Quasis-Experimental Methods: Regression Discontinuity Design

Impact Evaluation Workshop
Belgrade, Serbia
Key messages

- An RDD is a valid method for causal inference
- It requires a program assignment rule based on a cutoff score
- The estimators are unbiased around the cutoff score (Local average treatment effect)
- The assumptions are weak
- The score must not be manipulable: check your 3 graphs!
Why do we need another IE method?

- Simple OLS suffers from omitted variable bias
- RCTs are the golden standard, but they are costly
- Many programs rely on institutional rules for targeting, creating an opportunity for a RDD
What is an RDD?

• RDD are a valid way to estimate causal effects of programs

• The intuition is relatively simple:
  – Take advantage of an existing score that determines eligibility to a program
    • Poverty index
    • Exam test score
    • Minimum age
    • Weight at birth
    • Etc...
  – All those over (or below) a certain score are offered the program

• Presume those a little bit over the cutoff and those a little bit under the cutoff score are identical
RDD Example 1

- Scholarships are strongly correlated with outcomes of interest like academic performance.
- Scholarships are often offered based on strict rules based on academic scores.
- The difference between those that barely make it and those that barely miss it is equivalent to random noise the closer we get to the cutoff score.
- The difference of outcomes later in life between those that barely make it and those that barely miss it can be causally attributed to the scholarship program.
We want to improve school attendance for poor communities with a program of block grants to support local services

- Communities with a poverty index \(< 50\) are poor,
- Communities with a poverty index \(\geq 50\) are not poor

→ Communities just above and just below are good comparisons
RDD Example 2: first graph

- The score perfectly assigns treatment status
RDD Example 2: visualize the effect

- The outcome at baseline is continuous around the cutoff, but changes after the program.
RDD Example 2

- Note: the regressions on each side of the cutoff may not be parallel or even linear:

It is also common to restrict the observations to those close to the cutoff
RDD What if we have incomplete compliance?

- Sometimes the score does not perfectly treatment status; there can be never-takers and always-compliers

For fuzzy cutoffs: use the assignment variable as an instrument
RDD Continuity assumption

- The central assumption behind RDD is that the assignment is random very close on each side of the cutoff.

- This can be violated if:
  - Beneficiaries can manipulate the score variable (e.g. income)
  - Evaluators can manipulate the score (be helpful, give a chance)

- You can check if there are signs of violation with graphs
RDD What if we have score manipulation?

- Clumps are a sign that those above and below the cutoff are not the same
RDD What if we have score manipulation?
RDD What if we have score manipulation?
RDD What if we have score manipulation?
RDD What if we have score manipulation?

![Graph showing poverty index score distribution with a vertical line at 1998.](graph.png)
RDD What if we have score manipulation?
RDD What if we have score manipulation?
RDD What if we have score manipulation?
RDD What if we have score manipulation?
RDD What if we have score manipulation?

- RDD requires a continuous score!

![Graph showing poverty index score distribution with a peak in 2003.]
Concrete example: Cambodia CCT

- Eligibility is based on an index of the likelihood of dropping out of school.

- 2 cutoff points within each school:
  - Applicants with the highest dropout risk offered US $60 per year scholarship
  - Applicants with intermediate dropout risk offered US $45 per year scholarship

- Applicants with low dropout risk were not offered scholarship by the program

Likelihood of dropping out of school

Cutoff 1

Cutoff 2

No Scholarship
Concrete example: Cambodia CCT

No scholarship versus $45

$60 versus $45 scholarship

![Graphs showing relative ranking and probability for recipients and non-recipients.]

Estimate of impact

- Recipients
- Non-recipients

SIEF Strategic Impact Evaluation Fund

The World Bank
Cambodia CCT, in the author's words:

“In this paper, we use data from Cambodia, where a program known as the CESSP Scholarship Program (CSP) made cash transfers of different magnitudes to observationally very similar households, based on an index of the likelihood of dropping out of school. The program design lends itself to an identification strategy based on regression discontinuity. The intuition is that we compare the school enrollment of children `just’ above and just below the cutoff for receiving a `large’ versus a small transfer—US $60 versus US$45 per year, conditional on school enrollment—and then compare those just above and just below the cutoff for receiving a `small’ transfer—US $45 versus no scholarship.”
Other example: Serbia Subsidies for the self-employed

- Run by local NES offices
- Approx. $2,400 subsidy for a self-employment plan
- Applicants submit business plan to NES branch office
- NES meets min once per month to evaluate applications
- Highest number of points receives subsidy
- NES branch can exercise discretion in deciding cut-off
Other example: Serbia Subsidies for the self-employed

<table>
<thead>
<tr>
<th>Business Plan</th>
<th>85</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sector</td>
<td></td>
</tr>
<tr>
<td>Manufacturing</td>
<td>10</td>
</tr>
<tr>
<td>Services, trade, construction</td>
<td>10</td>
</tr>
<tr>
<td>Hotel</td>
<td>5</td>
</tr>
<tr>
<td>Analysis of customer product</td>
<td>5</td>
</tr>
<tr>
<td>Analysis of competition</td>
<td>5</td>
</tr>
<tr>
<td>Suppliers</td>
<td>5</td>
</tr>
<tr>
<td>Promotional activities</td>
<td>5</td>
</tr>
<tr>
<td>Office</td>
<td>10</td>
</tr>
<tr>
<td>Equipment needed to conduct business</td>
<td>10</td>
</tr>
<tr>
<td>Structure of employment</td>
<td>10</td>
</tr>
<tr>
<td>Investment needed</td>
<td>10</td>
</tr>
<tr>
<td>Sources of funding</td>
<td>10</td>
</tr>
<tr>
<td>Evaluation of Investment Efficiency</td>
<td>5</td>
</tr>
</tbody>
</table>
Other example: Serbia Subsidies for the self-employed

- Distribution of the assignment rule
Other example: Serbia Subsidies for the self-employed

- Imperfect compliance: fuzzy cutoff. We can estimate the effect of offering the program AND the effect of accepting the program.
Other example: Serbia Subsidies for the self-employed

- Outcome: jump at the cutoff!
Other example: Serbia Subsidies for the self-employed

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Employed(=1)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Approved(=1)</td>
<td>0.226</td>
<td>0.214</td>
<td>0.215</td>
<td>0.214</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.051)**</td>
<td>(0.050)**</td>
<td>(0.050)**</td>
<td>(0.035)**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Entered program(=1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.240</td>
<td>0.195</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.055)**</td>
<td>(0.119)*</td>
</tr>
<tr>
<td>Score</td>
<td>-0.008</td>
<td>-0.007</td>
<td>0.011</td>
<td>-0.003</td>
<td>-0.005</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.008)</td>
<td>(0.058)</td>
<td>(0.002)</td>
<td>(0.007)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>Score squared</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.000</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.000)</td>
<td></td>
</tr>
<tr>
<td>Background chars</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Decision meeting FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Estimation method</td>
<td>OLS</td>
<td>OLS</td>
<td>OLS</td>
<td>OLS</td>
<td>IV</td>
<td>IV</td>
</tr>
<tr>
<td>Sample window</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>15</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Treatment arms</td>
<td>All</td>
<td>All</td>
<td>All</td>
<td>All</td>
<td>All</td>
<td>Acc/Rej</td>
</tr>
<tr>
<td>Observations</td>
<td>846</td>
<td>846</td>
<td>846</td>
<td>1835</td>
<td>846</td>
<td>693</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.38</td>
<td>0.42</td>
<td>0.42</td>
<td>0.33</td>
<td>0.44</td>
<td>0.11</td>
</tr>
</tbody>
</table>

Standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Instruments (5): (a) Approved
Instruments (6): (a) No. applicants in DM, (b) No. of app. Squared, (c) Avg score in DM
Wrapping up: For a discontinuity design you need…

1. Continuous eligibility score, perhaps in an existing database

2. Clearly defined eligibility cut-off rule
   Observations with a score ≤ cutoff are eligible
   Observations with a score > cutoff are not-eligible
   (Or vice-versa)
Wrapping up: Potential pitfalls

1. The score may not be continuous

2. External validity: the effect is unbiased close to the cutoff, but may not be valid for the whole distribution

3. The results may be sensitive to the functional form of the regressions

4. Requires a lot of observations, especially in the bandwidth close to the cutoff
Questions?