widehat{U}]$ that would equal an exact valid interval $[L,U]$ except that the lower and upper optimization problems have certified absolute value gaps at most $\tau_-$ and $\tau_+$, respectively. Then reporting*

$$[\widehat{L} - \tau_-, \widehat{U} + \tau_+] \tag{46}$$

*contains $[L,U]$. The reported interval includes these certified value gaps as deterministic endpoint enlargements.*

### D.5 An explicit minimax lower bound

The lower bound uses an explicit stochastic block submodel of the mean-response framework, with fully specified statistically close models, dependence degree, and Kullback–Leibler distance.

**Theorem D.5** (Identification plus network-statistical lower bound)**.** *Fix $\alpha \in (0, 1/2)$, an exposure probability $p \in (0,1]$, a block size $q = 1 + \Delta$ dividing $n$, and $R \ge 1$. Let $N = nR$, and let $D_0 \in [0, M]$. Choose a constant $c > 0$ such that*

$$\kappa_c := 1 - 2\alpha - \sqrt{8}\, c > 0, \qquad \epsilon_N := c\sqrt{\frac{q}{Np}} \le \frac{1}{4}. \tag{47}$$

*There exists a pair of admissible block-dependent data-generating processes $P_+$ and $P_-$ with score dependency degree $\Delta$, supported-exposure probability $p$, and identified regions*

$$\Theta(P_\pm) = \left[ M\left(\frac{1}{2} \pm \epsilon_N\right), M\left(\frac{1}{2} \pm \epsilon_N\right) + D_0 \right], \tag{48}$$

*for which*

$$\mathrm{KL}(P_+, P_-) \le 16c^2, \qquad \mathrm{TV}(P_+, P_-) \le \sqrt{8}\, c. \tag{49}$$

*Consequently, every random interval $\widehat{I}_N$ satisfying*

$$\inf_{P \in \{P_+, P_-\}} \mathbb{P}_P\{\Theta(P) \subseteq \widehat{I}_N\} \ge 1 - \alpha \tag{50}$$

*obeys*

$$\sup_{P \in \{P_+, P_-\}} \mathbb{E}_P \operatorname{diam}(\widehat{I}_N) \ge (1-\alpha) D_0 + 2\kappa_c c M \sqrt{\frac{1+\Delta}{Np}}. \tag{51}$$

*If the procedure is constrained to have diameter at least $D_0$ almost surely, the first term improves from $(1-\alpha)D_0$ to $D_0$.*

**Corollary D.4** (Minimax order of honest interval length)**.** *Fix $\alpha \in (0, 1/2)$. For a model class $\mathcal{P}_N$, define the honest-length risk*

$$\mathfrak{R}_{N,\alpha}(\mathcal{P}_N) := \inf_{\widehat{I}:\, \inf_{P \in \mathcal{P}_N} \mathbb{P}_P\{\Theta(P) \subseteq \widehat{I}\} \ge 1-\alpha} \sup_{P \in \mathcal{P}_N} \mathbb{E}_P \operatorname{diam}(\widehat{I}). \tag{52}$$

*Fix $p_\star \in (0, 1/2]$, an integer $\bar{\Delta} \ge 0$, and constants $D_0 \in [0, M]$, $1 \le \bar{\Gamma} < \infty$. Let $\mathcal{P}_N(D_0, p_\star, \bar{\Delta}, \bar{\Gamma})$ have fixed exposure dimensions $K, s$, common outcome bound $M$, and satisfy, uniformly over its elements,*

*(i)* $p_{\min} \ge p_\star$ *and score dependency degree at most* $\bar{\Delta}$*;*

*(ii)* $\sup_{m \in \mathcal{M}_n} \omega_n(m) \le D_0$*;*

*(iii) for* $c \in \{\ell, -\ell\}$, $|\phi_{n,c}(m) - \phi_{n,c}(m')| \le \bar{\Gamma} \|m - m'\|_\infty$ *for all* $m, m' \in \mathcal{M}_n$*;*

*(iv) the class contains the pair of Theorem 5.3 with* $c_\alpha = (1-2\alpha)/(2\sqrt{8})$, $p = p_\star$, $\Delta = \bar{\Delta}$*, and diameter* $D_0$*.*

Table 2: Executed theorem–experiment map. Each entry is generated by an archived script and reported in a machine-readable summary.

| Theoretical object | Executed design | Archived primary output |
|---|---|---|
| Persistent nonidentification | 1–256 repeated waves, fixed supported cells | Width plateau and full-set coverage |
| Positivity phase | $N = 100, \dots, 30{,}000$, $p = 0, \dots, 0.30$ | Width/coverage surface and realized target counts |
| Target-specific identification | Three contrasts on one completion polytope | Exact sharp endpoints |
| Plug-in bias and projection | 50–2500 observations per supported cell | Signed bias, width, Wilson interval for coverage |
| Common-dual certificate | 30 distance-to-kink/radius cells, 600 trials each | Pass/fallback and false-certification rates |
| Dual bottleneck value | 150 independently weighted target contrasts | Controlled star-polytope algebra check |
| Metric and slack robustness | Euclidean/mechanism metrics; slack 0–0.15 | Population containment, coverage, threshold recovery |
| Dependency rate | $p = 0.02, \dots, 0.50$, $1 + \Delta = 1, \dots, 80$ | Exact and empirical variances, log–log slopes |
| End-to-end network chain | 100–800 independent six-node rings; 1,000 repetitions | Primitive/full-set coverage, width, Monte Carlo intervals |
| Real-topology audit | 34-node Zachary graph; 5000 assignments | Support atlas under a hub safety rule |
| Field-data stress test | Six leave-one-exposure-cell-out fits on 142 households | Interval containment, width, point error, dual binders |
| Paired-cell holdout | All 15 pairs of the six observed exposure cells | 30 held-out means, calibration sensitivity, LP checks |

*For all sufficiently large $N$ along the subsequence on which $1 + \bar{\Delta}$ divides $n$,*

$$\mathfrak{R}_{N,\alpha}\{\mathcal{P}_N(D_0, p_\star, \bar{\Delta}, \bar{\Gamma})\} \asymp_{\alpha,K,s,\bar{\Gamma}} D_0 + M\sqrt{\frac{1+\bar{\Delta}}{Np_\star}}, \tag{53}$$

*The coefficient on $D_0$ in the unrestricted lower bound is $1 - \alpha$; it becomes one under the pathwise-width condition in Theorem 5.3. Thus the scale*

$$\textit{unavoidable interval scale} = \textit{identification diameter} + \textit{network–overlap sampling radius} \tag{54}$$

*is minimax sharp in order.*

## E Experimental Design and Additional Evidence

The Monte Carlo suite uses Python 3.12.14, NumPy 2.5.3, pandas 3.0.5, SciPy 1.18.1, Matplotlib 3.11.2, and NetworkX 3.6.1. The dependency construction has $n = 2400$ units; its block score is duplicated $1 + \Delta$ times, yielding the exact reference variance. Coverage summaries include Wilson Monte Carlo intervals. The public Zachary topology is supplied by NetworkX under BSD-3-Clause. The processed insurance files come from commit `827632d0` of the BSD-3-Clause `spilled_t` replication repository; their archived SHA-256 values and the distinction between 343 non-self links and 142 stored diagonal entries appear in the manifest.

### E.1 Paired exposure-cell holdouts on observed insurance data

We extended the single-cell analysis by removing every unordered pair of exposure cells. The 15 fits yield 30 held-out cell means from the same 142 households. Own information assignment $z \in \{0, 1\}$ and treated-friend count $f \in \{0, 1, 2+\}$ define the six cells. Each fit uses outcomes from the four retained cells for calibration and regression. The held-out outcomes enter only the final comparison.

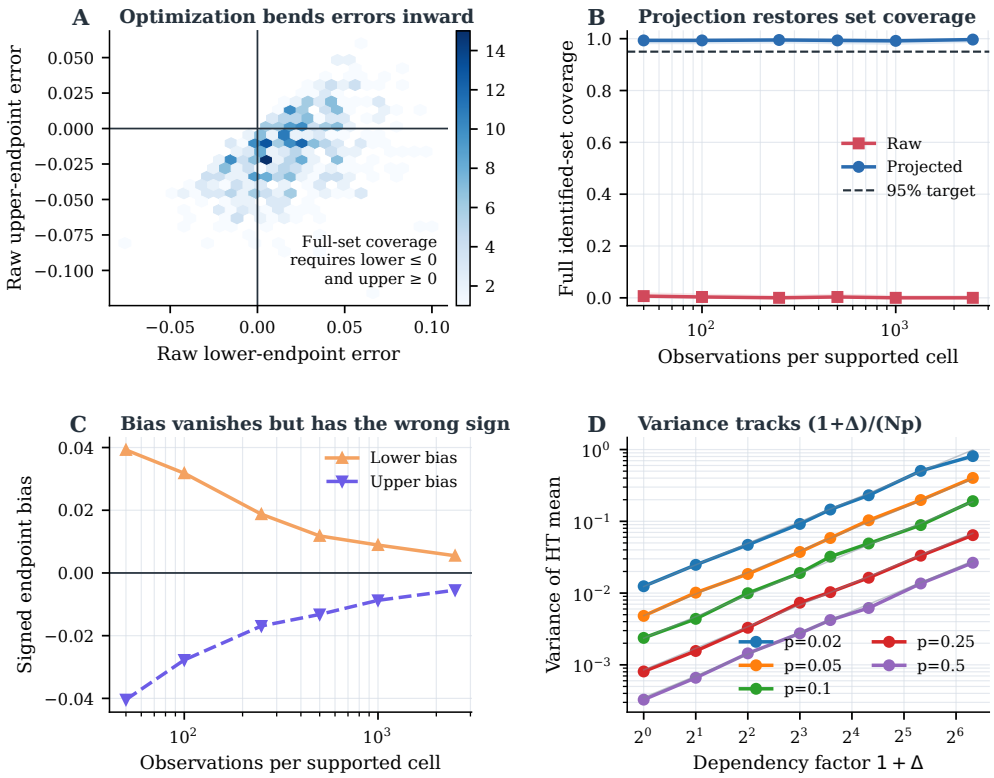


Figure 5: **Inference and dependence.** Raw optimization bends primitive estimation errors inward (A), while projection restores full-set coverage (B). The lower and upper signed biases approach zero from opposite directions (C). Empirical variance follows the exact block-construction law $(1+\Delta)/(Np)$ (D).

For each distance $d$, the retained empirical means $\bar{Y}_e$ determine

$$\widehat{L}_{-\mathcal{H}} = \max_{e,e' \notin \mathcal{H},\ e \neq e'} \frac{|\bar{Y}_e - \bar{Y}_{e'}|}{d(e,e')}, \qquad |\mathcal{H}| = 2.$$

We use both Manhattan and Euclidean distances on the existing exposure grid. The fixed multipliers $\zeta \in \{1, 1.5, 2\}$ give completion constraints with slope $\zeta \widehat{L}_{-\mathcal{H}}$. Raw intervals fix the four retained means. Wilson-based intervals project their binomial reference intervals, each computed with error level $0.05/4$. Table 3 summarizes every fixed multiplier; the supplementary files report every pair. Metrics and multipliers were fixed before these fits. Containment here measures agreement with masked observed adoption means on a dependent, reused dataset.

At the fitted slope, Manhattan projection contained 28 means with average width $0.845$. Both misses involved $(z, f) = (1, 2+)$, whose nine households had mean adoption $7/9$. Removing this cell together with $(0, 1)$ reduced the fitted slope to $0.0367$; the projected upper endpoint was $0.5426$, a gap of $0.2352$. Removing it together with $(0, 0)$ gave upper endpoint $0.7573$, a gap of $0.0205$. Doubling the slope increased containment to 29 and average width to $0.949$, with one remaining gap of $0.1617$. Fewer retained cells can therefore reduce the empirical slope enough to narrow an interval while worsening containment. Logistic regression with the archived ridge penalty $10^{-4}$ and clipped least squares gave mean absolute errors $0.242$ and $0.251$, respectively.

The protocol was fixed before this extension was run. All 21 single- and paired-cell fits passed the held-out-outcome perturbation check. The 864 endpoint LP solves agreed with the closed-form bounds to $1.2 \times 10^{-16}$, and the six single-cell fits reproduced the archived endpoints to the same tolerance. The supplementary directory `cai_pair_holdout/` contains the protocol, every result, and the execution audit; `code/run_cai_pair_holdout.py` reruns the analysis using the bundled inputs.

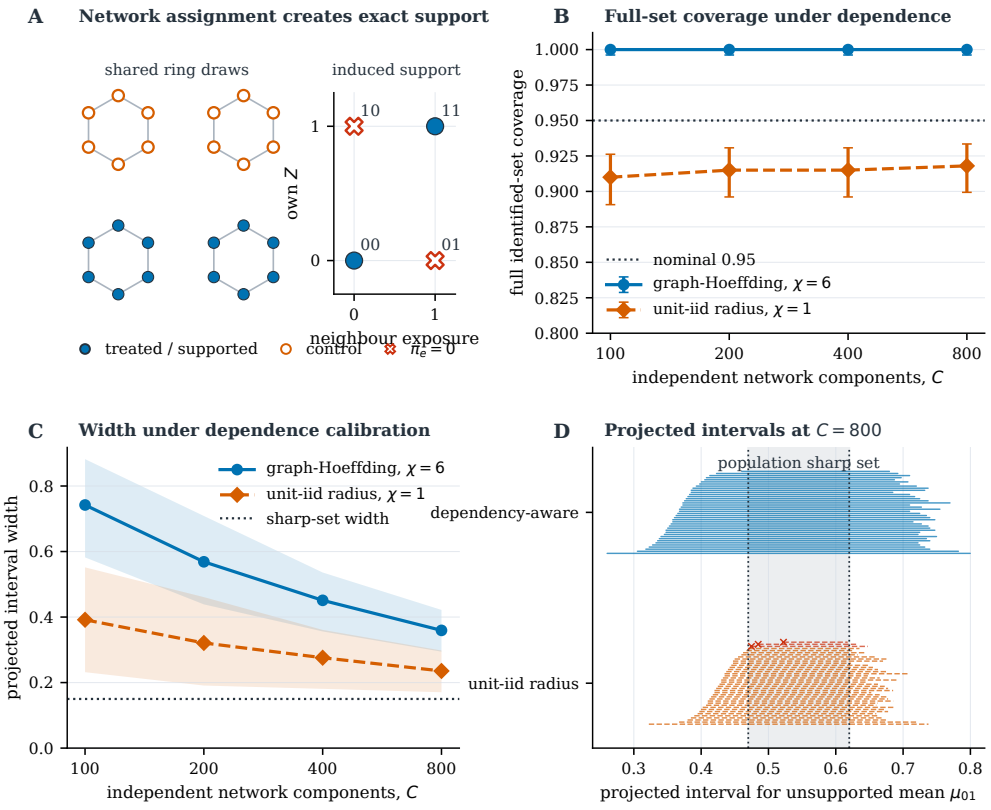


Figure 6: **End-to-end network randomization and full-set inference.** A valid graph-Hoeffding primitive region followed by deterministic LP projection covers the complete population sharp set, whereas the unit-iid radius is narrower but undercovers both endpoints jointly. Panel A shows one executed ring assignment and its exact support; B reports full-set coverage with 95% Wilson Monte Carlo intervals; C reports projected widths; D shows 40 deterministic repetition-level intervals at the largest network, with full-set misses marked in red.

Table 3: Paired-cell holdout results. Counts are out of 30 held-out cell evaluations, equally weighted. Widths average over those evaluations. Wilson-based projection uses retained-cell reference intervals.

| | | Raw completion | | Wilson projection | |
|---|---|---|---|---|---|
| Distance | $\zeta$ | Contained | Width | Contained | Width |
| Manhattan | 1.0 | 18 | 0.507 | 28 | 0.845 |
| | 1.5 | 22 | 0.732 | 29 | 0.924 |
| | 2.0 | 24 | 0.828 | 29 | 0.949 |
| Euclidean | 1.0 | 18 | 0.474 | 28 | 0.845 |
| | 1.5 | 22 | 0.727 | 29 | 0.924 |
| | 2.0 | 24 | 0.828 | 29 | 0.949 |
| Unrestricted | – | 30 | 1.000 | 30 | 1.000 |

## F Machine-Checked Deterministic Core

The bundled `LeanProofs` project pins Lean 4.33.0 and mathlib v4.33.0. The file `LeanProofs/StructuralSupport.lean` contains 13 manuscript-facing declarations: the midpoint lower bound and its positive-diameter corollary; compact-convex linear-image interval geometry and endpoint containment; the point-identification equivalence; four laws for the bidi-

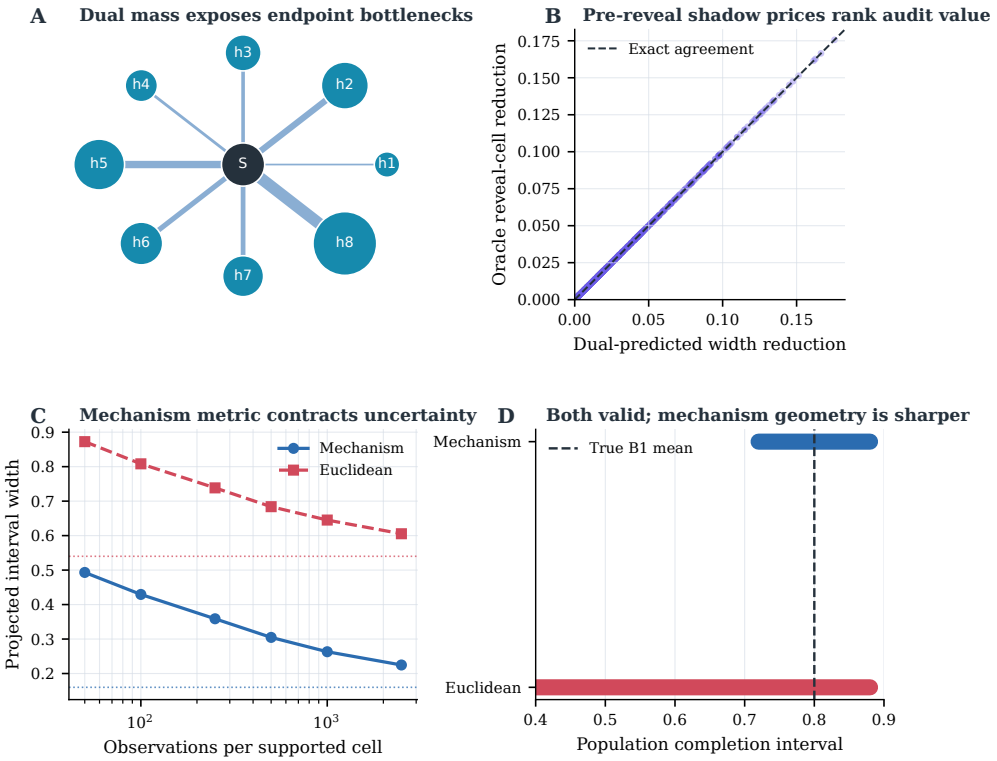


Figure 7: **Bottlenecks and structural geometry.** Pre-revelation base-LP edge duals rank reveal value in a controlled star-polytope check (A–B). With both Lipschitz constants population-calibrated on supported pairs, the mechanism distance yields a sharper valid interval than Euclidean distance at the held-out boundary (C–D).

rectional transformation pseudometric; conversion of paired directional inequalities to an absolute-difference restriction; primitive-region and true-interval containment under projection; endpoint and solver-gap enlargement; the deterministic dual-slack sandwich; and the duplicated-block variance identity. Each declaration is audited with `#print axioms`.

The verification commands `lake env lean --error=warning`, `lake build --wfail`, and the independent placeholder checker all exit successfully. A source-wide search finds no `sorry`, `admit`, user-declared `axiom`, or `unsafe`; the axiom reports contain only the standard Lean/mathlib foundations `propext`, `Classical.choice`, and `Quot.sound`. The project README maps every declaration to the corresponding manuscript result and records the exact toolchain commit.

The formalized layer is finite-dimensional and deterministic. LP strong duality and complementary slackness, Horvitz–Thompson randomization arguments, Cantelli calibration, dependency-graph central limit theory, stable-face selection, and the KL/total-variation minimax reduction use the conventional proofs below.

## G Notation and Implementation Objects

Table 4: Notation table.

| Symbol | Meaning |
|---|---|
| $G_n = (V_n, E_n)$ | Fixed network for a population of $n$ units |
| $K, s, r_A$ | Numbers of exposure labels, supported coordinates, and structural rows |

| Symbol | Meaning |
|---|---|
| $T^{(r)} \in \mathcal{T}_n$ | Feasible global assignment in randomization wave $r$ |
| $E_i^{(r)} = d_i(T^{(r)}, G_n, X_{1:n})$ | Local exposure label for unit $i$ in wave $r$ |
| $m_i(e)$ | Conditional mean response of unit $i$ under exposure $e$ |
| $\mu_{0e} = n^{-1}\sum_i m_i(e)$ | Finite-population mean response under exposure $e$ |
| $\ell$ | Pre-specified target contrast vector |
| $\theta_0 = \ell^\top \mu_0$ | Target network causal contrast |
| $p_{ir}(e)$ | Known exposure probability under the assignment design |
| $\mathcal{S}_n, \mathcal{U}_n$ | Uniformly supported and structurally absent exposure labels |
| $c_\to(e, e')$ | Directed shortest-path transformation cost |
| $d(e, e')$ | Two-way extended distance used by symmetric restrictions |
| $A\nu \le b$ | Outcome bounds, structural links, and auxiliary restrictions |
| $D\nu = m$ | Equalities fixing the design-identifiable supported means |
| $\mathcal{C}_n(m)$ | Structural completion polytope |
| $\mathcal{M}_n$ | Supported-mean vectors with at least one completion |
| $\Theta_n(m)$ | Population partial-identification region |
| $\phi_{n,c}(m)$ | Minimum value of objective $c$ over $\mathcal{C}_n(m)$ |
| $\theta_n^-, \theta_n^+$ | Lower and upper target endpoints |
| $\omega_n(m)$ | Population identified-set diameter |
| $\rho, \lambda$ | Dual multipliers for supported equalities and structural inequalities |
| $R_w, N_w = nR_w$ | Number of waves and unit–wave observations in block $w$ |
| $\Delta_w$ | Maximum degree of a dependency graph for unit–wave scores |
| $r_w(\eta)$ | Coordinate radius of the primitive confidence polytope |
| $\widehat{\mathcal{B}}_w(\eta)$ | Feasible supported-mean confidence polytope |
| $\mathcal{D}_c^*(B)$ | Dual pairs optimal for every supported mean in polytope $B$ |
| $\Gamma_{n,c}(K_0)$ | Relative polyhedral slope modulus on $K_0$ |
| $\mathsf{T}_n = \mathrm{lin}(\mathcal{M}_n - \mathcal{M}_n)$ | Tangent space of the supported-moment domain |
| $\sigma \in \{-, +\}$ | Endpoint sign; distinct from the supported-coordinate count $s$ |
| $\bar{\eta}_e, \eta_\ell$ | Mapping-error bounds for a population mean and the target contrast |

## H Auxiliary Lemmas

**Lemma H.1** (Compactness, interval image, and attainment)**.** *For every $m \in \mathcal{M}_n$, $\mathcal{C}_n(m)$ is a nonempty compact convex polytope. Every endpoint program is attained, and $\{\ell^\top \nu : \nu \in \mathcal{C}_n(m)\}$ is the closed interval $[\phi_{n,\ell}(m), -\phi_{n,-\ell}(m)]$.*

*Proof.* The box rows in $A\nu \le b$ imply $\mathcal{C}_n(m) \subseteq [0, M]^K$. The equality and weak-inequality constraints are closed, so $\mathcal{C}_n(m)$ is a closed subset of a compact cube and is therefore compact. It is convex because all constraints are affine, and nonemptiness is the definition of $m \in \mathcal{M}_n$. The extreme-value theorem gives endpoint attainment. The continuous linear image of a nonempty compact convex set in $\mathbb{R}$ is a nonempty compact convex set, hence the interval between its minimum and maximum. □

**Lemma H.2** (Convexity, concavity, and polyhedrality)**.** *For every objective $c$, $m \mapsto \phi_{n,c}(m)$ is a finite polyhedral convex function on $\mathcal{M}_n$. Consequently, $m \mapsto -\phi_{n,-\ell}(m)$ is concave.*

*Proof.* Take $m_1, m_2 \in \mathcal{M}_n$, $t \in [0, 1]$, and minimizers $\nu_j \in \mathcal{C}_n(m_j)$. Then $t\nu_1 + (1-t)\nu_2 \in \mathcal{C}_n(tm_1 + (1-t)m_2)$, so

$$\phi_{n,c}\{tm_1 + (1-t)m_2\} \le t\phi_{n,c}(m_1) + (1-t)\phi_{n,c}(m_2).$$

Thus the value is convex. By finite-dimensional LP duality, proved in Lemma H.3, it is the pointwise maximum of affine dual objectives. On any region where the value is finite, a parametric linear

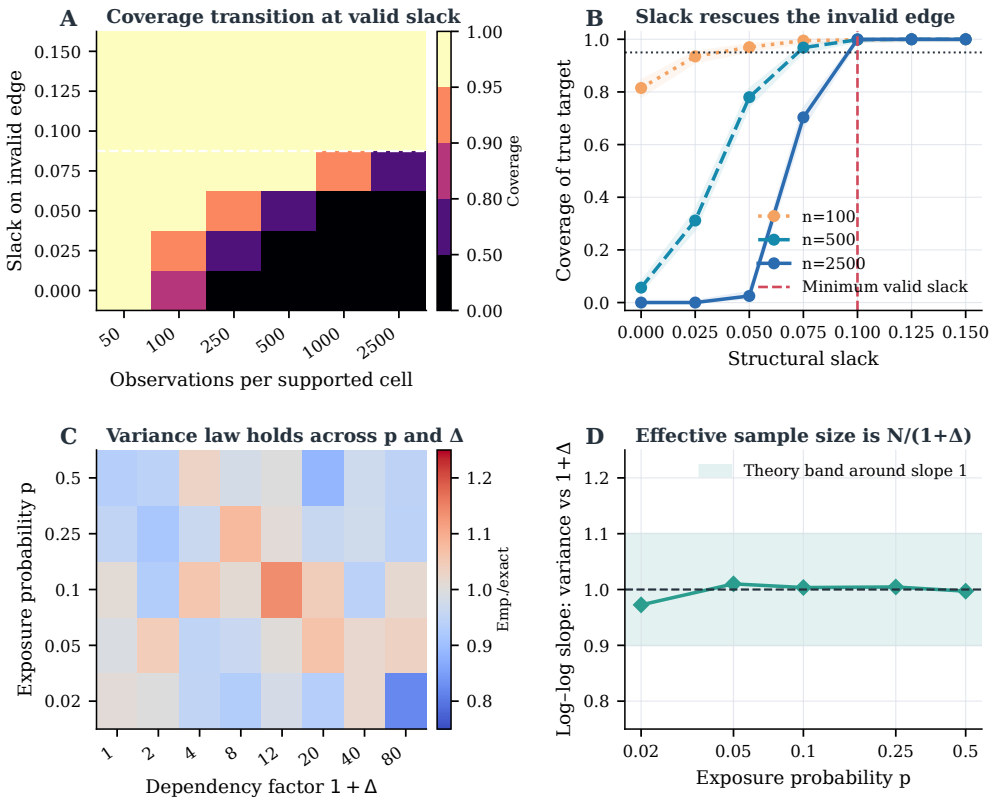


Figure 8: **Slack threshold and dependence scaling.** A restriction with known violation yields a coverage transition at its true slack (A–B). Across exposure probability and dependency degree, empirical variance tracks the exact variance (C), and the fitted log–log dependence slopes concentrate around one (D).

program has finitely many relevant bases, so the function is polyhedral. Applying the same argument to objective $-\ell$ proves concavity of the upper endpoint. □

**Lemma H.3** (Weak and strong LP duality)**.** *For every primal-feasible $\nu$ and dual-feasible $(\rho, \lambda) \in \mathcal{D}_c$,*

$$\rho^\top m - \lambda^\top b \le c^\top \nu. \tag{55}$$

*If $m \in \mathcal{M}_n$, both sides have optimizers and their optimal values agree.*

*Proof.* Dual feasibility gives $c = D^\top \rho - A^\top \lambda$. Therefore

$$c^\top \nu = \rho^\top D\nu - \lambda^\top A\nu = \rho^\top m - \lambda^\top A\nu \ge \rho^\top m - \lambda^\top b,$$

because $A\nu \le b$ and $\lambda \ge 0$. This proves weak duality. Lemma H.1 makes the primal feasible with finite attained value. Finite-dimensional LP strong duality then gives dual attainment and equality of optimal values. □

**Lemma H.4** (Polyhedral active-face derivative)**.** *Let $f$ be a proper polyhedral convex function on a polyhedron $M$, and suppose that near $x \in \mathrm{ri}(M)$,*

$$f(y) = \max_{j \in J} \{a_j^\top y + b_j\}$$

*for a finite set $J$. If $J^*(x) = \{j : a_j^\top x + b_j = f(x)\}$, then relative to $\mathrm{aff}(M)$,*

$$\partial f(x) = \mathrm{conv}\{\Pi_{\mathrm{lin}(M-M)} a_j : j \in J^*(x)\}, \qquad f'(x;h) = \max_{j \in J^*(x)} a_j^\top h. \tag{56}$$

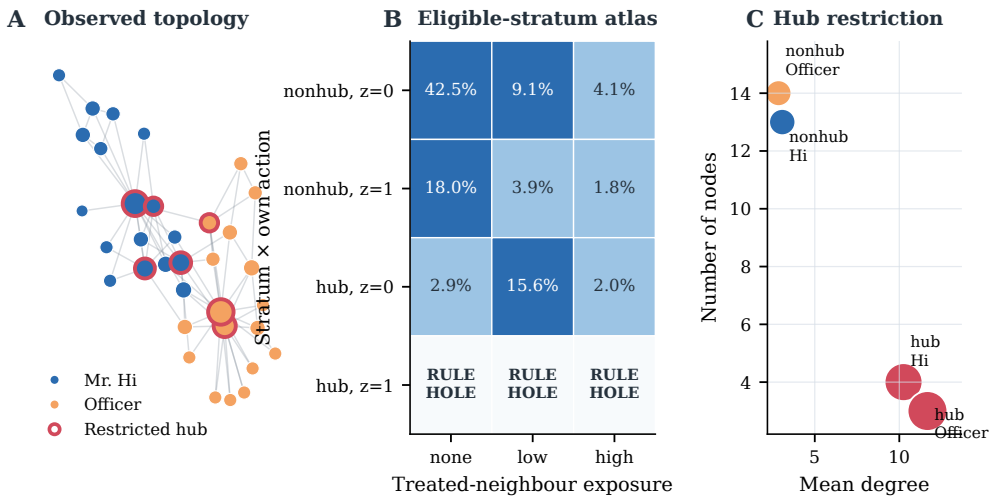


Figure 9: **Semi-synthetic public-network audit.** The observed Zachary topology and club labels (A) are combined with a pre-specified hub treatment restriction, producing three analytically implied holes in the eligible-stratum exposure atlas (B). Restricted hubs occupy the high-degree tail in both observed groups (C). Assignments, exposure labels, and outcomes are simulated; the topology is observed.

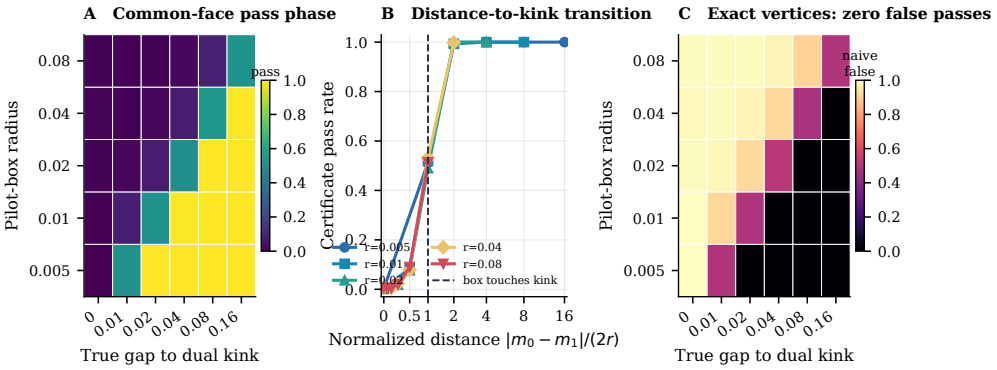


Figure 10: **Exact common-face certification.** Pass probability changes with pilot radius and distance to the dual kink (A–B). The exact vertex rule produced zero false certifications under independent dense-grid validation; the heat map in C records center-only misclassification.

*Proof.* Every active affine function globally minorizes $f$ and equals it at $x$, so each active projected slope is a relative subgradient, as is every convex combination. Conversely, suppose a relative subgradient $g$ lies outside their convex hull. A separating hyperplane in the tangent space supplies a direction $h$ with $g^\top h > \max_{j \in J^*(x)} a_j^\top h$. For sufficiently small positive $t$, inactive affine pieces remain below the active maximum, hence

$$\frac{f(x+th) - f(x)}{t} = \max_{j \in J^*(x)} a_j^\top h.$$

This contradicts the subgradient inequality $f(x+th) - f(x) \geq t g^\top h$. The same local representation yields the directional derivative formula. □

**Lemma H.5** (Primitive confidence polytope)**.** *Under Theorem 4.1, for every* $\eta \in (0, 1)$,

$$\mathbb{P}\{m_0 \in \widehat{\mathcal{B}}_w(\eta) \mid \mathcal{F}_0\} \geq 1 - \eta.$$

*Proof.* For a coordinate vector $a = e_j$, Theorem 4.1 gives

$$\mathrm{Var}(\widehat{m}_j^{(w)} - m_{0j} \mid \mathcal{F}_0) \le \frac{M^2(1+\Delta_w)}{N_w p_{\min,w}}.$$

By Chebyshev's inequality and the definition of $r_w(\eta)$,

$$\mathbb{P}\{|\widehat{m}_j^{(w)} - m_{0j}| > r_w(\eta) \mid \mathcal{F}_0\} \le \frac{\eta}{s}.$$

A union bound over the $s$ supported coordinates proves $\left\|\widehat{m}^{(w)} - m_0\right\|_\infty \le r_w(\eta)$ with probability at least $1-\eta$. Assumption A.3 gives $m_0 \in \mathcal{M}_n$, so the event is exactly $m_0 \in \widehat{\mathcal{B}}_w(\eta)$. □

**Lemma H.6** (One-sided Cantelli inequality). *If $\mathbb{E}X = 0$, $\mathrm{Var}(X) \le V < \infty$, and $t > 0$, then*

$$\mathbb{P}(X > t) \le \frac{V}{V + t^2}. \tag{57}$$

*For $\gamma \in (0,1)$, taking $t = \sqrt{V(1-\gamma)/\gamma}$ gives $\mathbb{P}(X > t) \le \gamma$ when $V > 0$. If $V = 0$, then $X = 0$ almost surely, so the same strict-tail bound holds with $t = 0$.*

*Proof.* For any $a \ge 0$, the event $X \ge t$ implies $(X+a)^2 \ge (t+a)^2$. Markov's inequality gives

$$\mathbb{P}(X \ge t) \le \frac{\mathbb{E}(X+a)^2}{(t+a)^2} \le \frac{V + a^2}{(t+a)^2}.$$

When $V > 0$, choose $a = V/t$ and simplify to obtain $V/(V+t^2)$. The case $V = 0$ is immediate. □

**Lemma H.7** (Relative Lipschitz property of endpoint values). *Let $K_0 \subset \mathrm{ri}(\mathcal{M}_n)$ be compact and convex. For every $m, m' \in K_0$,*

$$|\phi_{n,c}(m) - \phi_{n,c}(m')| \le \Gamma_{n,c}(K_0)\, \|m - m'\|_\infty . \tag{58}$$

*Proof.* A finite polyhedral convex function is piecewise affine on every line segment in its relative interior. Parameterize the segment by $m_t = m + t(m' - m)$, $t \in [0,1]$. Outside finitely many breakpoints, the derivative exists and equals $g_t^\top (m' - m)$ for some relative subgradient $g_t \in \partial_{\mathsf{T}_n} \phi_{n,c}(m_t)$. Hence

$$|\phi_{n,c}(m') - \phi_{n,c}(m)| \le \int_0^1 \|g_t\|_1 \, \|m' - m\|_\infty \, dt \le \Gamma_{n,c}(K_0)\, \|m' - m\|_\infty .$$

The finitely many breakpoints do not affect the integral. □

**Lemma H.8** (Bounded dependency-graph normal approximation). *For each $N$, let $X_{N1}, \dots, X_{NN}$ be centered random variables with a dependency graph of maximum degree $\Delta_N$. Let $S_N = \sum_j X_{Nj}$, $v_N^2 = \mathrm{Var}(S_N) > 0$, and $|X_{Nj}| \le B_N$ almost surely. If, for some fixed integer $k \ge 3$,*

$$N(1+\Delta_N)^{k-1}(B_N/v_N)^k \longrightarrow 0, \tag{59}$$

*then $S_N/v_N \rightsquigarrow N(0,1)$.*

*Proof.* This is the bounded dependency-graph theorem of Janson (1988), written with closed dependency neighborhoods and hence $1 + \Delta_N$. In Janson's notation, the number of summands is $N$, the uniform bound is $B_N$, the standard deviation is $v_N$, and the size of a closed dependency neighborhood is at most $1 + \Delta_N$. Condition (59) is exactly the required small-neighborhood condition. □

**Lemma H.9** (Two-point interval-length inequality). *Let $P_+, P_-$ be two laws with identified intervals*

$$\Theta(P_\pm) = [a \pm \delta, a \pm \delta + D], \qquad D, \delta \ge 0.$$

*If a random interval $\widehat{I}$ contains each corresponding identified interval with probability at least $1-\alpha$, then*

$$\mathbb{E}_{P_+} \mathrm{diam}(\widehat{I}) \ge (1-\alpha)D + 2\delta\{1 - 2\alpha - \mathrm{TV}(P_+, P_-)\}_+. \tag{60}$$

*Proof.* Let $A_+$ and $A_-$ be the events that $\widehat{I}$ contains $\Theta(P_+)$ and $\Theta(P_-)$, respectively. Honesty and total variation imply

$$P_+(A_+ \cap A_-) \geq P_+(A_+) + P_+(A_-) - 1 \geq 1 - 2\alpha - \mathrm{TV}(P_+, P_-).$$

On $A_+$, the diameter is at least $D$; on $A_+ \cap A_-$, it must span an additional distance $2\delta$. Thus

$$\mathrm{diam}(\widehat{I}) \geq D\mathbf{1}(A_+) + 2\delta\mathbf{1}(A_+ \cap A_-).$$

Taking $P_+$-expectations gives the claim. □

**Lemma H.10** (Bernoulli divergence bound)**.** *For* $0 \leq \epsilon \leq 1/4$,

$$\mathrm{kl}\{\mathrm{Bern}(1/2+\epsilon), \mathrm{Bern}(1/2-\epsilon)\} \leq 16\epsilon^2. \tag{61}$$

*Proof.* Direct calculation gives

$$\mathrm{kl} = 2\epsilon \log \frac{1+2\epsilon}{1-2\epsilon}.$$

For $x \in [0, 1/2]$, $\log\{(1+x)/(1-x)\} \leq 2x/(1-x) \leq 4x$. Taking $x = 2\epsilon$ yields $\mathrm{kl} \leq 16\epsilon^2$. □

# I Proofs of Main Results

## I.1 Proof of Theorem 3.1

*Proof.* By Lemma H.1, choose endpoint completions $\nu^-$ and $\nu^+$ in $\mathcal{C}_n(m_0)$ such that $\ell^\top \nu^- = \theta_n^-$ and $\ell^\top \nu^+ = \theta_n^+$. Construct two mean-response schedules by setting

$$m_i^\sigma(e) = \nu_e^\sigma, \qquad i \leq n,\ e \in \mathcal{E}, \quad \sigma \in \{-, +\}.$$

The box rows guarantee boundedness, and $A\nu^\sigma \leq b$ guarantees all stated mean restrictions. For every supported exposure $e \in \mathcal{S}_n$, $\nu_e^- = \nu_e^+ = m_{0e}$ because $D\nu^\sigma = m_0$. For every unsupported exposure $e \in \mathcal{U}_n$, the current design assigns probability zero. Let the observed response be degenerate at $m_i^\sigma(E_i^{(r)})$. Under either schedule, every realized supported response is therefore the same deterministic value, while unsupported responses are never observed. The assignment laws are also identical. Hence the joint observed-data law is identical under $P^-$ and $P^+$ for every $R$.

Let $Q$ denote this common law and set $a = \theta_n^-$, $b = \theta_n^+$. For any random variable $X$,

$$(X-a)^2 + (X-b)^2 = 2\left(X - \frac{a+b}{2}\right)^2 + \frac{(b-a)^2}{2}.$$

Taking $Q$-expectations and using that the maximum of two nonnegative numbers is at least their average gives

$$\max\{\mathbb{E}_Q(X-a)^2, \mathbb{E}_Q(X-b)^2\} \geq \frac{(b-a)^2}{4}.$$

Apply this inequality to $X = \widehat{\theta}_R$ and note that $b - a = \omega_n(m_0)$. □

## I.2 Proof of Theorem 3.2

*Proof.* Let $\vartheta \in \Theta_n(m_0)$. By definition, there exists $\nu \in \mathcal{C}_n(m_0)$ with $\ell^\top \nu = \vartheta$. Set $\widetilde{m}_i(e) = \nu_e$ for all $i, e$, and take a degenerate bounded response distribution at this mean. The box rows imply $\widetilde{m}_i(e) \in [0, M]$, the schedule depends on a global assignment only through the exposure label, and its exposure mean is $n^{-1}\sum_i \widetilde{m}_i(e) = \nu_e$. Thus it has supported mean vector $m_0$, satisfies $A\nu \leq b$, and has target $\vartheta$. This proves (i).

Conversely, every admissible schedule has an exposure-mean vector $\mu$ satisfying $D\mu = m_0$ and $A\mu \leq b$, so $\mu \in \mathcal{C}_n(m_0)$ and $\ell^\top \mu \in \Theta_n(m_0)$. This proves (ii).

For (iii), apply the same constant-schedule construction separately to $\nu$ and $\nu'$. Supported coordinates agree because both satisfy $D\nu = D\nu' = m_0$; unsupported coordinates are never realized under (4). The resulting observed-data laws are therefore identical, while the targets equal $\ell^\top \nu$ and $\ell^\top \nu'$. □

### I.3 Proof of Proposition 3.1

*Proof.* Primal compactness and attainment follow from Lemma H.1; dual attainment and equality follow from Lemma H.3. By Lemma H.2, the value is polyhedral and, locally on the relative interior of $\mathcal{M}_n$, is the maximum of finitely many affine dual objectives whose slopes are equality multipliers of dual-optimal bases. Lemma H.4 therefore yields (15) and (16), with projection because only relative perturbations in $\mathsf{T}_n$ are feasible.

Jointly in $(m, b)$, the dual affine objective has slope $(\rho, -\lambda)$. Applying the same active-face argument yields (17). Along the relaxation direction $(h, k) = (0, e_j)$, this derivative is $\max_{(\rho,\lambda)\in\mathcal{D}_c^*(m)}(-\lambda_j)$, which equals $-\min_{(\rho,\lambda)\in\mathcal{D}_c^*(m)} \lambda_j$. This proves (18) and shows why a single multiplier is insufficient under dual degeneracy. □

### I.4 Proof of Corollaries 3.1 and B.2

*Proof of Corollary 3.1.* By Lemma H.1, the target image is the interval $[\phi_{n,\ell}(m), -\phi_{n,-\ell}(m)]$. It is a singleton if and only if its endpoints agree, which is equivalent to $\phi_{n,\ell}(m) + \phi_{n,-\ell}(m) = 0$. The image is a singleton if and only if the linear functional $\ell^\top \nu$ is constant on $\mathcal{C}_n(m)$, which is exactly condition (iii). □

*Proof of Corollary B.2.* If $\mathcal{S}_n = \mathcal{E}$, then $D\nu = m_0$ fixes every coordinate and structural validity makes the unique vector $\mu_0$ feasible. For unrestricted unsupported box coordinates, supported entries are fixed. Minimizing $\ell_e\nu_e$ over $[0, M]$ selects $M$ when $\ell_e < 0$ and 0 otherwise; maximizing selects $M$ when $\ell_e > 0$ and 0 otherwise. Summing the coordinatewise extrema proves (19)–(20). □

### I.5 Proof of Theorem 4.1

*Proof.* Fix $e \in \mathcal{S}_n$. By iterated expectation and (2),

$$\mathbb{E}\left[\frac{\mathbf{1}\{E_i^{(r)} = e\}Y_i^{(r),\mathrm{obs}}}{p_{ir}(e)}\middle|\mathcal{F}_0\right] = \frac{p_{ir}(e)m_i(e)}{p_{ir}(e)} = m_i(e).$$

Averaging over $i$ and $r$ gives $\mathbb{E}(\widehat{\mu}_e^{(w)} \mid \mathcal{F}_0) = \mu_{0e}$.

For $a \in \mathbb{R}^s$, define

$$Z_{ir}(a) := \sum_{e\in\mathcal{S}_n} a_e \frac{\mathbf{1}\{E_i^{(r)} = e\}Y_i^{(r),\mathrm{obs}}}{p_{ir}(e)}.$$

At most one exposure occurs for a fixed unit–wave pair. Therefore

$$\mathbb{E}\{Z_{ir}(a)^2 \mid \mathcal{F}_0\} \le M^2 \sum_{e\in\mathcal{S}_n} \frac{a_e^2}{p_{ir}(e)} \le \frac{M^2}{p_{\min,w}} \|a\|_2^2 =: C_a.$$

Thus $\mathrm{Var}(Z_{ir}(a) \mid \mathcal{F}_0) \le C_a$. Nonadjacent score vectors have zero covariance. For every dependency edge, Cauchy–Schwarz gives an absolute covariance at most $C_a$. The dependency graph has $N_w$ vertices and at most $N_w\Delta_w/2$ undirected edges, so

$$\mathrm{Var}\left\{\frac{1}{N_w}\sum_{r\in\mathcal{R}_w}\sum_{i=1}^{n} Z_{ir}(a)\}\middle|\mathcal{F}_0\right) \le \frac{N_wC_a + N_w\Delta_wC_a}{N_w^2}.$$

The average inside braces equals $a^\top \widehat{m}^{(w)}$, and subtracting its nonrandom conditional mean does not change variance. This proves (23). Retaining the unit–wave-specific second-moment maximum gives the sharper statement. □

### I.6 Proof of Proposition C.1

*Proof.* Lemma H.2 and Jensen's inequality give

$$\phi_{n,\ell}(m) = \phi_{n,\ell}(\mathbb{E}\widehat{m}) \le \mathbb{E}\phi_{n,\ell}(\widehat{m}).$$

Applying Jensen to the convex function $\phi_{n,-\ell}$ and multiplying by $-1$ yields

$$\mathbb{E}\{-\phi_{n,-\ell}(\widehat{m})\} \le -\phi_{n,-\ell}(\mathbb{E}\widehat{m}).$$

The integrands are finite because $\widehat{m} \in \mathcal{M}_n$ almost surely and every completion lies in the compact cube $[0, M]^K$. □

### I.7 Proof of Theorem 4.2

*Proof.* Condition (i) immediately implies (ii). Conversely, suppose the dual-feasible pair $(\bar{\rho}, \bar{\lambda})$ is optimal at every vertex of $B$. For an arbitrary $m \in B$, write $m = \sum_{j=1}^{J} \alpha_j v_j$, where $v_j \in \mathrm{vert}(B)$, $\alpha_j \ge 0$, and $\sum_j \alpha_j = 1$. Vertex optimality and affine linearity give

$$\bar{\rho}^\top m - \bar{\lambda}^\top b = \sum_j \alpha_j \phi_{n,c}(v_j).$$

Convexity of $\phi_{n,c}$ implies $\phi_{n,c}(m) \le \sum_j \alpha_j \phi_{n,c}(v_j)$. Weak duality gives the reverse inequality $\bar{\rho}^\top m - \bar{\lambda}^\top b \le \phi_{n,c}(m)$. Hence equality holds and the pair is optimal at $m$.

For each vertex, the equality $\rho^\top v - \lambda^\top b = \phi_{n,c}(v)$, together with dual feasibility, is linear in $(\rho, \lambda)$. Their intersection is therefore the feasibility system (30). Any failure of common optimality is witnessed by a vertex because the nonnegative duality gap is convex and piecewise linear. The worst-case search may require examining exponentially many vertices. □

### I.8 Proof of Theorem 5.1

*Proof.* Let

$$G_P := \{m_0 \in \widehat{\mathcal{B}}_1(\beta)\}.$$

Lemma H.5 gives $\mathbb{P}(G_P^c \mid \mathcal{F}_0) \le \beta$. On $G_P$, every pilot-certified common dual pair is optimal at $m_0$ by Theorem 4.2.

Condition on the pilot data and $G_P$. If both endpoint objectives are certified, the selected dual pairs are fixed relative to the evaluation block. Strong duality at $m_0$ and Theorem 4.1 give

$$\mathbb{E}(\widetilde{\theta}^- \mid \text{pilot}, \mathcal{F}_0) = \theta_n^-, \qquad \mathbb{E}(\widetilde{\theta}^+ \mid \text{pilot}, \mathcal{F}_0) = \theta_n^+,$$

and conditional variances bounded by $V_{2,-}$ and $V_{2,+}$. Lemma H.6, applied to the lower error and to the negative upper error, yields

$$\mathbb{P}\{\widetilde{\theta}^- - t_{2,-}(\gamma_-) > \theta_n^- \mid \text{pilot}, G_P, \mathcal{F}_0\} \le \gamma_-,$$

$$\mathbb{P}\{\widetilde{\theta}^+ + t_{2,+}(\gamma_+) < \theta_n^+ \mid \text{pilot}, G_P, \mathcal{F}_0\} \le \gamma_+.$$

A union bound makes the active-branch failure probability at most $\gamma$.

If either objective is not certified, let

$$G_E := \{\left\|\widehat{m}^{(2)} - m_0\right\|_\infty \le r_2(\gamma)\}.$$

Lemma H.5 gives $\mathbb{P}(G_E^c \mid \mathcal{F}_0) \le \gamma$. On $G_E$, every $\nu \in \mathcal{C}_n(m_0)$ satisfies $A\nu \le b$ and

$$\left\|D\nu - \widehat{m}^{(2)}\right\|_\infty = \left\|m_0 - \widehat{m}^{(2)}\right\|_\infty \le r_2(\gamma).$$

Thus $\mathcal{C}_n(m_0) \subseteq \widehat{\mathcal{C}}_2(\gamma)$, and minimizing and maximizing the same target over the larger set covers the entire identified interval. If the projected set is empty, the structural-only fallback contains $\mathcal{C}_n(m_0)$ under Assumption A.3 and therefore also covers the identified set.

The branch indicator is pilot measurable. Conditional on $G_P$, whichever branch is selected fails with probability at most $\gamma$. Hence

$$\mathbb{P}\{\Theta_n(m_0) \not\subseteq \widehat{\mathcal{I}}_{n,1-\alpha}^{\mathrm{hyb}} \mid \mathcal{F}_0\} \le \beta + \gamma \le \alpha.$$

□

### I.9 Proof of Theorem 5.2 and Corollary D.1

*Proof of Theorem 5.2.* Write

$$L_B := \min_{m\in B} \phi_{n,\ell}(m), \qquad U_B := \max_{m\in B}\{-\phi_{n,-\ell}(m)\}.$$

These are exactly the endpoints of (35). Let $m_L, m_U \in B$ attain the extrema. By Lemma H.7,

$$U_B - L_B = \omega_n(m_L) + \{-\phi_{n,-\ell}(m_U) + \phi_{n,-\ell}(m_L)\} \le \sup_{m\in B} \omega_n(m) + \Gamma_{n,-\ell}(K_0)\,\|m_U - m_L\|_\infty .$$

The $\ell_\infty$-diameter of $B$ is at most $2r$, so this is bounded by $\sup_B \omega_n + 2r\Gamma_n(K_0)$, proving (36).

If $m_0 \in B$, then every $m \in B$ satisfies $\|m - m_0\|_\infty \le 2r$. Lemma H.7 gives

$$L_B \ge \phi_{n,\ell}(m_0) - 2r\Gamma_{n,\ell}(K_0), \quad U_B \le -\phi_{n,-\ell}(m_0) + 2r\Gamma_{n,-\ell}(K_0).$$

Subtracting proves (37). □

*Proof of Corollary D.1.* The two positive-part inequalities follow directly by rearranging (36) and (37). Lemma H.5 gives $\mathbb{P}(m_0 \in B \mid \mathcal{F}_0) \ge 1 - \eta$. Substituting $r_2(\eta)$ from (26), fixed $s, \eta$ and the uniform slope bounds make each statistical term a constant multiple of

$$M\sqrt{\frac{1+\Delta_2}{N_2 p_{\min,2}}}.$$

□

### I.10 Proof of Proposition D.1

*Proof.* Let $G_P = \{m_0 \in \widehat{\mathcal{B}}_1(\beta_{N_1})\}$. Lemma H.5 and $\beta_{N_1} \to 0$ give $\mathbb{P}(G_P) \to 1$. On $G_P$, every $m \in \widehat{\mathcal{B}}_1(\beta_{N_1})$ obeys

$$\|m - m_0\|_\infty \le \left\|m - \widehat{m}^{(1)}\right\|_\infty + \left\|\widehat{m}^{(1)} - m_0\right\|_\infty \le 2r_1(\beta_{N_1}).$$

For all sufficiently large $N_1$, this radius is smaller than $\min_c \delta_c$. The population pair from Condition D.1 is then optimal at every point, and hence every vertex, of the pilot polytope. Theorem 4.2 implies that the common-dual feasibility problem passes for both objectives.

It remains to identify the equality multiplier of every retained pair. The same variance–union calculation used in Lemma H.5 gives, for every fixed $\varepsilon > 0$,

$$\mathbb{P}\Big\{\|\widehat{m}^{(1)} - m_0\|_\infty > \varepsilon r_1(\beta_{N_1})\Big\} \le \beta_{N_1}/\varepsilon^2 = o(1).$$

Take $\varepsilon = 1/2$. With probability tending to one, $\mathcal{M}_n \cap B_\infty\{m_0, r_1(\beta_{N_1})/2\} \subseteq \widehat{\mathcal{B}}_1(\beta_{N_1})$. Local full dimensionality makes this subset full dimensional in $\mathbb{R}^s$. A retained common dual pair represents $\phi_{n,c}$ by its affine dual objective throughout that subset. Condition D.1 represents the same function there by the unique population affine slope. Two affine functions that agree on a full-dimensional set have identical linear coefficients, so every retained equality multiplier equals that slope. This proves both assertions. □

### I.11 Proof of Theorem D.3 and Corollary D.2

*Proof of Theorem D.3.* Fix a population affine direction $a$. The $N_2$ centered summands in (40) have the evaluation dependency graph. Since $0 \le Y \le M$, $0 \le m_i(e) \le M$, and at most one exposure occurs,

$$|\xi_{ir}(a)| \le \sum_e |a_e| \left\{\frac{M}{p_{ir}(e)} + M\right\} \le \frac{2M\,\|a\|_1}{p_{\min,2}} =: B_{N_2}.$$

The variance of their sum is $v_{N_2}^2 = N_2\sigma_{N_2}^2(a)$. Substitution into Lemma H.8 gives precisely condition (42), so the fixed-affine evaluation is asymptotically normal.

By Proposition D.1, with probability tending to one the pilot certificates select the unique population affine equality multipliers. On that event, the endpoint estimation errors are exactly the fixed-affine evaluations above. The complement probability tends to zero. At every continuity point of the standard normal distribution, the difference between the distribution functions of the actual and fixed-affine standardized errors is bounded by that complement probability. The two marginal limits follow. □

*Proof of Corollary D.2.* Theorem D.3 and consistent standard errors give one-sided asymptotic coverage $1-\alpha/2$ for each endpoint. A union bound gives simultaneous endpoint coverage $1-\alpha+o(1)$. On that event, the reported lower limit is no larger than $\theta_n^-$, and the reported upper limit is no smaller than $\theta_n^+$, so the entire population interval is contained. □

### I.12 Proofs of Theorem 5.4, Corollary D.3, and Proposition D.2

*Proof of Theorem 5.4.* Relaxing $b$ to $b+\varepsilon$ enlarges the primal feasible set, so $\phi_{n,c}^{\varepsilon}(m) \le \phi_{n,c}(m)$. The relaxed dual has the same feasibility constraints and objective $\rho^\top m - \lambda^\top(b+\varepsilon)$. Evaluating it at an original optimal pair and using relaxed strong duality gives

$$\phi_{n,c}^{\varepsilon}(m) \ge (\rho^*)^\top m - (\lambda^*)^\top(b+\varepsilon) = \phi_{n,c}(m) - (\lambda^*)^\top\varepsilon.$$

Combining inequalities proves (45). Sharpness applies to the model defined by the relaxed constraints. □

*Proof of Corollary D.3.* For each imposed link,

$$|\mu_{0e} - \mu_{0e'}| \le L d_0(e,e') \le L\widetilde{d}(e,e') + L\delta_{ee'}.$$

Thus adding slack $L\delta_{ee'}$ to that row makes the true mean feasible. The endpoint effect then follows from Theorem 5.4. □

*Proof of Proposition D.2.* A certified absolute gap $\tau_-$ means the reported lower endpoint differs from its exact counterpart by at most $\tau_-$, so subtracting $\tau_-$ is no larger than the exact lower limit. Similarly, adding $\tau_+$ is no smaller than the exact upper limit. Therefore the enlarged numerical interval contains the exact valid interval pointwise. □

### I.13 Proof of Theorem 5.3 and Corollary 5.1

*Proof of Theorem 5.3.* Write $n = Bq$. Partition the units into $B$ disjoint blocks of size $q$. In each wave $r$ and block $b$, draw $Z_{br} \sim \mathrm{Bern}(p)$, independently across block–wave pairs. If $Z_{br} = 1$, every unit in the block receives a supported exposure $e_s$; if $Z_{br} = 0$, every unit receives a reference exposure $e_0$. Introduce a third exposure $e_u$ that is structurally absent.

Under $P_\pm$, conditional on $Z_{br} = 1$, draw $W_{br} \sim \mathrm{Bern}(1/2 \pm \epsilon_N)$ and set the observed outcome of every unit in the block to $MW_{br}$. Under $Z_{br} = 0$, set all block outcomes to zero. The duplicated unit scores within a block form a clique of size $q$, while different blocks and waves are independent; hence the maximum dependency degree is $q-1 = \Delta$. Every unit has supported-exposure probability $p$.

Let the mean of the unsupported coordinate range freely over $[0, D_0]$, and take the target to be the sum of the supported mean under $e_s$ and this unsupported coordinate. The completion constraints consist of this box restriction together with the fixed supported mean. Therefore the identified sets are exactly (48).

The $q$ duplicated outcomes in a block contain no more information than $(Z_{br}, Z_{br}W_{br})$. The KL divergence for one block–wave pair is

$$p\,\mathrm{kl}\{\mathrm{Bern}(1/2+\epsilon_N), \mathrm{Bern}(1/2-\epsilon_N)\}.$$

There are $BR = N/q$ independent block–wave pairs. Lemma H.10 gives

$$\mathrm{KL}(P_+, P_-) \le 16\frac{N}{q}p\epsilon_N^2 = 16c^2.$$

Pinsker's inequality yields $\mathrm{TV}(P_+, P_-) \le \sqrt{\mathrm{KL}(P_+, P_-)/2} \le \sqrt{8}\, c$.

Apply Lemma H.9 with $a = M/2$, $\delta = M\epsilon_N$, and $D = D_0$. The positive testing margin is at least $\kappa_c$, giving

$$\mathbb{E}_{P_+} \operatorname{diam}(\widehat{I}_N) \ge (1-\alpha)D_0 + 2\kappa_c M \epsilon_N,$$

which is (51). Taking the supremum over the pair proves the claim. If the interval has diameter at least $D_0$ pathwise, use that deterministic baseline instead of $D_0 \mathbf{1}(A_+)$ in Lemma H.9. □

*Proof of Corollary 5.1.* Pool the $R$ waves into one HT estimate and use its simultaneous primitive box with error probability $\alpha$ and radius

$$r_N = M\sqrt{\frac{(1+\bar{\Delta})s}{Np_\star \alpha}}.$$

Let $\bar{m}$ be an $\ell_\infty$-projection of $\widehat{m}$ onto $\mathcal{M}_n$, using a fixed measurable tie-breaking rule, and report

$$\left[\phi_{n,\ell}(\bar{m}) - 2\bar{\Gamma} r_N,\, -\phi_{n,-\ell}(\bar{m}) + 2\bar{\Gamma} r_N\right].$$

On the event $\|\widehat{m} - m_0\|_\infty \le r_N$, projection optimality gives $\|\bar{m} - \widehat{m}\|_\infty \le r_N$, hence $\|\bar{m} - m_0\|_\infty \le 2r_N$. Lemma H.5 assigns this event probability at least $1-\alpha$, and condition (iii) then shows that the displayed interval contains both population endpoints. Its diameter is bounded pathwise by $D_0 + 4\bar{\Gamma} r_N$ using conditions (ii)–(iii). This proves the uniform expected-length upper bound.

Condition (iv) and Theorem 5.3 give a two-point submodel inside the class on which every honest interval has expected diameter at least $(1-\alpha)D_0 + 2\kappa_{c_\alpha} c_\alpha M\sqrt{(1+\bar{\Delta})/(Np_\star)}$, where $\kappa_{c_\alpha} = 1 - 2\alpha - \sqrt{8}c_\alpha > 0$. For fixed $\alpha < 1/2$, the upper and lower bounds are constant multiples of $D_0 + M\sqrt{(1+\bar{\Delta})/(Np_\star)}$. Under the pathwise-width condition, the lower coefficient on $D_0$ is one. □

## J Heterogeneous and Partially Supported Exposure Cells

Under the global-cell convention in the main text, an exposure template is either feasible for every unit or forbidden for every unit. In many networks, an exposure may be feasible only for an eligible subset. Let

$$V_e := \{i : \min_{1\le r\le R} p_{ir}(e) > 0\}, \qquad N_e := |V_e| > 0,$$

where eligibility is fixed before outcome analysis. Assume $p_{ir}(e) = 0$ for every $i \notin V_e$ and every wave $r$, so the eligible set is common to all waves. The corresponding design-identifiable coordinate is

$$\mu_{0e}^{\mathrm{elig}} := \frac{1}{N_e}\sum_{i\in V_e} m_i(e), \tag{62}$$

with replicated estimator

$$\widehat{\mu}_e^{\mathrm{elig}} := \frac{1}{RN_e}\sum_{r=1}^{R}\sum_{i\in V_e} \frac{\mathbf{1}\{E_i^{(r)} = e\} Y_i^{(r),\mathrm{obs}}}{p_{ir}(e)}. \tag{63}$$

Termwise expectation proves unbiasedness. With the same unit–wave dependency graph, a covariance argument gives, for a contrast over eligible means,

$$\mathrm{Var}\left\{\sum_e a_e(\widehat{\mu}_e^{\mathrm{elig}} - \mu_{0e}^{\mathrm{elig}})\right\} \le M^2(1+\Delta)\sum_e \frac{a_e^2}{RN_e p_{\min,e}}, \qquad p_{\min,e} := \min_{i\in V_e, r} p_{ir}(e). \tag{64}$$

The completion framework remains linear after its primitive coordinates are redefined as these eligible-population means. This changes the estimand. Moreover, a restriction comparing $\mu_{0e}^{\mathrm{elig}}$ and $\mu_{0e'}^{\mathrm{elig}}$ compares different finite populations when $V_e \ne V_{e'}$. Such a row requires a composition-invariance or transport assumption relating the two eligible populations.

A fully general alternative expands the completion vector to unit–exposure means $\{m_i(e) : i \le n,\ e \in \mathcal{E}\}$ and constrains supported unit–exposure pairs directly. The resulting program remains a sparse LP in $nK$ variables, but its structural assumptions must be stated at the unit level. The unit-level formulation therefore increases both computational cost and the scope of the required structural assumptions.

## K Exposure-Mapping Approximation

Let the analyst's exposure-indexed means satisfy

$$\sup_{T:d_i(T)=e} \left|\mathbb{E}\{Y_i^{\mathrm{obs}} \mid T, \mathcal{F}_0\} - m_i(e)\right| \le \eta_i(e), \qquad \bar{\eta}_e := n^{-1}\sum_i \eta_i(e). \tag{65}$$

Assume the outcome bound, known design, and score dependency graph remain valid, and the nominal means satisfy $A\mu_0 \le b$. Mapping error affects both estimation of supported means and the target represented by the exposure mapping.

**Supported-mean bias.** For a supported exposure, iterated expectation gives

$$\left|\mathbb{E}\{\widehat{\mu}_e^{(w)} \mid \mathcal{F}_0\} - \mu_{0e}\right| \le \bar{\eta}_e.$$

The second-moment argument in Theorem 4.1 still bounds variance around the estimator's own expectation. A coordinatewise Chebyshev bound therefore yields the bias-enlarged region

$$\widehat{\mathcal{B}}_w^{\mathrm{map}}(\eta) := \left\{m \in \mathcal{M}_n : |m_e - \widehat{\mu}_e^{(w)}| \le r_w(\eta) + \bar{\eta}_e \quad (e \in \mathcal{S}_n)\right\}, \tag{66}$$

with $\mathbb{P}\{m_0 \in \widehat{\mathcal{B}}_w^{\mathrm{map}}(\eta) \mid \mathcal{F}_0\} \ge 1-\eta$. Project this region through the completion constraints to obtain $\widehat{I}^{\mathrm{map}} = [L^{\mathrm{map}}, U^{\mathrm{map}}]$. If the region is empty, use the structural-only range. The resulting interval covers the nominal identified set with the same probability.

**Assignment-level target.** For each unit and meaningful exposure, fix a distribution over assignments with that exposure, and average the true conditional response under that distribution. Denote the resulting mean by $\widetilde{m}_i(e)$, and set $\widetilde{\mu}_e = n^{-1}\sum_i \widetilde{m}_i(e)$. Equation (65) implies

$$|\ell^\top \widetilde{\mu} - \ell^\top \mu_0| \le \eta_\ell := \sum_{e\in\mathcal{E}} |\ell_e|\bar{\eta}_e. \tag{67}$$

Thus $[L^{\mathrm{map}} - \eta_\ell, U^{\mathrm{map}} + \eta_\ell]$ covers the assignment-level contrast with probability at least $1-\eta$. The coordinate enlargement controls bias in the supported inputs; the final enlargement transfers the nominal contrast to the assignment-level target. If a candidate mapping changes support, recompute the support partition and its exposure probabilities before applying these steps.

## L Variance Estimation and Zero Joint-Inclusion Probabilities

The finite-sample interval in Theorem 5.1 uses the deterministic variance bound in Theorem 4.1. For the Gaussian refinement, define

$$\pi_{ij,r}(e,e') := \mathbb{P}\{E_i^{(r)} = e, E_j^{(r)} = e'\}.$$

The mean-exposure model specifies first moments and permits conditionally correlated outcomes. For every pair with $\pi_{ij,r}(e,e') > 0$, define the assignment-averaged conditional second moment

$$q_{ij,r}(e,e') := \frac{\mathbb{E}\Big[\mathbf{1}\{E_i^{(r)} = e, E_j^{(r)} = e'\} Y_i^{(r),\mathrm{obs}} Y_j^{(r),\mathrm{obs}} \mid \mathcal{F}_0\Big]}{\pi_{ij,r}(e,e')}. \tag{68}$$

For a fixed affine direction $a$, the exact conditional variance of the one-wave HT sum is

$$\sum_{i,j}\sum_{e,e'\in\mathcal{S}_n} a_e a_{e'} \left\{\frac{\pi_{ij,r}(e,e') q_{ij,r}(e,e')}{p_{ir}(e)p_{jr}(e')} - m_i(e)m_j(e')\right\}, \tag{69}$$

with the first term defined as zero when $\pi_{ij,r}(e,e') = 0$. The variance of the average adds the usual $(nR)^{-2}$ scale and sums over independent waves. Under deterministic exposure-indexed potential outcomes, or under the corresponding conditional product-moment restriction, $q_{ij,r}(e,e') = m_i(e)m_j(e')$, and (69) reduces to the familiar inclusion-covariance formula.

Positive joint-inclusion probabilities make the $q_{ij,r}$ term unbiasedly estimable by inverse-joint-probability observed products. A zero joint-inclusion probability leaves the cross-product of first

moments in (69) unavailable from one wave; distinct exposures for the same unit give the simplest example. Independent repeated waves can estimate such products by cross-wave pairing. Other valid routes are a proved conservative bound or an explicit second-moment restriction. The Gaussian interval in Corollary D.2 therefore requires a ratio-consistent estimator of each displayed marginal score variance under one of these routes. The finite-sample projected and Cantelli procedures continue to use the deterministic bound in Theorem 4.1.

## M Common-Dual Cutting-Plane Implementation

Direct vertex enumeration of $\widehat{\mathcal{B}}_1(\beta)$ can be expensive when the number of supported cells is large. The following finite constraint-generation scheme is exact whenever the implementation supplies an exact oracle for the worst violating vertex. Such an oracle may use parametric-LP basis enumeration or a verified mixed-integer formulation; worst-case complexity can be exponential.

**Algorithm 2** Constraint-generation common-dual certificate for objective $c$

**Require:** Pilot polytope $B$, endpoint objective $c$, tolerance $\tau_{\mathrm{cert}}$.

1: Choose any vertex $v_1 \in \mathrm{vert}(B)$, solve $\phi_{n,c}(v_1)$, and initialize $\mathcal{W}_1 = \{v_1\}$.
2: **for** $k = 1, 2, \ldots$ **do**
3: Solve the common-dual feasibility problem using equality rows for vertices in $\mathcal{W}_k$. If infeasible, return NOT CERTIFIED.
4: Let $(\rho_k, \lambda_k)$ be a feasible common pair and solve

$$\max_{m \in B} \left[\phi_{n,c}(m) - \{\rho_k^\top m - \lambda_k^\top b\}\right]$$

by an exact verified vertex-separation oracle.
5: **if** the maximum duality gap is at most $\tau_{\mathrm{cert}}$ **then**
6: Return CERTIFIED, the pair, and the worst-case gap.
7: **else**
8: Add a maximizing vertex to $\mathcal{W}_k$ to form $\mathcal{W}_{k+1}$, and continue.

With exact arithmetic and $\tau_{\mathrm{cert}} = 0$, the algorithm terminates after finitely many oracle calls because $B$ has finitely many vertices. The number of oracle calls depends on the vertex structure of $B$. In floating-point computation, the returned worst-case gap should enter the numerical enlargement in Proposition D.2.