Today’s lecture

The following lecture is based on the following readings:

▶ Duflo et al. article on Randomization

Supply-side interventions in education

- Last class, we focused on some experiments that sought to increase the demand for schooling among the poor.
Supply-side interventions in education

- Last class, we focused on some experiments that sought to increase the demand for schooling among the poor

- Children of poor households are credit constrained → they want to consume more school if they had the money
Supply-side interventions in education

- Last class, we focused on some experiments that sought to increase the demand for schooling among the poor.

- Children of poor households are credit constrained → they want to consume more school if they had the money.

- Progresa (and other CCT’s throughout the world) relaxes this constrain.
Supply-side interventions in education

- Last class, we focused on some experiments that sought to increase the demand for schooling among the poor.

- Children of poor households are credit constrained → they want to consume more school if they had the money.

- Progresa (and other CCT’s throughout the world) relaxes this constrain.

- We saw that such programs have been useful in increasing the demand for school.
Increasing demand is only part of the story.
Supply-side interventions in education

- Increasing demand is only part of the story.

- Another way of increasing schooling among the poor is to focus on the cost of schooling through the supply side.
Supply-side interventions in education

- Increasing demand is only part of the story.

- Another way of increasing schooling among the poor is to focus on the cost of schooling through the supply side
  - School construction
Supply-side interventions in education

- Increasing demand is only part of the story.

- Another way of increasing schooling among the poor is to focus on the cost of schooling through the supply side
  - School construction
  - Improve transportation
Supply-side interventions in education

- Increasing demand is only part of the story.

- Another way of increasing schooling among the poor is to focus on the cost of schooling through the supply side
  - School construction
  - Improve transportation
  - Eliminate school fees (text books, uniforms, etc)
Supply-side interventions in education

- Increasing demand is only part of the story.

- Another way of increasing schooling among the poor is to focus on the cost of schooling through the supply side
  - School construction
  - Improve transportation
  - Eliminate school fees (text books, uniforms, etc)
  - Reduce child labor (opportunity costs)
Do we have any evidence that supply-side interventions work?

▶ This is a hard question to answer
Do we have any evidence that supply-side interventions work?

▶ This is a hard question to answer

▶ For example, suppose we wanted to know whether improving access to schools (say by constructing more schools) will increase school enrollment?
Do we have any evidence that supply-side interventions work?

▶ This is a hard question to answer

▶ For example, suppose we wanted to know whether improving access to schools (say by constructing more schools) will increase school enrollment?

▶ How would we design such an intervention and evaluation?
Do we have any evidence that supply-side interventions work?

- This is a hard question to answer.

- For example, suppose we wanted to know whether improving access to schools (say by constructing more schools) will increase school enrollment?

- How would we design such an intervention and evaluation?

- The ideal design would be to randomize, of course! But how?
Unit of analysis

First, we need to define the unit of analysis

- Do we randomize at the village level or the district level?
Unit of analysis

First, we need to define the unit of analysis

- Do we randomize at the village level or the district level?
- What are some of the tradeoffs?
Unit of analysis

First, we need to define the unit of analysis

▶ Do we randomize at the village level or the district level?

▶ What are some of the tradeoffs?
  ▶ Sample size - the smaller the sample the less statistical power we will have to detect the impact
Unit of analysis

First, we need to define the unit of analysis

- Do we randomize at the village level or the district level?

- What are some of the tradeoffs?
  - Sample size - the smaller the sample the less statistical power we will have to detect the impact
  - Spillover - children in non-treated villages may still benefit from a new school if it is constructed in a nearby village → makes it harder to detect the effect
Assignment of the intervention

Suppose we decide that the unit of analysis is village, how do we determine which villages gets a school?

- random assignment
Assignment of the intervention

Suppose we decide that the unit of analysis is village, how do we determine which villages get a school?

- random assignment

- is randomization politically feasible?
Assignment of the intervention

Suppose we decide that the unit of analysis is the village, how do we determine which villages get a school?

- random assignment

- is randomization politically feasible?

- is randomization ethical?
Measuring the impact

Suppose that we are able to convince a government official to randomize across villages, how do we measure the impact? What kind of data do we collect?

▶ A baseline? Why would this be useful?
Measuring the impact

Suppose that we are able to convince a government official to randomize across villages, how do we measure the impact? What kind of data do we collect?

▶ A baseline? Why would this be useful?

▶ How long before the follow-up? How many follow ups?
Measuring the impact

Suppose that we are able to convince a government official to randomize across villages, how do we measure the impact? What kind of data do we collect?

▶ A baseline? Why would this be useful?

▶ How long before the follow-up? How many follow ups?

▶ Do we collect information at the households or the schools?
Measuring the impact

Suppose that we are able to convince a government official to randomize across villages, how do we measure the impact? What kind of data do we collect?

▶ A baseline? Why would this be useful?

▶ How long before the follow-up? How many follow ups?

▶ Do we collect information at the households or the schools?

▶ A well-designed randomized evaluation will take time
Difference-in-differences

- To conduct a randomized evaluation on these “bigger” questions is difficult, and often infeasible.
To conduct a randomized evaluation on these “bigger” questions is difficult, and often infeasible.

But these are important questions, we need to think of alternative ways in which to estimate the impact.
Difference-in-differences

- To conduct a randomized evaluation on these “bigger” questions is difficult, and often infeasible

- But these are important questions, we need to think of alternative ways in which to estimate the impact

- One alternative approach is called: difference-in-differences
Difference-in-differences

▶ To conduct a randomized evaluation on these “bigger” questions is difficult, and often infeasible

▶ But these are important questions, we need to think of alternative ways in which to estimate the impact

▶ One alternative approach is called: difference-in-differences
  ▶ Useful when we have a treatment and a control group, with data before and after the intervention
Difference-in-differences

Enrollment

2007 2008

Treatment Group
Difference-in-differences

Treatment Group

Control Group

Simple difference - Estimate of Impact

2007  2008

time
Difference-in-differences

Enrollment

Difference prior to intervention

2007

2008
Difference-in-differences

How do we control for this initial difference (selection bias)?

- Look at the change over time in the control group
Difference-in-differences

How do we control for this initial difference (selection bias)?

- Look at the change over time in the control group

- Assume that the same change over time would have happened in the treatment group
Difference-in-differences

How do we control for this initial difference (selection bias)?

- Look at the change over time in the control group
- Assume that the same change over time would have happened in the treatment group
- Adjust the difference post intervention by difference prior to the intervention (hence the name difference-in-differences)
Difference-in-differences

Enrollment

2007  2008

time

Diff-in-Diff Estimate of Impact
Difference-in-differences

Mathematically,

- \( Y_1^T \) potential outcome if treated in period 1 (after treatment occurs)
- \( Y_1^C \) potential outcome if untreated in period 1
- \( Y_0^T \) potential outcome if treated in period 0 (before treatment occurs)
- \( Y_0^C \) potential outcome if untreated in period 0

We are interested in an estimate of the average treatment effect (ATE): Which is?
Difference-in-differences

Mathematically,

- $Y_1^T$ potential outcome if treated in period 1 (after treatment occurs)
- $Y_1^C$ potential outcome if untreated in period 1
- $Y_0^T$ potential outcome if treated in period 0 (before treatment occurs)
- $Y_0^C$ potential outcome if untreated in period 0

We are interested in an estimate of the average treatment effect (ATE): Which is?

$$ATE = E[Y_1^T|T] - E[Y_1^C|T]$$

$T$ indicates group assignment. Recall we do not observe this difference?
Difference-in-differences

Average treatment effect

\[ ATE = E[Y_1^T | T] - E[Y_1^C | T] \]
Average treatment effect

\[ ATE = E[Y_1^T|T] - E[Y_1^C|T] \]

Let's add zero

Difference-in-differences

Let’s add zero again

Difference-in-differences

Let’s add zero again

\[
\]

One more time and rearrange

\[
ATE = \{E[Y_T^1|T] - E[Y_C^1|C]\} - \{E[Y_C^0|T] - E[Y_C^0|C]\} + E[Y_C^1|C] - E[Y_C^0|C] + E[Y_C^0|T] - E[Y_C^1|T]
\]
Difference-in-differences

What does this equation mean?

\[
ATE = \left\{ E[Y_1^T|T] - E[Y_1^C|C] \right\} - \left\{ E[Y_0^C|T] - E[Y_0^C|C] \right\} \\
\text{Difference-in-differences} \\
+ \quad E[Y_1^C|C] - E[Y_0^C|C] + E[Y_0^C|T] - E[Y_1^C|T]
\]
Difference-in-differences

What does this equation mean?

\[
ATE = \underbrace{\{E[Y_1^T|T] - E[Y_1^C|C]\} - \{E[Y_0^C|T] - E[Y_0^C|C]\}}_{\text{Difference-in-differences}} \\
+ \ E[Y_1^C|C] - E[Y_0^C|C] + E[Y_0^C|T] - E[Y_1^C|T]
\]

So for a causal interpretation we need the second quantity to equal zero, or

\[
E[Y_1^C|C] - E[Y_0^C|C] = E[Y_1^C|T] - E[Y_0^C|T]
\]

What does this mean?
Difference-in-differences

Counterfactual assumption of the difference-in-differences estimate:

- “Parallel trends” - change in outcome in the control group would be the same as the treatment group, if the treatment never took place
Difference-in-differences

Counterfactual assumption of the difference-in-differences estimate:

▶ “Parallel trends” - change in outcome in the control group would be the same as the treatment group, if the treatment never took place

▶ The method fails if the comparison group is on a different trajectory
Difference-in-differences

Counterfactual assumption of the difference-in-differences estimate:

- “Parallel trends” - change in outcome in the control group would be the same as the treatment group, if the treatment never took place

- The method fails if the comparison group is on a different trajectory
  - Suppose the government targeted the schools at villages with the poor school enrollment
Difference-in-differences

Counterfactual assumption of the difference-in-differences estimate:

- “Parallel trends” - change in outcome in the control group would be the same as the treatment group, if the treatment never took place

- The method fails if the comparison group is on a different trajectory
  - Suppose the government targeted the schools at villages with the poor school enrollment
  - Places with lowest enrollment rates will have a tendency to grow faster than high enrollment rate places (reversion to the mean)
Difference-in-differences

Enrollment

2007  2008

time
Another issue with D-D approach

Functional form

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>1.3</td>
<td>2.1</td>
</tr>
<tr>
<td>C</td>
<td>1.1</td>
<td>1.8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>ln(1.3)</th>
<th>ln(2.1)</th>
<th>0.479</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>ln(1.1)</td>
<td>ln(1.8)</td>
<td>0.492</td>
</tr>
<tr>
<td>C</td>
<td></td>
<td></td>
<td>-0.013</td>
</tr>
</tbody>
</table>
Difference-in-differences: Regression

How would we estimate the difference-in-differences in a regression?

\[ Y_{it} = \beta_0 + \beta_1 Post_t + \beta_2 Program_i + \beta_3 Post \times Program_{it} + \epsilon_{it} \]

- Post\(_t\) - indicator if observation is in post period
- Program\(_i\) - indicator if observation is in treatment group
- Post \(\times\) Program\(_{it}\) - interaction (multiplication) of the post indicator with the program indicator, i.e. indicator if observation is in the treatment group and in the post period

Claim: \(\beta_3\) is the difference-in-differences estimate. Why?
Suppose there are two time periods $t = 0, 1$ and the intervention happens in $t = 1$.

$$Y_{it} = \beta_0 + \beta_1 \text{Post}_t + \beta_2 \text{Program}_i + \beta_3 \text{Post} \times \text{Program}_{it} + \epsilon_{it}$$

- $E[Y_{i1}|\text{Program} = 1] = ?$
- $E[Y_{i0}|\text{Program} = 1] = ?$
- $E[Y_{i1}|\text{Program} = 0] = ?$
- $E[Y_{i0}|\text{Program} = 0] = ?$
Difference-in-differences: Regression

\[ Y_{it} = \beta_0 + \beta_1 \text{Post}_t + \beta_2 \text{Program}_i + \beta_3 \text{Post} \times \text{Program}_{it} + \epsilon_{it} \]

- \( E[ Y_{i1} | \text{Program} = 1 ] = \beta_0 + \beta_1 + \beta_2 + \beta_3 \)
- \( E[ Y_{i0} | \text{Program} = 1 ] = \beta_0 + \beta_2 \)
- \( E[ Y_{i1} | \text{Program} = 0 ] = \beta_0 + \beta_1 \)
- \( E[ Y_{i0} | \text{Program} = 0 ] = \beta_0 \)
Difference-in-differences: Regression

\[ Y_{it} = \beta_0 + \beta_1 Post_t + \beta_2 Program_i + \beta_3 Post \times Program_{it} + \epsilon_{it} \]

- \( E[Y_{i1}|\text{Program} = 1] = \beta_0 + \beta_1 + \beta_2 + \beta_3 \)
- \( E[Y_{i0}|\text{Program} = 1] = \beta_0 + \beta_2 \)
- \( E[Y_{i1}|\text{Program} = 0] = \beta_0 + \beta_1 \)
- \( E[Y_{i0}|\text{Program} = 0] = \beta_0 \)

Difference-in-differences estimator (DID):

\[ DID = E[Y_{i1}|\text{Program} = 1] - E[Y_{i0}|\text{Program} = 1] - [E[Y_{i1}|\text{Program} = 0] - E[Y_{i0}|\text{Program} = 0]] = \beta_3 \]
Figure 1: Program Evaluation Design