

# Regression Discontinuity Designs

James H. Steiger

Department of Psychology and Human Development  
Vanderbilt University

Multilevel Regression Modeling, 2009

# Regression Discontinuity Designs

- 1 Introduction
- 2 The Basics of Regression Discontinuity
- 3 Analyzing the Effect
- 4 What Can Go Wrong? (C.P.)

# The Regression Discontinuity Design

- Regression discontinuity designs have wide application in a variety of fields
- Under appropriate assumptions, they allow causal inferences in situations where they seem very counterintuitive
- Rather than being damaged by selection, the design capitalizes on it

# The Regression Discontinuity Design

- Regression discontinuity designs have wide application in a variety of fields
- Under appropriate assumptions, they allow causal inferences in situations where they seem very counterintuitive
- Rather than being damaged by selection, the design capitalizes on it

# The Regression Discontinuity Design

- Regression discontinuity designs have wide application in a variety of fields
- Under appropriate assumptions, they allow causal inferences in situations where they seem very counterintuitive
- Rather than being damaged by selection, the design capitalizes on it

## An Introductory Example

Shadish, Cook and Campbell (2002, p. 207) discuss the study by Berk and colleagues examining the effect of receiving unemployment compensation support on recidivism rates of newly released ex-convicts.

- Newly released prisoners received unemployment compensation support, but *only if they had worked more than 652 hours over the previous 12 months while in prison*
- Those who had worked fewer hours were ineligible
- There were no exceptions
- Berk and Rauma (1983) found that those receiving compensation had a recidivism rate 13% lower than controls

## An Introductory Example

Shadish, Cook and Campbell (2002, p. 207) discuss the study by Berk and colleagues examining the effect of receiving unemployment compensation support on recidivism rates of newly released ex-convicts.

- Newly released prisoners received unemployment compensation support, but *only if they had worked more than 652 hours over the previous 12 months while in prison*
- Those who had worked fewer hours were ineligible
- There were no exceptions
- Berk and Rauma (1983) found that those receiving compensation had a recidivism rate 13% lower than controls

## An Introductory Example

Shadish, Cook and Campbell (2002, p. 207) discuss the study by Berk and colleagues examining the effect of receiving unemployment compensation support on recidivism rates of newly released ex-convicts.

- Newly released prisoners received unemployment compensation support, but *only if they had worked more than 652 hours over the previous 12 months while in prison*
- Those who had worked fewer hours were ineligible
- There were no exceptions
- Berk and Rauma (1983) found that those receiving compensation had a recidivism rate 13% lower than controls



## An Introductory Example

Shadish, Cook and Campbell (2002, p. 207) discuss the study by Berk and colleagues examining the effect of receiving unemployment compensation support on recidivism rates of newly released ex-convicts.

- Newly released prisoners received unemployment compensation support, but *only if they had worked more than 652 hours over the previous 12 months while in prison*
- Those who had worked fewer hours were ineligible
- There were no exceptions
- Berk and Rauma (1983) found that those receiving compensation had a recidivism rate 13% lower than controls

## The Basic Design Structure

- Experimenter must control assignment of participants to 2 or more treatments
- The assignment is made on the basis of a strict *cutoff score* on a *treatment assignment variable*
- The assignment variable can be any measure taken *prior to treatment*

## The Basic Design Structure

- Experimenter must control assignment of participants to 2 or more treatments
- The assignment is made on the basis of a strict *cutoff score* on a *treatment assignment variable*
- The assignment variable can be any measure taken *prior to treatment*

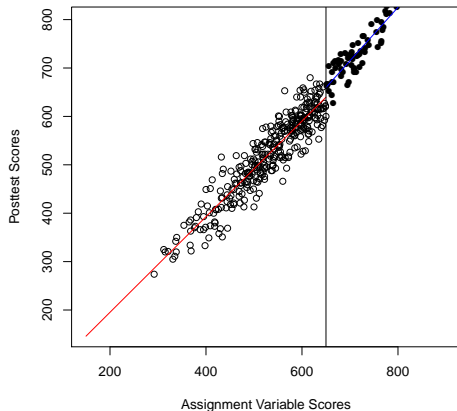
## The Basic Design Structure

- Experimenter must control assignment of participants to 2 or more treatments
- The assignment is made on the basis of a strict *cutoff score* on a *treatment assignment variable*
- The assignment variable can be any measure taken *prior to treatment*

## A Graphical Example

Centennial High is a high school in an upper middle class area of Philadelphia, PA. In 1997, every student at Centennial High took the English PSAT, and only those scoring above 650 were given a special training program in writing. Subsequently, all students took the Verbal SAT, and scores were recorded.

# A Graphical Example



## The Model

The simplest analysis measures the effect of the treatment with the model

$$y_i = \beta_0 + \beta_1 T_i + \beta_2(x_i - x_c) + \epsilon_i \quad (1)$$

$x_c$  is the cutoff score, and centering the  $x$  scores around the cutoff causes the equation to estimate the treatment effect at the cutoff score, where the groups are most similar.

## The analysis

```
> x.centered ← x-650  
> fit ← lm(y~T+x.centered)  
> summary(fit)
```

Call:

```
lm(formula = y ~ T + x.centered)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-84.9832	-19.5729	-0.7468	20.3798	98.0410

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	640.43862	2.89906	220.912	< 2e-16 ***
T	26.57391	5.52041	4.814	2.11e-06 ***
x.centered	0.99395	0.02013	49.375	< 2e-16 ***

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 30.19 on 397 degrees of freedom

Multiple R-squared: 0.9272, Adjusted R-squared: 0.9269

F-statistic: 2530 on 2 and 397 DF, p-value: < 2.2e-16



## Threats to Validity

Key assumptions in RD designs are

- The assignment mechanism is fixed and performed exactly according to  $X$  and the cutoff value
- The functional form of the regression model is correct

With the above in mind, take out a piece of paper and spend the next couple of minutes imagining one or two ways that the regression discontinuity design can mislead.