Boundary Discontinuity Designs: Theory and Practice*

Matias D. Cattaneo† Rocio Titiunik‡ Ruiqi (Rae) Yu§

November 3, 2025

Abstract

We review the literature on boundary discontinuity (BD) designs, a powerful non-experimental research methodology that identifies causal effects by exploiting a thresholding treatment assignment rule based on a bivariate score and a boundary curve. This methodology generalizes standard regression discontinuity designs based on a univariate score and scalar cutoff, and has specific challenges and features related to its multi-dimensional nature. We synthesize the empirical literature by systematically reviewing over 80 empirical papers, tracing the method's application from its formative uses to its implementation in modern research. In addition to the empirical survey, we overview the latest methodological results on identification, estimation and inference for the analysis of BD designs, and offer recommendations for practice.

Keywords: regression discontinuity, treatment effects estimation, causal inference.

*This chapter was prepared for the invited session on "Causal Inference and Statistical Decisions" at the 2025 World Congress of the Econometric Society, Seoul (Korea). We thank Alberto Abadie, Xiaohong Chen, Boris Hanin, Kosuke Imai, Xinwei Ma, Francesca Molinari, Jörg Stoye, and Jeff Wooldridge for comments and discussions. Financial support from the National Science Foundation through grants SES-2019432, DMS-2210561, and SES-2241575 is gratefully acknowledged. Yu was also partially supported through the Data-Driven Social Science initiative at Princeton University.

†Department of Operations Research and Financial Engineering, Princeton University.

‡Department of Politics, Princeton University.

§Department of Operations Research and Financial Engineering, Princeton University.

Contents

1 Introduction

The Regression Discontinuity (RD) design is one of the leading observational methods for program evaluation and causal inference (see Abadie and Cattaneo, 2018; Hernán and Robins, 2020, for overviews and further references). In its canonical form, a binary treatment is assigned to units whose value of a score is equal to or above a known cutoff, and not assigned to units whose score is below the cutoff. Under the key assumption that all observable and unobservable pretreatment characteristics vary smoothly around the cutoff, the discontinuous change generated by the hard-thresholding treatment assignment rule based on the univariate score and scalar cutoff can be used to learn about causal treatment effects. Lee and Lemieux (2010) offer an early review with a list of standard RD empirical applications, Cattaneo and Titiunik (2022) give a recent overview of the RD methodological literature, and Cattaneo et al. (2020, 2024) provide a practical introduction to modern RD methods and practice.

The Boundary Discontinuity (BD) design generalizes the RD design to the case where the score is bivariate, and the treatment is assigned to units according to the location of their score relative to a known boundary that splits its support into two disjoint regions. This setup is sometimes called the Multi-Score RD design (Papay et al., 2011; Reardon and Robinson, 2012; Wong et al., 2013), or the Geographic RD design (Keele and Titiunik, 2015; Keele et al., 2015; Keele and Titiunik, 2016; Keele et al., 2017; Galiani et al., 2017; Rischard et al., 2021; Diaz and Zubizarreta, 2023). In this chapter, we discuss the empirical strategies most commonly employed for the analysis and interpretation of BD designs, and review the most recent theoretical and methodological results characterizing their econometric properties (Cattaneo et al., 2025a,b,c).

We introduce the BD design formally in Section 2, and then outline our review of the empirical literature in Economics, Political Science, Education, and other disciplines. Table 1 reports over 80 empirical papers, which form the basis for the organization of this chapter. We find that the overwhelming majority of papers collapse the bivariate score into a univariate

distance score measuring the shortest distance to the treatment assignment boundary, and then report a single average treatment effect estimated by pooling all the observations that are close to the boundary, regardless of their specific location. In contrast, only a handful of empirical papers investigate heterogeneity by localizing to specific regions or points along the boundary, despite the potential importance of this rich information for policy evaluation and decision-making.

Following our review of the empirical literature, in Section 3 we discuss pooling-based methods. We begin by discussing the seminal work of Card and Krueger (1994), who appear to be the first to employ the BD design conceptually, although they did not localize to the assignment boundary for estimation purposes. We then discuss Holmes (1998) and Black (1999), who appear to be the first two papers to employ localization to the assignment boundary explicitly as their main empirical strategy. Subsequent empirical work has recognized the importance of localization to the boundary, but has considered different empirical strategies for treatment effect estimation via local flexible regression methods. One of the most influential papers is Dell (2010), who incorporated a polynomial expansion of the bivariate location score as part of the local regression specification. More recently, an alternative empirical approach views the estimation based on distance to the boundary as a pooled univariate RD design, and thus estimates treatment effects using a local regression including a polynomial expansion of the univariate distance score. Recent examples include Ito and Zhang (2020) and Dehdari and Gehring (2022). Many empirical papers also include boundary-segment fixed effects in their local regression specifications, and some include interactions with the treatment indicator and/or between the polynomial expansions of the univariate distance score and the bivariate location score. Furthermore, for simplicity, or because distinct treatments occur in different regions of the boundary, some papers reduce the BD design to boundary-segment-specific univariate RD designs—see Ou (2010), Londoño-Vélez et al. (2020) and Salti et al. (2022). Putting all together, Section 3 discusses eight distinct pooled local regression specifications commonly found in empirical work.

Building on our review of empirical strategies used for pooled estimation in BD designs, in Section 3.1 we also overview the main methodological results in Cattaneo et al. (2025c) for pooling-based estimation and inference. These methods are viewed as univariate RD designs based on the shortest distance to the assignment boundary, and hence are related to identification, estimation, and inference over a specific one-dimensional submanifold curve on the plane, defined by the assignment boundary. Recognizing this mathematical structure, Cattaneo et al. (2025c) establish their formal results leveraging concepts from geometric measure theory (Federer, 2014). After reviewing these ideas, we briefly discuss how estimation and inference can be conducted using the state-of-the-art methods from the RD design literature (Calonico et al., 2014, 2019, 2020). Some of this work is conceptually related to a recent paper by Chen and Gao (2025), who study estimation and inference for integral functionals on submanifolds when using nonparametric sieve estimation.

Section 4 discusses heterogeneity analysis and subsequent aggregation of causal treatment effects along the boundary. Although the boundary average treatment effect parameter emerging from pooling-based methods discussed in Section 3 is often of interest, and provides a useful aggregate measure of average treatment effects, the richness of the two-dimensional assignment in BD designs offer the opportunity to explore heterogeneous treatment effects along the assignment boundary. In geographic applications, this heterogeneity is directly linked to the geographic location of the units, providing useful information about how treatment effects vary in space; in non-geographic settings, the heterogeneity is directly linked to features of the units that are captured by the bivariate score variable, and hence the interpretation is necessarily application-specific. Thus, Section 4 begins by introducing the boundary average treatment effect curve, which captures the average treatment effect for each point along the assignment boundary. We also discuss how other aggregate causal effects can be recovered from that building block parameter, such as the weighted boundary average treatment effect and the largest average treatment effect.

After introducing the main ideas and causal parameter of interest, Section 4 reviews the

theoretical and methodological takeaways from Cattaneo et al. (2025a) and Cattaneo et al. (2025b). These two recent papers investigate identification, estimation, and inference for heterogeneous treatment effects along the assignment boundary, and transformations thereof, from two connected but distinct perspectives: Cattaneo et al. (2025a) studies boundary-point specific distance-based methods, which are directly motivated by the poloing-based approaches discussed in Section 3, while Cattaneo et al. (2025b) studies methods based on the bivariate location directly. For each approach, the estimation and inference results are valid both pointwise and uniformly along the assignment boundary, in addition to providing the basis for constructing other treatment effect estimators.

Finally, Section 5 offers recommendations for practice, and concludes. The methods discussed in this chapter can be implemented using the general-purpose software packages `rdrobust`, `rdhte`, `rdmulti`, and `rd2d`, located at `https://rdpackages.github.io/`, where replication files and related references are also available.

2 The Boundary Discontinuity Design

Each unit i in the study, $i = 1, \ldots, n$, has a continuous bivariate score $\mathbf{X}_i = (X_{1i}, X_{2i})$ that takes values in the set $\mathscr{X} \subseteq \mathbb{R}^2$. The assignment of units to the treatment or control condition depends on the location of their score $\mathbf{X}_i$ relative to a known one-dimensional boundary curve $\mathscr{B}$ that splits $\mathscr{X}$ into two disjoint regions: $\mathscr{X} = \mathscr{A}_0 \cup \mathscr{A}_1$, with $\mathscr{A}_0$ and $\mathscr{A}_1$ the control and treatment disjoint (connected) regions, respectively, and $\mathscr{B} = \mathtt{bd}(\mathscr{A}_0) \cap \mathtt{bd}(\mathscr{A}_1)$, where $\mathtt{bd}(\mathscr{A}_t)$ denotes the boundary of the set $\mathscr{A}_t$. (We follow convention and assume that the boundary belongs to the treatment group, that is, $\mathtt{bd}(\mathscr{A}_1) \subset \mathscr{A}_1$ and $\mathscr{B} \cap \mathscr{A}_0 = \emptyset$.) The observed outcome variable is Y_i.

An ancestor of the BD design was used by Card and Krueger (1994) in their seminal study of the effects of the minimum wage on employment. In their analysis, the treatment of interest is the increase of the state minimum wage in New Jersey that went into effect on

April 1, 1992, and rose the minimum wage to $5.05. The authors compared a sample of fast food restaurants in New Jersey to a sample of fast food restaurants in eastern Pennsylvania, a state that shares a border with New Jersey and where the minimum wage did not increase. After the increase in New Jersey, the average starting wage was $5.08 in the New Jersey sample and $4.62 in the Pennsylvania sample.

The rationale for comparing restaurants in New Jersey to restaurants in eastern Pennsylvania was that, by focusing on close geographic areas, these restaurants would be subject to similar economic conditions that would otherwise confound the effects of changes to the minimum wage. Although Card and Krueger (1994) also relied on a difference-in-differences design comparing both samples before and after the increase, they explicitly mentioned the close geographic proximity between the areas to justify their choice of comparison group, as they believed that fast-food stores in eastern Pennsylvania formed a "natural basis for comparison with the experiences of restaurants in New Jersey" due to Pennsylvania being a "nearby state[s]" (Card and Krueger, 1994, p. 773).

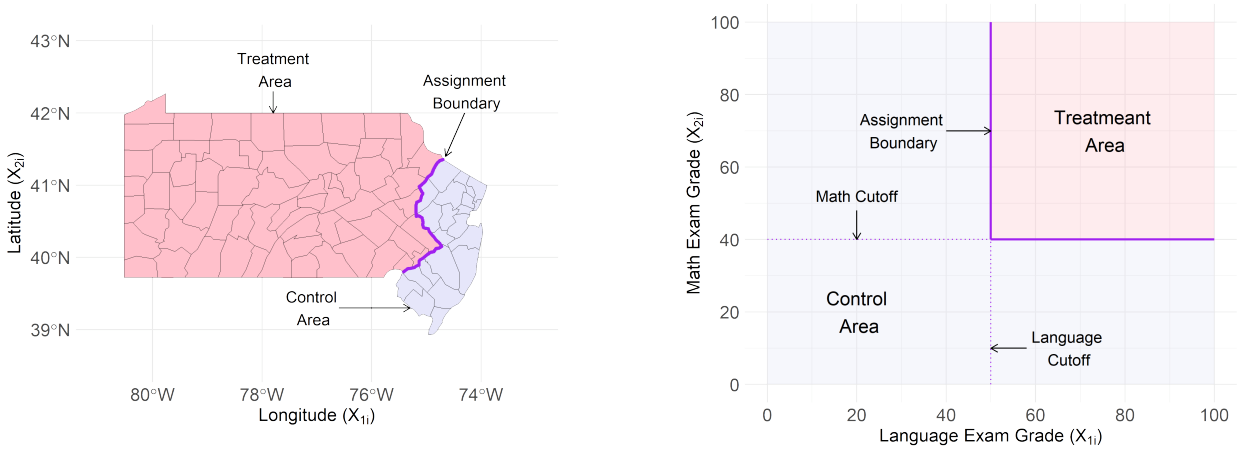
The use of close proximity as the basis for reducing potential confounders and enhancing the credibility of causal interpretations bears a direct connection to the BD design. The main difference is that Card and Krueger (1994) included non-adjacent areas in New Jersey and Pennsylvania (e.g., the New Jersey sample included restaurants in the New Jersey shore, far away from the Pennsylvania border), and did not directly use the unit's distance to the boundary in their analysis. The explicit use of distance to the assignment boundary is central to the BD design, which is why we consider Card and Krueger (1994) to be an ancestor rather than an instance of the BD design itself: while the authors leverage the change in treatment assignment induced by a geographic boundary, they did not explicitly localize to the boundary itself.

Holmes (1998) and Black (1999) are the earliest empirical examples that we could find employing the BD design. Holmes (1998) used a BD design to study the effect of pro-business policies on manufacturing, comparing states that adopted right-to-work laws with adjacent

states that did not. Using the longitude and latitude coordinates of each county's centroid, he calculated the minimum distance of the centroid to the assignment border, and focused the analysis on counties close to the border as a way to "control for differences across states in these various characteristics that are unrelated to policy" (Holmes, 1998, p. 671). Similarly, Black (1999) used a BD design to infer the quality of public schools from housing prices. She compared house prices on opposite sides of school attendance district boundaries, restricting the sample to houses close to the boundary. The justification for this localization was that houses on opposite sides of the boundary within a small area around it would be similar in all characteristics except for school quality—as the latter changes discontinuously at the attendance district boundary. In these examples, the bivariate score is the pair of latitude and longitude coordinates that determines the location of each unit (county or house), and the assignment boundary is the geographic border between two adjacent states or school attendance districts; the treatment is the policy or feature that changes abruptly at the border—labor policy for Holmes (1998) and school quality for Black (1999). This type of BD design is often referred to as a Geographic RD design.

One of the most common non-geographic applications of the BD design occurs in education when a treatment is given on the basis of two exam scores. For example, Ou (2010) studies the effect of a high school exit exam on the likelihood of dropping out of high school. The exam has a mathematics and a language arts component, each of which is graded separately; students must achieve a minimum proficiency in each section in order to pass the exam. The bivariate score in this case is the pair of mathematics and language test scores, and the treatment is passing the exam.

Figure 1 illustrates the boundary and treatment assignment regions in two stylized settings corresponding to the geographic and non-geographic examples. Figure 1a illustrates the geographic BD design case, where units are split into adjacent treated and control areas according to their location with respect to a geographic border, as in Black (1999) and Holmes (1998). (This figure shows the New Jersey-Pennsylvania border as considered by



(a) Geographic BD Design

(b) Non-geographic BD Design

Figure 1: BD Design Illustration

Card and Krueger (1994).) Figure 1b illustrates the non-geographic setting, using as example the study by Ou (2010) where students receive two exam scores and a treatment is given only to those individuals who score above a minimum cutoff in each exam. In geographic applications, the bivariate score is always composed of the geographic coordinates of the units. In non-geographic applications, the bivariate score could include different types of variables; in addition to test scores from multiple exams, non-geographic bivariate scores include shares of airline passengers in origin and destination cities (Snider and Williams, 2015), tax rates and income ratios (Egger and Wamser, 2015), systolic and diastolic blood pressure measurements (Dai et al., 2022), different component scores in a means testing formula (Salti et al., 2022; Kämpfen and Mosca, 2024), and population counts at different time periods (Hinnerich and Pettersson-Lidbom, 2014).

Figure 1 illustrates an important distinction between the geographic and non-geographic examples. In non-geographic examples, the function that describes the boundary tends to be known, and is typically linear with few irregularities and very few kink points. In contrast, in geographic BD designs the boundary is typically the border between political or administrative subdivisions; as a consequence, its shape is only available as coordinates on a map, and its geometry can be complicated. In the best case possible, it will be piecewise

linear with many kink points, but it will be more irregular in many applications. In his early study, Holmes (1998, p.681-682) already recognized this challenge and discussed explicitly the complex boundary shapes created by geographic borders. Indeed, since the influential work by Mandelbrot (1967, 1983), there has been an ongoing debate among mathematicians and philosophers trying to decide whether geographic borders (and other shapes in nature) are fractals (see Avnir et al., 1998, and references therein). From a methodological perspective, as recently formalized in Cattaneo et al. (2025a,b,c), the geometry of the assignment boundary in BD designs has fundamental implications for identification, estimation, and inference. We will review some of these recent findings in the upcoming sections.

In the standard RD design, the distance between the univariate score and the single cutoff is naturally always measured by the Euclidean distance, and observations are "close" to the cutoff c when their univariate score X_i implies a "small" distance $|X_i - c|$. In contrast, in the BD design, "closeness" to the boundary can be defined in different ways, depending on whether the parameter of interest is defined relative to a specific point on the boundary, or as some aggregate thereof along the entire boundary, such as an average or supremum. Furthermore, due its multivariate nature, the analysis can depend on the specific notion of distance used.

In the next sections, we discuss the two main approaches to the analysis of BD designs. The first approach uses the variation in treatment assignment induced by entire boundary simultaneously, and thus pools all observations sufficiently "close" the boundary in a single analysis, producing a single average treatment effect across all boundary points. This pooling approach relies on localization to $\mathscr{B}$ by first computing the distance between each observation's location and the nearest point on the boundary, denoted by D_i, and then retaining only those observations for which this distance is no larger than a specific bandwidth, denoted by h. The advantage of this approach is that it naturally mimics the univariate RD design, using the bivariate location of each unit to build a univariate score that is used directly in a "standard" univariate RD analysis. This approach thus produces a single average

treatment effect estimate.

In contrast, a second approach focuses on estimating the average treatment effect at each point on the boundary, performing localization relative to the specific point of analysis, thereby retaining observations whose distance to that point is no greater than a chosen bandwidth. Estimation and inference are conducted separately at each boundary point $\mathbf{x} \in \mathscr{B}$, which captures the heterogeneity of treatment effects along the boundary. Aggregate causal effects can then be recovered by averaging (or applying other transformations to) these boundary-point specific average treatment effects, including as a special case the average treatment effect along the entire boundary that is the focus of the pooling approach. The boundary-point-specific approach can be seen as focusing on heterogeneous average treatment effects that can then be used as building blocks to construct more aggregated causal parameters. The advantage of this approach is that it captures the full richness of the BD design by first learning the heterogeneity of the average treatment effect along the boundary, which usually summarizes unit characteristics of high policy relevance such as geographic location, academic performance, or economic need.

2.1 BD Designs in Practice

We searched the academic literature in multiple social sciences to learn how applied researchers implemented the BD design. We conducted our search in four general queries. In our initial Economics query, we included ten leading journals in Economics—*Journal of Political Economy*, *Review of Economic Studies*, *Econometrica*, *American Economic Review*, *Quarterly Journal of Economics*, *Review of Economics and Statistics*, *American Economic Journal: Applied Economics*, *American Economic Journal: Economic Policy*, *American Economic Journal: Macroeconomics*, and *American Economic Journal: Microeconomics*. In our initial Political Science query, we searched six leading Political Science journals—*American Political Science Review*, *American Journal of Political Science*, *Journal of Politics*, *Political Science Research and Methods*, *British Journal of Political Science*, and *Quarterly Journal of*

Political Science. In both cases we searched for terms like "boundary discontinuity", "multivariate regression discontinuity", and "multiscore regression discontinuity", and focused our attention on papers published since 2015. We also expanded our search to include other articles that were repeatedly cited by the articles collected in our initial queries.

We also searched in the field of Education, as methodological researchers in that area were among the earliest scholars who discussed the multi-score RD design in non-geographic settings. For this search, we started with the seminal methodological papers in Education by Reardon and Robinson (2012) and Wong et al. (2013), and collected papers that both cited this methodological papers and were cited by them. Finally, we also included some papers in environmental science, to illustrate uses of the design at the intersection between the natural and social sciences. Our review is not meant to provide an exhaustive list of the literature. This literature is vast, and cannot be fully captured by our approach since we excluded several journals and restricted the time period. Our goal is simply to provide an overview of recent empirical work using the BD design across the social sciences to gauge the most common strategies used for empirical analysis.

The results of our search are summarized in Table 1. Our search yielded 82 published papers. Of these, we found that the overwhelming majority (76 articles or approximately 93%) used a pooling approach, focusing on the average effect along the entire boundary and using the univariate closest distance to the boundary as the running variable. Fewer than 10 articles were explicit in accounting for the heterogeneity along the boundary; of these, only 5 reported effects for different points or segments along the boundary (Gonzalez, 2021; Grout et al., 2011; Keele and Titiunik, 2015; Snider and Williams, 2015; Velez and Newman, 2019). In sum, our review of the practical literature indicates that most applied researchers adopt a pooling approach, performing localization by grouping all observations sufficiently "close" the boundary and then estimating a single average treatment effect. These common practices motivate the taxonomy we used in the rest of this chapter.

3 Pooling-Based Methods

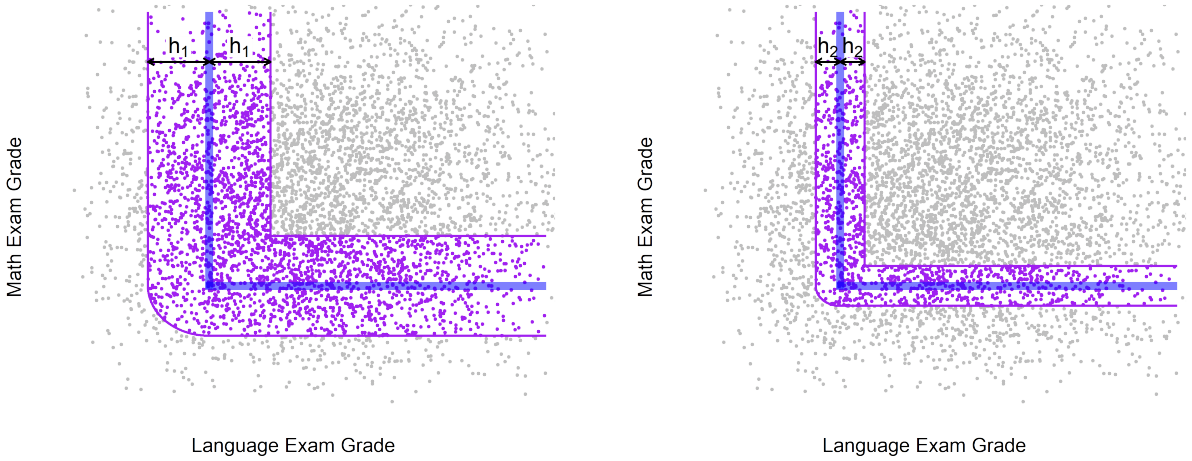
Given a sample of outcomes and bivariate scores, $(Y_1, \mathbf{X}_1), \ldots, (Y_n, \mathbf{X}_n)$, a distance function $\mathscr{d}(\cdot,\cdot)$ is used to measure the closeness between any two points in $\mathscr{X}$. The most common example is the Euclidean distance $\mathscr{d}(\mathbf{x}_1, \mathbf{x}_2) = \|\mathbf{x}_1 - \mathbf{x}_2\| = \sqrt{(x_{11} - x_{21})^2 + (x_{12} - x_{22})^2}$ for $\mathbf{x}_j = (x_{j1}, x_{j2})$, $j = 1, 2$, although in some applications this is not the most appropriate measure. For example, Ambrus et al. (2020) are interested in how close residences in a city are to each other in terms of how long it takes for a person to walk between them; this requires calculating a walking distance, which is different from the Euclidean distance which would not take into account, say, that the path must follow city blocks and cannot go through buildings. More broadly, in geographic settings it is common practice to employ specific geodetic distances (Banerjee, 2005).

Once a distance function $\mathscr{d}(\mathbf{X}_i, \mathbf{x})$ has been chosen, researchers who employ the pooling approach define the closest signed distance to the boundary,

$$D_i \equiv \big(\mathbb{1}(\mathbf{X}_i \in \mathscr{A}_1) - \mathbb{1}(\mathbf{X}_i \in \mathscr{A}_0)\big) \cdot \inf_{\mathbf{x} \in \mathscr{B}} \mathscr{d}(\mathbf{X}_i, \mathbf{x}),$$

for $i = 1, \ldots, n$. For a unit i, D_i measures how far i's location $\mathbf{X}_i$ is to the boundary point that is closest to $\mathbf{X}_i$, regardless of where that boundary point is located. Following the standard RD design logic, but now expanded to the BD design (a bivariate setting), localization occurs simultaneously along the entire assignment boundary because the univariate distance D_i to $\mathscr{B}$ is used to construct the region such that $|D_i| \leq h$, which covers the boundary $\mathscr{B}$, where h denotes a bandwidth parameter. This region, called a "tubular neighborhood" in mathematics (Federer, 2014), is a natural generalization of the standard localization approach used in univariate RD designs, where the region is simply $|X_i - c| \leq h$ when the score X_i is scalar and c is the univariate cutoff.

Figure 2 illustrates the idea graphically. In both cases, the tubular neighborhood represents



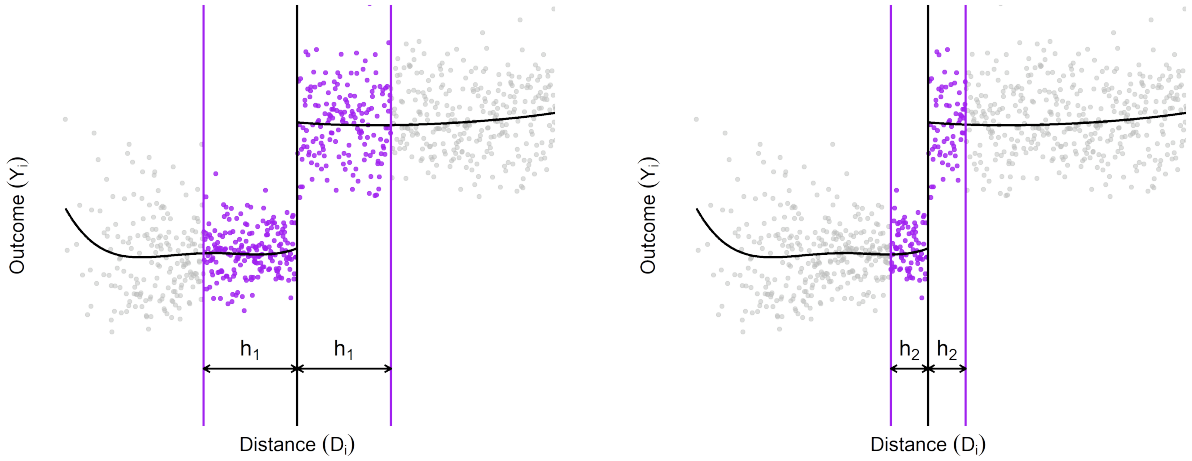
(a) Large Bandwidth (b) Small Bandwidth

Figure 2: Pooling-Based Method: Localization around Entire Assignment Boundary

the area that includes the observations used in the pooling-based analysis. Figure 2a shows a "large" bandwidth h_1, while Figure 2b shows a "small" bandwidth h_2. All units contribute simultaneously, regardless of their specific location. In particular, two units i and j with the same signed distance to the boundary, $D_i = D_j = d$, will be assigned to the same group (treatment or control, depending on the sign of d) and will be $|d|$ units away from the boundary, but could be arbitrarily far from each other.

Figure 3 presents the two resulting RD plots (Calonico et al., 2015) in the pooling approach, based on the outcome data Y_i and the univariate distance score D_i (the cutoff is $c = 0$ by construction of D_i). These plots include a fourth-order global polynomial fit (solid line) for visualization purposes only; those global approximations are not recommended for estimation of treatment effects because they tend to exhibit poor performance due to the Runge's phenomenon. Instead, as anticipated in Figure 2 and demonstrated in Figure 3, the idea is to localize around the "cutoff" (i.e., the one-dimensional boundary curve) to conduct estimation and inference.

In the pooling approach, the analysis is conducted using all the observations with score $\mathbf{X}_i$ within the tubular neighborhood determined by the bandwidth h, that is, for all units with $|D_i| \leq h$. However, for this subsample, the specific treatment effect estimation ap-



(a) Large Bandwidth (b) Small Bandwidth

Figure 3: Pooling-Based Method: RD Plots (with Global Quartic Polynomial Estimates)

proach can vary substantially. To describe and unify the most common approaches in the empirical literature illustrated in Table 1 as well as other approaches recommended from a methodological perspective, we define the following generic weighted least squares regression notation:

$$\texttt{reg} \quad Y_i \quad \texttt{on} \quad \mathbf{Z}_i^\top \quad [\, W_i \,] \qquad \Longleftrightarrow \qquad \min_{\boldsymbol{\beta}} \sum_{i=1}^{n} \left(Y_i - \mathbf{Z}_i^\top \boldsymbol{\beta}\right)^2 \cdot W_i$$

where Y_i is the outcome variable, $\mathbf{Z}_i$ is a (column) vector of independent variables, and W_i are weights. A localized regression analysis can be implemented by setting $W_i = \mathbb{1}(|D_i| \leq h)$ or, more generally, $W_i = K(D_i/h)$ for a kernel function $K(\cdot)$ determining the relative weight given to each observation. The bandwidth h determines the degree of localization, as illustrated in Figure 2. When $\mathbf{Z}_i$ contains polynomial expansions of D_i, the resulting estimation approach is often called nonparametric local polynomial regression but, as we will see shortly, empirical researchers often employ other variables in their local regression fit within the tubular determined by D_i.

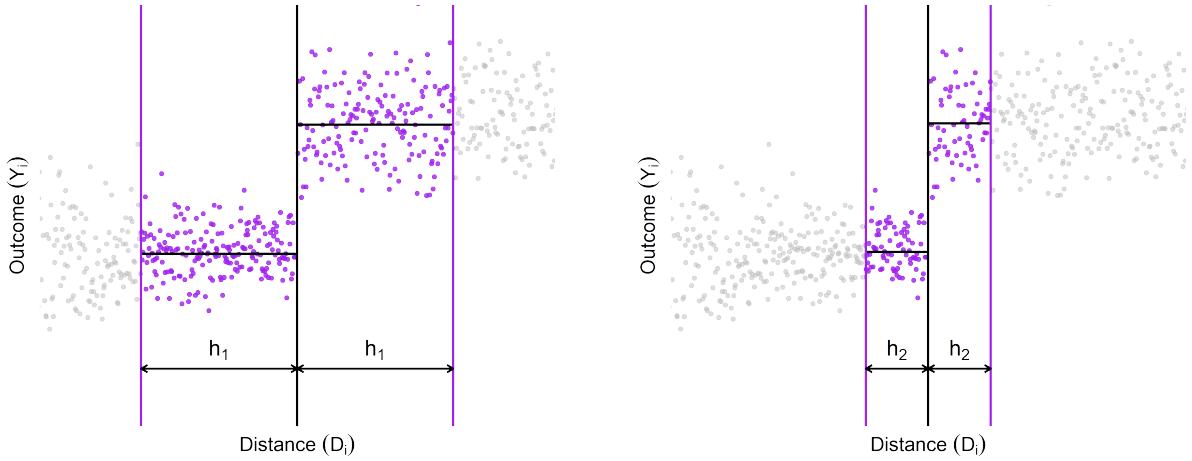
We also define the treatment assignment indicator

$$T_i = \mathbb{1}(D_i \geq 0) = \mathbb{1}(\mathbf{X}_i \in \mathscr{A}_1).$$

All empirical approaches naturally include T_i in $\mathbf{Z}_i$ to estimate treatment effects; the majority also include other variables such as D_i, $\mathbf{X}_i$, boundary-segment fixed effects (formally defined below), as well as certain transformations and interactions thereof. For simplicity, we omit discussing the inclusion of predetermined covariates, but we note that researchers often further augment the basic local regression specifications with pretreatment variables such as census characteristics, terrain ruggedness and elevation, time fixed effects, among other features. For a discussion of the role of preintervention covariates in the standard RD design see Calonico et al. (2019) for efficiency gains and Calonico et al. (2025) for treatment effect heterogeneity.

The pooling approach for analyzing BD designs begins by choosing the localization bandwidth h to determine the units with $|D_i| \leq h$ that will be used in the subsequent estimation. Until now, there were no formal methods for choosing the bandwidth h in an objective and data-driven way, so researchers often considered a range of "reasonable" values. For example, Black (1999) reported results in three different bandwidths, keeping observations within 0.35, 0.25, and 0.15 miles of the nearest point on the boundary. As in standard univariate RD designs, different choices of h trade off bias and variance in the estimation, with a larger h leading to more bias but less variance, and a smaller h having the opposite effect. Cattaneo et al. (2025c) recently established formal results characterizing this tradeoff and provided formal guidance for choosing the bandwidth when implementing pooling BD methods. We discuss these methods in Section 3.1.

Once the bandwidth h has been chosen, researchers pool all units with distance D_i within h and then compare those assigned to the treatment group ($D_i \geq 0$) to those assigned to the control group ($D_i < 0$). Because the bandwidth localizes the analysis to only a "small" region on either side of the boundary $\mathscr{B}$, the empirical analysis compares barely treated units to barely control units which, as in the standard univariate RD design, is the basis for the causal interpretation of the comparison, as all confounders are assumed to vary smoothly at the boundary while the treatment assignment changes abruptly from zero to one.



(a) Large Bandwidth (b) Small Bandwidth

Figure 4: Pooling-Based Method: Local-Constant Estimation ($p = 0$)

The most basic analysis proceeds by assuming that the units within the shrinking tubular of "width" $2h$ are as-if randomly assigned to treatment and control. This approach is akin to the "local randomization" framework in standard RD designs (Lee and Lemieux, 2010; Cattaneo et al., 2015, 2017). Thus, a natural treatment effect estimator is the difference-in-means of the outcome between treated and control units with distance to the boundary no larger than the bandwidth:

$$\texttt{reg} \quad Y_i \quad \texttt{on} \quad 1,\, T_i \quad [\, \mathbb{1}(|D_i| \leq h)\,], \tag{1}$$

where the coefficient on T_i is

$$\widehat{\tau}_{\texttt{DIM}} = \frac{1}{N_1(h)} \sum_{i=1}^{n} \mathbb{1}(0 \leq D_i \leq h) \cdot Y_i - \frac{1}{N_0(h)} \sum_{i=1}^{n} \mathbb{1}(-h \leq D_i < 0) \cdot Y_i,$$

with $N_0(h) = \sum_{i=1}^{n} \mathbb{1}(-h \leq D_i < 0)$ and $N_1(h) = \sum_{i=1}^{n} \mathbb{1}(0 \leq D_i \leq h)$. This specification can be interpreted as a local-constant polynomial fit on D_i; it does not include the score in the fit because under the assumption of local randomization this score is uncorrelated with the outcome.

For example, Black (1999, Table III) reported results using specification (1), calling it

a "rough, nonparametric estimate of the value of better schools" (p. 590). She also augmented the basic specification to include "segment fixed effects", that is, binary variables indicating whether an observation is closest to one of several segments partitioning the assignment boundary. To describe this approach formally, we let $\mathscr{B}_\ell$, $\ell = 1, \ldots, L$, be disjoint segments that partition the assignment boundary, that is, $\mathscr{B} = \sqcup_{1 \leq \ell \leq L} \mathscr{B}_\ell$. Then, $S_i = \arg\min_{1 \leq \ell \leq L} d(\mathbf{X}_i, \mathscr{B}_\ell)$ is the boundary-segment allocation random index for each unit $i = 1, \ldots, n$, where it is understood that when the arg min contains more than one element, the smallest item is chosen. The vector of segment fixed effects for observation i is $\boldsymbol{\iota}_L(S_i) = (\mathbb{1}(S_i = 1), \cdots, \mathbb{1}(S_i = L))$, which collects the L boundary segment binary indicators determined by i's location. The difference-in-means specification with segment fixed effects is

$$\texttt{reg} \quad Y_i \quad \texttt{on} \quad \boldsymbol{\iota}_L(S_i),\ T_i \quad [\ \mathbb{1}(|D_i| \leq h)\ ], \tag{2}$$

where the intercept is removed due to its perfect collinearity with the boundary-segment fixed effects. Setting $L = 1$ reduces $\boldsymbol{\iota}_L(S_i)$ to the common intercept in specification (1).

For example, Black (1999, Table II) used a modified version of specification (2) where she included the test scores of the school associated with observation i instead of the binary treatment assignment indicator T_i. The causal interpretation of this specification rests on the assumption that, after localizing the analysis to include only observations near the boundary of school attendance districts, the only factor that can explain a difference in average house prices between the treated and control areas is school quality, because school quality changes discontinuously at the boundary while all other determinants of house prices are equal on average. In this sense, using test scores instead of the treatment assignment indicator in specifications (1) and (2) is in the spirit of an instrumental variables strategy where T_i is used as an instrument for test scores (which proxy school quality), and the localization to a small bandwidth around the boundary ensures that units assigned to treatment and units

assigned to control are similar in all other characteristics that might affect house prices.

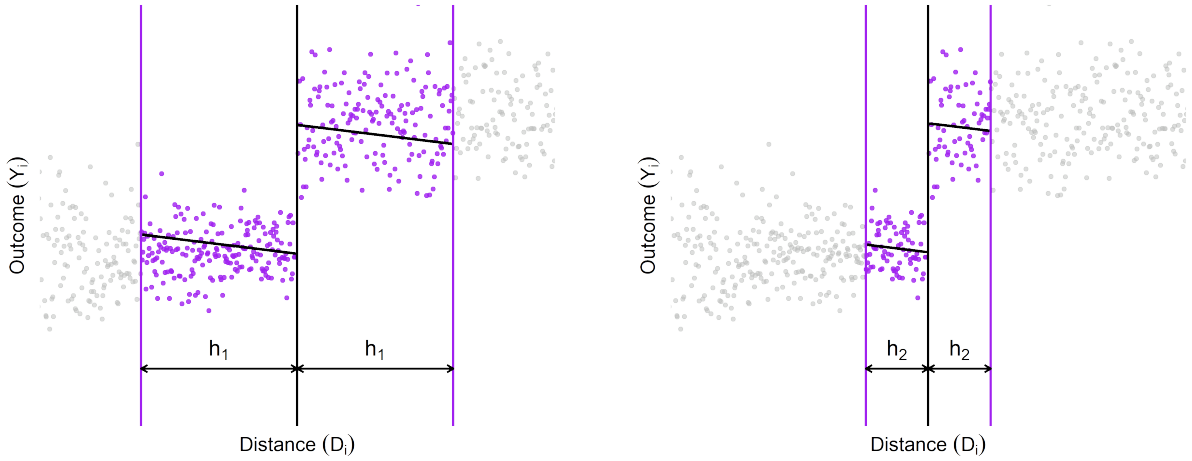
The seminal research design used by Black (1999) is prior to the development of modern methods for RD analysis, which occurred predominantly during the 2000s; see Cattaneo and Titiunik (2022) for an overview. In more recent practice, the simple specifications (1) and (2) are not commonly used because researchers tend to adopt a continuity-based approach rather than a local randomization approach for interpretation. Under the continuity-based approach, the local constant fit has poor control of the misspecification bias, and hence practitioners prefer more flexible local regression specifications.

Later empirical work using BD designs made more explicit the connection between geographic discontinuities in treatment assignment and the classical RD design, following more closely the methods of univariate continuity-based RD analysis, which emphasized including the RD score in the polynomial specification to allow for arbitrary dependence between the RD score and outcome even within the bandwidth. Although there is considerable diversity in the specifications used in more recent practice, most researchers that include the score adopted one of two main strategies: a polynomial fit of the outcome on the bivariate score $\mathbf{X}_i$, or a polynomial fit of the outcome on the univariate distance D_i.

In an influential contribution, Dell (2010) augmented specification (2) by including the bivariate score $\mathbf{X}_i$. She studied the long-run impact of the mita, a forced mining labor system instituted by the Spanish in Peru and Bolivia between 1573 and 1812, using the historical boundary that determined which communities were forced to send labor to define treated and control areas. She studied the effect of the mita on contemporary outcomes measured at the household level, including consumption and childhood stunting. Her main local regression specification was

$$\texttt{reg} \quad Y_i \quad \texttt{on} \quad \boldsymbol{\iota}_L(S_i),\ T_i,\ \mathbf{r}_p(\mathbf{X}_i) \quad [\ \mathbb{1}(|D_i| \leq h)\ ], \tag{3}$$

which includes the treatment indicator, boundary segment fixed effects, and a polynomial



(a) Large Bandwidth (b) Small Bandwidth

Figure 5: Pooling-Based Method: Local-Linear No-Interaction Estimation ($p = 1$)

expansion of the bivariate score $\mathbf{X}_i$ (with $p = 3$ in her case). The estimated coefficient of interest is always the coefficient that accompanies the treatment indicator T_i. In our notation, $\mathbf{r}_p(\mathbf{u}) = (u_1, u_2, u_1^2, u_1u_2, u_2^2, \ldots, u_1^p, u_1^{p-1}u_2, \ldots, u_1u_2^{p-1}, u_2^p)$ with $\mathbf{u} = (u_1, u_2)$. The inclusion of (a polynomial basis of) the bivariate location score $\mathbf{X}_i$ is meant to control for "the smooth effects of geographic location" Dell (2010, p. 1875). A notable characteristic of this specification is that it does not include interactions between the bivariate score and the treatment indicator, which imposes the restriction that the polynomials be identical in the treated and control areas. As for bandwidth selection, Dell (2010) considered 100km of the mita boundary first, but this choice is later reduced to 75km and 50km in subsequent analyses.

Specification (3) has become a common strategy to analyze BD designs. For example, Méndez and Van Patten (2022) use it to study the effect of private sector companies in the development of local amenities, using a land concession to the multinational United Fruit Company in Costa Rica; Paulsen et al. (2023) use it to estimate the effect of school funding on future earnings and political participation in New York, comparing areas that received higher funding to neighboring areas that did not; and De Kadt and Larreguy (2018) compare areas in South Africa where traditional leaders remain highly influential to areas where they

are not, and estimate the electoral impact of strong political ties between traditional leaders and the African National Congress (ANC). Other examples are listed in Table 1. All of these examples follow specification (2) by including a flexible polynomial expansion of latitude and longitude, as well as boundary-segment fixed effects, but without allowing for interactions with the treatment indicator T_i.

Another common approach in modern practice is to directly mimic the standard RD design. This approach treats the scalar (signed) distance to the closest point on the boundary for every unit, D_i, as the univariate score, and then employs canonical RD design methods based on local polynomial regression. The basic local regression specification in this context often takes the form

$$\texttt{reg} \quad Y_i \quad \texttt{on} \quad 1,\ T_i,\ \mathbf{r}_p(D_i) \qquad [\ \mathbb{1}(|D_i| \leq h)\ ], \tag{4}$$

with the polynomial expansion reducing to the simpler vector of the univariate distance $\mathbf{r}_p(D_i) = (D_i, D_i^2, \ldots, D_i^p)$. In practice, this specification is often implemented with $p = 1$ and augmented with boundary-segment fixed effects:

$$\texttt{reg} \quad Y_i \quad \texttt{on} \quad \boldsymbol{\iota}_L(S_i),\ T_i,\ \mathbf{r}_p(D_i) \qquad [\ \mathbb{1}(|D_i| \leq h)\ ]. \tag{5}$$

Furthermore, depending on whether standard software for RD design analysis is used, the specifications (4) and (5) may use a non-uniform weighting kernel (e.g., a triangular kernel that downweights observations as their distance from the boundary increases). In both cases, the treatment effect estimator is the coefficient estimate accompanying T_i. Figure 5 offers a graphical representation of this empirical strategy, abstracting from the boundary-segment fixed effects (e.g., setting $L = 1$ for simplicity).

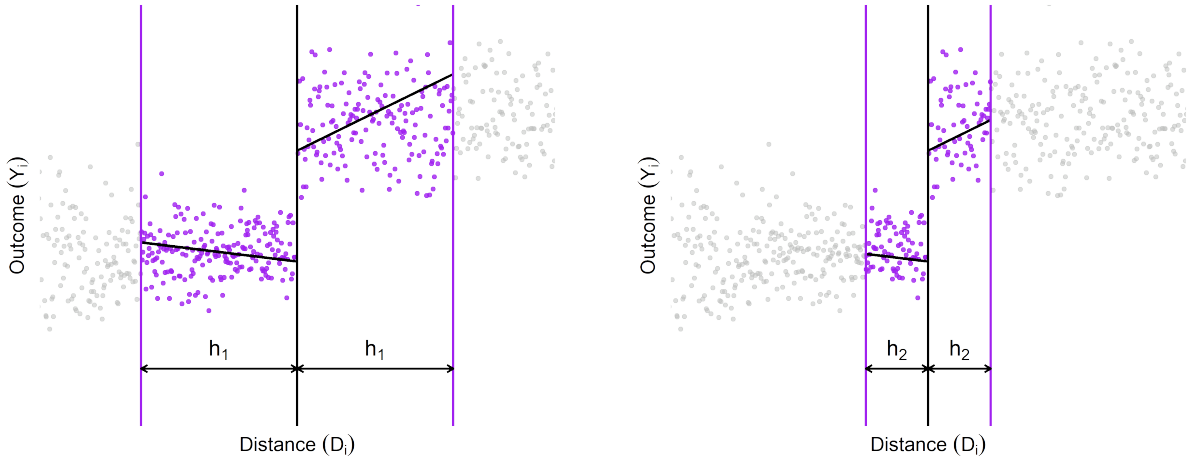
For example, Dehdari and Gehring (2022) use the border between the French regions of Alsace and Lorraine to study the effect of repressive nation-building policies on regional identity. The unit of analysis is the municipality, and the treatment is location in the

Lorraine area, where historical nation-building policies were more repressive than in Alsace. They find that regional identity is systematically stronger among citizens in the treated areas, as measured by the share of yes votes in a referendum. The distance D_i is calculated as the shortest distance from the municipality's centroid to the border. Their main analysis follows specification (5), which includes the treatment assignment indicator, a polynomial on the signed univariate distance D_i, and boundary segment fixed effects.

Although researchers following this specification aim to closely mimic the univariate RD design, their approach differs from best practices for univariate RD analysis in an important respect. In the standard RD design, the recommended specification is a polynomial fit of the outcome on an intercept, the treatment assignment indicator, a polynomial expansion of the score, and the interaction between this expansion and the treatment assignment indicator. The latter interaction is crucial to allow the polynomial fit for control units to be different from the polynomial fit for treated units. None of the specifications discussed so far allow for this flexibility; instead, by excluding the interaction, only the intercepts are allowed to be different among treated and control units, but not the slopes, thereby introducing potential misspecification bias whenever the treatment affects the average response functions near the boundary differently for treatment and control units.

The omission of the interaction is common in empirical practice, but we found some papers where it is included. For example, Ito and Zhang (2020) use the closest distance to the boundary in a univariate analysis to study households' willingness to pay for clean air in China, using the discontinuous change in pollution that occurs at the Huai River due to a long-standing heating policy that affects areas north of the river but not areas south of it. In their first-stage analysis, the authors study the effect of the heating policy on air pollution using the following specification (with $p = 1$)

$$\texttt{reg} \quad Y_i \quad \texttt{on} \quad \boldsymbol{\iota}_L(S_i),\ T_i,\ \mathbf{r}_p(D_i),\ T_i \cdot \mathbf{r}_p(D_i) \qquad [\ \mathbb{1}(|D_i| \leq h)\ ], \tag{6}$$



(a) Large Bandwidth (b) Small Bandwidth

Figure 6: Pooling-Based Method: Local-Linear Estimation ($p = 1$)

where the estimated coefficient of interest continues to be the one accompanying T_i.

Abstracting from the inclusion of boundary-segment fixed effects (or otherwise setting them to some specific value), Figure 6 illustrates the importance of including the interaction term $T_i \cdot \mathbf{r}_p(D_i)$ in the treatment effect estimation, as in specification (6). Comparing Figures 5 and 6, it is visually striking and methodologically clear that the two approaches can lead to different local approximations, with pooling approaches based on specifications (1) through (5) potentially exhibiting a larger bias than approaches based on specification (6). Following the same logic, specification (3) can also be augmented to include the interaction term between the treatment assignment and the polynomial expansion of the bivariate score $\mathbf{X}_i$:

$$\texttt{reg} \quad Y_i \quad \texttt{on} \quad \boldsymbol{\iota}_L(S_i),\ T_i,\ \mathbf{r}_p(\mathbf{X}_i),\ \ T_i \cdot \mathbf{r}_p(\mathbf{X}_i) \quad [\ \mathbb{1}(|D_i| \leq h)\ ], \tag{7}$$

where the estimated treatment effect coefficient accompanying T_i enjoys better bias properties due to the added flexibility in the local regression specification. This approach is not common in the empirical literature—exceptions are Gonzalez (2021), Egger and Wamser (2015), and Jones et al. (2022b).

In some applications, researchers estimate treatment effects in BD designs by considering

two separate univariate RD designs, that is, keeping one dimension fixed and studying the other dimension of the bivariate score. For example, Londoño-Vélez et al. (2020) investigate the effects of *Ser Pilo Paga* (SPP), a government subsidy in Colombia that provided tuition support for post-secondary students to attend a government-certified higher education institution. Eligibility was based on both merit and economic need: in order to qualify for the program, students had to obtain a high grade in Colombia's national standardized high school exit exam, *SABER 11*, and they also had to come from economically disadvantaged families, as measured by the survey-based wealth index *SISBEN*. The resulting BD design looks exactly as in Figure 1b with $\mathbf{X}_i = (X_{1i}, X_{2i}) = (SABER11_i, SISBEN_i)$. The authors analyzed it by considering each cutoff separately and then pooling all observations in the other dimension, leading to two univariate RD designs: one with score $SISBEN_i$ in the sample $SABER11_i \geq 0$, and the other with score $SABER11_i$ in the sample $SISBEN_i \geq 0$. This approach is equivalent to splitting the analysis along each of the two linear segments of $\mathscr{B}$ and then employing a pooling approach. This idea leads to the following local regression specification:

$$\texttt{reg} \quad Y_i \quad \texttt{on} \quad T_i \cdot \boldsymbol{\iota}_L(S_i),\ \boldsymbol{\iota}_L(S_i) \otimes \mathbf{r}_p(D_i),\ T_i \cdot \boldsymbol{\iota}_L(S_i) \otimes \mathbf{r}_p(D_i) \qquad [\ \mathbb{1}(|D_i| \leq h)\ ], \tag{8}$$

where the indicator variable T_i is removed due to the perfect collinearity with $T_i \cdot \boldsymbol{\iota}_L(S_i)$, since $\boldsymbol{\iota}_L(S_i)$ contains indicator variables capturing a partitioning of $\mathscr{B}$, and $\otimes$ denotes the Kronecker product. The treatment effect of interest is now the collection of coefficients accompanying $T_i \cdot \boldsymbol{\iota}_L(S_i)$, one for each segment. Specification (8) is roughly equivalent to employing specification (6) L times, in each case only using the data corresponding to one of L segments. (The only minor difference is whether observations are assigned to only one segment, or reused for different segment-specific treatment effect estimation implementation.) For instance, Ou (2010), Londoño-Vélez et al. (2020) and Salti et al. (2022) employ this boundary-segment specific pooling-based approach with $L = 2$, analyzing each of the two

linear segments as in Figure 1b. Another empirical example employing this approach is .

Finally, some researchers have also included both the univariate distance D_i and the bivariate score $\mathbf{X}_i$ in their local regression specifications—see, for example, Ehrlich and Seidel (2018) and Dehdari and Gehring (2022). In addition, it is possible to conceptualize more flexible specifications by including other interaction terms between T_i, $\mathbf{r}_p(\mathbf{X}_i)$, and $\boldsymbol{\iota}_L(S_i)$ analogous to specification (8). These approaches are uncommon in practice according to our literature review (Table 1), so we do not discuss them further.

3.1 Methodological Results

While there has been a proliferation of different empirical strategies relaying on the pooling of observations closest to the treatment assignment boundary, little is known about their properties and relative merits for the analysis of BD designs. Cattaneo et al. (2025c) recently investigated the main empirical approaches employing pooling-based methods, offering formal identification, estimation, and inference results. Their methodological results rely on techniques from geometric measure theory (Federer, 2014), which are used to characterize how the different treatment effect estimates emerging from specifications (1) through (6) behave in large samples, that is, when $h \to 0$ as $n \to 0$. These asymptotic approximations form the basis for the continuity-based approach in standard univariate RD designs (Hahn et al., 2001), and are nowadays the most common way of characterizing the properties of causal treatment effect estimators in all RD settings. This section summarizes some of the findings in Cattaneo et al. (2025c).

The main challenge when studying pooling-based methods for BD designs is that the localization is done using a sequence of shrinking tubular neighborhoods covering the one-dimensional submanifold $\mathscr{B}$, as illustrated in Figure 2. An important implication is that the geometry of the treatment assignment boundary and the specific distance function used can substantially affect statistical properties. In this section, we omit most technicalities, and focus instead on the main ideas and conclusions to aid empirical researchers interpret and

implement pooling-based methods.

Because the treatment assignment boundary $\mathscr{B}$ could be highly irregular, particularly in Geographic RD designs, it is necessary to restrict its geometry to establish formal properties. A minimal assumption commonly used in mathematics is to require $\mathscr{B}$ to be a rectifiable curve, that is, to assume that it has finite length. More formally, $\mathscr{B}$ is a rectifiable curve if and only if it can be reparameterized by a Lipschitz continuous function. This restriction allows for formally defining and computing integrals of functions along the one-dimensional domain $\mathscr{B}$. More importantly, under additional technical conditions, regularity on the geometry of $\mathscr{B}$ allows for those integrals to be computed over shrinking tubular neighborhoods and for limits to be well-defined, which is at the core of the results in Cattaneo et al. (2025c). Under regularity conditions, the authors establish the following result:

$$\lim_{h\downarrow 0}\int_{\mathscr{T}(h)}\frac{1}{h}g\Big(\frac{\mathscr{d}(\mathbf{x},\mathscr{B})}{h}\Big)m(\mathbf{x})\mathrm{d}\mathbf{x}=\mathsf{c}_{\mathscr{B}}\cdot\int_0^1 g(s)ds\cdot\int_{\mathscr{B}}\frac{m(\mathbf{x})}{\mathrm{J}_2\mathscr{d}(\mathbf{x},\mathscr{B})}\mathrm{d}\mathfrak{H}(\mathbf{x}), \tag{9}$$

where $\mathscr{T}(h)=\big\{\mathbf{x}\in\mathscr{X}:\mathscr{d}(\mathbf{x},\mathscr{B})\leq h\big\}$, for each $h\geq 0$, is a tubular covering $\mathscr{B}$,

$$\mathscr{d}(\mathbf{x},\mathscr{B})=\inf_{\mathbf{b}\in\mathscr{B}}\mathscr{d}(\mathbf{x},\mathbf{b})$$

and $\mathrm{J}_2\mathscr{d}(\mathbf{x},\mathscr{B})$ denotes its Jacobian, and $\mathfrak{H}(\mathbf{x})$ denotes the 1-dimensional Hausdorff measure. The constant $\mathsf{c}_{\mathscr{B}}$ can also be characterized, but it is not important for the upcoming discussion; for the case of pooling-based methods it can be set $\mathsf{c}_{\mathscr{B}}=1$. Note that $D_i=\big(\mathbb{1}(\mathbf{X}_i\in\mathscr{A}_1)-\mathbb{1}(\mathbf{X}_i\in\mathscr{A}_0)\big)\cdot\mathscr{d}(\mathbf{X}_i,\mathscr{B})$.

The integral $\int_{\mathscr{B}}\frac{m(\mathbf{x})}{\mathrm{J}_2\mathscr{d}(\mathbf{x},\mathscr{B})}\mathrm{d}\mathfrak{H}(\mathbf{x})$ is a natural generalization of the standard line integral, where $\mathrm{d}\mathfrak{H}$ serves as the rigorous "length element", enabling a robust framework for integration over complex one-dimensional domains in the plane. In particular, if $\mathscr{B}$ is piecewise linear (as in Figures 1b and 2), the integral can be reduced to the sum of integrals over the linear segments using a natural smooth curve parametrization and standard Riemann integration

(when $m(\mathbf{x})$ is Riemann integrable). For notational simplicity, we write

$$\int_{\mathscr{B}} m(\mathbf{b})\mathrm{d}\mathbf{x} = \int_{\mathscr{B}} \frac{m(\mathbf{x})}{\mathrm{J}_2 \mathscr{d}(\mathbf{x},\mathscr{B})}\mathrm{d}\mathfrak{H}(\mathbf{x}),$$

but with the understanding that the left-hand-side integral is just notation for the rigorously defined right-hand-side integral.

The limiting integral in (9) can be used to give precise meaning to the probability limit of the treatment effect estimates from any of the pooling specifications discussed previously. However, we need to introduce standard potential outcomes notation (see, e.g., Hernán and Robins, 2020, for a review) to give a causal interpretation to the probability limit of those parameters. For $i = 1, 2, \ldots, n$, let $Y_i(0)$ and $Y_i(1)$ denote the potential outcomes for unit i under control and treatment assignment, respectively, and hence the observed outcome is

$$Y_i = (1 - T_i) \cdot Y_i(0) + T_i \cdot Y_i(1).$$

To streamline the presentation, we consider only the simplest specification (1), leading to the local difference-in-means estimator $\widehat{\tau}_{\texttt{DIM}}$, to illustrate the methodological findings in Cattaneo et al. (2025c). The other more flexible local regression specifications will enjoy similar properties.

The probability limit of $\widehat{\tau}_{\texttt{DIM}}$ can be naturally characterized as a ratio of integrals along the boundary. As $n \to \infty$ and leveraging the logic of the law of large numbers, $\frac{1}{nh}N_1(h) = \frac{1}{nh}\sum_{i=1}^{n} \mathbb{1}(0 \leq D_i \leq h)/(nh)$ will be close in probability to $\frac{1}{h}\mathbb{E}[\mathbb{1}(0 \leq D_i \leq h)]$, and

$$\frac{1}{h}\mathbb{E}[\mathbb{1}(0 \leq D_i \leq h)] = \int_{\mathscr{T}(h)} \frac{1}{h} g\Big(\frac{\mathscr{d}(\mathbf{x},\mathscr{B})}{h}\Big) f(\mathbf{x})\mathrm{d}\mathbf{x} \to \int_0^1 g(s)ds \cdot \int_{\mathscr{B}} f(\mathbf{x})\mathrm{d}\mathbf{x},$$

where the limit is taken as $h \to 0$, $f(\mathbf{x})$ denotes the Lebesgue density of $\mathbf{X}_i$, and the conclusion follows by (9) upon setting $g(u) = \mathbb{1}(0 \leq u < 1)$ and $m(\mathbf{x}) = \mathbb{1}(\mathbf{x} \in \mathscr{A}_1)f(\mathbf{x})$. While $g(u)$ is redundant in this case, and $\int_0^1 g(s)ds = 1$, we nonetheless make it explicit to highlight that

a term will appear when a non-uniform kernel is used. Similarly, $N_0(h)/(nh) \to_{\mathbb{P}} \int_{\mathscr{B}} f(\mathbf{x})\mathrm{d}\mathbf{x}$. Furthermore, by the same logic, the two (properly rescaled) numerators of $\widehat{\tau}_{\texttt{DIM}}$ will be close in probability to their expectations, where

$$\begin{aligned}\frac{1}{h}\mathbb{E}[\mathbb{1}(0 \leq D_i \leq h) \cdot Y_i] &= \frac{1}{h}\mathbb{E}\Big[\mathbb{1}(0 \leq D_i \leq h) \cdot \mathbb{E}[Y_i(1)|D_i]\Big] \\ &= \int_{\mathscr{T}(h)} \frac{1}{h} g\Big(\frac{\mathscr{d}(\mathbf{x},\mathscr{B})}{h}\Big)\mathbb{E}[Y_i(1)|\mathscr{d}(\mathbf{X}_i,\mathscr{B}) = \mathscr{d}(\mathbf{x},\mathscr{B})]f(\mathbf{x})\mathrm{d}\mathbf{x} \\ &\to \int_{\mathscr{B}} \mathbb{E}[Y_i(1)|\mathbf{X}_i = \mathbf{x}]f(\mathbf{x})\mathrm{d}\mathbf{x},\end{aligned}$$

and, analogously,

$$\frac{1}{h}\mathbb{E}[\mathbb{1}(-h \leq D_i \leq 0)Y_i] \to \int_{\mathscr{B}} \mathbb{E}[Y_i(0)|\mathbf{X}_i = \mathbf{x}]f(\mathbf{x})\mathrm{d}\mathbf{x}.$$

The above calculations require precise regularity conditions and technical work: restrictions on the geometry of $\mathscr{B}$ and the distance function $\mathscr{d}(\cdot)$ are essential to ensure that $\mathbb{E}[Y_i(t)|\mathscr{d}(\mathbf{X}_i,\mathscr{B}) = \mathscr{d}(\mathbf{x},\mathscr{B})]$ is close to $\mathbb{E}[Y_i(t)|\mathbf{X}_i = \mathbf{x}]$ for $t = 0, 1$, as well as to guarantee that the limits are valid.

Putting the above calculations together,

$$\widehat{\tau}_{\texttt{DIM}} \to_{\mathbb{P}} \tau = \frac{\int_{\mathscr{B}} \tau(\mathbf{x})f(\mathbf{x})\mathrm{d}\mathbf{x}}{\int_{\mathscr{B}} f(\mathbf{x})\mathrm{d}\mathbf{x}}, \qquad \tau(\mathbf{x}) = \mathbb{E}[Y_i(1) - Y_i(0)|\mathbf{X}_i = \mathbf{x}].$$

The parameter τ is called the Boundary Average Treatment Effect (BATE) by Cattaneo et al. (2025c). This parameter was heuristically introduced in the education literature by Wong et al. (2013), who expressed it as

$$\tau = \int_{\mathscr{B}} \tau(\mathbf{x})f(\mathbf{x}|\mathbf{X}_i \in \mathscr{B})\mathrm{d}\mathbf{x} = \mathbb{E}[Y_i(1) - Y_i(0)|\mathbf{X}_i = \mathbf{x}, \mathbf{X}_i \in \mathscr{B}],$$

using the notation $f(\mathbf{x}|\mathbf{X}_i \in \mathscr{B}) = \frac{f(\mathbf{x})}{\int_{\mathscr{B}} f(\mathbf{x})\mathrm{d}\mathbf{x}}$. Around the same time, Keele and Titiunik (2015) also discussed the BATE parameter in the context of Geographic RD designs.

The parameter τ captures a density-weighted average of average treatment effects at each point along the boundary, and thus aggregates the potentially heterogeneous treatment effects captured by $(\tau(\mathbf{x}) : \mathbf{x} \in \mathscr{B})$. Cattaneo et al. (2025b) call $\tau(\mathbf{x})$ the Boundary Average Treatment Effect Curve. As discussed in Section 4, this functional causal parameter can be seen as a building block for several causal parameters of interest in BD designs.

Going beyond identification, Cattaneo et al. (2025c) also study estimation accuracy and uncertainty quantification for pooling-based methods. For point estimation, the authors shows that

$$|\widehat{\tau}_{\texttt{DIM}} - \tau|^2 \approx_{\mathbb{P}} \mathsf{B}_n^2 + \frac{\mathsf{V}}{nh},$$

when $h \to 0$ as $n \to \infty$, and where $\approx_{\mathbb{P}}$ denotes up to higher order terms in probability, B_n denotes a leading bias term, and V denotes the asymptotic variance that is bounded and bounded away from zero. Characterizing the precise order of the bias is difficult without additional restrictions on the treatment assignment boundary $\mathscr{B}$: the authors show that $\mathsf{B}_n = O(h)$ in general. Furthermore, as in the case of standard univariate RD designs, and under additional regularity conditions, the order of the bias may be improved when $\mathbf{r}_p(D_i)$ and $T_i \cdot \mathbf{r}_p(D_i)$ are included in the local regression specification, as in (6), leading to $\mathsf{B}_n = O(h^{p+1})$. Another interesting feature is that the rate of decay of the variance component is the same as in one-dimensional nonparametric estimation: this improvement in convergence rate comes from the aggregation along the boundary of the bivariate nonparametric function $\tau(\mathbf{x})$. See Cattaneo et al. (2025b) and Chen and Gao (2025) for more discussion.

For bandwidth selection, under regularity conditions on $\mathscr{B}$ and $\mathscr{d}(\cdot)$, the bias B_n may admit a valid expansion as in the standard RD design literature. Therefore, optimal bandwidth selection procedures based on mean square error (MSE) or coverage error minimization may be used for implementation of pooling methods. See Calonico et al. (2020) for a review.

For distribution theory and inference, consider the usual test statistic

$$\widehat{\mathrm{T}} = \frac{\widehat{\tau}_{\mathtt{DIM}} - \tau}{\sqrt{\widehat{\mathbb{V}}[\widehat{\tau}_{\mathtt{DIM}}]}},$$

where $\widehat{\mathbb{V}}[\widehat{\tau}_{\mathtt{DIM}}]$ denotes any of the usual variance estimators from (local, weighted) least squares regression methods. Under regularity conditions, and if $nh \to \infty$ and $nh\mathsf{B}_n^2 \to 0$, $\mathbb{P}[\widehat{\mathrm{T}} \leq u] \to \Phi(u)$, where $\Phi(u)$ denotes the standard Gaussian cumulative distribution function. Importantly, as discussed by Calonico et al. (2014) in the context of standard univariate RD designs, the usual confidence intervals

$$\widehat{\mathrm{I}}(\alpha) = \left[\widehat{\tau}_{\mathtt{DIM}} - \mathfrak{c}_\alpha \cdot \sqrt{\widehat{\mathbb{V}}[\widehat{\tau}_{\mathtt{DIM}}]} \ , \ \widehat{\tau}_{\mathtt{DIM}} + \mathfrak{c}_\alpha \cdot \sqrt{\widehat{\mathbb{V}}[\widehat{\tau}_{\mathtt{DIM}}]} \right]$$

with $\mathfrak{c}_\alpha = \Phi^{-1}(1-\alpha/2)$, will not be valid whenever there is local misspecification. More precisely, the validity of the usual confidence intervals $\widehat{\mathrm{I}}(\alpha)$ require the small bias condition $nh\mathsf{B}_n^2 \to 0$, which is not satisfied when the bandwidth is chosen to be MSE-optimal (or is otherwise "large"). Therefore, valid uncertainty quantification depends on the choice of bandwidth and regression specification, because the bias B_n would be of order h^p, under regularity conditions on $\mathscr{B}$, $\mathscr{d}(\cdot)$, and the underlying conditional expectation functions.

If the MSE-optimal bandwidth is used, which in the context of pooling-based methods would depend on the order of the bias B_n, then the small bias condition will not be satisfied, and thus the resulting inference procedures will over-reject the null hypothesis, sometimes substantially. Following Calonico et al. (2014), a simple and practical solution is to employ robust bias correction. The two-step procedure is simple and intuitive: first, the pooled treatment effect is estimated using a choice of polynomial approximation p and its associated MSE-optimal bandwidth; second, confidence intervals and hypothesis tests are constructed by estimating the pooled treatment effect and variance estimator using the same bandwidth choice but with a polynomial of larger order q. In practice, the most common choice is $p = 1$

for MSE-optimal treatment effect estimation, and $q = 2$ for robust bias-corrected uncertainty quantification. This inference approach has several theoretical advantages (Calonico et al., 2018, 2022), and has been validated empirically (Hyytinen et al., 2018; De Magalhães et al., 2025). Standard RD software for estimation and inference in univariate score settings (`rdrobust`, `rdhte`) available at `https://rdpackages.github.io/` can be used directly, provided the regularity conditions are satisfied.

4 Heterogeneity and Aggregation Along the Boundary

Our review of the literature indicates that most empirical researchers using BD designs employ pooling-based methods, thereby focusing on (density-weighted) average treatment effects along the entire (or a few segments of the) boundary $\mathscr{B}$. In most implementations, the estimation begins by calculating the distance to the nearest boundary point for every observation, keeping only observations for which this distance is less than a bandwidth, and then continues by pooling all observations in a single local regression analysis, mimicking standard univariate RD methods.

Although the focus on a single scalar parameter is a convenient way of summarizing the information provided by the BD design, this approach fails to exploit its full richness. In this section, we discuss a more general approach for the analysis of BD designs that starts by studying the average treatment effect at each point on the boundary, and then uses this functional causal parameter as the building block for aggregation along the boundary, hence recovering other causal treatment effects. This section summarizes the recent results on identification, estimation and inference discussed in Cattaneo et al. (2025a,b).

The starting point is to consider the Boundary Average Treatment Effect Curve (BATEC):

$$\tau(\mathbf{x}) = \mathbb{E}[Y_i(1) - Y_i(0)|\mathbf{X}_i = \mathbf{x}], \qquad \mathbf{x} \in \mathscr{B},$$

which was already implicitly introduced as part of the probability limit emerging from the

pooling approaches. However, the functional causal parameter $\tau(\mathbf{x})$ is of independent interest because it captures the average treatment effects at each point on the boundary. Despite capturing potentially interesting causal evidence of heterogeneous treatment effects along the assignment boundary, only a handful of empirical papers have investigated the BATEC or variations thereof. Exceptions include Keele et al. (2017), Velez and Newman (2019), and Gonzalez (2021).

Another feature of the BATEC is that it provides the key building block for constructing other causal parameters of interest by aggregation along the boundary. For example, Cattaneo et al. (2025b) discuss the following two parameters.

- Weighted Boundary Average Treatment Effect (WBATE):

$$\tau_{\texttt{WBATE}} = \frac{\int_{\mathscr{B}} \tau(\mathbf{x}) w(\mathbf{x}) \mathrm{d}\mathfrak{H}(\mathbf{x})}{\int_{\mathscr{B}} w(\mathbf{x}) \mathrm{d}\mathfrak{H}(\mathbf{x})},$$

  where $w(\mathbf{x})$ denotes a user-chosen weighing scheme.

- Largest Boundary Average Treatment Effect (LBATE):

$$\tau_{\texttt{LBATE}} = \sup_{\mathbf{x} \in \mathscr{B}} \tau(\mathbf{x}).$$

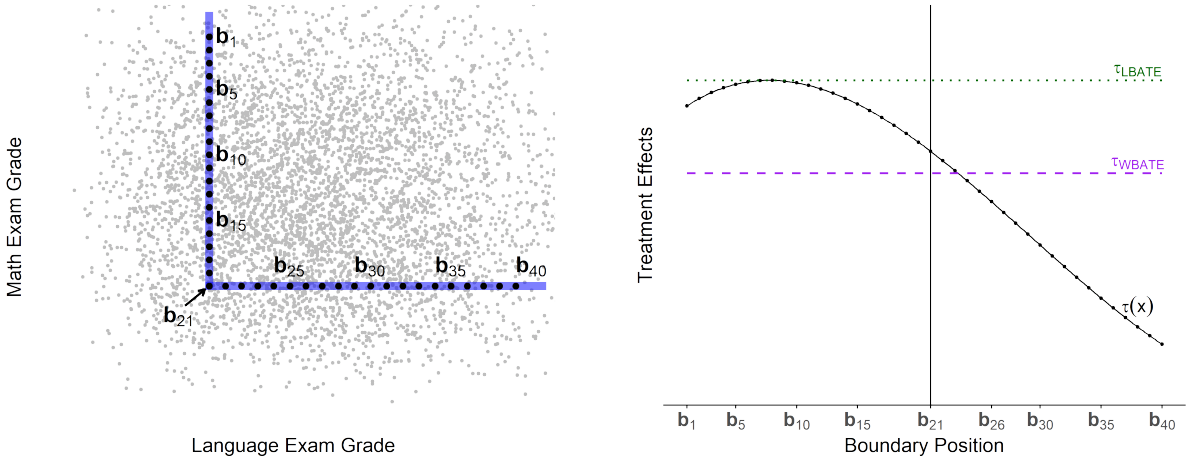
The WBATE parameter represents a weighted average of the (potentially heterogeneous) treatment effects $\tau(\mathbf{x})$ at each boundary point. In particular, setting $w(\mathbf{x}) = f(\mathbf{x})$ recovers the BATE parameter τ. In practice, ther weighting schemes may also be of interest. See Reardon and Robinson (2012) and Wong et al. (2013) for early methodological discussions, and Rischard et al. (2021) for other examples considered in the context of Bayesian methods for BD designs. The LBATE parameter captures the "best" causal treatment effect along the assignment boundary, and thus can be useful for evaluating or developing targeted policies. See, for example, Andrews et al. (2024).

Figure 7 gives a graphical representation of the BATEC, WBATE, and LBATE in the con-

text of the same stylized BD design used in Section 3. Because $\tau(\mathbf{x})$ is a function, estimation and inference proceeds by discretizing the boundary to construct local regression estimates for each point $\mathbf{x} \in \mathscr{B}$ on the grid of evaluation points chosen. Figure 7a demonstrates the idea with 40 evenly-spaced grid points, denoted by $\mathbf{b}_1, \ldots, \mathbf{b}_{40}$. As discussed in the upcoming subsections, distance-based and location-based methods employ these points to conduct estimation and inference, both pointwise and uniformly along the assignment boundary.

Figure 7b illustrates graphically the three target causal parameters $(\tau(\mathbf{x}) : \mathbf{x} \in \mathscr{B})$, $\tau_{\texttt{WBATE}}$, and $\tau_{\texttt{LBATE}}$. The functional causal parameter $\tau(\mathbf{x})$ indicates the presence of heterogeneity along a region of the boundary, while $\tau_{\texttt{WBATE}}$ and $\tau_{\texttt{LBATE}}$ give two distinct notions of aggregation of those heterogeneous treatment effects. This stylized example is motivated by the SPP study by Londoño-Vélez et al. (2020) introduced above, where $\mathbf{X}_i = (X_{1i}, X_{2i}) = (SABER11_i, SISBEN_i)$. Re-analyzing the SPP application, Cattaneo et al. (2025b) found little evidence of heterogeneous treatment effects along the X_{1i} dimension, but declining treatment effects along the X_{2i} dimension. This empirical finding indicates that the average treatment effects along the boundary are roughly similar for all students with low academic performance ($SABER11_i = 0$) and regardless of their wealth level ($SISBEN_i \geq 0$), while heterogeneity in treatment effects is present for the wealthier eligible students ($SISBEN_i = 0$) as their academic performance increases ($SABER11_i \geq 0$).

The discussion so far has focused on causal parameters defined along the entire $\mathscr{B}$, but in some applications researchers have considered different segments of $\mathscr{B}$. Specification (8), and its motivation from the implementation used by Londoño-Vélez et al. (2020) and others, gives one example. Employing the BATEC parameter as a building block, it is natural to define causal parameters as integrals or suprema of $\tau(\mathbf{x})$ over a specific region of $\mathscr{B}$. All the methods developed in Cattaneo et al. (2025a,b) immediately apply to those cases, offering estimation and inference methods for segment-specific causal parameters based on the core BATEC parameter.



(a) Boundary Discretization (b) Causal Parameters

Figure 7: Heterogeneity and Aggregation along the Boundary

4.1 Distance-Based Methods

The distance $D_i = \mathscr{d}(\mathbf{X}_i, \mathscr{B})$ used by all pooling approaches is based on the distance of each observation $\mathbf{X}_i$ to its nearest point $\mathbf{x} \in \mathscr{B}$ on the boundary, denoted by $\mathscr{d}(\mathbf{X}_i, \mathbf{x})$. Researchers sometimes first construct the signed distance to a pre-specified point on the boundary

$$D_i(\mathbf{x}) = \big(\mathbb{1}(\mathbf{X}_i \in \mathscr{A}_1) - \mathbb{1}(\mathbf{X}_i \in \mathscr{A}_0)\big) \cdot \mathscr{d}(\mathbf{X}_i, \mathbf{x}), \qquad \mathbf{x} \in \mathscr{B},$$

for each unit $i = 1, \ldots, n$, and then use this distance to construct a tubular covering $\mathscr{B}$ to implement pooling methods. One example of this approach is Ehrlich and Seidel (2018). While the resulting tubular will often be geometrically different from $\mathscr{T}(h) = \{\mathbf{x} \in \mathscr{X} : \mathscr{d}(\mathbf{x}, \mathscr{B}) \leq h\}$, all the ideas and results discussed in Section 3 continue to apply.

However, the intermediate boundary-point-specific signed distance $(D_i(\mathbf{x}) : \mathbf{x} \in \mathscr{B})$ for each unit can also be used directly to estimate and conduct inference for the BATEC. The idea is to view the outcomes and the boundary-point-specific scalar distance variables, $(Y_1, D_1(\mathbf{x})), \ldots, (Y_n, D_n(\mathbf{x}))$, as a one dimensional RD design for each point $\mathbf{x} \in \mathscr{B}$. For example, this approach was used by Keele and Titiunik (2015) and Velez and Newman

(2019) to study the effect of TV exposure on voter turnout, using reception and media market boundaries that determine whether citizens are exposed to specific television programming. In these examples, each unit's score $\mathbf{X}_i$ is a latitude-longitude pair that determines their place of residence, the outcome Y_i is whether the individual turned out to vote, and the boundary $\mathscr{B}$ is the media market or reception boundary that separates the treated from the control region. In Keele and Titiunik (2015), the treatment is exposure to presidential candidate television ads during the presidential election, and in Velez and Newman (2019) the treatment is exposure to Spanish-language television. In both cases, researchers first selected a grid $\mathbf{b}_1, \mathbf{b}_2, \dots \mathbf{b}_J$ on the boundary (e.g. Figure 7a), and then used the local regression specification:

$$\texttt{reg} \quad Y_i \quad \texttt{on} \quad 1,\ T_i,\ \mathbf{r}_p(D_i(\mathbf{b}_j)),\ T_i \cdot \mathbf{r}_p(D_i(\mathbf{b}_j)) \qquad [\ \mathbb{1}(|D_i(\mathbf{b}_j)| \leq h)\ ], \tag{10}$$

for each $j = 1, \dots, J$. The key difference is that estimation is conducted with localization for each evaluation point $\mathbf{b}_j$ on the boundary, as opposed to localization to the entire boundary as in specifications (1) through (8). Therefore, the estimated coefficient accompanying T_i gives a point estimate of $\tau(\mathbf{b}_j)$ for each $j = 1, \dots, J$

Cattaneo et al. (2025a) study the large sample properties of the distance-based estimation procedure for BATEC obtained from specification (10). They give necessary and/or sufficient conditions for identification, estimation, and inference in large samples, both pointwise and uniformly along the boundary. Their identification result requires minimal regularity of the boundary $\mathscr{B}$, the distance function $\mathscr{d}(\cdot)$, and the underlying data generating process, building on continuity-based identification in standard univariate RD designs (Hahn et al., 2001). Cattaneo et al. (2025a) also show that the bias of the distance-based point estimator will depend on the smoothness of the assignment boundary, while the variance of the estimator will be of order $(nh^2)^{-1}$. Putting these two results together, it follows that bandwidth selection methods for standard univariate RD designs will either minimize the mean square error

of the estimator or deliver undersmoothing, depending on the geometry of the assignment boundary and other assumptions. For inference, the authors establish asymptotic validity of confidence intervals and confidence bands, provided the bias of the distance-based point estimator is small enough (which, in turn, crucially depends on the geometry of $\mathscr{B}$).

4.2 Location-Based Methods

While the distance-based estimator constructed using specification (10) provides a natural bridge between the pooling methods and heterogeneity analysis, Cattaneo et al. (2025a) demonstrate that this approach would require a smooth assignment boundary for higher-order bias reduction. Otherwise, the distance-based approach will generate a consistent estimator with possibly a high bias, and therefore require a small bandwidth for estimation and inference validity. A solution to this problem is to use local bivariate regression methods.

Cattaneo et al. (2025b) study the properties of local regression treatment effect estimators directly employing the location information encoded in $\mathbf{X}_i$ for each unit. The authors develop pointwise and uniform estimation and inference methods for both the BATEC and transformations thereof, such as the WBATE and LBATE. Their local regression specification takes the form:

$$\texttt{reg} \quad Y_i \quad \texttt{on} \quad 1,\ T_i,\ \mathbf{r}_p(\mathbf{X}_i - \mathbf{b}_j),\ T_i \cdot \mathbf{r}_p(\mathbf{X}_i - \mathbf{b}_j) \qquad [\ K\big(\tfrac{\mathbf{X}_i - \mathbf{b}_j}{h}\big)\ ], \tag{11}$$

where $\mathbf{b}_1, \ldots, \mathbf{b}_J$ is the grid of points on the boundary (e.g., Figure 7a). This local polynomial specification crucially depends on $\mathbf{X}_i$ directly via the pth order polynomial expansion $\mathbf{r}_p(\mathbf{X}_i - \mathbf{x})$. The weights are determined by the bivariate kernel function $K(u_1, u_2)$, and the localization to each point $\mathbf{b}_j$ continues to be is controlled by the bandwidth h. In particular, $K\big(\tfrac{\mathbf{X}_i - \mathbf{b}_j}{h}\big) = \mathbb{1}(\|\mathbf{X}_i - \mathbf{b}_j\| \leq h) = \mathbb{1}(|D_i(\mathbf{b}_j)| \leq h)$ is a valid choice, demonstrating an interesting connection with distance-based methods as implemented via specification (10). The estimated coefficient accompanying T_i, denoted by $\widehat{\tau}(\mathbf{b}_j)$, gives the location-based point

estimate of $\tau(\mathbf{b}_j)$, for each $j = 1, \ldots, J$.

In the location-based setting, identification of $\tau(\mathbf{x})$ for $\mathbf{x} \in \mathscr{B}$ follows directly from a generalization of continuity-based identification results in the standard univariate RD design (see Hahn et al., 2001; Papay et al., 2011; Reardon and Robinson, 2012; Wong et al., 2013; Keele and Titiunik, 2015; Cattaneo et al., 2016). Under minimal regularity conditions on the boundary and data generating process, Cattaneo et al. (2025b) establishes (pointwise and) uniform convergence rates of the form $\sup_{\mathbf{x}\in\mathscr{B}} |\widehat{\tau}(\mathbf{x}) - \tau(\mathbf{x})|^2 = O_{\mathbb{P}}(\frac{\log(n)}{nh^2} + h^{2(p+1)})$, demonstrating that the location-based estimator does not require stringent smoothness restrictions on the boundary $\mathscr{B}$ to achieve higher order debiasing. Furthermore, the authors establish pointwise and integrated mean square error expansions for the estimator, and develop MSE-optimal bandwidth selection methods. By combining these results, they develop optimal point estimation of the BATEC, both pointwise and uniform, based on specification (11).

For uncertainty quantification, Cattaneo et al. (2025b) consider confidence intervals and confidence bands of the form:

$$\widehat{\mathrm{I}}(\mathbf{x}, \alpha) = \left[\widehat{\tau}(\mathbf{x}) - \mathfrak{c}_\alpha \cdot \sqrt{\widehat{\mathbb{V}}[\widehat{\tau}(\mathbf{x})]} \;\;,\;\; \widehat{\tau}(\mathbf{x}) + \mathfrak{c}_\alpha \cdot \sqrt{\widehat{\mathbb{V}}[\widehat{\tau}(\mathbf{x})]} \right], \qquad \mathbf{x} \in \mathscr{B},$$

where $\widehat{\mathbb{V}}[\widehat{\tau}(\mathbf{x})]$ denotes any of the usual variance estimators from (weighted) least squares regression methods, and $\mathfrak{c}_\alpha$ is the quantile used depending on inferential goal. For confidence intervals, the usual Gaussian quantile is asymptotically valid (i.e., $\mathfrak{c}_\alpha = \Phi^{-1}(1-\alpha/2)$), while for confidence bands, a new (larger) quantile needs to be constructed to capture the joint uncertainty features of the estimators $\widehat{\tau}(\mathbf{x})$ for different values along the boundary $\mathscr{B}$. As in the case of pooling-based and distance-based methods, a key condition needed for validity of these confidence intervals and bands is the small (misspecification) bias property. Cattaneo et al. (2025b) address this issue by relying on robust bias correction (Calonico et al., 2014, 2022), a construction that proceeds in two steps.

- *Step 1.* For a choice p (usually $p = 1$), the MSE-optimal bandwidth is computed,

and then for that choice of bandwidth the point estimator $\widehat{\tau}(\mathbf{x})$ is computed using specification (11). As a result, $\widehat{\tau}(\mathbf{x})$ is an asymptotically MSE-optimal point estimator of $\tau(\mathbf{x})$.

- *Step 2.* For a choice $q > p$ (usually $q = 2$), the point estimator $\widehat{\tau}(\mathbf{x})$ and variance estimator $\widehat{\mathbb{V}}[\widehat{\tau}(\mathbf{x})]$ are computed using specification (11) with the same bandwidth used in *Step 1*. As a result, inference procedures are robust bias-corrected, and hence asymptotically valid.

Cattaneo et al. (2025b) give technical and computational details, which we omit to conserve space, and because they do not fundamentally change the core methodological ideas.

Finally, $\widehat{\tau}(\mathbf{x})$ can also be used to estimate and conduct inference for WBATE and LBATE via plug-in methods. Cattaneo et al. (2025b) provide sufficient conditions and study the validity of such methods. In particular, the WBATE corresponds to an integral over a submanifold of the nonparametric estimator $\widehat{\tau}(\mathbf{x})$. See Chen and Gao (2025) for related theoretical results when employing series approximations instead of local polynomial regression as commonly done in RD settings.

The general-purpose `R` package `rd2d` implement both distance-based and location-based methods (`https://rdpackages.github.io/`). See Cattaneo et al. (2025d) for more discussion. Figure 8 gives a stylized example of the type of results that can be obtained via distance-based and location-based methods, building on the same setup as in Figure 1. These figures are motivated by empirical work using the SPP application (Londoño-Vélez et al., 2020).

5 Recommendations for Practice

Our review of the empirical literature employing BD designs showed that the overwhelming majority of studies focus exclusively on the identification, estimation and inference for the BATE using pooling-based methods (Section 3). This scalar parameter captures a density-

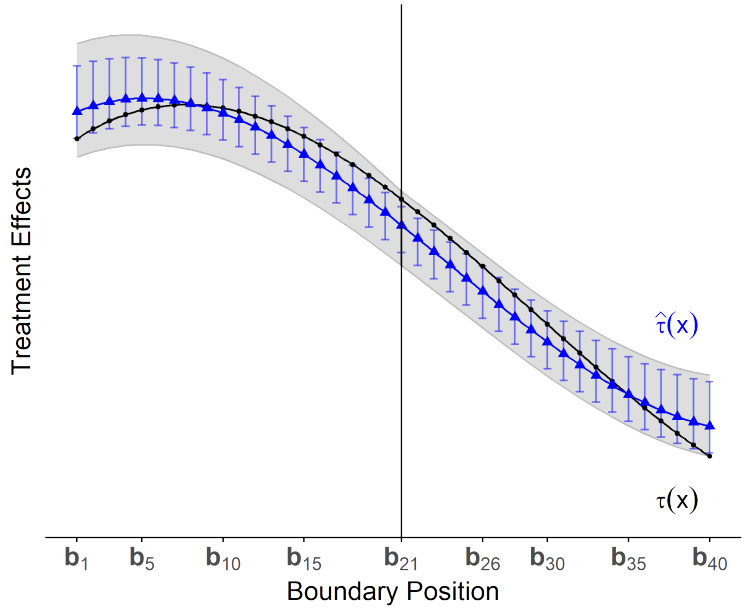


Figure 8: Estimation and Inference for BATEC

weighted average of potentially heterogeneous treatment effects along the assignment boundary. While modern empirical work has recognized the importance of localization and flexible local parametrization, the current literature often exhibits three key limitations: (i) local regression specifications frequently omit the interaction between the treatment assignment variable T_i and the flexible regression terms $\mathbf{r}_p(D_i)$ or $\mathbf{r}_p(\mathbf{X}_i)$, risking significant misspecification bias; (ii) the bandwidth h is often chosen in an ad hoc manner; and (iii) the role of misspecification bias from a "large" bandwidth, such as an MSE-optimal choice, is often ignored, which can lead to over-rejection of the null hypothesis. Cattaneo et al. (2025c) provide a formal study of pooling-based methods that explicitly addresses these limitations. Moving forward, we recommend that empirical researchers implementing pooling-based methods always use specifications (6) or (7), select an MSE-optimal bandwidth, and conducting inference using robust bias-corrected methods.

More broadly, as noted by Cattaneo et al. (2025a,b), a more general approach to analyzing and interpreting BD designs is to first focus on the BATEC parameter, and then consider other aggregate causal treatment effects as transformations thereof. As summarized in Sec-

tion 4, distance-based and location-based methods capture the full richness of the BD design by estimating average treatment effects at each point along the assignment boundary. These methods also recover the BATE parameter (Section 3) as a special case, as well as other interesting causal parameters such as the LBATE. These methods can be implemented using the general-purpose `R` package `rd2d` available at `https://rdpackages.github.io/`, and discussed in Cattaneo et al. (2025d). In particular, location-based methods are the most robust and general approach; we therefore recommend them for future empirical work leveraging BD designs.

Table 1: Selected Empirical Articles Using The Boundary Discontinuity Design

Article	Y_i	T_i	$\mathbf{X}_i$	Distance	Specifications	Bandwidth	Heter
Card and Krueger (1994)	Employment	Minimum wage increase	Location of towns/counties	NA	(1), (2)	NA	No
Holmes (1998)	Manufacturing activity	Right-to-work policies	Lat, Lon	Euclidean	(1), (2), (6)	Manual	Yes
Black (1999)	House prices	School quality	Lat, Lon	Euclidean	(1), (2)	Manual	No
Kane et al. (2003)	House prices	School quality	Lat, Lon	Euclidean	(1), (2)	Manual	No
Kane et al. (2006)	Housing prices	School quality	Lat, Lon	Euclidean	(1), (2)*	Manual	No
Bayer et al. (2007)	House prices	School quality	Lat, Lon	Euclidean	(1), (2)	Manual	No
Lalive (2008)	Unemployment duration	Unemployment benefits	Lat, Lon	Driving time	(6)	Manual	No
Dell (2010)	Consumption, childhood stunting	Forced mining labor	Lat, Lon	Euclidean	(3)	MSE	No
Ou (2010)	High school dropout	Failure of high school exit exam	Mathematics score, Language score	Euclidean	(8)	Manual	No
Grout et al. (2011)	Property values	Land-use regulations	Lat, Lon	Euclidean	(8)	Manual	Yes
Robinson (2011)	English proficiency	English proficiency reclassification	Scores on five language tests	Euclidean	Binding score	MSE	No
Hinnerich and Pettersson-Lidbom (2014)	Political regimes	Income redistribution	Population in two separate years	Euclidean	(8)	MSE	No
Ferwerda and Miller (2014)	Resistance activity	Devolution of governing authority	Lat, Lon	Euclidean	(6)	MSE, Manual	No

Continued on next page...

Table 1 – Continued from previous page

Article	Y_i	T_i	$\mathbf{X}_i$	Distance	Specifications	Bandwidth	Heter
Snider and Williams (2015)	Airline industry outcomes	Access to airport facilities granted to new airlines	Share of passengers in top two carriers in origin and destination city		(1), (11)	MSE, Manual	Yes
Barone et al. (2015)	Vote shares	Switch from analog to digital TV	Lat, Lon	Euclidean	(5)	Manual	No
Michalopoulos and Papaioannou (2016)	Civil conflict	Ethnic partitioning	Lat, Lon	Euclidean	(1), (2)	Manual	No
Egger and Wamser (2015)	Foreign investment	Controlled foreign company rules	Tax rate, Income measures	NA	(7) $L = 1$	MSE	No
Keele and Titiunik (2015)	Voter turnout	Political TV advertisement	Lat, Lon	Chordal	(10)	MSE	Yes
MacDonald et al. (2016)	Crime	Policing	Lat, Lon	Euclidean	(6)	MSE	No
Evans (2017)	STEM major	Financial incentive	GPA, Family contribution score	Euclidean	(8)	MSE	No
Kumar (2018)	Mortgage default	Limits on home equity borrowing	Lat, Lon	Euclidean	(4), (3)	Manual	No
Ehrlich and Seidel (2018)	Economic density	Subsidies	Lat, Lon	Euclidean	(4)+$\mathbf{r}_p(\mathbf{X}_i)$	MSE, Manual	No
Dell et al. (2018)	Economic outcomes	Dai Viet administrative institutions	Lat, Lon	Euclidean	(3), (5), (5)+$\mathbf{r}_p(\mathbf{X}_i)$	Manual	No
Dell and Querubin (2018)	Public goods, political attitudes	War strategies	Lat, Lon	Euclidean	(3)	Manual	No
Clinton and Sances (2018)	Voter registration and turnout	Medicaid expansion	Lat, Lon	Euclidean	(6)	Manual	No

Continued on next page...

Table 1 – Continued from previous page

Article	Y_i	T_i	$\mathbf{X}_i$	Distance	Specifications	Bandwidth	Heter
Spenkuch and Toniatti (2018)	Voter turnout	Television advertisement	Lat, Lon	Euclidean	(2), (6)	MSE, Manual	No
De Kadt and Larreguy (2018)	Party vote shares	Political alignment	Lat, Lon	Euclidean	(3)	Manual	No
Eugster and Parchet (2019)	Tax competition	Cultural preferences	Lat, Lon	Road	(1) in DID	Manual	No
Giuntella and Mazzonna (2019)	Sleep duration, Health outcomes	Timing of natural light	Lat, Lon	Euclidean	(6)	MSE, Manual	No
Johnson (2019)	High School graduation	English learner classification	Scores on two language tests	Euclidean	Binding score	MSE	No
Velez and Newman (2019)	Voter turnout	Spanish-language television station	Lat, Lon	Euclidean	(1), (2), (10)	MSE, Manual	Yes
Dupraz (2019)	Schooling	Colonial legacies	Lat, Lon	Euclidean	(6), (3)	MSE	No
Moscona et al. (2020)	Conflict	Segmentary lineage organization	Lat, Lon	Euclidean	(6)	Manual	No
Ambrus et al. (2020)	Property value	Cholera epidemic	Lat, Lon	Walking	(6) $L = 1$	MSE, Manual	No
Dell and Olken (2020)	Schools, education levels	Sugar cultivation for Dutch cultivation system	Lat, Lon	Euclidean	(3)	Manual	No
He et al. (2020)	Pollution	Water quality monitoring	Lat, Lon	Euclidean	(6) $L = 1$	MSE	No or NA
Ito and Zhang (2020)	Willingness to pay for clean air		Lat, Lon	Euclidean, road	(6)	MSE, Manual	No
Wuepper et al. (2020a)	Soil erosion	Country borders	Lat, Lon	Euclidean	(6)	MSE	No

Continued on next page...

Table 1 – Continued from previous page

Article	Y_i	T_i	$\mathbf{X}_i$	Distance	Specifications	Bandwidth	Heter
Albertus (2020)	Conflict	Land reform	Lat, Lon	Euclidean	(6)	MSE, Manual	No
Wuepper et al. (2020b)	Crop yield gaps, nitrogen pollution	Country borders	Lat, Lon	Euclidean	(6)	MSE	No
Schafer and Holbein (2020)	Voter turnout	Time zones	Lat, Lon	Chordal	(6)	MSE, Manual	No
Letsa and Wilfahrt (2020)	Attitudes towards local power	Colonial legacies	Lat, Lon	Euclidean	(3)	Manual	No
Londoño-Vélez et al. (2020)	College enrollment	Financial subsidy	High school exit exam score, wealth index	Euclidean	(8)	MSE	No
Gonzalez (2021)	Election fraud	Cell phone coverage	Lat, Lon	Euclidean	(6), (7)	MSE	Yes
Ayres et al. (2021)	Land value	Adjudication of groundwater rights	Lat, Lon	Euclidean	(6) $L = 1$	MSE	No
Laliberté (2021)	School quality	Educational achievement	Lat, Lon	Euclidean	(6)	MSE	No
Aaronson et al. (2021)	Characteristics of urban neighborhoods	Red lining maps	Lat, Lon	Euclidean	(2) in DID	Manual	No
Appau et al. (2021)	Agricultural productivity	Bombing intensity	Lat, Lon	Absolute difference between centroid and 17th parallel north latitude	(4) (first stage)	Manual	No
Lowes and Montero (2021)	Education, health, wealth	Concession to extract rubber	Lat, Lon	Euclidean	(6)	MSE, Manual	No
Sides et al. (2022)	Election outcomes	Television advertisement	Lat, Lon	Euclidean	(2)	Manual	No

Continued on next page...

Table 1 – Continued from previous page

Article	Y_i	T_i	$\mathbf{X}_i$	Distance	Specifications	Bandwidth	Heter
Zheng (2022)	House prices	Charter school entry	Lat, Lon	Euclidean	Event study in sample near boundary	Manual	No
Dehdari and Gehring (2022)	Regional identity	Actions of nation states	Lat, Lon	Euclidean	(5)	MSE, Manual	No
Méndez and Van Patten (2022)	Living standards	Land concession to multinational firm	Lat, Lon	Euclidean	(3)	Manual	No
Jones et al. (2022a)	Profits, irrigation efficiency	Irrigation	Lat, Lon	Euclidean	(6)	Manual	No
Henn (2023)	State capacity	Remoteness from administrative headquarters	Lat, Lon	Euclidean	(6)*	Manual	No
Jones et al. (2022b)	Enrollment, graduation	Tuition scholarship	High school GPA, SAT/ACT score	NA	(7)	Manual	No
Dai et al. (2022)	Lifestyle outcomes	Hypertension diagnosis	Systolic and Diastolic blood pressure	Euclidean	(8)	MSE	No
Salti et al. (2022)	Economic outcomes	Cash assistance	Two scores in means testing formula	NA	(8)*	MSE	No
Castro and Esposito (2022)	Teacher retention, student learning	Recruitment bonus	Population, distance to capital	Euclidean	(8)	MSE	No
Moussa et al. (2022)	Children health outcomes	Cash assistance	Two scores in means testing formula	NA	(8)*	MSE	No
Mangonnet et al. (2022)	Protected area designation	Political alignment	Lat, Lon	Euclidean	(2)	Manual	No

Continued on next page...

Table 1 – Continued from previous page

Article	Y_i	T_i	$\mathbf{X}_i$	Distance	Specifications	Bandwidth	Heter
Baragwanath et al. (2023)	Forest growth	Property rights	Lat, Lon	Euclidean	(6)	MSE	No
Larsen and Valant (2024)	Educational attainment	Grade retention	Mathematics and English Language Arts test scores	Euclidean	(8)	MSE	No
Woller et al. (2023)	Vote buying	Cost of voting	Lat, Lon	Euclidean	(4)	Manual	No
Paulsen et al. (2023)	Voter turnout, economic outcomes	State funding for common schools	Lat, Lon	Euclidean	(3)	Manual	No
Prillaman (2023)	Political participation	Participation in women-only credit groups	Lat, Lon	Euclidean	(2), (3)	Manual	No
McAlexander (2023)	Rebel actions	UN partition line of Palestine	Lat, Lon	Euclidean	(6) $L = 1$	MSE	No
Murphy and Johnson (2023)	Special education placement	English learner designation	Various scores of language assessment	NA	Binding score	MSE	No
Bjerre-Nielsen and Gandil (2024)	School enrollment	School change	Lat, Lon	Euclidean	(2) in DID	Manual	No
Kämpfen and Mosca (2024)	Blood pressure	Blood pressure diagnosis	Systolic and Diastolic blood pressure	NA	Binding score	MSE	No
Wuepper et al. (2024)	Deforestation	Country borders	Lat, Lon	Euclidean	(6)	MSE	No
Cox et al. (2024)	Voter mobilization	Social networks	Lat, Lon	Fastest driving	(1)*	NA	No
Doucette (2024a)	Urbanization, commercialization	Duchy of Württemberg	Lat, Lon	Euclidean	(4)+$\mathbf{r}_p(\mathbf{X}_i)$	MSE, Manual	No

Continued on next page...

Table 1 – Continued from previous page

Article	Y_i	T_i	$\mathbf{X}_i$	Distance	Specifications	Bandwidth	Heter
Doucette (2024b)	Party support	Inclusive institutions	Lat, Lon	Euclidean	(4)+$\mathbf{r}_p(\mathbf{X}_i)$	Manual	No
Grasse (2024)	Poverty, economic development	Mass repression	Lat, Lon	Euclidean	(5)	Manual	No
Jardim et al. (2024)	Wages, employment	Increase in minimum wage	Lat, Lon	Driving time	(1)*, (2)*	Manual	No
Ring (2025)	Savings	Wealth tax assessment	Lat, Lon	Euclidean	(6)	Manual	No
Yamagishi and Sato (2025)	Land price	Buraku neighborhoods	Lat, Lon	Euclidean	(1), (6) $L = 1$	MSE, Manual	No
Loumeau (2025)	Commuting flows, residential decisions	Departmental borders	Lat, Lon	Euclidean	(5)	MSE	No
Boix (2025)	Jewish National identity	Political emancipation	Lat, Lon	Euclidean	(6)	MSE	No
Müller-Crepon (2025)	Share of predominant ethnic group	Administrative borders	Lat, Lon	Euclidean	(6)	Manual	No

Note: MSE denotes optimal mean squared error bandwidth choice. Binding score refers to the approach of constructing a single univariate score as the minimum of all scores that determine treatment assignment, known as the binding score, and then conducting a univariate RD analysis using the binding score (see, e.g., Reardon and Robinson, 2012). A specification number with the symbol * indicates it has been modified in a some way; for details, we refer readers to the article. GPA denotes grade point average.

References

Aaronson, D., D. Hartley, and B. Mazumder (2021): "The effects of the 1930s HOLC "redlining" maps," *American Economic Journal: Economic Policy*, 13, 355–392.

Abadie, A. and M. D. Cattaneo (2018): "Econometric Methods for Program Evaluation," *Annual Review of Economics*, 10, 465–503.

Albertus, M. (2020): "Land reform and civil conflict: Theory and evidence from Peru," *American Journal of Political Science*, 64, 256–274.

Ambrus, A., E. Field, and R. Gonzalez (2020): "Loss in the time of cholera: Long-run impact of a disease epidemic on the urban landscape," *American Economic Review*, 110, 475–525.

Andrews, I., T. Kitagawa, and A. McCloskey (2024): "Inference on winners," *Quarterly Journal of Economics*, 139, 305–358.

Appau, S., S. A. Churchill, R. Smyth, and T.-A. Trinh (2021): "The long-term impact of the Vietnam War on agricultural productivity," *World Development*, 146, 105613.

Avnir, D., O. Biham, D. Lidar, and O. Malcai (1998): "Is the geometry of nature fractal?" *Science*, 279, 39–40.

Ayres, A. B., K. C. Meng, and A. J. Plantinga (2021): "Do environmental markets improve on open access? Evidence from California groundwater rights," *Journal of Political Economy*, 129, 2817–2860.

Banerjee, S. (2005): "On Geodetic Distance Computations in Spatial Modeling," *Biometrics*, 61, 617–625.

Baragwanath, K., E. Bayi, and N. Shinde (2023): "Collective property rights lead to secondary forest growth in the Brazilian Amazon," *Proceedings of the National Academy of Sciences*, 120, e2221346120.

Barone, G., F. D'Acunto, and G. Narciso (2015): "Telecracy: Testing for channels of persuasion," *American Economic Journal: Economic Policy*, 7, 30–60.

Bayer, P., F. Ferreira, and R. McMillan (2007): "A unified framework for measuring preferences for schools and neighborhoods," *Journal of political economy*, 115, 588–638.

Bjerre-Nielsen, A. and M. H. Gandil (2024): "Attendance boundary policies and the limits to combating school segregation," *American Economic Journal: Economic Policy*, 16, 190–227.

Black, S. E. (1999): "Do better schools matter? Parental valuation of elementary education," *The quarterly journal of economics*, 114, 577–599.

Boix, C. (2025): "Political emancipation and modern Jewish national identity," *American Political Science Review*, 1–21.

Calonico, S., M. D. Cattaneo, and M. H. Farrell (2018): "On the Effect of Bias Estimation on Coverage Accuracy in Nonparametric Inference," *Journal of the American Statistical Association*, 113, 767–779.

——— (2020): "Optimal Bandwidth Choice for Robust Bias Corrected Inference in Regression Discontinuity Designs," *Econometrics Journal*, 23, 192–210.

——— (2022): "Coverage Error Optimal Confidence Intervals for Local Polynomial Regression," *Bernoulli*, 28, 2998–3022.

Calonico, S., M. D. Cattaneo, M. H. Farrell, F. Palomba, and R. Titiunik (2025): "Treatment Effect Heterogeneity in Regression Discontinuity Designs," *arXiv preprint arXiv:2503.13696*.

Calonico, S., M. D. Cattaneo, M. H. Farrell, and R. Titiunik (2019): "Regression Discontinuity Designs using Covariates," *Review of Economics and Statistics*, 101, 442–451.

Calonico, S., M. D. Cattaneo, and R. Titiunik (2014): "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs," *Econometrica*, 82, 2295–2326.

——— (2015): "Optimal Data-Driven Regression Discontinuity Plots," *Journal of the American Statistical Association*, 110, 1753–1769.

Card, D. and A. B. Krueger (1994): "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania," *American Economic Review*, 84, 772–793.

Castro, J. F. and B. Esposito (2022): "The effect of bonuses on teacher retention and student learning in rural schools: A story of spillovers," *Education finance and policy*, 17, 693–718.

Cattaneo, M. D., B. Frandsen, and R. Titiunik (2015): "Randomization Inference in the Regression Discontinuity Design: An Application to Party Advantages in the U.S. Senate," *Journal of Causal Inference*, 3, 1–24.

Cattaneo, M. D., N. Idrobo, and R. Titiunik (2020): *A Practical Introduction to Regression Discontinuity Designs: Foundations*, Cambridge University Press.

——— (2024): *A Practical Introduction to Regression Discontinuity Designs: Extensions*, Cambridge University Press.

Cattaneo, M. D., L. Keele, R. Titiunik, and G. Vazquez-Bare (2016): "Interpreting Regression Discontinuity Designs with Multiple Cutoffs," *Journal of Politics*, 78, 1229–1248.

Cattaneo, M. D. and R. Titiunik (2022): "Regression Discontinuity Designs," *Annual Review of Economics*, 14, 821–851.

Cattaneo, M. D., R. Titiunik, and G. Vazquez-Bare (2017): "Comparing Inference Approaches for RD Designs: A Reexamination of the Effect of Head Start on Child Mortality," *Journal of Policy Analysis and Management*, 36, 643–681.

Cattaneo, M. D., R. Titiunik, and R. R. Yu (2025a): "Estimation and Inference in Boundary Discontinuity Designs: Distance-Based Methods," *Working paper*.

——— (2025b): "Estimation and Inference in Boundary Discontinuity Designs: Location-Based Methods," *Working paper*.

——— (2025c): "Estimation and Inference in Boundary Discontinuity Designs: Pooling-Based Methods," *Working paper*.

——— (2025d): "`rd2d`: Causal Inference in Boundary Discontinuity Designs," *Working paper*.

Chen, X. and W. Y. Gao (2025): "Semiparametric Learning of Integral Functionals on Submanifolds," *arXiv preprint arXiv:2507.12673*.

Clinton, J. D. and M. W. Sances (2018): "The politics of policy: The initial mass political effects of medicaid expansion in the states," *American Political Science Review*, 112, 167–185.

Cox, G. W., J. H. Fiva, and M.-E. M. King (2024): "Bound by borders: Voter mobilization through social networks," *British Journal of Political Science*, 54, 1198–1216.

Dai, T., S. Jiang, X. Liu, and A. Sun (2022): "The effects of a hypertension diagnosis on health behaviors: A two-dimensional regression discontinuity analysis," *Health Economics*, 31, 574–596.

De Kadt, D. and H. A. Larreguy (2018): "Agents of the regime? Traditional leaders and electoral politics in South Africa," *The Journal of Politics*, 80, 382–399.

De Magalhães, L., D. Hangartner, S. Hirvonen, J. Meriläinen, N. A. Ruiz, and J. Tukiainen (2025): "When Can We Trust Regression Discontinuity Design Estimates from Close Elections? Evidence from Experimental Benchmarks," *Political Analysis*.

Dehdari, S. H. and K. Gehring (2022): "The origins of common identity: Evidence from Alsace-Lorraine," *American Economic Journal: Applied Economics*, 14, 261–292.

Dell, M. (2010): "The persistent effects of Peru's mining mita," *Econometrica*, 78, 1863–1903.

Dell, M., N. Lane, and P. Querubin (2018): "The historical state, local collective action, and economic development in Vietnam," *Econometrica*, 86, 2083–2121.

Dell, M. and B. A. Olken (2020): "The development effects of the extractive colonial economy: The dutch cultivation system in java," *The Review of Economic Studies*, 87, 164–203.

Dell, M. and P. Querubin (2018): "Nation building through foreign intervention: Evidence from discontinuities in military strategies," *The Quarterly Journal of Economics*, 133, 701–764.

Diaz, J. D. and J. R. Zubizarreta (2023): "Complex Discontinuity Designs Using Covariates for Policy Impact Evaluation," *Annals of Applied Statistics*, 17, 67–88.

Doucette, J. (2024a): "Parliamentary Constraints and Long-Term Development: Evidence from the Duchy of W¨urttemberg," *American Journal of Political Science*, 68, 24–41.

Doucette, J. S. (2024b): "Pre-Modern Institutions and Later Support for Autocrats in Democratic Elections," *British Journal of Political Science*, 54, 1395–1403.

Dupraz, Y. (2019): "French and British colonial legacies in education: Evidence from the partition of Cameroon," *The Journal of Economic History*, 79, 628–668.

Egger, P. H. and G. Wamser (2015): "The impact of controlled foreign company legislation on real investments abroad. A multi-dimensional regression discontinuity design," *Journal of Public Economics*, 129, 77–91.

Ehrlich, M. v. and T. Seidel (2018): "The persistent effects of place-based policy: Evidence from the West-German Zonenrandgebiet," *American Economic Journal: Economic Policy*, 10, 344–374.

Eugster, B. and R. Parchet (2019): "Culture and taxes," *Journal of Political Economy*, 127, 296–337.

Evans, B. J. (2017): "SMART money: Do financial incentives encourage college students to study science?" *Education Finance and Policy*, 12, 342–368.

Federer, H. (2014): *Geometric measure theory*, Springer.

Ferwerda, J. and N. L. Miller (2014): "Political devolution and resistance to foreign rule: A natural experiment," *American Political Science Review*, 108, 642–660.

Galiani, S., P. J. McEwan, and B. Quistorff (2017): "External and Internal Validity of a Geographic Quasi-Experiment Embedded in a Cluster-Randomized Experiment," in *Regression Discontinuity Designs: Theory and Applications (Advances in Econometrics, volume 38)*, ed. by M. D. Cattaneo and J. C. Escanciano, Emerald Group Publishing, 195–236.

Giuntella, O. and F. Mazzonna (2019): "Sunset time and the economic effects of social jetlag: evidence from US time zone borders," *Journal of health economics*, 65, 210–226.

Gonzalez, R. M. (2021): "Cell phone access and election fraud: evidence from a spatial regression discontinuity design in Afghanistan," *American Economic Journal: Applied Economics*, 13, 1–51.

Grasse, D. (2024): "State terror and long-run development: The persistence of the Khmer Rouge," *American Political Science Review*, 118, 195–212.

Grout, C. A., W. K. Jaeger, and A. J. Plantinga (2011): "Land-use regulations and property values in Portland, Oregon: A regression discontinuity design approach," *Regional Science and Urban Economics*, 41, 98–107.

Hahn, J., P. Todd, and W. van der Klaauw (2001): "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design," *Econometrica*, 69, 201–209.

He, G., S. Wang, and B. Zhang (2020): "Watering down environmental regulation in China," *The quarterly journal of economics*, 135, 2135–2185.

Henn, S. J. (2023): "Complements or substitutes? How institutional arrangements bind traditional authorities and the state in Africa," *American Political Science Review*, 117, 871–890.

Hernán, M. A. and J. M. Robins (2020): *Causal Inference: What If*, Boca Raton: Chapman & Hall/CRC.

Hinnerich, B. T. and P. Pettersson-Lidbom (2014): "Democracy, redistribution, and political participation: Evidence from Sweden 1919–1938," *Econometrica*, 82, 961–993.

Holmes, T. J. (1998): "The effect of state policies on the location of manufacturing: Evidence from state borders," *Journal of political Economy*, 106, 667–705.

Hyytinen, A., J. Meriläinen, T. Saarimaa, O. Toivanen, and J. Tukiainen (2018): "When does regression discontinuity design work? Evidence from random election outcomes," *Quantitative Economics*, 9, 1019–1051.

Ito, K. and S. Zhang (2020): "Willingness to pay for clean air: Evidence from air purifier markets in China," *Journal of Political Economy*, 128, 1627–1672.

Jardim, E., M. C. Long, R. Plotnick, J. Vigdor, and E. Wiles (2024): "Local minimum wage laws, boundary discontinuity methods, and policy spillovers," *Journal of Public Economics*, 234, 105131.

Johnson, A. (2019): "The effects of English learner classification on high school graduation and college attendance," *AERA Open*, 5, 2332858419850801.

Jones, M., F. Kondylis, J. Loeser, and J. Magruder (2022a): "Factor market failures and the adoption of irrigation in rwanda," *American Economic Review*, 112, 2316–2352.

Jones, T. R., D. Kreisman, R. Rubenstein, C. Searcy, and R. Bhatt (2022b): "The effects of financial aid loss on persistence and graduation: A multi-dimensional regression discontinuity approach," *Education Finance and Policy*, 17, 206–231.

Kämpfen, F. and I. Mosca (2024): "Heterogeneous effects of blood pressure screening," *Economics Letters*, 242, 111845.

Kane, T. J., S. K. Riegg, and D. O. Staiger (2006): "School quality, neighborhoods, and housing prices," *American law and economics review*, 8, 183–212.

Kane, T. J., D. O. Staiger, G. Samms, E. W. Hill, and D. L. Weimer (2003): "School accountability ratings and housing values [with comments]," *Brookings-Wharton papers on urban Affairs*, 83–137.

Keele, L. and R. Titiunik (2016): "Natural experiments based on geography," *Political Science Research and Methods*, 4, 65–95.

Keele, L. J., S. Lorch, M. Passarella, D. Small, and R. Titiunik (2017): "An Overview of Geographically Discontinuous Treatment Assignments with an Application to Children's Health Insurance," in *Regression Discontinuity Designs: Theory and Applications (Advances in Econometrics, volume 38)*, ed. by M. D. Cattaneo and J. C. Escanciano, Emerald Group Publishing, 147–194.

Keele, L. J. and R. Titiunik (2015): "Geographic Boundaries as Regression Discontinuities," *Political Analysis*, 23, 127–155.

Keele, L. J., R. Titiunik, and J. Zubizarreta (2015): "Enhancing a Geographic Regression Discontinuity Design Through Matching to Estimate the Effect of Ballot Initiatives on Voter Turnout," *Journal of the Royal Statistical Society: Series A*, 178, 223–239.

Kumar, A. (2018): "Do restrictions on home equity extraction contribute to lower mortgage defaults? Evidence from a policy discontinuity at the Texas border," *American Economic Journal: Economic Policy*, 10, 268–297.

Laliberté, J.-W. (2021): "Long-term contextual effects in education: Schools and neighborhoods," *American Economic Journal: Economic Policy*, 13, 336–377.

Lalive, R. (2008): "How do extended benefits affect unemployment duration? A regression discontinuity approach," *Journal of econometrics*, 142, 785–806.

Larsen, M. F. and J. Valant (2024): "The long-term effects of grade retention: Evidence on persistence through high school and college," *Journal of Research on Educational Effectiveness*, 17, 615–646.

Lee, D. S. and T. Lemieux (2010): "Regression Discontinuity Designs in Economics," *Journal of Economic Literature*, 48, 281–355.

Letsa, N. W. and M. Wilfahrt (2020): "The mechanisms of direct and indirect rule: Colonialism and economic development in Africa," *Quarterly Journal of Political Science*, 15, 539–577.

Londoño-Vélez, J., C. Rodríguez, and F. Sánchez (2020): "Upstream and downstream impacts of college merit-based financial aid for low-income students: Ser Pilo Paga in Colombia," *American Economic Journal: Economic Policy*, 12, 193–227.

Loumeau, G. (2025): "Regional Borders, Commuting, and Transport Network Integration," *Review of Economics and Statistics*, 1–15.

Lowes, S. and E. Montero (2021): "Concessions, violence, and indirect rule: evidence from the Congo Free State," *The Quarterly Journal of Economics*, 136, 2047–2091.

MacDonald, J. M., J. Klick, and B. Grunwald (2016): "The effect of private police on crime: evidence from a geographic regression discontinuity design," *Journal of the Royal Statistical Society Series A: Statistics in Society*, 179, 831–846.

Mandelbrot, B. (1967): "How long is the coast of Britain? Statistical self-similarity and fractional dimension," *science*, 156, 636–638.

Mandelbrot, B. B. (1983): "The fractal geometry of nature," *New York*.

Mangonnet, J., J. Kopas, and J. Urpelainen (2022): "Playing politics with environmental protection: the political economy of designating protected areas," *The Journal of Politics*, 84, 1453–1468.

McAlexander, R. J. (2023): "How do international borders affect conflict processes? Evidence from the end of Mandate Palestine," *Journal of Peace Research*, 60, 823–838.

Méndez, E. and D. Van Patten (2022): "Multinationals, monopsony, and local development: Evidence from the united fruit company," *Econometrica*, 90, 2685–2721.

Michalopoulos, S. and E. Papaioannou (2016): "The long-run effects of the scramble for Africa," *American Economic Review*, 106, 1802–1848.

Moscona, J., N. Nunn, and J. A. Robinson (2020): "Segmentary Lineage Organization and Conflict in Sub-Saharan Africa," *Econometrica*, 88, 1999–2036.

Moussa, W., N. Salti, A. Irani, R. Al Mokdad, Z. Jamaluddine, J. Chaaban, and H. Ghattas (2022): "The impact of cash transfers on Syrian refugee children in Lebanon," *World Development*, 150, 105711.

Müller-Crepon, C. (2025): "Building tribes: How administrative units shaped ethnic groups in Africa," *American Journal of Political Science*, 69, 406–422.

Murphy, M. and A. Johnson (2023): "Dual identification? The effects of English learner (EL) status on subsequent special education (SPED) placement in an equity-focused district," *Educational Evaluation and Policy Analysis*, 45, 311–335.

Ou, D. (2010): "To leave or not to leave? A regression discontinuity analysis of the impact of failing the high school exit exam," *Economics of Education Review*, 29, 171–186.

Papay, J. P., J. B. Willett, and R. J. Murnane (2011): "Extending the regression-discontinuity approach to multiple assignment variables," *Journal of Econometrics*, 161, 203–207.

Paulsen, T., K. Scheve, and D. Stasavage (2023): "Foundations of a new democracy: schooling, inequality, and voting in the early republic," *American Political Science Review*, 117, 518–536.

Prillaman, S. A. (2023): "Strength in numbers: how women's groups close India's political gender gap," *American Journal of Political Science*, 67, 390–410.

Reardon, S. F. and J. P. Robinson (2012): "Regression Discontinuity Designs with Multiple Rating-Score Variables," *Journal of Research on Educational Effectiveness*, 5, 83–104.

Ring, M. A. (2025): "Wealth taxation and household saving: Evidence from assessment discontinuities in Norway," *Review of economic studies*, 92, 3375–3402.

Rischard, M., Z. Branson, L. Miratrix, and L. Bornn (2021): "Do School Districts Affect NYC House Prices? Identifying Border Differences using a Bayesian Nonparametric Approach to Geographic Regression Discontinuity Designs," *Journal of the American Statistical Association*, 116, 619–631.

Robinson, J. P. (2011): "Evaluating criteria for English learner reclassification: A causal-effects approach using a binding-score regression discontinuity design with instrumental variables," *Educational Evaluation and Policy Analysis*, 33, 267–292.

Salti, N., J. Chaaban, W. Moussa, A. Irani, R. Al Mokdad, Z. Jamaluddine, and H. Ghattas (2022): "The impact of cash transfers on Syrian refugees in Lebanon: Evidence from a multidimensional regression discontinuity design," *Journal of Development Economics*, 155, 102803.

Schafer, J. and J. B. Holbein (2020): "When time is of the essence: A natural experiment on how time constraints influence elections," *The Journal of Politics*, 82, 418–432.

Sides, J., L. Vavreck, and C. Warshaw (2022): "The effect of television advertising in United States elections," *American Political Science Review*, 116, 702–718.

Snider, C. and J. W. Williams (2015): "Barriers to entry in the airline industry: A multidimensional regression-discontinuity analysis of AIR-21," *Review of Economics and Statistics*, 97, 1002–1022.

Spenkuch, J. L. and D. Toniatti (2018): "Political advertising and election results," *The Quarterly Journal of Economics*, 133, 1981–2036.

Velez, Y. R. and B. J. Newman (2019): "Tuning in, not turning out: Evaluating the impact of ethnic television on political participation," *American Journal of Political Science*, 63, 808–823.

Woller, A., M. K. Justesen, and J. G. Hariri (2023): "The cost of voting and the cost of votes," *The Journal of Politics*, 85, 593–608.

Wong, V. C., P. M. Steiner, and T. D. Cook (2013): "Analyzing regression-discontinuity designs with multiple assignment variables: A comparative study of four estimation methods," *Journal of Educational and Behavioral Statistics*, 38, 107–141.

Wuepper, D., P. Borrelli, and R. Finger (2020a): "Countries and the global rate of soil erosion," *Nature sustainability*, 3, 51–55.

Wuepper, D., T. Crowther, T. Lauber, D. Routh, S. Le Clec'h, R. D. Garrett, and J. B¨orner (2024): "Public policies and global forest conservation: Empirical evidence from national borders," *Global Environmental Change*, 84, 102770.

Wuepper, D., S. Le Clech, D. Zilberman, N. Mueller, and R. Finger (2020b): "Countries influence the trade-off between crop yields and nitrogen pollution," *Nature Food*, 1, 713–719.

Yamagishi, A. and Y. Sato (2025): "Persistent Stigma in Space: 100 Years of Japan's Invisible Race and Their Neighborhoods," *Review of Economics and Statistics*, 1–46.

Zheng, A. (2022): "The valuation of local school quality under school choice," *American Economic Journal: Economic Policy*, 14, 509–537.