

Inference for Directional Population Sensitivity

Fragility Curves and Breakdown in Linear Projections

Lucas D. Konrad

WU Vienna

lucas.konrad@wu.ac.at

Nikolas Kuschnig

Monash University

nikolas.kuschnig@monash.edu

Preliminary Draft

October 19, 2026

Precise regression estimates may still strongly depend on the composition of the population itself. We develop $\sqrt{N}$ inference for a directional fragility curve that traces the worst-case change in a target contrast as an increasing share of the population is removed. The directional curve holds the full-sample contrast direction fixed and adjusts for the remaining leverage among retained observations, rather than refitting for each subset. We derive pointwise confidence intervals and simultaneous confidence bands for the curve, accounting for first-stage estimation of both the regression residuals and the design matrix. Inverting the curve yields inference for a breakdown fraction—the smallest removed population share needed to overturn a threshold of interest. We characterize when this directional curve coincides with unrestricted subset refitting and show that it provides a valid lower bound whenever deletion does not rotate the design direction.

1. Introduction

Empirical estimates have at least three basic attributes: magnitude, statistical precision, and sensitivity to population composition. Applied work routinely reports the first two through a point estimate and a standard error, but leaves the third largely implicit. A growing literature therefore seeks a “third number” that quantifies sensitivity to population composition.¹ A standard error measures sampling uncertainty about a fixed population parameter, but not how that parameter would change under a different population composition. Consequently, an estimate can become increasingly precise as the sample grows while remaining just as sensitive to the mix of people, firms, or places represented in the population (Giordano, Meager, and Broderick, 2026b). The practical importance of this distinction becomes apparent when empirical conclusions depend on seemingly minor choices about which observations enter an analysis.²

Consider the linear projection $\beta(P)$ of an outcome y on covariates x , with $Q = \mathbb{E}[xx'] > 0$, and a target contrast $\vartheta(P) = a'\beta(P)$ oriented so that a decrease is adverse.³ Removing a fixed population share $p \in (0, 1)$ and refitting asks how strongly this contrast depends on population composition. In practice, evaluating worst-case deletion sensitivity faces a trade-off between local approximations and combinatorial subset search. The approximate maximum influence perturbation (AMIP) of Giordano, Meager, and Broderick (2026a) linearizes the refit around the full sample, ranking observations by their estimated influence scores and summing the largest $k = \lfloor pN \rfloor$ contributions. At a fixed positive deletion share, however, deleting influential observations also changes the retained design leverage, potentially altering which observations are jointly most influential. Conversely, unrestricted least-squares refitting after deletion requires a multivariate subset search and allows deletion to rotate the covariate adjustment direction.

Our main object, the *directional sensitivity* $\delta(p)$, resolves this trade-off by measuring the largest adverse movement in the contrast from an exact least-squares refit

¹E.g., Giordano, Meager, and Broderick, 2026a,b; Masten and Poirier, 2020; Spini, 2026.

²Examples span development, macroeconomics, and climate economics. The growth–undernutrition association in Vollmer et al. (2014) prompted an exchange over the influence of specific survey waves (O’Connell and Smith, 2016), while anomalous country observations featured centrally in the reanalysis and subsequent retraction of a recent climate-damages estimate (Bearpark, Hogan, and Hsiang, 2025; Kotz, Levermann, and Wenz, 2024, 2025).

³The contrast can represent a regression coefficient, a linear combination of coefficients, or a fitted value at a prespecified covariate vector, and is defined without assuming a linear conditional mean.

constrained to the full-population contrast line $\beta + \gamma Q^{-1}a$. For a single regression coefficient, constraining the refit to this line freezes the full-population partialling-out of the controls. This restriction preserves the adjusted regressor across retained sub-populations, isolating compositional sensitivity and retained leverage from changes in the control adjustment.⁴ Crucially, this restriction makes exact finite-share leverage adjustment computable without combinatorial subset search. By reformulating the line-restricted ratio problem via a scalar root (Dinkelbach, 1967) and the superquantile (CVaR) threshold dual of Rockafellar and Uryasev (2000), we show that the empirical maximum $\hat{\delta}(p)$ is the unique root of a one-dimensional equation solvable in $O(N \log N)$ time per criterion evaluation, and is sandwiched between the first-order AMIP benchmark and a leverage-adjusted upper bound (Proposition 5).

As the deletion share varies, $p \mapsto \delta(p)$ traces a directional *fragility curve*. Given a substantive conclusion margin $c = c(P) > 0$ — for example, the movement required to reduce the population contrast to zero or across a threshold of economic significance — its first crossing defines the directional *breakdown fraction* (or *reversal radius*),

$$p_{\text{dir}}^{\dagger} := \inf\{p : \delta(p) \geq c\},$$

the smallest share whose directional removal overturns the conclusion. Reported alongside the point estimate and its standard error, the fragility curve and breakdown fraction summarize how strongly a conclusion depends on population composition. Because both objects must be estimated from the sample, valid inference is essential to distinguish population fragility from sampling variation.

Contributions. We develop $\sqrt{N}$ -asymptotic inference for the directional fragility curve and breakdown fraction under i.i.d. sampling and regularity conditions near the score cutoff. First, we establish pointwise asymptotic normality (Theorem 1) and a functional central limit theorem over compact intervals of positive deletion shares (Theorem 2). Our asymptotic expansion handles data-dependent changes in the selected observation set via a fixed-budget value sandwich that avoids differentiating

⁴The line-restricted refit requires positive retained variation only along $Q^{-1}a$ (or its sample analogue), remaining well defined even when deletion makes the full retained design singular. When deletion does not rotate the contrast direction at the directional deletion region, $\delta(p)$ provides a lower bound on unrestricted refitting; without that alignment, the ordering can reverse (Section 3.3). A companion manuscript, *Finite-Radius Population Sensitivity*, studies unrestricted refitting and its sampling theory for general estimating equations.

discrete indicators, and accounts for first-stage estimation of both the regression residuals and the contrast direction. Second, these results yield plug-in standard errors, multiplier-bootstrap simultaneous confidence bands for the fragility curve (Proposition 1), and paired simultaneous bands for the gap between the curve and a conclusion margin estimated from the same sample (Proposition 2). Third, an envelope characterization of the curve’s slope yields Wald inference for the breakdown fraction (Theorem 3) alongside confidence bounds from inverting the simultaneous bands (Corollary 1).

1.1. Composition stress testing and sample auditing

The directional fragility curve provides a worst-case stress test of population composition. Rather than prespecifying which groups or characteristics might shift, it asks how much an empirical conclusion could deteriorate under the most adverse removal of a given population share. Allowing selection to depend jointly on outcomes and covariates captures vulnerabilities that need not arise from treatment effect heterogeneity or covariate shift.⁵ The resulting breakdown fraction translates this sensitivity into an interpretable substantive unit — the smallest population share whose directional removal overturns the conclusion.

Our framework admits two complementary modes of analysis. Under i.i.d. sampling and regular score cutoffs, the estimated curve and breakdown fraction support population inference about compositional sensitivity. We verify these conditions for continuous or mixed controls under explicit moment and conditional-density bounds, and for a complementary finite-covariate class. Alternatively, for a realized dataset, the same computation provides a deterministic audit of sensitivity to deleting observations, without requiring a sampling model. The latter interpretation requires only a specified regression, deletion unit, and positive retained directional leverage.

Our empirical illustrations highlight both modes. First, a deterministic country-level audit of the institutions–income regression of Acemoglu, Johnson, and Robinson (2001) shows how finite-share leverage adjustment changes the sensitivity assessment

⁵Deleting the upper tail of $(a'Q^{-1}x)u$ induces correlation between x and u in the retained subpopulation, so $\delta(p)$ is generally positive even when $\mathbb{E}[y \mid x] = x'\beta$ and $u \perp\!\!\!\perp x$. In a randomized experiment with a constant treatment effect, for instance, trimming treated units with high outcomes and control units with low outcomes moves the difference in means without any treatment effect heterogeneity; when $\mathbb{E}[y \mid x]$ is nonlinear, covariate-only selection can also move the linear projection by reweighting local slopes.

relative to first-order influence approximations. Second, we illustrate population curve and band calculations under a maintained sampling model using a large cognitive-training trial (Hardy et al., 2015).

1.2. Relation to the literature

Our paper connects distributional sensitivity analysis with influence-function diagnostics, influential-subset search, and robust statistics.

Removing and renormalizing a population share p corresponds to the contaminated-sampling geometry of Horowitz and Manski (1995) in reverse, with the original distribution held fixed and the retained distribution subject to a density-ratio cap $(1 - p)^{-1}$. Related distributional sensitivity analyses optimize parameters over neighborhoods of the reference distribution (Christensen and Connault, 2023; Dewaskar et al., 2025), while the breakdown frontier of Masten and Poirier (2020) characterizes assumption relaxations that overturn a conclusion and develops uniform inference for that boundary. In the subpopulation literature, Jeong and Namkoong (2020) examine worst-case effects across covariate-defined subgroups, Kallus (2023) studies worst-fraction treatment effect risk via CVaR, and Spini (2026) proposes a covariate-shift breakdown measure for external validity. Our setting allows deletion to depend jointly on outcomes and covariates, and accounts for the resulting depletion of retained design leverage in a linear projection.

Classical influence functions characterize the directional derivative of an estimator as the perturbation share $p \rightarrow 0$ (Hampel, 1968, 1974; Huber, 1981; Ichimura and Newey, 2022). Recent work uses this local calculus to approximate the effect of dropping small subsets of observations (Fisher et al., 2023; Giordano, Meager, and Broderick, 2026a,b). Complementing local influence approximations, a large literature building on Cook (1977) and Cook (1979) and Lawrance (1995) develops subset-search heuristics, approximation bounds, and deterministic robustness certificates for fixed datasets (Freund and Hopkins, 2023; Hu et al., 2024; Huang et al., 2025; Konrad and Kuschnig, 2026a; Kuschnig, Zens, and Crespo Cuaresma, 2021; Moitra and Rohatgi, 2023; Rubinstein and Hopkins, 2024), as well as finite- k stability bounds (Azar, Feldman, and Nadler, 2026) and extreme-value tests for maximal influence (Konrad and Kuschnig, 2026b). Our approach complements these methods by retaining an exact finite-share leverage adjustment along a fixed contrast direction, while providing a population target and sampling theory for the resulting sensitivity curve.

In robust statistics and econometrics, breakdown points and maximum-bias curves are traditionally used to design or compare estimators that limit the influence of outliers, contamination, or local moment misspecification (Armstrong and Kolesár, 2021; Berrendero and Zamar, 2001; Bonhomme and Weidner, 2022; Čížek, 2008; Donoho and Huber, 1983; Kitamura, Otsu, and Evdokimov, 2013). Rather than replacing ordinary least squares with a trimmed or robust estimator, we hold the standard linear projection estimand fixed and conduct inference on how its population value moves across retained subpopulations.

1.3. Roadmap

Sections 2 and 3 define the directional target, scalar computation, and influence function, and explain the qualified connection to unrestricted refitting. Section 4 develops curve bands and breakdown inference and summarizes primitive applicability and practical reporting. Section 5 assesses finite-sample performance in Monte Carlo simulations. Section 6 presents an institutions–income sample audit and a conditional cognitive-training curve-and-band illustration. The Appendix gives the directional proofs; the Online Supplement collects primitive verifications, the unrestricted bridge, the matched signal-to-noise comparison and additional numerical details.

2. Projection Setup and Target Contrast

We study the linear projection of a scalar outcome y on a covariate vector $x \in \mathbb{R}^d$. Let $\{(x_i, y_i)\}_{i=1}^N \subset \mathbb{R}^d \times \mathbb{R}$ be i.i.d. draws from a population P with $\mathbb{E}y^2 < \infty$, $\mathbb{E}\|x\|^4 < \infty$, and a positive-definite design moment $Q := \mathbb{E}[xx'] \succ 0 \in \mathbb{R}^{d \times d}$. Define the population projection coefficient $\beta = \beta(P)$, its residual, and the conditional residual second moment by

$$\beta := Q^{-1}\mathbb{E}[xy], \quad u := y - x'\beta, \quad \sigma^2(x) := \mathbb{E}[u^2 | x], \quad (1)$$

with $\mathbb{E}[\|x\|^2 \sigma^2(x)] < \infty$. The population normal equation

$$\mathbb{E}[xu] = 0 \quad (2)$$

3. Directional Sensitivity and Computation

holds by construction and defines $\beta(P)$ without requiring the conditional mean $\mathbb{E}[y | x]$ to be linear. This setup allows nonlinear conditional expectations, dependence between x and u , and conditional heteroskedasticity. Our inference results and exact-refitting comparison impose additional conditions where stated.

The full-sample OLS estimator, design matrix, and residuals are

$$\hat{\beta} := \hat{Q}_N^{-1} \left(N^{-1} \sum_i x_i y_i \right) \text{ with } \hat{Q}_N := N^{-1} \sum_i x_i x_i', \quad \hat{u}_i := y_i - x_i' \hat{\beta}.$$

To index sample deletions of a fixed fraction $p \in (0, 1)$, set $k = \lfloor pN \rfloor$, $p_N = k/N \rightarrow p$, and let $\mathcal{S} := \{S \subset [N] : |S| = k\}$ denote all size- k subsets of $[N] := \{1, \dots, N\}$.

For a nonzero contrast vector $a \in \mathbb{R}^d$ fixed before observing the data, the target parameter is $\vartheta(P) = a' \beta(P)$. Define the contrast direction in the population design metric and its squared norm in that metric by

$$r_a := Q^{-1} a, \quad \mu_Z(a) := a' Q^{-1} a,$$

with sample analogues $\hat{r}_a, \hat{\mu}_Z(a)$ based on $\hat{Q}_N^{-1}$. Here, $(r_a' x_i) u_i = a' Q^{-1} x_i u_i$ is the standard full-population influence score for $a' \hat{\beta}$, and $\mu_Z(a) = \mathbb{E}[(r_a' x)^2]$ is the second moment of the projected regressor $r_a' x$.

3. Directional Sensitivity and Computation

The directional rule refits least squares along the fixed coefficient line $\beta + \gamma r_a$, where $r_a = Q^{-1} a$. For a deletion region D with $\mathbb{E}[(r_a' x)^2 \mathbf{1}\{D^c\}] > 0$, define

$$\gamma_D := \arg \min_{\gamma \in \mathbb{R}} \mathbb{E}[\{y - x'(\beta + \gamma r_a)\}^2 \mathbf{1}\{D^c\}], \quad \beta_{-D}^{\text{line}} := \beta + \gamma_D r_a.$$

The scalar normal equation, $\mathbb{E}[xu] = 0$ and $a' r_a = \mu_Z(a)$ give

$$\gamma_D = - \frac{\mathbb{E}[(r_a' x) u \mathbf{1}\{D\}]}{\mathbb{E}[(r_a' x)^2 \mathbf{1}\{D^c\}]}, \quad a'(\beta - \beta_{-D}^{\text{line}}) = \frac{\mathbb{E}[(r_a' x) u \mathbf{1}\{D\}]}{\mathbb{E}[(r_a' x)^2 \mathbf{1}\{D^c\}] / \mu_Z(a)}.$$

This contrast movement is the ratio of deleted profile scores to retained directional leverage, motivating the normalized pair

$$Y_i(a) := r_a' x_i u_i, \quad Z_i(a) := \frac{(r_a' x_i)^2}{\mu_Z(a)}, \quad \mathbb{E}[Z(a)] = 1. \quad (3)$$

Here $\mu_Z(a) = \mathbb{E}[(r'_a x)^2]$ is the unnormalized index second moment. For a coefficient on w with controls z and $a = e_1$, full-population partialling-out gives

$$x = (w, z')', \quad r'_a x = \frac{\tilde{w}}{\mathbb{E}[\tilde{w}^2]}, \quad \tilde{w} = w - z' \mathbb{E}[zz']^{-1} \mathbb{E}[zw].$$

The rule keeps this adjusted regressor fixed as observations are removed. Replacing population quantities $(Y_i(a), Z_i(a))$ with their full-sample OLS counterparts $\hat{s}_i(a) := \hat{r}'_a x_i \hat{u}_i$ and $\hat{Z}_i(a) := (\hat{r}'_a x_i)^2 / \hat{\mu}_Z(a)$ gives the *directional statistic*:

$$\hat{\delta}(p) := \max_{S \in \mathcal{S}} \frac{\sum_{i \in S} \hat{s}_i(a)}{\sum_{i \notin S} \hat{Z}_i(a)}. \quad (4)$$

At the population level, the directional sensitivity target is

$$\delta(p) := \sup_{P(D)=p, \mathbb{E}[Z(a)\mathbf{1}\{D^c\}]>0} \frac{\mathbb{E}[Y(a)\mathbf{1}\{D\}]}{\mathbb{E}[Z(a)\mathbf{1}\{D^c\}]},$$

which maximizes the line-restricted contrast movement $a'(\beta - \beta_D^{\text{line}})$ over measurable deletion regions of mass p . Since $\mathbb{E}[Z(a)] = 1$, any region D with positive retained leverage satisfies, for a candidate $\lambda \geq 0$,

$$\frac{\mathbb{E}[Y(a)\mathbf{1}\{D\}]}{\mathbb{E}[Z(a)\mathbf{1}\{D^c\}]} \geq \lambda \iff \mathbb{E}[\{Y(a) + \lambda Z(a)\}\mathbf{1}\{D\}] - \lambda \geq 0.$$

By Dinkelbach (1967)'s reduction, comparing the ratio to λ reduces to a linear selection problem in the modified score $W_a(\lambda) := Y(a) + \lambda Z(a)$ (see Konrad and Kuschnig, 2026a). Let $F_V^{-1}(u) := \inf\{t : F_V(t) \geq u\}$ denote the left-continuous quantile function of an integrable random variable V . By the CVaR dual of Rockafellar and Uryasev (2000), the upper- p tail integral of V is

$$\mathcal{T}_p(V) := \int_{1-p}^1 F_V^{-1}(u) du = \min_{\xi \in \mathbb{R}} \{\mathbb{E}[(V - \xi)_+] + p\xi\},$$

which equals the maximal expectation of V over relaxed deletion weights $0 \leq \pi \leq 1$ with $\mathbb{E}\pi = p$. Accordingly, the population root criterion is

$$f(\lambda) := \mathcal{T}_p\{Y(a) + \lambda Z(a)\} - \lambda = \min_{\xi \in \mathbb{R}} \mathbb{E}[(Y(a) + \lambda Z(a) - \xi)_+ + p\xi - \lambda Z(a)].$$

When $(Y(a), Z(a))$ has finite second moments, retained leverage is positive for every

mass- p deletion rule ($P(r'_a x = 0) < 1 - p$), and the upper- p score tail is positive ($\mathcal{T}_p\{Y(a)\} > 0$), Proposition 4 in Section A establishes that $f(\lambda) = 0$ has a unique positive root equal to the relaxed ratio value (Assumptions 2 to 4). If $W_a(\lambda)$ has no probability mass at its $(1 - p)$ -quantile ξ at the root, the optimum is attained by the pure threshold region $\{W_a(\lambda) \geq \xi\}$ and the root equals $\delta(p)$ (Assumption 5).

3.1. Computation

The sample problem mirrors the population reduction exactly. For $1 \leq k < N$, the sample normal equations imply $\sum_i \hat{s}_i(a) = 0$, and $N^{-1} \sum_i \hat{z}_i(a) = 1$ makes the retained denominator $1 - N^{-1} \sum_{i \in S} \hat{z}_i(a)$. Suppressing the argument a , let $T_{N,p}(V) := N^{-1} \sum_{j=1}^k V_{(N-j+1)}$ denote the normalized sum of the k largest values of an array V , and set $h_Z := T_{N,p}(\hat{Z})$. Every size- $(N - k)$ retained directional denominator is strictly positive if and only if $h_Z < 1$. On that event, Propositions 3 and 5 characterize $\hat{\delta}(p)$ as the unique root of $f_N(\lambda) = 0$ and sandwich it between the first-order AMIP approximation and a leverage-adjusted upper bound:

$$f_N(\lambda) := T_{N,p}(\hat{s} + \lambda \hat{Z}) - \lambda, \quad T_{N,p}(\hat{s}) \leq \hat{\delta}(p) \leq \frac{T_{N,p}(\hat{s})}{1 - h_Z}.$$

Evaluating $f_N(\lambda)$ at a candidate λ requires sorting the modified scores $\hat{s}_i + \lambda \hat{z}_i$ and summing the largest k (which costs $O(N \log N)$ operations; see Konrad and Kuschnig, 2026a). Because the relative ranking of observations changes with λ , a single initial sort at $\lambda = 0$ does not generally solve the finite-share problem. Bisection on the bracket of width $H = T_{N,p}(\hat{s})h_Z/(1 - h_Z)$ computes the unique root to absolute tolerance ε in $O(1 + \log(1 + H/\varepsilon))$ evaluations.⁶ The final sort returns a candidate subset S^* and cutoff; by Proposition 3, exact optimality is certified whenever S^* remains a top- k set of $\hat{s}_i + \lambda \hat{z}_i$ at its own attained ratio λ . Any boundary ties at the cutoff leave $\hat{\delta}(p)$ invariant (Proposition 3 and Lemma 17), although finite-sample tie counts do not by themselves verify population cutoff regularity (Remark 1).

⁶Our asymptotic theorems apply to the exact root $\hat{\delta}(p)$; choosing a numerical tolerance $\varepsilon = o(N^{-1/2})$ ensures that root-finding error is asymptotically negligible.

3.2. Influence function

At a regular cutoff the directional rule deletes the upper- p region of the modified score; we suppress budget arguments at a fixed p :

$$W_a^* := Y(a) + \delta(p) Z(a), \quad \xi_a^* := F_{W_a^*}^{-1}(1 - p), \quad (5)$$

yielding the population *active set* $A^* := \{W_a^* \geq \xi_a^*\}$. Three population quantities govern the oracle expansion of the root around A^* :

$$\begin{aligned} \nu_{a,Z}^* &:= \mathbb{E}[Z(a) \mathbf{1}\{A^*\}], \\ B_a^* &:= 1 - \nu_{a,Z}^* = \mathbb{E}[Z(a) \mathbf{1}\{A^{*,c}\}], \end{aligned} \quad (6)$$

$$\psi_{a,i}^* := (W_{a,i}^* - \xi_a^*)_+ + p \xi_a^* - \delta(p) Z_i(a). \quad (7)$$

Here $\nu_{a,Z}^*$ and B_a^* measure the deleted and retained fractions of directional leverage; $B_a^* > 0$ is also the magnitude of the population criterion's slope, $-\partial_\lambda f(\delta(p)) = B_a^*$. Evaluating the dual integrand at the root and optimal cutoff gives the mean-zero oracle score $\psi_{a,i}^*$; the term $-\delta(p) Z_i(a)$ preserves the empirical leverage normalization of the ratio. Implicit differentiation of the root divides this score by B_a^* , so the oracle root's influence function is $\psi_{a,i}^*/B_a^*$, before the first-stage corrections below.

Because the feasible statistic $\hat{\delta}(p)$ is constructed from estimated residuals $\hat{u}_i$ and the estimated direction $\hat{r}_a$, its asymptotic expansion acquires two first-stage corrections through a fixed-budget value sandwich that avoids differentiating the discrete selected set (Lemma 8). First, estimating $\hat{\beta}$ shifts the numerator scores and introduces a residual correction governed by the active-set cross-moment

$$m_a := \mathbb{E}[x(r'_a x) \mathbf{1}\{A^*\}].$$

Second, when $d \geq 2$, estimating $\hat{r}_a = \hat{Q}_N^{-1} a$ introduces a design-direction correction governed by the gradient $g_a := \nabla_r \Phi_a(r)|_{r=r_a}$ of the population directional functional $\Phi_a(r) := (a'r) \delta^{\text{sc}}(r)$, where $\delta^{\text{sc}}(r)$ solves $\mathcal{T}_p\{(r'x)u + \lambda(r'x)^2\} = \lambda(r'Qr)$. Writing $\delta^{\text{sc}} := \delta(p)/\mu_Z(a)$ and $g_{A^*} := \mathbb{E}[xu \mathbf{1}\{A^*\}]$, Theorem 5 gives the closed form

$$g_a = \delta^{\text{sc}} a + \frac{1}{B_a^*} [g_{A^*} + 2\delta^{\text{sc}}(m_a - a)]. \quad (8)$$

By degree-zero homogeneity of $\Phi_a(r)$, g_a is orthogonal to the contrast direction

($g'_a r_a = 0$), so it vanishes identically when $d = 1$. The design correction need not vanish in general; the appendix gives sufficient simplifications and explains why mean-linearity of the index regression alone is insufficient (Corollary 2 and remark 3).

Conditions for inference. Pointwise inference requires sufficient moments, a regular score cutoff, and positive retained directional leverage and influence variance. Simultaneous inference additionally requires these conditions to hold uniformly over the deletion-share interval. The following assumption summarizes the requirements; Assumptions 6 and 7 state the detailed technical clauses, and the Supplement gives primitive verifications.

Assumption 1 (Regularity for directional inference).

- (a) **Pointwise regularity.** At a fixed $p \in (0, 1)$, observations are i.i.d., the projection is well-defined with sufficient fourth and mixed moments, and the directional root exists and is nondegenerate. The modified score has a unique cutoff with positive, continuous density and a differentiable cutoff path near the root. Retained directional leverage and influence variance are positive.
- (b) **Uniform regularity.** On a compact interval $\mathcal{P} \subset (0, 1)$, the pointwise conditions hold uniformly: score densities are continuous and bounded above and away from zero near their cutoffs, and retained leverage and influence variance are bounded away from zero.

Estimated margins additionally require Assumption S1, paired bands require a uniform positive gap-variance floor (Proposition 2), and crossing Wald inference requires an interior crossing and positive variance there (Theorem 3).

Theorem 1 (CLT for $\hat{\delta}(p)$). Under Assumption 1(a),

$$\sqrt{N}(\hat{\delta}(p) - \delta(p)) = \frac{1}{B_a^* \sqrt{N}} \sum_{i=1}^N \Psi_{a,i} + o_p(1) \xrightarrow{d} \mathcal{N}(0, \omega_a^2), \quad \omega_a^2 = \frac{\text{Var}(\Psi_a)}{B_a^{*2}}.$$

The influence function is

$$\Psi_{a,i} = \psi_{a,i}^* - m'_a Q^{-1} x_i u_i - B_a^* (r'_a x_i) (\zeta'_a x_i), \quad \zeta_a = Q^{-1} g_a. \quad (9)$$

The three terms capture first-order variation from oracle selection, estimation of the projection residuals, and estimation of the contrast direction. The last term vanishes

under the sufficient conditions of Corollary 2; its general form and sample-splitting interpretation are discussed in Section A.

3.3. Relation to unrestricted refitting

Whereas $\delta(p)$ restricts the refitted coefficient vector to the full-population contrast line $\beta + \gamma Q^{-1}a$, unrestricted refitting allows deletion to rotate the covariate adjustment direction:

$$\delta^{\text{ex}}(p) := \sup_{P(D)=p} \{\vartheta(P) - \vartheta(P_{-D})\}, \quad (10)$$

where P_{-D} is the retained distribution conditional on D^c (assuming the retained projection is non-singular). At the directional active set A_p^* , let $\Sigma_p = \mathbb{E}[xx' \mathbf{1}\{A_p^{*,c}\}] > 0$ denote the retained design matrix. If deleting A_p^* rescales rather than rotates the contrast direction, i.e., $\Sigma_p^{-1}a \propto Q^{-1}a$, then $\Sigma_p^{-1}a = B_a^*(p)^{-1}r_a$, and an unrestricted OLS refit on $A_p^{*,c}$ produces the exact same contrast movement as the line-restricted refit. Under this no-rotation condition, $\delta(p)$ is an attained unrestricted movement and therefore a lower bound on worst-case unrestricted sensitivity, so $\delta(p) \leq \delta^{\text{ex}}(p)$.

When deletion rotates the contrast direction, the two targets are not generally ordered. Because OLS minimizes squared prediction error, relaxing the line restriction on a given subset can amplify or attenuate the movement in $a'\beta$. In practice, whenever the retained design matrix on the directionally selected complement $(S^*)^c$ is invertible, one can directly compute the unrestricted OLS refit $\hat{\beta}_{-S^*}$ as a diagnostic check without imposing population alignment. The Supplement formalizes this bridge, provides an example in which the ordering reverses, and characterizes stationarity and global target agreement under elliptical symmetry.

4. Inference for the Fragility Curve and Breakdown Fraction

At a fixed deletion share $p \in (0, 1)$, Theorem 1 yields a direct plug-in standard error for the directional sensitivity target $\delta(p)$. Let $S^* \in \mathcal{S}$ denote the optimal size- k subset returning $\hat{\delta}(p)$ in (4), with boundary ties broken by observation index, and let $\hat{\xi}(p) = \hat{W}_{(N-k+1)}^*$ be the k -th largest fitted modified score $\hat{W}_i^* = \hat{s}_i(a) + \hat{\delta}(p)\hat{Z}_i(a)$. Set $p_N = k/N$, $\hat{v}_i = \hat{r}_a' x_i$, $\delta^{\text{sc}} = \hat{\delta}(p)/\hat{\mu}_Z(a)$ and $\hat{B}_a^* = N^{-1} \sum_{i \notin S^*} \hat{Z}_i(a)$. Averaging the selected moments over all N observations gives $\hat{m}_a = N^{-1} \sum_{i \in S^*} x_i \hat{v}_i$ and $\hat{g}_{S^*} = N^{-1} \sum_{i \in S^*} x_i \hat{u}_i$,

4. Inference for the Fragility Curve and Breakdown Fraction

and hence

$$\hat{g}_a = \delta^{\text{sc}} a + \hat{B}_a^{*-1} [\hat{g}_{S^*} + 2\delta^{\text{sc}}(\hat{m}_a - a)], \quad \hat{\zeta}_a = \hat{Q}_N^{-1} \hat{g}_a. \quad (11)$$

The fitted influence score is

$$\hat{\Psi}_{a,i} = (\hat{W}_i^* - \hat{\xi})_+ + p_N \hat{\xi} - \delta(p) \hat{Z}_i(a) - (x_i' \hat{Q}_N^{-1} \hat{m}_a) \hat{u}_i - \hat{B}_a^* \hat{v}_i(\hat{\zeta}_a' x_i). \quad (12)$$

Here p_N matches the fitted top- k criterion; using p instead adds a common offset to the scores and leaves their sample variance unchanged. Writing $\widehat{\text{Var}}(\hat{\Psi}_a)$ for the sample variance of $\{\hat{\Psi}_{a,i}\}_{i=1}^N$, the sandwich standard error is

$$\widehat{\text{SE}}(\hat{\delta}(p)) = \left\{ N^{-1} \hat{B}_a^{*-2} \widehat{\text{Var}}(\hat{\Psi}_a) \right\}^{1/2}. \quad (13)$$

Under Assumption 1(a) at this single prespecified share, the fitted scores are empirically L^2 consistent and the standard error consistently estimates $\omega_a/\sqrt{N}$ (Proposition 7). Thus $\hat{\delta}(p) \pm z_{1-\alpha/2} \widehat{\text{SE}}(\hat{\delta}(p))$ has asymptotic coverage $1 - \alpha$ (Corollary 3), without uniform-in- p conditions. Reporting a path or choosing a share after inspecting it requires the simultaneous inference below; comparing it with an estimated margin also requires the paired score in (17).

4.1. Functional central limit theorem

To report sensitivity across a range of deletion shares, fix a compact interval $\mathcal{P} = [\underline{p}, \bar{p}] \subset (0, 1)$. The directional fragility curve is $\delta(p)$, and its empirical counterpart $\hat{\delta}(p)$ is computed at the integer budget $k(p) = \lfloor pN \rfloor$. The cutoff conditions identify the region target with the root at each share. All active-set quantities now carry an explicit budget argument: the modified score $W_a^*(p) = Y(a) + \delta(p)Z(a)$, its $(1-p)$ -quantile cutoff $\xi_a^*(p)$, the active set $A_p^* = \{W_a^*(p) \geq \xi_a^*(p)\}$, the retained-leverage slope $B_a^*(p)$, and the influence score $\Psi_a(p)$.

Under Assumption 1(b), Section B derives the uniform first-stage expansion (Lemma 14) and the fitted-score consistency and entropy bounds (Lemmas 9 and 15) needed for the multiplier bootstrap.

Lemma 1 (The fragility curve). *Under Assumption 1(b), $p \mapsto \delta(p)$ is continuously*

differentiable on the interior of $\mathcal{P}$, with

$$\partial_p \delta(p) = \frac{\xi_a^*(p)}{B_a^*(p)}, \quad (14)$$

and is strictly increasing on $\mathcal{P}$.

By the envelope theorem, increasing the deletion share p raises directional sensitivity in proportion to the marginal modified score at the deletion cutoff, $\xi_a^*(p)$, scaled by the retained directional leverage $B_a^*(p)$. Projection orthogonality ($\mathbb{E}[Y(a)] = 0$) and the root identity ($\mathbb{E}[W_a^*(p)\mathbf{1}\{A_p^*\}] = \delta(p) = \mathbb{E}[W_a^*(p)]$) imply that retained modified scores average to zero. Because retained scores lie below $\xi_a^*(p)$ and are nondegenerate, the cutoff must be strictly positive at every interior share (Lemma 11), making the population fragility curve strictly increasing. Its empirical counterpart $\hat{\delta}(p)$ is a nondecreasing step path in $k = \lfloor pN \rfloor$ whenever retained directional leverage is positive (Lemma 17; see Section B.9 for increment bounds, tie invariance, and crossing search). Numerical root error at most ε permits downward steps of at most 2ε .

Theorem 2 (Functional central limit theorem). *Under Assumption 1(b),*

$$\sqrt{N}\{\hat{\delta}(p) - \delta(p)\} = \frac{1}{B_a^*(p)\sqrt{N}} \sum_{i=1}^N \Psi_{a,i}(p) + o_p(1) \quad \text{uniformly over } p \in \mathcal{P}.$$

In $\ell^\infty(\mathcal{P})$, the process converges weakly to a tight mean-zero Gaussian process $\mathbb{G}$ with almost surely uniformly continuous paths and covariance kernel

$$K(p, q) = \frac{\mathbb{E}[\Psi_a(p)\Psi_a(q)]}{B_a^*(p)B_a^*(q)}, \quad K(p, p) = \omega_a^2(p). \quad (15)$$

The proof in Section B first establishes the uniform Bahadur expansion for the oracle root path, and then applies Lemma 14 to incorporate the generated-residual and design-direction corrections uniformly over $p \in \mathcal{P}$, with budget discretization $k(p) = \lfloor pN \rfloor$ contributing $O(N^{-1})$. At any fixed deletion share p , the marginal limit recovers Theorem 1. Across shares, estimation errors share observations and common first-stage estimates; Theorem 2 characterizes this joint variation via $K(p, q)$, providing the basis for simultaneous confidence bands and breakdown-fraction inference.

4.2. Simultaneous bands and the breakdown fraction

The functional limit in Theorem 2 supports both simultaneous confidence bands for the fragility curve and inference for the deletion share at which the curve crosses a conclusion threshold. Both procedures use the fitted influence scores of Section A.3.8.

Simultaneous confidence bands. For each $p \in \mathcal{P}$, compute $\hat{\Psi}_{a,i}(p)$, $\hat{B}_a^*(p)$, and $\hat{\omega}_a(p) = \{\widehat{\text{Var}}(\hat{\Psi}_a(p))\}^{1/2}/\hat{B}_a^*(p)$ at $k(p) = \lfloor pN \rfloor$. Draw independent standard-normal multipliers $e_1, \dots, e_N$, independent of the data, and form the multiplier process

$$\hat{\mathbb{G}}_N^*(p) = \frac{1}{\hat{B}_a^*(p)\sqrt{N}} \sum_{i=1}^N e_i \{\hat{\Psi}_{a,i}(p) - \bar{\Psi}_a(p)\}, \quad \bar{\Psi}_a(p) = \frac{1}{N} \sum_{i=1}^N \hat{\Psi}_{a,i}(p). \quad (16)$$

Let $\hat{q}_{1-\alpha}$ be the conditional $(1 - \alpha)$ -quantile of $\sup_{p \in \mathcal{P}} |\hat{\mathbb{G}}_N^*(p)|/\hat{\omega}_a(p)$ given the data.

Proposition 1 (Simultaneous band). *Under Assumption 1(b),*

$$\mathbb{P}\left(\delta(p) \in [\hat{\delta}(p) \pm \hat{q}_{1-\alpha} \hat{\omega}_a(p)/\sqrt{N}] \ \forall p \in \mathcal{P}\right) \longrightarrow 1 - \alpha.$$

In Section B, Lemma 15 establishes uniform plug-in score consistency and verifies the fitted-class entropy condition (Lemma 9), which Lemma 16 uses to prove conditional multiplier validity. The fitted centered scores, slopes, and studentizers depend on p only through $k(p) = \lfloor pN \rfloor$ under the fixed tie-breaking rule. Consequently, evaluating every integer budget $k = \lfloor N\underline{p} \rfloor, \dots, \lfloor N\bar{p} \rfloor$ computes the fitted multiplier supremum over $\mathcal{P}$ exactly. The critical value and band inversion require all these budgets because the budget-dependent studentizer can make band endpoints nonmonotone. The point curve, by contrast, can be searched or bracketed using the monotonicity of Lemma 17, as in Corollary 4. Directional inference requires positive retained leverage along the contrast direction, allowing the curve band, gap band and breakdown procedures to apply under their respective conditions even when a retained design matrix is singular.

Estimated conclusion margins. In applications, the conclusion margin $c = c(P) > 0$ may be estimated from the same sample as the directional curve. For example, the distance of a prespecified contrast $a'_0\beta$ from a threshold b gives

$$c = |a'_0\beta - b|, \quad \hat{c} = |a'_0\hat{\beta} - b|,$$

with the adverse direction oriented toward b . Inference for the gap $g(p) = \delta(p) - c$ must then account for the covariance between the estimated curve and margin.

Suppose $\hat{c}$ is asymptotically linear with influence function φ_c and fitted scores $\hat{\varphi}_{c,i}$ satisfying Assumption S1 in the Supplement. The population and fitted gaps and their paired influence scores are

$$\begin{aligned} g(p) &= \delta(p) - c, & \hat{g}(p) &= \hat{\delta}(p) - \hat{c}, \\ h_i(p) &= \frac{\Psi_{a,i}(p)}{B_a^*(p)} - \varphi_c(O_i), & \hat{h}_i(p) &= \frac{\hat{\Psi}_{a,i}(p)}{\hat{B}_a^*(p)} - \hat{\varphi}_{c,i}, \end{aligned} \quad (17)$$

where $O_i = (x_i, y_i)$. For the contrast margin above, $\varphi_c(x, y) = (r'_a x)u$ when $a = \text{sign}(a'_0 \beta - b)a_0$. Computations use $\hat{a} = \text{sign}(a'_0 \hat{\beta} - b)a_0$ for the fitted curve and $\hat{\varphi}_{c,i} = (\hat{r}'_{\hat{a}} x_i) \hat{u}_i$. A known margin sets both margin scores to zero. Section S1 in the Supplement gives the orientation argument, alternative margin specifications, and pointwise gap inference.

Proposition 2 (Simultaneous gap band). *Under Assumption 1(b), Assumption S1, and $\inf_{p \in \mathcal{P}} \sigma_g^2(p) > 0$, where $\sigma_g^2(p) = \text{Var}(h(p))$, let $\hat{\sigma}_g^2(p) = \widehat{\text{Var}}(\hat{h}(p))$. Using the standard-normal multipliers $e_1, \dots, e_N$ from (16), shared by the curve and margin across all $p \in \mathcal{P}$, form*

$$\hat{\mathbb{G}}_{g,N}^*(p) = \frac{1}{\sqrt{N}} \sum_{i=1}^N e_i \{\hat{h}_i(p) - \bar{\hat{h}}(p)\}, \quad \bar{\hat{h}}(p) = \frac{1}{N} \sum_{i=1}^N \hat{h}_i(p),$$

and let $\hat{q}_{1-\alpha}^g$ be the conditional $(1 - \alpha)$ -quantile of $\sup_{p \in \mathcal{P}} |\hat{\mathbb{G}}_{g,N}^*(p)| / \hat{\sigma}_g(p)$. Then

$$\mathbb{P}\left(g(p) \in [\hat{g}(p) \pm \hat{q}_{1-\alpha}^g \hat{\sigma}_g(p) / \sqrt{N}] \ \forall p \in \mathcal{P}\right) \longrightarrow 1 - \alpha.$$

The proof is given in Section B.

The breakdown fraction. Given a fixed threshold $c > 0$ or an estimated positive margin $c = c(P)$ satisfying Assumption S1, the directional breakdown fraction and its empirical counterpart are

$$p_{\text{dir}}^\dagger := \inf\{p \in \mathcal{P} : \delta(p) \geq c\}, \quad \hat{p}_{\text{dir}}^\dagger := \inf\{p \in \mathcal{P} : \hat{\delta}(p) \geq \hat{c}\}.$$

4. Inference for the Fragility Curve and Breakdown Fraction

We abbreviate these as $p^\dagger$ and $\hat{p}^\dagger$. On a prespecified reporting interval $\mathcal{P} = [\underline{p}, \bar{p}]$, we right-censor both quantities at $\bar{p}$ via the convention $\inf \emptyset = \bar{p}$. Like the breakdown frontier of Masten and Poirier (2020), $p^\dagger$ summarizes a sensitivity curve by the threshold at which a conclusion is first overturned, here indexed by the deleted population share along the contrast direction.

For direct Wald inference on $p^\dagger$, suppose the population curve crosses the margin in the interior of the reporting interval: $\delta(\underline{p}) < c < \delta(\bar{p})$. By Lemmas 1 and 11, the crossing $p^\dagger \in (\underline{p}, \bar{p})$ is unique, satisfies $\delta(p^\dagger) = c$, and has strictly positive slope $\partial_p \delta(p^\dagger) = \xi_a^*(p^\dagger)/B_a^*(p^\dagger) > 0$.

Theorem 3 (Asymptotic normality of the breakdown fraction). *Under Assumption 1(b), Assumption S1, the interior-crossing condition $\delta(\underline{p}) < c < \delta(\bar{p})$, and $\sigma_g^2(p^\dagger) > 0$, the directional breakdown fraction satisfies*

$$\begin{aligned} \sqrt{N}(\hat{p}^\dagger - p^\dagger) &= -\frac{\sqrt{N}(\hat{g}(p^\dagger) - g(p^\dagger))}{\partial_p \delta(p^\dagger)} + o_p(1) \xrightarrow{d} \mathcal{N}(0, \sigma_\dagger^2), \\ \sigma_\dagger^2 &= \frac{\sigma_g^2(p^\dagger)}{(\partial_p \delta(p^\dagger))^2} = \frac{\text{Var}(\Psi_a(p^\dagger) - B_a^*(p^\dagger)\varphi_c)}{\xi_a^*(p^\dagger)^2}. \end{aligned}$$

A consistent plug-in standard error is

$$\widehat{\text{SE}}(\hat{p}^\dagger) = \frac{\widehat{\text{Var}}(\hat{\Psi}_a(\hat{p}^\dagger) - \hat{B}_a^*(\hat{p}^\dagger)\hat{\varphi}_c)^{1/2}}{\sqrt{N} \hat{\xi}(\hat{p}^\dagger)},$$

where $\hat{\xi}(\hat{p}^\dagger)$ is the empirical cutoff at $k(\hat{p}^\dagger)$, and the Wald interval $\hat{p}^\dagger \pm z_{1-\alpha/2} \widehat{\text{SE}}(\hat{p}^\dagger)$ has asymptotic coverage $1 - \alpha$.

The proof is in Section B. Notably, the retained-leverage divisor $B_a^*(p^\dagger)$ cancels between the gap score $h_i(p^\dagger) = \Psi_{a,i}(p^\dagger)/B_a^*(p^\dagger) - \varphi_{c,i}$ and the curve slope $\partial_p \delta(p^\dagger) = \xi_a^*(p^\dagger)/B_a^*(p^\dagger)$. Holding the numerator variance fixed, a larger marginal modified score at the cutoff, $\xi_a^*(p^\dagger) > 0$, steepens the fragility curve and identifies the breakdown fraction more precisely.

Band inversion. When the crossing lies near a boundary or no crossing is observed in the sample, inverting the simultaneous gap band of Proposition 2 remains valid without requiring an interior crossing. Writing $L_{g,N}(p)$ and $U_{g,N}(p)$ for the lower and

upper endpoints of the simultaneous gap band in Proposition 2, define the first-crossing bounds

$$\hat{p}_{\text{low}} := \inf\{p \in \mathcal{P} : U_{g,N}(p) \geq 0\}, \quad \hat{p}_{\text{up}} := \inf\{p \in \mathcal{P} : L_{g,N}(p) \geq 0\},$$

with the convention $\inf \emptyset = \bar{p}$. Whereas the point crossing $\hat{p}^\dagger$ can be found by bisection on k , computing $\hat{p}_{\text{low}}$ and $\hat{p}_{\text{up}}$ requires evaluating all integer budgets in $\mathcal{P}$ because studentized band endpoints need not be monotone (Section B.9).

Corollary 1 (Confidence interval by band inversion). *Under the conditions of Proposition 2,*

$$\liminf_{N \rightarrow \infty} \mathbb{P}\{\hat{p}_{\text{low}} \leq p^\dagger \leq \hat{p}_{\text{up}}\} \geq 1 - \alpha.$$

Moreover, on the simultaneous coverage event, $\sup_{p \in \mathcal{P}} U_{g,N}(p) < 0$ certifies that $\delta(p) < c$ throughout $\mathcal{P}$.

Proof. On the simultaneous coverage event of Proposition 2, $L_{g,N}(p) \leq g(p) \leq U_{g,N}(p)$ for all $p \in \mathcal{P}$. Hence, $\{L_{g,N} \geq 0\} \subseteq \{g \geq 0\} \subseteq \{U_{g,N} \geq 0\}$, and taking infima preserves the ordering $\hat{p}_{\text{low}} \leq p^\dagger \leq \hat{p}_{\text{up}}$, including when any of the sets is empty under $\inf \emptyset = \bar{p}$. Similarly, $\sup_{p \in \mathcal{P}} U_{g,N}(p) < 0$ implies $g(p) < 0$ for all $p \in \mathcal{P}$ on the same event. $\square$

Under the no-rotation condition of Section 3.3, with $\Sigma_p \succ 0$ and $\Sigma_p^{-1}a \propto Q^{-1}a$, a nonnegative lower gap band $L_{g,N}(p) \geq 0$ also certifies unrestricted fragility on the simultaneous coverage event, whereas $U_{g,N}(p) < 0$ certifies directional stability only.

4.3. Primitive conditions and practical reporting

The Supplement verifies the pointwise and uniform inference conditions for linear projections with continuous or mixed controls, including intercepts and nonlinear conditional means. The sufficient conditions impose standard moment and nonsingularity requirements, bounded positive conditional error scales, and positive continuous conditional error densities with suitable tail bounds. They cover Gaussian, logistic, and centered Student- t errors with more than two degrees of freedom. Global verification excludes population mass at the zero contrast index $x'Q^{-1}a = 0$, while a localized result permits such mass under an explicit deletion-budget restriction and local cutoff regularity. A complementary finite-covariate class requires nonzero contrast indices and positive continuous conditional outcome densities with finite second moments.

5. Simulations

For these classes, a fixed nonzero sign or threshold margin yields uniformly positive paired gap variance on compact deletion intervals. This supports gap-band inference and inversion; crossing Wald inference additionally requires an interior crossing. A calibrated comparison in the Supplement shows that populations with the same target contrast and conventional asymptotic variance can nevertheless have markedly different fragility curves and breakdown fractions.

Practical reporting. Applied reporting starts by specifying the target contrast $a'\beta$, the substantive conclusion margin $c > 0$, and a compact deletion range $\mathcal{P} = [\underline{p}, \bar{p}]$. Pointwise intervals apply to a single prespecified deletion share p , whereas simultaneous confidence bands cover the entire fragility curve over $\mathcal{P}$. When the conclusion margin $\hat{c}$ is estimated from the same sample, paired inference accounts for its covariance with the estimated curve. Band-inversion endpoints at the reporting limit are right-censored; stability throughout $\mathcal{P}$ is certified only when the upper gap band remains strictly negative. Finally, when i.i.d. sampling or score-cutoff regularity is not maintained, $\hat{\delta}(p)$ and the attained unrestricted OLS refit on $(S^*)^c$ can be reported together as deterministic sample-audit diagnostics.

5. Simulations

We assess finite-sample inference in two experiments. The first examines pointwise directional inference across four linear projection designs, varying tail behavior, heteroskedasticity, and asymmetry. The second evaluates simultaneous bands and breakdown inference when the conclusion margin is estimated from the same sample.

Simulation designs. For the pointwise experiment, let $x = (1, x_2, x_3)'$, $a = e_2$, and $p = 0.10$. We consider four designs, each satisfying $\mathbb{E}[xu] = 0$:

- (i) *Gaussian*: independent $\mathcal{N}(0, 1)$ regressors and error;
- (ii) *Elliptical- t* : bivariate Student- t_5 regressors and $\mathcal{N}(0, 1)$ error;
- (iii) *Skewed*: independent centered $\text{Exp}(1)$ regressors and centered χ_3^2 error;
- (iv) *Heteroskedastic*: independent $\mathcal{N}(0, 1)$ regressors and $u = \sqrt{1 + x_2^2} \varepsilon$, where $\varepsilon \sim \mathcal{N}(0, 1)$ is independent of x .

The first, second, and fourth designs are centrally symmetric, implying $g_a(p) = 0$, while the skewed design has a nonzero design correction. We use 5,000 replications at sample sizes ranging from 125 to 4,000. The Supplement provides further design details, verification of regularity conditions, and additional calibration experiments.

5.1. Consistency, normality, and pointwise coverage

Table 1 reports the pointwise results at $p = 0.10$. Bias is small relative to sampling variation, the scaled standard deviation stabilizes with sample size, and plug-in standard errors track empirical dispersion. Wald coverage is below nominal in smaller samples but approaches 95% across all four designs, ranging from 0.948 to 0.952 at $N = 4,000$. As predicted, the design correction has negligible effect under central symmetry but improves coverage by 1–2 percentage points in the skewed design.

TABLE 1: Pointwise inference at $p = 0.10$.

N	Bias	$\sqrt{N}$ SD	SE/SD	Coverage	Correction gain
<i>Gaussian</i>					
125	-0.010	0.446	1.027	0.913	+0.002
500	-0.001	0.471	0.990	0.940	0.000
2000	0.000	0.476	0.981	0.944	0.000
4000	0.000	0.461	1.012	0.952	0.000
<i>Elliptical-t</i>					
125	-0.011	0.467	1.009	0.902	+0.001
500	-0.001	0.497	0.996	0.943	0.000
2000	0.000	0.505	0.993	0.947	+0.001
4000	0.000	0.512	0.981	0.948	0.000
<i>Skewed</i>					
125	-0.009	1.821	0.935	0.892	+0.017
500	0.000	1.844	0.976	0.929	+0.015
2000	0.000	1.841	0.994	0.951	+0.013
4000	+0.001	1.825	1.009	0.948	+0.012
<i>Heteroskedastic</i>					
125	-0.027	0.959	0.960	0.867	-0.002
500	-0.004	1.025	0.965	0.924	-0.001
2000	-0.001	1.027	0.986	0.941	0.000
4000	0.000	1.013	1.002	0.948	0.000

Notes. Each row uses 5,000 replications. SE/SD is the mean estimated standard error relative to empirical dispersion; coverage refers to nominal 95% Wald intervals. Correction gain is the change in coverage from including the design correction. Coverage Monte Carlo standard errors are at most 0.005.

5.2. The fragility curve and breakdown fraction

We next examine paired inference in two asymmetric four-cell regression designs with normal or t_3 errors. The conclusion margin and its orientation are estimated from the same sample as the directional curve, making the dependence accounted for by the paired score in Proposition 2 relevant. We use 1,600 replications at $N \in \{200, 800, 3,200\}$ and evaluate simultaneous coverage over $p \in [0.02, 0.20]$.

Table 2 shows that the paired procedure delivers coverage close to nominal, although band coverage falls to 0.927 in the smallest heavy-tailed design. Coverage of the inverted breakdown intervals ranges from 0.948 to 0.967. These results support the use of paired inference for estimated margins, with some finite-sample deterioration under heavy tails. The Supplement reports implementation details, additional diagnostics, and fixed-margin functional calibration.

TABLE 2: Simultaneous gap bands and breakdown inference.

Law	N	Gap band	Breakdown interval	Crossing Wald	Wald availability
Normal	200	0.932	0.962	0.899	0.937
	800	0.936	0.956	0.937	0.999
	3 200	0.952	0.967	0.956	1.000
t_3	200	0.927	0.948	0.883	0.914
	800	0.943	0.964	0.942	0.993
	3 200	0.946	0.966	0.951	1.000

Notes. Nominal coverage is 95%; 1,600 replications per design and sample size. Gap-band coverage is evaluated over the full interval $p \in [0.02, 0.20]$. Breakdown intervals invert the paired band. Wald availability is the fraction of replications yielding a crossing-Wald interval; unavailable intervals count as noncoverage.

6. Empirical Illustrations

We illustrate the directional diagnostic in two settings: a deterministic sample audit comparing finite-share and first-order sensitivity, and a conditional population analysis with simultaneous inference.

6.1. A sample audit of institutions and income

We revisit the 64-country base sample of Acemoglu, Johnson, and Robinson (2001), using their Table 2, column 6 OLS specification: log GDP per capita in 1995 on pro-

tection against expropriation, absolute latitude, continent indicators and an intercept. The institutions slope is 0.4013. We compare its directional sensitivity to country deletions with the first-order AMIP approximation as a deterministic sample audit.

Figure 1 and Table 3 show that at three deletions, the directional and AMIP movements are 0.0833 and 0.0683, respectively (20.8% and 17.0% of the observed slope). Their paths diverge at larger budgets: the directional diagnostic first crosses the observed slope at 27 deletions (42.2%), while AMIP remains below it throughout the reported domain, up to 32 deletions. Re-estimating all six coefficients after the respective selected deletions gives institutions slopes of -0.0023 and 0.0774 at $k = 27$. These are attained refits, not unrestricted worst-case values; the crossing requires a substantial change in sample composition.

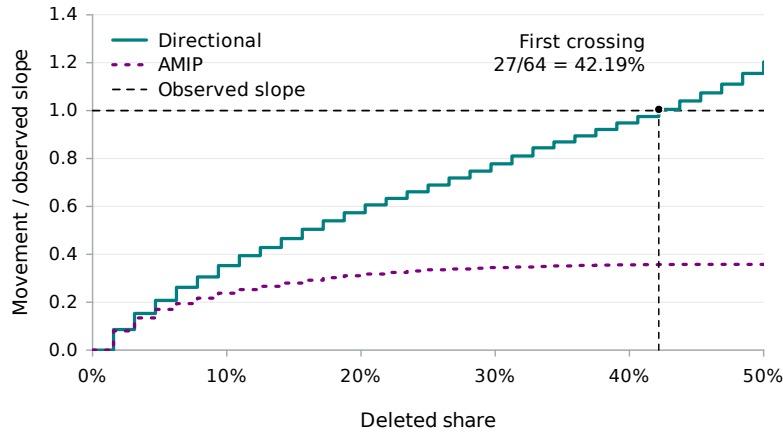


FIGURE 1: Sensitivity to country deletions with latitude and continent controls (AJR Table 2, column 6). Directional (solid teal) and AMIP (dashed purple) movements are divided by the observed institutions slope (0.4013); one marks the sign margin. The directional curve first crosses at 27/64 deletions, while AMIP remains below one through 32/64. Both curves are deterministic sample diagnostics.

6.2. Fragility curve and simultaneous bands in a cognitive-training trial

To illustrate population fragility curves and simultaneous confidence bands (Proposition 1 and corollary 1), we analyze the cognitive-training trial of Hardy et al. (2015) among $N = 5,045$ post-test completers. In an unadjusted regression of Grand-Index change on treatment assignment, the estimated difference between the Lumosity training group and the crossword-puzzle active control group is 3.0972 points (the

7. Conclusion

TABLE 3: Institutions and income with latitude and continent controls.

k	Deletions	Diagnostic movement		Refitted OLS slope	
	Share (%)	Directional	AMIP	Directional set	AMIP set
3	4.69	0.0833	0.0683	0.3104	0.3104
12	18.75	0.2301	0.1246	0.1621	0.1866
26	40.62	0.3913	0.1432	0.0024	0.0799
27	42.19	0.4033	0.1434	−0.0023	0.0774
32	50.00	0.4829	0.1435	−0.0442	0.0447

Notes. AJR Table 2, column 6: unweighted OLS of log GDP per capita on protection against expropriation, absolute latitude, continent indicators and an intercept in the 64-country base sample. The full-sample institutions slope is 0.4013; movements are in coefficient units. Each refit re-estimates all six coefficients after deleting the corresponding diagnostic-selected set; these are attained refits, not unrestricted optima. The directional curve first crosses the observed slope at 27/64 deletions; AMIP never crosses over $k = 0, \dots, 32$. The exercise is a deterministic sample audit and supplies no sampling inference.

standard error is 0.355). Maintaining i.i.d. sampling from a completer superpopulation and regular score cutoffs over $\mathcal{P} = [0.02, 0.20]$, we trace the directional sensitivity of this contrast against the fixed benchmark threshold $c = 3.0972$.

Figure 2 plots the retained contrast $c - \hat{\delta}(p)$ alongside 95% and 99% multiplier-bootstrap simultaneous confidence bands over the 2%–20% deletion range. Despite a t -statistic exceeding 8.7, the retained contrast crosses zero after removing 6.48% of the sample (327 of 5,045 participants). Inverting the 95% simultaneous curve band against the fixed benchmark $c = 3.0972$ yields a tight breakdown bracket of [6.16%, 6.78%]. Thus, a highly precise treatment-effect estimate can nonetheless be overturned by adversely removing roughly one in fifteen observations. In this illustration, $c = 3.0972$ is treated as a fixed reference threshold to display the curve band of Proposition 1; accounting simultaneously for sampling variation in the estimated baseline contrast itself uses the paired gap band of Proposition 2. The Supplement provides further details on the trial design, completer selection, differences from the published specification, and additional diagnostics.

7. Conclusion

Precision and compositional stability answer different questions. Standard errors describe sampling variation around a fixed population parameter, while directional fragility measures the largest adverse movement under population deletions when

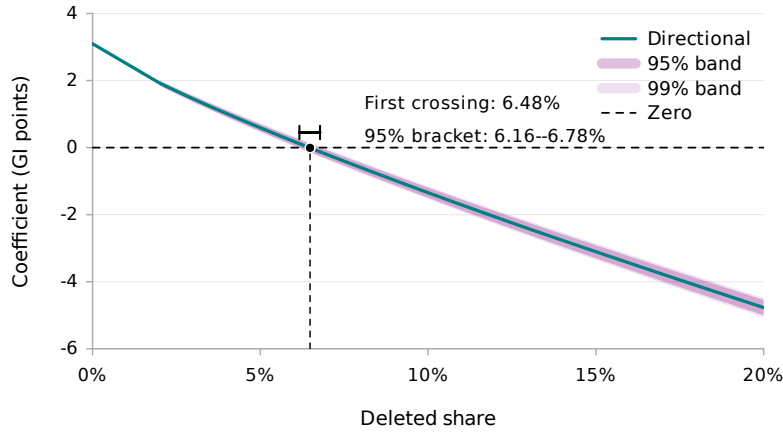


FIGURE 2: Directional sensitivity among post-test completers. The curve shows the observed treatment contrast (3.0972 points) minus directional movement, with 95% and 99% simultaneous bands over the 2–20% deletion range. The first sign crossing occurs at 6.48%; the 95% conditional crossing bracket is [6.16%, 6.78%].

refitting is restricted to the full-population contrast line. The first crossing of a conclusion margin gives the directional breakdown fraction. By reformulating the fixed-direction ratio problem as a scalar root backed by a superquantile threshold dual, we obtain an $O(N \log N)$ criterion evaluation that captures finite-share leverage depletion exactly and admits a tractable asymptotic expansion. Under i.i.d. sampling and score regularity near the deletion cutoff, we establish a functional central limit theorem for the empirical fragility curve that incorporates first-stage corrections for both generated residuals and the estimated contrast direction. These results yield pointwise confidence intervals, multiplier-bootstrap simultaneous bands, and valid inference for the breakdown fraction.

The directional target isolates selection and retained leverage along a fixed adjustment direction. Its relation to unrestricted refitting requires additional conditions; the institutions–income audit illustrates how a directional selection can instead be checked through an attained actual refit. The cognitive-training illustration highlights conditional curve-band and breakdown inference against a fixed benchmark threshold, complementing our paired gap bands for estimated margins. Primitive verifications establish applicability for stated continuous/mixed and finite-covariate classes under i.i.d. sampling, while extensions to dependent sampling and broader outcome distributions offer promising directions for further work.

References

- Acemoglu, D., S. Johnson, and J. A. Robinson (2001). “The Colonial Origins of Comparative Development: An Empirical Investigation”. In: *American Economic Review* 91.5, pp. 1369–1401. DOI: [10.1257/aer.91.5.1369](https://doi.org/10.1257/aer.91.5.1369).
- Armstrong, T. B. and M. Kolesár (2021). “Sensitivity Analysis Using Approximate Moment Condition Models”. In: *Quantitative Economics* 12.1, pp. 77–108. DOI: [10.3982/QE1609](https://doi.org/10.3982/QE1609).
- Azar, E., M. J. Feldman, and B. Nadler (2026). *Robustness of OLS to Sample Removals: Theoretical Analysis and Implications*. DOI: [10.48550/arXiv.2512.23069](https://doi.org/10.48550/arXiv.2512.23069). URL: <http://arxiv.org/abs/2512.23069> (visited on 09/22/2026). Pre-published.
- Bearpark, T., D. Hogan, and S. Hsiang (2025). “Data Anomalies and the Economic Commitment of Climate Change”. In: *Nature* 644.8075, E7–E11. DOI: [10.1038/s41586-025-09320-4](https://doi.org/10.1038/s41586-025-09320-4).
- Berrendero, J. R. and R. H. Zamar (2001). “Maximum Bias Curves for Robust Regression with Non-Elliptical Regressors”. In: *The Annals of Statistics* 29.1, pp. 224–251. DOI: [10.1214/aos/996986507](https://doi.org/10.1214/aos/996986507).
- Bonhomme, S. and M. Weidner (2022). “Minimizing Sensitivity to Model Misspecification”. In: *Quantitative Economics* 13.3, pp. 907–954. DOI: [10.3982/QE1930](https://doi.org/10.3982/QE1930).
- Christensen, T. and B. Connault (2023). “Counterfactual Sensitivity and Robustness”. In: *Econometrica* 91.1, pp. 263–298. DOI: [10.3982/ECTA17232](https://doi.org/10.3982/ECTA17232).
- Čížek, P. (2008). “General Trimmed Estimation: Robust Approach to Nonlinear and Limited Dependent Variable Models”. In: *Econometric Theory* 24.6, pp. 1500–1529. DOI: [10.1017/S0266466608080596](https://doi.org/10.1017/S0266466608080596).
- Cook, R. D. (1977). “Detection of Influential Observation in Linear Regression”. In: *Technometrics* 19.1, pp. 15–18. DOI: [10.1080/00401706.1977.10489493](https://doi.org/10.1080/00401706.1977.10489493).
- Cook, R. D. (1979). “Influential Observations in Linear Regression”. In: *Journal of the American Statistical Association* 74.365, pp. 169–174. DOI: [10.2307/2286747](https://doi.org/10.2307/2286747).
- Dewaskar, M. et al. (2025). “Robustifying Likelihoods by Optimistically Re-weighting Data”. In: *Journal of the American Statistical Association* 120.551, pp. 1787–1798. DOI: [10.1080/01621459.2025.2468012](https://doi.org/10.1080/01621459.2025.2468012).
- Dinkelbach, W. (1967). “On Nonlinear Fractional Programming”. In: *Management Science* 13.7, pp. 492–498. DOI: [10.1287/mnsc.13.7.492](https://doi.org/10.1287/mnsc.13.7.492).
- Donoho, D. L. and P. J. Huber (1983). “The Notion of Breakdown Point”. In: *A Festschrift for Erich L. Lehmann*. Ed. by P. J. Bickel, K. A. Doksum, and J. L. Hodges Jr. Belmont, CA: Wadsworth, pp. 157–184.
- Fisher, J. et al. (2023). “Influence Diagnostics under Self-Concordance”. In: *Proceedings of The 26th International Conference on Artificial Intelligence and Statistics*. International Conference on Artificial Intelligence and Statistics. PMLR, pp. 10028–10076.
- Freund, D. and S. B. Hopkins (2023). “Towards Practical Robustness Auditing for Linear Regression”. In: *ArXiv e-prints*. DOI: [10.48550/arXiv.2307.16315](https://doi.org/10.48550/arXiv.2307.16315).

- Giordano, R., R. Meager, and T. Broderick (2026a). “An Automatic Finite-Sample Robustness Metric: When Can Dropping a Little Data Change Conclusions? Part I: Definitions and Experiments”. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 384.2321, p. 20250001. DOI: [10.1098/rsta.2025.0001](https://doi.org/10.1098/rsta.2025.0001).
- (2026b). “An Automatic Finite-Sample Robustness Metric: When Can Dropping a Little Data Change Conclusions? Part II: Theory and Intuition”. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 384.2321, p. 20240614. DOI: [10.1098/rsta.2024.0614](https://doi.org/10.1098/rsta.2024.0614).
- Hampel, F. R. (1968). “Contributions to the Theory of Robust Estimation”. PhD thesis. Berkeley: University of California, Berkeley.
- (1974). “The Influence Curve and Its Role in Robust Estimation”. In: *Journal of the American Statistical Association* 69.346, pp. 383–393. DOI: [10.1080/01621459.1974.10482962](https://doi.org/10.1080/01621459.1974.10482962).
- Hardy, J. L. et al. (2015). “Enhancing Cognitive Abilities with Comprehensive Training: A Large, Online, Randomized, Active-Controlled Trial”. In: *PLOS ONE* 10.9, e0134467. DOI: [10.1371/journal.pone.0134467](https://doi.org/10.1371/journal.pone.0134467).
- Horowitz, J. L. and C. F. Manski (1995). “Identification and Robustness with Contaminated and Corrupted Data”. In: *Econometrica* 63.2, p. 281. DOI: [10.2307/2951627](https://doi.org/10.2307/2951627).
- Hu, Y. et al. (2024). “Most Influential Subset Selection: Challenges, Promises, and Beyond”. In: *Conference on Neural Information Processing Systems (NeurIPS 2024)*. Vol. 38. DOI: [10.48550/arXiv.2409.18153](https://doi.org/10.48550/arXiv.2409.18153).
- Huang, J. Y. et al. (2025). *Approximations to Worst-Case Data Dropping: Unmasking Failure Modes*. DOI: [10.48550/arXiv.2408.09008](https://doi.org/10.48550/arXiv.2408.09008).
- Huber, P. J. (1981). *Robust Statistics*. 1st ed. Wiley Series in Probability and Statistics. Wiley. DOI: [10.1002/0471725250](https://doi.org/10.1002/0471725250).
- Ichimura, H. and W. K. Newey (2022). “The Influence Function of Semiparametric Estimators”. In: *Quantitative Economics* 13.1, pp. 29–61. DOI: [10.3982/QE826](https://doi.org/10.3982/QE826).
- Jeong, S. and H. Namkoong (2020). “Robust causal inference under covariate shift via worst-case subpopulation treatment effects”. In: *Conference on Learning Theory (COLT 2020)*. 33rd Annual Conference on Learning Theory. Vol. 125. Proceedings of Machine Learning Research. PMLR, pp. 2079–2084. DOI: [10.48550/arxiv.2007.02411](https://doi.org/10.48550/arxiv.2007.02411).
- Kallus, N. (2023). “Treatment Effect Risk: Bounds and Inference”. In: *Management Science* 69.8, pp. 4579–4590. DOI: [10.1287/mnsc.2023.4819](https://doi.org/10.1287/mnsc.2023.4819).
- Kitamura, Y., T. Otsu, and K. Evdokimov (2013). “Robustness, Infinitesimal Neighborhoods, and Moment Restrictions”. In: *Econometrica* 81.3, pp. 1185–1201. DOI: [10.3982/ECTA8617](https://doi.org/10.3982/ECTA8617).
- Konrad, L. D. and N. Kuschnig (2026a). “Finding Most Influential Sets”. In: *Proceedings of the 43rd International Conference on Machine Learning (ICML)*. Vol. 306. Proceedings of Machine Learning Research. PMLR, pp. 60188–60203. DOI: [10.48550/arXiv.2606.05919](https://doi.org/10.48550/arXiv.2606.05919).
- (2026b). “Testing Most Influential Sets”. In: *Proceedings of the 14th International Conference on Learning Representations (ICLR)*. DOI: [10.48550/arXiv.2510.20372](https://doi.org/10.48550/arXiv.2510.20372).

References

- Kotz, M., A. Levermann, and L. Wenz (2024). “The Economic Commitment of Climate Change”. In: *Nature* 628.8008. Retracted, pp. 551–557. DOI: [10.1038/s41586-024-07219-0](https://doi.org/10.1038/s41586-024-07219-0).
- (2025). “Retraction Note: The Economic Commitment of Climate Change”. In: *Nature* 648.8094, p. 764. DOI: [10.1038/s41586-025-09726-0](https://doi.org/10.1038/s41586-025-09726-0).
- Kuschnig, N., G. Zens, and J. Crespo Cuaresma (2021). “Hidden in Plain Sight: Influential Sets in Linear Regression”. In: *CESifo Working Paper* 8981. DOI: [10.13140/RG.2.2.30682.47042](https://doi.org/10.13140/RG.2.2.30682.47042).
- Lawrance, A. J. (1995). “Deletion Influence and Masking in Regression”. In: *Journal of the Royal Statistical Society: Series B (Methodological)* 57.1, pp. 181–189. DOI: [10.1111/j.2517-6161.1995.tb02023.x](https://doi.org/10.1111/j.2517-6161.1995.tb02023.x).
- Masten, M. A. and A. Poirier (2020). “Inference on Breakdown Frontiers”. In: *Quantitative Economics* 11.1, pp. 41–111. DOI: [10.3982/QE1288](https://doi.org/10.3982/QE1288).
- Moitra, A. and D. Rohatgi (2023). “Provably Auditing Ordinary Least Squares in Low Dimensions”. In: *International Conference on Learning Representations (ICLR 2023)*. DOI: [10.48550/arXiv.2205.14284](https://doi.org/10.48550/arXiv.2205.14284).
- O’Connell, S. A. and C. Smith (2016). “Economic Growth and Child Undernutrition”. In: *The Lancet Global Health* 4.12, e901–e902. DOI: [10.1016/S2214-109X\(16\)30250-9](https://doi.org/10.1016/S2214-109X(16)30250-9).
- Rockafellar, R. T. and S. Uryasev (2000). “Optimization of Conditional Value-at-Risk”. In: *The Journal of Risk* 2.3, pp. 21–41. DOI: [10.21314/jor.2000.038](https://doi.org/10.21314/jor.2000.038).
- Rubinstein, I. and S. B. Hopkins (2024). “Robustness Auditing for Linear Regression: To Singularity and Beyond”. In: *ArXiv e-prints*. DOI: [10.48550/arXiv.2410.07916](https://doi.org/10.48550/arXiv.2410.07916).
- Spini, P. E. (2026). *A Simple Measure of Robustness for External Validity under Covariate Shifts*. Pre-published.
- Tsirel’son, V. S. (1976). “The Density of the Distribution of the Maximum of a Gaussian Process”. In: *Theory of Probability & Its Applications* 20.4, pp. 847–856. DOI: [10.1137/1120092](https://doi.org/10.1137/1120092).
- van der Vaart, A. W. and J. A. Wellner (1996). *Weak Convergence and Empirical Processes: With Applications to Statistics*. Springer Series in Statistics. Springer. DOI: [10.1007/978-1-4757-2545-2](https://doi.org/10.1007/978-1-4757-2545-2).
- Vollmer, S. et al. (2014). “Association between Economic Growth and Early Childhood Undernutrition: Evidence from 121 Demographic and Health Surveys from 36 Low-Income and Middle-Income Countries”. In: *The Lancet Global Health* 2.4, e225–e234. DOI: [10.1016/S2214-109X\(14\)70025-7](https://doi.org/10.1016/S2214-109X(14)70025-7).

Section A develops the directional threshold engine, population target, pointwise expansion and plug-in formulas; Section B proves the functional limit, multiplier validity and crossing results. The Supplement collects primitive regression verifications, the unrestricted-refitting bridge and additional numerical detail.

A. Asymptotic theory

Section A.1 reduces the combinatorial selection to a scalar Dinkelbach root and a top- k threshold criterion. Section A.2 establishes the directional population root, and Section A.3 derives the directional sampling expansion, including the generated-residual and design-direction corrections.

A.1. The Dinkelbach–threshold reformulation

The directional statistic is a maximum over $\binom{N}{k}$ subsets. The Dinkelbach root and the top- k threshold representation reduce it to a scalar root and an empirical-average criterion, the input to the directional asymptotic theory below. In practice r_a and u_i are replaced by $\hat{r}_a = \hat{Q}_N^{-1}a$ and $\hat{u}_i$, and it is convenient not to normalize the denominator inside the optimization loop: with $\hat{v}_i := a' \hat{Q}_N^{-1} x_i$, the *computational scalarized statistic* is

$$\hat{\delta}^{\text{sc}} := \max_{S \in \mathcal{S}} \frac{\sum_{i \in S} \hat{v}_i \hat{u}_i}{\sum_{i \notin S} \hat{v}_i^2}. \quad (18)$$

Since $\hat{\mu}_Z(a) > 0$ is common to every subset, maximizing and rescaling gives

$$\hat{\delta} = \hat{\mu}_Z(a) \hat{\delta}^{\text{sc}}. \quad (19)$$

Thus $\hat{\delta}^{\text{sc}}$ is only a computational normalization.

For $p \in (0, 1)$, write $\tau := 1 - p$. Fix an i.i.d. pair (Y_i, Z_i) with $Z_i \geq 0$, and write

$$Y_S = \sum_{i \in S} Y_i, \quad Z_{-S} = \sum_{i \notin S} Z_i, \quad W_i(\lambda) := Y_i + \lambda Z_i, \quad F_N(\lambda) := \max_{S \in \mathcal{S}} (Y_S - \lambda Z_{-S}),$$

with $\hat{\delta} := \max_{S \in \mathcal{S}} Y_S / Z_{-S}$ on the event $Z_{-S} > 0$. In the scalar engine, δ and $\hat{\delta}$ denote the population and sample roots for the specified generic pair (Y, Z) ; the normalized directional pair in (3) recovers the directional target and its oracle statistic. This is the form of $\hat{\delta}^{\text{sc}}$ of (18) and $\hat{\delta}$ of (4): freezing the metric at $\hat{Q}_N$ leaves the fixed

quadratic-form denominator $\sum_{i \notin S} \hat{v}_i^2$, and $\hat{\delta} = \hat{\mu}_Z(a) \hat{\delta}^{\text{sc}}$. The maximization is a fractional program which the Dinkelbach transform converts into a root-finding problem, and the threshold representation below then exposes the empirical-average structure that drives the asymptotics. With $\mu_Z := \mathbb{E}[Z]$, the normalized directional pair in (3) has $\mu_Z = 1$.

We maintain the following throughout. Let $F_V^{-1}(u) := \inf\{t : F_V(t) \geq u\}$ denote the left-continuous quantile function of a random variable V . The next three assumptions give a well-defined statistic with probability tending to one and a positive population root: i.i.d. data with finite second moments, positive retained leverage and a positive upper-tail numerator.

Assumption 2 (Data-generating process). *(Y_i, Z_i) are i.i.d. with common distribution; $Z_i \geq 0$ a.s.; and $\mathbb{E}[Y^2] < \infty$, $\mathbb{E}[Z^2] < \infty$.*

For the normalized directional pair in (3), $\mu_Z := \mathbb{E}[Z] = 1$; in general, the scalar engine uses $\mu_Z = \mathbb{E}[Z]$.

Assumption 3 (Denominator non-degeneracy). *With $\mu_Z := \mathbb{E}[Z]$, assume $\mu_Z > 0$ and $F_Z^{-1}(\tau) > 0$. Since $Z \geq 0$ by Assumption 2, $F_Z^{-1}(\tau) > 0$, $\int_0^\tau F_Z^{-1}(u) du > 0$, and $P(Z = 0) < \tau$ are equivalent; the quantile condition also implies $\mu_Z > 0$.*

Assumption 4 (Positive upper-tail numerator). $\mu_p^Y := \int_\tau^1 F_Y^{-1}(u) du > 0$.

For the directional pair these reduce to mild conditions on the design and the score; e.g. Assumption 3 holds whenever $P(r'_a x = 0) < \tau$.

Proposition 3 (Fractional representation). *On the event $\mathcal{D}_N := \{Z_{-S} > 0 \text{ for all } S \in \mathcal{S}\}$:*

- (i) F_N is finite, continuous, convex, and strictly decreasing;
- (ii) $\hat{\delta} = \max_{S \in \mathcal{S}} Y_S / Z_{-S}$ is the unique root of $F_N(\lambda) = 0$;
- (iii) any maximizer $S^* \in \arg \max_S Y_S / Z_{-S}$ also solves $S^* \in \arg \max_S \sum_{i \in S} W_i(\hat{\delta})$, hence consists of the k observations with the largest values of $W_i(\hat{\delta})$, ties at the cutoff being allocated arbitrarily (a fixed convention, such as the break-by-index rule of Section A.3.8, selects one maximizer); conversely, every such top- k set attains the ratio $\hat{\delta}$.

Moreover $\mathbb{P}(\mathcal{D}_N) \rightarrow 1$ under Assumption 3.

Remark 1 (Population atoms and empirical tie counts). *If $P(W_a^*(p) = \xi_a^*(p)) > 0$ and $Z(a)$ varies within the cutoff cell, different fractional tie allocations can give different retained slopes and influence variances; the Gaussian limit theorems do not apply. Empirical tie counts among fitted scores depend on sampling and numerical precision and neither diagnose nor rule out population atoms.*

Lemma 2 (Top- k threshold representation). *For reals $V_1, \dots, V_N$ and $1 \leq k \leq N$,*

$$\sum_{j=1}^k V_{(N-j+1)} = \min_{\xi \in \mathbb{R}} \left\{ \sum_{i=1}^N (V_i - \xi)_+ + k\xi \right\},$$

with minimizers $\{\xi : \#\{i : V_i > \xi\} \leq k \leq \#\{i : V_i \geq \xi\}\}$.

The population counterpart of Lemma 2 is the Rockafellar–Uryasev dual representation (Rockafellar and Uryasev, 2000):

$$\mathcal{T}_p(V) = \min_{\xi \in \mathbb{R}} \{ \mathbb{E}(V - \xi)_+ + p\xi \}. \quad (20)$$

Thus $\mathcal{T}_p(V)/p$ is the upper- p tail mean.

The empirical formula is (20) at $p = k/N$, multiplied by N ; its subgradient gives the stated minimizer interval. Applying Lemma 2 to $V_i = W_i(\lambda)$ and normalizing defines

$$\Gamma_N(\lambda, \xi) := \frac{1}{N} \sum_{i=1}^N [(Y_i + \lambda Z_i - \xi)_+ + p_N \xi - \lambda Z_i], \quad f_N(\lambda) := \frac{F_N(\lambda)}{N} = \min_{\xi \in \mathbb{R}} \Gamma_N(\lambda, \xi). \quad (21)$$

A.2. Population targets

The population criterion replaces the empirical average by expectation. With $W(\lambda) = Y + \lambda Z$, define $\Gamma(\lambda, \xi) := \mathbb{E}[(W(\lambda) - \xi)_+ + p\xi - \lambda Z]$ and $f(\lambda) := \inf_{\xi} \Gamma(\lambda, \xi)$. Equivalently,

$$f(\lambda) = \mathcal{T}_p(Y + \lambda Z) - \lambda \mu_Z, \quad f_N(\lambda) = \frac{1}{N} \sum_{j=1}^k W_{(N-j+1)}(\lambda) - \lambda \bar{Z}_N,$$

where $\bar{Z}_N = N^{-1} \sum_i Z_i$. The root is the relaxed population value, and equals the region target under the cutoff condition in Proposition 4; its local expansion under the regularity conditions below gives the influence function (Section A.3).

Lemma 3 (Population threshold characterization). *For fixed λ , ξ minimizes $\Gamma(\lambda, \cdot)$ iff $P(W(\lambda) > \xi) \leq p \leq P(W(\lambda) \geq \xi)$, equivalently $F_{W(\lambda)}(\xi-) \leq \tau \leq F_{W(\lambda)}(\xi)$. The minimizer is unique when the τ -quantile is unique (e.g. $W(\lambda)$ has a positive density near that quantile), and is then denoted $\xi(\lambda) = F_{W(\lambda)}^{-1}(\tau)$.*

This is Rockafellar and Uryasev (2000, Thm. 1): the minimizer set of the dual (20) is the interval of $(1 - p)$ -quantiles of $W(\lambda)$. We omit the proof.

The threshold is not needed to establish existence or uniqueness of the population root in Proposition 4; the tail-functional inequality holds even at a tied cutoff. The unique-cutoff condition below is used by subsequent cutoff-based arguments and to define the score ψ in the central limit theorem.

Assumption 5 (Population threshold regularity). *For every $\lambda \geq 0$, $W(\lambda) = Y + \lambda Z$ has a unique τ -quantile $\xi(\lambda) = F_{W(\lambda)}^{-1}(\tau)$ with $P(W(\lambda) \geq \xi(\lambda)) = p$. (A sufficient condition on compact λ -sets is a density positive and continuous near $F_{W(\lambda)}^{-1}(\tau)$, uniformly over the set.)*

Proposition 4 (Existence and uniqueness of the population root). *Under Assumptions 2 to 4, $f(\lambda) = \mathcal{T}_p(Y + \lambda Z) - \lambda \mu_Z$ is continuous and strictly decreasing on $[0, \infty)$, with $f(0) = \mu_p^Y > 0$ and $\lim_{\lambda \rightarrow \infty} f(\lambda) = -\infty$. Hence it has a unique root $\delta \in (0, \infty)$. It is the relaxed ratio value*

$$\delta = \sup_{0 \leq \pi \leq 1, \mathbb{E}\pi = p} \frac{\mathbb{E}[Y\pi]}{\mathbb{E}[Z(1 - \pi)]}.$$

If at the root there is a cutoff ξ with $P(W(\delta) \geq \xi) = p$, the supremum is attained by the region $D = \{W(\delta) \geq \xi\}$ and equals the supremum over regions of mass p . In particular, for $(Y(a), Z(a))$ with $\mu_Z = 1$, this root equals δ under Assumption 5.

Proof. We have $f(0) = \mathcal{T}_p(Y) = \mu_p^Y > 0$, and by Assumption 2 and Assumption 3, $c_Z := \mu_Z - \mathcal{T}_p(Z) = \int_0^\tau F_Z^{-1}(u) du > 0$. Subadditivity and positive homogeneity of $\mathcal{T}_p$ imply, for $0 \leq \lambda_1 < \lambda_2$,

$$f(\lambda_2) \leq f(\lambda_1) - (\lambda_2 - \lambda_1)c_Z.$$

Thus f is strictly decreasing and tends to $-\infty$. The bound $|\mathcal{T}_p(V) - \mathcal{T}_p(B)| \leq \mathbb{E}|V - B|$ gives $|f(\lambda_1) - f(\lambda_2)| \leq 2\mu_Z|\lambda_1 - \lambda_2|$, so f is continuous. The intermediate value theorem gives a unique positive root. For any feasible intensity, $\mathbb{E}[Z(1 - \pi)] \geq c_Z > 0$, and the upper-tail allocation identity gives

$$\mathbb{E}[Y\pi] - \delta \mathbb{E}[Z(1 - \pi)] = \mathbb{E}[W(\delta)\pi] - \delta \mu_Z \leq f(\delta) = 0.$$

The upper-tail intensity $\pi^* = \mathbf{1}\{W(\delta) > \xi\} + \alpha \mathbf{1}\{W(\delta) = \xi\}$, where ξ is a $(1-p)$ -quantile of $W(\delta)$ and $\alpha \in [0, 1]$ makes $\mathbb{E}\pi^* = p$, attains equality, so its ratio is δ . If $P(W(\delta) \geq \xi) = p$, one can take $\pi^* = \mathbf{1}\{W(\delta) \geq \xi\}$, which proves the region identification. $\square$

Proposition 5 (Relation to the first-order approximation). *For the directional pair, assume $\mathbb{E}[Z(a)] = 1$ and let $\delta \in [0, \infty)$ satisfy $\mathcal{T}_p(Y(a) + \delta Z(a)) = \delta$. If $\mathcal{T}_p(Z(a)) < 1$, then*

$$\mathcal{T}_p(Y(a)) \leq \delta \leq \frac{\mathcal{T}_p(Y(a))}{1 - \mathcal{T}_p(Z(a))}.$$

The lower bound $\mathcal{T}_p(Y(a))$ is the first-order approximation to the maximum deletion effect: its sample analogue is $N^{-1} \sum_{j=1}^k \hat{s}_{(N-j+1)}(a)$, the normalized sum of the k largest directional scores with the design held fixed.

Proof. Since $Z(a) \geq 0$ and $\delta \geq 0$, monotonicity gives $\delta = \mathcal{T}_p(Y + \delta Z) \geq \mathcal{T}_p(Y)$. Subadditivity and positive homogeneity give $\delta \leq \mathcal{T}_p(Y) + \delta \mathcal{T}_p(Z)$. Because $1 - \mathcal{T}_p(Z) > 0$, rearrangement yields the upper bound. $\square$

The upper-to-lower ratio tends to one as $p \rightarrow 0$ whenever $\mathcal{T}_p(Y(a)) > 0$, since $\mathcal{T}_p(Z(a)) \rightarrow 0$.

Throughout, $W_a^* := Y(a) + \delta Z(a)$ denotes the modified score at the directional population root and $\xi_a^* := F_{W_a^*}^{-1}(1-p)$ its $(1-p)$ -quantile. These define the active set $A^* = \{W_a^* \geq \xi_a^*\}$ and the retained slope $B_a^* = 1 - \nu_{a,Z}^* = \mathbb{E}[Z(a) \mathbf{1}\{A^{*,c}\}]$ of (6).

A.3. Directional expansion

We expand f_N around the population root to obtain the oracle influence function, then add the first-stage corrections for estimating $\hat{\beta}$ and, when $d \geq 2$, $\hat{r}_a$.

A.3.1. Regularity

Assumption 6 (CLT regularity). *In addition to Assumption 2–Assumption 5, impose conditions C1–C3 and C6 below; the intermediate properties C4 (local equicontinuity) and C5 (design differentiability) follow from C1–C3 and are established in Lemma 5.*

(C1) **Moments and design.** $\mathbb{E}[Y^2] < \infty$, $\mathbb{E}[Z^2] < \infty$; in the multivariate case the model (1) holds with $\mathbb{E}[xu] = 0$, $0 < \mathbb{E}u^2 < \infty$, $\mathbb{E}[u^2 | x] = \sigma^2(x)$ with $\mathbb{E}[\|x\|^2 \sigma^2(x)] < \infty$, $Q > 0$, $\mathbb{E} \|x\|^4 < \infty$, $a \neq 0$ fixed; independence of x and u is not required.

(C2) **Quantile regularity.** $W = Y + \delta Z$ has a density f_W positive and continuous at $\xi = F_W^{-1}(\tau)$, $0 < f_W(\xi) < \infty$; in the directional case the same density condition holds for $W(\lambda)$ near δ , ensuring uniqueness of the cutoff. We additionally assume that $\lambda \mapsto \xi(\lambda)$ is C^1 near δ . For the stated results, only local upper and lower bounds on the probability mass immediately on each side of the cutoff enter the proofs; continuity of the density is not otherwise used.

(C3) **Nonzero denominator slope.** $B = \mu_Z - \nu_p^Z > 0$ (directionally $B_a^* > 0$).

(C6) **Non-degenerate influence variance.** $\text{Var}(\Psi_a) > 0$, so that $\omega_a > 0$ in Theorem 1; in the oracle or scalar case, $\text{Var}(\psi) > 0$ in Theorem 4.

The independent pointwise inputs are C1–C3 and C6, in addition to the standing assumptions. Lemma 5 derives local criterion equicontinuity and population direction differentiability, retaining the names C4–C5 for these properties. Theorem 5 computes the gradient; it vanishes under the sufficient conditions in Corollary 2 and remark 4. The functional appendix separately requires U1–U2; the fixed-budget implication does not establish uniform-in-budget regularity. Condition C6 is required for nondegenerate Wald inference, not for the expansions. The Online Supplement states and proves primitive continuous/mixed-control and finite-covariate verifications of these conditions.

Two standing facts are used throughout. The envelope $H := 1 + \|x\| |u| + \|x\|^2$ is square-integrable at every dimension under the model of Section 2— $\|x\|^2 \in L^2$ is $\mathbb{E} \|x\|^4 < \infty$, and $\|x\| |u| \in L^2$ is $\mathbb{E} [\|x\|^2 \sigma^2(x)] < \infty$ —and every moment bound below is dominated by H or H^2 . Second, writing $\hat{\theta}_2 := (\hat{\beta}, \hat{r}_a)$ and $\theta_0 := (\beta, r_a)$, the normal equation and (1) give $\hat{\beta} - \beta = \hat{Q}_N^{-1} N^{-1} \sum_i x_i u_i$ and $\hat{r}_a - r_a = -\hat{Q}_N^{-1} (\hat{Q}_N - Q) r_a$ with $\hat{Q}_N := N^{-1} \sum_i x_i x_i'$; since $\mathbb{E}[xu] = 0$ holds by construction, (2), and $\mathbb{E} \|x\|^2 u^2 = \mathbb{E} [\|x\|^2 \sigma^2(x)] < \infty$, while $N^{-1} \sum_i \|x\|_i^4 = O_p(1)$ and $Q \succ 0$, both averages are $O_p(N^{-1/2})$ and

$$\begin{aligned} \|\hat{\theta}_2 - \theta_0\| &= O_p(N^{-1/2}), \\ \sqrt{N}(\hat{\beta} - \beta) &= Q^{-1} N^{-1/2} \sum_i x_i u_i + o_p(1), \\ \sqrt{N}(\hat{r}_a - r_a) &= -Q^{-1} N^{-1/2} \sum_i (x_i x_i' - Q) r_a + o_p(1). \end{aligned} \tag{22}$$

A.3.2. Shared class bounds and nuisance notation

Lemma 4 (Compact parameter meshes). *Let $\Theta \subset \mathbb{R}^m$ be compact and $(\theta, w) \mapsto h_\theta(w)$ jointly measurable, with*

$$|h_\theta| \leq M, \quad |h_\theta - h_\eta| \leq M \|\theta - \eta\|, \quad M \in L^2(P).$$

The class is pointwise measurable and P -Donsker, with polynomial $L^2(P)$ bracketing numbers. For a constant C_Θ , its empirical covering numbers satisfy, simultaneously for every $\varepsilon > 0$,

$$\log \mathcal{N}(\varepsilon, \{h_\theta : \theta \in \Theta\}, \|\cdot\|_{N,2}) \leq m \log(1 + C_\Theta \|M\|_{N,2} / \varepsilon), \quad \|h\|_{N,2}^2 := P_N h^2.$$

Squares and pairwise products of this family obey uniform laws of large numbers.

Proof. A countable dense parameter set gives pointwise measurability by the Lipschitz inequality. Euclidean nets of radius ε yield brackets $h_\theta \pm \varepsilon M$, giving polynomial bracketing and the Donsker conclusion (van der Vaart and Wellner, 1996, Theorem 2.7.11). Independently, nets of radius $\varepsilon / \|M\|_{N,2}$ give the displayed empirical covers; if this norm is zero the empirical class is a singleton. The continuous image of the compact parameter set in $\mathbb{R}^N$ is a measurable random compact set, so these covering numbers are measurable. For products, the same finite nets have errors bounded by a constant times εM^2 ; the law of large numbers and $PM^2 < \infty$ give their uniform laws. $\square$

For the regression calculations write $\theta = (b, r)$, $\theta_0 = (\beta, r_a)$, $v = r'_a x$ and $c_0 = a' r_a = \mu_Z(a)$. On compact neighborhoods with $a' r \geq c_0/2$, define

$$Y_\theta = (r'x)(y - x'b), \quad Z_\theta = \frac{(r'x)^2}{a'r}, \quad V_\theta(\lambda) = Y_\theta + \lambda Z_\theta. \quad (23)$$

At θ_0 their derivative coordinates are

$$D_b Y_{\theta_0} = -vx', \quad D_r Y_{\theta_0} = xu, \quad D_r Z_{\theta_0} = \frac{2vx}{c_0} - \frac{v^2 a}{c_0^2}. \quad (24)$$

The criterion coefficients, including the derivative of the full-sample subtraction, are

$$\begin{aligned}\mathbb{E}[D_b Y_{\theta_0} \mathbf{1}\{A^*\}] &= -m'_a, \\ D_a &:= \mathbb{E}[(D_r Y_{\theta_0} + \delta D_r Z_{\theta_0}) \mathbf{1}\{A^*\}] - \delta \mathbb{E}[D_r Z_{\theta_0}] \\ &= g_{A^*} + 2\delta^{\text{sc}} m_a - \delta^{\text{sc}} (\nu_{a,Z}^* + 1)a, \quad \delta^{\text{sc}} = \delta/c_0.\end{aligned}\tag{25}$$

Here $\mathbb{E}[D_r Z_{\theta_0}] = a/c_0$; since $B_a^* = 1 - \nu_{a,Z}^*$, D_a/B_a^* is the expression in (8). Its identification as the population gradient is proved below, separately from this algebra.

A.3.3. Derived pointwise regularity

Lemma 5 (Pointwise equicontinuity and direction differentiability). *Under Assumption 2–Assumption 5 and C1–C3, the following properties hold at the fixed budget p :*

(C4) **Local equicontinuity.** *For every fixed $M > 0$,*

$$\sup_{|h| \leq M} \left| \sqrt{N} \{ (f_N - f)(\delta + h/\sqrt{N}) - (f_N - f)(\delta) \} \right| \xrightarrow{p} 0.$$

(C5) **Design regularity.** *In regression, Φ_a of (28) is differentiable at r_a . Write $t = (b, r, \lambda)$, $t_0 = (\beta, r_a, \delta)$ and $F_N(t) = f_{N,(b,r)}(\lambda)$ for the nuisance-indexed criterion of (23), with $F(t)$ its population counterpart. For every deterministic $\varepsilon_N \downarrow 0$,*

$$\sup_{\|t - t_0\| \leq \varepsilon_N} \left| \sqrt{N} \{ (F_N - F)(t) - (F_N - F)(t_0) \} \right| \xrightarrow{p} 0.\tag{26}$$

Proof. An indexed hinge class. Using (23) on a compact neighborhood with $a'r \geq c_0/2$, write $Y_{b,r} = Y_\theta$, $Z_r = Z_\theta$ and put

$$g_{t,\xi} = (Y_{b,r} + \lambda Z_r - \xi)_+ + p\xi - \lambda Z_r, \quad F(t) = \inf_{\xi} \mathbb{E} g_{t,\xi}, \quad F_N(t) = \inf_{\xi} \{ P_N g_{t,\xi} + (p_N - p)\xi \}.$$

Here P_N is the empirical average and $p_N = \lfloor pN \rfloor / N$. For bounded λ, ξ , the envelope and parameter Lipschitz modulus are bounded by $CH \in L^2$:

$$|g_{t,\xi} - g_{s,\eta}| \leq CH(\|t - s\| + |\xi - \eta|).$$

This follows from $y - x'b = u - x'(b - \beta)$, the positive lower bound on $a'r$ and contraction

of the positive-part map. Lemma 4 gives a pointwise measurable Donsker class. For the general scalar pair use $g_{\lambda,\xi} = (Y + \lambda Z - \xi)_+ + p\xi - \lambda Z$ instead; its envelope and Lipschitz modulus are bounded by $C(1 + |Y| + Z) \in L^2$. No regression representation is needed in this case.

Profiling. The score envelope and strict tail-count bounds localize all population and empirical threshold minimizers in a deterministic compact interval, with probability tending to one for the latter, as in the threshold localization argument of Proposition 6. Let ξ_t and $\hat{\xi}_t$ be population and empirical minimizers. For every shrinking neighborhood of t_0 , all these minimizers converge uniformly to $\xi_0 = \xi_a^*$: the base population objective has a strictly positive gap outside any fixed neighborhood of its unique minimizer, and uniform continuity and the uniform law of large numbers preserve that gap. This uses uniqueness at the base point only, not a density or a quantile rate at fitted directions. With $\mathbb{G}_N = \sqrt{N}(P_N - P)$, the profiling correction obeys

$$(P_N - P)(g_{t,\hat{\xi}_t} - g_{t,\xi_t}) + (p_N - p)(\hat{\xi}_t - \xi_t) \leq R_N(t) := \Gamma_N(t, \hat{\xi}_t) - \Gamma_N(t, \xi_t) \leq 0,$$

where $\Gamma_N(t, \xi) = P_N g_{t,\xi} + (p_N - p)\xi$. The two inequalities are population and empirical optimality, respectively. Donsker equicontinuity and uniform threshold convergence make $\sqrt{N}R_N(t) = o_p(1)$ uniformly on shrinking neighborhoods; rounding contributes only $O(N^{-1/2})$ on the compact threshold interval. Consequently

$$\sqrt{N}\{F_N(t) - F(t)\} = \mathbb{G}_N g_{t,\xi_t} + o_p(1)$$

uniformly there. The L^2 distance from g_{t,ξ_t} to g_{t_0,ξ_0} tends uniformly to zero, proving (26). The identical scalar argument, restricted to $\lambda = \delta + h/\sqrt{N}$, proves C4 over its full scalar scope. Endpoint quantiles give measurable minimizer choices; profiled values are continuous by value contraction, so suprema can be taken over countable dense parameter sets.

Population direction differentiability. Fix $b = \beta$, let $q = (r, \lambda)$, $q_0 = (r_a, \delta)$, $V_q = Y_{\beta,r} + \lambda Z_r$, and $A_0 = \{V_{q_0} \geq \xi_0\}$. Use the tail-value representation $\mathcal{T}_p(V) = \sup_{0 \leq w \leq 1, \mathbb{E}w=p} \mathbb{E}[Vw]$; fractional cutoff weights permit neighboring atoms without changing the value. At the base point the maximizing weight is uniquely $w_0 = \mathbf{1}\{A_0\}$ almost surely. All neighboring quantiles tend to ξ_0 by the same compact argmin argument. Since $|V_q - V_{q_0}| \leq C \|q - q_0\| H$, every maximizing weight w_q differs from w_0 only in a shrink-

ing base-cutoff band or where H is large. Truncation and cutoff atomlessness give $\mathbb{E}[H|w_q - w_0|] \rightarrow 0$. Thus the population sandwich

$$\mathbb{E}[(V_q - V_{q_0})w_0] \leq \mathcal{T}_p(V_q) - \mathcal{T}_p(V_{q_0}) \leq \mathbb{E}[(V_q - V_{q_0})w_q]$$

has gap $o(\|q - q_0\|)$. The score map has an L^1 Taylor expansion, so $F(\beta, r, \lambda)$ is differentiable at the base point with $\partial_\lambda F = -B_a^*$ and

$$D_r F = D_a.$$

The standing denominator assumption gives $b_0 = \int_0^\tau F_{Z_{r_a}}^{-1}(s) ds > 0$. Every nearby retained weight of mass τ has leverage at least $b_0 - \mathbb{E}|Z_r - Z_{r_a}| \geq b_0/2$. Hence the population criterion decreases in λ by at least $b_0/2$; continuity preserves the opposite signs at two fixed points bracketing δ , giving a unique nearby root $\lambda(r)$. Evaluating the criterion at the base root first gives $|\lambda(r) - \delta| = O(\|r - r_a\|)$. The joint differentiability just proved then gives

$$\lambda(r) - \delta = B_a^{*-1} D'_a(r - r_a) + o(\|r - r_a\|).$$

Substituting $\lambda = (a'r)\delta^{\text{sc}}$ in the root equation gives exactly (27); thus $\lambda(r) = \Phi_a(r)$ and $\nabla \Phi_a(r_a) = D_a/B_a^*$. The full subtraction $-\lambda \mathbb{E}Z_r = -\lambda r' Q r / (a'r)$ is essential off the truth. Neither this pointwise root argument nor the empirical argument uses C4–C5, the oracle CLT or uniform-in-budget assumptions as premises. $\square$

A.3.4. Consistency

Proposition 6 (Uniform convergence and consistency). *Under Assumptions 2 to 5, for every compact $[0, \Lambda]$, $\sup_{\lambda \in [0, \Lambda]} |f_N(\lambda) - f(\lambda)| \xrightarrow{P} 0$, and consequently $\hat{\delta} \xrightarrow{P} \delta$. In the directional case this gives $\hat{\delta} \xrightarrow{P} \delta$.*

Proof. Uniform convergence. Let $R := |Y| + \Lambda Z$ and choose K so that $q := P(R > K) < \min(p, \tau)$. Since $|W(\lambda)| \leq R$ uniformly, the population minimizers lie in $[-K, K]$. Also, the law of large numbers and $p_N \rightarrow p$ imply $P_N \mathbf{1}\{R > K\} < \min(p_N, 1 - p_N)$ with probability tending to one. Then, uniformly in λ , fewer than k observations have $W_i(\lambda) > K$ and fewer than $N - k$ have $W_i(\lambda) < -K$. The top- k threshold condition in Lemma 2, $\#\{W_i > \xi\} \leq k \leq \#\{W_i \geq \xi\}$, then rules out every minimizer above K or below $-K$, including when there are ties. On $[0, \Lambda] \times [-K, K]$, the criterion class is

VC-subgraph with integrable envelope proportional to $R + K$, so the uniform law of large numbers gives $\sup_{\lambda, \xi} |\Gamma_N - \Gamma| \xrightarrow{p} 0$ there (the $p_N - p$ term is uniformly $o(1)$). Taking infima over $\xi \in [-K, K]$ yields the claim.

Consistency. The population root is positive by Proposition 4. For every sufficiently small $\eta > 0$, strict decrease gives $f(\delta - \eta) > 0 > f(\delta + \eta)$. Uniform convergence gives the same signs for f_N with probability tending to one, and its unique decreasing root lies between those endpoints. Thus $\hat{\delta} \xrightarrow{p} \delta$.

For the feasible directional statistic, first apply this argument to the oracle pair $(Y(a), Z(a))$. The standing regression moments give $(\hat{\beta}, \hat{r}_a) \xrightarrow{p} (\beta, r_a)$. On compact root intervals, the top- k value contraction bounds the feasible-minus-oracle criterion by $C\{\|\hat{\beta} - \beta\| + \|\hat{r}_a - r_a\|\}P_N H = o_p(1)$, with H the envelope defined after Assumption 6. The oracle bottom- $(N - k)$ average of $Z(a)$ tends to $\int_0^\tau F_{Z(a)}^{-1}(s) ds > 0$; replacing $Z(a)$ by its fitted version costs at most $P_N |\hat{Z} - Z| = o_p(1)$. Hence every feasible retained denominator is positive with probability tending to one, and the same sign-bracketing argument gives $\hat{\delta} \xrightarrow{p} \delta$. $\square$

A.3.5. The oracle central limit theorem

Theorem 4 (Scalar/oracle CLT). *Under Assumption 6, with $\psi(Y, Z) := (W - \xi)_+ + p\xi - \delta Z$ ($\mathbb{E}\psi = 0$),*

$$\sqrt{N}(\hat{\delta} - \delta) = \frac{1}{B\sqrt{N}} \sum_{i=1}^N \psi(Y_i, Z_i) + o_p(1) \xrightarrow{d} \mathcal{N}(0, \omega_p^2), \quad \omega_p^2 = \frac{\text{Var}(\psi)}{B^2}.$$

Equivalently $\psi(Y, Z) = Y\mathbf{1}\{W \geq \xi\} - \delta Z\mathbf{1}\{W < \xi\} + \xi(\mathbf{1}\{W < \xi\} - \tau)$ a.s. Applied to $(Y(a), Z(a))$ this is the oracle directional CLT for the statistic computed as though u_i and r_a were known— $\hat{\delta}^{\text{orc}}$ of Lemma 8—with influence function ψ_a and slope B_a^* ; it is not yet the limit for the feasible $\hat{\delta}$, which is Theorem 1.

Proof. Step 1 (linearize $f_N(\delta)$). The scalar profiling argument in Lemma 5, at the base point, gives $f_N(\delta) = \Gamma_N(\delta, \xi) + o_p(N^{-1/2})$. As $p_N - p = O(N^{-1})$,

$$f_N(\delta) = \frac{1}{N} \sum_{i=1}^N \psi(Y_i, Z_i) + o_p(N^{-1/2}).$$

Since $\mathbb{E}\psi = \Gamma(\delta, \xi) = f(\delta) = 0$ and $\mathbb{E}\psi^2 < \infty$, the i.i.d. CLT gives $\sqrt{N}f_N(\delta) \xrightarrow{d} \mathcal{N}(0, \text{Var}(\psi))$.

Step 2 (population slope). By the envelope theorem and uniqueness of the minimizer at δ , $f'(\delta) = \mathbb{E}[Z\mathbf{1}\{W \geq \xi\}] - \mu_Z = -(\mu_Z - \nu_p^Z) = -B < 0$, so $f(\delta + t) = f(\delta) - Bt + o(t)$.

Step 3 (Z-estimator expansion). On $\mathcal{D}_N$, $f_N(\hat{\delta}) = 0$ and $\hat{\delta} \xrightarrow{p} \delta$. Local linearity of f , C4 and tightness of $\sqrt{N}f_N(\delta)$ make the probability of $f_N(\delta - M/\sqrt{N}) > 0 > f_N(\delta + M/\sqrt{N})$ arbitrarily close to one by choosing M sufficiently large and then N large. Since f_N is decreasing and vanishes at $\hat{\delta}$, this gives $\sqrt{N}(\hat{\delta} - \delta) = O_p(1)$. Writing $\hat{\delta} = \delta + h_N/\sqrt{N}$ with $h_N = O_p(1)$ and applying C4 together with $f(\delta + t) = f(\delta) - Bt + o(t)$, $f_N(\hat{\delta}) - f_N(\delta) = -B(\hat{\delta} - \delta) + o_p(N^{-1/2})$. With $f_N(\hat{\delta}) = 0$ this gives $\sqrt{N}(\hat{\delta} - \delta) = \sqrt{N}f_N(\delta)/B + o_p(1)$; substituting Step 1 and Slutsky completes the proof. $\square$

A.3.6. The fixed-budget value sandwich and the moving selected set

For (23), at the single budget $k = \lfloor pN \rfloor$, write

$$T_N(W) = \max_{|S|=k} P_N[W\mathbf{1}\{S\}], \quad f_{N,\theta}(\lambda) = T_N(V_\theta(\lambda)) - \lambda P_N Z_\theta.$$

The oracle root at θ_0 is $\hat{\delta}^{\text{orc}}$. The following two elementary tools separate value algebra from stochastic selected-set convergence.

Lemma 6 (Value and root stability). *For $1 \leq k < N$ and arrays V, W with cardinality- k top sets S_V, S_W ,*

$$P_N[(V-W)\mathbf{1}\{S_W\}] \leq T_N(V) - T_N(W) \leq P_N[(V-W)\mathbf{1}\{S_V\}], \quad |T_N(V) - T_N(W)| \leq P_N|V - W|.$$

For $f(\lambda) = T_N(Y + \lambda Z) - \lambda P_N Z$, $Z \geq 0$, and top sets S_j at λ_j , $\lambda_1 < \lambda_2$,

$$-P_N[Z\mathbf{1}\{S_1^c\}] \leq \frac{f(\lambda_2) - f(\lambda_1)}{\lambda_2 - \lambda_1} \leq -P_N[Z\mathbf{1}\{S_2^c\}].$$

If every retained denominator is at least $b > 0$, f has a unique root ℓ , and for any comparison root ℓ_0 ,

$$|\ell - \ell_0| \leq |f(\ell_0)|/b.$$

For two leverage arrays, uniformly over $|S| = k$,

$$P_N[Z_\theta \mathbf{1}\{S^c\}] \geq N^{-1} \sum_{j \leq N-k} Z_{(j)}(\theta_0) - P_N|Z_\theta - Z_{\theta_0}|.$$

Proof. Evaluate each maximum at the other maximizing set to obtain the value sandwich. Substitute $V = Y + \lambda_2 Z$ and $W = Y + \lambda_1 Z$ and subtract $(\lambda_2 - \lambda_1)P_N Z$ to obtain the slope bounds. A positive denominator floor gives the unique ratio root by Proposition 3, and the slope bound between that root and ℓ_0 gives root stability. The last inequality follows by replacing Z_θ by Z_{θ_0} on S^c and minimizing the latter retained sum. $\square$

Lemma 7 (Cutoff trapping and weighted selection). *Index oracle scores and cutoffs by $t \in K$, with $A_t = \{W_t \geq \xi_t\}$, $P(A_t) = p_t$ and $k_N(t)/N = p_t + O(N^{-1})$ uniformly. Suppose their level and band indicators obey uniform laws of large numbers and, for each small $\eta > 0$,*

$$\inf_t \min\{p_t - P(W_t \geq \xi_t + \eta), P(W_t > \xi_t - \eta) - p_t\} > 0, \quad \sup_t P(|W_t - \xi_t| \leq \eta) \longrightarrow 0$$

as $\eta \downarrow 0$ in the second expression. Let $H \geq 1$, $PH^2 < \infty$ and $\varepsilon_N \downarrow 0$ be deterministic. Uniformly over all perturbed arrays with $|V_i - W_t(w_i)| \leq C\varepsilon_N H_i$, every measurable cardinality- $k_N(t)$ top set S has $k_N(t)$ -th largest cutoff $\hat{\xi}$ satisfying

$$\sup |\hat{\xi} - \xi_t| \xrightarrow{P} 0, \quad \sup P_N[L\mathbf{1}\{S \Delta A_t\}] \xrightarrow{P} 0, \quad L \in \{1, H, H^2\}.$$

Proof. For fixed $\eta > 0$, the empirical fraction with $|V - W_t| > \eta/2$ is at most $2C\varepsilon_N P_N H / \eta = o_p(1)$. The oracle count gaps at $\xi_t \pm \eta/2$, their uniform laws and budget rounding therefore give fewer than $k_N(t)$ scores at or above $\xi_t + \eta$ and more than $k_N(t)$ strictly above $\xi_t - \eta$. These inequalities trap $\hat{\xi}$ uniformly. Membership in a top set then gives

$$S \Delta A_t \subseteq \{|W_t - \xi_t| \leq \eta + C\varepsilon_N H\},$$

irrespective of tie allocation. For fixed M and eventually $C\varepsilon_N M \leq \eta$, its weighted mass is at most

$$M^2 P_N\{|W_t - \xi_t| \leq 2\eta\} + P_N[H^2 \mathbf{1}\{H > M\}].$$

Use the band uniform law and vanishing band mass, first letting $\eta \downarrow 0$ for fixed M , then $M \rightarrow \infty$ by integrability of H^2 . $\square$

Lemma 8 (Fixed-budget value sandwich, selected-set consistency, and root transfer). *Let Assumption 6 hold at the budget p fixed above; in particular $B_a^* > 0$ by C3.*

- (a) **Exact sandwiches.** For any arrays V, W with top- k sets S_V, S_W , $P_N[(V-W)\mathbf{1}\{S_W\}] \leq T_N(V) - T_N(W) \leq P_N[(V-W)\mathbf{1}\{S_V\}]$ and $|T_N(V) - T_N(W)| \leq P_N|V - W|$. Hence, for θ, θ' in a compact set with $a'r, a'r' \geq \mu_Z(a)/2$ and every $|\lambda| \leq \Lambda$, $|f_{N,\theta}(\lambda) - f_{N,\theta'}(\lambda)| \leq C_\Lambda \|\theta - \theta'\| P_N H$ with $C_\Lambda \leq C(1+\Lambda)$, and for $\lambda_1 < \lambda_2$ with top- k sets S_1, S_2 of $V_\theta(\lambda_1)$ and $V_\theta(\lambda_2)$,

$$-(\lambda_2 - \lambda_1)P_N[Z_\theta \mathbf{1}\{S_1^c\}] \leq f_{N,\theta}(\lambda_2) - f_{N,\theta}(\lambda_1) \leq -(\lambda_2 - \lambda_1)P_N[Z_\theta \mathbf{1}\{S_2^c\}].$$

- (b) **Weighted selected-set consistency.** Let $\varepsilon_N \downarrow 0$ be deterministic, small enough that $\|\theta - \theta_0\| \leq \varepsilon_N$ implies $a'r \geq \mu_Z(a)/2$, and let $\hat{\xi}_N(\theta, \lambda)$ be the k -th largest value of $V_\theta(\lambda)$ and $S_\theta(\lambda)$ a top- k set. Uniformly over $\|\theta - \theta_0\| \leq \varepsilon_N$ and $|\lambda - \delta| \leq \varepsilon_N$,

$$\sup |\hat{\xi}_N(\theta, \lambda) - \xi_a^*| \xrightarrow{P} 0 \text{ and, for } L \in \{1, H, H^2\}, \sup P_N[L \mathbf{1}\{S_\theta(\lambda) \triangle A^*\}] \xrightarrow{P} 0.$$

Consequently $\sup |P_N[Z_\theta \mathbf{1}\{S_\theta(\lambda)^c\}] - B_a^*| \xrightarrow{P} 0$ and, uniformly over pairs λ_1, λ_2 in the same tube, $f_{N,\theta_0}(\lambda_2) - f_{N,\theta_0}(\lambda_1) = -(\lambda_2 - \lambda_1)\{B_a^* + o_p(1)\}$.

- (c) **Joint first-order expansion of the criterion.** If $\|\theta_N - \theta_0\| = O_p(N^{-1/2})$ and $\lambda_N \xrightarrow{P} \delta$ with $\theta_N = (b_N, r_N)$, then

$$f_{N,\theta_N}(\lambda_N) - f_{N,\theta_0}(\lambda_N) = -m'_a(b_N - \beta) + D'_a(r_N - r_a) + o_p(1) \|\theta_N - \theta_0\| + O_p(1) \|\theta_N - \theta_0\|^2,$$

where $D_a = B_a^* g_a$ by (25) and Lemma 5.

- (d) **Retained denominator and uniform descent.** With $b_0 := \int_0^\tau F_{Z(a)}^{-1}(s) ds > 0$, on an event whose probability tends to one, we have $P_N[Z_\theta \mathbf{1}\{S^c\}] \geq b_0/2$ simultaneously over the tube of part (b) and over all S with $|S| = k$. Consequently the denominator event of Proposition 3 holds simultaneously for every pair in the tube, each $f_{N,\theta}$ there is strictly decreasing with slope at most $-b_0/2$ and has a unique root $\hat{\delta}_N(\theta)$, and every sequence θ_N with $\|\theta_N - \theta_0\| = O_p(N^{-1/2})$ obeys

$$|\hat{\delta}_N(\theta_N) - \hat{\delta}^{\text{orc}}| \leq \frac{2\delta_N}{b_0}, \quad \delta_N := C(|\delta| + 2) \|\theta_N - \theta_0\| P_N H,$$

so that $\hat{\delta}_N(\theta_N) - \hat{\delta}^{\text{orc}} = O_p(N^{-1/2})$ and $\hat{\delta}_N(\theta_N) \xrightarrow{P} \delta$.

- (e) **Root transfer.** If, in addition to part (c), $f_{N,\theta_N}(\lambda_N) = 0$, then $\lambda_N = \hat{\delta}_N(\theta_N)$,

$$\lambda_N \xrightarrow{p} \delta \text{ and}$$

$$\sqrt{N}(\lambda_N - \hat{\delta}^{\text{orc}}) = -B_a^{*-1} m'_a \sqrt{N}(b_N - \beta) + B_a^{*-1} D'_a \sqrt{N}(r_N - r_a) + o_p(1),$$

which, under the identification $D_a = B_a^* g_a$ of (8), reads $-B_a^{*-1} m'_a \sqrt{N}(b_N - \beta) + g'_a \sqrt{N}(r_N - r_a) + o_p(1)$.

Proof. A constant C , changing line to line, depends only on θ_0 , a , $\mu_Z(a)$ and the compact parameter sets.

(a). The value and slope inequalities are Lemma 6. On the stated compact, the product and quotient rules in (23) give

$$|V_\theta(\lambda) - V_{\theta'}(\lambda)| \leq C(1 + |\lambda|) \|\theta - \theta'\| H, \quad |Z_\theta - Z_{\theta'}| \leq C \|\theta - \theta'\| \|x\|^2.$$

Value contraction gives the criterion bound.

(b). Apply Lemma 7 with the singleton oracle index (W_a^*, ξ_a^*, p) . Condition C2 gives positive count gaps and $P(|W_a^* - \xi_a^*| \leq \eta) \leq 2C_f \eta$ locally. For each fixed band width, only finitely many oracle level and band sets occur, so their ordinary laws of large numbers suffice; no uniform-in-budget assumptions enter. Part (a) gives $|V_\theta(\lambda) - W_a^*| \leq C \varepsilon_N H$ over the stated tube, proving all three weighted conclusions. Since $Z(a) \leq CH$, the bound

$$|P_N[Z_\theta \mathbf{1}\{S^c\}] - B_a^*| \leq P_N|Z_\theta - Z(a)| + P_N[Z(a) \mathbf{1}\{S \Delta A^*\}] + |(P_N - P)[Z(a) \mathbf{1}\{A^{*,c}\}]|$$

vanishes uniformly. Insert this into the slope sandwich of part (a).

(c). Put $\Delta\theta = (\Delta b, \Delta r) = \theta_N - \theta_0$. Choose a deterministic shrinking tube containing (θ_N, λ_N) with probability tending to one. On the compact nuisance/root domain, (24) gives

$$V_{\theta_N}(\lambda_N) - V_{\theta_0}(\lambda_N) = D_b Y_{\theta_0} \Delta b + (D_r Y_{\theta_0} + \lambda_N D_r Z_{\theta_0})' \Delta r + \rho_N, \quad |\rho_N| \leq CH \|\Delta\theta\|^2.$$

Both bounding selected sets in the value sandwich may be replaced by A^* , by part (b), at cost $o_p(1) \|\Delta\theta\|$. The ordinary laws for the fixed derivative weights, all bounded by CH , replace empirical coefficients by expectations. Replacing λ_N by δ costs $o_p(1) \|\Delta\theta\|$, and $P_N |\rho_N| = O_p(1) \|\Delta\theta\|^2$. Taylor-expand $-\lambda_N P_N Z_{\theta_N}$ as well; its direction coefficient is $-\delta \mathbb{E}[D_r Z_{\theta_0}]$. The resulting coefficients are exactly (25); Lemma 5 identifies $D_a = B_a^* g_a$.

(d). The bottom-trimmed-sum bound in Lemma 6 converges to $b_0 > 0$ by the

trimmed-sum law and Assumption 3. Its perturbation is at most $C\varepsilon_N P_N \|x\|^2 = o_p(1)$, giving the simultaneous floor $b_0/2$. By Proposition 6, $|\hat{\delta}^{\text{orc}}| \leq |\delta| + 1$ with probability tending to one. Evaluate the criterion contraction of part (a) at this oracle root, then apply Lemma 6 with $b = b_0/2$ to obtain $|\hat{\delta}_N(\theta_N) - \hat{\delta}^{\text{orc}}| \leq 2\delta_N/b_0$. This proves consistency and the root perturbation rate without using a first-stage expansion.

(e). On the denominator event of (d) the root of f_{N,θ_N} is unique, so $f_{N,\theta_N}(\lambda_N) = 0$ forces $\lambda_N = \hat{\delta}_N(\theta_N)$; part (d) and the oracle CLT put both endpoints at distance $O_p(N^{-1/2})$ from δ , hence in the tube of (b) with $\varepsilon_N = N^{-1/3}$ with probability tending to one. The slope identity of (b), applied at $\theta = \theta_0$ between $\hat{\delta}^{\text{orc}}$ and λ_N , expresses $f_{N,\theta_0}(\lambda_N) = -\{B_a^* + o_p(1)\}(\lambda_N - \hat{\delta}^{\text{orc}})$, and the expansion of (c) with $f_{N,\theta_N}(\lambda_N) = 0$ expresses the same quantity as $m'_a \Delta b - D'_a \Delta r + o_p(N^{-1/2})$; equating the two gives the representation, division by $B_a^* + o_p(1)$ being legitimate as C3 asserts $B_a^* > 0$, and the second form is the identification of D_a with $B_a^* g_a$.

□

To identify the first-order contribution, express the statistic as a function of the plug-in direction and differentiate its population counterpart at r_a . Writing

$$\hat{\delta} = \max_s \frac{(a' \hat{r}_a) \sum_{i \in S} (\hat{r}'_a x_i) \hat{u}_i}{\sum_{i \in S} (\hat{r}'_a x_i)^2},$$

numerator and denominator are homogeneous of degree two in $\hat{r}_a$, so $\hat{\delta}$ is homogeneous of degree zero, meaning the component of $\sqrt{N}(\hat{r}_a - r_a)$ along r_a drops out, but the orthogonal component need not. For a generic direction r (not just the true r_a) define the scalar root $\delta^{\text{sc}}(r)$, the value the frozen-metric statistic would target if the contrast were scalarized along r , as the unique solution of

$$f(\lambda; r) := \mathcal{T}_p \{(r'x)u + \lambda(r'x)^2\} - \lambda(r'Qr) = 0, \quad (27)$$

which is just the population Dinkelbach criterion of Proposition 4 read in direction r , and the population directional functional

$$\Phi_a(r) := (a'r) \delta^{\text{sc}}(r), \quad \Phi_a(r_a) = \delta, \quad \delta^{\text{sc}}(r_a) =: \delta^{\text{sc}} = \delta/\mu_Z(a). \quad (28)$$

$\Phi_a(r)$ is the population limit of $\hat{\delta}$ seen as a function of the plug-in direction alone, so the design correction is nothing but its gradient at the truth r_a . Since $\mathcal{T}_p$ is positively

homogeneous of degree one, Φ_a is homogeneous of degree zero.

Theorem 5 computes the design correction in closed form and shows that it is orthogonal to r_a . Under index sufficiency, Remark 2 further expresses it through departures of $\mathbb{E}[x \mid r'_a x]$ from mean-linearity.

Theorem 5 (Design influence function). *Under Assumption 6, $\hat{r}_a$ contributes to $\hat{\delta}$ the first-order term $g'_a \sqrt{N}(\hat{r}_a - r_a)$, where $g_a := \nabla_r \Phi_a(r)|_{r=r_a}$ has the closed form (8). The gradient is orthogonal to the efficient direction, $g'_a r_a = 0$ (Euler's relation), and the per-observation design influence is*

$$d_{a,i} := g'_a \dot{r}_i = -(r'_a x_i)(\zeta'_a x_i), \quad \dot{r}_i := r_a - (r'_a x_i)Q^{-1}x_i, \quad \zeta_a := Q^{-1}g_a, \quad (29)$$

where $\dot{r}_i$ is the influence function of $\hat{r}_a$, ζ_a re-expresses the design gradient g_a in the Q -metric, and $\mathbb{E}[d_{a,i}] = 0$. The contribution of $d_{a,i}$ to Ψ_a in (9) carries an explicit factor B_a^* , since Ψ_a is normalized by $1/B_a^*$ in Theorem 1.

Proof. Lemma 5 gives the derivative at the truth directly, without a neighborhood implicit-function theorem: $g_a = D_a/B_a^*$. The coefficient calculation (25) therefore gives (8). The first-order contribution follows from Lemma 8, whose joint expansion has this same direction coefficient. For $g'_a r_a = 0$: the root identity $\mathbb{E}[W_a^* \mathbf{1}\{W_a^* \geq \xi_a^*\}] = \delta$ with $W_a^* = Y(a) + \delta Z(a)$ gives $r'_a g_{A^*} = \delta B_a^*$ and $r'_a m_a = \mu_Z(a)\nu_{a,Z}^*$; substituting and using $\delta^{\text{sc}} \mu_Z(a) = \delta$ collapses $g'_a r_a$ to zero. Finally $\dot{r}_i = -Q^{-1}(x_i x'_i - Q)r_a = r_a - (r'_a x_i)Q^{-1}x_i$ is the influence of $\hat{Q}_N^{-1}a$; using $g'_a r_a = 0$, $g'_a \dot{r}_i = -(r'_a x_i)(g'_a Q^{-1}x_i) = -(r'_a x_i)(\zeta'_a x_i)$ with $\zeta_a = Q^{-1}g_a$. $\square$

Remark 2 (The design's nonlinearity correction). *Conditioning on $t = r'_a x$ and using index sufficiency $\mathbb{E}[x \mid t, u] = \mathbb{E}[x \mid t]$ (weaker than $x \perp\!\!\!\perp u$), $B_a^* g_a = \mathbb{E}[\rho(t)(u + 2\delta^{\text{sc}} t) \mathbf{1}\{W_a^* \geq \xi_a^*\}]$ with $\rho(t) := \mathbb{E}[x \mid t] - \mu_Z(a)^{-1}a$ the nonlinear part of the regression of x on the index t . Under index sufficiency the design correction is driven entirely by departures of $\mathbb{E}[x \mid r'_a x]$ from linearity, and vanishes when that conditional mean is linear. Neither condition can be dropped: the identity above is derived under index sufficiency, so mean-linearity of $\mathbb{E}[x \mid r'_a x]$ on its own does not force $g_a = 0$, and index sufficiency on its own does not either. Remark 3 gives a two-regressor design in which $\rho \equiv 0$, $\mathbb{E}[xu] = 0$ and $B_a^* > 0$ all hold while $g_a \neq 0$, because u carries information about the second regressor that the index does not.*

Remark 3 (Mean-linearity of the index regression). Let $x = (x_1, x_2)' \in \mathbb{R}^2$ have independent standard-normal coordinates, take $a = e_1$, and set $u = x_2(x_1^2 - 1)$. Then $Q = I_2$, $\mathbb{E}[xu] = 0$, and $\mathbb{E}[x | x_1] = (x_1, 0)' = \mu_Z(a)^{-1}a x_1$, so $\rho \equiv 0$. Index sufficiency fails: given $(x_1 = t, u)$ with $t^2 \neq 1$ the second regressor is pinned to $x_2 = u/(t^2 - 1)$, so $\mathbb{E}[x | x_1, u] \neq \mathbb{E}[x | x_1]$. The modified score is $W_a^* = x_1 u + \delta^{\text{sc}} x_1^2$, a function of (x_1, u) ; since $u = x_2(x_1^2 - 1)$, its selection event depends on x_2 given x_1 , so the active set carries information about x_2 beyond the index. Evaluating the right-hand side of (8) at $p = 0.10$ gives $B_a^* = 0.665 > 0$ and a strictly positive second component, $g_{a,2} \approx 0.60$; a direct Monte Carlo and a scrambled Sobol evaluation of the population directional derivative agree to within Monte-Carlo error. The design gradient therefore does not vanish on the paper's advertised generality, and Corollary 2 charges index sufficiency explicitly.

The formula above tells us when the design layer can be dropped; the operative condition is the vanishing of the design gradient itself, Corollary 2.

Corollary 2 (Partial adaptivity). If $g_a = 0$, then $\hat{r}_a$ has no first-order effect and $\hat{\delta}$'s influence reduces to the generated-residual influence numerator $\tilde{\psi}_{a,i} := \psi_{a,i}^* - m_a' Q^{-1} x_i u_i$, so the normalized influence is $\tilde{\psi}_{a,i}/B_a^*$ as in (9). A convenient primitive sufficient condition is index sufficiency together with mean-linearity of the index regression: $\mathbb{E}[x | r_a' x, u] = \mathbb{E}[x | r_a' x]$ and $\mathbb{E}[x | r_a' x] = \mu_Z(a)^{-1}a (r_a' x)$, which is the condition under which Remark 2 rewrites $B_a^* g_a$ as the display above. Both hold under mean-zero elliptic symmetry of x with $u \perp\!\!\!\perp x$, and both hold at $d = 1$, where x is measurable with respect to its own index; at $d = 1$ the vanishing also follows directly from Euler's relation $g_a' r_a = 0$ in Theorem 5. Mean-linearity alone does not suffice; see Remark 3.

Remark 4 (Sign symmetry as an alternative route to $g_a = 0$). Suppose the law of (x, u) is invariant under $(x, u) \mapsto (Jx, -u)$ for a linear involution J satisfying $JQJ = Q$ and $Ja = -a$ (with an intercept and slope contrast, $J = \text{diag}(1, -1, \dots, -1)$ represents joint central symmetry of the non-constant regressors and u). Then $Jr_a = -r_a$, so $t = r_a' x$ and u flip sign together while $W_a^* = tu + \mu_Z(a)^{-1} \delta t^2$ and A^* remain invariant. Consequently,

$$g_{A^*} = \mathbb{E}[xu \mathbf{1}\{A^*\}] = -J g_{A^*}, \quad m_a = \mathbb{E}[x t \mathbf{1}\{A^*\}] = -J m_a,$$

so g_{A^*} , m_a , a , and hence g_a in (8) all lie in the (-1) -eigenspace of J and have zero intercept component. When ρ in Remark 2 is constant (so the intercept is the sole source of nonlinearity in $\mathbb{E}[x | t]$) and index sufficiency holds, Remark 2 restricts g_a to $\text{span}\{\rho\} = \text{span}\{e_1\}$, forcing $g_a = 0$.

With all three influence contributions in hand (oracle, residual, and design) together with the conditions under which each is present or absent, the limit for the feasible statistic is now just their sum.

A.3.7. The central limit theorem for the feasible statistic $\hat{\delta}$

Assembling the oracle limit with the two corrections gives the distribution of the feasible directional statistic. Its influence function has the oracle, generated-residual, and design layers; the design layer disappears under the stated adaptivity conditions of Corollary 2. Sample splitting alone does not remove it from the unconditional limit: it makes the design layer an orthogonal offset computed on the training sample, so $g'_a \sqrt{N_{\text{train}}}(\hat{r}_a - r_a)$ remains first order whenever $g_a \neq 0$ and both samples grow at comparable rates; it enters the reported variance additively rather than through the test-sample influence.

Proof of Theorem 1. Expand the feasible criterion jointly in the two estimated objects rather than one at a time. Lemma 8(e), evaluated at $\theta_N = \hat{\theta}_2 = (\hat{\beta}, \hat{r}_a)$ where $\lambda_N = \hat{\delta}$, together with the first-stage expansion (22), gives

$$\sqrt{N}(\hat{\delta} - \hat{\delta}^{\text{orc}}) = -\frac{m'_a Q^{-1} N^{-1/2} \sum_i x_i u_i}{B_a^*} + \frac{N^{-1/2} \sum_i B_a^* g'_a \dot{r}_i}{B_a^*} + o_p(1),$$

$$g'_a \dot{r}_i = -(r'_a x_i)(\zeta'_a x_i),$$

where the direction coefficient of the lemma is identified with $B_a^* g_a$ by (8), and the identity for $g'_a \dot{r}_i$ is Theorem 5. Theorem 4 at the pair $(Y(a), Z(a))$ supplies the oracle layer, $\sqrt{N}(\hat{\delta}^{\text{orc}} - \delta) = \{B_a^* \sqrt{N}\}^{-1} \sum_i \psi_{a,i}^* + o_p(1)$. Adding the two expansions, the three layers sum to $\{B_a^* \sqrt{N}\}^{-1} \sum_i \Psi_{a,i}$ with $\Psi_{a,i}$ as in (9), the design layer carrying its factor B_a^* against the common B_a^{*-1} . No cross term arises, because the lemma expands the criterion in the single joint perturbation $\theta_N - \theta_0$ with a remainder that is $o_p(\|\theta_N - \theta_0\|)$: the residual and design errors enter only through one linear term of one expansion. Each layer is centered— $\mathbb{E}\psi_a^* = 0$ by (7), the residual layer by (2), the design layer by $g'_a r_a = 0$ in Theorem 5—and all three are square-integrable by the standing moment bounds, so the Lindeberg–Lévy theorem applied to the i.i.d. summand $\Psi_{a,i}$ and Slutsky's theorem give the limit with $\omega_a^2 = \text{Var}(\Psi_a)/B_a^{*2}$, strictly positive by C6. $\square$

Theorem 1 provides the pointwise inferential foundation for the directional target, with a variance estimator that retains all three layers. The next subsection makes that estimator explicit and proves its pointwise consistency under the same conditions.

A.3.8. Plug-in formulas for the sandwich standard error

Lemma 9 (Empirical covering by a parameter mesh). *On a fixed compact domain for $\theta = (e, r, \lambda, c, \xi, p, B, t, \zeta)$, with $e, r, t, \zeta \in \mathbb{R}^d$ and c, B bounded away from zero, define*

$$h_\theta(x, u) = \frac{((r'x)(u - x'e) + \lambda(r'x)^2/c - \xi)_+ + p\xi - \lambda(r'x)^2/c}{B} - (x't)(u - x'e) - (r'x)(\zeta'x).$$

This family has envelope and parameter Lipschitz modulus $M = C_\theta H \in L^2(P)$. Its population and empirical class bounds are those of Lemma 4, with parameter count $d_\theta = 4d + 5$. If fitted parameters lie uniformly in this domain with probability tending to one, denote their score class by $\hat{\mathcal{G}}_{\Psi, N}$. On events $\mathcal{C}_N$ with probability tending to one it has the deterministic entropy profile

$$\log \mathcal{N}(\varepsilon, \hat{\mathcal{G}}_{\Psi, N}, \|\cdot\|_{N, 2}) \leq \psi(\varepsilon) := d_\theta \log(1 + C_\theta''/\varepsilon), \quad 0 < \varepsilon \leq 1.$$

The same profile, after enlarging its constant, covers the corresponding population score class.

Proof. Positive-part contraction and the product and quotient rules give $|h_\theta| \leq M$ and $|h_\theta - h_\eta| \leq M \|\theta - \eta\|$. Apply Lemma 4; intersect fitted localization with $\{\|M\|_{N, 2} \leq 2(PM^2)^{1/2}\}$ to obtain $\mathcal{C}_N$. The strong law makes this one event valid simultaneously at every covering radius. For the standardized population score $\Psi_a(p)/B_a^*(p)$ take

$$(e, r, \lambda, c, \xi, p, B, t, \zeta) = (0, r_a, \delta(p), \mu_Z(a), \xi_a^*(p), p, B_a^*(p), Q^{-1}m_a(p)/B_a^*(p), \zeta_a(p)).$$

For the fitted score take the corresponding fitted quantities, with $e = \hat{\beta} - \beta$ and budget coordinate $p_N(p)$. Both λ slots use the same root, and all fitted coefficients use the same cardinality- $k(p)$ selection. Then $u - x'e = \hat{u}$ and division by $\hat{B}_a^*$ cancels the design-layer factor in (12), so the parameter reading is exact, including at $d = 1$. At a fixed budget the parameters localize by pointwise consistency; over $\mathcal{P}$, compact population paths and fitted localization are verified in Lemma 15. Selected-set changes thus enter through finitely many coefficients, not an additional observation-specific indicator. $\square$

Proposition 7 (Pointwise plug-in consistency). *Under Assumption 6 at the reported deletion share p ,*

$$\hat{B}_a^* \xrightarrow{p} B_a^*, \quad \frac{1}{N} \sum_{i=1}^N (\hat{\Psi}_{a,i} - \Psi_{a,i})^2 \xrightarrow{p} 0, \quad \sqrt{N} \widehat{\text{SE}}(\hat{\delta}(p)) \xrightarrow{p} \omega_a.$$

Corollary 3 (Directional confidence interval). *Under Assumption 6, including the positive influence variance in (C6),*

$$\text{CI}_{1-\alpha} = [\hat{\delta}(p) \pm z_{1-\alpha/2} \widehat{\text{SE}}(\hat{\delta}(p))], \quad \mathbb{P}\{\delta(p) \in \text{CI}_{1-\alpha}\} \rightarrow 1 - \alpha.$$

Proof of Proposition 7. By Lemma 8(b)–(e) and (22), $(\hat{\beta}, \hat{r}_a, \hat{\delta}, \hat{\xi}) \xrightarrow{p} (\beta, r_a, \delta, \xi_a^*)$ and $P_N[H\mathbf{1}\{S^* \triangle A^*\}] \xrightarrow{p} 0$. The fixed-set law of large numbers and this weighted selected-set convergence give $(\hat{m}_a, \hat{g}_{S^*}, \hat{B}_a^*) \xrightarrow{p} (m_a, g_{A^*}, B_a^*)$, with $B_a^* > 0$. Hence $(\hat{g}_a, \hat{\zeta}_a) \xrightarrow{p} (g_a, \zeta_a)$ by (11) and $\hat{Q}_N \xrightarrow{p} Q > 0$.

The score Lipschitz bound in Lemma 9, applied to these convergent parameters, gives $|\hat{\Psi}_{a,i} - \Psi_{a,i}| \leq o_p(1)H(w_i)$, also accounting for $p_N - p = O(N^{-1})$. Since $P_N H^2 \xrightarrow{p} \mathbb{E} H^2 < \infty$, empirical- L^2 convergence follows, hence variance and standard-error consistency. $\square$

B. Functional limit theory in the deletion fraction

This appendix proves the results of Section 4.1 and Section 4.2. Throughout, $\mathcal{P} = [p, \bar{p}] \subset (0, 1)$, Assumption 7 is in force, $(Y_i, Z_i) = (Y_i(a), Z_i(a))$ is the oracle pair, and constants are uniform over $\mathcal{P}$. At budget p , let $k(p) = \lfloor pN \rfloor$, $p_N(p) = k(p)/N$, use the criteria of (21), and write $\xi(\lambda, p) = F_{W(\lambda)}^{-1}(1-p)$ and $\mathbb{G}_N \phi = N^{-1/2} \sum_i \{\phi(Y_i, Z_i, x_i, u_i) - \mathbb{E}\phi\}$. Set $\|h\|_{N,2} = \{N^{-1} \sum_i h(w_i)^2\}^{1/2}$ and let $\mathcal{N}(\varepsilon, \mathcal{G}, \|\cdot\|_{N,2})$ denote the empirical covering number. Steps 1–2 use the oracle root $\hat{\delta}^{\text{orc}}(p)$ and its threshold minimizer $\hat{\xi}_N^{\text{orc}}(p)$; Step 3 transfers the expansion to feasible $\hat{\delta}(p)$ through the conclusion U3 of Lemma 14. Probabilities involving the multipliers of (16) are joint over data and multipliers, with outer values written $\mathbb{P}^*$ and $\mathbb{E}^*$ where measurability is not asserted.

The uniform regularity assumption in force throughout is the following, the uniform-in- p strengthening of Assumption 6.

Assumption 7 (Uniform regularity in p). *Assumption 6 holds at every $p \in \mathcal{P}$ and conditions U1–U2 below hold.*

- (U1) **Uniform quantile regularity.** *There are a neighborhood $\mathcal{U}_0 \subset \mathbb{R} \times (0, 1)$ of the root path $\{(\delta(p), p) : p \in \mathcal{P}\}$ and constants $0 < c_f \leq C_f < \infty$ such that for every $(\lambda, p) \in \mathcal{U}_0$ the score $W(\lambda) = Y(a) + \lambda Z(a)$ has a density $f_{W(\lambda)}$, continuous and taking values in $[c_f, C_f]$ on a fixed neighborhood of its $(1 - p)$ -quantile $\xi(\lambda, p)$; consequently the margin bound $\sup_{(\lambda, p) \in \mathcal{U}_0} P(|W(\lambda) - \xi(\lambda, p)| \leq t) \leq 2C_f t$ holds for all small $t > 0$.*
- (U2) **Uniformly nonzero slope and variance.** $\inf_{p \in \mathcal{P}} B_a^*(p) > 0$ and $\inf_{p \in \mathcal{P}} \text{Var}(\Psi_a(p)) > 0$, the uniform counterparts of C3 and C6.

Under U1 and C1 the first-stage coefficients $p \mapsto m_a(p)$ and (for $d \geq 2$) $p \mapsto \zeta_a(p)$ are automatically continuous on $\mathcal{P}$, by dominated convergence across the Lebesgue-null boundary $\{W_a^*(p) = \xi_a^*(p)\}$, so no separate condition is imposed on them.

B.1. An indexed Donsker class

Lemma 10 (Parametric Lipschitz class). *Let $\Theta \subset \mathbb{R}^2 \times \mathcal{P}$ be compact and*

$$\mathcal{G}_\Theta := \{g_{\lambda, \xi, p} : (y, z) \mapsto (y + \lambda z - \xi)_+ + p\xi - \lambda z : (\lambda, \xi, p) \in \Theta\}.$$

Then $\mathcal{G}_\Theta$ is pointwise measurable with square-integrable envelope and is Lipschitz in its parameter,

$$|g_{\theta_1}(y, z) - g_{\theta_2}(y, z)| \leq L(y, z) \|\theta_1 - \theta_2\|, \quad L(y, z) = 2z + 2 + \sup_{\xi} |\xi| \in L^2,$$

hence P -Donsker. The same holds for the indicator layer $\{(y, z) \mapsto z \mathbf{1}\{y + \lambda z \geq \xi\} : (\lambda, \xi)\}$ (VC-subgraph with envelope $z \in L^2$) and for any class formed by adding to $\mathcal{G}_\Theta$ finitely many fixed square-integrable functions with continuous parametric coefficients.

Proof. Positive-part contraction gives the stated modulus and an envelope $C(1 + |y| + z) \in L^2$. Apply Lemma 4 to the hinge family. The indicator layer is a fixed L^2 weight times a VC class of half-planes, hence Donsker. Adding finitely many fixed L^2 functions with bounded coefficients preserves the Donsker property. $\square$

B.2. Proof of Lemma 11

Lemma 11 (Positive threshold). *Let $p \in (0, 1)$, suppose $\mathbb{E}[xu] = 0$ and $B_a^*(p) > 0$, and let $W_a^*(p)$ be non-degenerate on the retained region $A_p^{*,c}$. Then $\xi_a^*(p) > 0$.*

Proof of Lemma 11. Since $\mathbb{E}[Y(a)] = r'_a \mathbb{E}[xu] = 0$ and $\mathbb{E}[Z(a)] = 1$, the modified score has mean $\mathbb{E}[W_a^*(p)] = \delta(p)$. The identity defining the root is $\mathcal{T}_p\{W_a^*(p)\} = \mathbb{E}[W_a^*(p) \mathbf{1}\{A_p^*\}] = \delta(p)$, so subtracting the two gives

$$\mathbb{E}[W_a^*(p) \mathbf{1}\{A_p^{*,c}\}] = 0.$$

On $A_p^{*,c}$ we have $W_a^*(p) < \xi_a^*(p)$, and $P(A_p^{*,c}) = 1 - p > 0$, so the displayed expectation is strictly below $\xi_a^*(p)(1 - p)$ unless $W_a^*(p)$ is a.s. equal to $\xi_a^*(p)$ on $A_p^{*,c}$. Hence $0 < \xi_a^*(p)(1 - p)$, i.e. $\xi_a^*(p) > 0$. $\square$

B.3. Proof of Lemma 1

Proof of Lemma 1. Recall $f(\lambda; p) = \mathcal{T}_p(W(\lambda)) - \lambda\mu_Z$.

Derivative in p . For fixed λ , $p \mapsto \int_{1-p}^1 F_{W(\lambda)}^{-1}(u) du$ has a moving lower limit; since $F_{W(\lambda)}^{-1}$ is continuous at $1 - p$ for $(\lambda, p) \in \mathcal{U}_0$ (U1), Lebesgue differentiation gives $\partial_p f(\lambda; p) = F_{W(\lambda)}^{-1}(1 - p) = \xi(\lambda, p)$, jointly continuous on $\mathcal{U}_0$.

Derivative in λ . $\Gamma(\lambda, \xi; p) = \mathbb{E}[(W(\lambda) - \xi)_+ + p\xi - \lambda Z]$ is convex in (λ, ξ) , differentiable in λ with $\partial_\lambda \Gamma = \mathbb{E}[Z \mathbf{1}\{W(\lambda) \geq \xi\}] - \mu_Z$, and its inner minimizer $\xi(\lambda, p)$ is unique on $\mathcal{U}_0$ (U1, Lemma 3); Danskin's theorem then yields $\partial_\lambda f(\lambda; p) = \mathbb{E}[Z \mathbf{1}\{W(\lambda) \geq \xi(\lambda, p)\}] - \mu_Z = -B(\lambda, p)$, jointly continuous, with $B(\delta(p), p) = B_a^*(p) \geq B_{\min} > 0$ by U2.

Implicit function theorem. $f(\delta(p); p) = 0$ for each p (Proposition 4); f is C^1 on $\mathcal{U}_0$ with $\partial_\lambda f \leq -B_{\min}/2 < 0$ on a possibly smaller neighborhood of the root path. The implicit function theorem gives $p \mapsto \delta(p) \in C^1$ with $\partial_p \delta(p) = -\partial_p f / \partial_\lambda f = \xi_a^*(p) / B_a^*(p)$, which is (14).

Monotonicity. By Lemma 11, $\xi_a^*(p) > 0$ for every $p \in \mathcal{P}$, and $B_a^*(p) \geq B_{\min} > 0$ by U2; hence $\partial_p \delta(p) > 0$ on $\mathcal{P}$ and the curve is strictly increasing. No condition on $P(Y(a) > 0)$ or on the location of $\bar{p}$ is required. (The non-degeneracy of $W_a^*(p)$ on the retained region used by Lemma 11 is implied by the positive density of U1.) $\square$

B.4. Uniform localization and first-stage expansion

Write P_N for the empirical average, $r_0 = r_a$, $c_0 = a'r_0 = \mu_Z(a) > 0$ and $\theta_0 = (\beta, r_0)$. For $\theta = (b, r)$ on a compact neighborhood of θ_0 with $a'r \geq c_0/2$, the nuisance pair Y_θ, Z_θ and the criterion $f_{N,\theta}$ are those of (23), read at the deletion budget $k(p)$ with $p_N(p) = k(p)/N$; we write $f_{N,\theta}(\lambda; p)$ for that criterion at budget p . Here $\mathbf{1}\{S\}$ denotes

the observation-index indicator, whereas $\mathbf{1}\{A_p^*\}$ below denotes a population event evaluated on the observations. At $\hat{\theta} = (\hat{\beta}, \hat{\tau}_a)$ the root is precisely $\hat{\delta}(p)$; away from that fitted point $P_N Z_{\theta}$ need not equal one and must remain in the criterion. Every empirical selected set below has exactly $k(p)$ elements: take the largest scores and break boundary ties by observation index. The proofs also apply to any other measurable cardinality- $k(p)$ maximizing selection.

Lemma 12 (Uniform consistency and feasible localization). *Under Assumption 6 at every $p \in \mathcal{P}$ and U1–U2, before deriving the first-stage expansion U3,*

$$\sup_p \left| \hat{\delta}^{\text{orc}}(p) - \delta(p) \right| \xrightarrow{p} 0, \quad \sup_p \left| \hat{\delta}(p) - \hat{\delta}^{\text{orc}}(p) \right| = O_p(N^{-1/2}).$$

Proof. Use the nuisance Lipschitz bound of Lemma 8(a) and the deterministic stability bounds of Lemma 6, whose constants do not depend on p . Let $b_0 = \int_0^{1-\bar{p}} F_{Z_{\theta_0}}^{-1}(s) ds > 0$, using Assumption 3 at the largest budget. The empirical bottom- $(N - k(\bar{p}))$ sum b_N of Z_{θ_0} , divided by N , tends to b_0 by the trimmed-sum law of large numbers. The bottom-sum inequality in Lemma 6, with $p \leq \bar{p}$, and (22) give every oracle and feasible retained denominator the floor $b_0/2$ with probability tending to one. The oracle root is the maximum over S of the ratios $P_N[Y_{\theta_0} \mathbf{1}\{S\}]/P_N[Z_{\theta_0} \mathbf{1}\{S^c\}]$, each of which lies between $-P_N|Y_{\theta_0}|/b_N$ and $P_N|Y_{\theta_0}|/b_N$. Thus $\sup_p |\hat{\delta}^{\text{orc}}(p)| \leq P_N|Y_{\theta_0}|/b_N$ even when $f_{N,\theta_0}(0; p) < 0$; the laws of large numbers put the roots in a deterministic bounded interval with probability tending to one. No monotonicity of the top- k numerator in k is used.

On that interval $|Y_{\theta_0}| + \Lambda Z_{\theta_0}$ is an integrable score envelope. Choose a deterministic threshold bound whose envelope-tail probability is smaller than half of $\min(\underline{p}, 1 - \bar{p})$. The law of large numbers and the strict count inequalities in Lemma 2 localize every empirical threshold minimizer in this bound, uniformly over λ, p ; the population minimizers are localized there as well. The compact positive-part class of Lemma 10 is Glivenko–Cantelli, so taking threshold infima gives uniform convergence $f_{N,\theta_0} \rightarrow f$, with budget rounding contributing at most C/N . Population retained denominators are at least b_0 , which supplies uniform root identification; the localized threshold dual also gives joint continuity of the population criterion. Thus the oracle roots converge uniformly to the continuous population root path. The population arguments of Lemma 1, which use only U1–U2, give continuity of its threshold path as well.

Finally, top- k values are Lipschitz in the empirical L^1 distance between scores, so

on the deterministic root interval

$$\sup_{\lambda, p} |f_{N, \hat{\theta}}(\lambda; p) - f_{N, \theta_0}(\lambda; p)| \leq C \|\hat{\theta} - \theta_0\| P_N H = O_p(N^{-1/2}).$$

Evaluate at each oracle root and use the feasible decreasing-slope bound to obtain the stated root perturbation rate, hence feasible consistency without using the subsequent first-stage expansion U3. $\square$

Lemma 13 (Uniform weighted selected-set consistency). *Under the conditions of Lemma 12, let $\varepsilon_N \downarrow 0$ be deterministic and let $S_\theta(\lambda, p)$ be a cardinality- $k(p)$ top set of $Y_\theta + \lambda Z_\theta$. Uniformly over $\|\theta - \theta_0\| \leq \varepsilon_N$, $|\lambda - \delta(p)| \leq \varepsilon_N$ and $p \in \mathcal{P}$, its threshold converges to $\xi_a^*(p)$ and*

$$P_N[L\mathbf{1}\{S_\theta(\lambda, p) \Delta A_p^*\}] \xrightarrow{p} 0$$

for $L = H$ and $L = H^2$. The conclusions hold for the fitted and oracle root paths of Lemma 12.

Proof. Apply Lemma 7 with $t = p \in \mathcal{P}$ and oracle pair $(W_a^*(p), \xi_a^*(p))$. Oracle superlevel sets and bands are half-planes and intersections of half-planes in $(Y_{\theta_0}, Z_{\theta_0})$, hence VC. Their uniform laws also hold after multiplication by fixed integrable weights, by bounded simple-function approximation and tail truncation; this covers H, H^2 and all derivative coordinates used below. Condition U1 supplies uniform count gaps and the vanishing band bound, while the nuisance/root tube gives $|V_\theta(\lambda) - W_a^*(p)| \leq C\varepsilon_N H$. The template proves threshold and weighted selection consistency without fitted-density or quantile-rate assumptions, for every measurable cardinality- $k(p)$ tie allocation. The same strict counts trap both endpoints of the threshold-minimizer interval. For the random consistent paths of Lemma 12, choose a deterministic tube decreasing slowly enough to contain them with probability tending to one. $\square$

Lemma 14 (Uniform first-stage expansion). *The conditions of Lemma 12 imply the following uniform-in- p strengthening of Theorems 1 and 5, including the design term when $g_a(p) \neq 0$:*

(U3) **Uniform first-stage expansions.** With $S_N := N^{-1/2} \sum_i x_i u_i$ and $\Delta_N := \hat{Q}_N - Q$,

$$\sup_{p \in \mathcal{P}} \left| \sqrt{N} \left\{ \hat{\delta}(p) - \hat{\delta}^{\text{orc}}(p) \right\} + \frac{m_a(p)' Q^{-1} S_N}{B_a^*(p)} + \sqrt{N} \zeta_a(p)' \Delta_N r_a \right| \xrightarrow{p} 0,$$

the last term being absent when $d = 1$ or, more generally, when $g_a = 0$ by Corollary 2.

Proof. The exact value sandwich is established in Lemma 8; apply it here uniformly over p and the shrinking nuisance/root tubes. The observation-wise nuisance Taylor remainder is bounded by $CH \|\theta - \theta_0\|^2$. Lemma 13 replaces both selected sets in the first derivative by A_p^* with coefficient error $o_p(1)$, uniformly on shrinking nuisance and root tubes. The weighted uniform law replaces empirical derivative coefficients by their population values, and replacing λ by $\delta(p)$ changes them by $o_p(1)$. Taylor-expanding the subtraction $-\lambda P_N Z_\theta$ as well gives

$$\begin{aligned} f_{N,\theta}(\lambda; p) - f_{N,\theta_0}(\lambda; p) &= -m_a(p)'(b - \beta) + B_a^*(p)g_a(p)'(r - r_0) \\ &\quad + o_p(1) \|\theta - \theta_0\| + O_p(1) \|\theta - \theta_0\|^2. \end{aligned}$$

The coefficient identity is Lemma 8(c), identified with $B_a^*(p)g_a(p)$ by (8).

The slope sandwich in Lemma 8, combined with Lemma 13 and the weighted uniform law, gives retained averages $B_a^*(p) + o_p(1)$ uniformly for $\lambda_2 > \lambda_1$ in the shrinking root tube. Consequently, across empirical kinks and ties,

$$f_N(\lambda_2; p) - f_N(\lambda_1; p) = -B_a^*(p)(\lambda_2 - \lambda_1) + o_p(1) |\lambda_2 - \lambda_1|. \quad (30)$$

The reversed ordering follows by changing signs. Evaluate the nuisance expansion at the feasible root and this slope expansion between the feasible and oracle roots. Both roots are consistent and their difference is already $O_p(N^{-1/2})$ by Lemma 12, so their root equations yield

$$\sqrt{N}(\hat{\delta}(p) - \hat{\delta}^{\text{orc}}(p)) = -\frac{m_a(p)'\sqrt{N}(\hat{\beta} - \beta)}{B_a^*(p)} + g_a(p)'\sqrt{N}(\hat{r}_a - r_0) + o_p(1)$$

uniformly in p . Substitute $\sqrt{N}(\hat{\beta} - \beta) = Q^{-1}S_N + o_p(1)$ and $\sqrt{N}(\hat{r}_a - r_0) = -Q^{-1}\sqrt{N}\Delta_N r_0 + o_p(1)$ from (22). Bounded coefficient paths and the slope floor give U3 with $\zeta_a(p) = Q^{-1}g_a(p)$, without a uniform second-order assumption on the population envelope. $\square$

B.5. Proof of Theorem 2

Proof of Theorem 2. Steps 1–2 expand the oracle root, Step 3 adds the first-stage corrections and Step 4 establishes weak convergence.

Step 1: uniform oracle criterion expansion on the root path. Write $\psi^0(Y, Z; \lambda, \xi, p) = (Y + \lambda Z - \xi)_+ + p\xi - \lambda Z$, which at $(\delta(p), \xi_a^*(p), p)$ is the mean-zero oracle score $\psi_a^*(p)$. Apply the profiling sandwich in the proof of Lemma 5: threshold consistency is now uniform by Lemma 13, and Donsker equicontinuity holds on the compact budget-indexed class of Lemma 10. Thus profiling costs $o_p(N^{-1/2})$ uniformly. The rounding terms are bounded by C/N on the localized threshold interval, since $\sup_p |p_N(p) - p| \leq N^{-1}$. Consequently

$$\sqrt{N}f_N(\delta(p); p) = \mathbb{G}_N \psi^0(\cdot; \delta(p), \xi_a^*(p), p) + o_p(1) \quad \text{uniformly in } p \in \mathcal{D}. \quad (31)$$

Step 2: uniform root expansion. The uniform retained-denominator bound in Lemma 12 and (31) first give $\sup_p \sqrt{N} |\hat{\delta}^{\text{orc}}(p) - \delta(p)| = O_p(1)$. Both arguments lie in a shrinking root tube, where the empirical slope sandwich (30) applies even across kinks and ties. Using $f_N(\hat{\delta}^{\text{orc}}(p); p) = 0$ and the positive slope floor therefore gives

$$\sqrt{N}(\hat{\delta}^{\text{orc}}(p) - \delta(p)) = \frac{\sqrt{N}f_N(\delta(p); p)}{B_a^*(p)} + o_p(1) = \frac{\mathbb{G}_N \psi^0(\cdot; \delta(p), \xi_a^*(p), p)}{B_a^*(p)} + o_p(1), \quad (32)$$

uniformly in p .

Step 3: first-stage corrections, uniformly in p . The feasible criterion is built from $(\hat{s}_i, \hat{Z}_i)$, i.e. from $\hat{u}_i$ and $\hat{r}_a$ rather than (u_i, r_a) . The pointwise decompositions of Theorem 1 and Theorem 5 express the resulting shift in (32) as

$$-\frac{1}{B_a^*(p)} \mathbb{G}_N [m_a(p)' Q^{-1} x u] - \mathbb{G}_N [(r_a' x)(\zeta_a(p)' x)] + o_p(1),$$

with a remainder uniform in p by Lemma 14. The design term carries no $1/B_a^*(p)$ because its entry in (9) already includes $B_a^*(p)$. It is centered since $\mathbb{E}[(r_a' x)(\zeta_a(p)' x)] = g_a(p)' r_a = 0$ by Theorem 5. Combining these corrections with (32) gives $\mathbb{G}_N [\Psi_a(\cdot; p)]/B_a^*(p)$ and the uniform expansion of the theorem.

Step 4: weak convergence. The leading term is $\nu_N(p) := \mathbb{G}_N [\Psi_a(\cdot; p)]/B_a^*(p)$, where $p \mapsto \Psi_a(\cdot; p)$ traces a continuous path through the fixed Donsker class assembled in Lemma 10: its parameters $(\delta(p), \xi_a^*(p), p, m_a(p), \zeta_a(p))$ are continuous in p by

Lemma 1 and Assumption 7. Finite-dimensional convergence to the Gaussian vector with covariance (15) is the multivariate CLT; tightness in $\ell^\infty(\mathcal{P})$ follows because ν_N is the restriction of $\mathbb{G}_N$ to a Donsker class scaled by $1/B_a^*(p)$, which is continuous and bounded above and below. The intrinsic semimetric $\rho(p, q) := \{\mathbb{E}[(\Psi_a(p)/B_a^*(p) - \Psi_a(q)/B_a^*(q))^2]\}^{1/2}$ is continuous on $\mathcal{P} \times \mathcal{P}$: the parameter path is continuous, ψ^0 is almost surely continuous in its parameters, and the envelope of Lemma 10 dominates, so dominated convergence applies. Hence the limit $\mathbb{G}$ is tight with almost surely uniformly ρ -continuous—and therefore uniformly continuous—sample paths.

Budget rounding is included in Step 1; the denominator floor makes its root effect $O(N^{-1})$, negligible at the scale of the theorem. $\square$

B.6. Proof of Proposition 1

For the multiplier approximation write the standardized population and fitted score classes as

$$\mathcal{G}_\Psi := \left\{ \frac{\Psi_a(\cdot; p)}{B_a^*(p)} : p \in \mathcal{P} \right\}, \quad \hat{\mathcal{G}}_{\Psi, N} := \left\{ \frac{\hat{\Psi}_a(\cdot; p)}{\hat{B}_a^*(p)} : p \in \mathcal{P} \right\},$$

whose fitted-class entropy bound is the conclusion of Lemma 9. Let $\mathbb{G}_N^\circ$ denote the multiplier process (16) built with the population scores $\Psi_{a,i}(p)$ and slope $B_a^*(p)$, and write $\Delta_i(p) := \Psi_{a,i}(p)/B_a^*(p)$. Two facts about this oracle process are proved once here and used twice below. First, $\mathcal{G}_\Psi$ is Donsker: its oracle layer is the class of Lemma 10, and the residual and design layers of (9) are that class augmented by finitely many fixed functions with continuous parametric coefficients, which the lemma's final clause allows, with envelope

$$\tilde{L} := C(1 + \|x\| \|u\| + \|x\|^2) \in L^2(P),$$

where C depends on the bounded population coefficient paths and $\underline{B} := \inf_p B_a^*(p) > 0$. This is a genuine score envelope, including the Y dependence, not just the parameter Lipschitz modulus of Lemma 10; its square-integrability is that of the envelope H noted after Assumption 6. The fixed compact score family of Lemma 9 has a Glivenko–Cantelli product class with integrable envelope $\tilde{L}^2$, by a finite parameter mesh and the law of large numbers. Thus the empirical second moments $P_N[\Delta(p)\Delta(q)]$ converge uniformly to the kernel K of (15). Since $P\Delta(p) = 0$, the score uniform law also

gives $\sup_p |P_N \Delta(p)| = o_p(1)$, so the centered conditional covariances $P_N[\Delta(p)\Delta(q)] - P_N \Delta(p)P_N \Delta(q)$ have the same limit. The canonical semimetric is therefore uniformly close to ρ . The conditional multiplier central limit theorem for Donsker classes (van der Vaart and Wellner, 1996, Theorem 2.9.6) gives, in probability,

$$\begin{aligned} & \mathbb{G}_N^\circ \rightsquigarrow \mathbb{G} \quad \text{in } \ell^\infty(\mathcal{P}) \text{ conditionally,} \\ \lim_{\varrho \downarrow 0} \limsup_N \mathbb{P}^* \left\{ \sup_{\rho(p,q) < \varrho} |\mathbb{G}_N^\circ(p) - \mathbb{G}_N^\circ(q)| > \varepsilon \right\} &= 0 \quad \text{for every } \varepsilon > 0, \end{aligned} \quad (33)$$

the second display being the asymptotic equicontinuity of the multiplier process, which the conditional convergence in $\ell^\infty(\mathcal{P})$ delivered by that theorem entails, with ρ the semimetric of Step 4 of the proof of Theorem 2.

Lemma 15 (Uniform directional plug-ins). *Under the conditions of Lemma 12, the fitted root, threshold, slope and moment coefficients in (12) converge uniformly to their population counterparts. In particular,*

$$\sup_p |\hat{B}_a^*(p) - B_a^*(p)| \xrightarrow{p} 0, \quad \sup_p P_N[(\hat{\Psi}_a(p) - \Psi_a(p))^2] \xrightarrow{p} 0,$$

and the fitted parameter localization required by Lemma 9 holds.

Proof. Lemmas 12 and 13 give root and threshold consistency and weighted selected-set consistency. The oracle selected moment weights $x(r'_a x)$, xu and Z are integrable and bounded by a constant times H . Their weighted uniform laws and selected-set consistency give uniform convergence of $\hat{m}_a$, $\hat{g}_{a*}$ and $\hat{B}_a^*$; replacing the residual and direction inside these moments costs at most $o_p(1)P_N H$. The slope floor in U2 and the formulas (11) then give uniform consistency of $\hat{g}_a$ and $\hat{\zeta}_a$. The first-stage errors obey $\hat{\beta} - \beta = O_p(N^{-1/2})$ and $\hat{r}_a - r_a = O_p(N^{-1/2})$ by the expansion (22), so both coefficient blocks of θ are stochastically bounded. All population coefficient paths are bounded and continuous, $\hat{Q}_N \rightarrow Q \succ 0$ and $a'\hat{r}_a \rightarrow \mu_Z(a) > 0$, so the fitted parameters lie uniformly in a deterministic compact with both denominators bounded away from zero on events of probability tending to one. The pointwise Lipschitz inequality of Lemma 9 and uniform parameter consistency give a common $o_p(1)H$ bound for the difference of standardized scores. Since $P_N H^2 \rightarrow \mathbb{E}H^2$, this proves empirical- L^2 convergence; slope consistency and the bounded score envelope give the displayed unstandardized convergence as well. The same compact localization and parameter mesh verify the localization hypothesis of Lemma 9. Squares and pairwise products

of the fixed score family obey uniform laws, using its integrable H^2 envelope and a finite parameter mesh, so the variance plug-ins are uniformly consistent too. $\square$

Lemma 16 (Fitted multiplier class). *Write $\hat{\Delta}_i(p) := \hat{\Psi}_{a,i}(p)/\hat{B}_a^*(p)$, $\eta_i(p) := \hat{\Delta}_i(p) - \Delta_i(p)$, $\bar{\Delta}(p) := N^{-1} \sum_j \hat{\Delta}_j(p)$, $\bar{\Delta}(p) := N^{-1} \sum_j \Delta_j(p)$ and $\bar{\eta}(p) := N^{-1} \sum_j \eta_j(p)$, so that the difference of the two multiplier processes of (16) is $Z_N(p) := \hat{\mathbb{G}}_N^*(p) - \mathbb{G}_N^\circ(p) = N^{-1/2} \sum_i e_i \{\eta_i(p) - \bar{\eta}(p)\}$. Under Lemma 9, U2, and the uniform consistencies delivered by Lemma 15:*

$$(i) \sup_{p \in \mathcal{P}} \|\eta(p)\|_{N,2} \xrightarrow{P} 0.$$

$$(ii) \sup_{p \in \mathcal{P}} \left| \hat{\mathbb{G}}_N^*(p) - \mathbb{G}_N^\circ(p) \right| \xrightarrow{P} 0.$$

(iii) *the conditional law of $\hat{\mathbb{G}}_N^*$ given the data inherits from that of $\mathbb{G}_N^\circ$ its asymptotic ρ -equicontinuity, for the semimetric ρ of Step 4 of the proof of Theorem 2:*
 $\lim_{\varepsilon \downarrow 0} \limsup_N \mathbb{P}^* \{ \sup_{\rho(p,q) < \varepsilon} \left| \hat{\mathbb{G}}_N^*(p) - \hat{\mathbb{G}}_N^*(q) \right| > \varepsilon \} = 0$ for every $\varepsilon > 0$.

Proof. Throughout, conditional statements are given the sample, with the outer-probability convention of the appendix. The population path is ρ -continuous by Step 4 of the proof of Theorem 2.

(i). The standardized empirical- L^2 convergence follows directly from Lemma 15, its slope floor and the common L^2 score envelope. In particular $\sup_p \|\hat{\Delta}(p)\|_{N,2} = O_p(1)$.

(ii). Conditional on the sample, Z_N is a centered Gaussian process, being the linear image of the standard normal vector e with coefficients fixed by the data. Its canonical semimetric is the empirical variance of the coefficient increment,

$$d_N^2(p, q) = \frac{1}{N} \sum_i \left[\eta_i(p) - \eta_i(q) - \{\bar{\eta}(p) - \bar{\eta}(q)\} \right]^2 \leq \|\eta(p) - \eta(q)\|_{N,2}^2,$$

so centering can only shrink the metric, and $\text{diam}(\mathcal{P}, d_N) \leq 2D_N$, where $D_N := \sup_p \|\eta(p)\|_{N,2} \xrightarrow{P} 0$ by (i). The linear map $h \mapsto h - (N^{-1} \sum_i h_i) \mathbf{1}$ is an empirical- L^2 contraction. Consequently a cover of the algebraic difference of the fitted and population classes gives a cover of the centered difference class, and

$$\begin{aligned} \log \mathcal{N}(\varepsilon, \{\eta(p) - \bar{\eta}(p) \mathbf{1} : p \in \mathcal{P}\}, \|\cdot\|_{N,2}) \\ \leq \log \mathcal{N}(\varepsilon/4, \hat{\mathcal{G}}_{\Psi,N}, \|\cdot\|_{N,2}) + \log \mathcal{N}(\varepsilon/4, \mathcal{G}_\Psi, \|\cdot\|_{N,2}). \end{aligned}$$

The fixed factor four allows recentering at a member of the indexed class if covers are required to have internal centers. The first term is bounded by $\psi(\varepsilon/4)$ on the

event of Lemma 9; the second is bounded directly by the population parameter mesh of Lemma 9, not by transferring population brackets. Intersect these events with $\{2D_N \leq 1\}$ to obtain $\mathcal{C}_N^\dagger$ with probability tending to one. On this event both entropies are dominated at every $\varepsilon \in (0, 1]$ by a deterministic non-increasing function $\psi^\dagger$ with integrable square root. For $\eta > 0$ the conditional Markov inequality splits the target probability at $\mathcal{C}_N^\dagger$,

$$\mathbb{P}^* \left\{ \sup_p |Z_N(p) - Z_N(p_0)| > \eta \right\} \leq \frac{1}{\eta} \mathbb{E} \left[\mathbf{1}_{\{\mathcal{C}_N^\dagger\}} \mathbb{E}^* \sup_p |Z_N(p) - Z_N(p_0)| \right] + \mathbb{P}\{\mathcal{C}_N^{\dagger c}\},$$

and on $\mathcal{C}_N^\dagger$ Dudley's entropy bound for centered Gaussian processes (van der Vaart and Wellner, 1996, Chapter 2), applied conditionally, bounds the inner supremum, for any fixed $p_0 \in \mathcal{P}$, by

$$K \int_0^{u_N} \{1 + \psi^\dagger(\varepsilon)\}^{1/2} d\varepsilon, \quad u_N := 2D_N \wedge 1.$$

On $\mathcal{C}_N^\dagger$, $u_N = 2D_N$ bounds the actual diameter; off that event only its vanishing probability is used. The integral is a bounded function of u_N tending to zero as $u_N \rightarrow 0$, so its expectation vanishes and conditional Markov gives $\sup_p |Z_N(p) - Z_N(p_0)| \xrightarrow{P} 0$. At p_0 the conditional variance is at most $D_N^2 \xrightarrow{P} 0$, hence $Z_N(p_0) \xrightarrow{P} 0$ by conditional Chebyshev. This proves (ii).

(iii). Every fitted increment is bounded by the corresponding oracle increment plus $2 \sup_p |Z_N(p)|$. Apply (33) and part (ii); joint negligibility also gives conditional negligibility in probability by Markov's inequality. No continuity bound for the fitted budget path is needed. $\square$

Proof of Proposition 1. By (33) and the sup-norm replacement in Lemma 16, conditional Slutsky in the bounded-Lipschitz metric gives $\hat{\mathbb{G}}_N^* \rightsquigarrow \mathbb{G}$ in $\ell^\infty(\mathcal{P})$ conditionally on the data, in probability.

Studentization is uniformly consistent by Lemmas 15 and 16: the plug-in lemma supplies uniform score and variance consistency, while U2 bounds the population slope and variance away from zero. Hence $\sup_p |\hat{\mathbb{G}}_N^*(p)| / \hat{\omega}_a(p)$ converges conditionally in distribution to $\sup_p |\mathbb{G}(p)| / \omega_a(p)$, the limit law in Theorem 2.

That limit is the supremum of the absolute value of a tight continuous Gaussian process whose standard deviation is bounded away from zero by U2, so its distribution function F is continuous (Tsirel'son, 1976). Conditional weak convergence in

probability and Pólya's theorem therefore give uniform convergence in probability of the conditional distribution functions to F , hence $F(\hat{q}_{1-\alpha}) \rightarrow 1 - \alpha$ in probability. For any $0 < \varepsilon < \min(\alpha, 1 - \alpha)$ choose deterministic $q_- < q_+$ with $F(q_-) = 1 - \alpha - \varepsilon$ and $F(q_+) = 1 - \alpha + \varepsilon$; with probability tending to one $q_- \leq \hat{q}_{1-\alpha} \leq q_+$, so the coverage event is sandwiched between the events using these fixed critical values, whose probabilities tend to $F(q_-)$ and $F(q_+)$ by Theorem 2, the continuous mapping theorem and Slutsky. Sending $\varepsilon \downarrow 0$ gives

$$\mathbb{P}\left(\sup_{p \in \mathcal{P}} \sqrt{N} \left| \hat{\delta}(p) - \delta(p) \right| / \hat{\omega}_a(p) \leq \hat{q}_{1-\alpha}\right) \rightarrow 1 - \alpha,$$

which is the band statement. Reading $\hat{q}_{1-\alpha}$ as an estimate of the population quantile $\inf\{x : F(x) \geq 1 - \alpha\}$ is a stronger conclusion, needing in addition that F be strictly increasing at that point; the coverage statement above does not ask for it. $\square$

B.7. Proof of Proposition 2

Proof of Proposition 2. Write $\Delta(p) = \Psi_a(p)/B_a^*(p)$. The population class $\{\Delta(p) : p \in \mathcal{P}\}$ is Donsker by the proof of Theorem 2; adjoining the single mean-zero L^2 function φ_c gives joint convergence of the curve process and $\mathbb{G}_N \varphi_c$. Consequently $\sqrt{N}(\hat{g} - g)$ converges in $\ell^\infty(\mathcal{P})$ to the continuous Gaussian process $\mathbb{G}_g(p) = \mathbb{G}(p) - Z_c$, where $(\mathbb{G}, Z_c)$ has the covariance of $(\Delta(p), \varphi_c)$. In particular its covariance is $\mathbb{E}[h(p)h(q)]$ and its variance path $\sigma_g^2(p)$ is continuous.

The oracle multiplier process based on $h(p)$ has the same conditional joint limit by the multiplier theorem for this augmented Donsker class. The fitted curve replacement is uniformly negligible by Lemma 16. The margin replacement is constant across budgets and its conditional variance is at most $P_N(\hat{\varphi}_c - \varphi_c)^2 = o_p(1)$, where P_N is the empirical average. Conditional Chebyshev therefore makes the fitted gap replacement uniformly negligible as well; no entropy condition on the single fitted margin score is needed.

By Lemma 15 and assumption S1, $\sup_p P_N(\hat{h}(p) - h(p))^2 = o_p(1)$. Uniform second-moment laws for $h(p)$ follow from the curve's compact score family: its squared envelope is integrable, and the cross-products with φ_c have an integrable envelope by Cauchy–Schwarz. A finite parameter mesh and dominated convergence give their uniform laws. Another application of Cauchy–Schwarz now gives $\sup_p |\hat{\sigma}_g^2(p) - \sigma_g^2(p)| = o_p(1)$. The assumed positive variance floor permits uniform studentization.

The supremum of $|\mathbb{G}_g(p)|/\sigma_g(p)$ has a continuous distribution by the same Gaussian-supremum result used in the proof of Proposition 1. The conditional distribution-function convergence and deterministic critical-value sandwich in that proof apply to the gap process and give the stated coverage. $\square$

B.8. Proof of Theorem 3

Proof of Theorem 3. Write $\kappa_{\dagger} = \partial_p \delta(p^{\dagger}) > 0$ and $H_N(p) = \sqrt{N}\{\hat{g}(p) - g(p)\}$. The augmented Donsker argument in the proof of Proposition 2 gives tightness and asymptotic equicontinuity of H_N , independently of the uniform variance floor needed for studentized bands. The crossing theorem requires positive gap variance only at $p^{\dagger}$.

Uniform consistency of $\hat{g}$ and the strict interior crossing first give $\hat{p}^{\dagger} \xrightarrow{p} p^{\dagger}$. On a fixed neighborhood of $p^{\dagger}$, the positive continuous derivative bounds $|\dot{g}(p)|$ below by $(\kappa_{\dagger}/2)|p - p^{\dagger}|$, with the sign of $p - p^{\dagger}$. Since $\sup_p |H_N(p)| = O_p(1)$, the first crossing lies within $O_p(N^{-1/2})$ of $p^{\dagger}$. For every fixed M , differentiability and equicontinuity yield

$$\sup_{|t| \leq M} \left| \sqrt{N} \hat{g}(p^{\dagger} + t/\sqrt{N}) - \{\kappa_{\dagger}t + H_N(p^{\dagger})\} \right| = o_p(1).$$

The affine approximation has positive slope. On an event where $|H_N(p^{\dagger})|$ and the crossing displacement are bounded, this uniform approximation places the first crossing between $-H_N(p^{\dagger})/\kappa_{\dagger} - \varepsilon$ and $-H_N(p^{\dagger})/\kappa_{\dagger} + \varepsilon$, for every fixed $\varepsilon > 0$, with probability tending to one. Tightness lets the bound tend to infinity and then ε tend to zero, proving

$$\sqrt{N}(\hat{p}^{\dagger} - p^{\dagger}) = -\frac{H_N(p^{\dagger})}{\kappa_{\dagger}} + o_p(1) \xrightarrow{d} \mathcal{N}\left(0, \frac{\sigma_g^2(p^{\dagger})}{\kappa_{\dagger}^2}\right).$$

This argument allows the empirical step curve, whose monotonicity in the budget is deterministic (Lemma 17), and does not require its first crossing to solve an exact equality. Using (14) gives the stated crossing score and variance.

The uniform empirical- L^2 consistency of $\hat{h}$, its variance consistency and Lemma 15 give consistent variance and cutoff estimates at the random budget $\hat{p}^{\dagger}$. The population paths are continuous, and Lemma 11 bounds the cutoff away from zero near $p^{\dagger}$. Thus $\sqrt{N} \widehat{\text{SE}}(\hat{p}^{\dagger}) \xrightarrow{p} \sigma_{\dagger}$; the positive crossing variance and Slutsky give Wald coverage. $\square$

B.9. A deterministic step path in the budget

Lemma 17 (The empirical curve is nondecreasing in the budget). *On the normalized pair of (3), write $\hat{s}_i = \hat{s}_i(a)$, $\hat{Z}_i = \hat{Z}_i(a)$, $\hat{W}_i(\lambda) = \hat{s}_i + \lambda \hat{Z}_i$ and $h_{Z,j} := T_{N,j/N}(\hat{Z})$, and let $\hat{W}_{(1)} \leq \dots \leq \hat{W}_{(N)}$ denote the order statistics of a modified score array. For $k \geq 1$ with $h_{Z,k+1} < 1$, hence also $h_{Z,k} < 1$, Proposition 3 makes $\hat{\delta}(j/N)$ the unique root of $f_{N,j}(\lambda) := T_{N,j/N}(\hat{s} + \lambda \hat{Z}) - \lambda$ at $j = k, k+1$; with $\hat{m}_k := \hat{W}_{(N-k)}(\hat{\delta}(k/N))$, the largest modified score among the observations retained at a budget- k optimum,*

$$\hat{\delta}((k+1)/N) \geq \hat{\delta}(k/N), \quad \frac{\hat{m}_k}{N} \leq \hat{\delta}((k+1)/N) - \hat{\delta}(k/N) \leq \frac{\hat{m}_k}{N(1 - h_{Z,k+1})},$$

with $\hat{m}_k \geq 0$, and the step is flat if and only if every retained modified score at the budget- k root is zero. Neither the statistic nor its increments depends on how cutoff ties are allocated among observations.

Proof of Lemma 17. The pair $(\hat{s}_i, \hat{Z}_i)$ satisfies $\hat{Z}_i \geq 0$ and $N^{-1} \sum_i \hat{s}_i = N^{-1} \hat{r}'_a \sum_i x_i \hat{u}_i = 0$ by the sample normal equations, and $N^{-1} \sum_i \hat{Z}_i = 1$ by the normalization in (3); hence $N^{-1} \sum_i \hat{W}_i(\lambda) = \lambda$ for every λ . Write $\lambda_j := \hat{\delta}(j/N)$ and $f_j := f_{N,j}$, and let S_k be any size- k set of largest modified scores at λ_k . The root identity $f_k(\lambda_k) = 0$ says $N^{-1} \sum_{i \in S_k} \hat{W}_i(\lambda_k) = \lambda_k$, so subtraction gives $N^{-1} \sum_{i \notin S_k} \hat{W}_i(\lambda_k) = 0$: the retained modified scores average zero, their maximum is nonnegative, $\hat{m}_k \geq 0$, and that maximum is zero exactly when all $N - k$ of them vanish. Adding the next order statistic gives $f_{k+1}(\lambda_k) = \hat{m}_k/N$.

As the maximum of affine functions of λ , $T_{N,(k+1)/N}(\hat{s} + \lambda \hat{Z})$ is convex and piecewise linear; its active subgradients are $N^{-1} \sum_{i \in S} \hat{Z}_i$ over the size- $(k+1)$ sets S that rank first at some λ , hence lie between zero and $h_{Z,k+1}$. So f_{k+1} is convex, continuous and strictly decreasing with subgradients in $[-1, -(1 - h_{Z,k+1})]$: under $h_{Z,k+1} < 1$, part (ii) of Proposition 3 read at budget $k+1$ gives a unique root λ_{k+1} , and $0 = f_{k+1}(\lambda_{k+1}) \leq f_{k+1}(\lambda_k) = \hat{m}_k/N$ implies $\lambda_{k+1} \geq \lambda_k$. The secant slope of f_{k+1} over $[\lambda_k, \lambda_{k+1}]$ is bounded by the extreme active subgradients, so evaluating $f_{k+1}(\lambda_k) - f_{k+1}(\lambda_{k+1})$ two ways gives

$$\frac{\hat{m}_k}{N} \leq \lambda_{k+1} - \lambda_k \leq \frac{\hat{m}_k}{N(1 - h_{Z,k+1})},$$

and the left inequality makes the step flat only if $\hat{m}_k = 0$, while a zero $\hat{m}_k$ makes λ_k itself the budget- $(k+1)$ root. Nothing in the argument depends on the allocation of

ties, since f_j and $\widehat{m}_k$ are functions of order statistics alone, and $h_{z,j}$ is nondecreasing in j , so the premise at $k + 1$ delivers the denominator event of Proposition 3 at every compared budget. The largest admissible budget is the last one with $h_{z,j} < 1$, at most $N - 1$: at $j = N$ nothing is retained and the criterion slope bound collapses. $\square$

Remark 5 (Discrete slope bounds and bisection over k). *The increment bound in Lemma 17 is the discrete counterpart of the population slope $\partial_p \delta(p) = \xi_a^*(p)/B_a^*(p)$ in (14): $\widehat{m}_k \leq \hat{\xi}(k/N)$, the k -th largest modified score, and $1 - h_{z,k+1}$ lower-bounds the retained directional leverage. The exact path $k \mapsto \hat{\delta}(k/N)$ is nondecreasing on its admissible budgets, starting from $\hat{\delta}(0) = 0$, so its first crossing of a budget-invariant threshold $\hat{c}$ among $n_{\mathcal{P}}$ integer budgets in $\mathcal{P}$ can be found by binary search in at most $\lceil \log_2 n_{\mathcal{P}} \rceil + 1$ root evaluations; each evaluation still requires solving the ratio problem of Section A.1. Absolute root error at most ε at each budget permits downward numerical steps of at most 2ε and leaves comparisons within ε of $\hat{c}$ unresolved without further refinement. By contrast, the band endpoints $\hat{\delta}(k/N) \pm \hat{q}_{1-\alpha} \hat{\omega}_a(k/N)/\sqrt{N}$ need not be monotone because the studentizer varies with the budget, so computing the multiplier critical value and inverting the simultaneous band still require evaluating every integer budget in $\mathcal{P}$.*

B.10. Crossing brackets from a finite reporting grid

When simultaneous intervals are evaluated only on a finite grid, monotonicity of $\delta(p)$ still yields a conservative bracket for the breakdown fraction without between-grid interpolation.

Corollary 4 (Finite-grid crossing bracket). *Let $p_1 < \dots < p_K$ ($K \geq 2$) be a fixed grid, $c > 0$ be fixed, and $\delta(p)$ be nondecreasing on $[p_1, p_K]$. Given simultaneous intervals $[L_j, U_j]$ satisfying $\liminf_{N \rightarrow \infty} \mathbb{P}\{L_j \leq \delta(p_j) \leq U_j \ \forall j\} \geq 1 - \alpha$, define $p^\dagger = \inf\{p \in [p_1, p_K] : \delta(p) \geq c\}$ (with $\inf \emptyset = p_K$) and*

$$\ell_N = \max(\{p_1\} \cup \{p_j : U_j < c\}), \quad u_N = \min(\{p_K\} \cup \{p_j : L_j \geq c\}).$$

Then $\liminf_{N \rightarrow \infty} \mathbb{P}\{\ell_N \leq p^\dagger \leq u_N\} \geq 1 - \alpha$, and $U_K < c$ certifies $\delta(p) < c$ throughout $[p_1, p_K]$ on the coverage event.

Proof. On the joint coverage event, $U_j < c \implies \delta(p_j) < c \implies p^\dagger \geq p_j$ by monotonicity, and $L_j \geq c \implies \delta(p_j) \geq c \implies p^\dagger \leq p_j$. Taking the maximum lower bound and

minimum upper bound (respecting the endpoint conventions when either set is empty) yields $\ell_N \leq p^\dagger \leq u_N$. Similarly, $U_K < c$ implies $\delta(p) \leq \delta(p_K) \leq U_K < c$ on $[p_1, p_K]$. $\square$

Under Assumption 7, restricting the multiplier maximum in Proposition 1 to $\{p_1, \dots, p_K\}$ yields valid simultaneous grid intervals $[L_j, U_j]$, and Lemma 1 supplies monotonicity. Note that ℓ_N is the last grid point with $U_j < c$ (rather than the first with $U_j \geq c$), which guards against crossings occurring between grid points. For the reported step path itself the analogue warrant is Lemma 17: two reported budgets bracket its crossing of a budget-independent threshold.

Online Supplement: Inference for Directional Population Sensitivity

Lucas D. Konrad and Nikolas Kuschnig

October 19, 2026

The supplement states and proves the primitive regression verifications, develops the unrestricted-refitting bridge and reports additional Monte Carlo, sample-audit and calibrated population-geometry detail. References to results outside this supplement refer to the main paper.

S1 Estimated conclusion margins

The paired procedure allows the conclusion margin and the directional curve to be estimated from the same observations $O_i = (x_i, y_i)$. The following assumption states the additional conditions on the margin estimator and its fitted influence score.

Assumption S1 (Asymptotic linearity of the conclusion margin). *For a population margin $c = c(P) > 0$, the estimator $\hat{c}$ and fitted influence scores $\hat{\varphi}_{c,i}$ satisfy*

$$\begin{aligned} \sqrt{N}(\hat{c} - c) &= \frac{1}{\sqrt{N}} \sum_{i=1}^N \varphi_c(O_i) + o_p(1), & \mathbb{E}\varphi_c &= 0, & \mathbb{E}\varphi_c^2 &< \infty, \\ \frac{1}{N} \sum_{i=1}^N \{\hat{\varphi}_{c,i} - \varphi_c(O_i)\}^2 &= o_p(1). \end{aligned} \tag{S1}$$

A fixed, known margin is the special case $\hat{c} = c$ and $\varphi_c = \hat{\varphi}_c = 0$.

Contrast margins and empirical orientation. For a prespecified contrast $a_0'\beta \neq b$, let $\iota = \text{sign}(a_0'\beta - b)$ orient the adverse direction as $a = \iota a_0$. The population margin, its

plug-in estimate, and its influence score are

$$c(P) = |a'_0\beta - b|, \quad \hat{c} = |a'_0\hat{\beta} - b|, \quad \varphi_c(x, y) = \iota(a'_0Q^{-1}x)u = (r'_a x)u.$$

In sample computation, use $\hat{\iota} = \text{sign}(a'_0\hat{\beta} - b)$, $\hat{a} = \hat{\iota}a_0$, and $\hat{\varphi}_{c,i} = (\hat{r}'_{\hat{a}}x_i)\hat{u}_i$, evaluating the fitted curve and its scores in direction $\hat{a}$. Because $c(P) > 0$ and OLS consistency imply $\mathbb{P}(\hat{\iota} = \iota) \rightarrow 1$, this empirical orientation shares the same asymptotic law as the fixed population direction a . The maintained moment and OLS consistency conditions deliver the empirical- L^2 consistency in (S1). A zero population margin ($a'_0\beta = b$) lies outside this regular regime: the absolute-value map is not differentiable there, and the estimated sign need not converge to a fixed orientation.

Other margin specifications. Conclusion margins can also measure the distance to a nominal significance boundary, either overturning or creating apparent evidence: with a fixed reference scale s_0 and critical value z , the downward and upward creation margins for a coefficient β_T inside $[-zs_0, zs_0]$ are $\beta_T + zs_0$ and $zs_0 - \beta_T$. These margins inherit the coefficient's influence score with the corresponding sign; an estimated reference scale requires its own joint expansion, and crossing a nominal boundary after outcome-dependent deletion does not establish selection-adjusted significance. Only strictly positive margins satisfying Assumption S1 are covered by the regular paired procedure.

Pointwise gap inference. For the gap and paired score in (17), Theorem 2 and assumption S1 give

$$\sqrt{N}\{\hat{g}(p) - g(p)\} = N^{-1/2} \sum_{i=1}^N h_i(p) + o_p(1)$$

uniformly on $\mathcal{P}$. At a single prespecified share p , the pointwise expansion instead requires only Assumption 1(a) and Assumption S1. With $\text{Var}(h(p)) > 0$, the pointwise gap standard error is $\hat{\sigma}_g(p)/\sqrt{N}$, where $\hat{\sigma}_g^2(p) = \widehat{\text{Var}}(\hat{h}(p))$ incorporates the covariance between the fragility curve and the estimated margin. This gives the Wald interval $\hat{g}(p) \pm z_{1-\alpha/2}\hat{\sigma}_g(p)/\sqrt{N}$ and a one-sided test of $g(p) \leq 0$, rejecting when $\sqrt{N}\hat{g}(p)/\hat{\sigma}_g(p) > z_{1-\alpha}$. Choosing a share after inspecting the path instead requires simultaneous inference, as in Proposition 2.

S2 Primitive Regression Verifications

S2.1 Continuous and mixed covariate laws

Proposition S1 (Continuous- and mixed-control primitive verification). *Let (x, y) be i.i.d. in fixed dimension, fix $a \neq 0$, and suppose $Q = \mathbb{E}[xx'] > 0$, $\mathbb{E}\|x\|^4 < \infty$, and $P(x'Q^{-1}a = 0) = 0$. Write $y = m(x) + \sigma_0(x)\varepsilon$ with measurable m, σ_0 satisfying*

$$\mathbb{E}[(1 + \|x\|^2)m(x)^2] < \infty, \quad 0 < \sigma_- \leq \sigma_0(x) \leq \sigma_+ < \infty.$$

Suppose $\mathbb{E}[\varepsilon \mid x] = 0$, and $\varepsilon \mid x$ has a jointly measurable regular conditional density $h_x(t)$, strictly positive and continuous in t for almost every x , with deterministic constants $M_h, V < \infty$ such that

$$\sup_{t \in \mathbb{R}} (1 + t^2)h_x(t) \leq M_h, \quad t^2 h_x(t) \longrightarrow 0 \quad (|t| \longrightarrow \infty), \quad \mathbb{E}[\varepsilon^2 \mid x] \leq V \quad \text{a.s.}$$

The tail limit is pointwise for almost every x , not necessarily uniform in x . With $\beta = Q^{-1}\mathbb{E}[xm(x)]$, the standing assumptions Assumptions 2 to 5 and Assumption 6 hold at every $p \in (0, 1)$, and Assumption 7 holds on every compact $\mathcal{P} \subset (0, 1)$. Consequently the pointwise intervals and functional multiplier bands apply.

Proof of Proposition S1. Write $r_0 = Q^{-1}a$, $\mu = a'r_0$, $v = r_0'x$, $Y = vu$, and $Z = v^2/\mu$. The bounded positive scales, fourth design moment and exact mixed moment $\mathbb{E}[(1 + \|x\|^2)m(x)^2] < \infty$ imply

$$\beta = Q^{-1}\mathbb{E}[xm(x)], \quad \mathbb{E}[xu] = 0, \quad 0 < \mathbb{E}u^2 < \infty, \quad \mathbb{E}[\|x\|^2 u^2] < \infty.$$

Thus $\mathbb{E}Y = 0$, $\mathbb{E}Z = 1$, $\mathbb{E}(Y^2 + Z^2) < \infty$, and $H = 1 + \|x\| |u| + \|x\|^2 \in L^2$, proving the standing moments and C1.

Leverage and roots before cutoff smoothness. Fix $\mathcal{P} = [\underline{p}, \bar{p}] \subset (0, 1)$ and put

$$b_0 = \int_0^{1-\bar{p}} F_Z^{-1}(t) dt > 0.$$

Here positivity follows from $P(v = 0) = 0$. Bottom-quantile allocation gives $\mathbb{E}[ZR] \geq b_0$ for every intensity $0 \leq R \leq 1$ with $\mathbb{E}R = 1 - p$, $p \leq \bar{p}$. This is all-rule scalar leverage, not retained-matrix rank. Conditional full support makes Y nonconstant with both tails; its mean zero implies $\mathcal{T}_p(Y) > 0$ for every proper upper tail. Hence Assumptions 3

and 4 and Proposition 4 give a unique positive relaxed root $\delta(p) \leq \mathbb{E}|Y|/b_0$. Quantile integration is continuous in p for integrable scores, and $|\mathcal{T}_p(W_\lambda) - \mathcal{T}_p(W_{\lambda'})| \leq |\lambda - \lambda'| \mathbb{E}Z$. Joint continuity and strict descent by b_0 therefore give root continuity without cutoff regularity. For every finite λ , the oracle score $W_\lambda = \nu u + \lambda Z$ mixes full-support absolutely continuous laws, since $P(\nu = 0) = 0$. Its CDF is continuous and strictly increasing, proving Assumption 5 for every $\lambda \geq 0$. At the root,

$$\mathbb{E}W_\delta = \delta = \mathcal{T}_p(W_\delta), \quad \int_0^{1-p} F_{W_\delta}^{-1}(t) dt = 0.$$

There is strictly negative score mass. A nonpositive $(1-p)$ -quantile would make the last integral strictly negative. Thus $\xi(p) = \xi_a^*(p) > 0$, and the root region has exactly mass p , identifying $\delta(p) = \delta(p)$.

Joint nuisance regularity through zero indices. Let $\theta = (b, r)$, $c = a'r$, $s = r'x$, $d = m(x) - x'b$, $k = d + \lambda s/c$, and

$$W_{\theta, \lambda} = s(y - x'b) + \lambda s^2/c, \quad t = (w/s - k)/\sigma_0(x) \quad (s \neq 0).$$

Work on compact parameter sets with $c \geq \mu/2$, bounded λ , and $w \in [\eta, M]$, $\eta > 0$. Holding x fixed, the conditional score CDF is $F_x(t)$ for $s > 0$, $1 - F_x(t)$ for $s < 0$, and 1 for $s = 0$. For $s \neq 0$ its derivatives are

$$\begin{aligned} \partial_w F &= h_x(t)/(s|\sigma_0|), & \partial_b F &= \text{sign}(s)h_x(t)x/\sigma_0, \\ \partial_\lambda F &= -|s|h_x(t)/(c\sigma_0), & \nabla_r F &= \frac{\text{sign}(s)h_x(t)}{\sigma_0} \left[\left(-\frac{w}{s^2} - \frac{\lambda}{c} \right) x + \frac{\lambda s}{c^2} a \right]. \end{aligned}$$

They extend continuously by zero through $s = 0$, using $t^2 h_x(t) \rightarrow 0$ pointwise in x . Actual differentiability there follows since, for fixed x and locally small s , the CDF differs from 1 by at most $P(|\varepsilon| > C_x/|s| \mid x) = O(s^2) = o(|s|)$. No density derivative is assumed. The identity $w/s = k + \sigma_0 t$ gives

$$\frac{h_x(t)}{|s|\sigma_0} \leq C(1 + |m(x)| + \|x\|)/\eta, \quad \frac{w}{s^2} h_x(t) \leq C(1 + m(x)^2 + \|x\|^2)/\eta.$$

All first CDF derivatives therefore have the integrable envelope $C\{1 + |m(x)| + \|x\|^3 + \|x\| m(x)^2\}$, even for unbounded controls. Dominated differentiation makes the unconditional CDF jointly C^1 . Its density $\mathbb{E}[h_x(t)/(s|\sigma_0|)]$, with the integrand extended

by zero at $s = 0$, is finite, continuous and positive on positive-score domains: $c > 0$ and $Q > 0$ imply $P(s \neq 0) > 0$. No global density bound at score zero is asserted; nearby nuisance directions may have zero-index atoms.

The quantile implicit function theorem gives locally C^1 nuisance quantiles about each positive root cutoff. Root continuity implies cutoff continuity, hence compactness gives $0 < \min_p \xi(p) \leq \max_p \xi(p) < \infty$. A sufficiently small common compact nuisance/root/cutoff tube lies entirely at positive score coordinates, with density bounded above and away from zero. Quantile bracketing and uniqueness join its local quantile maps, proving C2 and U1.

Leverage, differentiability and empirical expansions. Uniform L^1 continuity of $Z_\theta = (r'x)^2/(a'r)$ preserves retained leverage at least $b_0/2$ on a small nuisance neighborhood, proving C3 and the slope clause of U2. The moments and cutoff regularity already established give C1–C3; Lemma 5 then supplies the criterion equicontinuity and direction differentiability conclusions C4–C5, with the full gradient in (8).

Positive variances and feasible inference. For fixed p , put $\alpha(x) = m_a(p)'Q^{-1}x$. Conditional on x with $v \neq 0$, (9) is continuous and piecewise affine in u , with retained and active slopes $-\alpha(x)$ and $v - \alpha(x)$; the design correction is constant given x . On $\{v \neq 0\}$ both branches have positive conditional probability and their slopes differ by $v \neq 0$, so conditional variance is positive there. Since this event has positive probability, the unconditional variance is positive. The oracle score has active slope v , proving its positivity too. Continuous root/cutoff paths, boundary atomlessness and the common L^2 envelope give continuous selected moments and influence variances. Compactness supplies their positive variance floors, proving C6 and U2. The common cutoff tube, leverage and variance floors and $H \in L^2$ meet the hypotheses of Lemmas 12 and 13. Now Lemma 14 supplies U3, and Lemmas 9, 15 and 16 supply fitted empirical- L^2 score consistency, uniform variance consistency, studentization and multipliers. $\square$

Proposition S2 (Localized primitive verification with zero-index atoms). *Maintain the moment, conditional-mean, scale and conditional-error-density conditions of Proposition S1, except replace $P(x'Q^{-1}a = 0) = 0$ by the following condition. For a compact budget interval $\mathcal{P} = [\underline{p}, \bar{p}] \subset (0, 1)$, let $\rho_0 = P(x'Q^{-1}a = 0)$ and assume $\rho_0 < 1 - \bar{p}$. In addition, assume the threshold is unique and has exact upper-tail mass on a neighborhood of the population root path; on a common neighborhood of its cutoff path, the score laws have jointly continuous densities bounded above and bounded away from zero. These are local conditions only and need not hold at other multiplier values or score levels.*

Then the directional root is attained by a region of mass p for each $p \in \mathcal{P}$. The pointwise CLT and studentization, and the curve functional CLT and multiplier band, hold on $\mathcal{P}$ with these local conditions replacing the global threshold and cutoff-density requirements in their proofs. The band-inversion crossing interval also applies; the crossing Wald result additionally requires an interior crossing. For paired gap-band conclusions, retain the fixed nonzero-margin conditions of Corollary S1 and the stated conditions of Proposition 2. Inference remains directional: no unrestricted-refit agreement, exact-refit certificate or finite-sample coverage guarantee is implied. If $p \geq 1 - \rho_0$, the conclusion is not asserted.

Proof of Proposition S2. Write $r_0 = Q^{-1}a$, $v = r'_0x$, $\mu = a'r_0$, $Y = vu$, and $Z = v^2/\mu$. The moment calculation at the start of the preceding proof is unchanged. For $\tau_1 = 1 - \bar{p}$, define

$$b_0 = \int_0^{\tau_1} F_Z^{-1}(t) dt.$$

Since $Z = 0$ exactly on the zero-index event, $\rho_0 < \tau_1$ implies $b_0 > 0$. The bottom-quantile allocation bound gives $\mathbb{E}[ZR] \geq b_0$ for every retained intensity of mass $1 - p$, uniformly over $p \in \mathcal{P}$. Also $P(v \neq 0) > 0$ because $Q > 0$ and $a \neq 0$; conditional full support on this event makes Y nonconstant with both tails. The population root therefore exists, is unique and positive by Proposition 4, and its path is continuous by strict descent with slope bounded away from zero.

At the root, $\mathbb{E}W_\delta = \mathcal{T}_p(W_\delta) = \delta$, so the lower-tail integral of W_δ through its $1 - p$ quantile is zero. A nonpositive quantile would make this integral strictly negative: the score has positive probability of being negative, while the zero-index atom is at zero. Hence the root cutoff is strictly positive. Conditional full support for $v \neq 0$ implies a continuous, strictly increasing score distribution at every positive score; consequently the root cutoff is unique and has exact upper-tail mass p . The region target is thus attained and equals the relaxed root.

The positive-score nuisance-CDF argument in the preceding proof already includes $s = r'x = 0$: its derivatives extend continuously by zero through this event and have the same integrable envelope. It therefore applies unchanged near the positive root cutoffs here, despite the oracle zero-index atom.

The assumed local exact-mass thresholds and density bounds supply a common positive-cutoff tube about the compact root path, and the retained-leverage floor persists under small nuisance perturbations. The proofs of Lemmas 5, 12 and 13

use precisely these cutoff-local properties, the L^2 envelope CH and weighted laws for score half-spaces on this tube. Their profiling, value-sandwich and root-transfer arguments therefore apply with the local replacements, without imposing the global threshold condition at multipliers outside the tube. After the variance floors below are established, the uniform first-stage and fitted-multiplier arguments yield the stated inference conclusions on $\mathcal{P}$.

Finally, conditional on x with $v \neq 0$, the feasible influence is piecewise affine in u with branch slopes $-\alpha(x)$ and $v - \alpha(x)$, where $\alpha(x) = m_a(p)'Q^{-1}x$. Their difference is $v \neq 0$, and both branches have positive conditional probability, so the conditional variance is positive on a set of positive probability. For the paired gap the slope difference is v/B ; it is nonzero on $\{v \neq 0\}$, not almost surely when $\rho_0 > 0$. The root and cutoff paths are continuous, and the local boundary atomlessness and common L^2 envelope give continuity of selected moments and influence variances, including the paired gap. Hence the pointwise variances have positive minima on the compact budget interval. This supplies the variance floors when the paired-gap result's other stated margin conditions hold; the curve inference conclusions do not require a margin. No claim is made at or above the budget cap. $\square$

Corollary S1 (Paired sign-margin inference). *Under Proposition S1 for a fixed $a_0 \neq 0$ and threshold b , suppose $c = |a'_0\beta - b| > 0$ and orient $a = \iota a_0$, where $\iota = \text{sign}(a'_0\beta - b)$. Then (S1) holds with $\varphi_c = (r'_a x)u$ and its fitted score, and*

$$\inf_{p \in \mathcal{P}} \text{Var} \left\{ \frac{\Psi_a(p)}{B_a^*(p)} - (r'_a x)u \right\} > 0.$$

Thus Proposition 2 and corollary 1 apply; Theorem 3 additionally requires an actual interior crossing. Validity is pointwise in the population, not uniform as c approaches zero.

Proof of Corollary S1. Let $c = |a'_0\beta - b| > 0$, $\iota = \text{sign}(a'_0\beta - b)$, and $a = \iota a_0$. OLS asymptotic linearity and sign consistency give (S1) with $\varphi_c = vu$ and $\hat{\varphi}_{c,i} = (\hat{r}'_a x_i)\hat{u}_i$. Indeed $\mathbb{P}(\hat{\iota} = \iota) \rightarrow 1$; on this event the fitted-minus-population margin score is bounded by $o_p(1)H$, proving empirical- L^2 consistency since $P_N H^2 = O_p(1)$. The paired gap is $h(p) = \Psi_a(p)/B_a^*(p) - vu$, as in (17). Put $B = B_a^*(p)$ and $\alpha(x) = m_a(p)'Q^{-1}x$. Conditional on x , its retained and active slopes are

$$-\alpha(x)/B - v, \quad (v - \alpha(x))/B - v.$$

On $\{v \neq 0\}$ both branches have positive probability and their slopes differ by $v/B \neq 0$; this event has positive probability, though it need not have probability one. Thus $\text{Var}(h(p)) > 0$; continuity and compactness give $\inf_{p \in \mathcal{P}} \text{Var}(h(p)) > 0$. The margin-score bound, Lemma 15, and the slope floor give $\sup_p P_N[\{\hat{h}(p) - h(p)\}^2] = o_p(1)$. The fixed compact score family has an integrable squared envelope; its uniform law and Cauchy–Schwarz for the fitted replacement give uniform gap-variance consistency. Hence Proposition 2 and corollary 1 apply. Crossing Wald inference additionally requires an actual interior crossing; band inversion does not. This is a fixed nonzero-margin result, not uniform validity as margins approach zero. $\square$

Remark S1. *The independent symmetric density proportional to $\{(1 + t^2)^{3/2}[\log(e + t^2)]^2\}^{-1}$ satisfies these density and variance bounds but has no moment above order two and no polynomial density envelope of exponent greater than three. The fourth-design and mixed-mean moments remain required.*

S2.2 Finite covariate support

Proposition S3 (Finite-covariate primitive verification). *Let x take fixed values $x_1, \dots, x_J \in \mathbb{R}^d$ with probabilities $\pi_j > 0$, where J and d are fixed and $Q = \sum_j \pi_j x_j x_j' > 0$. For each j , suppose that $y \mid x = x_j$ has a strictly positive continuous density h_j on $\mathbb{R}$ and a finite second moment. Fix $a \neq 0$ and suppose $v_j = x_j' Q^{-1} a \neq 0$ for every j . Under i.i.d. sampling of (x, y) , with β the linear projection coefficient of (1), Assumption 6 holds at every $p \in (0, 1)$, and Assumption 7 holds on every compact $\mathcal{P} \subset (0, 1)$. In particular, with $\mu = a' Q^{-1} a$,*

$$B_a^*(p) \geq (1 - p) \min_{1 \leq j \leq J} \frac{v_j^2}{\mu} > 0.$$

Proof of Proposition S3. Write $c_j = \mathbb{E}[y \mid x = x_j]$, $\mu = a' Q^{-1} a$, and $r_a = Q^{-1} a$. The projection coefficient is $\beta = Q^{-1} \sum_j \pi_j x_j c_j$, so $\mathbb{E}[xu] = 0$; the finite support and finite conditional second moments give C1 and the standing model moments. Put $z_j = v_j^2 / \mu > 0$. Every retained region has leverage at least $(1 - p) \min_j z_j$, so Assumption 3 holds. The conditional densities h_j imply that $Y(a) = v_j u$ has a continuous, nondegenerate full-support distribution and mean zero, hence its upper- p tail integral is positive. Thus Assumption 4 and Proposition 4 give a positive finite root.

Joint cutoff regularity. For $\theta = (b, r)$ near (β, r_a) , let $s_j = r'x_j$, $c(r) = a'r$, and

$$W_{\theta, \lambda} = (r'x)(y - x'b) + \lambda \frac{(r'x)^2}{a'r}, \quad \gamma_j = -s_j x_j' b + \lambda s_j^2 / c(r).$$

Choose a compact neighborhood where $c(r) \geq \mu/2$ and each s_j is bounded away from zero with the sign of v_j . In cell j , the score is $s_j y + \gamma_j$, with mixture density

$$f_{\theta, \lambda}(w) = \sum_{j=1}^J \frac{\pi_j}{|s_j|} h_j \left(\frac{w - \gamma_j}{s_j} \right). \quad (\text{S2})$$

It is positive and jointly continuous in (θ, λ, w) . On compact parameter and cutoff sets it is bounded above and away from zero. The cellwise CDFs are $F_j((w - \gamma_j)/s_j)$ for $s_j > 0$ and $1 - F_j((w - \gamma_j)/s_j)$ for $s_j < 0$; since $F_j' = h_j$ is continuous, the mixture CDF is C^1 and strictly increasing. This verifies Assumption 5 and the local density regularity.

Fix $\mathcal{P} = [p, \bar{p}] \subset (0, 1)$. The retained-slope floor is $b_0 = (1 - \bar{p}) \min_j z_j > 0$, so the root is bounded by $\mathbb{E}|Y(a)|/b_0$. Joint continuity of the tail criterion on bounded root sets and this uniform descent give continuity of the root path; strict quantile monotonicity gives continuity of its cutoff path. On a common band about that compact path, the density in (S2) has uniform positive lower and finite upper bounds. Quantile bracketing and the implicit function theorem therefore give a C^1 cutoff map locally, verifying C2, U1, and the slope clauses of C3 and U2.

Derived process regularity. The established moments, leverage and cutoff conditions give C1–C3, so Lemma 5 supplies C4–C5.

Influence variance. For fixed p , set $\alpha_j = m_a(p)'Q^{-1}x_j$. Conditional on cell j , (9) is continuous and piecewise affine in u , with slopes $-\alpha_j$ on the retained half-line and $v_j - \alpha_j$ on the active half-line. Both half-lines have positive probability under $h_j > 0$, and the slopes cannot both vanish since $v_j \neq 0$. Thus the conditional variance, and hence C6, is positive; the oracle score also has positive variance because its active-region slope is $v_j \neq 0$. The root and cutoff paths and selected moments vary continuously in p ; a common $C(1 + |y|)$ square-integrable envelope gives continuity of the influence variance. Its minimum on $\mathcal{P}$ is positive, establishing U2. The finite-support envelope $C(1 + |y|) \in L^2$, common cutoff tube and these floors give Lemmas 12 and 13; Lemma 14 then supplies U3. $\square$

S2.3 Gaussian regressors through the origin

Proposition S4 (Primitive Gaussian verification). *Let $x \sim \mathcal{N}(0, I_d)$ and $u \sim \mathcal{N}(0, 1)$ independently, let $y = x'\beta + u$, and fix $a = e_1$. Then Assumption 6 holds at every $p \in (0, 1)$ and Assumption 7 on every compact $\mathcal{P} \subset (0, 1)$. The design gradient is zero; with $T = x_1$ and $\nu(p) = \mathbb{E}[T^2 \mathbf{1}\{A_p^*\}]$,*

$$\Psi_a(p) = (Tu + \delta(p)T^2 - \xi_a^*(p))_+ + p\xi_a^*(p) - \delta(p)T^2 - \nu(p)Tu.$$

Proof of Proposition S4. Apply Proposition S1 with $m(x) = x'\beta$, $\sigma_0(x) = 1$, and $\varepsilon = u$. The Gaussian design has $Q = I_d$ and all fourth and mixed-mean moments; the independent standard normal conditional density is jointly measurable, positive and continuous, has variance one, obeys $\sup_t (1 + t^2)\phi(t) < \infty$ and $t^2\phi(t) \rightarrow 0$, and has conditional mean zero. Finally $r_a = e_1$ and $P(x_1 = 0) = 0$. Thus all standing assumptions, Assumption 6, and Assumption 7 follow from that proposition.

For $r \neq 0$, rotational invariance and independence give

$$((r'x)u, (r'x)^2) \stackrel{d}{=} (\|r\| Tu, \|r\|^2 T^2).$$

Positive homogeneity of the tail criterion therefore yields

$$\delta^{\text{sc}}(r; p) = \delta(p) / \|r\|, \quad \Phi_a(r; p) = \frac{a'r}{\|r\|} \delta(p).$$

The latter has zero gradient at $r_a = a = e_1$. Since A_p^* depends only on (T, u) , each coordinate x_j , $j > 1$, is independent of that event and of T and has mean zero. Consequently $\mathbb{E}[x_j T \mathbf{1}\{A_p^*\}] = 0$ for $j > 1$, whereas its first coordinate is $\mathbb{E}[T^2 \mathbf{1}\{A_p^*\}] = \nu(p)$: hence $m_a(p) = \nu(p)e_1$. In (9), the design layer vanishes and the residual layer is $-m'_a Q^{-1}xu = -\nu(p)Tu$. Substituting $Y = Tu$, $Z = T^2$ gives the stated score. $\square$

S3 Unrestricted-refitting bridge

The exact-refitting target provides the unrestricted comparison for the directional sensitivity developed in the main text. We define its population value and fixed-point characterization here; exact-statistic sampling theory is developed separately in the companion manuscript named in Section 1. The remaining results focus on the geometry, bounds, and conditions relating the exact and directional targets.

S3.1 Sample definitions and scalar equivalence

The exact maximum-influence statistic. The unrestricted sample benchmark maximizes the decrease in the estimated contrast $a'\hat{\beta}$ over all size- k deletion sets $S \in \mathcal{S}$. By the full-sample normal equations $\sum_{i=1}^N x_i \hat{u}_i = 0$, the retained residual sum satisfies $-\sum_{i \notin S} x_i \hat{u}_i = \sum_{i \in S} x_i \hat{u}_i$, which yields the standard leave- S -out identity.

Lemma S1 (Leave- S -out identity). *For every $S \in \mathcal{S}$, define the retained design matrix and deleted score sum*

$$\hat{Q}_{-S} := \frac{1}{N} \sum_{i \notin S} x_i x_i', \quad \hat{g}_S := \frac{1}{N} \sum_{i \in S} x_i \hat{u}_i.$$

Both use full- N normalization; the retained-sample design average is $\hat{Q}_{-S}/(1 - p_N)$. Whenever $\hat{Q}_N$ and $\hat{Q}_{-S}$ are invertible,

$$\Delta_a(S) := a'(\hat{\beta} - \hat{\beta}_{-S}) = a' \left(\sum_{i \notin S} x_i x_i' \right)^{-1} \sum_{i \in S} x_i \hat{u}_i = a' \hat{Q}_{-S}^{-1} \hat{g}_S,$$

where $\hat{\beta}_{-S}$ is the OLS estimator computed on the retained sample $[N] \setminus S$.

At the population level, removing a measurable region D of mass $P(D) = p$ obeys the analogous identity $\vartheta(P) - \vartheta(P_{-D}) = a' \Sigma(D)^{-1} g(D)$, where $\Sigma(D) := \mathbb{E}[xx' \mathbf{1}\{D^c\}]$ is the retained design moment and $g(D) := \mathbb{E}[xu \mathbf{1}\{D\}]$ is the deleted score moment (Section S3). Evaluating the exact sample maximum over all $\binom{N}{k}$ subsets requires every retained design matrix to be invertible.

Assumption S2 (Sample non-degeneracy). *At the reported deletion budget,*

$$\kappa_N := \min_{S \in \mathcal{S}} \lambda_{\min}(\hat{Q}_{-S}) > 0.$$

Definition S1. *The exact maximum-influence statistic in direction a , defined on the event of Assumption S2, is*

$$\hat{\delta}^{\text{ex}}(p) := \max_{S \in \mathcal{S}} a' \hat{Q}_{-S}^{-1} \hat{g}_S. \quad (\text{S3})$$

Assumption S4(E1) states the population counterpart of Assumption S2. Because the retained inverse $\hat{Q}_{-S}^{-1}$ varies with the deletion set S , subset deletion both rescales leverage and rotates the effective contrast direction $\hat{Q}_{-S}^{-1}a$, coupling all d coordinates of the design. The directional statistic in (4) holds the full-sample contrast direction fixed while adjusting for retained leverage along that direction, requiring a much

weaker scalar denominator condition; Section 3.3 compares the resulting population targets.

Lemma S2 (Scalar reduction when both statistics are defined). *Suppose $\hat{Q}_N$ is invertible and the sample satisfies Assumption S2, so that both statistics are defined. Let $\mathcal{D}_N$ denote the event of Proposition 3 that every retained denominator is strictly positive, read at the feasible pair $(\hat{s}_i(a), \hat{Z}_i(a))$. If $d = 1$ and $a = 1$, then on $\mathcal{D}_N$ the two statistics coincide exactly,*

$$\hat{\delta}(p) = \hat{\delta}^{\text{ex}}(p).$$

For $d = 1$ and $a = 1$, $\hat{s}_i(1) = x_i \hat{u}_i / \hat{Q}_N$ and $\hat{Z}_i(1) = x_i^2 / \hat{Q}_N$, so for each size- k set S ,

$$\frac{\sum_{i \in S} \hat{s}_i(1)}{\sum_{i \notin S} \hat{Z}_i(1)} = \frac{\sum_{i \in S} x_i \hat{u}_i}{\sum_{i \notin S} x_i^2}.$$

Maximizing over S gives $\hat{\delta}^{\text{ex}}(p)$, which establishes the scalar reduction directly. In one dimension, the retained inverse $\hat{Q}_S^{-1}$ changes only by a scalar factor, so the directional and unrestricted statistics coincide whenever defined. Under the conditions of Lemma 8(d), $\mathbb{P}(\mathcal{D}_N) \rightarrow 1$. For $d \geq 2$, unrestricted refitting may also rotate the contrast direction; Section 3.3 analyzes the resulting population target difference.

S3.2 Population comparison

The directional target optimizes exact retained least-squares movements on a fixed coefficient line; unrestricted refitting allows the displacement to leave that line. This restriction does not by itself order adverse contrast movements, because least squares minimizes prediction error rather than the target contrast. To compare its population target $\delta^{\text{ex}}(p)$ with the directional target $\delta(p)$ at a fixed deletion share $p \in (0, 1)$, the population normal equations imply that removing a measurable region D with $P(D) = p$ shifts the projection contrast by

$$\Phi(D) := \vartheta(P) - \vartheta(P_{-D}) = a' \Sigma(D)^{-1} g(D), \quad \Sigma(D) := \mathbb{E}[xx' \mathbf{1}\{D^c\}], \quad g(D) := \mathbb{E}[xu \mathbf{1}\{D\}],$$

so that $\delta^{\text{ex}}(p) = \sup_{P(D)=p} \Phi(D)$. For any D with $\Sigma(D) > 0$, $\Phi(D)$ is an attained movement; with $B_D = r_a' \Sigma(D) r_a / \mu_Z(a) > 0$, its difference from the line-restricted movement is

$$\Phi(D) - \frac{r_a' g(D)}{B_D} = g(D)' \{ \Sigma(D)^{-1} a - B_D^{-1} r_a \}. \quad (\text{S4})$$

At $D = A^*$ the line term is $\delta(p)$; this same-set rotation increment can have either sign. An admissible actual refit supplies a lower witness for the unrestricted supremum even without alignment.

Evaluating the retained design and deleted score moments at the directional active set $A^* = \{W_a^* \geq \xi_a^*\}$ defines the directional *retained-design metric* Q_{-A^*} and moment pair (Σ_{-A^*}, g_{A^*}) :

$$Q_{-A^*} := (1 - p)^{-1} \mathbb{E}[xx' \mathbf{1}\{A^{*,c}\}], \quad \Sigma_{-A^*} := (1 - p)Q_{-A^*}, \quad g_{A^*} := \mathbb{E}[xu \mathbf{1}\{A^*\}]. \quad (\text{S5})$$

Two alignments govern comparison with the directional refit. The first requires the retained design to rescale, rather than rotate, the contrast direction:

$$Q_{-A^*}^{-1}a \propto Q^{-1}a. \quad (\text{SN1})$$

The second requires the refitted shift $b = \Sigma_{-A^*}^{-1}g_{A^*}$ to align with the contrast direction:

$$\Sigma_{-A^*}^{-1}g_{A^*} \propto Q^{-1}a \quad (\text{equivalently, } g_{A^*} \propto a \text{ when Equation (SN1) holds}). \quad (\text{SN2})$$

Each condition imposes $d - 1$ directional restrictions rather than full matrix equality $Q_{-A^*} = Q$. The first restricts a retained second moment of x ; the second restricts an active-set cross-moment of x and u , and by Lemma S4 coincides under Equation (SN1) with the vanishing of the first-stage design correction ($g_a = 0$). No contrast rotation alone supplies the one-sided comparison:

Proposition S5 (No contrast rotation orders the population targets). *Under Assumption S4(E1)–(E2), Assumption 6, and Equation (SN1),*

$$\Sigma_{-A^*}^{-1}a = B_a^{*-1}r_a \quad \text{and} \quad \Phi(A^*) = \delta(p) \leq \delta^{\text{ex}}(p),$$

with strict inequality $\delta(p) < \delta^{\text{ex}}(p)$, under Assumption S4(E3), whenever Equation (SN2) fails and the design is boundary-regular.

Here boundary-regularity means positive Lebesgue density of (x, u) on an open ball in $\mathbb{R}^{d+1}$ intersecting $\{W_a^* = \xi_a^*\}$. It excludes intercept and discrete-covariate specifications, but is needed only for the strict-gap clause, not the sufficient lower bound.

Because $\delta(p) \leq \delta^{\text{ex}}(p)$ under Equation (SN1), directional fragility directly implies unrestricted fragility.

Corollary S2 (Transfer of evidence of exact fragility). *Let the conditions of Proposition 2 hold, and suppose Assumption S4(E1)–(E2) and Equation (SN1) hold at every $p \in \mathcal{P}$. For the lower gap band $L_{g,N}$ of Proposition 2 and margin $c = c(P) > 0$,*

$$\liminf_{N \rightarrow \infty} \mathbb{P}\{L_{g,N}(p) \geq 0 \implies \delta^{\text{ex}}(p) \geq c \text{ for every } p \in \mathcal{P}\} \geq 1 - \alpha.$$

Proof. On simultaneous gap coverage, $L_{g,N}(p) \leq \delta(p) - c \leq \delta^{\text{ex}}(p) - c$ at every budget, by Proposition S5. The implication therefore holds on an event whose probability tends to $1 - \alpha$. $\square$

The simultaneous band and its inverse concern the directional population target. Under no contrast rotation throughout $\mathcal{P}$, a nonnegative lower gap band transfers evidence of fragility to unrestricted refitting; a negative upper gap band certifies directional stability, not unrestricted stability. The latter requires target agreement or a valid upper confidence bound for $\delta^{\text{ex}}(p)$.

Remark S2 (The qualification is sharp: an audit where the ordering reverses). *Equation (SN1) restricts the retained design at the directional active set, and when it fails the two targets are ordered in neither direction. In the intercept-and-slope regression with covariate values $w = (0, 1, 2, 3, 4)$ and outcomes $y = (1, 2, -1, 4, -4)$, so that $N = 5$ and the full-sample fit is $\hat{\beta} = (2, -0.8)'$, the slope contrast $a = (0, 1)'$ gives $\hat{r}_a = (-1, 0.5)'$ and $\hat{\mu}_Z(a) = 0.5$. At $k = 2$ ($p = 2/5$) all ten retained designs are non-singular ($\min_S \det \hat{Q}_{-S} = 0.24$) and every retained directional denominator is positive ($h_Z = T_{N,p}(\hat{Z}) = 0.8 < 1$). The directional statistic, read either as the ratio in (4) or as the unique root of the threshold dual $f_N(\lambda) = T_{N,p}(\hat{s} + \lambda \hat{Z}) - \lambda$, equals 1.2800 and is attained uniquely at $S^* = \{1, 4\}$, which deletes the observations with $w = 0$ and $w = 3$. The exact maximum-influence statistic over the same ten sets equals 1.1286, likewise attained at S^* , and the first-order benchmark is $T_{N,p}(\hat{s}) = 0.6400$ on that same set, so all three rules agree on the deletion and the reversal lies in the criterion value rather than the selection. The mechanism is the increment above: at S^* the retained design rotates the contrast, $\hat{Q}_{-S^*}^{-1}a = (-2.5, 1.0714)'$ is not proportional to $\hat{r}_a$ (angle 3.37°), and the attained refit falls short of the line-restricted movement by $\hat{g}_{S^*}'\{\hat{Q}_{-S^*}^{-1}a - \hat{B}_{S^*}^{-1}\hat{r}_a\} = -0.1514$, exactly the observed gap (Table S1). Proposition S5 is illustrated rather than contradicted, since*

its sufficient condition fails at S^* ; reporting $\hat{\delta}(p)$ alone overstates the largest realised refit by 13.4% at this budget, whereas the same-set check step of Section 4.3 reveals the rotation directly. The audit is finite-sample and discrete: boundary-regularity and Assumption S4(E2) fail here, and the reversal is specific to this contrast orientation and budget, the ordinary ordering holding at $k = 1$ and $k = 3$.

TABLE S1: Directional and unrestricted refits in a five-observation audit where the directional statistic exceeds the unrestricted optimum.

Statistic	Value	Exact	Deleted set
Directional refitting, $\hat{\delta}(2/5)$	1.2800	32/25	{1, 4}
Unrestricted refitting, $\hat{\delta}^{\text{ex}}(2/5)$	1.1286	79/70	{1, 4}
Attained refit on the $\hat{\delta}$ set	1.1286	79/70	{1, 4}
First-order benchmark, $T_{N,p}(\hat{s})$	0.6400	16/25	{1, 4}
Gap, $\hat{\delta}(2/5) - \hat{\delta}^{\text{ex}}(2/5)$	0.1514	53/350	—
Same-set rotation increment	−0.1514	−53/350	—
Contrast rotation at S^*	3.37°	—	—

Notes. Intercept-and-slope ordinary least squares on $N = 5$ observations with $w = (0, 1, 2, 3, 4)$ and $y = (1, 2, -1, 4, -4)$, slope contrast $a = (0, 1)'$, budget $k = 2$, so $p = 2/5$. Directional refitting is the maximum over the ten size-two deletion sets of the line-restricted movement $\sum_{i \in S} \hat{s}_i(a) / \sum_{i \in S} \hat{Z}_i(a)$ (4); unrestricted refitting is the exact maximum-influence statistic $\max_S a' \hat{Q}_S^{-1} \hat{g}_S$ (S3), that is, the largest realised decrease in the fitted slope, whose refitted value is -1.9286 against a full-sample slope of -0.8 ; the first-order benchmark is $T_{N,p}(\hat{s})$. All ten retained designs are nonsingular and all ten retained directional denominators are positive. This discrete-design sample calculation illustrates finite-sample behavior; it is neither a population nor coverage result, and does not furnish a smooth-law counterexample to Proposition S5, whose sufficient condition fails at the selected set.

Both alignments make the directional region an exact fixed point, but target agreement additionally requires global optimality (Proposition S7). Index linearity supplies both alignments; mean-zero ellipticity with independent errors also yields global agreement (Corollary S3), a through-origin benchmark rather than a result for ordinary intercept specifications. The remaining subsections develop these distinctions and the population envelope bound, which has no sampling theory here.

S3.3 The exact statistic

The exact population target is defined directly, as the best that deletion of a p -mass can do. Identify a deletion set with a measurable region $D \subset \mathbb{R}^d \times \mathbb{R}$ and write

$$\Phi(D) := a' \Sigma(D)^{-1} g(D), \quad \delta^{\text{ex}}(p) := \sup_{P(D)=p} \Phi(D), \quad (\text{S6})$$

with $\Sigma(D) = \mathbb{E}[xx' \mathbf{1}\{D^c\}]$ the design second moment over the survivors and $g(D) = \mathbb{E}[xu \mathbf{1}\{D\}]$ the score mass the deleted region carries, using the population normal equations. Taking the supremum over *all* regions of mass p matches the unrestricted scope of the sample statistic, which maximizes over all $\binom{N}{k}$ subsets. Optimality yields the fixed-point characterization below, although finding a global maximizer remains difficult.

Proposition S6 (Attainment and the fixed-point characterisation). *Under the maintained moments $\mathbb{E} \|x\|^2 < \infty$ and $\mathbb{E}[\|x\| |u|] < \infty$ of Section 2, and Assumption S4 (E1) and (E2), the supremum in (S6) is attained, and every maximiser is almost surely a threshold set of its own deletion score, hence an exact fixed point in the sense of (S7).*

Every population maximizer satisfies the following self-consistency system, where $r_{a,-p} = \Sigma(D)^{-1}a$, $b = \Sigma(D)^{-1}g(D)$, and $s = (r'_{a,-p}x)(u + x'b)$:

$$\begin{aligned} D &= \{s \geq \xi\}, & P(D) &= p, \\ \Sigma(D) &= \mathbb{E}[xx' \mathbf{1}\{D^c\}] = (1-p)Q_{-D}, & g(D) &= \mathbb{E}[xu \mathbf{1}\{D\}]. \end{aligned} \quad (\text{S7})$$

Definition S2 (Exact fixed point). *An exact fixed point is a region and its moment pair $(D, \Sigma(D), g(D))$ satisfying (S7), with attained value $\Phi(D) = a' \Sigma(D)^{-1}g(D)$. This value equals $\delta^{\text{ex}}(p)$ only if D is globally maximizing; the fixed-point equations alone do not ensure this. The retained metric $Q_{-D} = \Sigma(D)/(1-p)$ differs in general from the directional retained-design metric Q_{-A^*} of (S5).*

S3.4 The exact target and index linearity

The exact statistic generally has a different population target because $\hat{Q}_s^{-1}$ is evaluated on the complement of the active set. The retained-design second moment at the directional optimum is the metric Q_{-A^*} of (S5), the population analogue of $\tilde{Q}_s :=$

$\hat{Q}_S/(1 - p_N)$. The exact statistic instead inverts the retained moment $\Sigma(D)$ at its maximizing region D , so replacing Q by Q_{-A^*} alone need not make the targets coincide.

Full design balance, $Q_{-A^*} = Q$, is strictly stronger than needed: it implies condition [SN1](#) but neither implies condition [SN2](#) nor holds in the independent-error elliptical case under the support conditions of Lemma [S6\(ii\)](#), so we work with the pair of directional conditions throughout.

The primitive single-index condition delivering both alignments is the following.

Assumption S3 (Index linearity). *The regressors are mean-linear in the contrast index, conditionally on the error,*

$$\mathbb{E}[x \mid r'_a x, u] = \frac{a}{\mu_Z(a)} r'_a x \quad a.s. \quad (\text{S8})$$

Assumption [S3](#) restricts the design in the *single* direction r_a and delivers both alignments in one line, $\Sigma_{-A^*} r_a = B_a^* a$ and $g_{A^*} = \mu_Z(a)^{-1} \delta B_a^* a$, hence condition [SN1](#) and condition [SN2](#); it is implied by, but strictly weaker than, mean-zero ellipticity with $x \perp\!\!\!\perp u$.

The exact population criterion depends on a region through the retained-design inverse. Its first variation gives the deletion score; combined with attainment, this yields the fixed-point characterization in Definition [S2](#).

Lemma S3 (Deletion score). *Adding mass ε at (x, u) to D changes $\Phi(D)$ by $\varepsilon s_D(x, u) + O(\varepsilon^2)$, with*

$$s_D(x, u) = (r'_{a,-D} x)(u + x' b_D), \quad r_{a,-D} := \Sigma(D)^{-1} a, \quad b_D := \Sigma(D)^{-1} g(D). \quad (\text{S9})$$

If a region D satisfies the relaxed first-order variational inequality $\mathbb{E}[(\pi - \mathbf{1}\{D\})s_D] \leq 0$ for every intensity $0 \leq \pi \leq 1$ with $\mathbb{E}\pi = p$, then $\{s_D > \xi\} \subseteq D \subseteq \{s_D \geq \xi\}$ up to null sets for a cutoff ξ . Under cutoff atomlessness this gives $D = \{s_D \geq \xi\}$ almost surely.

Proof. Adding mass ε at (x, u) sends $g \mapsto g + \varepsilon x u$ and $\Sigma \mapsto \Sigma - \varepsilon x x'$, so by Sherman–Morrison $\Sigma^{-1} \mapsto \Sigma^{-1} + \varepsilon \Sigma^{-1} x x' \Sigma^{-1} + O(\varepsilon^2)$. Substituting into [\(S6\)](#) and collecting the ε term gives $\dot{\Phi} = \varepsilon[(a' \Sigma^{-1} x)u + (a' \Sigma^{-1} x)(x' \Sigma^{-1} g)] = \varepsilon(r'_{a,-D} x)(u + x' b_D)$. The stated variational inequality makes $\mathbf{1}\{D\}$ maximize the linear functional $\pi \mapsto \mathbb{E}[\pi s_D]$ over the mass- p intensity class. Upper-tail allocation gives the threshold inclusions, allowing allocation on a tied cutoff; cutoff atomlessness removes that allocation. $\square$

Proof of Proposition S6. Attainment. Relax the region to a deletion intensity $\pi : \mathbb{R}^d \times \mathbb{R} \rightarrow [0, 1]$ with $\mathbb{E}\pi = p$, and write $\Sigma(\pi) = \mathbb{E}[xx'(1 - \pi)]$, $g(\pi) = \mathbb{E}[xu\pi]$, $\Phi(\pi) = a'\Sigma(\pi)^{-1}g(\pi)$; every region D induces a feasible indicator $\pi = \mathbf{1}\{D\}$. The feasible set is weak-* compact in $L^\infty(P) = (L^1(P))^*$ by Banach–Alaoglu, since it is a weak-* closed subset of the unit ball. The moment bounds make $\Sigma(\pi)$ and $g(\pi)$ weak-* continuous, while E1 gives $\Sigma(\pi) \succeq \kappa_p I$ for every feasible intensity. Thus inversion and Φ are continuous on this compact set, so a maximizer π^* exists.

Bang-bang. Let η be any feasible direction at π^* ($\mathbb{E}\eta = 0$ and $\pi^* + t\eta$ feasible for small $t > 0$). Differentiating as in Lemma S3,

$$\left. \frac{d}{dt} \Phi(\pi^* + t\eta) \right|_{t=0} = a'\Sigma^{-1}\mathbb{E}[xx'\eta]\Sigma^{-1}g + a'\Sigma^{-1}\mathbb{E}[xu\eta] = \mathbb{E}[\eta s(x, u)],$$

where $s = (r'_{a,-p}x)(u + x'b)$, evaluated at $(\Sigma, g) = (\Sigma(\pi^*), g(\pi^*))$. Optimality of π^* therefore makes it a maximiser of the *linear* functional $\pi \mapsto \mathbb{E}[\pi s]$ over $\{0 \leq \pi \leq 1, \mathbb{E}\pi = p\}$ —a fractional knapsack, whose solutions are $\pi = \mathbf{1}\{s > \xi\}$ together with arbitrary mass on $\{s = \xi\}$, where ξ is the $(1 - p)$ -quantile of s . By E2 the score s has no atom at ξ , so $\pi^* = \mathbf{1}\{s \geq \xi\}$ almost surely. Thus the relaxed maximiser is an honest region, the relaxed and integral suprema coincide, and π^* satisfies the three equations of Definition S2: it is an exact fixed point. The same argument applies to every relaxed maximizer and therefore to every maximizing region. $\square$

Assumption S4 (Exact-target regularity).

(E1) *Restricted eigenvalue.*

$$\kappa_p := \inf \{ \lambda_{\min}(\mathbb{E}[xx'\rho]) : 0 \leq \rho \leq 1, \mathbb{E}\rho = \tau \} > 0,$$

equivalently $\inf_{\|v\|=1} \int_0^\tau F_{(v'x)^2}^{-1}(s) ds > 0$, i.e. $\sup_{\|v\|=1} P(v'x = 0) < \tau$. This is the multivariate counterpart of Assumption 3, stated for retention intensities so it applies directly to the relaxation in Proposition S6. It bounds every retained region; under a nonatomic population, the intensity infimum also equals the infimum over regions of mass p , by splitting any tied leverage level.

(E2) *Score regularity at the cutoff.* For every multiplier (Σ, g) generated by a maximizer of the relaxed problem in the proof of Proposition S6, and at the directional pair (Σ_{-A^*}, g_{A^*}) , the cutoff $\xi(\Sigma, g)$ is a continuity point of the score distribution:

$P(s_{\Sigma,g} = \xi(\Sigma, g)) = 0$. (Absolute continuity of the joint law of (x, u) is sufficient, and it delivers the condition at every nonsingular (Σ, g) at once by Lemma S5.)

(E3) **Well-separated global target.** The value functional Φ of (S6) has a unique global maximiser region D_p among all measurable deletion regions of mass p , separated in value: for every $\epsilon > 0$,

$$\Phi(D_p) - \sup \{ \Phi(D') : P(D') = p, P(D' \Delta D_p) > \epsilon \} > 0.$$

By Proposition S6 the supremum is attained and every maximiser is an exact fixed point, so this is no stronger than it needs to be: restricting the separation to fixed points would not by itself connect the population target to the empirical maximum.

To verify the equivalences in E1, put $h(v) := \inf_{0 \leq \rho \leq 1, \mathbb{E}\rho = \tau} \mathbb{E}[(v'x)^2 \rho]$ for $\|v\| = 1$. The lower-tail allocation identity, allowing fractional allocation at a tied cutoff, gives $h(v) = \int_0^\tau F_{(v'x)^2}^{-1}(s) ds$, which is positive exactly when $P(v'x = 0) < \tau$. Under the maintained $\mathbb{E} \|x\|^2 < \infty$,

$$|h(v) - h(w)| \leq \mathbb{E} |(v'x)^2 - (w'x)^2| \leq 2\mathbb{E} \|x\|^2 \|v - w\|.$$

Thus h is continuous on the compact unit sphere and, if positive everywhere, has positive minimum κ_p . Also $v \mapsto P(v'x = 0)$ is upper semicontinuous by reverse Fatou, since $v_n \rightarrow v$ gives $\limsup_n \mathbf{1}\{v'_n x = 0\} \leq \mathbf{1}\{v'x = 0\}$; its maximum is attained, so pointwise strictness is equivalent to the displayed strict supremum bound. On an atomic population, the region infimum need not equal the intensity infimum; the intensity formulation avoids assuming that a tied atom can be split.

Condition E1 keeps retained designs nonsingular, as required by the first display of Proposition S8. Condition E2 supplies the no-atom premise used in the bang-bang step of Proposition S6, Proposition S7(iii), and the strictness clause of Proposition S5; the additional boundary regularity those results require is stated at their use. Condition E3 supplies the separated global target and, as a standing premise of Proposition S7, licenses the necessity clause read by the strictness clause of Proposition S5; index linearity alone gives fixed-point alignment, not global optimality.

S3.5 Alignment, stationarity and global agreement

Corresponding deviations from Equations (SN1) and (SN2) are measured by the Euclidean *contrast-rotation* and *numerator-rotation* angles

$$\varrho_p := \angle(\Sigma_{-A^*}^{-1}a, r_a), \quad \varsigma_p := \angle(\Sigma_{-A^*}^{-1}g_{A^*}, r_a), \quad (\text{S10})$$

so that Equation (SN1) holds if and only if $\varrho_p = 0$, and Equation (SN2) holds if and only if $\varsigma_p = 0$. Both conditions hold under Assumption S3, in particular under mean-zero ellipticity with $\mathbb{E}[x \mid r'_a x, u] = \mathbb{E}[x \mid r'_a x]$ (Lemma S6); ellipticity alone is not sufficient. These angles record departures from the two alignment conditions; the results below use the alignments themselves and make no quantitative claim bounding the target gap by either angle.

Recall that boundary-regularity requires positive ambient Lebesgue density on an open ball intersecting the directional cutoff surface. For any candidate direction $v \neq 0$, let $\Phi_a^{\text{sup}}(v)$ be the directional envelope in (S11) and $\Phi_a^* := \sup_{v \neq 0} \Phi_a^{\text{sup}}(v)$. Then $\Phi_a^{\text{sup}}(r_a) = \delta(p)$ and $\delta^{\text{ex}}(p) \leq \Phi_a^*$ by Proposition S8.

Proposition S7 (Stationarity, bounds, and target agreement). *Let Assumption 6 and Assumption S4(E1)–(E3) hold.*

- (i) (Stationarity.) *If Equations (SN1) and (SN2) hold, then A^* is an exact fixed point of (S7) with*

$$r_{a,-p} = B_a^{*-1} r_a, \quad b = \mu_Z(a)^{-1} \delta r_a, \quad \xi_p = B_a^{*-1} \xi_a^*, \quad \Phi(A^*) = \delta.$$

Conversely, if the design is boundary-regular, Equations (SN1) and (SN2) are necessary for A^* to be an exact fixed point.

- (ii) (Two-sided envelope bound.) *Under Equation (SN1) alone, $\Phi(A^*) = \delta$, yielding the sandwich*

$$\delta = \Phi(A^*) \leq \delta^{\text{ex}} \leq \Phi_a^*.$$

- (iii) (Global agreement.) *Suppose Equation (SN1), boundary-regularity, and the separated-maximum condition E3 hold, and let D_p denote the unique global maximizer of $\Phi(D)$ in (S6). Then*

$$\delta^{\text{ex}} = \delta \quad \Longleftrightarrow \quad A^* = D_p \text{ up to a } P\text{-null set,}$$

and either statement implies Equation (SN2).

- (iv) (Sufficient globality via the envelope.) Under Equation (SN1), if r_a maximizes $\Phi_a^{\text{sup}}(v)$ over $\mathbb{R}^d \setminus \{0\}$, then $\Phi_a^* = \delta$, the sandwich in (ii) closes with $\delta^{\text{ex}} = \delta$, and Equation (SN2) follows automatically.

Corollary S3 (A shared target). Under Assumption S3, Assumption 6, and Assumption S4(E1), A^* is an exact fixed point with value δ . If, in addition, x is mean-zero elliptical with $u \perp x$, then r_a globally maximizes Φ_a^{sup} , the sandwich of Proposition S8 closes, A^* is the global maximizer in (S6), and $\delta^{\text{ex}} = \delta$.

S3.5.1 Proofs for the population exact-refitting bridge

Proof of Proposition S7. Write $t = r'_a x$, so $Y(a) = tu$, $Z(a) = t^2/\mu_Z(a)$ and $W_a^* = t\{u + \mu_Z(a)^{-1}\delta t\}$, and recall $A^* = \{W_a^* \geq \xi_a^*\}$.

(i) *Stationarity: sufficiency.* Assume condition SN1 and condition SN2. By condition SN1, $\Sigma_{-A^*}^{-1}a = c r_a$ for some $c > 0$; multiplying $a = c \Sigma_{-A^*} r_a$ by r'_a gives $\mu_Z(a) = a' r_a = c r'_a \Sigma_{-A^*} r_a = c \mathbb{E}[t^2 \mathbf{1}\{A^{*,c}\}] = c \mu_Z(a) B_a^*$, so $c = B_a^{*-1}$ and $r'_{a,-p} x = B_a^{*-1} t$. By condition SN2, $b = \Sigma_{-A^*}^{-1} g_{A^*} = \gamma r_a$ for a scalar γ , so $x'b = \gamma t$ and $s = (r'_{a,-p} x)(u + x'b) = B_a^{*-1} t(u + \gamma t)$. The value at A^* is

$$\Phi(A^*) = a' \Sigma_{-A^*}^{-1} g_{A^*} = B_a^{*-1} r'_a g_{A^*} = B_a^{*-1} \mathbb{E}[Y(a) \mathbf{1}\{A^*\}] = \delta,$$

the last step by the root identity $\mathcal{F}_p(W_a^*) = \delta$, which reads $\mathbb{E}[Y(a) \mathbf{1}\{A^*\}] + \delta \nu_{a,Z}^* = \delta$, i.e. $\mathbb{E}[Y(a) \mathbf{1}\{A^*\}] = \delta B_a^*$ by (6). Also $\Phi(A^*) = a'b = \gamma a' r_a = \gamma \mu_Z(a)$, so $\gamma = \delta/\mu_Z(a)$ and $s = B_a^{*-1} W_a^*$. Setting $\xi_p = B_a^{*-1} \xi_a^*$ therefore gives $\{s \geq \xi_p\} = A^*$ and $P(s \geq \xi_p) = p$, while $\Sigma_{-A^*} = \mathbb{E}[xx' \mathbf{1}\{A^{*,c}\}]$ and $g_{A^*} = \mathbb{E}[xu \mathbf{1}\{A^*\}]$ hold by definition: all three equations of Definition S2 are satisfied at A^* , with value δ .

Necessity. Suppose A^* is an exact fixed point, i.e. $A^* = \{s \geq \xi_p\}$ up to a null set, with $s = (r'_{a,-p} x)(u + x'b)$, $r_{a,-p} = \Sigma_{-A^*}^{-1} a$, $b = \Sigma_{-A^*}^{-1} g_{A^*}$, and let

$$\begin{aligned} P_1(x, u) &:= s - \xi_p = (r'_{a,-p} x)u + (r'_{a,-p} x)(x'b) - \xi_p, \\ P_2(x, u) &:= W_a^* - \xi_a^* = (r'_a x)u + \mu_Z(a)^{-1} \delta (r'_a x)^2 - \xi_a^*. \end{aligned}$$

The key structural fact is that *both are affine in u* , so it is enough to compare them section by section; no appeal to the geometry of quadric surfaces is needed.

Let $\mathcal{U}$ be the ball supplied by boundary-regularity. Since $\{P_1 \geq 0\}$ and $\{P_2 \geq 0\}$ agree up to a P -null set and the density is positive on $\mathcal{U}$, the two open sets $\{P_1 > 0 > P_2\} \cap \mathcal{U}$ and $\{P_2 > 0 > P_1\} \cap \mathcal{U}$ are null, hence empty: on $\mathcal{U}$ the two polynomials never take strictly opposite signs. Fix $(x_0, u_0) \in \mathcal{U}$ with $P_2(x_0, u_0) = 0$. Then $r'_a x_0 \neq 0$, for otherwise $W_a^*(x_0, u_0) = 0 = \xi_a^*$, contradicting $\xi_a^* > 0$ (Lemma 11); and $r_{a,-p} \neq 0$ because $\Sigma_{-A^*} > 0$ and $a \neq 0$. The root $u_2(x) = \{\xi_a^* - \mu_Z(a)^{-1} \delta(r'_a x)^2\} / (r'_a x)$ is continuous near x_0 , so its graph lies in $\mathcal{U}$ over an open neighborhood of x_0 . Because $r_{a,-p} \neq 0$, that neighborhood is not contained in the hyperplane $\{x : r'_{a,-p} x = 0\}$; choose a boundary graph point outside that hyperplane. Shrinking to a ball around this point, we may assume $r'_a x \neq 0$ and $r'_{a,-p} x \neq 0$ throughout, with $u_2(x)$ in the ball's u -section for every x in an open set $O \subset \mathbb{R}^d$.

Fix $x \in O$. Both $P_1(x, \cdot)$ and $P_2(x, \cdot)$ are affine with nonzero slopes $r'_{a,-p} x$ and $r'_a x$, and $P_2(x, \cdot)$ changes sign at $u_2(x)$, an interior point of the section. Were $P_1(x, \cdot)$ to vanish elsewhere, or to have the opposite slope sign, the two would take strictly opposite signs on one side of $u_2(x)$. Hence the roots and the slope signs agree:

$$\frac{\xi_p - (r'_{a,-p} x)(x'b)}{r'_{a,-p} x} = \frac{\xi_a^* - \mu_Z(a)^{-1} \delta(r'_a x)^2}{r'_a x}, \quad \text{sign}(r'_{a,-p} x) = \text{sign}(r'_a x).$$

Cross-multiplying gives

$$(r'_a x)[\xi_p - (r'_{a,-p} x)(x'b)] = (r'_{a,-p} x)[\xi_a^* - \mu_Z(a)^{-1} \delta(r'_a x)^2]$$

for every $x \in O$; both sides are polynomials in x agreeing on an open set, hence identically on $\mathbb{R}^d$. Matching homogeneous parts, the degree-three part reads $(r'_a x)(r'_{a,-p} x)[x'b - \mu_Z(a)^{-1} \delta r'_a x] \equiv 0$, and since $\mathbb{R}[x]$ is an integral domain and $(r'_a x)(r'_{a,-p} x) \neq 0$,

$$b = \mu_Z(a)^{-1} \delta r_a,$$

which is condition SN2. The degree-one part reads $\xi_p(r'_a x) = \xi_a^*(r'_{a,-p} x)$ for all x , i.e. $\xi_p r_a = \xi_a^* r_{a,-p}$; as $\xi_a^* > 0$ and $r_{a,-p} \neq 0$ this forces $\xi_p \neq 0$ and

$$r_{a,-p} = \kappa r_a, \quad \kappa = \xi_p / \xi_a^*,$$

which is condition SN1, with $\kappa > 0$ by the slope-sign agreement.

(ii) *The one-sided bound.* The computation $c = B_a^{*-1}$ and $\Phi(A^*) = \delta$ in the sufficiency

part used condition [SN1](#) only. Since A^* is a feasible region of mass p , $\delta^{\text{ex}} \geq \Phi(A^*) = \delta$. The upper bound $\delta^{\text{ex}} \leq \Phi_a^*$ is the first display of Proposition [S8](#), which uses [E1](#) alone.

(iii) *Agreement.* Suppose $\delta^{\text{ex}} = \delta$. By (ii), $\Phi(A^*) = \delta = \delta^{\text{ex}} = \sup\{\Phi(D) : P(D) = p\}$, so A^* attains the supremum, and the uniqueness in [E3](#) identifies it with D_p up to a null set. Conversely, if $A^* = D_p$ up to a null set then $\delta^{\text{ex}} = \Phi(D_p) = \Phi(A^*) = \delta$ by (ii). For the implication, a maximiser is an exact fixed point by Proposition [S6](#), so under boundary-regularity the necessity clause of (i) applies to A^* and yields $b = \mu_Z(a)^{-1} \delta r_a$, which is condition [SN2](#). The two alignments do not settle the converse implication: conditions [SN1](#) and [SN2](#) place A^* at a stationary point of the exact problem and leave the global comparison open, as Remark [S3](#) states.

(iv) *A sufficient global condition.* If r_a maximises Φ_a^{sup} then $\Phi_a^* = \Phi_a^{\text{sup}}(r_a) = \Phi_a(r_a) = \delta$ by the identification above, and the sandwich of (ii) gives $\delta \leq \delta^{\text{ex}} \leq \delta$. Condition [SN2](#) is not needed as a hypothesis: $a'r_a = \mu_Z(a) > 0$, so on the neighbourhood $\{a'v > 0\}$ of r_a the truncation in ([S11](#)) can only raise the ratio, whence $\Phi_a(v) \leq \Phi_a^{\text{sup}}(v) \leq \Phi_a^{\text{sup}}(r_a) = \Phi_a(r_a)$ for every v there; r_a is therefore a local maximiser of Φ_a , which is differentiable at r_a by Lemma [5\(C5\)](#), so $g_a = \nabla_r \Phi_a(r_a) = 0$, and given condition [SN1](#) that is condition [SN2](#) by Lemma [S4](#). $\square$

Remark S3 (Stationarity is not optimality). *Conditions [SN1](#) and [SN2](#) place the directional set at a fixed point of the exact rule, but a fixed point need not be the global maximiser. This is why the agreement result states its globality condition separately.*

Proof of Proposition [S5](#). The computation $c = B_a^{*-1}$ and $\Phi(A^*) = \delta$ in the sufficiency part above used condition [SN1](#) only, not condition [SN2](#). (More generally, the same-set identity of Section [3.3](#) at $D = A^*$ reads $\Phi(A^*) = \delta + \Delta_{\text{rot}}$ with $\Delta_{\text{rot}} := g_{A^*}'\{\Sigma_{-A^*}^{-1}a - B_a^{*-1}r_a\}$, so the displayed ordering holds whenever this single scalar increment is nonnegative, condition [SN1](#) being the zero-increment special case.) Since A^* is a feasible region of mass p , $\delta^{\text{ex}} = \sup_{P(D)=p} \Phi(D) \geq \Phi(A^*) = \delta$. For strictness, suppose condition [SN2](#) fails and the design is boundary-regular. If A^* were equal to $\{s \geq \xi\}$ up to null sets, A^* would be an exact fixed point and the necessity clause of Proposition [S7](#), available under [E1](#)–[E3](#), would force $b = \mu_Z(a)^{-1} \delta r_a$, i.e. condition [SN2](#); so $A^* \neq \{s \geq \xi\}$ up to null sets. (Pointwise inequality of s and W_a^* is *not* enough on its own: two scores generate the same super-level sets whenever one is a strictly increasing transform of the other, and it is the polynomial argument in that necessity clause, and hence boundary-regularity, that rules this out.) Since s has no atom at its $(1-p)$ -quantile ξ by [E2](#), it follows that $P(A^* \cap \{s < \xi\}) > 0$ and $P(A^{*,c} \cap \{s > \xi\}) > 0$. Transfer

mass $\epsilon > 0$ from the first event to the second: by the first variation of Lemma S3, Φ increases by $\epsilon \mathbb{E}[s \mid A^{*,c} \cap \{s > \xi\}] - \epsilon \mathbb{E}[s \mid A^* \cap \{s < \xi\}] + O(\epsilon^2) > 0$ for small ϵ . This is a feasible intensity perturbation; by Proposition S6 under E1–E2, the relaxed and region suprema agree, so the improved relaxed value implies $\delta^{\text{ex}} > \Phi(A^*) = \delta$. Finally, in the strict-gap case just established, where condition SN2 fails and the design is boundary-regular, one has $\delta^{\text{ex}} > \delta$, so $\sqrt{N}(\hat{\delta} - \delta^{\text{ex}}) \rightarrow -\infty$ in probability and any interval of width $O_p(N^{-1/2})$ centered at $\hat{\delta}$ has coverage of δ^{ex} tending to zero. $\square$

The envelope bound compares deletion regions, so the scalarisation it needs is a supremum over regions rather than the root of (27). For every $v \neq 0$ define the truncated directional supremum

$$\Phi_a^{\text{sup}}(v) := (a'v) \max \left\{ 0, \sup \left\{ \mathbb{E}[(v'x)u \mathbf{1}\{D\}] / \mathbb{E}[(v'x)^2 \mathbf{1}\{D^c\}] : P(D) = p \right\} \right\}. \quad (\text{S11})$$

By E1 every competitor has denominator at least $\kappa_p \|v\|^2 > 0$, while Cauchy–Schwarz and the standing model of Section 2, which gives $\mathbb{E} \|x\| |u| \leq (\mathbb{E} \|x\|^2 \mathbb{E} u^2)^{1/2} < \infty$, bound its numerator by $\|v\| \mathbb{E} \|x\| |u|$; hence $\Phi_a^{\text{sup}}(v)$ is well defined for every $v \neq 0$ under E1 alone, with no quantile or root condition charged on the direction-perturbed score, and, being of degree zero, obeys $\Phi_a^{\text{sup}}(v) \leq \|a\| \mathbb{E} \|x\| |u| / \kappa_p$ at every $v \neq 0$. It is homogeneous of degree zero under positive rescaling of v , non-negative on $\{a'v > 0\}$ and non-positive on $\{a'v < 0\}$. At any direction r for which the pair $((r'x)u, (r'x)^2)$ satisfies Assumption 2–Assumption 5, the Rockafellar–Uryasev dual (20) rewrites the criterion as $f(\lambda; r) = \sup_{P(D)=p} \{\mathbb{E}[(r'x)u \mathbf{1}\{D\}] - \lambda \mathbb{E}[(r'x)^2 \mathbf{1}\{D^c\}]\}$, and at a non-negative root the two-sided bound $\kappa_p \|r\|^2 \leq \mathbb{E}[(r'x)^2 \mathbf{1}\{D^c\}] \leq r'Qr$ forces that root to equal the inner supremum of (S11); a root is automatically non-negative, so the truncation does not bind there. At r_a these are the hypotheses of Proposition 4, so $\Phi_a^{\text{sup}}(r_a) = (a'r_a)\delta^{\text{sc}} = \delta = \Phi_a(r_a)$.

Proposition S8 (Directional envelope). *Take an arbitrary direction $v \neq 0$ and let Φ_a^{sup} be as in (S11), so that $\Phi_a^{\text{sup}}(r_a) = \Phi_a(r_a) = \delta$. Under Assumption S4(E1), for every region D with $P(D) = p$,*

$$\Phi(D) \leq \Phi_a^{\text{sup}}(\Sigma(D)^{-1}a), \quad \text{and therefore} \quad \delta^{\text{ex}} \leq \sup_{v \neq 0} \Phi_a^{\text{sup}}(v).$$

Under condition [SN1](#) the two bounds sandwich the exact target,

$$\delta = \Phi_a^{\text{sup}}(r_a) \leq \delta^{\text{ex}} \leq \sup_{v \neq 0} \Phi_a^{\text{sup}}(v) = \Phi_a^*.$$

Proof of Proposition [S8](#). Fix D with $P(D) = p$ and set $v := \Sigma(D)^{-1}a$, which is well defined and nonzero because $\Sigma(D) \succeq \kappa_p I$ by [E1](#). Then

$$\begin{aligned} \Phi(D) &= a' \Sigma(D)^{-1} g(D) = v' g(D) = \mathbb{E}[(v'x) u \mathbf{1}\{D\}], \\ a'v &= v' \Sigma(D) v = \mathbb{E}[(v'x)^2 \mathbf{1}\{D^c\}] > 0. \end{aligned}$$

Hence

$$\Phi(D) = (a'v) \cdot \frac{\mathbb{E}[(v'x)u \mathbf{1}\{D\}]}{\mathbb{E}[(v'x)^2 \mathbf{1}\{D^c\}]} \leq \Phi_a^{\text{sup}}(v),$$

because D is one competitor in the inner supremum of [\(S11\)](#) and $a'v > 0$, so the truncation $\max\{0, \cdot\}$ dominates the ratio. No root identity is invoked here, and therefore no quantile or tail condition is charged on the direction-perturbed score. Taking the supremum over D gives $\delta^{\text{ex}} \leq \sup_{v \neq 0} \Phi_a^{\text{sup}}(v) = \Phi_a^*$; the lower bound is Proposition [S5](#), together with $\Phi_a^{\text{sup}}(r_a) = \Phi_a(r_a) = \delta$ from [\(S11\)](#). $\square$

The same alignment governs the design layer of Theorem [1](#) and the fixed-point property of Proposition [S7\(i\)](#).

Lemma S4 (The design gradient is the numerator misalignment). Write $\delta^{\text{sc}} = \delta/\mu_Z(a)$ for the scalarised target as stated in [\(28\)](#). Without any assumption on the design,

$$m_a = a - \Sigma_{-A^*} r_a, \quad \text{hence} \quad g_a = \delta^{\text{sc}} a + B_a^{*-1} g_{A^*} - 2\delta^{\text{sc}} B_a^{*-1} \Sigma_{-A^*} r_a. \quad (\text{S12})$$

Under condition [SN1](#) these collapse to

$$m_a = \nu_{a,Z}^* a \quad \text{and} \quad g_a = B_a^{*-1} g_{A^*} - \delta^{\text{sc}} a,$$

so that, given condition [SN1](#), $g_a = 0$ if and only if condition [SN2](#) holds.

Proof of Lemma [S4](#). By definition $m_a = \mathbb{E}[x(r'_a x) \mathbf{1}\{A^*\}] = \mathbb{E}[xx' \mathbf{1}\{A^*\}] r_a = (Q - \Sigma_{-A^*}) r_a = a - \Sigma_{-A^*} r_a$, using $Q r_a = a$; substituting $m_a - a = -\Sigma_{-A^*} r_a$ into [\(8\)](#) gives the second identity in [\(S12\)](#). Under condition [SN1](#) the computation in the sufficiency part of the proof of Proposition [S7](#) gives $\Sigma_{-A^*} r_a = B_a^* a$, whence $m_a = (1 - B_a^*) a = \nu_{a,Z}^* a$

by (6) and

$$g_a = \delta^{\text{sc}} a + B_a^{*-1} g_{A^*} - 2\delta^{\text{sc}} a = B_a^{*-1} g_{A^*} - \delta^{\text{sc}} a.$$

Finally, the root identity $\mathcal{T}_p(W_a^*) = \delta$ gives $r'_a g_{A^*} = \mathbb{E}[Y(a)\mathbf{1}\{A^*\}] = \delta B_a^*$, so if $g_{A^*} = c a$ then $c \mu_Z(a) = \delta B_a^*$, i.e. $c = \delta^{\text{sc}} B_a^*$ and $g_a = 0$; conversely $g_a = 0$ forces $g_{A^*} = \delta^{\text{sc}} B_a^* a \propto a$. Under condition SN1, $g_{A^*} \propto a$ is exactly condition SN2. $\square$

Proof of Corollary S3. Under (S8) the active set A^* is a function of (t, u) alone, so $\Sigma_{-A^*} r_a = \mathbb{E}[x t \mathbf{1}\{A^{*,c}\}] = \mathbb{E}[\mathbb{E}[x | t, u] t \mathbf{1}\{A^{*,c}\}] = \mu_Z(a)^{-1} a \mathbb{E}[t^2 \mathbf{1}\{A^{*,c}\}] = B_a^* a$ and $g_{A^*} = \mathbb{E}[x u \mathbf{1}\{A^*\}] = \mu_Z(a)^{-1} a \mathbb{E}[t u \mathbf{1}\{A^*\}] = \mu_Z(a)^{-1} \delta B_a^* a$, which are condition SN1 and condition SN2 because $\Sigma_{-A^*} > 0$ by E1. The stationarity computation of Proposition S7(i) is direct here: $r_{a,-p} = B_a^{*-1} r_a$, $b = \mu_Z(a)^{-1} \delta r_a$ and $s = B_a^{*-1} W_a^*$, so the cutoff $B_a^{*-1} \xi_a^*$ recovers A^* and $\Phi(A^*) = \delta$, without invoking E2 or E3. For the elliptical statement write $x = Q^{1/2} \epsilon$ with ϵ spherical and $u \perp \epsilon$, and put $w = Q^{1/2} v$, so that $v'x = w'\epsilon$ and $v'Qv = \|w\|^2$. For every unit vector w the joint law of $(w'\epsilon, u)$ is the same, so the inner supremum in (S11) depends on w only through $\|w\|$; it is homogeneous of degree -1 under positive rescaling of v , and it is unchanged when v is replaced by $-v$ (the pair $((v'x)u, (v'x)^2)$ is then equal in law to the original, since $v'x$ and $-v'x$ are equidistributed and $u \perp x$), so it equals $\phi_p / \|w\|$ for a constant $\phi_p > 0$ — positive by the regular-root identification at r_a , where $\delta > 0$, and rotational invariance — and the truncation of (S11) is inactive. Therefore

$$\Phi_a^{\text{sup}}(v) = (a'v) \frac{\phi_p}{\|w\|} = \phi_p \frac{(Q^{-1/2}a)'w}{\|w\|} \leq \phi_p \|Q^{-1/2}a\| = \phi_p \sqrt{\mu_Z(a)},$$

with equality if and only if $w \propto Q^{-1/2}a$, i.e. $v \propto Q^{-1}a = r_a$. Hence $\Phi_a^* = \sup_{v \neq 0} \Phi_a^{\text{sup}}(v) = \Phi_a^{\text{sup}}(r_a) = \delta$, and the sandwich of Proposition S8 forces $\delta^{\text{ex}} = \delta$; since $\Phi(A^*) = \delta = \delta^{\text{ex}}$, A^* attains the supremum in (S6). $\square$

Lemma S5 (Atomlessness of the deletion score). *For a multiplier $\varpi = (r_{a,-p}, b)$, write $s_{\varpi}(x, u) = (r'_{a,-p}x)(u + x'b)$. If the law of (x, u) is absolutely continuous on $\mathbb{R}^{d+1}$, then $P(s_{\varpi} = w) = 0$ for every $w \in \mathbb{R}$ and every multiplier with $r_{a,-p} \neq 0$; in particular E2 holds at every score parameter the bridge evaluates.*

Proof. Fix ϖ with $r_{a,-p} \neq 0$ and $w \in \mathbb{R}$. The polynomial $(x, u) \mapsto (r'_{a,-p}x)(u + x'b) - w$ has $r'_{a,-p}x$ as the coefficient of u , and $r'_{a,-p}x$ is not the zero polynomial, so the polynomial is not identically zero. Its zero set is a proper algebraic hypersurface and

has Lebesgue measure zero, and absolute continuity of the law of (x, u) turns this into $P(s_{\overline{w}} = w) = 0$. $\square$

Lemma S6 (No rotation under elliptical designs). *Let Assumption 6 and Assumption S4(E1) hold, so that $B_a^* > 0$ and $\Sigma_{-A^*} > 0$. Suppose x is elliptically symmetric about the origin with $\mathbb{E}[xx'] = Q \succ 0$ and $\mathbb{E}[x \mid r_a'x, u] = \mathbb{E}[x \mid r_a'x]$ (in particular when $u \perp x$). Write $t := r_a'x$ and $\Sigma_{\perp} := Q - \mu_Z(a)^{-1}aa'$, so that $\Sigma_{\perp}r_a = 0$ and $Q = \Sigma_{\perp} + \mu_Z(a)^{-1}aa'$. Then:*

(i) $\mathbb{E}[x \mid t] = \mu_Z(a)^{-1}at$ and $\mathbb{E}[xx' \mid t] = w(t)\Sigma_{\perp} + \mu_Z(a)^{-2}t^2aa'$ for a scalar $w(\cdot) \geq 0$, with $w \equiv 1$ in the Gaussian case.

(ii) If additionally $u \perp x$, then $Q_{-A^*} = Q + c_{\perp}\Sigma_{\perp} + c_a aa'$ with $(Q_{-A^*} - Q)r_a = c_a\mu_Z(a)a$ and

$$c_a = \frac{1}{\mu_Z(a)} \left(\frac{B_a^*}{1-p} - 1 \right) \leq 0, \quad (\text{S13})$$

where $c_{\perp} = 0$ in the Gaussian case (so $\text{rank}(Q_{-A^*} - Q) \leq 1$). If moreover $|t|$ is non-degenerate and u has a strictly positive density on $\mathbb{R}$, then $B_a^* < 1 - p$ and $c_a < 0$, so $Q_{-A^*} \neq Q$ while Equation (SN1) ($Q_{-A^*}^{-1}a \propto Q^{-1}a$) holds identically.

Proof. (i) Elliptical symmetry gives $\mathbb{E}[x \mid t] = \text{Cov}(x, t)\text{Var}(t)^{-1}t = \mu_Z(a)^{-1}at$ and decomposes the conditional second moment into its r_a component and a scalar multiple $w(t)\Sigma_{\perp}$ on the orthogonal complement; Gaussianity gives $w \equiv 1$.

(ii) If $u \perp x$, put $\pi(\tau) = P(A^* \mid |t| = \tau)$. Integrating (i) gives the aa' coefficient of Q_{-A^*} as $B_a^*/\{\mu_Z(a)(1-p)\}$ and hence c_a in (S13); the $\Sigma_{\perp}$ coefficient is $(1-p)^{-1}\mathbb{E}[w(t)(1 - \mathbf{1}\{A^*\})]$. It equals one when $w \equiv 1$, so $c_{\perp} = 0$ in the Gaussian case.

For $t > 0$, selection is $u \geq \varphi(|t|)$ and for $t < 0$ it is $u \leq -\varphi(|t|)$, where $\varphi(\tau) = \{\xi_a^* - \mu_Z(a)^{-1}\delta\tau^2\}/\tau$ is strictly decreasing because $\xi_a^* > 0$ (Lemma 11) and $\delta \geq 0$. Independence implies that $\pi(\tau)$ is nondecreasing, strictly increasing when u has a positive density. Therefore $\mathbb{E}[t^2\mathbf{1}\{A^*\}]/\mu_Z(a) = \mathbb{E}[Z(a)\mathbf{1}\{A^*\}] \geq p$, strictly if $|t|$ is nondegenerate and the density is positive. Hence $B_a^* \leq 1 - p$ and $c_a \leq 0$, with strict inequality under the stated strictness conditions. $\square$

S4 Monte Carlo evidence

This section of the Online Supplement presents supporting exhibits for the four simulation designs in Section 5.

S4.1 Wald coverage across the budget

Figure S1 traces empirical Wald coverage across deletion shares $p \in [0.01, 0.20]$, complementing the pointwise calibration in Table 1.

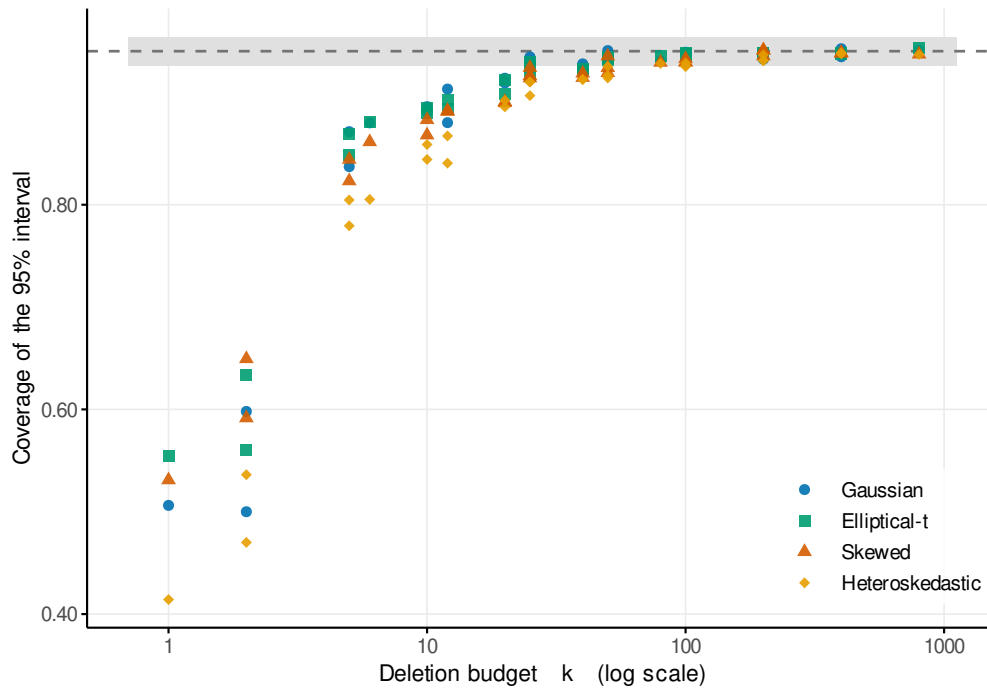


FIGURE S1: Empirical Wald coverage across deletion budgets $k = \lfloor pN \rfloor$ for $p \in [0.01, 0.20]$ (Corollary 3). The $N = 125, p = 0.01$ ($k = 1$) cell illustrates small- k undercoverage where the fixed-fraction asymptotic approximation deteriorates. Shading marks nominal coverage plus or minus twice the maximum Monte Carlo standard error across cells, not a simultaneous uncertainty band.

S4.2 The fragility curve and the breakdown fraction

Table S2 reports simultaneous grid coverage and breakdown-fraction Wald intervals on the eleven-point grid $0.02, 0.03, \dots, 0.12$ at $N = 1,500$, over 20,000 replications. The

functional simulations evaluate the general feasible influence score at every design, so the fitted design correction is retained even where its population counterpart vanishes (Remark 4); the residual-corrected score is kept as an ablation. The paired coverage difference, general feasible minus residual-corrected, is $+0.0003$ in the Gaussian design and -0.0001 in the elliptical- t design (both inside one paired standard error), and $+0.0129$ with a paired standard error of 0.0008 in the skewed design. Band coverage is below nominal in every cell, and the shortfall is concentrated at the smallest deletion budgets: marginal coverage generally rises across the grid from about 0.93 – 0.95 at $p = 0.02$ to about 0.96 – 0.97 at $p = 0.12$. Repeating the design at $N = 3,000$ raises band coverage to 0.934 , 0.937 , 0.922 and 0.924 on the Gaussian, elliptical- t , skewed and heteroskedastic laws, so the deficit shrinks with sample size but remains visible at $N = 3,000$. Both conclusion thresholds are fixed oracle-reference levels, chosen at $p^\dagger = 0.06$ and 0.10 ; this experiment does not calibrate the paired estimated-margin band of Proposition 2 or continuum coverage on $\mathcal{P}$. Crossing estimates linearly interpolate between grid points rather than search all integer budgets; the reported coverage concerns this numerical procedure, with no asserted interpolation-error rate. The printed lengths are mean crossing-Wald interval widths, not mean curve-band widths.

The earlier single-sample illustration in Figure S2 is distinct from both this 20,000-replication table and the paired estimated-margin campaign below. Its saved coordinates come from the prior fixed-margin study; it illustrates the curve and band rather than supplying a new coverage result. Across the four fixed-margin designs, the envelope slope $\partial_p \delta(p) = \xi_a^*(p)/B_a^*(p)$ closely matches finite differences of the oracle curve.

S4.3 Estimated-margin paired-gap campaign

The separate campaign underlying Table 2 uses $x = (1, T, C)'$ with cell probabilities $(.30, .35, .10, .25)$, conditional means $(0, .2, .9, 2.3)$ and scales $(.5, 1.2, .8, 1.5)$ in the order $(T, C) = (0, 0), (0, 1), (1, 0), (1, 1)$. Errors are independent standard normal or t_3 draws; the contrast is $a_0 = (0, 1, 0)'$ and the conclusion threshold is $b = 1$. The sign and margin $|a_0' \hat{\beta} - b|$ are estimated per sample, with all 1,600 attempts per cell retained and scored against the fixed population orientation. Conditional-tail integrals supply population curves and crossings, rather than an empirical oracle draw; numerical root brackets are floating checks, not outward arithmetic certificates.

TABLE S2: Inference on the fragility curve: simultaneous-band coverage and grid-interpolated crossing Wald intervals.

Design	Curve band (simult. cov.)		Crossing	Breakdown fraction $\hat{p}^\dagger$		
	Multiplier	Pointwise		$p^\dagger$	Coverage	Length
Gaussian	0.929	0.882	interior	0.060	0.944	0.0155
			near-bound	0.100	0.948	0.0231
Elliptical-t	0.930	0.883	interior	0.060	0.949	0.0205
			near-bound	0.100	0.950	0.0292
Skewed	0.912	0.872	interior	0.060	0.935	0.0379
			near-bound	0.100	0.978	0.0450
Heteroskedastic	0.914	0.870	interior	0.060	0.938	0.0209
			near-bound	0.100	0.944	0.0297

Notes. 20,000 replications at $N = 1,500$, nominal $1 - \alpha = 0.95$, all designs evaluated with the general feasible influence score; the residual-corrected score is retained as an ablation and reported in the online supplement. *Curve band*: simultaneous coverage of the whole fragility curve over the grid $\mathcal{D}$ by the multiplier-bootstrap band vs. the naive pointwise band. *Breakdown fraction*: the simulation linearly interpolates the fitted grid curve to estimate the crossing, then uses the plug-in Wald standard error from Theorem 3, at an interior and a near-boundary reference level c . This numerical procedure does not search all integer budgets; its interpolation error is not bounded here. Coverage is of $p^\dagger$; Length is the mean Wald-interval width. The binomial Monte Carlo standard error of a reported coverage $\hat{r}$ is $\{\hat{r}(1 - \hat{r})/20,000\}^{1/2}$, about 0.002 at nominal coverage. Reference targets are computed on an oracle draw of 300,000 observations; their approximation error is not included in these Monte Carlo standard errors.

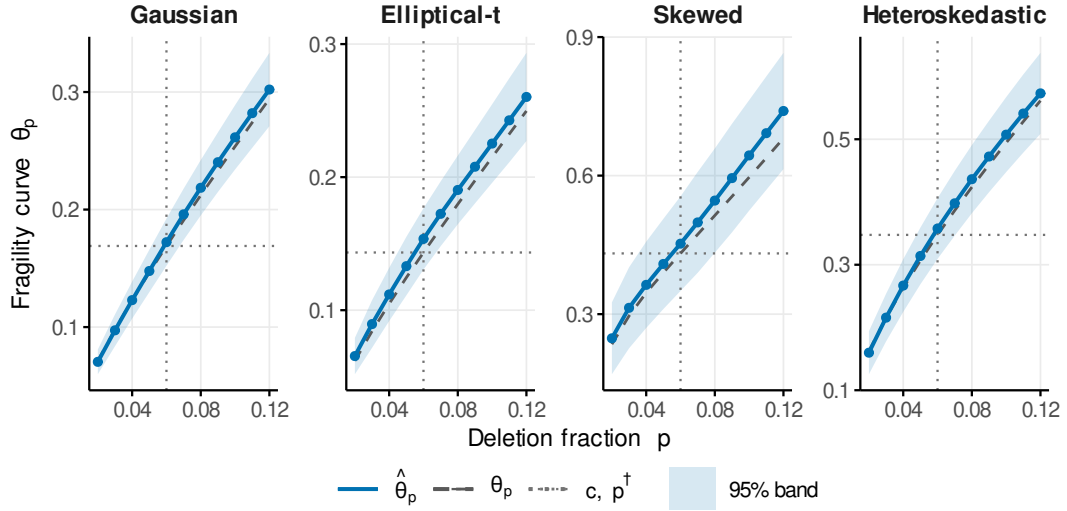


FIGURE S2: Earlier fixed-margin directional-curve illustration, from the study whose coverage experiment used 300 replications, not the later Table S2 or paired-gap campaign. A single-sample estimate and multiplier band on the eleven-point grid are compared with the numerical population reference; dotted lines mark the fixed conclusion threshold and crossing. These saved coordinates are not continuum or estimated-margin calibration.

Every integer count intersecting $[0.02, 0.20]$ uses the full paired influence and the same 999 multipliers. Because the fitted band is constant within a count cell while the population curve is monotone, checking both ends, including one-sided limits, evaluates coverage over the continuous interval without interpolating the empirical crossing. Grid-state coverage exceeds this continuum coverage by 0.0013–0.0088; the latter is the main-text column. Table S3 retains censoring and available-only summaries separately from the unconditional coverage denominator in the main table. At $N \geq 800$, conditional crossing-Wald coverage ranges from 0.937 to 0.956; small-sample heavy-tail undercoverage is not tuned away.

S5 Additional Sample-Audit Detail

The following calculations extend the deterministic institutions–income audit in Section 6.1, using the original AJR Table 2, column 6 specification with latitude and continent controls. The authors’ replication command requests heteroskedasticity-robust standard errors; this exercise reproduces its OLS point estimates and supplies

TABLE S3: Paired-gap campaign: censoring, availability and conditional widths.

Law	N	Censored L/R	Wald available	Wald cond.	Inv. width	Wald width	Sign errors
Normal	200	408/1141	1499	0.960	0.141	0.163	0
Normal	800	0/162	1599	0.937	0.089	0.083	0
Normal	3200	0/0	1600	0.956	0.045	0.042	0
t_3	200	1175/638	1462	0.966	0.143	0.158	26
t_3	800	455/3	1589	0.948	0.086	0.081	0
t_3	3200	7/0	1600	0.951	0.045	0.041	0

Notes. Counts are out of 1,600 attempts; left and right inversion censoring can overlap. Wald conditional coverage and width use available intervals only; inversion width includes every fit. Sign errors are scored against the fixed population orientation, not a sample-redefined target. There were no fit refusals or unavailable orientations; no attempt was replaced.

no standard errors or significance claims. The directional statistic freezes the full-sample adjusted contrast direction, while each attained OLS refit re-estimates all six coefficients on its retained sample.

TABLE S4: Adjusted institutions–income audit: movements, refits and retained leverage.

k	$100k/N$	$\hat{\delta}$	AMIP	$\hat{\beta}_{1,-S_{\text{dir}}}$	$\hat{\beta}_{1,-S_{\text{AMIP}}}$	$\hat{B}_a^*$
3	4.69	0.0833	0.0683	0.3104	0.3104	0.819
6	9.38	0.1416	0.0954	0.2493	0.2547	0.624
12	18.75	0.2301	0.1246	0.1621	0.1866	0.524
24	37.50	0.3697	0.1424	0.0278	0.0951	0.325
26	40.62	0.3913	0.1432	0.0024	0.0799	0.310
27	42.19	0.4033	0.1434	−0.0023	0.0774	0.271
32	50.00	0.4829	0.1435	−0.0442	0.0447	0.177

Notes. AJR Table 2, column 6; $N = 64$, observed slope 0.4013. AMIP is $T_{N,p}(\hat{s})$; both refit columns re-estimate all six coefficients on their retained samples. Every selected retained design has rank six over $k = 0, \dots, 32$. The directional denominator is $\hat{B}_a^* = N^{-1} \sum_{i \in S^*} \hat{Z}_i$; the all-set diagnostic $h_Z = T_{N,p}(\hat{Z})$ is at most 0.9017 on this domain.

Table S5 separates the two ranking rules from their own objective values. At $k = 27$, the sets share 19 of 27 countries, yet the directional criterion evaluated on the AMIP set is 0.3167, versus its own optimum 0.4033; AMIP evaluated on the directional set is 0.1092, versus 0.1434 on its own set. Thus substantial overlap does not make the rules interchangeable near the directional crossing.

At $k = 3$, both rules select Singapore, Gabon and Bangladesh, giving the same attained refit slope of 0.3104, compared with the baseline 0.4013. At the directional crossing, deleting its 27 selected countries gives an actual OLS slope of −0.0023; the

TABLE S5: Cross-evaluation of directional and AMIP criteria with latitude and continent controls.

k	Shared rows	Directional on		AMIP on	
		S_{dir}	S_{AMIP}	S_{AMIP}	S_{dir}
3	3/3	0.0833	0.0833	0.0683	0.0683
12	10/12	0.2301	0.1974	0.1246	0.1206
27	19/27	0.4033	0.3167	0.1434	0.1092

Notes. The directional value on S is $N^{-1} \sum_{i \in S} \hat{s}_i / (1 - N^{-1} \sum_{i \in S} \hat{z}_i)$; AMIP is $N^{-1} \sum_{i \in S} \hat{s}_i$. Each criterion is evaluated on both fixed selected sets at the same budget; shared rows count common deletions.

AMIP-selected deletion leaves it at 0.0774. The attained movement is approximately 0.4036, versus the directional 0.4033: a same-set rotation increment of about +0.0003 using the unrounded results as defined in (S4). The directional-selected refit is still positive at $k = 26$ (0.0024), and the AMIP-selected refits remain positive throughout the reported path. These are attained sample movements from two computable selection rules, not maximizations over unrestricted refits.

At the crossing, the directional-selected deletion leaves 27.07% of directional leverage. The rule ranks the fitted scores $\hat{s}_i + \hat{\delta}(p)\hat{z}_i$, combining residual contributions with directional leverage, and recomputes the selected set at each budget. The cross-evaluation isolates criterion-specific ranking losses on these two sets, but does not decompose their actual-refit coefficient gap into selection and refitting components.

Finite-share adjustment changes the sample curve substantially, but the crossing requires deleting 42.2% of the countries. The 27-country set is a feasible actual reversal, not a certified smallest reversing set, and the small negative refit is not a significance claim. This deterministic OLS audit is not a robustness analysis of the IV estimand or population inference, which requires the conditions in Section 4.

S6 Conditional Illustration

S6.1 Cognitive-training completers

The source is the official supplementary participant CSV of Hardy et al. (2015). The observed frame comprises 5,045 post-test completers (2,667 Lumosity and 2,378 crosswords) of 9,919 randomized participants. No Grand-Index change value is missing within this release, but completion is post-randomization; observed-case

completeness does not identify an all-randomized ITT. Trial registration NCT02367898 was first posted on February 20, 2015, after primary completion in April 2014. The unadjusted assignment contrast was self-declared for this calculation, not justified by prospective registration of that analysis. The published ANCOVA is reconstructible using the released pre-test index on the crossover-removed frame (coefficient 2.7677, $t(4712) = 8.7324$); absence of that covariate is not the reason for the unadjusted illustration.

The working model maintains iid sampling from a completer superpopulation and the LP cutoff conditions. Released scores are rounded; all observed change values being distinct does not verify positive population cutoff density. Exact reconstruction of percentile-rank norm tables is unavailable. The directional curve is not the unrestricted mean-contrast refit curve: actual selected-set refits are separate checks, and the directional and unrestricted-optimal arm allocations differ at the six checked budgets. The calculation supplies neither randomization-based coverage nor a missingness correction.

The displayed bands use integer counts 101, ..., 1009 and 99,999 common Gaussian multipliers with seed 20261019; the 95% and 99% critical values are 2.293468954 and 2.899102585, respectively. The zero-budget anchor is excluded from band inference. At the observed frozen contrast 3.09716777, the directional movement is 3.09701265 at $k = 326$ and 3.10480894 at $k = 327$, so the first integer sign crossing is 327/5045, not 326/5045. Monotonicity-based grid inversion of the 95% curve band reports [311/5045, 342/5045] without endpoint censoring. This holds the reported threshold fixed: it is not the paired estimated-margin procedure of Proposition 2, and no calibrated interval for the fixed population sign-margin radius is claimed. The crossover-removed ablation has $n = 4,715$, contrast 3.14867537 and bracket [289/4715, 322/4715]. The earlier 999-draw primary had 95% critical value 2.3604202182 and bracket [310/5045, 343/5045]; its separate multiplier-precision ablation used 9,999 draws and critical value 2.2657193861. Neither ablation validates empirical sampling coverage. The supplementary-file redistribution license is not established by the article's CC BY license; the computed figure does not redistribute the participant CSV.

S7 Population Geometry at Matched Signal-to-Noise

Standardized signal-to-noise does not determine the directional population curve or its zero-margin crossing, even within the finite-covariate class verified in the Online Supplement. Table S6 compares four deliberately calibrated specifications, not samples or a coverage experiment.

There are four strata s and a binary treatment T , with $x = (1, T, \mathbf{1}\{s = 2\}, \mathbf{1}\{s = 3\}, \mathbf{1}\{s = 4\})'$ and contrast $a = (0, 1, 0, 0, 0)'$. Write $w_s = P(s)$ and $q_s = P(T = 1 \mid s)$, so the eight cell probabilities are $w_s(1 - q_s)$ and $w_s q_s$. Outcomes obey $y = b_s + \kappa g_s T + \sigma_s e$, where e is independent of (s, T) , has mean zero and variance one, and $b = (.40, -.25, .70, -.35)$. For D0, D2 and D3, set $w = (.30, .30, .20, .20)$, $q = (.5, .3, .7, .4)$ and $g = (.6, -.1, .9, -.4)$. D0 uses standard Gaussian e and $\sigma = (1, 1, 1, 1)$; D2 changes only e to a logistic law with scale $\sqrt{3}/\pi$; D3 uses Gaussian e and $\sigma = (.55, .8, 1.25, 1.85)$. D4 uses Gaussian e , $\sigma = (1, 1, 1, 1)$, $w = (.43, .43, .07, .07)$, $q = (.52, .48, .97, .03)$ and $g = (.35, -.05, 4.5, -4.5)$. In each specification, β is the population linear projection, $u = y - x'\beta$, $\vartheta = a'\beta$ and $V = \mathbb{E}[(a'Q^{-1}xu)^2]$. Choose the positive effect scale κ to solve $\vartheta/\sqrt{V/1000} = 3$: approximately .810110894316 for D0 and D2, .918218422370 for D3 and 2.410023043738 for D4. The reference size 1000 expresses a population calibration, not a realized t -statistic.

All cell probabilities are positive and the support spans $\mathbb{R}^5$. Writing $H = \sum_s w_s q_s (1 - q_s) > 0$, the cell index is $x'Q^{-1}a = (T - q_s)/H \neq 0$. Gaussian and logistic conditional densities are continuous and strictly positive on $\mathbb{R}$, with finite second moments; hence all four laws satisfy that finite-covariate verification, including its positive retained-denominator guarantee. The entries solve the population modified-score cutoff and ratio equations of Section A, up to the reported numerical optimization tolerance.

D0 and D2 share the covariate law, conditional means and conditional variances, matching ϑ and V separately within 10^{-12} , yet D2 has a larger curve at 0.02 and 0.10, a smaller curve at 0.20 and an earlier zero-margin crossing. D3 and D4 match only standardized signal-to-noise; D4 has the largest movement at .10 but the latest crossing among these specifications, whose margins differ. D4 jointly changes stratum weights, treatment shares and effect patterns and recalibrates the scale; it illustrates joint geometry rather than isolating a leverage mechanism. These calibrated directional-population examples establish non-sufficiency of standardized signal-to-noise, not unrestricted-refit ordering, sampling coverage or empirical prevalence.

TABLE S6: Directional population geometry at matched standardized signal-to-noise.

Design	ϑ	$\sqrt{V}$	$\delta(0.02)$	$\delta(0.10)$	$\delta(0.20)$	$p_{\text{dir}}^\dagger$
D0 Gaussian baseline	0.2036	2.1461	0.1211	0.4523	0.7987	0.0375
D2 logistic twin	0.2036	2.1461	0.1293	0.4540	0.7839	0.0356
D3 heteroskedastic	0.2308	2.4325	0.1534	0.4985	0.8282	0.0342
D4 leverage-concentrated	0.3548	3.7396	0.2023	0.5523	0.9293	0.0523

Notes: $\vartheta = a\beta$ and $V = \mathbb{E}[(aQ^{-1}xu)^2]$. The zero-margin crossing is $p_{\text{dir}}^\dagger = \inf\{p : \delta(p) \geq \vartheta\}$. All rows are deliberately calibrated to $\vartheta/\sqrt{V/1000} = 3$ (absolute tolerance 10^{-12}); 1000 is a reference size, not a realized sample or t -statistic. D2 also matches D0 in ϑ and V separately within 10^{-12} ; D3 and D4 match only the ratio. Values are numerical population optimizations, not arithmetic enclosures or sampling uncertainty.

References

Hardy, J. L. et al. (2015). “Enhancing Cognitive Abilities with Comprehensive Training: A Large, Online, Randomized, Active-Controlled Trial”. In: *PLOS ONE* 10.9, e0134467. doi: [10.1371/journal.pone.0134467](https://doi.org/10.1371/journal.pone.0134467).