Impacts of an Early Education Intervention on Students’ Learning Achievements in the Philippines
Application of Propensity Score Matching with Panel Data

Futoshi Yamauchi (IFPRI)
June 9, 2011
SEARCA, Los Banos
Objectives

- Introduce impact evaluation concepts and method - counterfactual
- Example: Third Elem Educ Project (TEEP) from the Philippines
The Archetypal Evaluation Problem (1)

- Impact evaluation assesses outcomes for a specific program *relative to* the situation in the absence of the program.
- Let $P$ denote program participation status of unit (household, student, school, etc) $i$. With $Pi = 1$ if unit $i$ receives program (treated) and $Pi = 0$ if not.
- Let $S$ be observable outcome, e.g., school performance. Two potential outcomes of each program participant $i$ are
  - $S1i = $ outcome with the program
  - $S0i = $ outcome without the program
The Archetypal Evaluation Problem (2)

• The impact of the program for unit $i$ is

$$IMPACT = E(S_{1i} \mid P_i=1) - E(S_{0i} \mid P_i=1)$$

• The first term is the *actual outcome* for participant

• The second term is what would have happened to participant without the program (a *counterfactual outcome*)

• The impact is the difference between the actual outcome and the counterfactual outcome
The Fundamental Issue and a Naïve Solution

• The fundamental issue in impact evaluation is that the counterfactual outcome $E(S_{0i}|P_i=1)$ is not observable.

• One naïve solution is to use outcome of the non-participants: $E(S_{1i}|P_i=1) - E(S_{0i}|P_i=0)$.
Does the Naïve Solution Work?

• Whether the naïve solution works or not depends on whether $E(S1i|\mathbf{P}=1) - E(S0i|\mathbf{P}=0)$ can approximate IMPACT.

• In other words, **Is the counterfactual outcome the same as the outcome of non-participants?**

• The difference between the counterfactual outcome and the outcome of non-participants is called **Selection bias**: $\text{Bias} = E(S0i|\mathbf{P}=1) - E(S0i|\mathbf{P}=0)$.
A Graphic Representation of the Counterfactual and the Impact

- Actual outcome for participants after program
- Actual outcome for non-participants after program
- Counterfactual outcome for participants after program

before and after comparison

participants

non-participants

before program

program implementation

program evaluation
Is It Hard to Find the Counterfactual?

- Theoretically, yes. Because no one can be in two different conditions at the same time.
- But if the participate are randomly selected then it is NOT hard to find the counterfactual.
- **Randomization**: the assignment of the program is independent of the characteristics of the recipients.
- Such a design is called a *social experiment*. The naïve solution works!
When Randomization is Not Implemented

• Why not?
  – Targeting: Donors want to target the most needed, eg. TEEP targeted areas with poor school infrastructure.
  – Agents decision: Eligible units make their own participation decision.
  – ……

• We have to understand how the participants are selected into the project.
Selection Problem

• **Selection problem:** Participants are different from nonparticipants in many ways.

We cannot simply assume that the outcome of nonparticipants provides a good estimate for the counterfactual.

• Two sources of selection:
  (i) Selection on observables (to researcher)
  (ii) Selection on unobservables (to researcher)

• We have to use econometric methods to take care of the counterfactual.
Econometric Methods to Deal with Selection on Observables

**Idea:** Develop a *comparison group* (a group of non-participants) that is similar to the treatment group in observable characteristics.
Propensity Score Matching

- Propensity score: the probability of participating in the project conditional on observed characteristics: $\text{Prob}(P_i=1 \mid X_i)$

- Compare participants and non-participants that share the same $\text{Prob}(P_i=1 \mid X_i)$

  - Rosenbaum and Rubin (1983) show that, matching on propensity score is as good as matching on $X_i$.

  - Reducing a multiple dimension problem into a single dimension problem
How about Selection on Unobservables?

• More challenging task is to deal with selection on unobservables (participants and non-participants are different in unobservable characteristics)

• Two types of unobservables
  – Unobservables that are fixed over time
  – Unobservables that are changing over time

• Econometric methods
  – Double-difference (DD) method
  – Discontinuity design
  – Pipeline comparison
  – Instrumental variable (IV) method
Double-Difference (DD) Method

• DD requires **panel data**, which include
  – **baseline data** collected before the program started
  – a **follow-up survey** that collect data after the program was implemented

• **Before intervention**: \( S_{i0} = a_0 + cX_{i0} + v_i + u_{i0} \)
• **After intervention**: \( S_{i1} = a_1 + bG_i + cX_{i1} + v_i + u_{i1} \)
• **DD**: \( dS_i = (a_1-a_0) + bG_i + c \ dX_i + (u_{i1}-u_{i0}) \)
• Impact estimate: \( b \)
• Time-invariant or fixed unobservable term \( v_i \) disappears!
Combination of Two or More Methods: Eg. DD+PSM

- DD method assumes selection bias doesn’t change over time.
- However, participants and non-participants have different trends.
- Solution: match the initial condition between the participants and non-participants before doing the DD
- More on DD+PSM later → We take this approach in TEEP IE study
Example: School Interventions
TEEP in the Philippines (1)

• Historically large intervention to poor divisions in 2001-2006
• Integrated package of reforms and inputs to schools
• Both hard and soft components

• Not randomized: Targeted to poor divisions (the most depressed)
• Initially 3 batch plan, but implemented sequentially if division is ready (esp. batches 1 & 2 were mixed)
Example: School Interventions
TEEP in the Philippines (2)

Components:

- School building construction & renovations
- Teacher training: instructional & subject-based
- Textbooks
- School-based management (parents, barangay, school: localized school governance)
- Equipment
• Luzon Sample
• Visayas Sample
• Mindanao Sample
Our Strategy

• Outcome variables: Change in test score
  – *Difference between NAT Grade 4 (2002/03) and Grade 6 (2004/05)*
  – *Overall score, mathematics score*

• Conditioning variable for TEEP in PSM
  – Municipality (school district) income class

• *Intuition*: Compare TEEP and Non-TEEP schools in under the similar initial condition (municipality income level) within each region
Income Class Dist in Visayas
School Districts by TEEP

Non-TEEP

TEEP

Histograms by teep

Income_class

Histograms by teep
Income Class Dist in Visayas
Schools by TEEP

Non-TEEP

TEEP

Histograms by teep
income_class

Fraction

0

.402526

teep==0

1 5

0

1 5

income_class

Histograms by teep
Logit Results: P(Z)

<table>
<thead>
<tr>
<th>Teep</th>
<th>Coef.</th>
<th>Std. Err.</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Region 6</td>
<td>-2.161</td>
<td>0.211</td>
<td>***</td>
</tr>
<tr>
<td>Region 8</td>
<td>-2.518</td>
<td>0.226</td>
<td>***</td>
</tr>
<tr>
<td>Income 2</td>
<td>1.341</td>
<td>0.308</td>
<td>***</td>
</tr>
<tr>
<td>Income 3</td>
<td>1.702</td>
<td>0.370</td>
<td>***</td>
</tr>
<tr>
<td>Income 4</td>
<td>0.306</td>
<td>0.190</td>
<td></td>
</tr>
<tr>
<td>Income 5</td>
<td>0.141</td>
<td>0.186</td>
<td></td>
</tr>
<tr>
<td>Region 6 * Income 2</td>
<td>-1.337</td>
<td>0.419</td>
<td>***</td>
</tr>
<tr>
<td>Region 6 * Income 3</td>
<td>-1.097</td>
<td>0.425</td>
<td>***</td>
</tr>
<tr>
<td>Region 6 * Income 4</td>
<td>0.330</td>
<td>0.259</td>
<td></td>
</tr>
<tr>
<td>Region 6 * Income 5</td>
<td>-1.980</td>
<td>0.388</td>
<td>***</td>
</tr>
<tr>
<td>Region 8 * Income 2</td>
<td>-0.784</td>
<td>0.397</td>
<td>**</td>
</tr>
<tr>
<td>Region 8 * Income 3</td>
<td>-0.911</td>
<td>0.426</td>
<td>**</td>
</tr>
<tr>
<td>Region 8 * Income 4</td>
<td>1.325</td>
<td>0.264</td>
<td>***</td>
</tr>
<tr>
<td>Region 8 * Income 5</td>
<td>0.954</td>
<td>0.312</td>
<td>***</td>
</tr>
<tr>
<td>Pupil teacher ratio (both local and national)</td>
<td>-0.008</td>
<td>0.004</td>
<td>*</td>
</tr>
<tr>
<td>Grade 4 total enrollment (in ages 6 to 11)</td>
<td>-0.008</td>
<td>0.001</td>
<td>***</td>
</tr>
<tr>
<td>Number of multi-grade classes</td>
<td>-0.042</td>
<td>0.040</td>
<td></td>
</tr>
<tr>
<td>Proportion of local funded teachers</td>
<td>0.203</td>
<td>0.596</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>1.304</td>
<td>0.212</td>
<td>***</td>
</tr>
<tr>
<td>Number of obs.</td>
<td>4208</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LR chi2(18)</td>
<td>1258</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prob &gt; chi2</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log likelihood</td>
<td>-2236</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pseudo R2</td>
<td>0.22</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Balancing Property

<table>
<thead>
<tr>
<th></th>
<th>diff1</th>
<th>se1</th>
<th>sig1</th>
<th>diff2</th>
<th>se2</th>
<th>sig2</th>
<th>diff3</th>
<th>se3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Region 6</td>
<td>-0.287</td>
<td>0.047</td>
<td>***</td>
<td>-0.004</td>
<td>0.046</td>
<td>0.010</td>
<td>0.046</td>
<td></td>
</tr>
<tr>
<td>Region 8</td>
<td>-0.144</td>
<td>0.050</td>
<td>***</td>
<td>0.000</td>
<td>0.055</td>
<td>-0.003</td>
<td>0.057</td>
<td></td>
</tr>
<tr>
<td>Income 2</td>
<td>0.012</td>
<td>0.032</td>
<td>0.002</td>
<td>0.017</td>
<td>0.004</td>
<td>0.022</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income 3</td>
<td>-0.012</td>
<td>0.040</td>
<td>0.000</td>
<td>0.035</td>
<td>-0.004</td>
<td>0.034</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income 4</td>
<td>0.108</td>
<td>0.050</td>
<td>**</td>
<td>0.004</td>
<td>0.062</td>
<td>0.022</td>
<td>0.060</td>
<td></td>
</tr>
<tr>
<td>Income 5</td>
<td>0.021</td>
<td>0.039</td>
<td>-0.001</td>
<td>0.054</td>
<td>0.000</td>
<td>0.041</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Region 6 * Income 2</td>
<td>-0.024</td>
<td>0.015</td>
<td>0.000</td>
<td>0.010</td>
<td>-0.002</td>
<td>0.011</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Region 6 * Income 3</td>
<td>-0.026</td>
<td>0.026</td>
<td>-0.001</td>
<td>0.025</td>
<td>-0.002</td>
<td>0.028</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Region 6 * Income 4</td>
<td>-0.048</td>
<td>0.033</td>
<td>-0.002</td>
<td>0.032</td>
<td>0.001</td>
<td>0.038</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Region 6 * Income 5</td>
<td>-0.101</td>
<td>0.020</td>
<td>***</td>
<td>0.000</td>
<td>0.005</td>
<td>-0.002</td>
<td>0.005</td>
<td></td>
</tr>
<tr>
<td>Region 8 * Income 2</td>
<td>-0.032</td>
<td>0.019</td>
<td>*</td>
<td>0.000</td>
<td>0.014</td>
<td>-0.004</td>
<td>0.014</td>
<td></td>
</tr>
<tr>
<td>Region 8 * Income 3</td>
<td>-0.041</td>
<td>0.027</td>
<td>0.000</td>
<td>0.025</td>
<td>-0.003</td>
<td>0.027</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Region 8 * Income 4</td>
<td>0.026</td>
<td>0.038</td>
<td>0.001</td>
<td>0.047</td>
<td>0.003</td>
<td>0.044</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Region 8 * Income 5</td>
<td>-0.008</td>
<td>0.014</td>
<td>-0.001</td>
<td>0.014</td>
<td>0.004</td>
<td>0.014</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pupil teacher ratio</td>
<td>-2.254</td>
<td>0.758</td>
<td>***</td>
<td>-1.101</td>
<td>0.847</td>
<td>-1.306</td>
<td>0.930</td>
<td></td>
</tr>
<tr>
<td>Grade 4 total enrollment</td>
<td>-7.475</td>
<td>1.325</td>
<td>***</td>
<td>0.687</td>
<td>1.198</td>
<td>0.511</td>
<td>1.257</td>
<td></td>
</tr>
<tr>
<td>Number of multi-grade classes</td>
<td>0.134</td>
<td>0.050</td>
<td>***</td>
<td>-0.037</td>
<td>0.077</td>
<td>-0.038</td>
<td>0.090</td>
<td></td>
</tr>
<tr>
<td>Proportion of local funded teachers</td>
<td>-0.005</td>
<td>0.003</td>
<td>-0.001</td>
<td>0.004</td>
<td>0.000</td>
<td>0.004</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of observations</td>
<td>4208</td>
<td></td>
<td></td>
<td>3949</td>
<td></td>
<td></td>
<td>3949</td>
<td></td>
</tr>
</tbody>
</table>
Propensity and Trimming: Teep and Non-TEEP

Graphs by teep
PSM Results

<table>
<thead>
<tr>
<th></th>
<th>Treated diff</th>
<th>Control diff</th>
<th>DD</th>
<th>se</th>
<th>sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall score</td>
<td>16.737</td>
<td>15.348</td>
<td>1.389</td>
<td>0.874</td>
<td></td>
</tr>
<tr>
<td>Math score</td>
<td>17.645</td>
<td>16.385</td>
<td>1.260</td>
<td>1.090</td>
<td></td>
</tr>
<tr>
<td>Number of obs.</td>
<td>1774</td>
<td>2434</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Treated diff</th>
<th>Control diff</th>
<th>DD</th>
<th>se</th>
<th>sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall score</td>
<td>16.074</td>
<td>12.139</td>
<td>3.934</td>
<td>1.129</td>
<td>***</td>
</tr>
<tr>
<td>Math score</td>
<td>16.961</td>
<td>11.719</td>
<td>5.242</td>
<td>1.473</td>
<td>***</td>
</tr>
<tr>
<td>Number of obs.</td>
<td>1541</td>
<td>2408</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Treated diff</th>
<th>Control diff</th>
<th>DD</th>
<th>se</th>
<th>sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall score</td>
<td>16.074</td>
<td>12.260</td>
<td>3.813</td>
<td>1.172</td>
<td>***</td>
</tr>
<tr>
<td>Math score</td>
<td>16.961</td>
<td>11.961</td>
<td>5.000</td>
<td>1.442</td>
<td>***</td>
</tr>
<tr>
<td>Number of obs.</td>
<td>1541</td>
<td>2408</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Downward bias in DD: TEEP was allocated to schools with a lower trend in NAT change over time
- Confirming that TEEP was targeted to areas/schools that have constraints on growth
## Component Effects

<table>
<thead>
<tr>
<th></th>
<th>TEEP &amp; Non-TEEP</th>
<th>TEEP only</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade 4 textbooks (per pupil)</td>
<td>0.014</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>-0.003</td>
<td>0.004</td>
</tr>
<tr>
<td>Grade 5 textbooks (per pupil)</td>
<td>-0.004</td>
<td>0.003</td>
</tr>
<tr>
<td>Instructional training (man-hours per pupil)</td>
<td>0.339</td>
<td>0.192</td>
</tr>
<tr>
<td>Subject training (man-hours per pupil)</td>
<td>-0.582</td>
<td>0.260</td>
</tr>
<tr>
<td>New constructions (SY2003/04)</td>
<td>2.287</td>
<td>1.199</td>
</tr>
<tr>
<td>New renovations (SY 2003/04)</td>
<td>0.235</td>
<td>0.292</td>
</tr>
<tr>
<td>Region 6</td>
<td>0.388</td>
<td>2.672</td>
</tr>
<tr>
<td>Region 8</td>
<td>-2.808</td>
<td>2.716</td>
</tr>
<tr>
<td>Income 2</td>
<td>5.629</td>
<td>2.906</td>
</tr>
<tr>
<td>Income 3</td>
<td>-0.036</td>
<td>2.864</td>
</tr>
<tr>
<td>Income 4</td>
<td>-0.424</td>
<td>2.711</td>
</tr>
<tr>
<td>Income 5</td>
<td>1.666</td>
<td>2.526</td>
</tr>
<tr>
<td>Region 6 * Income 2</td>
<td>-2.378</td>
<td>3.584</td>
</tr>
<tr>
<td>Region 6 * Income 3</td>
<td>-1.943</td>
<td>3.791</td>
</tr>
<tr>
<td>Region 6 * Income 4</td>
<td>-0.373</td>
<td>3.314</td>
</tr>
<tr>
<td>Region 6 * Income 5</td>
<td>0.467</td>
<td>3.156</td>
</tr>
<tr>
<td>Region 8 * Income 2</td>
<td>-1.671</td>
<td>3.738</td>
</tr>
<tr>
<td>Region 8 * Income 3</td>
<td>-0.382</td>
<td>3.349</td>
</tr>
<tr>
<td>Region 8 * Income 4</td>
<td>0.066</td>
<td>3.187</td>
</tr>
<tr>
<td>Region 8 * Income 5</td>
<td>2.788</td>
<td>3.473</td>
</tr>
<tr>
<td>Pupil teacher ratio (both local and national)</td>
<td>-0.101</td>
<td>0.037</td>
</tr>
<tr>
<td>Grade 4 total enrollment (in ages 6 to 11)</td>
<td>0.050</td>
<td>0.008</td>
</tr>
<tr>
<td>Number of multi-grade classes</td>
<td>-0.533</td>
<td>0.284</td>
</tr>
<tr>
<td>Proportion of local funded teachers</td>
<td>-10.257</td>
<td>5.170</td>
</tr>
<tr>
<td>Constant</td>
<td>17.540</td>
<td>2.624</td>
</tr>
<tr>
<td>Number of obs</td>
<td>4186.000</td>
<td></td>
</tr>
<tr>
<td>F (25, 446)</td>
<td>5.870</td>
<td></td>
</tr>
<tr>
<td>R-squared</td>
<td>0.046</td>
<td></td>
</tr>
</tbody>
</table>
Conclusions

• TEEP Average Effect
  – Significantly positive impact
  – 12 to 15 NAT score point increase over 6 years of elementary school (compared to non-TEEP)
  – Larger impact on mathematics

• Component Effects
  – Textbook: Early stage investment has dynamic positive effect on performance (cumulative effect)
  – Training: Methodology/theory training has a positive effect, while subject-wise training showed a negative effect
  – School Building: New constructions have a large positive effect (One new building/classroom = 3 to 4 NAT score increase in 2 years)
  – SBM: “Funding” does not show positive effect (however, SBM is thought to increase the above component effects)
Long-term impact study (On-going)

3500 hhs/students from 8 divisions [TEEP divisions]:
Ifugao*, Neuva Viscaya
Antique*, Iloilo
Negros Oriental*, Cebu
Leyte* and Western Samar

Gr-6 SY 1999/00 [Pre-TEEP cohort]
Gr-6 SY 2004/05, 2005/06 [In-TEEP cohort]

Gr-6 NEAT/NAT score data
Siblings data of the 3500 students (3500 * 6 = 21000)
Tracking 3500 students to capture schooling and work history