

2016/25

External Validity in Fuzzy Regression Discontinuity Designs

MARINHO BERTANHA AND GUIDO W. IMBENS



50 YEARS OF
CORE
DISCUSSION PAPERS

CORE

Voie du Roman Pays 34, L1.03.01

Tel (32 10) 47 43 04

Fax (32 10) 47 43 01

Email: immaq-library@uclouvain.be

<http://www.uclouvain.be/en-44508.html>

External Validity in Fuzzy Regression Discontinuity Designs*

Marinho Bertanha[†] Guido W. Imbens[‡]

First Draft: November 2014

This Draft: February 2016

Abstract

Many empirical studies use Fuzzy Regression Discontinuity (FRD) designs to identify treatment effects when the receipt of treatment is potentially correlated to outcomes. Existing FRD methods identify the local average treatment effect (LATE) on the subpopulation of compliers with values of the forcing variable that are equal to the threshold. We develop methods that assess the plausibility of generalizing LATE to subpopulations other than compliers, and to subpopulations other than those with forcing variable equal to the threshold. Specifically, we focus on testing the equality of the distributions of potential outcomes for treated compliers and always-takers, and for untreated compliers and never-takers. We show that equality of these pairs of distributions implies that the expected outcome conditional on the forcing variable and the treatment status is continuous in the forcing variable at the threshold, for each of the two treatment regimes. Our main recommendation is that researchers, as a matter of routine, present graphs with estimates of these two conditional expectations in addition to graphs with estimates of the expected outcome conditional on the forcing variable alone. We illustrate our methods using data on the academic performance of students attending the summer school program in two large school districts in the US.

Keywords: Fuzzy Regression Discontinuity Designs, Treatment Effects, Potential Outcomes, Exogeneity, External Validity

*We are grateful for comments by Joshua Angrist, Brigham Frandsen and Arthur Lewbel on an earlier version of this paper. We are also grateful to Brian Jacob, Lars Lefgren, and Jordan Matsudaira for making available the data used in this paper.

[†]Center for Operations Research and Econometrics - Université catholique de Louvain, Voie du Roman Pays 34, 1348 Louvain-la-Neuve, Belgium. Email: marinho.bertanha@uclouvain.be. Website: www.uclouvain.be/marinho.bertanha. Bertanha gratefully acknowledges the financial support received as a B.F. Haley and E.S. Shaw Fellow from the Stanford Institute for Economic Policy Research.

[‡]The Applied Econometrics Professor, and Professor of Economics, and Graduate School of Business Trust Faculty Fellow 2013-2014, Graduate School of Business, Stanford University, and NBER. Email: imbens@stanford.edu. Website: faculty-gsb.stanford.edu/imbens.

1 Introduction

In empirical studies in economics and other social sciences researchers are often concerned about the possible endogeneity of key variables. Recently, many studies have used regression discontinuity (RD) designs, originated in the psychology literature (Thistlewaite and Campbell, 1960) to address these concerns. See Van Der Klaauw (2008), Imbens and Lemieux (2008), and Lee and Lemieux (2010) for recent surveys, and Cook (2008) for a historical perspective. In Sharp Regression Discontinuity settings, treatment is received if and only if the (pre-treatment) forcing variable is on the right (or left) of a fixed policy threshold. In Fuzzy Regression Discontinuity (FRD) settings, the probability of receiving the treatment changes discontinuously at this threshold. Hahn, Todd, and Van Der Klaauw (2001, HTV from hereon) show that under mild smoothness conditions the local average treatment effect on the subpopulation of compliers with forcing variable equal to the threshold is identified, and propose an estimator. However, researchers are often interested in average treatment effects for subpopulations other than compliers. In the current paper we study assumptions under which the average effect for compliers at the threshold is generalizable to other subpopulations.

The FRD estimand is analogous to the binary instrumental variable (IV) estimand in the conventional treatment effects setting with endogeneity of treatment receipt of Imbens and Angrist (1994). In models assuming constant treatment effects across individuals, endogeneity of treatment receipt is typically assessed with Hausman tests (Hausman, 1978). It consists of comparing the IV estimator with the ordinary least squares (OLS) estimator because both are consistent for the treatment effect under exogeneity. However, constant treatment effects is an unrealistic assumption in many applications. For models assuming varying treatment effects across individuals, Angrist (2004) discusses the interpretation of the Hausman test and proposes a more attractive test for exogeneity of treatment receipt behavior. In this paper, we extend the Hausman and Angrist tests to the FRD setting, and we propose an alternative test.

In a Hausman test, the OLS estimator is compared to the IV estimator. In the regression discontinuity context, the analogue of the OLS estimator identifies the average treatment effect at the threshold under exogeneity or unconfoundedness assumptions, whereas the second estimator identifies the average treatment effect at the threshold for

the subpopulation of compliers under the FRD identification assumptions. We show that in settings with heterogeneous effects, both in the standard IV setting and in the fuzzy regression discontinuity setting, this comparison is difficult to interpret, and thus not attractive. Angrist’s improved version of this test compares the estimates of the average treatment effect for compliers with the difference between the average outcome for always-takers and the average outcome for never-takers. Here we use the compliance terminology of Imbens and Angrist (1994), extended to the FRD case by HTV. Our first point is that, instead of testing the single equality of either the Hausman or Angrist approaches, researchers should test jointly a pair of restrictions on the potential outcomes of compliers, always-takers, and never-takers. The pair of restrictions are: (i) the equality between the average outcome of always-takers and treated compliers at the threshold; (ii) the equality between the average outcome for never-takers and untreated compliers at the threshold. If both equalities hold, we argue that this lends support to the hypothesis that one can extrapolate the average effect for compliers to other subpopulations at the threshold; that is, it is more likely that the estimates have external validity. Moreover, external validity also allows for identification of treatment effects on subpopulations with values of the forcing variable that are different than the threshold.

We show that this pair of restrictions is equivalent to continuity of the conditional expectation of the outcome as a function of the forcing variable at the threshold, separately for each of the two treatment regimes. Currently, researchers applying fuzzy regression discontinuity designs typically present a graph containing the estimated conditional expectation of the outcome given the forcing variable to illustrate the identification strategy. One of our two main recommendations is that researchers present two additional graphs: the estimated conditional expectation of the outcome as a function of the forcing variable, separately for the subpopulation of treated and untreated individuals, and adjusted for exogenous covariates. A discontinuity at the threshold in either or both of these two additional graphs provides evidence against exogeneity or unconfoundedness assumptions, and, thereby, evidence against external validity of the estimates. The second recommendation is to test the null hypothesis that the two restrictions hold.

In the recent causal literature there have been a number of alternative proposals for assessing and improving the external validity of regression discontinuity estimates and instrumental variables estimates more broadly. The focus and the applicability of

the methods varies, with some applying only to fuzzy regression discontinuity designs, and some requiring the presence of additional exogenous covariates. In an interesting approach, Dong and Lewbel (2014) point out that at the threshold one cannot only estimate the magnitude of the discontinuity, but also the change in the first (or even higher order) derivatives of the regression function. Under smoothness of the two conditional mean functions, knowledge of the higher order derivatives would allow the researcher to extrapolate at least locally away from the threshold. The Dong and Lewbel methods apply both in the sharp and in the fuzzy regression discontinuity design, without requiring the presence of additional covariates. Angrist and Rokkanen (2012) exploit the presence of additional exogenous covariates, and assess whether conditional on these covariates the correlation between the forcing variable and the outcome vanishes. This would imply that the assignment can be thought of as good as random conditional on the additional covariates, and it would allow for extrapolation away from the threshold. The Angrist-Rokkanen methods apply both in the case of sharp and fuzzy regression discontinuity designs. Angrist and Fernandez-Val (2010), in a conventional instrumental variables setting that can be generalized to the fuzzy regression discontinuity setting, consider extrapolating local average treatment effects by exploiting the presence of other exogenous covariates. Their key assumption, which Angrist and Fernandez-Val label “conditional effect ignorability,” is that conditional on these additional covariates the average effect for compliers is identical to the average effect for never-takers and always-takers. These three approaches are all complementary to ours. In contrast to the Dong-Lewbel and Angrist-Rokkanen methods our approach requires the regression discontinuity design to be fuzzy rather than sharp. Unlike the Angrist-Rokkanen and Angrist-Fernandez-Val methods our approach does not require additional covariates, although it is more powerful if covariates are available.

The remainder of this paper is organized as follows. In Section 2, we introduce the notation for the FRD setting, define parameters of interest, and state regularity and identification conditions. Section 3 discusses methods to assess the plausibility of exogeneity and external validity in the simplest setting without additional pretreatment variables. In Section 4 we extend the analysis to settings with additional covariates. Section 5 illustrates our proposed methods using two data sets previously used to estimate the effect of summer school programs on academic performance; the first data set was

originally analyzed by Jacob and Lefgren (2004), and the second data set was previously studied by Matsudaira (2008). Section 6 concludes. An appendix presents proofs for the results in this paper.

2 Setup

In this section, we set up the framework for analyzing fuzzy regression discontinuity designs. We follow the widely used Rubin Causal Model or Potential Outcome setup, extended to the fuzzy regression discontinuity design by HTV. Recent theoretical contributions include Porter (2003), McCrary (2008), Imbens and Kalyanaraman (2012), Calonico, Cattaneo and Titiunik (2014ab), Dong (2014), Dong and Lewbel (2014), Bertanha (2014), and Gelman and Imbens (2014). Influential applications include Black (1999), Berk and Rauma (1983), Angrist and Lavy (1999), Van Der Klaauw (2002), Jacob and Lefgren (2004), Battistin and Rettore (2008), Lee (2008), Lalive (2008), Lee, Moretti and Butler (2004), and Matsudaira (2008). Surveys of regression discontinuity methods include Van Der Klaauw (2008), Imbens and Lemieux (2008), and Lee and Lemieux (2010).

2.1 Notation

We consider a setting where we have a random sample from a large population, with the units in the sample indexed by $i = 1, \dots, N$. Let W_i^{obs} be the binary treatment of interest, which may be endogenous. We are interested in the causal effect of the treatment on an outcome Y_i . Let $Y_i(0)$ and $Y_i(1)$ denote the potential outcomes. For units in the sample the realized and observed outcome is

$$Y_i^{\text{obs}} = \begin{cases} Y_i(0) & \text{if } W_i^{\text{obs}} = 0, \\ Y_i(1) & \text{if } W_i^{\text{obs}} = 1. \end{cases}$$

In addition, we observe a covariate for each unit in the sample, which we refer to as the forcing variable, denoted by X_i . The support of the distribution of the forcing variable X_i is denoted by $\mathbb{X}$. The forcing variable is a fixed characteristic of the individual that is not affected by the treatment. In Section 4, we also consider the case where additional pre-treatment variables are observed. The incentives to participate in the treatment change with the threshold, that is, a particular value for the forcing variable denoted by

T^* . We view the value of this threshold as a quantity that can be potentially changed by the policy maker. We assume that changing the value of the threshold does not change the values of the potential outcomes.

Our setup for the determination of the treatment received is slightly different from that in HTV, and more in line with Dong (2015) and Dong and Lewbel (2015). We explicitly allow potential treatment participation to vary with the threshold. Let $W_i : \mathbb{X} \mapsto \{0, 1\}$ be the potential outcome denoting whether individual i would participate in the treatment if the threshold were set equal to t . Consider the application we are using in this paper, that is, the assignment to summer school. The forcing variable X_i is a test-score prior to the summer program. The school district sets a threshold t , and students scoring below the threshold, that is, students with a score X_i such that $X_i \leq t$ are in general required to participate in the summer program. If everyone strictly followed the district's policy, we would have $W_i(t) = \mathbf{1}_{X_i \leq t}$, and we could use sharp regression discontinuity design methods. However, it need not be the case that all students follow this rule.

We define four types of individuals for a given policy threshold T^* . They are denoted $G_i = a$ (always-takers), $G_i = c$ (compliers), $G_i = n$ (never-takers), and $G_i = d$ (defiers), depending on the behavior of $W_i(t)$ locally to the actual threshold T^* . For a small fixed $\varepsilon > 0$

$$G_i = \begin{cases} a & \text{if } W_i(t) = 1 & \text{for } \forall t \in [T^* - \varepsilon, T^* + \varepsilon], \\ c & \text{if } W_i(t) = \mathbf{1}_{X_i \leq t} & \text{for } \forall t \in [T^* - \varepsilon, T^* + \varepsilon], \\ n & \text{if } W_i(t) = 0 & \text{for } \forall t \in [T^* - \varepsilon, T^* + \varepsilon], \\ d & \text{otherwise.} \end{cases} \quad (2.1)$$

where $\text{pr}(G_i = d) = 0$ is assumed. The observed treatment status is $W_i^{\text{obs}} = W_i(T^*)$.

We focus on conditional expectations given X_i , and look at conditioning sets that shrink to $X_i = x \in \mathbb{X}$. In our notation, a conditional expectation given $\{X_i = x_+\}$ denotes the limit of that conditional expectation given $\{x < X_i < x + h\}$ as $h \downarrow 0$. Similarly, a conditional expectation given $\{X_i = x_-\}$ denotes the limit of that conditional expectation given $\{x - h < X_i \leq x\}$ as $h \downarrow 0$. Finally, a conditional expectation given $\{X_i = x_\pm\}$ denotes the limit of that conditional expectation given $\{x - h < X_i < x + h\}$ as $h \downarrow 0$. For example, $\mathbb{E}[Y_i^{\text{obs}} | X_i = T_+^*] = \lim_{h \downarrow 0} \mathbb{E}[Y_i^{\text{obs}} | T^* < X_i < T^* + h]$, $\mathbb{E}[Y_i^{\text{obs}} | X_i = T_-^*] = \lim_{h \downarrow 0} \mathbb{E}[Y_i^{\text{obs}} | T^* - h < X_i \leq T^*]$, and $\mathbb{E}[Y_i^{\text{obs}} | X_i = T_\pm^*] = \lim_{h \downarrow 0} \mathbb{E}[Y_i^{\text{obs}} | T^* - h < X_i < T^* + h]$.

2.2 Causal Estimands

We define the main causal estimands considered in this paper. In RD settings, identification is typically obtained only at the cutoff value. The average effect for all individuals with $X_i = x$ is denoted

$$\tau(x) = \mathbb{E}[Y_i(1) - Y_i(0)|X_i = x].$$

The value of $\tau(x)$ at $x = T^*$, $\tau(T^*)$, plays a particularly important role. The overall average effect is

$$\tau^{\text{ate}} = \mathbb{E}[Y_i(1) - Y_i(0)] = \mathbb{E}[\tau(X_i)]$$

In addition, we define the local average treatment effect (LATE) as the average effect for compliers with $X_i = T^*$

$$\tau^{\text{late}} = \mathbb{E}[Y_i(1) - Y_i(0)|G_i = c, X_i = T^*]$$

Below, we present standard regression discontinuity assumptions that are sufficient for identification of τ^{late} . Our approach to external validity looks at the relationships between these estimands, and assumptions that allow for identification of these quantities. We wish to assess the possibility of generalizing τ^{late} to the identification of $\tau(x)$ and τ^{ate} .

2.3 FRD and Exogenous Estimands

Next, we define the probability limits of estimators that rely on either FRD assumptions, or on exogeneity/unconfoundedness type of assumptions. First, the typical FRD estimator is obtained by dividing the estimated difference in the side-limits of the average of Y_i at the cutoff by the estimated difference in the side-limits of the average of W_i at the cutoff. Such estimator is consistent for

$$\tau^{\text{frd}} = \frac{\mathbb{E}[Y_i^{\text{obs}}|X_i = T_+^*] - \mathbb{E}[Y_i^{\text{obs}}|X_i = T_-^*]}{\mathbb{E}[W_i^{\text{obs}}|X_i = T_+^*] - \mathbb{E}[W_i^{\text{obs}}|X_i = T_-^*]}$$

Second, the estimator that relies on local exogeneity of the treatment status conditional on $X_i = T^*$ simply compares treated and untreated units for the subpopulation with forcing variable close to the threshold. This estimator is consistent for

$$\tau^{\text{exo}} = \mathbb{E}[Y_i^{\text{obs}}|W_i^{\text{obs}} = 1, X_i = T_\pm^*] - \mathbb{E}[Y_i^{\text{obs}}|W_i^{\text{obs}} = 0, X_i = T_\pm^*].$$

2.4 Assumptions

We make the following assumptions, which are similar to standard assumptions in the regression discontinuity literature. See HTV, Dong and Lewbel (2014), Imbens and Lemieux (2008), Porter (2003), and Calonico, Cattaneo and Titiunik (2014).

Assumption 1. *The sample is a random sample from a large population.*

Assumption 2. *For every $w \in \{0, 1\}$ and $g \in \{n, a, c\}$: (i) the conditional distribution of $Y_i(w)|X_i = x$, $G_i = g$ is continuous in x at all $x \in \mathbb{X}$, and all of its moments are finite; (ii) the conditional probability $\text{pr}(G_i = g|X_i = x)$ is continuous in x for all $x \in \mathbb{X}$.*

Assumption 3. *The distribution of X_i has a continuous density function $f_X(x)$ and support $\mathbb{X}$. The probability $\text{pr}(W_i^{\text{obs}} = 1|X_i = x)$ is strictly between zero and one at $x = T^*$.*

Assumption 4. *The actual threshold is T^* , with the density $f_X(T^*) > 0$ and $\text{pr}(G_i = c|X_i = T^*) > 0$.*

Assumption 2 implies continuity of the expectation of potential outcomes conditional on the compliance type and the forcing variable as a function of the forcing variable; as well as, continuity of the probability of the compliance type conditional on the forcing variable as a function of the forcing variable. These two implications along with Assumptions 1, 3, and 4 are sufficient for the identification of τ^{late} . See Dong (2015). The following result is due to HTV, and it is stated without proof.

Lemma 1. *Suppose Assumptions 1-4 hold. Then*

$$\tau^{\text{frd}} = \tau^{\text{late}}.$$

3 Testing for External Validity in FRD

In this section, we briefly discuss two existing approaches in the literature used for testing the exogeneity of treatment receipt. We show that the null hypothesis of each of these approaches is one linear combination of two testable restrictions implied by the exogeneity of treatment receipt behavior (or compliance type) with respect to potential

outcomes. We then propose a more attractive procedure that consists of testing both of these restrictions jointly. Exogeneity of treatment receipt behavior is also shown to imply identification of average treatment effects for subpopulations with forcing variable different than the threshold.

3.1 The Hausman Test in FRD

To set the stage, let us first consider the conventional instrumental variables setting (i.e. non-regression discontinuity), with a single binary endogenous regressor and a single binary instrument. In a constant coefficient model we have

$$Y_i^{\text{obs}} = \alpha + \mu \cdot W_i^{\text{obs}} + \varepsilon_i,$$

with a binary instrument Z_i . The instrument Z_i denotes eligibility for treatment of individual i and is assumed to be exogenous with respect to potential outcomes. The Hausman test compares the OLS estimator $\hat{\mu}^{\text{ols}}$ with the IV estimator $\hat{\mu}^{\text{iv}}$ using Z_i as an instrument for W_i^{obs} . In large samples the test compares the two quantities,

$$\text{plim}(\hat{\mu}^{\text{ols}}) = \mathbb{E}[Y_i^{\text{obs}} | W_i^{\text{obs}} = 1] - \mathbb{E}[Y_i^{\text{obs}} | W_i^{\text{obs}} = 0],$$

with

$$\text{plim}(\hat{\mu}^{\text{iv}}) = \frac{\mathbb{E}[Y_i^{\text{obs}} | Z_i = 1] - \mathbb{E}[Y_i^{\text{obs}} | Z_i = 0]}{\mathbb{E}[W_i^{\text{obs}} | Z_i = 1] - \mathbb{E}[W_i^{\text{obs}} | Z_i = 0]}.$$

The first result explores the interpretation of the restriction $\text{plim}(\hat{\mu}^{\text{ols}}) = \text{plim}(\hat{\mu}^{\text{iv}})$ in the LATE setting with heterogenous treatments, under monotonicity, exogeneity of the instrument, and the exclusion restriction (Imbens and Angrist, 1994; Angrist, Imbens and Rubin, 1996). Let $G_i \in \{n, c, a\}$ denote the compliance type for unit i in the standard IV setting, let $\pi_n = \text{pr}(G_i = n)$, $\pi_c = \text{pr}(G_i = c)$, and $\pi_a = \text{pr}(G_i = a)$ denote their population shares, and let $p_z = \text{pr}(Z_i = 1)$ denote the population probability that the instrument takes on the value one. Suppose that Z_i is independent of $(Y_i(0), Y_i(1), G_i)$. Then the Hausman null hypothesis

$$H_0^{H, \text{iv}} : \text{plim}(\hat{\mu}^{\text{ols}}) = \text{plim}(\hat{\mu}^{\text{iv}}),$$

is equivalent to the null hypothesis

$$H_0^{H, \text{iv}'} : \frac{\pi_a}{\pi_a + \pi_c \cdot p_z} \cdot \left(\mathbb{E}[Y_i(1) | G_i = a] - \mathbb{E}[Y_i(1) | G_i = c] \right) \quad (3.1)$$

$$= \frac{\pi_n}{\pi_n + \pi_c \cdot (1 - p_z)} \cdot \left(\mathbb{E}[Y_i(0)|G_i = n] - \mathbb{E}[Y_i(0)|G_i = c] \right).$$

In the FRD setting, the natural analogue of the Hausman test is the null hypothesis that the FRD estimand is identical to the local comparison of treated and control units:

$$H_0^{H,\text{frd}} : \tau^{\text{exo}} = \tau^{\text{frd}}. \quad (3.2)$$

The following lemma shows the equivalence between the equality $\tau^{\text{exo}} = \tau^{\text{frd}}$ and an equality of weighted averages of potential outcomes of compliers, always-takers and never-takers. Now, let $\pi_n = \text{pr}(G_i = n|X_i = T^*)$, $\pi_a = \text{pr}(G_i = a|X_i = T^*)$, and $\pi_c = \text{pr}(G_i = c|X_i = T^*)$ be shorthand for the compliance-type probabilities at the threshold.

Lemma 2. *Suppose that Assumptions 1-4 hold. Then the null hypothesis*

$$H_0^{H,\text{frd}} : \tau^{\text{exo}} = \tau^{\text{frd}},$$

is equivalent to the null hypothesis

$$\begin{aligned} H_0^{H,\text{frd}'} : & \frac{\pi_a}{\pi_a + \pi_c/2} \cdot \left(\mathbb{E}[Y_i(1)|G_i = a, X_i = T^*] - \mathbb{E}[Y_i(1)|G_i = c, X_i = T^*] \right) \\ &= \frac{\pi_n}{\pi_n + \pi_c/2} \cdot \left(\mathbb{E}[Y_i(0)|G_i = n, X_i = T^*] - \mathbb{E}[Y_i(0)|G_i = c, X_i = T^*] \right). \end{aligned}$$

The slight change in the weights relative to the standard IV case comes from assuming that the distribution of the forcing variable is continuous. Hence, the probability of being to the left or to the right of the threshold, conditional on being very close to the threshold, is equal to 1/2. Note that, if the design is sharp rather than fuzzy ($\pi_a = \pi_n = 0$), then $H_0^{H,\text{frd}}$ is always true; that is, a simple difference of means between treated and untreated near the cutoff yields the treatment effect on the entire population local to the cutoff.

COMMENT 1: The first point of the paper is that the Hausman null hypothesis, which was originally developed for the setting with constant treatment effects, is unattractive in the setting with heterogeneous treatment effects (see $H_0^{H,\text{iv}'}$ or $H_0^{H,\text{frd}'}$). The Hausman null hypothesis allows for differences in average outcomes between always-takers and compliers with the treatment, and between never-takers and compliers without the treatment, as long as the particular weighted average of those differences cancel out. The weights depend on the shares of the compliance types and do not have a substantive interpretation. $\square$

3.2 The Angrist (2004) Restrictions

Angrist (2004) considers the conventional instrumental variables setting with heterogeneous treatment effects and suggests testing the following null hypothesis of exogeneity

$$\begin{aligned} H_0^{A,iv} : \quad & \frac{\mathbb{E}[Y_i^{\text{obs}}|Z_i = 1] - \mathbb{E}[Y_i^{\text{obs}}|Z_i = 0]}{\mathbb{E}[W_i^{\text{obs}}|Z_i = 1] - \mathbb{E}[W_i^{\text{obs}}|Z_i = 0]} \\ & = \mathbb{E}[Y_i^{\text{obs}}|W_i^{\text{obs}} = 1, Z_i = 0] - \mathbb{E}[Y_i^{\text{obs}}|W_i^{\text{obs}} = 0, Z_i = 1]. \end{aligned}$$

This is equivalent to the null hypothesis

$$H_0^{A,iv'} : \mathbb{E}[Y_i(1) - Y_i(0)|G_i = c] = \mathbb{E}[Y_i(1)|G_i = a] - \mathbb{E}[Y_i(0)|G_i = n],$$

or, to stress the difference with the Hausman test,

$$H_0^{A,iv''} : \mathbb{E}[Y_i(1)|G_i = a] - \mathbb{E}[Y_i(1)|G_i = c] = \mathbb{E}[Y_i(0)|G_i = n] - \mathbb{E}[Y_i(0)|G_i = c].$$

Compared to the Hausman null hypothesis $H^{H,iv'}$, in the Angrist null hypothesis $H^{A,iv''}$ the weights have been dropped, and we simply compare the average effect for compliers to the difference in average outcomes for always-takers and never-takers, which appears to be a more natural and attractive comparison.

In the fuzzy regression discontinuity setting, the natural analogue of Angrist's null hypothesis is

$$\begin{aligned} H_0^{A,\text{frd}} : \quad & \frac{\mathbb{E}[Y_i^{\text{obs}}|X_i = T_+^*] - \mathbb{E}[Y_i^{\text{obs}}|X_i = T_-^*]}{\mathbb{E}[W_i^{\text{obs}}|X_i = T_+^*] - \mathbb{E}[W_i^{\text{obs}}|X_i = T_-^*]} \\ & = \mathbb{E}[Y_i^{\text{obs}}|W_i^{\text{obs}} = 1, X_i = T_+^*] - \mathbb{E}[Y_i^{\text{obs}}|W_i^{\text{obs}} = 0, X_i = T_-^*]. \end{aligned}$$

Lemma 3. *Suppose that Assumptions 1-4 hold. Then the null hypothesis $H_0^{A,\text{frd}}$ is equivalent to the null hypotheses*

$$\begin{aligned} H_0^{A,\text{frd}'} : \quad & \mathbb{E}[Y_i(1)|G_i = c, X_i = T^*] - \mathbb{E}[Y_i(0)|G_i = c, X_i = T^*] \\ & = \mathbb{E}[Y_i(1)|G_i = a, X_i = T^*] - \mathbb{E}[Y_i(0)|G_i = n, X_i = T^*], \end{aligned}$$

or

$$\begin{aligned} H_0^{A,\text{frd}''} : \quad & \mathbb{E}[Y_i(1)|G_i = a, X_i = T^*] - \mathbb{E}[Y_i(1)|G_i = c, X_i = T^*] \\ & = \mathbb{E}[Y_i(0)|G_i = n, X_i = T^*] - \mathbb{E}[Y_i(0)|G_i = c, X_i = T^*]. \end{aligned}$$

Angrist (2004) motivates his proposed test primarily by arguments concerning the statistical power. In our view his test is also conceptually more attractive than the Hausman test because it has a more intuitive interpretation.

3.3 External Validity

Under the assumptions we laid out, regression discontinuity estimates are valid locally, where the qualifier “locally” limits their generalizability in two aspects. These estimates are valid only for units with the value of the forcing variable X_i close to the threshold T^* , and they are valid only for compliers. Often we are interested in causal effects also for non-compliers, and for units with values for the forcing variable away from the threshold. In this section, we explore the assumptions that validate extrapolation (external validity) of treatment effects to individuals of different compliance types and to individuals with different values of the forcing variable.

The following assumption allows for the generalization to other compliance types within the subpopulation of units with values of the forcing variable close to the threshold.

Assumption 5. (LOCAL EXTERNAL VALIDITY)

The study has local external validity if,

$$G_i \perp\!\!\!\perp (Y_i(0), Y_i(1)) \mid X_i. \quad (3.3)$$

This assumption is related to what Angrist (2004) calls the “no-selection bias” condition in the conventional instrumental variables setting, and what Dong and Lewbel call the “local invariance assumption”.

The last assumption restricts the heterogeneity of the potential outcome functions $W_i(t)$ to be fully described by the three compliance types in the definition of G_i (Equation 2.1). That in combination and Assumption 5 allow for identification of treatment effects on units with values of the forcing variable away from the threshold.

Assumption 6. $W_i(t)$ takes on three forms. Either $W_i(t) = 0$ for all t , $W_i(t) = 1$ for all t , or $W_i(t) = \mathbf{1}_{X_i \leq t}$.

Given this assumption, individuals with $W_i(t) = 0$ for all t are never-takers, individuals with $W_i(t) = 1$ for all t are always-takers, and individuals with $W_i(t) = \mathbf{1}_{X_i < t}$ are compliers.

Lemma 4. *Suppose Assumptions 1-5 hold. Then $\tau(T^*)$ is identified,*

$$\tau(T^*) = \mathbb{E}[Y_i^{\text{obs}} | W_i^{\text{obs}} = 1, X_i = T^*_{\pm}] - \mathbb{E}[Y_i^{\text{obs}} | W_i^{\text{obs}} = 0, X_i = T^*_{\pm}] \quad (3.4)$$

If in addition Assumption 6 holds, then τ^{ate} is identified:

$$\tau^{\text{ate}} = \mathbb{E} [\mathbb{E}[Y_i^{\text{obs}} | W_i^{\text{obs}} = 1, X_i] - \mathbb{E}[Y_i^{\text{obs}} | W_i^{\text{obs}} = 0, X_i]] \quad (3.5)$$

In the FRD case, we observe treated and untreated individuals with values for the forcing variable that are different than the threshold. In general, comparing treated and untreated outcomes away from the threshold does not identify treatment effects because treated and untreated individuals belong to different compliance subpopulations. Under external validity, the average potential outcome does not vary across compliance subpopulations, and the comparison of treated and untreated outcomes away from the threshold does identify treatment effects.

Let us now explore the testable restrictions implied by the external validity assumption. External validity implies the following conditional independence restriction:

$$G_i \perp\!\!\!\perp Y_i(1) \mid X_i = T^*.$$

This in turn implies

$$G_i \perp\!\!\!\perp Y_i(1) \mid G_i \in \{a, c\}, X_i = T^*,$$

so that

$$\mathbb{E}[Y_i(1) | G_i = c, X_i = T^*] = \mathbb{E}[Y_i(1) | G_i = a, X_i = T^*], \quad (3.6)$$

and, by a similar argument,

$$\mathbb{E}[Y_i(0) | G_i = c, X_i = T^*] = \mathbb{E}[Y_i(0) | G_i = n, X_i = T^*], \quad (3.7)$$

These two restrictions are still in terms of unobserved variables. However, they imply a pair of restrictions on observed variables, as summarized in the following lemma.

Lemma 5. *Suppose Assumptions 1-5 hold. Then:*

$$\mathbb{E} [Y_i^{\text{obs}} \mid W_i^{\text{obs}} = w, X_i = T_+^*] = \mathbb{E} [Y_i^{\text{obs}} \mid W_i^{\text{obs}} = w, X_i = T_-^*], \quad (3.8)$$

for $w = 0, 1$.

COMMENT 2: The two restrictions (3.6) and (3.7) together imply the restriction tested by the Hausman test, $H_0^{H, \text{frd}'}$, as well as the restriction tested by Angrist's test, $H_0^{A, \text{frd}''}$. However, the two restrictions are stronger than the restrictions tested by either the Hausman or the Angrist test on their own. We view testing the pair of restrictions (3.6) and (3.7) jointly as more attractive than simply one linear combination of them. For example, suppose the Angrist null hypothesis $H_0^{A, \text{frd}''}$ holds, but not the pair of restrictions (3.6) and (3.7). In that case the Angrist null hypothesis would no longer hold after a transformation of the outcome, say from Y_i^{obs} to $\ln(Y_i^{\text{obs}})$. $\square$

COMMENT 3: Equation (3.8) is the key equation in the paper, and there is a simple graphical representation of it. Equation (3.8) says that the expectation of Y_i^{obs} conditional on $W_i^{\text{obs}} = w$ and $X_i = x$ is continuous in x at the threshold T^* for $w = 0, 1$. This pair of restrictions can be assessed by estimating two functions of x

$$\begin{aligned} \mathbb{E} [Y_i^{\text{obs}} | W_i^{\text{obs}} = 1, X_i = x] \\ \mathbb{E} [Y_i^{\text{obs}} | W_i^{\text{obs}} = 0, X_i = x] \end{aligned}$$

and checking for a discontinuity at $x = T^*$ in both of them. The implementation of the test is very similar to estimating a treatment effect in a regression discontinuity design. Note, however, that the density of the forcing variable X_i conditional on $W_i = w$ may be discontinuous at the threshold, whereas in comparisons of $\mathbb{E} [Y_i^{\text{obs}} | X_i = x]$ we typically expect the density of X_i to be continuous at T^* . This amounts to a slight change in the asymptotic variance of the non-parametric side-limit estimator. $\square$

4 External Validity and Covariates

Empirical studies using FRD often have data on pre-treatment variables beyond the forcing variable X_i . For estimating τ^{frd} these variables typically matter very little. In most cases, the distribution of the pre-treatment variables, like that of the forcing variable, is similar on both sides of the threshold and adjusting for differences in these pre-treatment variables is only an issue for efficiency: the failure to adjust for them does not introduce any bias. This is fundamentally different for the question of external validity. We may well find that the conditional expectation of Y_i^{obs} given $X_i = x$ and $W_i^{\text{obs}} = w$ is discontinuous at the cutoff because of differences across compliance groups in pre-treatment

variables other than X_i . As a result, in many settings, external validity becomes a more plausible assumption when the researcher is able to perform the analysis conditional on these pre-treatment characteristics of individuals. In the summer school example, always-taker students may be less skilled than compliers and never-takers because always-takers attend summer school irrespective of the threshold for attending summer school. In this case, treatment effects on compliers are likely to be different than treatment effects on always-takers if we only condition on the forcing variable X_i . A richer data-set may contain an additional vector of pre-treatment characteristics like parental education, household income, or even other test-scores measures different than the forcing variable. For example, potential outcomes for always-takers could differ from never-takers and compliers solely because always-takers come from lower socio-economic status homes. In this case, the study is externally valid conditional on the forcing variable and the additional pre-treatment characteristics.

Let V_i be a vector of q covariates describing the characteristics of individual i before the treatment. The support of V_i is denoted $\mathbb{V} \subset \mathbb{R}^q$. We modify Assumptions 2 and 5 to accommodate the analysis conditional on covariates V_i .

Assumption 7. *For every $w \in \{0, 1\}$, $g \in \{\text{n, a, c}\}$, and $v \in \mathbb{V}$, (i) the conditional distribution of $Y_i(w)|X_i = x, G_i = g, V_i = v$ is continuous in x at all $x \in \mathbb{X}$, and all of its moments are finite; (ii) the conditional probability $\text{pr}(G_i = g|X_i = x, V_i = v)$ is continuous in x at $x = T^*$.*

In this section, we are interested in testing the plausibility of external validity conditional on X_i and V_i .

Assumption 8. *The study has local external validity if,*

$$G_i \perp\!\!\!\perp (Y_i(0), Y_i(1)) \mid X_i, V_i \quad (4.1)$$

This is a generalized version of the ‘conditional effect ignorability’ assumption of Angrist and Fernandez-Val (2010). Given these assumptions, we have a version of Lemma 5 conditional on the additional pre-treatment variables.

Lemma 6. *Suppose Assumptions 1, 3, 4, 7 and 8 hold. Then:*

$$\mathbb{E}[Y_i^{\text{obs}} \mid W_i^{\text{obs}} = w, X_i = T_+^*, V_i = v] = \mathbb{E}[Y_i^{\text{obs}} \mid W_i^{\text{obs}} = w, X_i = T_-^*, V_i = v] \quad (4.2)$$

for $w = 0, 1$ and $v \in \mathbb{V}$.

In the same spirit of Lemma 4, adding Assumption 6 identifies $\mathbb{E}[Y_i(w)|X_i = x, V_i = v]$ by $\mathbb{E}[Y_i^{\text{obs}}|W_i^{\text{obs}} = w, X_i = x, V_i = v]$. Integration over the distribution (V_i, X_i) gives the global average treatment effect τ^{ate} .

The result in Lemma 6 leads us to look at the conditional expectation of Y_i^{obs} given $X_i = x$, $W_i = w$, and $V_i = v$ as a function of x around the threshold T^* for all values of v . Doing so may be difficult in practice because of small sample sizes for each value of v . A practical solution is to marginalize over a given distribution of V_i . A natural choice is the distribution of V_i conditional on X_i being close to the threshold since this is the average of the distributions $V_i|X_i = T^*, G_i = g$ for $g \in \{c, a, n\}$. Define the function $\mu : \{0, 1\} \times \mathbb{X} \rightarrow \mathbb{R}$ by

$$\mu(w, x) = \mathbb{E} \left[\mathbb{E} [Y_i^{\text{obs}} | W_i^{\text{obs}} = w, X_i = x, V_i] \middle| X_i = T_{\pm}^* \right] \quad (4.3)$$

We are interested in testing for a discontinuity at $x = T^*$ in the functions $\mu(1, x)$ and $\mu(0, x)$. Non-parametric estimation of $\mathbb{E} [Y_i^{\text{obs}} | W_i^{\text{obs}} = w, X_i = x, V_i = v]$ requires bandwidth choices for both X_i and V_i and presents slow convergence rates when the dimension of V_i is large. To facilitate the implementation of the test, we assume that

$$\mathbb{E}[Y_i(w)|G_i = g, X_i = x, V_i = v] = \psi(w, g, x) + \gamma(w, g, x)'v \quad (4.4)$$

for $w \in \{0, 1\}$, $g \in \{c, a, n\}$, $x \in \mathbb{X}$, $v \in \mathbb{V}$; where $\psi : \{0, 1\} \times \{c, a, n\} \times \mathbb{X} \rightarrow \mathbb{R}$ and $\gamma : \{0, 1\} \times \{c, a, n\} \times \mathbb{X} \rightarrow \mathbb{R}^q$ are smooth functions of x for every (w, g) . Under local external validity (Assumption 8), the conditional mean of potential outcomes above does not depend on g :

$$\mathbb{E}[Y_i(w)|G_i = g, X_i = x, V_i = v] = \psi(w, x) + \gamma(w, x)'v$$

which in turn implies that

$$\begin{aligned} \mu(w, x) &= \mathbb{E} \left[\psi(w, x) + \gamma(w, x)'V_i \middle| X_i = T_{\pm}^* \right] = \psi(w, x) + \gamma(w, x)'\mathbb{E}[V_i|X_i = T_{\pm}^*] \\ \mu(w, x) &= \mathbb{E} \left[Y_i^{\text{obs}} \middle| W_i^{\text{obs}} = w, X_i = x, V_i = \mathbb{E}[V_i|X_i = T_{\pm}^*] \right] \end{aligned}$$

which is continuous in x for $w \in \{0, 1\}$. Restriction (4.4) implies that estimation of $\mu(w, x)$ is done consistently by local linear regression of Y_i^{obs} on the sample of X_i local

to x but not necessarily V_i local to $\mathbb{E}[V_i|X_i = T_\pm^*]$. Therefore, there is no need to choose additional bandwidths for V_i , and the convergence rate of the discontinuity estimator is faster under restriction (4.4).

First we estimate $\mathbb{E}[V_i|X_i = T_\pm^*]$. Choose a bandwidth h for X_i and a symmetric kernel density function $k(\cdot)$. In practice, we use the IK bandwidth (Imbens and Kalyanaraman, 2012) or the Calonico, Cattaneo and Titiunik (2014) bandwidth, along with the edge kernel $k(u) = \mathbf{1}_{|u| \leq 1} \cdot (1 - |u|)$. Next, regress V_i on a constant and $(X_i - T^*)$ weighting observations by $k\left(\frac{X_i - T^*}{h}\right)$. Denote $\bar{V}$ the intercept estimator of this regression.

Second, we estimate $\mu(w, T_+^*)$ and $\mu(w, T_-^*)$ for $w \in \{0, 1\}$. To estimate $\mu(1, T_+^*)$, simply regress Y_i^{obs} on a constant, $(X_i - T^*)$ and $(V_i - \bar{V})$ using observations with $W_i^{\text{obs}} = 1$, $X_i > 0$, and weighting by $k\left(\frac{X_i - T^*}{h}\right)$. Denote $\hat{\mu}(1, T_+^*)$ the intercept estimator of this regression. Do the same to obtain $\hat{\mu}(1, T_-^*)$ except to use $X_i \leq 0$ observations instead of $X_i > 0$. A similar idea applies to obtain $\hat{\mu}(0, T_+^*)$ and $\hat{\mu}(0, T_-^*)$. To test jointly for discontinuities, simply compare these side-limits inversely weighted by an estimator of their asymptotic variance.

5 Two Applications

In this section, we illustrate our methods using data from Jacob and Lefgren (2004) and Matsudaira (2008). These papers estimate the effect of remedial summer school programs on academic performance of students using fuzzy regression discontinuity methods.

5.1 The Jacob and Lefgren (2004) Data

Jacob and Lefgren (2004) use administrative data from the Chicago Public Schools which instituted in 1996 an accountability policy that tied summer school attendance and promotional decisions to performance on standardized tests. This policy was followed by other school districts in the country. In section 5.2, we apply our methods to Matsudaira (2008) who studied a similar policy adopted by a district in Northeastern US. The standard rule is to send a student to summer school if the minimum between his reading and math test-score is below a certain threshold. The final decision is up to teacher's discretion which leads to a FRD design.

There are reasons to expect those who do not attend summer school despite scor-

ing below the threshold (never-takers) to be different from compliers. Never-takers may have been judged to have scored below expectations and viewed as not in need of summer school, and thus better than compliers with similar pre-summer school scores; alternatively, never-takers could also be reluctant to participate in summer school and actually be worse in terms of academic ability than compliers. Similarly, it could be that always-takers are students viewed as particularly in need of the extra instruction, and thus worse in terms of academic ability than compliers with the same scores prior to summer school; alternatively, always-takers could be students that are particularly eager for additional educational experiences and who would have done better than compliers regardless of their summer school participation.

Jacob and Lefgren use observations around a test-score cutoff to estimate the impact of attending summer school on academic performance. They find a positive impact on achievement for third graders. We use the data for third graders in years 1997-99. The forcing variable X_i is the minimum between reading and math score before the summer school minus the threshold (2.75 in this application). The outcome variable Y_i^{obs} is measured by the math score after the summer school. Summary statistics are presented in Table 1.

Following the work by Hahn, Todd and Van Der Klaauw (2001) and Porter (2003), we fit a local linear regression on each side of the cutoff using the edge kernel $k(u) = \mathbf{1}_{|u| \leq 1} \cdot (1 - |u|)$. The optimal bandwidth was computed based on the Imbens and Kalyanaraman (2012) optimal bandwidth rule. Estimates are presented in Table 2, and conditional mean plots, in Figure 1. In Figure 1a we plot the probability of attending summer school given the value of the forcing variable. Most of the students that are eligible for summer school based on the minimum test-score do attend summer school, and the change in the probability of attending summer school is estimated at 0.894 (s.e. 0.006). Next, in Figure 1b, we plot the conditional mean of the math test-score after summer school given the minimum test-score prior to summer school. Conditional on having the minimum test-score close to the threshold, there is a significant increase in the outcome for those who are required to attend summer school compared to those who are not required to attend summer school. The causal effect of summer school on subsequent academic performance for the subpopulation of compliers is around 0.20 (s.e. 0.03).

In Figures 1c and 1d, we present estimates of the expected value of the outcome

conditional on both the forcing variable and treatment status $\mathbb{E}[Y_i|X_i, W_i = w]$ for $w \in \{0, 1\}$. We find that there is not a substantial difference between untreated never-takers and compliers, with the difference in average outcomes estimated at 0.06 (s.e. 0.05). On the other hand, there is a substantial difference in average outcomes between treated always-takers and compliers, at 0.35 (s.e. 0.12). Always-takers perform substantially worse than treated compliers, consistent with the notion that the always-takers are guided towards the summer program even if they score slightly above the threshold. In this application, the Hausman and Angrist null hypotheses are rejected at conventional levels, suggesting heterogeneity of treatment effects across compliance groups. Our approach finds evidence supporting homogeneity of treatment effects between compliers and never-takers. However, the test of the pair of restrictions for always-takers and compliers and for never-takers and compliers leads to a clear rejection, with an F-statistic equal to 9.5. Adjusting for prior test scores reduces the F-statistic to 1.8, suggesting that the data are consistent with the hypothesis that always-takers and compliers, and never-takers and compliers, are similar after adjusting for observed covariates (Figures 1e and 1f).

5.2 The Matsudaira (2008) Data

Matsudaira(2008) uses administrative data from a large urban school district in North-eastern United States to evaluate the impact of attending summer school on students' academic performance. One of the promotion criteria in this school district requires students in third grade or above to score above a given cutoff score on year-end examinations in both math and reading in order to pass to the next grade. Students who fail to score above the cutoff are more likely to be required to attend the summer school program. Besides passing the test-score in reading and math, there are other criteria for promotion like attendance which is not recorded in these data. This makes the assignment to summer school fuzzy around the test cutoff, similar to the Jacob-Lefgren study.

We use individual level data for fifth grade students with reading and math test-scores taken in 2001 and 2002. The forcing variable X_i is the minimum of the 2001 reading score and the 2001 mathematics score, minus the threshold for passing. The outcome variable Y_i is the standardized mathematics score in 2002. Summary statistics are presented in Table 3.

We present the same plots as in the Jacob-Lefgren application in Figure 2. Estimates are shown in Table 4. First, in Figure 2a we present the probability of attending summer school given the value of the forcing variable. Next, in Figure 2b, we present the conditional mean of the math test-score after summer school given the minimum test-score before summer school. Conditional on having the minimum test-score close to the threshold, there is a significant increase in the outcome for those who are required to attend summer school compared to those who are not required to attend summer school, implying a positive causal effect of summer school on compliers estimated at 0.20 (s.e. 0.03). We report the graph for the expected value of the outcome conditional on both the forcing variable and treatment status $\mathbb{E}[Y_i|X_i, W_i = w]$ for $w \in \{0, 1\}$ in Figures 2c and 2d. The average potential outcome if treated is statistically different between always-takers and compliers, and the average potential outcome if not treated is also statistically different between never-takers and compliers. In contrast to the findings for the Jacob-Lefgren data, we find in the Matsudaira data substantial evidence of heterogeneity in achievement between students that never need summer school (never-takers) and those that might need depending on their test-scores (compliers). Even after adjusting for observed differences in prior test scores, these differences between always-takers and compliers remain substantial. Compared to the unadjusted differences the F-statistic goes down from 31.1 to 14.1, still far above conventional significance levels.

6 Conclusion

External validity of fuzzy regression discontinuity analyses is often a concern because identification is credible only around the threshold of the forcing variable and for the subpopulation of compliers. In many cases, however, researchers are also interested in generalizing the findings to subpopulations with values of the forcing variable away from the threshold and to subpopulations other than compliers. In this paper we explore assumptions that allow for such generalizations, and we derive testable implications of such assumptions that are easy and more attractive to test compared to existing approaches in the literature. We show that these implications can be assessed by inspecting the conditional expectation of the outcome given the forcing variable separately by treatment status. As a matter of routine, we recommend that researchers present graphs containing

estimates of these conditional expectations and test for continuity at the threshold.

It should be noted that, unlike in most regression discontinuity analyses, exogenous covariates play an important role in our testing procedures. Adjusting for observed differences in covariates between control and treatment groups can make a substantial difference. In the first of our applications, evidence of external validity is only found after our proposed adjustment for covariates is applied.

Appendix

Proof of Lemma 2:

By the arguments in Hahn, Todd and Van Der Klaauw (2001),

$$\tau^{\text{frd}} = \mathbb{E}[Y_i(1)|G_i = c, X_i = T^*] - \mathbb{E}[Y_i(0)|G_i = c, X_i = T^*].$$

Define the following events and probabilities:

$$A_{i,h} = \{T^* - h < X_i < T^* + h\},$$

$$A_{i,h}^+ = \{T^* < X_i < T^* + h\},$$

$$A_{i,h}^- = \{T^* - h < X_i \leq T^*\},$$

$$\pi_a = \pi_a(T^*) = \text{pr}(G_i = a|X_i = T^*),$$

$$\pi_c = \pi_c(T^*) = \text{pr}(G_i = c|X_i = T^*),$$

and

$$\pi_n = \pi_n(T^*) = \text{pr}(G_i = n|X_i = T^*).$$

Next, consider the first term in $\tau_{\text{exo}}, \lim_{h \downarrow 0} \mathbb{E}[Y_i^{\text{obs}}|W_i^{\text{obs}} = 1, A_{i,h}]$ without the limit operator. For the ε of definition (2.1), pick $h \in (0, \varepsilon)$.

$$\begin{aligned} \mathbb{E}[Y_i^{\text{obs}}|W_i^{\text{obs}} = 1, A_{i,h}] &= \mathbb{E}[Y_i(1)|W_i^{\text{obs}} = 1, A_{i,h}] \\ &= \mathbb{E}[Y_i(1)|W_i^{\text{obs}} = 1, G_i = a, A_{i,h}] \cdot \text{pr}(G_i = a|W_i^{\text{obs}} = 1, A_{i,h}) \\ &\quad + \mathbb{E}[Y_i(1)|W_i^{\text{obs}} = 1, G_i = c, A_{i,h}] \cdot \text{pr}(G_i = c|W_i^{\text{obs}} = 1, A_{i,h}) \\ &= \mathbb{E}[Y_i(1)|G_i = a, A_{i,h}] \cdot \text{pr}(G_i = a|W_i^{\text{obs}} = 1, A_{i,h}) \\ &\quad + \mathbb{E}[Y_i(1)|G_i = c, A_{i,h}^-] \cdot \text{pr}(G_i = c|W_i^{\text{obs}} = 1, A_{i,h}) \\ &= \mathbb{E}[Y_i(1)|G_i = a, A_{i,h}] \cdot \text{pr}(G_i = a|W_i^{\text{obs}} = 1, A_{i,h}^-) \cdot \text{pr}(A_{i,h}^-|W_i^{\text{obs}} = 1, A_{i,h}) \\ &\quad + \mathbb{E}[Y_i(1)|G_i = a, A_{i,h}] \cdot \text{pr}(G_i = a|W_i^{\text{obs}} = 1, A_{i,h}^+) \cdot \text{pr}(A_{i,h}^+|W_i^{\text{obs}} = 1, A_{i,h}) \\ &\quad + \mathbb{E}[Y_i(1)|G_i = c, A_{i,h}^-] \cdot \text{pr}(G_i = c|W_i^{\text{obs}} = 1, A_{i,h}^-) \cdot \text{pr}(A_{i,h}^-|W_i^{\text{obs}} = 1, A_{i,h}) \\ &\quad + \mathbb{E}[Y_i(1)|G_i = c, A_{i,h}^-] \cdot \text{pr}(G_i = c|W_i^{\text{obs}} = 1, A_{i,h}^+) \cdot \text{pr}(A_{i,h}^+|W_i^{\text{obs}} = 1, A_{i,h}) \\ &= \mathbb{E}[Y_i(1)|G_i = a, A_{i,h}] \cdot \text{pr}(G_i = a|G_i \in \{a, c\}, A_{i,h}^-) \\ &\quad \cdot \frac{\text{pr}(W_i^{\text{obs}} = 1|A_{i,h}^-)\text{pr}(A_{i,h}^-)}{\sum_{s \in \{-, +\}} \text{pr}(W_i^{\text{obs}} = 1|A_{i,h}^s)\text{pr}(A_{i,h}^s)} \\ &\quad + \mathbb{E}[Y_i(1)|G_i = a, A_{i,h}] \cdot \text{pr}(G_i = a|G_i = a, A_{i,h}^+) \end{aligned}$$

$$\begin{aligned}
& \cdot \frac{\text{pr}(W_i^{\text{obs}} = 1 | A_{i,h}^+) \text{pr}(A_{i,h}^+)}{\sum_{s \in \{-, +\}} \text{pr}(W_i^{\text{obs}} = 1 | A_{i,h}^s) \text{pr}(A_{i,h}^s)} \\
& + \mathbb{E}[Y_i(1) | G_i = c, A_{i,h}^-] \cdot \text{pr}(G_i = c | G_i \in \{a, c\}, A_{i,h}^-) \\
& \cdot \frac{\text{pr}(W_i^{\text{obs}} = 1 | A_{i,h}^-) \text{pr}(A_{i,h}^-)}{\sum_{s \in \{-, +\}} \text{pr}(W_i^{\text{obs}} = 1 | A_{i,h}^s) \text{pr}(A_{i,h}^s)} \\
& + \mathbb{E}[Y_i(1) | G_i = c, A_{i,h}^-] \cdot \text{pr}(G_i = c | G_i = a, A_{i,h}^+) \\
& \cdot \frac{\text{pr}(W_i^{\text{obs}} = 1 | A_{i,h}^+) \text{pr}(A_{i,h}^+)}{\sum_{s \in \{-, +\}} \text{pr}(W_i^{\text{obs}} = 1 | A_{i,h}^s) \text{pr}(A_{i,h}^s)} \\
& = \mathbb{E}[Y_i(1) | G_i = a, A_{i,h}] \cdot \text{pr}(G_i = a | G_i \in \{a, c\}, A_{i,h}^-) \\
& \cdot \frac{\text{pr}(G_i \in \{a, c\} | A_{i,h}^-)}{\text{pr}(G_i \in \{a, c\} | A_{i,h}^-) + \text{pr}(G_i = a | A_{i,h}^+) \text{pr}(A_{i,h}^+) / \text{pr}(A_{i,h}^-)} \\
& + \mathbb{E}[Y_i(1) | G_i = a, A_{i,h}] \cdot 1 \\
& \cdot \frac{\text{pr}(G_i = a | A_{i,h}^+)}{\text{pr}(G_i \in \{a, c\} | A_{i,h}^-) \text{pr}(A_{i,h}^-) / \text{pr}(A_{i,h}^+) + \text{pr}(G_i = a | A_{i,h}^+)} \\
& + \mathbb{E}[Y_i(1) | G_i = c, A_{i,h}^-] \cdot \text{pr}(G_i = c | G_i \in \{a, c\}, A_{i,h}^-) \\
& \cdot \frac{\text{pr}(G_i \in \{a, c\} | A_{i,h}^-)}{\text{pr}(G_i \in \{a, c\} | A_{i,h}^-) + \text{pr}(G_i = a | A_{i,h}^+) \text{pr}(A_{i,h}^+) / \text{pr}(A_{i,h}^-)} \\
& + 0
\end{aligned}$$

Now, take the limit as $h \downarrow 0$. By Assumption 2, the conditional mean of potential outcomes given each compliance type and the forcing variable is a continuous function of the forcing variable; also, the conditional probability of each compliance type given the forcing variable is a continuous function of the forcing variable; lastly, continuity of the density of the forcing variable at the threshold implies $\lim_{h \downarrow 0} \text{pr}(A_{i,h}^-) / \text{pr}(A_{i,h}^+) = 1$. Therefore,

$$\begin{aligned}
& \lim_{h \downarrow 0} \mathbb{E}[Y_i^{\text{obs}} | W_i^{\text{obs}} = 1, A_{i,h}] \\
& = \mathbb{E}[Y_i(1) | G_i = a, X_i = T^*] \cdot \frac{\pi_a}{\pi_a + \pi_c} \cdot \frac{\pi_a + \pi_c}{2\pi_a + \pi_c} \\
& \quad + \mathbb{E}[Y_i(1) | G_i = a, X_i = T^*] \cdot \frac{\pi_a}{2\pi_a + \pi_c} \\
& \quad + \mathbb{E}[Y_i(1) | G_i = c, X_i = T^*] \cdot \frac{\pi_c}{\pi_a + \pi_c} \cdot \frac{\pi_a + \pi_c}{2\pi_a + \pi_c} \\
& = \mathbb{E}[Y_i(1) | G_i = a, X_i = T^*] \cdot \frac{\pi_a}{\pi_a + \pi_c/2} + \mathbb{E}[Y_i(1) | G_i = c, X_i = T^*] \cdot \frac{\pi_c/2}{\pi_a + \pi_c/2}
\end{aligned}$$

By a similar argument

$$\begin{aligned} & \lim_{h \downarrow 0} \mathbb{E}[Y_i^{\text{obs}} | W_i^{\text{obs}} = 0, A_{i,h}] \\ &= \mathbb{E}[Y_i(0) | G_i = \text{n}, X_i = T^*] \cdot \frac{\pi_n}{\pi_n + \pi_c/2} + \mathbb{E}[Y_i(0) | G_i = \text{c}, X_i = T^*] \cdot \frac{\pi_c/2}{\pi_n + \pi_c/2} \end{aligned}$$

Finally,

$$\begin{aligned} & \tau^{\text{exo}} - \tau^{\text{frd}} \\ &= \mathbb{E}[Y_i(1) | G_i = \text{a}, X_i = T^*] \cdot \frac{\pi_a}{\pi_a + \pi_c/2} + \mathbb{E}[Y_i(1) | G_i = \text{c}, X_i = T^*] \cdot \frac{\pi_c/2}{\pi_a + \pi_c/2} \\ &\quad - \mathbb{E}[Y_i(0) | G_i = \text{n}, X_i = T^*] \cdot \frac{\pi_n}{\pi_n + \pi_c/2} - \mathbb{E}[Y_i(0) | G_i = \text{c}, X_i = T^*] \cdot \frac{\pi_c/2}{\pi_n + \pi_c/2} \\ &\quad - \mathbb{E}[Y_i(1) | G_i = \text{c}, X_i = T^*] + \mathbb{E}[Y_i(0) | G_i = \text{c}, X_i = T^*] \end{aligned}$$

$$\begin{aligned} & \tau^{\text{exo}} - \tau^{\text{frd}} \\ &= \frac{\pi_a}{\pi_a + \pi_c/2} (\mathbb{E}[Y_i(1) | G_i = \text{a}, X_i = T^*]) - \mathbb{E}[Y_i(1) | G_i = \text{c}, X_i = T^*] \\ &\quad - \frac{\pi_n}{\pi_n + \pi_c/2} (\mathbb{E}[Y_i(0) | G_i = \text{n}, X_i = T^*] - \mathbb{E}[Y_i(0) | G_i = \text{c}, X_i = T^*]) \end{aligned}$$

□

Proof of Lemma 3: The equality of $H_0^{A, \text{frd}'}$ and $H_0^{A, \text{frd}''}$ is immediate. The proof therefore focuses on the equality of $H_0^{A, \text{frd}'}$ and $H_0^{A, \text{frd}}$. As shown before,

$$\tau^{\text{frd}} = \mathbb{E}[Y_i(1) | G_i = \text{c}, X_i = T^*] - \mathbb{E}[Y_i(0) | G_i = \text{c}, X_i = T^*],$$

so all that remains to be shown is

$$\begin{aligned} & \lim_{h \downarrow 0} \mathbb{E}[Y_i^{\text{obs}} | W_i^{\text{obs}} = 1, T^* < X_i < T^* + h] - \mathbb{E}[Y_i^{\text{obs}} | W_i^{\text{obs}} = 0, T^* - h < X_i \leq T^*] \\ &= \mathbb{E}[Y_i(1) | G_i = \text{a}, X_i = T^*] - \mathbb{E}[Y_i(0) | G_i = \text{n}, X_i = T^*], \end{aligned}$$

By definition,

$$\begin{aligned} & \lim_{h \downarrow 0} \mathbb{E}[Y_i^{\text{obs}} | W_i^{\text{obs}} = 1, T^* < X_i < T^* + h] - \mathbb{E}[Y_i^{\text{obs}} | W_i^{\text{obs}} = 0, T^* - h < X_i \leq T^*] \\ &= \lim_{h \downarrow 0} \mathbb{E}[Y_i(1) | W_i^{\text{obs}} = 1, T^* < X_i < T^* + h] - \mathbb{E}[Y_i(0) | W_i^{\text{obs}} = 0, T^* - h < X_i \leq T^*]. \end{aligned}$$

This is equal to

$$\lim_{h \downarrow 0} \mathbb{E}[Y_i(1) | W_i^{\text{obs}} = 1, G_i = \text{a}, T^* < X_i < T^* + h] - \mathbb{E}[Y_i(0) | W_i^{\text{obs}} = 0, G_i = \text{n}, T^* - h < X_i \leq T^*].$$

$$\begin{aligned}
&= \lim_{h \downarrow 0} \mathbb{E}[Y_i(1)|G_i = a, T^* < X_i < T^* + h] - \mathbb{E}[Y_i(0)|G_i = n, T^* - h < X_i \leq T^*]. \\
&= \mathbb{E}[Y_i(1)|G_i = a, X_i = T^*] - \mathbb{E}[Y_i(0)|G_i = n, X_i = T^*],
\end{aligned}$$

which finishes the proof. $\square$

Proof of Lemma 4:

First Result: we show Equation (3.4). For an arbitrary $x \in \mathbb{X}$, by external validity (Assumption 5), it follows that :

$$\mathbb{E}[Y_i(0)|G_i = g, X_i = x] = \mathbb{E}[Y_i(0)|X_i = x] \quad \forall g \in \{n, a, c\} \quad (\text{A.1})$$

$$\mathbb{E}[Y_i(1)|G_i = g, X_i = x] = \mathbb{E}[Y_i(1)|X_i = x] \quad \forall g \in \{n, a, c\} \quad (\text{A.2})$$

According to the proof of Lemma 2, (A.1) and (A.2):

$$\begin{aligned}
&\mathbb{E}[Y_i^{\text{obs}}|W_i^{\text{obs}} = 1, X_i = T_{\pm}^*] \\
&= \mathbb{E}[Y_i(1)|G_i = a, X_i = T^*] \cdot \frac{\pi_a}{\pi_a + \pi_c/2} + \mathbb{E}[Y_i(1)|G_i = c, X_i = T^*] \cdot \frac{\pi_c/2}{\pi_a + \pi_c/2} \\
&= \mathbb{E}[Y_i(1)|X_i = T^*]
\end{aligned}$$

$$\begin{aligned}
&\mathbb{E}[Y_i^{\text{obs}}|W_i^{\text{obs}} = 0, X_i = T_{\pm}^*] \\
&= \mathbb{E}[Y_i(0)|G_i = n, X_i = T^*] \cdot \frac{\pi_n}{\pi_n + \pi_c/2} + \mathbb{E}[Y_i(0)|G_i = c, X_i = T^*] \cdot \frac{\pi_c/2}{\pi_n + \pi_c/2} \\
&= \mathbb{E}[Y_i(0)|X_i = T^*]
\end{aligned}$$

Therefore, $\tau(T^*) = \mathbb{E}[Y_i^{\text{obs}}|W_i^{\text{obs}} = 1, X_i = T_{\pm}^*] - \mathbb{E}[Y_i^{\text{obs}}|W_i^{\text{obs}} = 0, X_i = T_{\pm}^*]$.

Second Result: we show Equation (3.5). Note that

$$\tau^{\text{ate}} = \mathbb{E}[\mathbb{E}[Y_i(1) - Y_i(0)|X_i]]$$

so that it suffices to show that $\mathbb{E}[Y_i(1) - Y_i(0)|X_i = x]$ is identified for every $x \in \mathbb{X}$ (support of X_i) because the distribution of X_i is identified.

Case I: $x \leq T^*$

Using the definitions of G_i and W_i^{obs} , and Assumption 6:

$$G_i \in \{c, a\} \iff W_i^{\text{obs}} = 1 \mid X_i = x \quad (\text{A.3})$$

$$G_i = n \iff W_i^{\text{obs}} = 0 \mid X_i = x \quad (\text{A.4})$$

Consider

$$\mathbb{E}[Y_i^{\text{obs}}|W_i^{\text{obs}} = 0, X_i = x]$$

$$\begin{aligned}
&= \mathbb{E}[Y_i(0)|W_i^{\text{obs}} = 0, X_i = x] \\
&= \mathbb{E}[Y_i(0)|G_i = \text{n}, X_i = x] \\
&= \mathbb{E}[Y_i(0)|X_i = x]
\end{aligned}$$

where we used (A.4) in the second equality, and (A.1) in the third equality.

Next, consider

$$\begin{aligned}
&\mathbb{E}[Y_i^{\text{obs}}|W_i^{\text{obs}} = 1, X_i = x] \\
&= \mathbb{E}[Y_i(1)|W_i^{\text{obs}} = 1, X_i = x] \\
&= \mathbb{E}[Y_i(1)|G_i \in \{\text{c}, \text{a}\}, X_i = x] \\
&= \mathbb{E}[Y_i(1)|G_i = \text{c}, X_i = x] \cdot \text{pr}(G_i = \text{c}|G_i \in \{\text{c}, \text{a}\}, X_i = x) \\
&\quad + \mathbb{E}[Y_i(1)|G_i = \text{a}, X_i = x] \cdot \text{pr}(G_i = \text{a}|G_i \in \{\text{c}, \text{a}\}, X_i = x) \\
&= \mathbb{E}[Y_i(1)|X_i = x]
\end{aligned}$$

where we used (A.3) in the second equality, and (A.2) in the fourth equality.

Therefore, $\mathbb{E}[Y_i(1) - Y_i(0)|X_i = x]$ is identified in case I.

Case II: $x > T^*$

Again, by the definitions of G_i and W_i^{obs} , and Assumption 6:

$$G_i = \text{a} \iff W_i^{\text{obs}} = 1 \mid X_i = x \tag{A.5}$$

$$G_i \in \{\text{n}, \text{c}\} \iff W_i^{\text{obs}} = 0 \mid X_i = x \tag{A.6}$$

Consider

$$\begin{aligned}
&\mathbb{E}[Y_i^{\text{obs}}|W_i^{\text{obs}} = 0, X_i = x] \\
&= \mathbb{E}[Y_i(0)|W_i^{\text{obs}} = 0, X_i = x] \\
&= \mathbb{E}[Y_i(0)|G_i \in \{\text{n}, \text{c}\}, X_i = x] \\
&= \mathbb{E}[Y_i(0)|G_i = \text{n}, X_i = x] \cdot \text{pr}(G_i = \text{n}|G_i \in \{\text{n}, \text{c}\}, X_i = x) \\
&\quad + \mathbb{E}[Y_i(0)|G_i = \text{c}, X_i = x] \cdot \text{pr}(G_i = \text{c}|G_i \in \{\text{n}, \text{c}\}, X_i = x) \\
&= \mathbb{E}[Y_i(0)|X_i = x]
\end{aligned}$$

where we used (A.6) in the second equality, and (A.1) in the fourth equality.

Next, consider

$$\begin{aligned}
&\mathbb{E}[Y_i^{\text{obs}}|W_i^{\text{obs}} = 1, X_i = x] \\
&= \mathbb{E}[Y_i(1)|W_i^{\text{obs}} = 1, X_i = x] \\
&= \mathbb{E}[Y_i(1)|G_i = \text{a}, X_i = x]
\end{aligned}$$

$$= \mathbb{E}[Y_i(1)|X_i = x]$$

where we used (A.5) in the second equality, and (A.2) in the third equality.

Therefore, $\mathbb{E}[Y_i(1) - Y_i(0)|X_i = x]$ is identified in case II.

□

Proof of Lemma 5: The claims follow directly from the proof of the second result in Lemma 4 above. Simply consider Case I with $x \in [T^* - h, T^*]$ and Case II with $x \in (T^*, T^* + h]$ for $h \in (0, \varepsilon)$ with ε of Definition (2.1). Then, take $h \rightarrow 0$.

□

Proof of Lemma 6: It follows directly from the proof of Lemma 5 above conditional on $V_i = v$ and using Assumptions 7 and 8.

□

References

- ANGRIST, J., (2004), “Treatment Effect Heterogeneity in Theory and Practice,” , *The Economic Journal* 114(494), C52-C83.
- ANGRIST, J., AND I. FERNANDEZ-VAL, (2010), “ExtrapoLATE-ing: External Validity and Overidentification in the LATE Framework”, NBER Working Paper 16566.
- ANGRIST, J.D., G.W. IMBENS, AND D.B. RUBIN, (1996), Identification of Causal Effects Using Instrumental Variables, *Journal of the American Statistical Association* 91, 444–472.
- ANGRIST, J.D. AND A.B. KRUEGER, (1991), Does Compulsory School Attendance Affect Schooling and Earnings?, *Quarterly Journal of Economics* 106, 979-1014.
- ANGRIST, J.D., AND V. LAVY, (1999), Using Maimonides’ Rule to Estimate the Effect of Class Size on Scholastic Achievement”, *Quarterly Journal of Economics* 114, 533-575.
- ANGRIST, J., AND M. ROKKANEN, (2012), “Wanna Get Away? RD Identification Away from the Cutoff”, NBER Working Paper 18662.
- BATTISTIN, E. AND E. RETTORE, (2008), “Ineligibles and Eligible Non-Participants as a Double Comparison Group in Regression-Discontinuity Designs,” *Journal of Econometrics* 142(2), 715-730
- BERK, R., AND D. RAUMA, (1983), “Capitalizing on Nonrandom Assignment to Treatments: A Regression-Discontinuity Evaluation of a Crime-Control Program,” *Journal of the American Statistical Association*, Vol. 78(381): 21-27.
- BERTANHA, M., (2015), “Regression Discontinuity Design with Many Thresholds,” Unpublished Manuscript, Department of Economics, Stanford.
- BLACK, S., (1999), Do Better Schools Matter? Parental Valuation of Elementary Education, *Quarterly Journal of Economics* 114, 577-599.
- CALONICO, S., M. CATTANEO, AND R. TITIUNIK, (2014a), “Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs” , *Econometrica*, Vol. 82(6): 2295-2326.
- CALONICO, S., M. CATTANEO, AND R. TITIUNIK, (2014b), “Robust Data-Driven Inference in the Regression-Discontinuity Design” forthcoming, *Stata Journal*.
- CALONICO, S., M. CATTANEO, AND R. TITIUNIK, (2015), “Optimal Data-Driven Regression Discontinuity Plots.” *Journal of the American Statistical Association*, forthcoming.
- CARD, D., D. LEE, Z. PEI, AND A. WEBER, (2015), “Inference on Causal Effects in a Generalized Regression Kink Design” , *Econometrica*, Vol. 83(6): 2453-2483.
- CATTANEO, M., B. FRANSEN, AND R. TITIUNIK, (2014), “Randomization Inference in the Regression Discontinuity Design: An Application to Party Advantages in the U.S. Senate” forthcoming, *Journal of Causal Inference*.

- COOK, T., (2008), "Waiting for Life to Arrive: A History of the Regression-Discontinuity Design in Psychology, Statistics and Economics." *Journal of Econometrics*, 142(2): 636-54.
- DONG, Y., (2015), "An Alternative Assumption to Identify LATE in Regression Discontinuity Designs," Unpublished Manuscript, UC Irvine.
- DONG, Y., AND A. LEWBEL (2015): "Identifying the Effect of Changing the Policy Threshold in Regression Discontinuity Models" forthcoming, *Review of Economics and Statistics*.
- GELMAN, A., AND G. IMBENS, , (2014), "Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs," NBER Working Paper 20405.
- HAHN, J., P. TODD AND W. VAN DER KLAUW, (2001), Identification and Estimation of Treatment Effects with a Regression Discontinuity Design, *Econometrica* 69, 201-209.
- HOTZ J., G. IMBENS, AND J. MORTIMER (2005), "Predicting the Efficacy of Future Training Programs Using Past Experiences," *Journal of Econometrics*, Vol. 125: 241270.
- HUBER, M., (2013), "A simple test for the ignorability of non-compliance in experiments", *Economic Letters* 120: 389-391.
- IMBENS, G., AND J. ANGRIST, (1994), "Identification and Estimation of Local Average Treatment Effects", *Econometrica* 61, 467-476.
- IMBENS, G., AND K. KALYANARAMAN, (2012), "Optimal Bandwidth Choice for the Regression Discontinuity Estimator," *Review of Economic Studies* 79(3), 933-959
- IMBENS, G., AND T. LEMIEUX, (2008), "Regression Discontinuity Designs: A Guide to Practice," *Journal of Econometrics*, Vol. 142(2): 615-35.
- JACOB, B.A., AND L. LEFGREN, (2004), Remedial Education and Student Achievement: A Regression-Discontinuity Analysis, *Review of Economics and Statistics* 68, 226-244.
- LALIVE, R., (2008), How do Extended Benefits affect Unemployment Duration? A Regression Discontinuity Approach, *Journal of Econometrics*, 142(2), 785-806
- LEE, D., (2008), "Randomized Experiments from Non-random Selection in U.S. House Elections," *Journal of Econometrics*, 142(2) 675697
- LEE, D., AND T. LEMIEUX, (2010), "Regression Discontinuity Designs in Economics," *Journal of Economic Literature*, Vol. 48, 281-355.
- LEE, D., E. MORETTI, AND M. BUTLER, (2010), "Do Voters Affect or Elect Policies? Evidence from the U.S. House," *Quarterly Journal of Economics*, Vol. 48, 807-859.
- LUDWIG, J., AND D. MILLER, (2007), Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design, *Quarterly Journal of Economics* 122(1), 159-208.
- MATSUDAIRA, J., (2008), "Mandatory Summer School and Student Achievement", *Journal of Econometrics*, 142(2) 829850.

- MCCRARY, J., (2008), “Testing for Manipulation of the Running Variable in the Regression Discontinuity Design,” *Journal of Econometrics*, 142(2) 698714.
- PORTER, J., (2003), Estimation in the Regression Discontinuity Model,” mimeo, Department of Economics, University of Wisconsin, http://www.ssc.wisc.edu/jporter/reg_discont_2003.pdf.
- RUBIN, D., (1974), Estimating Causal Effects of Treatments in Randomized and Non-randomized Studies, *Journal of Educational Psychology* 66, 688-701.
- THISTLEWAITE, D., AND D. CAMPBELL, (1960), Regression-Discontinuity Analysis: An Alternative to the Ex-Post Facto Experiment, *Journal of Educational Psychology* 51, 309-317.
- TROCHIM, W., (1984), *Research Design for Program Evaluation; The Regression-Discontinuity Design* (Sage Publications, Beverly Hills, CA).
- TROCHIM, W., (2001), Regression-Discontinuity Design, in N.J. Smelser and P.B Baltes, eds., *International Encyclopedia of the Social and Behavioral Sciences* 19 (Elsevier North-Holland, Amsterdam) 12940-12945.
- VAN DER KLAAUW, W., (2002), Estimating the Effect of Financial Aid Offers on College Enrollment: A Regression-Discontinuity Approach, *International Economic Review* 43, 1249-1287.
- VAN DER KLAAUW, W., (2008) “Regression-Discontinuity Analysis: A Survey of Recent Developments in Economics.” *Labour*, Vol. 22(2): 21945.

Table 1: Jacob-Lefgren Data

	Overall	Treated	Untreated
<u>Mean and S.D.</u>			
Outcome Y_i	4.503 (1.069)	3.811 (0.831)	4.961 (0.957)
Forcing X_i	0.234 (0.979)	-0.660 (0.527)	0.825 (0.728)
Treatment W_i	0.398 (0.489)	1.000 (0.000)	0.000 (0.000)
Covariate V_i	-1.107 (1.065)	-1.903 (0.725)	-0.581 (0.917)
<u>Number of Obs</u>			
Overall	70831	28180	42651
Within Bandwidth	26180	11777	14403

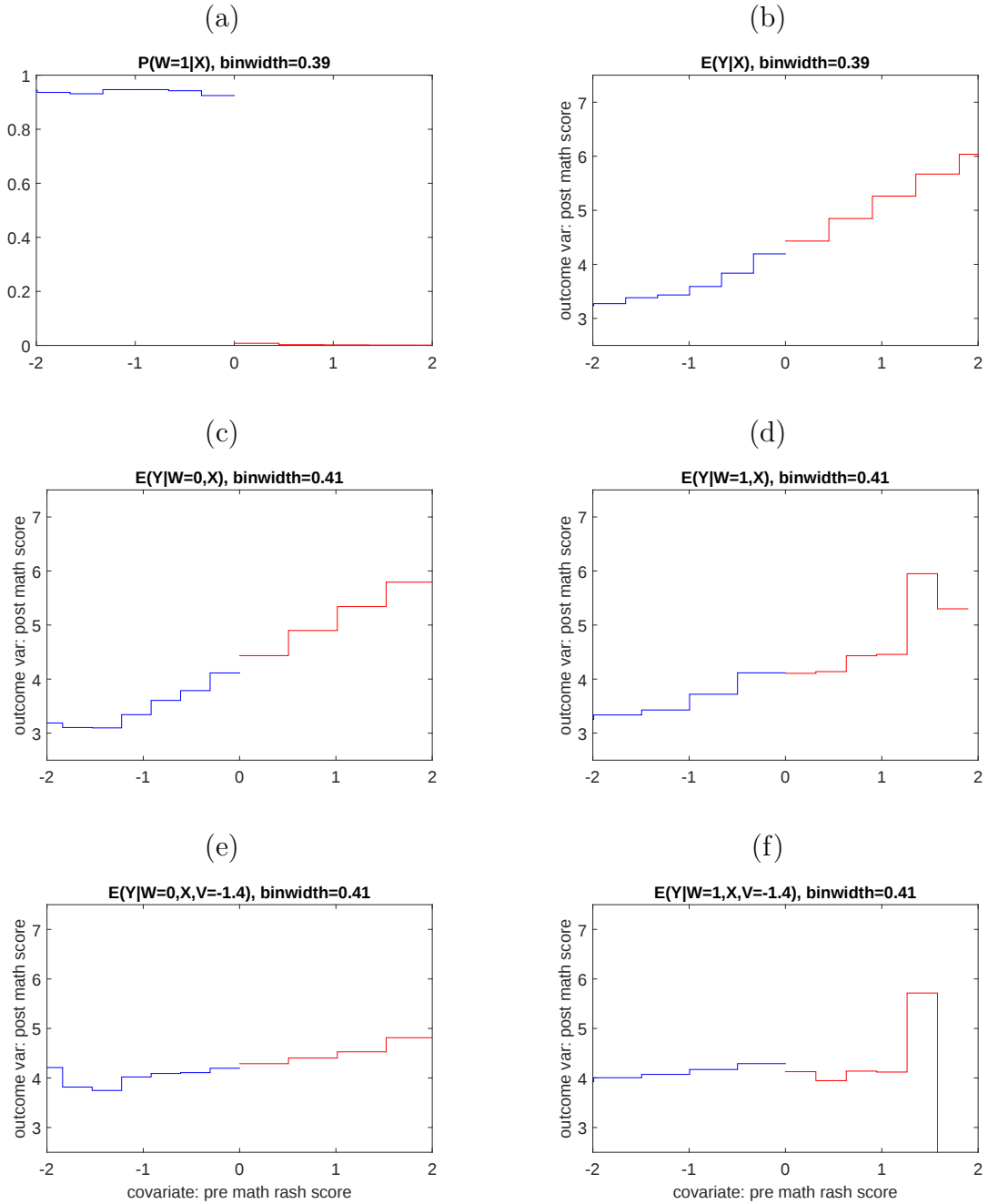
Notes: The outcome variable Y is the math score (after the summer school); the forcing variable X is the minimum between the reading and math score (before the summer school) minus the cutoff; W is the participation indicator; and the covariate V is the math Rasch score (before the summer school). Non-parametric estimates are obtained by local linear regression and optimal bandwidth $h = 0.54$ by Imbens and Kalyanaraman (2012).

Table 2: Jacob-Lefgren Data

Estimates	Estimand	Estimate	s.e.	p-value
Participation	π_c	0.894	(0.006)	
	π_a	0.016	(0.002)	
	π_n	0.090	(0.005)	
Cond. Means of Potential Outcomes	$\mathbb{E}[Y_i(1) X_i = T^*, G_i = c]$	4.394	(0.016)	
	$\mathbb{E}[Y_i(1) X_i = T^*, G_i = a]$	4.037	(0.123)	
	$\mathbb{E}[Y_i(0) X_i = T^*, G_i = c]$	4.197	(0.017)	
	$\mathbb{E}[Y_i(0) X_i = T^*, G_i = n]$	4.267	(0.047)	
LATE	$\frac{\mathbb{E}[Y_i^{\text{obs}} X_i=T_+^*] - \mathbb{E}[Y_i^{\text{obs}} X_i=T_-^*]}{\mathbb{E}[W_i^{\text{obs}} X_i=T_+^*] - \mathbb{E}[W_i^{\text{obs}} X_i=T_-^*]}$	0.197	(0.025)	
OLS	$\mathbb{E}[Y_i^{\text{obs}} W_i^{\text{obs}} = 1, X_i = T_\pm^*] - \mathbb{E}[Y_i^{\text{obs}} W_i^{\text{obs}} = 0, X_i = T_\pm^*]$	0.173	(0.021)	
<u>Tests</u>				
Hausman	$\frac{\mathbb{E}[Y_i^{\text{obs}} X_i=T_+^*] - \mathbb{E}[Y_i^{\text{obs}} X_i=T_-^*]}{\mathbb{E}[W_i^{\text{obs}} X_i=T_+^*] - \mathbb{E}[W_i^{\text{obs}} X_i=T_-^*]} - \mathbb{E}[Y_i^{\text{obs}} W_i^{\text{obs}} = 1, X_i = T_\pm^*] + \mathbb{E}[Y_i^{\text{obs}} W_i^{\text{obs}} = 0, X_i = T_\pm^*]$	0.024	(0.011)	0.024
Angrist	$\frac{\mathbb{E}[Y_i^{\text{obs}} X_i=T_+^*] - \mathbb{E}[Y_i^{\text{obs}} X_i=T_-^*]}{\mathbb{E}[W_i^{\text{obs}} X_i=T_+^*] - \mathbb{E}[W_i^{\text{obs}} X_i=T_-^*]} - \mathbb{E}[Y_i^{\text{obs}} W_i^{\text{obs}} = 1, X_i = T_+^*] + \mathbb{E}[Y_i^{\text{obs}} W_i^{\text{obs}} = 0, X_i = T_-^*]$	0.427	(0.138)	0.002
Conditional Jumps	$\mathbb{E}\left[Y_i^{\text{obs}} \mid W_i^{\text{obs}} = 0, X_i = T_-^*\right] - \mathbb{E}\left[Y_i^{\text{obs}} \mid W_i^{\text{obs}} = 0, X_i = T_+^*\right]$	0.064	(0.050)	0.203
	$\mathbb{E}\left[Y_i^{\text{obs}} \mid W_i^{\text{obs}} = 1, X_i = T_-^*\right] - \mathbb{E}\left[Y_i^{\text{obs}} \mid W_i^{\text{obs}} = 1, X_i = T_+^*\right]$	0.351	(0.123)	0.004
Joint F-test		9.503		0.009
<u>Adj. for Covariate</u>				
Conditional	$\mathbb{E}\left[\mathbb{E}\left[Y_i^{\text{obs}} \mid W_i^{\text{obs}} = 0, X_i = T_-^*, V_i\right] \mid X_i = T_\pm^*\right]$	-0.014	(0.041)	0.726
	$-\mathbb{E}\left[\mathbb{E}\left[Y_i^{\text{obs}} \mid W_i^{\text{obs}} = 0, X_i = T_+^*, V_i\right] \mid X_i = T_\pm^*\right]$			
Jumps	$\mathbb{E}\left[\mathbb{E}\left[Y_i^{\text{obs}} \mid W_i^{\text{obs}} = 1, X_i = T_-^*, V_i\right] \mid X_i = T_\pm^*\right]$	0.170	(0.131)	0.194
	$-\mathbb{E}\left[\mathbb{E}\left[Y_i^{\text{obs}} \mid W_i^{\text{obs}} = 1, X_i = T_+^*, V_i\right] \mid X_i = T_\pm^*\right]$			
Joint F-test		1.815		0.404

Notes: The outcome variable Y is the math score (after the summer school); the forcing variable X is the minimum between the reading and math score (before the summer school) minus the cutoff; W is the participation indicator; and the covariate V is the math Rasch score (before the summer school). Non-parametric estimates are obtained by local linear regression and optimal bandwidth $h = 0.54$ by Imbens and Kalyanaraman (2012). The last three estimates are conditional on $V_i = v = -1.378$. The standard errors of all estimates are computed using 1000 bootstrap iterations. The F-test is a χ^2 statistic computed as a quadratic form of the 2×1 vector of the conditional jumps inversely weighted by the covariance matrix of such vector.

Figure 1: Jacob and Lefgren



Notes: The outcome variable Y is the math score (after the summer school); the forcing variable X is the minimum between the reading and math score (before the summer school) minus the cutoff; W is the participation indicator; and the covariate V is the math Rasch score (before the summer school). Step functions are made of averages within each bin, and plots have the same number of equal sized bins on each side of the cutoff. For (e) and (f), the average within each bin is the average of predicted Y at $V = v$ using a linear regression of Y on a constant, X and V .

Table 3: Matsudaira Data

	Overall	Treated	Untreated
<u>Mean and S.D.</u>			
Outcome Y_i	0.000 (1.000)	-0.610 (0.852)	0.176 (0.970)
Forcing X_i	10.320 (41.266)	-20.003 (38.308)	19.079 (37.797)
Treatment W_i	0.224 (0.417)	1.000 (0.000)	0.000 (0.000)
Covariate V_i	633.463 (38.529)	607.403 (31.744)	640.991 (37.005)
<u>Number of Obs</u>			
Overall	68798	15420	53378
Within Bandwidth	37580	10379	27201

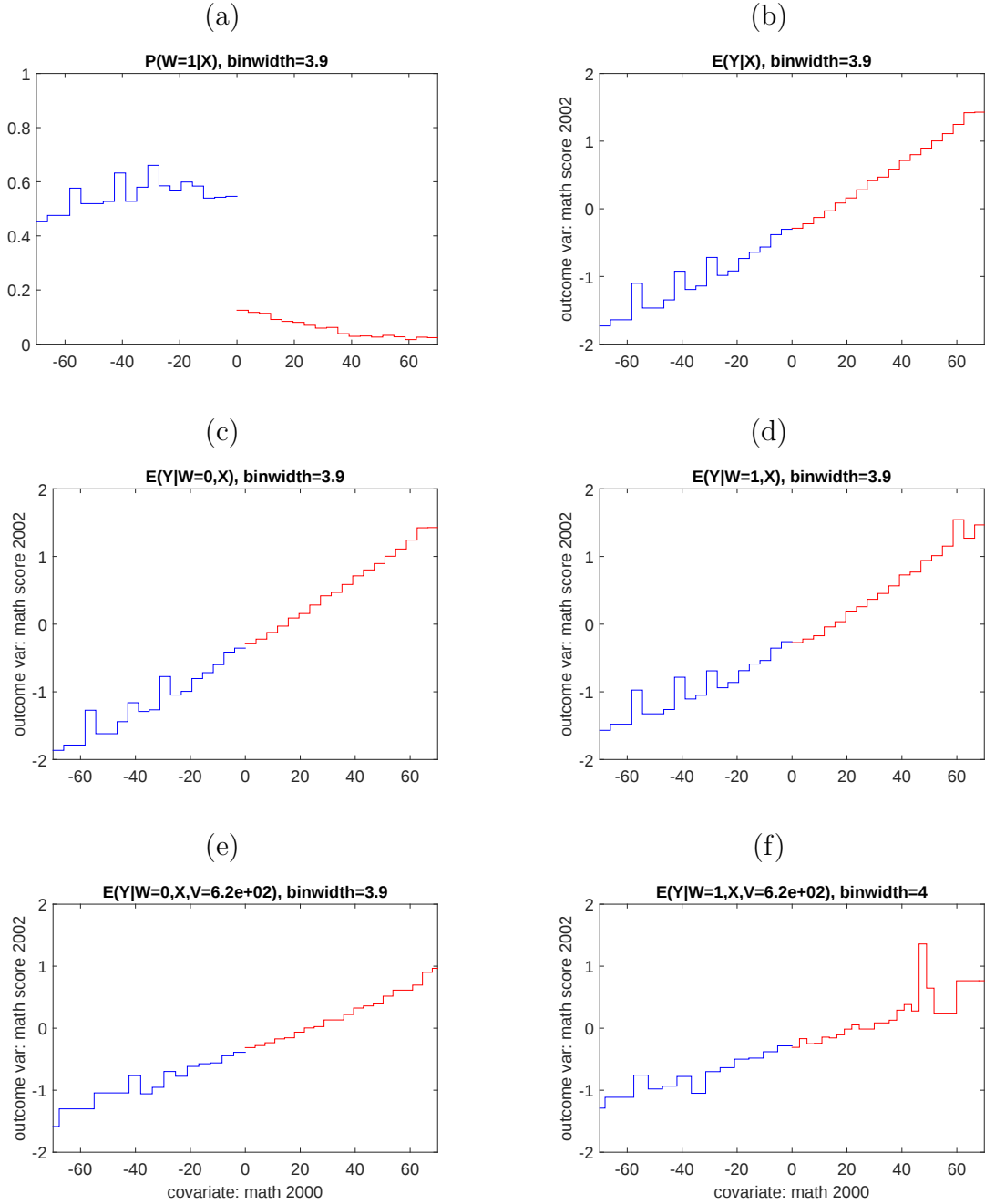
The forcing variable X is the 2001 minimum score between reading and math minus the cutoff for passing; the outcome variable Y is the standardized math score in 2002, and W is the participation indicator; the covariate V is the math score in 2000.

Table 4: Matsudaira Data

Estimates	Estimand	Estimate	s.e.	p-value
Participation	π_c	0.400	(0.010)	
	π_a	0.131	(0.005)	
	π_n	0.469	(0.008)	
Cond. Means of Potential Outcomes	$\mathbb{E}[Y_i(1) X_i = T^*, G_i = c]$	-0.213	(0.016)	
	$\mathbb{E}[Y_i(1) X_i = T^*, G_i = a]$	-0.359	(0.022)	
	$\mathbb{E}[Y_i(0) X_i = T^*, G_i = c]$	-0.413	(0.023)	
	$\mathbb{E}[Y_i(0) X_i = T^*, G_i = n]$	-0.325	(0.012)	
LATE	$\frac{\mathbb{E}[Y_i^{\text{obs}} X_i=T_+^*] - \mathbb{E}[Y_i^{\text{obs}} X_i=T_-^*]}{\mathbb{E}[W_i^{\text{obs}} X_i=T_+^*] - \mathbb{E}[W_i^{\text{obs}} X_i=T_-^*]}$	0.200	(0.026)	
OLS	$\mathbb{E}[Y_i^{\text{obs}} W_i^{\text{obs}} = 1, X_i = T_\pm^*] - \mathbb{E}[Y_i^{\text{obs}} W_i^{\text{obs}} = 0, X_i = T_\pm^*]$	0.080	(0.012)	
<u>Tests</u>				
Hausman	$\frac{\mathbb{E}[Y_i^{\text{obs}} X_i=T_+^*] - \mathbb{E}[Y_i^{\text{obs}} X_i=T_-^*]}{\mathbb{E}[W_i^{\text{obs}} X_i=T_+^*] - \mathbb{E}[W_i^{\text{obs}} X_i=T_-^*]} - \mathbb{E}[Y_i^{\text{obs}} W_i^{\text{obs}} = 1, X_i = T_\pm^*] + \mathbb{E}[Y_i^{\text{obs}} W_i^{\text{obs}} = 0, X_i = T_\pm^*]$	0.120	(0.025)	0.000
Angrist	$\frac{\mathbb{E}[Y_i^{\text{obs}} X_i=T_+^*] - \mathbb{E}[Y_i^{\text{obs}} X_i=T_-^*]}{\mathbb{E}[W_i^{\text{obs}} X_i=T_+^*] - \mathbb{E}[W_i^{\text{obs}} X_i=T_-^*]} - \mathbb{E}[Y_i^{\text{obs}} W_i^{\text{obs}} = 1, X_i = T_+^*] + \mathbb{E}[Y_i^{\text{obs}} W_i^{\text{obs}} = 0, X_i = T_-^*]$	0.234	(0.043)	0.000
Conditional Jumps	$\mathbb{E}\left[Y_i^{\text{obs}} \middle W_i^{\text{obs}} = 0, X_i = T_-^*\right] - \mathbb{E}\left[Y_i^{\text{obs}} \middle W_i^{\text{obs}} = 0, X_i = T_+^*\right]$	0.041	(0.014)	0.004
	$\mathbb{E}\left[Y_i^{\text{obs}} \middle W_i^{\text{obs}} = 1, X_i = T_-^*\right] - \mathbb{E}\left[Y_i^{\text{obs}} \middle W_i^{\text{obs}} = 1, X_i = T_+^*\right]$	0.110	(0.024)	0.000
Joint F-test		31.083		0.000
<u>Adj. for Covariate</u>				
Conditional	$\mathbb{E}\left[\mathbb{E}\left[Y_i^{\text{obs}} \middle W_i^{\text{obs}} = 0, X_i = T_-^*, V_i\right] \middle X_i = T_\pm^*\right]$	0.022	(0.014)	0.104
	$-\mathbb{E}\left[\mathbb{E}\left[Y_i^{\text{obs}} \middle W_i^{\text{obs}} = 0, X_i = T_+^*, V_i\right] \middle X_i = T_\pm^*\right]$			
Jumps	$\mathbb{E}\left[\mathbb{E}\left[Y_i^{\text{obs}} \middle W_i^{\text{obs}} = 1, X_i = T_-^*, V_i\right] \middle X_i = T_\pm^*\right]$	0.074	(0.022)	0.001
	$-\mathbb{E}\left[\mathbb{E}\left[Y_i^{\text{obs}} \middle W_i^{\text{obs}} = 1, X_i = T_+^*, V_i\right] \middle X_i = T_\pm^*\right]$			
Joint F-test		14.118		0.001

The forcing variable X is the 2001 minimum score between reading and math minus the cutoff for passing; the outcome variable Y is the standardized math score in 2002, and W is the participation indicator; the covariate V is the math score in 2000. Non-parametric estimates are obtained by local linear regression and optimal bandwidth $h = 28$ by Imbens and Kalyanaraman (2012). The last three estimates are conditional on $V_i = v = 620.3$. The standard errors of all estimates are computed using 1000 bootstrap iterations. The F-test is a χ^2 statistic computed as a quadratic form of the 2×1 vector of the conditional jumps inversely weighted by the covariance matrix of such vector.

Figure 2: Matsudaira



Notes: The forcing variable X is the 2001 minimum score between reading and math minus the cutoff for passing; the outcome variable Y is the standardized math score in 2002, and W is the participation indicator; the covariate V is the math score in 2000. Step functions are made of averages within each bin, and plots have the same number of equal sized bins on each side of the cutoff. For (e) and (f), the average within each bin is the average of predicted Y at $V = v$ using a linear regression of Y on a constant, X and V .