Credible Extrapolation in Regression Discontinuity Designs with Covariates

Samuel G.Z. Asher

Ph.D. Student, Stanford Graduate School of Business

Overview

Setting
- **Sharp RD:** ubiquitous causal inference tool; identifies local estimands
- **Common goal:** estimate non-local effects using covariates (often lagged Y)
- **Tension:** we ran the RD *because* we don't trust selection on those covariates!

Contribution
- **Bias-aware CIs** for linear estimators of non-local estimands in sharp RDs
- **No ignorability:** assumptions only on smoothness of $\mathbb{E}[Y(0) \mid X, Z]$
- **Adaptive:** bounds tighten with covariate quality and overlap

estimands · estimators / weights · smoothness / bias bounds · features $\phi(X)$ · running variable Z

Motivating Example: Testing Mechanisms by Extrapolating

Setting: Urvoy (2025, AER) *"Elections and Public Funding of Nonprofits"*

- **Question:** Do French mayors steer grants to ideologically aligned nonprofits?
- **Data:** 5,592 municipal elections & data on gov. transfers to local nonprofits

Potential Mechanisms

Clientelism	Strategic fundraising
Reward loyal groups *regardless* of the race	Buy future mobilization *where it is pivotal*
⇒ effect **persists** far from the cutoff	⇒ effect **fades** away from the cutoff

- **Current practice** (Angrist & Rokkanen 2015): residualize Y on X, scan windows until residuals look flat in Z, then assume unconfoundedness inside the window

Methodological Problem:

- **Unconfoundedness:** Implausible even within a cutoff window
- **Ad Hoc Selection:** Author selects a 15pp margin without principled justification
- **Uncertainty Reporting:** Extrapolation CIs *75% narrower* than RD CIs

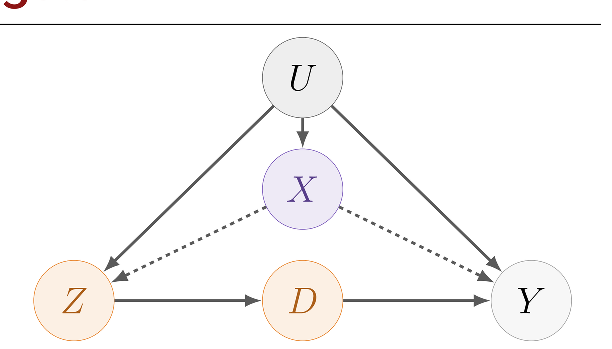
Q1: Can we extrapolate in a principled way?
Q2: Can we report both sampling and identification uncertainty?

Model & Targets

$Y(0) = g(U) + e^Y, \quad Y(1) = Y(0) + \tau(U)$

$Z = r(U) + e, \qquad D = \mathbf{1}\{Z \geq 0\}$

$X = G(U) + e^X, \quad U$ latent



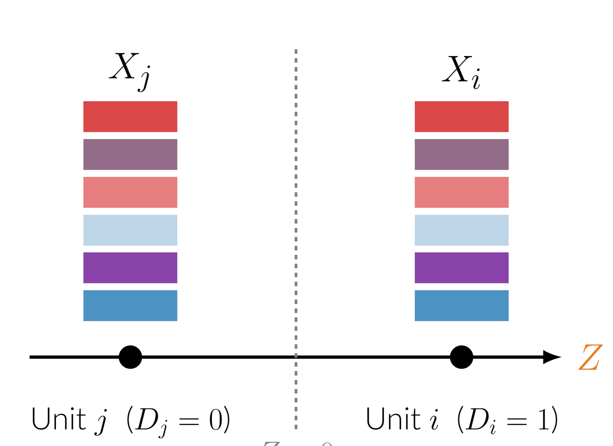
Targets $\tau_{\mathrm{RD}} = \mathbb{E}[\tau(U) \mid Z{=}0]$ $\quad \tau_{\mathrm{ATT}} = \mathbb{E}[\tau(U) \mid D{=}1]$ $\quad \tau_h = \mathbb{E}[\tau(U) \mid |Z| < h]$

Assumption: f_e continuous with one bounded derivative $\implies \tau_{\mathrm{RD}}$ identified

Two Types of Treated–Control Comparisons

- **Estimator class:** $\hat{\tau}_w = w^\top Y$ (e.g. covariate balancing; local linear regression)

Balancing view
- Nearly identical X
- Gap plausibly noise
- *High weight*



RD view
- Far apart in Z
- Likely far in latent U
- *Low weight*

Approach: Smoothness Bound on Extrapolation Bias

(1) Balance covariates to handle the portion of confounding they can see
(2) Smoothness bound for the remaining confounding
(3) Residual smoothness measures covariate quality & overlap

- **Bias of** $\hat{\tau}_w - \tau_{\mathrm{ATT}}$: a function of the CEF of $Y(0)$ in X and Z

Assumption: $m_\phi := \mathbb{E}[Y(0) \mid \phi(X) = x_\phi] = \beta^\top \phi(X)$

$$m(x_\phi, z) = \mathbb{E}\big[Y(0) \mid \phi(X) = x_\phi,\ Z = z\big] = m_\phi(x_\phi) + m^\perp(x_\phi, z)$$

- **After covariate balancing:** $\mathrm{Bias}(w) = \mathbb{E}[m^\perp \mid D{=}1] - \mathbb{E}_w[m^\perp \mid D{=}0]$
- **X-balancing** handles the component of the CEF explained by X
- **Residual problem:** $m^\perp$ is remaining *extrapolation* in the running variable

Lemma. If f_e has a bounded first derivative, then

$$\big|\mathrm{Bias}(w)\big| \leq \Gamma(w;\phi,L) := \sup_{q \in \mathcal{Q}(\phi,L)} \big|\mathbb{E}[q \mid D{=}1] - \mathbb{E}_w[q \mid D{=}0]\big|$$

$$\mathcal{Q}(\phi, L) = \Big\{q :\ \mathbb{E}[q \mid \phi(X)] = 0,\ \ \mathrm{Lip}_z(q) \leq L\Big\}$$

- **Uncertainty Reporting:** $L,\ \hat{\tau}_w \pm c_\alpha\big(\widehat{se},\ \Gamma(w;\phi,L)\big)$, with normal critical values
- **Uniform Finite-Sample Coverage:** given ex-ante L (Armstrong & Kolesár 2018)
- **Pointwise Asymptotic Coverage:** for consistently estimated $\hat{L}$ (Ghosh et al. 2025)

Estimating the Smoothness Class

- **Feature grid:** partition $\phi(X)$ into bins b_k, grid grows finer as $n \to \infty$
- **Within-bin:** residual bias is a 1-D transport problem in Z

$$\Gamma(w;\phi,L) = \sum_k \pi_k\, L(b_k)\, W_1\big(\mu^{Z}_{1|b_k},\ \mu^{Z,w}_{0|b_k}\big)$$

Worst-Case Bias = **extrapolation cost** $L(b_k)$ × **extrapolation amount** W_1

- **Estimating** $L(b_k)$: control-side bin quadratics, shrunk to pooled slope

Assumption: worst-case smoothness $L(b_k)$ attained on the control side

- **Sensitivity analysis:** report honest intervals for $t \cdot \hat{L}$ across a range of t

Smoothness Measures Identification Uncertainty

Lemma. $L(x_\phi) \leq \sqrt{V(x_\phi) \cdot I(x_\phi)}$

$$V = \sup_z \mathrm{Var}\big(Y(0) \mid \phi(X), Z{=}z\big), \qquad I = \sup_z \mathbb{E}\big[s_z^2 \mid \phi(X), Z{=}z\big], \qquad s_z = \partial_z \log p(U \mid X, Z)$$

- **No residual variation:** $Y(0)$ flat given $\phi(X) \Rightarrow V = 0$
- **Covariates explain** U: Z adds nothing about U given $X \Rightarrow I = 0$
- **Latent overlap:** running variable determined idiosyncratically ⇒ small I

Need high-quality covariates and good latent overlap for tight bounds

Computing the Estimator

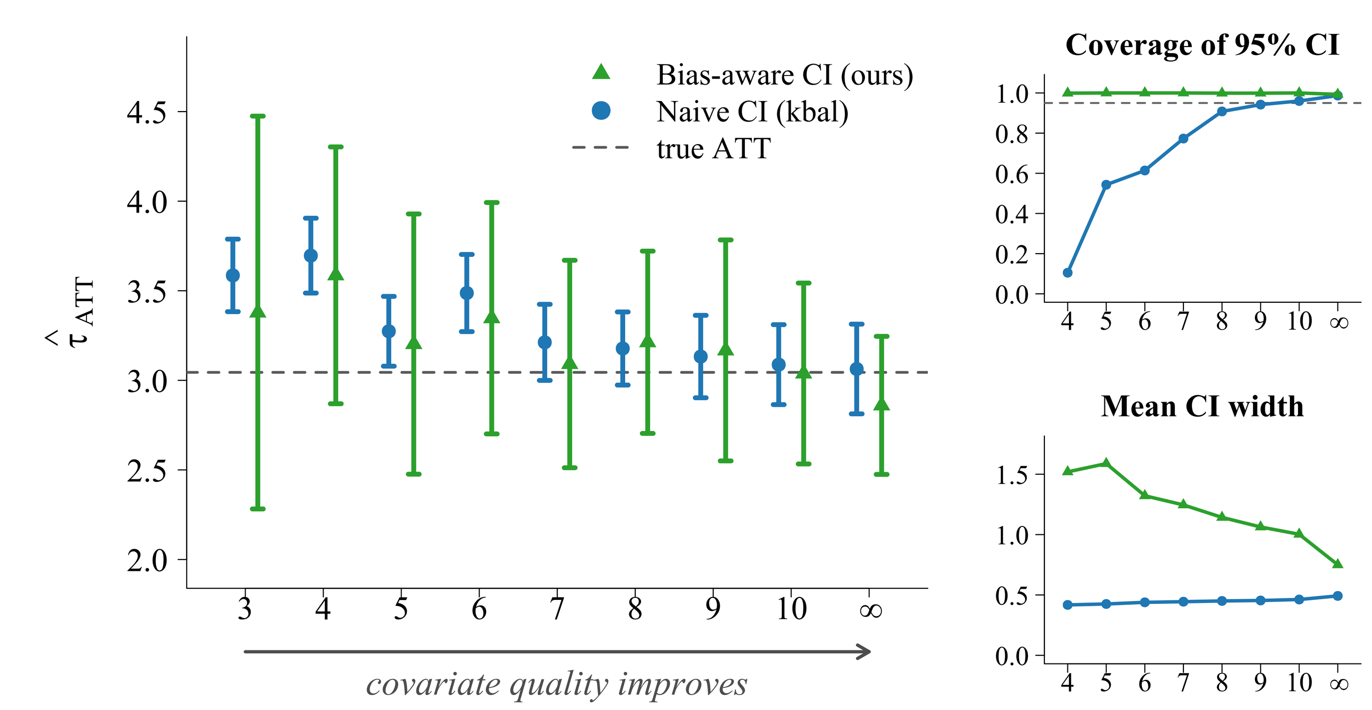
$$\min_{w \in \mathcal{W}} \underbrace{\Gamma(w;\phi,L)^2}_{\text{worst-case bias}} + \underbrace{\lambda\Big\|\bar{\phi}_T - \textstyle\sum_{D_i=0} w_i \phi(X_i)\Big\|^2}_{\text{covariate balance}} + \underbrace{\sigma^2\|w\|_2^2}_{\text{variance}}$$

- **Balancing space:** Researcher-chosen; e.g. linear, Gaussian RKHS
- **Targeting** τ_h: fix treated weights uniform on $\{|Z| < h\}$; solve for control weights
- **Computation:** a single convex program; select λ to minimize half-width

Simulations: Valid Inference at Every Covariate Quality

Setting
- **DGP:** latent $U \to (Z, X, Y)$; nonlinear selection and $\tau(U)$; N=1,500; 1,000 sims
- **Covariates:** $X = G(U$ coarsened to K bins)
- **Quality Improves:** K=1 covariates useless → K=∞ means τ_{ATT} identified

Competitor: Kernel balancing (kbal) (Hazlett 2020)
- Balancing weights on kernel features $\phi(X)$ + standard CI

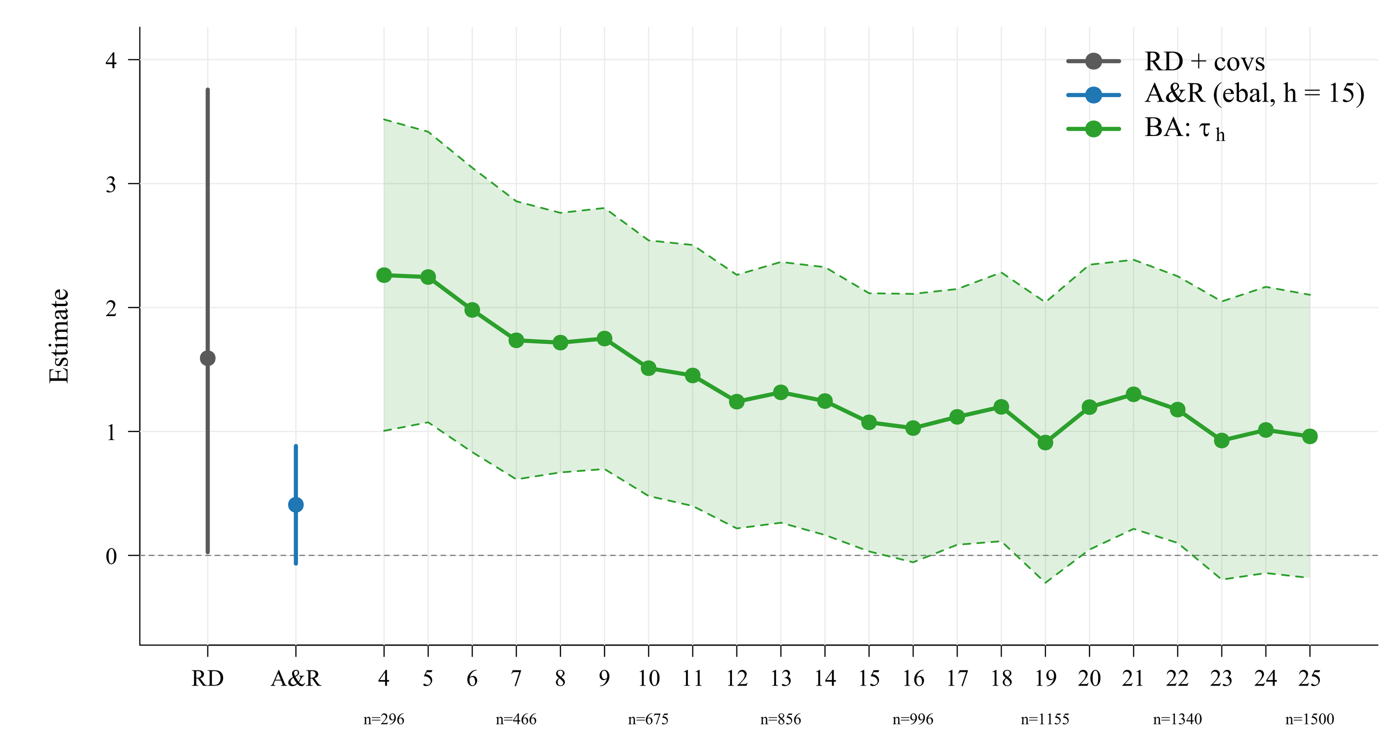


Left: one simulation draw per K; dashed line = true τ_{ATT}. Right: averages over 1,000 sims

Results
- Naive CI: **11%** coverage at K=4; nominal only once X is near-lossless
- Bias-aware CI: covers at every K; half-width **shrinks** as X improves

Application: Electoral Motives for Public Funding

Setting (Urvoy 2025, AER)
- **Mechanism test:** does the effect fade or persist away from the cutoff?
- **Estimates:** mean features; A&R at eyeballed 15pp vs. honest band for every τ_h



Findings
- Effect fades away from the cutoff ⇒ evidence for strategic fundraising
- Traditional extrapolation **dramatically understates** identification uncertainty

Which controls get weighted? (h = 15; weights against running variable)