# Bounds on Treatment Effects in Regression Discontinuity Designs with a Manipulated Running Variable

François Gerard, Miikka Rokkanen, and Christoph Rothe

**Abstract**

The key assumption in regression discontinuity analysis is that the distribution of potential outcomes varies smoothly with the running variable around the cutoff. In many empirical contexts, however, this assumption is not credible; and the running variable is said to be *manipulated* in this case. In this paper, we show that while causal effects are not point identified under manipulation, one can derive sharp bounds under a general model that covers a wide range of empirical patterns. The extent of manipulation, which determines the width of the bounds, is inferred from the data in our setup. Our approach therefore does not require making a binary decision regarding whether manipulation occurs or not, and can be used to deliver manipulation-robust inference in settings where manipulation is conceivable, but not obvious from the data. We use our methods to study the disincentive effect of unemployment insurance on (formal) reemployment in Brazil, and show that our bounds remain informative, despite the fact that manipulation has a sizable effect on our estimates of causal parameters.

First Version: August 15, 2015. This Version: February 7, 2020. We thank the editor, the referees, Joshua Angrist, Wilbert van der Klaauw, Henrik Kleven, Ismael Mourifié, Johannes Schmieder, Edward Vytlacil and audiences at several seminars and conferences for their helpful comments. We gratefully acknowledge the research support of the National Science Foundation via grant SES-1559058, and of the European Research Council via grant SH1-772021. Authors' contact information: François Gerard, Queen Mary, University of London, Mile End Road, London E1 4NS, UK, f.gerard@qmul.ac.uk; Miikka Rokkanen, Columbia University, Department of Economics, 420 West 118th Street, New York, NY 10027; Christoph Rothe, University of Mannheim, Department of Economics, L7 3-5, 68161 Mannheim, Germany, rothe@vwl.uni-mannheim.de

## 1. Introduction

In a regression discontinuity (RD) design, treatment assignment is determined by whether a special covariate, the running variable, falls to the left or the right of a fixed cutoff value. The treatment's average causal effect among units at the cutoff is then identified by what effectively amounts to a comparison of the average outcomes (and treatment probabilities, in the case of a fuzzy design) of units in small neighborhoods on either side of the cutoff. The key assumption for the validity of such an analysis is that the distribution of units' potential outcomes varies continuously with the running variable around the cutoff, because this ensures that the only systematic difference between units that are close to but on different sides of the cutoff is their treatment assignment.

Continuity of the potential outcome distribution given the running variable, however, may not be a credible assumption in many empirical settings where the running variable is not exogenously determined. Consider, for instance, studying the effect of a program that offers financial aid to students who score above a certain threshold on a test. Since the program affects incentives, it likely affects the running variable, i.e., test scores. This fact alone does not invalidate the key identifying assumption for an RD analysis, and published empirical papers in which the running variable is not exogenous abound in the literature (e.g., Solis, 2017). Problems arise in such settings if, for instance, students whose score came up short bargain with their teacher for extra points, or teachers might proactively give extra points to certain students with scores below the threshold. If the potential outcomes of students who become eligible for financial aid through such channels differ from those of the overall student population close to the cutoff, a conventional RD analysis is generally biased.[1]

Using now standard terminology, we refer to all setups in which such violations occur as *RD designs with a manipulated running variable*.[2] The practical importance of this issue is widely recognized in the literature. Following McCrary (2008), who argues that a jump in the density of the running variable at the cutoff is a strong indication of manipulation, it has become common empirical practice to test for the presence of such a jump. If the corresponding null hypothesis is not rejected, researchers typically proceed with their RD analysis under the assumption that continuity of the potential outcome distribution is satisfied. In contrast, the cutoff is often no longer used for inference on treatment effects if the null hypothesis is

[1]Evidence for violations of the continuity condition on the distribution of potential outcomes has been documented in many contexts. See, among many others, Urquiola and Verhoogen (2009), Camacho and Conover (2011), Scott-Clayton (2011), Card and Giuliano (2014), or Dee, Dobbie, Jacob, and Rockoff (2019).

[2]This terminology is not unproblematic, as it can be understood as suggesting that observational units are engaging in a form of wrongdoing. While this might be the case in some settings, there could also be other actors within the respective institutional contexts that are violating rules, and manipulated running variables can even occur if no rules are violated at all.

rejected.[3] This practice is problematic for at least two reasons. First, a non-rejection may not be due to the absence of manipulation but to a lack of statistical power, e.g., due to a small sample size. Units just to the left and right of the cutoff could still differ in their unobservable characteristics in this case, and estimates ignoring this possibility may be severely biased. Second, even if one correctly rejects the null hypothesis of no manipulation, the extent of the problem may be modest, and the data may remain informative. In this paper, we propose a systematic approach to dealing with the issue of potentially manipulated running variables in RD designs, which addresses both of these concerns.

We begin by laying out a simple yet general model that posits the existence of two unobservable types of units: *always-assigned* units, for which the realization of the running variable is always on one side of the cutoff (normalized to be the right side); and *potentially-assigned* units, for which the standard assumptions of an RD design are valid. The standard RD framework is a special case of our model in which always-assigned units are absent. This setup is able to capture a wide range of empirical scenarios of manipulation by appropriately assigning the two labels to specific groups of units. The only substantial requirement is that manipulation of the running variable occurs through a form of "one-sided" selection.

We then pursue a partial identification approach (e.g., Manski, 2003, 2009) that avoids making a binary decision about whether the RD design is affected by manipulation (i.e., whether always-assigned units are present). Instead, we let the data decide about the extent and "worst case" impact of the issue. This line of reasoning leads to bounds on causal parameters in two steps. First, we use the magnitude of the discontinuity in the density of the running variable at the cutoff to identify the proportion of always-assigned units among all units close to the cutoff. Second, we use this information to bound treatment effects by finding those "worst case" scenarios in which the distribution of outcomes among always-assigned units takes its "highest" and "lowest" feasible value (in a stochastic dominance sense). For sharp RD designs, the bounds are simply obtained by trimming the tails of the outcome distribution among units just to the right of the cutoff.[4] For fuzzy RD designs, the bounds are more elaborate in structure due to the various shape restrictions implied by our model. As extensions of our main results, we show that the bounds can be sharpened by using covariate

[3]Some studies also rely on ad-hoc "fixes." For instance, a "doughnut-hole" approach is sometimes used in the existing literature to estimate causal parameters in cases of potential manipulation. This method excludes observations around the cutoff somewhat heuristically, and then relies on extrapolation outside the range of the remaining data to recover estimates of treatment effects at the cutoff for a population of units that may or may not be actually observed at the cutoff under any circumstances. As we discuss below, this approach is problematic in several ways and goes against the spirit of the usual RD identification argument.

[4]This result shares similarities with that of Horowitz and Manski (1995) or Lee (2009); and some applied papers have used heuristic arguments to arrive at some version of this strategy (e.g., Card, Dobkin, and Maestas, 2009; Sallee, 2011; Anderson and Magruder, 2012; Schmieder, von Wachter, and Bender, 2012). Our contribution with regard to the sharp design is thus mainly to formalize this approach.

information, or by imposing further assumptions about the behavior of economic agents. We also show that one can identify the distribution of covariates among always-assigned and potentially-assigned units at the cutoff, which is helpful to characterize these groups.

To implement our identification results in practice, we describe computationally convenient sample analogue estimators of our bounds, and confidence intervals for the causal parameters of interest based on recent methods from the literature on set inference (e.g., Imbens and Manski, 2004; Stoye, 2009; Andrews and Soares, 2010). Software packages that implement our methods in R and Stata are available on the authors' websites. Our confidence intervals provide reliable inference on treatment effects in cases where manipulation clearly occurs. However, they are also valid in applications where it seems unclear whether the standard RD assumptions are satisfied, and we recommend their use in such settings in order to ensure that inference is robust against the possibility of manipulation.

Lastly, we illustrate the use of our approach through a study of the effect of unemployment insurance (UI) around an eligibility cutoff in Brazil. We find significant evidence of manipulation and selection at the cutoff, and our bounds imply that the magnitude of naïve RD estimates may be heavily affected by selection. Nevertheless, we are able to infer that UI takeup increases the covered UI duration by at least 35.4 days or at least .236 month per month of potential UI duration. This estimate is almost twice as large as estimates around another discontinuity, and thus for another group of workers, in Brazil (Gerard and Gonzaga, 2016). Behavioral responses to UI benefits are thus relatively large in our sample.

The rest of the paper is organized as follows. Section 2 introduces our general framework for RD designs with a manipulated running variable. Section 3 contains our main partial identification results, and Section 4 discusses estimation and inference. Section 5 presents our empirical application. Section 6 concludes. Proofs and additional material can be found in the Appendix. Throughout the paper, we use the notation that $g(c^+) = \lim_{x \downarrow c} g(x)$ and $g(c^-) = \lim_{x \uparrow c} g(x)$ for a generic function $g(\cdot)$. We also follow the convention that whenever we take a limit we implicitly assume that this limit exists and is finite. Similarly, whenever an expectation or some other moment of a random variable is taken, it is implicitly assumed that the corresponding object exists and is finite.

## 2. Model and Parameters of Interest

In this section, we introduce a general model for RD designs in which manipulation possibly occurs, discuss its applicability, and clarify the interpretation of the parameters of interest.

### 2.1. **Model.**

The general structure of the data is the same as in conventional RD designs in our setup. We observe independent data points $(X_i, Y_i, D_i)$, $i = 1, \ldots, n$, where $X_i$ is the running variable, $Y_i$ is the outcome of interest, and $D_i$ is the actual treatment status, with $D_i = 1$ if unit $i$ receives the treatment, and $D_i = 0$ otherwise. Units are assigned to the treatment group if $X_i \geq c$ for some fixed cutoff value $c$. Our RD design is said to be sharp if $D_i = \mathbb{I}(X_i \geq c)$, and said to be fuzzy otherwise.

The main structural feature of our model is that the population under study can be partitioned into two groups with membership indicated by an unobservable dummy variable $M_i \in \{0, 1\}$. In a sense made precise below, units with $M_i = 0$, which we call *potentially-assigned*, behave as prescribed by a standard RD framework, while units with $M_i = 1$, which we call *always-assigned*, are only restricted to have realization of the running variable on one side of the cutoff (which we normalize to be the right side without loss of generality). Potentially-assigned units also have potential outcomes $Y_i(d)$, for $d \in \{0, 1\}$, corresponding to the outcome unit $i$ would have experienced had it received treatment $d$; and potential treatment states $D_i(x)$, for $x \in \text{supp}(X_i)$, corresponding to the treatment status unit $i$ would have experienced if the value $x$ had been used to determine its treatment assignment. For potentially-assigned units we thus have $Y_i = Y_i(D_i)$ and $D_i = D_i(X_i)$, respectively, and we also put $D_i^+ = D_i(c^+)$ and $D_i^- = D_i(c^-)$. Always-assigned units are not even required to have potential outcomes and treatment states. Manipulation then exists in this setup by definition whenever always-assigned units exist in the population. The exact behavior of the units is restricted through the following three assumptions.

**Assumption 1.** *(i)* $\text{P}(D = 1|X = c^+, M = 0) > \text{P}(D = 1|X = c^-, M = 0)$*; (ii)* $\text{P}(D^+ \geq D^-|X = c, M = 0) = 1$*; (iii)* $\text{P}(Y(d) \leq y|D^+ = d^1, D^- = d^0, X = x, M = 0)$*,* $\text{E}(Y(d)|D^+ = d^1, D^- = d^0, X = x, M = 0)$ *and* $\text{P}(D^+ = d^1, D^- = d^0|X = x, M = 0)$ *are continuous in $x$ at $c$ for $d, d^0, d^1 \in \{0, 1\}$ and all $y$; (iv) $F_{X|M=0}(x)$ is differentiable in $x$ at $c$, and the derivative is strictly positive.*

This assumption implies that the standard conditions from the RD literature are satisfied among potentially-assigned units.[5] Assumption 1(i)–(iii) impose, respectively, a non-zero first stage, a monotonicity or "no defiers" condition, and a key continuity condition which requires the distributions of potential outcomes and potential treatment states to be the same on both sides of the cutoff. Note that Assumptions 1(i)-(iii) simplify to the condition that $\text{E}(Y(d)|X = x, M = 0)$ is continuous in $x$ at $c$ for $d \in \{0, 1\}$ for the special case of a sharp RD design. Assumption 1(iv) ensures that there are potentially-assigned units close to

[5]We formalize the notion of a RD design in terms of continuity conditions on the distributions of potential outcomes and treatment states as in Frandsen, Frölich, and Melly (2012), Dong (2018) or Bertanha and Imbens (2019). This leads to the same identification results as directly imposing the local independence condition that the treatment effect is independent of the treatment status conditional on the running variable near the cutoff, as in Hahn, Todd, and Van der Klaauw (2001).

the cutoff on either side, which is crucial for any identification argument based on comparing units just to the left and right of the cutoff.

**Assumption 2.** *The derivative of* $F_{X|M=0}(x)$ *is continuous in* $x$ *at* $c$.

This is a weak regularity condition on the distribution of the running variable among potentially-assigned units. Together with Assumption 1(iv), it implies that the density of $X_i$ among potentially-assigned units is smooth and strictly positive over some open neighborhood of $c$. Continuity of the running variable's density around the cutoff is a reasonable condition in applications, and is generally considered to be an indication for the absence of manipulation in the applied literature (Lee, 2008; McCrary, 2008).

**Assumption 3.** *(i)* $\mathrm{P}(X \geq c|M=1) = 1$, *(ii)* $F_{X|M=1}(x)$ *is right-differentiable in* $x$ *at* $c$.

Assumption 3 is the only restriction we impose on always-assigned units. Its first part implies that the running variable only takes on values to the right of the cutoff among these units. This (local) *one-sided manipulation* assumption is key for the identification argument in the next section as it allows us to identify the proportion of always-assigned units among all units close to the cutoff. The second part rules out that the running variable is exactly equal to the cutoff value for some (or all) always-assigned units. If this was the case, one could easily identify the units who are problematic for the validity of the RD design through their value of the running variable, and simply remove them from the analysis. Finally, together with Assumption 2, Assumption 3 also implies that the running variable is continuously distributed in the full population, with a density that is potentially discontinuous at $c$.

2.2. **Discussion.** Our model is able to capture a wide range of empirical scenarios in which validity of the standard RD design could be achieved by excluding a group of problematic units just to the right of the cutoff from the data. As an example, suppose students need to achieve a certain test score in order to be admitted to a prestigious school. The test is taken by all students, even those who do not plan to attend the prestigious school (e.g., because tuition is too high), and those who would be admitted even if their score falls below the cutoff (e.g., because of legacy admissions). After a preliminary round of grading, the teacher decides to "bump up" the scores of those students below the admission cutoff to some value above the cutoff if she believes that the student would highly benefit from attending the prestigious school. We then only observe the final score assigned by the teacher.

In such a scenario, there is manipulation as long as the teacher's decision to "bump up" students is related to their potential outcomes. Students whom the teacher believes would highly benefit from attending the prestigious school are always-assigned: their final test score is always above the cutoff, either because they were "bumped up", or because they already scored above the cutoff in the preliminary round of grading. All other students are potentially-assigned. Moreover, in this scenario the RD design is likely fuzzy for both groups, in the sense that we can expect to see treated and untreated units among the potentially-assigned (below and above the cutoff) and the always-assigned ones (above the cutoff).

Through similar reasoning, one can fit a wide range of settings into our model, including settings in which no agent engages in any form of wrongdoing, by assigning the labels of always-assigned and potentially-assigned appropriately to specific groups of units. We illustrate this point in more detail in Appendix D.

2.3. **Parameter of Interest.** We focus on causal effects among potentially-assigned units as our parameter of interest in this paper. Specifically, we study identification of

$$\Gamma = \mathrm{E}(Y(1) - Y(0)|X = c, D^+ > D^-, M = 0).$$

which can be understood as the local average treatment effect for the subgroup of potentially-assigned "compliers", who receive the treatment if and only if their value of the running variable $X_i$ is above the cutoff (Imbens and Angrist, 1994). It is the natural analogue to the full population local average treatment effect $\mathrm{E}(Y(1) - Y(0)|X = c, D^+ > D^-)$ typically considered in RD design without manipulation, in that both capture the causal effect for units for which the RD design is valid.[6]

The parameter $\Gamma$ also retains a notion of policy relevance similar to the parameter of interest in setups without manipulation. Specifically, it represents the causal effect for units whose treatment status would change following marginal changes in the level of the cutoff. This can be illustrated using the example from the previous subsection. If the admission cutoff for the prestigious school increases, the teacher might still "bump up" the scores of students she believes would highly benefit from attending. Treatment assignment might thus only change for potentially-assigned "complier" students, whose effect is measured by $\Gamma$.

One should note that $\Gamma$ is the causal effects for a population that is actually observed at the cutoff, and not some hypothetical population that one would observe at the cutoff under some circumstances. In particular, our approach does not require assuming the existence of a hypothetical "true" value of the running variable that one would supposedly observe if one could for instance "close" the institutional channel that causes manipulation. It also avoids making assumptions about how such a "true" value and the observed value of the running variable are related. We see this as an advantage relative to "doughnut hole" RD designs, for example, which are sometimes used in applications where manipulation is a concern.[7]

[6]Since our model imposes hardly any restrictions on the behavior of always-assigned units, it is not possible to derive meaningful conclusions about the causal effect of the treatment on them from observable quantities.

[7]Doughnut hole RD designs exclude observations around the cutoff and extrapolate trends estimated

## 3. Identification

In this section, we derive our main results regarding the identification of $\Gamma$. We first state some preliminary results, then consider the relatively simple case of a sharp RD design, before finally analyzing the general case of a fuzzy RD design. Proofs are given in Appendix A. We also give an overview of a number of extensions to our main identification results (e.g., quantile treatment effects), which are collected in Appendix C. To present the results, it will be useful to have the following shorthand notation to categorize various types of units:

$$
\begin{aligned}
C_0 &= \{D^+ > D^-, M = 0\}, \text{ potentially-assigned compliers;} \\
A_0 &= \{D^+ = D^- = 1, M = 0\}, \text{ potentially-assigned always-takers;} \\
N_0 &= \{D^+ = D^- = 0, M = 0\}, \text{ potentially-assigned never-takers;} \\
T_1 &= \{D = 1, M = 1\}, \text{ always-assigned treated units;} \\
U_1 &= \{D = 0, M = 1\}, \text{ always-assigned untreated units.}
\end{aligned}
$$

3.1. **Preliminaries.** Since it is not possible to determine whether a specific unit is of the always-assigned or the potentially-assigned type, $\Gamma$ is generally not point identified under manipulation of the running variable. We therefore derive sharp lower and upper bounds on this parameter for both sharp and fuzzy RD designs. Our general strategy is to first obtain sharp lower and upper bounds, in a first-order stochastic dominance sense, on the c.d.f.s $F_{Y(d)|X=c,C_0}$ for $d \in \{0,1\}$. That is, we derive c.d.f.s $F_d^U$ and $F_d^L$ that are feasible candidates for $F_{Y(d)|X=c,C_0}$ in the sense that they are compatible with our assumptions and the population distribution of observable quantities, and that are such that $F_d^U \succeq F_{Y(d)|X=c,C_0} \succeq F_d^L$, where $\succeq$ denotes first-order stochastic dominance. Once these c.d.f. bounds have been obtained, it follows from Stoye (2010, Lemma 1) that sharp upper and lower bounds on $\Gamma$ are given, respectively, by

$$\Gamma^U \equiv \int y dF_1^U(y) - \int y dF_0^L(y) \quad \text{and} \quad \Gamma^L \equiv \int y dF_1^L(y) - \int y dF_0^U(y).$$

An advantage of this approach is that, given bounds on the c.d.f.s of potential outcomes, it is straightforward to consider quantile treatment effects as well. For notational convenience, all results in this section are stated for the case of a continuously distributed outcome variable; we extend our results to outcomes whose distribution has mass points in Appendix C.

outside of the excluded range to the cutoff. The result is commonly interpreted as a causal effect for a population that would be observed at the cutoff if the distribution of potential outcomes there would follow its trend from outside the excluded range. This hypothetical population is often considered to be the one that would be observed in a counterfactual in which the channel leading to a manipulated running variable was "closed." This interpretation requires strong assumptions regarding how manipulation occurs, and statistical assumptions implying that extrapolation biases are small. No such assumptions are required in our case.

Our analysis repeatedly uses an important intermediate quantity, the proportion of always-assigned units among all units just to the right of the cutoff, which we denote by

$$\tau \equiv \mathrm{P}(M = 1|X = c^+). \tag{3.1}$$

While we cannot observe or infer the type of any given unit, under our assumptions we can point identify $\tau$ from the size of the discontinuity in the density $f_X$ of the observed running variable at the cutoff. We formally state this insight in the following Lemma.

**Lemma 1.** *If Assumptions 1–3 hold, then $\tau = 1 - f_X(c^-)/f_X(c^+)$ is point identified.*

3.2. **Sharp RD Designs.** In a sharp RD design every unit receives the treatment if and only if its value of the running variable is to the right of the cutoff. Since every unit just to the left of the cutoff is potentially-assigned, the distribution of $Y$ in this subpopulation coincides with the distribution of $Y(0)$ among potentially-assigned compliers ($C_0$) at the cutoff:

$$F_{Y(0)|X=c,C_0}(y) = F_{Y|X=c^-}(y).$$

To bound $\Gamma$, we therefore only need to bound the distribution of $Y(1)$ among potentially-assigned compliers at the cutoff. Information about $Y(1)$ is only contained in the subpopulation of treated units, which contains potentially-assigned compliers and always-assigned treated units ($C_0$ and $T_1$). Sharpness of the RD design then implies that $\mathrm{P}(T_1|X = c^+) = \tau$. Since this quantity is point identified by Lemma 1, we proceed analogously to Horowitz and Manski (1995) or Lee (2009) to obtain a bound on $F_{Y(1)|X=c,C_0}(y)$. In particular, a sharp upper bound on $F_{Y(1)|X=c,C_0}(y)$, in a first-order stochastic dominance sense, is obtained by truncating the distribution $F_{Y|X=c^+}(y)$ below its $\tau$-quantile. A sharp lower bound is obtained analogously by truncating $F_{Y|X=c^+}(y)$ above its $(1-\tau)$-quantile. That is, the bounds on $F_{Y(1)|X=c,C_0}(y)$ are given, respectively, by

$$F^U_{1,SRD}(y) = F_{Y|X=c^+,Y\geq Q_{Y|X=c^+}(\tau)}(y) \quad \text{and} \quad F^L_{1,SRD}(y) = F_{Y|X=c^+,Y\leq Q_{Y|X=c^+}(1-\tau)}(y).$$

These bounds correspond to the "extreme" scenarios in which the proportion $1-\tau$ of units just to the right of the cutoff with either the highest or the lowest outcomes are the potentially-assigned units. These bounds are sharp because both "extreme" scenarios are empirically feasible. The following theorem translates these findings into explicit bounds on $\Gamma$.

**Theorem 1.** *Suppose Assumptions 1–3 hold, that $P(D^+ > D^-) = 1$, and that $F_{Y|X=c^+}(y)$ is*

Table 1: Allocation of Units' Types in the Fuzzy RD Design

| Subset of population | Types of units present |
|---|---|
| $X = c^+, D = 1$ | $C_0,\ A_0,\ T_1$ |
| $X = c^-, D = 1$ | $A_0$ |
| $X = c^+, D = 0$ | $N_0,\ U_1$ |
| $X = c^-, D = 0$ | $C_0,\ N_0$ |

*Note:* See Section 2.1 for a definition of units' types.

*continuous in $y$. Then sharp lower and upper bounds on $\Gamma$ are given by*

$$\Gamma^L_{SRD} = \mathrm{E}(Y|X = c^+, Y \leq Q_{Y|X=c^+}(1-\tau)) - \mathrm{E}(Y|X = c^-) \quad \text{and}$$
$$\Gamma^U_{SRD} = \mathrm{E}(Y|X = c^+, Y \geq Q_{Y|X=c^+}(\tau)) - \mathrm{E}(Y|X = c^-),$$

*respectively.*

3.3. **Fuzzy RD Designs.** In a fuzzy RD design with a manipulated running variable, the population of potentially-assigned units might contain always-takers and never-takers in addition to compliers, and always-assigned untreated units might exist in addtion to treated ones. As shown in Table 1, there are thus five different types of units and four possible combinations of treatment assignments and treatment decisions that are relevant for our analysis. To derive bounds on the distributions of the two potential outcomes among potentially-assigned compliers ($C_0$) at the cutoff, we begin by introducing the following notation for the proportion of always-assigned units among those units with treatment status $d \in \{0, 1\}$ just to the right of the cutoff:

$$\tau_d \equiv \mathrm{P}(M = 1|X = c^+, D = d), \quad d \in \{0, 1\}. \tag{3.2}$$

We then proceed in three steps. In Step 1 and 2 we obtain bounds on the distribution of potential outcomes under treatment and non-treatment, respectively, for the hypothetical case in which the true values of $\tau_1$ and $\tau_0$ are known. In Step 3, we then derive our final bounds on $\Gamma$, given that the true values of $\tau_1$ and $\tau_0$ are in fact unknown.

**Step 1: Distribution of Potential Outcome under Treatment.** We begin by considering bounds on $F_{Y(1)|X=c,C_0}$. Information about the distribution of $Y(1)$ is only contained in the data on treated units. From Table 1, we see that the subpopulation of treated units just to the left of the cutoff consists exclusively of potentially-assigned always-takers ($A_0$). The c.d.f. $F_{Y(1)|X=c,A_0}$ is therefore point identified:

$$F_{Y(1)|X=c,A_0}(y) = F_{Y|X=c^-,D=1}(y).$$

Using simple algebra, we find that the proportion of $A_0$ units among treated units just to the right of the cutoff, which we denote by $\kappa_1$, is point identified as well:

$$\kappa_1 \equiv \mathrm{P}(A_0|X = c^+, D = 1) = (1 - \tau) \cdot \frac{\mathrm{E}(D|X = c^-)}{\mathrm{E}(D|X = c^+)}. \tag{3.3}$$

To simplify the notation, we also define

$$G(y) \equiv F_{Y(1)|X=c,C_0 \cup T_1}(y).$$

It then follows from the law of total probability that this c.d.f. is also point identified:[8]

$$G(y) = \frac{1}{1 - \kappa_1} \left( F_{Y|X=c^+,D=1}(y) - \kappa_1 F_{Y|X=c^-,D=1}(y) \right).$$

The c.d.f. $F_{Y(1)|X=c,C_0}$ can now be bounded sharply by considering the two "extreme" scenarios in which potentially-assigned compliers ($C_0$) are those units just to the right of the cutoff in the subpopulation $C_0 \cup T_1$ with either the highest or the lowest outcomes. The share of $C_0$ units in this subpopulation is

$$\mathrm{P}(C_0|X = c^+, C_0 \cup T_1) = 1 - \frac{\tau_1}{1 - \kappa_1}.$$

Given knowledge of $\tau_1$, we therefore obtain a sharp upper bound on $F_{Y(1)|X=c,C_0}$, in a first-order stochastic dominance sense, by truncating the distribution $G$ below its $\tau_1/(1 - \kappa_1)$ quantile. Analogously, we obtain a sharp lower bound by truncating $G$ above its $1-\tau_1/(1-\kappa_1)$ quantile. With some algebra, these bounds on $F_{Y(1)|X=c,C_0}$ given knowledge of $(\tau_1, \tau_0)$ can be written, respectively, as

$$F^U_{1,FRD}(y, \tau_1, \tau_0) = \frac{(1 - \kappa_1) \cdot G(y) - \tau_1}{1 - \kappa_1 - \tau_1} \cdot \mathbb{I}\left\{ y \geq G^{-1}\left( \frac{\tau_1}{1 - \kappa_1} \right) \right\} \quad \text{and}$$
$$F^L_{1,FRD}(y, \tau_1, \tau_0) = \frac{(1 - \kappa_1) \cdot G(y)}{\tau_1} \cdot \mathbb{I}\left\{ y \leq G^{-1}\left( 1 - \frac{\tau_1}{1 - \kappa_1} \right) \right\},$$

**Step 2: Distribution of Potential Outcome under Non-Treatment.** Next, we consider bounds on $F_{Y(0)|X=c,C_0}$. Information about the distribution of $Y(0)$ is only contained in the data on untreated units. From Table 1, we see that untreated potentially-assigned compliers ($C_0$) are never observed in isolation just to the left of the cutoff, but only together with

[8]The quantity on the right-hand-side of the following equation is guaranteed to be a proper c.d.f. in our model. If that were not to be the case empirically, this would mean that our model is rejected by the data.

potentially-assigned never-takers ($N_0$). Given knowledge of $\tau_0$, the share of the latter type of units, which we denote by $\kappa_0 \cdot (1-\tau_0)$, is point identified:

$$\mathrm{P}(N_0|X=c^-, D=0) = \kappa_0 \cdot (1-\tau_0), \qquad \kappa_0 = \frac{1}{1-\tau} \cdot \frac{1-\mathrm{E}(D|X=c^+)}{1-\mathrm{E}(D|X=c^-)}. \tag{3.4}$$

If we were to use only information from untreated units just to the left of the cutoff, we could therefore obtain lower and upper bounds on $F_{Y(0)|X=c,C_0}(y)$ by truncating the distribution $F_{Y|X=c^-,D=0}(y)$ below its $\kappa_0 \cdot (1-\tau_0)$ quantile and above its $1-\kappa_0 \cdot (1-\tau_0)$ quantile, respectively. However, such bounds are generally not sharp. This is because they correspond to "extreme" scenarios in which potentially-assigned never-takers ($N_0$) have either the highest or the lowest outcomes among untreated units just to the left of the cutoff. By Assumption 1, however, the c.d.f. $F_{Y(0)|X=x,N_0}(y)$ varies continuously in $x$ around the cutoff, and thus these two "extreme" scenarios might be at odds with the distribution of outcomes that we observe among untreated units just to the right of the cutoff. Indeed, from Table 1, we see that the subpopulation of untreated units just to the right of the cutoff also contains potentially-assigned never-takers, together with always-assigned untreated units ($U_1$), and their share in this subpopulation is

$$\mathrm{P}(N_0|X=c^+, D=0) = 1-\tau_0.$$

We can thus write the density $f_{Y(0)|X=c,N_0}(y)$ in two different ways using information from either side of the cutoff (assuming $\kappa_0 > 0$ and $\tau_0 < 1$):

$$f_{Y(0)|X=c,N_0}(y) = \frac{f_{Y|X=c^-,D=0}(y) - (1-\kappa_0 \cdot (1-\tau_0)) f_{Y(0)|X=c,C_0}(y)}{\kappa_0 \cdot (1-\tau_0)} \quad \text{and} \tag{3.5}$$

$$f_{Y(0)|X=c,N_0}(y) = \frac{f_{Y|X=c^+,D=0}(y) - \tau_0 f_{Y(0)|X=c,U_1}(y)}{1-\tau_0}. \tag{3.6}$$

To be compatible with the distribution of $Y$ among untreated units on either side of the cutoff, any feasible candidate for $f_{Y(0)|X=c,N_0}(y)$ thus has to be such that

$$f_{Y(0)|X=c,N_0}(y) \leq s(y,\tau_0)$$

for all $y \in \mathbb{R}$, where

$$s(y,\tau_0) \equiv \frac{1}{1-\tau_0} \cdot \min\left\{ \frac{1}{\kappa_0} \cdot f_{Y|X=c^-,D=0}(y), f_{Y|X=c^+,D=0}(y) \right\}.$$

This is because otherwise one of the density functions $f_{Y(0)|X=c,C_0}(y)$ or $f_{Y(0)|X=c,U_1}(y)$ would have to take a negative value in order for equations (3.5)–(3.6) to be satisfied. The most "extreme" feasible candidates for $F_{Y(0)|X=c,N_0}(y)$, which put as much probability mass as possible to one of the tail regions of the support of the outcome variable, are then given by

$$F^U_{Y(0)|X=c,N_0}(y) = \int_{-\infty}^{y} s(t,\tau_0)\mathbb{I}\left\{t \geq q_U(\tau_0)\right\} dt \quad \text{and}$$
$$F^L_{Y(0)|X=c,N_0}(y) = \int_{-\infty}^{y} s(t,\tau_0)\mathbb{I}\left\{t \leq q_L(\tau_0)\right\} dt,$$

respectively, where $q_U(\tau_0)$ and $q_L(\tau_0)$ are constants such that

$$\int_{q_U(\tau_0)}^{\infty} s(t,\tau_0)dt = \int_{-\infty}^{q_L(\tau_0)} s(t,\tau_0)dt = 1. \tag{3.7}$$

We illustrate this construction in Figure 1. The "extreme" candidates for $F_{Y(0)|X=c,N_0}(y)$ directly correspond to "extreme" candidates for the density $f_{Y(0)|X=c,C_0}(y)$ through the relationship (3.5), which in turn yields the following sharp upper and lower bounds, in a first-order stochastic dominance sense, on the c.d.f. $F_{Y(0)|X=c,C_0}$ given knowledge of $(\tau_1,\tau_0)$:

$$F^U_{0,FRD}(y,\tau_1,\tau_0) = \frac{F_{Y|X=c^-,D=0}(y) - \kappa_0\cdot(1-\tau_0)F^L_{Y(0)|X=c,N_0}(y)}{1-\kappa_0\cdot(1-\tau_0)} \quad \text{and}$$
$$F^L_{0,FRD}(y,\tau_1,\tau_0) = \frac{F_{Y|X=c^-,D=0}(y) - \kappa_0\cdot(1-\tau_0)F^U_{Y(0)|X=c,N_0}(y)}{1-\kappa_0\cdot(1-\tau_0)}.$$

If the envelope function $s(\cdot,\tau_0)$ is a proper density these two bounds coincide, and the c.d.f. $F_{Y(0)|X=c,C_0}$ is point identified. There are two main scenarios in which this would be the case. First, there exist no untreated always-assigned units just to the right of the cutoff, and thus $\tau_0 = 0$. Second, there exist no untreated units of any type just to the right of the cutoff, and thus $\mathrm{E}(D|X=c^+) = 1$.

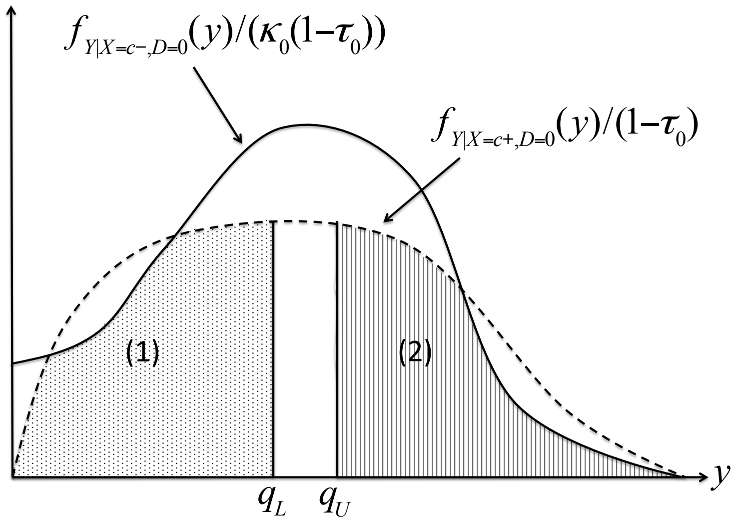
**Step 3: Bounds on Parameter of Interest.** The analysis in Steps 1 and 2 shows that if we knew the values of $\tau_1$ and $\tau_0$, sharp upper and lower bounds on the local average treatment effect $\Gamma$ would be given by

$$\begin{aligned}\Gamma^U_{FRD}(\tau_1,\tau_0) &\equiv \int y dF^U_{1,FRD}(y,\tau_1,\tau_0) - \int y dF^L_{0,FRD}(y,\tau_1,\tau_0) \quad \text{and}\\ \Gamma^L_{FRD}(\tau_1,\tau_0) &\equiv \int y dF^L_{1,FRD}(y,\tau_1,\tau_0) - \int y dF^U_{0,FRD}(y,\tau_1,\tau_0),\end{aligned} \tag{3.8}$$

respectively. However, these bounds are not directly feasible, as the population values of $\tau_1$ and $\tau_0$ are generally unknown. Nevertheless, the two values can be shown to be partially identified. To see this, note that there are four logical restrictions on the range of their plausible values. First, since $\tau_1$ and $\tau_0$ are probabilities, it has to be the case that

$$(\tau_1,\tau_0) \in [0,1]^2 \tag{3.9}$$

Figure 1: Illustration of the construction of our upper and lower bounds for $F_{Y(0)|X=c,N_0}$




Notes: The solid and dotted lines represent the graph of the functions $f_{Y|X=c^-,D=0}(y)/((1-\tau_0)\kappa_0)$ and $f_{Y|X=c^+,D=0}(y)/(1-\tau_0)$, respectively. The function $s(y,\tau_0)$ is the pointwise minimum of these two functions. The upper contours of the shaded areas (1) and (2) then correspond to the densities of $F^L_{Y(0)|X=c,N_0}$ and $F^U_{Y(0)|X=c,N_0}$, respectively, as the constants $q_L(\tau_0)$ and $q_U(\tau_0)$ are chosen such that the surface of the shaded areas is equal to 1. Note that it is not necessarily the case that $q_L(\tau_0) < q_U(\tau_0)$.

Second, by the law of total probability, it must hold that

$$\tau = \tau_1 \cdot \mathrm{E}(D|X = c^+) + \tau_0 \cdot (1 - \mathrm{E}(D|X = c^+)). \tag{3.10}$$

Third, our monotonicity condition in Assumption 1(i) implies that

$$\mathrm{E}(D|X = c^+) \cdot \frac{1-\tau_1}{1-\tau} > \mathrm{E}(D|X = c^-). \tag{3.11}$$

Note that this condition can be equivalently stated as $\tau_1 < 1 - \kappa_1$, and ensures that the c.d.f. $G$ in Step 1 is truncated at a proper quantile level. Finally, requiring the terms $q_U(\tau_0)$ and $q_L(\tau_0)$, defined in (3.7), to be well-defined implies that

$$\int s(y,\tau_0)dy \geq 1. \tag{3.12}$$

These four conditions exhaust the informational content of our model regarding the possible values of $(\tau_1,\tau_0)$. Therefore the set $\mathcal{T}$ of candidates that satisfy these four restrictions, formally given by

$$\mathcal{T} \equiv \{(\tau_1,\tau_0) : \text{conditions (3.9)–(3.12) are satisfied}\},$$

is the sharp identified set for $(\tau_1, \tau_0)$. Using this result, we can now find sharp bounds on $\Gamma$ by finding those values of $(\tau_1, \tau_0) \in \mathcal{T}$ that lead to the most extreme values of the quantities defined in (3.8).[9] These bounds on $\Gamma$ are sharp because they are based on assigning "worst case" distributions of the potential outcomes to each of the six groups mentioned in Table 1 that satisfy our assumptions and are compatible with the distribution of observables.

**Theorem 2.** *Suppose that Assumptions 1–3 hold, and that $F_{Y|X=c^+,D=d}(y)$ and $F_{Y|X=c^-,D=d}(y)$ are continuous in $y$ for $d \in \{0, 1\}$. Then sharp lower and upper bounds on $\Gamma$ are given by*

$$\Gamma^L_{FRD} = \inf_{(t_1,t_0)\in\mathcal{T}} \Gamma^L_{FRD}(t_1, t_0) \quad \text{and} \quad \Gamma^U_{FRD} = \sup_{(t_1,t_0)\in\mathcal{T}} \Gamma^U_{FRD}(t_1, t_0),$$

*respectively.*

3.4. **Additional Results.** We present here a brief overview of a number of extensions to our main identification results; for the sake of brevity, details are relegated to Appendix C.

**Quantile Treatment Effects.** Quantile treatment effects can be an attractive alternative to average effects in applications because they are less sensitive to variation in the outer tails of the outcome distribution. Since our identification results for $\Gamma$ are based on first-order stochastic dominance bounds on the respective conditional c.d.f. of potential outcomes, they are straightforward to extend to quantile counterparts of these parameters, such as $\Psi(u) \equiv Q_{Y(1)|X=c^-,D^+>D^-}(u) - Q_{Y(0)|X=c^-,D^+>D^-}(u)$, where $u \in (0, 1)$ is some quantile level.

**Non-Continuously Distributed Outcomes.** Theorem 1 and 2 are stated for the special case of a continuously distributed outcome variable. This is for simplicity only, and our results immediately generalize to the case of a discrete outcome variable, which occurs frequently in empirical applications. Discrete outcomes do not pose any conceptual challenges, but some care needs to be taken when defining truncated distributions.

**Behavioral Assumptions.** In some applications, it seems plausible that the probability of actually receiving the treatment conditional on being eligible should be relatively high in some appropriate sense for always-assigned units. For instance, if manipulation results from some units making a conscious effort to locate to the right of the cutoff, they will likely want to receive the treatment conditional on being eligible. This could be modeled for example by assuming that always-assigned units are more likely to be treated than potentially-assigned ones, or by assuming that all always-assigned units are treated. We show that imposing assumptions of this kind can narrow the bounds in Theorem 2 by shrinking the set $\mathcal{T}$.

[9] Note that under the model in Section 2.1 the set $\mathcal{T}$ *has* to be non-empty. If that were not to be the case empirically, this would mean that our model is rejected by the data.

**Covariates.** Following arguments similar to those in Lee (2009), covariates that are measured prior to treatment assignment can also be used to narrow the bounds in Theorem 1 and 2. The idea is that, if the outcome distribution or the proportion of always-assigned units just to the right of the cutoff changes with the covariates, trimming units based on their position in the outcome distribution conditional on the covariates leads to units with less extreme values in the overall outcome distribution being trimmed. Additionally, we show that one can also identify the distribution of covariates among always-assigned and potentially-assigned units. This could be useful for targeting policies aimed at mitigating manipulation for instance.

## 4. Estimation and Inference

While our main focus in this paper is on deriving identification results for causal effects in RD designs with a manipulated running variable, this section also discusses some methods for estimation and inference, based on the results in Section 3.[10]

4.1. **Estimation of the Bounds.** We describe the construction of our final estimates of the bounds on $\Gamma$ for the general case of a Fuzzy RD design. Bounds for the sharp case can be obtained in a more simple fashion. First, note that the set $\mathcal{T}$ is a straight line in the unit square, and can therefore be represented in terms of the location of the endpoints of the line:

$$\mathcal{T} = \{(\eta_1(t), \eta_0(t)) : t \in [0,1]\} \quad \text{with} \quad \eta_d(t) = \tau_d^L + t \cdot (\tau_d^U - \tau_d^L)$$

for $d \in \{0,1\}$, where

$$\tau_1^L = \max\left\{0, 1 - \frac{1-\tau}{g^+}\right\}, \quad \tau_1^U = \min\left\{1 - \frac{(1-\tau)\cdot g^-}{g^+}, \frac{\tau - \max\{0, 1 - \int \tilde{s}(y)dy\}(1-g^+)}{g^+}\right\},$$
$$\tau_0^U = \min\left\{1, \frac{\tau}{1-g^+}\right\}, \qquad \tau_0^L = \max\left\{0, \tau - \frac{(1-\tau)\cdot(g^+ - g^-)}{1-g^+}, 1 - \int \tilde{s}(y)dy\right\},$$

with $\tilde{s}(y) = \min\{f_{Y|X=c^-,D=0}(y)/\kappa_0, f_{Y|X=c^+,D=0}(y)\}$, $g^+ = \mathrm{E}(D_i|X_i = c^+)$ and $g^- = \mathrm{E}(D_i|X_i = c^-)$. Dropping the "FRD" subscript to simplify the notation, the bounds on $\Gamma$ from Theorem 2 can then be written as

$$\Gamma^L = \inf_{t\in[0,1]} \Gamma^L(\eta_1(t), \eta_0(t)) \quad \text{and} \quad \Gamma^U = \sup_{t\in[0,1]} \Gamma^U(\eta_1(t), \eta_0(t)).$$

This expression is convenient because it makes the area over which optimization takes place free of unknown quantities that have to be estimated.

With this notation, estimates of our bounds on $\Gamma$ can be obtained through a "plug-in"

[10]Our approach uses a number of different techniques that are well-understood individually, but whose combination requires a tedious theoretical analysis. We therefore do not present any formal results here.

approach that replaces unknown population quantities with suitable sample counterparts. Specifically, our estimates of the lower and upper bounds on $\Gamma$ are then given, respectively, by

$$\widehat{\Gamma}^L = \inf_{t\in[0,1]} \widehat{\Gamma}^L(\widehat{\eta}_1(t), \widehat{\eta}_0(t)) \quad \text{and} \quad \widehat{\Gamma}^U = \sup_{t\in[0,1]} \widehat{\Gamma}^U(\widehat{\eta}_1(t), \widehat{\eta}_0(t)),$$

where our software package uses grid search to solve these two optimization problems. Here

$$\widehat{\Gamma}^U(t_1, t_0) = \int y d\widehat{F}_1^U(y, t_1, t_0) - \int y d\widehat{F}_0^L(y, t_1, t_0).$$
$$\widehat{\Gamma}^L(t_1, t_0) = \int y d\widehat{F}_1^L(y, t_1, t_0) - \int y d\widehat{F}_0^U(y, t_1, t_0);$$

the function $\widehat{F}_d^j(y, t_1, t_0)$ is a sample analogue estimator of the function $F_{d,FRD}^j(y, t_1, t_0)$ for $d \in \{0, 1\}$; and $\widehat{\eta}_d(t)$ is a sample analogue estimator of the function $\eta_d(t)$ introduced above. The precise definition of these estimates is given in Appendix B. Following the recent RD literature, we focus on flexible nonparametric methods, and in particular local polynomial smoothing (Fan and Gijbels, 1996), for their construction.

4.2. **Inference.** In order to quantify sampling uncertainty about $\Gamma$, we construct confidence intervals that are "manipulation-robust" in the sense that they are valid irrespective of the true value of $\tau$. Such a construction involves a number of complications that we describe in this subsection. We focus again on the general case of a Fuzzy RD design in Theorem 2, as the procedure works analogously for the sharp case.

The first conceptual complication is due to the presence of an optimization operator in the definition of the bounds.[11] We address this as follows. Suppose that for every $t \in [0, 1]$ we had a $1 - \alpha$ confidence interval $\mathcal{C}_{1-\alpha}^{FRD}(t)$ for $\Gamma$ that was valid if the true value of $(\tau_1, \tau_0)$ was equal to $(\eta_1(t), \eta_0(t))$. Then the intersection-union principle (Berger, 1982) implies that $\mathcal{C}_{1-\alpha}^{FRD} = \cup_{t\in[0,1]} \mathcal{C}_{1-\alpha}^{FRD}(t)$ is a $1 - \alpha$ confidence interval for $\Gamma$. That is, a candidate value for $\Gamma$ is outside of $\mathcal{C}_{1-\alpha}^{FRD}$ if and only if it is outside of $\mathcal{C}_{1-\alpha}^{FRD}(t)$ for all $t \in [0, 1]$. Both the "fixed $t$" and the overall confidence interval have level $1 - \alpha$: there is no need for an adjustment to account for the fact that we are implicitly testing a continuum of hypotheses.

We are thus left with the problem of constructing a "fixed $t$" confidence interval, which is our second main complication. If the estimates $\widehat{\Gamma}^L(\widehat{\eta}_1(t), \widehat{\eta}_0(t))$ and $\widehat{\Gamma}^U(\widehat{\eta}_1(t), \widehat{\eta}_0(t))$ were jointly asymptotically normal irrespective of the true value of $\tau$, one could use the approach proposed by Imbens and Manski (2004) and Stoye (2009) for this purpose. However, our bound estimates are only jointly asymptotically normal (under appropriate regularity conditions) if $\tau > 0$.

[11] Our problem differs from the one in Chernozhukov, Lee, and Rosen (2013), who study inference on *intersection bounds* of the form $[\sup_v \theta(v), \inf_v \theta(v)]$. It is more accurately described as an example of *union bounds*, as the role of the inf and the sup operator in the definition of the identified set is reversed.

For $\tau = 0$, their limiting distribution is non-Gaussian, as the estimated level of manipulation $\widehat{\tau} = \max\{0, 1 - \widehat{f}^- / \widehat{f}^+\}$, formally defined in Appendix B, fails to be asymptotically normal.[12] A Gaussian approximation to the distribution of the "fixed $t$" estimates is thus typically poor in finite samples if $\tau$ is not well-separated from zero, and the standard bootstrap is unable to fix this issue (Andrews, 2000).

We therefore propose an approach similar to moment selection in the moment inequality literature (e.g. Andrews and Soares, 2010; Andrews and Barwick, 2012). Roughly speaking, we estimate the limiting distribution of the estimated bounds for a level of manipulation that is "tilted" away from zero, with the amount of tilting vanishing if $\widehat{\tau}$ is very large relative to its standard error. Since $\tau$ determines the extent to which certain distributions are trimmed, the estimated bounds $\widehat{\Gamma}^L(\widehat{\eta}_1(t), \widehat{\eta}_0(t))$ and $\widehat{\Gamma}^U(\widehat{\eta}_1(t), \widehat{\eta}_0(t))$ are stochastically increasing in $\tau$. By potentially "tilting" the value of $\tau$ away from zero, we simultaneously guarantee asymptotic normality of the bounds estimates and correct coverage of the corresponding confidence interval. For convenience, we construct such a confidence interval $\mathcal{C}_{1-\alpha}^{FRD}$ via the bootstrap; the formal algorithm is described in Appendix B.

4.3. **"Fixed $\tau$" Inference.** The confidence interval construction above takes a deliberately agnostic view about the true value of $\tau$. This view can be overly pessimistic in certain contexts. Suppose that a researcher strongly believes that manipulation is either fully absent or at least of negligible magnitude in a particular setting, and that this belief is confirmed by a point estimate of $\tau$ that is close to zero. Now, if the corresponding standard error is large, the confidence interval $\mathcal{C}_{1-\alpha}^{FRD}$ can be rather wide, as the data by themselves do not rule out a high level of manipulation. In such a scenario, the researcher could consider an alternative confidence interval for $\Gamma$ that is computed under the assumption that the value of $\tau$ is known to be some specific $\tau^* \geq 0$. Such an interval $\mathcal{C}_{1-\alpha}(\tau^*)$ can be calculated through a modified bootstrap algorithm described in Appendix B. For $\tau^* = 0$, this algorithm yields the usual "no manipulation" confidence interval, and generally $\mathcal{C}_{1-\alpha}(\tau^*)$ widens as $\tau^*$ increases.

To see how this is useful, suppose that the researcher's main goal is testing the hypothesis that $\Gamma = 0$ against the alternative that $\Gamma \neq 0$. Remember that $\Gamma$ corresponds to the usual "no manipulation" RD parameter if $\tau = 0$ (i.e., always-assigned units are absent). The researcher can plot the upper and lower boundary of $\mathcal{C}_{1-\alpha}(\tau^*)$ as a function of $\tau^*$, and check graphically for which levels of manipulation the value of 0 is contained in the confidence interval. The largest value of $\tau^*$ for which $0 \notin \mathcal{C}_{1-\alpha}(\tau^*)$ is then called the *breakdown point* of the null hypothesis that $\Gamma = 0$ (cf. Horowitz and Manski, 1995; Masten and Poirier, 2020). For

[12]Under standard regularity conditions $\sqrt{nh}(\widehat{\tau} - \tau) \xrightarrow{d} \max\{0, Z\}$ if $\tau = 0$, where $Z$ is a Gaussian random variable with mean zero.

example, suppose that $0 \notin \mathcal{C}_{1-\alpha}(0)$, but that $0 \in \mathcal{C}_{1-\alpha}(\tau^*)$ for $\tau^* \geq 0.1$. Then the researcher can report that in his preferred "no manipulation" specification the null hypothesis $\Gamma = 0$ is rejected at the critical level $\alpha$, and that at least a 10% level of manipulation around the cutoff would be needed to reverse this result. The researcher can then argue why such a high value of $\tau$ is implausible in her setting, even if it is not formally rejected by the data. We believe that such an exercise is a useful robustness check for every RD study, including those in which manipulation is generally not believed to be an issue.

## 5. Empirical Application

In this section, we apply the methods developed above to bound treatment effects of unemployment insurance (UI) on (formal) reemployment around an eligibility cutoff in Brazil.

UI programs often feature discontinuities in the level or duration of potential UI benefits based on the value of some running variable, such as age at layoff or the number of months of employment prior to layoff. RD designs are thus natural empirical strategies to estimate this effect. At the same time, the possibility of manipulation of the running variable is a common concern in the UI context (e.g., Card, Chetty, and Weber, 2007; Schmieder, von Wachter, and Bender, 2012). For instance, the net value of a match (compared to the outside option) may decrease once workers are eligible for UI, leading to more separations (see, e.g., Feldstein, 1976).[13] Our key identifying assumption ("one-sided manipulation") is likely to apply in this context, as displaced workers are likely to have a weak preference for being eligible for UI benefits (they always have the choice to not take up UI). Moreover, in most countries (the US being a notable exception), employers have no incentive to lay off their workers before they become eligible for UI as UI benefits are not experience-rated.

The setting of our application is also interesting in itself. UI programs have been adopted in a number of developing countries. Yet, the existing evidence for countries with high informality remains limited. One reason is that the concern of manipulation around discontinuities in potential UI benefits may be more severe in these countries, complicating the estimation of treatment effects. The costs of being formally laid off when eligible for UI may be relatively lower for some workers if they can work informally while drawing UI benefits.

5.1. **Institutional Details, Data, and Sample Selection.** Our empirical exercise focuses on an eligibility cutoff in the Brazilian UI program. In the interest of space, we present the institutional details and the data succinctly. For more details, see Gerard and Gonzaga

[13]The manipulation in our application may also be due to other types of behaviors that likely fit under our general model in Section 2: some workers may provoke their layoff or ask their employer to report their quit as a layoff once they are eligible for UI (Hopenhayn and Nicolini, 2009), workers laid off with a value of the running variable to the left of the cutoff may lobby their employers to lay them off on a later date, etc.

(2016), who study other aspects of the Brazilian UI program.

**Institutional Details.** In Brazil, a worker who is reported as involuntarily laid off from a private-sector formal job is eligible for UI under two conditions. First, she must have at least six months of continuous job tenure at layoff. Second, there must be at least 16 months between the date of her layoff and the date of the last layoff after which she applied for and drew UI benefits. We focus on the eligibility cutoff created by the second condition. The 16-month cutoff is more arbitrary and thus less likely to coincide with other possible discontinuities.[14] Workers who satisfy the two conditions can withdraw monthly UI payments after a 30-day waiting period and until they are formally reemployed or exhaust their potential UI duration. The potential UI duration is equal to three, four, or five months of UI benefits if workers accumulated more than 6, 12, or 24 months of formal employment in the 36 months prior to layoff, respectively. The benefit level depends on workers' average wage in the three months prior to layoff. The replacement rate is 100% at the bottom of the wage distribution but is already down to 60% for a worker who earned three times the minimum wage (see Appendix E for the full schedule). Finally, UI benefits are not experience-rated in Brazil.

**Data.** Our empirical analysis relies on two administrative datasets. The first one is a longitudinal matched employee-employer dataset covering by law the universe of formal employees. Every year, firms must report all workers formally employed at some point during the previous calendar year. The data include information on wage, tenure, age, gender, education, and sector of activity. The data also include hiring and separation dates, as well as the reason for separation. The second dataset is the registry of all UI payments. Individuals can be matched in both datasets as they are identified through the same ID number. Combining the datasets (we have both from 2002 to 2010), we can study the effect of UI on the time it takes for displaced *formal* workers to find a new *formal* job. Gerard and Gonzaga (2016) show that it is the relevant outcome to study in order to measure the efficiency cost from the usual moral hazard of UI in a context of high informality.

**Sample selection.** Our sample of analysis is constructed as follows. First, we consider all workers, between 18 and 55 years old, who lost a private-sector full-time formal job between 2004 and 2008. We start in 2004 to identify workers who were displaced from another formal job about 16 months earlier. We end in 2008 to observe two years after layoff for all workers. Second, we keep workers who had more than six months of job tenure at layoff (the other eligibility condition). Third, we restrict attention to workers for whom the difference between

[14]For instance, six months of job tenure may be a salient milestone for evaluating employees' performance. Gerard and Gonzaga (2016) show evidence of manipulation around the six-month cutoff as well. This has been confirmed recently by Carvalho, Corbi, and Narita (2017).

the layoff date and the date of their previous layoff fell within 50 days of the 16-month eligibility cutoff. Finally, we limit the sample to workers who exhausted their UI benefits after the previous layoff such that the change in eligibility at the 16-month cutoff is sharp.[15] Our sample ultimately consists of 169,575 workers with a relatively high attachment to the formal labor force, high turnover rate, and high ability to find a new formal job rapidly.[16] These are not the characteristics of the average displaced formal employee or UI taker in Brazil, but characteristics of workers for whom the 16-month cutoff may be binding.

5.2. **Graphical Evidence.** Figure 2 displays some patterns in our data. Observations are aggregated by day between the layoff date and the 16-month cutoff. Panels A and B provide some evidence of potential manipulation of the running variable. The density of the running variable and the average statutory UI replacement rate (statutory UI benefit/wage) appear to increase at the cutoff, highlighting the possibility of selection at the cutoff.[17] Panel C suggests that workers were partially aware of the eligibility rule. The share of workers applying for UI benefits jumps at the cutoff. Panel D shows that the eligibility rule was enforced. The share of workers drawing some UI benefits is close to zero to the left of the cutoff, but takeup jumps to 73% at the cutoff. Eligible workers drew on average 3.1 months of UI benefits (panel E); UI takers thus drew on average 3.1/.73 =4.25 months of UI benefits. Finally, Panel F shows that the average duration without a formal job (censored at two years) jumps from about 220 days to about 280 days at the cutoff. The average duration is high on both sides of the cutoff because the distribution of this variable has a long upper tail: about 15% of workers remain without a formal job two years after layoff (see the full distribution in Appendix E).

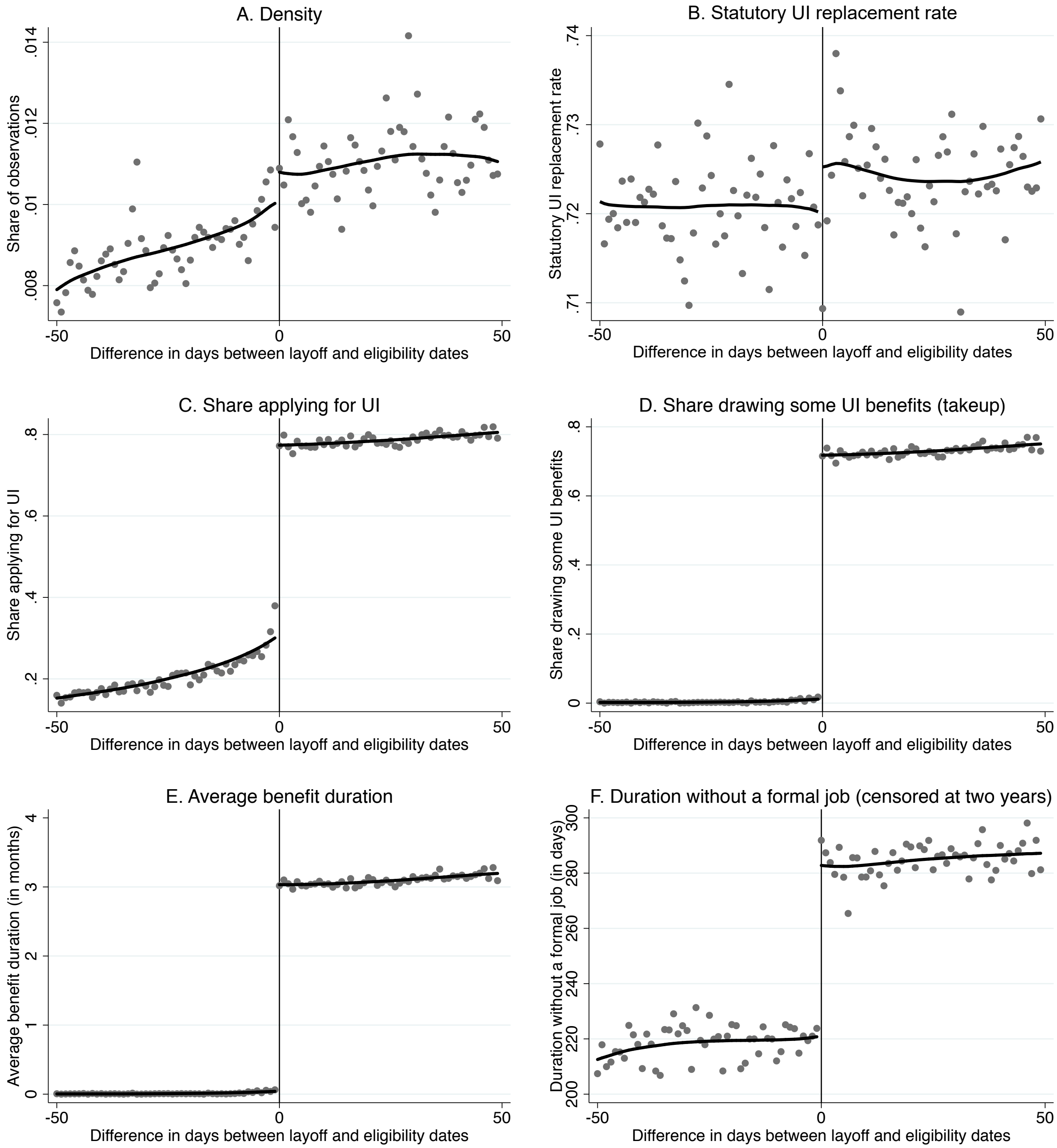
5.3. **Estimates.** The discontinuity in average duration without a formal job in Figure 2 could be due to a treatment effect, but also to a selection bias. Workers on each side of the cutoff may have different potential outcomes in the presence of manipulation. Our methods allow us to bound treatment effects, despite the possibility of selection effects. We present results from using our methods in Table 2 for an edge kernel (Cheng, Fan, and Marron, 1997)

[15] Workers who find a new formal job before exhausting their benefits are entitled to draw the remaining benefits after a new layoff, even if it occurs before the 16-month cutoff. To implement this restriction, we select workers who drew the maximum number of benefits after the previous layoff (about 40% of cases) because we measure the number of UI benefits a worker is eligible for imprecisely in the data. We also drop workers previously laid off after the 28$^{\text{th}}$ of a month. Otherwise, there is bunching in the layoff density at the 16-month cutoff even in the absence of manipulation (because February has only 28 days).

[16] They were previously eligible for five months of UI, so they accumulated 24 months of formal employment within a 36-month window. They were laid off again within 16 months and had at least six months of continuous tenure at layoff, so they found a job relatively quickly after their previous layoff (50% of workers eligible for five months of UI benefits remain without a formal job one year after layoff).

[17] The replacement rate in panel B is calculated for all workers, including those who are not actually eligible for UI, based on their wage at layoff and the UI benefit schedule.

Figure 2: Graphical evidence for our empirical application.




Notes: The figure displays the mean of different variables on each side of the cutoff by day between the layoff and eligibility dates, as well as local linear regressions on each side of the cutoff using an edge kernel and a bandwidth of 30 days. The figure is based on a RD sample of 169,575 displaced formal workers. The statutory UI replacement rate in panel B is calculated for all workers, including those who are not actually eligible for UI, based on their wage at layoff and the UI benefit schedule.

and a bandwidth of 30 days around the cutoff.[18] For bounds in the Fuzzy RD case that involve numerical optimization, we use a grid search to look for the infimum and supremum using 51 values for $t \in [0, 1]$. Confidence intervals are based on 500 bootstrap samples.[19]

Panel A reports estimates of key inputs for our bounds. First, the increase in the density documented in panel A of Figure 2 is estimated to reach 6.5% and to be statistically different from zero at conventional levels.[20] This implies that always-assigned units account for $\tau = 6.5\%$ of observations just to the right of the cutoff. The value of $\tau$ appears well-separated from zero, so the safeguards that ensure uniform validity of the confidence intervals for our bounds in case of small and imprecisely estimated values of $\tau$ are of no practical importance here. Second, UI takeup is estimated to increase by 71%-points at the cutoff.

Panels B-D then report the results from two types of exercises. First, we consider a Sharp RD design (SRD), in which *UI eligibility* is defined as the treatment. The causal effect on the outcome can be interpreted as an intention-to-treat (ITT) parameter in this case. Second, we consider the Fuzzy RD design (FRD) with *UI takeup* as the treatment. In each case, we display both Naïve RD estimates that assume no manipulation and estimates of our bounds for the treatment effects. We present results for the average effect on the duration without a formal job censored at 6 and 24 months after layoff in panels B and C, respectively. The 6-month duration proxies for the covered UI duration (up to 5 months after a 30-day waiting period); Gerard and Gonzaga (2016) show that the increase in the covered duration caused by changes in benefits is the main source of efficiency cost for UI programs. Considering both the 6-month and 24-month durations allows us to illustrate how our bounds for average treatment effects are affected by long tails in the distribution of the outcome variable. Relatedly, we present results for the estimated effects at the median using the outcome censored at 24 months after layoff in panel D, which allows us to illustrate the usefulness of looking at quantile treatment effects, as these are rather insensitive to long tails.

Naïve RD estimates that assume no manipulation yield an average increase in the duration without a formal job from UI eligibility (SRD) of 29.4 and 61.9 days for censoring points of 6 and 24 months, respectively. The corresponding figures are 41.6 and 87.7 days for the effect of UI takeup (FRD). For the duration censored at 24 months, naïve treatment effects at the median are larger, at 86 days (SRD) and 99 days (FRD). The median worker is reemployed within a year, and is thus more likely to respond to UI given the short potential duration.

[18]We do not have theoretical results on the optimal bandwidth for the estimation of our bounds. Our estimates are similar if we use bandwidths of 10 or 50 days around the cutoff (available upon request).

[19]Due to the censoring of the outcome variable, we use identification results for non-continuously distributed outcomes described in Appendix C.

[20]The increase in the average UI replacement rate in panel B of Figure 2 is also statistically different from zero at conventional levels (see Appendix E), highlighting the possibility of selection.

Table 2: Estimated treatment effects of UI on the duration without a formal job

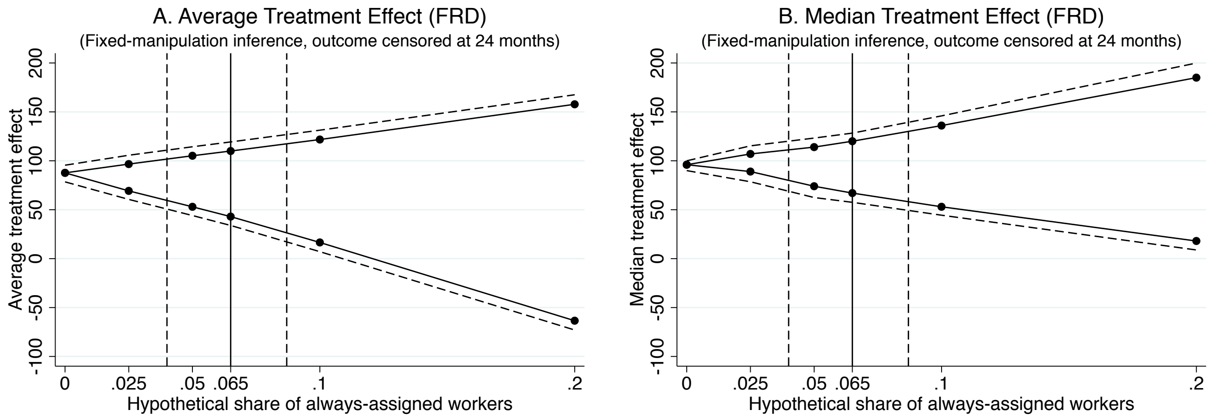
| | Estimate | 95% CI |
|---|---|---|
| **A. Basic Inputs** | | |
| Share of always-assigned workers | 0.065 | [0.040; 0.087] |
| Increase in UI takeup at the cutoff | 0.706 | [0.698; 0.713] |
| **B. Average effect: duration without a formal job censored at 6 months** | | |
| ITT/SRD: Ignoring manipulation | 29.4 | [27.6; 31.0] |
| ITT/SRD: Bounds for $\Gamma$ | [26.4; 38.8] | [24.2; 42.6] |
| LATE/FRD: Ignoring manipulation | 41.6 | [39.4; 43.7] |
| LATE/FRD: Bounds for $\Gamma$ | [35.4; 51.7] | [32.6; 55.7] |
| **C. Average effect: duration without a formal job censored at 24 months** | | |
| ITT/SRD: Ignoring manipulation | 61.9 | [55.3; 67.5] |
| ITT/SRD: Bounds for $\Gamma$ | [31.0; 81.2] | [16.8; 90.7] |
| LATE/FRD: Ignoring manipulation | 87.7 | [78.5; 95.5] |
| LATE/FRD: Bounds for $\Gamma$ | [42.9; 110.0] | [23.3; 121.7] |
| **D. Median effect: duration without a formal job censored at 24 months** | | |
| ITT/SRD: Ignoring manipulation | 86 | [80.9; 91.1] |
| ITT/SRD: Bounds for $\Gamma$ | [75; 99] | [67.7; 106.8] |
| LATE/FRD: Ignoring manipulation | 99 | [91.0; 107.0] |
| LATE/FRD: Bounds for $\Gamma$ | [67; 120] | [52.4; 131.2] |

Notes: Total number of observations within our bandwidth of 30 days around the cutoff: 102,791 displaced formal workers. Confidence intervals have nominal level of 95% and are based on 500 bootstrap samples.

A few points are useful to highlight for the behavior of our bounds in this application. First, the bounds for the average treatment effects among potentially-assigned units ($\Gamma$) are relatively tight for the duration without a formal job censored at 6 months after layoff. The lower bounds, in particular, are close to the naïve RD estimates, with point estimates of 26.4 days (SRD) and 35.4 days (FRD). Second, the bounds for the average treatment effects become wider on both sides of the naïve estimates when we consider higher censoring points. This difference comes from the fact that the distribution of the outcome becomes more dispersed and has less probability mass at the censoring point when we increase the censoring threshold. Third, bounds for quantile treatment effects, which are less sensitive to tails of the outcome distribution, can be tighter than bounds on average treatment effects in these cases. When we censor the outcome at 24 months, we obtain bounds for the average treatment effect between 31 and 81.2 days, but between 75 and 99 days for the treatment effect at the median (SRD; the bounds are also tighter in the FRD case).

Finally, we illustrate the alternative strategy for inference that we recommend when

Figure 3: Fixed-manipulation inference for our empirical application




Notes: The figure displays point estimates of our bounds and confidence intervals for the respective parameter of interest under fixed levels of the degree of manipulation. We consider LATE/FRD estimates for the average treatment effect and the quantile treatment effect at the 50th percentile for the outcome censored at 24 months. The solid vertical line (resp. dashed vertical lines) corresponds to our point estimate (resp. confidence interval) for the extent of manipulation (see Table 2).

researchers have strong beliefs that manipulation is unlikely in their setting. After all, it is not obvious from Figure 2 that there is manipulation in our data. Figure 3 displays point estimates and confidence intervals for our bounds in the Fuzzy RD case for various fixed levels of the extent of manipulation (hypothetical values of $\tau$). Panel A shows that inference on the average treatment effect can be quite sensitive to the extent of manipulation. The width of the confidence intervals doubles when we assume a small degree of manipulation ($\tau = .025$) rather than no manipulation. This illustrates the importance of taking into account the possibility of manipulation even when the McCrary (2008) test fails to reject the null hypothesis of no manipulation. The width of the confidence intervals grows quickly with larger degrees of manipulation. Panel B shows that inference on quantile treatment effects is less sensitive to the extent of manipulation. Inference may remain meaningful, even for large degrees of manipulation, illustrating the usefulness of looking at quantile treatment effects.

In sum, we find significant evidence of manipulation at the cutoff, and our bounds imply that the magnitude of naïve RD estimates may be heavily affected by selection. Nevertheless, we can still draw useful conclusions from this empirical exercise. For instance, we estimate a lower bound for the effect of UI takeup on the duration covered by UI (i.e., the outcome censored at 6 months) to be around 35.4 days. This corresponds to an increase of at least $35.4/(5 \cdot 30) = .236$ month per month of potential UI duration (given a maximum potential UI duration of 5 months). In comparison, Gerard and Gonzaga (2016) find an increase of only .126 month in the covered UI duration per additional month of potential UI duration among UI takers. Behavioral responses to UI benefits are thus relatively large in our setting, which is consistent with the composition of our sample (high attachment to the formal labor force, high turnover rate, and high ability to find a new formal job rapidly).

## 6. Conclusions

In this paper, we propose a partial identification approach to deal with the issue of potentially manipulated running variables in RD designs. We show that while the data are unable to uniquely pin down treatment effects if a running variable is subject to manipulation, they are generally still informative in the sense that they imply bounds on the value of causal parameters in both sharp and fuzzy RD designs. Our main contribution is to derive and explicitly characterize these bounds. We also propose methods to estimate our bounds in practice, and discuss how to construct confidence intervals. The approach is illustrated with an application to the Brazilian UI program. We recommend the use of our approach in applications irrespective of the outcome of McCrary's (2008) test for manipulation. Software packages that implement our methods in R and Stata are available on our websites.

## A. Proofs

A.1. **Proof of Lemma 1.** Since the density of the running variable is continuous around the cutoff among potentially-assigned units by Assumption 2, we have that $f_{X|M=0}(c^-) = f_{X|M=0}(c^+)$, and therefore $f_X(c^+) = (1 - \mathrm{P}(M = 1)) f_{X|M=0}(c^-) + \mathrm{P}(M = 1) f_{X|M=1}(c^+)$. Since there are no always-assigned units below the cutoff by Assumption 3, we have $f_{X|M=1}(x) = 0$ for $x < c$, and thus $f_X(c^-) = (1 - \mathrm{P}(M = 1)) f_{X|M=0}(c^-)$. Hence $(f_X(c^+) - f_X(c^-))/f_X(c^+) = f_{X|M=1}(c^+)\mathrm{P}(M = 1)/f_X(c^+) = \tau$, where the last equality follows from Bayes' Theorem. $\square$

A.2. **Proof of Theorem 1.** The result is a minor variation of results in Horowitz and Manski (1995) and Lee (2009). $\square$

A.3. **Proof of Theorem 2.** It follows from the arguments presented in the main body of the paper that the bounds on $\Gamma$ given knowledge $(\tau_1, \tau_0)$, formally stated in equation (3.8), are valid and sharp. That is, any value of $\Gamma$ outside of these bounds is clearly incompatible with the distribution of $(Y, D, X)$; and every value within the bounds is feasible. Moreover, it is clear that any value of $(\tau_1, \tau_0) \notin \mathcal{T}$ is incompatible with the distribution of observable quantities. It thus remains to be shown that any point $(\tau_1, \tau_0) \in \mathcal{T}$ is compatible with our model and the observed joint distribution of the data.

To show this, we proceed by constructing for any $(\tau_1, \tau_0) \in \mathcal{T}$ the distribution of a random vector $(\tilde{Y}(1), \tilde{Y}(0), \tilde{D}^+, \tilde{D}^-, \tilde{M}, \tilde{X})$ in such a way that the assumptions of our model are satisfied, and that the distribution of $(\tilde{Y}, \tilde{D}, \tilde{X})$, where $\tilde{D} = \tilde{D}^+\mathbb{I}(\tilde{X} \geq c) + \tilde{D}^-\mathbb{I}(\tilde{X} < c)$ and $\tilde{Y} = \tilde{Y}(\tilde{D})$, for $\tilde{X} \in (c - \epsilon, c + \epsilon)$ for some $\epsilon > 0$, is the same as that of $(Y, D, X)$ for $X \in (c - \epsilon, c + \epsilon)$. Note that it suffices to restrict attention to an $\epsilon$-neighborhood around the cutoff because our model has no implications for the distribution of observables outside of that range. Also note that our construction defines the notion of potential treatment states and potential outcomes for always-assigned units. This is not a concern because our model does not require such notions to be well-defined, but does not rule out that case either,

We now construct a distribution of $(\tilde{Y}(1), \tilde{Y}(0), \tilde{D}^+, \tilde{D}^-, \tilde{M}, \tilde{X})$ for $\tilde{X} \in (c - \epsilon, c + \epsilon)$. For $x \in (c - \epsilon, c + \epsilon)$, let

$$f_{\tilde{X}}(x) = f_X(x) \quad \text{and} \quad \mathrm{P}(\tilde{M} = 1 | \tilde{X} = x) = \begin{cases} 1 - f_X(c^-)/f_X(x) & \text{if } x \geq c \\ 0 & \text{if } x < c. \end{cases}$$

Moreover, let

$$\begin{aligned}
\mathrm{P}(\tilde{D}^- = 0, \tilde{D}^+ = 1 | \tilde{X} = x, \tilde{M} = 0) &= \begin{cases} \mathrm{P}(D = 1 | X = x) \cdot \frac{1-\tau_1}{1-\tau} - \mathrm{P}(D = 1 | X = c^-) \\ \qquad \text{if } x \geq c, \\ \mathrm{P}(D = 1 | X = c^+) \cdot \frac{1-\tau_1}{1-\tau} - \mathrm{P}(D = 1 | X = x) \\ \qquad \text{if } x < c, \end{cases} \\
\mathrm{P}(\tilde{D}^- = 1, \tilde{D}^+ = 1 | \tilde{X} = x, \tilde{M} = 0) &= \begin{cases} \mathrm{P}(D = 1 | X = c^-) & \text{if } x \geq c, \\ \mathrm{P}(D = 1 | X = x) & \text{if } x < c, \end{cases} \\
\mathrm{P}(\tilde{D}^- = 0, \tilde{D}^+ = 0 | \tilde{X} = x, \tilde{M} = 0) &= 1 - \mathrm{P}(\tilde{D}^- = 0, \tilde{D}^+ = 1 | \tilde{X} = x, \tilde{M} = 0) \\
&\quad - \mathrm{P}(\tilde{D}^- = 1, \tilde{D}^+ = 1 | \tilde{X} = x, \tilde{M} = 0), \\
\mathrm{P}(\tilde{D}^- = 1, \tilde{D}^+ = 0 | \tilde{X} = x, \tilde{M} = 0) &= 0,
\end{aligned}$$

and

$$
\begin{aligned}
\mathrm{P}(\tilde{D}^- = 0, \tilde{D}^+ = 1 | \tilde{X} = x, \tilde{M} = 1) &= \begin{cases} \mathrm{P}(D = 1 | X = x) \cdot \frac{\tau_1}{\tau} - h(x) & \text{if } x \geq c, \\ \mathrm{P}(D = 1 | X = c^+) \cdot \frac{\tau_1}{\tau} - h(c^+) & \text{if } x < c, \end{cases} \\
\mathrm{P}(\tilde{D}^- = 1, \tilde{D}^+ = 1 | \tilde{X} = x, \tilde{M} = 1) &= \begin{cases} h(x) & \text{if } x \geq c, \\ h(c^+) & \text{if } x < c, \end{cases} \\
\mathrm{P}(\tilde{D}^- = 0, \tilde{D}^+ = 0 | \tilde{X} = x, \tilde{M} = 1) &= 1 - \mathrm{P}(\tilde{D}^- = 0, \tilde{D}^+ = 1 | \tilde{X} = x, \tilde{M} = 1), \\
&\quad - \mathrm{P}(\tilde{D}^- = 1, \tilde{D}^+ = 1 | \tilde{X} = x, \tilde{M} = 1), \\
\mathrm{P}(\tilde{D}^- = 1, \tilde{D}^+ = 0 | \tilde{X} = x, \tilde{M} = 1) &= 0,
\end{aligned}
$$

where $h(\cdot)$ is an arbitrary continuous function satisfying that $0 \leq h(x) \leq \mathrm{P}(D = 1 | X = x) \cdot \tau_1/\tau$. With these choices, the implied distribution of $(\tilde{D}, \tilde{X}) | \tilde{X} \in (c - \epsilon, c + \epsilon)$ is the same as that of $(D, X) | X \in (c - \epsilon, c + \epsilon)$ for every $(\tau_1, \tau_0) \in \mathcal{T}$. It thus remains to be shown that one can construct a distribution of $(\tilde{Y}(1), \tilde{Y}(0))$ given $(\tilde{D}^+, \tilde{D}^-, \tilde{X}, \tilde{M})$ that is compatible with our assumptions, and such that the distribution of $\tilde{Y}$ given $(\tilde{D}, \tilde{X})$ for $\tilde{X} \in (c - \epsilon, c + \epsilon)$ is the same as the distribution of $Y$ given $(D, X)$ for $X \in (c - \epsilon, c + \epsilon)$ for every $(\tau_1, \tau_0) \in \mathcal{T}$. But this is possible by setting $(\tilde{Y}(1), \tilde{Y}(0))$ as independent of $(\tilde{D}^+, \tilde{D}^-, \tilde{X}, \tilde{M})$, and then assigning one of the respective extreme distributions derived in the main body of the text to the respective marginals. This completes our proof. $\square$

## B. Additional Notation for Estimation and Inference

In this section, we give further details on the construction of the estimators and confidence intervals described in the main body of the paper. To simplify the exposition, we use the same polynomial order $p$, bandwidth $h$ and kernel function $K(\cdot)$ in all intermediate estimation steps in this paper. We also use the notation that $\boldsymbol{\pi}_p(x) = (1/0!, x/1!, x^2/2!, \ldots, x^p/p!)'$ and $K_h(x) = K(x/h)/h$ for any $x \in \mathbb{R}$, and define the $(p+1)$-vector $\mathbf{e}_1 = (1, 0, \ldots, 0)'$. The data are an independent sample $\{(Y_i, D_i, X_i), i = 1, \ldots, n\}$ of size $n$.

Following the result in Lemma 1, estimating $\tau$ requires estimates of the right and left limits of the density at the cutoff. There are a number of nonparametric estimators that can be used to estimate densities at boundary points; see for example Lejeune and Sarda (1992), Jones (1993), Cheng (1997) or Cattaneo, Jansson, and Ma (2019). Here we use a minor variation of the procedure in Cheng (1997), which also forms the basis for the McCrary

(2008) test, and estimate $f_X(c^+)$ and $f_X(c^-)$ by

$$\widehat{f}^+ = \mathbf{e}_1' \operatorname*{argmin}_{\beta\in\mathbb{R}^{p+1}} \sum_{i=1}^n (\widehat{f}(X_i) - \boldsymbol{\pi}_p(X_i - c)'\beta)^2 K_h(X_i - c)\mathbb{I}\{X_i \geq c\}, \text{ and}$$
$$\widehat{f}^- = \mathbf{e}_1' \operatorname*{argmin}_{\beta\in\mathbb{R}^{p+1}} \sum_{i=1}^n (\widehat{f}(X_i) - \boldsymbol{\pi}_p(X_i - c)'\beta)^2 K_h(X_i - c)\mathbb{I}\{X_i < c\},$$

respectively, where $\widehat{f}(X_i) = (1/n)\sum_{j=1}^n K_h(X_j - X_i)$. Since by assumption the proportion of always-assigned units among units just to the right of the cutoff has to be non-negative, our estimate of $\tau$ is then given by

$$\widehat{\tau} = \max\{\widetilde{\tau}, 0\}, \quad \text{with} \quad \widetilde{\tau} = 1 - \widehat{f}^-/\widehat{f}^+.$$

Local polynomial regression estimates of $g^+ = \mathrm{E}(D_i|X_i = c^+)$ and $g^- = \mathrm{E}(D_i|X_i = c^-)$, the conditional treatment probabilities on either side of the cutoff, are given by

$$\widehat{g}^+ = \mathbf{e}_1' \operatorname*{argmin}_{\beta\in\mathbb{R}^{p+1}} \sum_{i=1}^n (D_i - \boldsymbol{\pi}_p(X_i - c)'\beta)^2 K_h(X_i - c)\mathbb{I}\{X_i \geq c\}, \text{ and}$$
$$\widehat{g}^- = \mathbf{e}_1' \operatorname*{argmin}_{\beta\in\mathbb{R}^{p+1}} \sum_{i=1}^n (D_i - \boldsymbol{\pi}_p(X_i - c)'\beta)^2 K_h(X_i - c)\mathbb{I}\{X_i < c\},$$

respectively (Fan and Gijbels, 1996). The conditional c.d.f.s $F_{Y|X=c^+,D=d}(y)$ and $F_{Y|X=c^-,D=d}(y)$ are estimated by

$$\widehat{F}_{Y|X=c^+,D=d}(y) = \mathbf{e}_1' \operatorname*{argmin}_{\beta\in\mathbb{R}^{p+1}} \sum_{i=1}^n (\mathbb{I}\{Y_i \leq y\} - \boldsymbol{\pi}_p(X_i - c)'\beta)^2 K_h(X_i - c)\mathbb{I}\{X_i \geq c\}, \text{ and}$$
$$\widehat{F}_{Y|X=c^-,D=d}(y) = \mathbf{e}_1' \operatorname*{argmin}_{\beta\in\mathbb{R}^{p+1}} \sum_{i=1}^n (\mathbb{I}\{Y_i \leq y\} - \boldsymbol{\pi}_p(X_i - c)'\beta)^2 K_h(X_i - c)\mathbb{I}\{X_i < c\},$$

respectively, which for every $y \in \mathbb{R}$ corresponds to a local polynomial regression with $\mathbb{I}\{Y_i \leq y\}$ as the dependent variable (Hall, Wolff, and Yao, 1999). We then estimate the conditional p.d.f.s $f_{Y|X=c^+,D=d}(y)$ and $f_{Y|X=c^-,D=d}(y)$ by

$$\widehat{f}_{Y|X=c^+,D=d}(y) = \mathbf{e}_1' \operatorname*{argmin}_{\beta\in\mathbb{R}^{p+1}} \sum_{i=1}^n (K_h(Y_i - y) - \boldsymbol{\pi}_p(X_i - c)'\beta)^2 K_h(X_i - c)\mathbb{I}\{X_i \geq c\}, \text{ and}$$
$$\widehat{f}_{Y|X=c^-,D=d}(y) = \mathbf{e}_1' \operatorname*{argmin}_{\beta\in\mathbb{R}^{p+1}} \sum_{i=1}^n (K_h(Y_i - y) - \boldsymbol{\pi}_p(X_i - c)'\beta)^2 K_h(X_i - c)\mathbb{I}\{X_i < c\}$$

respectively, which for every $y \in \mathbb{R}$ corresponds to a local polynomial regression with $K_h(Y_i - y)$ as the dependent variable (Fan, Yao, and Tong, 1996).

Next, we put

$$\widehat{F}_1^U(y,t_1,t_0) = \frac{(1-\widehat{\kappa}_1)\widehat{G}(y) - t_1}{1-\widehat{\kappa}_1 - t_1} \cdot \mathbb{I}\left\{y \geq \widehat{G}^{-1}\left(\frac{t_1}{1-\widehat{\kappa}_1}\right)\right\},$$
$$\widehat{F}_0^U(y,t_1,t_0) = \frac{\widehat{F}_{Y|X=c^-,D=0}(y) - \widehat{\kappa}_0 \cdot (1-t_0)\widehat{F}^L_{Y(0)|X=c,N_0}(y,t_0)}{1-\widehat{\kappa}_0 \cdot (1-t_0)},$$

and define the functions $\widehat{F}_1^L$ and $\widehat{F}_0^L$ analogously. We use the notation that

$$\widehat{G}(y) = \frac{\widehat{F}_{Y|X=c^+,D=1}(y) - \widehat{\kappa}_1 \widehat{F}_{Y|X=c^-,D=1}(y)}{1-\widehat{\kappa}_1},$$
$$\widehat{F}^L_{Y(0)|X=c,N_0}(y,t_0) = \int_{-\infty}^{y} \widehat{s}(u,t_0)\mathbb{I}\left\{u \geq \widehat{q}_L(t_0)\right\} du,$$
$$\widehat{s}(y,t_0) = \frac{\min\left\{\widehat{f}_{Y|X=c^-,D=0}(y)/\widehat{\kappa}_0, \widehat{f}_{Y|X=c^+,D=0}(y)\right\}}{1-t_0},$$
$$\widehat{\kappa}_1 = \frac{(1-\widehat{\tau})\widehat{g}^-}{\widehat{g}^+}, \quad \widehat{\kappa}_0 = \frac{1-\widehat{g}^+}{(1-\widehat{\tau})(1-\widehat{g}^-)};$$

with $\widehat{q}_L(t_0)$ the value that satisfies $\int_{-\infty}^{\widehat{q}_L(t_0)} \widehat{s}(y,t_0)dy = 1$. Finally, we define the functions

$$\widehat{\eta}_d(t) = \widehat{\tau}_d^L + t \cdot (\widehat{\tau}_d^U - \widehat{\tau}_d^L), \quad d \in \{0,1\},$$

where for $j \in \{U, L\}$ and $d \in \{0,1\}$ the term $\widehat{\tau}_d^j$ is the obvious sample analogue estimator of the point $\tau_d^j$ introduced above.

We now introduce notation and details regarding the construction of confidence intervals. The confidence interval $\mathcal{C}_{1-\alpha}^{FRD}$ is constructed using a bootstrap distribution under which the bootstrap analogue of $\widetilde{\tau} = 1 - \widehat{f}^-/\widehat{f}^+$ is centered around $\max\{\widehat{\tau}, \kappa_n \widehat{\sigma}_{\widetilde{\tau}}\}$, where $\widehat{\sigma}_{\widetilde{\tau}}$ is the standard error of $\widetilde{\tau}$, and $\kappa_n$ is a sequence of constants that slowly tends to infinity. Following much of the moment inequality literature, we choose $\kappa_n = \log(n)^{1/2}$. The algorithm for our bootstrap is as follows.

1. Generate bootstrap samples $\{Y_{i,b}, D_{i,b}, X_{i,b}\}_{i=1}^n$, $b = 1, \ldots, B$ by sampling with replacement from the original data $\{Y_i, D_i, X_i\}_{i=1}^n$; for some large integer $B$.

2. Calculate $\widetilde{\tau}_b^* = 1 - \widehat{f}_b^-/\widehat{f}_b^+$, and put $\widehat{\sigma}_{\widetilde{\tau}}$ as the sample standard deviation of $\{\widetilde{\tau}_b^*\}_{b=1}^B$.

3. Calculate $\widetilde{\tau}_b = \widetilde{\tau}_b^* - \widetilde{\tau} + \max\{\widehat{\tau}, \kappa_n \widehat{\sigma}_{\widetilde{\tau}}\}$ and $\widehat{\tau}_b = \max\{\widetilde{\tau}_b, 0\}$.

4. For $j \in \{U, L\}$, calculate $\widehat{\Gamma}^j(\widehat{\eta}_1(t), \widehat{\eta}_0(t))$ using the redefined estimate $\widehat{\tau}_b$ from the previous step, and put $\widehat{\sigma}^j(t)$ as the sample standard deviation of $\{\widehat{\Gamma}^j(\widehat{\eta}_1(t), \widehat{\eta}_0(t)\}_{b=1}^B$.

Now define $\widehat{\Gamma}^{L*}(t)$ and $\widehat{\Gamma}^{U*}(t)$ exactly as $\widehat{\Gamma}^L(\widehat{\eta}_1(t), \widehat{\eta}_0(t))$ and $\widehat{\Gamma}^U(\widehat{\eta}_1(t), \widehat{\eta}_0(t))$, with the exception that $\widehat{\tau}^* = \max\{\widetilde{\tau}, \kappa_n \widehat{\sigma}_{\widetilde{\tau}}\}$ is used instead of $\widehat{\tau}$. Following Imbens and Manski (2004) and

Stoye (2009), our "fixed $t$" confidence interval for $\Gamma$ with level $1-\alpha$ is then given by

$$\mathcal{C}_{1-\alpha}^{FRD}(t) = \left[\widehat{\Gamma}^{L*}(t) - r_\alpha(t)\cdot\widehat{\sigma}^L(t),\ \widehat{\Gamma}^{U*}(t) + r_\alpha(t)\cdot\widehat{\sigma}^U(t)\right],$$

where $r_\alpha(t)$ is the value that solves the equation

$$\Phi\left(r_\alpha(t) + \frac{\widehat{\Gamma}^{U*}(t) - \widehat{\Gamma}^{L*}(t)}{\max\{\widehat{\sigma}^L(t), \widehat{\sigma}^U(t)\}}\right) - \Phi(-r_\alpha(t)) = 1-\alpha,$$

and $\Phi(\cdot)$ is the CDF of the standard normal distribution. The final intersection-union confidence interval for $\Gamma$ is then given by

$$\mathcal{C}_{1-\alpha}^{FRD} = \left[\inf_{t\in[0,1]}\left(\widehat{\Gamma}^L(t) - r_\alpha(t)\cdot\widehat{\sigma}^L(t)\right),\ \sup_{t\in[0,1]}\left(\widehat{\Gamma}^U(t) + r_\alpha(t)\cdot\widehat{\sigma}^U(t)\right)\right].$$

Note that this construction does not account for discontinuities in the limiting distribution of the "fixed $t$" estimates at those values of $\tau$ under which one of the various max and min operators in the definition of the function $\eta_d(\cdot)$ becomes binding. We expect this to have only minor importance in practice, and therefore do not include any "safeguards" against such cases into our procedure. Our construction also implicitly assumes that the two functions involved in the definition of the term $s(y,\tau_0)$ cross at a finite number of points. If that was not the case the presence of the max operator would generate a bias, which could be removed using techniques analogous to those in Anderson, Linton, and Whang (2012).

The confidence interval $\mathcal{C}_{1-\alpha}(\tau^*)$ can be calculated through the following modified bootstrap algorithm:

1. For $\tau^* \in [0,1]$ and $t \in [0,1]$, define $\widehat{\Gamma}^L(\tau^*, t)$ and $\widehat{\Gamma}^U(\tau^*, t)$ exactly as $\widehat{\Gamma}^L(\widehat{\eta}_1(t), \widehat{\eta}_0(t))$ and $\widehat{\Gamma}^U(\widehat{\eta}_1(t), \widehat{\eta}_0(t))$, with the exception that $\tau^*$ is used instead of $\widehat{\tau}$.

2. Generate bootstrap samples $\{Y_{i,b}, D_{i,b}, X_{i,b}\}_{i=1}^n$, $b = 1,\ldots,B$ by sampling with replacement from the original data $\{Y_i, D_i, X_i\}_{i=1}^n$; for some large integer $B$.

3. For $j \in \{U, L\}$, calculate $\widehat{\Gamma}_b^j(\tau^*, t)$, and put $\widehat{\sigma}^j(\tau^*, t)$ as the sample standard deviation of $\{\widehat{\Gamma}_b^j(\tau^*, t)\}_{b=1}^B$.

4. Compute the $1-\alpha$ confidence interval $\mathcal{C}_{1-\alpha}^{FRD}(\tau^*)$ as

$$\left[\inf_{t\in[0,1]}\left(\widehat{\Gamma}^L(\tau^*, t) - r_\alpha(\tau^*, t)\cdot\widehat{\sigma}^L(\tau^*, t)\right),\ \sup_{t\in[0,1]}\left(\widehat{\Gamma}^U(\tau^*, t) + r_\alpha(\tau^*, t)\cdot\widehat{\sigma}^U(\tau^*, t)\right)\right],$$

where $r_\alpha(\tau^*, t)$ is the value that solves the equation

$$\Phi\left(r_\alpha(\tau^*, t) + \frac{\widehat{\Gamma}^U(\tau^*, t) - \widehat{\Gamma}^L(\tau^*, t)}{\max\{\widehat{\sigma}^L(\tau^*, t), \widehat{\sigma}^U(\tau^*, t)\}}\right) - \Phi(-r_\alpha(\tau^*, t)) = 1 - \alpha.$$

For $\tau^* = 0$ this algorithm yields the usual "no manipulation" confidence interval, and generally $\mathcal{C}_{1-\alpha}(\tau^*)$ becomes wider as $\tau^*$ increases.

## C. Identification: Further Results and Extensions

The results in Section 3 in the main body of the paper can be extended in various ways. In this section, we show how our results can be extended to quantile treatment effects, how our bounds on $\Gamma$ change if we allow for non-continuously distributed outcomes, that additional behavioral assumptions can lead to narrower bounds on $\Gamma$, that covariates can be used to tighten the bounds as well, and that the distribution of covariates among always-assigned and potentially-assigned units is point identified in our model.

C.1. **Quantile Treatment Effects.** It is straightforward to generalize our identification analysis for average treatment effects to their quantile counterparts. For example, instead of the parameter $\Gamma$ that we focus on in Section 3, one could consider the quantile treatment effect among compliers at the cutoff, formally defined as

$$\Psi(u) \equiv Q_{Y(1)|X=c^-,D^+>D^-}(u) - Q_{Y(0)|X=c^-,D^+>D^-}(u).$$

This extension is straightforward because our general strategy in Section 3 is to first obtain sharp lower and upper bounds, in a first-order stochastic dominance sense, on the c.d.f.s $F_{Y(d)|X=c,C_0}$ for $d \in \{0, 1\}$. Once these have been obtained, it follows from Stoye (2010, Lemma 1) that sharp upper and lower bounds on on any functional of the form $\theta(F_{Y(d)|X=c,C_0})$ are given, respectively, by $\theta(F_d^U(y))$ and $\theta(F_d^L(y))$ as long as $\theta(\cdot)$ increases with first-order stochastic dominance. Quantiles are easily seen to fall into this class. To define the resulting bounds on $\Psi(u)$ in the general context of Theorem 2, for example, let

$$\begin{aligned}
\Psi^U_{FRD}(u, t_1, t_0) &= Q^U_{1,FRD}(u, \tau_1, \tau_0) - Q^L_{0,FRD}(u, \tau_1, \tau_0) \quad \text{and} \\
\Psi^L_{FRD}(u, t_1, t_0) &= Q^L_{1,FRD}(u, \tau_1, \tau_0) - Q^U_{0,FRD}(u, \tau_1, \tau_0),
\end{aligned}$$

where $Q^U_{1,FRD}(\cdot, \tau_1, \tau_0)$ is the inverse of $F^U_{1,FRD}(\cdot, \tau_1, \tau_0)$, and the other terms in the previous equation are defined similarly. Then sharp lower and upper bounds on $\Psi(u)$ are given by

$$\Psi^L_{FRD}(u) = \inf_{(t_1,t_0)\in\mathcal{T}} \Psi^L_{FRD}(u, t_1, t_0) \quad \text{and} \quad \Psi^U_{FRD}(u) = \sup_{(t_1,t_0)\in\mathcal{T}} \Psi^U_{FRD}(u, t_1, t_0),$$

respectively. Bounds on quantile treatment effects in sharp designs, or under the various refinements studied in the remainder of this section, can be obtained analogously.

C.2. **Non-Continuously Distributed Outcomes.** Theorem 1 and 2 are stated for the case in which the outcome variable is continuously distributed. This is for notational convenience only, and our results immediately generalize to the case of a discrete outcome variable, which occurs frequently in empirical applications. Suppose that $\text{supp}(Y)$ is a finite set. Then in the case of a Sharp RD design our sharp upper and lower bounds on $F_{Y(1)|X=c,C_0}$ are

$$\begin{aligned} F^U_{1,SRD}(y) &= (1-\theta^U)F_{Y|X=c^+,Y>Q_{Y|X=c^+}(\tau)}(y) + \theta^U \mathbb{I}\left\{y \geq Q_{Y|X=c^+}(\tau)\right\} \quad \text{and} \\ F^L_{1,SRD}(y) &= (1-\theta^L)F_{Y|X=c^+,Y<Q_{Y|X=c^+}(1-\tau)}(y) + \theta^L \mathbb{I}\left\{y \geq Q_{Y|X=c^+}(1-\tau)\right\}, \end{aligned}$$

where

$$\theta^L = \frac{\mathrm{P}(Y \geq Q_{Y|X=c^+}(1-\tau)|X=c^+)-\tau}{1-\tau} \quad \text{and} \quad \theta^U = \frac{\mathrm{P}(Y \leq Q_{Y|X=c^+}(\tau)|X=c^+)-\tau}{1-\tau}.$$

The following Corollary uses these bounds to obtain explicit sharp bounds on the local average treatment effect $\Gamma$.

**Corollary 1.** *Suppose that the assumptions of Theorem 1 hold, and that $\text{supp}(Y)$ is a finite set. Then sharp lower and upper bounds on $\Gamma$ are given by*

$$\begin{aligned} \Gamma^L_{SRD} &= (1-\theta^L)\mathrm{E}(Y|X=c^+, Y<Q_{Y|X}(1-\tau|c^+)) + \theta^L Q_{Y|X}(1-\tau|c^+) \\ &\quad - \mathrm{E}(Y|X=c^-) \quad \text{and} \\ \Gamma^U_{SRD} &= (1-\theta^U)\mathrm{E}(Y|X=c^+, Y>Q_{Y|X}(\tau|c^+)) + \theta^U Q_{Y|X}(\tau|c^+) \\ &\quad - \mathrm{E}(Y|X=c^-), \end{aligned}$$

*respectively.*

In a Fuzzy RD design, we modify the expressions for the sharp upper and lower bounds on $F_{Y(1)|X=c,C_0}$ and $F_{Y(0)|X=c,N_0}$ for known values of $\tau_1$ and $\tau_0$ as follows:

$$\begin{aligned} F^U_{1,FRD}(y,\tau_1,\tau_0) &= (1-\theta^U_1)G_{Y|Y>Q_G\left(\frac{\tau_1}{1-\kappa_1}\right)}(y) + \theta^U_1 \mathbb{I}\left\{y \geq Q_G\left(\frac{\tau_1}{1-\kappa_1}\right)\right\} \quad \text{and} \\ F^L_{1,FRD}(y,\tau_1,\tau_0) &= (1-\theta^L_1)G_{Y|Y<Q_G\left(1-\frac{\tau_1}{1-\kappa_1}\right)}(y) + \theta^L_1 \mathbb{I}\left\{y \geq Q_G\left(1-\frac{\tau_1}{1-\kappa_1}\right)\right\}, \end{aligned}$$

where

$$\theta^U_1 = \frac{\mathrm{P}_G\left(Y \leq Q_G\left(\frac{\tau_1}{1-\kappa_1}\right)\right) - \frac{\tau_1}{1-\kappa_1}}{1-\frac{\tau_1}{1-\kappa_1}} \quad \theta^L_1 = \frac{\mathrm{P}_G\left(Y \geq Q_G\left(1-\frac{\tau_1}{1-\kappa_1}\right)\right) - \frac{\tau_1}{1-\kappa_1}}{1-\frac{\tau_1}{1-\kappa_1}}.$$

The modified expressions for bounds on $F_{Y(0)|X=c,N_0}$ are given by

$$F^U_{Y(0)|X=c,N_0}(y) = \int_{-\infty}^{y} s(t,\tau_0)\mathbb{I}\left\{t \le q_U(\tau_0)\right\} dt + \theta_0^U \mathbb{I}\left\{y > q_U(\tau_0)\right\} \quad \text{and}$$
$$F^L_{Y(0)|X=c,N_0}(y) = \int_{-\infty}^{y} s(t,\tau_0)\mathbb{I}\left\{t \ge q_L(\tau_0)\right\} dt + \theta_0^L \mathbb{I}\left\{y > q_L(\tau_0)\right\},$$

where

$$\begin{aligned}
\theta_0^U &= 1 - \int_{-\infty}^{q_U(\tau_0)} s(t,\tau_0)\mathbb{I}\left\{t \le q_U(\tau_0)\right\} dt,\\
\theta_0^L &= 1 - \int_{q_L(\tau_0)}^{\infty} s(t,\tau_0)\mathbb{I}\left\{t \ge q_L(\tau_0)\right\} dt,\\
q_L(\tau_0) &= \inf\{y \in \mathrm{supp}(Y) : \int_y^{\infty} s(t,\tau_0)dt \le 1\}, \text{ and}\\
q_U(\tau_0) &= \sup\{y \in \mathrm{supp}(Y) : \int_{-\infty}^{y} s(t,\tau_0)dt \le 1\}.
\end{aligned}$$

We then obtain the following expressions for sharp bounds on the local average treatment effect $\Gamma$ given knowledge of $\tau_1$ and $\tau_0$:

$$\begin{aligned}
\Gamma^U_{FRD}(\tau_1,\tau_0) &\equiv \int y dF^U_{1,FRD}(y,\tau_1,\tau_0) - \int y dF^L_{0,FRD}(y,\tau_1,\tau_0),\\
\Gamma^U_{FRD}(\tau_1,\tau_0) &\equiv \int y dF^L_{1,FRD}(y,\tau_1,\tau_0) - \int y dF^U_{0,FRD}(y,\tau_1,\tau_0).
\end{aligned}$$

The following Corollary finally states the sharp bounds on $\Gamma$ given that the values of $\tau_1$ and $\tau_0$ are only partially identified.

**Corollary 2.** *Suppose that the assumptions of Theorem 2 hold, and that supp$(Y)$ is a finite set. Then sharp lower and upper bounds on $\Gamma$ are given by*

$$\Gamma^L_{FRD} = \inf_{(t_1,t_0)\in\mathcal{T}} \Gamma^L_{FRD}(t_1,t_0) \quad \text{and} \quad \Gamma^U_{FRD} = \sup_{(t_1,t_0)\in\mathcal{T}} \Gamma^U_{FRD}(t_1,t_0),$$

*respectively.*

C.3. **Adding Behavioral Assumptions in Fuzzy RD Designs.** The bounds in Theorem 2 can be narrowed by imposing stronger assumptions on the units' behavior, which relate to behavioral restrictions that arise naturally in certain empirical contexts. Consider for instance settings where always-assigned units obtain values of the running variable to the right of the cutoff by taking conscious actions. Since such units actively choose to be eligible for the treatment, it seems plausible to assume that their probability of actually receiving the treatment conditional on being eligible is relatively high in some appropriate sense.

First, one might be willing to assume that always-assigned units are at least as likely to get treated as eligible potentially-assigned units, implying the following corollary:

**Corollary 3.** *Suppose that the conditions of Theorem 2 hold, and that* $\mathrm{E}(D|X=c^+,M=1)\geq \mathrm{E}(D|X=c^+,M=0)$*. Then sharp lower and upper bounds on* $\Gamma$ *are given by*

$$\Gamma^L_{FRD(a)}=\inf_{(t_1,t_0)\in\mathcal{T}_a}\Gamma^L_{FRD}(t_1,t_0)\quad\text{and}\quad\Gamma^U_{FRD(a)}=\sup_{(t_1,t_0)\in\mathcal{T}_a}\Gamma^U_{FRD}(t_1,t_0),$$

*respectively, where* $\mathcal{T}_a\equiv\{(t_1,t_0):(t_1,t_0)\in\mathcal{T}\text{ and }t_1\geq\tau\}$*.*

We see that the additional restriction of Corollary 3 relative to Theorem 2 increases the lowest possible value of $\tau_1$ from $\max\{0,1+(\tau-1)/\mathrm{E}(D|X=c^+)\}$ to $\tau$, and correspondingly decreases the largest possible value for $\tau_0$ from $\min\{1,\tau/(1-\mathrm{E}(D|X=c^+))\}$ to $\tau$. This follows from a simple application of Bayes' Rule, and means that $\mathcal{T}_a\subset\mathcal{T}$. We then obtain bounds on $\Gamma$ that are (weakly) narrower, as optimization is carried out over a smaller set.

Second, in some cases, it may be reasonable to drive this line of reasoning further and assume that always-assigned units *always* receive the treatment. This implies the following corollary:

**Corollary 4.** *Suppose that the conditions of Theorem 2 hold, and that* $\mathrm{E}(D|X=c^+,M=1)=1$*. Then* $\tau_1=\tau/\mathrm{E}(D|X=c^+)$ *and* $\tau_0=0$ *are point identified; and sharp lower and upper bounds on* $\Gamma$ *are given by*

$$\Gamma^L_{FRD(b)}=\Gamma^L_{FRD}\left(\frac{\tau}{\mathrm{E}(D|X=c^+)},0\right)\quad\text{and}\quad\Gamma^U_{FRD(b)}=\Gamma^U_{FRD}\left(\frac{\tau}{\mathrm{E}(D|X=c^+)},0\right),$$

*respectively.*

Under the conditions of Corollary 4, the set of feasible values of $(\tau_1,\tau_0)$ shrinks to a singleton, which means that sharp bounds on our parameter of interest can be defined without invoking an optimization operator. Moreover, we can see from Table 1 that due to the absence of always-assigned untreated units the distributions $F_{Y(0)|X=c,N_0}$ and $F_{Y(0)|X=c,C_0}$ are point identified in this case.

C.4. **Using Covariates to Tighten the Bounds.** Following arguments similar to those in Lee (2009), covariates that are measured prior to treatment assignment can also be used to narrow the bounds on $\Gamma$ that we derived above. Let $W$ be a vector of such covariates, and denote its support by $\mathcal{W}$. The idea is that, if the outcome distribution or the proportion of always-assigned units varies with $W$, trimming units based on their position in the outcome distribution conditional on $W$ leads to units with less extreme values in the overall outcome distribution being trimmed, which narrows the bounds.

For the sharp RD design, the sharp upper and lower bounds on $F_{Y(1)|X=c,C_0}$ become:

$$F^U_{1,SRD(W)}(y) = \int F_{Y|X=c^+,W=w,Y\geq Q_{Y|X=c^+,W=w}(\tau(w))}(y)dF_{W|X=c^-}(w) \quad \text{and}$$
$$F^L_{1,SRD(W)}(y) = \int F_{Y|X=c^+,W=w,Y\leq Q_{Y|X=c^+,W=w}(1-\tau(w))}(y)dF_{W|X=c^-}(w),$$

where $\tau(w) = \mathrm{P}(M = 1|X = c^+, W = w)$ is a conditional version of $\tau$ defined as in (3.1), which is point identified as $\tau(w) = 1 - f_{X|W}(c^-, w)/f_{X|W}(c^+, w)$ through arguments analogous to those used in the proof of Lemma 1, conditioning on $W = w$ throughout. The next corollary gives the resulting sharp lower and upper bounds on $\Gamma$.

**Corollary 5.** *Suppose that the assumptions of Theorem 1 hold, mutatis mutandis, with conditioning on the covariates $W$. Then sharp lower and upper bounds on $\Gamma$ are given by*

$$\begin{aligned}\Gamma^L_{SRD(W)} = &\int \mathrm{E}(Y|X = c^+, W = w, Y \leq Q_{Y|X=c^+,W=w}(1 - \tau(w)))dF_{W|X=c^-}(w)\\ &- \mathrm{E}(Y|X = c^-) \quad \text{and}\\ \Gamma^U_{SRD(W)} = &\int \mathrm{E}(Y|X = c^+, W = w, Y_i \geq Q_{Y|X=c^+,W=w}(\tau(w)))dF_{W|X=c^-}(w)\\ &- \mathrm{E}(Y|X = c^-),\end{aligned}$$

*respectively.*

To state a similar result for the fuzzy RD design, we need to define conditional versions of $\tau_1$, $\tau_0$, $\mathcal{T}$, $\kappa_1$ and $\kappa_0$ in the same fashion. We denote the resulting quantities by $\tau_1(w)$, $\tau_0(w)$, $\mathcal{T}(w)$, $\kappa_1(w)$ and $\kappa_0(w)$, respectively. We then define conditional versions of $F^U_{d,FRD}(y, \tau_1, \tau_0)$ and $F^L_{d,FRD}(y, \tau_1, \tau_0)$, denoted by $F^U_{d,FRD|W=w}(y, \tau_1(w), \tau_0(w))$ and $F^L_{d,FRD|W=w}(y, \tau_1(w), \tau_0(w))$, respectively, for $d \in \{0, 1\}$. These objects are constructed following the steps in the previous section by conditioning on $W = w$ throughout. We also define the set $\mathcal{T}_\mathcal{W} = \{(t_1(\cdot), t_1(\cdot)) : (t_1(w), t_1(w)) \in \mathcal{T}(w) \text{ for all } w \in \mathcal{W}\}$. Finally, we denote the proportion of potentially-assigned compliers ($C_0$) conditional on $W = w$ just to the left of the cutoff by

$$\begin{aligned}\mathrm{P}(C_0|X = c^-, W = w) &= \frac{1 - \tau_1(w)}{1 - \tau(w)}\mathrm{E}(D|X = c^+, W = w) - \mathrm{E}(D|X = c^-, W = w)\\ &\equiv \Pi_{W=w}(\tau_1(w), \tau_0(w)).\end{aligned}$$

With this notation, we can then construct sharp upper and lower bounds on $F_{Y(1)|X=c,C_0}$ and $F_{Y(0)|X=c,C_0}$ given (hypothetical) knowledge of the function $w \mapsto (\tau_1(w), \tau_0(w))$. These bounds are given by

$$F^U_{d,FRD(W)}(y,\tau_1(\cdot),\tau_0(\cdot)) = \int F^U_{d,FRD|W=w}(y,\tau_1(w),\tau_0(w))\omega(w,\tau_1(w),\tau_0(w))dF_{W|X=c^-}(w)$$
$$F^L_{d,FRD(W)}(y,\tau_1(\cdot),\tau_0(\cdot)) = \int F^L_{d,FRD|W=w}(y,\tau_1(w),\tau_0(w))\omega(w,\tau_1(w),\tau_0(w))dF_{W|X=c^-}(w),$$

for $d \in \{0,1\}$, where

$$\omega(w,\tau_1(w),\tau_0(w)) \equiv \frac{\Pi_{W=w}(\tau_1(w),\tau_0(w))}{\int \Pi_{W=w}(\tau_1(w),\tau_0(w))dF_{W|X=c^-}(w)}.$$

The resulting sharp upper and lower bounds on the local average treatment effect $\Gamma$ given (hypothetical) knowledge of the function $w \mapsto (\tau_1(w),\tau_0(w))$ are given by

$$\begin{aligned}
&\Gamma^U_{FRD(W)}(\tau_1(\cdot),\tau_0(\cdot)) \\
&\quad \equiv \int y dF^U_{1,FRD(W)}(y,\tau_1(\cdot),\tau_0(\cdot)) - \int y dF^L_{0,FRD(W)}(y,\tau_1(\cdot),\tau_0(\cdot)) \quad \text{and} \\
&\Gamma^L_{FRD(W)}(\tau_1(\cdot),\tau_0(\cdot)) \\
&\quad \equiv \int y dF^L_{1,FRD(W)}(y,\tau_1(\cdot),\tau_0(\cdot)) - \int y dF^U_{0,FRD(W)}(y,\tau_1(\cdot),\tau_0(\cdot)),
\end{aligned}$$

respectively. The following corollary gives the feasible sharp bounds on $\Gamma$, using the fact that the function $w \mapsto (\tau_1(w),\tau_0(w))$ is partially identified.

**Corollary 6.** *Suppose that the assumptions of Theorem 2 hold, mutatis mutandis, with conditioning on the covariates $W$. Then sharp lower and upper bounds on $\Gamma$ are given by*

$$\Gamma^L_{FRD(W)} = \inf_{(t_1(\cdot),t_0(\cdot))\in\mathcal{T}_\mathcal{W}} \Gamma^L_{FRD}(t_1(\cdot),t_0(\cdot)) \quad \text{and}$$
$$\Gamma^U_{FRD(W)} = \sup_{(t_1(\cdot),t_0(\cdot))\in\mathcal{T}_\mathcal{W}} \Gamma^U_{FRD}(t_1(\cdot),t_0(\cdot)),$$

*respectively.*

C.5. **Characteristics of Always- and Potentially-Assigned Units.** It is not possible to determine whether any given unit belongs to the group of always-assigned or potentially-assigned units in our model. This does not mean, however, that it is impossible to give any further characterization of these two groups. In particular, if the data include a vector $W$ of covariates that are measured prior to treatment assignment, and whose conditional distribution given the running variable does not change discontinuously at $c$ among potentially-assigned units, one can identify the distribution of these covariates among both always-assigned and potentially-assigned units. This information could be useful, for instance, for targeting policies aimed at mitigating manipulation. The following corollary formally states this result.

**Corollary 7.** *Suppose that Assumptions 1–2 hold, that* $\mathrm{P}(W \leq w|X = x, M = m)$ *is continuous in* $x$ *at* $c$ *for* $m \in \{0, 1\}$*. Then*

$$\begin{aligned}\mathrm{P}(W \leq w|X = c, M = 1) &= \frac{1}{\tau}(\mathrm{P}(W \leq w|X = c^{+}) - \mathrm{P}(W \leq w|X = c^{-})) \\ &\quad + \mathrm{P}(W \leq w|X = c^{-}) \quad \text{and} \\ \mathrm{P}(W \leq w|X = c, M = 0) &= \mathrm{P}(W \leq w|X = c^{-}).\end{aligned}$$

Of course, identification of the distribution of $W$ immediately implies identification of moments, quantiles, and related summary statistics.

## D. Applicability of our Model

Our model, developed in Section 2, is able to capture a wide range of empirical scenarios of manipulation by appropriately assigning the labels of always-assigned and potentially-assigned to specific groups of units. To illustrate this point, consider a transfer program for which eligibility is based on a cutoff value of a poverty score, and the formula that creates the score takes as inputs household characteristics and assets recorded during home visits by local administrators. There might also be other criteria that make a household (in-)eligible irrespective of the poverty score, so that the resulting RD design could in principle be fuzzy. These types of programs are common in developing countries, and various types of manipulation have been documented for them (e.g., Camacho and Conover, 2011).

The following examples illustrate how various empirical scenarios are accommodated by our model. They also show why it may be necessary to allow always-assigned units to be treated or untreated in some settings, while in others it can be reasonable to assume that all of them are treated. One can easily construct further variants of these examples that also fit into our model, and these examples also have natural analogues in other contexts. For instance, Example 3 and Example 4 below are similar to the two manipulation scenarios for the financial aid example in the introduction, respectively.

**Example 1** ("Unsystematic" Misreporting)**.** There might be concerns of manipulation whenever a running variable can be affected by some agents' behaviors. Running variables are commonly endogenous, misreported, or mismeasured in the empirical literature, and this may certainly affect the composition of the units observed around the cutoff. However, it is not sufficient to create a manipulated running variable in the sense used in this paper. Suppose for example that the formula for the poverty score is not publicly known. Then, even if households might misreport or genuinely modify their input variables (within reasonable bounds), they may not be able to ensure program assignment. All households are potentially-assigned in this case; households just above and below the cutoff are still comparable; and a standard RD analysis could estimate causal parameters for those households with realized poverty scores at the cutoff. This is a special case of our model in which always-assigned units are absent.

**Example 2** (“Systematic” Misreporting)**.** Suppose that some households know the poverty score formula, and local administrators are unwilling or unable to recognize whether a household reports inaccurate information as long as it is within reasonable bounds. Some households with knowledge of the formula, and whose poverty score would otherwise fall to the left of the cutoff, may then be able to misreport their inputs such that their score is to the right of the cutoff. The assumption of one-sided manipulation is likely to hold, e.g., if program assignment is weakly desirable for all households (they can always refuse to participate). They might also have an incentive to report data that put them barely above the cutoff but not exactly at the cutoff, e.g., in order to avoid detection. This makes the assumption of a continuously distributed running variable among always-assigned units palatable. If these misreporting households are systematically different from the other households with poverty scores in the vicinity of the cutoff, the distribution of potential outcomes may be discontinuous at the cutoff, and conventional RD analysis is invalid. In our model, the households with knowledge of the formula that would be willing to misreport their inputs in order to avoid having a poverty score below the cutoff are always-assigned: all of them will have a score above the cutoff, either by misreporting inputs if their score would otherwise fall to the left of the cutoff, or by scoring above the cutoff even without misreporting (they would have misreported inputs if needed, but it is simply not needed in that case). In that sense, they are “always assigned”. All other households are potentially-assigned. Given that always-assigned households are willing to actively violate the rules of the program to ensure that their poverty score is above the cutoff, it may be reasonable to assume that all of them end up treated.

**Example 3** (“Systematic” Misreporting with Partial Verification Checks)**.** Suppose that the same households as above misreport their data to try to ensure program assignment, but that some local administrators now thoroughly verify the information provided to them. As a result, only a fraction of the households is able to carry out its intended misreporting. Those households with knowledge of the formula that would be willing to misreport their inputs in order to avoid having a poverty score below the cutoff, and that would be successful in doing so, are now the always-assigned units in our setup. The households with knowledge of the formula that would be willing to misreport their inputs in order to avoid having a poverty score below the cutoff, but that would be unsuccessful in doing so, are classified as potentially-assigned along with all other households, provided that local administrators simply enter the correct information if they detect misreporting. Indeed, this type of households –

that would unsuccessfully misreport their information if their correct score fell to the left of the cutoff – also exists on the right of the cutoff; they just did not need to try to misreport any data given that they were already on the right of the cutoff. Suppose instead that local administrators apply a penalty by removing households from the data if they detect misreporting. In that case, the same type of households will not be observed on the left of the cutoff anymore; it will only exist on the right of the cutoff and will thus be classified as always-assigned. In both cases, it seems reasonable to assume again that all always-assigned units are treated.

**Example 4** ("Systematic" Misreporting by Administrators)**.** Suppose that all households report their information truthfully, but that local administrators sometimes misreport the information that they receive. This may lead to a manipulated running variable even though the observational units, i.e., the households, do not engage in any manipulation themselves. For instance, local administrators may increase the score of households who support the local government to ensure their program assignment in case their score would otherwise fall to the left of the cutoff. Conventional RD analysis is invalid in this case too, e.g., if a household's political leanings correlates with the effect of program participation. Our general model also likely applies. Manipulation is likely to be one-sided and local administrators are unlikely to misreport information such that the modified scores are equal to the cutoff (e.g., to avoid detection by central administrators). Households who support the local government are then always-assigned, and all others are potentially-assigned. Note that some always-assigned households might now refuse to participate in the program (e.g., if it comes with social stigma), or might have qualified even with a lower poverty score. Alternatively, suppose that local administrators also decrease the score of political opponents whose score would otherwise fall to the right of the cutoff. This would be a situation in which our model does not apply because of two-sided manipulation.

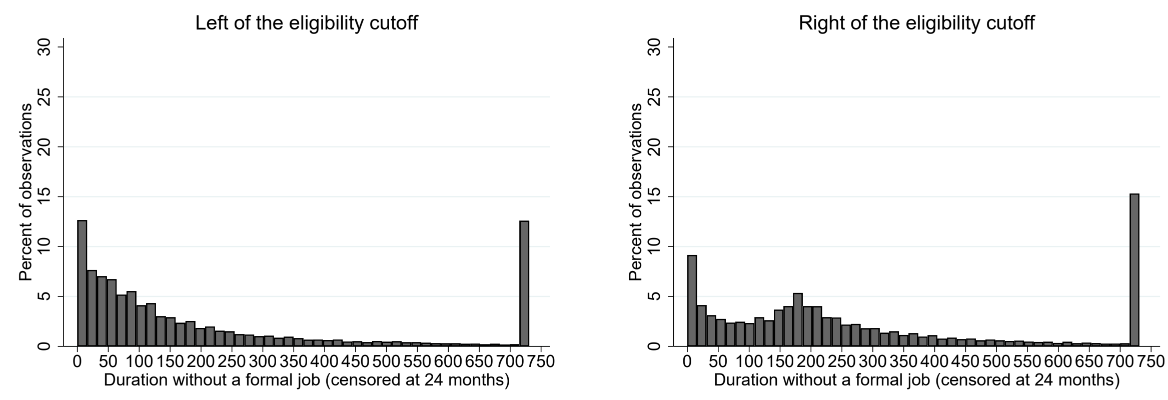
**Example 5** (Manipulation through Location Selection)**.** Manipulation of the running variable does not require that any agent engages in some form of wrongdoing. Suppose that there is no misreporting whatsoever, but that the program only exists in some localities. Households in other localities may then choose to move to become eligible for the program. If the formula is known, the probability of moving may increase discontinuously for households whose poverty score would fall above the cutoff conditional on living in an eligible locality. As a result, the density of the poverty score in eligible localities may be discontinuous at the cutoff and, to the extent that the potential outcomes of movers differ from those of incumbent residents observed around the cutoff, a conventional RD analysis may be invalid. Moreover, the assumptions of one-sided manipulation and of a continuously distributed running variable among always-assigned units are reasonable if the program is weakly desirable. Those households who move because they know that they are eligible for the program at destination are then the always-assigned units in our model (they are responsible for the discontinuity in the moving probability) and they are all likely to be treated in this setting.

**Example 6** (Second Home Visit)**.** Finally, here is another example in which manipulation of the running variable does not require that any agent engages in some form of wrongdoing. Suppose that households' information is measured with some error in any given home visit, and that households can request a second home visit after learning the value of their score by arguing that their information was mismeasured in the first visit. Additionally, only the score based on the most recent visit, which determines program eligibility, is observed by the econometrician. Let $X_{ji}$ be the poverty score for household $i$ based on visit $j$, which is assumed to be smoothly distributed at the cutoff, and suppose that households request a second visit if and only if they were ineligible based on the first visit. The observed poverty score is then: $X_i = X_{1i} \cdot \mathbb{I}\left(X_{1i} \geq c\right) + X_{2i} \cdot \mathbb{I}\left(X_{1i} < c\right)$. Its density is discontinuous at the cutoff as long as error terms are imperfectly correlated across visits. The excess density is due to households whose score fell on the right side of the cutoff in the first visit; those are the always-assigned units in our model. Moreover, to the extent that their potential outcomes differ from those of households observed on the left of the cutoff (whose poverty score fell on the left in both visits), a conventional RD analysis is invalid. In contrast, if feasible, a RD analysis based on $X_{1i}$ or $X_{2i}|X_{1i} < c$ could be valid in this setting. Depending on the details of the program, this is also a case in which it may be reasonable to allow always-assigned households to be treated or untreated.

## E. Additional Tables and Graphs for the Empirical Application

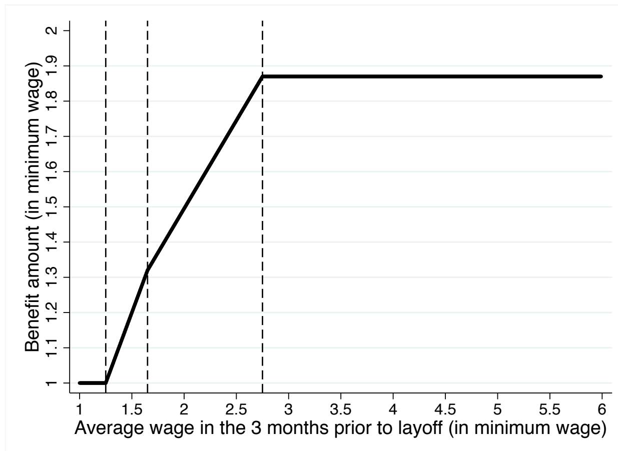
We present here some supporting graphs for our empirical application. Figure 4 displays the distribution of our outcome variable (duration without a formal job, censored at two years after layoff) on the left and on the right of the cutoff (30-day window around the cutoff). Figure 5 displays the full schedule of the UI benefit level, which is a function of a beneficiary's average monthly wage in the three years prior to her layoff. Figure 6 displays the mean of different covariates on each side of the cutoff by day between the layoff and eligibility dates. Finally, Table 3 displays estimates of the change in the average value of these covariates at the cutoff, including the UI replacement rate, and of the average value of these covariates among potentially-assigned and always-assigned units, based on the results in Corollary 7.

Figure 4: Distribution of our outcome variable on each side of the cutoff



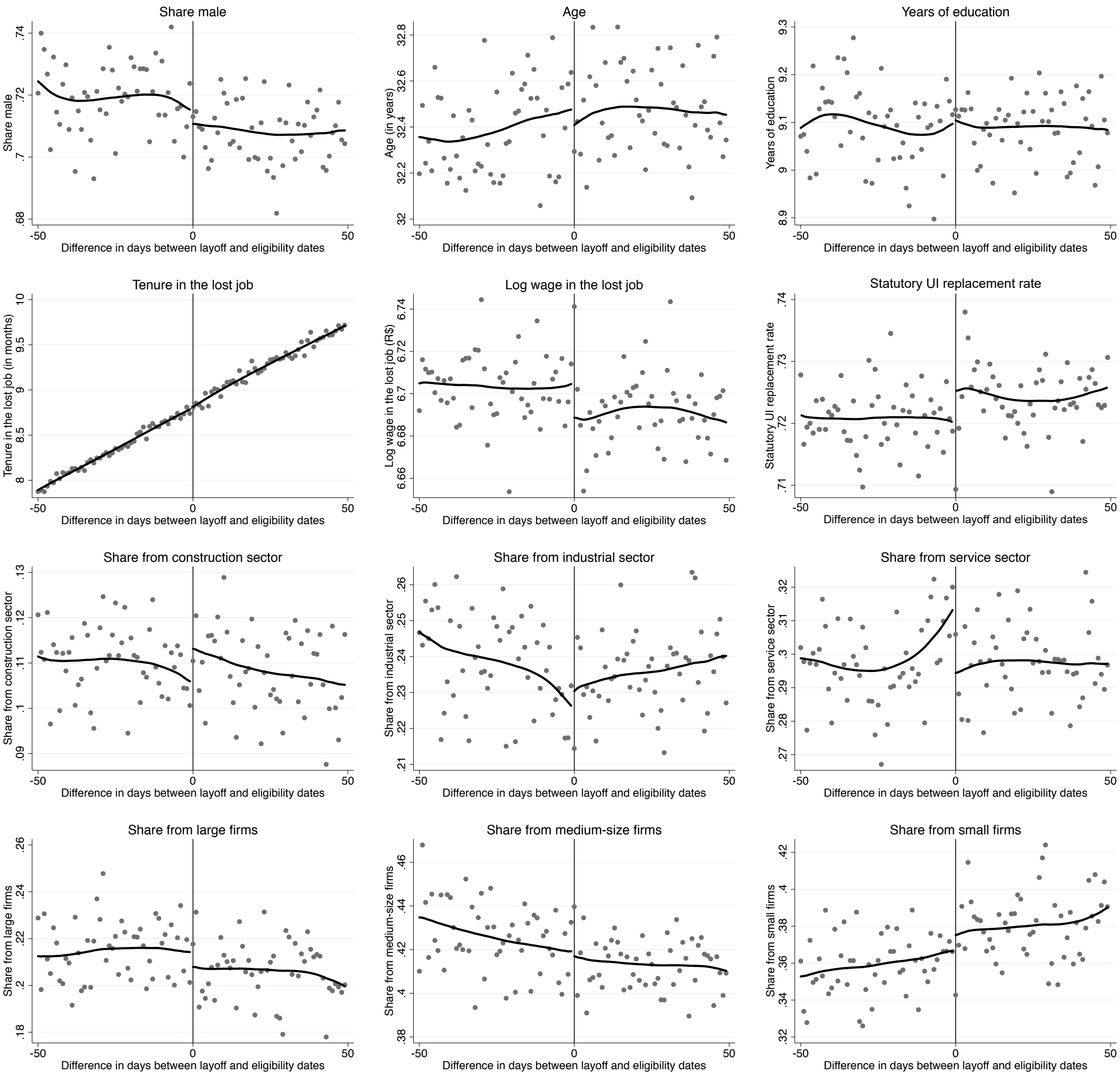

Notes: The figure displays the distribution of our outcome variable (duration without a formal job, censored at two years after layoff) on the left and on the right of the cutoff (30-day window on each side of the cutoff). The figure is based on a sample of 102,791 displaced formal workers.

Figure 5: Monthly UI benefit amount




The figure displays the relationship between a UI beneficiary's average monthly wage in the three months prior to her layoff and her monthly UI benefit level. All monetary values are indexed to the federal minimum wage, which changes every year. The replacement rate is 100% at the bottom of the wage distribution as the minimum benefit level is equal to one minimum wage. The graph displays a slope of 0% until 125% of the minimum wage, then of 80% until 165% of the minimum wage, and finally of 50% until 275% of the minimum wage. The maximum benefit level is equal to 187% of the minimum wage.

Figure 6: Graphical evidence of potential selection at the cutoff for our empirical application




Notes: The figure displays the mean of different covariates on each side of the cutoff by day between the layoff and eligibility dates, as well as local linear regressions on each side of the cutoff using an edge kernel and a bandwidth of 30 days. The figure is based on a RD sample of 169,575 displaced formal workers.

Table 3: Characteristics of always- and potentially-assigned workers

| | Difference at the cutoff | Potentially-assigned | Always-assigned |
|---|---|---|---|
| Share male | -0.0031 | 0.714 | 0.665 |
| | [-0.0168;0.0105] | [0.704;0.724] | [0.439;0.892] |
| Average age | -0.0729 | 32.475 | 31.345 |
| | [-0.3091;0.1633] | [32.304;32.645] | [27.627;35.063] |
| Average years of education | 0.0011 | 9.104 | 9.121 |
| | [-0.0803;0.0825] | [9.049;9.160] | [7.836;10.406] |
| Average tenure | 0.0103 | 8.802 | 8.961 |
| | [-0.0418;0.0623] | [8.771;8.833] | [8.100;9.821] |
| Average log wage | -0.016 | 6.704 | 6.456 |
| | [-0.0308;-0.0012] | [6.693;6.716] | [6.208;6.704] |
| Average UI replacement rate | 0.0051 | 0.720 | 0.800 |
| | [0.0005;0.0098] | [0.717;0.724] | [0.722;0.878] |
| Share from commercial sector | 0.0071 | 0.355 | 0.465 |
| | [-0.0059;0.02] | [0.346;0.365] | [0.264;0.665] |
| Share from construction sector | 0.0073 | 0.106 | 0.218 |
| | [-0.0015;0.0161] | [0.099;0.112] | [0.079;0.358] |
| Share from industrial sector | 0.0061 | 0.225 | 0.319 |
| | [-0.006;0.0182] | [0.216;0.234] | [0.131;0.507] |
| Share from service sector | -0.0204 | 0.314 | -0.002 |
| | [-0.0332;-0.0077] | [0.305;0.324] | [-0.201;0.197] |
| Share from small firm | 0.0083 | 0.367 | 0.496 |
| (<10 employees) | [-0.0057;0.0224] | [0.357;0.377] | [0.268;0.730] |

Notes: Total number of observations within our bandwidth of 30 days around the cutoff: 102,791 displaced formal workers. Numbers in square brackets are 95% confidence intervals calculated by adding $\pm 1.96\times$standard error to the respective point estimate, where standard errors are calculated via the bootstrap with 500 replications. The characteristics of potentially-assigned and always-assigned units are obtained using the results in Corollary 7.

## References


Anderson, G., O. Linton, and Y.-J. Whang (2012): "Nonparametric estimation and inference about the overlap of two distributions," *Journal of Econometrics*, 171(1), 1–23.

Anderson, M., and J. Magruder (2012): "Learning from the Crowd: Regression Discontinuity Estimates of the Effects of an Online Review Database," *Economic Journal*, 122(563), 957–989.

Andrews, D. (2000): "Inconsistency of the bootstrap when a parameter is on the boundary of the parameter space," *Econometrica*, 68(2), 399–405.

Andrews, D., and P. Barwick (2012): "Inference for parameters defined by moment inequalities: A recommended moment selection procedure," *Econometrica*, 80(6), 2805–2826.

Andrews, D., and G. Soares (2010): "Inference for parameters defined by moment inequalities using generalized moment selection," *Econometrica*, 78(1), 119–157.

Berger, R. (1982): "Multiparameter hypothesis testing and acceptance sampling," *Technometrics*, 24(4), 295–300.

Bertanha, M., and G. W. Imbens (2019): "External Validity in Fuzzy Regression Discontinuity Designs," *Journal of Business & Economic Statistics, forthcoming.*

Camacho, A., and E. Conover (2011): "Manipulation of Social Program Eligibility," *American Economic Journal: Economic Policy*, 3(2), 41–65.

Card, D., R. Chetty, and A. Weber (2007): "Cash-on-Hand and Competing Models of Intertemporal Behavior: New Evidence from the Labor Market," *Quarterly Journal of Economics*, 122(4), 1511–1560.

Card, D., C. Dobkin, and N. Maestas (2009): "Does Medicare Save Lives?," *The Quarterly Journal of Economics*, 124(2), 597–636.

Card, D., and L. Giuliano (2014): "Does Gifted Education Work? For Which Students?," *NBER Working Paper*, 20453.

Carvalho, C., C. Corbi, and R. Narita (2017): "Unintended consequences of unemployment insurance: Evidence from stricter eligibility criteria in Brazil," *Economic Letters.*

Cattaneo, M. D., M. Jansson, and X. Ma (2019): "Simple Local Polynomial Density Estimators," *Journal of the American Statistical Association, forthcoming.*

Cheng, M.-Y. (1997): "A bandwidth selector for local linear density estimators," *Annals of Statistics*, 25(3), 1001–1013.

Cheng, M.-Y., J. Fan, and J. Marron (1997): "On Automatic Boundary Corrections," *Annals of Statistics*, 25, 1691–1708.

Chernozhukov, V., S. Lee, and A. Rosen (2013): "Intersection Bounds: estimation and inference," *Econometrica*, 81(2), 667–737.

Dee, T., W. Dobbie, B. Jacob, and J. Rockoff (2019): "The Causes and Consequences of Test Score Manipulation: Evidence from the New York Regents Examinations," *American Economic Journal: Applied Economics*, 11(3), 382–423.

DONG, Y. (2018): "Alternative Assumptions to Identify LATE in Fuzzy Regression Discontinuity Designs," *Oxford Bulletin of Economics and Statistics*, 88(5), 1020–1027.

FAN, J., AND I. GIJBELS (1996): *Local Polynomial Modelling and Its Applications*. Chapman & Hall, London.

FAN, J., Q. YAO, AND H. TONG (1996): "Estimation of conditional densities and sensitivity measures in nonlinear dynamical systems," *Biometrika*, 83(1), 189–206.

FELDSTEIN, M. (1976): "Temporary Layoffs in the Theory of Unemployment," *Journal of Political Economy*, 84, 937–958.

FRANDSEN, B. R., M. FRÖLICH, AND B. MELLY (2012): "Quantile treatment effects in the regression discontinuity design," *Journal of Econometrics*, 168(2), 382–395.

GERARD, F., AND G. GONZAGA (2016): "Informal Labor and the Efficiency Cost of Social Programs: Evidence from the Brazilian Unemployment Insurance Program," *NBER Working Paper*, 22608.

HAHN, J., P. TODD, AND W. VAN DER KLAAUW (2001): "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design," *Econometrica*, 69(1), 201–209.

HALL, P., R. C. WOLFF, AND Q. YAO (1999): "Methods for estimating a conditional distribution function," *Journal of the American Statistical Association*, 94(445), 154–163.

HOPENHAYN, H., AND J. P. NICOLINI (2009): "Optimal Unemployment Insurance and Employment History," *Review of Economic Studies*, 76, 1049–1070.

HOROWITZ, J. L., AND C. F. MANSKI (1995): "Identification and robustness with contaminated and corrupted data," *Econometrica*, 63(2), 281–302.

IMBENS, G., AND J. ANGRIST (1994): "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 62(2), 467–475.

IMBENS, G., AND C. MANSKI (2004): "Confidence Intervals for Partially Identified Parameters," *Econometrica*, 72(6), 1845–1857.

JONES, M. C. (1993): "Simple boundary correction for kernel density estimation," *Statistics and Computing*, 3(3), 135–146.

LEE, D. (2009): "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects," *Review of Economic Studies*, 76(3), 1071–1102.

LEE, D. S. (2008): "Randomized experiments from non-random selection in US House elections," *Journal of Econometrics*, 142(2), 675–697.

LEJEUNE, M., AND P. SARDA (1992): "Smooth estimators of distribution and density functions," *Computational Statistics & Data Analysis*, 14(4), 457–471.

MANSKI, C. F. (2003): *Partial identification of probability distributions*. Springer Science & Business Media.

——— (2009): *Identification for prediction and decision*. Harvard University Press.

Masten, M. A., and A. Poirier (2020): "Inference on Breakdown Frontiers," *Quantitative Economics*, 11(1), 41–111.

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

Sallee, J. (2011): "The Surprising Incidence of Tax Credits for the Toyota Prius," *American Economic Journal: Economic Policy*, 3, 189–219.

Schmieder, J., T. von Wachter, and S. Bender (2012): "The Effects of Extended Unemployment Insurance Over the Business Cycle: Evidence from Regression Discontinuity Estimates over Twenty Years," *Quarterly Journal of Economics*, 127(2), 701–752.

Scott-Clayton, J. (2011): "On Money and Motivation: A Quasi-Experimental Analysis of Financial Incentives for College Achievement," *Journal of Human Resources*, 46(3), 614–646.

Solis, A. (2017): "Credit Access and College Enrollment," *Journal of Political Economy*, 125(2), 562–622.

Stoye, J. (2009): "More on confidence intervals for partially identified parameters," *Econometrica*, 77(4), 1299–1315.

——— (2010): "Partial identification of spread parameters," *Quantitative Economics*, 1(2), 323–357.

Urquiola, M., and E. Verhoogen (2009): "Class-Size Caps, Sorting, and the Regression Discontinuity Design," *American Economic Review*, 99(1), 179–215.