Testing Continuity of a Density via g-order statistics in the Regression Discontinuity Design[*]

Federico A. Bugni
Department of Economics
Duke University
federico.bugni@duke.edu

Ivan A. Canay
Department of Economics
Northwestern University
iacanay@northwestern.edu

September 11, 2018

Abstract

In the regression discontinuity design (RDD), it is common practice to assess the credibility of the design by testing the continuity of the density of the running variable at the cut-off, e.g., McCrary (2008). In this paper we propose a new test for continuity of a density at a point based on the so-called g-order statistics, and study its properties under a novel asymptotic framework. The asymptotic framework is intended to approximate a *small* sample phenomenon: even though the total number n of observations may be large, the number of *effective* observations local to the cut-off is often small. Thus, while traditional asymptotics in RDD require a growing number of observations local to the cut-off as $n \to \infty$, our framework allows for the number q of observations local to the cut-off to be *fixed* as $n \to \infty$. The new test is easy to implement, asymptotically valid under weaker conditions than those used by competing methods, exhibits finite sample validity under stronger conditions than those needed for its asymptotic validity, and has favorable power properties against certain alternatives. In a simulation study, we find that the new test controls size remarkably well across designs. We finally apply our test to the design in Lee (2008), a well-known application of the RDD to study incumbency advantage.

KEYWORDS: Regression discontinuity design, g-ordered statistics, sign tests, continuity, density.

JEL classification codes: C12, C14.

1   Introduction

The regression discontinuity design (RDD) has been extensively used in recent years to retrieve causal treatment effects - see Lee and Lemieux (2010) and Imbens and Lemieux (2008) for exhaustive surveys. The design is distinguished by its unique treatment assignment rule where individuals receive treatment when an observed covariate, known as the running variable, crosses a known cut-off. Such an assignment rule allows nonparametric identification of the average treatment effect (ATE) at the cut-off, provided that potential outcomes have continuous conditional expectations at the cut-off (Hahn et al., 2001). The credibility of this identification strategy along with the abundance of such discontinuous rules have made RDD increasingly popular in empirical applications.

While the continuity assumption that is necessary for nonparametric identification of the ATE at the cut-off is fundamentally untestable, researchers routinely assess the plausibility of their RDD by exploiting two testable implications of a stronger identification assumption proposed by Lee (2008). We can describe the two implications as follows: (i) individuals have imprecise control over the running variable, which translates into the density of the running variable being continuous at the cut-off; and (ii) the treatment is locally randomized at the cut-off, which translates into the distribution of all observed baseline covariates being continuous at the cut-off. The practice of judging the reliability of RDD applications by assessing either of the two above stated implications (commonly referred to as manipulation, or falsification, or placebo tests) is ubiquitous in the empirical literature. Indeed, Table 4 surveys RDD empirical papers in four leading applied economic journals during the period 2011-2015. Out of 62 papers, 43 of them include some form of manipulation, falsification, or placebo test.

This paper proposes a novel test for the null hypothesis on the first testable implication, i.e., the density of the running variable is continuous at the cut-off.[1] The new test has a number of distinctive attractive properties relative to existing methods. First, the test does not require consistent non-parametric estimators of densities and simply exploits the fact that a certain functional of order statistics of the data is approximately binomially distributed under the null hypothesis. Second, our test controls the limiting null rejection probability under fairly mild conditions that, in particular, do not require existence of derivatives of the density of the running variable. In addition, our test is valid in finite samples under stronger, yet plausible, conditions. Third, the asymptotic validity of our test holds under two alternative asymptotic frameworks; one in which the number q of observations local to the cut-off is *fixed* as the sample size n diverges to infinity, and another one where q is allowed to grow as n diverges to infinity. Importantly, both frameworks require the same mild assumptions. Fourth, our test is arguably simple to implement as it only involves computing order statistics and a constant critical value. This contrasts with existing alternatives that require

---

[1]It is important to emphasize that the null hypothesis we test in this paper is neither necessary nor sufficient for identification of the ATE at the cut-off. See Section 2 for a discussion on this.

local polynomial estimation of some order and either bias correction or under-smoothed bandwidth choices. Finally, we have developed a companion `Stata` package to facilitate the adoption of our test.[2]

The construction of our test is based on the simple intuition that, when the density of the running variable is continuous at the cut-off, the fraction of units under treatment and control local to the cut-off should be roughly the same. This means that the number of treated units, out of the q observations closest to the cut-off, is approximately distributed as a binomial random variable with sample size q and probability $\frac{1}{2}$. In order to formalize this intuition, we exploit and develop properties of the so-called g-order statistics (see, e.g., Kaufmann and Reiss, 1992; Reiss, 1989) and employ two asymptotic frameworks that aim at capturing the *small* sample nature of the problem. In the first framework, the number of observations q local to the cut-off is *fixed* as $n \to \infty$. This framework is similar to the one in Canay and Kamat (2018), who in turn exploit results from Canay, Romano and Shaikh (2017), and is the one that we prefer for the testing problem under consideration. However, it is worth noting that the hypotheses we test, the test statistic, the critical value, and most of the formal arguments are different from those in Canay and Kamat (2018) or Canay, Romano and Shaikh (2017). In the second framework, we let q to slowly diverge to infinity with n, in the sense that we require $\frac{q}{n} \to 0$. Notably, we further show that if q diverges to infinity at a faster rate, then our test would fail to control the limiting rejection probability under the null hypothesis, see Remark 4.4. The asymptotic framework where $q \to \infty$ as $n \to \infty$ is similar to the one in McCrary (2008); Otsu et al. (2013); Cattaneo et al. (2017a), among others, and is in line with more traditional asymptotic arguments in non-parametric tests.

From a technical standpoint, this paper has several contributions relative to the existing literature. To start, our results exhibit two important differences relative to Canay and Kamat (2018) that go beyond the difference in the null hypotheses. First, we do not study our test as an approximate randomization test but rather as an approximate two-sided sign-test. This not only requires different analytical tools, but also by-passes some of the challenges that would arise if we were to characterize our test as an approximate randomization test; see Remark 4.2 for a discussion on this. In addition, our approach in turn facilitates the analysis for the second asymptotic framework where q diverges to infinity. Second, we develop results on g-order statistics as important intermediate steps towards our main results. Some of them may be of independent interest; e.g., Theorem 4.1. In addition, relative to the results in McCrary (2008); Otsu et al. (2013); Cattaneo et al. (2017a); our test does not involve consistent estimators of density functions to either side of the cut-off and does not require conditions involving existence of derivatives of the density of the running variable local to the cut-off. Finally, we note that related binomial tests have been recently presented in the RDD context by Cattaneo et al. (2016) and Cattaneo et al. (2017b); see Remark 3.3 for a detailed description. While these papers rely on finite sample arguments to justify their

---

test construction for the hypothesis of local randomization, here we provide a rigorous asymptotic analysis under the two aforementioned asymptotic frameworks for the hypothesis of continuity of a density. To the best of our knowledge, the formal asymptotic results we present are original to this paper.

The remainder of the paper is organized as follows. Section 2 introduces the notation and describes the null hypothesis of interest. Section 3 defines q-order statistics, formally describes the test we propose, and discusses all aspects related to its implementation including a data dependent way of choosing q. Section 4 presents the main formal results of the paper, dividing those results according to the two alternative asymptotic frameworks we employ. In Section 5, we examine the relevance of our asymptotic analysis for finite samples via a simulation study. Finally, Section 6 implements our test to reevaluate the validity of the design in Lee (2008) and Section 7 concludes. The proofs of all results can be found in the Appendix.

2   Setup and notation

Let $Y \in \mathbf{R}$ denote the (observed) outcome of interest for an individual or unit in the population and $A \in \{0, 1\}$ denote an indicator for whether the unit is treated or not. Further denote by $Y(1)$ the potential outcome of the unit if treated and by $Y(0)$ the potential outcome if not treated. As usual, the (observed) outcome and potential outcomes are related to treatment assignment by the relationship

$$Y = Y(1)A + Y(0)(1 - A) . \tag{1}$$

The treatment assignment in the (sharp) RDD follows a discontinuous rule,

$$A = I\{Z \geq \bar{z}\} ,$$

where $Z \in \mathcal{Z} \equiv \operatorname{supp}(Z)$ is an observed scalar random variable known as the running variable and $\bar{z}$ is the known threshold or cut-off value. For convenience we normalize $\bar{z} = 0$, which is without loss of generality as we can always redefine Z as $Z - \bar{z}$. This treatment assignment rule allows us to identify the average treatment effect (ATE) at the cut-off; i.e.,

$$E[Y(1) - Y(0)|Z = 0] .$$

In particular, Hahn et al. (2001) establish that identification of the ATE at the cut-off relies on the discontinuous treatment assignment rule and the assumption that

$$E[Y(1)|Z = z] \quad \text{and} \quad E[Y(0)|Z = z] \quad \text{are both continuous in } z \text{ at } z = 0 . \tag{2}$$

Reliability of the RDD thus depends on whether the mean outcome for units marginally below the cut-off identifies the true counterfactual for those marginally above the cut-off.

The continuity assumption in (2) is arguably weak, but fundamentally untestable. In practice, researchers routinely employ two specification checks in RDD that, in turn, are implications of a stronger sufficient condition proposed by Lee (2008, Condition 2b). The first check involves testing whether the distribution of pre-determined characteristics (conditional on the running variable) is continuous at the cut-off. See Shen and Zhang (2016) and Canay and Kamat (2018) for a recent treatment of this problem. The second check involves testing the continuity of the density of the running variable at the cut-off, an idea proposed by McCrary (2008). This second check is particularly attractive in settings where pre-determined characteristics are not available or where these characteristics are likely to be unrelated to the outcome of interest. Formally, we can state this hypothesis testing problem as

$$H_0 : f_Z^+(0) = f_Z^-(0) \quad \text{vs.} \quad H_1 : f_Z^+(0) \neq f_Z^-(0) \ , \tag{3}$$

where

$$f_Z^+(0) \equiv \lim_{\epsilon \downarrow 0} \frac{1}{\epsilon} P\{Z \in [0, \epsilon)\} \quad \text{and} \quad f_Z^-(0) \equiv \lim_{\epsilon \downarrow 0} \frac{1}{\epsilon} P\{Z \in (-\epsilon, 0)\} \ , \tag{4}$$

provided these limits exist. In RDD empirical studies, the aforementioned specification checks are often implemented (with different levels of formality) and referred to as falsification, manipulation, or placebo tests (see Table 4 for a survey).

In this paper we propose a test for the null hypothesis of continuity in the density of the running variable Z at the cut-off $\bar{z} = 0$, i.e., (3). The new test has two attractive features compared to existing approaches (see, e.g., McCrary, 2008; Otsu et al., 2013; Cattaneo et al., 2017a). First, it does not require commonly imposed smoothness conditions on the density of Z, as it does not involve non-parametric estimation of such a density. Second, it exhibits finite sample validity under certain (stronger) easy to interpret conditions. We discuss these features further in Section 4.

Remark 2.1. Gerard et al. (2016) study the consequences of discontinuities in the density of Z at the cut-off. In particular, the authors consider a situation in which manipulation occurs only for a subset of participants and use the magnitude of the discontinuity of $f(z)$ at $z = 0$ to identify the proportion of always-assigned units among all units close to the cut-off. Using this setup, Gerard et al. (2016) show that treatment effects in RDD are not point identified but that the model still implies informative bounds. ∎

Remark 2.2. It is important to emphasize that a running variable with a continuous density is neither necessary nor sufficient for identification of the average treatment effect at the cut-off. For a discussion of this and some intuitive examples, see Lee (2008) and McCrary (2008). ∎

3   A test based on g-ordered statistics

Let P be the distribution of Z and $Z^{(n)} = \{Z_i : 1 \leq i \leq n\}$ be a random sample of n i.i.d. observations from P. Let q be a small (relative to n) positive integer and $g : \mathcal{Z} \to \mathbf{R}$ be a

measurable function such that $g(Z)$ has a continuous distribution function. For any $z, z' \in \mathcal{Z}$ define $\leq_g$ as

$$z \leq_g z' \quad \text{if} \quad g(z) \leq g(z') .$$

The ordering defined by $\leq_g$ is called a g-ordering on $\mathcal{Z}$. The g-order statistics $Z_{g,(i)}$ corresponding to $Z^{(n)}$ have the property

$$Z_{g,(1)} \leq_g \cdots \leq_g Z_{g,(n)} ,$$

see, e.g., Reiss (1989, Section 2.1) and Kaufmann and Reiss (1992).

To construct our test statistic, we use the sign of the q values of $\{Z_i : 1 \leq i \leq n\}$ that are induced by the q smallest values of $\{g(Z_i) = |Z_i| : 1 \leq i \leq n\}$. That is, for $Z_{g,(1)}, \ldots, Z_{g,(q)}$, let

$$A_{g,(j)} \equiv I\{Z_{g,(j)} \geq 0\} \text{ for } 1 \leq j \leq q , \tag{5}$$

and

$$S_n \equiv \sum_{j \leq q} A_{g,(j)} . \tag{6}$$

The test statistic of our test only depends on the data via S_n and is defined as

$$T_q(S_n) \equiv \sqrt{q} \left| \frac{1}{q} S_n - \frac{1}{2} \right| . \tag{7}$$

In order to describe the critical value of our test it is convenient to recall that the cumulative distribution function (CDF) of a binomial random variable with q trials and probability of success $\frac{1}{2}$ is given by

$$\Psi_q(b) \equiv \frac{1}{2^q} \sum_{x=0}^{\lfloor b \rfloor} \binom{q}{x} I\{b \geq 0\} , \tag{8}$$

where $\lfloor x \rfloor$ is the largest integer not exceeding x. Using this notation the critical value for a significance level $\alpha \in (0, 1)$ is given by

$$c_q(\alpha) \equiv \sqrt{q} \left(\frac{1}{2} - \frac{b_q(\alpha)}{q} \right) , \tag{9}$$

where $b_q(\alpha)$ is the unique value in $\{0, 1, \ldots, \lfloor \frac{q}{2} \rfloor\}$ satisfying

$$\Psi_q(b_q(\alpha) - 1) \leq \frac{\alpha}{2} < \Psi_q(b_q(\alpha)) . \tag{10}$$

The test we propose is then given by

$$\phi(S_n) = \begin{cases} 1 & \text{if } T_q(S_n) > c_q(\alpha) \\ a_q(\alpha) & \text{if } T_q(S_n) = c_q(\alpha) \\ 0 & \text{if } T_q(S_n) < c_q(\alpha) \end{cases} , \tag{11}$$

where

$$a_q(\alpha) \equiv 2^{q-1} \binom{q}{b_q(\alpha)}^{-1} [\alpha - 2\Psi_q(b_q(\alpha) - 1)] \ . \tag{12}$$

Intuitively, the test $\phi(S_n)$ exploits the fact that, under the null hypothesis in (3), the distribution of the treatment assignment should be locally the same to either side of the cut-off. That is, local to the cut-off the treatment assignment behaves as purely randomized under the null hypothesis, and so the fraction of units under treatment and control should be similar.

Remark 3.1. The test in (11) is possibly randomized. The non-randomized version of the test that rejects when $T_q(S_n) > c_q(\alpha)$ is also asymptotically level α by Theorem 4.2. In our simulations, and for our data dependent choice of q we describe in the next section, the randomized and non-randomized versions perform similarly. ∎

Remark 3.2. The value of $b_q(\alpha) \in \{0, 1, \ldots, \lfloor \frac{q}{2} \rfloor\}$ solving (10) is well-defined and unique for all $q \geq 1$ and $\alpha \in (0, 1)$. To see this, let

$$q^*(\alpha) \equiv 1 - \frac{\log \alpha}{\log 2} \ . \tag{13}$$

When $q < q^*(\alpha)$, (10) uniquely holds for $b_q(\alpha) = 0$. In this case, $\phi(S_n)$ in (11) is purely randomized. When $q \geq q^*(\alpha)$, the uniqueness of the solution is guaranteed by $\Psi_q(b)$ being strictly increasing over $\{0, 1, \ldots, \lfloor \frac{q}{2} \rfloor\}$, $\Psi_q(0) = \frac{1}{2^q}$, and $\Psi_q(\frac{q}{2}) \geq \frac{1}{2}$. In this case, $\phi(S_n)$ in (11) deterministically rejects with positive probability. This shows that in order for the non-randomized version of the test to be non-trivial (see Remark 3.1), q needs to be larger than $q^*(\alpha)$. In order to better appreciate these magnitudes, note that for $\alpha = 5\%$ this requires $q \geq 6$ while for $\alpha = 1\%$ this requires $q \geq 8$. Similarly, and given $b_q(\alpha)$, the value of $a_q(\alpha)$ in (12) is also uniquely defined and taking values in $[0, 1)$ by the same properties of $\Psi_q(\cdot)$. ∎

Given q, the implementation of our test proceeds in the following five steps.

Step 1. Find the q observations closest to the cut-off, i.e., $Z_{g,(1)}, \ldots, Z_{g,(q)}$.

Step 2. Count the number of non-negative observations in $Z_{g,(1)}, \ldots, Z_{g,(q)}$, i.e., S_n as in (6).

Step 3. Compute test statistic $T_q(S_n)$ as in (7), $c_q(\alpha)$ as in (9), and $a_q(\alpha)$ as in (12).

Step 4. Compute the p-value of the non-randomized version of the test as

$$p_{\text{value}} = 2 \min \left\{ \Psi_q(S_n), \Psi_q(q - S_n) \right\} \ . \tag{14}$$

Step 5. Reject the null hypothesis in (3) using $\phi(S_n)$ in (11). If a non-randomized test is preferred, reject the null hypothesis if $p_{\text{value}} < \alpha$.

Remark 3.3. As it follows from Theorems 4.1 and 4.2 in Section 4, the test $\phi(S_n)$ is an approximate two-sided sign test (or approximate binomial test). As we mentioned in the introduction, related binomial tests have been recently presented in the RDD context by Cattaneo et al. (2016) and Cattaneo et al. (2017b). There, the authors use a binomial test based on the number of "successes" in a window $Z \in [-h, h]$ for a varying bandwidth h. The authors propose to vary h until a "breakdown" window size h^* is found, which is defined as the largest window such that the minimum p-value of the binomial test is larger than α for all nested (smaller) windows. The justification provided for the validity of such a test involves a finite sample argument: under the hypothesis of "local randomization/random assignment" in $[-h^*, h^*]$, a binomial test with probability π is exact (Cattaneo et al., 2017b, p. 650, note that $\pi = \frac{1}{2}$ is the most natural choice in the absence of additional information). Contrary to Cattaneo et al. (2016) and Cattaneo et al. (2017b), here we do not aim at testing a "local random assignment" hypothesis but rather the continuity hypothesis in (3). As a result of this, we cannot exploit finite sample arguments and rather need to provide a rigorous asymptotic analysis of the test we propose under an arguably weak assumption (i.e., Assumption 4.1). The formal results in Theorems 4.1, 4.2, and 4.3 are novel to this paper and, to the best of our knowledge, they provide the first formal results about approximate two-sided sign tests for the hypothesis in (3) in the RDD (or any other) framework. ∎

3.1 Data dependent rule for q

In this section we discuss the practical considerations involved in the implementation of our test, highlighting how we addressed these considerations in the companion `Stata` package. Additional computational details are presented in Appendix C.

The only tuning parameter of our test is the number q of observations closest to the cutoff. In this paper we propose a data dependent way to choose q that combines a rule of thumb with a local optimization. We call this data-dependent rule the "informed rule of thumb" and its computation requires the following two steps. For the sake of clarity, in this section we do not use the normalization $\bar{z} = 0$.

In the first step, we compute an initial rule of thumb. Concretely, we propose

$$q_{\mathrm{rot}} = C_\phi \frac{n}{\log n} \ , \tag{15}$$

where

$$C_\phi = \frac{1}{\max\left\{25 \left|\sigma^2 \phi'_{\mu,\sigma}(\bar{z})\right|, 1\right\}} \times \frac{\phi_{\mu,\sigma}(\bar{z})}{\phi_{\mu,\sigma}(\mu)} \ , \tag{16}$$

μ is the mean of Z, σ^2 is the variance of Z, $\phi_{\mu,\sigma}(\cdot)$ is the density of a normally distributed random variable with mean μ and variance σ^2, and

$$\sigma^2 \phi'_{\mu,\sigma}(\bar{z}) = \frac{1}{\sqrt{2\pi}} \left(\frac{\bar{z} - \mu}{\sigma}\right) \exp\left(-\frac{1}{2}\left(\frac{\bar{z} - \mu}{\sigma}\right)^2\right) \ , \tag{17}$$

7

is the (scale invariant) derivative of the density $\phi_{\mu,\sigma}(\cdot)$ at $\bar{z}$. The rate of q_{rot} satisfies the properties in Section 4 and is such that $q_{\text{rot}}/n \propto 1/\log n \to 0$. The constant C_ϕ intends to capture the idea that a steeper density at the cut-off should be associated with a smaller value of q. Intuitively, the steeper the density, the more it resembles a density that is discontinuous (Figure 1.(c) illustrates this in Section 5). Thus, C_ϕ is inversely proportional to such a derivative in the normal case. It also captures the idea that q should be small if the cut-off is a point of low density relative to the mode. Intuitively, when $f_Z(\bar{z})$ is low, the q closest observations to $\bar{z}$ are likely to be "far" from $\bar{z}$ if q is too large (Figure 1.(a) with $\mu = -2$ illustrates this in Section 5). Thus, C_ϕ is proportional to this ratio in the normal case, i.e., $\frac{\phi_{\mu,\sigma}(\bar{z})}{\phi_{\mu,\sigma}(\mu)}$. One could alternatively replace the normality assumption with non-parametric estimators of these densities and derivatives, but being just a rule of thumb we prefer to prioritize its simplicity.

The second step involves a local maximization of the asymptotic rejection probability under the null hypothesis of the non-randomized version of the test. In particular, based on Theorem 4.2 we propose

$$q_{\text{irot}} = \underset{q \in \mathcal{N}(q_{\text{rot}})}{\operatorname{argmax}} \ \Psi_q(b_q(\alpha) - 1) \ , \tag{18}$$

where $\Psi_q(\cdot)$ is the CDF defined in (8), $b_q(\alpha)$ is defined in (10), and $\mathcal{N}(q_{\text{rot}})$ is a discrete neighborhood of q_{rot} that we describe in Appendix C. This second step is important for the performance of the non-randomized version of the test (see Remark 3.1) as $\Psi_q(b_q(\alpha) - 1)$ is not monotonic in q, see Figure 3. In practice, we replace μ and σ with sample analogs to deliver a feasible informed rule of thumb that we denote by $\hat{q}_{\text{irot}}$.

Remark 3.4. The recommended choice of q in (15) is simply a sensible rule of thumb that exploits the shape of the limiting null-rejection probability of the non-randomized version of the test to derive a better choice of q. However, this rule of thumb is not "optimal" in any formal sense. Given the asymptotic framework considered in this paper, where q may be fixed as $n \to \infty$, it is difficult, and out of the scope of this paper, to derive "optimal" rules for choosing q. ∎

4 Asymptotic framework and formal results

In this section we derive the asymptotic properties of the test in (11) using two alternative asymptotic frameworks. The first one, and the most interesting from our point of view, is one where q is fixed as $n \to \infty$. This framework is similar to that in Canay and Kamat (2018) and intends to capture a situation where only few of the observations available to the researcher contain good information about the properties "at" the cut-off. The second one is in line with more traditional arguments and requires $q \to \infty$ and $\frac{q}{n} \to 0$ as $n \to \infty$.

There are three main features of our results that are worth highlighting: (i) our test exhibits similar properties under both asymptotic frameworks, (ii) the implementation of the test *does not*

depend on which asymptotic framework one has in mind, and (iii) all formal results require the same, arguably weak, assumption. We start by introducing this assumption.

Assumption 4.1. *The distribution P satisfies*

(i) $\exists \delta > 0$ such that Z has a continuous density on $(-\delta, 0) \cup (0, \delta)$.

(ii) $f_Z^-(0) + f_Z^+(0) > 0$.

Assumption 4.1(i) allows for the distribution of Z to be discontinuous, both at the cut-off $\bar{z} = 0$ and outside a neighborhood of the cut-off. It allows us to study alternative hypotheses with a mass point at the cut-off. More importantly, it does not require the density of Z to be differentiable anywhere. This is in contrast to McCrary (2008), who requires three continuous and bounded derivatives of the density of Z (everywhere except possibly at $\bar{z} = 0$), and Cattaneo et al. (2017a) and Otsu et al. (2013), who require the density of Z to be twice continuously differentiable local to the cut-off (in the case of a local-quadratic approximation). Assumption 4.1(ii) rules out a situation where $f_Z^-(0) = f_Z^+(0) = 0$, which is implicitly assumed in McCrary (2008) and Otsu et al. (2013) and is weaker than assuming a positive density of Z in a neighborhood of the cut-off as in Cattaneo et al. (2017a). In Section 5 we explore the sensitivity of our results to violations of these conditions.

4.1 Results for fixed q

In this section we present two main results. The first result, Theorem 4.1, describes the asymptotic properties of S_n in (6) when q is fixed as $n \to \infty$. This result about g-order statistics with $g(\cdot) = |\cdot|$ represents an important milestone in proving the asymptotic validity of our test. The second result, Theorem 4.2, exploits the result in Theorem 4.1 to show that the test in (11) controls the limiting rejection probability under the null hypothesis.

Theorem 4.1. *Let Assumption 4.1 hold and let $q > 1$ be fixed. If $P\{Z = 0\} = 0$, then*

$$S_n \xrightarrow{d} S \sim \mathrm{Bi}(q, \pi_f)$$

as $n \to \infty$, where $\mathrm{Bi}(q, \pi_f)$ denotes the Binomial distribution with q trials and probability of success

$$\pi_f \equiv \frac{f_Z^-(0)}{f_Z^-(0) + f_Z^+(0)} \ .$$

If $P\{Z = 0\} > 0$, then $S = q$ with probability one.

Theorem 4.1, although fairly intuitive, does not follow from standard arguments. The case when $P\{Z = 0\} > 0$ is relatively simple, so we focus our discussion on the case $P\{Z = 0\} = 0$. First, the random variables $\{A_{g,(j)} : 1 \leq j \leq q\}$ are not necessarily i.i.d. by virtue of being indicators

of g-order statistics. In general, they are neither independent nor identically distributed. Second, applying results from the literature on g-order statistics (e.g., Kaufmann and Reiss, 1992, Theorem 1) requires $g(Z) = |Z|$ to have a continuous distribution function everywhere on its domain. Under Assumption 4.1 this is only true in $(0, \delta)$, and mass points are allowed, both at zero and outside $(0, \delta)$. In the proof of Theorem 4.1 we use a smoothing transformation of Z as an intermediate step and then accommodate the results in Kaufmann and Reiss (1992, Theorem 1) to reach the desired conclusion.

The following result, which heavily relies on Theorem 4.1, is the main result of this section and characterizes the asymptotic properties of our test.

Theorem 4.2. *Let Assumption 4.1 hold and let $q > 1$ be fixed. Then, under H_0 in (3),*

$$\lim_{n \to \infty} E[\phi(S_n)] = 2\Psi_q(b_q(\alpha) - 1) + \frac{a_q(\alpha)}{2^{q-1}} \binom{q}{b_q(\alpha)} = \alpha .$$

In addition, under H_1 in (3), $\lim_{n \to \infty} E[\phi(S_n)] > \alpha$.

Theorem 4.2 shows that $\phi(S_n)$ behaves asymptotically, as $n \to \infty$, as the two-sided sign-test in an experiment with $S \sim \text{Bi}(q, \pi)$ and the hypotheses $H_0 : \pi = \frac{1}{2}$ versus $H_1 : \pi \neq \frac{1}{2}$. This test is not only among the oldest significance tests in statistics (see, e.g., Arbuthnott (1710)), but it is also the uniformly most powerful test among the class of unbiased test for such hypothesis testing problem; see Lehmann and Romano (2005, Section 4.2) and Lemma B.5 in the appendix.

Remark 4.1. Theorem 4.2 shows that $\lim_{n \to \infty} E[\phi(S_n)] > \alpha$ under H_1 in (3). It is worth noticing that $\lim_{n \to \infty} E[\phi(S_n)] = 1$ for any alternative hypothesis with a mass point at the cut-off, i.e., a distribution such that $P\{Z = 0\} > 0$. ∎

Remark 4.2. The test $\phi(S_n)$ could be alternatively characterized as an "approximate" randomization test, see Canay et al. (2017) for a general description of such tests. However, such a characterization would make the analysis of the formal properties of the test more complicated and, in particular, the results in Canay et al. (2017) would not immediately apply due to two fundamental challenges. First, Assumption 3.1(iii) in Canay et al. (2017) is immediately violated in our setting. Second, such an approach would require an asymptotic approximation to the joint distribution of $\{A_{g,(j)} : 1 \leq j \leq q\}$, which in turn would require a strengthening of Lemma B.4. Our proof approach avoids both of these technicalities by directly exploiting the binary nature of $\{A_{g,(j)} : 1 \leq j \leq q\}$ and by simply approximating the distribution of S_n, which is a scalar, as in Theorem 4.1. ∎

Remark 4.3. It is possible to show that $\phi(S_n)$ in (11) is level α in finite samples whenever the distribution of Z is continuous and symmetric local to the cut-off. In this case, the fundamental result in Lemma B.4 holds for S_n with $P\{Z > 0 \mid |Z| < r\} = \frac{1}{2}$ for any $r > 0$, and the proof of Theorem 4.2 can in turn be properly modified to show $E[\phi(S_n)] = \alpha$ for all $n \geq 1$. ∎

4.2 Results for large q

In this section we study the properties of $\phi(S_n)$ in (11) in an asymptotic framework where q diverges to infinity as $n \to \infty$. We further restrict the rate at which q is allowed to grow by requiring that $\frac{q}{n} \to 0$ as $n \to \infty$; a condition that turns out to be necessary for our results to hold (see Remark 4.4). Importantly, the results in this section follow from Assumption 4.1 as well, and so the asymptotic properties of our test under small and large q require the same mild conditions.

Theorem 4.3. *Let Assumption 4.1 hold and assume $P\{Z = 0\} = 0$. Let $q \to \infty$ and $\frac{q}{n} \to 0$ as $n \to \infty$. Then,*

$$\sqrt{q}\left(\frac{1}{q}S_n - \pi_f\right) \xrightarrow{d} N\left(0, \pi_f(1 - \pi_f)\right) ,$$

where π_f is as in Theorem 4.1. In addition, the following results hold for $\alpha \in (0, 1)$:

(a) Under H_0 in (3), $\lim_{n \to \infty} E[\phi(S_n)] = \alpha$.

(b) Under a sequence of alternative distributions local to H_0 satisfying $\sqrt{q}(\pi_f - \frac{1}{2}) \to \Delta \neq 0$,

$$\lim_{n \to \infty} E[\phi(S_n)] = P\{|\zeta + 2\Delta| > z_{\alpha/2}\} > \alpha ,$$

where $\zeta \sim N(0, 1)$ and $z_{\alpha/2}$ is the $(1 - \frac{\alpha}{2})$-quantile of ζ.

Theorem 4.3, although fairly intuitive again, does not follow from standard arguments. In particular, given that the random variables $\{A_{g,(j)} : 1 \leq j \leq q\}$ are neither independent nor identically distributed, the result does not follow from a simple application of the central limit theorem. We instead adapt Kaufmann and Reiss (1992, Theorem 1) and prove the result using first principles and the normal approximation to the binomial distribution.

Remark 4.4. Lemma B.6 in the Appendix shows that $\liminf_{n \to \infty} E[\phi(S_n)]$ may exceed α under H_0 in (3) when $\frac{q}{n} \nrightarrow 0$. This illustrates the sense in which, even when q is allowed to grow, it is required that q remains small relative to n in order for the test to have good properties under the null hypothesis. ∎

Remark 4.5. It may be tempting to use the first part of Theorem 4.3 to consider a variation of the test we propose; namely the test that rejects H_0 when $T_q(S_n) > \frac{1}{2}z_{\alpha/2}$ and $z_{\alpha/2}$ is the $(1 - \frac{\alpha}{2})$-quantile of a standard normal random variable. However, we do not recommend this variation as it provides no theoretical advantages over $\phi(S_n)$ in the asymptotic framework where $q \to \infty$, and it is not formally justified in the asymptotic framework where q is fixed (in particular, such a variation will not inherit the properties discussed in Remark 4.3). ∎

Remark 4.6. As pointed out by a referee, in the asymptotic framework where $q \to \infty$, the test statistic $T_q(S_n)$ can be shown to be proportional to a Wald-type statistic

$$W_n = |\hat{f}_Z(h_n) - \hat{f}_Z(-h_n)| ,$$

where $\hat{f}_Z(z)$ is a non-parametric kernel density estimator of f_Z (implemented with a uniform on $[-1, 1]$ kernel and bandwidth h_n). Under some conditions, standard asymptotic arguments could be used to show that the Wald-type statistic above is asymptotically normally distributed. A test for H_0 could therefore be constructed by using the quantile of a normal distribution and a consistent estimator of the asymptotic variance. However, the equivalence between an approach like this and the test we propose only holds in the asymptotic framework where $q \to \infty$, and for the same reasons as those discussed in Remark 4.5, we do not recommend a variation like this. $\blacksquare$

5 Simulations

In this section, we examine the finite-sample performance of the test we propose in this paper with a simulation study. Instead of just presenting designs where our proposed test excels relative to competing ones, we present an array of data generating processes that hopefully illustrate the relative strengths and weaknesses of the test we propose. The data for the study is simulated as an i.i.d. sample from the following designs, where below $\text{Beta}(a, b)$ denotes the Beta distribution with parameters (a, b).

Design 1: For $\mu \in \{-2, -1, 0\}$, $Z \sim N(\mu, 1)$.

Design 2: For $\lambda \in \{\frac{1}{3}, 1\}$,

$$
Z \sim \begin{cases} V_1 & \text{with prob. } \lambda \\ V_2 & \text{with prob. } (1 - \lambda) \end{cases},
$$

where $V_1 \sim 2\text{Beta}(2, 4) - 1$ and $V_2 \sim 1 - 2\text{Beta}(2, 8)$.

Design 3: For $(\lambda_1, \lambda_2, \lambda_3) = (0.4, 0.1, 0.5)$,

$$
Z \sim \begin{cases} V_1 & \text{with prob. } \lambda_1 \\ V_2 & \text{with prob. } \lambda_2 \\ V_3 & \text{with prob. } \lambda_3 \end{cases},
$$

where $V_1 \sim N(-1, 1)$, $V_2 \sim N(-0.2, 0.2)$, and $V_3 \sim N(3, 2.5)$.

Design 4: For $\kappa \in \{0.05, 0.10, 0.25\}$, the density of Z is given by

$$
f_Z(z) = \begin{cases} 0.75 & \text{if } z \in [-1, -\kappa] \\ 0.75 - \frac{1}{4\kappa}(z + \kappa) & \text{if } z \in [-\kappa, \kappa] \\ 0.25 & \text{if } z \in [\kappa, 1] \end{cases}.
$$

Design 5: For $\kappa \in \{0.05, 0.10, 0.25\}$, the density of Z is given by

$$f_Z(z) = \begin{cases} 0.25 & \text{if } z \in [-1, -\kappa] \\ 0.50 & \text{if } z \in [-\kappa, \kappa] \\ 0.75 & \text{if } z \in [\kappa, 1] \end{cases} .$$

Design 6: We first non-parametrically estimate the density of the running variable in Lee (2008, see Section 6 for details) and then take i.i.d. draws from such a density.

Design 1 in Figure 1(a) is the canonical normal case and, by Remark 4.3, our test is expected to control size in finite samples when $\mu = 0$ but not when $\mu \in \{-2, -1\}$. Indeed, $\mu = -2$ is a challenging case due to the low probability of getting observations to the right of the cut-off. Design 2 in Figure 1(b) is taken from Canay and Kamat (2018). Design 3 in Figure 1(c) is a parametrization of the taxable income density in Saez (2010, Figure 8). This design exhibits a spike (almost a kink) to the left of the cut-off which is essentially a violation of the smoothness assumptions required by McCrary (2008) and Cattaneo et al. (2017a). It also exhibits a steep density at the cut-off, which also makes it a difficult case in general. Similar to Design 3, Design 4 in Figure 1(b) also illustrates the difficulty in distinguishing a discontinuity from a very steep slope; see Kamat (2017) for a formal discussion. Here we can study the sensitivity to the slope by changing the value of κ. Design 5 in Figure 1(e) requires δ in Assumption 4.1(a) to be such that $\delta < \kappa$ in order for our approximations to be accurate, but as opposed to Design 4, it is locally symmetric around the cut-off. As κ gets smaller, we expect our test to perform worse if q is not chosen carefully. Finally, Design 6 in Figure 1(f) draws data i.i.d. from the non-parametric density estimator of the running variable in Lee (2008), i.e., Z is the difference in vote shares between Democrats and Republicans.

We consider sample sizes $n \in \{1,000; 5,000\}$, a nominal level of $\alpha = 10\%$, and perform $10,000$ Monte Carlo repetitions. Designs 1 to 6 satisfy the null hypothesis in (3). We additionally consider the same models under the alternative hypothesis by randomly changing the sign of observations in the interval $[0, 0.1]$ with probability 0.1. We report results for the following tests.

AS-NR and **AS-R**: the approximate sign-test we propose in this paper in its two versions. The randomized version (AS-R) in (11) and the non-randomized version (AS-NR) that rejects when p_{value} in (14) is below α, see Remark 3.1. We include the randomized version in order to illustrate the differences between the randomized and non-randomized versions of the test. The tuning parameter q is set to

$$q \in \{20, 50, 75, \hat{q}_{\text{irot}}\} ,$$

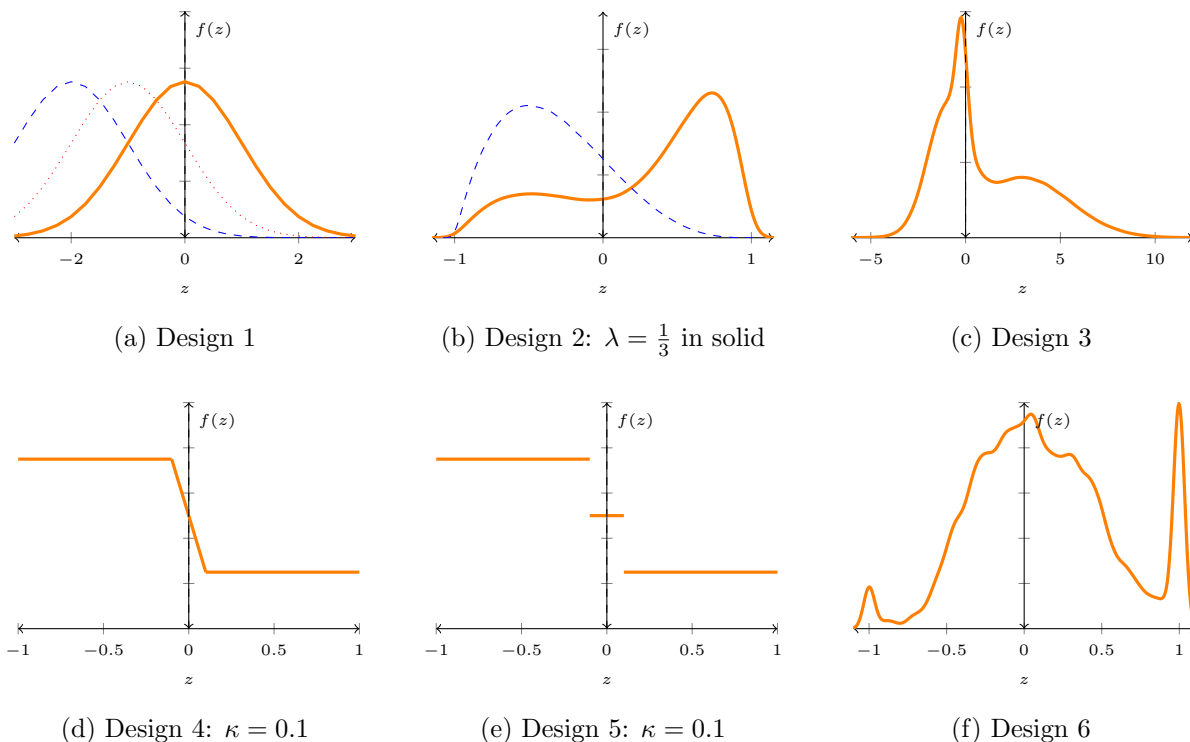where $\hat{q}_{\text{irot}}$ is the feasible informed rule of thumb described in Section 3 and Appendix C.

Figure 1: Density functions $f(z)$ for Designs 1 to 6 used in the Monte Carlo simulations

McC: the test proposed by McCrary (2008). We implement this test using the function `DCdensity` from the R package `rdd`, with the default choices for the bandwidth parameter and kernel type.

CJM: the test proposed by Cattaneo et al. (2017a). We implement this test using the `rddensity` function from the R package `rddensity`. We use jackknifed standard errors and bias correction, as these are the default choices in the paper.

Tables 1 and 2 report rejection probabilities under the null and alternative hypotheses for the six designs we consider and for sample sizes of $n = 1,000$ and $n = 5,000$, respectively. We start by discussing the results under the null hypothesis. AS-NR delivers rejection probabilities under the null hypothesis closer to the nominal level than those delivered by McC and CJM in most of the designs. The two empirically motivated designs (Designs 3 and 6) illustrate this feature clearly. Designs 4 and 5 also show big differences in performance, both in cases where AS-NR delivers rejection rates equal to the nominal level (Design 5) and McC and CJM severely over-reject; as well as in cases where all tests over-reject (Design 4, $\kappa = 0.05$) but AS-NR is relatively closer to the nominal level. The relatively most difficult case for AS-NR is Design 1 with $\mu = -2$, where the probability of getting observations to the right of the cut-off is below 2%. Tables 1 and 2 also show negligible differences between the randomized (AS-R) and non-randomized (AS-NR) versions of our test, consistent with our discussion in Remark 3.1.

14

	Rejection Rate under H_0							Rejection Rate under H_1						
	AS-NR				AS-R	McC	CJM	AS-NR				AS-R	McC	CJM
	q				q			q				q		
Design	20	50	75	$\hat{q}_{\mathrm{irot}}$	$\hat{q}_{\mathrm{irot}}$			20	50	75	$\hat{q}_{\mathrm{irot}}$	$\hat{q}_{\mathrm{irot}}$		
D1: $\mu = 0$	4.5	6.4	6.3	10.0	10.1	9.5	8.2	6.1	13.2	15.5	16.9	17.0	16.2	12.8
D1: $\mu = -1$	4.2	8.3	11.9	9.3	9.7	11.6	8.2	6.9	19.5	25.3	12.9	13.4	19.4	10.2
D1: $\mu = -2$	11.8	84.2	99.7	12.9	13.7	11.4	7.7	16.2	87.8	99.8	17.1	18.2	11.5	7.9
D2: $\lambda = 1$	4.2	7.2	7.2	10.4	10.5	11.2	10.2	6.1	10.9	9.4	14.5	14.7	22.5	12.7
D2: $\lambda = \frac{1}{3}$	4.2	7.7	10.2	9.9	10.2	10.4	7.9	6.7	19.4	30.5	13.8	14.1	26.5	15.0
D3	4.8	17.9	39.3	15.4	15.6	100.0	86.0	8.4	36.5	59.9	30.1	30.3	100.0	90.0
D4: $\kappa = 0.25$	4.1	8.7	12.7	11.6	11.7	12.1	9.8	7.4	21.4	35.5	21.6	21.8	28.1	17.8
D4: $\kappa = 0.1$	4.8	18.3	41.8	16.1	16.3	47.8	24.8	8.5	37.4	69.8	31.0	31.1	73.6	40.8
D4: $\kappa = 0.05$	6.5	47.8	86.6	31.4	31.6	85.0	60.2	12.6	69.4	95.8	50.9	51.0	95.0	76.4
D5: $\kappa = 0.25$	3.8	6.8	6.3	9.8	9.9	20.8	12.9	5.9	13.1	16.0	15.5	15.7	14.2	29.0
D5: $\kappa = 0.1$	4.0	6.7	6.4	9.8	10.0	17.6	32.9	6.4	13.2	17.0	16.3	16.5	20.1	16.2
D5: $\kappa = 0.05$	4.1	7.8	35.3	9.4	9.6	40.6	25.6	5.7	18.4	63.4	16.0	16.1	67.5	39.8
D6	4.3	6.0	6.4	9.6	9.7	9.2	10.1	6.1	11.6	13.0	15.4	15.5	28.0	17.6

Table 1: Rejection probabilities (in %) under H_0 and H_1 across Designs 1-6 and for $n = 1,000$.

	Rejection Rate under H_0							Rejection Rate under H_1						
	AS-NR				AS-R	McC	CJM	AS-NR				AS-R	McC	CJM
	q				q			q				q		
Design	20	50	75	$\hat{q}_{\mathrm{irot}}$	$\hat{q}_{\mathrm{irot}}$			20	50	75	$\hat{q}_{\mathrm{irot}}$	$\hat{q}_{\mathrm{irot}}$		
D1: $\mu = 0$	3.9	6.6	6.9	10.1	10.2	9.6	9.0	6.1	13.0	16.7	51.2	51.3	45.6	39.9
D1: $\mu = -1$	4.0	6.7	6.8	10.1	10.4	12.2	9.0	6.2	14.4	19.5	20.0	20.3	43.4	20.9
D1: $\mu = -2$	4.3	12.3	25.7	13.3	13.4	11.9	9.1	8.2	30.1	46.0	26.4	26.6	15.0	13.1
D2: $\lambda = 1$	4.1	7.0	6.8	10.7	10.9	10.8	9.4	6.3	13.1	15.8	26.5	26.7	49.8	13.4
D2: $\lambda = \frac{1}{3}$	3.9	6.3	6.4	9.8	10.0	10.6	8.1	6.5	14.2	19.4	20.3	20.5	71.1	37.5
D3	4.4	7.0	8.1	23.6	23.8	100.0	93.7	6.8	17.1	25.2	65.6	65.7	100.0	98.1
D4: $\kappa = 0.25$	4.2	6.4	6.4	11.1	11.3	12.0	10.2	6.3	15.1	19.8	39.8	40.0	64.0	40.4
D4: $\kappa = 0.1$	4.0	6.7	7.7	17.8	18.0	69.1	23.4	6.0	15.6	24.3	53.3	53.5	97.8	64.8
D4: $\kappa = 0.05$	4.3	8.4	12.0	41.1	41.3	99.8	94.1	6.8	20.7	34.2	76.2	76.4	100.0	99.4
D5: $\kappa = 0.25$	4.1	6.2	6.0	9.5	9.6	31.7	17.1	6.1	13.5	16.9	32.9	33.1	38.9	68.6
D5: $\kappa = 0.1$	4.2	6.5	6.1	9.4	9.5	51.0	44.4	6.2	12.8	16.1	30.1	30.4	20.0	18.8
D5: $\kappa = 0.05$	3.9	6.8	6.6	10.3	10.4	46.2	31.3	6.2	13.2	16.4	29.9	30.1	92.0	55.8
D6	3.8	6.3	6.4	9.9	9.9	13.2	13.7	6.0	12.6	15.8	38.1	38.1	72.1	48.1

Table 2: Rejection probabilities (in %) under H_0 and H_1 across Designs 1-6 and for $n = 5,000$.

To describe the performance of the different tests under the alternative hypothesis, we focus on designs where the rejection under the null hypothesis is close to the nominal level for all tests. In those cases, we see that AS-NR has competitive power, although it is rarely the test with the highest rejection under the alternative hypothesis. In Design 1 with $\mu = 0$ AS-NR indeed delivers the highest rejection probability under the alternative hypothesis both for $n = 1,000$ and $n = 5,000$.

	D1	D1	D1	D2	D2	D3	D4	D5	D6
	$\mu = 0$	$\mu = -1$	$\mu = -2$	$\lambda = 1$	$\lambda = \frac{1}{3}$		all κ	all κ	
$n = 1000$	147	18	13	37	18	37	37	37	37
$n = 5000$	562	53	37	119	53	147	131	125	144

Table 3: Mean values of $\hat{q}_{\text{irot}}$ in Designs 1-6 and for $n = 1,000$ and $n = 5,000$.

(a) Design 1: $\mu = 0$

(b) Design 6

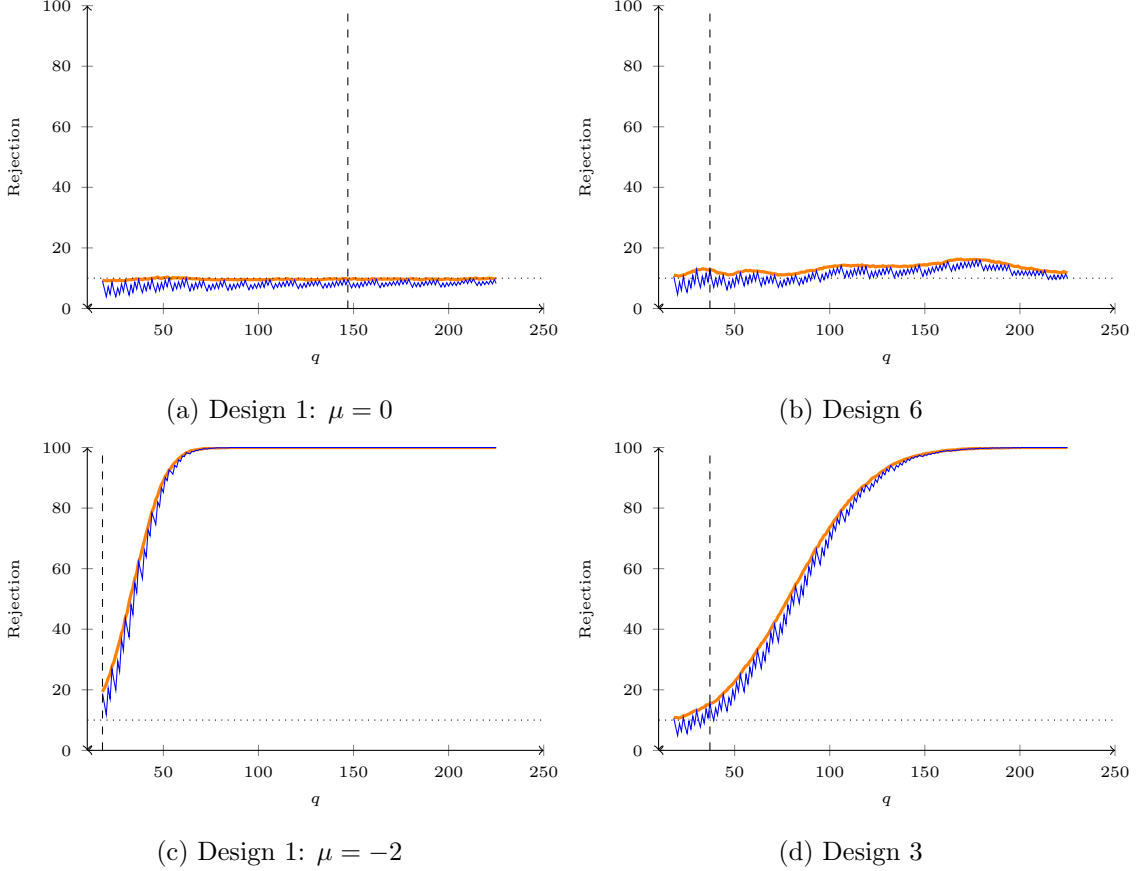(c) Design 1: $\mu = -2$

(d) Design 3

Figure 2: Rejection probability of AS-NR (blue line) and AS-R (orange line) as a function of q. The vertical dashed line denotes the value of $\hat{q}_{\text{irot}}$ and the horizontal dotted line the value of α.

In the rest of the cases McC exhibits the highest power, sometimes followed by AS-NR (i.e., Design 1 with $\mu = -1$, Design 2 with $\lambda = 1$, Design 4 with $\kappa = 0.25$ when $n = 1,000$), and sometimes followed by CJM (i.e., Design 2 with $\lambda = \frac{1}{3}$, Design 4 with $\kappa = 0.25$ when $n = 5,000$). Overall, the best design in terms of power for AS-NR is Design 1 with $\mu = 0$, while the worst one is Design 2 with $\lambda = \frac{1}{3}$.

Table 3 shows the mean values of $\hat{q}_{\text{irot}}$ across simulations for all designs and sample sizes. As described in Section 3, $\hat{q}_{\text{irot}}$ takes into account both the slope and the magnitude of the density at the cut-off. As a result, $\hat{q}_{\text{irot}}$ is relatively high in designs with flat density at the cut-off and high $f_Z(0)$ (e.g., Design 1 with $\mu = 0$) and relatively low in designs with steep slopes or low $f_Z(0)$ (e.g.,

16

Design 1 with $\mu = -2$ or Design 2 with $\lambda = \frac{1}{3}$). To gain further insight on the sensitivity of our test to the choice of q, Figure 2 displays the rejection probabilities of AS-NR and AS-R as a function on q in two types of designs. In the top row we illustrate two designs where the rejection probability is mostly insensitive to the choice of q (Design 1 with $\mu = 0$ and Design 6). These are designs where the density is rather flat around the cut-off and so increasing q does not deteriorate the performance of our test. In the bottom row we illustrate two designs where the rejection probability is highly sensitive to the choice of q (Design 1 with $\mu = -2$ and Design 3). These are designs that feature a steep density at the cut-off (also low in Design 1), and so increasing q very quickly deteriorates the performance of the test under the null hypothesis. The rule of thumb $\hat{q}_{\text{irot}}$ is displayed in each case with a vertical dashed line and seems to be doing a good job at choosing relatively smaller values in the sensitive cases.

We conclude this section by highlighting how one could compare the results in Tables 1 and 2 for a fix value of q to appreciate the results in Section 4.1. For example, taking $q = 75$, the rejection probability in Design 1 with $\mu = -2$ and Design 3 are 84.3 and 17.9, respectively, when $n = 1,000$. The same numbers when $n = 5,000$ are 12.3 and 7.0, respectively, which are closer to the nominal level as predicted by our results.

6   Empirical Illustration

In this section we briefly reevaluate the validity of the design in Lee (2008). Lee studies the benefits of incumbency on electoral outcomes using a discontinuity constructed with the insight that the party with the majority wins. Specifically, the running variable Z is the difference in vote shares between Democrats and Republicans at time t; see Figure 1(f) for a graphical illustration of the density of Z. The assignment rule then takes a cutoff value of zero that determines the treatment of incumbency to the Democratic candidate, which is used to study their election outcomes in time $t+1$. The total number of observations is 6,559 with 2,740 below the cutoff. The dataset is publicly available at http://economics.mit.edu/faculty/angrist/data1/mhe.

Lee assessed the credibility of the design in this application by inspecting discontinuities in means of the baseline covariates, but mentions in footnote 19 the possibility of using the test proposed by McCrary (2008). Here, we frame the validity of the design in terms of the hypothesis in (3) and use the newly developed test as described in Section 3, using $\hat{q}_{\text{irot}}$ as our default choice for the number of observations q. The new test delivers a p-value of 0.71 for $S_n = 137$ out of $\hat{q}_{\text{irot}} = 267$ observations. The null hypothesis of continuity is therefore not rejected.

7 Concluding remarks

This paper presents a new tests for testing the continuity of a density at a point in RDD. The test can be interpreted as an approximate two-sided sign test and is based on the so-called g-order statistics. We study its properties under a novel asymptotic framework where the number q of observations employed by the test are allowed to be fixed as the sample size $n \to \infty$. Our new test is easy to implement, asymptotically valid under weaker conditions than those used by competing methods, exhibits finite sample validity under stronger conditions than those needed for its asymptotic validity, and delivers competitive power properties in simulations.

A final aspect we would like to highlight of our test is its simplicity. The test only requires to count the number of non-negative observations out of the q observations closest to the cut-off (note that this is all we need to compute the p-value in (14)), and does not involve kernels, local polynomials, bias correction, or bandwidth choices. Importantly, we have developed the `rdcont` `Stata` package that allow for effortless implementation of the test we propose in this paper.

A Proof of the main results

A.1 Proof of Theorem 4.1

Consider the case $P\{Z = 0\} = 0$ first. In this case, by Assumption 4.1(i) the distribution of Z is continuous in $(-\delta, \delta)$ for some $\delta > 0$. Throughout the proof we repeatedly use $\{Z_i^* : 1 \leq i \leq n\}$ as defined in Lemma B.1, which in turn allow us to apply Kaufmann and Reiss (1992, Theorem 1) later in the proof, when invoking Lemma B.4.

Let $Z_{g,(1)}^*, \ldots, Z_{g,(q)}^*$ denote the q values of $\{Z_i^* : 1 \leq i \leq n\}$ that are induced by the q smallest values of $\{g(Z_i^*) = |Z_i^*| : 1 \leq i \leq n\}$ and let

$$A_{g,(j)}^* \equiv I\{Z_{g,(j)}^* \geq 0\} \text{ for } 1 \leq j \leq q$$

and

$$S_n^* \equiv \sum_{j=1}^{q} A_{g,(j)}^* . \tag{A-19}$$

Next consider S_n in (6) and note that S_n takes values in $\mathbf{N}_q \equiv \{0, 1, \ldots, q\}$. By the Portmanteau's theorem, see e.g. van der Vaart and Wellner (1996, Theorem 1.3.4(iii)), it follows that if $S_n \overset{d}{\to} S$ for some random variable S, then

$$1 = \liminf_{n \to \infty} P\{S_n \in \mathbf{N}_q\} \leq P\{S \in \mathbf{N}_q\} ,$$

since $\mathbf{N}_q$ is a closed subset of $\mathbf{R}$. Thus, it must be that S_n and S take values in $\mathbf{N}_q$, and so by Durrett (2010, Exercise 3.2.11), convergence in distribution is equivalent to convergence of the probability mass function (pmf) for all $s \in \mathbf{N}_q$. To establish this result, let $p_q(s|\pi)$ be the pmf of a Binomial random variable with q trials and probability of success $\pi \in [0, 1]$, i.e.

$$p_q(s|\pi) = \binom{q}{s} \pi^s (1 - \pi)^{q-s} . \tag{A-20}$$

It suffices to show that for any $\eta > 0$, there exists N such that $\forall n \geq N$,

$$|P\{S_n = s\} - p_q(s|\pi_f)| \leq \eta ,$$

with

$$\pi_f = \frac{f_Z^-(0)}{f_Z^-(0) + f_Z^+(0)} . \tag{A-21}$$

To this end, first note that $p_q(s|\pi)$ is continuous in π and so there exists $\mu > 0$ such that

$$\sup_{|\pi - \pi_f| \leq \mu} |p_q(s|\pi) - p_q(s|\pi_f)| \leq \frac{\eta}{2} . \tag{A-22}$$

For such μ, we can find $\varepsilon \in (0, \frac{\delta}{2})$ such that

$$\sup_{r \leq \varepsilon} \left| P\{Z^* \geq 0 \mid |Z^*| < r\} - \pi_f \right| = \sup_{r \leq \varepsilon} \left| P\{Z \geq 0 \mid |Z| < r\} - \pi_f \right| \leq \mu , \tag{A-23}$$

where the first equality holds by Lemma B.1(b) and the second equality holds for by Lemma B.2. The rest of the argument will make repeated reference to ε determined by (A-23).

19

Next consider the following decomposition for $s \in \mathbf{N}_q$,

$$P\{S_n = s\} = R_{n,1} + R_{n,2} + R_{n,3} ,$$

with

$$R_{n,1} \equiv P\{S_n = s\} - P\{S_n^* = s\}$$

$$R_{n,2} \equiv \int_\epsilon^\infty P\{S_n^* = s \mid |Z_{g,(q+1)}^*| = r\} dP\{|Z_{g,(q+1)}^*| = r\}$$

$$R_{n,3} \equiv \int_0^\epsilon P\{S_n^* = s \mid |Z_{g,(q+1)}^*| = r\} dP\{|Z_{g,(q+1)}^*| = r\} .$$

First, Lemma B.3(b) implies that $R_{n,1} = o(1)$. Second, note that

$$0 \le R_{n,2} \le \int_\epsilon^\infty dP\{|Z_{g,(q+1)}^*| = r\} = P\{|Z_{g,(q+1)}^*| \ge \epsilon\} = o(1) , \tag{A-24}$$

where the last equality follows from Lemma B.3(a). Finally, note that for $\pi(r) = P\{Z \ge 0 \mid |Z| < r\}$,

$$
\begin{aligned}
R_{n,3} &= \int_0^\epsilon p_q(s|\pi(r)) dP\{|Z_{g,(q+1)}^*| = r\} \\
&\ge P\{|Z_{g,(q+1)}^*| \le \epsilon\} \inf_{r \le \epsilon} p_q(s|\pi(r)) \\
&\ge P\{|Z_{g,(q+1)}^*| \le \epsilon\} \inf_{|\pi - \pi_f| \le \mu} p_q(s|\pi) \\
&\ge P\{|Z_{g,(q+1)}^*| \le \epsilon\} p_q(s|\pi_f) - \frac{\eta}{2} , \tag{A-25}
\end{aligned}
$$

where the first line follows from Lemma B.4, the third line follows from (A-23), and the fourth line follows from (A-22). By the analogous arguments,

$$R_{n,3} \le P\{|Z_{g,(q+1)}^*| \le \epsilon\} p_q(s|\pi_f) + \frac{\eta}{2} \le p_q(s|\pi_f) + \frac{\eta}{2} . \tag{A-26}$$

Combining (A-25) and (A-26) we obtain

$$|R_{n,3} - p_q(s|\pi_f)| \le \frac{\eta}{2} + (1 - P\{|Z_{g,(q+1)}^*| \le \epsilon\}) = \frac{\eta}{2} + o(1) ,$$

where the equality follows from Lemma B.3(a). We conclude that $|P\{S_n = s\} - p_q(s|\pi_f)| \le \frac{\eta}{2} + o(1)$ for all $s \in \mathbf{N}_q$ and this completes the proof for the case where $P\{Z = 0\} = 0$.

Now consider the case where $P\{Z = 0\} > 0$. Take the following decomposition,

$$P\{S_n \neq q\} = \tilde{R}_{n,1} + \tilde{R}_{n,2} + \tilde{R}_{n,3} ,$$

with

$$\tilde{R}_{n,1} \equiv P\{S_n \neq q\} - P\{S_n^* \neq q\}$$

$$\tilde{R}_{n,2} \equiv P\{\{S_n^* \neq q\} \cap \{|Z_{g,q+1}^*| \neq 0\}\}$$

$$\tilde{R}_{n,3} \equiv P\{\{S_n^* \neq q\} \cap \{|Z_{g,q+1}^*| = 0\}\} .$$

First, Lemma B.3(b) implies that $\tilde{R}_{n,1} = o(1)$. Second, $\tilde{R}_{n,2} \le P\{|Z_{g,q+1}^*| \neq 0\} = o(1)$, where the last equality follows from Lemma B.3(c). Finally, $\{|Z_{g,q+1}^*| = 0\}$ implies that $\{S_n^* = q\}$ by definition, and so $\tilde{R}_{n,3} = 0$. We conclude that $P\{S_n = q\} \to 1$ as $n \to \infty$ and this completes the proof. ∎

20

A.2 Proof of Theorem 4.2

By the definition of $\phi(S_n)$ in (11) and the expressions of $T(S_n)$ in (7) and $c_q(\alpha)$ in (9),

$$E[\phi(S_n)] = P\{S_n < b_q(\alpha)\} + P\{S_n > q - b_q(\alpha)\} + a_q(\alpha)\left(P\{S_n = b_q(\alpha)\} + P\{S_n = q - b_q(\alpha)\}\right) .$$

Theorem 4.1 shows that $P\{S_n = s\} = P\{S = s\} + o(1)$ for all $s \in \mathbf{N}_q \equiv \{0, 1, \ldots, q\}$, where $S \sim \mathrm{Bi}(q, \pi_f)$ and π_f is as in (A-21). It follows from this result and the above display that $E[\phi(S_n)] \to E[\phi(S)]$ as $n \to \infty$, where

$$E[\phi(S)] = P\{S < b_q(\alpha)\} + P\{S > q - b_q(\alpha)\} + a_q(\alpha)\left(P\{S = b_q(\alpha)\} + P\{S = q - b_q(\alpha)\}\right) . \qquad \text{(A-27)}$$

We complete the proof by analyzing each term on the right hand side of (A-27) under H_0 and H_1 in (3).

Under H_0 in (3), $\pi_f = \frac{1}{2}$. In this case, it follows immediately that

$$P\{S < b_q(\alpha)\} + P\{S > q - b_q(\alpha)\} = 2\Psi_q(b_q(\alpha) - 1) ,$$

where we used that $b_q(\alpha) \in \{0, 1, \ldots, \frac{q}{2}\}$ and $P\{S < b\} = P\{S > q - b\}$ for any $b \in \{0, \ldots, \frac{q}{2}\}$ when $\pi_f = \frac{1}{2}$. In addition,

$$a_q(\alpha)\left(P\{S = b_q(\alpha)\} + P\{S = q - b_q(\alpha)\}\right) = 2a_q(\alpha)\frac{1}{2^q}\binom{q}{b_q(\alpha)} ,$$

where we used that $\binom{q}{C} = \binom{q}{q-C}$ for any $C \in \{0, \ldots, q\}$. We conclude that, under H_0 in (3),

$$\lim_{n\to\infty} E[\phi(S_n)] = 2\Psi_q(b_q(\alpha) - 1) + \frac{a_q(\alpha)}{2^{q-1}}\binom{q}{b_q(\alpha)} = \alpha , \qquad \text{(A-28)}$$

where the last equality follows by definition of $a_q(\alpha)$.

To claim that $E[\phi(S)] > \alpha$ under H_1 in (3), let $S \sim \mathrm{Bi}(q, \pi)$ and consider testing $H_0 : \pi = \frac{1}{2}$ against a simple alternative $H_1 : \pi = \bar{\pi} \neq \frac{1}{2}$. By Lehmann and Romano (2005, Theorem 3.2.1), the most powerful test rejects for large values of the likelihood ratio

$$\frac{\bar{\pi}^S(1 - \bar{\pi}^{q-S})}{\frac{1}{2^q}} .$$

That is, if $\bar{\pi} > \frac{1}{2}$ the most powerful test rejects for small values of S, while if $\bar{\pi} < \frac{1}{2}$ the most powerful test rejects for large values of S. Since $E[\phi(S)] = \alpha$ by (A-28), it follows that $\phi(S)$ is most powerful for this simple testing problem. The result then follows from Lehmann and Romano (2005, Corollary 3.2.1) as the alternative $\bar{\pi} \neq \frac{1}{2}$ was arbitrary. ∎

A.3 Proof of Theorem 4.3

Let $\xi_n \equiv \sqrt{q}\left(\frac{1}{q}S_n - \pi_f\right)$ and $\xi_n^* \equiv \sqrt{q}\left(\frac{1}{q}S_n^* - \pi_f\right)$ with S_n^* as in (A-19). For any $\pi \in (0, 1)$ denote by $J_q(x|\pi)$ the cdf of $\sqrt{q}\left(\frac{1}{q}S - \pi\right)$, where $S \sim \mathrm{Bi}(q, \pi)$, and by $J(x|\pi)$ the cdf of the normal distribution with mean zero and variance $\pi(1 - \pi)$. It suffices to show that for any $\eta > 0$, there exists N such that $\forall n \geq N$,

$$|P\{\xi_n \leq x\} - J(x|\pi_f)| \leq \eta .$$

To this end, first note that $J(x|\pi)$ is continuous in π and so there exists $\mu > 0$ such that

$$\sup_{|\pi - \pi_f| \leq \mu} |J(x|\pi) - J(x|\pi_f)| \leq \frac{\eta}{2} \ .$$

For such a μ, we can find $\varepsilon \in (0, \frac{\delta}{2})$ such that

$$\sup_{r \leq \varepsilon} \left| P\{Z^* \geq 0 \mid |Z^*| < r\} - \pi_f \right| = \sup_{r \leq \varepsilon} \left| P\{Z \geq 0 \mid |Z| < r\} - \pi_f \right| \leq \mu \ , \tag{A-29}$$

where the first equality holds by Lemma B.1(b) and the second equality holds for by Lemma B.2. The rest of the argument will make repeated reference to the ε determined by (A-29).

Next consider the following decomposition for $x \in \mathbf{R}$,

$$P\{\xi_n \leq x\} = \bar{R}_{n,1} + \bar{R}_{n,2} + \bar{R}_{n,3} \ , \tag{A-30}$$

with

$$\bar{R}_{n,1} \equiv P\{\xi_n \leq x\} - P\{\xi_n^* \leq x\}$$

$$\bar{R}_{n,2} \equiv \int_\epsilon^\infty P\{\xi_n^* \leq x \mid |Z_{g,(q+1)}^*| = r\} dP\{|Z_{g,(q+1)}^*| = r\}$$

$$\bar{R}_{n,3} \equiv \int_0^\epsilon P\{\xi_n^* \leq x \mid |Z_{g,(q+1)}^*| = r\} dP\{|Z_{g,(q+1)}^*| = r\} \ .$$

First, Lemma B.3(b) implies that $\bar{R}_{n,1} = o(1)$. Second, $\bar{R}_{n,2} = o(1)$ by the same arguments as those in (A-24). Finally,

$$|\bar{R}_{n,3} - J_q(x|\pi_f)| \leq \frac{\eta}{2} + (1 - P\{|Z_{g,(q+1)}^*| \leq \epsilon\}) = \frac{\eta}{2} + o(1) \ , \tag{A-31}$$

invoking the same arguments as those in (A-25) and (A-26). In particular, the result relies again on Lemma B.4. We conclude that

$$|P\{\xi_n \leq x\} - J(x|\pi_f)| \leq |P\{\xi_n \leq x\} - J_q(x|\pi_f)| + |J_q(x|\pi_f) - J(x|\pi_f)| \leq \frac{\eta}{2} + o(1) \ , \tag{A-32}$$

where the last inequality follows from (A-30), (A-31), and the normal approximation to the binomial distribution, i.e., $J_q(x|\pi_f) \to J(x|\pi_f)$ as $q \to \infty$. This completes the proof of the first statement of the theorem.

Next, note that by definition of $c_q(\alpha)$ in (9), it follows that for any $\alpha \in (0,1)$

$$c_q(\alpha) \to \frac{1}{2} z_{\alpha/2} \tag{A-33}$$

as $q \to \infty$, where $z_{\alpha/2}$ be the $(1 - \frac{\alpha}{2})$-quantile of a standard normal random variable. Under the null hypothesis in (3) it follows that $\pi_f = \frac{1}{2}$, so $\lim_{n \to \infty} E[\phi(S_n)] = \alpha$ directly follows from (A-32) and (A-33). This proves part (a) of the Theorem. In addition, since

$$2\sqrt{q} \left(\frac{1}{q} S_n - \frac{1}{2} \right) = 2\sqrt{q} \left(\frac{1}{q} S_n - \pi_f \right) + 2\sqrt{q} \left(\pi_f - \frac{1}{2} \right) \ ,$$

part(c) follows again from (A-32) and (A-33). ∎

B   Auxiliary Results

Lemma B.1. *Let $\delta > 0$ be as in Assumption 4.1 and $\{v_i : 1 \le i \le n\}$ be an i.i.d. sample such that $v_i \sim U(-\frac{\delta}{2}, \frac{\delta}{2})$ independent of $Z^{(n)}$. Define the sequence of i.i.d. random variables $\{Z_i^* : 1 \le i \le n\}$ as*

$$Z_i^* \equiv Z_i + v_i I\{|Z_i| \ge \delta\} \ .$$

The following statements follow from Assumption 4.1 whenever $P\{Z = 0\} = 0$.

(a) The distribution function of $|Z^*|$ is continuous on $\mathbf{R}$.

(b) For any $r \in (0, \frac{\delta}{2})$,
$$P\{Z^* \ge 0 \mid |Z^*| < r\} = P\{Z \ge 0 \mid |Z| < r\} \ . \tag{B-34}$$

(c) For any $r \in \mathbf{R}$, $P\{|Z^*| < r\} > 0$.

Proof. Note that the distribution of Z is continuous in $(-\delta, \delta)$ by Assumption 4.1(i) and $P\{Z = 0\} = 0$. Then, to prove part (a), let $E \subset \mathbf{R}$ be a set of zero Lebesgue measure and note that

$$
\begin{aligned}
P\{|Z^*| \in E\} &= P\{|Z + v I\{|Z| \ge \delta\}| \in E\} \\
&= P\{|Z + v I\{|Z| \ge \delta\}| \in E \cap |Z| \ge \delta\} + P\{|Z + v I\{|Z| \ge \delta\}| \in E \cap |Z| < \delta\} \\
&= P\{|Z + v| \in E \cap |Z| \ge \delta\} + P\{|Z| \in E \cap |Z| < \delta\} \\
&\le P\{|Z + v| \in E\} + P\{|Z| \in E \cap (0, \delta)\} = 0
\end{aligned}
$$

where the last equality holds because $|Z + v|$ is continuously distributed and $E \cap (0, \delta)$ is a subset of zero Lebesgue measure in the region where $|Z|$ is continuously distributed.

For part (b) note that for any $r \in (0, \frac{\delta}{2})$, $|Z^*| < r$ implies that $Z = Z^*$ and (B-34) follows.

For part (c) use again that $P\{|Z^*| < r\} = P\{|Z| < r\}$ whenever $r \in (0, \frac{\delta}{2})$. In addition, Assumption 4.1 implies that the distribution of Z is continuous in $(-\delta, \delta)$ when $P\{Z = 0\} = 0$. It follows that for any $0 < \varepsilon < \delta$,

$$\frac{1}{\varepsilon} P\{|Z| < \varepsilon\} = \frac{1}{\varepsilon} P\{Z \in [0, \varepsilon)\} + \frac{1}{\varepsilon} P\{Z \in (-\epsilon, 0)\} \ .$$

Taking limits as $\varepsilon \downarrow 0$, using the definitions of $f_Z^+(0)$ and $f_Z^-(0)$ in (4), and invoking Assumption 4.1(ii) shows that $\lim_{\epsilon \downarrow 0} \frac{1}{\varepsilon} P\{|Z| < \varepsilon\} = f_Z^+(0) + f_Z^-(0) > 0$. Thus, there exists $\bar{\varepsilon} < \delta$ such that $P\{|Z| < \epsilon\} > 0$ for all $\varepsilon \in (0, \bar{\varepsilon})$ and so $P\{|Z| < r\} > 0$ for all $r \in \mathbf{R}$. This completes the proof. $\blacksquare$

Lemma B.2. *Let Assumption 4.1 hold, π_f be defined as in (A-21), and assume $P\{Z = 0\} = 0$. Then, for all $\mu > 0$, there exists $\epsilon > 0$ s.t.*

$$\sup_{r \le \epsilon} |P\{Z \ge 0 \mid |Z| < r\} - \pi_f| \le \mu \ .$$

Proof. First note that, under Assumption 4.1(ii), the proof of Lemma B.4 shows that $P\{|Z| < r\} > 0$ for all $r \in \mathbf{R}$. It follows that

$$
\begin{aligned}
P\{Z \geq 0 \mid |Z| < \epsilon\} &\equiv \frac{P\{Z \in [0, \epsilon)\}}{P\{Z \in (-\epsilon, \epsilon)\}} \\
&= \frac{\frac{1}{\epsilon} P\{Z \in [0, \epsilon)\}}{\frac{1}{\epsilon} P\{Z \in [0, \epsilon)\} + \frac{1}{\epsilon} P\{Z \in (-\epsilon, 0)\}} \\
&= \frac{f_Z^+(0)}{f_Z^+(0) + f_Z^-(0)} + \delta_\epsilon \;,
\end{aligned}
$$

where $\delta_\epsilon \to 0$ as $\epsilon \to 0$ and in the last equality we used the definitions of $f_Z^+(0)$ and $f_Z^-(0)$ in (4) and Assumption 4.1(ii) again. The result then follows by definition of π_f. ∎

Lemma B.3. *Let Assumption 4.1 hold and $q < n$ be such that $\frac{q}{n} \to 0$ as $n \to \infty$. Then,*

(a) For any $\epsilon \in (0, \frac{\delta}{2})$, $P\{\liminf_{n\to\infty}\{|Z_{g,(q+1)}^*| \leq \epsilon\}\} = P\{\liminf_{n\to\infty}\{|Z_{g,(q+1)}| \leq \epsilon\}\} = 1$.

(b) $P\{\liminf_{n\to\infty}\{S_n = S_n^*\}\} = 1$, where S_n is as in (6) and S_n^* is as in (A-19).

(c) If $P\{Z = 0\} > 0$, (a) also holds with $\epsilon = 0$.

Proof. Fix $\epsilon \in (0, \frac{\delta}{2})$ arbitrarily and set $N_n \equiv \sum_{i=1}^n I\{|Z_i| \leq \epsilon\}$. Note that $N_n \geq q+1$ implies that $Z_i^* = Z_i$ and $Z_{g,(j)}^* = Z_{g,(j)}$ for at least $q+1$ observations that are within an ϵ-neighborhood of zero. It follows that for all these observations, $A_{g,(j)} = A_{g,(j)}^*$, $Z_{g,(j)}^* \leq \epsilon$, and $Z_{g,(j)} \leq \epsilon$. We conclude that $N_n \geq q+1$ implies

$$
S_n = S_n^*, \quad Z_{g,(q+1)}^* \leq \epsilon, \quad \text{and} \quad Z_{g,(q+1)} \leq \epsilon \;.
$$

Results (a)-(b) thus follow from proving that $P\{\liminf_{n\to\infty}\{N_n \geq q+1\}\} = 1$. In order to show this, note that $N_n \sim \mathrm{Bi}(n, P\{|Z| \leq \epsilon\})$. Now set $\mu \equiv \frac{1}{2} P\{|Z| \leq \epsilon\}$, which is strictly positive by the proof of Lemma B.1. It follows that

$$
\begin{aligned}
P\{\liminf_{n\to\infty}\{N_n \geq q+1\}\} &= P\left\{\liminf_{n\to\infty}\left\{\frac{1}{n} N_n \geq \frac{1}{n}(q+1)\right\}\right\} \\
&\geq P\left\{\liminf_{n\to\infty}\left\{\frac{1}{n} N_n \geq \mu\right\}\right\} = 1
\end{aligned}
$$

where the inequality holds for all $n > \frac{1}{\mu}(q+1)$, and the last equality follows by the strong law of large numbers, i.e., $N_n/n \overset{a.s.}{\to} 2\mu > 0$. This completes the proof for parts (a)-(b).

Finally, part (c) holds by repeating the same argument for $\epsilon = 0$. The result extends in this case because $\mu \equiv \frac{1}{2} P\{|Z| = 0\} > 0$. ∎

Lemma B.4. *Let Assumption 4.1 hold and assume $P\{Z = 0\} = 0$. Fix $r \in (0, \frac{\delta}{2})$ and $q \in \{1, \ldots, n-1\}$ arbitrarily. Then, for all $s \in \mathbf{N}_q \equiv \{0, 1, \ldots, q\}$,*

$$
P\{S_n^* = s \mid |Z_{g,(q+1)}^*| = r\} = p_q(s|\pi(r))
$$

where $p_q(s|\pi(r))$ is the pmf defined in (A-20) with $\pi(r) \equiv P\{Z \geq 0 \mid |Z| < r\}$.

Proof. Let $X \equiv (|Z^*|, A^*)$ with $A^* = I\{Z^* \geq 0\}$ and note that the g-order statistics we defined in Section 3 using $g = |\cdot|$, could be alternatively obtained using X and $\tilde{g}$-order statistics where $\tilde{g}$ is now the projection into the first component of X, i.e.

$$\tilde{g}(X) = |Z^*| .$$

In this way, and for this particular choice of $\tilde{g}$, $\tilde{g}$-order statistics on X deliver

$$X_{\tilde{g},(1)} \equiv (|Z^*|_{(1)}, A^*_{[1]}) \leq_{\tilde{g}} (|Z^*|_{(2)}, A^*_{[2]}) \leq_{\tilde{g}} \cdots \leq_g (|Z^*|_{(n)}, A^*_{[n]}) \equiv X_{\tilde{g},(q)} ,$$

where the random variables $(A^*_{[1]}, \ldots, A^*_{[n]})$ are called *induced order statistics* or *concomitants* of order statistics, see David and Galambos (1974); Bhattacharya (1974).

Let $\tilde{X}_1, \ldots, \tilde{X}_q$ be i.i.d. bivariate random variables such that $\tilde{X} \stackrel{d}{=} \{X \mid \tilde{g}(X) < r\}$. Theorem 1 in Kaufmann and Reiss (1992) states that

$$\{(X_{\tilde{g},(1)}, \ldots, X_{\tilde{g},(q)}) \mid \tilde{g}(X_{\tilde{g},(q+1)}) = r\} \stackrel{d}{=} \{\tilde{X}_{\tilde{g},(1)}, \ldots, \tilde{X}_{\tilde{g},(q)}\} , \tag{B-35}$$

with $\tilde{X}_{\tilde{g},(1)}, \ldots, \tilde{X}_{\tilde{g},(q)}$ being the $\tilde{g}$-order statistics of $\tilde{X}_1, \ldots, \tilde{X}_q$, provided that (i) $\tilde{g}(X)$ has a continuous distribution and (ii) $P\{\tilde{g}(X) < r\} > 0$. Since $\tilde{g}(X) = |Z^*|$ has a continuous distribution by Lemma B.1(a) and $P\{\tilde{g}(X) < r\} = P\{|Z^*| < r\} > 0$ by Lemma B.1(c), we use (B-35) to prove our result.

Next, note that we can re-write S_n^* in (A-19) as a function of $(X_{\tilde{g},(1)}, \ldots, X_{\tilde{g},(q)})$ by using the function h that projects into the second component of X, i.e.

$$S_n^* = \sum_{j=1}^{q} A^*_{g,(j)} = \sum_{j=1}^{q} A^*_{[j]} = \sum_{j=1}^{q} h(X_{\tilde{g},(j)}) ,$$

where in the second equality we used that $A^*_{g,(j)} = A^*_{[j]}$ by definition. Using this characterization, it follows that

$$P\{S_n^* = s \mid |Z^*_{g,(q+1)}| = r\} = P\Big\{\sum_{j=1}^{q} h(X_{\tilde{g},(j)}) = s \mid |\tilde{g}(X_{\tilde{g},(q+1)})| = r\Big\}$$

$$= P\Big\{\sum_{j=1}^{q} h(\tilde{X}_{\tilde{g},(j)}) = s\Big\}$$

$$= P\Big\{\sum_{j=1}^{q} h(\tilde{X}_j) = s\Big\}$$

$$= p_q(s|\pi(r)) ,$$

where the second equality follows from (B-35), the third equality follows from $\sum_{j=1}^{q} h(\tilde{X}_{\tilde{g},(j)}) = \sum_{j=1}^{q} h(\tilde{X}_j)$, and the last equality follows from $h(\tilde{X}_1), \ldots, h(\tilde{X}_q)$ being i.i.d. bivariate random variables such that $h(\tilde{X}) \stackrel{d}{=} \{h(X) \mid \tilde{g}(X) < r\}$ and $\{h(X) \mid \tilde{g}(X) < r\} = \{I\{Z^* \geq 0\} \mid |Z^*| < r\}$ being distributed Bernoulli with parameter $\pi(r) = P\{Z^* \geq 0 \mid |Z^*| < r\}$. Since $P\{Z^* \geq 0 \mid |Z^*| < r\} = P\{Z \geq 0 \mid |Z| < r\}$ for $r \in (0, \frac{\delta}{2})$ by Lemma B.1(b), this completes the proof. ∎

Lemma B.5. *Let $S \sim \mathrm{Bi}(q, \pi)$ for $q > 1$ and consider testing $H_0 : \pi = \frac{1}{2}$ versus $H_1 : \pi \neq \frac{1}{2}$ at level $\alpha \in (0, 1)$ using the test $\phi(S)$ in (11). It follows that (i) $\phi(S)$ is a level α test, (ii) $\phi(S)$ is unbiased, (iii) $\phi(S)$ is uniformly most powerful unbiased.*

Proof. This results follows from Lehmann and Romano (2005, Example 4.2.1) after noticing that Equations (4.5) and (4.6) in Lehmann and Romano (2005) reduce to (12) in this paper (which jointly determines $a_q(\alpha)$ and $b_q(\alpha)$) under $H_0 : \pi = \frac{1}{2}$. ∎

Lemma B.6. *Let Assumption 4.1 hold and assume $P\{Z = 0\} = 0$. Let $q \to \infty$ and $\frac{q}{n} \not\to 0$ as $n \to \infty$. Then, under H_0 in (3), $\liminf_{n\to\infty} E[\phi(S_n)] > \alpha$.*

Proof. Let $Z \sim N(\mu, 1)$ with $\mu > 0$ and note that the distribution of Z is continuous everywhere. This creates two major simplifications relative to the results in Theorems 4.1 and 4.3. First, we do not need to introduce the smoothing transformation Z^* as in Lemma B.1. Second, the result in Lemma B.4 is no longer restricted to values of $r \in (0, \frac{\delta}{2})$, i.e., it holds for any $r > 0$. We exploit these two simplifications below.

Assume that $q = \tau n$ for some $\tau \in (0, 1]$ (this is without loss as $\limsup q/n > 0$ implies this must hold for a subsequence). In turn, $q = \tau n$ implies that $|Z_{g,(q+1)}|$ converges in probability to the τ-th quantile of $|Z|$, which we denote by $Q > 0$. If we now let $\pi(r) \equiv P\{Z \geq 0 \mid |Z| < r\}$, $Z \sim N(\mu, 1)$ implies that

$$\lim_{r\downarrow 0} \pi(r) = \frac{1}{2} \quad \text{and} \quad \pi(Q) > \frac{1}{2} \ ,$$

where the second statement follows from $\pi(r)$ being continuous and strictly increasing in r. Let $\xi_n \equiv \sqrt{q}\left(\frac{1}{q}S_n - \frac{1}{2}\right)$. For any $x \in \mathbf{R}$ and $\eta \in (0, Q/2)$, consider the following derivation

$$P\{\xi_n \leq x\} = \underline{R}_{n,1}(x) + \underline{R}_{n,2}(x) + \underline{R}_{n,3}(x) \ ,$$

with

$$\underline{R}_{n,1}(x) \equiv \int_0^{Q-\eta} P\{\xi_n \leq x \mid |Z_{g,(q+1)}| = r\} dP\{|Z_{g,(q+1)}| = r\}$$

$$\underline{R}_{n,2}(x) \equiv \int_{Q+\eta}^{\infty} P\{\xi_n \leq x \mid |Z_{g,(q+1)}| = r\} dP\{|Z_{g,(q+1)}| = r\}$$

$$\underline{R}_{n,3}(x) \equiv \int_{Q-\eta}^{Q+\eta} P\{\xi_n \leq x \mid |Z_{g,(q+1)}| = r\} dP\{|Z_{g,(q+1)}| = r\} \ .$$

Below we argue that $\underline{R}_{n,j}(x) = o(1)$ for $j \in \{1, 2, 3\}$ and $x \in \mathbf{R}$. This, together with the fact that $c_q(\alpha) \to \frac{1}{2}z_{\alpha/2} > 0$ as $q \to \infty$, implies that $P\{\xi_n > c_q(\alpha)\} \to 1$ and completes the proof.

First, note that

$$\sup_{x\in\mathbf{R}} |\underline{R}_{n,1}(x)| \leq \int_0^{Q-\eta} dP\{|Z_{g,(q+1)}| = r\} = P\{|Z_{g,(q+1)}| \leq Q - \eta\} \to 0 \ ,$$

which follows from $|Z_{g,(q+1)}| \xrightarrow{P} Q$ and $\eta > 0$. Similarly, $\sup_{x\in\mathbf{R}} |\underline{R}_{n,2}(x)| \leq P\{|Z_{g,(q+1)}| \geq Q + \eta\} \to 0$. Finally, for $\underline{R}_{n,3}(x)$ we need to introduce additional notation. Let $\bar{\xi}(\pi) \equiv \sqrt{q}\left(\frac{1}{q}\bar{S}(\pi) - \frac{1}{2}\right)$ where $\bar{S}(\pi) \sim$ Bi(q, π). Next, note that

$$\underline{R}_{n,3}(x) = \int_{Q-\eta}^{Q+\eta} P\{\bar{\xi}(\pi(r)) \leq x\} dP\{|Z_{g,(q+1)}| = r\}$$

$$\leq \int_{Q-\eta}^{Q+\eta} P\{\bar{\xi}(\pi(Q-\eta)) \leq x\} dP\{|Z_{g,(q+1)}| = r\}$$

$$= P\{\bar{\xi}(\pi(Q-\eta)) \leq x\} P\{|Z_{g,(q+1)}| \in [Q-\eta, Q+\eta]\} \leq P\{\bar{\xi}(\pi(Q-\eta)) \leq x\}$$

$$= P\left\{\sqrt{q}\left(\frac{1}{q}\bar{S}(\pi) - \pi(Q-\eta)\right) + \sqrt{q}\left(\pi(Q-\eta) - \frac{1}{2}\right) \leq x\right\} \to 0 \ ,$$

26

where the first line follows from Lemma B.4, and the second line follows from $\pi(r)$ being an increasing function of r and $P\{\bar{\xi}(\pi) \le x\}$ being a decreasing function of π, and the last line follows from the normal approximation to the Binomial and the fact that $\sqrt{q}\left(\pi(Q-\eta) - \frac{1}{2}\right) \to \infty$ by $\pi(Q-\eta) > \frac{1}{2}$. This completes the proof. ■

C Computational details on the data-dependent rule for q

In the simulations of Section 5 and in the companion `Stata` package, the feasible informed rule of thumb is computed as follows. First, we compute

$$\hat{q}_{\text{rot}} = \left\lceil \max\left\{q^*(\alpha), \hat{C}_\phi \frac{n}{\log n}\right\} \right\rceil ,$$

where $q^*(\alpha)$ is defined in (13),

$$\hat{C}_\phi = \frac{1}{\max\left\{25\left|\hat{\sigma}^2\phi'_{\hat{\mu},\hat{\sigma}}(\bar{z})\right|, 1\right\}} \times \frac{\phi_{\hat{\mu},\hat{\sigma}}(\bar{z})}{\phi_{\hat{\mu},\hat{\sigma}}(\hat{\mu})} ,$$

$\hat{\mu}$ is the sample mean of $\{Z_1, \ldots, Z_n\}$, $\hat{\sigma}^2$ is the sample variance of $\{Z_1, \ldots, Z_n\}$, $\bar{z}$ is the cut-off point, and n is the sample size. In principle, the value $\hat{q}_{\text{rot}}$ could be used to implement our test. However, this would ignore the non-monotonicity of the limiting null rejection probability of the non-randomized version of our test, which according to Theorem 4.2, equals $2\Psi_q(b_q(\alpha) - 1)$ with $b_q(\alpha)$ defined in (10). Figure 3 displays $2\Psi_q(b_q(\alpha) - 1)$ for $\alpha = 5\%$ as a function of q. The figure shows that $2\Psi_q(b_q(\alpha) - 1)$ takes values very close to α for q as low as 17 (i.e., 4.9%), but could be far from α for $q = 19$ (i.e., 1.9%). We therefore propose an additional layer in the data-dependent way of choosing q that guarantees that such a value delivers a local "peak" of $2\Psi_q(b_q(\alpha) - 1)$ in Figure 3.
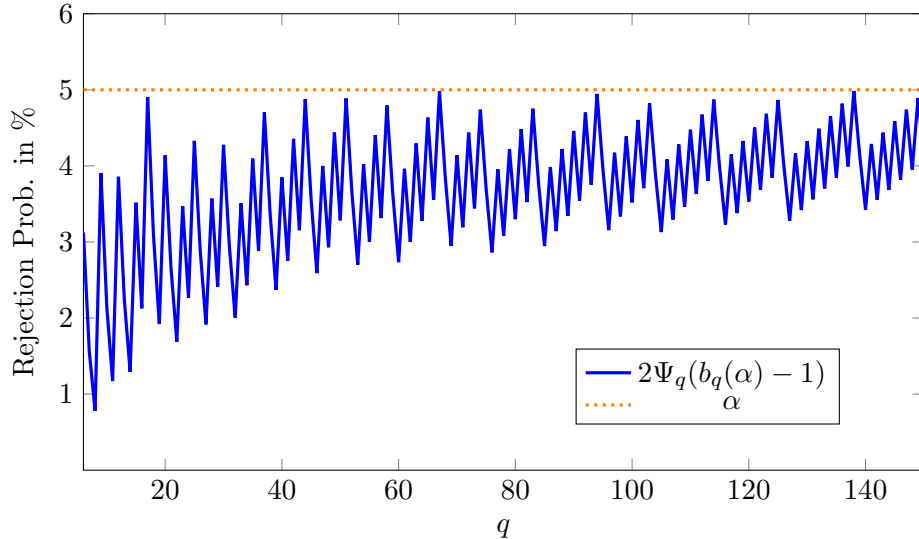

Figure 3: Limiting null rejection probability (in %) of the non-randomized version of the test, $2\Psi_q(b_q(\alpha) - 1)$, as a function of q (red solid line). Nominal level of the test (dotted orange line).

To be concrete, we define $\hat{q}_{\text{irot}}$ as

$$\hat{q}_{\text{irot}} = \underset{q \in \mathcal{N}(\hat{q}_{\text{rot}})}{\text{argmax}} \ \Psi_q(b_q(\alpha) - 1) \ , \tag{C-36}$$

where $\mathcal{N}(\hat{q}_{\text{rot}}) \equiv \{q \in \mathbf{N} : \max\{q^*(\alpha), \hat{q}_{\text{rot}} - \lceil 4\log(\hat{q}_{\text{rot}})\rceil\} \leq q \leq \hat{q}_{\text{rot}} + \lceil 4\log(\hat{q}_{\text{rot}})\rceil\}$. The value of window size $\lceil 4\log(\hat{q}_{\text{rot}})\rceil$ is the minimum number of points that are required to reach a local peak of $2\Psi_q(b_q(\alpha) - 1)$ for values of $\alpha \in \{1\%, 5\%, 10\%\}$ and is such that, for large values of q_{rot}, the window gets larger to improve the chances of getting one of the peaks closer to α as q_{rot} increases. A smaller window size may not guarantee one actually reaches a local peak. The value $\hat{q}_{\text{irot}}$ defined in (C-36) is the one we use in the simulations of Section 5 and the default value in the companion `Rdcont Stata` package.

D   Surveyed papers on RDD

Table 4 displays the list of papers we surveyed in leading journals that use regression discontinuity designs. For a description on the criteria used to compile the list of papers in Table 4, see Canay and Kamat (2018, Appendix E).

Authors (Year)	Journal	Density Test	Mean Test	Authors (Year)	Journal	Density Test	Mean Test
Schmieder et al. (2016)	AER	✓	✓	Miller et al. (2013)	AEJ:AppEcon	✓	✓
Feldman et al. (2016)	AER	✓	✓	Litschig and Morrison (2013)	AEJ:AppEcon	✓	✓
Jayaraman et al. (2016)	AER	×	×	Dobbie and Skiba (2013)	AEJ:AppEcon	✓	✓
Dell (2015)	AER	✓	✓	Kazianga et al. (2013)	AEJ:AppEcon	✓	✓
Hansen (2015)	AER	✓	✓	Magruder (2012)	AEJ:AppEcon	×	×
Anderson (2014)	AER	×	×	Dustmann and Schnberg (2012)	AEJ:AppEcon	×	×
Martin et al. (2014)	AER	×	×	Clots-Figueras (2012)	AEJ:AppEcon	✓	✓
Dahl et al. (2014)	AER	✓	✓	Manacorda et al. (2011)	AEJ:AppEcon	✓	✓
Shigeoka (2014)	AER	×	✓	Chetty et al. (2014)	QJE	✓	✓
Crost et al. (2014)	AER	×	✓	Michalopoulos and Papaioannou (2014)	QJE	×	✓
Kostol and Mogstad. (2014)	AER	✓	✓	Fredriksson et al. (2013)	QJE	✓	✓
Clark and Royer (2013)	AER	×	✓	Schmieder et al. (2012)	QJE	✓	✓
Brollo et al. (2013)	AER	✓	✓	Lee and Mas (2012)	QJE	×	×
Bharadwaj et al. (2013)	AER	✓	✓	Saez et al. (2012)	QJE	×	×
Pop-Eleches and Urquiola (2013)	AER	✓	✓	Barreca et al. (2011)	QJE	×	×
Lacetera et al. (2012)	AER	✓	×	Almond et al. (2011)	QJE	✓	✓
Duflo et al. (2012)	AER	×	×	Malamud and Pop-Eleches (2011)	QJE	✓	✓
Gopinath et al. (2011)	AER	✓	✓	Fulford (2015)	ReStat	×	✓
Auffhammer and Kellogg (2011)	AER	×	×	Snider and Williams (2015)	ReStat	×	×
Duflo et al. (2011)	AER	×	×	Doleac and Sanders (2015)	ReStat	×	×
Ferraz and Finan (2011)	AER	×	×	Coşar et al. (2015)	ReStat	×	×
McCrary and Royer (2011)	AER	×	✓	Avery and Brevoort (2015)	ReStat	×	×
Beland (2015)	AEJ:AppEcon	✓	✓	Carpenter and Dobkin (2015)	ReStat	×	✓
Buser (2015)	AEJ:AppEcon	✓	✓	Black et al. (2014)	ReStat	✓	✓
Fack and Grenet (2015)	AEJ:AppEcon	✓	✓	Anderson et al. (2014)	ReStat	×	×
Cohodes and Goodman (2014)	AEJ:AppEcon	✓	✓	Alix-Garcia et al. (2013)	ReStat	×	✓
Haggag and Paci (2014)	AEJ:AppEcon	✓	✓	Albouy (2013)	ReStat	×	×
Dobbie and Fryer (2014)	AEJ:AppEcon	✓	✓	Garibaldi et al. (2012)	ReStat	✓	✓
Sekhri (2014)	AEJ:AppEcon	✓	✓	Manacorda (2012)	ReStat	✓	✓
Schumann (2014)	AEJ:AppEcon	✓	✓	Martorell and McFarlin (2011)	ReStat	✓	✓
Lucas and Mbiti (2014)	AEJ:AppEcon	✓	✓	Grosjean and Senik (2011)	ReStat	×	×

Table 4: Papers using manipulation/placebo tests from $2011 - 2015$.

References

ARBUTHNOTT, J. (1710). An argument for divine providence, taken from the constant regularity observed in the births of both sexes. *Philosophical Transactions*, **27** 186–190.

BHATTACHARYA, P. (1974). Convergence of sample paths of normalized sums of induced order statistics. *The Annals of Statistics* 1034–1039.

CANAY, I. A. and KAMAT, V. (2018). Approximate permutation tests and induced order statistics in the regression discontinuity design. *The Review of Economic Studies*, **85** 1577–1608.

CANAY, I. A., ROMANO, J. P. and SHAIKH, A. M. (2017). Randomization tests under an approximate symmetry assumption. *Econometrica*, **85** 1013–1030.

CATTANEO, M. D., JANSSON, M. and MA, X. (2017a). Simple local polynomial density estimators. Tech. rep., Working Paper.

CATTANEO, M. D., TITIUNIK, R. and VAZQUEZ-BARE, G. (2016). Inference in regression discontinuity designs under local randomization. *Stata Journal*, **16** 331–367.

CATTANEO, M. D., TITIUNIK, R. and VAZQUEZ-BARE, G. (2017b). Comparing inference approaches for rd designs: A reexamination of the effect of head start on child mortality. *Journal of Policy Analysis and Management*, **36** 643–681.

DAVID, H. and GALAMBOS, J. (1974). The asymptotic theory of concomitants of order statistics. *Journal of Applied Probability* 762–770.

DURRETT, R. (2010). *Probability: theory and examples*. Cambridge university press.

GERARD, F., ROKKANEN, M. and ROTHE, C. (2016). Bounds on treatment effects in regression discontinuity designs with a manipulated running variable, with an application to unemployment insurance in brazil. Working paper.

HAHN, J., TODD, P. and KLAAUW, W. V. D. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, **69** pp. 201–209.

IMBENS, G. W. and LEMIEUX, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of econometrics*, **142** 615–635.

KAMAT, V. (2017). On nonparametric inference in the regression discontinuity design. *Econometric Theory,* forthcoming 1–10.

KAUFMANN, E. and REISS, R.-D. (1992). On conditional distributions of nearest neighbors. *Journal of Multivariate Analysis*, **42** 67–76.

LEE, D. S. (2008). Randomized experiments from non-random selection in u.s. house elections. *Journal of Econometrics*, **142** 675 – 697. The regression discontinuity design: Theory and applications.

LEE, D. S. and LEMIEUX, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature*, **48** 281–355.

LEHMANN, E. and ROMANO, J. P. (2005). *Testing Statistical Hypotheses.* 3rd ed. Springer, New York.

McCRARY, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, **142** 698 – 714. The regression discontinuity design: Theory and applications.

OTSU, T., XU, K.-L. and MATSUSHITA, Y. (2013). Estimation and inference of discontinuity in density. *Journal of Business & Economic Statistics*, **31** 507–524.

REISS, R.-D. (1989). *Approximate distributions of order statistics: with applications to nonparametric statistics.* Springer-Verlag, New York.

SAEZ, E. (2010). Do taxpayers bunch at kink points? *American Economic Journal: Economic Policy*, **2** 180–212.

SHEN, S. and ZHANG, X. (2016). Distributional tests for regression discontinuity: Theory and empirical examples. *Review of Economics and Statistics.* Forthcoming.

VAN DER VAART, A. W. and WELLNER, J. A. (1996). *Weak Convergence and Empirical Processes.* Springer-Verlag, New York.