Regression Discontinuity Designs with Nonclassical Measurement Error

Takahide Yanagi*

Hitotsubashi University

October 27, 2015

Abstract

This paper develops a nonparametric identification analysis in regression discontinuity (RD) designs where each observable may contain measurement error. Our analysis allows the measurement error to be nonclassical in the sense that it can be arbitrarily dependent of the unobservables as long as the joint distribution satisfies a few smoothness conditions. We provide formal identification conditions under which the standard RD estimand based on the observables identifies a local weighted average treatment effect parameter. We also show that our identifying conditions imply a testable implication of the continuous density of the observable assignment variable.

Keywords: regression discontinuity; measurement error; nonparametric identification; local average treatment effect; treatment effect heterogeneity.

JEL Classification: C14; C21; C25.

*Graduate School of Economics, Hitotsubashi University, Naka 2-1, Kunitachi, Tokyo, 186-8601, Japan. Email: t.yanagi@r.hit-u.ac.jp

1 Introduction

Regression discontinuity (RD) designs are quasi-experimental designs to evaluate the impact of possibly endogenous binary policy variables. RD designs allow econometricians to nonparametrically identify a local average treatment effect (LATE) based on policy rules that generate discontinuous conditional treatment probabilities (Hahn et al., 2001). Thanks to their usefulness, RD designs have been extensively utilized in the microeconometric literature. See, for example, Imbens and Lemieux (2008) and Lee and Lemieux (2010) for excellent reviews on theoretical and empirical contributions in the RD literature.

Empirical applications based on RD designs are often forced to utilize data that may contain measurement error. This is particularly common when econometricians use survey data to conduct their analyses (Bound et al., 2001); there are at least two reasons why RD applications often require survey data. First, econometricians may not be able to use data gathered by policy makers due to data unavailability. Second, survey data tends to contain more abundant variables such as demographical characteristics than policy maker's data does. There are several empirical studies based on RD designs with survey data, such as Card and Shore-Sheppard (2004), Ludwig and Miller (2007), Card et al. (2008), Battistin et al. (2009), Schanzenbach (2009), and Hullegie and Klein (2010). Some of them pay attention to the possibility of the presence of measurement error in their observables, and of these, several theoretical studies develop identification, estimation, and inference procedures for RD designs with measurement error (see the section "Related literature" below).

The present paper contributes to the literature by providing a nonparametric identification analysis in general RD designs with measurement error. Our analysis is developed for the general fuzzy RD design, including the sharp RD design as a special case. The novelty of the present paper is to allow situations in which any observables can contain measurement error. Our analysis permits each of the observed outcome, treatment, and assignment variable to contain measurement error. Furthermore, we allow the measurement error to be nonclassical in the sense that the measurement error can be arbitrarily dependent of all the unobservables except for a few smoothness conditions. We do not assume any independence conditions on the joint distribution of the measurement error and unobservables. Since independence conditions can be restrictive in several empirical situations, it is a prominent case in the literature that our analysis works without them.

This paper aims to show the parameter that is identified by the standard RD estimand based on the possibly mismeasured variables under plausible conditions instead of independence conditions. To this end, we introduce two key identifying conditions. The first is the continuity of the joint distribution of measurement error and unobservables. The continuity is considerably weaker than independence conditions and could be unrestrictive in many empirical situations. The second is also an essential condition under which measurement error on the assignment variable for the compliers has the positive probability mass at zero. It requires that there are compliers whose assignment values are observed without measurement error near the cutoff point. The validity of the assumption depends on the accuracy of the data. This assumption guarantees that the conditional treated probability is discontinuous even when the assignment variable contains measurement error.

We show that the standard RD estimand based on the possibly mismeasured observables identifies a local weighted average treatment effect (LWATE) parameter under a set of identification assumptions. While the weight depends on the conditional probability that the assignment variable for the complier is correctly reported, it does not depend on the distribution of the measurement error for the outcome and treatment. This observation emphasizes that an accurate record of the assignment variable is especially important in RD applications.

We prove that the LWATE is not identical to the standard RD LATE parameter in general. Accordingly, econometricians should pay careful attention to interpreting RD estimates when their observables may contain measurement error. However, we also discuss that the LWATE gives econometricians meaningful information on the standard RD LATE parameter. In particular, the LWATE is equal to the standard RD LATE if the individual treatment effect and the potential treatment status are jointly independent of measurement error on the assignment variable. Further, the LWATE is equal to the standard RD LATE if the treatment effect is homogeneous across economic units at the cutoff. Otherwise, in some situations, the LWATE may tell us the sign of the standard RD LATE.

We also show that the continuity of the density of the observed assignement variable is a testable implication for the identification, in the spirit of Lee (2008) and McCrary (2008), even in the presence of measurement error. This suggests the empirical importance of testing the continuous density of the observable assignment variable even when the assignment variable contains measurement error. In other words, we justify the enforcement of McCrary's continuous density test in general RD designs with measurement error.

Related literature: The current paper most closely relates to the theoretical contribution of Battistin et al. (2009), who consider the partially fuzzy RD design (Battistin and Rettore, 2008) in which the assignment variable contains measurement error with the positive probability mass at zero. They assume that conditional on the "true" assignment variable, the indicator for the presence of the measurement error and the mismeasured assignment variable are jointly independent of the outcome and the treatment status. Under the independence assumption, the average treatment effect on the treated at the cutoff is shown to be identified by the standard RD estimand. We emphasize that the present paper and the theoretical analysis of Battistin et al. (2009) are different in at least five aspects. First and foremost, the analysis in the present paper does not need any independence conditions. It does not restrict the measurement error for the assignment variable to be independent of the unobservables. Second, we allow the presence of measurement error for any observable variables not just for the assignment variable. Battistin et al. (2009) do not allow the presence of measurement error for the treatment and outcome, while they advert to the possibility of measurement error for the treatment status. Third, we consider the general fuzzy RD design, which includes the partially fuzzy RD design in Battistin et al. (2009) as a special case. Fourth, we provide a set of formal conditions under which we derive an exact form of the parameter that is identified by the standard RD estimand with nonclassical measurement error. Battistin et al. (2009) do not develop such a consideration. Fifth, we provide the testable implication led by our identification conditions unlike Battistin et al. (2009).

The present paper also closely relates to Card et al. (2015) who develop analyses in the sharp and fuzzy regression kink (RK) designs. Their fuzzy RK design allows the presence of measurement error for the assignment and policy variables. In RK designs, the policy variable is a continuous random variable with the kinked conditional mean at a cutoff point. They show that the fuzzy RK estimand identifies a weighted average of the policy effect even when the measurement error is nonclassical. The key identifying conditions in their paper as well as the present study are the presence of individuals who correctly report their assignment values and a smooth density condition. Our analysis is thus related to the analysis in Card et al. (2015) in terms of techniques. However, we stress that the situations considered in both papers are different. The policy variable in the RD design is binary with the discontinuous conditional mean, although the policy variable in the RK design is continuous with the kinked conditional mean. As a result, the fuzzy RK estimand is different from the standard RD estimand. In

addition, Card et al. (2015) do not consider the situation in which the outcome may contain measurement error. The statements of the main theorems as well as the settings, estimands, identification conditions, and identifying parameters in both papers are different.

Several other papers contribute to identification and estimation of RD designs with measurement error, although the settings in these are not the same as ours. Hullegie and Klein (2010), Yu (2012), Davezies and Le Barbanchon (2014), and Yanagi (2014) consider RD designs where the assignment variable may contain continuous measurement error. They point out that the continuous measurement error invalidates the RD research design. Specifically, they find that given the observed assignment variable, the conditional treated probability becomes continuous due to the continuous measurement error even when given the true assignment variable, the conditional treated probability is discontinuous at the cutoff. Further, their papers propose methodologies to overcome such problems due to the continuous measurement error. We stress that the situations in the present paper and the abovementioned papers are distinct because the measurement error in the latter is continuously distributed without any probability mass. Furthermore, to our knowledge, no existing study develops an analysis for RD designs in which measurement error can be arbitrarily dependent of the unobservables.

There are other studies that consider particular types of measurement error in RD designs. Dong (2015c) develops an analysis where the observed assignment variable contains rounding error. She proposes a parametric approach to identify the LATE parameter with the rounding error. We note that the analysis in the present paper does not allow rounding error, because the rounding error for a continuous assignment variable would not have any probability mass. Elsewhere, Pei (2011) proposes a parameteric analysis for RD designs where the assignment variable and its measurement error are discretely distributed. Our analysis is also different from his analysis, since ours requires that the assignment variable and the measurement error are continuously distributed with a probability mass.

Organization of the paper: Section 2 introduces the general RD design with measurement error. Section 3 develops our identification analysis. Section 4 discusses the testable implication led by our identification conditions. Section 5 concludes. Appendix contains the proofs of all the theorems and technical lemmas.

Notation: For a real function $g(\cdot)$, $g(a^+) := \lim_{x \downarrow a} g(x)$ and $g(a^-) := \lim_{x \uparrow a} g(x)$ are the right and left limits, respectively, at $a \in \mathbb{R}$. We denote the Lebesgue measure as $\lambda(\cdot)$, the

5

indicator function as $\mathbf{1}(\cdot)$, and the Cartesian product of sets $\mathcal{S}_A$ and $\mathcal{S}_B$ as $\mathcal{S}_{AB}$. For a generic random vector Z and an event E, we write the conditional cumulative distribution function and conditional density (with respect to the Lebesgue measure) of Z given E as $F_{Z|E}(\cdot)$ and $f_{Z|E}(\cdot)$, respectively. For ease of exposition, we omit the interval of the integration unless there is confusion.

2 Setting

This section introduces the setting considered in this paper. We first explain the general RD design and summarize the identification analyses developed in the literature. We then introduce measurement error for each observable variable.

2.1 Regression discontinuity design

The outcome Y is a random variable. The treatment $T \in \{0, 1\}$ is binary: $T = 1$ if the individual is treated and $T = 0$ otherwise. The assignment variable X is a continuous covariate, which can affect both Y and T. The random variables are defined on a probability space $(\Omega, \mathcal{F}, P)$.

In RD designs, T is (perfectly or partly) determined by X with a jump in the conditional treatment probability. For example, $T = \mathbf{1}(X \geq x_0)$ in a sharp RD design where $x_0 \in \mathbb{R}$ is a known constant cutoff. More generally, in the fuzzy RD designs, T is partly determined by X and $E(T|X = x_0^+) \neq E(T|X = x_0^-)$. Since the sharp design is a special case of the fuzzy design, this paper concentrates on the fuzzy design only. Without loss of generality, we normalize $x_0 = 0$. Let $\mathcal{S}_0$ be a neighbourhood around the cutoff 0.

We are interested in the causal effect of T on Y. To be specific, we introduce the potential outcome notations. Y_0 is the potential outcome when the individual is untreated ($T = 0$), and Y_1 is the potential outcome when the individual is treated ($T = 1$). For notational simplicity, let $\alpha := Y_0$ and $\beta := Y_1 - Y_0$. The causal effect for the individual is β. The treatment effect is allowed to be heterogeneous and can depend on the assignment value. Thus, we can write that $Y = TY_1 + (1 - T)Y_0 = \alpha + \beta T$.

We then introduce the notion of potential treatment status as in Dong (2015a,b). Let $T(x)$ be the random variable that represents the potential treatment status when the assignment value for the invididual equals x. Moreover, for individuals with $X \geq 0$, let $T_1 := T(X)$ and $T_0 := T(0^-)$ when the limit exists almost surely. Similarly, for individuals with $X < 0$, let

$T_0 := T(X)$ and $T_1 := T(0^+)$ when the limit exists almost surely. In this manner, the potential treatment status (T_0, T_1) for any individuals can be well-defined. Note that (T_0, T_1) is a random vector with the support $\{0,1\}^2 = \{0,1\} \times \{0,1\}$. It can vary across individuals and can depend on the assignment value. In this notation, we can write $T = T_0 + \mathbf{1}(X \geq 0)(T_1 - T_0)$.

For ease of exposition, we define random vectors $\tilde{Y} := (Y_0, Y_1)$ and $\tilde{T} := (T_0, T_1)$. We denote $(Y_0, Y_1) = (y_0, y_1)$ and $(T_0, T_1) = (t_0, t_1)$ as $\tilde{Y} = \tilde{y}$ and $\tilde{T} = \tilde{t}$, respectively.

We define the following events for four types of individuals in the probability space as in Angrist et al. (1996) and Dong (2015a):

$$A := \{\omega \in \Omega : T_1(\omega) = T_0(\omega) = 1\},$$

$$C := \{\omega \in \Omega : T_1(\omega) = 1,\ T_0(\omega) = 0\},$$

$$D := \{\omega \in \Omega : T_1(\omega) = 0,\ T_0(\omega) = 1\},$$

$$N := \{\omega \in \Omega : T_1(\omega) = T_0(\omega) = 0\}.$$

Here, A is the set of the always takers, C is that of the compliers, D is that of the defiers, and N is that of the never takers. The always takers are treated irrespective of whether their assignment is above or below the cutoff. The compliers are treated if and only if their assignment is above the cutoff. The defiers are treated if and only if their assignment is below the cutoff. The never takers deny the treatment irrespective of their assinment values.

The standard RD estimand is the following ratio of the discontinuity at the cutoff:

$$\tau := \frac{E(Y|X = 0^+) - E(Y|X = 0^-)}{E(T|X = 0^+) - E(T|X = 0^-)}.$$

It is known in the literature that, τ identifies a local average treatment effect (LATE) under a set of identification assumptions. Specifically, Hahn et al. (2001) and Dong (2015a) show that

$$\tau = E(\beta|C, X = 0). \tag{1}$$

The right-hand side is the standard LATE parameter in RD designs, which is the average treatment effect for the compliers at the cutoff. The LATE is thus identified if (Y, T, X) is observed without measurement error.

Hahn et al. (2001) show that the LATE is identified under three notable assumptions, summarized as follows:

(HTV1): $E(Y_t|X = \cdot)$ is continuous at 0 for $t \in \{0, 1\}$.

(HTV2): $P(T_1 < T_0|X = x) = P(D|X = x) = 0$ for any $x \in \mathcal{S}_0$.

(HTV3): $(\beta, \tilde{T})$ is jointly independent of X conditional on $X = x$ for any $x \in \mathcal{S}_0$.

The first assumption restricts the continuity of the conditional means. Intuitively, it requires that the other factors affecting the outcome except for the treatment are balanced near the cutoff. It also implies that any individual cannot perfectly manipulate his/her assignment value (Lee, 2008; McCrary, 2008). The second assumption requires that the defiers do not exist. It is a popular assumption to identify LATE parameters (Angrist et al., 1996) in the microeconometric literature. The third asssumption requires that the policy effect and the potential treatment status are independent of the assignment variable near the cutoff. It implies that the LATE parameter does not vary near the cutoff, that is, $E(\beta|C, X = x)$ is constant for $x \in \mathcal{S}_0$.

Recently, Dong (2015a) points out that the independece assumption (HTV3) can be restrictive in several econometric applications. Instead of the independence assumption, she provides more plausible identification assumptions implied by the smoothness of the conditional density. Specifically, she introduces the following continuity:

(D1): $f_{X|\tilde{Y}=\tilde{y},\tilde{T}=\tilde{t}}(\cdot)$ is continuous at 0 for any $(\tilde{y}, \tilde{t})$ and $f_X(\cdot)$ is strictly positive on $\mathcal{S}_0$.

(D1) implies that any individual cannot perfectly manipulate his/her assignment value. It also restricts that any discontinuous changes affecting the potential outcomes do not exist at the cutoff except for the discontinuous treatment probability. Dong (2015a) shows that this condition implies the following continuity:

(D2): $P(\tilde{T} = \tilde{t}|X = \cdot)$ is continuous at 0 for any $\tilde{t} \in \{0, 1\}^2$.

(D3): $E(Y_t|\tilde{T} = \tilde{t}, X = \cdot)$ is continuous at 0 for any $t \in \{0, 1\}$ and $\tilde{t} \in \{0, 1\}^2$.

(D2) means the conditional probabilities of the individual types are continuous at the cutoff. (D3) implies that the conditional means of the potential outcomes are also continuous at the cutoff. Intuitively, like (HTV1), these imply that *ceteris paribus* conditions are balanced near the cutoff. Dong (2015a) then proves that (D2) and (D3) with (HTV2) lead to identification result (1) without independence conditions.

As in Dong (2015a), the present paper develops an analysis based on such plausible smoothness conditions instead of independence conditions.

2.2 Measurement error

The present paper focuses on RD designs in which each observable possibly contains measurement error. We consider the situation where the true outcome, treatment, and assignment variable (Y, T, X) may not be observed due to the presence of measurement error.

We introduce measurement error for (Y, T, X) in the following manner.

$$Y^* = Y + U_Y,$$

$$T^* = T + U_T,$$

$$X^* = X + U_X,$$

where U_Y is the measurement error for Y, U_T for T, and U_X for X. We note that U_T is dependent of T unless $U_T = 0$ almost surely because the support of U_T given $T = 1$ is $\{-1, 0\}$ and that given $T = 0$ is $\{0, 1\}$.

Econometricians cannot observe (Y, T, X), but they observe (Y^*, T^*, X^*) that may contain measurement error. The distribution of the observables (Y^*, T^*, X^*) is determined by the distribution of the unobservables $(\tilde{Y}, \tilde{T}, X, U_Y, U_T, U_X)$.

The representation for measurement is general enough, unless we restrict the relationship between $(\tilde{Y}, \tilde{T}, X, U_Y, U_T, U_X)$. Our analysis allows the measurement error to be nonclassical in the sense that the measurement error can be arbitrarily dependent of the unobservables as long as the distribution satisfies a few smoothness conditions. We need only a certain extent of continuity of the distribution of $(\tilde{Y}, \tilde{T}, X, U_Y, U_T, U_X)$. We expound specific conditions to establish our identification in the following section.

Importantly, our setting may allow the situation where policy makers may not observe the true variables (Y, T, X) as well as econometricians do. For example, policy makers may observe not the true X but the mismeasured X^* when some individuals may make the false report on their assignment values by accident or by partly (not perfectly) manipulating their reporting. Then, policy makers may decide to treat the individuals based on their misreported assignment values and observe the outcome after the treatment. In this situation, the policy makers may observe only (Y^*, T^*, X^*) as well as the econometricians do. Worse, in such situations, the measurement error for X may be dependent of the individual treatment effect. For example, this would happen when an individual who cannot receive the treatment based on his/her true assignment value tends to make a false report if he/she makes a high estimate of his/her

treatment effect. Our analysis could be informative even in such situations.

We focus on showing the parameter that econometricians identify based on observables (Y^*, T^*, X^*) in the general RD design. Specifically, our interest is to prove the parameter that is identified by the following standard RD estimand based on the observables:

$$\tau^* := \frac{E(Y^*|X^* = 0^+) - E(Y^*|X^* = 0^-)}{E(T^*|X^* = 0^+) - E(T^*|X^* = 0^-)}.$$

τ^* is identified by data even when the observables contain measurement error. We note that τ^* can be estimated with a random sample $\{(Y_i^*, T_i^*, X_i^*)\}_{i=1}^n$ based on standard estimation techniques such as local polynomial regressions developed in the RD literature (see, e.g., Imbens and Kalyanaraman, 2012; Calonico et al., 2014; Arai and Ichimura, 2015a,b).

3 Identification with measurement error

This section develops an identification analysis under the setting of the previous section. Our identification analysis is composed of three steps. First, we show that the continuity of conditional probabilities of individual types and that of conditional means of potential outcomes such as (D2) and (D3) are established even in the presence of measurement error. Second, we prove that the conditional probability of the observable treatment is discontinuous at the cutoff under a set of assumptions. Finally, we show that these results naturally lead to the statement that τ^* identifies a LATE parameter.

3.1 Continuity of conditional probabilities and conditional means

This subsection shows that conditional probabilities and conditional means such as those in (D2) and (D3) are continuous with respect to the assignment variable even when the observable variables may contain measurement error. To this end, we introduce two key assumptions on the measurement error for X and on a joint density. We also require a regularity condition. Let $\mathcal{S}_{VX\tilde{Y}} := \mathcal{S}_V \times \mathcal{S}_X \times \mathcal{S}_{Y_0} \times \mathcal{S}_{Y_1} \subset \mathbb{R}^4$ be an arbitrary large compact set.

Assumption 1 (X's error). *U_X is represented by*

$$U_X = G \cdot V,$$

where $G \in \{0, 1\}$ is a binary random variable indicating the presence of measurement error with

$P(G = 0|C, X = 0) > 0$ *and V is continuously distributed error.* $P(G = 0|\tilde{T} = \tilde{t}, V = \cdot, X = \cdot, \tilde{Y} = \cdot)$ *is continuous on $\mathcal{S}_{VX\tilde{Y}}$ for any $\tilde{t} \in \{0, 1\}^2$.*

Assumption 2 (smooth density). *The conditional density $f_{VX\tilde{Y}|\tilde{T}=\tilde{t}}(\cdot, \cdot, \cdot)$ is continuous on $\mathcal{S}_{VX\tilde{Y}}$ for any $\tilde{t} \in \{0, 1\}^2$ and $f_X(\cdot)$ is strictly positive on $\mathcal{S}_0$.*

Assumption 3 (regularity). *The support of $(V, X, \tilde{Y})$ is a subset of $\mathcal{S}_{VX\tilde{Y}}$.*

Assumption 1 restricts the distribution of the measurement error for X. It requires that the measurement error is continuously distributed except at zero. The measurement error must have the positive probability mass at zero, which is essential for our identification analysis. This modeling for the measurement error is essentially identical to the contaminated sampling model considered in Battistin et al. (2009). It also assumes the continuity of the conditional probability of no measurement error. We stress that Assumption 1 does not require independence between the measurement error and true unobservable variables, which may be undesirable in empirical applications.

Whether Assumption 1 is valid for the survey data at hand would depend on the accuracy of the data. The measurement error can have positive probability mass if there are compliers who correctly report their assignment values. Otherwise, the measurement error would be continuously distributed without the probability mass. In such situations, we need other strategies to identify a local average treatment effect such as in Hullegie and Klein (2010), Yu (2012), Davezies and Le Barbanchon (2014), or Yanagi (2014).

Assumption 2 requires that the conditional density of $(V, X, \tilde{Y})$ is smooth and that the density of X is strictly positive near the cutoff point. It is similar to the smooth density assumption (D1). Therefore, intuitively, Assumption 2 implies that no economic units can perfectly manipulate their true assignment variable and continuous error. However, we note that Assumption 2 conditions stronger smoothness than (D1) does. This is because we study the effect of the measurement error. We stress that although Assumption 2 implicitly requires that $\tilde{Y}$ is continuously distributed, we can allow $\tilde{Y}$ to be a discrete random vector. The same results in this paper can be established under a different smoothness condition in such a situation[1].

Assumption 3 is a regularity condition to use the dominated convergence theorem. Strictly speaking, it is stronger than necessary, and hence, it can be substituted with weaker conditions. However, we require Assumption 3 for its perspicuity. We note that the support of X^*, $\mathcal{S}_{X^*}$, is

---

[1]If $\tilde{Y}$ is discretely distributed with a finite support $\mathcal{S}_{\tilde{Y}}$, then it suffices that $P(G = 0|\tilde{T} = \tilde{t}, V = \cdot, X = \cdot, \tilde{Y} = \tilde{y})$ and $f_{VX|\tilde{T}=\tilde{t}, \tilde{Y}=\tilde{y}}(\cdot, \cdot)$ is continuous for any $\tilde{t} \in \{0, 1\}^2$ and $\tilde{y} \in \mathcal{S}_{\tilde{Y}}$ in Assumptions 1 and 2.

also a subset of a compact set under the assumption. For this reason, we consider that $\mathcal{S}_{X^*}$ is a subset of $\mathcal{S}_X$ without loss of generality.

The continuity of conditional probabilities and conditional means such as (D2) and (D3) is established under the assumptions.

Lemma 1. *Suppose that Assumptions 1–3 hold. Then, (i) $P(\tilde{T} = \tilde{t}|G = 0, X^* = \cdot) = P(\tilde{T} = \tilde{t}|G = 0, X = \cdot)$ is continuous at 0 for any $\tilde{t} \in \{0,1\}^2$ and (ii) $E(Y_t|G = 0, \tilde{T} = \tilde{t}, X^* = \cdot) = E(Y_t|G = 0, \tilde{T} = \tilde{t}, X = \cdot)$ is continuous at 0 for any $t \in \{0,1\}$ and $\tilde{t} \in \{0,1\}^2$.*

Lemma 1 shows that, conditional on $G = 0$, the conditional probabilities of individual types and the conditional means of potential outcomes are continuous with respect to the observable assignment variable. We note that the smoothness of the conditional probabilities given $G = 0$ is satisfactory for our identification analysis, as shown in the later subsections. Therefore, according to the identification result in Dong (2015a), we conjecture that a LATE parameter is identified by the observable variables if $E(T^*|X^*)$ is discontinuous at the cutoff even in the presence of measurement error. The discontinuity is proved in the following subsection.

3.2 Discontinuity of the conditional treated probability

We show that the conditional probability of the observed treatment is discontinuous at the cutoff even in the presence of measurement error. We need the following assumptions to guarantee the discontinuity of the conditional treated probability.

Assumption 4 (*T*'s error). *$P(U_T = u_T|G = g, V = \cdot, X = \cdot)$ is continuous on $\mathcal{S}_{VX}$ for any $u_T = -1, 1$ and $g = 0, 1$.*

Assumption 5 (RD). *$E(T|X = 0^+) > E(T|X = 0^-)$.*

Assumption 6 (monotonicity). *$P(T_1 < T_0|X = x) = P(D|X = x) = 0$ for any $x \in \mathcal{S}_0$.*

Assumption 4 assumes that the conditional probabilities of the measurement error for T are continuous with respect to V and X. In particular, the conditional probabilities need to be continuous in X near the cutoff. Assumption 4 allows U_T to be arbitrarily dependent of the unobservables as long as the smoothness is satisfied. However, we need to note that this assumption may not be satisfied in the sharp RD design. For example, when $T = \mathbf{1}(X \geq 0)$, it must be satisfied that $P(U_T = 1|G = g, V = v, X = x) = 0$ for any $x \geq 0$. Therefore,

the probability can be discontinuous at the cutoff if it is strictly positive just below the cutoff. However, we stress that Assumption 4 may be satisfied even in the sharp design.

Assumptions 5 and 6 are standard identifying conditions in RD designs. Assumption 5 requires that the conditional probability of the true treatment is discontinuous at the cutoff, that is, it guarantees the presence of compliers. We note that the direction of the inequality is normalized. Assumption 6 is the same as (HTV2), which restricts that there are no defiers.

It is established that the conditional treated probability is discontinuous at the cutoff.

Lemma 2. *Under Assumptions 1–6, $E(T^*|X^* = \cdot)$ is discontinuous at the cutoff. Specifically, it holds that*

$$E(T^*|X^* = 0^+) - E(T^*|X^* = 0^-)$$
$$= (E(T|G = 0, X = 0^+) - E(T|G = 0, X = 0^-))P(G = 0|X^* = 0)$$
$$= P(C|G = 0, X = 0)P(G = 0|X^* = 0) > 0.$$

Lemma 2 shows that the conditional probability of the observable treatment is discontinuous at the cutoff even when measurement error exists. It implies that the research design can be a valid RD design regardless of the presence of measurement error. However, the jump size of $E(T^*|X^*)$ is not identical to that of $E(T|X)$ in general. For example, if T is independent of G conditional on X, the former is less than the latter, owing to the effect of G. In particular, even if the true design is sharp RD, $E(T^*|X^*)$ may not have sharp discontinuity.

The intuition behind the proof of Lemma 2 can be understood by the following observation. The law of iteratred expectations implies that

$$E(T^*|X^* = x^*) = E(T|X^* = x^*) + E(U_T|X^* = x^*)$$
$$= E(T|G = 0, X = x^*)P(G = 0|X^* = x^*)$$
$$\quad + E(T|G \neq 0, X^* = x^*)P(G \neq 0|X^* = x^*) + E(U_T|X^* = x^*).$$

In this expansion, $P(G = 0|X^* = \cdot)$, $P(G \neq 0|X^* = \cdot)$, and $E(U_T|X^* = \cdot)$ are continuous at the cutoff under the introduced assumptions. The key to show Lemma 2 is to prove that $E(T|G = 0, X = \cdot)$ is discontinuous at the cutoff under Assumption 5, while $E(T|G \neq 0, X^* = \cdot)$ is continuous even when Assumption 5 is satisfied. Intuitively, since the continuous error V smooths the conditional mean of T given $G \neq 0$, the conditional mean becomes continuous even

when the true treated probability is discontinuous at the cutoff. Similar findings are pointed out by Yu (2012), Davezies and Le Barbanchon (2014), and Yanagi (2014).

Our conjecture is that Lemmas 1 and 2 may imply that the RD estimand based on the observables identifies a LATE parameter. The following subsection establishes this conjecture.

3.3 Identifying a local weighted average treatment effect

We present the main theorem showing that τ^* identifies a local weighted average treatment effect (LWATE) parameter. To this end, we introduce a condition for the measurement error of Y.

Assumption 7 (Y's error). $E(U_Y|G = g, V = \cdot, X = \cdot)$ *is continuous on* $\mathcal{S}_{VX}$ *for any* $g = 0, 1$.

Assumption 7 restricts the smoothness of the conditional mean of the measurement error for Y. In particular, it requires that the conditional mean is continuous in X near the cutoff. It is similar to Assumption 4 for the measurement error on T. Assumption 7 allows the measurement error to be dependent of the true variables as long as the smoothness holds. We note that U_Y does not require the probability mass at zero unlike U_X, but we allow such a possibility.

We provide the main theorem for our identification.

Theorem 1. *Suppose that Assumptions 1–7 are satisfied. Then* τ^* *identifies a local weighted average treatment effect:*

$$\tau^* = \frac{E(Y|G = 0, X = 0^+) - E(Y|G = 0, X = 0^-)}{E(T|G = 0, X = 0^+) - E(T|G = 0, X = 0^-)}$$

$$= E(\beta|C, G = 0, X = 0)$$

$$= \int b \cdot \psi(b) dF_{\beta|C, X=0}(b),$$

where $\psi(\cdot)$ *is a weight function defined as*

$$\psi(b) := \frac{P(G = 0|C, X = 0, \beta = b)}{P(G = 0|C, X = 0)}$$

$$= \frac{P(G = 0|C, X = 0, \beta = b)}{\int P(G = 0|C, X = 0, \beta = b') dF_{\beta|C, X=0}(b')}.$$

Theorem 1 shows that τ^* identifies the average treatment effect for the compliers at the cutoff conditional on $G = 0$. It also shows that the parameter is equal to the weighted average of the treatment effect for the compliers at the cutoff. In this sense, the standard RD estimand

based on the possibly mismeasured observables identifies a local weighted average treatment effect. Remarkably, if X does not contain measurement error, then τ^* is equal to the standard RD LATE irrespective of the presence of the measurement error for Y and T. Accordingly, Theorem 1 implies that an accurate record of X is particularly important in RD applications.

The weight function $\psi(\cdot)$ depends on the conditional probabilities of no measurement error for the assignment variable. The weight function is not constant if the treatment effect β is dependent of the measurement error for X. Therefore, the LWATE is not identical to the standard RD LATE parameter in general. This result points out that econometricians should be careful in interpreting estimates in RD designs in which the assignment variable may contain nonclassical measurement error. However, the estimates continue to manifest an averaged treatment effect for a subpopulation.

The LWATE is equal to the standard RD LATE parameter if β and $\tilde{T}$ are independent of G conditional on $X = 0$. This result is consistent with the identification analysis developed in Battistin et al. (2009). This situation would be achieved when the assignment value for the individual is correctly reported regardless of his/her treatment effect and potential treatment status. In particular, the LWATE is identical to the standard RD LATE and the average treatment effect at the cutoff if the treatment effect is homogeneous across economic units at the cutoff, that is, if β is constant at the cutoff.

Although the LWATE is not identical to the standard RD LATE, the LWATE may tell us the sign of the standard RD LATE or that of the average treatment effect. For example, the sign of the LWATE is equal to that of the standard RD LATE if β for the compliers is positive (or negative) almost surely conditional on $X = 0$. More generally, we could examine the sign of the standard RD LATE under some conditions on β and $\psi(\cdot)$. Therefore, identifying the LWATE can provide econometricians considerably meaningful policy implications if economic theory can give information on β and/or $\psi(\cdot)$.

Theorem 1 also implies the intuition that the LWATE can be close to the standard RD LATE if the dependence between the measurement error U_X and the individual treatment effect β is weak. If the dependence between U_X and β is considerably weak, then $\psi(b)$ would be relatively insensitive to the value of b. Therefore, the LWATE estimate could be close to the standard RD LATE estimate in such cases.

4   Testable implication

This section discusses the testable implication led by the identification conditions introduced in the previous section.

The validity of RD designs is often examined by testing whether economic units can perfectly manipulate their assignment values (see, e.g., Lee and Lemieux, 2010). To investigate the presence of such manipulation, in the spirit of Lee (2008) and McCrary (2008), the analyses based on RD designs often examine the continuity of the density of the assignment variable.

The following theorem shows that testing the continuity of the observed assignment variable is meaningful even when the assignement variable contains nonclassical measurement error.

Theorem 2. *Under Assumptions 1, 2, and 3, $f_{X^*}(\cdot)$ is continuous on $\mathcal{S}_{X^*}$.*

Theorem 2 shows that the density of the observable assignment variable is continuous on the support under the identification conditions. In particular, the density must be continuous at the cutoff point. This implies that the discontinuous density of the assignment variable indicates the invalidity of some identification conditions. Accordingly, McCrary's continuous density test is justified even in the presence of measurement error for the assignment variable. Thus, econometricians should examine the continuity of the density of the observable assignment variable even when the assignment variable may contain measurement error. As usual, the test can be conducted with methdologies developed by McCrary (2008) and Otsu et al. (2013).

5   Conclusion

This paper develops an identification analysis in RD designs in which each observable can contain measurement error. Our analysis allows nonclassical measurement error in the sense that the measurement error can be arbitrarily dependent of the unobservables as long as the joint distribution satisfies a few plausible smoothness conditions.

We provide a set of identification conditions under which the standard RD estimand based on the possibly mismeasured observables identifies a local weighted average treatment effect (LWATE). The weight depends on the conditional probability that the assignment variable for the complier is correctly reported. The key identifying conditions are a smooth density condition and the positive probability that the assignment variable for the complier can be observed without measurement error near the cutoff point. We discuss situations under which

the LWATE provides meaningful information on the standard RD local average treatment effect (LATE) parameter.

We also show that the continuous density of the possibly mismeasured assignment variable is a testable implication for our identification. This justifies McCrary's continuous density test even when the assignment variable contains nonclassical measurement error.

Several extentions of the analysis in this paper would be interesting. First, identification of distributional treatment effects such as quantile treatment effects (Frandsen et al., 2012) may be possible in RD designs with nonclassical measurement error. Second, identification of a treatment effect derivative parameter (Dong and Lewbel, 2014) may be acheived even in our setting with stronger smoothness conditions. Third, we may be able to establish identification of a LATE parameter in RD designs with measurement error where the conditional treated probability is not discontinuous but kinked (Dong, 2015b).

A   Appendix

This appendix presents the proofs of the lemmas and the theorems stated in the main body of the paper and the technical lemmas.

A.1   Proofs of the lemmas and the theorems

A.1.1   Proof of Lemma 1

We first show statement (i). Consider any $\tilde{t} \in \{0, 1\}^2$. Bayes' theorem leads to

$$
\begin{aligned}
P(\tilde{T} = \tilde{t}|G = 0, X^* = x^*) &= P(\tilde{T} = \tilde{t}|G = 0, X = x^*) \\
&= \frac{P(G = 0|\tilde{T} = \tilde{t}, X = x^*)P(\tilde{T} = \tilde{t}|X = x^*)}{P(G = 0|X = x^*)}.
\end{aligned}
$$

Because $P(G = 0|\tilde{T} = \tilde{t}, X = \cdot)$ and $P(G = 0|X = \cdot)$ are continuous at 0 by Lemma A4 and $P(\tilde{T} = \tilde{t}|X = \cdot)$ is continuous at 0 by Lemma A3, statement (i) holds.

We next show statement (ii). For any $t \in \{0, 1\}$ and $\tilde{t} \in \{0, 1\}^2$, we have

$$
\begin{aligned}
&E(Y_t|G = 0, \tilde{T} = \tilde{t}, X^* = x^*) \\
=&E(Y_t|G = 0, \tilde{T} = \tilde{t}, X = x^*) \\
=& \int y_t dF_{Y_t|G=0, \tilde{T}=\tilde{t}, X=x^*}(y_t)
\end{aligned}
$$

17

$$= \int y_t \frac{P(G=0|\tilde{T}=\tilde{t}, X=x^*, Y_t=y_t)P(\tilde{T}=\tilde{t}|X=x^*, Y_t=y_t)f_{X|Y_t=y_t}(x^*)}{P(G=0|\tilde{T}=\tilde{t}, X=x^*)P(\tilde{T}=\tilde{t}|X=x^*)f_X(x^*)} dF_{Y_t}(y_t),$$

where the last equality follows from Bayes' theorem. We note that $P(G=0|\tilde{T}=\tilde{t}, X=\cdot, Y_t=\cdot)$ and $P(G=0|\tilde{T}=\tilde{t}, X=\cdot)$ are continuous by Lemma A4, $P(\tilde{T}=\tilde{t}|X=\cdot, Y_t=\cdot)$ and $P(\tilde{T}=\tilde{t}|X=\cdot)$ are continuous by Lemma A3, and $f_{X|Y_t=\cdot}(\cdot)$ and $f_X(\cdot)$ are continuous by Lemma A2. We thus have the desired result by Lemma A1.

$\square$

A.1.2 Proof of Lemma 2

It holds that

$$E(T^*|X^*=x^*) = E(T|X^*=x^*) + E(U_T|X^*=x^*).$$

$E(U_T|X^*=\cdot)$ is continuous at 0 by Lemma A7. The first equailty of the statement is thus established if we show that

$$
\begin{aligned}
& E(T|X^*=0^+) - E(T|X^*=0^-) \\
& = \big(E(T|G=0, X=0^+) - E(T|G=0, X=0^-)\big) P(G=0|X^*=0).
\end{aligned}
\tag{2}
$$

The law of iterated expectations implies that

$$
\begin{aligned}
& E(T|X^*=x^*) \\
& = E(T|G=0, X^*=x^*)P(G=0|X^*=x^*) + E(T|G\neq 0, X^*=x^*)P(G\neq 0|X^*=x^*) \\
& = E(T|G=0, X=x^*)P(G=0|X^*=x^*) + E(T|G\neq 0, X^*=x^*)P(G\neq 0|X^*=x^*).
\end{aligned}
$$

The continuity of $P(G=0|X^*=\cdot)$ and $P(G\neq 0|X^*=\cdot) = 1 - P(G=0|X^*=\cdot)$ at 0 is shown by Lemma A6. The continuity of $E(T|G\neq 0, X^*=\cdot)$ at 0 is shown by Lemma A9. Therefore, (2) holds.

Next, we show the second equality of the statement. To this end, it suffices to show that

$$E(T|G=0, X=0^+) - E(T|G=0, X=0^-) = P(C|G=0, X=0). \tag{3}$$

For any arbitrary small $\delta > 0$, we have

$$
\begin{aligned}
E(T|G = 0, X = \delta) =& P(T = 1|G = 0, X = \delta) \\
=& P(T_1 = 1|G = 0, X = \delta) \\
=& P(T_0 = 0, T_1 = 1|G = 0, X = \delta) + P(T_0 = 1, T_1 = 1|G = 0, X = \delta) \\
=& P(C|G = 0, X = \delta) + P(A|G = 0, X = \delta),
\end{aligned}
$$

and

$$
\begin{aligned}
E(T|G = 0, X = -\delta) =& P(T = 1|G = 0, X = -\delta) \\
=& P(T_0 = 1|G = 0, X = -\delta) \\
=& P(T_0 = 1, T_1 = 0|G = 0, X = -\delta) + P(T_0 = 1, T_1 = 1|G = 0, X = -\delta) \\
=& P(D|G = 0, X = -\delta) + P(A|G = 0, X = -\delta) \\
=& P(A|G = 0, X = -\delta),
\end{aligned}
$$

where the last equality follows from Assumption 6. Therefore, we have proven (3) by the continuity of $P(A|G = 0, X = \cdot)$ and $P(C|G = 0, X = \cdot)$ at 0 which is implied by Lemma 1.

We show the inequality in the statement. Bayes' theorem leads to

$$
\begin{aligned}
P(C|G = 0, X = 0) =& \frac{P(G = 0|C, X = 0)P(C|X = 0)}{P(G = 0|X = 0)} \\
=& \frac{P(G = 0|C, X = 0)(E(T|X = 0^+) - E(T|X = 0^-))}{P(G = 0|X = 0)},
\end{aligned}
$$

because the law of iterated expectations, Lemma A3, and Assumption 6 mean that

$$
\begin{aligned}
E(T|X = 0^+) - E(T|X = 0^-) =& P(T_1 = 1|X = 0^+) - P(T_0 = 1|X = 0^-) \\
=& P(C|X = 0^+) + P(A|X = 0^+) - (P(D|X = 0^-) + P(A|X = 0^-)) \\
=& P(C|X = 0).
\end{aligned}
$$

We thus have $P(C|G = 0, X = 0) > 0$ under Assumptions 1 and 5. We also note that

$$
P(G = 0|X^* = 0) = P(G = 0|X = 0)\frac{f_X(0)}{f_{X^*}(0)} > 0,
$$

by Assumptions 1 and 2. Therefore, we have the desired result.

$\square$

A.1.3 Proof of Theorem 1

We first note that the law of iterated expectations leads to

$$E(Y^*|X^* = x^*)$$
$$=E(Y|X^* = x^*) + E(U_Y|X^* = x^*)$$
$$=E(Y|G = 0, X = x^*)P(G = 0|X^* = x^*) + E(Y|G \neq 0, X^* = x^*)P(G \neq 0|X^* = x^*)$$
$$+ E(U_Y|X^* = x^*).$$

The continuity of $P(G = 0|X^* = \cdot)$, $P(G \neq 0|X^* = \cdot)$, $E(Y|G \neq 0, X^* = \cdot)$, and $E(U_Y|X^* = \cdot)$ at 0 is shown by Lemmas A6, A8, and A10. We thus have

$$E(Y^*|X^* = 0^+) - E(Y^*|X^* = 0^-)$$
$$=(E(Y|G = 0, X = 0^+) - E(Y|G = 0, X = 0^-))P(G = 0|X^* = 0).$$

Therefore, it holds that

$$\tau^* = \frac{E(Y|G = 0, X = 0^+) - E(Y|G = 0, X = 0^-)}{E(T|G = 0, X = 0^+) - E(T|G = 0, X = 0^-)}$$
$$= \frac{E(Y|G = 0, X = 0^+) - E(Y|G = 0, X = 0^-)}{P(C|G = 0, X = 0)}, \tag{4}$$

by Lemma 2. The first line shows the first equality in the statement.

To indicate the second equality of the statement, we show that the numerator in the last line of (4) equals to $E(Y_1 - Y_0|C, G = 0, X = 0)P(C|G = 0, X = 0)$. Remember that $\alpha := Y_0$, $\beta := Y_1 - Y_0$, and $Y = \alpha + \beta T$. For any arbitrarily small $\delta > 0$, we have

$$E(Y|G = 0, X = \delta)$$
$$=E(\alpha + \beta T|G = 0, X = \delta)$$
$$=E(\alpha + \beta T_1|G = 0, X = \delta)$$
$$=E(\alpha|G = 0, T_0 = 0, T_1 = 0, X = \delta)P(T_0 = 0, T_1 = 0|G = 0, X = \delta)$$
$$+ E(\alpha|G = 0, T_0 = 1, T_1 = 0, X = \delta)P(T_0 = 1, T_1 = 0|G = 0, X = \delta)$$

20

$$+ E(\alpha + \beta | G = 0, T_0 = 0, T_1 = 1, X = \delta)P(T_0 = 0, T_1 = 1 | G = 0, X = \delta)$$

$$+ E(\alpha + \beta | G = 0, T_0 = 1, T_1 = 1, X = \delta)P(T_0 = 1, T_1 = 1 | G = 0, X = \delta)$$

$$= E(\alpha | N, G = 0, X = \delta)P(N | G = 0, X = \delta) + E(\alpha | D, G = 0, X = \delta)P(D | G = 0, X = \delta)$$

$$+ E(\alpha + \beta | C, G = 0, X = \delta)P(C | G = 0, X = \delta) + E(\alpha + \beta | A, G = 0, X = \delta)P(A | G = 0, X = \delta)$$

$$= E(\alpha | N, G = 0, X = \delta)P(N | G = 0, X = \delta)$$

$$+ E(\alpha + \beta | C, G = 0, X = \delta)P(C | G = 0, X = \delta) + E(\alpha + \beta | A, G = 0, X = \delta)P(A | G = 0, X = \delta),$$

where the third equality follows from the law of iterated expectations and the last equality follows from Assumption 6. By the same procedure, it holds that

$$E(Y | G = 0, X = -\delta)$$

$$= E(\alpha | N, G = 0, X = -\delta)P(N | G = 0, X = -\delta)$$

$$+ E(\alpha | C, G = 0, X = -\delta)P(C | G = 0, X = -\delta) + E(\alpha + \beta | A, G = 0, X = -\delta)P(A | G = 0, X = -\delta).$$

Therefore, Lemma 1 leads to

$$E(Y | G = 0, X = 0^+) - E(Y | G = 0, X = 0^-) = E(\beta | C, G = 0, X = 0)P(C | G = 0, X = 0).$$

Now, we show the third equality of the statement. We have

$$E(\beta | C, G = 0, X = 0) = \int b \, dF_{\beta | C, G = 0, X = 0}(b)$$

$$= \int b \cdot \frac{P(G = 0 | C, X = 0, \beta = b)}{P(G = 0 | C, X = 0)} dF_{\beta | C, X = 0}(b)$$

$$= \int b \cdot \psi(b) dF_{\beta | C, X = 0}(b),$$

where the second equality follows from Bayes' theorem. We thus have the desired result.

□

A.1.4 Proof of Theorem 2

The statement follows immediately from Lemma A5.

□

A.2 Technical lemmas

We first provide a technical lemma to show the continuity of functions, which is repeatedly used in the proofs of the following lemmas.

Lemma A1. *Let $\phi(\cdot, \cdot) : \mathcal{S}_1 \times \mathcal{S}_2 \to \mathbb{R}$ be a function where $\mathcal{S}_1 \subset \mathbb{R}^{k_1}$ and $\mathcal{S}_2 \subset \mathbb{R}^{k_2}$ are compact sets with positive integers k_1 and k_2. Suppose that $\phi(\cdot, \cdot)$ is continuous on $\mathcal{S}_1 \times \mathcal{S}_2$. Further, suppose that $\alpha(\cdot)$ is a measure satisfying $0 < \alpha(\mathcal{S}_3) < \infty$ with a set $\mathcal{S}_3 \subset \mathcal{S}_2$. It holds that $m(\cdot) := \int_{\mathcal{S}_3} \phi(\cdot, u) d\alpha(u)$ is continuous on $\mathcal{S}_1$.*

Proof. For any $d_1, d_2 \in \mathcal{S}_1$, the modulus inequality leads to

$$
\begin{aligned}
|m(d_1) - m(d_2)| &= \left| \int_{\mathcal{S}_3} \phi(d_1, u) d\alpha(u) - \int_{\mathcal{S}_3} \phi(d_2, u) d\alpha(u) \right| \\
&\leq \int_{\mathcal{S}_3} |\phi(d_1, u) - \phi(d_2, u)| d\alpha(u).
\end{aligned}
\tag{5}
$$

Because $\phi(\cdot, \cdot)$ is continuous on the compact set $\mathcal{S}_1 \times \mathcal{S}_2$, $\phi(\cdot, \cdot)$ is uniformly continuous on $\mathcal{S}_1 \times \mathcal{S}_2$. It thus holds that, for any $\varepsilon > 0$, there exists some $\delta > 0$ such that $\|d_1 - d_2\| < \delta$ means $|\phi(d_1, u) - \phi(d_2, u)| < \varepsilon/\alpha(\mathcal{S}_3)$ for every $u \in \mathcal{S}_3$ (note that δ and ε do not depend on u). This fact and (5) imply that, for any $\varepsilon > 0$, there is some $\delta > 0$ such that $\|d_1 - d_2\| < \delta$ implies $|m(d_1) - m(d_2)| < \varepsilon$. $\square$

Suppose that W_1 is a sub-vector of (V, X, Y_0, Y_1) and that W_2 is the vector of the remaining elements of (V, X, Y_0, Y_1) excluded from W_1. For example, if $W_1 = (V, X)$, then $W_2 = (Y_0, Y_1)$ and we write $(V, X) = (v, x)$ and $(Y_0, Y_1) = (y_0, y_1)$ as $W_1 = w_1$ and $W_2 = w_2$, respectively, and we denote $\mathcal{S}_{VX}$ and $\mathcal{S}_{Y_0 Y_1}$ as $\mathcal{S}_{W_1}$ and $\mathcal{S}_{W_2}$, respectively. The supports $\mathcal{S}_{W_1}$ and $\mathcal{S}_{W_2}$ are compact sets under Assumption 3. We note that we allow the situation in which $W_1 = (V, X, Y_0, Y_1)$ and W_2 is empty. In this case, the integration with respect to W_2 developed below can be omitted.

Lemma A2. *Suppose that Assumptions 2 and 3 are satisfied. Then, (i) $f_{W_1 | \tilde{T} = \tilde{t}}(\cdot)$ and $f_{W_1}(\cdot)$ are continuous on $\mathcal{S}_{W_1}$ for any $\tilde{t} \in \{0, 1\}^2$ and (ii) $f_{W_1 | \tilde{T} = \tilde{t}, W_2 = \cdot}(\cdot)$ and $f_{W_1 | W_2 = \cdot}(\cdot)$ are continuous on $\mathcal{S}_{W_1 W_2}$ for any $\tilde{t} \in \{0, 1\}^2$.*

Proof. Consider any $\tilde{t} \in \{0, 1\}^2$. For statement (i), we note that

$$
f_{W_1 | \tilde{T} = \tilde{t}}(w_1) = \int f_{W_1 W_2 | \tilde{T} = \tilde{t}}(w_1, w_2) d\lambda(w_2).
$$

Thus, $f_{W_1|\tilde{T}=\tilde{t}}(\cdot)$ is continuous by Lemma A1 under Assumptions 2 and 3. We then note that

$$f_{W_1}(w_1) = \sum_{\tilde{t}\in\{0,1\}^2} f_{W_1|\tilde{T}=\tilde{t}}(w_1)P(\tilde{T}=\tilde{t}).$$

Thus, the continuity of $f_{W_1}(\cdot)$ also holds by the continuity of $f_{W_1|\tilde{T}=\tilde{t}}(\cdot)$.

For statement (ii), we note that

$$f_{W_1|\tilde{T}=\tilde{t},W_2=w_2}(w_1) = \frac{f_{VX\tilde{Y}|\tilde{T}=\tilde{t}}(v,x,\tilde{y})}{f_{W_2|\tilde{T}=\tilde{t}}(w_2)},$$
$$f_{W_1|W_2=w_2}(w_1) = \frac{f_{VX\tilde{Y}}(v,x,\tilde{y})}{f_{W_2}(w_2)}.$$

Therefore, statement (ii) is established by statement (i). □

Lemma A3. *Under Assumptions 2 and 3, $P(\tilde{T}=\tilde{t}|W_1=\cdot)$ is continuous on $\mathcal{S}_{W_1}$ for any $\tilde{t}\in\{0,1\}^2$.*

Proof. Consider any $\tilde{t}\in\{0,1\}^2$. Bayes' theorem leads to

$$P(\tilde{T}=\tilde{t}|W_1=w_1) = \frac{f_{W_1|\tilde{T}=\tilde{t}}(w_1)P(\tilde{T}=\tilde{t})}{f_{W_1}(w_1)}.$$

Therefore, the statement holds by Lemma A2. □

Lemma A4. *Suppose that Assumptions 1, 2, and 3 hold. Then, (i) $P(G=0|\tilde{T}=\tilde{t},W_1=\cdot)$ is continuous on $\mathcal{S}_{W_1}$ for any $\tilde{t}\in\{0,1\}^2$ and (ii) $P(G=0|W_1=\cdot)$ is continuous on $\mathcal{S}_{W_1}$.*

Proof. We first prove statement (i). Consider any $\tilde{t}\in\{0,1\}^2$. We note that

$$P(G=0|\tilde{T}=\tilde{t},W_1=w_1) = \int P(G=0|\tilde{T}=\tilde{t},W_1=w_1,W_2=w_2)f_{W_2|\tilde{T}=\tilde{t},W_1=w_1}(w_2)d\lambda(w_2).$$

Therefore, the statement holds by Lemmas A1 and A2 under Assumptions 1 and 3.

Statement (ii) follows immediately from statement (i) and Lemma A3 because

$$P(G=0|W_1=w_1) = \sum_{\tilde{t}\in\{0,1\}^2} P(G=0|\tilde{T}=\tilde{t},W_1=w_1)P(\tilde{T}=\tilde{t}|W_1=w_1).$$

□

Suppose that Q_1 is a sub-vector of (V,Y_0,Y_1) at least including V and that Q_2 is the vector

of the elements of (V, Y_0, Y_1) excluded from Q_1. For example, if $Q_1 = V$, then $Q_2 = (Y_0, Y_1)$ and we write $V = v$ and $(Y_0, Y_1) = (y_0, y_1)$ as $Q_1 = q_1$ and $Q_2 = q_2$, respectively, and we denote $\mathcal{S}_V$ and $\mathcal{S}_{Y_0 Y_1}$ as $\mathcal{S}_{Q_1}$ and $\mathcal{S}_{Q_2}$, respectively. The supports $\mathcal{S}_{Q_1}$ and $\mathcal{S}_{Q_2}$ are compact sets under Assumption 3. Note that we allow the situation in which $Q_1 = (V, Y_0, Y_1)$ and Q_2 is empty. In such a situation, the integration with respect to Q_2 developed below can be omitted.

Lemma A5. *Suppose that Assumptions 1, 2, and 3 hold. Then, (i) $f_{X^*|Q_1=\cdot}(\cdot)$ is continuous on $\mathcal{S}_{Q_1 X}$ and (ii) $f_{X^*}(\cdot)$ is continuous on $\mathcal{S}_X$.*

Proof. We first prove statement (i). It holds that

$$
\begin{aligned}
&f_{X^*|Q_1=q_1}(x^*) \\
=&f_{X^*|G=0,Q_1=q_1}(x^*)P(G=0|Q_1=q_1) + f_{X^*|G\neq 0,Q_1=q_1}(x^*)P(G\neq 0|Q_1=q_1) \\
=&f_{X|G=0,Q_1=q_1}(x^*)P(G=0|Q_1=q_1) + f_{X|G\neq 0,Q_1=q_1}(x^*-v)P(G\neq 0|Q_1=q_1) \\
=&P(G=0|X=x^*,Q_1=q_1)\frac{f_{X|Q_1=q_1}(x^*)}{P(G=0|Q_1=q_1)}P(G=0|Q_1=q_1) \\
&+ P(G\neq 0|X=x^*-v,Q_1=q_1)\frac{f_{X|Q_1=q_1}(x^*-v)}{P(G\neq 0|Q_1=q_1)}P(G\neq 0|Q_1=q_1) \\
=&P(G=0|X=x^*,Q_1=q_1)f_{X|Q_1=q_1}(x^*) + P(G\neq 0|X=x^*-v,Q_1=q_1)f_{X|Q_1=q_1}(x^*-v),
\end{aligned}
$$

where the third equality follows from Bayes' theorem. Therefore, we have statement (i) by Lemmas A2 and A4.

For statement (ii), we have

$$
f_{X^*}(x^*) = \int f_{X^*|Q_1=q_1}(x^*)dF_{Q_1}(q_1),
$$

and the continuity of $f_{X^*}(\cdot)$ holds by statement (i) and Lemma A1 under Assumption 3. $\square$

Lemma A6. *Suppose that Assumptions 1, 2, and 3 hold. Then, (i) $P(G=0|Q_1=\cdot, X^*=\cdot)$ is continuous on $\mathcal{S}_{Q_1 X}$ and (ii) $P(G=0|X^*=\cdot)$ is continuous on $\mathcal{S}_X$.*

Proof. For statement (i), Bayes' theorem leads to

$$
P(G=0|Q_1=q_1, X^*=x^*) = P(G=0|Q_1=q_1, X=x^*)\frac{f_{X|Q_1=q_1}(x^*)}{f_{X^*|Q_1=q_1}(x^*)}.
$$

Thus, the statement holds because $P(G=0|Q_1=\cdot, X=\cdot)$ is continuous by Lemma A4, $f_{X|Q_1=\cdot}(\cdot)$ is continuous by Lemma A2, and $f_{X^*|Q_1=\cdot}(\cdot)$ is continuous by Lemma A5.

For statement (ii), we note that

$$P(G = 0|X^* = x^*) = \int P(G = 0|Q_1 = q_1, X^* = x^*)dF_{Q_1|X^*=x^*}(q_1)$$

$$= \int P(G = 0|Q_1 = q_1, X^* = x^*)\frac{f_{X^*|Q_1=q_1}(x^*)}{f_{X^*}(x^*)}dF_{Q_1}(q_1),$$

where the second equality follows from Bayes' theorem. The statement holds by statement (i) and Lemmas A1 and A5 under Assumption 3. □

Lemma A7. *Under Assumptions 1, 2, 3, and 4, $E(U_T|X^* = \cdot)$ is continuous on S_X.*

Proof. We first note that

$$E(U_T|X^* = x^*) = P(U_T = 1|X^* = x^*) - P(U_T = -1|X^* = x^*).$$

We thus focus on showing the continuity of $P(U_T = u_T|X^* = \cdot)$ for $u_T \in \{-1, 1\}$.

For any $u_T \in \{-1, 1\}$, we have

$$P(U_T = u_T|X^* = x^*) = P(U_T = u_T|G = 0, X = x^*)P(G = 0|X^* = x^*)$$

$$+ P(U_T = u_T|G \neq 0, X^* = x^*)P(G \neq 0|X^* = x^*).$$

Because $P(G = 0|X^* = \cdot)$ and $P(G \neq 0|X^* = \cdot)$ are continuous by Lemma A6, we shall show the continuity of $P(U_T = u_T|G = 0, X = \cdot)$ and $P(U_T = u_T|G \neq 0, X^* = \cdot)$ below.

We have

$$P(U_T = u_T|G = 0, X = x^*)$$

$$= \int P(U_T = u_T|G = 0, V = v, X = x^*)dF_{V|G=0,X=x^*}(v)$$

$$= \int P(U_T = u_T|G = 0, V = v, X = x^*)\frac{P(G = 0|V = v, X = x^*)f_{X|V=v}(x^*)}{P(G = 0|X = x^*)f_X(x^*)}dF_V(v).$$

Because $P(G = 0|V = \cdot, X = \cdot)$ and $P(G = 0|X = \cdot)$ are continuous by Lemma A4 and $f_{X|V=.}(\cdot)$ and $f_X(\cdot)$ are continuous by Lemma A2, $P(U_T = u_T|G = 0, X = \cdot)$ is continuous by Lemma A1 under Assumptions 3 and 4. By the same procedure, we have

$$P(U_T = u_T|G \neq 0, X^* = x^*)$$

$$= \int P(U_T = u_T|G \neq 0, V = v, X = x^* - v)\frac{P(G = 0|V = v, X^* = x^*)f_{X^*|V=v}(x^*)}{P(G = 0|X^* = x^*)f_{X^*}(x^*)}dF_V(v).$$

Thus, $P(U_T = u_T | G \neq 0, X^* = \cdot)$ is continuous by Lemmas A1, A5, and A6 under Assumptions 3 and 4.

Therefore, we have shown the desired result. $\qquad\square$

Lemma A8. *Under Assumptions 1, 2, 3, and 7, $E(U_Y | X^* = \cdot)$ is continuous on $\mathcal{S}_X$.*

Proof. The law of iterated expectations leads to

$$
\begin{aligned}
E(U_Y | X^* = x^*) =\ & E(U_Y | G = 0, X = x^*) P(G = 0 | X^* = x^*) \\
& + E(U_Y | G \neq 0, X^* = x^*) P(G \neq 0 | X^* = x^*).
\end{aligned}
$$

Because $P(G = 0 | X^* = \cdot)$ and $P(G \neq 0 | X^* = \cdot)$ are continuous by Lemma A6, we focus on proving the continuity of $E(U_Y | G = 0, X = \cdot)$ and $E(U_Y | G \neq 0, X^* = \cdot)$.

The law of iterated expectations and Bayes' theorem imply

$$
\begin{aligned}
E(U_Y | G = 0, X = x^*) &= \int E(U_Y | G = 0, V = v, X = x^*) dF_{V|G=0, X=x^*}(v) \\
&= \int E(U_Y | G = 0, V = v, X = x^*) \frac{P(G = 0 | V = v, X = x^*) f_{X|V=v}(x^*)}{P(G = 0 | X = x^*) f_X(x^*)} dF_V(v).
\end{aligned}
$$

We thus have the continuity of $E(U_Y | G = 0, X = \cdot)$ by Lemmas A1, A2, and A4 under Assumptions 3 and 7. Similarly, we have

$$
\begin{aligned}
& E(U_Y | G \neq 0, X^* = x^*) \\
&= \int E(U_Y | G \neq 0, V = v, X^* = x^*) dF_{V|G \neq 0, X^* = x^*}(v) \\
&= \int E(U_Y | G \neq 0, V = v, X = x^* - v) \frac{P(G \neq 0 | V = v, X^* = x^*) f_{X^*|V=v}(x^*)}{P(G \neq 0 | X^* = x^*) f_{X^*}(x^*)} dF_V(v).
\end{aligned}
$$

Thus, $E(U_Y | G \neq 0, X^* = \cdot)$ is continuous by Lemmas A1, A5, and A6 under Assumptions 3 and 7.

Therefore, we have proven the statement. $\qquad\square$

Lemma A9. *Under Assumptions 1, 2, and 3, $E(T | G \neq 0, X^* = \cdot)$ is continuous on $\mathcal{S}_X$.*

Proof. Suppose that v_l and v_u are the lower and upper bounds of $\mathcal{S}_V$, respectively. The law of

iterated expectations and Bayes' theorem lead to

$$E(T|G \neq 0, X^* = x^*)$$

$$= \int E(T|G \neq 0, V = v, X^* = x^*)dF_{V|G\neq0,X^*=x^*}(v)$$

$$= \int E(T|G \neq 0, V = v, X = x^* - v)dF_{V|G\neq0,X^*=x^*}(v)$$

$$= \int_{v_l}^{x^*} E(T|G \neq 0, V = v, X = x^* - v)dF_{V|G\neq0,X^*=x^*}(v)$$

$$+ \int_{x^*}^{v_u} E(T|G \neq 0, V = v, X = x^* - v)dF_{V|G\neq0,X^*=x^*}(v)$$

$$= \int_{v_l}^{x^*} P(T_1 = 1|G \neq 0, V = v, X = x^* - v)\frac{P(G \neq 0|V = v, X^* = x^*)f_{X^*|V=v}(x^*)}{P(G \neq 0|X^* = x^*)f_{X^*}(x^*)}dF_V(v)$$

$$+ \int_{x^*}^{v_u} P(T_0 = 1|G \neq 0, V = v, X = x^* - v)\frac{P(G \neq 0|V = v, X^* = x^*)f_{X^*|V=v}(x^*)}{P(G \neq 0|X^* = x^*)f_{X^*}(x^*)}dF_V(v),$$

where the fourth equality follows from the definition of (T_0, T_1). We note that for any $d \in \{0, 1\}$

$$P(T_d = 1|G \neq 0, V = v, X = x^* - v)$$

$$= P(T_d = 1|V = v, X = x^* - v)\frac{P(G \neq 0|T_d = 1, V = v, X = x^* - v)}{P(G \neq 0|V = v, X = x^* - v)},$$

by Bayes' theorem. Therefore, the statement holds by Lemmas A1, A3, A4, A5, and A6 under Assumption 3. □

Lemma A10. *Under Assumptions 1, 2, and 3, $E(Y|G \neq 0, X^* = \cdot)$ is continuous on $\mathcal{S}_X$.*

Proof. By the same procedure in the proof of Lemma A9, we have

$$E(Y|G \neq 0, X^* = x^*)$$

$$= \int_{v_l}^{x^*} E(Y|G \neq 0, V = v, X = x^* - v)\frac{P(G \neq 0|V = v, X^* = x^*)f_{X^*|V=v}(x^*)}{P(G \neq 0|X^* = x^*)f_{X^*}(x^*)}dF_V(v)$$

$$+ \int_{x^*}^{v_u} P(Y|G \neq 0, V = v, X = x^* - v)\frac{P(G \neq 0|V = v, X^* = x^*)f_{X^*|V=v}(x^*)}{P(G \neq 0|X^* = x^*)f_{X^*}(x^*)}dF_V(v).$$

All the above terms except $E(Y|G \neq 0, V = v, X = x^* - v)$ are continuous by Lemmas A3, A4, A5, and A6. Thus, if we show the continuity of $E(Y|G \neq 0, V = v, X = x^* - v)$ for $v \in [v_l, x^*)$ and $v \in (x^*, v_u]$, we have the desired result by Lemma A1 under Assumption 3.

We expand $E(Y|G \neq 0, V = v, X = x^* - v)$. For $v \in [v_l, x^*)$, the law of iterated expectations

leads to

$$
E(Y|G \neq 0, V = v, X = x^* - v)
$$

$$
= \sum_{\tilde{t} \in \{0,1\}^2} E(Y|G \neq 0, \tilde{T} = \tilde{t}, V = v, X = x^* - v) P(\tilde{T} = \tilde{t}|G \neq 0, V = v, X = x^* - v)
$$

$$
= E(Y_0|N, G \neq 0, V = v, X = x^* - v) P(N|G \neq 0, V = v, X = x^* - v)
$$

$$
+ E(Y_1|A, G \neq 0, V = v, X = x^* - v) P(A|G \neq 0, V = v, X = x^* - v)
$$

$$
+ E(Y_1|C, G \neq 0, V = v, X = x^* - v) P(C|G \neq 0, V = v, X = x^* - v)
$$

$$
+ E(Y_0|D, G \neq 0, V = v, X = x^* - v) P(D|G \neq 0, V = v, X = x^* - v).
$$

For each $(t, R) \in \{(0, N), (1, A), (1, C), (0, D)\}$, we have

$$
E(Y_t|R, G \neq 0, V = v, X = x^* - v)
$$

$$
= \int y_t dF_{Y_t|R, G \neq 0, V = v, X = x^* - v}(y_t)
$$

$$
= \int y_t \frac{P(G \neq 0|R, V = v, X = x^* - v, Y_t = y_t) P(R|V = v, X = x^* - v, Y_t = y_t)}{P(G \neq 0|R, V = v, X = x^* - v) P(R|V = v, X = x^* - v)} dF_{Y_t}(y_t),
$$

which is continuous in (x^*, v) by Lemmas A1, A3, and A4 under Assumption 3. In addition, Bayes' theorem leads to

$$
P(R|G \neq 0, V = v, X = x^* - v)
$$

$$
= \frac{P(G \neq 0|R, V = v, X = x^* - v) P(R|V = v, X = x^* - v)}{P(G \neq 0|V = v, X = x^* - v)},
$$

which is continuous in (x^*, v) by Lemmas A3 and A4. Thus, $E(Y|G \neq 0, V = v, X = x^* - v)$ is continuous in (v, x^*) for $v \in [v_l, x^*)$. Similarly, we can show that $E(Y|G \neq 0, V = v, X = x^* - v)$ is continuous in (x^*, v) for $v \in (x^*, v_u]$.

Therefore, we have shown the desired result. $\qquad \square$

Acknowledgements

References

J. D. Angrist, G. W. Imbens, and D. B. Rubin. Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, 91(434):444–455, 1996. doi: 10.1080/01621459.1996.10476902. URL http://dx.doi.org/10.1080/01621459.1996.10476902.

Y. Arai and H. Ichimura. Optimal bandwidth selection for the fuzzy regression discontinuity estimator. mimeo, 2015a.

Y. Arai and H. Ichimura. Simultaneous selection of optimal bandwidths for the sharp regression discontinuity estimator. mimeo, 2015b.

E. Battistin and E. Rettore. Ineligibles and eligible non-participants as a double comparison group in regression-discontinuity designs. *Journal of Econometrics*, 142(2):715–730, 2008. doi: 10.1016/j.jeconom.2007.05.006. URL http://dx.doi.org/10.1016/j.jeconom.2007.05.006.

E. Battistin, A. Brugiavini, E. Rettore, and G. Weber. The retirement consumption puzzle: Evidence from a regression discontinuity approach. *The American Economic Review*, 99(5): 2209–2226, 2009. doi: 10.1257/aer.99.5.2209. URL http://dx.doi.org/10.1257/aer.99.5.2209.

J. Bound, C. Brown, and N. Mathiowetz. Measurement error in survey data. *Handbook of Econometrics*, 5:3705–3843, 2001. doi: 10.1016/S1573-4412(01)05012-7. URL http://dx.doi.org/10.1016/S1573-4412(01)05012-7.

S. Calonico, M. D. Cattaneo, and R. Titiunik. Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6):2295–2326, 2014. doi: 10.3982/ECTA11757. URL http://dx.doi.org/10.3982/ECTA11757.

D. Card and L. D. Shore-Sheppard. Using discontinuous eligibility rules to identify the effects of the federal medicaid expansions on low-income children. *The Review of Economics and Statistics*, 86(3):752–766, 2004. doi: 10.1162/0034653041811798. URL http://dx.doi.org/10.1162/0034653041811798.

D. Card, C. Dobkin, and N. Maestas. The impact of nearly universal insurance coverage on health care utilization: Evidence from Medicare. *The American Economic Review*, 98(5):

2242–2258, 2008. doi: 10.1257/aer.98.5.2242. URL http://dx.doi.org/10.1257/aer.98.5.2242.

D. Card, D. Lee, Z. Pei, and A. Weber. Inference on causal effects in a generalized regression kink design. *Forthcoming in Econometrica*, 2015.

L. Davezies and T. Le Barbanchon. Regression discontinuity design with continuous measurement error in the running variable. mimeo, 2014.

Y. Dong. An alternative assumption to identify LATE in regression discontinuity designs. mimeo, 2015a.

Y. Dong. Jump or kink? Identification of binary treatment regression discontinuity design without the discontinuity. mimeo, 2015b.

Y. Dong. Regression discontinuity applications with rounding errors in the running variable. *Journal of Applied Econometrics*, 30(3):422–446, 2015c. doi: 10.1002/jae.2369. URL http://dx.doi.org/10.1002/jae.2369.

Y. Dong and A. Lewbel. Identifying the effect of changing the policy threshold in regression discontinuity models. *The Review of Economics and Statistics*, 2014. doi: 10.1162/REST_a_00510. URL http://dx.doi.org/10.1162/REST_a_00510.

B. R. Frandsen, M. Frölich, and B. Melly. Quantile treatment effects in the regression discontinuity design. *Journal of Econometrics*, 168(2):382–395, 2012. doi: 10.1016/j.jeconom.2012.02.004. URL http://dx.doi.org/10.1016/j.jeconom.2012.02.004.

J. Hahn, P. Todd, and W. van der Klaauw. Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, 69(1):201–209, 2001. doi: 10.1111/1468-0262.00183. URL http://dx.doi.org/10.1111/1468-0262.00183.

P. Hullegie and T. J. Klein. The effect of private health insurance on medical care utilization and self-assessed health in Germany. *Health Economics*, 19(9):1048–1062, 2010. doi: 10.1002/hec.1642. URL http://dx.doi.org/10.1002/hec.1642.

G. Imbens and K. Kalyanaraman. Optimal bandwidth choice for the regression discontinuity estimator. *The Review of Economic Studies*, 79:933–959, 2012. doi: 10.1093/restud/rdr043. URL http://dx.doi.org/10.1093/restud/rdr043.

G. W. Imbens and T. Lemieux. Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, 142(2):615–635, 2008. doi: 10.1016/j.jeconom.2007.05.001. URL http://dx.doi.org/10.1016/j.jeconom.2007.05.001.

D. S. Lee. Randomized experiments from non-random selection in US house elections. *Journal of Econometrics*, 142(2):675–697, 2008. doi: 10.1016/j.jeconom.2007.05.004. URL http://dx.doi.org/10.1016/j.jeconom.2007.05.004.

D. S. Lee and T. Lemieux. Regression discontinuity designs in economics. *Journal of Economic Literature*, 48(2):281–355, 2010. doi: 10.1257/jel.48.2.281. URL http://dx.doi.org/10.1257/jel.48.2.281.

J. Ludwig and D. L. Miller. Does head start improve children's life chances? Evidence from a regression discontinuity design. *The Quarterly Journal of Economics*, 122(1):159–208, 2007. doi: 10.1162/qjec.122.1.159. URL http://dx.doi.org/10.1162/qjec.122.1.159.

J. McCrary. Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142(2):698–714, 2008. doi: 10.1016/j.jeconom.2007.05.005. URL http://dx.doi.org/10.1016/j.jeconom.2007.05.005.

T. Otsu, K.-L. Xu, and Y. Matsushita. Estimation and inference of discontinuity in density. *Journal of Business & Economic Statistics*, 31(4):507–524, 2013. doi: 10.1080/07350015.2013.818007. URL http://dx.doi.org/10.1080/07350015.2013.818007.

Z. Pei. Regression discontinuity design with measurement error in the assignment variable. mimeo, 2011.

D. W. Schanzenbach. Do school lunches contribute to childhood obesity? *Journal of Human Resources*, 44(3):684–709, 2009. doi: 10.3368/jhr.44.3.684. URL http://dx.doi.org/10.3368/jhr.44.3.684.

T. Yanagi. The effect of measurement error in the sharp regression discontinuity design. mimeo, 2014.

P. Yu. Identification of treatment effects in regression discontinuity designs with measurement error. mimeo, 2012.