Noise-Induced Randomization in Regression Discontinuity Designs

Dean Eckles, Nikolaos Ignatiadis, Stefan Wager, Han Wu

Presented by: Sai Zhang

November 18, 2022

Outline

Discussion

Literature: Continuity-Based RD

Most popular: local linear regression (Hahn et al., 2001; G. W. Imbens and Lemieux, 2008)

$$\hat{\tau}_c = \arg\min_{\tau} \left\{ \sum_{i=1}^{n} \underbrace{K}_{\text{weighting}} \left(\frac{|Z_i - c|}{\underbrace{h_n}_{\text{bandwidth}}} \right) \left(Y_i - a - \tau W_i - \beta_- (Z_i - c)_- - \beta_+ (Z_i - c)_+ \right)^2 \right\}$$

;

Literature: Continuity-Based RD

Most popular: local linear regression (Hahn et al., 2001; G. W. Imbens and Lemieux, 2008)

$$\hat{\tau}_c = \arg\min_{\tau} \left\{ \sum_{i=1}^{n} \underbrace{K}_{\text{weighting}} \left(\frac{|Z_i - c|}{\underbrace{h_n}_{\text{bandwidth}}} \right) \left(Y_i - a - \tau W_i - \beta_- (Z_i - c)_- - \beta_+ (Z_i - c)_+\right)^2 \right\}$$

;

Literature: Continuity-Based RD

Most popular: local linear regression (Hahn et al., 2001; G. W. Imbens and Lemieux, 2008)

$$\hat{\tau}_c = \arg\min_{\tau} \left\{ \sum_{i=1}^{n} \underbrace{K}_{\text{weighting}} \left(\frac{|Z_i - c|}{\underbrace{h_n}_{\text{bandwidth}}} \right) \left(Y_i - a - \tau W_i - \beta_- (Z_i - c)_- - \beta_+ (Z_i - c)_+ \right)^2 \right\}$$

- $\mu_{(w)}(z) = \mathbb{E}[Y(w) \mid Z = z]$ is **smooth**
- h_n decays at an **appropriate** rate

;

Literature: Continuity-Based RD

Most popular: local linear regression (Hahn et al., 2001; G. W. Imbens and Lemieux, 2008)

$$\hat{\tau}_c = \arg\min_{\tau} \left\{ \sum_{i=1}^{n} \underbrace{K}_{\text{weighting}} \left(\frac{|Z_i - c|}{\underbrace{h_n}_{\text{bandwidth}}} \right) \left(Y_i - a - \tau W_i - \beta_- (Z_i - c)_- - \beta_+ (Z_i - c)_+ \right)^2 \right\}$$

- $\mu_{(w)}(z) = \mathbb{E}[Y(w) \mid Z = z]$ is **smooth**
- h_n decays at an **appropriate** rate

Robust CIs (Armstrong and Kolesár, 2020; Calonico et al., 2014; Kolesár and Rothe, 2018);

Literature: Continuity-Based RD

Most popular: local linear regression (Hahn et al., 2001; G. W. Imbens and Lemieux, 2008)

$$\hat{\tau}_c = \arg\min_{\tau} \left\{ \sum_{i=1}^{n} \underbrace{K}_{\text{weighting}} \left(\frac{|Z_i - c|}{\underbrace{h_n}_{\text{bandwidth}}} \right) \left(Y_i - a - \tau W_i - \beta_- (Z_i - c)_- - \beta_+ (Z_i - c)_+\right)^2 \right\}$$

- $\mu_{(w)}(z) = \mathbb{E}[Y(w) \mid Z = z]$ is **smooth**
- h_n decays at an **appropriate** rate

Robust CIs (Armstrong and Kolesár, 2020; Calonico et al., 2014; Kolesár and Rothe, 2018); Data-adaptive bandwidths (G. Imbens and Kalyanaraman, 2012)

Literature: Continuity-Based RD extended

$$\mu_{(w)}(z) = \mathbb{E}\left[Y(w) \mid Z = z\right]$$

If further assume **convexity** of $\mu_{(w)}(z)$, e.g.:

$$\left|\mu''_{(w)}(z)\right| \leq B, \forall z \in \mathbb{R}$$

Literature: Continuity-Based RD extended

$$\mu_{(w)}(z) = \mathbb{E}\left[Y(w) \mid Z = z\right]$$

If further assume **convexity** of $\mu_{(w)}(z)$, e.g.:

$$\left|\mu''_{(w)}(z)\right| \leq B, \forall z \in \mathbb{R}$$

Optimization-based RD: the treatment effect τ_c can be estimated (minimax linear estimation) via **numerical convex optimization** (Armstrong and Kolesár, 2018; G. Imbens and Wager, 2019)

Link Noise-Induced RD and Continuity-Based RD

$$
\begin{aligned}
\mu_{(w)}(z) &= \mathbb{E}\left[Y(w) \mid Z = z\right] \\
&= \frac{\int \mathbb{E}\left[Y(w) \mid U = u\right] p(z \mid u)\, \mathrm{d}G(u)}{f_G(z)} = \frac{\int \alpha_{(w)}(u)\, p(z \mid u) \mathrm{d}G(u)}{\int p(z \mid u) \mathrm{d}G(u)}
\end{aligned}
$$

Link Noise-Induced RD and Continuity-Based RD

$$\mu_{(w)}(z) = \mathbb{E}[Y(w) \mid Z = z]$$

$$= \frac{\int \mathbb{E}[Y(w) \mid U = u]\, p(z \mid u)\, \mathrm{d}G(u)}{f_G(z)} = \frac{\int \boxed{\alpha_{(w)}(u)}\, p(z \mid u) \mathrm{d}G(u)}{\int p(z \mid u) \mathrm{d}G(u)}$$

Link Noise-Induced RD and Continuity-Based RD

$$\mu_{(w)}(z) = \mathbb{E}[Y(w) \mid Z = z]$$
$$= \frac{\int \mathbb{E}[Y(w) \mid U = u]\, p(z \mid u)\, \mathrm{d}G(u)}{f_G(z)} = \frac{\int \boxed{\alpha_{(w)}(u)}\, p(z \mid u) \mathrm{d}G(u)}{\int p(z \mid u) \mathrm{d}G(u)}$$

Then the worst-case possible curvature is:

$$\mathrm{Curv}(z, \rho, p) = \sup \left\{ \left| \frac{\mathrm{d}^2 \mu_{(w)}(z)}{\mathrm{d}z^2} \right| : f_G(z) = \int p(z \mid u)\, \mathrm{d}G(u) \geq \rho > 0, \alpha_{(w)}(\cdot) \in [0, 1] \right\}$$

Link Noise-Induced RD and Continuity-Based RD

$$\mu_{(w)}(z) = \mathbb{E}[Y(w) \mid Z = z]$$
$$= \frac{\int \mathbb{E}[Y(w) \mid U = u]\, p(z \mid u)\, \mathrm{d}G(u)}{f_G(z)} = \frac{\int \boxed{\alpha_{(w)}(u)}\, p(z \mid u) \mathrm{d}G(u)}{\int p(z \mid u) \mathrm{d}G(u)}$$

Then the worst-case possible curvature is:

$$\operatorname{Curv}(z, \rho, p) = \sup \left\{ \left| \frac{\mathrm{d}^2 \mu_{(w)}(z)}{\mathrm{d}z^2} \right| : f_G(z) = \int p(z \mid u)\, \mathrm{d}G(u) \geq \rho > 0, \alpha_{(w)}(\cdot) \in [0, 1] \right\}$$

Armstrong and Kolesár (2020): fit 4th-degree polynomials to $\mu_{(0)}(z)$ and $\mu_{(1)}(z)$, and take the largest estimated curvature obtained anywhere

Literature: Randomization Inference RD

Posit a non-trivial interval $\mathcal{I}$ with $c \in \mathcal{I}$ s.t.

$$\{Y_i(0), Y_i(1)\} \perp Z_i \mid \{Z_i \in \mathcal{I}\}$$

then focus on this interval, perform classical **randomized** study inference

Literature: Randomization Inference RD

Posit a non-trivial interval $\mathcal{I}$ with $c \in \mathcal{I}$ s.t.

$$\{Y_i(0), Y_i(1)\} \perp Z_i \mid \{Z_i \in \mathcal{I}\}$$

then focus on this interval, perform classical **randomized** study inference

- Design-based approach (Rubin, 2008)

Literature: Randomization Inference RD

Posit a non-trivial interval $\mathcal{I}$ with $c \in \mathcal{I}$ s.t.

$$\{Y_i(0), Y_i(1)\} \perp Z_i \mid \{Z_i \in \mathcal{I}\}$$

then focus on this interval, perform classical **randomized** study inference

- Design-based approach (Rubin, 2008)
- Strong assumption
  No **data-driven way** of choosing $\mathcal{I}$
  If the interval $\mathcal{I}$ is known a-priori, the problem collapses to a **RCT**

Measurement Error Induced RD

Rokkanen (2015) considers a similar approach, assuming:

Measurement Error Induced RD

Rokkanen (2015) considers a similar approach, assuming:

- noisy running variables **(A2)** and exogeneity of the noise **(A3)**
- **NOT** assuming prior knowledge of the noise distribution $p(\cdot \mid u)$

Measurement Error Induced RD

Rokkanen (2015) considers a similar approach, assuming:

- noisy running variables **(A2)** and exogeneity of the noise **(A3)**
- **NOT** assuming prior knowledge of the noise distribution $p(\cdot \mid u)$
- A stronger assumption: observing at least 3 noisy measurements of the latent variable U_i, $\{Z_i, Z_i', Z_i''\}$

Measurement Error Induced RD

Rokkanen (2015) considers a similar approach, assuming:

- noisy running variables **(A2)** and exogeneity of the noise **(A3)**
- **NOT** assuming prior knowledge of the noise distribution $p(\cdot \mid u)$
- A stronger assumption: observing at least 3 noisy measurements of the latent variable U_i, $\{Z_i, Z_i', Z_i''\}$

Measurement Error Induced RD

Rokkanen (2015) considers a similar approach, assuming:

- noisy running variables **(A2)** and exogeneity of the noise **(A3)**
- **NOT** assuming prior knowledge of the noise distribution $p(\cdot \mid u)$
- A stronger assumption: observing at least 3 noisy measurements of the latent variable U_i, $\{Z_i, Z_i', Z_i''\}$
  - (U_i, Z_i, Z_i', Z_i'') is **joint normal**
  - $\alpha_{(w)}(u) = \mathbb{E}\left[Y_i(w) \mid U_i = u\right]$ is **linear** w.r.t. u

RD with Ordinal Running Variables

Similarly, ordinal Z_i (bond rating, custody security score, etc.) can be seen as a noisy measurement of a latent variable U_i.
Li et al. (2021) assume

$$U_i = \mathbf{X}_i \beta$$

then use **inverse-propensity weighting** with estimated propensities $e(u) = \mathbb{P}\left[Z_i \geq c \mid U_i = u\right]$ for inference.

RD with Ordinal Running Variables

Similarly, ordinal Z_i (bond rating, custody security score, etc.) can be seen as a noisy measurement of a latent variable U_i.

Li et al. (2021) assume

$$U_i = \mathbf{X}_i \beta$$

then use **inverse-propensity weighting** with estimated propensities $e(u) = \mathbb{P}\left[Z_i \geq c \mid U_i = u\right]$ for inference.

Assuming: U_i can be observed, and well predicted by $\mathbf{X}_i$

Measurement Errors

- The **running variable** is unobserved, only a noisy measurement is observed
  Bartalotti et al. (2021), Davezies and Le Barbanchon (2017), Dong and Kolesár (2021), and Pei and Shen (2017)
- Measurement error in causal inference beyond RD
  Jiang and Ding (2020), Kuroki and Pearl (2014), and Pearl (2012)

A Comparison

RD designs	Assumptions
Noise-induced RD	a **known distribution** of the measurement error $p(\cdot \mid u)$
Noise-induced RD (Rokkanen, 2015)	**multiple joint-normal** noisy measurements (U_i, Z_i, Z_i', Z_i'')
	linear $\alpha_{(w)}(u) = \mathbb{E}\left[Y_i(w) \mid U_i = u\right]$
Continuity-based RD	$\mu_{(w)} = \mathbb{E}\left[Y(w) \mid Z = z\right]$ is **smooth**
OPtimization-based RD	**convexity** of $\mu_{(w)}(z)$: $\left\lvert\mu''_{(w)}(z)\right\rvert \leq B, \forall z \in \mathbb{R}$
Randomization inference RD	an **"RCT"** interval $\mathcal{I}$: $\{Y_i(0), Y_i(1)\} \perp Z_i \mid \{Z_i \in \mathcal{I}\}$
RD with ordinal Z_i	U_i can be observed, and well predicted by $\mathbf{X}_i$

References I

Armstrong, T. B., & Kolesár, M. (2018). Optimal inference in a class of regression models. *Econometrica, 86*(2), 655–683.

Armstrong, T. B., & Kolesár, M. (2020). Simple and honest confidence intervals in nonparametric regression. *Quantitative Economics, 11*(1), 1–39.

Bartalotti, O., Brummet, Q., & Dieterle, S. (2021). A correction for regression discontinuity designs with group-specific mismeasurement of the running variable. *Journal of Business & Economic Statistics, 39*(3), 833–848.

Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica, 82*(6), 2295–2326.

Davezies, L., & Le Barbanchon, T. (2017). Regression discontinuity design with continuous measurement error in the running variable. *Journal of econometrics, 200*(2), 260–281.

Dong, Y., & Kolesár, M. (2021). When can we ignore measurement error in the running variable? *arXiv preprint arXiv:2111.07388.*

References II

Eckles, D., Ignatiadis, N., Wager, S., & Wu, H. (2020). Noise-induced randomization in regression discontinuity designs. *arXiv preprint arXiv:2004.09458*.

Hahn, J., Todd, P., & Van der Klaauw, W. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, *69*(1), 201–209.

Imbens, G. W., & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of econometrics*, *142*(2), 615–635.

Imbens, G., & Kalyanaraman, K. (2012). Optimal bandwidth choice for the regression discontinuity estimator. *The Review of economic studies*, *79*(3), 933–959.

Imbens, G., & Wager, S. (2019). Optimized regression discontinuity designs. *Review of Economics and Statistics*, *101*(2), 264–278.

Jiang, Z., & Ding, P. (2020). Measurement errors in the binary instrumental variable model. *Biometrika*, *107*(1), 238–245.

Kolesár, M., & Rothe, C. (2018). Inference in regression discontinuity designs with a discrete running variable. *American Economic Review*, *108*(8), 2277–2304.

References III

Kuroki, M., & Pearl, J. (2014). Measurement bias and effect restoration in causal inference. *Biometrika, 101*(2), 423–437.

Li, F., Mercatanti, A., Mäkinen, T., & Silvestrini, A. (2021). A regression discontinuity design for ordinal running variables: Evaluating central bank purchases of corporate bonds. *The Annals of Applied Statistics, 15*(1), 304–322.

Pearl, J. (2012). On measurement bias in causal inference. *arXiv preprint arXiv:1203.3504*.

Pei, Z., & Shen, Y. (2017). The devil is in the tails: Regression discontinuity design with measurement error in the assignment variable. In *Regression discontinuity designs*. Emerald Publishing Limited.

Rokkanen, M. A. (2015). Exam schools, ability, and the effects of affirmative action: Latent factor extrapolation in the regression discontinuity design.

Rubin, D. B. (2008). For objective causal inference, design trumps analysis. *The annals of applied statistics, 2*(3), 808–840.

Thank you!