

# Regression Discontinuity

Francis J. DiTraglia

Econ 224

## “Sharp” Regression Discontinuity

Suppose we are interested in learning the causal effect of a binary treatment  $D$  on an outcome  $Y$ . In some special settings, whether or not a person is treated is solely determined by a special covariate  $x$ , called the *running variable*

$$D_i = \begin{cases} 0 & \text{if } x_i < x_0 \\ 1 & \text{if } x_i \geq x_0 \end{cases}$$

The preceding expression says that  $D$  is a *deterministic function* of  $x$ : everyone who has  $x \geq x_0$  is treated, and no one who has  $x < x_0$  is treated. This setting is called a *sharp regression discontinuity design* and it provides us with a powerful tool for causal inference. We’ll distinguish this from another kind of regression discontinuity setup called a *fuzzy regression discontinuity design* below. I’ll use the shorthand RD to refer to regression discontinuity in these notes.

When we previously used regression to carry out causal inference, the idea was to compare two groups of people who had been *matched* using a set of covariates  $\mathbf{x}$ . One group was treated and the other was not, but both groups had exactly the same values of  $\mathbf{x}$ . Under the assumption that treatment is “as good as randomly assigned” after conditioning on  $\mathbf{x}$ , we could learn the causal effect of treatment by comparing the mean outcomes of the two groups.

Sharp RD is very different since there is no way to carry out matching using the running variable  $x$ . This is because everyone who has  $x < x_0$  is untreated while everyone who has  $x \geq x_0$  is treated. Instead of matching people who have the same covariate values, sharp RD *extrapolates* by comparing people with *different* covariate values. The basic idea is very simple: we compare people whose  $x$  is close to but slightly *below* the cutoff  $x_0$  to people whose  $x$  is close to but slightly *above* the cutoff. The idea here is that both  $D$  and  $x$  could effect  $Y$ , but at the cutoff  $x_0$ ,  $D$  “jumps.” As long as the direct effect of  $x$  on  $Y$  (i.e. the effect that doesn’t go through  $D$ ) doesn’t jump too, we’ll be able to notice the jump, or discontinuity, in our dataset. This is how we’ll isolate the causal effect of  $D$  on  $Y$ .

Now we’ll be a little more precise about the intuition from the preceding paragraph by thinking about exactly how  $x$  and  $Y$  are related. The key will be to create a link to potential outcomes, since this is our main tool for thinking about causality. To make things simple, let’s imagine that the running variable  $x$  is between zero and one, and that the cutoff  $x_0$  equals 0.5. This means that anyone with  $x < 0.5$  is untreated while anyone with  $x \geq 0.5$  is treated.

To begin, suppose we had a data for a large random sample of people who all had  $x_i = 0.3$ . If we took the mean  $Y$  for these people we would get an unbiased estimate of  $E[Y_i|x_i = 0.3]$ . The crucial point about sharp RD is that treatment is *completely determined* by  $x$ . This means that there is no way that someone with  $x_i = 0.3$  could have chosen to be treated: everyone with  $x < 0.5$  is untreated. This means that there is no selection bias, so  $E[Y_i|x_i = 0.3] = E[Y_{0,i}|x_i = 0.3]$ . Note that this is *not* the same thing as  $E[Y_{0,i}]$  since the running variable  $x$  could *directly* affect  $Y$ . For example,  $E[Y_{0,i}|x_i = 0.3]$  may not equal  $E[Y_{0,i}|x_i = 0.4]$  even though neither someone with  $x$  equal to 0.3 nor someone with  $x$  equal to 0.4 is treated. But this is fine, since we know how to use *predictive* modeling tools to estimate a conditional mean function. Here is the key point: since there is no selection into treatment for people

Let’s start by looking at a very simple example. Suppose we happened to know that the potential outcome when untreated,  $Y_0$  was linearly related to  $x$ , namely

$$E[Y_{0,i}|x_i] = \alpha + \beta x_i$$

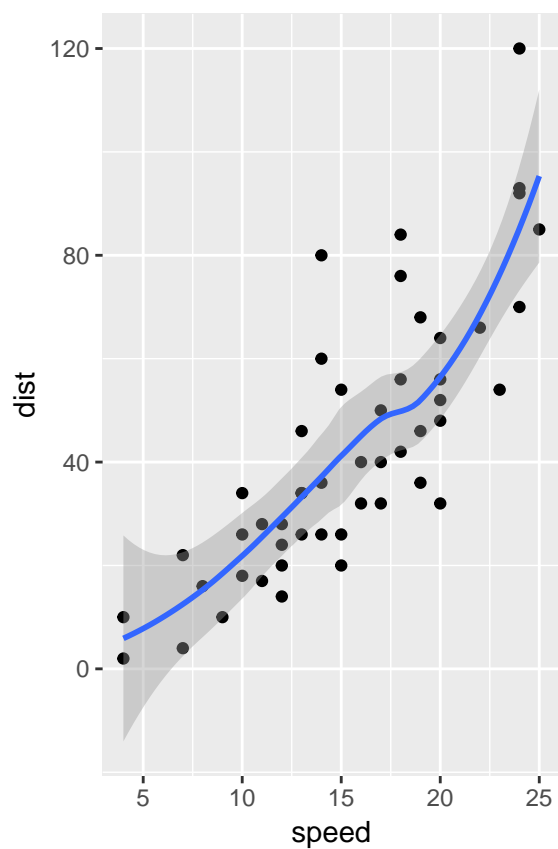
Now suppose we also knew that the potential outcome when treated was equal to the potential outcome when untreated, plus a constant shift  $\rho$ , that is

$$Y_{1,i} = Y_{0,i} + \rho$$

```
library(ggplot2)
summary(cars)
```

```
##      speed      dist
##  Min.   : 4.0    Min.   :  2.00
##  1st Qu.:12.0    1st Qu.: 26.00
##  Median :15.0    Median : 36.00
##  Mean   :15.4    Mean   : 42.98
##  3rd Qu.:19.0    3rd Qu.: 56.00
##  Max.   :25.0    Max.   :120.00
```

```
qplot(speed, dist, data=cars) +
  geom_smooth()
```



This is a nice example 4.