A Note on Spatial Regression Discontinuity Designs

A Preprint

Alexander Lehner
University of Chicago
preliminary and incomplete, comments appreciated

lehner@uchicago.edu

2024-03-01

Abstract

Implementations of spatial regression discontinuity estimation and inference vary considerably in the literature. I show that many commonly used estimators for spatial RDDs, albeit working well in common scenarios, generally do not identify the local average treatment effect at the RD boundary. Second, I propose a way to report heterogeneous treatment effects alongside the RD cutoff. Third, I introduce randomization inference to the spatial RD framework by creating a set of functions that allow to randomly shift borders. These tools might be interesting for other identification strategies that rely on the shift of boundaries. A companion R-package, SpatialRDD, includes all the tools necessary to carry out spatial RD estimation, including the proposed improvements. The package makes such applications more transparent, and, more importantly ensures easy and straightforward replicability of all necessary steps.

Keywords Geographic Regression Discontinuity Design • Spatial RDD • Causal Inference • R

[Click here for latest version]

lehner@uchicago.edu, Becker Friedman Institute, Development Innovation Lab, 5730 S Woodlawn Ave, Chicago, IL 60637.

Acknowledgements: So far I benefited from comments & conversations (on earlier versions) with Luc Anselin, Roger Bivand, Melissa Dell, Margherita Fort, Nicolà Gennaioli, Julia Koschinsky, Trevon D. Logan, Robson Tigre, Rocío Titiunik, Andrea Weber, David Weil, seminar participants at the University of Oxford (Nuffield, Graduate Seminar Series), the University of Bologna, the Macro Lunch at Brown University, and the Center for Spatial Data Science at

the University of Chicago. A shoutout also goes to the attendants of the PhD course in advanced spatial data analysis at the University of Bologna who were to first ones to test-run the `SpatialRDD` package and gave useful suggestions. The `SpatialRDD` package can be found on Github and CRAN.

1 Introduction

The regression discontinuity design (RDD) is an integral part of the modern causal inference toolbox and is being applied across many disciplines. Dating back to [Thistlethwaite and Campbell, 1960], it tremendously gained in popularity and importance throughout the last two decades [see e.g. Cattaneo et al., 2019, for a recent overview]. When applicable, it allows the researcher to obtain credible causal estimates from purely observational data. In a traditional RDD, the units of observation have a one-dimensional score, sometimes referred to as a running variable. Treatment and control status is assigned based on a single cutoff value of the score that depends only on external factors: all units in the sample with a score above or below a known cutoff are assigned treatment while the rest is not. Under certain straightforward identifying assumptions, the observed jump in the dependent variable of units near the cutoff then represents a (local) average treatment effect [Hahn et al., 2001, Imbens and Lemieux, 2008].

Recently, there has been a growing interest in applications that extend this approach to two or more running variables. For the general case of multivariate RD designs (MRDD) several methodological contributions have been made [Matsudaira, 2008, Papay et al., 2011, Imbens and Zajonc, 2011, Zajonc, 2012, Reardon and Robinson, 2012, Wong et al., 2013], while [Keele and Titiunik, 2015] explicitly focus on geographic RD designs and juxtapose several ways for estimation. Implementations of spatial regression discontinuity estimation and inference, however, vary considerably in the literature, with many researchers fully relying on approaches from the classic RD literature and others focusing on flexibly controlling for geographic variation or simply compare averages of units sufficiently close to the RD boundary.

In this paper, I first juxtapose commonly used spatial RD estimators and show that some are biased when the underlying data generating process (DGP) exhibits a spatial pattern - sometimes not being able to detect a discontinuity and sometimes reporting an RD effect when there is none. Second, I build on Zajonc [2012] and highlight ways to leverage the two-dimensional nature of spatial RDs for visualization, estimation, and inference. With simulations, I illustrate that, in theory, one can trace out a boundary-wide treatment effect curve. Third, I introduce randomization inference to the spatial RD framework and propose a way to obtain Fisher exact p-values by randomly creating counterfactual RD boundaries. Everything that is being proposed has been implemented in the (geo-)statistical R-package `SpatialRDD` [Lehner, 2020]. This software will also make it easy for researchers to carry out robustness checks and make their spatial RD designs fully replicable without having to intransparently rely on point-and-click GIS software.

By comparing all commonly encountered ways of spatial RD estimation in the literature on simulated data, I establish they deliver approximately similar point estimates when there is no spatial component in the DGP. When there are spatial trends, however, some estimating equations do not estimate the correct RD effect with the direction of bias being ambiguous. Depending on the DGP imposed, they either over- or under estimate the effect and sometimes even fail to detect an effect. I establish that it is crucial to allow the conditional expectation function to have different slopes on either side of the RD boundary - as is common practice in the standard RD literature.

By explicitly taking into account the two-dimensional nature of geographic RDs, I follow Zajonc [2012] and show how one can estimate heterogeneous RD effects alongside the boundary. While explicitly integrating/summing estimates over the entire boundary will usually not differ from reducing the problem to one dimension and using distance to the nearest boundary as a scalar forcing variable, I stress two appealing advantages to this approach. First, it makes it possible to leverage the spatial dimension in visualizations. In standard RD designs, credibility often stems in large part from communicating the main results using simple graphs. In practice, however, researchers so far collapse the two-dimensional running variable into a single score and only visualize that. Showing (possibly differing) RD estimates alongside a cutoff gives researchers an opportunity to resort to "visual RD" [Angrist and Pischke, 2014] also in spatial applications of discontinuity designs. Second, such visualizations can be informative for researchers by illustrating heterogeneity and allow to investigate specific channels through which a discontinuously induced treatment operates. While the computation of all of the above in reality is infeasible with common GIS software, I illustrate how it can be done in a time-saving and replicable way with the functionalities of the `SpatialRDD` package.

I also clarify some misconceptions when it comes to coordinate references systems - an aspect that usually gets neglected in applied research but deserves more attention. Researchers have to make a choice whether they want to work with angular coordinates (longitude and latitude) or whether they want to map these angles onto $\mathbb{R}^2$ with a projection system. I will argue that the latter is preferred for a variety of reasons. If one decides to work with angles, however, they have to make sure that computations are actually being carried out on a sphere. Conventional GIS software, when provided with spherical coordinates, "under the hood" typically maps these onto a plane by simply treating the two angles as x- and y-coordinates (Plate carrée, the equidistant cylindrical projection) [see e.g., Chrisman, 2017, Tobler, 2002]. This leads to incorrect geometric operations and wrong distance calculations, the degree of bias increasing with distance to the equator[1].

Furthermore, the paper presents and implements an approach to easily shift and move borders, making robustness checks and replication of spatial RDDs more transparent. I further extend it to allow for the creation of random borders. This will be useful more generally for research that relies on moving and shifting geographic features for identification or robustness checks. It can also be easily extended towards classic multi-cutoff RDD applications by just moving the coordinates into a different coordinate system. To give an example from the education literature, one could replace the x-y coordinates in space of a projected system simply with the range of test scores for two separate entrance exams - allowing to work with lines and polygons in $\mathbb{R}^2$.

Specifically, I use the functionalities to randomly create (counterfactual) borders and introduce randomization inference to the spatial RDD literature. Distribution free inference provides an alternative to conventional hypothesis testing that might be appealing in this context due to looming concerns of spatial dependence of potentially unknown form.

Notable early papers that exploited spatial discontinuities for identification are Holmes [1998], Black [1999], and Bayer et al. [2007]. Dell [2010] became a classic reference in the Economics literature, using insights from trend-

[1] R is an exception, where calculations are being carried out on the ellipsoid when geographic coordinates are provided since the implementation of s2 [Dunnington et al., 2023] when `sf` is used [Pebesma, 2018].

surface modelling to flexibly control for space by including polynomials in the x-y coordinates as additional regressors [Legendre, 1993, Schabenberger and Gotway, 2004]. Even though the strong RD identifying restrictions often fail in geographic contexts, most importantly due to sorting, the number of applications keeps growing as the availability of disaggregated geospatial data is steadily increasing. For example, scholars have used discontinuities in land use regulations [Turner et al., 2014] or urban planning [Michaels et al., 2021], while others link historical discontinuities with modern-day data to establish persistence [e.g., Ambrus et al., 2020, Fontana et al., 2023]. All approaches and suggestions discussed in the paper naturally extend to applications where the adoption of treatment status is regarded as fuzzy [e.g., Basten and Betz, 2013] or where the sorting assumption is violated [e.g., the "border contrast" design of Eugster et al., 2017].

When it comes to the randomization of borders, the functions of `SpatialRDD` make it possible to implement tests without distributional assumptions to carry out inference with geographic vector data (lines and polygons). These can potentially be useful for many applications in the social sciences, where randomization inference [e.g., Imbens and Rubin, 2015, Cattaneo et al., 2015, Canay et al., 2017, Borusyak and Hull, 2020, Alvarez et al., 2022, MacKinnon et al., 2023] and spatial "placebo-type" inference [e.g. Donaldson, 2018, Dell and Olken, 2020] have experienced a recent surge in popularity. Randomization inference with spatial boundaries aligns with a longer tradition in spatial statistics, where permutation based inference plays a central role already for a long time [e.g., for the global Moran's I Cliff and Ord, 1981; the Local Indicators of Spatial Association (LISA) Anselin, 1995; Ripley's K Diggle et al., 2000; Hahn, 2012, among others]. Broadly speaking, the randomization inference approach can also be viewed through the lens of a growing literature in causal inference that focuses on the assignment process of observed exogenous shocks [e.g., Borusyak and Hull, 2020, Shaikh and Toulis, 2021].

Section 2 quickly reviews the conventional RD design and Section 3 discusses it in the context of geographic applications. In Section 3.1 I try to make the coordinate reference systems and spatial projections more salient for researchers. After introducing some additional notation, in Section 3.4 we will explore RD designs where the two dimensions are explicitly accounted for. Section 4 discusses Poisson line processes and introduces randomization inference with borders. In Section 6 I will simulate different (spatial) data generating processes (DGPs) in order to illustrate all of the above.

2 The Regression Discontinuity Design

An RD design has three core components: a score, a cutoff, and a treatment with known assignment. The score is often referred to as running- or forcing variable and for every unit of observation $i = 1, 2, ..., n$ which has a score above the cutoff value, we say that treatment was assigned. Formally, we can define the treatment indicator, D_i, as $D_i = D(X_i) = \mathbf{1}(X_i \geq c)$, meaning that a unit is treated if the score exceeds the ex-ante known threshold c. In addition, each unit i has two potential outcomes, $Y_i(0)$ and $Y_i(1)$, where the former is the value of the outcome variable in the absence of treatment and the latter having received treatment. This can be written as

$$Y_i = f(X_i) + \varepsilon_i,$$

where $f(X_i) = (1 - D_i) \cdot Y_i(0) + D_i \cdot Y_i(1)$ and ε_i is an i.i.d. error term with an expected value of 0. The outcomes Y_i are called potential because even though every unit i is assumed to have both, in reality we will only observe one of the two [the so-called "fundamental problem of causal inference", Holland, 1986] - depending on whether treatment was assigned or not. This is why it is often said that causal inference is inherently a missing data problem [e.g. Ding and Li, 2018]. For the so-called sharp RD design, it has been shown that if the regression functions $\mathbb{E}[Y_i(0) \mid X_i = x]$ and $\mathbb{E}[Y_i(1) \mid X_i = x]$ are continuous at the cutoff, $x = c$, one can obtain valid estimates of the local average treatment effect:

$$\tau_{\text{RD}} = \mathbb{E}\left[Y_i(1) - Y_i(0) \mid X = c\right] = \lim_{x \downarrow c} \mathbb{E}\left[Y_i \mid X_i = x\right] - \lim_{x \uparrow c} \mathbb{E}\left[Y_i \mid X_i = x\right].$$

The idea was first formalized by Hahn et al. [2001]. An accessible and extensive treatment of the many methodological innovations in the RD literature over the last years can be found in Cattaneo et al. [2019]. See Cunningham [2021] for a very intuitive textbook introduction to the method.

In an RD design, the assumption of overlap is violated since it is not possible to observe units with either $D = 0$ or $D = 1$ for a given value of the forcing variable X. To compensate for the failure of the overlap condition, a continuity assumption is required. In other words, we cannot observe treatment and non-treatment for the same value of X, but we can observe the two outcomes for values of X around the cutoff point that are arbitrarily close to each other. A central goal of empirical RD analysis is therefore to adequately perform (local) extrapolation in order to compare control and treatment units [Cattaneo et al., 2019].

In order to reduce the bias of such an extrapolation exercise, estimation is not carried out globally, but locally within a small bandwidth h around the cutoff. The practical challenge is then to come up with a heuristic for how small or large to make this bandwidth. Generally speaking, the preferred method for boundary estimation is local linear regression, separately on either side of the cutoff and taking the difference between the adjoining estimated endpoints. This is the same as estimating the conditional expectations function nonparametrically, allowing for a discontinuity [Hansen, 2022, Ch19]. The usage of polynomials should be discouraged because series estimators have high variance at the boundary [see e.g., Gelman and Imbens, 2019, Hansen, 2022]. For efficient estimation at boundary points, a Triangular kernel is recommended[2]. Several heuristics for data driven bandwidth selection have been proposed [e.g., Imbens and Kalyanaraman, 2012, Calonico et al., 2014]. Cattaneo et al. [2019] offer a comprehensive discussion of virtually all approaches and their recommendations are implemented in the package `rdrobust` [Calonico et al., 2014, 2015].

[2]The Epanechnikov and Gaussian have similar efficiencies, however, and some authors have made a case for the Rectangular kernel (resulting in an efficiency loss of about 3% in root AMSE)[Hansen, 2022, Ch19.10]

2.1 A different approach to estimation

Rather than estimating a regression function on either side of the cutoff and taking their difference when $x = c$, one can obtain the same point estimate by fitting a single linear regression that includes an interaction term between the treatment indicator and the score. These approaches are algebraically equivalent [see e.g. Cattaneo et al., 2019, Hansen, 2022]. Crucially, for the correct interpretation of β_1, the score has to be re-centered such that the cutoff occurs at 0 and the following standard regression estimated for the subsample of observations when $|X - c| \leq h$:

$$Y_i = \beta_0 + \beta_1 \, D_i + \beta_2 \, (X_i - c) + \beta_3 \, D_i \times (X_i - c) + \varepsilon_i. \tag{1}$$

In this specification, β_1 is simply the effect of D_i on Y_i when the re-centered score is 0, that is, at the cutoff - which is exactly our RD parameter of interest. Kernel weights have to be computed ex-ante with the respective formula and can simply be plugged in to obtain weighted least squares estimates. While it is computationally more convenient to incorporate Equation 1 as an OLS regression, one still has to resort to RD software for the bandwidth computations (the excellent `rdrobust` should be the first choice [Calonico et al., 2015]).

3 The Spatial Regression Discontinuity Design

This section illustrates how the conventional RD framework can be applied in geographic settings. In general there are two ways of estimation that are commonly encountered in applied research. Both of them are treating space in a uni-dimensional fashion by computing the shortest distance to the geographic cutoff for every point, stacking all observations, and then use the shortest distance as a one-dimensional score.

The first commonly encountered approach regresses the outcome within a bandwidth on a treatment indicator, a set of boundary segment categories, control variables, and sometimes (polynomial) controls for the position in space - similar to what is done in trend surface analysis. The second uni-dimensional approach applies the non-parametric estimation outlined above, often including fixed effects and controls to increase precision.

In Section 6, I illustrate with simulations that the former specifications - albeit working well in conventional contexts - generally do not recover the (local) average treatment effect. Their bias increases with the degree of the spatial trend of the outcome variable. The non-parametric procedure outlined above, instead, estimates the correct effect size. Its equivalent, allowing for both a shift in intercepts and slopes, can be conveniently estimated with OLS and the appropriate interaction term - as outlined above.

A drawback of these uni-dimensional approaches - even though they typically utilize segment fixed effects as an appropriate way to ensure only units close to each-other are compared - is that they mask potential heterogeneity. As a further step in Section 3.4 I will illustrate a way to get at this heterogeneity through a straightforward approach [Imbens and Zajonc, 2011, Zajonc, 2012, Keele and Titiunik, 2015]. In Section 8, we will also see that this approach can - at least in theory - be used to trace out a whole treatment effect curve alongside the RD boundary. In practice,

however, this can be prove to be infeasible due to a sparsity of observations close to the cutoff that researchers often encounter.

3.1 Some Notes on Choosing a Coordinate System

In general, there are a few important considerations for choosing a coordinate references system (CRS) when working with spatial data. The two options at a high level are to either work with geographic - sometimes referred to as ellipsoidal - coordinates (expressed as degrees latitude and longitude, pointing to locations on a sphere approximating the Earth's shape) or projected coordinates (on a flat, two-dimensional space)[3].

Because of the way conventional GIS software typically handles computations with (unprojected) data on the sphere, it is generally recommended to transform the data using a (local) projection system, which implies working with distances in meters and geographic operations on the plane[4].

When location data in terms of angular units (represented by longitude and latitude in a geographic coordinate space) are provided, they need to have an attached datum in order to uniquely describe points on the Earth's surface. WGS84, a "one-size-fits" all global datum, is usually the most common choice and often implicitly assumed when no specific information is provided[5].

In order to project angular data, one has to define a set of rules to translate the curved surface of the Earth onto a two-dimensional plane. Each projected CRS is a combination of an underlying geographic CRS and a projection that tells the map how to distort the earth onto a flat surface. A natural choice for a study relying on a spatial RDD would be the Universal Transverse Mercator (UTM) projection. It divides the globe into 60 zones and one has to find the right one for the specific study area at hand. UTM is conformal and thus preserves angles and shapes across small regions. Distances and areas, however, will be distorted. Projecting ellipsoidal coordinates always comes with a trade-off and means that shapes, directions, areas, or even all three, are distorted [Iliffe and Lott, 2008]. Since most spatial RDDs are carried out over relatively confined geographic spaces, this is less of a concern if an appropriate projection has

[3] Longitude is the location in the East-West direction in angular distance from the Prime Meridian plane. Latitude is the angular distance North or South of the equatorial plane. Angular longitude and latitude coordinates may be expressed in degrees, minutes, and seconds, or in decimal degrees. An arcminute is 1/60th of a degree and an arcsecond is 1/3600th of a degree. For example, 40 degrees and 30 minutes is the same as 40.5 degrees. For reference, a degree is roughly 111 kilometers (or 69 miles) long at the equator. In a projected CRS we simply can work with xy coordinates in the Euclidean plane, $\mathbb{R}^2$.

[4] Most GIS software does not carry out calculations on the ellipsoid when geographic coordinates are provided but models the Earth with flat map projections [see e.g., Chrisman, 2017]. Usually this is done by treating longitude and latitude as the xy coordinates on a flat space (Plate Carrée, also called the equirectangular projection), leading to wrong distances and geometric operations [see e.g. Bivand and Pebesma, 2023, for a discussion]. Even though many of the issues when it comes to computations with data that are declared in an ellipsoidal CRS have been largely resolved in R due to the implementation of s2 [Dunnington et al., 2023, Pebesma, 2018], I still recommend working in a projection because many operations and packages still only work in $\mathbb{R}^2$ - even in R.

[5] Often it is being said that "coordinates are in WGS84" which is not entirely correct and refers to a situation where long/lat is relative to the WGS84 datum. It lays out a set of rules about the shape of the Earth, to which point of the Earth the origin is associated, and how it is directed. WGS84 is often referred to as a (geocentric) geographic CRS, which is unambiguously determined only by its corresponding EPSG code, `4326`. In this special case, WGS84 refers to both the ellipsoid/spheroid and the datum at the same time. As the shape of the Earth is not a perfect ellipsoid, several ellipsoids with different shape parameters and bound to the Earth in different ways are being used. Precision can be increased for specific areas by choosing a local datum, which shifts the ellipsoidal surface to align with the surface of a particular location. For example, the Australian Geodetic Datum 1984 is designed to fit the Earth around Australia whereas `NAD27` increases precision when working on the continental US and `NAD83` (`EPSG:3435`) is very specific to the state of Illinois alone. It might be advisable to change the datum according to the area the data is located in order to ensure higher precision [see, e.g., Ince et al., 2019; or Lovelace et al., 2020, chapter 2.4, for discussions].

been selected. Researchers can get assistance in choosing an appropriate CRS on CRS-explorer.proj.org or with the R package `crsuggest` [Walker, 2022]. Comprehensive discussions of coordinate systems and map projections can be found in Maling [1992], Arlinghaus and Kerski [2013], Lovelace et al. [2020], Bivand and Pebesma [2023] and many others.

3.2 Some Notation

Every observation i receives a description of its location in space as an additional attribute. $\mathbf{x_i} \in \mathcal{X}$ is a vector of length two and its elements, $(x_{1i}, x_{2i})'$, refer to the x- and y-coordinates in $\mathbb{R}^2$ in case the data are projected or, in case of unprojected data, to the angles - longitude and latitude - pointing to locations on a sphere, $\mathbb{S}^2$.

To proceed further, we now need to introduce geometries and formalize the required operations on them. In addition to the already described points, we will need lines and polygons. To describe spatial relations of two geometries ("binary predicates"), we will follow a well-established approach in geospatial topology by Egenhofer and Franzosa [1991]. Let $\mathcal{P}^t$ and $\mathcal{P}^c$ be the treatment and control polygons which fully tessellate the study area (usually a rectangular bounding box). All points that are located in $\mathcal{P}^t$ are part of the "treatment group", $\mathcal{X} \cap \mathcal{P}^t = \mathcal{X}^t$, and all points that are located inside $\mathcal{P}^c$ are thus part of the "control group", $\mathcal{X} \cap \mathcal{P}^c = \mathcal{X}^c$. We obtain the RD boundary by intersecting the boundaries of the two polygons, $\mathcal{B} = B(\mathcal{P}^t) \cap B(\mathcal{P}^c)$, where $\mathcal{B}$ is a line with $dim(\mathcal{B}) = 1$. Depending on whether a researcher decides to work with Cartesian coordinates in a projected coordinate reference system (CRS) or with angles - longitude and latitude - in an unprojected ("geographic") CRS, the geometries are defined either in $\mathbb{R}^2$ or on a sphere, $\mathbb{S}^2$. Treatment status is assigned if a point lies within the treatment polygon:

$$D_i = \mathbb{1}\{\mathbf{x_i} \cap \mathcal{P}^t\} = \mathbb{1}\{\mathbf{x_i} \in \mathcal{X}^t\}.$$

3.3 Estimation of Uni-Dimensional Specifications

The most intuitive and widely adopted way to carry out a spatial RD estimation is by treating the distance to the RD boundary as a uni-dimensional score, just as in the generic setting in the previous section. This is what Cattaneo et al. [2023] call the estimation of normalized-and-pooled treatment effects. To obtain the score we simply have to compute the shortest distance to the line $\mathcal{B}$ for every observation i:

$$dist_i = d(\mathbf{x_i}, \mathcal{B}),$$

where in the case of projected data with Cartesian coordinates this is simply the Euclidean distance. When the coordinates refer to angles, it is the spherical distance. The cutoff, c, is then naturally when $dist = 0$. Finally, we have to make the distance to the cutoff of one of the two groups negative by just multiplying all distances of either the treated or the control units by -1.

RD boundaries in real-world applications can be of substantial length. When estimating Equation 1, this could lead to some undesirable comparisons in practice. For example, units can have a very small score and thus be very close to

the uni-dimensional cutoff, but in reality be geographically very far away from each other. To ensure that we compare only units in close proximity, we can first split the cutoff into several different segments, $\mathcal{B} = \bigcup\{\mathcal{B}_s\}_{s=1}^{S}$. Then, for every point $\mathbf{x_i}$, we compute which segment is closest, $\underset{s}{\operatorname{argmin}}\, d(\mathbf{x_i}, \mathcal{B}_s)$. The resulting categorical variable is then used to create border segment fixed effects. These ensure that only within-segment variation is exploited and thus rule out comparisons of units that are geographically far away from each other. In practice, it is important to visualize these border segments in a map for full transparency. The estimating equation can then be changed to:

$$\begin{aligned} y_i = \beta_0 + \beta_1\, D_i + \beta_2\, dist_i + \beta_3\, D_i \times dist_i \\ + \sum_{s=1}^{S} \gamma_b\, seg_i^s + \mathbf{v_i'}\, \phi + \varepsilon_i, \end{aligned} \tag{2}$$

where the variable seg^s equals one if unit i is closest to segment $\mathcal{B}_s$ and zero otherwise. In addition, one can add a set of control variables, $\mathbf{v_i'}$. These controls, however, should not be used to correct for a discontinuity in an important pre-treatment covariate - i.e., to restore the validity of the desing [see, e.g., Cattaneo et al., 2019].

In practice, researchers often estimate an OLS specification without the interaction term or include a polynomial in location x_{1i} and x_{2i} as a control variable to the above regression[6]. In a specification where the interaction between $dist_i$ and the treatment indicator D_i is included, this addition changes little. And, more importantly, in a specification where the interaction is replaced with the polynomial in location, it can lead to biased estimates when, for example, the dependent variable exhibits a spatial trend.

In section Section 6.3 I will illustrate that it is paramount to allow for different slopes on both sides of the cutoff, i.e. to estimate Equation 1. With simulations I show that running an OLS regression of the outcome on treatment status within a small bandwidth or controlling for space with polynomials in xy-coordinates can recover the correct effect size in common situations. However, when there are for example spatial trends in the data - which are often to be expected - these estimators fall short of recovering the RD estimate. I will illustrate one situation where they report an effect even if there is none, and a situation where they fail to detect an effect even if there is one in the underlying simulated data. It should therefore be encouraged to only rely on estimating regressions as outlined in Equation 1 or Equation 2 when it comes to uni-dimensional specifications.

Triangular kernel weight in practice, show formula. Kernel choice has little impact in practice; hence some scholars suggest to use a rectangular kernels can be used for simplicity and convenience [e.g., Lee and Lemieux, 2010].

3.4 The Spatial RDD as Two-Dimensional Design

A spatial RD design is a special case of a multi-score RD where treatment assignment is determined based on the x-y position in space and the discontinuity is represented by a geographic boundary. A key difference is that the treatment assignment rule cannot be based on multiple scalar cutoff rules, e.g., being above the treshhold simultaneously in two

[6]For example, for a second order polynomial this would entail adding $\alpha_1 * x_{1i} + \alpha_2 * x_{2i} + \alpha_3 x_{1i} \times x_{2i} + \alpha_4 x_{1i}^2 + \alpha_5 x_{2i}^2$ to the regression - very much in line with the literature on trend-surface modelling [see, e.g., Legendre, 1993]

entrance exams, because treatment polygons in practice are rarely convex. In practice, so far most researchers abstract from the two-dimensionality and collapse the assignment rule into a one-dimensional score by computing the shortest distance to the cutoff for every unit i. This approach was illustrated in the previous section. We will see in Section 6 that it can recover the average effect at the cutoff.

Border discontinuity designs, however, allow to identify a wider set of parameters by estimating an effect along the treatment boundary[7]. This has already been suggested by Holmes [1998] and was formalized by Imbens and Zajonc [2011] and Zajonc [2012]. I build on this approach, illustrate its workings with Monte Carlo simulations, and implement it into (geo-)statistical software for easy replication [Lehner, 2020].

Specifically, for every single point $\mathbf{b_b} \in \mathcal{B}$ alongside the boundary we can estimate a treatment effect, $\tau(\mathbf{b_b})$. Every boundary point is defined by two elements, $(b_{1b}, b_{2b})'$, referring to the x- and y-coordinates in $\mathbb{R}^2$ in case the data are projected or, in case of unprojected data, to the angles - longitude and latitude - pointing to locations on a sphere, $\mathbb{S}^2$. Additionally, we need a two-dimensional score, defined in the same respective space - $\mathbf{X}$. The RD treatment effect at every point on the boundary can then be written as

$$\tau_{\text{BRD}}(\mathbf{b_b}) = \mathbb{E}\left[Y_i(1) - Y_i(0) \mid \mathbf{X} = \mathbf{b}\right] \tag{3}$$

and a boundary-wide treatment effect, τ_{BRD}, can be obtained by taking the expectation over the whole boundary.

Let $N^t_{h_b}(\mathbf{b_b}) \equiv N_{h_b}(\mathbf{b_b}) \cap \mathcal{P}^t$ be the neighborhood of points that receive treatment and $N^c_{h_b}(\mathbf{b_b}) \equiv N_{h_b}(\mathbf{b_b}) \cap \mathcal{P}^c$ the neighborhood of points that are in the control group.

Using the approach outlined for the uni-dimensional case, we can then estimate a treatment effect by (locally) extrapolating at every boundary point:

$$\tau_{\text{BRD}}(\mathbf{b_b}) = \lim_{\mathbf{X} \to \mathbf{b_b}} \mathbb{E}\left[Y_i \mid \mathbf{X} \in N^t_{h_b}(\mathbf{b_b})\right] - \lim_{\mathbf{X} \to \mathbf{b_b}} \mathbb{E}\left[Y_i \mid \mathbf{X} \in N^c_{h_b}(\mathbf{b_b})\right]. \tag{4}$$

In addition to the continuity assumption, we also have to impose a boundary positivity assumption [Imbens and Zajonc, 2011], ensuring that both treated and untreated units exist for every boundary point. In practice, one will often be forced to take this one step further and make sure to select boundary points with a sufficiently large number of observations in their vicinity in order to avoid obtaining overly imprecise estimates[8]. For identification, we can rely on the standard RD result from Hahn et al. [2001]. Estimating Equation 4 as outlined in Section 3.3 ensures that the lines have different slopes on both sides of the cutoff for every boundary point and identify the local average boundary effect correctly.

In order to obtain the boundary-wide treatment effect, instead of integrating over the whole boundary, we can approximate it with

[7] See – for applications for the general case of discontinuity designs with multiple scores

[8] As pointed out by Cattaneo et al. [2023], it is important to check the density of the distance measure because a very low number of observations with distances near zero will result in excessive extrapolation in RD estimation.

$$\hat{\tau}_{\text{BRD}} = \frac{\sum_{b=1}^{B} \hat{\tau}_{\text{BRD}}(\mathbf{b_b})\, w_b}{\sum_{b=1}^{B} w_b}. \tag{5}$$

As the number of boundary points increases, this will approach an integral. In the simulation section we will see that a modest number of points suffices to estimate $\hat{\tau}_{\text{BRD}}$. Zajonc [2012] proposes to use the estimated bivariate density of coordinate vectors for every boundary point as weights. However, the varying bandwidths for every boundary point imply that the number of observations included in the local estimation is changing flexibly. Therefore, a more intuitive and transparent approach would be to weigh boundary point estimates by the resulting sample size that was chosen by the bandwidth selection algorithm [Calonico et al., 2015] for estimation. This will ensure that more precise estimates receive a higher weight in the computation of the boundary-wide effect.

Inference on the boundary-wide effect can be done using two straightforward methods. First, plugging into the formula for the nonparametric delta method as proposed by Zajonc [2012]. Second, simply rely on resampling methods to calculate the standard errors.

Finally, as pointed out in Tristan Zajonc's thesis [Zajonc, 2012], explicitly integrating/summing estimates over the entire boundary will often not differ from reducing the problem to one dimension and using distance to the nearest boundary as a scalar forcing variable - the approach we illustrated in Section 3.3. However, I argue that estimating several RD coefficients alongside the boundary is useful for two main reasons. First, it helps researchers understand potential heterogeneity that can occur in terms of treatment effects which, in turn, might give useful insights for the research question at hand. For example, it could be that an RD effect is driven by only a few segments alongside the border while there is no jump in the outcome variable in other parts. Understanding why a jump only occurs on one part of the cutoff and not on the other can help in narrowing down potential channels through which a treatment operates and whether there are mediating variables. Another motivation for considering the two-dimensional estimation is for visual reasons alone. One of the main reasons why regression discontinuity in general is a very popular and intuitive estimation method is that it can be visualized in a very straightforward and simple way ("visual RD", Angrist and Pischke [2014]). Even though in spatial implementations of RDDs there is a second dimension, that it, space, it gets usually thrown out.

4 Randomization Inference for Boundary Designs

As noted in the introduction, tests based on randomization inference have experienced a surge in popularity in recent years. The idea dates back to Fisher [1935] and was first formalized by Pitman [1937; see David, 2008, for a brief history] and offers a distribution-free way to perform hypothesis tests. A formal introduction to the concept can be found in Lehmann and Romano [2005], Chapter 15.2. The idea of randomization inference boils down to comparing a test statistic that is actually observed with a set of test statistics that might have been observed if treatment had been assigned differently. This procedure works for any choice of test statistic, $T(\cdot)$, and any outcome-generating process.

The empirical distribution of the values of a given statistic recomputed over transformations of the data then serve as a null distribution and we can approximate a p-value:

$$\hat{p}_{RI} = R^{-1} \sum_{r=1}^{R} \mathbf{1}\{T(\mathcal{B}^r, Y) > T(\mathcal{B}, Y)\}.$$

In other words, we count the number of times the counterfactual test statistic under randomization was larger than the observed original test statistic and divide by the number of randomizations. Note that the randomness of the test statistic comes solely from the randomization of the RD boundary, $\mathcal{B}^r$, which in turn determines the counterfactual treatment assignment.

It is only an approximation, because in practice the number of all potential permutations is too large, and it is computationally infeasible to enumerate all possible values of the treatment vector. We therefore approximate $\hat{p}_{RI}$ by Monte Carlo.

The fact that treatment follows a spatially contiguous assignment mechanism in the spatial RD framework complicates matters for practical implementation. Randomly re-assigning treatment status - like in a "conventional" RI exercise - will thus not yield a valid counterfactual randomization distribution because the treatment was not independently assigned. When there is a spatial assignment mechanism, the transformations of the data over which the statistic of interest has to be recomputed in order to serve as a null distribution has to follow a valid assignment process. RI procedures are valid only when the distribution of the test statistic is invariant to the realization of the re-randomizations across permutations of assigned treatments [Lehmann and Romano, 2005, Section 15.2]. It is therefore important to incorporate all available information about treatment assignment in conducting the re-randomization [**?**].

To be more specific, we randomly draw a counterfactual discontinuity in space, $\mathcal{B}^r$, use it to create two polygons, $\mathcal{P}^{tr}$ and $\mathcal{P}^{cr}$ that tesselate the bounding box of the study area, assign treatment and control status based on these polygons, and then estimate an RD specification and store $T(\cdot)$, our test statistic of interest. We repeat this procedure $|R|$ times and if the original value of the test statistic is far in the tails of the simulated null distribution, we then have grounds to reject the null that $\beta_1 = 0$.

Following MacKinnon and Webb [2020], in this note we will focus on randomization inference based on two test statistics, RI-β and RI-t. The former uses the conventional test statistic τ_{RD} that has been introduced in Section 3, or rather its regression equivalent $\hat{\beta}_1$, and RI-t uses the robust $t-$statistic. It has been shown that the latter approach is valid under stronger assumptions in finite samples, and also asymptotically valid under weaker assumptions [e.g. Janssen, 1997, Chung and Romano [2013], Bugni et al. [2018], Wu and Ding [2021]].

This procedure, in a nutshell, requires the following steps:

1. Create a random discontinuity
2. Assign all datapoints to either the randomization treatment or control group

3. Compute the distance of all datapoints to this randomization discontinuity to obtain a randomization inference score
4. Carry out the RD estimation

Repeat this procedure R times. The R package `SpatialRDD` [Lehner, 2023] allows researchers to unify all of these tasks in a replicable and streamlined way. The crucial part of this exercise is the creation of a random line, $\mathcal{B}^r$, in every iteration $r \in \{1, \dots, R\}$ within the bounding box of a given research study. Each random line is then used to create the two randomization polygons, $\mathcal{P}^{tr}$ and $\mathcal{P}^{cr}$, in order to assign (placebo) treatment and control status.

4.1 Creating Random Lines and Polygons

In this note I propose to rely on a poisson line process due to its computational simplicity and because it can be easily implemented in R with package[9]

In a nutshell, lines can be parametrized by just two parameters, a (perpendicular) distance $r \in (0, \infty)$ from the origin and an angle $\theta \in [0, \pi)$ with the x-axis. Each point (r, θ) can be generated in a so-called representation space from a Poisson process on $(0, \infty) \times [0, \pi)$ with intensity $\frac{1}{2}\lambda \, \mathrm{d}r \, \mathrm{d}\theta$. Each of the generated points in the representation space can be mapped onto a line in the original space. In other words, a point process in the representation space gives a line process in the original space - the bounding box of a study. The number of lines, i.e. the intensity of the process, will be a Poisson random variable. In this context that's not a crucial parameter as long as it leads to the creation of more than one line since we only need to create a single random boundary in every iteration.

More details can be found in texts on poisson processes Kingman [1993] and stochastic geometry more generally Chiu et al. [2013].

In practice, poisson line processes can be simulated in R via `spatstat` [Baddeley and Turner, 2005, Baddeley et al., 2019]. The necessary functions have been adapted and nested into `SpatialRDD` and `SpatialInference` which build upon these and carry out all the needed further steps outlined above, i.e. create randomization polygons and re-assign treatment status[10]. In Section 7 I illustrate how to operationalize all of the above in practice on simulated data.

5 Threats to Identification

The key regression discontinuity identifying assumption is that all relevant factors besides treatment vary smoothly at the cutoff. Balancing checks are the appropriate way to assess the plausibility of this assumption[11]. Baseline covariates are not needed for identification, but they can improve the precision of estimates when included in RD regressions [Lee, 2008, Imbens and Lemieux, 2008].

[9] I am grateful to Roger Bivand for encouraging me to rely on poisson line processes instead of trying create the randomization borders from the original border of the research study. Even though in practice this works and can be implemented with the functionalities of `SpatialRDD`. It is much harder, however, to establish true randomness of these borders.

[10] See the vignette for details

[11] Lee [2008] shows that under certain conditions the distribution of baseline covariates in the RDD must be continuous at the cut-off.

The assumption implies that all possible determinants of the outcome, Y_i, must be a continuous function of distance to the RD boundary. This guarantees that untreated units that are close to the boundary serve as plausible counterfactuals for units that were treated. The inclusion of boundary segment fixed effects ensures that not only units close to each other in terms of distance to the cutoff, but also in terms of distance alongside the cutoff are being compared.

In the context of spatial RDDs, one of the biggest threats for the validity of the design is endogenous sorting around the RD boundary.

5.1 Useful Robustness Checks to Ensure Validity

- illustrating that before the treatment occured, there was no jump in the outcome variable (this is often not possible due to the unavailability of data)
- showing balancedness in key covariates (e.g., geography)
- showing balancedness of placebo outcomes
- shifting borders in different directions and demonstrate there is no effect at placebo cutoffs

6 A Simulation Study to Illustrate Which Estimators Work

In this section, I demonstrate the estimation techniques described above on a set of spatially simulated data. The vector data used is from [Lehner2019a], representing the Indian state of Goa, and projected in the local UTM Zone43 projection system (`EPSG:32643`)[12].

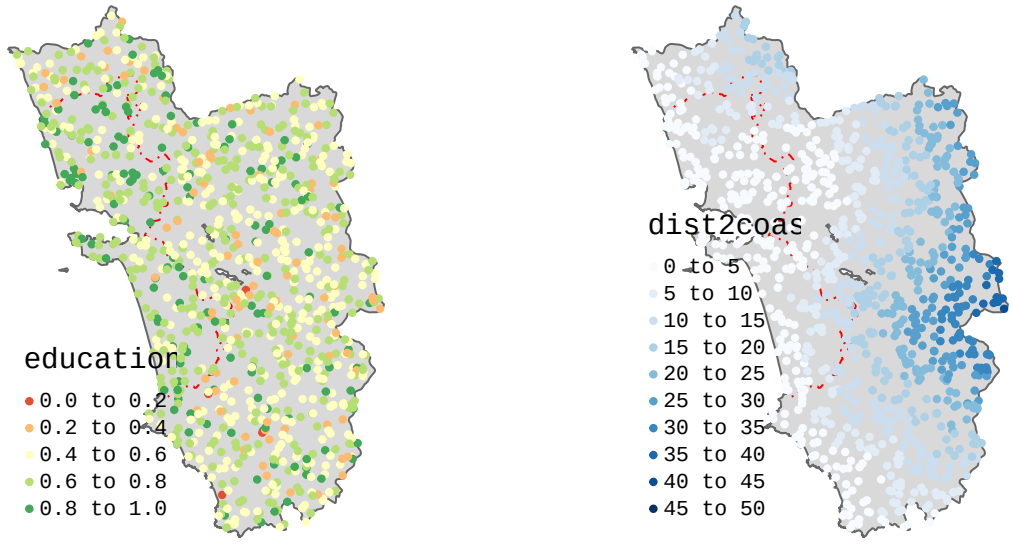
6.1 Simulations

The dataset contains a "treated" polygon, a polygon for the full extent that also defines the bounding box for our study, and a line that represents the discontinuity in space. It comes together with the `SpatialRDD` package. The cutoff - a historical border in Goa, India - is 129 kilometers long and the treatment area is 802 square kilometers large.

I randomly simulate 1,000 datapoints within this geography using R-package `sf` [Pebesma, 2018]. Each point i has thus an assigned location $\mathbf{x_i} \in \mathcal{X}$. Each of these points in space then get assigned four different variables that we are going to use throughout:

1. A variable bound between 0 and 1 with a discontinuity of 0.1: `education` as measured by e.g. literacy rates
2. A variable that exhibits a uniform spatial trend: (euclidean) distance to coastline measured in kilometers (`dist2coast`)
3. The same distance to coast variable with a jump of 5 (km) - artificially introduced at the cutoff (`D_dist2coast`)

[12]The Universal Transverse Mercator (UTM) coordinate system is a set of coordinate reference systems (CRSs) that divides the Earth into 60 longitudinal wedges and 20 latitudinal segments. The transverse Mercator projection used by UTM CRSs is conformal but distorts areas and distances with increasing severity with distance from the center of the UTM zone.



(a) discontinuous: education

(b) spatial trend: distance to coast

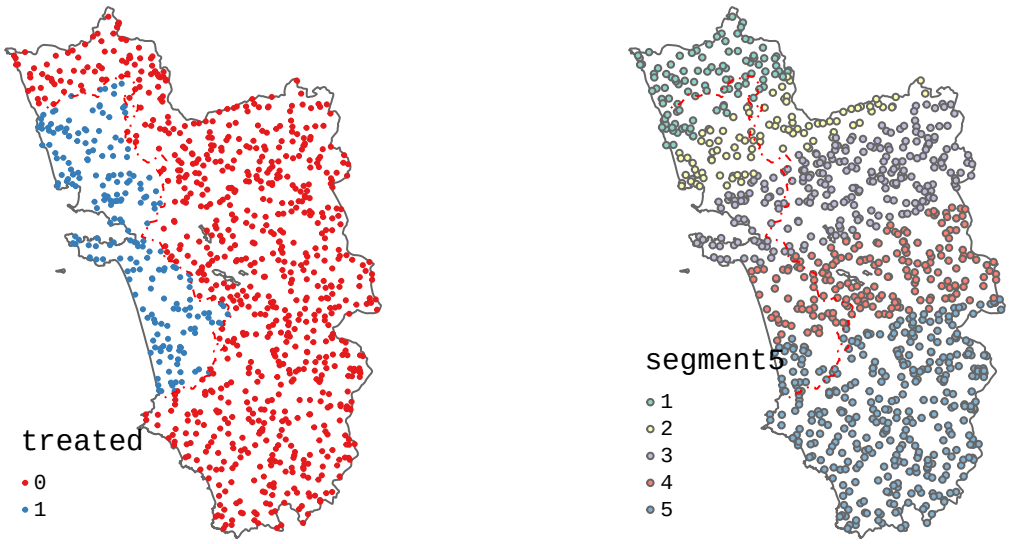
Figure 1: All c. 1000 (spatially) simulated observations and the values for the dependent variable. The red dashed line depicts the discontiunity in space.

4. A variable with a spatial trend that is different for treated and control units (`D_dist2cutoff`) - exhibiting a constructed discontinuity at the cutoff of 1

These variables represent extreme cases of the most common scenarios that researchers face when looking at real world geographic data. Ideally, one would want to encounter scenario 1, with a clean discontinuity and no other disturbances. In reality, however, some type of spatial dependence - in addition to a discontinuity - has to be expected. This is illustrated with an extreme form of spatial autocorrelation by modelling a distance. This distance variable alone is then illustrating the problems that can arise in estimation when a variable exhibits a spatial trend but no discontinuity. Such a data generating process is common in practice, albeit in a less extreme version. For example, many variables of interest for social scientists such as income, crime, or health are functions of first nature geography (e.g., climate or elevation/ruggedness) or other geographic features such as the distance to a capital city or a central business district. This can induce a variable to exhibit some type of spatial decay. The fourth variable introduces a special case where we allow the spatial gradients to differ between treated and control units. This situation serves to illustrate that RD estimation that does not allow for differential slopes can fail to detect a discontinuity in the data. Figure 1 visualizes all observations in space. For our `education` variable, treated observations were assigned a mean of 0.7 by default and received a random draw from a normal distribution with mean zero and a standard deviation of 0.15. The few observations that ended up having values above 1 were assigned 1.0 by hand.

A default set of five border segments has been created for the parametric estimations using the `border_segment()` function. These are later going to be used as fixed effects categories in the regressions, ensuring that only locations that are within the same segment are being compared to each other. In our simulation study, though, the segment indicators/fixed-effects are not going to impact the point estimate of our treatment indicator due to the fact that we

introduced a data generating process (DGP) that is spatially uniform. In this simulation section, we on purpose create data with such a homogeneous and very straightforward structure in order to illustrate the functioning and performance of the different spatial RDD specifications that were outlined earlier. Figure 2 visualizes the treatment assignment next to the border segment categories.



(a) treatment assignment 0/1

(b) fixed effects categories

Figure 2: Visualisation of the 'treated' region and the created border segments (that will be used as fixed effects later).

6.2 Uni-Dimensional Specifications

In this section we are comparing the uni-dimensional estimators that were outlined in Section 2 and Section 3.3 to illustrate that they are able to identify the treatment effect on the simulated discontinuous variables - `education`, `D_dist2coast`, `D_dist2cutoff` - and do not pick up an effect on the variable with only a spatial trend, `dist2coast`. Specifically, this exercise also serves to demonstrate the numerical equivalence between running two local linear regressions (with polynomial of order one) on either side of the cutoff, as, e.g., done by the package `rdrobust`, and by running an OLS regression allowing for different slopes on either side of the cutoff by including an interaction term [see, e.g., Cattaneo et al., 2019, Hansen, 2022, and the expostion in Section 3.3].

Table 1 illustrates these results with a triangular kernel and an `mserd` bandwidth chosen by `rdrobust`. We can see that the point estimates between an OLS regression with a treatment indicator, the running variable centered at the cutoff, and their interaction is identical to the RD estimate from local linear regressions on each side of the cutoff (carried out with `rdrobust`). Specifically, we can see that the estimate for `education` is 0.164, implying a 16.4 percentage point increase in that variable given that it was intended to measure a percentage. Column 2 shows that the estimator correctly did not detect an effect significantly different from zero for the `distance2coast` variable which only exhibits a spatial trend. In the last two columns, we can see that the discontinuity was identified with the correct magnitudes of 5.749 and 1.126, respectively. Figure 7 visualizes these estimates on binnscatter plots.

Table 1: Regression output for one-dimensional specifications with simulated data (education)

	education	dist2coast	D_dist2coast	D_dist2cutoff
I(treated)	0.164	0.749	5.749	1.126
	(0.045)	(1.684)	(1.684)	(0.251)
dist	0.000	−0.001	−0.001	−0.001
	(0.000)	(0.001)	(0.001)	(0.000)
I(treated)*dist	0.000	−0.001	−0.001	0.001
	(0.000)	(0.001)	(0.001)	(0.000)
Observations	213	185	185	204
Moran's I [y]	2.669	38.999	35.834	1.165
Moran's I [resid]	1.053	36.606	36.606	3.270
Mean [y]	0.647	6.915	4.051	1.500
Std.Dev. [y]	0.177	5.429	5.470	1.123

Table 2: RD estimates through rdrobust

	education	dist2coast	D_dist2coast	D_dist2cutoff
I(treated)	0.164	0.749	5.749	1.126
	(0.044)	(1.812)	(1.812)	(0.268)

6.3 Comparison of Uni-Dimensional Estimators

Having illustrated the equivalence of the two RD estimators above. We will juxtapose it to two different ways of estimation that are commonly encountered in research papers. These include either polynomials in the x-y coordinates to control for space or just the indicator variable for treatment alone - in addition to border segment fixed effects and additional control variables. Both of them do not allow the slope to vary on either side of the cutoff. In many situations these perform well and deliver similar point estimates to the RD specification. However, there are situations where the bias of these estimators gets big enough to arrive at wrong conclusions. I illustrate the performance of these estimators on the simulated data described above. For DGPs with spatial trends, there are situations where these estimators conclude there is an effect when there is none, and a where they conclude that there is no effect when in reality there is one.

We will investigate the following specifications:

1. $y_i = \beta_0 + \beta_1\, D_i + \beta_2\, dist_i + \beta_3\, D_i \times dist_i + \varepsilon_i$ ["RD"],
2. $y_i = \beta_0 + \beta_1\, D_i + f(\mathbf{x_i}) + \varepsilon_i$ ["OLS_xy"],[13]
3. $y_i = \beta_0 + \beta_1\, D_i + \varepsilon_i$["OLS"],

all of which will be estimated on a subset of the data within a bandwidth (`mserd`) selected by `rdrobust`. In practice, as illustrated in Section 3.3, researchers add border segment fixed effects and control variables to these specifications. In this section we can omit them for simplicity as they will have no impact due to the way the data was simulated.

Specification 1) is the textbook regression discontinuity specification with a one-dimensional score. Specification 2) is inspired by Dell [2010], who follows the trend surface modelling literature [see, e.g., Legendre, 1993, Schabenberger

[13] We let $f(\mathbf{x_i})$ be a second order polynomial in the geographic coordinates, which is the most commonly seen version in research papers: $\alpha_1 * x_{1i} + \alpha_2 * x_{2i} + \alpha_3 x_{1i} \times x_{2i} + \alpha_4 x_{1i}^2 + \alpha_5 x_{2i}^2$.

Table 3: Comparing the performance of the estimation methods on all four simulated DGPs.

	education			dist2coast			D_dist2coast			D_dist2cutoff		
	RD	OLS_xy	OLS	RD	OLS_xy	OLS	RD	OLS_xy	OLS	RD	OLS_xy	OLS
I(treated)	0.164	0.136	0.142	0.749	2.319	−1.796	5.749	7.319	3.204	1.126	0.711	0.450
	(0.045)	(0.027)	(0.024)	(1.684)	(0.303)	(0.902)	(1.684)	(0.303)	(0.902)	(0.251)	(0.187)	(0.168)
Observations	213	213	213	185	185	185	185	185	185	204	204	204
Moran's I [y]	2.669	2.669	2.669	38.999	38.999	38.999	35.834	35.834	35.834	1.165	1.165	1.165
Moran's I [resid]	1.053	1.780	0.870	36.606	39.527	36.798	36.606	39.527	36.798	3.270	2.253	2.967
Mean [y]	0.647	0.647	0.647	6.915	6.915	6.915	4.051	4.051	4.051	1.500	1.500	1.500
Std.Dev. [y]	0.177	0.177	0.177	5.429	5.429	5.429	5.470	5.470	5.470	1.123	1.123	1.123

Notes: All specifications use a triangular kernel with the respective mserd bandwidth chosen by rdrobust.
[1] It can be seen that 'OLS_xy' and 'OLS' deliver erratic results, sometimes even in the wrong direction.
Text of the second note.

and Gotway, 2004] and smoothly controls for space by including (second-order) polynomials in the x- and y-coordinate for every observation as control variables. Specification 3) simply compares units on opposite sides of the cutoff - within their respective border segments - within a fixed bandwidth. For all specifications an identical `mserd` bandwidth - selected by `rdrobust` - has been used.

Table 3 juxtaposes the previously outlined estimating equations independently for all simulated DGPs. The first column for every variable anchors the correct local average treatment effect estimate at the cutoff discussed in the previous section.

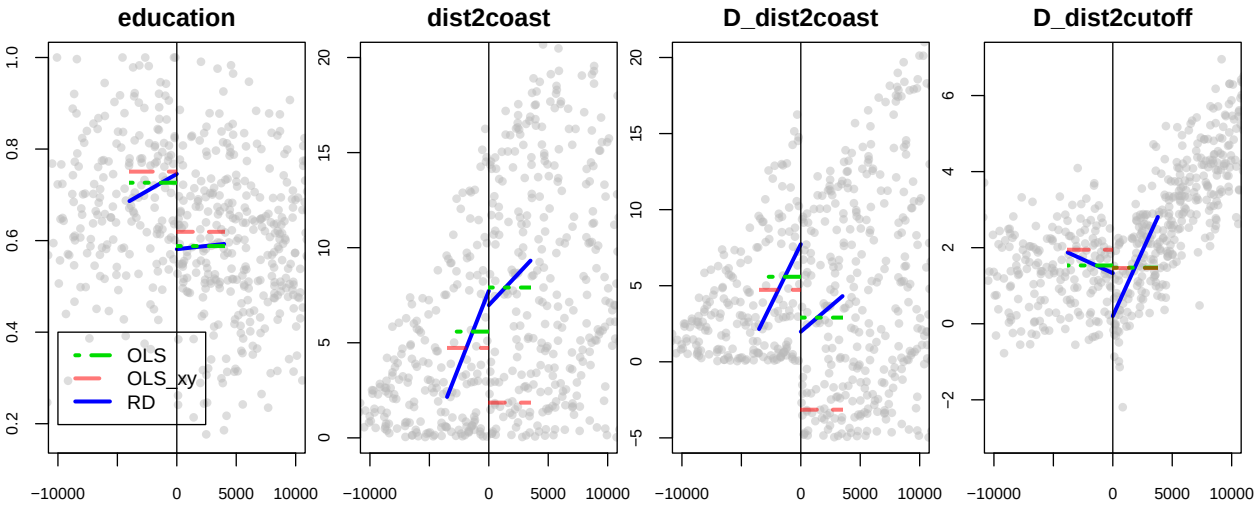
For the conventional scenario with a discontinuity and no other spatial element in the DGP - simulated as `education` - it can be seen that all estimators arrive at comparable conclusions. Albeit the estimates are slightly different, all of the confidence intervals are overlapping.

Block two, where `dist2coast` exhibits a plain spatial trend but no discontinuity, demonstrates that only the conventional RD estimator correctly does not report a discontinuity. Interestingly, `OLS_xy` and `OLS` not only report an RD effect but their estimates also have opposing signs.

In block three, `OLS_xy` and `OLS` again correctly detect a discontinuity at the cutoff, however, the magnitudes differ quite substantially from the simulated discontinuity in the data. Specifically, controlling for the x-y coordinates in space delivers a higher point estimate while the plain `OLS` specification delivers a lower one. The fourth block illustrates that when there is a break in the spatial trend approximately around the cutoff, the conventionally used estimating equations `OLS_xy` and `OLS` almost fail to detect a discontinuity in the data even if there is one.

The bottom line of this exercise is that it is not ex-ante clear in which direction bias goes as it highly depends on the spatial DGP. It can be seen that it is paramount to allow for estimation with different slopes on both sides of the cutoff. Specification `OLS_xy` - inspired by trend surface modelling - does control for space but it does not, however, guarantee correct extrapolation of the conditional expectation function close to the cutoff.

Figure 3 visually compares all different scenarios in the RD scatterplots and illustrates how much the RD point estimates differ across the outlined estimating equations. The plots clearly show that not allowing for different slopes on either side of the cutoff is less of an issue in non-spatial DGPs. Any sort of spatial trend, however, leads to erratic



(a) 'conventional' discontinuity (b) a spatial trend alone (c) spatial trend + discontinuity (d) spatial trend + discontinuity

Figure 3: RD clouds visualized together with the different RD point estimates. They can differ substantially. The underlying code heavily borrows from the excellent rdrobust package.

estimates in specifications that do not allow the conditional expectation function to have different slopes on either side of the RD boundary. The sign and magnitude of the bias is ex-ante ambiguous and depends on the underlying DGP.

7 Randomization Inference in Practice

To carry out distribution free inference for spatial RD designs, in Section 4 I proposed a way to obtain p-values through randomization inference by creating a large number of counterfactual RD boundaries. In this section, I demonstrate its practical implementation in connection with the `SpatialRDD` package.

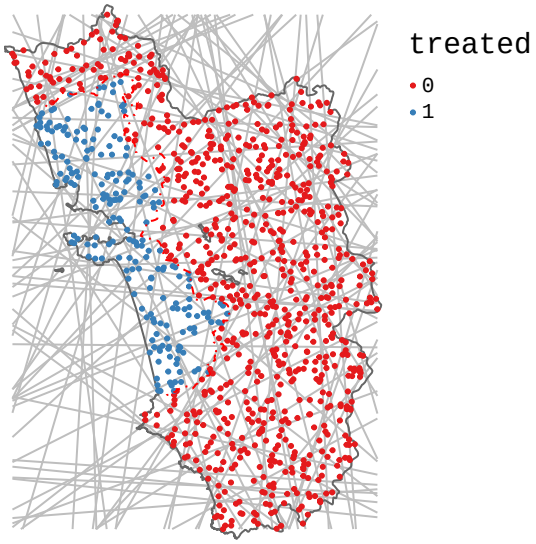


Figure 4: A subset of the random borders that were created with the SpatialRDD package to carry out randomization inference.

The following necessary steps are carried out "under-the-hood" by the function `randinf()` for each iteration, drawing upon other functions in `SpatialRDD` and the well-known R package `spatstat` [Baddeley and Turner, 2005, Baddeley et al., 2019]. For every iteration $r \in \{1, ..., R\}$:

1. Create a random discontinuity from a poisson line process, $\mathcal{B}^r$
2. Use $\mathcal{B}^r$ to create the two corresponding randomization polygons, $\mathcal{P}^{tr}$ and $\mathcal{P}^{cr}$
3. Assign all datapoints to treatment and control. In practice the labeling does not matter as we are looking at absolute values and therefore it is the same if the "randomization treatment effect" is positive or negative
4. Compute the distance of all datapoints to $\mathcal{B}^r$ and obtain a randomization inference RD score
5. Carry out the RD estimation on the uni-dimensional score as illustrated in previous chapters
6. Store the point estimate, β_1^r

In practice, the number of Monte Carlo draws, R, should be several thousand. For the purpose of easier illustration, Figure 4 only shows 100 randomly drawn lines, however. A simulation for the `education` variable leads to a randomization p-value of 0.036. Carrying out this randomization inference procedure on the variable `distance2cutoff` delivers a p-value of 0.758, clearly indicating that the null-hypothesis of there being no discontinuous effect at the cutoff cannot be rejected.

8 A Spatial Monte Carlo Exercise with Two-Dimensional RDDs

In this section I illustrate the two-dimensional approach to RD described in Section 3.4. Specifically, I will use the simulated variable `education` from above to showcase the RD estimation at several boundary points. As stated earlier, the boundary-wide effect obtained by summing over all individual boundary point effects will rarely differ from the estimation when the space is collapsed into a one-dimensional score and estimated with standard RD procedures - including segment fixed effects. However, having estimates alongside the RD boundary allows to visualize the design in space instead of disregarding the second dimension and rely on classic RD plots with only one score. Furthermore, possible heterogeneity can be used to shed further light on the research question at hand.

As a first step, we have to draw a set of boundarypoints from the existing RD border. We choose 30 equally spaced points via `discretise_border()`. With the option `random = TRUE` the points could have been selected at random, but for the purpose of this exposition it would have changed little.

Figure 5 shows the point estimates for every boundarypoint, $\mathbf{b_b}$. Note that some points have been automatically excluded because there were not enough observations in the vicinity of these points in order to guarantee reliable estimation and inference. A good rule of thumb is to have at least 30 observations on either side of the RD cutoff. Also note that many of the point estimates' confidence intervals are overlapping with zero. This is not much of a concern because they are not powered to detect the effect at hand with such a small amount of observations in the surroundings of a boundarypoint.

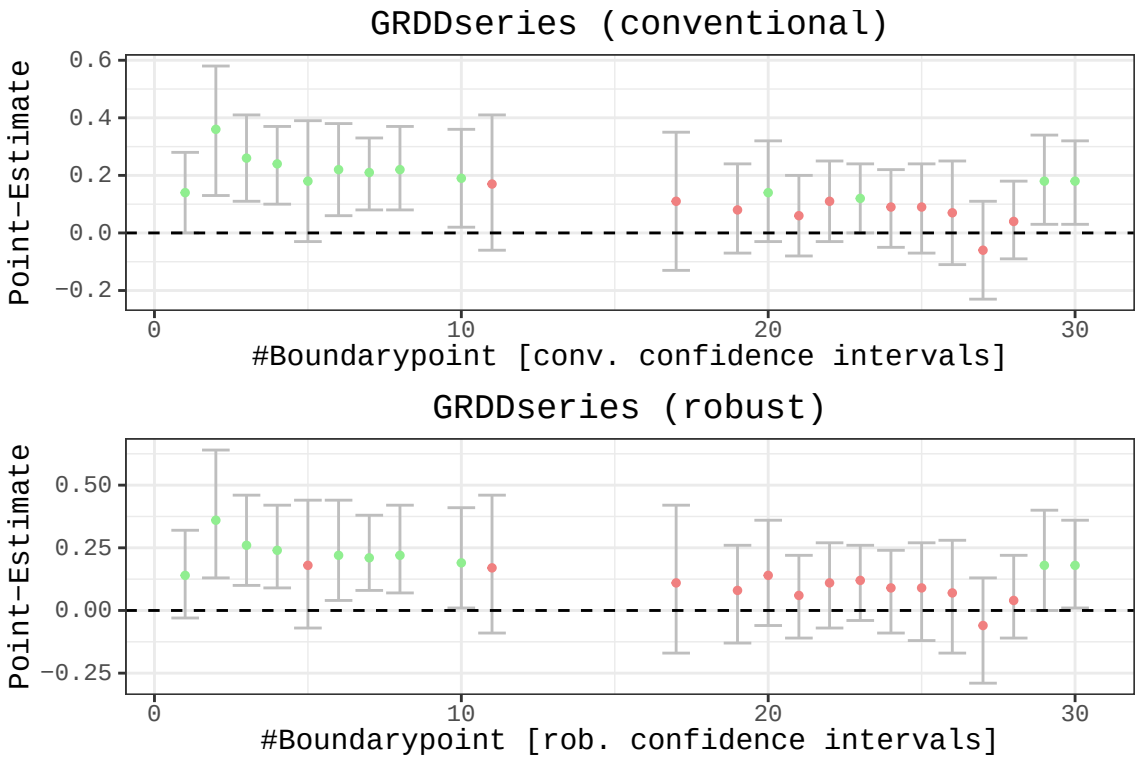


Figure 5: The RD point estimates for every boundary point visualized. As the sample size for many points is quite low, the estimation is often underpowered, leading to wide confidence intervals at times overlapping with zero. This has no bearing, however, on inference for the overall boundary-wide treatment effect.

The overall boundary-wide effect estimated with Equation 5 amounts to 0.14 which is very much in line with the estimate of 0.164 when we collapsed the score into one dimension. Inference on the boundary-wide effect can be done by plugging into the formula for the delta method or by bootstrapping over boundary segments.

8.1 Simulating a Boundary-Wide Treatment Effect Curve

To illustrate that the two-dimensional approach is able to estimate a boundary-wide treatment effect curve I rely on Monte Carlo simulations. Specifically, I simulate only 20 draws of the DGP that created the `education` variable. For every draw I store the boundarypoint estimates and finally approximate the line with a simple LOESS fit.

It can be easily seen from Figure 6 that the effect curve has the right magnitude. Also note that approximately in the center the point estimates get noisier. This is by construction as the boundary gets close to the ocean and therefore naturally limits the amount of observations that can be close to these points. Geographic features limiting the number of observations in certain sections of an RD boundary is often encountered in applied research. Insignificant point estimates in such areas do not necessarily imply that there is no effect at the cutoff. It is just an indication for the effect size not being big enough in order to be picked up with estimation using such a small sample size. In other words, it is a problem of statistical power.

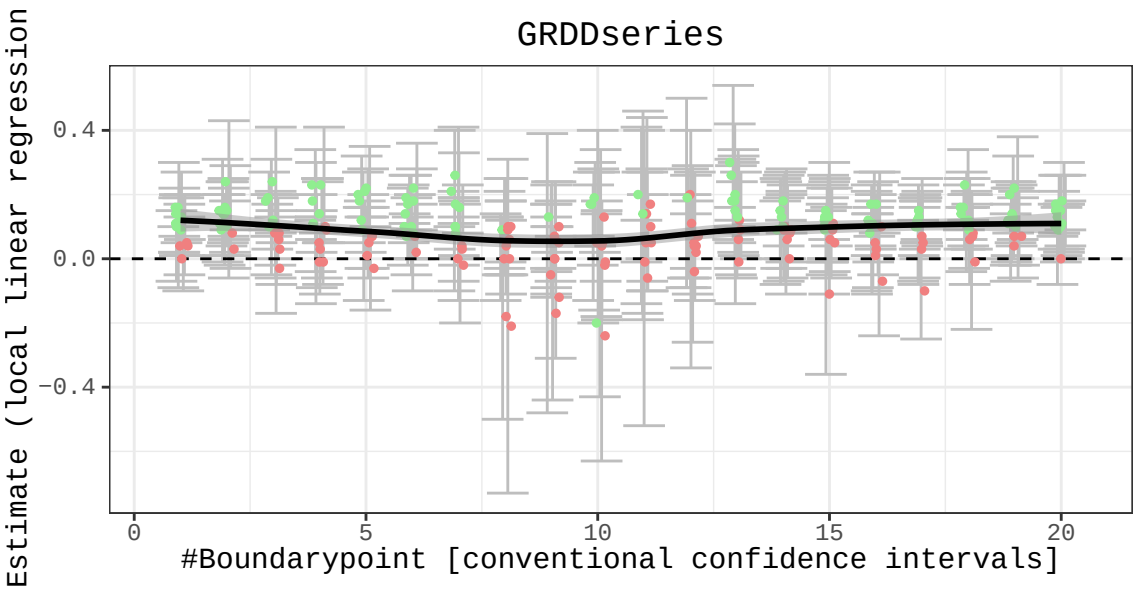


Figure 6: Estimates for all boundary points over many Monte Carlo simulations of the same DGP visualized. Treatment effect curve is drawn as a LOESS fit.

9 Conclusion

In this paper I highlighted the limitations of commonly used estimators used for spatial RDD estimation. I have demonstrated that while they all provide approximately accurate estimates in many scenarios, the performance of some of them deteriorates when the DGP exhibits spatial patterns, leading to biased estimations of the RD effect. This finding underscores the necessity of adjusting the conditional expectation function to accommodate different slopes on either side of the RD boundary, aligning with best practices in the standard RD literature.

Further, by embracing the two-dimensional nature of geographic RDs, I have illustrated the potential to estimate heterogeneous RD effects along the boundary. This approach not only facilitates the incorporation of spatial dimensions into visualizations, enhancing the credibility and interpretability of findings but also opens up avenues for investigating the mechanisms through which discontinuously induced treatments exert their effects.

The `SpatialRDD` R-package, developed as part of this research, incorporates a set of functions which enables researchers to carry out all necessary steps for spatial RD estimation in an easily replicable way, thus circumventing the limitations of conventional GIS software relying on graphical user interfaces. The package also includes all proposed novelties and offers useful features for research outside of boundary RD designs.

Addressing an often overlooked aspect in practice, I have argued for the preferential use of projected coordinate reference systems over angular coordinates for boundary RD designs. This recommendation is based on the fact that most GIS software, when provided with angular coordinates (longitude, latitude), implicitly projects these on a flat space (Plate Carrée), leading to wrong distances and geometric operations. Even within R, where this problem has

been solved and calculations are being carried out on the sphere using s2, I still recommend using projected coordinates because important operations such as shifting lines are only implemented in $\mathbb{R}^2$.

References

Luis Alvarez, Bruno Ferman, and Raoni Oliveira. Randomization Inference Tests for Shift-Share Designs, June 2022. URL http://arxiv.org/abs/2206.00999.

Attila Ambrus, Erica Field, and Robert Gonzalez. Loss in the Time of Cholera: Long-Run Impact of a Disease Epidemic on the Urban Landscape. *American Economic Review*, 110(2):475–525, February 2020. ISSN 0002-8282. doi: 10.1257/aer.20190759. URL https://www.aeaweb.org/articles?id=10.1257/aer.20190759.

Joshua D. Angrist and Jörn-Steffen Pischke. *Mastering 'Metrics: The Path from Cause to Effect*. Princeton University Press, Princeton ; Oxford, with french flaps edition edition, December 2014. ISBN 978-0-691-15284-4.

Luc Anselin. Local Indicators of Spatial Association—LISA. *Geographical Analysis*, 27(2):93–115, 1995. ISSN 1538-4632. doi: 10.1111/j.1538-4632.1995.tb00338.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1538-4632.1995.tb00338.x.

Sandra Lach Arlinghaus and Joseph J. Kerski. *Spatial Mathematics: Theory and Practice through Mapping*. CRC Press, Boca Raton, 1 edition edition, June 2013. ISBN 978-1-4665-0532-2.

Adrian Baddeley and Rolf Turner. Spatstat: An R Package for Analyzing Spatial Point Patterns. *Journal of Statistical Software*, 12(1):1–42, January 2005. ISSN 1548-7660. doi: 10.18637/jss.v012.i06. URL https://www.jstatsoft.org/index.php/jss/article/view/v012i06.

Adrian Baddeley, Rolf Turner, and Ege Rubak. Spatstat: Spatial Point Pattern Analysis, Model-Fitting, Simulation, Tests, December 2019. URL https://CRAN.R-project.org/package=spatstat.

Christoph Basten and Frank Betz. Beyond Work Ethic: Religion, Individual, and Political Preferences. *American Economic Journal: Economic Policy*, 5(3):67–91, August 2013. ISSN 1945-7731. doi: 10.1257/pol.5.3.67. URL https://www.aeaweb.org/articles?id=10.1257/pol.5.3.67.

Patrick Bayer, Fernando Ferreira, and Robert McMillan. A Unified Framework for Measuring Preferences for Schools and Neighborhoods. *Journal of Political Economy*, 115(4):588–638, 2007. URL https://ideas.repec.org/a/ucp/jpolec/v115y2007i4p588-638.html.

Roger Bivand and Edzer Pebesma. *Spatial Data Science: With Applications in R*. Chapman and Hall/CRC, New York, 05 2023. doi: 10.1201/9780429459016. URL https://r-spatial.org/book/. DOI: 10.1201/9780429459016.

Sandra E. Black. Do Better Schools Matter? Parental Valuation of Elementary Education. *The Quarterly Journal of Economics*, 114(2):577–599, May 1999. ISSN 0033-5533. doi: 10.1162/003355399556070. URL https://academic.oup.com/qje/article/114/2/577/1844232.

Kirill Borusyak and Peter Hull. Non-Random Exposure to Natural Experiments: Theory and Applications. 2020.

Federico A. Bugni, Ivan A. Canay, and Azeem M. Shaikh. Inference under Covariate-Adaptive Randomization. *Journal of the American Statistical Association*, 113(524):1784–1796, 2018. ISSN 0162-1459. doi: 10.1080/01621459.2017.1375934.

Sebastian Calonico, Matias D. Cattaneo, and Rocío Titiunik. Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs. *Econometrica*, 82(6):2295–2326, 2014.

Sebastian Calonico, Matias D. Cattaneo, and Rocío Titiunik. Rdrobust: An R package for Robust Nonparametric Inference in Regression-Discontinuity Designs. *R Journal*, 7(1):38–51, 2015.

Ivan A. Canay, Joseph P. Romano, and Azeem M. Shaikh. Randomization Tests Under an Approximate Symmetry Assumption. *Econometrica*, 85(3):1013–1030, 2017. ISSN 1468-0262. doi: 10.3982/ECTA13081. URL https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA13081.

Matias D. Cattaneo, Brigham R. Frandsen, and Rocío Titiunik. Randomization Inference in the Regression Discontinuity Design: An Application to Party Advantages in the U.S. Senate. *Journal of Causal Inference*, 3(1):1–24, March 2015. ISSN 2193-3685. doi: 10.1515/jci-2013-0010. URL https://www.degruyter.com/document/doi/10.1515/jci-2013-0010/html.

Matias D. Cattaneo, Nicolás Idrobo, and Rocío Titiunik. A practical introduction to regression discontinuity designs: Foundations. *Elements in Quantitative and Computational Methods for the Social Sciences*, 11 2019. doi: 10.1017/9781108684606. URL https://www.cambridge.org/core/elements/practical-introduction-to-regression-discontinuity-designs/F04907129D5C1B823E3DB19C31CAB905. ISBN: 9781108684606 9781108710206 Publisher: Cambridge University Press.

Matias D. Cattaneo, Nicolas Idrobo, and Rocio Titiunik. A Practical Introduction to Regression Discontinuity Designs: Extensions, January 2023. URL http://arxiv.org/abs/2301.08958.

Sung Nok Chiu, Dietrich Stoyan, Wilfrid S. Kendall, and Joseph Mecke. *Stochastic Geometry and Its Applications*. Wiley, Chichester, 1st edition edition, September 2013. ISBN 978-0-470-66481-0.

Nicholas R. Chrisman. Calculating on a round planet. *International Journal of Geographical Information Science*, 31(4):637–657, April 2017. ISSN 1365-8816. doi: 10.1080/13658816.2016.1215466. URL https://doi.org/10.1080/13658816.2016.1215466.

EunYi Chung and Joseph P. Romano. Exact and asymptotically robust permutation tests. *The Annals of Statistics*, 41(2):484–507, April 2013. ISSN 0090-5364, 2168-8966. doi: 10.1214/13-AOS1090. URL https://projecteuclid.org/journals/annals-of-statistics/volume-41/issue-2/Exact-and-asymptotically-robust-permutation-tests/10.1214/13-AOS1090.full.

A. D. Cliff and J. K. Ord. *Spatial Processes: Models and Applications*. Pion Limited, London, UK, 1981.

Scott Cunningham. *Causal Inference: The Mixtape*. Yale University Press, New Haven ; London, 01 2021.

Herbert A David. The Beginnings of Randomization Tests. *The American Statistician*, 62(1):70–72, February 2008. ISSN 0003-1305. doi: 10.1198/000313008X269576. URL https://doi.org/10.1198/000313008X269576.

Melissa Dell. The persistent effects of peru's mining mita. *Econometrica*, 78(6):1863–1903, 2010. tex.owner: crunchbangax Citation Key: Dell2010.

Melissa Dell and Benjamin A Olken. The Development Effects of the Extractive Colonial Economy: The Dutch Cultivation System in Java. *The Review of Economic Studies*, 87(1):164–203, January 2020. ISSN 0034-6527. doi: 10.1093/restud/rdz017. URL http://www.nber.org/papers/w24009.

Peter J. Diggle, Jorge Mateu, and Helen E. Clough. A comparison between parametric and non-parametric approaches to the analysis of replicated spatial point patterns. *Advances in Applied Probability*, 32(2):331–343, June 2000. ISSN 0001-8678, 1475-6064. doi: 10.1239/aap/1013540166. URL https://www.cambridge.org/core/journals/advances-in-applied-probability/article/abs/comparison-between-parametric-and-nonparametric-approaches-to-the-analysis-of-replicated-spatial-point-patterns/71AAE5CFE60B44F0988DBE0775DA1D40.

Peng Ding and Fan Li. Causal Inference: A Missing Data Perspective. *Statistical Science*, 33(2):214–237, May 2018. ISSN 0883-4237, 2168-8745. doi: 10.1214/18-STS645. URL https://projecteuclid.org/journals/statistical-science/volume-33/issue-2/Causal-Inference-A-Missing-Data-Perspective/10.1214/18-STS645.full.

Dave Donaldson. Railroads of the Raj: Estimating the Impact of Transportation Infrastructure. *American Economic Review*, 108(4-5):899–934, April 2018. ISSN 0002-8282. doi: 10.1257/aer.20101199. URL https://www.aeaweb.org/articles?id=10.1257/aer.20101199.

Dewey Dunnington, Edzer Pebesma, and Ege Rubak. *S2: Spherical Geometry Operators Using the S2 Geometry Library*, 2023.

Max J. Egenhofer and Robert D. Franzosa. Point-set topological spatial relations. *International journal of geographical information systems*, 5(2):161–174, January 1991. ISSN 0269-3798. doi: 10.1080/02693799108927841. URL https://doi.org/10.1080/02693799108927841.

Beatrix Eugster, Rafael Lalive, Andreas Steinhauer, and Josef Zweimüller. Culture, Work Attitudes, and Job Search: Evidence from the Swiss Language Border. *Journal of the European Economic Association*, 15(5):1056–1100, October 2017. ISSN 1542-4766. doi: 10.1093/jeea/jvw024. URL https://doi.org/10.1093/jeea/jvw024.

Ronald A. Fisher. *The Design of Experiments*. Olivier and Boyd, 1935.

Nicola Fontana, Tommaso Nannicini, and Guido Tabellini. Historical roots of political extremism: The effects of Nazi occupation of Italy. *Journal of Comparative Economics*, 51(3):723–743, September 2023. ISSN 0147-5967. doi: 10.1016/j.jce.2023.05.006. URL https://www.sciencedirect.com/science/article/pii/S0147596723000537.

Andrew Gelman and Guido Imbens. Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business & Economic Statistics*, 37(3):447–456, 07 2019. doi: 10.1080/07350015.2017.1366909. URL https://doi.org/10.1080/07350015.2017.1366909. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/07350015.2017.1366909.

Jinyong Hahn, Petra Todd, and Wilbert van der Klaauw. Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, 69(1):201–209, 2001. doi: 10.1111/1468-0262.00183. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/1468-0262.00183.

Ute Hahn. A Studentized Permutation Test for the Comparison of Spatial Point Patterns. *Journal of the American Statistical Association*, 107(498):754–764, June 2012. ISSN 0162-1459. doi: 10.1080/01621459.2012.688463. URL https://doi.org/10.1080/01621459.2012.688463.

Bruce E. Hansen. *Econometrics*. Princeton University Press, 08 2022. URL https://press.princeton.edu/books/hardcover/9780691235899/econometrics.

Paul W. Holland. Statistics and Causal Inference. *Journal of the American Statistical Association*, 81(396):945–960, 1986. ISSN 0162-1459. doi: 10.2307/2289064. URL https://www.jstor.org/stable/2289064.

Thomas Holmes. The Effect of State Policies on the Location of Manufacturing: Evidence from State Borders. *Journal of Political Economy*, 106(4):667–705, 1998. URL https://econpapers.repec.org/article/ucpjpolec/v_3a106_3ay_3a1998_3ai_3a4_3ap_3a667-705.htm.

Jonathan Iliffe and Roger Lott. *Datums and Map Projections: For Remote Sensing, GIS and Surveying*. Whittles Publishing, Dunbeath, 2nd edition edition, April 2008. ISBN 978-1-904445-47-0.

Guido W. Imbens and Karthik Kalyanaraman. Optimal bandwith choice for the regression discontinuity estimator. *Review of Economic Studies*, 79(3):127–155, 2012. tex.owner: crunchbangax Citation Key: ImbensKalyanaraman2012.

Guido W. Imbens and Thomas Lemieux. Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, 142:615–635, 2008.

Guido W. Imbens and Donald B. Rubin. *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge University Press, Cambridge, 2015. ISBN 978-0-521-88588-1. doi: 10.1017/CBO9781139025751. URL https://www.cambridge.org/core/books/causal-inference-for-statistics-social-and-biomedical-sciences/71126BE90C58F1A431FE9B2DD07938AB.

Guido W. Imbens and Tristan Zajonc. Regression Discontinuity Design with Multiple Forcing Variables. *mimeo*, 2011.

E. Sinem Ince, Franz Barthelmes, Sven Reißland, Kirsten Elger, Christoph Förste, Frank Flechtner, and Harald Schuh. ICGEM – 15 years of successful collection and distribution of global gravitational models, associated services, and future plans. *Earth System Science Data*, 11(2):647–674, May 2019. ISSN 1866-3508. doi: 10.5194/essd-11-647-2019. URL https://essd.copernicus.org/articles/11/647/2019/.

Arnold Janssen. Studentized permutation tests for non-i.i.d. hypotheses and the generalized Behrens-Fisher problem. *Statistics & Probability Letters*, 36(1):9–21, November 1997. ISSN 0167-7152. doi: 10.1016/S0167-7152(97)00043-6. URL https://www.sciencedirect.com/science/article/pii/S0167715297000436.

Luke J. Keele and Rocío Titiunik. Geographic Boundaries as Regression Discontinuities. *Political Analysis*, 23(1):127–155, 2015. doi: 10.1093/pan/mpu014. URL +http://dx.doi.org/10.1093/pan/mpu014.

J. F. C. Kingman. *Poisson Processes*. Clarendon Press, Oxford : New York, 1st edition edition, January 1993. ISBN 978-0-19-853693-2.

David S. Lee. Randomized experiments from non-random selection in U.S. House elections. *Journal of Econometrics*, 142(2):675–697, February 2008. ISSN 0304-4076. doi: 10.1016/j.jeconom.2007.05.004. URL https://www.sciencedirect.com/science/article/pii/S0304407607001121.

David S. Lee and Thomas Lemieux. Regression Discontinuity Designs in Economics. *Journal of Economic Literature*, 48(2):281–355, 2010.

Pierre Legendre. Spatial Autocorrelation: Trouble or New Paradigm? *Ecology*, 74(6):1659–1673, 1993. ISSN 1939-9170. doi: 10.2307/1939924. URL https://esajournals.onlinelibrary.wiley.com/doi/abs/10.2307/1939924.

Erich L. Lehmann and Joseph P. Romano. *Testing Statistical Hypotheses*. Springer, New York, third edition edition, January 2005. ISBN 978-0-387-98864-1.

Alexander Lehner. SpatialRDD: An R-package for Spatial Regression Discontinuity Designs, 2020. URL https://github.com/axlehner/SpatialRDD.

Alexander Lehner. SpatialRDD: Conduct Multiple Types of Geographic Regression Discontinuity Designs, August 2023. URL https://cran.r-project.org/web/packages/SpatialRDD/index.html.

Robin Lovelace, Jakub Nowosad, and Jannes Muenchow. *Geocomputation with R*. Chapman and Hall/CRC, Boca Raton London New York, 1st edition edition, December 2020. ISBN 978-0-367-67057-3.

James G. MacKinnon and Matthew D. Webb. Randomization inference for difference-in-differences with few treated clusters. *Journal of Econometrics*, 218(2):435–450, October 2020. ISSN 0304-4076. doi: 10.1016/j.jeconom.2020.04.024. URL https://www.sciencedirect.com/science/article/pii/S0304407620301445.

James G. MacKinnon, Morten Ørregaard Nielsen, and Matthew D. Webb. Cluster-robust inference: A guide to empirical practice. *Journal of Econometrics*, 232(2):272–299, February 2023. ISSN 0304-4076. doi: 10.1016/j.jeconom.2022.04.001. URL https://www.sciencedirect.com/science/article/pii/S0304407622000781.

D. H. Maling. *Coordinate Systems and Map Projections*. Pergamon, Oxford New York, 2nd edition edition, January 1992. ISBN 978-0-08-037233-4.

Jordan Matsudaira. Mandatory summer school and student achievement. *Journal of Econometrics*, 142(2):829–850, 2008. ISSN 0304-4076. URL https://econpapers.repec.org/article/eeeeconom/v_3a142_3ay_3a2008_3ai_3a2_3ap_3a829-850.htm.

Guy Michaels, Dzhamilya Nigmatulina, Ferdinand Rauch, Tanner Regan, Neeraj Baruah, and Amanda Dahlstrand. Planning Ahead for Better Neighborhoods: Long Run Evidence from Tanzania. *Journal of Political Economy*, February 2021. ISSN 0022-3808. doi: 10.1086/714119. URL https://www.journals.uchicago.edu/doi/abs/10.1086/714119.

John P. Papay, John B. Willett, and Richard J. Murname. Extending the Regression-Discontinuity Approach to Multiple Assignment Variables. *Journal of Econometrics*, 161(2):203–207, 2011.

Edzer Pebesma. Simple features for r: Standardized support for spatial vector data. *The R Journal*, 10(1):439–446, 2018. doi: 10.32614/RJ-2018-009. URL https://doi.org/10.32614/RJ-2018-009.

E. J. G. Pitman. Significance Tests Which May be Applied to Samples From any Populations. *Supplement to the Journal of the Royal Statistical Society*, 4(1):119–130, 1937. ISSN 1466-6162. doi: 10.2307/2984124. URL https://www.jstor.org/stable/2984124.

Sean F. Reardon and Joseph P. Robinson. Regression Discontinuity Designs With Multiple Rating-Score Variables. *Journal of Research on Educational Effectiveness*, 5(1):83–104, January 2012. ISSN 1934-5747. doi: 10.1080/19345747.2011.609583. URL https://doi.org/10.1080/19345747.2011.609583.

Oliver Schabenberger and Carol A. Gotway. *Statistical Methods for Spatial Data Analysis*. Chapman and Hall/CRC, Boca Raton, 1 edition edition, December 2004. ISBN 978-1-58488-322-7.

Azeem M. Shaikh and Panos Toulis. Randomization Tests in Observational Studies With Staggered Adoption of Treatment. *Journal of the American Statistical Association*, 116(536):1835–1848, October 2021. ISSN 0162-1459, 1537-274X. doi: 10.1080/01621459.2021.1974458. URL https://www.tandfonline.com/doi/full/10.1080/01621459.2021.1974458.

Donald L. Thistlethwaite and Donald T. Campbell. Regression-discontinuity analysis: An alternative to the ex post facto experiment. *Journal of Educational Psychology*, 51(6):309–317, 1960. ISSN 1939-2176(Electronic),0022-0663(Print). doi: 10.1037/h0044319.

Waldo Tobler. Global spatial analysis. *Computers, Environment and Urban Systems*, 26(6):493–500, November 2002. ISSN 0198-9715. doi: 10.1016/S0198-9715(02)00017-0. URL https://www.sciencedirect.com/science/article/pii/S0198971502000170.

Matthew A. Turner, Andrew Haughwout, and Wilbert van der Klaauw. Land Use Regulation and Welfare. *Econometrica*, 82(4):1341–1403, 2014. ISSN 1468-0262. doi: 10.3982/ECTA9823. URL https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA9823.

Kyle Walker. Crsuggest: Obtain Suggested Coordinate Reference System Information for Spatial Data, July 2022. URL https://cran.r-project.org/web/packages/crsuggest/index.html.

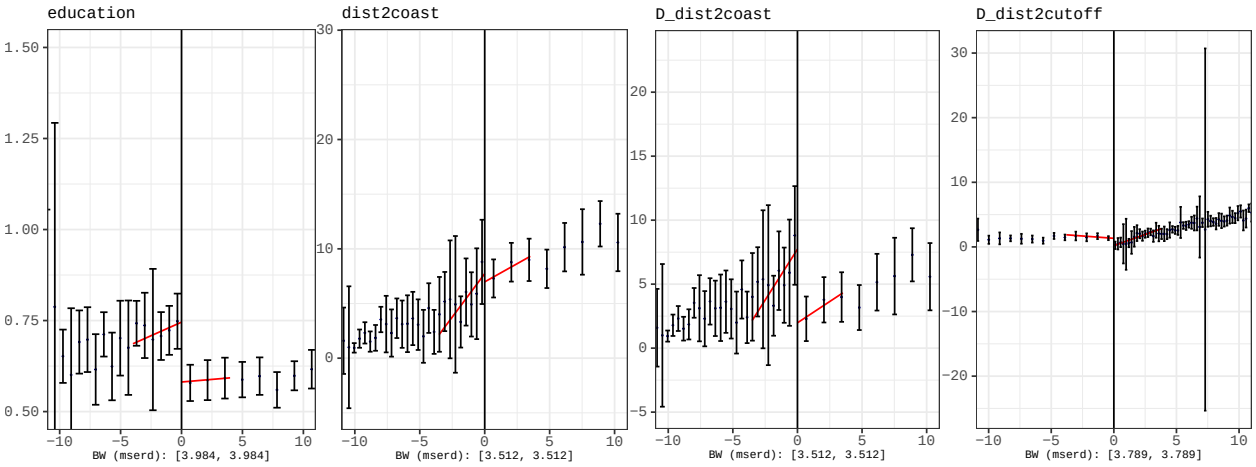
Vivian C. Wong, Peter M. Steiner, and Thomas D. Cook. Analyzing Regression-Discontinuity Designs With Multiple Assignment Variables: A Comparative Study of Four Estimation Methods. *Journal of Educational and Behavioral Statistics*, 38(2):107–141, April 2013. ISSN 1076-9986, 1935-1054. doi: 10.3102/1076998611432172. URL http://journals.sagepub.com/doi/10.3102/1076998611432172.

Jason Wu and Peng Ding. Randomization Tests for Weak Null Hypotheses in Randomized Experiments. *Journal of the American Statistical Association*, 116(536):1898–1913, October 2021. ISSN 0162-1459, 1537-274X. doi: 10.1080/01621459.2020.1750415. URL https://www.tandfonline.com/doi/full/10.1080/01621459.2020.1750415.

Tristan Zajonc. Essays on Causal Inference for Public Policy. August 2012. URL https://dash.harvard.edu/handle/1/9368030.

10 Appendix

10.1 Estimation With Uni-Dimensional Score



(a) 'conventional' discontinuity (b) a spatial trend alone (c) spatial trend + discontinuity (d) spatial trend + discontinuity

Figure 7: RD clouds visualized together with the different RD point estimates. They can differ substantially. The underlying code heavily borrows from the excellent rdrobust package.

10.2 Illustration of Treatment Effect Heterogeneity in the Boundary Design

10.3 Shifting Borders in `SpatialRDD`