

Mehta, Nirav

**Working Paper**

## An economic approach to generalizing findings from regression-discontinuity designs

CIBC Working Paper, No. 2015-6

**Provided in Cooperation with:**

Centre for Human Capital & Productivity (CHCP), Department of Economics, University of Western Ontario

*Suggested Citation:* Mehta, Nirav (2016) : An economic approach to generalizing findings from regression-discontinuity designs, CIBC Working Paper, No. 2015-6, The University of Western Ontario, CIBC Centre for Human Capital and Productivity, London (Ontario)

This Version is available at:

<https://hdl.handle.net/10419/180874>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

**An Economic Approach to Generalizing  
Findings from Regression-Discontinuity  
Designs**

by

**Nirav Mehta**

**Working Paper # 2015-6**

**November 2015**

Updated May 21, 2016



***CIBC Working Paper Series***

Department of Economics  
Social Science Centre  
Western University  
London, Ontario, N6A 5C2  
Canada

This working paper is available as a downloadable pdf file on our website  
<http://economics.uwo.ca/centres/cibc/>

# An Economic Approach to Generalizing Findings from Regression-Discontinuity Designs

Nirav Mehta\*

May 21, 2016

## **Abstract**

Regression-discontinuity (RD) designs estimate treatment effects at a cutoff. This paper shows what can be learned about average treatment effects for the treated (ATT), untreated (ATUT), and population (ATE) if the cutoff was chosen to maximize the net gain from treatment. The ATT must be positive. Without capacity constraints, the RD estimate bounds the ATT from below and the ATUT from above, implying bounds for the ATE. Optimality of the cutoff rules out constant treatment effects. Testable implications of cutoff optimality are derived. Bounds are looser if the capacity constraint binds. The results are applied to existing RD studies.

---

\*Department of Economics, University of Western Ontario. Email: nirav.mehta@uwo.ca.

# 1 Introduction

Regression-discontinuity (RD) designs are a popular tool for program evaluation due to the ubiquity of cutoff-based treatment assignment and agnosticism they afford researchers. However, it is not always clear how best to use such estimates to inform policy. Suppose an impact evaluation of a program using an RD design found the treatment effect at the cutoff to be positive, but small. Is this evidence the program should be terminated? Because treatment effects likely vary, it is useful to extend findings from RD designs to people away from the cutoff.

The goal of this paper is to demonstrate how combining an RD estimate with a simple economic model can deliver useful information about treatment effects in certain contexts. Researchers using RD designs typically focus on treatment effects at the cutoff, an approach that has the ostensible benefit of imposing minimal structure. Though such estimates can help decide whether to extend the treatment at the margin, this approach implicitly assumes that the treatment effect at the cutoff is completely uninformative about treatment effects elsewhere in the distribution. This assumption may not always be appropriate. This paper considers a program administrator interested in maximizing the gain from treatment, net of treatment costs, but who, as is often the case in real-world applications, has been constrained to assign treatment using a cutoff rule. By imposing structure not on treatment effects, but on the economic environment, I show that we can learn about important characteristics of the distribution of treatment effects in cases where there is reason to believe the administrator has information about the costs and benefits of treatment. That is, the choice of cutoff may reveal information key to understanding the overall costs and benefits of a program.<sup>1</sup>

Combining an estimate of the treatment effect at the cutoff with a simple model of cutoff choice yields many insights. The most basic inference we can draw about treatment effects is that the average effect of treatment on the treated (ATT) must be positive if the marginal cost of treatment is positive. There is also a basic testable implication of the model: if the treatment effect at the cutoff is negative, then we can reject cutoff optimality. Several results obtain if the cutoff does not reflect a binding capacity constraint. Optimality implies that the treatment effect is increasing at the cutoff, an implication that can be tested using results from Dong and Lewbel (2015). Additional results relate the treatment effect at the cutoff to the distribution of treatment effects. First, the RD estimate at the treatment cutoff provides a lower bound for the ATT; if this were not true the administrator could have obtained higher utility by moving the cutoff. Intuitively, the administrator will not

---

<sup>1</sup>See Heckman et al. (1997) for a discussion of heterogeneous treatment effects in the context of an experimental setting. See Heckman and Smith (1998) for a discussion of how to link information about program benefits with conventional cost-benefit analysis and welfare calculations.

place the cutoff where the gain from treatment is very large if, as is commonly assumed by practitioners, treatment effects are smooth. Second, the fact that the administrator chose not to extend treatment to certain students provides an upper bound for the average effect of treatment on the untreated (ATUT). These bounds only require the RD estimate, not actual estimates of the cost of treatment. Additionally, we can rule out a constant treatment effect, a finding that relates to the literature comparing findings from RD studies with those from experiments (e.g., Black et al. (2007) and Buddelmeyer and Skoufias (2004)). Finally, the bounds on the ATT and ATUT provide informative bounds on the average treatment effect (ATE). Bounds are looser if the chosen cutoff reflects a binding capacity constraint. However, a new testable implication also emerges: if the program is subsequently expanded until the constraint no longer binds, the RD estimate will be lower than when it had been when the constraint was binding.

These results have implications for the use of RD estimates in policymaking. Perhaps the most striking result is that, because an unconstrained administrator is unlikely to choose a cutoff where the gain is quite large, one may incorrectly surmise from RD estimates that certain programs are ineffective and eliminate them, even though in reality they are quite effective for the treated population. In fact, such a mistake would be more likely for a program with a very low marginal cost, holding constant the ATT, because an unconstrained optimizing administrator would extend treatment to units until the gain, i.e., marginal benefit, equaled this low marginal cost. If the cutoff reflects a binding capacity constraint, then the RD estimate will exceed the marginal cost of treatment, which may help explain why it is sometimes difficult to “scale up” successful interventions to larger populations (see, e.g., Elmore (1996) and Sternberg et al. (2006)).

To make these findings more concrete, the results are illustrated using three recent empirical applications. The studies used in the applications, Hoekstra (2009), Lindo et al. (2010), and Bui et al. (2014), all exploit discontinuities in treatment assignment rules to study questions in the economics of education, and cover a mix of cases where the capacity constraint likely does and does not bind. I first show that bounds obtained for the sharp RD design can be extended to the “fuzzy” design used in two of the applications. I also formally test the necessary conditions of optimality, and find that I cannot reject the model assumptions that the cutoffs were chosen optimally by informed program administrators.<sup>2</sup>

There is a long tradition in economics, starting with Roy (1951), of using revealed preferences to inform empirical work about information unobservable to the econometrician. This

---

<sup>2</sup>Applications need not be restricted to the economics of education. For example, the findings from this paper might apply to a job training program in which the program officer receives a bonus based on the increase in wages. I reiterate that one could test whether the environment studied in this paper was applicable for this, or any other context by checking that the treatment effect is nondecreasing at the cutoff.

paper simply makes clear what we could learn by embedding the choice of treatment cut-off within a larger decision problem. To most clearly demonstrate what can be learned by taking into account the administrator’s context, I assume she knows both the distribution of treatment effects and her cost function. Though for a different context, the insights of this paper are similar to those in Heckman and Vytlacil (2007), who build on the work of Björklund and Moffitt (1987) by using the optimality of individual decision-making in the context of a Roy model and the marginal treatment effect (MTE) to bound treatment effects for non-marginal units.

By embedding an RD design within a simple economic model, this paper contributes to several literatures. First, it adds to the literature examining technical features of RD designs (Hahn et al. (2001), Van der Klaauw (2008)) by demonstrating how inferences from RD designs can be generalized by using a simple theoretical framework. This paper also relates to the debate about the usefulness of discontinuity and other LATE estimators of treatment effects (Heckman et al. (1999), Heckman and Urzua (2010), Imbens (2010)). In terms of goals, this paper is most similar to Manski and Pepper (2000), which uses monotone instrumental variables to bound treatment effects. This paper takes a different approach by assuming optimality of assignment to treatment status, while making minimal assumptions about the responses of agents to the treatment.

This paper also contributes to a literature seeking to extend results from RD designs. Dong and Lewbel (2015) show that the differentiability assumptions typically invoked to estimate RD models can be exploited to estimate the derivative of the treatment effect. In a similar vein, DiNardo and Lee (2011) show how a Taylor expansion around the cutoff can be used to estimate the ATT. Angrist and Rokkanen (2013) invoke a conditional independence assumption to generalize findings from RD studies. Due to the different type of assumption made (i.e., statistical versus economic), this paper complements their work. There is also an extensive literature studying the validity of RD designs.<sup>3</sup> This paper treats the RD design as valid and instead examines how findings from such studies can be generalized to other parts of the population.

Section 2 lays out the model of the administrator’s problem, which serves as the foundation for the results in Section 3. Section 4 illustrates the results using empirical applications. Section 5 shows how a model of human capital formation can help interpret the results, and discusses policy implications as well as variations on the informational assumptions made in this paper.

---

<sup>3</sup>This literature is reviewed in Imbens and Lemieux (2008). See Lee and Card (2008) and McCrary (2008) for examples.

## 2 Model

Consider a program administrator who can assign students to a training program. The administrator knows how effective the program would be for any given student and also knows the cost of enrolling students in the program. Due to institutional reasons, she is constrained to choose a cutoff rule for assigning the treatment, above which students are enrolled. The choice of cutoff-based treatment assignment captures the fact that many real-world policies are discrete in nature (Ferrall and Shearer (1999)).

There is a measure one of students, uniformly distributed over  $[0, 1]$ . The students are indexed by  $x \in [0, 1]$ ; this index doubles as the running variable in the discontinuity design. Let  $\tau(x) = 1$  if  $x$  is given the treatment and 0 otherwise. The administrator is constrained to choose a cutoff rule where  $\tau(x) = 1$  if and only if  $x \geq \kappa$  for some  $\kappa \in (0, 1)$ .<sup>4</sup> To simplify exposition, I assume a “sharp” RD design and perfect compliance, which means that students with indices of  $\kappa$  or greater receive the treatment (i.e., participate in the program) and students with indices less than  $\kappa$  don’t receive the treatment (i.e., don’t participate in the program).<sup>5</sup> Let  $\kappa^*$  denote the treatment cutoff chosen by the administrator. The measure of students receiving treatment is  $\mu = \int_0^1 \tau(x)dx$ . The administrator faces a cost of treating  $\mu$  students,  $c(\mu)$ , which may capture an implicit budget constraint. Results from the administrator’s unconstrained problem, the model for which is presented in this section, are presented in Section 3.1. Section 3.2 introduces capacity constraints and then analyzes that problem.

Let  $Y_\tau(x)$  denote student  $x$ ’s outcome under treatment group  $\tau$ . For example, this may be the wage earned as a function of being enrolled in a training program. The treatment effect is  $\Delta(x) \equiv Y_1(x) - Y_0(x)$ , and, as is made explicit shortly in Assumption 1, is known by the administrator. The administrator need not be perfectly informed about students’ potential outcomes; so long as the administrator has an unbiased signal of  $\Delta(x)$ , uncertainty about treatment effects does not affect the analysis.<sup>6</sup> In general, there may be multiple students with the index  $x$ , and a distribution of treatment effects among these students at  $x$ . In this case, one could think of  $\Delta(x)$  as the expected gain from treating students at  $x$ , i.e.  $\Delta(x) = \int \Delta(z)f(z|x)dz$ , where  $f(\cdot)$  describes heterogeneity at  $x$ . As is common in studies employing

---

<sup>4</sup>If the administrator could choose precisely which  $x \in [0, 1]$  to treat we can make the inference that the gain from serving those students was at least as large as the marginal cost of serving them, point-wise. This stronger informational assumption would imply that gains were positive for *all* treated students, as opposed to positive *on average*. Therefore, we could also bound from below the share of students who would gain from treatment:  $\int_0^1 \mathbf{1}\{\Delta(x) > 0\}dx \geq \mu$ . This would be relevant if, say, the population voted on whether to implement the treatment.

<sup>5</sup>Section 4 shows that the theoretical results are identical under a “fuzzy” design where, instead of perfect compliance, the probability of participation discontinuously changes at the treatment cutoff.

<sup>6</sup>This is shown in Appendix B.

discontinuity designs, the stable-unit-treatment-value-assumption (SUTVA) is maintained here (Rubin (1980)), ruling out general equilibrium effects and other interactions between other units' treatment status and one's own treatment effect, such as endogenous social interactions.

The fundamental problem of causal inference is that we only observe each student in one treatment condition, making it difficult to recover the entire function  $\Delta(\cdot)$ . What can we say about  $\Delta(\cdot)$  knowing 1) that  $\kappa^*$ , the treatment cutoff, was chosen by the administrator and 2) the value of  $\Delta(\kappa^*)$ , from a RD design? Though I find that we can not say much about  $\Delta(\cdot)$  for particular  $x$  who are not at the cutoff, we will be able to bound averages of  $\Delta(\cdot)$  over different intervals, listed in Definition 1.

**Definition 1** (Treatment effects of interest). *I focus on the:*

- *Average Effect of Treatment on the Treated (ATT):*  $\int_{\kappa^*}^1 \frac{\Delta(x)}{1-\kappa^*} dx$ ,
- *Average Effect of Treatment on the Untreated (ATUT):*  $\int_0^{\kappa^*} \frac{\Delta(x)}{\kappa^*} dx$ , and
- *Average Treatment Effect (ATE):*  $\int_0^1 \Delta(x) dx$ .

The administrator's problem is to choose a cutoff to maximize the total treatment effect, net treatment cost:<sup>7</sup>

$$\max_{\tilde{\kappa}} \beta \left( \int_{\tilde{\kappa}}^1 \Delta(x) dx \right) - c(1 - \tilde{\kappa}), \quad (1)$$

where  $\beta$  measures how much the administrator values the effect of the program in terms of the cost of treatment. Though in principle identified when the cost function is known to the researcher,  $\beta$  is normalized to one to simplify exposition. The administrator has an outside option of zero. This objective function is similar to those studied in Manski (2003, 2004, 2011), where a utilitarian social planner takes an action to maximize expected welfare (i.e., the gain net the cost of treatment), as well as those in studies of statistical discrimination such as Knowles et al. (2001), Anwar and Fang (2006), and Brock et al. (2011), where police officers face a cost of pulling over motorists to maximize expected hit rates. The economic rationale for studying a utilitarian social planner is that a system of lump-sum transfers could then be designed to redistribute total output in such a manner as the social planner saw fit; that is, the utilitarian objective corresponds to the efficient allocation.<sup>8</sup>

<sup>7</sup>For an example of a slightly different objective function, see Heinrich et al. (2002), who study treatment decisions when administrators face performance standards.

<sup>8</sup>In the baseline case presented here, the administrator is a utilitarian who weighs gains for all students equally. See Appendix C for a case where gains are not weighed equally; the results derived in Section 3.1 are also obtained there.



**Assumption 1.** *The following assumptions about costs and benefits of treatment are maintained throughout this section:*

- (i) *The cost of treatment is known by the administrator, and is strictly increasing and linear in the number of units treated, i.e.  $c(\mu) = \chi\mu$ , where  $\chi = c'(\cdot) > 0$  denotes the constant marginal cost of treatment.*
- (ii) *Treatment effects  $\Delta(\cdot)$  are differentiable in  $x$  and known by the administrator.*
- (iii) *There exist finite lower and upper bounds of  $\Delta(\cdot)$ . Denote these by  $\underline{\Delta} \in \mathcal{R}$  and  $\overline{\Delta} \in \mathcal{R}$ , respectively.*

Assumption 1(i) implies that the marginal cost of providing treatment is known and strictly positive; the assumption of a linear cost function is made to simplify exposition. The assumption that the cost function is linear is also made in Manski (2011)’s analysis of optimal treatment choices, who assumes costs are separable across treated units. All of the following results would obtain in the more general case where the marginal cost of treatment was nonincreasing in  $\mu$ , i.e., where the cost of treatment is weakly concave in  $\mu$  (see Appendix D for details).<sup>9</sup> The first part of Assumption 1(ii), i.e. differentiability of  $\Delta(\cdot)$ , is typically invoked in applications of RD designs, which control for a smooth (typically polynomial or smoothed non-parametric) function of the running variable.<sup>10</sup> The second part of Assumption 1(ii), that  $\Delta(\cdot)$  is known by the administrator, produces a testable implication (as is shown in the next section). Assumption 1(iii) means that the set of outcomes  $Y_\tau(\cdot)$  has finite support, which makes sense for outcomes such as wages, test scores, or probabilities.<sup>11</sup>

### 3 Results

Section 3.1 develops results for the administrator’s problem when there is no capacity constraint. Section 3.2 develops results for the administrator’s problem in the presence of a capacity constraint.

---

<sup>9</sup>If one thought marginal costs were increasing in a context of interest, the bounds would have to be adjusted accordingly (see Appendix D for details).

<sup>10</sup>Note that  $\Delta(\cdot)$  only needs to be smooth local to the chosen cutoff. The assumption that it is globally smooth is only made to simplify exposition.

<sup>11</sup>Note that  $Y_\tau(\cdot)$  may contain negative values, which may capture a negative treatment effect or a positive opportunity cost of taking an ineffective treatment.

### 3.1 Results without capacity constraints

The goal of this paper is to link the chosen cutoff  $\kappa^*$  to treatment effects  $\Delta(\cdot)$ . Therefore, I first provide necessary and sufficient conditions to characterize  $\kappa^*$ , in terms of  $\Delta(\cdot)$  and the cost function  $c(\mu) = \chi\mu$ .

**Condition 1** (Necessity). *The following necessary conditions must hold for  $\kappa^*$ :*

$$(i) \text{ MB=MC: } \Delta(\kappa^*) = c'(1 - \kappa^*) = \chi$$

$$(ii) \text{ Increasing MB: } \Delta'(\kappa^*) \geq 0.$$

*Proof.* Differentiate the administrator's problem (1) with respect to  $\tilde{\kappa}$  to obtain (i). Note that if the derivative is negative at a candidate solution satisfying (i), the administrator would gain by not treating students just above  $\kappa^*$ , thereby obtaining (ii).  $\square$

Condition 1(i) will play a key role in bounding treatment effects, and also provides testable implication of cutoff optimality, in that a negative LATE at the cutoff (i.e.,  $\Delta(\kappa^*) < 0$ ) would contradict Assumption 1, because  $\chi > 0$ . Condition 1(ii) is another testable implication of the model's assumptions that the administrator is acting optimally and with knowledge of  $\Delta(\cdot)$ . It can be tested using methods developed in Dong and Lewbel (2015). That is, the model's maintained Assumption 1 would be falsified if one rejected that  $\Delta'(\kappa^*) \geq 0$ . Condition 1 need not be sufficient; there can be multiple cutoffs satisfying it.

**Assumption 2** (Unique maximand).  $\kappa^*$  uniquely maximizes the administrator's problem (1).

Assumption 2 implies that  $\Delta(\cdot)$  crosses  $c'(\cdot)$  finitely many times and is made to simplify exposition. Note that uniqueness of  $\kappa^*$  implies that Condition 1(ii) should be strict (i.e.  $\Delta'(\kappa^*) > 0$ ). To guarantee uniqueness, inspection of (1) implies two additional conditions sufficient for characterizing  $\kappa^*$ .

**Condition 2** (Sufficiency). *The fact the program was implemented implies that the total gain from treating those students was at least as large as the total costs, i.e.:*

$$\text{Participation: } \int_{\kappa^*}^1 \Delta(x)dx \geq c(1 - \kappa^*) = \chi(1 - \kappa^*). \quad (2)$$

*The fact the program was not extended to  $\hat{\kappa} < \kappa^*$  implies that treating these students would be sub-optimal, i.e.:*

$$\int_{\hat{\kappa}}^{\kappa^*} \Delta(x)dx < c(1 - \hat{\kappa}) - c(1 - \kappa^*) = \chi(\kappa^* - \hat{\kappa}). \quad (3)$$

Intuitively, Condition 2 uses revealed preferences to make statements about the gains and costs of treating students who are either treated or untreated. It must be worthwhile to have treated the treated students, and it could not have been worthwhile to treat the untreated. Two corollaries immediately follow.

**Corollary 1.**  $\Delta(\cdot)$  cannot be constant.

**Corollary 2.**  $\Delta(\cdot)$  is not on average decreasing in  $x$ .

*Proof.* Condition 1(ii) already rules out  $\Delta(\cdot)$  decreasing at  $\kappa^*$ . Consider the behavior of  $\Delta(\cdot)$  for positive measures of students away from the cutoff. The second part of Condition 2 says that it must be the case that  $\int_{\hat{\kappa}}^{\kappa^*} (\Delta(x) - \chi) dx < 0$ . Moreover, we know that  $\Delta(\kappa^*) = \chi$ . Therefore, if  $\Delta(\cdot)$  were on average decreasing in  $x$ , this inequality would be violated.  $\square$

It is often said that there is no reason to believe treatment effects would be the same for students away from the cutoff, though this notion is not always reflected in empirical implementations. Corollary 1 strengthens this statement by ruling out constant treatment effects. Corollary 2 tells us that the ATT will be larger than the ATUT. Intuitively, if the administrator has decided to provide the treatment to students above, as opposed to below,  $\kappa^*$ , the former must have gained more from the treatment.

Corollaries 1-2 are fairly weak statements about the global behavior of  $\Delta(\cdot)$ . Therefore, I next examine what can be deduced about averages of treatment effects for subsets of students.

**Proposition 1.** *The ATT is bounded below by the LATE at the treatment cutoff.*

*Proof.* Divide (2) by the measure of treated students  $(1 - \kappa^*)$  and combine this with Condition 1(i) to obtain

$$\underbrace{\frac{\int_{\kappa^*}^1 \Delta(x) dx}{1 - \kappa^*}}_{\text{ATT}} \geq \frac{\chi(1 - \kappa^*)}{1 - \kappa^*} = \chi = \underbrace{\Delta(\kappa^*)}_{\text{LATE at } \kappa^*} . \quad \square$$

A corollary immediately follows.

**Corollary 3.** *The ATT is positive.*

*Proof.* This follows directly from Proposition 1 because  $\chi > 0$ , by Assumption 1(i).  $\square$

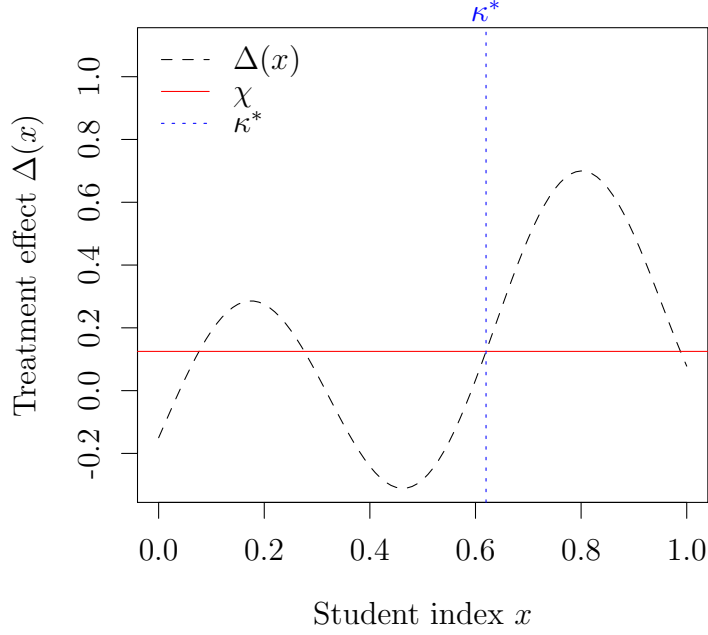


Figure 1: Example  $\Delta(\cdot)$  with optimal cutoff  $\kappa^*$

Proposition 1 shows that the discontinuity-based estimate provides a lower bound for the average effect of treatment on the treated. *In other words, estimates from discontinuity-based designs will understate the effect of treatment on the treated.* Intuitively, the administrator chooses  $\kappa^*$  to set the marginal benefit from providing the treatment equal to the marginal cost, which is lower than the average cost of providing treatment to treated students (i.e., the marginal cost). The fact that the administrator chose to implement the program, however, implies that the gain to treating those students must have been at least as large as the total cost of treating them. Note that the level of the marginal cost does not need to be known by the researcher.

However, although we have quite a bit of information about averages of treatment effects  $\Delta(\cdot)$  over some intervals of interest, we cannot make statements about  $\Delta(x)$  for students  $x \neq \kappa^*$ . Figure 1 plots an example treatment effect function  $\Delta(x)$  (dashed black curve) and the marginal cost of treatment (solid red horizontal line) against the student index  $x$ , which ranges from 0 to 1, and the optimal cutoff  $\kappa^*$  (dotted blue vertical line). This figure shows a case satisfying Conditions 1-2 where there are also untreated students with gains greater than their cost of treatment and treated students with gains smaller than their cost of treatment. Although Corollary 2 rules out a decreasing treatment effect, it could be the

case that  $\Delta(\cdot)$  increases for some  $x < \kappa^*$ . Therefore, it is useful to make a statement about the average effect of extending treatment to the untreated. In particular, we can bound averages of  $\Delta(\cdot)$  itself for subsets of untreated students.

**Proposition 2.** *There exists an informative upper bound for  $\int_a^b \Delta(x)dx$  for  $0 \leq a < b \leq \kappa^*$ .*

*Proof.* Suppose we would like to characterize  $\Delta(\cdot)$  for values less than  $\hat{x} < \kappa^*$ . Let  $\hat{\mu}$  be the measure of students under consideration and split (3) into two parts at  $\hat{x}$  and rearrange terms:

$$\int_{\hat{x}-\hat{\mu}}^{\hat{x}} \Delta(x)dx < c(1-(\hat{x}-\hat{\mu}))-c(1-\kappa^*)-\int_{\hat{x}}^{\kappa^*} \Delta(x)dx \Rightarrow \int_{\hat{x}-\hat{\mu}}^{\hat{x}} \Delta(x)dx < \chi(\kappa^*-(\hat{x}-\hat{\mu}))- \underline{\Delta}(\kappa^*-\hat{x}), \quad (4)$$

□

where the implication follows from Assumption 1(iii).

The right hand side of the expression in Proposition 2 provides an upper bound for the gain from treating students  $x \in [\hat{x}-\hat{\mu}, \hat{x}]$ . Because we do not know  $\Delta(\cdot)$ , by assuming the worst possible treatment effect ( $\underline{\Delta}$ ) we can find an upper bound for how large it could be for a measure of students  $\hat{\mu}$  and satisfying the individual rationality constraint from Condition 2. Intuitively, this upper bound grows the further below the cutoff we go. To gain more intuition for Proposition 2, rearrange (4) and divide by the measure of students under consideration  $\hat{\mu}$  to obtain:

$$\int_{\hat{x}-\hat{\mu}}^{\hat{x}} \frac{\Delta(x)}{\hat{\mu}} dx < (\chi - \underline{\Delta}) \left( \frac{\kappa^* - \hat{x}}{\hat{\mu}} \right) + \chi. \quad (5)$$

The left side of (5) is the average treatment effect among students  $x \in [\hat{x}-\hat{\mu}, \hat{x}]$ . First, consider the extreme scenario where we want an upper bound for the treatment effect for student  $\hat{x}$ ,  $\Delta(\hat{x})$ . Take the limit of (5) as the additional treated students go to zero:

$$\underbrace{\lim_{\hat{\mu} \rightarrow 0} \left( \int_{\hat{x}-\hat{\mu}}^{\hat{x}} \frac{\Delta(x)}{\hat{\mu}} dx \right)}_{\Delta(\hat{x})} < \lim_{\hat{\mu} \rightarrow 0} \left( (\chi - \underline{\Delta}) \left( \frac{\kappa^* - \hat{x}}{\hat{\mu}} \right) + \chi \right) = \infty,$$

i.e. the expression becomes completely uninformative when we evaluate it for measure zero of students to bound  $\Delta(\cdot)$  at a point. However, consider the other extreme where  $\hat{\mu} = \hat{x}$ , i.e.

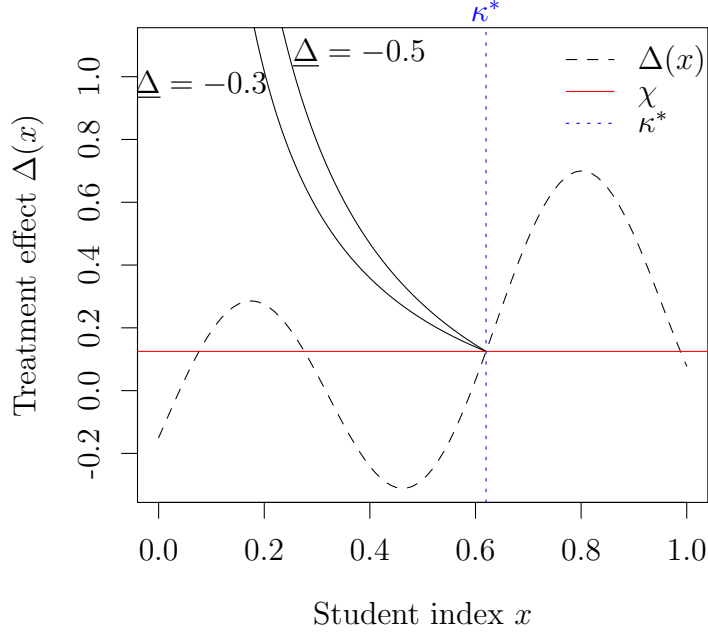


Figure 2: Upper bounds for the remaining ATUT, for different values of  $\underline{\Delta}$

the administrator is considering extending treatment to all students below  $\hat{x}$ :

$$\int_0^{\hat{x}} \frac{\Delta(x)}{\hat{x}} dx < (\chi - \underline{\Delta}) \left( \frac{\kappa^*}{\hat{x}} \right) + \underline{\Delta}. \quad (6)$$

Equation (6) says that the upper bound on the average treatment effect among students  $x \leq \hat{x}$  (the left side) grows the further  $\hat{x}$  goes below  $\kappa^*$ , the higher is the marginal cost  $\chi$ , and the lower is the lower bound  $\underline{\Delta}$ . This relationship is depicted in Figure 2, which overlays the solid black curves showing the upper bound for the average gain to treating students  $x < \hat{x}$  on top of Figure 1, for different values of  $\underline{\Delta}$ .

Setting the measure of students to whom the treatment is extended equal to  $\kappa^*$  provides the following result about the ATUT.

**Corollary 4.** *The ATUT is bounded above by the LATE at the treatment cutoff.*

*Proof.* Let  $\hat{x} = \hat{\mu} = \kappa^*$  in (4) and divide through by  $\kappa^*$  to obtain the result:

$$\underbrace{\int_0^{\kappa^*} \frac{\Delta(x)}{\kappa^*} dx}_{\text{ATUT}} < \underbrace{\frac{c(1) - c(1 - \kappa^*)}{\kappa^*}}_{>0, < \infty} = \chi, \quad (7) \quad \square$$

where the right hand side is positive from Assumption 1(i).

Analogously to the upper bound for the ATT, although Corollary 4 bounds the average of treatment effects for all untreated students, there is no informative (i.e. greater than  $\underline{\Delta}$ ) lower bound.

Finally, the next result shows how the prior results can be used to bound the average treatment effect (ATE).

**Corollary 5.** *The ATE has informative bounds.*

*Proof.* To form the lower bound for the ATE, note that measure  $\kappa^*$  students are untreated, and, by Assumption 1(iii), the treatment effect for each student cannot be worse than  $\underline{\Delta}$ . Analogously,  $1 - \kappa^*$  students are treated, and (2) implies the ATT is no smaller than  $\chi$ . Integrate and sum the two parts to form  $\Delta_{\kappa^*\chi}^{LB} \equiv \underline{\Delta}\kappa^* + \chi(1 - \kappa^*)$ . To form the upper bound for the ATE, note that (3) implies that the ATUT is no larger than  $\chi$ . By Assumption 1(iii), the treatment effect for any treated student cannot exceed  $\overline{\Delta}$ . Integrate and sum to form  $\Delta_{\kappa^*\chi}^{UB} \equiv \chi\kappa^* + \overline{\Delta}(1 - \kappa^*)$ .  $\square$

Corollary 5 shows that higher values of  $\kappa^*$  tighten the upper bound on the ATE while loosening the lower bound on the ATE. Intuitively, treating fewer students increases the share of untreated students, who have an upper bound of  $\chi$ , while increasing the share of students with very low values of  $\Delta(\cdot)$ , i.e.,  $\underline{\Delta}$ .

To summarize, optimality of the treatment cutoff  $\kappa^*$  implies a lower bound on the average effect of treatment on the treated (ATT) and an upper bound on the average effect of treatment on the untreated (ATUT). The average treatment effect for the population (ATE) combines the above bounds.<sup>12</sup> Put another way, we can exploit information revealed by the optimizing behavior of the administrator to extrapolate from the LATE at the treatment cutoff, which is often available but can apply only to measure zero of the population, to obtain bounds for the treated, untreated, and the entire population. Optimality further implies that  $ATUT < \Delta(\kappa^*) \leq ATT$ . Moreover, we can contrast what can be said about  $\Delta(\cdot)$  for students with indices  $x > \kappa^*$  and  $\Delta(\cdot)$  for student with indices  $x < \kappa^*$ . Because the treatment is being provided to all students in the treated group, we cannot separate how treatment effects accumulate for students  $x > \kappa^*$ . But the fact that the administrator is not choosing to extend (i.e., decrease) the cutoff to student  $\hat{x} < \kappa^*$  provides us information for how large treatment effects can possibly be for students  $x \in [\hat{x}, \kappa^*)$ .

---

<sup>12</sup>Most bounds remain the same if, instead of being constant, the marginal cost of treatment is decreasing. Bounds for the ATE are looser when the marginal cost of providing treatment is increasing and tighter when it is decreasing. Results are in Appendix D.

### 3.2 Results with capacity constraints

Suppose now that the administrator faces a capacity constraint,  $\bar{\mu}$ . Then, the administrator solves

$$\begin{aligned} \max_{\tilde{\kappa}} \quad & \beta \left( \int_{\tilde{\kappa}}^1 \Delta(x) dx \right) - c(1 - \tilde{\kappa}) \\ \text{s.t.} \quad & 1 - \tilde{\kappa} \leq \bar{\mu}, \end{aligned} \tag{1}$$

where we continue to maintain Assumptions 1-2 from problem (1).

If the desired (i.e., unconstrained) measure of treated students does not exceed capacity, i.e.,  $\mu^*(\equiv 1 - \kappa^*) \leq \bar{\mu}$ , then the constraint does not bind, and the optimal cutoff  $\kappa^*$  and resulting analysis are unaffected. This means the results from Section 3.1 apply here as well. By definition, if the capacity constraint is binding the measure of students treated must be strictly less than the desired measure of students treated, meaning the optimal cutoff and results may differ from those in Section 3.1. Let  $\bar{\kappa}^*$  denote the binding solution to the constrained problem (1). This section focuses on averages of  $\Delta(\cdot)$  when the capacity constraint binds, defined as follows.

**Definition 1** (Treatment effects of interest). *When the capacity constraint is binding, define:*

- *Average Effect of Treatment on the Treated ( $\overline{ATT}$ ):  $\int_{\bar{\kappa}^*}^1 \frac{\Delta(x)}{1 - \bar{\kappa}^*} dx$ ,*
- *Average Effect of Treatment on the Untreated ( $\overline{ATUT}$ ):  $\int_0^{\bar{\kappa}^*} \frac{\Delta(x)}{\bar{\kappa}^*} dx$ , and*
- *Average Treatment Effect ( $\overline{ATE}$ ):  $\int_0^1 \Delta(x) dx$ .*

The conditional treatment effects in Definition 1 differ from those in Definition 1 because they use the capacity-constrained-optimal cutoff  $\bar{\kappa}^*$ . However, the capacity-constrained ATE is the same as the unconstrained one, i.e.,  $\overline{ATE} = ATE$ .

Begin by characterizing the constrained-optimal cutoff,  $\bar{\kappa}^*$ . Condition 1 must be adapted, resulting in Condition 1.

**Condition 1** (Necessity). *If  $\bar{\kappa}^*$  solves the administrator's capacity-constrained problem (1), then  $\Delta(\bar{\kappa}^*) > \chi$ .*

*Proof.* Ignoring the measure-zero case(s) where  $\Delta(\bar{\kappa}^*) = \chi$ , this follows from Condition 1(i) and the fact that the capacity constraint is binding, i.e., the administrator would have liked to treat more students.  $\square$



Condition  $\bar{1}$  has the same testable implication as that derived from Condition 1(i)—that the model can be falsified if the RD estimate is negative. Note, however, that Condition 1(ii), that the treatment effect derivative is nondecreasing at the cutoff, need not apply when the constraint binds. However, an alternative testable implication of optimality can be deduced by combining Condition 1(i) with Condition  $\bar{1}$ :  $\Delta(\bar{\kappa}^*) > \Delta(\kappa^*)$ . This could be tested by using data from two years where one thought the model parameters  $\Delta(\cdot)$  and  $\chi$  remained constant, one where the constraint was binding (allowing estimation of  $\Delta(\bar{\kappa}^*)$ ) and another where the administrator’s budget increased, say, due to a large RD estimate stemming from the binding constraint in the first year (allowing estimation of  $\Delta(\kappa^*)$ ). What may look like a lack of “scale-up” for a program may simply reflect that the marginal benefit at the treatment cutoff is smaller if the constraint is no longer binding. Further note that one could use the variation in capacity constraints to trace out  $\Delta(\cdot)$ . Interestingly, these last two results are only implementable if the constraint binds in at least one period—otherwise the cutoff, and resulting LATE, would be the same for both periods. This is perhaps counterintuitive, as we may actually learn more when the constraint binds in at least one year.

**Condition  $\bar{2}$  (Sufficiency).** *The fact the program was implemented implies that the total gain from treating those students was larger than the total costs, i.e.:*

$$\text{Participation: } \int_{\bar{\kappa}^*}^1 \Delta(x)dx > \chi(1 - \bar{\kappa}^*), \quad (\bar{2})$$

where the strict inequality follows from combining Condition  $\bar{1}$  with (2), from Condition 2.

Note that Condition  $\bar{2}$  does not have an analogue to the second part of Condition 2. This is because the administrator would have wanted to treat the inframarginal students (i.e., those between the unconstrained- and constrained-optimal cutoffs); otherwise the constraint would not have been binding.

Next, I examine what can be learned about  $\Delta(\cdot)$  for subsets of students, starting with the treated.

**Corollary  $\bar{3}$ .** *The  $\overline{ATT}$  is positive.*

*Proof.* Because the marginal cost of treatment is positive (Assumption 1(i)),  $(\bar{2})$  implies that

$$\int_{\bar{\kappa}^*}^1 \Delta(x)dx > \chi(1 - \bar{\kappa}^*) > 0.$$

Divide through by  $(1 - \bar{\kappa}^*)$  to obtain the result:

$$\underbrace{\int_{\bar{\kappa}^*}^1 \frac{\Delta(x)}{(1 - \bar{\kappa}^*)} dx}_{\overline{\text{ATT}}} > \underbrace{\chi}_{\text{avg. cost of treating treated}} > 0. \quad \square$$

Corollary 3 shows that we can bound the  $\overline{\text{ATT}}$  from below by zero. This bound is looser than that for the ATT (i.e., the unconstrained analogue of  $\overline{\text{ATT}}$ ). This is because, though we know that both  $\Delta(\bar{\kappa}^*)$  and the  $\overline{\text{ATT}}$  are nonnegative, we do not have enough information to order them. This is in contrast to Proposition 1, which showed that the unconstrained RD estimate bounded the unconstrained ATT from below.<sup>13</sup> For example, consider  $\Delta(\cdot)$  such that  $\Delta'(x) < 0$ , for  $x \geq \bar{\kappa}^*$ ; in this case the RD estimate would bound the  $\overline{\text{ATT}}$  from above. However, we could use (2) to tighten this bound, if information about  $\chi$ , the marginal cost of treatment, were available.<sup>14</sup>

Next, I focus on what can be learned about  $\Delta(\cdot)$  for subsets of untreated students.

**Proposition 2.** *There exists an informative upper bound for  $\int_a^b \Delta(x)dx$  for  $0 \leq a < b \leq \bar{\kappa}^*$ , if  $\kappa^*$  is known and  $a < \kappa^*$ .*

*Proof.* If  $b \leq \kappa^*$ , then by Proposition 2 the upper bound is  $\chi(\kappa^* - a) - \underline{\Delta}(\kappa^* - b)$ ; if  $\chi$  is unknown then apply Condition 1 to form the upper bound  $\Delta(\bar{\kappa}^*)(\kappa^* - a) - \underline{\Delta}(\kappa^* - b)$ .

If  $b > \kappa^*$ , then we can split the integral into two parts:  $\int_a^{\kappa^*} \Delta(x)dx$ , which, applying the result for  $b \leq \kappa^*$ , has an upper bound of  $\Delta(\bar{\kappa}^*)(\kappa^* - a)$ , and  $\int_{\kappa^*}^b \Delta(x)dx$ , which has an upper bound of  $\overline{\Delta}(b - \kappa^*)$ . Combine to form the upper bound  $\Delta(\bar{\kappa}^*)(\kappa^* - a) + \overline{\Delta}(b - \kappa^*)$ .

Note that if  $a \geq \kappa^*$ , then all students  $x \in [a, b]$  are inframarginal, meaning the upper bound would be  $\overline{\Delta}(b - a)$ , i.e., uninformative.  $\square$

Proposition 2 applies to subsets of untreated students, and can be applied to two subsets of particular interest.

**Corollary 4.** *Upper bounds can be obtained for treatment effects for groups of untreated students:*

- (i) *The ATUT is bounded from above by the LATE at the treatment cutoff.*
- (ii) *There is an informative upper bound for the  $\overline{\text{ATUT}}$ , if  $\kappa^*$  is known.*

<sup>13</sup>Also note that  $\overline{\text{ATT}}$  is bounded below by the ATT. However, as we do not know the ATT—at best, we would know its lower bound if the constraint did not bind—this result is of limited use.

<sup>14</sup>Assuming  $\beta$  was also known.

*Proof.* Part (i) follows by combining Condition  $\bar{1}$  with Corollary 4 and Condition 1(i).

To show part (ii), apply Proposition  $\bar{2}$ , setting  $a = 0$  and  $b = \bar{\kappa}^*$ , and divide by  $\bar{\kappa}^*$ , the measure of untreated students, obtaining  $(\kappa^*/\bar{\kappa}^*)\Delta(\bar{\kappa}^*) + (\bar{\kappa}^* - \kappa^*)/\bar{\kappa}^*)\bar{\Delta}$ .  $\square$

Corollary  $\bar{4}$ (i) is the analogue of Corollary 4, delivering an upper bound to the unconstrained ATUT. Corollary  $\bar{4}$ (ii) shows that knowledge of how severe the capacity constraint is, i.e.,  $(\kappa^*/\bar{\kappa}^*)$ , tightens the upper bound on the  $\overline{\text{ATUT}}$ . The extent to which it can be tightened depends on  $\kappa^*$ , but this bound is still looser than in the unconstrained case, because  $\Delta(\bar{\kappa}^*) > \Delta(\kappa^*)$ . Intuitively, if the desired measure of treated students is not known then we would have to set  $\kappa^* = 0$ , to maximize the upper bound, returning the uninformative upper bound of  $\bar{\Delta}$ . It is reasonable to think we may be able to determine how severely the constraint binds in real-world applications, even when we don't know  $\kappa^*$  or the marginal cost of treatment. Specifically, one could compare an administrator's budget request with the actual amount expended and exploit the fact that the fraction,  $(\kappa^*/\bar{\kappa}^*)$ , which is required to compute the bound, is a known function of the ratio of requested and realized budgets, because the unknown marginal cost of treatment cancels when computing the requested/realized budget ratio.<sup>15</sup>

Finally, we can combine the previous results to bound the ATE.

**Corollary  $\bar{5}$ .** *The ATE has an informative lower bound when the capacity constraint binds, if  $\underline{\Delta} < 0$ . If  $\kappa^*$  is known, then the ATE also has an informative upper bound.*

*Proof.* By Corollary  $\bar{3}$ , the lower bound for students  $x \geq \bar{\kappa}^*$  is 0. This increases the ATE lower bound from  $\underline{\Delta}$  to  $\underline{\Delta}\bar{\kappa}^*$ .

If  $\kappa^*$  is known, then by Corollary  $\bar{4}$ (ii) the upper bound on the total gain for the untreated is  $\kappa^*\Delta(\bar{\kappa}^*) + (\bar{\kappa}^* - \kappa^*)\bar{\Delta}$ , reducing the ATE upper bound from  $\bar{\Delta}$  to  $\kappa^*\Delta(\bar{\kappa}^*) + (1 - \kappa^*)\bar{\Delta}$ .  $\square$

There is no informative upper bound (i.e., less than  $\bar{\Delta}$ ) for the ATE if  $\kappa^*$  is unknown. Note that, if  $\chi$  and  $\kappa^*$  are known, then the ATE bounds would be the same as in Section 3.1.

In summary, bounds on  $\Delta(\cdot)$  are looser when the capacity constraint binds. The lower bound on the  $\overline{\text{ATT}}$  is lower—it is zero instead of the RD-estimate LATE. The RD estimate bounds the ATUT, but not the  $\overline{\text{ATUT}}$ , from above. Knowledge of how severe the capacity constraint is—which in principle could be obtained by comparing budget requests and realized allocations—significantly tightens the upper bound on the  $\overline{\text{ATUT}}$  and, consequently,

---

<sup>15</sup>Let the requested budget be  $B^* = (1 - \kappa^*)\chi$  and the realized budget be  $\bar{B}^* = (1 - \bar{\kappa}^*)\chi$ . We can compute  $B^*/\bar{B}^* = ((1 - \kappa^*)\chi)/((1 - \bar{\kappa}^*)\chi) = (1 - \kappa^*)/(1 - \bar{\kappa}^*)$ , which permits the solution for  $\kappa^*$ , even when marginal cost  $\chi$  is unknown.

the ATE. Additionally, though the derivative-based test of cutoff optimality does not apply when the constraint binds, the nonnegative LATE testable implication still does apply and there is a new testable implication of optimality (which could be applied if one had access to data for binding and non-binding periods), and a new policy-relevant result that may help explain why it is difficult to “scale up” successful programs to larger populations (see, e.g., Elmore (1996) and Sternberg et al. (2006)).

## 4 Applications

This section shows how this paper’s theoretical results can be used to extend findings from regression-discontinuity designs. There are three applications, all of which happen to be in the economics of education and examine contexts where it seems reasonable to expect that program administrators had information about the gains and costs of treatment.

Recall that the administrator’s objective depends on the total gain from treatment in the baseline model presented in Section 2. This specification is a good fit for many applications of interest, in particular, the applications studied here, which all study either wages directly, or measures of human capital, such as GPA or standardized test scores. This is because, given a rental rate for human capital, maximizing human capital, maximizing wages, and maximizing output may be viewed as equivalent, meaning the objective considered here corresponds to the efficient allocation.

To most fully illustrate the theoretical results, I first examine a context where it seems likely that the administrator’s capacity constraint binds; this is followed by two contexts where the constraint is likely not binding. I check the model implication that the treatment effect at the cutoff is nonnegative for all three applications; this test follows from Condition 1(i) when the capacity constraint is not binding and Condition  $\bar{1}$  when the constraint does bind. For the latter two applications, I also conduct the test of model assumptions implied by Condition 1(ii), i.e., that the treatment effect is increasing at the cutoff.<sup>16</sup> Reassuringly, we cannot reject that the cutoff was chosen optimally by an informed administrator in each of the (five) falsification tests.

Two of the studies employ fuzzy designs, so I first show how the earlier results pertaining to sharp designs generalize here.<sup>17</sup> Some new notation is necessary. Let  $\omega(x)$  denote the administrator’s intended treatment group for student with index  $x$ .<sup>18</sup> For example, if students

---

<sup>16</sup>Recall that this condition does not hold for the binding-constraint case.

<sup>17</sup>These results are derived for when the capacity constraint is not binding; analogous results obtain when the constraint binds.

<sup>18</sup>Recall that students are distributed uniformly over  $[0, 1]$ .

with indices  $x$  above cutoff  $\kappa$  are targeted for a program, then  $\omega(x) = 1$  for  $x \geq \kappa$  and  $\omega(x) = 0$  for  $x < \kappa$ . The probability of being treated ( $\tau = 1$ ) depends on  $\omega$  according to  $\rho_\omega = \Pr\{\tau = 1|\omega\}$ .<sup>19</sup> In a fuzzy design,  $0 \leq \rho_0 < \rho_1 < 1$ , i.e. not all students targeted for treatment are treated. It may also be the case that some students not targeted for treatment (in the above example, a student with index  $x < \kappa$ ), receive the treatment, in which case  $\rho_0 > 0$ . The fuzzy design requires the probability of treatment to increase discontinuously at the cutoff  $\kappa$  (Hahn et al. (2001)). This notation can also capture a sharp design when  $0 = \rho_0 < \rho_1 = 1$ . In a fuzzy design, the administrator chooses the treatment cutoff  $\kappa$  to maximize her expected objective:<sup>20</sup>

$$\max_{\tilde{\kappa}} \left( \rho_0 \int_0^{\tilde{\kappa}} (\Delta(x) - \chi) dx + \rho_1 \int_{\tilde{\kappa}}^1 (\Delta(x) - \chi) dx \right). \quad (8)$$

The optimal cutoff  $\kappa^*$  is characterized by  $\rho_1 \Delta(\kappa^*) = \rho_1 \chi$ , implying that  $\Delta(\kappa^*) = \chi$ . Note that this condition is identical to Condition 1(i) for the sharp design. Moreover, multiplying through by  $\rho_\omega$  shows that the fuzzy design returns exactly the same bounds for the ATT and ATUT as does the sharp design when  $\rho_\omega$  are constant within treatment status.<sup>21</sup>

The tests of model assumptions can be described using a sharp design without any loss of generality.<sup>22</sup> Suppose students with index  $x \geq \kappa^*$  were treated. In this context, the assumptions that the administrator knows  $\Delta(\cdot)$  and is acting optimally would be rejected if we found that either  $\Delta'(\kappa^*) < 0$ , because the administrator would gain by increasing the cutoff and avoid treating inframarginal students with gains lower than that for students at the cutoff, or  $\Delta(\kappa^*) < 0$ , because the marginal cost of treatment is positive, contradicting optimality of treating students at  $\kappa^*$ . As commonly assumed in regression-discontinuity designs, assume the expected outcome for a student with index  $x$ ,  $Y_{\tau(x)}(x)$ , depends on treatment status  $\tau(x)$  and the running variable  $(x - \kappa^*)$  according to the following statistical relationship:<sup>23</sup>

$$Y_{\tau(x)}(x) = \alpha_0 + \alpha_1(x - \kappa^*) + \alpha_2\tau(x) + \alpha_3\tau(x)(x - \kappa^*),$$

---

<sup>19</sup>That is, treatment probability only depends on  $x$  through  $\omega(x)$ .

<sup>20</sup>As was the case in the theoretical model,  $\beta$  has been set to 1. Recall that the marginal costs of treatment are assumed to be constant in the model. Though estimates of cost functions are not widely available, I was able to find evidence supporting this assumption for the applications studying university outcomes. This evidence is presented on page 21.

<sup>21</sup>Appendix A shows results when treatment probabilities  $\rho_\omega$  depend on  $x$ .

<sup>22</sup>Dong and Lewbel (2015) show that a similar result holds for fuzzy designs. In particular, the derivative sign test implied by Condition 1(ii) is the same.

<sup>23</sup>This relationship only needs to be approximately true in a neighborhood around  $\kappa$  for the argument made here. However, if this were instead thought to be a reasonable approximation to the *global* behavior of  $Y_{\tau(x)}(x)$ , and, therefore,  $\Delta(x)$ , then Appendix B shows that inclusion of an additive independent error  $\epsilon$  does not affect the choice of  $\kappa^*$  or theoretical results. Some studies also use polynomial functions of the running variable, which affects how to estimate  $\Delta'(\kappa^*)$  but, does not affect the test results for these applications.

and the observed outcome for student  $i$ ,  $\tilde{Y}_i$ , measures  $Y_{\tau(x_i)}(x_i)$  with an independent measurement error  $\epsilon_i$  according to:

$$\tilde{Y}_i = Y_{\tau(x)}(x) + \epsilon_i = \alpha_0 + \alpha_1(x - \kappa^*) + \alpha_2\tau(x) + \alpha_3\tau(x)(x - \kappa^*) + \epsilon_i. \quad (9)$$

The estimate of the LATE at the treatment cutoff is  $\hat{\Delta}(\kappa^*) = \hat{\alpha}_2$ . Dong and Lewbel (2015) show that the estimate of the treatment effect derivative at the cutoff here would be  $\hat{\Delta}'(\kappa^*) = \hat{\alpha}_3$ . Therefore, the model has a testable implication, i.e., is falsifiable, because using  $\hat{\alpha}_3$  to test the null hypothesis  $H_0 : \alpha_3 \geq 0$ , versus the alternative hypothesis  $H_1 : \alpha_3 < 0$ , amounts to a test of the model assumptions. Evidence strong enough to reject the null that  $\alpha_3 \geq 0$  would cast doubt on the validity of Assumption 1. Moreover, evidence strong enough to reject the null hypothesis that  $\alpha_2 \geq 0$  would also cast doubt on the validity of the model assumptions.

#### 4.1 Hoekstra (2009): “The Effect of Attending the Flagship State University on Earnings: A Discontinuity-based Approach”

This section applies this paper’s results to Hoekstra (2009), who studies the effect of attending a flagship public university on subsequent mean wages for a sample of white males between the ages of 28 and 33. The objective considered in (1), where the administrator seeks to maximize the amount gained (i.e., increase in wages) net cost of treatment (i.e., having a student attend a high-quality public university) may be a good fit for this environment because a public university likely has the education of the state’s denizens at heart, especially if these students become more productive and stay in the state upon graduation ( 70% of applicants to the flagship eventually earn wages in the same state).<sup>24</sup>

Hoekstra uses a fuzzy design in which treatment was targeted to students at or above a covariate-adjusted SAT score, i.e.  $\omega(x) = 1 \Leftrightarrow x \geq \kappa^*$ . The intended treated students ( $\omega(x) = 1$ ) were offered admission to the flagship and, for the most part, attended it. The intended untreated students ( $\omega(x) = 0$ ) represent a combination of students who do not pursue any higher education, students who attend some other institution of higher education, and a small number of students who attend the flagship university, though the author provides evidence that most likely attend another institution. Let  $Y_0(x)$  and  $Y_1(x)$  denote the expected

---

<sup>24</sup>Epple et al. (2006) find that a model where universities optimize student achievement can explain the data. Because students’ achievement measures their human capital, which itself augments wages, one could therefore view universities as wanting to maximize future wages. Similarly, maximizing students’ completing college or finding (or keeping) jobs would naturally be captured by having the administrator maximize wages, as these schooling and labor market outcomes are all positively related to wages. That is, though it admittedly abstracts from alternative dimensions universities may care about, modeling public universities as maximizing student wages may reasonably approximate their objectives.

wage for student with index  $x$  in treatment group  $\tau$ . Variation in  $\Delta(x) \equiv Y_1(x) - Y_0(x)$  then represents the oft-studied heterogeneity in the returns to education, for the case of selective public universities. The inferential problem with extrapolating from the RD estimate is that the gain may vary between students.

Research by Epple et al. (2006) shows that universities admit students until their capacity constraints bind, meaning it is likely that Hoekstra (2009) was implemented under a binding capacity constraint. The marginal cost of treatment is assumed to be constant and, as in the model, is denoted by  $\chi$ . This assumption has supported by Izadi et al. (2002), who estimate a CES cost function for universities. Based on parameter estimates provided in that paper, one cannot reject that university cost functions are linear in the number of students served.<sup>25</sup> This assumption is also supported by other work, such as Epple et al. (2006), who estimate a model of the higher education market for private colleges and do not find evidence that the cost of serving students is nonlinear.<sup>26</sup>

Recall that the fuzzy design returns exactly the same bounds as does the sharp design, when  $\rho_\omega$  are constant within treatment status. Therefore, the lower bound for the average effect of treatment on the treated ( $\overline{ATT}$ ) and upper bound for the average effect of treatment on the untreated ( $\overline{ATUT}$ ) developed earlier also apply here. The main result reported in Hoekstra (2009) is that attending the flagship university increases log wages by 20% for students at the treatment cutoff.<sup>27</sup> This positive estimate means this context is consistent with (constrained-) optimality of the cutoff, implied by Condition  $\bar{1}$ ; that is, we cannot falsify the model assumptions. Note that when we use the bounds for the  $\overline{ATT}$  when the capacity constraint binds, we can only surmise that the effect of treatment on the treated is positive, i.e.,  $\overline{ATT} > 0$ . As noted by Hoekstra, this estimate seems fairly high; for example, Ashenfelter and Rouse (1998) estimate that an additional year of schooling increases wages by 9%, while Behrman et al. (1996) find that an additional year of schooling increases wages by 6-8% and that there is a 20% increase in wages from attending a large public college versus only graduating from high school. However, the relatively large estimated effect in Hoekstra (2009) are quite intuitive when viewed through the lens of Condition  $\bar{1}$  (that the RD estimate exceeds the marginal cost of treatment when the capacity constraint binds).

---

<sup>25</sup>Specifically, I test whether the second derivative of the cost of serving arts and science students is zero in the number of that type of student, and find that even an 80% confidence interval for the second derivative contains zero for both student types. Izadi et al. (2002) use data from the UK.

<sup>26</sup>See Table II on page 907 of Epple et al. (2006).

<sup>27</sup>Conservatively, this corresponds to an average increase of \$8,000 in the annual wage. Based on correspondence with the author, the average baseline annual income for the untreated was between \$40,000 and \$50,000.

## 4.2 Lindo et al. (2010): “Ability, Gender, and Performance Standards: Evidence from Academic Probation”

This section applies this paper’s results to Lindo et al. (2010), which studies how being placed on academic probation affects subsequent outcomes for university students. They exploit a sharp discontinuity design, where students with GPAs below a chosen cutoff are assigned to academic probation, i.e.,  $\tau(x) = 1 \Leftrightarrow x \leq \kappa^*$ , where  $x$  is the student’s GPA last semester. Students on academic probation must keep their GPAs above a certain standard, else they will be placed on academic suspension. The estimation sample comprises students from three campuses of a public university in Canada.

As with Hoekstra (2009), the fact that the university is public means it is reasonable to expect that it would value student achievement. Therefore, I focus on effect of being placed on probation on subsequent GPA, which means that the treatment effect  $\Delta(x)$  is the expected gain in subsequent GPA if student with prior GPA  $x$  were placed on academic probation. Lindo et al. (2010) use a simplified version of Bénabou and Tirole (2000) to motivate why there may be heterogeneity in the affect of probation on student outcomes; the takeaway being that students far above the cutoff naturally perform well in their classes, and, therefore, would gain little from being put on probation. The university faces a cost of placing students on probation, which captures the fact that students are offered additional counseling and support services to help them improve their achievement. Therefore, assigning all students to probation would mean incurring costs for students who have little expected gain. Because only a subset of students are placed on probation and the effects of probation likely depend on student ability, it is useful to think about how we can extrapolate away from the treatment cutoff. The university could have treated more students by sending out more probation letters and hiring the counselor/tutor for more hours, which means it is reasonable to assume the capacity constraint was not binding in this application.

I begin by conducting the falsification test on the treatment effect derivative, implied by Condition 1(ii). The sign of  $\Delta'(\kappa^*)$ , and therefore the rejection region for the falsification test, is reversed here because treatment is offered to students below  $\kappa^*$ , meaning that extending treatment to students above  $\kappa^*$  should not improve the administrator’s objective. Using the information made available by the journal’s replication policy, I ran regression (9), the results of which are presented in Table 1. I find that  $\hat{\alpha}_3 = 0.047$ , with a standard error of 0.094, which means that there is not strong evidence that the treatment effect is increasing at the cutoff (p-value 0.31 that the treatment effect derivative is greater than zero).<sup>28</sup> Moreover, the positive

---

<sup>28</sup>The data are available at <https://www.aea-net.org/articles.php?doi=10.1257/app.2.2.95>; R code for replication can be found at Chi and Dow (2014).



RD estimate (see below) means the model passes the nonnegative LATE falsification test, implied by Condition 1(i). That is, we cannot reject that the model assumptions hold here. Therefore, the first result is that we can rule out constant treatment effects, by Corollary 1.

Table 1: Results from main specification in Lindo et al. (2010)

<i>Regressor:</i>	<i>Dependent variable: GPA next semester (<math>\check{Y}_i</math>)</i>
Intercept ( $\hat{\alpha}_0$ )	0.312 (0.019)
Running variable - cutoff ( $\hat{\alpha}_1$ )	0.699 (0.053)
Treatment indicator ( $\hat{\alpha}_2$ )	0.233 (0.031)
Treatment indicator $\times$ (running variable - cutoff) ( $\hat{\alpha}_3$ )	0.047 (0.094)
Obs.	11,258
R <sup>2</sup>	0.035

*Note:* Standard errors are in parentheses

The main result of Lindo et al. (2010) is that the estimated treatment effect of being placed on academic probation on the next term’s grade performance for the full sample is  $\hat{\Delta}(\kappa^*) = 0.233$  higher GPA points.<sup>29</sup> By extending this finding using the results of this paper, we can bound the ATT and ATUT according to:  $ATT \geq \hat{\Delta}(\kappa^*) = 0.233 > ATUT$ . In other words, placing students below the treatment cutoff on academic probation would, on average, increase their GPA the next term by 0.233 points, while doing so for students above the treatment cutoff would increase their GPA next term by at most 0.233 points, on average. Intuitively, on average, academic probation may be more useful for students at the bottom of the grade distribution, by providing them with an external commitment to increase their performance above some minimal level.<sup>30</sup>

### 4.3 Bui et al. (2014): “Is Gifted Education a Bright Idea? Assessing the Impact of Gifted and Talented Programs on Students”

Bui et al. (2014) study the effects of a gifted and talented program on the academic achieve-

<sup>29</sup>This can be found in Table 5, in column (1) of panel A of Lindo et al. (2010).

<sup>30</sup> This is a conjecture. Proposition 2 shows that we can obtain bounds for subsets of untreated students. However, no such bounds can be obtained for subsets of treated students, such as those with the lowest GPAs.

ment of 7th graders in a large school district. They exploit a fuzzy eligibility cutoff, in which students with prior achievement above the cutoff were more likely to be found eligible (i.e.,  $\omega(x) = 1 \Leftrightarrow x \geq \kappa^*$ , and  $0 < \rho_0 < \rho_1 < 1$ ), to compute how the program affected achievement of marginal students near the cutoff. Students eligible for the gifted and talented programs could additionally apply to special gifted and talented programs at district-wide magnet schools. The gifted and talented program is comprised of peers and curricula; the authors show that teachers are essentially the same between gifted and marginal non-gifted students. Therefore, the school district could have switched more teachers into gifted classes, meaning it is reasonable to assume the capacity constraint does not bind in this application. Two of these magnet schools were oversubscribed so the district assigned applicants by lottery. The authors exploit this to conduct a lottery-based design, comparing the achievement of applicants who had been randomized into these magnet schools with those who had been randomized out (and into a “regular”, or non-magnet, gifted and talented program).

The data contain many standardized test score measures of both incoming and outcome human capital. Following the authors, I set the noisy measure of incoming human capital  $x$  equal to the student’s “gifted and talented identification matrix” score, which converts each student’s prior achievement, teacher ratings, and socio-economic status into one scale that serves as the running variable in their discontinuity design. The authors report treatment effects for outcome scores in many subjects  $s$ ,  $\Delta^s(\kappa^*)$ , where  $s$  can be math, reading, language, science, or social studies. Let  $\hat{\Delta}^s(\kappa^*)$  denote the estimate of subject-specific treatment effect  $\Delta^s(\kappa^*)$ . These are then aggregated into a composite score, i.e.,  $\hat{\Delta}(\kappa^*) = \sum_s \lambda_s \hat{\Delta}^s(\kappa^*)$  measures the composite gain  $\Delta(\kappa^*) = \sum_s \lambda_s \Delta^s(\kappa^*)$ , where the weights  $\lambda_s$  satisfy  $\sum_s \lambda_s = 1$  are chosen to correspond to different composite treatment effects of interest. Note that, in addition to being used to calculate bounds, the weighted composite scores can also be used in the treatment-effect derivative test, because derivatives are linear operators.

The top panel of Table 2 summarizes RD estimates from Bui et al. (2014), Table 4b (the authors’ preferred specification), by subject. Columns (1) and (2) present treatment effect derivatives and standard errors, respectively. Columns (3) and (4) present RD estimates of treatment effects and their standard errors, respectively. The bottom panel presents two different averages across subject-specific estimates.<sup>31</sup> First, based on the notion that the school district desires well-rounded students, the unweighted average across subjects is presented. Looking at the first two columns, we can see that the (unweighted) average treatment effect derivative is -0.002, with a standard error of 0.002, which is not significantly

---

<sup>31</sup>Standard errors for averages are computed assuming the subject-specific measures are independently distributed. This is conservative in the sense that, because it provides a lower bound for the standard error of  $\hat{\Delta}^s(\kappa^*)$ , it increases the likelihood of rejecting the null hypothesis that the model assumptions hold.

Table 2: Summary of subject-specific treatment effect derivatives and RD estimates in Bui et al. (2014)

Subject	Treatment-effect derivative, $\widehat{\Delta}'^s(\kappa^*)$ ( $\widehat{\alpha}_3$ )		Treatment effect, $\widehat{\Delta}^s(\kappa^*)$ ( $\widehat{\alpha}_2$ )	
	Estimate (1)	Std. error (2)	Estimate (3)	Std. error (4)
Math	0.004	0.005	-0.037	0.074
Reading	-0.008	0.004	0.049	0.068
Language	-0.007	0.005	-0.015	0.066
Science	-0.004	0.004	0.003	0.084
Social studies	0.003	0.006	-0.025	0.084
Average (unweighted)	-0.002	0.002	-0.005	0.034
Average (weighted)	0.001	0.003	-0.020	0.046

*Note:* Standard errors for averages were computed assuming independence across subjects.

less than zero with a p-value of 0.13. That is, one cannot reject that the eligibility cutoff for gifted and talented programs is optimal in this school district. One should also note that, notwithstanding general equilibrium effects, this implies that there is no evidence that extending gifted and talented programs to students just below the threshold would increase achievement.<sup>32</sup> Consistent with the subject-specific estimates, the top row of the bottom panel shows that there is no discernible effect on the unweighted average achievement for the marginal student of being assigned from their regular program to a gifted and talented program, with a point estimate of -0.005 that has a standard error of 0.034, resulting in a p-value of 0.88 that the treatment effect is different from zero.

Second, weights were computed by relating, or “anchoring” standardized test scores to adult wages.<sup>33</sup> As one can see in the second row of the bottom panel, here too, one cannot reject optimality of the chosen cutoff (the composite treatment effect derivative is positive), and the composite marginal effect is indistinguishable from zero (with a point estimate of -0.020 with standard error 0.046, resulting in a p-value of 0.66 that the treatment effect is different from zero). Note that both the unweighted and weighted treatment effects are non-negative, meaning the model also passes this test of cutoff optimality (implied by Condition

<sup>32</sup>Dong and Lewbel (2015) refer to this assumption as “local policy invariance” and exploit it to estimate the Marginal Threshold Treatment Effect, or MTTE.

<sup>33</sup>This idea of “anchoring” skills in an outcome of interest has been advocated by Cunha et al. (2010) and Bond and Lang (2013), among others; Cawley et al. (1999) do so using adult wages. Sanders (2015) regresses the completed years of education on ASVAB scores for Science, Vocabulary, Paragraph Comprehension, and Math (see his Table 3). I computed weights corresponding to the partial correlation of that subject with completed years of education, divided by the sum of partial correlations across all subjects (Reading, Language, and Social studies were given the weight of the average of Vocabulary and Paragraph Comprehension); the motivation for this is that expected wages (and output produced) have generally been found to increase in one’s education.

1(i)).

Therefore, under either weighting scheme, we cannot reject that the treatment cutoff was chosen optimally and the lower bound for the ATT and the upper bound for the ATUT are both approximately zero. Additionally, as was the case with Lindo et al. (2010), by Corollary 1 we can rule out constant treatment effects.

The finding that the effect of treating the marginal student is very close to (and not discernible from) zero may seem surprising at first, but is very natural when viewed in the context of an endogenous treatment assignment cutoff. If treatment effects  $\Delta(\cdot)$  are smooth around the cutoff, an assumption upon which RD estimates are typically predicated,<sup>34</sup> then the school district would have little reason to place the cutoff at a point where there was a very large gain in achievement for the marginal student. This would be consistent with a human capital production function in which incoming human capital and school inputs were complements, in which case returns from assignment to the program would then be increasing in  $x$ , the measure of incoming human capital.

Without further information, one cannot surmise more about the average effect of treatment on the treated or untreated, other than the ATT cannot be negative and the ATUT cannot be positive. However, Bui et al. (2014) also features a complementary analysis that exploits the fact that students above the treatment cutoff were not only eligible for gifted and talented programs, but could also apply to attend oversubscribed magnet gifted and talented programs. The authors find no discernible difference in achievement across most subjects between students randomized into magnet and non-magnet gifted and talented programs.<sup>35</sup> If one thought that such magnet programs produced higher achievement than regular public schools, the fact that achievement was the same for eligible students in both magnet and non-magnet gifted and talented programs would mean that non-magnet gifted and talented programs also produced gains for eligible students. For example, Burke and Sass (2013) and Hoxby and Weingarth (2005) find that the effect of increases in peer quality is increasing in one's own ability. This would imply that the ATT, i.e., the average effect of attending gifted and talented programs for students above the assignment cutoff, is positive, which is consistent with the bounds obtained above.

---

<sup>34</sup>See Hahn et al. (2001) and Dong and Lewbel (2015) for discussions of this point.

<sup>35</sup>The one exception is that students randomized to attend magnet schools had higher outcomes on science tests.

## 5 Discussion

This paper represents a first step towards showing how one can use plausibly available information and a simple economic model to generalize findings from RD designs.<sup>36</sup> Perhaps the most intuitive findings relate to the case where the capacity constraint does not bind: i) if treating students is costly and the treatment cutoff has been chosen optimally, the ATT must be positive and treatment effects cannot be constant; ii) RD-based estimates provide a lower bound, or understate, the ATT; and iii) RD-based estimates provide an upper bound for the ATUT. Notably for applying these results, the model generates testable implications: if the treatment effect at the cutoff is negative or treatment effects are decreasing in the direction of treatment at the cutoff, then we can reject that the cutoff was chosen optimally by an administrator informed about the distribution of treatment effects. If the capacity constraint does bind, then the treatment-effect sign test still allows one to falsify the model, and treatment effects are generally looser. The treatment-effect-derivative test no longer applies, but there emerges a new testable implication of cutoff optimality, as well as an intuitive explanation for why program “scale-up” can be difficult in real-life applications.

The theoretical results were then demonstrated using three applications. We cannot reject that the cutoff was chosen optimally by an informed administrator in every one of the falsification tests conducted, which may increase confidence that the contexts studied here constitute reasonable applications of the theoretical results. The applications all study topics in the economics of education, which means it is useful to think about them using a standard model of human capital formation. The bounds obtained by applying this paper’s results to Hoekstra (2009) and Bui et al. (2014) are consistent with a model of human capital production in which incoming student human capital, which may have been measured by student SAT scores or prior academic achievement, and additional inputs, such as access to a flagship university or a gifted and talented program, are complements in the production of human capital. Not only would this imply then that treatment effects would be heterogeneous—which the model shows must be the case when the marginal cost of treatment is constant—but also that the ATT would be higher than the RD estimate, which itself would be higher than the ATUT.

Put another way, suppose the program could be assigned either to students above or below the cutoff. If student human capital and program inputs were complements, then the average gain would be higher when students above the cutoff received additional inputs than it would be if instead students below the cutoff were given those inputs. Therefore, the fact that students *above* the cutoff were treated, as opposed to those *below* the cutoff, is

---

<sup>36</sup>This paper also relates to work non-parametrically estimating the distribution of treatment effects.

consistent with student human capital and program inputs being complements. In contrast, Lindo et al. (2010) studied a treatment, academic probation, where students below the cutoff received the treatment. In terms of program inputs, being put on academic probation comprises mostly access to counseling and support required to have students meet a minimal performance standard, which likely serve more as substitutes for low ability students who may not meet performance standards on their own, as opposed to something like access to more rigorous curricula. The former is clearly much more useful to lower ability students than higher ability ones.

The findings in this paper have several implications for the use of RD results in policy. Perhaps most novel is that we may incorrectly surmise that some programs are ineffective and eliminate them, even though in reality they are quite effective for the treated population. Strikingly, *such a mistake would be more likely for a program with a very low marginal cost, holding constant the ATT*. This point is illustrated in Figure 3, which plots treatment effects associated with hypothetical programs at two sites,  $A$  and  $B$ , with respective distributions of treatment effects  $\Delta_A(\cdot)$  (solid black line) and  $\Delta_B(\cdot)$  (dashed black line).<sup>37</sup> The programs have different marginal costs of treatment, where  $\chi_B > \chi_A$ , and happen to have the same cutoff  $\kappa^*$  and the same ATT. The difference in marginal costs means that optimization by respective site administrators implies that  $\Delta_B(\kappa^*) > \Delta_A(\kappa^*)$ . If only based on these RD estimates, a policymaker would likely fund  $B$  over  $A$  because it has a higher LATE, even though  $A$  provides the same gain on the treated, *at a lower cost*. Additional policy-relevant results obtain if we can relate the policymaker’s objective with that of the administrator. First, if a policymaker knew that their valuation of treatment gain in terms of treatment cost (i.e.,  $\beta$ ) was at least as high as the administrator’s then he should definitely treat those units treated by the administrator. Second, the upper bound on the ATUT, which increases in distance from the cutoff, can help rule out whether it would be worthwhile to extend treatment to subsets of units below the cutoff.

Though, in this setting, estimates of the LATE at the treatment cutoff must be positive if treating students is costly, we cannot compare them with the ATE, in the manner of LaLonde (1986), Dehejia and Wahba (1999), or Smith and Todd (2005), without further information. There is some work comparing findings from RD and experimental designs (Buddelmeyer and Skoufias (2004), Black et al. (2007), Cook and Wong (2008), Gleason et al. (2012), Barrera-Osorio et al. (2014)), but unfortunately, none consider the case of a program where the treatment cutoff seems to have been chosen by an administrator with institutional knowledge of the environment (e.g., in Barrera-Osorio et al. (2014) the evaluators were external, and choose a poverty index as the threshold for treatment). However, the results here do suggest

---

<sup>37</sup>Note that capacity constraints do not bind in this example.

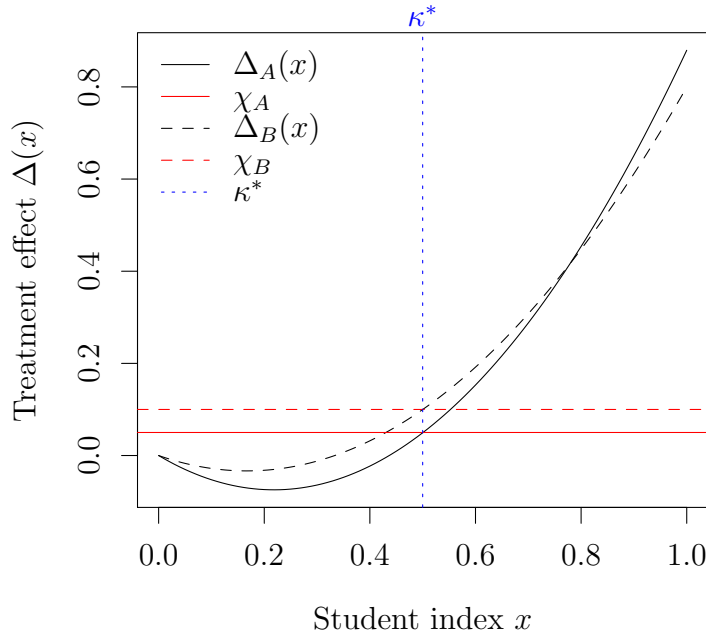


Figure 3: Example of when policy based only on the RD estimate could cause termination of the wrong program

that RD estimates may be higher when cutoffs are chosen by external evaluators without institutional knowledge, hence less information about treatment effects. Related to this point, an unconstrained optimizing administrator would not choose to place the cutoff where they know the gain from treatment is quite large. Because RD estimates may understate the ATT, there may be RD studies of useful programs that are simply not published because they lack statistically significant findings.

One practical variation of the environment considered here would introduce a more substantial form of uncertainty, for example, featuring learning about the distribution of treatment effects, into the administrator's problem. Such uncertainty would pervade to the bounds obtained here, perhaps motivating a Bayesian approach. A more formal approach could also combine bounds for a particular treatment that had been implemented across multiple sites, to build up a picture of the population-level (as opposed to site-specific) distribution of treatment effects. Another variation would investigate what could be learned if the administrator only knew certain moments of the distribution of treatment effects, say, the ATE, along with the RD-based estimate. Such variations could be worthwhile ways to build on the basic point made in this paper: revealed preferences can provide quite a bit of useful information about treatment effects away from the cutoff in regression-discontinuity

designs.

## References

- Angrist, J. D. and M. Rokkanen, “Wanna Get Away? RD Identification Away from the Cutoff,” 2013.
- Anwar, S. and H. Fang, “An Alternative Test of Racial Prejudice in Motor Vehicle Searches: Theory and Evidence,” *American Economic Review*, 96(1):127–151, 2006.
- Ashenfelter, O. C. and C. E. Rouse, “Income, Schooling, and Ability: Evidence from a New Sample of Identical Twins,” *Quarterly Journal of Economics*, 113(1):253–284, 1998.
- Barrera-Orsorio, F., D. Filmer and J. McIntyre, “Randomized Controlled Trials and Regression-Discontinuity Estimations: An Empirical Comparison,” 2014.
- Behrman, J. R., M. R. Rosenzweig and P. Taubman, “College Choice and Wages: Estimates Using Data on Female Twins,” *The Review of Economics and Statistics*, pp. 672–685, 1996.
- Bénabou, R. and J. Tirole, “Self-Confidence and Social Interactions,” *NBER Working Paper*, 2000.
- Björklund, A. and R. Moffitt, “The Estimation of Wage Gains and Welfare Gains in Self-Selection Models,” *The Review of Economics and Statistics*, pp. 42–49, 1987.
- Black, D. A., J. Galdo and J. A. Smith, “Evaluating the Worker Profiling and Reemployment Services System Using a Regression Discontinuity Approach,” *The American Economic Review*, pp. 104–107, 2007.
- Bond, T. N. and K. Lang, “The Evolution of the Black-White Test Score Gap in Grades K–3: The Fragility of Results,” *Review of Economics and Statistics*, 95(5):1468–1479, 2013.
- Brock, W., J. Cooley, S. Durlauf and S. Navarro, “On the Observational Implications of Taste-Based Discrimination in Racial Profiling,” *Journal of Econometrics*, 2011.
- Buddelmeyer, H. and E. Skoufias, *An Evaluation of the Performance of Regression Discontinuity Design On PROGRESA*, vol. 827, World Bank Publications, 2004.
- Bui, S. A., S. G. Craig and S. A. Imberman, “Is Gifted Education a Bright Idea? Assessing the Impact of Gifted and Talented Programs On Students,” *American Economic Journal: Economic Policy*, 6(3):30–62, 2014.
- Burke, M. A. and T. R. Sass, “Classroom Peer Effects and Student Achievement,” *Journal of Labor Economics*, 31(1):51–82, 2013.
- Cawley, J., J. Heckman and E. Vytlačil, “On Policies to Reward the Value Added by Educators,” *Review of Economics and Statistics*, 81(4):720–727, 1999.
- Chi, O. and A. Dow, “Replication Data For: Improving Balance in Regression Discontinuity Through Matching: Estimating the Effect of Academic Probation After the First Year of College,” Tech. rep., Harvard Dataverse, 2014.
- Cook, T. D. and V. C. Wong, “Empirical Tests of the Validity of the Regression Discontinuity Design,” *Annales d’Economie et de Statistique*, pp. 127–150, 2008.
- Cunha, F., J. Heckman and S. Schennach, “Estimating the Technology of Cognitive and Noncognitive Skill Formation,” *Econometrica*, 78(3):883–931, 2010.



- Dehejia, R. H. and S. Wahba, “Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs,” *Journal of the American statistical Association*, 94(448):1053–1062, 1999.
- DiNardo, J. and D. S. Lee, “Program Evaluation and Research Designs,” *Handbook of Labor Economics*, 4:463–536, 2011.
- Dong, Y. and A. Lewbel, “Identifying the Effect of Changing the Policy Threshold in Regression Discontinuity Models,” *Review of Economics and Statistics*, (0), 2015.
- Elmore, R., “Getting to Scale with Good Educational Practice,” *Harvard Educational Review*, 66(1):1–27, 1996.
- Epple, D., R. Romano and H. Sieg, “Admission, Tuition, and Financial Aid Policies in the Market for Higher Education,” *Econometrica*, 74(4):885–928, 2006.
- Ferrall, C. and B. Shearer, “Incentives and Transactions Costs Within the Firm: Estimating an Agency Model Using Payroll Records,” *The Review of Economic Studies*, 66(2):309–338, 1999.
- Gleason, P. M., A. M. Resch and J. A. Berk, “Replicating Experimental Impact Estimates Using a Regression-Discontinuity Approach,” Tech. rep., Mathematica Policy Research, 2012.
- Hahn, J., P. Todd and W. Van der Klaauw, “Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design,” *Econometrica*, 69(1):201–209, 2001.
- Heckman, J., J. Smith and N. Clements, “Making the Most Out of Programme Evaluations and Social Experiments: Accounting for Heterogeneity in Programme Impacts,” *The Review of Economic Studies*, 64(4):487–535, 1997.
- Heckman, J. J., R. J. LaLonde and J. A. Smith, “The Economics and Econometrics of Active Labor Market Programs,” *Handbook of Labor Economics*, 3:1865–2097, 1999.
- Heckman, J. J. and J. A. Smith, “Evaluating the Welfare State,” in “Econometrics and Economic Theory in the 20th Century: The Ragnar Frisch Centennial Symposium,” 31, p. 241, Cambridge University Press, 1998.
- Heckman, J. J. and S. Urzua, “Comparing IV with Structural Models: What Simple IV Can and Cannot Identify,” *Journal of Econometrics*, 156(1):27–37, 2010.
- Heckman, J. J. and E. J. Vytlačil, “Econometric Evaluation of Social Programs, Part II: Using the Marginal Treatment Effect to Organize Alternative Econometric Estimators to Evaluate Social Programs, and to Forecast Their Effects in New Environments,” *Handbook of Econometrics*, 6:4875–5143, 2007.
- Heinrich, C., J. J. Heckman and J. Smith, “The Performance of Performance Standards,” *Journal of Human Resources*, 37(4):778–811, 2002.
- Hoekstra, M., “The Effect of Attending the Flagship State University On Earnings: A Discontinuity-Based Approach,” *The Review of Economics and Statistics*, 91(4):717–724, 2009.
- Hoxby, C. M. and G. Weingarth, “Taking Race Out of the Equation: School Reassignment and the Structure of Peer Effects,” *NBER Working Paper*, 2005.
- Imbens, G. W., “Better LATE Than Nothing: Some Comments On Deaton (2009) and Heckman and Urzua (2009),” *Journal of Economic Literature*, pp. 399–423, 2010.

- Imbens, G. W. and T. Lemieux, "Regression Discontinuity Designs: A Guide to Practice," *Journal of Econometrics*, 142(2):615–635, 2008.
- Izadi, H., G. Johnes, R. Oskrochi and R. Crouchley, "Stochastic Frontier Estimation of a CES Cost Function: The Case of Higher Education in Britain," *Economics of Education Review*, 21(1):63–71, 2002.
- Knowles, J., N. Persico and P. Todd, "Racial Bias in Motor Vehicle Searches: Theory and Evidence," *Journal of Political Economy*, 109(1), 2001.
- LaLonde, R. J., "Evaluating the Econometric Evaluations of Training Programs with Experimental Data," *The American economic review*, pp. 604–620, 1986.
- Lee, D. S. and D. Card, "Regression Discontinuity Inference with Specification Error," *Journal of Econometrics*, 142(2):655–674, 2008.
- Lindo, J. M., N. J. Sanders and P. Oreopoulos, "Ability, Gender, and Performance Standards: Evidence From Academic Probation," *American Economic Journal: Applied Economics*, 2(2):95–117, 2010.
- Manski, C. F., *Partial Identification of Probability Distributions*, Springer Science & Business Media, 2003.
- , "Statistical Treatment Rules for Heterogeneous Populations," *Econometrica*, 72(4):1221–1246, 2004.
- , "Choosing Treatment Policies Under Ambiguity," *Annu. Rev. Econ.*, 3(1):25–49, 2011.
- Manski, C. F. and J. V. Pepper, "Monotone Instrumental Variables: With an Application to the Returns to Schooling," *Econometrica*, 68(4):997–1010, 2000.
- McCrary, J., "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test," *Journal of Econometrics*, 142(2):698–714, 2008.
- Roy, A., "Some Thoughts On the Distribution of Earnings," *Oxford Economic Papers*, 3(2):135, 1951.
- Rubin, D. B., "Comment," *Journal of the American Statistical Association*, 75(371):591–593, 1980.
- Sanders, C., "Reading Skills and Earnings: Why Do Doing Words Good Hurt You're Wages," 2015.
- Smith, J. A. and P. E. Todd, "Does Matching Overcome LaLonde's Critique of Nonexperimental Estimators?" *Journal of Econometrics*, 125(1):305–353, 2005.
- Sternberg, R. J., D. Birney, L. Jarvin, A. Kirlik, S. Stemler and E. L. Grigorenko, "Scaling Up Educational Interventions," *Translating educational theory and research into practice: Developments in content domains, large scale reform, and intellectual capacity*, 2006.
- Van der Klaauw, W., "Regression–Discontinuity Analysis: A Survey of Recent Developments in Economics," *Labour*, 22(2):219–245, 2008.

## Appendix

### A Allowing the probability of enrollment to vary by $x$

Consider (8), where  $\rho_\omega$  are no longer assumed to be uniform among treated students, but instead depends on  $x$ . The administrator's problem becomes:

$$\max_{\tilde{\kappa}} \left( \int_0^{\tilde{\kappa}} \rho_0(x)(\Delta(x) - \chi)dx \right) + \left( \int_{\tilde{\kappa}}^1 \rho_1(x)(\Delta(x) - \chi)dx \right).$$

Note that a necessary condition for  $\kappa^*$  being optimal is still  $\rho_1(\kappa^*)\Delta(\kappa^*) = \rho_1(\kappa^*)\chi \Rightarrow \Delta(\kappa^*) = \chi$ , which is identical to Condition 1(i) for the sharp design. Optimality of  $\kappa^*$  further implies:

$$\begin{aligned} \int_{\tilde{\kappa}}^1 \rho_1(x)\Delta(x)dx &\geq \int_{\tilde{\kappa}}^1 \rho_1(x)\chi dx = \bar{\rho}_1\chi(1 - \kappa^*) \Leftrightarrow \\ \underbrace{\frac{\int_{\tilde{\kappa}}^1 \rho_1(x)\Delta(x)dx}{1 - \kappa^*}}_{\text{ITT}} &\geq \underbrace{\bar{\rho}_1\chi}_{\text{expected cost of treating the treated}} = \bar{\rho}_1\hat{\Delta}(\kappa^*), \end{aligned} \quad (10)$$

where the second equality follows because  $\Delta(\kappa^*) = \chi$ . Equation (10) shows that if the average attendance probability among the treated ( $\bar{\rho}_1$ ) were known then the RD estimate of the treatment effect can again be used to provide a lower bound for the mean effect of intending-to-treat among the treated (ITT).

### B Uncertainty about the treatment effect

Suppose the administrator is uncertain about the treatment effect for particular students, but has observed  $\check{\Delta}(x)$ , an unbiased signal of  $\Delta(x)$ . Let  $\check{\Delta}(x) = \Delta(x) + \epsilon_i$ , where  $\epsilon$  is distributed independently from  $x$ , denote the administrator's noisy signal of the treatment effect for student  $i$  who has index  $x$ . Because the administrator has unbiased beliefs about  $\Delta(x)$  at every point  $x$ , it must be the case that  $E[\epsilon_i] = 0$ . The administrator chooses a

cutoff to maximize her expected objective:

$$\begin{aligned}
\max_{\tilde{\kappa}} \mathbb{E} \left[ \beta \left( \int_{\tilde{\kappa}}^1 \check{\Delta}(x) dx \right) - c(1 - \tilde{\kappa}) \right] &\Leftrightarrow \max_{\tilde{\kappa}} \beta \mathbb{E} \left[ \left( \int_{\tilde{\kappa}}^1 (\Delta(x) + \epsilon) dx \right) \right] - c(1 - \tilde{\kappa}) \\
&\Leftrightarrow \max_{\tilde{\kappa}} \beta \left( \int_{\tilde{\kappa}}^1 \Delta(x) dx \right) + \beta \mathbb{E}[\epsilon] - c(1 - \tilde{\kappa}) \\
&\Leftrightarrow \max_{\tilde{\kappa}} \beta \left( \int_{\tilde{\kappa}}^1 \Delta(x) dx \right) - c(1 - \tilde{\kappa}). \tag{11}
\end{aligned}$$

The first equivalence follows from the fact that the measure of students treated ( $c(1 - \tilde{\kappa})$ ) is known because it is chosen by the administrator. The second follows from the independence assumption and the third from unbiasedness. The last expression is the administrator's original problem, (1). Therefore, the analysis for this case is identical. Intuitively, uncertainty does not affect the administrator's problem because it is linear in the amount gained.

## C Weighted objective

The administrator's original problem (1) was utilitarian, i.e., it weighed gains for all students equally. The most natural alternative to the unweighted objective would be a redistributive policy, which assigned people with lower running variable indices larger weights. For example, if  $x$  measured incoming human capital, then putting more weight on gains for students with lower indices allows the administrator to place additional value on students' becoming proficient. In this case, we can adapt equation (1) to allow the administrator to weigh gains for students depending on their index  $x$  by using weights  $\phi(x)$ , where  $\phi' \leq 0$ :

$$\max_{\tilde{\kappa}} \left( \int_{\tilde{\kappa}}^1 \phi(x) \Delta(x) dx \right) - \chi(1 - \tilde{\kappa}), \tag{\hat{1}}$$

and proceed with the analysis.

**Condition  $\hat{1}$  (Necessity).** *For problem  $(\hat{1})$ , the following necessary conditions must hold for  $\kappa^*$ :*

$$(i) \text{ MB=MC: } \phi(\kappa^*) \Delta(\kappa^*) = \chi$$

$$(ii) \text{ Increasing MB: } \Delta'(\kappa^*) \geq 0.$$

*Proof.* Differentiate the administrator's problem ( $\hat{1}$ ) with respect to  $\hat{\kappa}$  to obtain (i). Note that if the derivative is negative at a candidate solution satisfying (i), the administrator would gain by not treating students just above  $\kappa^*$ , thereby obtaining (ii). The inequality is strict if  $\phi' < 0$ .  $\square$

**Condition  $\hat{2}$**  (Sufficiency). *The fact the program was implemented implies that the total gain from treating those students was at least as large as the total costs, i.e.:*

$$\text{Participation: } \int_{\kappa^*}^1 \phi(x)\Delta(x)dx \geq \chi(1 - \kappa^*). \quad (\hat{2})$$

*The fact the program was not extended to  $\hat{\kappa} < \kappa^*$  implies that treating these students would be sub-optimal, i.e.:*

$$\int_{\hat{\kappa}}^{\kappa^*} \phi(x)\Delta(x)dx < \chi(\kappa^* - \hat{\kappa}). \quad (\hat{3})$$

Proposition 1 remains true when  $\phi' \leq 0$ . To see this, divide ( $\hat{2}$ ) by the measure of treated students and combine with Condition  $\hat{1}$ (i) to obtain  $\left(\int_{\kappa^*}^1 \phi(x)\Delta(x)dx\right) / (1 - \kappa^*) \geq \phi(\kappa^*)\Delta(\kappa^*)$ . Because  $\phi' \leq 0$ , this implies that  $\left(\int_{\kappa^*}^1 \Delta(x)dx\right) / (1 - \kappa^*) \geq \Delta(\kappa^*)$ , where the inequality is strict if  $\phi' < 0$ . Intuitively, the gains for treating the treated must be even larger than the LATE if the administrator values such gains less. Analogous reasoning applied to Corollary 4 shows that the ATUT is bounded above by the LATE when  $\phi' \leq 0$ , and that this bound is strict when  $\phi' < 0$ . Therefore, the corollaries, in particular Corollary 5 bounding the ATE, also still obtained with the weighted problem ( $\hat{1}$ ). In summary, all of the bounds from the unweighted problem, including Corollary 5, which bounds the ATE, are also obtained for the weighted problem ( $\hat{1}$ ).

## D Variable marginal cost of treatment

Begin by relaxing Assumption 1(i), replacing it with

**Assumption 1'.** (i) *The cost function  $c(\cdot)$  is known and is non-negative, strictly increasing, and differentiable. The marginal cost function  $c'(\cdot)$  is monotonic.*

Note that Assumption 1'(i) still implies that the marginal cost of providing treatment is strictly positive. The second part of Assumption 1'(i) relaxes the constant marginal cost assumption.

I first adapt the conditions characterizing  $\kappa^*$ , in terms of  $\Delta(\cdot)$  and qualitative features of the cost function  $c(\cdot)$ . Specifically, I consider three cases for Assumption 1'(i): where the marginal cost is constant, decreasing, and increasing; these correspond to linear, concave, and convex cost functions, respectively. I then provide results bounding treatment effects of interest.

**Condition 1'** (Necessity). *The following necessary conditions must hold for  $\kappa^*$ :*

- (i) *MB=MC:  $\Delta(\kappa^*) = c'(1 - \kappa^*)$  for any cost function  $c(\cdot)$  satisfying Assumption 1'*
- (ii) *Increasing MB:  $\Delta'(\kappa^*) \geq 0$  if the marginal cost is constant or decreasing; this inequality is strict if the marginal cost is decreasing.*

*Proof.* Differentiate the administrator's problem (1) with respect to  $\tilde{\kappa}$  to obtain (i). Note that if the derivative is negative at a candidate solution satisfying (i) but the marginal cost is nonincreasing, the administrator would gain by not treating students just above  $\kappa^*$ , thereby obtaining (ii).  $\square$

Condition 1' is similar to Condition 1, except that Condition 1'(ii) has a strict inequality if the marginal cost of treatment is decreasing. As before, to guarantee uniqueness, inspection of (1) implies two additional conditions sufficient for characterizing  $\kappa^*$ . These conditions are identical to those in Condition 2, the only difference being that  $\chi$  no longer enters either expression.

**Condition 2'** (Sufficiency). *The fact the program was implemented implies that the total gain from treating those students was at least as large as the total costs, i.e.:*

$$\text{Participation: } \int_{\kappa^*}^1 \Delta(x) dx \geq c(1 - \kappa^*). \quad (2')$$

*The fact the program was not extended to  $\hat{\kappa} < \kappa^*$  implies that treating these students would be sub-optimal, i.e.:*

$$\int_{\hat{\kappa}}^{\kappa^*} \Delta(x) dx < c(1 - \hat{\kappa}) - c(1 - \kappa^*). \quad (3')$$

As before, two corollaries immediately follow.

**Corollary 1'.**  *$\Delta(\cdot)$  cannot be constant if the marginal cost of treatment is nonincreasing.*

**Corollary 2'.**  $\Delta(\cdot)$  is nondecreasing in  $x$  if the marginal cost of treatment is nonincreasing.

*Proof.* First consider the case with constant marginal cost of treatment  $\chi$ . The second part of Condition 2' says that it must be the case that  $\int_{\hat{\kappa}}^{\kappa^*} (\Delta(x) - \chi) dx < 0$ . Moreover, we know that  $\Delta(\kappa^*) = \chi$ . Therefore, if  $\Delta(\cdot)$  were decreasing in  $x$  then this inequality would be violated. If the marginal cost is decreasing then  $\Delta(\cdot)$  would have to be even more strongly increasing in  $x$  to satisfy the above inequality.  $\square$

As before, I next examine what can be deduced about averages of treatment effects for subsets of students.

**Corollary 3'.** The ATT is positive for any cost function  $c(\cdot)$  satisfying Assumption 1'.

*Proof.* The left side of (2') in Condition 2' is the total effect of treatment on the treated, i.e.  $(\int_{\kappa^*}^1 \frac{\Delta(x)}{(1-\kappa^*)} dx) (1 - \kappa^*)$ . Because the marginal cost of treatment is positive (Assumption 1'(i)), (2') implies that

$$\int_{\kappa^*}^1 \Delta(x) dx \geq c(1 - \kappa^*) > 0.$$

Divide through by  $(1 - \kappa^*)$  to obtain the result:

$$\underbrace{\int_{\kappa^*}^1 \frac{\Delta(x)}{(1 - \kappa^*)} dx}_{\text{ATT}} \geq \underbrace{\frac{c(1 - \kappa^*)}{(1 - \kappa^*)}}_{\text{avg. cost of treating treated}} > 0. \quad \square$$

Although Corollary 3' provides a lower bound for the average effect of treatment on the treated, there is no informative (i.e. lower than  $\bar{\Delta}$ ) upper bound. Corollary 3' makes no further assumptions about the shape of the cost function. However, if the marginal cost of treating students is nonincreasing, the lower bound on the average effect of treatment on the treated increases.

**Proposition 1'.** If the marginal cost of treatment is nonincreasing, the ATT is bounded below by the LATE at the treatment cutoff.

*Proof.* If the marginal cost of treatment is nonincreasing then  $c'(1 - \kappa^*) \leq \frac{c(1 - \kappa^*)}{1 - \kappa^*}$ , i.e. the marginal cost of treatment for  $1 - \kappa^*$  is no greater than the average cost of providing treatment for treated students. Insert this inequality into (2') and combine with this with

Condition 1'(i) to obtain

$$\underbrace{\frac{\int_{\kappa^*}^1 \Delta(x)dx}{1 - \kappa^*}}_{\text{ATT}} \geq \frac{c(1 - \kappa^*)}{1 - \kappa^*} \geq c'(1 - \kappa^*) = \underbrace{\Delta(\kappa^*)}_{\text{LATE at } \kappa^*}. \quad \square$$

As with Proposition 1, Proposition 1' shows that, if the marginal cost of treatment is nonincreasing, the discontinuity-based estimate provides a lower bound for the average effect of treatment on the treated. One should note that only qualitative information about the shape, not the level, of the marginal cost of treatment is all that is required for this result.

Although Corollary 2' rules out a decreasing treatment effect (if the marginal cost of treatment is nonincreasing), it could be the case that  $\Delta(\cdot)$  increases for some  $x < \kappa^*$ . Therefore, as before, it is useful to bound averages of  $\Delta(\cdot)$  itself for strict subsets of untreated students.

**Proposition 2'.** *There exists an informative upper bound for  $\int_a^b \Delta(x)dx$  for  $0 \leq a < b \leq \kappa^*$ .*

*Proof.* Suppose we would like to characterize  $\Delta(\cdot)$  for values less than  $\hat{x} < \kappa^*$ . Let  $\hat{\mu}$  be the measure of students under consideration and split (3') into two parts at  $\hat{x}$  and rearrange terms:

$$\int_{\hat{x}-\hat{\mu}}^{\hat{x}} \Delta(x)dx < c(1-(\hat{x}-\hat{\mu}))-c(1-\kappa^*)-\int_{\hat{x}}^{\kappa^*} \Delta(x)dx \Rightarrow \int_{\hat{x}-\hat{\mu}}^{\hat{x}} \Delta(x)dx < c(1-(\hat{x}-\hat{\mu}))-c(1-\kappa^*)-\underline{\Delta}(\kappa^* - \hat{x}), \quad (4')$$

where the implication follows from Assumption 1(iii).  $\square$

Setting the measure of students to whom the treatment is extended equal to  $\kappa^*$  provides the following result about the ATUT.

**Corollary 4'.** *The ATUT has an informative upper bound. If the marginal cost of treatment is nonincreasing, this upper bound is the LATE at the treatment cutoff.*

*Proof.* Let  $\hat{x} = \hat{\mu} = \kappa^*$  in (4') and divide through by  $\kappa^*$  to obtain the first result:

$$\underbrace{\int_0^{\kappa^*} \frac{\Delta(x)}{\kappa^*} dx}_{\text{ATUT}} < \underbrace{\frac{c(1) - c(1 - \kappa^*)}{\kappa^*}}_{>0, < \infty}, \quad (7')$$



where the right hand side is positive from Assumption 1'(i). For the second result, note that a nonincreasing marginal cost implies

$$\frac{c(1) - c(1 - \kappa^*)}{\kappa^*} < c'(1 - \kappa^*) = \underbrace{\Delta(\kappa^*)}_{\text{LATE at } \kappa^*},$$

where the equality follows from Proposition 1'(i).  $\square$

Analogously to the upper bound for the ATT, although Corollary 4' bounds the average of treatment effects for all untreated students, there is no informative (i.e. greater than  $\underline{\Delta}$ ) lower bound.

To summarize, optimality of  $\kappa^*$  implies a lower bound on the ATT and an upper bound on the ATUT. If the marginal cost of treatment is constant or decreasing then it must be the case that  $\text{ATUT} < \Delta(\kappa^*) \leq \text{ATT}$ . Though the ATT and ATUT are respectively bounded below and above by the cutoff LATE when marginal costs are nonincreasing, the LATE does not bound these moments when the marginal cost of treatment is *increasing*. Also note that, though the ATT is larger for treated students than the untreated when the marginal cost of treatment is nonincreasing, if the marginal cost is increasing, the ATUT could be larger than the ATT when the administrator is constrained to treat students starting from the top instead of the bottom.

## D.1 Bounding the ATE

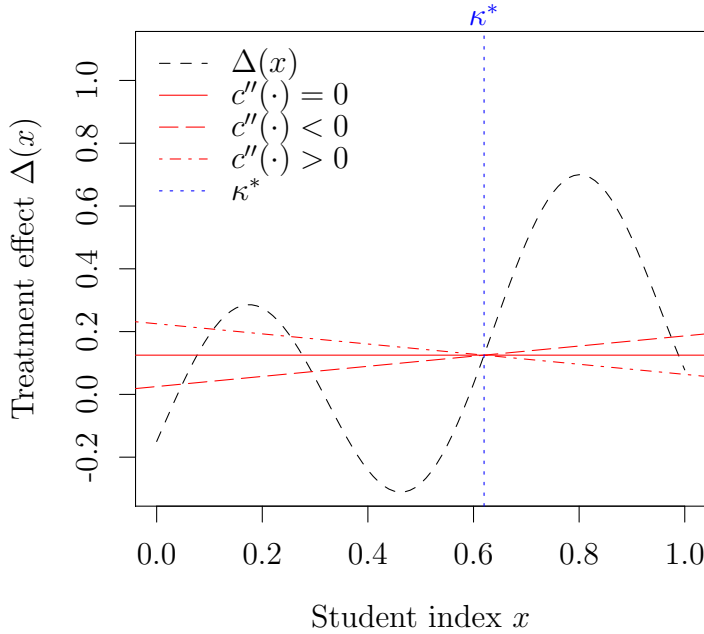
Corollary 5 bounds the ATE when the marginal cost of treatment is constant. This section studies the interplay between qualitative features of the cost of treatment and inferences about the distribution of treatment effects, by comparing three cases: constant, decreasing, and increasing marginal cost of treatment, where each marginal cost curve passes through the point  $(\kappa^*, \Delta(\kappa^*))$ . A decreasing marginal cost ( $c'' < 0$ ) might result from economies of scale, while an increasing marginal cost ( $c'' > 0$ ) might result from congestion effects, say if it becomes increasingly difficult to find a good fit for the program.

To begin, suppose the cost function is  $c(\mu) = \mu\chi$ . Then, as was shown in Section 3.1, the ATE lower bound is  $\Delta_{\kappa^*\chi}^{LB} \equiv \underline{\Delta}\kappa^* + \chi(1 - \kappa^*)$  and the ATE upper bound is  $\Delta_{\kappa^*\chi}^{UB} \equiv \chi\kappa^* + \overline{\Delta}(1 - \kappa^*)$ . Figure 4 builds on the example in Figure 1 to provide intuition for how the marginal cost of treatment bounds the ATE. Start with the solid red line representing a constant marginal cost of treatment, and rotate the cost function counterclockwise about the point  $(\kappa^*, \Delta(\kappa^*))$  to represent a decreasing marginal cost of treatment (long-dashed red line).<sup>38</sup> This rotation implies the ATT must be higher than the case corresponding to the

---

<sup>38</sup>Recall that this line is decreasing in  $x$  because the treatment is being extended from  $x = 1$  downwards.

constant marginal cost in order to satisfy (2'). Analogously, the maximum ATUT must be lower when marginal costs are decreasing; were they the same as with constant marginal costs, the administrator might gain from extending treatment to untreated units given that they now have a lower cost of being treated, violating (3'). The opposite holds true for when we rotate the cost curve clockwise about the point  $(\kappa^*, \Delta(\kappa^*))$ , to reflect an increasing marginal cost of treatment (dot-dashed red line). Table 3 summarizes these results, showing that when the marginal cost of treatment is decreasing, bounds on the ATE are tighter than they would be with a constant marginal cost, while when marginal cost is increasing, bounds on the ATE are looser.



Marginal cost	ATE bounds	
	Lower	Upper
Const. ( $c'' = 0$ )	$= \Delta_{\kappa^* \chi}^{LB}$	$= \Delta_{\kappa^* \chi}^{UB}$
Dec. ( $c'' < 0$ )	$> \Delta_{\kappa^* \chi}^{LB}$	$< \Delta_{\kappa^* \chi}^{UB}$
Inc. ( $c'' > 0$ )	$< \Delta_{\kappa^* \chi}^{LB}$	$> \Delta_{\kappa^* \chi}^{UB}$

Table 3: Summary of bounds on ATE

Figure 4: Example with different cost functions