SIEF Kenya Impact Evaluation Workshop

Difference-in-Difference Estimation

May 6, 2015

Instructor: Pamela Jakiela
University of Maryland, College Park, USA
Overview

- Review: false counterfactuals
- Difference-in-differences: the intuition
- Difference-in-differences: the Stata code
- Checking the common trends assumption
- A practical example
Motivation: False Counterfactuals
What Is an Impact Evaluation?

“An impact evaluation assesses changes in the well-being of individuals, households, communities or firms that can be attributed to a particular project, program or policy. The central impact evaluation question is what would have happened to those receiving the intervention if they had not in fact received the program. Since we cannot observe this group both with and without the intervention, the key challenge is to develop a counterfactual — that is, a group which is as similar as possible (in observable and unobservable dimensions) to those receiving the intervention. This comparison allows for the establishment of definitive causality — attributing observed changes in welfare to the program, while removing confounding factors.”
What Is an Impact Evaluation?

Goal: measure **causal** impacts of policy on participants

- We did $A$; as a result, $B$ happened
What Is an Impact Evaluation?

Goal: measure **causal** impacts of policy on participants

- We did $A$; as a result, $B$ happened
- $A$ is a policy or intervention
- $B$ is an outcome of interest (what we hope to impact)

Examples:
- We gave out insecticide-treated bednets, and fewer children under the age of 5 got sick with or died from malaria as a result
- We distributed free lunches in elementary schools, and school attendance and/or academic performance went up as a result
What Is an Impact Evaluation?

Goal: measure causal impacts of policy on participants

- We did $A$; as a result, $B$ happened
- $A$ is a policy or intervention
- $B$ is an outcome of interest (what we hope to impact)
- Examples:
  - We gave out insecticide-treated bednets, and fewer children under the age of 5 got sick with or died from malaria as a result
  - We distributed free lunches in elementary schools, and school attendance and/or academic performance went up as a result
Establishing Causality

Goal: measure *causal* impacts of policy on participants

- We want to be able to say $B$ happened because of $A$
  - We need to rule out other possible causes of $B$
- If we can say this, then we can also say: if we did $A$ again (in another place), we think that $B$ would happen there as well
Establishing Causality

Goal: measure **causal** impacts of policy on participants

- We want to be able to say $B$ happened because of $A$
  - We need to rule out other possible causes of $B$
- If we can say this, then we can also say: if we did $A$ again (in another place), we think that $B$ would happen there as well

In an ideal world (research-wise), we could clone each program participant and observe the impacts of our program on their lives
Establishing Causality

In an ideal world (research-wise), we could clone each program participant and observe the impacts of our program on their lives.

What is the impact of giving Lisa a book on her test score?

- Impact = Lisa’s score with a book - Lisa’s score without a book
Establishing Causality

In an ideal world (research-wise), we could clone each program participant and observe the impacts of our program on their lives.

What is the impact of giving Lisa a book on her test score?

- Impact = Lisa’s score with a book - Lisa’s score without a book

In the real world, we either observe Lisa with a book or without.

- We never observe the counterfactual.
Establishing Causality

To measure the causal impact of giving Lisa a book on her test score, we need to find a comparison group that did not receive a book.

Our estimate of the impact of the book is then the difference in test scores between the **treatment group** and the **comparison group**:

- \( \text{Impact} = \text{Lisa’s score with a book} - \text{Bart’s score without a book} \)
Establishing Causality

To measure the causal impact of giving Lisa a book on her test score, we need to find a comparison group that did not receive a book.

Our estimate of the impact of the book is then the difference in test scores between the treatment group and the comparison group:

- Impact = Lisa’s score with a book - Bart’s score without a book

As this example illustrates, finding a good comparison group is hard.
The Potential Outcomes Framework

Two potential outcomes for each individual, community, etc:

Potential outcome = \[ \begin{cases} Y_{0i} & P_i = 0 \\ Y_{1i} & P_i = 1 \end{cases} \]
The Potential Outcomes Framework

Two potential outcomes for each individual, community, etc:

Potential outcome = \begin{cases} Y_{0i} & P_i = 0 \\ Y_{1i} & P_i = 1 \end{cases}

The problem: we only observe one of $Y_{1i}$ and $Y_{0i}$

- Each individual either participates in the program or not
- The causal impact of program ($P$) on $i$ is: $Y_{1i} - Y_{0i}$
The Potential Outcomes Framework

Two potential outcomes for each individual, community, etc:

\[
\text{Potential outcome} = \begin{cases} 
Y_{0i} & P_i = 0 \\
Y_{1i} & P_i = 1 
\end{cases}
\]

The problem: we only observe one of \( Y_{1i} \) and \( Y_{0i} \)

- Each individual either participates in the program or not
- The causal impact of program (\( P \)) on \( i \) is: \( Y_{1i} - Y_{0i} \)

We observe \( i \)'s actual outcome:

\[
Y_i = Y_{0i} + (Y_{1i} - Y_{0i}) P_i
\]

impact
Defining the Counterfactual

To estimate the impact of a program, we need to know what would have happened to every participant $i$ in the absence the program.

- We call this the counterfactual.
Defining the Counterfactual

To estimate the impact of a program, we need to know what would have happened to every participant $i$ in the absence the program

- We call this the **counterfactual**

Of course, we can’t actually clone our participants and see what happens to the clones if they don’t participate in the program

- Instead, we estimate the counterfactual using a **comparison group**
Defining the Counterfactual

To estimate the impact of a program, we need to know what would have happened to every participant $i$ in the absence the program

- We call this the **counterfactual**

Of course, we can’t actually clone our participants and see what happens to the clones if they don’t participate in the program

- Instead, we estimate the counterfactual using a **comparison group**

The comparison group needs to:

- Look identical to the treatment group prior to the program
- Not be impacted by the program in anyway
Defining the Counterfactual

To estimate the impact of a program, we need to know what would have happened to every participant $i$ in the absence of the program.

- We call this the **counterfactual**

Of course, we can’t actually clone our participants and see what happens to the clones if they don’t participate in the program.

- Instead, we estimate the counterfactual using a **comparison group**

The comparison group needs to:

- Look identical to the treatment group prior to the program
- Not be impacted by the program in anyway

**YOU CANNOT HAVE A GOOD IMPACT EVALUATION WITHOUT A CREDIBLE, CONVINCING COMPARISON GROUP**
The Moving Parts of an Impact Evaluation

A policy or program of interest (aka the “treatment”)

- \( P_i = 1 \) if individual/community \( i \) participated in the program
- \( P_i = 0 \) otherwise

The treatment group: a group of people for whom \( P_i = 1 \)

The comparison group: a group of people for whom \( P_i = 0 \)

The outcome of interest: the dependent variable in our analysis

- Something that we care about
- Something that we expect to be impacted by the treatment

An impact evaluation compares values of the outcome of interest in the treatment group to values in the comparison group

- We attribute the difference to the impact of treatment
The Moving Parts of an Impact Evaluation

A policy or program of interest (aka the “treatment”)

- \( P_i = 1 \) if individual/community \( i \) participated in the program
- \( P_i = 0 \) otherwise

- The **treatment group**: a group of people for whom \( P_i = 1 \)
- The **comparison group**: a group of people for whom \( P_i = 0 \)
The Moving Parts of an Impact Evaluation

A policy or program of interest (aka the “treatment”)

- $P_i = 1$ if individual/community $i$ participated in the program
- $P_i = 0$ otherwise

- The **treatment group**: a group of people for whom $P_i = 1$
- The **comparison group**: a group of people for whom $P_i = 0$

The **outcome of interest**: the dependent variable in our analysis

- Something that we care about
- Something that we expect to be impacted by the treatment
The Moving Parts of an Impact Evaluation

A policy or program of interest (aka the “treatment”)

- \( P_i = 1 \) if individual/community \( i \) participated in the program
- \( P_i = 0 \) otherwise

- The **treatment group**: a group of people for whom \( P_i = 1 \)
- The **comparison group**: a group of people for whom \( P_i = 0 \)

The **outcome of interest**: the dependent variable in our analysis

- Something that we care about
- Something that we expect to be impacted by the treatment

An impact evaluation compares values of the outcome of interest in the treatment group to values in the comparison group

- We attribute the difference to the impact of treatment
False Counterfactuals

Two types of false counterfactuals:

- Before vs. After Comparisons
- Participant vs. Non-Participant Comparisons
False Counterfactuals

Two types of **false counterfactuals**:

- Before vs. After Comparisons
- Participant vs. Non-Participant Comparisons

Consider these false counterfactuals in context of a simple example:

- Problem: poor academic performance
- Program: extra training for teachers, materials for classrooms
- Outcome: student test scores
- Strategy: baseline and endline (before and after) data collection
Before vs. After Comparisons

Impact of the program: $B - A$?
Before vs. After Comparisons

Impact of the program: \( B - A \)?

Before vs. after analysis assumes test scores would not have changed between \( t = 0 \) and \( t = 1 \) in the absence of the program.
Before vs. After Comparisons

Impact of the program: $B - A$?
Before vs. After Comparisons

Impact of the program: $B - A$?

Before vs. after analysis assumes test scores would not have changed between $t = 0$ and $t = 1$ in the absence of the program.
Before vs. After Comparisons

What if the parents’ income, or students’ overall level of learning, or the teacher, or the weather, or some other thing(s) changed?
Before vs. After Comparisons

The perils of pre vs. post analysis should be obvious...

...as we all recall from reading the famous paper: “Does graduating from college cause women get pregnant? A pre-vs-post analysis of the impacts of education on fertility”
Before vs. After Comparisons

The perils of pre vs. post analysis should be obvious...

...as we all recall from reading the famous paper: “Does graduating from college cause women get pregnant? A pre-vs-post analysis of the impacts of education on fertility”

A slightly more subtle example of the perils of pre vs. post analysis comes from the mid-term report evaluating the Millennium Villages

- The report highlights the fourfold increase in mobile phone ownership between 2005 and 2008 among households in Bar Sauri
### Before vs. After Comparisons

**AVERAGE PROGRESS ON OTHER KEY MDG INDICATORS**

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Baseline</th>
<th>Year Three</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chronic malnutrition (stunting among children under two)</td>
<td>35%</td>
<td>54%</td>
</tr>
<tr>
<td>Gross attendance ratio in primary education</td>
<td>115%</td>
<td>121%</td>
</tr>
<tr>
<td>Births delivered by skilled health personnel</td>
<td>33%</td>
<td>47%</td>
</tr>
<tr>
<td>HIV testing in last year (15–49 year olds)</td>
<td>11%</td>
<td>29%</td>
</tr>
<tr>
<td>Access to improved drinking water (households)</td>
<td>26%</td>
<td>72%</td>
</tr>
<tr>
<td>Access to improved sanitation (households)</td>
<td>6%</td>
<td>41%</td>
</tr>
<tr>
<td>Mobile phone ownership (households)</td>
<td>5%</td>
<td>31%</td>
</tr>
</tbody>
</table>
Clemens and Demombynes (2010) compare changes in mobile phone ownership in Bar Sauri (rectangles) to trends in Kenya (red), rural Kenya (green), and rural areas in Nyanza Province (blue).

- The problem is obvious: before vs. after analysis assumes that there is no time trend in mobile phone ownership in Kenya.
Participants vs. Non-Participants

What if we compare (post-intervention) test scores in program schools to test scores in nearby schools that did not participate in the program?

Can we estimate the impact of the program by calculating $T - C$?
Participants vs. Non-Participants

\[ E \left[ Y_i \mid P_i = Z \right] \] denotes the population (or large sample) average of the outcome variable \( Y \) (test scores) in schools with \( P_i = 0 \) or \( P_i = 1 \)

- \( E[Y_i] \) = average test score in school \( i \)
- \( P_i = 1 \) program school, \( P_i = 0 \) in nearby (comparison) school
- Average outcome in program schools: \( E \left[ Y_i \mid P_i = 1 \right] = Y \)
- Average outcome in neighboring schools: \( E \left[ Y_i \mid P_i = 0 \right] = Z \)
Participants vs. Non-Participants

\( E [Y_i|P_i = Z] \) denotes the population (or large sample) average of the outcome variable \( Y \) (test scores) in schools with \( P_i = 0 \) or \( P_i = 1 \)

- \( E[Y_i] = \) average test score in school \( i \)
- \( P_i = 1 \) program school, \( P_i = 0 \) in nearby (comparison) school
- Average outcome in program schools: \( E [Y_i|P_i = 1] = Y \)
- Average outcome in neighboring schools: \( E [Y_i|P_i = 0] = Z \)

Our estimate of the impact of the program \( (P) \) is:

\[
\text{Impact} = E [Y_i|P_i = 1] - E [Y_i|P_i = 0]
\]
Participants vs. Non-Participants

\[ E[Y_i|P_i = Z] \] denotes the population (or large sample) average of the outcome variable \( Y \) (test scores) in schools with \( P_i = 0 \) or \( P_i = 1 \)

- \( E[Y_i] \) = average test score in school \( i \)
- \( P_i = 1 \) program school, \( P_i = 0 \) in nearby (comparison) school
- Average outcome in program schools: \( E[Y_i|P_i = 1] = Y \)
- Average outcome in neighboring schools: \( E[Y_i|P_i = 0] = Z \)

Our estimate of the impact of the program \((P)\) is:

\[
\text{Impact} = E[Y_i|P_i = 1] - E[Y_i|P_i = 0]
\]

In a regression framework: \( E[Y_i] = \alpha + \beta \cdot P_i \)

- When we regress \( Y \) on an indicator, \( P \): \( \hat{\beta} = \bar{Y}_{P_i=1} - \bar{Y}_{P_i=0} \)
Wait!! Why weren’t the neighboring schools included in the program?

- Maybe they had low quality head teachers (who didn’t bother to fill out the paperwork to enroll in the program)
- Maybe they already had high test scores
- Those who aren’t eligible and those who choose not to participate may have different outcomes in the absence of the program
- This is **selection bias**
Participants vs. Non-Participants

Wait!! Why weren’t the neighboring schools included in the program?

- Maybe they had low quality head teachers (who didn’t bother to fill out the paperwork to enroll in the program)
- Maybe they already had high test scores
- Those who aren’t eligible and those who choose not to participate may have different outcomes in the absence of the program
- This is selection bias

Remember: the causal impact of program on $i$ is: $Y_{1i} - Y_{0i}$

- Assuming that outcomes in program schools in the absence of the program would look like outcomes observed in the comparison schools
Participants vs. Non-Participants

The slide illustrates the comparison between observed and unobserved counterfactual outcomes for participants and non-participants. The graphs show the expected values under different conditions for comparison schools and program schools.
Our estimate of the impact of a training program ($P$) is:

$$\text{Impact} = E[Y_i|P_i = 1] - E[Y_i|P_i = 0]$$

$$= E[Y_1i|P_i = 1] - E[Y_0i|P_i = 1] + E[Y_0i|P_i = 1] - E[Y_0i|P_i = 0]$$

- program impact
- selection bias
Our estimate of the impact of a training program \((P)\) is:

\[
\text{Impact} = E[Y_i|P_i = 1] - E[Y_i|P_i = 0]
\]

\[
= E[Y_{1i}|P_i = 1] - E[Y_{0i}|P_i = 1] + E[Y_{0i}|P_i = 1] - E[Y_{0i}|P_i = 0]
\]

program impact \hspace{2cm} selection bias

When \(E[Y_{0i}|P_i = 1] - E[Y_{0i}|P_i = 0] \neq 0\), we have a problem.

- The treatment and comparison groups would not have looked the same in the absence of the program. Why might this occur?
Summary: False Counterfactuals

Before vs. After Comparisons:

- **Compares:** same individuals/communities before and after program
- **Drawback:** things (besides the program) may happen over time

Participant vs. Non-Participant Comparisons:

- **Compares:** participants to those not in the program
- **Drawback:** selection bias — why aren’t they in the program?
Difference-in-Difference Estimation: Intuition
Difference-in-Difference Estimation

Difference-in-difference (or “diff-in-diff” or “DD”) impact evaluations combine the pre vs. post and enrolled vs. not enrolled approaches

- This can **sometimes** overcome the twin problems of [1] selection bias and [2] time trends in the outcome of interest

- The basic idea is to observe the treatment group and a comparison group (for example, the not enrolled) before and after the program

The diff-in-diff estimator is:

$$DD = \bar{Y}_{treatment\ post} - \bar{Y}_{treatment\ pre} - (\bar{Y}_{comparison\ post} - \bar{Y}_{comparison\ pre})$$
Difference-in-difference (or “diff-in-diff” or “DD”) impact evaluations combine the pre vs. post and enrolled vs. not enrolled approaches

- This can sometimes overcome the twin problems of [1] selection bias and [2] time trends in the outcome of interest
- The basic idea is to observe the treatment group and a comparison group (for example, the not enrolled) before and after the program

The diff-in-diff estimator is:

\[
DD = \bar{Y}_{treatment\ post} - \bar{Y}_{treatment\ pre} - (\bar{Y}_{comparison\ post} - \bar{Y}_{comparison\ pre})
\]
**Difference-in-Difference Estimation**

<table>
<thead>
<tr>
<th></th>
<th>Treatment</th>
<th>Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pre-Program</strong></td>
<td>$\bar{Y}_{treatment\ pre}$</td>
<td>$\bar{Y}_{comparison\ pre}$</td>
</tr>
<tr>
<td><strong>Post-Program</strong></td>
<td>$\bar{Y}_{treatment\ post}$</td>
<td>$\bar{Y}_{comparison\ post}$</td>
</tr>
</tbody>
</table>

Intuitively, diff-in-diff estimation is just a comparison of 4 cell-level means.
**Difference-in-Difference Estimation**

<table>
<thead>
<tr>
<th></th>
<th>Treatment</th>
<th>Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pre-Program</strong></td>
<td>$\bar{Y}_{treatment pre}$</td>
<td>$\bar{Y}_{comparison pre}$</td>
</tr>
<tr>
<td><strong>Post-Program</strong></td>
<td>$\bar{Y}_{treatment post}$</td>
<td>$\bar{Y}_{comparison post}$</td>
</tr>
</tbody>
</table>

Only one of the 4 cells is **treated** (has received the program)
### Difference-in-Difference Estimation

#### Table

<table>
<thead>
<tr>
<th></th>
<th>Treatment</th>
<th>Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pre-Program</strong></td>
<td>$\bar{Y}_{\text{treatment pre}}$</td>
<td>$\bar{Y}_{\text{comparison pre}}$</td>
</tr>
<tr>
<td><strong>Post-Program</strong></td>
<td>$\bar{Y}_{\text{treatment post}}$</td>
<td>$\bar{Y}_{\text{comparison post}}$</td>
</tr>
</tbody>
</table>

Comparing treatment vs. comparison pre-program measures selection bias
Only one of the 4 cells is **treated** (has received the program)
The assumption underlying diff-in-diff estimation is that, in the absence of the program, individual $i$’s outcome at time $t$ is given by:

$$E[Y_i|P_i = 0, t] = \gamma_i + \lambda_t$$
Difference-in-Difference Estimation

The assumption underlying diff-in-diff estimation is that, in the absence of the program, individual $i$’s outcome at time $t$ is given by:

$$E[Y_i|P_i = 0, t] = \gamma_i + \lambda_t$$

There are two implicit identifying assumptions here:

- Selection bias relates to fixed characteristics of individuals ($\gamma_i$)
  - The magnitude of the selection bias term isn’t changing over time
- Time trend ($\lambda_t$) same for treatment and control groups
The assumption underlying diff-in-diff estimation is that, in the absence of the program, individual \( i \)’s outcome at time \( t \) is given by:

\[
E[Y_i|P_i = 0, t] = \gamma_i + \lambda_t
\]

There are two implicit identifying assumptions here:

- Selection bias relates to fixed characteristics of individuals \((\gamma_i)\)
  - The magnitude of the selection bias term isn’t changing over time
- Time trend \((\lambda_t)\) same for treatment and control groups

These two necessary conditions for identification in diff-in-diff estimation are often referred to (collectively) as the **common trends** assumption.
Difference-in-Difference Estimation

In the absence of the program, \( i \)'s outcome at time \( t \) is:

\[
E[Y_{0i}|P_i = 0, t] = \gamma_i + \lambda_t
\]
In the absence of the program, \( i \)'s outcome at time \( t \) is:

\[
E[Y_{0i}|P_i = 0, t] = \gamma_i + \lambda_t
\]

Outcomes in the comparison group:

\[
E[\bar{Y}_{comparison}^{pre}] = E[Y_{0i}|P_i = 0, t = 1] = E[\gamma_i|P_i = 0] + \lambda_1
\]

\[
E[\bar{Y}_{comparison}^{post}] = E[Y_{0i}|P_i = 0, t = 2] = E[\gamma_i|P_i = 0] + \lambda_2
\]
In the absence of the program, $i$’s outcome at time $t$ is:

$$E[Y_{0i} | P_i = 0, t] = \gamma_i + \lambda_t$$

Outcomes in the comparison group:

$$E[\bar{Y}_{\text{comparison}}^{\text{pre}}] = E[Y_{0i} | P_i = 0, t = 1] = E[\gamma_i | P_i = 0] + \lambda_1$$

$$E[\bar{Y}_{\text{comparison}}^{\text{post}}] = E[Y_{0i} | P_i = 0, t = 2] = E[\gamma_i | P_i = 0] + \lambda_2$$

Time trend:

$$E[\bar{Y}_{\text{post}}^{\text{comparison}}] - E[\bar{Y}_{\text{pre}}^{\text{comparison}}] = E[\gamma_i | P_i = 0] + \lambda_2 - (E[\gamma_i | P_i = 0] + \lambda_1)$$

$$= \lambda_2 - \lambda_1$$
Difference-in-Difference Estimation

Let $\delta$ denote the true impact of the program:

$$\delta = E[Y_{1i}|P_i = 1, t] - E[Y_{0i}|P_i = 1, t]$$

which does not depend on the time period or $i$’s characteristics.
Difference-in-Difference Estimation

Let $\delta$ denote the true impact of the program:

$$\delta = E[Y_{1i}|P_i = 1, t] - E[Y_{0i}|P_i = 1, t]$$

which does not depend on the time period or $i$’s characteristics
Difference-in-Difference Estimation

Let $\delta$ denote the true impact of the program:

$$\delta = E[Y_{1i}|P_i = 1, t] - E[Y_{0i}|P_i = 1, t]$$

which does not depend on the time period or $i$’s characteristics

Outcomes in the treatment group:

$$E[\hat{Y}_{pre}^{treatment}] = E[Y_{0i}|P_i = 1, t = 1] = E[\gamma_i|P_i = 1] + \lambda_1$$

$$E[\hat{Y}_{post}^{treatment}] = E[Y_{1i}|P_i = 1, t = 2] = E[\gamma_i|P_i = 1] + \delta + \lambda_2$$
Difference-in-Difference Estimation

Let $\delta$ denote the true impact of the program:

$$\delta = E[Y_{1i}|P_i = 1, t] - E[Y_{0i}|P_i = 1, t]$$

which does not depend on the time period or $i$’s characteristics.

Outcomes in the treatment group:

$$E[\bar{Y}_{\text{pre}}^{\text{treatment}}] = E[Y_{0i}|P_i = 1, t = 1] = E[\gamma_i|P_i = 1] + \lambda_1$$

$$E[\bar{Y}_{\text{post}}^{\text{treatment}}] = E[Y_{1i}|P_i = 1, t = 2] = E[\gamma_i|P_i = 1] + \delta + \lambda_2$$

If we were to calculate a pre-vs-post estimator, we’d have:

$$E[\bar{Y}_{\text{post}}^{\text{treatment}}] - E[\bar{Y}_{\text{pre}}^{\text{treatment}}] = E[\gamma_i|P_i = 1] + \delta + \lambda_2 - (E[\gamma_i|P_i = 1] + \lambda_1)$$

$$= \delta + \lambda_2 - \lambda_1$$

\[\text{timetrend}\]
If we calculated a treatment vs. comparison estimator, we’d have:

\[
E[\tilde{Y}_{\text{treatment, post}}] - E[\tilde{Y}_{\text{comparison, post}}] = E[\gamma_i | P_i = 1] + \delta + \lambda_2 - (E[\gamma_i | P_i = 0] + \lambda_2)
\]

\[
= \delta + E[\gamma_i | P_i = 1] - E[\gamma_i | P_i = 0]
\]

selection bias
Difference-in-Difference Estimation

If we calculated a treatment vs. comparison estimator, we’d have:

\[
E[\bar{Y}_{treatment}^{post}] - E[\bar{Y}_{comparison}^{post}] = E[\gamma_i | P_i = 1] + \delta + \lambda_2 - (E[\gamma_i | P_i = 0] + \lambda_2)
\]

\[
= \delta + E[\gamma_i | P_i = 1] - E[\gamma_i | P_i = 0]
\]

The diff-in-diff estimator removes the selection bias, time trend:

\[
DD = \bar{Y}_{treatment}^{post} - \bar{Y}_{treatment}^{pre} - (\bar{Y}_{comparison}^{post} - \bar{Y}_{comparison}^{pre})
\]
Difference-in-Difference Estimation

Substituting in the terms from our model:

\[ DD = E[Y_{1i}|P_i = 1, t = 2] - E[Y_{1i}|P_i = 1, t = 1] \]

\[ - \left( E[Y_{1i}|P_i = 0, t = 2] - E[Y_{1i}|P_i = 0, t = 1] \right) \]

\[ = E[\gamma_i|P_i = 1] + \delta + \lambda_2 - (E[\gamma_i|P_i = 1] + \lambda_1) \]

\[ - \left[ E[\gamma_i|P_i = 0] + \lambda_2 - \left( E[\gamma_i|P_i = 0] + \lambda_1 \right) \right] \]

\[ = \delta \]

the true impact of the program on participants
Example: Supply vs. Demand for Education

The supply side of education (provision of quality schools, teachers):

- Are there enough schools?
- Have teachers received enough training?
- Are teachers present in the classroom?
- Are class sizes too large?
The supply side of education (provision of quality schools, teachers):

- Are there enough schools?
- Have teachers received enough training?
- Are teachers present in the classroom?
- Are class sizes too large?

**Supply constraints** are related to school quality

- Main problem: governments need to provide more, better schools
Example: Supply vs. Demand for Education

The supply side of education (provision of quality schools, teachers):

- Are there enough schools?
- Have teachers received enough training?
- Are teachers present in the classroom?
- Are class sizes too large?

Supply constraints are related to school quality

- Main problem: governments need to provide more, better schools

Research question: if the government builds more schools, how much will education levels, human capital increase?
Example: Supply vs. Demand for Education

The demand for education: would parents send their kids to school in the absence of compulsory schooling laws? Would kids exert sufficient effort?

- How large is the return to education?
  - Increase in wages resulting from an additional year of school
- Do parents understand the return to education?
- Can HHs afford to pay for children to go to school?
  - What is the opportunity cost of education?
- Do HHs need children on the farm, working at home, etc?
Example: Supply vs. Demand for Education

The demand for education: would parents send their kids to school in the absence of compulsory schooling laws? Would kids exert sufficient effort?

- How large is the return to education?
  - Increase in wages resulting from an additional year of school
- Do parents understand the return to education?
- Can HHs afford to pay for children to go to school?
  - What is the opportunity cost of education?
- Do HHs need children on the farm, working at home, etc?

Demand constraints are likely to be critical determinants of educational outcomes if the return to education (in terms of wages) are relatively low.
A “Natural” Experiment in Education

In a famous paper in the *American Economic Review*, Esther Duflo examines the impacts of a large wave of school construction in Indonesia:

Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment

*By Esther Duflo*

Between 1973 and 1978, the Indonesian government engaged in one of the largest school construction programs on record. Combining differences across regions in the number of schools constructed with differences across cohorts induced by the timing of the program suggests that each primary school constructed per 1,000 children led to an average increase of 0.12 to 0.19 years of education, as well as a 1.5 to 2.7 percent increase in wages. This implies estimates of economic returns to education ranging from 6.8 to 10.6 percent. (JEL I2, J31, O15, O22)
A “Natural” Experiment in Education


- Oil crisis creates large windfall for Indonesia
- Suharto uses oil money to fund school construction
- Close to 62,000 schools built by national gov’t
  - Approximately 1 school built per 500 school-age children
- More schools built in areas which started with less
- Schools intended to promote national identity
The Return to Education in Indonesia

Strategy: difference-in-difference estimation

- Data on children born before and after program (pre vs. post)
- Data on children born in communities where many schools were build (treatment), those where few schools were built (comparison)
- Difference-in-difference estimate of program impact compares pre vs. post differences in treatment vs. comparison communities

Intuitively, difference-in-difference estimation asks:

After controlling for time trends and unchanging differences between treatment and control communities, do children who were born into areas with more newly built INPRES schools get more education?
In practice, the difference-in-difference estimator is:

$$DD = \bar{Y}_{treatment}^{post} - \bar{Y}_{treatment}^{pre} - (\bar{Y}_{comparison}^{post} - \bar{Y}_{comparison}^{pre})$$
The Return to Education in Indonesia

In practice, the difference-in-difference estimator is:

\[ DD = \bar{Y}_{\text{treatment\ post}} - \bar{Y}_{\text{treatment\ pre}} - \left( \bar{Y}_{\text{comparison\ post}} - \bar{Y}_{\text{comparison\ pre}} \right) \]

<table>
<thead>
<tr>
<th>Dependent Variable: Years of Schooling</th>
</tr>
</thead>
<tbody>
<tr>
<td>Many Schools Built</td>
</tr>
<tr>
<td>---------------------</td>
</tr>
<tr>
<td>Over 11 in 1974</td>
</tr>
<tr>
<td>Under 7 in 1974</td>
</tr>
<tr>
<td>Difference</td>
</tr>
</tbody>
</table>

Younger children (reached school age after INPRES) in areas where the program built a large number of schools are the treatment group.

- Who is the comparison group?
The Return to Education in Indonesia

In practice, the difference-in-difference estimator is:

\[
DD = \bar{Y}_{treatment\ post} - \bar{Y}_{treatment\ pre} - \left( \bar{Y}_{comparison\ post} - \bar{Y}_{comparison\ pre} \right)
\]
The Return to Education in Indonesia

In practice, the difference-in-difference estimator is:

$$DD = \bar{Y}_{treatment\ post} - \bar{Y}_{treatment\ pre} - \left( \bar{Y}_{comparison\ post} - \bar{Y}_{comparison\ pre} \right)$$

<table>
<thead>
<tr>
<th>Dependent Variable: Years of Schooling</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Many Schools Built</td>
</tr>
<tr>
<td>-------------------</td>
</tr>
<tr>
<td>Over 11 in 1974</td>
</tr>
<tr>
<td>Under 7 in 1974</td>
</tr>
<tr>
<td>Difference</td>
</tr>
</tbody>
</table>

Pre vs. post analysis does not control for time trends

- Indonesia is getting wealthier over time, so younger children (those entering school after the program) may get more education anyway
In practice, the difference-in-difference estimator is:

\[
DD = \bar{Y}_{treatment}^{post} - \bar{Y}_{treatment}^{pre} - \left( \bar{Y}_{comparison}^{post} - \bar{Y}_{comparison}^{pre} \right)
\]
The Return to Education in Indonesia

In practice, the difference-in-difference estimator is:

\[ DD = \bar{Y}_{treatment \ post} - \bar{Y}_{treatment \ pre} - (\bar{Y}_{comparison \ post} - \bar{Y}_{comparison \ pre}) \]

<table>
<thead>
<tr>
<th>Dependent Variable: Years of Schooling</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Many Schools Built</strong></td>
</tr>
<tr>
<td>-------------------------</td>
</tr>
<tr>
<td>Over 11 in 1974</td>
</tr>
<tr>
<td>Under 7 in 1974</td>
</tr>
<tr>
<td>Difference</td>
</tr>
</tbody>
</table>

Treatment vs. comparison analysis does not control for selection bias

- More schools were built in those areas that were initially lagging behind — poorer, more remote, less developed communities
In practice, the difference-in-difference estimator is:

\[
DD = \bar{Y}_{\text{treatment post}} - \bar{Y}_{\text{treatment pre}} - \left( \bar{Y}_{\text{comparison post}} - \bar{Y}_{\text{comparison pre}} \right)
\]
In practice, the difference-in-difference estimator is:

\[
DD = \bar{Y}_{\text{treatment post}} - \bar{Y}_{\text{treatment pre}} - \left( \bar{Y}_{\text{comparison post}} - \bar{Y}_{\text{comparison pre}} \right)
\]

### Dependent Variable: Years of Schooling

<table>
<thead>
<tr>
<th></th>
<th>Many Schools Built</th>
<th>Few Schools Built</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Over 11 in 1974</td>
<td>8.02</td>
<td>9.40</td>
<td>-1.38</td>
</tr>
<tr>
<td>Under 7 in 1974</td>
<td>8.49</td>
<td>9.76</td>
<td>-1.27</td>
</tr>
<tr>
<td>Difference</td>
<td>0.47</td>
<td>0.36</td>
<td>0.12</td>
</tr>
</tbody>
</table>

Difference-in-difference estimation compares the change in years of schooling (i.e. the pre vs. post estimate) in treatment, control areas

- Program areas increased faster than comparison areas
In practice, the difference-in-difference estimator is:

\[ DD = \bar{Y}_{treatment\ post} - \bar{Y}_{treatment\ pre} - \left( \bar{Y}_{comparison\ post} - \bar{Y}_{comparison\ pre} \right) \]
The Return to Education in Indonesia

In practice, the difference-in-difference estimator is:

$$DD = \bar{Y}_{treatment}^{post} - \bar{Y}_{treatment}^{pre} - (\bar{Y}_{comparison}^{post} - \bar{Y}_{comparison}^{pre})$$

Dependent Variable: Years of Schooling

<table>
<thead>
<tr>
<th></th>
<th>Many Schools Built</th>
<th>Few Schools Built</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Over 11 in 1974</td>
<td>8.02</td>
<td>9.40</td>
<td>-1.38</td>
</tr>
<tr>
<td>Under 7 in 1974</td>
<td>8.49</td>
<td>9.76</td>
<td>-1.27</td>
</tr>
<tr>
<td>Difference</td>
<td>0.47</td>
<td>0.36</td>
<td>0.12</td>
</tr>
</tbody>
</table>

Diff-in-diff estimate suggests program increased educational attainment by 0.12 years
In practice, the difference-in-difference estimator is:

\[
DD = \bar{Y}_{treatment\ post} - \bar{Y}_{treatment\ pre} - \left( \bar{Y}_{comparison\ post} - \bar{Y}_{comparison\ pre} \right)
\]
In practice, the difference-in-difference estimator is:

\[ DD = \bar{Y}_{\text{treatment} \ post} - \bar{Y}_{\text{treatment} \ pre} - (\bar{Y}_{\text{comparison} \ post} - \bar{Y}_{\text{comparison} \ pre}) \]

Dependent Variable: Log (Wages)

<table>
<thead>
<tr>
<th></th>
<th>Many Schools Built</th>
<th>Few Schools Built</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Over 11 in 1974</td>
<td>6.87</td>
<td>7.02</td>
<td>-0.15</td>
</tr>
<tr>
<td>Under 7 in 1974</td>
<td>6.61</td>
<td>6.73</td>
<td>-0.12</td>
</tr>
<tr>
<td>Difference</td>
<td>-0.26</td>
<td>-0.29</td>
<td>0.026</td>
</tr>
</tbody>
</table>

Diff-in-diff estimate suggests program increased (log) adult wages by 0.026
The Return to Education in Indonesia

A1: Experiment of interest: education

A2: Experiment of interest: log(wages)

B1: Control experiment: education

B2: Control experiment: log(wages)
Another way of looking at the data:

- Interact year of birth with schools built per 1000 students (intensity)
- We observe impacts for children under 10 when program started
The Return to Education in Indonesia

How much additional schooling did impacted children complete?
The Return to Education in Indonesia

- Total education attainment and adult wages grew faster in areas where more schools were built as part of the INPRES program
  - Schools caused an increase in education
  - Increases in education caused an increase in wages
The Return to Education in Indonesia

- Total education attainment and adult wages grew faster in areas where more schools were built as part of the INPRES program
  - Schools caused an increase in education
  - Increases in education caused an increase in wages
- Results suggest that each additional year of primary schooling leads to about an 8 percentage point increase in adult wages
- Returns to education are large, supply side interventions can work!
Diff-in-Diff in a Regression Framework
To implement diff-in-diff in a regression framework, we estimate:

\[ Y_i = \alpha + \beta P_i + \zeta Later + \delta (P_i \ast Later) + \epsilon_i \]

where:

- \( Later_i \) is an indicator equal to 1 if \( t = 2 \)
To implement diff-in-diff in a regression framework, we estimate:

$$Y_i = \alpha + \beta P_i + \zeta Later_t + \delta (P_i \times Later_t) + \varepsilon_i$$

where:

- $Later_i$ is an indicator equal to 1 if $t = 2$
- $\delta$ is the coefficient of interest (the treatment effect)
To implement diff-in-diff in a regression framework, we estimate:

\[ Y_i = \alpha + \beta P_i + \zeta \text{Later}_t + \delta (P_i \times \text{Later}_t) + \varepsilon_i \]

where:

- \( \text{Later}_i \) is an indicator equal to 1 if \( t = 2 \)
- \( \delta \) is the coefficient of interest (the treatment effect)
- \( \alpha = E[\gamma_i|P_i = 0] + \lambda_1 \) — pre-program mean in comparison group
- \( \beta = E[\gamma_i|P_i = 1] - E[\gamma_i|P_i = 0] \) — selection bias
- \( \zeta = \lambda_2 - \lambda_1 \) — time trend
The data set SIEFIE_DD_data1.dta contains observations of student test scores (score) for students in Standard 7, normalized to be out of 100 percent. The data set includes observations from two years, 2008 and 2009, for two schools. In 2009, school 1 was selected to receive a package of educational inputs (textbooks, flipcharts, and study guides) from a local NGO. School 2 was not selected to participate in the program.
Practice Problems 1

The data set SIEFIE_DD_data1.dta contains observations of student test scores (score) for students in Standard 7, normalized to be out of 100 percent. The data set includes observations from two years, 2008 and 2009, for two schools. In 2009, school 1 was selected to receive a package of educational inputs (textbooks, flipcharts, and study guides) from a local NGO. School 2 was not selected to participate in the program.

First steps:

- Open Stata
- Open SIEFIE_DD_data1.dta
- Open the do file diff-in-diff-part1-problems.do
### Practice Problems 1

#### Dependent Variable: Test Scores

<table>
<thead>
<tr>
<th>Year</th>
<th>Program School</th>
<th>Comparison School</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008</td>
<td>63.93</td>
<td>66.59</td>
<td>-2.66</td>
</tr>
<tr>
<td>2009</td>
<td>70.21</td>
<td>67.75</td>
<td>2.46</td>
</tr>
<tr>
<td>Difference</td>
<td>6.28</td>
<td>1.16</td>
<td>5.16</td>
</tr>
</tbody>
</table>

**Diff-in-diff estimate suggests program improved test scores**
Common Trends Assumption
Common Trends Assumption

Your fantastic research assistant has stumbled across test score data for both schools from 2005 through 2007. This data is stored in the data set SIEFIE_data2.dta.
Common Trends Assumption

Your fantastic research assistant has stumbled across test score data for both schools from 2005 through 2007. This data is stored in the data set SIEFIE_data2.dta.

What does this additional data buy us?

- Additional data allow us to plot pre-program trends by school
- Does the common trends assumption hold?
- How do we check?
- We can append the new data to find out
Common Trends Assumption

Average Test Score, by School

- Program School
- Comparison School

Year:
2004 2005 2006 2007 2008 2009 2010

Test Score:
55 60 65 70

SIEF IE Workshop: Difference-in-Difference Estimation  Slide 57
Common Trends Assumption

Diff-in-diff estimation goes wrong when treatment and comparison groups were not on the same trajectory prior to the program

- This is the common trends assumption
Common Trends Assumption

Diff-in-diff estimation goes wrong when treatment and comparison groups were not on the same trajectory prior to the program

- This is the common trends assumption

Remember the assumptions underlying diff-in-diff estimation:

- Selection bias relates to fixed characteristics of individuals ($\gamma_i$)
- Time trend ($\lambda_t$) same for treatment and control groups
Common Trends Assumption

Diff-in-diff estimation goes wrong when treatment and comparison groups were not on the same trajectory prior to the program

- This is the common trends assumption

Remember the assumptions underlying diff-in-diff estimation:

- Selection bias relates to fixed characteristics of individuals \((\gamma_i)\)
- Time trend \((\lambda_t)\) same for treatment and control groups

These assumptions guarantee that the common trends assumption is satisfied, but they cannot be tested directly — we have to trust!

- As with any identification strategy, it is important to think carefully about whether it checks out both econometrically and intuitively
Evidence of different pre-treatment differences in trends means that the assumptions underlying diff-in-diff are not reasonable.
Testing the Common Trends Assumption

Evidence of different pre-treatment differences in trends means that the assumptions underlying diff-in-diff are not reasonable.

With enough data, we can:

- Test for pre-program trend differences
  - Trends may, for example, exist in levels but not logs
- Include separate linear trends for treatment, comparison groups
- Include controls for other factors that might be driving differences in pre-existing trends (in treatment, comparison groups)
Diff-in-Diff in Practice
Malaria kills about 800,000 people per year

- Most are African children
- Repeated bouts of malaria may also reduce overall child health
- Countries with malaria are substantially poorer than other countries, but it is not clear whether malaria is the cause or the effect
Organized efforts to eradicate malaria are a natural experiment

- First the US (1920s) and then many Latin American countries (1950s) launched major (and successful) eradication campaigns

- Compare trends in adult income by birth cohort in regions which did, did not see major reductions in malaria because of campaigns
Malaria Eradication in the Americas: A Retrospective Analysis of Childhood Exposure†

By Hoyt Bleakley∗

This study uses the malaria-eradication campaigns in the United States (circa 1920) and in Brazil, Colombia, and Mexico (circa 1955) to measure how much childhood exposure to malaria depresses labor productivity. The campaigns began because of advances in health technology, which mitigates concerns about reverse causality. Malarious areas saw large drops in the disease thereafter. Relative to non-malarious areas, cohorts born after eradication had higher income as adults than the preceding generation. These cross-cohort changes coincided with childhood exposure to the campaigns rather than to pre-existing trends. Estimates suggest a substantial, though not predominant, role for malaria in explaining cross-region differences in income. (JEL I12, I18, J13, O15)
Colombia’s malaria eradication campaign began in the late 1950s...
Colombia’s malaria eradication campaign began in the late 1950s... and led to a huge decline in malaria morbidity
Areas with highest pre-program prevalence saw largest declines in malaria

Panel B. Highly infected areas saw greater declines in malaria

Figure 1. Malaria Incidence Before and After the Eradication Campaign, Colombia
Estimation Strategy

In this framework, treatment is a continuous variable

- Areas with higher pre-intervention malaria prevalence were, in essence “treated” more intensely by the eradication program
- Malaria-free areas should not benefit from eradication
- They can be used (implicitly) to measure the time trend
Estimation Strategy

In this framework, treatment is a continuous variable

- Areas with higher pre-intervention malaria prevalence were, in essence “treated” more intensely by the eradication program
- Malaria-free areas should not benefit from eradication
- They can be used (implicitly) to measure the time trend

Exposure (during childhood) also depends on one’s year of birth

- Colombians born after 1957 were fully exposed to program
  - Did not suffer from chronic malaria in their early childhood
  - Did not miss school because of malaria
- Colombians born before 1940 were adults by the time the eradication campaign began, serve as the comparison group
Regression specification:

\[ Y_{j, post} - Y_{j, pre} = \alpha + \beta M_{j, pre} + \delta X_{j, pre} + \varepsilon_j \]

where

- \( Y_{j,t} \) is an outcome of interest (e.g., literacy)
- \( M_{j,pre} \) is pre-eradication malaria prevalence
- \( X_{j,pre} \) is a vector of region-level controls
- \( \varepsilon_{ipt} \) is the noise term
Estimation Strategy

Regression specification:

\[ Y_{j, post} - Y_{j, pre} = \alpha + \beta M_{j, pre} + \delta X_{j, pre} + \varepsilon_j \]

where

- \( Y_{j,t} \) is an outcome of interest (e.g., literacy)
- \( M_{j,pre} \) is pre-eradication malaria prevalence
- \( X_{j,pre} \) is a vector of region-level controls
- \( \varepsilon_{ipt} \) is the noise term

As we saw in the practice problems, this specification ("in changes") is equivalent to a standard diff-in-diff regression specification in levels.
Higher pre-eradication malaria exposure (on the x-axis) is associated with larger increases in income across birth cohorts (on the y-axis).
Regression specification:

\[ Y_{j,\text{post}} - Y_{j,\text{pre}} = \alpha + \beta M_{j,\text{pre}} + \delta X_{j,\text{pre}} + \varepsilon_{ipt} \]

### Table 3—Cross-Cohort Differences and Malaria: Colombia

<table>
<thead>
<tr>
<th>Dependent variables: Differences across cohorts in...</th>
<th>Malaria ecology (Poveda)</th>
<th>Malaria ecology (Mellinger)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Literacy</td>
<td>Years of schooling</td>
</tr>
<tr>
<td>Panel A. Alternative controls</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Additional controls: None (basic specification)</td>
<td>0.035***</td>
<td>0.168*</td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.088)</td>
</tr>
</tbody>
</table>
The Impact of Childhood Exposure to Malaria

Regression specification:

\[ Y_{j,post} - Y_{j,pre} = \alpha + \beta M_{j,pre} + \delta X_{j,pre} + \epsilon_{ipt} \]

Open the do file `diff-in-diff-bleakley1.do` to replicate the analysis.
Exploiting (More) Variation by Birth Cohort

We gain statistical power by exploiting all of the variation in childhood exposure to treatment (eradication) across regions and birth cohorts.
Exploiting (More) Variation by Birth Cohort

We gain statistical power by exploiting all of the variation in childhood exposure to treatment (eradication) across regions and birth cohorts.
Exploiting (More) Variation by Birth Cohort

We gain statistical power by exploiting all of the variation in childhood exposure to treatment (eradication) across regions and birth cohorts.
Panel Data Analysis

We gain statistical power by exploiting all of the variation in childhood exposure to treatment (eradication) across regions and birth cohorts

- Between 0 and 18 years of childhood post-eradication
- Interact exposure with pre-program malaria prevalence
- Treatment impacts should be larger for birth cohorts who spent more years of childhood malaria-free, areas with more initial malaria
Panel Data Analysis

We gain statistical power by exploiting all of the variation in childhood exposure to treatment (eradication) across regions and birth cohorts

- Between 0 and 18 years of childhood post-eradication
- Interact exposure with pre-program malaria prevalence
- Treatment impacts should be larger for birth cohorts who spent more years of childhood malaria-free, areas with more initial malaria

Treat data set as a (YOB×location) panel

- Control for region, YOB fixed effects
Regression specification:

\[ Y_{jkt} = \beta (M_j \times EXP_k) + \delta_k + \delta_j + \delta_t + \varepsilon_{jkt} \]

where

- \( M_j \) is pre-eradication malaria prevalence (by region)
- \( EXP_k \) is proportion of childhood post-eradication (by YOB)
- \( \delta_k \) is a YOB fixed effect
- \( \delta_j \) is a region of birth effect
- \( \delta_t \) is a census year fixed effect
- \( \varepsilon_{jkt} \) is a conditionally mean-zero error term
Panel Data Analysis

We can also look at the relationship between log (adult) wages and pre-eradication malaria rates separately by birth cohort.

Figure 4. Cohort-Specific Relationships: Income and Pre-Campaign Malaria
Summary and Review


• Is a credible impact evaluation strategy when the common trends assumption can be tested (i.e. you have a long panel)

• Outcomes can be defined in levels or transformed levels

• Treatment need not be binary: often involves an interaction between pre-intervention conditions (eg malaria) and intervention timing