Estimating Unknown Cut-points in Regression Discontinuity & Kink Designs

Benjamin F. Reese*

Abstract

Regression discontinuity designs (RDD) and regression kink designs (RKD) are popular identification strategies across the social sciences. The relatively weak assumptions required for identifying a causal effect with RDD/RKDs make the designs attractive quasi-experimental methods for causal inference. One limitation of RDD/RKDs is that they rely on an exogenously given and previously known treatment rule. To overcome this limitation, economists (Porter & Yu 2015, Hansen 2017, Boehnke & Bonaldi 2019, and Tanu 2020) and computer scientists (Herlands et al. 2018) have begun to develop methods and tests that can find unknown discontinuities. However, these tests are either largely theoretically focused, computationally intensive, inconvenient to implement, or otherwise do not fit the specific needs of applied political scientists. Accordingly, this paper presents a novel, more flexible, and extremely easy to implement approach to estimate unknown cut-points in both RDDs and RKDs. It is the first method that can detect unknown thresholds in both RDDs and RKDs. I call this method the unknown cut-point regression discontinuity/kink design (UCRDD/UCRKD). It works by uniformly dividing an assignment variable into quantiles to create a distribution of "candidate" cut-points. Each candidate threshold is then tested in a RDD model (Imbens & Kalyanaraman 2012) to find the "best" cut-point, the cut-point that has the largest substantive effect and highest degree of statistical significance. Researchers can use UCRDD/UCRKD to find "tipping-points" in behavior; to determine obscured or non-public policy criteria; and as a diagnostic tool to determine if there are other significant thresholds in their traditional RDD/RKD. In the application section, I apply UCRKD to assess how to define the concept of a minority district and to estimate the treatment effect of minority districts on electoral support for minority candidates.

---

*Department of Government, Georgetown University, bfr11@georgetown.edu. This is a working draft and I appreciate any and all comments.

1. Introduction

First introduced in Thistlethwaite and Campbell (1960), regression discontinuity designs (RDDs), though sparsely employed in the decades following their inception (Cook 2008), have become an essential feature of applied researchers' econometric toolkits after being popularized in Angrist & Lavy (1999), Black (1999), and Hahn et al. (2001). The relatively weak assumptions required to identify a causal effect with RDDs makes the approach an attractive quasi-experimental research design (Imbens & Lemieux 2008; Lee & Lemieux 2010; Skovron & Titiunik 2015). The three core ingredients of a RDD are: an assignment or running/forcing variable, a treatment rule that determines at what value of the running variable indicates a unit's treatment status, and a dichotomous treatment variable indicating whether or not a unit is treated. The key effect of interest is the coefficient on the binary treatment variable. A substantively large and statistically significant coefficient on the treatment variable, as long as the identifying assumptions are met, indicates a causal effect in the relationship between X and Y. Theoretically, since treatment assignment is determined by an exogenous rule, treatment is uncorrelated with any possible confounders, allowing for an unbiased estimate of the local average treatment effect (LATE) of X on Y. RDDs as an identification strategy rely on either the as-if randomized assumption that assumes units close to the cut-point are sufficiently interchangeable such that treatment assignment is as-if randomized, or, the more common and realistic continuity assumption which assumes that there is no discontinuity in the error term, or that no potential covariates have discontinuities at the threshold. As long as the as-if randomized or the continuity assumption is met, researchers can confidently estimate an unbiased effect of their treatment of interest. A cognate method, regression kink designs (RKDs), follow the same logic as a RDD, but the regression function is continuous, meaning there is no discontinuity, or "jump," at the threshold (Nielsen et al. 2010). Instead, the slope of the regression line changes at the threshold because of some treatment rule. The cornerstone of RDDs and RKDs that make them such a powerful method for causal inference is the exogenously given treatment rule. Typical treatment rules include scholastic eligibility criteria like grade point average (GPA) cutoffs for college admissions or scholarship eligibility; majority vote thresholds in elections; and any instance where policy eligibility criteria is known to a researcher. The requirement of a known discontinuity is a challenge and major limitation of RDD/RKDs which limits their usefulness. Overcoming this challenge is the focus of this paper.

To remedy the requirement of a known cut-point, economists (Porter & Yu 2015, Hansen 2017, Boehnke & Bonaldi 2019, and Tanu 2020) and computer scientists (Herlands et al. 2018) have begun to develop tests to find unknown discontinuities. These tests, however, are either largely theoretically focused, computationally intensive, inconvenient to implement, or otherwise do not fit the specific needs of political scientists. This paper presents a novel, more flexible, and extremely easy to implement approach to estimate unknown cut-points based, in part, on the theoretical framework of Porter & Yu (2015). I call this method unknown cut-point regression discontinuity/kink design (UCRDD/UCRKD), and it is the first method proposed that is able to estimate unknown thresholds in both RDDs and RKDs. The principal

of the method is to uniformly partition a specified assignment variable into sufficiently small quantiles to create a distribution of "candidate" cut-points. Each candidate discontinuity is then tested in a triangular kernel localized linear RDD model (Imbens & Kalyanaraman 2012) or a uniform kernel localized linear RKD model (Card et al. 2016) to find the "best" cut-point, the cut-point that has the largest substantive effect and the highest degree of statistical significance. This "best" cut-point found in the sample is the true discontinuity in the population. Shown in section 4, UCRDD/UCRKD is highly accurate and can recover known discontinuities in both simulation and replication data with a high degree of precision, making it a more precise method than those mentioned above.

Researchers can use UCRDD/UCRKD to find "tipping-points," or discontinuities, in behavior when treatment rules are unknown; to determine obscured or non-public policy criteria; and as a diagnostic tool to determine if there are other thresholds in their traditional RDD/RKD. Further, UCRDD/UCRKD can be used to calculate a distribution of LATEs from differing possible thresholds that allow researchers operating with a known cut-point to compare LATEs across different discontinuities and thresholds, providing a more robust discussion of their treatment effects. Crucially, the approach outlined in this paper can be flexibly used in sharp RDDs/RKDs, fuzzy RDDs/RKDs, and is overall agnostic about the inclusion of covariates and the functional form of the relationship between the assignment variable and the dependent variable, another advantage as compared to previous methods. The application in this paper is primarily concerned with using an UCRKD approach to find discontinuities in behavior, or tipping points/thresholds at which behavior changes independent of a known cut-point. This application is inspired by Card et al. (2008), who used a variety of threshold detection methods to empirically test the Schelling (1971) model and find the exact proportion of African-American residents in a white community that would cause de facto segregation. A similar example exploring at what threshold of minority voters in a district results in minority candidates being elected is the focus of the application section. Using data from the 2012 presidential election, this paper finds that once a district passes 57.4% white, there is a significant decrease in electoral support for Barack Obama.

This paper makes two key contributions: the development of a more flexible and easy to implement method to detect unknown discontinuities in RDDs and RKDs, and it introduces, summarizes, and highlights the merits of applications of RDDs and RKDs with unknown discontinuities to political scientists. The remainder of this paper is organized as follows: the first section reviews the related literature on the estimation of unknown thresholds, how its been applied, and its cognate methodologies; the second section outlines the theoretical underpinnings of UCRDD/UCRKD; the third section assesses the accuracy of UCRDD with both simulation and replication data; and the final section considers the application of UCRKD to a topic of relevance for political scientists: how to define and assess the effects of minority voting districts.

2. The Estimation of Unknown Discontinuities & Thresholds

The approach used to detect unknown cut-points outlined in this paper employs a series of models with differing "candidate" cut-points that are tested and assessed by maximizing a specific selection criteria. The logic underpinning this detection of unknown discontinuities highly correlates with the rich tradition of machine learning for model and feature selection. Machine learning dates back at least to McCulloch and Pitts (1943), but its use in political science is far more recent. Grimmer et al. (2021), in their discussion of how machine learning can be used in political science, cite the most common machine learning methods employed in political science as neural networks (Beck et al. 2000; Williams et al. 2020), decision tree and random forest models (Stewart & Zhukov 2009; Hill & Jones 2014; Montgomery & Olivella 2018; Kaufman et al. 2019; Acharya et al. 2021), support vector machines (Hillard et al. 2008), and naïve Bayes (Nelsen 2017). The central aspect of each of these approaches is that researchers split their data into a training set and implementation set and, through a chosen estimation procedure and selection criteria - usually in the form of minimizing some error metric such as mean squared error or mean absolute error - find the "best" model that fits their data. In other words, repeatedly testing and assessing a model by some error metric allows researchers to make better predictions and find the factors that are most important in explaining a dependent variable of interest. I use this same basic logic to create an algorithm to find unknown discontinuities and thresholds in RDDs/RKDs. Machine learning, therefore, offers a powerful approach of learning directly from empirical data which is essential to finding unknown cut-points and forms the foundation of the approach presented in this paper.

Within the realm of political science, RDDs are most commonly associated with analyzing close elections, an approach first developed by Lee et al. (2004) and Lee (2008). Since then, the utilization of RDDs for analyzing closely contested elections has become increasingly common. See de la Cuesta & Imai(2016) for a more complete review and associated methodological challenges. The relatively myopic set of applications of RDDs in political science is surely due, in some part, to the constraint of needing a known treatment rule. The constraint of needing a known treatment assignment rule in RDDs and RKDs is one limitation that causal inference scholars have begun addressing.

The first method devised to estimate unknown cut-points in RDDs specifically was presented in Porter & Yu (2015). Porter & Yu (2015) advance a two-step procedure wherein the first stage the discontinuity is estimated with a difference kernel estimator and then researchers proceed with the estimated cut-point as if the actual discontinuity was known in the second stage. Porter & Yu's (2015) approach requires the exclusion of the possibility of a null treatment effect before it can probe for the presence of a selection effect. The method proposed in this paper is not limited by this constraint and can test for both selection effects, or significant discontinuities, and treatment effects, or the actual size of the jump at the threshold. Further, Porter & Yu's (2015) method is limited only to RDDs and is less flexible in terms of including covariates than UCRDD/UCRKD.

Boehnke & Bonaldi (2019) designed a more general approach that they call "synthetic regression discontinuity design" which creates a score, akin to a propensity score in certain matching methods, to predict a unit's treatment status. With the score calculated, researchers can use the predicted synthetic score as the assignment variable in a RDD model. The method proposed in this paper is far more computationally efficient, easily implemented, and more general and flexible as it does not require the creation or analysis of a latent score. Further, UCRDD/UCRKD can handle both RDD and RKDs.

Herlands et al. (2018) propose an alternative method to estimate unknown cut-points even further rooted in the machine learning literature that employs the use of an iterative k-nearest neighbors search algorithm which divides the neighborhood into two equal partitions and checks for a significant discontinuity with a log-likelihood ratio test. This process is repeated with different sizes of neighborhoods and any statistically significant candidate cut-points are determined and saved. Tanu (2020), in a yet to be published work, offers a slightly different approach that requires some knowledge of a possible cut-off for optimal efficiency. Tanu's (2020) method relies on the process of machine learning rooted in structural break detection (Pretis et al. 2018) or Andrews' Test for unknown change points first developed in Andrews (1993). Both of these methods are far more computationally intensive, less flexible, and do not work in RKDs as the method proposed in this paper does.

Beyond the context of unknown discontinuity estimation in RDDs, econometricians and statisticians have devoted significant scholarship to detecting unknown thresholds in a variety of other settings. Hansen (2017) shows how unknown kinks in RKDs - a more novel method than RDDs - can be estimated. Hansen (2017) is an application and extension of the inference methods devised in Fang & Santos (2014) and Hong & Li (2015) to the context of RKDs, and extends the method proposed in Chan & Tsay (1998) to specifically find the point in which slopes shift in an unknown RKD. Other than RKDs, Chan (1990 & 1991) and Caner & Hansen (2004) have offered approaches to detecting threshold effects in panel data settings and instrumental variable contexts, respectively. Testing for structural breaks in time series analyses, or when there is an abrupt change at one point in a time series, is also a form of estimating an unknown threshold.[1] Even more generally, threshold regression refers to a class of regression models where the relationship between the causal variable and the outcome variable in threshold dependent.[2] Threshold regression requires testing several candidate thresholds to determine when changes in slope are most significant. In each of these contexts, there is some threshold that is theorized to exist, but there is no clear indication of exactly where the threshold is. The detection of unknown thresholds is, therefore, essential for researchers to apply more general and powerful methodologies in contexts without clear indication of where cut-points are.

Despite the rich tradition of research detecting unknown discontinuities and thresholds in

---

[1]See Perron (2006 & 2008) and Casini & Perron (2018) for reviews of structural breaks in the time series literature.

[2]See Fong et al. (2017) for a brief review of threshold regression models and the software packages needed to implement them.

the econometrics and statistics literature, its prevalence in political science is basically non-existent. Since these approaches are well suited for modeling non-linear relationships which are common in political science, it is perhaps surprising that political methodologists have not yet attempted to adapt these and cognate methods for applications of interest to political researchers. This paper hopes to, in part, remedy this by presenting some of the possibilities of overlooked methods like RKDs and threshold regression in political science settings.

3. Unknown Cut-point Regression Discontinuity/Kink Design

Broadly, UCRDD/UCRKD relies on partitioning an assignment variable into n^{th} quantiles to create "candidate" cut-points. Each quantile - these could be percentiles or even smaller subdivisions - is assumed to be equally likely to be the true cut-point. The second step then estimates a basic RDD/RKD with each candidate cut-point. While this step is agnostic about the exact estimation procedure, when focusing on a RDD, a triangular kernel localized linear regression (Imbens & Kalyanaraman 2012) approach is used for this paper. For RKDs, a uniform kernel is the current best practice (Card et al. 2016). As novel estimation methods are developed, or researchers have some theoretical justification to prefer another estimation technique, UCRDD/UCRKD can easily be adapted by simply changing the estimation procedure, ensuring its continued applicability in the face of a rapidly developing literature. The third step is to save the LATE and p-value from each iterated RDD/RKD model. The candidate cut-point with the highest ratio of LATE to p-value, assuming the LATE achieves statistical significance at conventional levels (e.g. $p < .05$), is the best candidate for the true cut-point. If no candidate cut-point has an associated p-value below .05, then researchers cannot be confident there is an actual discontinuity, at least at the 95% level. Any candidate cut-points that have nearly equivalent selection criteria statistics can be differentiated by looking at the Akaike information criterion (AIC), or some other model selection criteria, of each iterated model and selecting the model with the lowest AIC.

For an example of how UCRDD/UCRKD can be applied, consider the common RDD application of determining the effect of education on future earnings potential. A scholarship granting body may award students with funding for college education based on a not publicly released rule, such as surpassing some threshold in a synthetic score based on academic achievement (i.e. GPA, SAT, etc.). If a researcher has data on students who did and did not get the scholarship, the synthetic score, and their future earnings, then, with UCRDD, no information about the exact cutoffs in a criteria are actually necessary. The researcher can simply test all of the values of the score iteratively and recover the treatment rule based on an optimal selection criterion. In sum, UCRDD/UCRKD partitions an assignment variable into small quantiles and tests each quantile as a candidate cut-point to determine the "best" candidate for the true threshold.

3.1 UCRDD/UCRKD Assumptions, Inference, & Limitations

UCRDD/UCRKD is an incredibly simple method that can be implemented extremely quickly with modern statistical computing and only rests upon three weak assumptions:

Assumption 1: The true cut-point takes a value between the minimum and maximum values of the assignment variable.

Assumption 2: There exists some form of treatment rule or discontinuity in behavior based on some value of the assignment variable.

Assumption 3: There exists some relationship of interest between the assignment variable and the outcome variable such that the traditional RDD/RKD assumptions are satisfied.

In other words, a cut-point must exist, take a possible value of the running variable, and there must be a relationship of theoretical or empirical interest between the assignment and outcome variables. The first two assumptions can only be argued qualitatively and never validated quantitatively. Researchers need to thoroughly understand their assignment variable and possible thresholds to know that a discontinuity or treatment rule actually exists. Put differently, deep substantive knowledge about policy criteria or some behavior is required to know that, in fact, a discontinuity/threshold exists despite its threshold not being known to the researcher. **Assumption 3** can be met more easily by simply satisfying the identification assumptions of RDD/RKDs. **Assumption 3** is also a limitation as the only way to detect a cut-point is to have a previously specified relationship in mind, again showing how central substantive expertise is to the UCRDD/UCRKD approach.

Explicit in the identification assumptions for the traditional RDD/RKD is the requirement of no sorting around the cut-point, or, in other words, individuals are not changing their behavior to move themselves just beyond the threshold.[3] Sorting at the threshold indicates that units are "gaming-the-system," or otherwise changing their behavior to achieve a value of the assignment variable just beyond the cutoff, thus questioning the exogeniety of the treatment rule. An example of this is a student who knows they can receive a scholarship if they attain a 3.0 GPA so they employ any strategy available to them such as submitting extra credit or bargaining with a teacher to push themselves past the threshold, drawing into question the exogeneity of treatment assignment. A treatment rule no longer being exogenous is not a concern when the public is unaware of an eligibility criteria.

For a RDD to be a valid research design, the treatment assignment must be akin to random treatment assignment and the units at the cut-point must be as-if interchangeable.

---

[3]McCrary (2008) offers a method to test for sorting around the cutoff in RDDs.

In instances where RDDs are applied in political science, mainly election outcomes, this assumption of no sorting is most likely to be violated (de la Cuesta & Imai 2016). In such cases, a different set of assumptions, which are difficult to satisfy, are required. While finding an unknown discontinuity for a RDD or an unknown threshold for a RKD is far from trivial and much harder than hunting for the traditional known cut-points, if a researcher can find and exploit an opportunity for a UCRDD/UCRKD, the result is a better identified and "purer", or closer to a randomized controlled trial, RDD/RKD than any known cut-point could provide. This is because the cut-off, if unknown to a researcher with considerable substantive expertise, is convincingly unknown to units in the treatment and control groups. In other words, UCRDD/UCRKD avoids the pitfall of manipulation around the cutoff. Therefore, the resulting treatment effect estimated in the RDD/RKD will be entirely free from sorting in the running variable. The trade-off is that UCRDD/UCRKD is more data intensive and requires an even higher degree of expert knowledge to find an applicable scenario. While at first glance these may seem to be major weaknesses of this approach, they are, in fact, a core strength of it as RDDs in political science are notoriously under-powered (Stommes et al. 2023). The risk of sorting at the cut-point, even if detectable through a McCrary density test (McCrary 2008), is also ever present and greater substantive knowledge of a topic can only benefit a research design. If the RDD/RKD assumptions are not met for causal identification, however, UCRDD and UCRKD can still be a powerful descriptive tool. If an assignment variable is unknown, a researcher can use UCRDD/UCRKD to test a range of different candidate running variables to determine if possible cut-points exist along different potential assignment variables. Researchers can then test their models and make inferences with new data sources.

Considering the increasingly rigorous standards for causal identification, it is important to remember the problematic nature of making inferences about a population given an estimated cut-point is both detected and tested with the same sample. To remedy this, as long as their is sufficient data, researchers can split their dataset into a training set and an implementation set, akin to the standard for predictive modeling in supervised machine learning models. Once the unknown cut-point is detected in the training set, one can conduct the usual RDD/RKD estimation exclusively on the implementation dataset. This splitting of the data necessitates having a large enough sample to maintain sufficient statistical power to estimate the LATE at different thresholds in a trimmed down training dataset. Accordingly, the key limitation of UCRDD/UCRKD is the considerable amount of data required to check each candidate cut-point as not finding a significant discontinuity is potentially a function of the lack of data rather than there being no threshold. With the assumptions, methods of inference, and limitations discussed, the remainder of this section outlines the step-by-step procedure of implementing UCRDD/UCRKD.

3.2 Creating Candidate Cut-points

Once a researcher has theorized that there does, in fact, exist an unknown treatment assignment rule based on some cutoff value existent in a running variable, the first step of

UCRDD/UCRKD is to partition the assignment variable into n^{th} quantiles. The values of the assignment variable at these different quantiles are the candidate cut-points that will each be tested in a subsequent step. Because the goal is to recover the true cutoff in the population, which is assumed to take a value between the minimum and the maximum of the assignment variable, the largest number of quantiles (e.g. the largest number of candidate cut-points) possible is optimal. It is more likely that a researcher can find the exact true cut-point with a high degree of precision given that they tested 500 different values of an assignment variable versus only testing 50 values. Modern computer software makes testing a large number of candidate cut-points trivial, and, with such a simple estimation procedure, UCRDD/UCRKD with a high number of candidate cut-points is still far less computationally intensive as Herlands et al. (2018) and Boehnke & Bonaldi (2019).

As the number of quantiles goes to infinity, the distance between numerical values of the assignment variable collapses to zero, meaning every possible value the assignment variable could take will be tested. When the assignment variable is divided into infinitely small partitions, one of those cut-points will, in the limit, exactly match the value of the true cutoff in the population.

$$\lim_{n->\infty} P(c_i = C) = 1 \tag{1}$$

In other words, as the number of quantiles goes to infinity, the probability that a candidate cut-point, c_i will equal the true cut-point in the population, C, is 100%. While the limitations of working with actual data may preclude the use of a massive number of quantiles - specifically if the variance or n of the assignment variable is small -, the largest feasible number of quantiles is optimal. Since one of the candidate cut-points will be the true threshold, all that is needed is a selection criteria to recover it. The optimal selection criteria is discussed in section 3.4.

3.3 Modeling Decisions

Once the assignment variable is divided into quantiles and the candidate cut-points are determined, the second step is to specify the model to be tested. The goal of this step is to simply determine an assignment and outcome variable of interest and to specify their relationship. The second step is where researchers must make determinations in regards to selecting between a RDD or a RKD; if they want a fuzzy or sharp design; if they want to include covariates; how they wish to model the functional form of the relationship between the assignment and outcomes variables; and any other modeling decisions. UCRDD/UCRKD can accommodate any decision related to model specification. While the basic RDD model is estimated with OLS and is most simply depicted as:

$$Y_i = \beta_0 + \beta_1 T_i + \beta_2 X_i + \beta_3 (X_i - C) + \epsilon_i \tag{2}$$

where $T_i = 0$ if $X < C$ and $T_i = 1$ if $X \geq C$, UCRDD - and its UCRKD counterpart - does not require nor assume a discrete model specification. Rather, UCRDD/UCRKD allows for any estimation procedure to be used. The flexibility of the method affords researchers supreme freedom in this step, as whichever estimation procedure they choose can be accommodated. UCRDD/UCRKD is completely agnostic about the functional form of the model, a benefit over Porter & Yu (2015), Hansen (2017), and Boehnke & Bonaldi (2019). While UCRDD/UCRKD allows for the use of any model, the RDD and RKD literatures have recommendations for each stage of an analysis from estimation and covariate inclusion to window size. It is important that researchers follow these modeling recommendations, or at least justify any deviations, in this step.[4] Since UCRDD/UCRKD is both a descriptive and inferential method, it is important to delineate the difference in modeling decisions here.

If UCRDD/UCRKD is intended for inference, the data should be split into training and implementation datasets and a theory and model should be specified before any analysis is undertaken. Best practices for causal inference necessitate that there should be no model fishing for what assignment variable goes with what outcome variable in order to find some significant discontinuity/threshold that lacks theoretical justification. Instead, a theorized relationship about an unknown threshold should be discussed in advance. When the cut-point is found with the training data, it can then be tested on the implementation data and inferences made as in the typical sequence of events for RDDs and RKDs. If a researcher's interest is not inferential, rather it is descriptive in nature, splitting into training and implementation datasets is optional. For descriptive purposes, researchers can utilize the full dataset and consider testing different assignment variables and different model specifications to learn about potential treatment rules and their effects on an outcome variable of interest. The descriptive approach allows researchers to build theories that can be tested with other datasets following the normal RDD or RKD procedures. As mentioned above and later in the application section, the descriptive approach can further be used to find tipping-points in political and social behavior that are of interest to scholars of politics.

3.4 Selection Criteria & Testing Candidate Cut-points

The final step is to test each candidate discontinuity and see which candidate cut-point maximizes the selection criteria. This is as simple as running each model from step two with each respective cut-point from step one and saving the results. For this third step, a uniform prior is assumed,

$$c_i \sim U(c_{min}, c_{max}) \tag{3}$$

where C_i is the vector of cut-points to be tested, C_{min} is the first cut-point, the minimum value of the assignment variable, and C_{max} is the last cut-point, the maximum value of the

---

[4]Cattaneo et al. (2018) provide "best practice" guidelines for how to implement RDDs while Card et al. (2016) provide a similar guide for RKDs.

assignment variable. Assuming a uniform prior, the probability distribution function (PDF) of determining the true cut-point is:

$$f(c_i) = 1 \ for \ \ c_{min} \leq c \leq c_{max} \tag{4}$$

and the cumulative distribution function (CDF) is:

$$F(c_i) = c \ for \ \ c_{min} \leq c \leq c_{max} \tag{5}$$

The PDF and CDF of finding the true cut-point in the population based on the candidate cut-points is depicted in *Figure 1* and *Figure 2*, respectively.
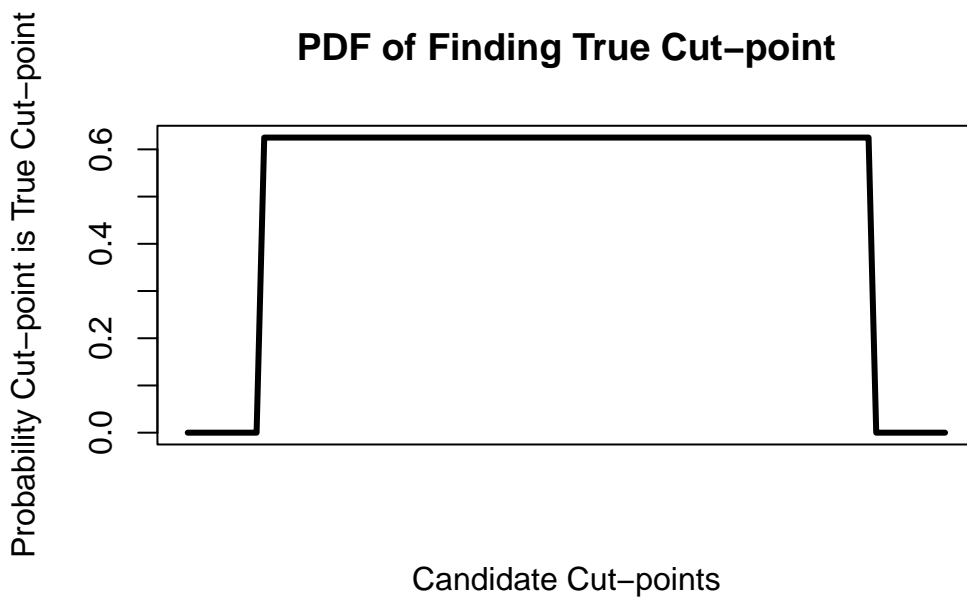

Figure 1: PDF of Finding True Cut-point

Put plainly, each potential cut-point from the sample, c_i, has an equal probability of being the true cutoff, C, in the population. This is perhaps an unnecessarily stringent assumption given the exact distribution of an assignment variable. For example, scholars working with heavily skewed data may not find much utility in testing cut-points that have few near to the threshold observations with which to provide enough statistical power to detect an effect. These can be dropped without any effect on performance as most software packages will even fail to estimate a model that lacks observations near the specified cut-point. The uniform prior, while not the most efficient, is the most agnostic about where the unknown threshold may be.
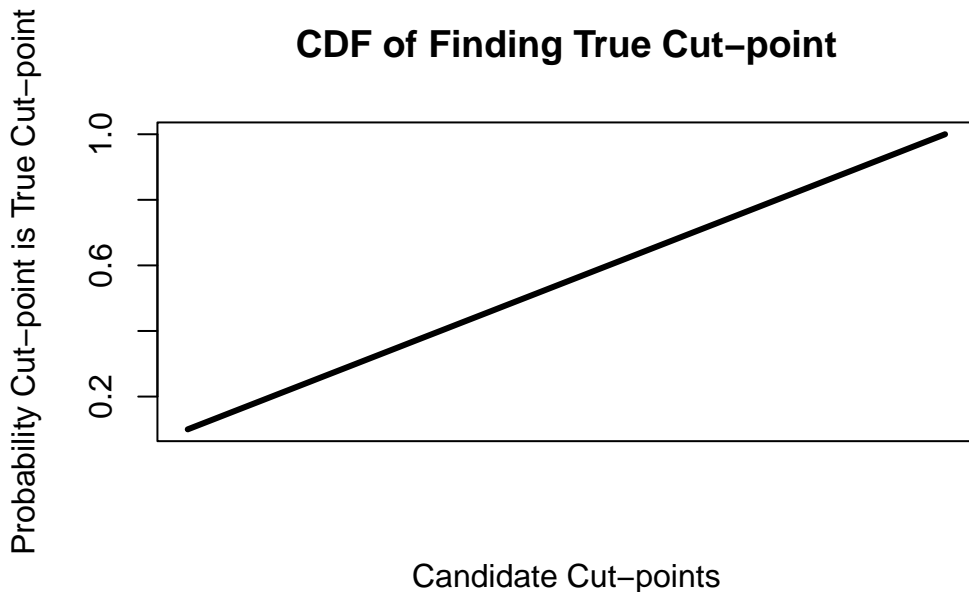
CDF of Finding True Cut-point

Probability Cut-point is True Cut-point

Candidate Cut-points

Figure 2: CDF of Finding True Cut-point

Each model iteration returns both a LATE and a p-value to be evaluated. The ratio of the magnitude of the treatment effect to the p-value, which I denote as λ_i, provides the necessary information about where the true discontinuity is. The candidate cut-point with the largest λ_i is the true cutoff in the population.[5]

$$\lambda_i = \left| \frac{\beta_{c_i}}{p_{c_i}} \right| \tag{6}$$

In words, the unknown threshold is determined by finding which candidate cut-point has the substantively largest and most statistically significant treatment effect. If the relationship between the assignment and outcome variables has no discontinuity, no model will be statistically significant at conventional thresholds and all of the λ_i's will be statistically indistinguishable from one another as the LATEs will be extremely small and the p-values large. When a discontinuity actually exists, the LATE will be larger than the other candidate cut-points' LATEs and statistically significant. In all accuracy tests, true cut-points have selection criteria massively larger than all other candidates. Thus, the true cut-point in the population is recovered by solving:

---

[5]See section 4 for a demonstration of UCRDD/UCRKD's accuracy.

$$max \left| \frac{\beta_{c_i}}{p_{c_i}} \right| \qquad (7)$$

This is the point in the procedure where the choice of the number of quantiles becomes extremely important. If one was to use percentiles, thus testing 100 candidate cut-points, they will be able to more easily differentiate between the candidate cut-points and find the one with the highest λ. However, the lack of granularity may lead to the best candidate cut-point being slightly off from the true cutoff in the population since the quantiles are not sufficiently small. If one was to divide the data into 1,000 qauntiles, though, they may see a grouping of candidate cut-points that have roughly equivalent λ's. If they are statistically indistinguishable from one another and have equivalent AICs, then the highest λ, which maximizes the selection criteria, will be the closest to the true cutoff in the population. Thus, there is a trade-off between comprehensiveness and distinguishability that needs to be considered in order to provide the best results. Starting with percentiles and then increasing granularity as needed is recommended.

The issue of quantile selection highlights how probabilistic a method to find unknown discontinuities is, as in some cases the exact cut-point will never be known unless infinite quantiles are used with extremely rich and expansive datasets. Lacking that, UCRDD/UCRKD, as well as any other method seeking to estimate an unknown quantity, can only provide an approximation. The key to selecting a threshold under such conditions, as is the key to statistical inference more broadly, is substantive expertise.

3.5 Description, Inference, & the Distribution of LATEs

Finally, with the best cut-point determined, researchers can implement their model with their implementation data and make a statistical inference about the effect of some causal variable. This is as simple as presenting the results from the best candidate cut-point model and making an inference based on the size of the LATE and its statistical significance. If inference is not the primary concern, researchers can simply report the finding of an unknown discontinuity/threshold which can inform future research or clarify existing theories. Researchers that already have a known discontinuity or threshold can also benefit from the use of UCRDD/UCRKD to provide further context and assess the internal validity of their RDD/RKD models. For example, a researcher interested in the effect of college attendance on future income may know that admissions criteria for a certain university is a 3.0 GPA in High School. Using the 3.0 figure as the cut-point may return a statistically significant result, but there exists the possibility that other GPA cutoffs - not publicly discussed by a university's admissions office - have an even greater chance of resulting in an admission. Thus, it may be of interest to researchers to estimate the LATE at those thresholds and compare the different effect sizes. Significant discontinuities beyond or below the known exogenous treatment rule should be a concern for applied researchers using RDD/RKDs as it indicates the treatment

rule is not neat and well-behaved. UCRDD/UCRKD provides an easy means with which to check this.

4. Accuracy of UCRDD/UCRKD

With the method explained, this paper now turns towards showing the accuracy and applications of UCRDD/UCRKD. As UCRDD/UCRKD is far more analytically tractable, more flexible, and less computationally intensive than existing methods (e.g. Porter & Yu 2015; Hansen 2017; Herlands et al. 2018; Boehnke & Bonaldi 2019; Tanu 2020), ensuring a high degree of accuracy is essential. Because there is no way to absolutely verify if an estimated threshold is, in fact, the true threshold without the true threshold being known, testing UCRDD/UCRKD with simulated data as well as with existing known cut-points is the only avenue to assess its accuracy. This section tests the accuracy of UCRDD/UCRKD with both simulation and replication data by finding simulated and known cut-points.

4.1 Simulation Data

In this section, I simulate assignment and outcome variables, force a discontinuity in the outcome variable at at a specified value in the assignment variable, and recover that threshold with UCRDD. I create a simulated assignment variable, X, which is 1000 draws from a uniformly distributed random variable ranging from (-1,1).

$$X \sim U(-1, 1) \tag{8}$$

I then create a dependent variable, Y, that is a function of the assignment variable and some random noise with a forced discontinuity at 0 in the assignment variable. The joint distribution is depicted in *Figure 3*.

I next partition X into percentiles. Since the discontinuity is quite large, I am willing to trade some granularity for more distinguishability. For this simulation, I use a triangular kernel localized linear RDD model:

$$Y_i = \beta_0 + \beta_1 T_i + f(X_i) + \epsilon_i \tag{9}$$

where $f(X_i)$ is a flexible function of the running variable. I follow the procedure outlined in section 3 exactly.[6] I run 100 models to test each of the 100 candidate cut-points. Examining the distribution of λ's, or selection criteria, the maximum in this case is candidate cut-point 51, which takes a value of 0.0003, a number that is extremely close to zero. Recall that zero

---

[6]Since I am not worried about causal inference, I do not split the data into training and testing datasets for this example.

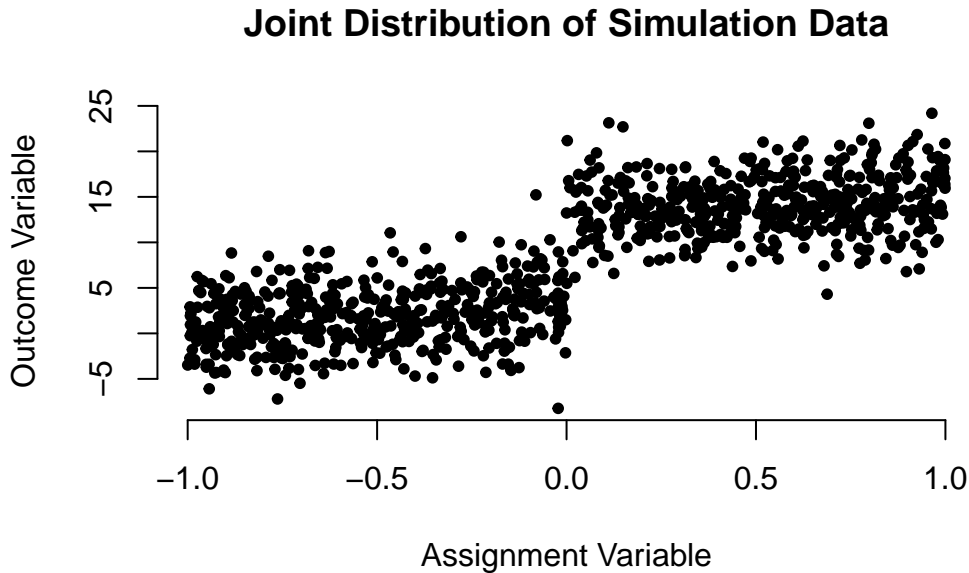## Joint Distribution of Simulation Data

Figure 3: Scatterplot of Simulation Data

is the true cut-point that we are trying to recover. The histogram in *Figure 4* illustrates the distribution of cut-points.

The candidate cut-points that are far from the actual true cut-point have very low relative values of λ as compared to the "best" candidate cut-point. As can be seen in the histogram depicting the distribution of λ's in *Figure 4*, there is clearly a "best" cut-point of the candidate cut-points. Its value of selection criteria is far above any other candidate threshold. It's $\lambda = 3.702656e + 24$, a massive number, compared to the other candidate cut-points that have λ's close to zero. UCRDD/UCRKD was able to recover this true cut-point by iterating through a series of candidate models and saving the ratio of the coefficient to the p-value in each model. *Figure 5* depicts the final analysis with the large discontinuity clear.

4.2 Replication Data

In this section, I further demonstrate the accuracy of UCRDD/UCRKD by using the approach to detect a known cut-point in a published study. In other words, I see if the method can accurately detect an already known threshold. I use replication data from Fowler & Hall (2016) as an example because it relates to elections - the most common substantive field RDDs are applied to in political science - and it offers the opportunity to detect a series of cut-points as there are multiple RDDs in the paper.
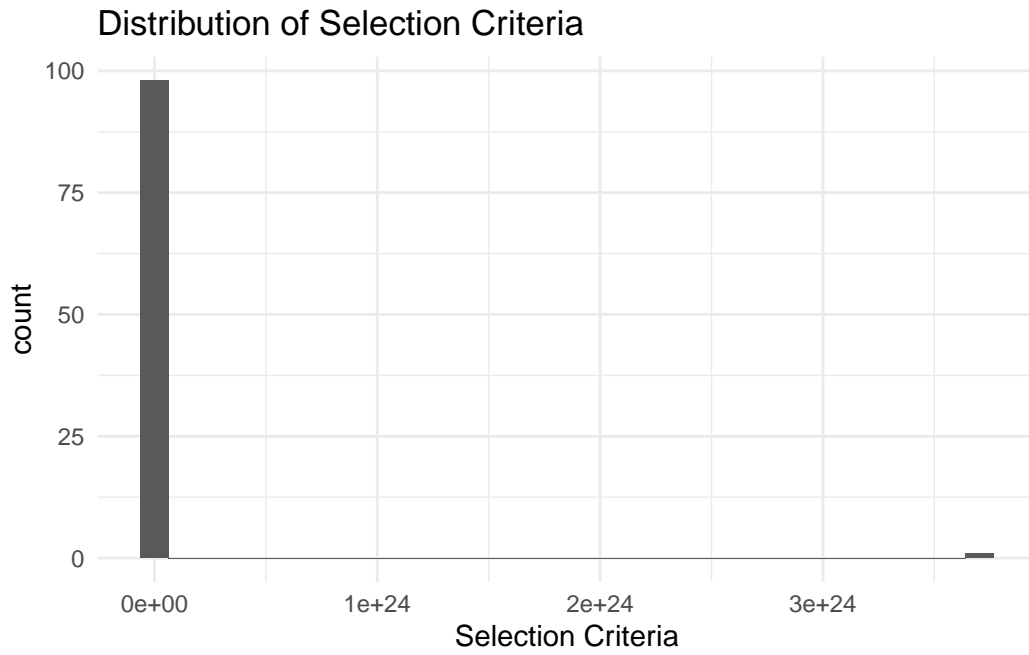
Distribution of Selection Criteria

Figure 4: Distribution of Selection Criteria with Simulation Data

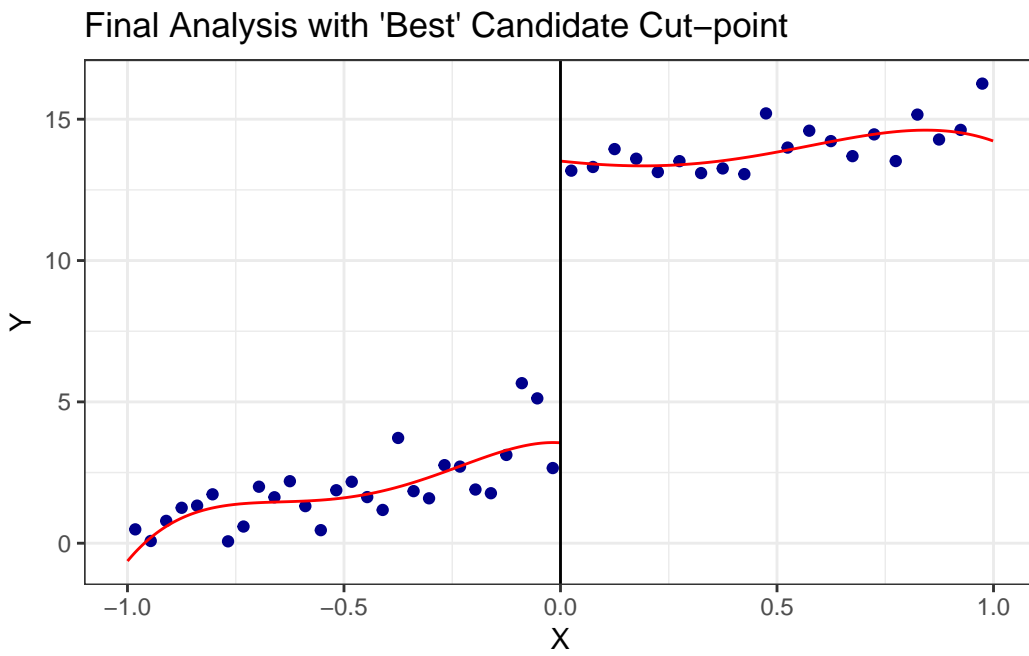## Final Analysis with 'Best' Candidate Cut−point

Figure 5: Final Analysis with Simulation Data

Fowler & Hall (2016) test the convergence hypothesis, or the expectations of Black (1948) and Downs (1957) that the preferences of political parties in single-member plurality electoral systems should converge on the median voter's preferences. If ideological convergence of elites is true, then the voting behavior of Republican and Democratic Members of Congress in more competitive districts should be more similar to one another than their colleagues from safer districts. Fowler & Hall (2016) find that this is not the case, and, in fact, there is a substantively large and statistically significant discontinuity in the probability that a member casts roll calls conservatively at the vote threshold needed to win an election. UCRDD/UCRKD finds that discontinuity with minimal candidate cut-points. I show this with five different policy areas as discussed by Fowler & Hall (2016): defense, welfare, education, public transportation, and energy. In each of these five scenarios, as shown in *Table 1*, the true-cut point is recovered. *Figure 6* through *Figure 10* show the distribution of selection criteria for each policy domain. For each policy domain, the cut-point is .5, and UCRDD/UCRKD finds .5 for each. The 50% two-party vote threshold is identified in each instance, and only 100 candidate cut-points were tested. UCRDD/ UCRKD is extremely accurate as it is able to identify true cut-points in scenarios where we ex ante know the exact true threshold. Thus, in scenarios where a cut-point is unknown, or a discontinuity in behavior exists without some clear treatment rule, UCRDD/UCRKD will return a reliable estimate of the true cut-point in the population. One such example is considered in the next section.
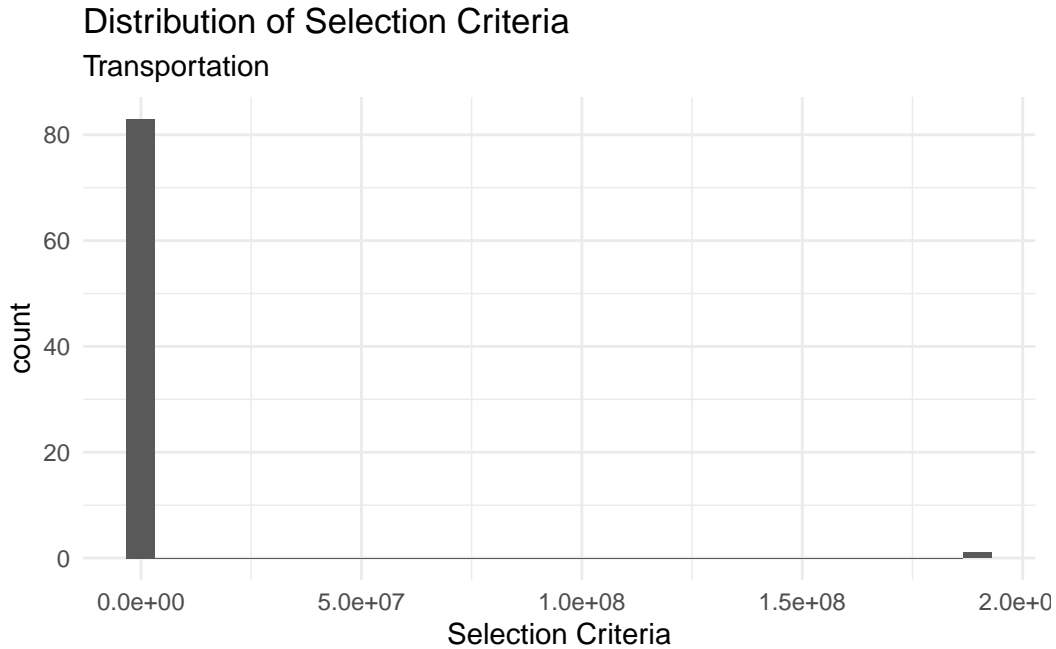

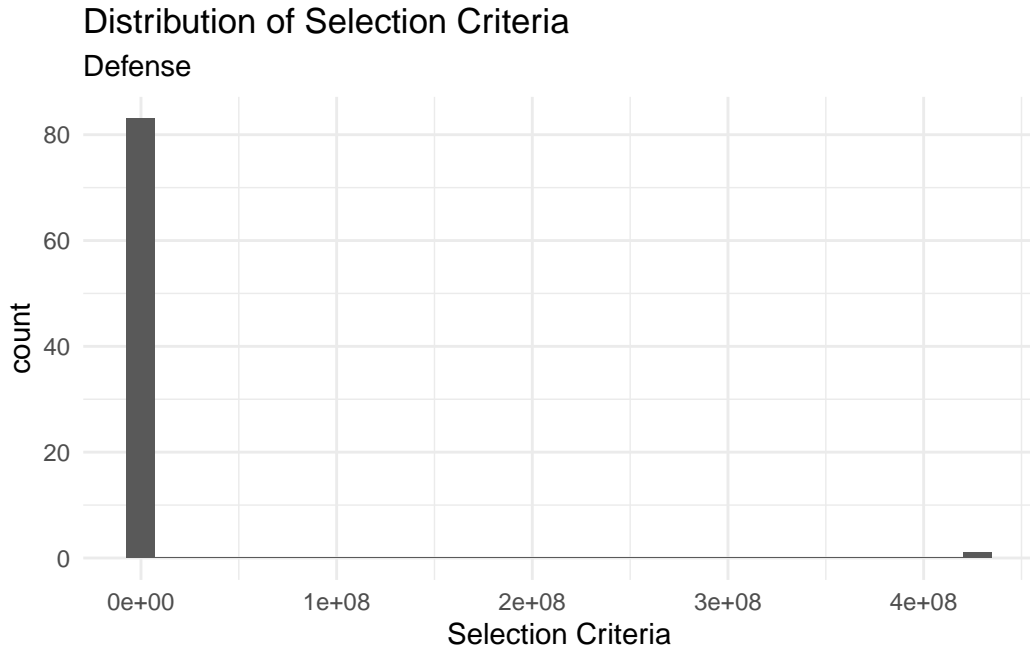Figure 6: Distribution of Selection Criteria for Transportation

17

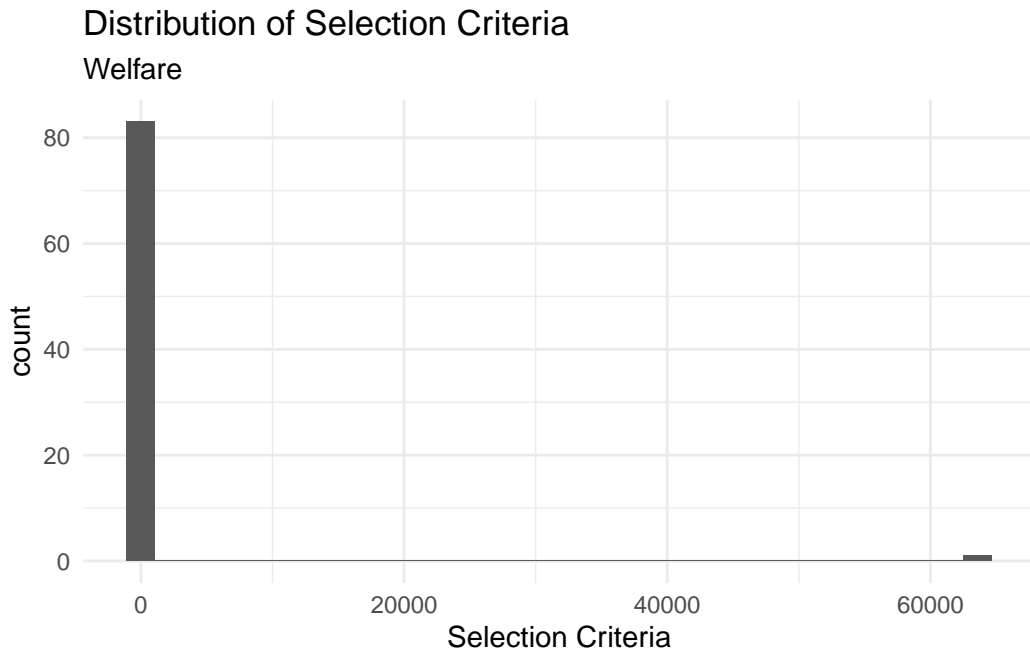Figure 7: Distribution of Selection Criteria for Defense

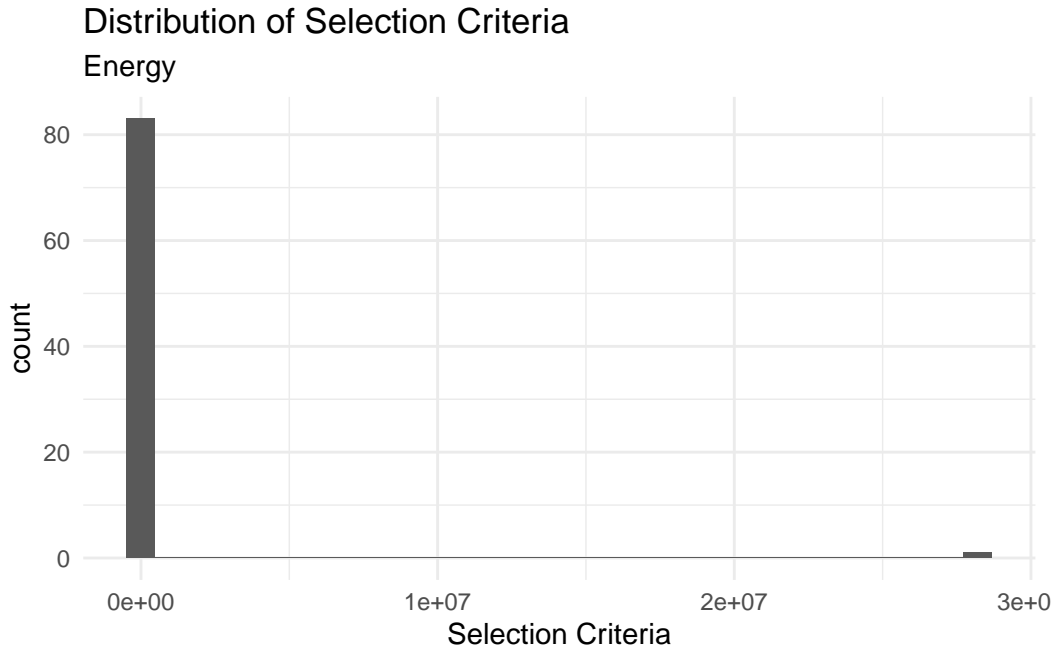Figure 8: Distribution of Selection Criteria for Welfare

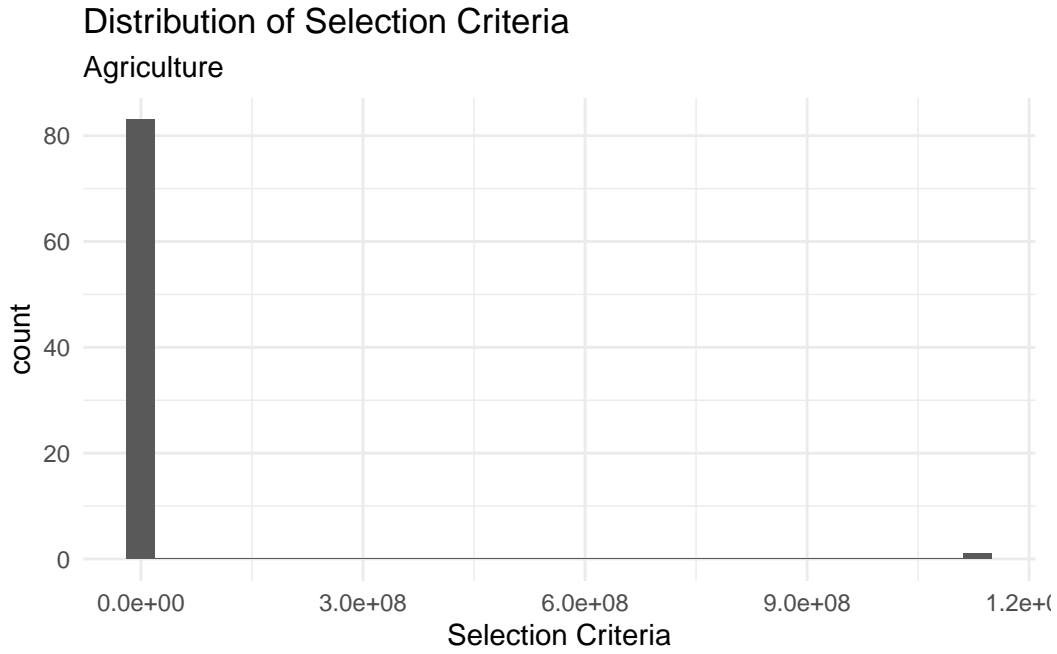Figure 9: Distribution of Selection Criteria for Energy

Figure 10: Distribution of Selection Criteria for Agriculture

Table 1

Policy Domain	Estimated Cut-point	Method Correct
Transportation	.5	✓
Defense	.5	✓
Welfare	.5	✓
Energy	.5	✓
Education	.5	✓

5. Minority Districts & Minority Candidate Performance

The key applications of estimating unknown discontinuities/thresholds for political scientists are: (1) to find discontinuities in behavior without knowledge of an exact threshold; (2) finding sorting, or a discontinuity/threshold, before or after a cut-point when a threshold is known; and (3) estimating the causal effects of policies with concealed or otherwise unobserved eligibility criteria. I consider in this section the first contribution as a largely descriptive tool to find threshold dependent, or tipping-point like, behavior with an unknown treatment rule. Specifically, I use the UCRKD methodology illustrated in this paper to both determine at what continuous value of the proportion of minority voters in a district indicates treatment as a minority district and the effect of treatment on support for minority candidates. I use county level demographic estimates and voting return data from the 2012 presidential election to determine at what percentage of a district that is minority most significantly increases the chances of the district voting for Barack Obama.

One extremely useful application of UCRDD/UCRKD is to determine how to define treatment status when a clear rule is not readily apparent. The causal inference literature consistently (e.g. Morgan & Winship 2007; Angrist & Pischke 2008; Cunningham 2021) characterizes research questions in terms of dichotomous treatment indicators. An issue arises, however, when considering a large portion of all empirical research in political science and economics utilizes continuous measures, rather than categorical. Continuous measures can make clean assignment into treatment and control groups extremely difficult. For example, if a researcher is interested in questions of retrospective/economic voting in the same vein as Fiorina (1981) or Achen & Bartels (2016), the main hypothesis is that high real disposable income leads to more support for the incumbent party, but when does real disposable income become "high"? How can we reliably separate "low" and "high" real disposable income figures? The same issue arises when considering minority districts. Does a district need to be above 50% minority to be a minority district as the majority-minority distinction indicates? Or is some other proportion better for declaring treatment status? UCRKD provides an answer.

5.1 Background & Related Literature

In the classic Schelling model of segregation (Schelling 1971), economist Thomas Schelling presented a simple model to explain segregation in a society independent of direct government interventions such as the Jim Crow-era policies of the US. Schelling (1971) argued that there is some "tipping-point" of minority residents in a community that causes white citizens to leave. Thus, de facto segregation sets in as whites voluntarily leave the community to live in a less diverse neighborhood. Card et al. (2008) utilized a variety of methodological approaches approximating UCRDD/UCRKD and methods like it to estimate this exact threshold. In the same vein, political scientists have long been interested in the voting behavior of minority, or majority-minority, districts, raising the empirical question of at what proportion of minority voters significantly increases the prospects of a minority candidate being elected.

Even before the passage of the Voting Rights Act in 1965, there has been much attention both in the media and in academia about how to ensure normative representation of minority voters in American political institutions. Early research from Key (1949), who famously argues that race is the single most defining aspect of southern politics, concludes that the counties that had the largest proportions of African-American voters favored Smith over Hoover in the 1928 presidential contest and were less likely to support Strom Thurmond's State's Rights campaign in 1948. Keech (1968) argues that at low levels of minority voters representatives take little notice to the interests of minorities until the percentage reaches around 30% or so, when polarization becomes more likely and the representative actively disfavors the minority voter preferences. Black (1978) and Bullock (1981) generally find support for Keech's theory. Using a more sophisticated methodology, Cameron et al. (1996) find that candidates of color can have a non-trivial chance of securing elected office running in districts that have less than 50% minority voters, and that, outside of the South, districts that are 47% Black maximizes substantive Black representation. This sampling of research focuses on the descriptive-substantive representation connection and would be greatly benefited by having a more clear distinction between minority and non-minority districts. The relationship between minority voters and a district's political behavior is a long-standing point of interest for scholars of politics, and scholars have also found the descriptive-substantive representation connection incentive to be particularly salient to minority voters as more minority districts tend to elect minority candidates.

Williams (1990), Terkildsen (1993), and Reeves (1997) all focus on how more African-American voters can increase the probability that African-American candidates are elected. The effect is not limited to African-Americans, however, as many scholars have found identity connections to be a leading determinant of vote choice (Barreto (2010); Fraga (2016); McConnaughy et al. 2010; Perez 2015). Boudreau et al. (2019) argue that when voters share a racial/ethnic identity with a candidate, ideological concerns are less important. The key thread at the center of this research into how the prevalence of minority voters influences minority candidate performance is the question of what a minority district is. More specifically, at what percent minority voters does a district become a "minority district"? This question

is at the focal point of intense academic scrutiny, but, considering the wealth of research that finds that Black members of Congress tend to be more responsive to Black constituents and are more likely to cast roll call votes in their favor than white Members (Canon 1999 & Tate 2003), it is clear why African-American voters would prefer to have an African-American representative. Further, Dawson (1994) argues that notions of linked fate, or that members of the African-American community see their personal fate as tied to the fate of their race, drive Black political behavior. In the framework of Mansbridge (1999), minority representatives bring unique and beneficial perspectives to political decision-making and increase the legitimacy of the government in the eyes of minority citizens. These findings illustrate the strong incentives minority voters have to support minority candidates.

5.2 Data & Methodology

This section uses the method proposed in this paper to offer an answer to the question of how to define a minority district. UCRKD provides the opportunity to both determine a tipping point and the LATE of minority districts on minority candidate vote shares. I use UCRKD to determine the threshold of minority voters that leads to significantly better performance of minority candidates. I use Obama's vote share in 2012 as a function of the percentage of white voters in a district to determine at what threshold should districts be classified as a minority district. To find the discontinuity, I use basic US Census demographic and official voting return data, as prepared by the University of Chicago's Center for Spatial Data Science.[7] I first subset the data to 2012 and estimate the basic regression kink model:

$$Dem\_Vote_i = \beta_0 + \beta_1(White\%_i - \gamma)_- + \beta_2(White\%_i - \gamma)_+ + f(White\%_i) + \epsilon_i \qquad (10)$$

Where Dem_Vote_i is the vote for Barack Obama in county i, $White\%_i$ is the percentage of whites in a county, and γ_i is the threshold to be estimated. In equation (10), the slope in terms of $White\%_i$ is β_1 for any value of $White\%_i$ less than γ, and is β_2 for any value of $White\%_i$ greater than γ. As equation (10) depicts a basic regression kink model, the regression function is continuous with respect to $White\%_i$, but is discontinuous - changing in slope - at $White\%_i = \gamma$. UCRKD treats γ as a unknown parameter to be estimated, when in the traditional RKD, γ is exogenously given, fixed, and known. With the inclusion of $f(White\%_i)$, the model above is agnostic about the functional form of the running variable. Unlike current best practices for RDDs, Card et al. (2016) argue that a uniform kernel, not a triangular kernel, is optimal for RKDs (Card et al. 2016). Therefore, I utilize a uniform kernel density estimator for model estimation.

For the purposes of this illustration, I omit the inclusion of any covariates. To ensure a high degree of precision, I use 1000 candidate cut-points. Finally, the choice of the 2012 presidential election is deliberate as 90+% of African-American voters and 70+% of Latino

---

[7]The full dataset can be found at https://geodacenter.github.io/data-and-lab/county_election_2012_2016-variables/

voters supported Obama over Mitt Romney.[8]. The re-election of the first Black president was extremely salient for minority voters. Thus, I can leverage this to see at what threshold of minority population did districts significantly increase in likelihood of voting for Obama.
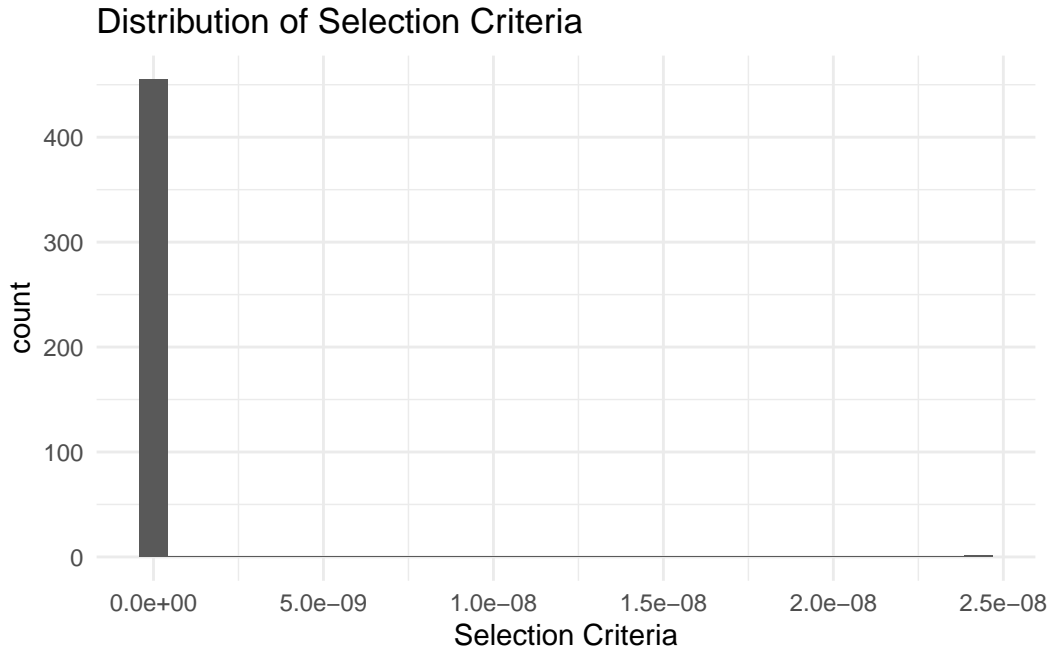
5.3 Results

Figure 11: Distribution of Selection Criteria

Of the 1000 candidate cut-points tested, the 57.4% white threshold had by far the highest λ_i. *Figure 11* showcases the full distribution of selection criteria. The size of slope change and degree of statistical significance was highest at the 57.4% threshold as compared to all other possible values that the assignment variable could take. Visually, as shown in *Figure 12*, there is a clear kink in the slope of the continuous regression function that shows, once districts exceed 57.4% white, vote shares for Barack Obama significantly decrease. Once districts pass the threshold of 57.4% white voters, the likelihood of minority representatives being elected drastically decreases. In other words, districts that exceed 42.6% minority voters see a significant increase in support for Obama. Thus, the estimated treatment rule suggests that the districts which have less than 57.4% white voters should be classified as a minority district, with these data. This is 7.6 percentage points lower than a simple 50% threshold district and 4.4 percentage points lower than the figure Cameron et al. (1996) argue maximizes minority substantive representation. Assuming minority districts can be determined by some binary

---

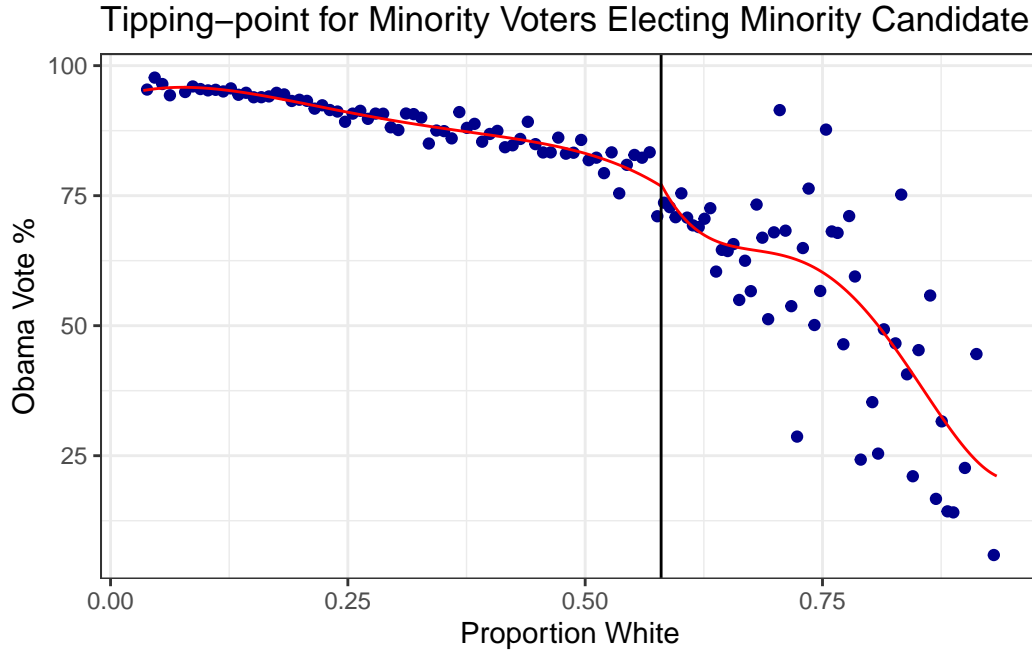## Tipping–point for Minority Voters Electing Minority Candidate

Figure 12: Tipping-point for Minority Districts

treatment rule, districts only need to be 42.6% minority to significantly increase the likelihood that a minority candidate is elected. These results can be useful for understanding how Black candidates perform in major American elections, as districts that just pass the 42.6% minority threshold will be far more likely to support minority candidates than less diverse districts. The coefficient on the treatment indicator is .026. In RKDs, this coefficient is interpreted as the change in slope at the threshold. Therefore, there is a 2.6 percentage point change in slope at the 57.4% white threshold. This coefficient is statistically significant at the $p < .05$ level. *Table 2* displays the results of the final regression kink analysis.

Table 2

Coefficient	Standard Error	p-value	95% CI	N
0.026	.013	0.044	[0.001 , 0.052]	2881

The next step - beyond the scope of this paper - is to apply this rule in a number of different elections with a wider range of minority candidates. While a more robust analysis is beyond the scope of this paper, this section has highlighted the flexibility and usefulness of UCRDD/UCRKD to answer important questions in political science.

6. Discussion & Conclusion

Regression discontinuity designs and regression kink designs are held in high regard as essential instruments in the econometrics toolkit. Each offers a reliable way, under a weak set of assumptions, to identify a causal effect given a known exogenous threshold. With a known treatment rule, the causal effects of policy participation or electoral victories can be reliably identified. However, the requirement of a known threshold limits the possible applications of RDDs and RKDs. UCRDD/UCRKD offers the most flexible, analytically tractable, and easily implemented method to estimate unknown cut-points in both RDDs and RKDs. Further, the method can handle variations of RDD/RKDs such as fuzzy and sharp models and models with covariates with minimal changes in procedure. The flexibility and ease of implementation provides an advantage over Porter & Yu (2016), Hansen (2017), Herlands et al. (2018), and Boehnke & Bonaldi (2019).

There are numerous instances in which this method can be applied to better understand political behavior that have not been discussed in this paper. Some examples include determining the exact democracy score that limits the probability that a country engages in war; determining the debt-to-GDP ratio that causes GDP-per-capita to drastically decrease; or even a re-examining of Schelling (1971) and Card et al. (2008). Accordingly, it is the goal of the author for researchers to extend UCRDD/UCRKD to a diverse range of research questions throughout political science and the social sciences more broadly. The method provided herein this paper allows scholars to find unknown behavioral patterns by learning directly from their data. Determining what a minority district is, for example, would be extraordinarily difficult, and perhaps arbitrary, without UCRDD/UCRKD. Using the fact that minority districts supported Barack Obama at unprecedented levels, we can determine what a minority district is in the context of presidential electoral politics. Other potential applications of UCRDD/UCRKD are numerous.

References

Acharya A, Bansak K, Hainmueller J. 2021. "Combining Outcome-based and Preference-based Matching: the G-constrained Priority Mechanism. *Political Analysis.*

Achen, Christopher & Larry Bartels. 2016. *Democracy for Realists.* Princeton University Press.

Andrews, D. W. K. 1993. "Tests for Parameter Instability and Structural Change with Unknown Change Point. *Econometrica.* 61(4), 821–856.

Angrist, J. D., & Lavy, V. 1999. "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement." *The Quarterly Journal of Economics.*

114(2), 533–575.

Angrist, J. D. & Jorn-Steffen Pischke. 2008. *Mostly Harmless Econometrics.* Princeton University Press.

Barreto, Matt. 2010. *Ethnic Cues.* Ann Arbor: University of Michigan Press

Beck N, King G, Zeng L. 2000. "Improving Quantitative Studies of International Conflict. *American Political Science Review.* 94:21–36.

Black, Sandra E. 1999. "Do Better Schools Matter? Parental Valuation of Elementary Education." *Quarterly Journal of Economics.* 114 (2): 577–99.

Black, D. 1948. "On the Rationale of Group Decision-making". *Journal of Political Economy,* 56(1), 23–34.

Black, Merle. 1978. "Racial Composition of Congressional Districts and Support for Federal Voting Rights in the American South." *Social Science Quarterly* 59:435-50.

Boehnke, Jorn and Pietro Bonaldi. 2019. "Synthetic Regression Discontinuity Design: Estimating Treatment Effects Using Machine Learning." *Manuscript.*

Boudreau, C., Elmendorf, C. S., & MacKenzie, S. A. 2019. "Racial or Spatial Voting? The Effects of Candidate Ethnicity and Ethnic Group Endorsements in Local Elections." *American Journal of Political Science.* 63(1), 5–20.

Bullock, Charles. 1981. "Congressional Voting and the Mobilization of a Black Electorate in the South." *Journal of Politics* 43:662-82.

Cameron, C., Epstein, D., & O'Halloran, S. 1996. "Do Majority-Minority Districts Maximize Substantive Black Representation in Congress?" *American Political Science Review.* 90(4), 794–812.

Caner, M., and Hansen, B. E. 2004. "Instrumental Variable Estimation of a Threshold Model." *Econometric Theory.* 20, 813–843.

Canon, David T. 1999. *Race, Redistricting, and Representation: The Unintended Consequences of Black Majority Districts.* Chicago, IL: University of Chicago Press.

Card, D., Mas, A., & Rothstein, J. 2008. "Tipping and the Dynamics of Segregation."

The Quarterly Journal of Economics. 123(1), 177–218.

Card, David, David Lee, Zhuan Pei, & Andrea Weber. 2016. "Regression Kink Design: Theory and Practice." Working Paper 22781. *NBER.*

Casini, Alessandro and Pierre Perron. 2018. "Structural Breaks in Time Series." *Manuscript.* Boston University.

Chan, K.-S. 1990. "Testing for Threshold Autoregression." *The Annals of Statistics.* 18, 1886–1894.

——— 1991. "Percentage Points of Likelihood Ratio Tests for Threshold Autoregression." *Journal of the Royal Statistical Society.* Series B, 53, 691-696.

Chan, K.-S., and Tsay, R. S. 1998. "Limiting Properties of the Least Squares Estimator of a Continuous Threshold Autoregressive Model." *Biometrika.* 45, 413–426.

Cook, Thomas D. "'Waiting for Life to Arrive': A history of the Regression Discontinuity Design in Psychology, Statistics and Economics." *Journal of Econometrics.* 142 (2), 636-654.

Cunningham, Scott. 2021. *Causal Inference: The Mixtape.* Yale University Press.

de la Cuesta, Brandon. & Kosuke Imai. 2016. "Misunderstandings About the Regression Discontinuity Design in the Study of Close Elections." *Annual Review of Political Science.* 19:375-96.

Downs, Anthony. *An Economic Theory of Democracy.* Harper & Row.

Fang, Z., and Santos, A. 2014. "Inference on Directionally Differentiable Functions." Working Paper, *University of California San Diego.*

Fiorina, Morris. 1981. *Retrospective Voting in American National Elections.* New Haven: Yale University Press.

Fowler, Anthony and Andrew Hall. 2016. "The Elusive Quest for Convergence." *Quarterly Journal of Political Science.* 11: 131-149.

Fong, Y., Huang, Y., Gilbert, P.B., Permar, S. 2017. "chngpt: Threshold Regression Model Estimation and Inference. *BMC Bioinformatics* 18, 454.

Fraga, Bernard L. 2016. "Redisricting and the Causal Impact of Race on Voter Turnout." *Journal of Politics.* 78(1): 19-34.

Grimmer, Justin, Margaret E. Roberts, and Brandon M. Stewart. 2021. "Machine Learning for Social Science: An Agnostic Approach." *Annual Review of Political Science.* 24. 395-419.

Hansen, B. E. 1996. "Inference When a Nuisance Parameter is not Identified Under the Null Hypothesis." *Econometrica.* 64, 413–430.

Hansen, Bruce E. 2017. "Regression Kink With an Unknown Threshold." *Journal of Business & Economic Statistics.* 35:2, 228-240.

Hahn, Jinyong, Petra Todd, and Wilbert van der Klaauw. 2001. "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design." *Econometrica* 69 (1): 201–9.

Herlands, William, Edward McFowland III, Andrew Gordon Wilson, Daniel B. Neill. "Automated Local Regression Discontinuity Designs Discovery. *KDD '18: The 24th ACM SIGKDD International Conference on Knowledge Discovery Data Mining*

Hillard D, Purpura S, Wilkerson J. 2008. "Computer-assisted Topic Classification for Mixed-methods Social Science Research. *Journal of Information Technology Politics* 4:31–46.

Hill DW Jr., Jones ZM. 2014. "An Empirical Evaluation of Explanations for State Repression." *American Political Science Review* 108:661–87.

Hong, H., and Li, J. 2015. "The Numerical Directional Delta Method," Working Paper. Stanford University

Imbens, Guido W. and Karthik Kalyanaraman. 2012. "Optimal Bandwidth Choice for the Regression Discontinuity Estimator. *The Review of Economic Studies.* 79(3):933–959.

Imbens, Guido W. and Thomas Lemieux. 2008. "Regression Discontinuity Designs: A Guide to Practice. *Journal of Econometrics.* 142(2):615–635, 2008.

Kaufman AR, Kraft P, Sen M. 2019. "Improving Supreme Court Forecasting Using Boosted Decision Trees. *Political Analysis* 27:381–87.

Key, V. O. 1949. *Southern Politics in State and Nation.* University of Tennessee Press.

Keech, William R. 1968. *The Impact of Negro Voting.* Chicago: Rand McNally.

Lee D, Moretti E, Butler M. 2004. "Do Voters Affect or Elect Policies? Evidence

From the US House. *Quarterly Journal of Economics.* 119(3):807–59.

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

Lee David S. 2008. "Randomized Experiments from Non-Random Selection in U.S. House
Elections." *Journal of Econometrics* 142(2):675–97.

Lee, S., Seo, M. H., and Shin, Y. 2011. "Testing for Threshold Effects in Regression
Models." *Journal of the American Statistical Association.* 106, 220–231.

Mansbridge, Jane. 1999. Should Blacks Represent Blacks and Women Represent Women?
A Contingent "Yes." *Journal of Politics.* 61:628–57.

McCrary J. 2008. "Manipulation of the Running variable in the Regression Discontinuity
Design: A Density Test. *Journal of Econometrics.* 142(2):698–714.

McConnaughy, Corrine M., Ismail K. White, David Leal, and Jason Casellas. 2010.
"A Latino on the Ballot: Explaining Co-Ethnic Voting among Latinos and White
Americans' Response." *Journal of Politics* 72(4): 1 199-1211.

McCulloch, Warren S. and Walter Pitts. 1943. "A Logical Calculus of the Ideas Immanent
in Nervous Activity. *The Bulletin of Mathematical Biophysics* 5(4):115–133.

Montgomery JM, Olivella S. 2018. "Tree-based Models for Political Science Data.
American Journal of Political Science. 62:729–44.

Morgan, Stephen L. & Christopher Morgan. 2007. *Counterfactuals and Causal Inference*
Cambridge University Press.

Nelson LK. 2017. "Computational Grounded Theory: A Methodological Framework.
Social Methods Research. 49:3–42.

Nielsen, Helena Skyt, Torben Sorensen, and Christopher R. Taber. 2010. "Estimating
the Effect of Student Aid on College Enrollment: Evidence from a Government
Grant Policy Reform." *American Economic Journal: Economic Policy.* 2 (2), 185–215.

Perez, Efren 0. 2015. "Xenophobic Rhetoric and Its Political Effects on Immigrants
and Their Co-Ethnics." *American Journal of Political Science* 59(3): 549-64.

Perron, P. 2006." Dealing with structural breaks." In *Palgrave Handbook of Econometrics
Vol. 1: Econometric Theory*, K. Patterson and T.C. Mills (eds.), Palgrave Macmillan,

278 - 352.

Perron, P. 2008. "Structural change." In *The New Palgrave Dictionary of Economics*, 2nd ed, S. Durlauf and L. Blume (eds.).

Porter, Jack and Ping Yu. 2015. "Regression Discontinuity Designs with Unknown Discontinuity Points: Testing and Estimation." *Journal of Econometrics.* 189(1):132–147.

Pretis, F., Reade, J. J., & Sucarrat, G. 2018. Automated General-to-specific (GETS) Regression Modeling and Indicator-saturation for Outliers and Structural Breaks. *Journal of Statistical Software.* 86(3).

Reeves, Keith. 1997. *Voting Hopes Or Fears?: White Voters, Black Candidates And Racial Politics In America.* Oxford University Press.

Schelling, Thomas C. 1971. "Dynamic Models of Segregation." *The Journal of Mathematical Sociology.* 1(2), 143-186.

Skovron, C. and Titiunik, R. 2015. "A Practical Guide to Regression Discontinuity. Working Paper, *Department of Political Science, University of Michigan.*

Stewart BM, Zhukov Y. 2009. "Use of Force and Civil-Military Relations in Russia: An Automated Content Analysis." *Small Wars Insurg.* 20:319–43.

Stommes, D., Aronow, P. M., & Sävje, F. 2023. "On the Reliability of Published Findings using the Regression Discontinuity Design in Political Science. *Research & Politics*, 10(2).

Tanu, Tanvir Ahmed Khan. 2020. "Regression Discontinuity Design with Unknown Cutoff: Cutoff Detection & Effect Estimation." Manuscript. *East West University.*

Tate, Katherine. 2003. *Black Faces in the Mirror: African Americans and Their Representatives in the U.S. Congress.* Princeton, NJ: Princeton University Press.

Terkildsen, Nadia. 1993. "When White Voters Evaluate Black Candidates: The Processing Implications of Candidate Skin Color, Prejudice, and Self-Monitoring." *American Journal of Political Science* 37:1032-53.

Thistlethwaite, Donald L. and Donald T. Campbell. "Regression-discontinuity Analysis: An Alternative to the ex-post-facto Experiment." *Journal of Educational Psychology.*

51(6):309–317.

Williams, Linda F. 1990. "White/Black Perceptions of the Electability of Black

Political Candidates." *National Political Science Review* 2:145-64.

Williams NW, Casas A, Wilkerson JD. 2020. *Images as Data for Social Science Research:*

An Introduction to Convolutional Neural Nets for Image Classification.

Cambridge, UK: Cambridge University Press.

Appendix A

This appendix displays the code to apply the method and will form the basis of an R package that will allow applied researchers to easily implement the procedure.

```r
## Packages
library(rdd)
library(rdpower)
library(tidyverse)
library(tidymodels)

## Function to Find Percentiles
percentile <- function(x) {

  unique(quantile(x, probs = c(seq(.01, 1, by = .01)), na.rm = T))

}

## Saving Percentiles (Candidate Cut-points) - Only saving 100, but can save more
## In the final package, this number will be subject to user input

g <- percentile(x)

## Initializing Data Frame
f <- data.frame(1)

## Creates the Candidate (x-c)'s
for (i in 1:length(g)) {

  p <- paste0("X_", i)
  c <- assign(p, (x - g[i]))
  f <- data.frame(c, f)
```

```r
}

## Drop the initialized column
f <- f %>%
  select(-X1)

## Initializing Variables for local linear regression loop
models <- list()
f <- f %>%
  select(-c)
variables <- setdiff(names(f), c("row_id"))

## Local Linear Regression Loop
for (var in variables) {
  models[[var]] = RDestimate(
    as.formula(paste0("y ~ ", var)),
    data = f
  )
}

## Selection Function
selection_criteria <- function(x) {
  d <- abs(mean(models[[x]]$est)/mean(models[[x]]$p))
  print(d)
}

## Creating Cut-point Names
cuts <- colnames(f)

## Finding Optimal Cut-point
s <- sapply(cuts, selection_criteria, USE.NAMES = TRUE)

s[s== max(s)]
```