PPOL 503-03, PPOL 503-04, Fall 2016
Course Notes #18: Regression Discontinuity Design

A. Overview

1. In its simplest most traditional form, the RD design is a pre-test-post-test comparison group strategy. The unique characteristic which sets RD designs apart from other pre-post group designs is the method by which research participants are assigned to conditions.

2. RD designs—participants are assigned to program or comparison groups solely on the basis of a cutoff score on a pre-program measure.

3. The RD design is distinguished from randomized experiments and other quasi-experimental strategies by its unique method of assignment. The cutoff criterion is the major advantage of the RD design; it is appropriate when we wish to target a program or treatment to those who most deserve it.

4. RD design does not require us to assign potentially needy individuals to a no-program comparison group in order to evaluate the effectiveness of the program.

5. Until recently, RD design was not widely used because at first glance it does not make sense. The typical design would have a comparison group that is equivalent to program groups on pre-program indicators so that one can attribute post-program differences to the program itself. Yet, because of the cutoff criterion, program and comparison groups are deliberately and maximally different on pre-program characteristics.

6. Inferences drawn from a well-implemented RD design are comparable in internal validity to conclusions drawn from randomized experiments.

B. The Basic Design

1. The “basic” RD design is a pretest-posttest two group design. The term “pretest-posttest” implies that the same measure is administered before and after some program or treatment.

2. A cutoff value on the pretest or pre-program measure is being used to assign persons or other units of the program. The basic design as a two group design implies that a single pretest cutoff score is used to assign participants to either the program or comparison group.

3. With RD design a program effects is suggested if we observe a jump or discontinuity in the regression lines at the cutoff point.

C. Logic of the RD Design

1. Key feature of the RD design: assignment based on a cutoff value on a pre-program measure. Cutoff rule for a simple two group case is: a) all persons on one side of the cutoff are assigned to one group, b) all persons on the other side of the cutoff are assigned to the other, and c) need a continuous pre-program measure.

2. Selection of the Cutoff: based on two factors:
a) program resources that are available;
for example, if a program can only handle 25 persons and 70 individuals apply, one can choose a cutoff that distinguishes the 25 most needy persons.
b) cutoff chosen on substantive grounds;
if the pre-program assignment measure is an indication of severity of illness measured on a scale of 1 to 7 and physicians contend that patients scoring 5 or more are critical and fit the criteria defined for program participation then a cutoff value of 5 may be used.

D. Role of the Comparison Group in RD Designs

1. In experimental or other quasi-experimental designs one either assumes or tries to provide evidence that the program and comparison groups are equivalent prior to the program so that the post-program differences can be attributed to manipulation.

2. In contrast, the RD design assumes that in the absence of the program, the pre-post relationship would be equivalent for the two groups.

3. Strength of the RD design depends on two factors: a) there is no spurious discontinuity in the pre-post relationship which coincides with the

cutoff point; b) the degree to which we can know and correctly model the pre-post relationship.

E. Internal Validity of the RD Design

1. Internal validity refers to whether one can infer that the treatment or program being investigated caused a change in outcome indicators.

2. The central question is whether any observed outcome differences between groups can be attributed to the program or some other factor.

3. In designs that do not use random assignment, the central internal validity concern revolves around the possibility that groups may not be equivalent prior to the program. The term “selection bias” is used to describe the case where the pre-program differences between groups are responsible for post-program differences.

4. In RD designs, because of the deliberate pre-program differences between groups, there are several selection threats to internal validity which might appear to be problematic.

5. Selection-maturation threat implies that different rates of maturation between the groups might explain outcome differences. Example: pre-post distribution with a linear relationship having a slope equal to two. If a person has a pretest score of 10, his posttest score would be 20 (absolute gain of 10). For a person with a pretest score of 50, one would predict a posttest score of 100 (absolute gain of 50). In the RD design, we expect that all participants may mature and in absolute terms this maturation rate may be different for the two groups on average.

Key Point: Program effect in the RD design is not indicated by the posttest averages of the groups, but rather by a change in the pre-post relationship at the cutoff point. Thus, in order to selection-maturation to pose a threat to internal validity in RD designs, it must induce a discontinuity in the pre-post relationship which happens to coincide with the cutoff point—an unlikely scenario.

6. Another selection threat to internal validity concerns the possibility of differential regression to the mean. The phenomenon of regression to the mean arises when we asymmetrically sample groups from a distribution.

On any subsequent measure, the obtained sample mean will be closer to the population mean for that measure than the sample mean from the original distribution is to its population mean. In RD designs we deliberately create asymmetric samples and expect regression towards the mean in both groups. In general, we expect the low-scoring pretest group to evidence a relative gain on the post-test and the high-scoring pretest group to show a relative loss. We do not expect that regression to the mean will result in a discontinuity in the bivariate relationship coincidental with the cutoff point.

7. Although the RD design may initially seem susceptible to selection bias, it is not. These considerations indicate that only factors that would naturally induce a discontinuity in the pre-post relationship could be considered threats to internal validity.

F. Ethics and the RD Design

1. The RD designs are not as statistically powerful as randomized experiments. In order to achieve the same level of statistical accuracy, an RD design needs as much as 2.75 times the participants as a randomized experiment. If a RE design needs 100 participants to achieve a certain level of power, the RD design might need as many as 275.

2. Why would one ever use the RD design instead of a randomized one? ANSWER: the RD design allows us to assign the treatment program to those who most need or deserve it.

G. Assumptions of RD Analysis

1. The cutoff criterion must be followed without exception.

2. The pre-post distribution: It is assumed that the pre-post distribution is describable as a polynomial function. If the true pre-post relationship is logarithmic, exponential or some other function, the model is misspecified and the program effects may be biased.

3. Comparison Group Pretest Variance: need to have sufficient number of pretest values in the comparison group to estimate the pre-post regression line.

4. Continuous Pretest Distribution: Both groups must come from a single continuous distribution with the division determined by the cutoff point.

5. Program Implementation: It is assumed that the program is uniformly delivered to all recipients.

H. Model Specification

1. Major problem in analyzing data from the RD design is model misspecification. If the true model is curved and we assume it is linear, we are likely to wrongly conclude that the treatment made a difference when it did not.

2. If we estimate a model that includes unnecessary variables, the estimate of the treatment effect is unbiased, but inefficient. So we may conclude that the treatment does not work when in fact it does.

3. Omitted variables cause the treatment effect to be biased and inconsistent.

I. Steps in the Analysis

a) Transform the pretest: Subtract the cutoff value from each pretest score.
b) Examine the relationship visually: 1) Determine if there is any visually discernable discontinuity in the relationship at the cutoff. The discontinuity could be a change in the level vertically (main effect), a change in the slope (interaction effect) or both. 2) Determine the degree of the polynomial function if there flexion points or bends in the function.
c) Specify higher-order terms
d) Estimate the initial model: Regress the posttest scores Y, on the modified pretest score X*, the treatment variable Z and other covariates.

$$Y_i = \beta_0 + \beta_1 X_i^* + \beta_2 Z_i + \beta_3 X_i * Z_i + \beta_5 (X_i^*)^2$$

$$+ \beta_3 (X_i^*)^2 Z_i + e_i$$

Z_i = dummy variable for treatment group (1 = treatment, 0 = control)

$X_i^* = X_i - X_c$ = transformed pre-test score