Universidad del Desarrollo
Facultad de Economía y Negocios

Longitudinal research designs and control groups

Vesna Mandakovic , PhD
Assistant Professor
FEN-UDD

Getting started

- Programs and policies are designed to achieve a certain goal (or set of goals).
- A program evaluation helps us determine if our goal is actually being achieved as intended.

- Impact evaluations gauge the success of a program—where success can be **broadly** or **narrowly** defined.

- They help us weed out less effective interventions from successful ones and also help us improve existing programs.

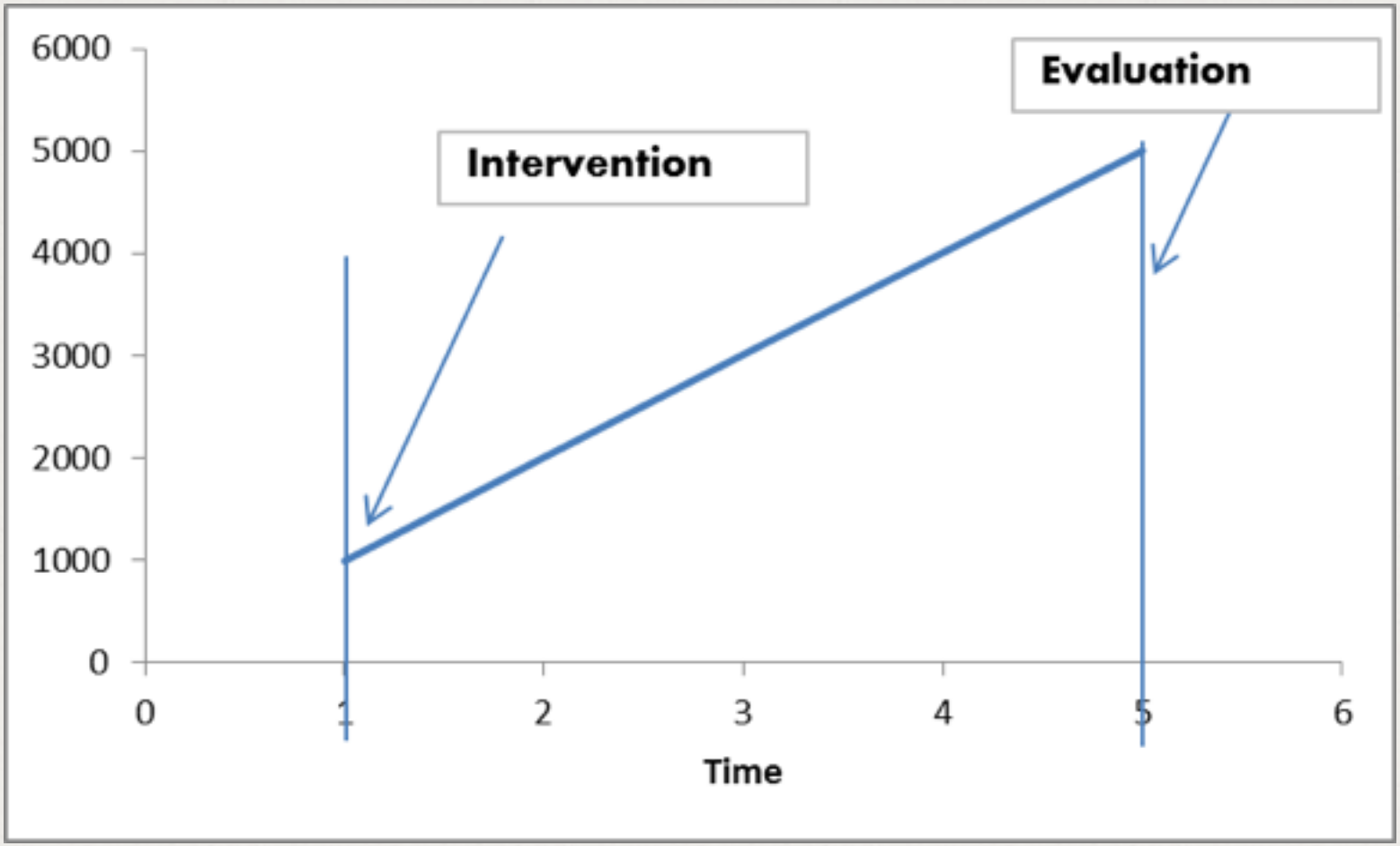
Impact Evaluation

- The primary purpose of impact evaluation is to determine whether a program has an impact (on a few key outcomes), and more specifically, to quantify *how large* that impact is. What is impact?

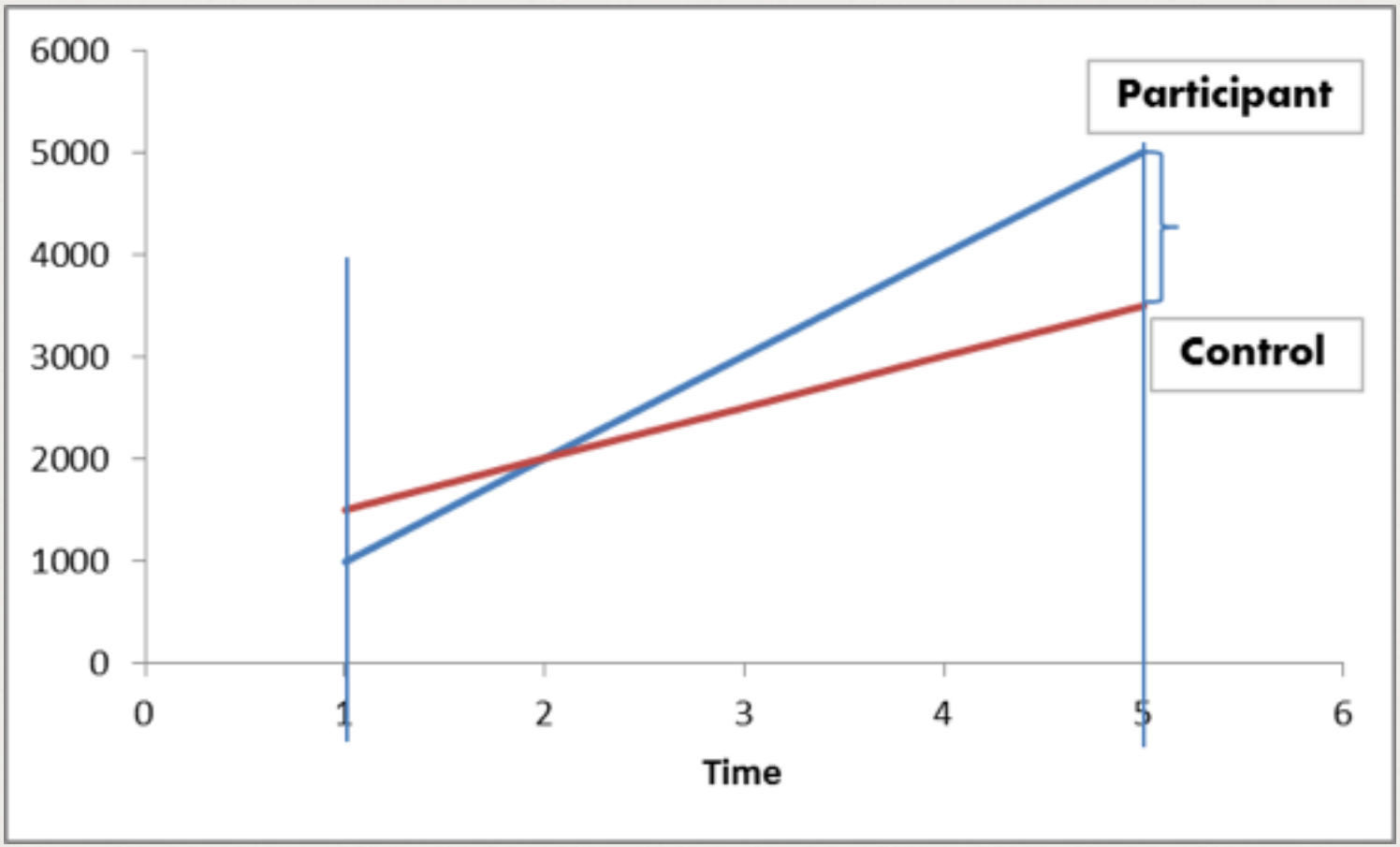
- Getting this number correct is more difficult than it sounds.

Example

- Consider the case of Grameen Bank's beneficiaries in Bangladesh.
- Grameen Bank offers credit to poor women to improve their food consumption.



Reflexive Method: compares before and after results of participants

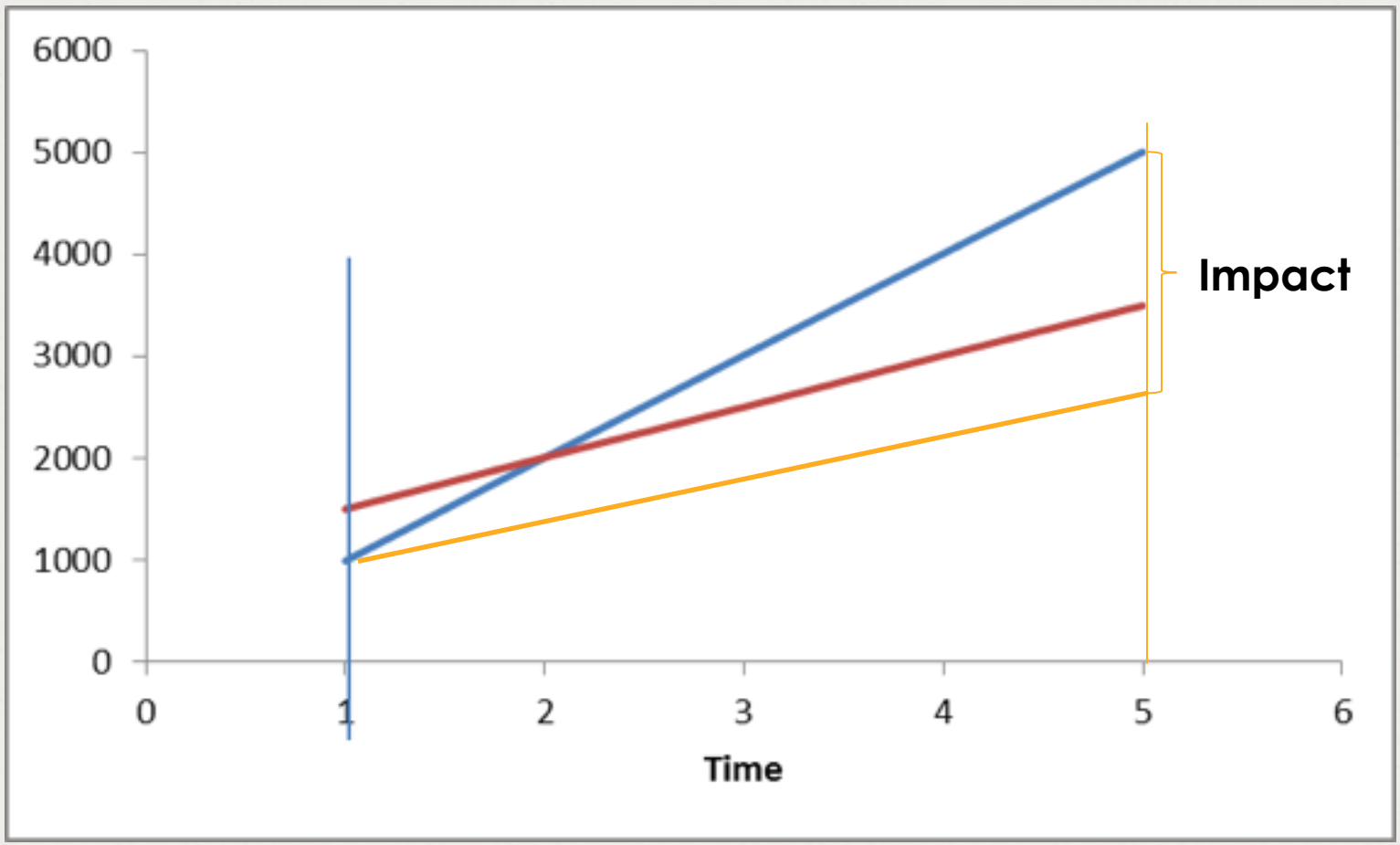

6000
5000
4000
3000
2000
1000
0
0
1
2
3
4
5
6
Time
Participant
Control

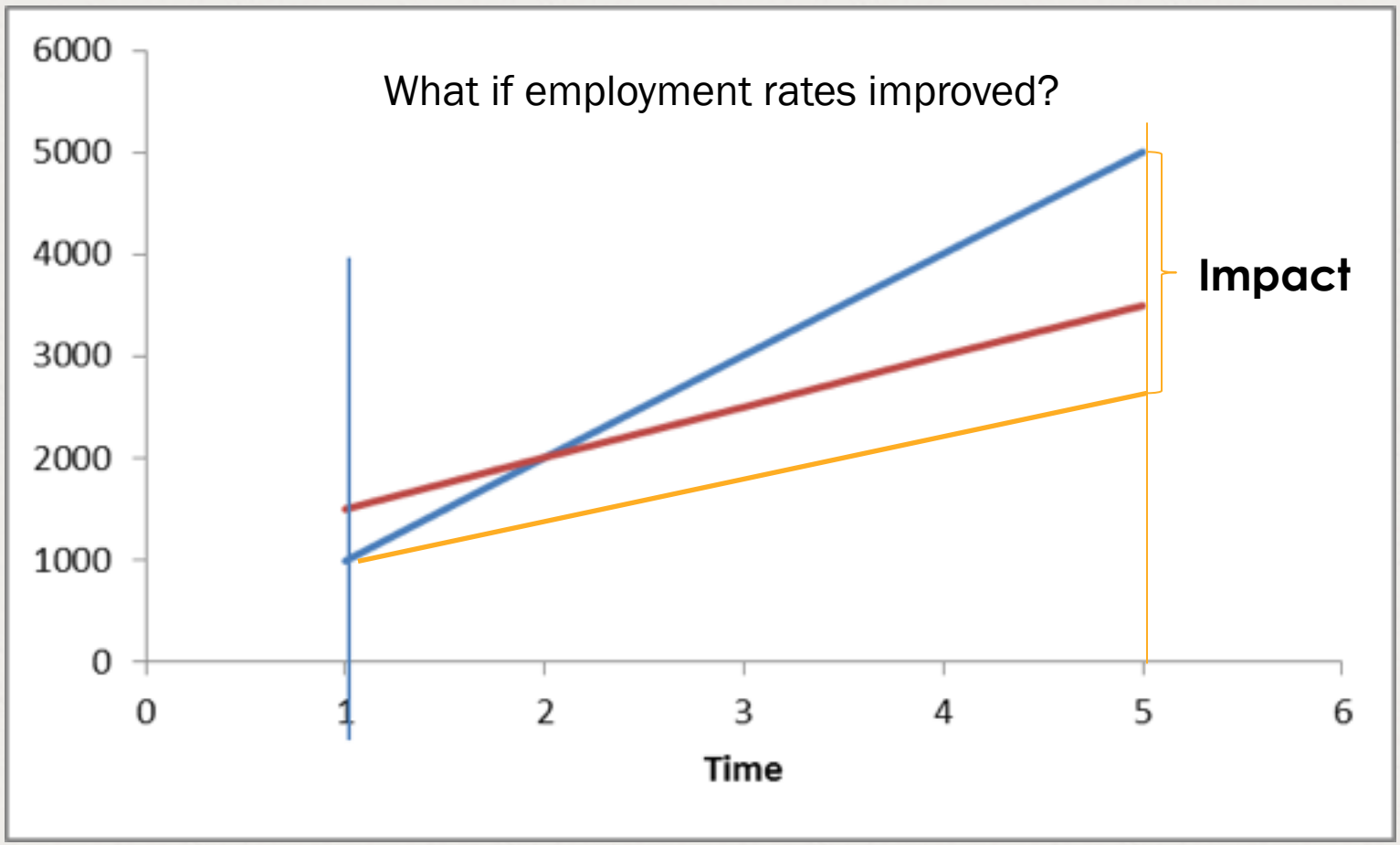
Impact Evaluation

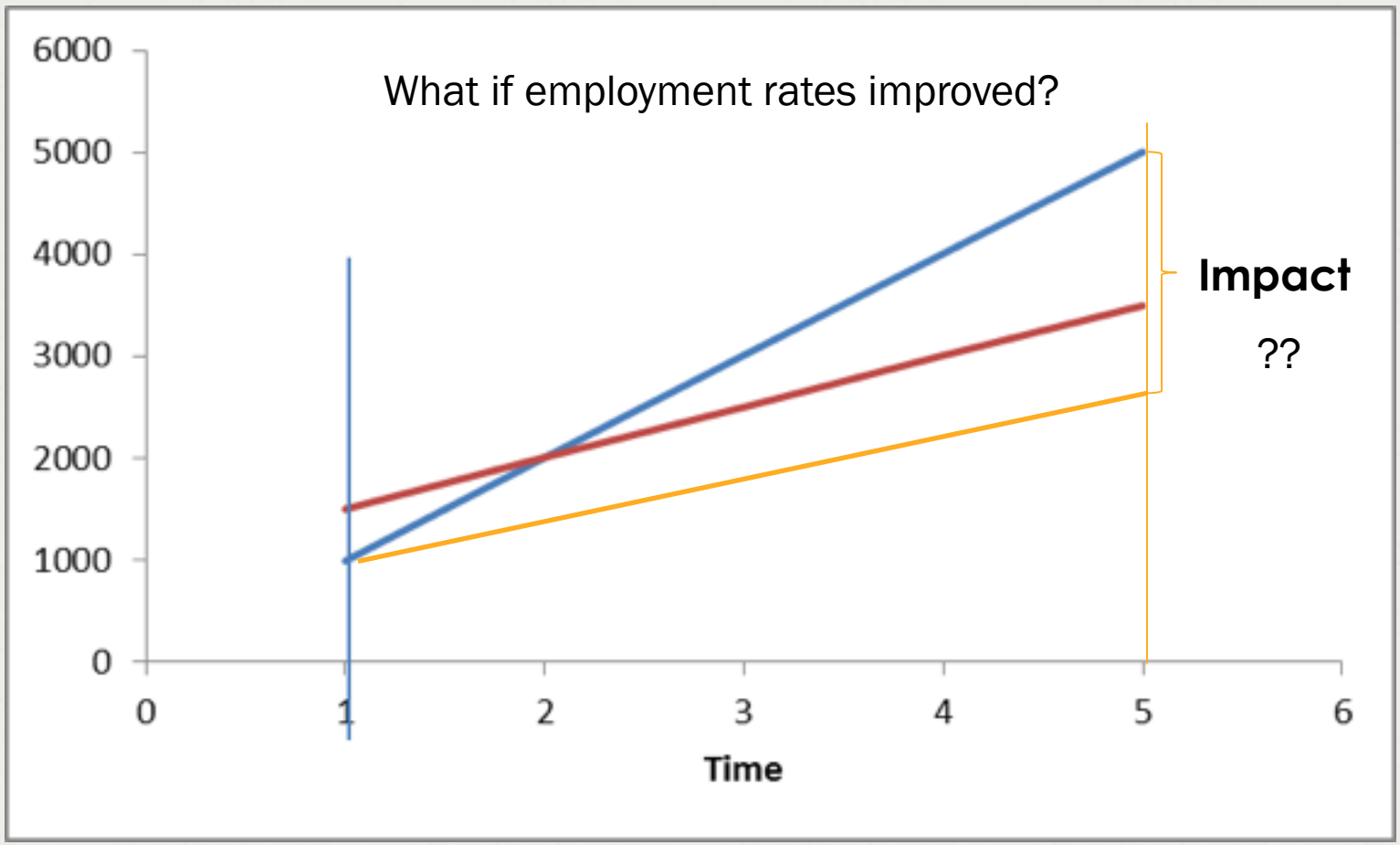
- Impact evaluations estimate program effectiveness usually by comparing outcomes of those who participated in the program against those who did not participate.

- The key challenge in impact evaluation is finding a group of people who did not participate, but closely resemble the participants had those participants not received the program.

- Measuring outcomes in this comparison group is as close as we can get to measuring "how participants would have been otherwise."


6000
5000
4000
3000
2000
1000
0
0
1
2
3
4
5
6
Time
Impact


What if employment rates improved?
6000
5000
4000
3000
2000
1000
0
0
1
2
3
4
5
6
Time
Impact


What if employment rates improved?
6000
5000
4000
3000
2000
1000
0
0
1
2
3
4
5
6
Time
Impact
??

The task is to isolate the effect of the program from other factors and potential selection bias.

Selection bias

Selection bias

- The treatment asignment could not be random because:

Selection bias

- The treatment asignment could not be random because:
  - purposive program placement: Programs are placed according to the need of the communities and individuals,
  - self-selection into the program: Based on observables and non observables

- The traditional way to measure the direct effect of the ε_i program *T* on outcomes *Y*.

$$Y_i = \alpha X_i + \beta T_i + \varepsilon_i$$

- X_i: Observable characteristics
- ε_i: Unobservable characteristics
- T_i: if treated *T=1* if not treated *T=0*
- If *COV*(ε_i, $T_i \neq 0$)….*Selection bias*

Methodology

Methodology

Quasi-Experimental Methods

Methodology

Quasi-Experimental Methods

Experimental Method

Methodology

Quasi-Experimental Methods

Experimental Method

- Difference in difference

Methodology

Quasi-Experimental Methods

- Difference in difference
- Statistical matching

Experimental Method

Methodology

Quasi-Experimental Methods

- Difference in difference
- Statistical matching
- Regression discontinuity design

Experimental Method

Methodology

Quasi-Experimental Methods

- Difference in difference
- Statistical matching
- Regression discontinuity design
- Instrumental variables

Experimental Method

Methodology

Quasi-Experimental Methods

- Difference in difference
- Statistical matching
- Regression discontinuity design
- Instrumental variables

Experimental Method

- Randomized Evaluation

DD: difference in difference

DD: difference in difference

- The unobserved heterogeneity exists, but this differences are time invariant.

DD: difference in difference

- The unobserved heterogeneity exists, but this differences are time invariant.

DD: difference in difference

- The unobserved heterogeneity exists, but this differences are time invariant.

- $\Delta Yit = \alpha + \beta Ti + \gamma \Delta Xit + \varepsilon it$, β = DD

DD: difference in difference

- The unobserved heterogeneity exists, but this differences are time invariant.

- $\Delta Yit = \alpha + \beta Ti + \gamma \Delta Xit + \varepsilon it$, β = DD

DD: difference in difference

- The unobserved heterogeneity exists, but this differences are time invariant.
- $\Delta Yit = \alpha + \beta Ti + \gamma \Delta Xit + \varepsilon it$, β = DD
- The treatment effect is determined by taking the difference in outcomes across treatment and control units before and after the program intervention.

Statistical Matching

Statistical Matching

Individuals in control group are compared to similar individuals.

Statistical Matching

Individuals in control group are compared to similar individuals.

EXACT MATCHING

Statistical Matching

Individuals in control group are compared to similar individuals.

EXACT MATCHING

PROPENSITY SCORE MATCHING

Statistical Matching

Individuals in control group are compared to similar individuals.

EXACT MATCHING

PROPENSITY SCORE MATCHING

Statistical Matching

Individuals in control group are compared to similar individuals.

EXACT MATCHING

PROPENSITY SCORE MATCHING

Statistical Matching

Individuals in control group are compared to similar individuals.

EXACT MATCHING

For each participant, at least one non-participant who is identical on selected characteristics

PROPENSITY SCORE MATCHING

Statistical Matching

Individuals in control group are compared to similar individuals.

EXACT MATCHING

For each participant, at least one non-participant who is identical on selected characteristics

PROPENSITY SCORE MATCHING

Non-participants who have a mix of characteristics which predict that they would be as likely to participate as participants

Regression discontinuity design

Regression discontinuity design

- They exploit exogenous program rules (such as eligibility requirements) to compare participants and nonparticipants in a close neighborhood around the eligibility cutoff.

Regression discontinuity design

- They exploit exogenous program rules (such as eligibility requirements) to compare participants and nonparticipants in a close neighborhood around the eligibility cutoff.

Regression discontinuity design

- They exploit exogenous program rules (such as eligibility requirements) to compare participants and nonparticipants in a close neighborhood around the eligibility cutoff.

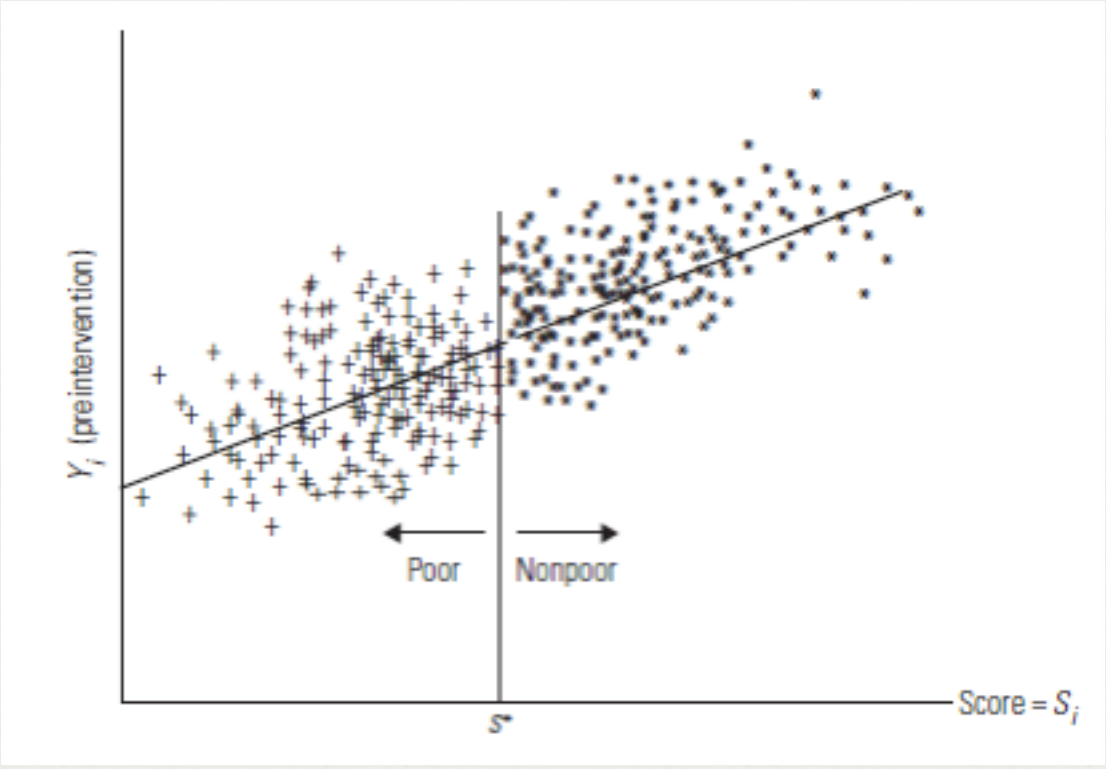
- Individuals who are close to the cutoff, but fall on the wrong side of that cutoff, and therefore do not get the program are the comparison group.

Ex. RDD

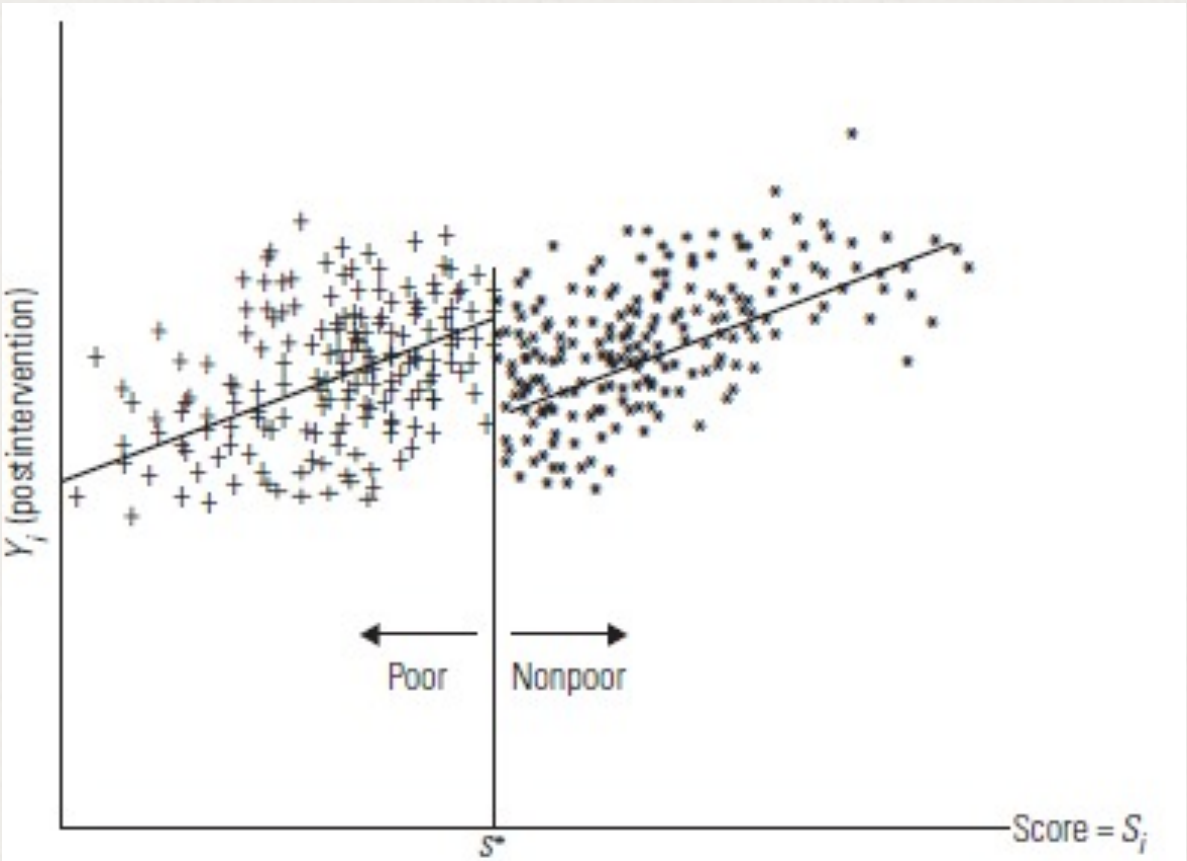
- The intuition behind the RDD is well illustrated using the evaluation of merit-based scholarships. The main problem with estimating the causal effect of such an intervention is the endogeneity of assignment to treatment (e.g. scholarship award): Since high-performing students are more likely to be awarded the merit scholarship and continue performing well at the same time, comparing the outcomes of awardees and non-recipients would lead to an upward bias of the estimates. Even if the scholarship did not improve grades at all, awardees would have performed better than non-recipients, simply because scholarships were given to students who were performing well ex ante.

- Despite the absence of an experimental design, a RDD can exploit exogenous characteristics of the intervention to elicit causal effects. If all students above a given grade—for example 80%—are given the scholarship, it is possible to elicit the local treatment effect by comparing students around the 80% cut-off: The intuition here is that a student scoring 79% is likely to be very similar to a student scoring 81%—given the pre-defined threshold of 80%, however, one student will receive the scholarship while the other will not. Comparing the outcome of the awardee (treatment group) to the counterfactual outcome of the non-recipient (control group) will hence deliver the local treatment effect

Pre intervention



Post intervention



Instrumental variables

- Participation can be predicted by an incidental factor or instrument variable.
- The IV variable is uncorrelated with the outcome, but predicts participation.
- Using IV makes the $COV(\varepsilon_i, T_i = 0)$ eliminating the selection bias

Randomized Evaluation

Randomized Evaluation

- *The PACES (Plan de Ampliación de Cobertura de la Educación Secundaria, or Plan for Increasing Secondary Education Coverage*) school voucher program, established by the Colombian government in late 1991, granted private secondary school vouchers to 125,000 children from poor neighborhoods who were enrolled in public primary schools. These vouchers covered about half of entering students' schooling expenses and were renewable depending on student performance.

Randomized Evaluation

- *The PACES (Plan de Ampliación de Cobertura de la Educación Secundaria, or Plan for Increasing Secondary Education Coverage*) school voucher program, established by the Colombian government in late 1991, granted private secondary school vouchers to 125,000 children from poor neighborhoods who were enrolled in public primary schools. These vouchers covered about half of entering students' schooling expenses and were renewable depending on student performance.

Randomized Evaluation

- *The PACES (Plan de Ampliación de Cobertura de la Educación Secundaria, or Plan for Increasing Secondary Education Coverage*) school voucher program, established by the Colombian government in late 1991, granted private secondary school vouchers to 125,000 children from poor neighborhoods who were enrolled in public primary schools. These vouchers covered about half of entering students' schooling expenses and were renewable depending on student performance.

- However, the program faced oversubscription because the number of eligible households (living in neighborhoods falling in the lowest two of six socioeconomic strata spanning the population) exceeded the number of

- To measure the impact of this school voucher program, Angrist and others (2002) surveyed **lottery winners and losers from three groups of applicants**. They administered an academic test to both groups, initially finding limited differences in performance for voucher recipients.

- To measure the impact of this school voucher program, Angrist and others (2002) surveyed **lottery winners and losers from three groups of applicants**. They administered an academic test to both groups, initially finding limited differences in performance for voucher recipients.

- To measure the impact of this school voucher program, Angrist and others (2002) surveyed **lottery winners and losers from three groups of applicants**. They administered an academic test to both groups, initially finding limited differences in performance for voucher recipients.

- One reason for this outcome, they suggest, is that about 10 percent of lottery winners did not end up using the voucher or other scholarship, whereas about 25 percent of nonrecipients obtained other scholarships or funding.

- To measure the impact of this school voucher program, Angrist and others (2002) surveyed **lottery winners and losers from three groups of applicants**. They administered an academic test to both groups, initially finding limited differences in performance for voucher recipients.

- One reason for this outcome, they suggest, is that about 10 percent of lottery winners did not end up using the voucher or other scholarship, whereas about 25 percent of nonrecipients obtained other scholarships or funding.

- To measure the impact of this school voucher program, Angrist and others (2002) surveyed **lottery winners and losers from three groups of applicants**. They administered an academic test to both groups, initially finding limited differences in performance for voucher recipients.

- One reason for this outcome, they suggest, is that about 10 percent of lottery winners did not end up using the voucher or other scholarship, whereas about 25 percent of nonrecipients obtained other scholarships or funding.

- Angrist and others (2002) therefore used the lottery receipt as an instrument for participation, calculating an intention-to-treat estimate that revealed much larger (50 percent greater) program effects on grade completion and reduced repetitions for lottery winners than in a simple comparison of winners and losers.

Closing Ideas…

Closing Ideas…

- During project identification and preparation, the importance and objectives of the evaluation need to be outlined clearly.

Closing Ideas…

- During project identification and preparation, the importance and objectives of the evaluation need to be outlined clearly.

Closing Ideas…

- During project identification and preparation, the importance and objectives of the evaluation need to be outlined clearly.

- Additional concerns one should time and structure impact evaluations beforehand to help program officials assess and update targeting, as well as other guidelines for implementation, during the course of the intervention.