Chapter 6

Getting a Little Jumpy: Regression Discontinuity Designs

> But when you start exercising those rules, all sorts of processes start to happen and you start to find out all sorts of stuff about people . . . Its just a way of thinking about a problem, which lets the shape of the problem begin to emerge. The more rules, the tinier the rules, the more arbitrary they are, the better.
>
> Douglas Adams, *Mostly Harmless* (1995)

Regression discontinuity (RD) research designs exploit precise knowledge of the rules determining treatment. RD identification is based on the idea that in a highly rule-based world, some rules are arbitrary and therefore provide good experiments. RD comes in two styles, fuzzy and sharp. The sharp design can be seen as a selection-on-observables story. The fuzzy design leads to an instrumental-variables-type setup.

6.1 Sharp RD

Sharp RD is used when treatment status is a deterministic and discontinuous function of a covariate, x_i. Suppose, for example, that

$$\mathrm{D}_i = \begin{cases} 1 & \text{if } x_i \geq x_0 \\ 0 & \text{if } x_i < x_0 \end{cases} . \tag{6.1.1}$$

where x_0 is a known threshold or cutoff. This assignment mechanism is a deterministic function of x_i because once we know x_i we know D_i. It's a discontinuous function because no matter how close x_i gets to x_0, treatment is unchanged until $x_i = x_0$.

This may seem a little abstract, so here is an example. American high school students are awarded National Merit Scholarship Awards on the basis of PSAT scores, a test taken by most college-bound high school juniors, especially those who will later take the SAT. The question that motivated the first discussions of RD is whether students who win these awards are more likely to finish college (Thistlewaithe and Campbell, 1960; Campbell, 1969). Sharp RD compares the college completion rates of students with PSAT scores just above and just below the National Merit Award thresholds. In general, we might expect students with higher PSAT scores to be more likely to finish college, but this effect can be controlled by fitting a regression to the relationship between college completion and PSAT scores, at least in the neighborhood of the award cutoff. In this example, jumps in the relationship between PSAT scores and college attendance in the neighborhood of the award threshold are taken as evidence of a treatment effect. It is this jump in regression lines that gives RD its name.[1]

An interesting and important feature of RD, highlighted in a recent survey of RD by Imbens and Lemieux (2008), is that there is *no* value of x_i at which we get to observe both treatment and control observations. Unlike full-covariate matching strategies, which are based on treatment-control comparisons conditional on covariate values where there is some overlap, the validity of RD turns on our willingness to extrapolate across covariate values, at least in a neighborhood of the discontinuity. This is one reason why sharp RD is usually

[1] The basic structure of RD designs appears to have emerged simultaneously in a number of disciplines but has only recently become important in applied econometrics. Cook (2008) gives an intellectual history. In a recent paper using Lalonde (1986) style within-study comparisons, Cook and Wong (2008) find that RD generally does a good job of reproducing the results from randomized trials.

seen as distinct from other control strategies. For this same reason, we typically cannot afford to be as agnostic about regression functional form in the RD world as in the world of Chapter 3.

Figure 6.1.1 illustrates a hypothetical RD scenario where those with $x_i \geq 0.5$ are treated. In Panel A, the trend relationship between Y_i and x_i is linear, while in Panel B, it's nonlinear. In both cases, there is a discontinuity in the relation between $E[\text{Y}_{0i}|x_i]$ and x_i around the point x_0.

A simple model formalizes the RD idea. Suppose that in addition to the assignment mechanism, (6.1.1), potential outcomes can be described by a linear, constant-effects model

$$\begin{aligned} E[\text{Y}_{0i}|x_i] &= \alpha + \beta x_i \\ \text{Y}_{1i} &= \text{Y}_{0i} + \rho \end{aligned}$$

This leads to the regression,

$$\text{Y}_i = \alpha + \beta x_i + \rho \text{D}_i + \eta_i, \tag{6.1.2}$$

where ρ is the causal effect of interest. The key difference between this regression and others we've used to estimate treatment effects (e.g., in Chapter 3) is that D_i, the regressor of interest, is not only correlated with x_i, it is a deterministic function of x_i. RD captures causal effects by distinguishing the nonlinear and discontinuous function, $1(x_i \geq x_0)$, from the smooth and (in this case) linear function, x_i.

But what if the trend relation, $E[\text{Y}_{0i}|x_i]$, is nonlinear? To be precise, suppose that $E[\text{Y}_{0i}|x_i] = f(x_i)$ for some reasonably smooth function, $f(x_i)$. Panel B in Figure 6.1.1 suggests there is still hope even in this more general case. Now we can construct RD estimates by fitting

$$\text{Y}_i = f(x_i) + \rho \text{D}_i + \eta_i, \tag{6.1.3}$$

where again, $\text{D}_i = 1(x_i \geq x_0)$ is discontinuous in x_i at x_0. As long as $f(x_i)$ is continuous in a neighborhood of x_0, it should be possible to estimate a model like (6.1.3), even with a flexible functional form for $f(x_i)$. For example, modeling $f(x_i)$ with a p^{th}-order polynomial, RD estimates can be constructed from the regression

$$\text{Y}_i = \alpha + \beta_1 x_i + \beta_2 x_i^2 + ... + \beta_p x_i^p + \rho \text{D}_i + \eta_i. \tag{6.1.4}$$

A generalization of RD based on (6.1.4) allows different trend functions for $E[\text{Y}_{0i}|x_i]$ and $E[\text{Y}_{1i}|x_i]$. Modeling both of these CEFs with p^{th}-order polynomials, we have

$$\begin{aligned} E[\text{Y}_{0i}|x_i] &= f_0(x_i) = \alpha + \beta_{01}\tilde{x}_i + \beta_{02}\tilde{x}_i^2 + ... + \beta_{0p}\tilde{x}_i^p \\ E[\text{Y}_{1i}|x_i] &= f_1(x_i) = \alpha + \rho + \beta_{11}\tilde{x}_i + \beta_{12}\tilde{x}_i{}^2 + ... + \beta_{1p}\tilde{x}_i{}^p, \end{aligned}$$

where $\tilde{x}_i \equiv x_i - x_0$. Centering x_i at x_0 is just a normalization; it ensures that the treatment effect at $x_i = x_0$ is still the coefficient on D_i in the regression model with interactions.

To derive a regression model that can be used to estimate the effects interest in this case, we use the fact that D_i is a deterministic function of x_i to write

$$E[\text{Y}_i|x_i] = E[\text{Y}_{0i}|x_i] + E[\text{Y}_{1i} - \text{Y}_{0i}|x_i]\text{D}_i.$$

Substituting polynomials for conditional expectations, we then have

$$\begin{aligned} \text{Y}_i &= \alpha + \beta_{01}\tilde{x}_i + \beta_{02}\tilde{x}_i^2 + ... + \beta_{0p}\tilde{x}_i^p \\ &\quad + \rho \text{D}_i + \beta_1^* \text{D}_i\tilde{x}_i + \beta_2^* \text{D}_i\tilde{x}_i{}^2 + ... + \beta_p^* \text{D}_i\tilde{x}_i{}^p + \eta_i, \end{aligned} \tag{6.1.6}$$

where $\beta_1^* = \beta_{11} - \beta_{01}$, $\beta_2^* = \beta_{12} - \beta_{02}$, and $\beta_p^* = \beta_{1p} - \beta_{0p}$ and the error term, η_i, is the CEF residual.

Equation (6.1.4) is a special case of (6.1.6) where $\beta_1^* = \beta_2^* = \beta_p^* = 0$. In the more general model, the treatment effect at $x_i - x_0 = c > 0$ is $\rho + \beta_1^* c + \beta_2^* c^2 + ... + \beta_p^* c^p$, while the treatment effect at x_0 is ρ. The model with interactions has the attraction that it imposes no restrictions on the underlying conditional mean functions But in our experience, RD estimates of ρ based on the simpler model, (6.1.4), usually turn out to be similar to those based on (6.1.6).

The validity of RD estimates based on (6.1.4) or (6.1.6) turns on whether polynomial models provide an adequate description of $E[\text{Y}_{0i}|\text{X}_i]$. If not, then what looks like a jump due to treatment might simply be an unaccounted-for nonlinearity in the counterfactual conditional mean function. This possibility is illustrated in Panel C of Figure 6.1.1, which shows how a sharp turn in $E[\text{Y}_{0i}|x_i]$ might be mistaken for a jump from one regression line to another. To reduce the likelihood of such mistakes, we can look only at data in a

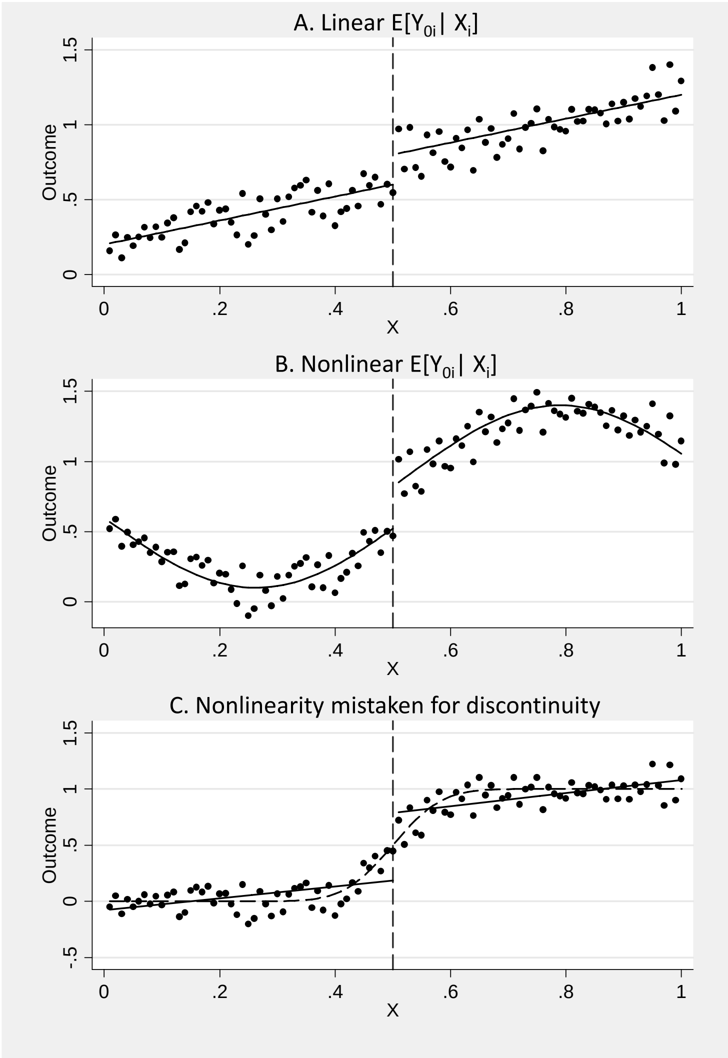

Figure 6.1.1: The sharp regression discontinuity design