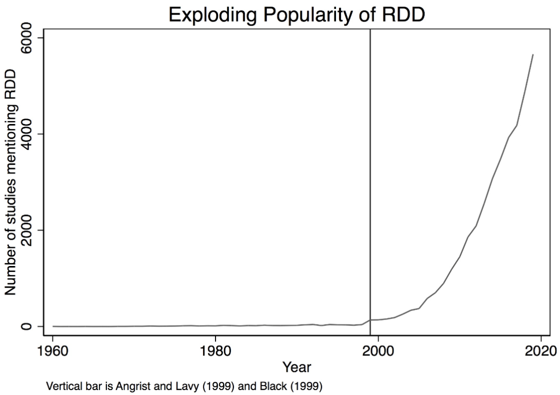
The Impacts of Optimal Bandwidths in Regression Discontinuity Design

Jack Fitzgerald

Vrije Universiteit Amsterdam and Tinbergen Institute

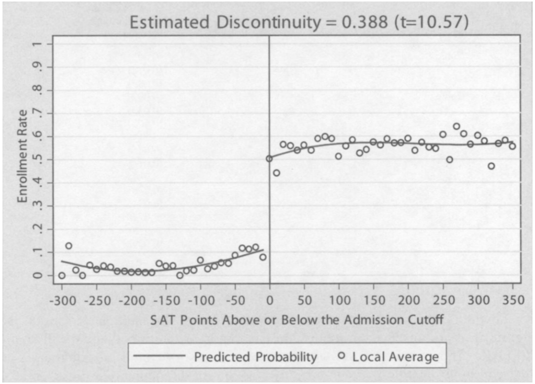
November 13, 2025

Regression Discontinuity Design (RDD)



Cunningham (2021) documents 5600 RDD papers published in 2019 alone

RDD's 'Experimental Appeal'



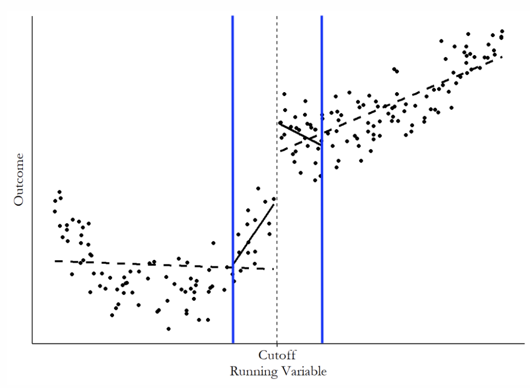
Source: Hoekstra (2009)

In principle, when an agent's running variable (RV) crosses the assignment cutoff, the agent should be effectively randomized into or out of treatment

- **Intuition:** Within a small bandwidth of a treatment cutoff, the question of whether you're one point below or above that cutoff is often as random as a coin flip

If we can accurately trace relationships between an outcome and the RV on both sides of the cutoff, then the jump in that outcome at the cutoff can be treated as the effect of a random treatment

Bandwidths



Modified from Huntington-Klein (2022)

RDD effects are only identified for observations whose RV values are right at the treatment cutoff

- To accurately trace relevant outcome-RV relationships as closely as possible, it is common in RDD to specify **bandwidths**
- A bandwidth restricts the sample to observations with RV values within a small window of the treatment cutoff

Bandwidth selection can make big differences for RDD estimates

- Bandwidth selection was historically (and still often is) *ad hoc*
- However, the last decade has seen an explosion in the development and availability of optimal, data-driven bandwidths

Popularity of Optimal Bandwidth Methods

Publication	Contribution	Scopus Citations (13-05-25)
Imbens & Kalyanaraman (2012)	Imbens-Kalyanaraman bandwidth	1147
Calonico, Cattaneo, & Titiunik (2014a)	`rdrobust` Stata package	371
Calonico, Cattaneo, & Titiunik (2014b)	MSE-optimal bandwidth	1461
Calonico, Cattaneo, & Titiunik (2015)	`rdrobust` R package	100
Calonico et al. (2017)	`rdrobust` R and Stata packages	467
Calonico et al. (2019)	MSE-optimal bandwidth with covariates	281
Calonico, Cattaneo, & Farrell (2020)	Coverage error-optimal bandwidth	178

Optimal bandwidth methods are now quite popular, being used in hundreds of RDD papers, but are far from universally-applied

- Compare these figures to the 5600 RDD papers Cunningham (2021) documented in 2019

From 2015-2018, around 53% of RDD publications in top political science journals used optimal bandwidths (Stommes, Aronow, & Sävje 2023)

This Project

I examine how optimal bandwidths affect RDD estimation in a large-scale reanalysis

- I leverage replication data on 36 RDD publications from 2009-2018 and compare full-sample to optimal-bandwidth specifications (Stommes, Aronow, & Sävje 2023)
- I employ *ceteris paribus* specifications which vary only one choice in the estimation process at a time, which lets me isolate the bandwidths' effects

I quantify the extent to which optimal bandwiths affect power, precision, and robustness in RDD

- In the median specification, optimal bandwidths slash sample sizes by over 60%, though treatment-control sampling ratios become significantly more balanced
- On average, MSE-optimal bandwidths increase TE SEs by over 59-74%, and coverage error-optimal bandwidths increase them by nearly 116%
- Optimal bandwidth specification changes conclusions for 24-31% of RDD estimates; over 70% of this effect remains after holding variance constant

Results imply that optimal bandwidths are necessary for robust RDD estimation, but that researchers must reckon with substantial power costs

Why Should We Care?

Implications for standard practice in RDD estimation

- Given their necessity for robustness, you should insist upon optimal bandwidths in RDD as a researcher, colleague, and reviewer
- You should also insist upon *effective* sample size reporting – can be very different from *nominal* effect sizes

Useful for understanding the power crisis in social science

- Now well-documented that social science is dramatically underpowered (Ioannidis, Stanley, & Doucouliagos 2017; Arel-Bundock et al. 2024; Askarov et al. 2024)
- A key ingredient is power limitations in causal inference methodologies popularized by the credibility revolution

Helps researchers anticipate effective sample sizes in RDD

- Before committing to a costly data source, expect to lose over half your observations after specifying optimal bandwidths, and think about what this means for power

Setup

I consider a parametric RDD framework

$$Y_i = \alpha + \underbrace{\sum_{j=1}^{p} \beta_j Z_i^j}_{\text{OLS fit before cutoff}} + \underbrace{\sum_{k=1}^{p} \gamma_k \left(Z_i^k \times \mathbb{1}\left[Z_i > c\right]\right)}_{\text{OLS fit after cutoff}} + \underbrace{\tau \mathbb{1}\left[Z_i > c\right]}_{\text{Jump at cutoff}} + \epsilon_i, \tag{1}$$

Y_i is the outcome of interest, c is the cutoff, Z_i is the recentered running variable (RV), and p is the polynomial order

- τ is the treatment effect (TE) parameter of interest

Like most software, I consider symmetric bandwidths of the form $[c - h,\ c + h]$, where $h > 0$

- Specifying a bandwidth of width h amounts to running Equation 1 only on the subsample of observations for whom $Z_i \in [c - h,\ c + h]$

Why Specify Bandwidths?

It's well-established that RDD identifies a LATE – the effect of a treatment for compliers (Angrist, Imbens, & Rubin 1996; Hahn, Todd, & Van der Klaauw 2001)

- Here compliers are those whose treatment status changes (in the right way) because their RV goes from just below to just above the treatment cutoff

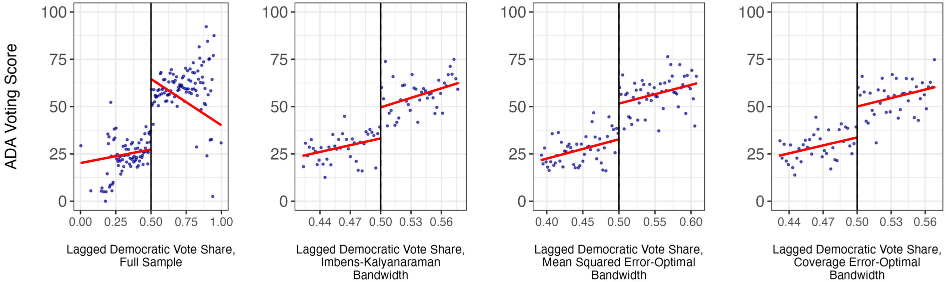
RDD estimates of TEs are thus not generalizable for outliers with RV values far from the cutoff

- If the relationship between the outcome and RV is different for these outliers than for compliers, then including these outliers in the estimation will yield biased estimates of identifiable TEs

Bandwidth specification in RDD attempts to mitigate outlier-driven bias on TE estimates by only estimating TEs on the most relevant observations – those closest to the cutoff

- Classic bias-variance tradeoff; restricting the sample gives us less bias, but also less power

Applied Example



Outliers can skew estimated relationships between the outcome and the RV, affecting where those estimated relationships intersect the cutoff

- In this example from Lee, Moretti, & Butler (2004) and Cunningham (2021), optimal bandwidths restore an appropriate regression slope above the cutoff, but require cutting 60-74% of the sample

The Imbens-Kalyanaraman (IK) Bandwidth

Imbens & Kalyanaraman (2012) first proposed the IK optimal bandwidth estimator, which attempts to minimize the mean squared error (MSE) of the RDD estimate

- Useful target to minimize for optimizing the bias-variance tradeoff, considering that MSE $\approx$ Variance $+$ Bias^2

`rdrobust`'s implementation of the symmetric IK bandwidth can be written as follows, with three key ingredients (Calonico, Cattaneo, & Titiunik 2014a; 2014b)

$$\hat{h}_{\text{IK}} = \left[\frac{\hat{V}_{\text{IK}}}{2(p+1)\hat{B}^2_{\text{IK}} + \hat{R}_{\text{IK}}}\right]^{\frac{1}{2p+3}} n^{-\frac{1}{2p+3}} \quad (2)$$

1. $\hat{V}_{\text{IK}}$: Estimated asymptotic variance of RDD estimate. *Ceteris paribus*, more variance $\rightarrow$ wider bandwidths
2. $\hat{B}_{\text{IK}}$: Estimated bias of RDD estimate. *Ceteris paribus*, more bias $\rightarrow$ tighter bandwidths
3. $\hat{R}_{\text{IK}}$: Regularization term to prevent small denominators

The MSE-Optimal Bandwidth

Calonico, Cattaneo, & Titiunik (2014b) propose an alternative MSE-optimal bandwidth

- Similar to the IK bandwidth, but with asymptotically optimal preliminary bandwidths

`rdrobust`'s implementation of the symmetric MSE-optimal bandwidth can be written as follows (Calonico et al. 2017; 2019)

$$\hat{h}_{\text{MSE}} = \left[\frac{\hat{V}_{\text{MSE}}}{2(p+1)\hat{B}^2_{\text{MSE}}}\right]^{\frac{1}{2p+3}} n^{-\frac{1}{2p+3}} \tag{3}$$

This MSE-optimal bandwidth also has no $\hat{R}_{\text{IK}}$ regularization term

- This is the default bandwidth specified in `rdrobust`

The CE-Optimal Bandwidth

Calonico, Cattaneo, & Farrell (2020) propose bandwidths that try to minimize coverage error (CE) instead of MSE

- I.e., CE-optimal bandwidths try to minimize the probability that the RDD estimate's confidence interval fails to cover the true TE
- Easy way to accomplish this is just to make the confidence interval really wide

`rdrobust`'s implementation of the symmetric CE-optimal bandwidth can be written as follows (Calonico et al. 2017)

$$\hat{h}_{CE} = \hat{h}_{MSE} \times n^{-\frac{p}{(3+p)(3+2p)}} \tag{4}$$

I.e., the CE-optimal bandwidth just shrinks the MSE-optimal bandwidth in Equation 3 by a fractional constant functionally given by the sample size and polynomial order

Replication Data

I leverage replication data from Stommes, Aronow, & Sävje (2023), who run robustness checks on 36 published RDD papers in *AJPS*, *APSR*, and *JOP* from 2009-2018

- Some papers use multiple datasets and outcome variables; I run analyses on each outcome-dataset combination, giving me 62 total RDD estimates to analyze

Used in other large-scale analyses of RDD findings (Fitzgerald 2025)

- Designs in this dataset include close election designs, spatial discontinuities, and age discontinuities
- 42% of historical economics RDD papers leverage one of these RVs (Lee & Lemieux 2010)
- Bandwidth properties observed in this data likely translate to many RDD applications in economics and other social sciences

Ceteris Paribus Specifications

My reference estimates are the published estimates recorded by Stommes, Aronow, & Sävje (2023)

- These are compared with estimates I obtain from the underlying studies' replication data

I vary 4x2x2 features of the specifications:

- **Bandwidths:** None, IK, MSE, and CE; last three computed using the `rdbwselect` command in R (Calonico, Cattaneo, & Titiunik 2014a; Calonico et al. 2017)
- **Functional form:** Linear or quadratic, based on recommendation of Gelman & Imbens (2018)
- **Estimator:** Conventional or bias-corrected (Calonico et al. 2014)

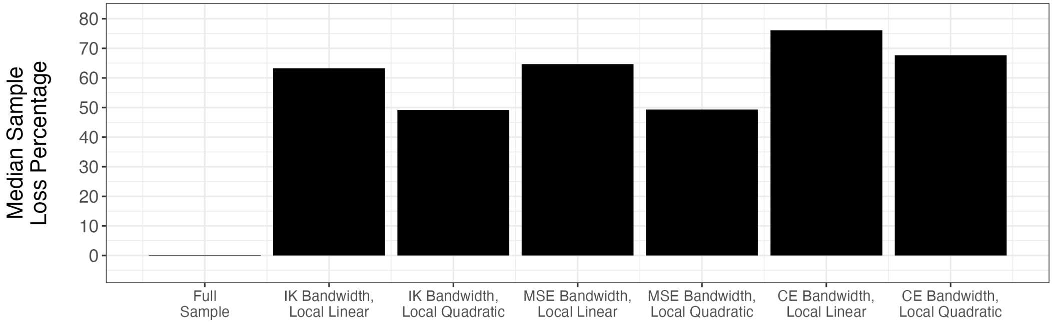
I thus have results from 17 specifications for each estimate

- Because I vary each specification choice one at a time, my **ceteris paribus specifications** can identify the specific effect of each specification choice on a study's results

All 16 of my robustness specifications are estimated using the `rdrobust` package in R (Calonico, Cattaneo, & Titunik 2014; Calonico et al. 2017)

- All estimation is done on uniform kernels, ensuring there's no interaction with kernel weighting

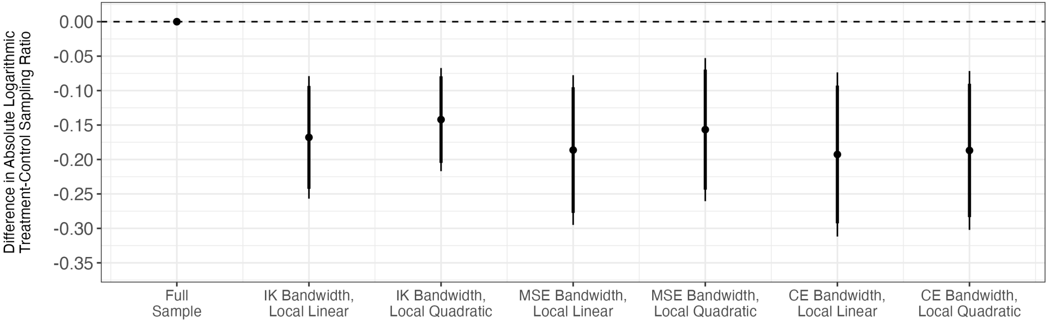
Sample Size Loss



Across all specifications and bandwidths, optimal bandwidths cut over 60% of the median specification's sample

- Across bandwidths, the within-bandwidth median sample loss rate is 49-76%

Treatment Imbalance

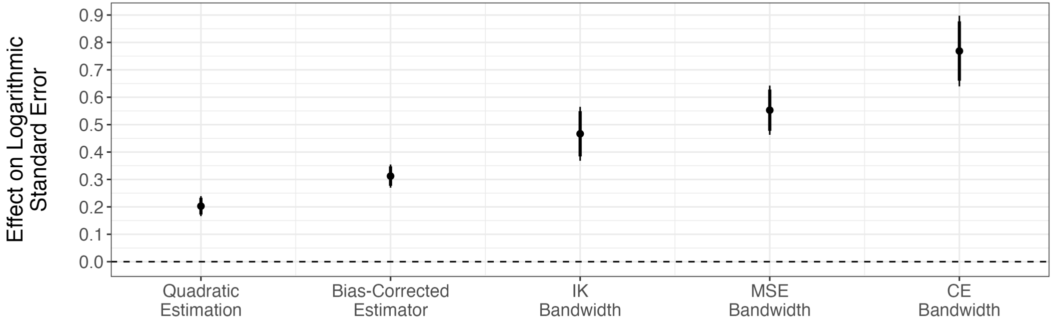


Contrary to my expectations, on average, specifying optimal bandwidths decreases absolute sampling ratios by 0.14-0.19 log points (13-18%)

- I.e., specifying optimal bandwidths significantly *decreases* treatment imbalance

Measuring Treatment Imbalance

Standard Errors



IK, MSE, and CE bandwidths increase SEs by 59%, 74%, and 116% on average

- Bandwidths have much greater effects on precision than quadratic vs. linear estimation (22%) and bias-corrected vs. conventional estimation (37%)

Robustness, Bias, and Variance

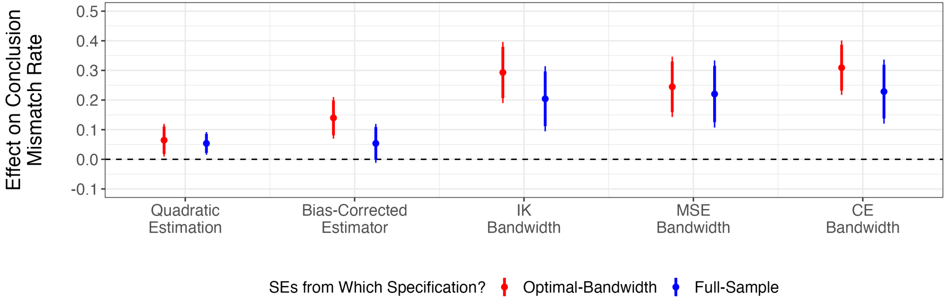
Robustness: Do you get the same conclusion from the optimal-bandwidth specification as you do from the published result?

- I define 'conclusion' using the standard tripartite standard NHST classification for a 5% α
- The main effect of interest is that on the **conclusion mismatch rate**

But conclusions may change from reductions in both bias + power; **which effect dominates?**

- We can isolate bias effects from variance effects by examining **'hybrid estimates'**, where a given specification's coefficient is assigned the published estimate's SE
- **Intuition:** How does a specification choice change results if we hold variance constant?
- What's left over can be attributed to bias reduction

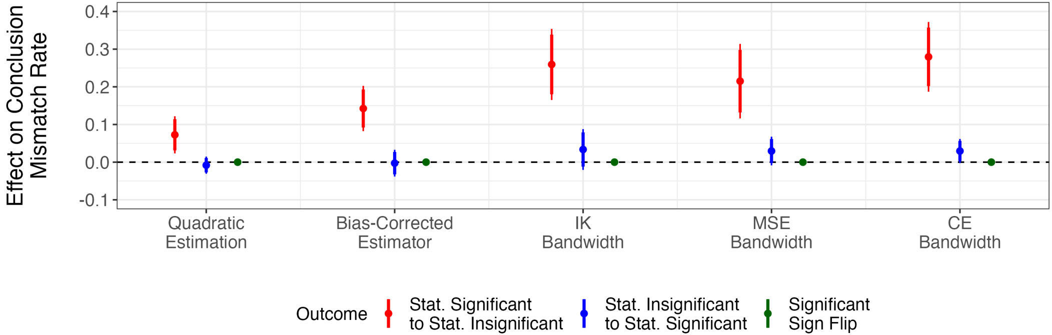
Robustness



Optimal bandwidth specification changes conclusions for 24-31% of RDD papers

- At least 70% of this effect remains after holding the published SE constant
- **Bandwidths are again dominant:** Differences in bandwidth specification change conclusions much more often than functional form (6%) or bias correction (14%)

Robustness Results Decomposed



Over 86% of these specification effects on conclusion agreement are driven by originally statistically significant results losing their statistical significance

- No significant sign flips are observed

Takeaways

Optimal bandwidths are necessary for robustness in RDD

- You cannot get close to generalizable, unbiased RDD estimates without bandwidths
- Over 24% of published RDD results in top journals change when optimal bandwidths are specified even after holding full-sample SEs constant, way more than other specification choices

Most RDD analyses are underpowered, even more than they first seem

- This is not optimal bandwidths' fault
- Optimal bandwidths *reveal* a lack of power to detect generalizable effects; they don't *cause* it

Before starting RDD analyses, plan to lose 60% of your sample

- Important consideration before committing to expensive administrative data, and when conducting power analysis

Insist on reporting effective sample sizes in RDD

- Probably less than half the full sample size!

References I

Angrist, J. D., G. W. Imbens, and D. B. Rubin (1996).
Identification of causal effects using instrumental variables.
Journal of the American Statistical Association 91(434), 444–455.

Arel-Bundock, V., R. C. Briggs, H. Doucouliagos, M. Mendoza Aviña, and T. Stanley (2024).
Quantitative political science research is greatly underpowered.
The Journal of Politics Forthcoming.

Calonico, S., M. D. Cattaneo, and M. H. Farrell (2020).
Optimal bandwidth choice for robust bias-corrected inference in regression discontinuity designs.
The Econometrics Journal 23(2), 192–210.

Calonico, S., M. D. Cattaneo, M. H. Farrell, and R. Titiunik (2017).
rdrobust: Software for regression-discontinuity designs.
The Stata Journal 17(2), 372–404.

References II

Calonico, S., M. D. Cattaneo, M. H. Farrell, and R. Titiunik (2019).
Regression discontinuity designs using covariates.
The Review of Economics and Statistics 101(3), 442–451.

Calonico, S., M. D. Cattaneo, and R. Titiunik (2014a).
Robust data-driven inference in the regression-discontinuity design.
The Stata Journal 14(4), 909–946.

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

Calonico, S., M. D. Cattaneo, and R. Titiunik (2015).
Rdrobust: An R package for robust nonparametric inference in regression-discontinuity designs.
The R Journal 7(1), 38–51.

References III

Cunningham, S. (2021, Aug).
Causal inference: The mixtape (1 ed.).
Yale University Press.

Fitzgerald, J. (2025).
Manipulation tests in regression discontinuity design: The need for equivalence testing.
MetaArXiv.

Gelman, A. and G. Imbens (2018, Jul).
Why high-order polynomials should not be used in regression discontinuity designs.
Journal of Business & Economic Statistics 37(3), 447–456.

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

References IV

Hoekstra, M. (2009).
The effect of attending the flagship state university on earnings: A discontinuity-based approach.
Review of Economics and Statistics 91(4), 717–724.

Huntington-Klein, N. (2022, Aug).
The Effect: An introduction to research design and causality (1 ed.).
CRC Press.

Imbens, G. and K. Kalyanaraman (2012).
Optimal bandwidth choice for the regression discontinuity estimator.
The Review of Economic Studies 79(3), 933–959.

Lee, D. S., E. Moretti, and M. J. Butler (2004).
Do voters affect or elect policies? evidence from the u. s. house.
The Quarterly Journal of Economics 119(3), 807–859.

References V

Stommes, D., P. M. Aronow, and F. Sävje (2023, Apr).
On the reliability of published findings using the regression discontinuity design in political science.
Research & Politics 10(2), 205316802311664.

Measuring Treatment Imbalance Extremity

A natural measure of treatment imbalance is the (inverse) ratio of treated to untreated observations

- **Complication:** Ratios are asymmetric about one
- **Solution:** Their natural logarithms are symmetric about zero

I thus measure treatment imbalance using the absolute logarithmic sampling ratio

$$\text{ALSR} = \left| \ln \left(\frac{N_1}{N_0} \right) \right|,$$

where N_1 and N_0 are respectively the number of treated and untreated observations

- Similar to logarithmic density ratio proposed in Fitzgerald (2025)

Back