Regression Discontinuity Design

Motivating example

- Many districts have summer school to help kids improve outcomes between grades
  - Enrichment, or
  - Assist those lagging
- Research question: does summer school improve outcomes
- Variables:
  - $x=1$ is summer school after grade $g$
  - $y$ = test score in grade $g+1$

LUSDINE

- To be promoted to the next grade, students need to demonstrate proficiency in math and reading
  - Determined by test scores
- If the test scores are too low – mandatory summer school
- After summer school, re-take tests at the end of summer, if pass, then promoted

Equation of interest

$y_i = \beta_0 + x_i \beta_1 + \epsilon_i$

Problem: what do you anticipate is $\text{cov}(x_i, \epsilon_i)$?
### Situation

- Let Z be test score – Z is scaled such that
  - \( Z \geq 0 \) not enrolled in summer school
  - \( Z < 0 \) enrolled in summer school
- Consider two kids
  - #1: \( Z = \varepsilon \)
  - #2: \( Z = -\varepsilon \)
  - Where \( \varepsilon \) is small

### Intuitive understanding

- Participants in SS are very different
- However, at the margin, those just at \( Z = 0 \) are virtually identical
- One with \( z = -\varepsilon \) is assigned to summer school, but \( z = \varepsilon \) is not
- Therefore, we should see two things

### Variable Definitions

- \( y_i \) = outcome of interest
- \( x_i = 1 \) if NOT in summer school, \( = 1 \) if in
- \( D_i = I(z_i \geq 0) \) -- I is indicator function that equals 1 when true, \( = 0 \) otherwise
- \( z_i \) = running variable that determines eligibility for summer school. \( z \) is re-scaled so that \( z_i = 0 \) for the lowest value where \( D_i = 1 \)
- \( w_i \) are other covariates
Initial equation

\[ x_i = \theta_0 + D_i \theta_1 + h_j(z_i) + w_i \theta_2 + u_i \]

\[ h_j(z_i) = \text{polynomial in } z \]

\[ h_j(z_i) = 0 \text{ at } z = 0 \]

\[ \hat{x}_i \text{ just at } z_i = 0 \text{ with summer school option} \]
\[ \hat{x}_i^1 = \hat{\theta}_0 + \hat{\theta}_1 + w_i \hat{\theta}_2 \]
\[ \hat{x}_i \text{ just at } z_i = 0 \text{ without summer school} \]
\[ \hat{x}_i^0 = \hat{\theta}_0 + w_i \hat{\theta}_2 \]

therefore
\[ \hat{\theta}_1 - \hat{\theta}_1^0 = \hat{\Delta}_1 \]

If \( \hat{\Delta}_1 = 1 \) Sharp design

If \( \hat{\Delta}_1 < 1 \) fuzzy design

\[ \text{ RDD System} \]

\textbf{Structural equation:}
\[ y_i = \beta_0 + x_i \beta_1 + h(z_i) + w_i \beta_2 + \varepsilon_i \]

\textbf{First stage:}
\[ x_i = \theta_0 + D_i \theta_1 + h_j(z_i) + w_i \theta_2 + u_i \]

\textbf{reduced – form}
\[ y_i = \pi_0 + D_i \pi_1 + h_j(z_i) + w_i \pi_2 + v_i \]

Note that
\[ \beta_1 = \pi_1 / \theta_1 \]
RDD Equation

\[ \hat{y} \text{ just at } z_i = 0 \text{ with treatment} \]
\[ \hat{y}_i^1 = \hat{\pi}_0 + \hat{\pi}_1 + w_i\hat{\pi}_2 \]

\[ \hat{y} \text{ just at } z_i = 0 \text{ without treatment} \]
\[ \hat{y}_i^0 = \hat{\pi}_0 + w_i\hat{\pi}_2 \]

therefore
\[ \hat{y}_i^1 - \hat{y}_i^0 = \hat{\pi}_1 \]

Order of polynomial

\[ h(z_i) = \text{polynomial in } z \]

First order: \[ h(z_i) = Dz_i\gamma_i + (1-D)z_i\alpha_i \]

Third order: \[ h(z_i) = Dz_i\gamma_i + Dz_i^2\gamma_i + Dz_i^3\gamma_i + (1-D)z_i\alpha_i + (1-D)z_i^2\alpha_i + (1-D)z_i^3\alpha_i \]

---

Key assumption of RDD models

- People right above and below \( Z_0 \) are functionally identical
- Random variation puts someone above \( Z_0 \) and someone below
- However, this small different generates big differences in treatment (x)
- Therefore any difference in Y right at \( Z_0 \) is due to x

Limitation

- Treatment is identified for people at the \( z_i=0 \)
- Therefore, model identifies the effect for people at that point
- Does not say whether outcomes change when the critical value is moved
Table 1

<table>
<thead>
<tr>
<th>Grade 3</th>
<th>Attendance SS</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td><strong>Outcome</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002 math score</td>
<td>641.8</td>
<td>650.4</td>
</tr>
<tr>
<td>(0.12)</td>
<td>(0.16)</td>
<td>(0.16)</td>
</tr>
<tr>
<td>2002 reading score</td>
<td>649.7</td>
<td>621.6</td>
</tr>
<tr>
<td>(0.176)</td>
<td>(0.24)</td>
<td>(0.204)</td>
</tr>
<tr>
<td>Summer school attendance</td>
<td>24</td>
<td>1</td>
</tr>
<tr>
<td>(0.02)</td>
<td>(0)</td>
<td>(0)</td>
</tr>
<tr>
<td>Days attended</td>
<td>4,373</td>
<td>18,200</td>
</tr>
</tbody>
</table>

Table 2

<table>
<thead>
<tr>
<th>Grade</th>
<th>Effect of being randomized</th>
<th>Effect of SS attendance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Math</td>
<td>Reduced Form</td>
</tr>
<tr>
<td>Group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1.09</td>
<td>0.049</td>
</tr>
<tr>
<td>5</td>
<td>0.38</td>
<td>0.090</td>
</tr>
<tr>
<td>6</td>
<td>0.11</td>
<td>0.010</td>
</tr>
</tbody>
</table>

Math:

- Grade 3: 0.049/0.383 = 0.128
- Grade 5: 0.093/0.385 = 0.241
- Grade 6: 0.061/0.320 = 0.190
Example: Selective High Schools
Alcohol and Mortality

• Alcohol use
  – Reduces inhibition, increases aggression, compromises motor skills, blurs vision
• Use is associated with increased
  – Motor vehicle accidents, suicides, homicides, falls, burns, drowning
• Between 1975-95 Alcohol was involved in
  – 40% traffic deaths, 47% homicides, 30% suicides

Alcohol Abuse among Young Adults

• 4 million adults reported driving impaired in 2010
  – 112 million episodes
  – 81% due to men
  – Men aged 21-34 1/3 of all episodes
• Drunk driving deaths in 2012
  – 10,322 (1/3 of all traffic deaths)
  – In fatal crashes, 1/3 of drunk drivers are aged 21-24

Binge Drinking

• Definition
  – Men: 5+ drinks in a row one sitting
  – Women: 4+
• 30-day Prevalence by age
  – 18-24: 28.2%
  – All ages: 17.1%
• Frequency (among binge drinkers)
  – 18-24: 4.2 times
  – All ages: 4.4 times
• Intensity (max number of drinks among bingers)
  – 18-24: 9.3
  – All ages: 7.9

• Easy to establish
  – $\Pr(\text{Drinking} \mid \text{MV death}) > \Pr(\sim \text{Drinking} \mid \text{MV Death})$

• Much harder to establish
  $\frac{\partial (\text{MV Death})}{\partial (\text{Alcohol use})}$

• What is required to identify this derivative?
  – A change in alcohol use

• Best option: variation in use generated by state policies

State alcohol control policies

• MLDA
• Price/taxes
• Retail sales restrictions
  – Date/time, Dram shop rules
• Drunk driving laws
  – BAC thresholds
  – Per se license revocation
  – Checkpoints
  – Mandatory minimum sentences

MLDA

• Used to vary across states
• In 1983, 35 states had MLDA<21
• National Minimum Drinking Age Act 1984
  – Passed July 17, 1984
  – Reduced federal highway funds for states by 10% if they had MLDA < 21
  – All states now have MLDA 21
  – US one of 4 countries with MLDA of 21
Previous research

• Difference in difference models
• 1983 law as the impetus
• MLDA <21 increases
  – Drinking, binge drinking, MV fatalities
  – MLDA 18 real problematic because it gets beer into high schools

This paper

• How does aging into drinking age impact use?
  – Estimated by RDD
  – sharp increase in use right at 21
• Given the change in use – is there a corresponding change in mortality outcomes

Nat. Health Interview Survey

• 1997-2005
• Random sample of US households
• Have date of birth and date of survey
• Measures drinking participation, heavy drinking over past week, month, year
  – Why is past-year drinking problematic for this question?
  – 71% use last month or week as reference period

Mortality detail files

• Annual data – authors use 1997-2005
• Contain census of deaths in the US (2.7 million/year)
• Variables: demographics, place, date, cause
• Restricted use data has date of birth
• Place people into months of age
Two groups of measures for alcohol use

- Participation
  - Any drinking in lifetime
  - 12 or more drinking in a year
  - Any heavy drinking past year
- Intensity
  - Proportion of days drinking
  - Proportion days heavy drinking
  - Drinks/day

Table 1: Participation

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportion of days drinking</td>
<td>0.0448</td>
<td>0.0422</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Over 21</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>16,107</td>
<td>16,107</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prob &gt; Chi-Squared</td>
<td>0.02</td>
<td>0.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 or more drinks in one year</td>
<td>0.0776</td>
<td>0.0716</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Over 21</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>16,107</td>
<td>16,107</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prob &gt; Chi-Squared</td>
<td>0.02</td>
<td>0.02</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Intensity

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportion of days drinking</td>
<td>0.0120</td>
<td>0.0060</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Over 21</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>15,825</td>
<td>15,825</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prob &gt; Chi-Squared</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drinks per day on days drinking</td>
<td>0.2347</td>
<td>0.2347</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Over 21</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>0.9096</td>
<td>0.9096</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prob &gt; Chi-Squared</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Covariates</td>
<td>N</td>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weights</td>
<td>N</td>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quadratic terms</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Curv terms</td>
<td>N</td>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LLR</td>
<td>N</td>
<td>N</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 3. Age Profiles for Death Rates

Note: Deaths from the National Vital Statistics Records. Includes all deaths that occurred in the United States between 1985-2003. The population denominators are derived from the census. See Table C.1 for a list of causes of death.

Figure 4. Age Profiles for Death Rates by External Cause

Note: See notes for Figure 3. The categories are mutually exclusive. The order of precedence is homicide, suicide, MXA, deaths with a mention of alcohol, and deaths with a mention of drugs. The IC9 and IC10 codes are in Appendix C.

Table 5. Decomposition in Loss Deaths at Age 21 Due to External Causes

<table>
<thead>
<tr>
<th>Cause</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alcohol</td>
<td>0.388</td>
<td>0.346</td>
<td>0.401</td>
<td>0.401</td>
</tr>
<tr>
<td>Homicide</td>
<td>0.370</td>
<td>0.128</td>
<td>0.061</td>
<td>0.061</td>
</tr>
<tr>
<td>Suicide</td>
<td>0.301</td>
<td>0.234</td>
<td>0.245</td>
<td>0.245</td>
</tr>
<tr>
<td>Other external</td>
<td>0.082</td>
<td>0.021</td>
<td>0.021</td>
<td>0.021</td>
</tr>
</tbody>
</table>

Note: See notes from Table 4. These are 7N observations where there are N deaths coded as due to alcohol. For this variable, C was added to the dependent variable before taking the log. There are 11 observations where there are N deaths coded as due to drug use, for this variable C was added to the count before taking the log.
Estimating RDD models

- All states moved to MLDA 21 by 1988
- Use data on deaths among people with Social Security Numbers from 1989-2008
- Generate monthly counts of deaths by age/months – from age=19, month=0 through age=21, month=11
- 48 observations
• * generate ln death counts
  gen deathsl=ln(deaths)

• * rescale the running variable so that
  index = 0 in the month someone turns 21
  gen rv=index-25

• * treatment dummy
  gen treatment=index>=25

• * generate separate running variables before and
  after the discontinuity
  gen rv_after1=treat*rv
  gen rv_after2=rv_after1*rv_after1
  gen rv_after3=rv_after2*rv_after1
  gen rv_before1=(1-treat)*rv
  gen rv_before2=rv_before1*rv_before1
  gen rv_before3=rv_before2*rv_before1

** Medicare **

• Introduced in 1963

• Federal health insurance programs for
  - the elderly
  - Disabled

• Among elderly – become eligible at age 65

• Two things happen at age 65
  - More become insured
  - Insurance is more generous
Medicare

- 2007
- 44.1 million recipients
- $432 bill. exp.
- 3.2% of GDP
- 16% of fed. budget

- 2040
- 87 million recipients
- 7.6% of GDP
- 30% of fed. budget

This paper

- Change in eligibility at age 65
- We should see
  - Greater levels of insurance
  - Greater use of medical services
- If health insurance improves health, we should also see a reduction in mortality

Sample

- CA hospital admissions 1992-2002

- Restrict sample to those admitted through emergency department
  - e.g., Chronic bronchitis, heart attack, stroke
  - Why?
TABLE V
Regression Discontinuity Estimates of Changes in Mortality Rates

<table>
<thead>
<tr>
<th>Death rate in</th>
<th>7 days</th>
<th>14 days</th>
<th>28 days</th>
<th>90 days</th>
<th>180 days</th>
<th>365 days</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fully included patients with no additional controls</td>
<td>-1.0</td>
<td>-1.0</td>
<td>-1.1</td>
<td>-1.1</td>
<td>-1.1</td>
<td>-1.0</td>
</tr>
<tr>
<td>Fully included patients with additional controls</td>
<td>-0.0</td>
<td>-0.0</td>
<td>-0.0</td>
<td>-0.0</td>
<td>-0.0</td>
<td>-0.0</td>
</tr>
<tr>
<td>Fully included patients plus additional controls</td>
<td>-0.0</td>
<td>-0.0</td>
<td>-0.0</td>
<td>-0.0</td>
<td>-0.0</td>
<td>-0.0</td>
</tr>
<tr>
<td>Logit regression with separate left- and right-hand side estimates</td>
<td>-0.0</td>
<td>-0.0</td>
<td>-0.0</td>
<td>-0.0</td>
<td>-0.0</td>
<td>-0.0</td>
</tr>
<tr>
<td>Max. of dependent variable (%)</td>
<td>5.0</td>
<td>7.1</td>
<td>6.6</td>
<td>14.7</td>
<td>16.4</td>
<td>15.0</td>
</tr>
</tbody>
</table>

Elder and Lubotsky
The downside of being the youngest in your class

- Suggestive evidence that children “young” for their class perform worse in school
  - Lower test scores/more repeated grades/more disciplinary problems/more ADHD diagnoses
- This has lead to two trends
  - Academic “red shirting”
  - States have moved the “age of entry” earlier
    - 1980, 10% of 5 years olds not in k-garten
    - 2002, this number was 21%

• Suppose all schools start September 1
• Consider the youngest possible kid in the class
• Three state laws — to start k-garten, a kid must turn 5 by: December 1, September 1 or June 1
• In these three states, at school start, the ages of the youngest kids in class are
  - 4 years, 9 months at start (12/1)
  - 5 years (9/1)
  - 5 years, 3 months (6/1)
Evidence to date

- Most of the evidence on the problems of being the youngest in your class is regression-based
- Outcome is regressed on age of child
- Control for other covariates

Consider a regression
\[ y_i = \beta_0 + \text{EA}_i \beta_1 + w_i \beta_2 + \epsilon_i \]
- Is the estimate for \( \beta_1 \) unbiased?
  - Can be biased up for down

Research strategy

- Suppose a state has a September 1 cutoff
- Consider two kids
  - One born August 31
  - One born September 2\textsuperscript{nd}
- One average – do we expect these kids to differ systematically?
- Yet – they will differ when they start school
  - August 31\textsuperscript{st} birth will start at age 5
  - September 2\textsuperscript{nd} birth will start at age 6

- Look on either side of cutoff date
- Should see a large change in age at school entry
- If this impacts outcomes, should see change in test scores at the cutoff as well
- Is the assumption that kids born 3 days apart a good assumption?
Early Childhood Longitudinal Study

- 20 kids from each of 1,000 schools
- Kindergarten class of 1988/89
- Students re-sampled in 1st, 3rd, 5th grade
- Obtain detailed information about the kids/parents/schools/teachers

Structural equation

\[ y_i = \beta_0 + EA_i \beta_1 + w_i \beta_2 + \epsilon_i \]

- EA is entry age

First stage

\[ EA_i = \theta_0 + PEA_i \theta_1 + w_i \theta_2 + \upsilon_i \]

- PEA = predicted entry age – age you would be at the start of kindergarten if you followed the state law to the letter

---

**Figure 2**

Average Predicted and Actual Entrance Ages by Birth Month in States with September 1 Cutoffs, ECLS-K

---

**Figure 3**

Kindergarten Math Scores by Month of Birth, ECLS-K
Test in fall of kindergarten

Test in spring of 5th grade

Panel A: Fall 1998 Test Scores

Panel B: Spring 2004 Test Scores

Panel C: Grade Repetition and Learning Disability Diagnoses

Table 1: Estimates of the Effect of Kindergarten Entrance Age on Reading Test Scores
### Table 8
**The Effect of Kindergarten Birth Year on Grade Retention and Learning Disabilities in the Full NCES/NCES and ECLS-K Samples and by Family Background Quartile**

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Mean</th>
<th>OLS (1)</th>
<th>OLS (2)</th>
<th>IV (3)</th>
<th>IV (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ECLS-K</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diagnosis of learning disability/ADD/ADHD</td>
<td>-0.088</td>
<td>0.008</td>
<td>0.005</td>
<td>-0.026</td>
<td>-0.025</td>
</tr>
<tr>
<td>Diagnosis of ADHD</td>
<td>12.860</td>
<td>(0.386)</td>
<td>(0.009)</td>
<td>(0.011)</td>
<td>(0.017)</td>
</tr>
<tr>
<td>Diagnosis of non-ADD/ADHD learning disability</td>
<td>0.045</td>
<td>0.012</td>
<td>0.044</td>
<td>-0.064</td>
<td>0.004</td>
</tr>
<tr>
<td>Diagnosis of non-ADD/ADHD learning disability</td>
<td>12.860</td>
<td>(0.386)</td>
<td>(0.009)</td>
<td>(0.011)</td>
<td>(0.017)</td>
</tr>
<tr>
<td>In 1st or 2nd grade</td>
<td>0.088</td>
<td>-0.112</td>
<td>-0.112</td>
<td>-0.116</td>
<td>-0.131</td>
</tr>
<tr>
<td>In Spring 2002</td>
<td>10.431</td>
<td>(0.010)</td>
<td>(0.011)</td>
<td>(0.013)</td>
<td>(0.015)</td>
</tr>
</tbody>
</table>

![Graph 1](image1)

![Graph 2](image2)

![Graph 3](image3)

![Graph 4](image4)
Average Birth weight by Birth Month

Month

Birth weight in grams

JAN FEB MAR APR MAY JUN JUL AUG SEP OCT NOV DEC

3280 3290 3300 3310 3320 3330 3340

3280 3290 3300 3310 3320 3330 3340