



THE WORLD BANK



# **Technical Track**

## **Session II:**

# **Randomized Experiments**

Plamen Nikolov  
Sarajevo, Bosnia and Herzegovina,  
2009

# Outline

---

- I. Landscape of Methodologies
- II. Randomized Experiments
- III. Examples of Experiments

# Empirical Methods

---

- Difference in Differences
  - E.g. Laws and policies
- Matching methods (like Propensity Score Estimation).
  - E.g. Effect of Microfinance on profit
- Instrumental Variables (IV)
  - E.g. Effect of health on income
- Randomized trials (field and lab experiments)
- Regression Discontinuity
  - E.g. Conditional cash transfer programs
- Structural Estimation

# A Field Experiment Bridge

---

## □ Harrison & List (2004)

Controlled Data

Naturally-Occurring Data

---

Lab	AFE	FFE	NFE	NE, PSM, IV, STR
-----	-----	-----	-----	------------------

- Lab: Lab experiment
- AFE: Artefactual field experiment
- FFE: Framed field experiment
- NFE: Natural field experiment
- NE: Natural experiment (eg. exogenous policy change)
- PSM: Propensity score estimation
- IV: Instrumental variables estimation (using non-experimental data)
- STR: Structural modeling

# A Field Experiment Bridge

---

## **Conventional Lab experiment:**

- ❑ Employs a standard pool of subjects, an abstract framing, and an imposed set of rules. Characteristics of lab subjects (i.e. students) might be self-selected in some way, so that they are a sample that excludes certain individuals with characteristics that are important determinants of underlying population behavior.

**Example: Contributions towards a public good.**

## **Artefactual field experiment:**

- ❑ Similar as a conventional lab experiment but with a nonstandard subject pool.

**Example: Particular socio-economics group.**

## **Framed field experiment:**

- ❑ Similar to artefactual field experiment but with field context in either the commodity, task, or information set that the subjects can use

**Example: Do females compete less than males? But in a natural environment rather than a lab.**

## **Natural field experiment:**

- ❑ Framed field experiment but subjects naturally undertake these tasks and where the subjects do not know that they are in an experiment

**Example: Moving to Opportunity**

# Randomized Trials

---

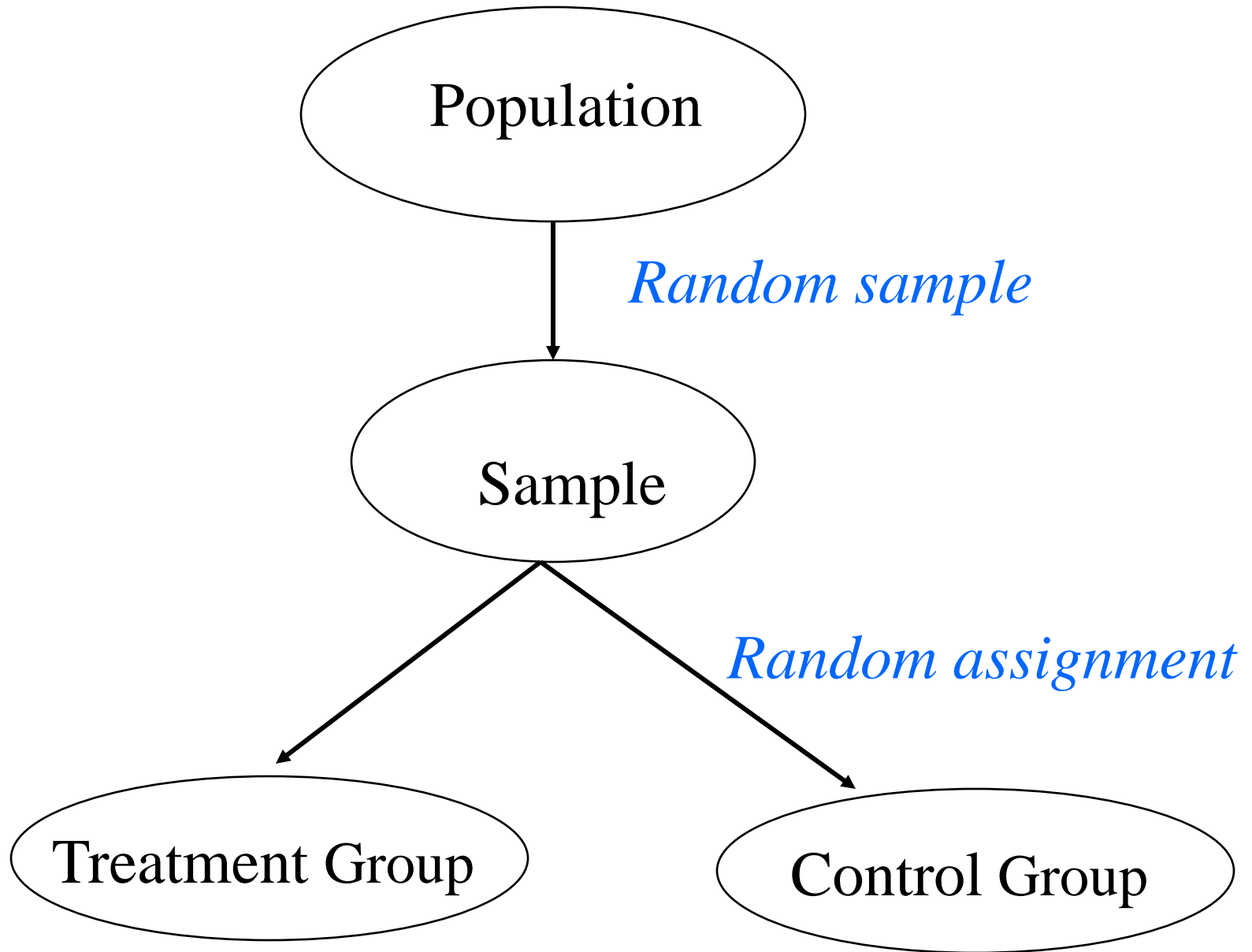
- ❑ How do researchers learn about counterfactual states of the world in practice?
- ❑ In many fields, and especially in medical research, evidence about counterfactuals is generated by randomized trials.
- ❑ Under certain conditions, randomized trials ensure that outcomes in the control group really do capture the **counterfactual** for a treatment group.

# Randomization for causal inference

---

Statisticians recommend a formal two-stage randomization model:

- ❑ **First stage:** a random sample of units is selected from a defined population.
- ❑ **Second stage:** this sample of units is randomly assigned to treatment and control groups.



# Why 2 stages of randomization?

---

## □ **First Stage: for External Validity**

- I.e. ensure that the results in the sample will represent the results in the population within a defined level of sampling error

## □ **Second Stage: for Internal Validity**

- I.e. ensure that the observed effect on the dependent variable is due to the treatment rather than to other confounding factors

# Random Assignment

---

Back to Evaluation Problem:

$$\text{Total Effect} = E[Y_1|D=1] - E[Y_0|D=1] - E[Y_0|D=0] + E[Y_0|D=1]$$

- Treatment Effect (ATT) =  $E[Y_1|D=1] - E[Y_0|D=1]$
- Selection Bias =  $E[Y_0|D=1] - E[Y_0|D=0]$

# Two-Stage Randomized Trials

---

In large samples, two-stage randomized trials ensure that:

$$[\bar{Y}_1 | D = 1] = [\bar{Y}_1 | D = 0] \quad \text{and} \quad [\bar{Y}_0 | D = 1] = [\bar{Y}_0 | D = 0]$$

Why is this true....

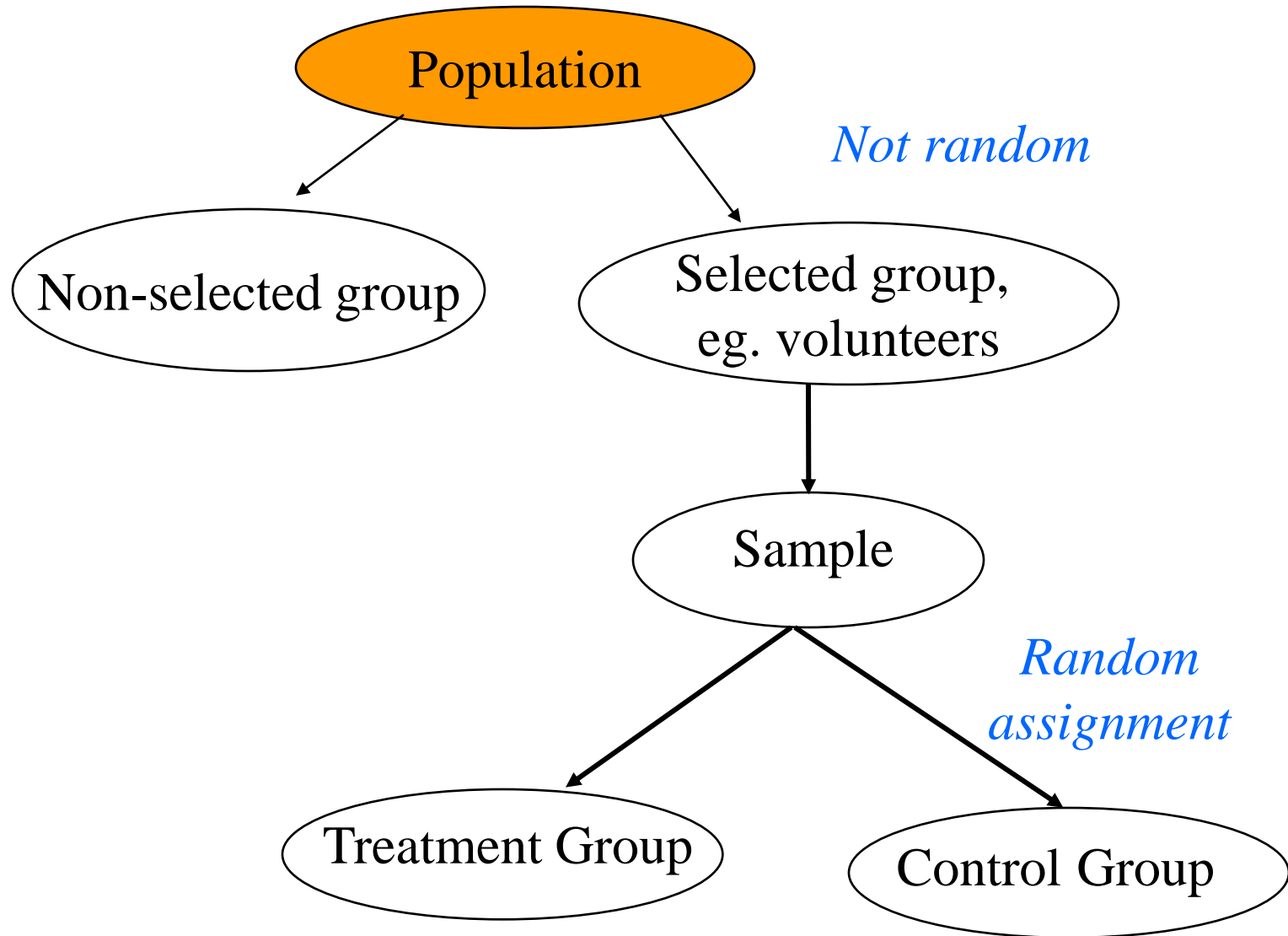
Thus, the estimator

$$\hat{\delta} = [\hat{Y}_1 | D = 1] - [\hat{Y}_0 | D = 0]$$

consistently estimates the Average Treatment Effect *ATE*

# Population vs. selected group?

---



# Population vs. selected group?

---

- If the randomization takes place on a selected group of units
- we'll be estimating ????
- The treatment effect on that selected group of units!

# Randomized Trials: Caveats

---

- **Non-compliance**
  - Not all treatment units will receive the treatment (non-compliance)
  - Some control units may receive treatment (non-compliance)
    - **Example:** Duncan Iron Deficiency Study
- **Attrition:** We may not be able to observe what happens to all units
  - **Example:** Tanzania study
- **Hawthorne effect:** just observing units makes them behave differently
  - **Example:** Lighting studies
- **John Henry effect:** the “controls” work harder to compensate
  - **Example:** Teachers in the comparison group for an evaluation may “compete” with the treatment teachers or, on the contrary, decide to slack off

# Randomized vs. Non-Randomized Trials

---

## □ **Randomized experiments**

- Assumptions play a minor role
- Or no role at all when testing the hypothesis of “no treatment effect”
- In the absence of difficulties such as noncompliance or attrition....

## □ **Non-randomized methods**

- Requires strong assumptions in order to validate the counterfactual
- Attrition is equally a problem as in randomized trials

# Example 1: Randomized Trial of “Computers for Education”, Colombia

---

- Research question: what is the impact of computers on educational outcomes?
- Program activities:
  - re-furbish computers donated by private firms and installs them in public schools.
  - train teachers in the pedagogic uses of computers with the help of a local university.
- 2006: 100 schools were subject to a randomization:
  - 50 of them received computers
  - 50 did not receive computers

## Appendix B. BASE LINE, COLOMBIA: SOME SELECTED STATISTICS

	Mean		Difference	Mean	Difference
	Urban	Rural	Urban/Rural	Sample	Treatment/Control
<b>SCHOOL INFORMATION</b>					
Number of teaches	16 (12.25)	7.55 (5.86)	8.45*** (1.84)	9.64 (8.67)	0.26 (1.34)
Student/Teach. ratio	14.1 (13.53)	13.53 (11.26)	0.57 (2.87)	13.67 (11.79)	-0.73 (1.77)
More than 10 years as teacher	0.88 (0.13)	0.64 (0.25)	0.24*** (0.02)	0.7 (0.25)	0.03 (0.03)
Total number of students	200.08 (217.21)	105.48 (140.87)	94.60*** (24.21)	128.89 (166.94)	-26.17 (21.17)
Number of repeating grade	14.08 (24.88)	4.96 (6.22)	9.12*** (3.04)	7.22 (13.89)	5.14*** (0.59)
Number of drop-outs	17.21 (24.37)	12.3 (15.15)	4.91** (2.42)	13.52 (17.86)	-1.23 (1.47)
Classrooms	10.83 (7.04)	6.48 (3.54)	4.35*** (1.10)	7.56 (4.99)	0.38 (0.90)
Libraries	0.54 (0.51)	0.6 (0.55)	-0.06 (0.06)	0.59 (0.54)	-0.01 (0.06)
Central component of school	0.83 (0.38)	0.86 (0.35)	-0.03 (0.06)	0.86 (0.35)	-0.17** (0.08)
Usefulness	1.08 (0.28)	1.04 (0.26)	0.04 (0.05)	1.05 (0.27)	-0.02 (0.03)

## INDIVIDUAL INFORMATION

Gender	0.55 (0.0067)	0.52 (0.0056)	0.03 (0.05)	0.54 (0.0043)	-0.04 (0.04)
Age	11 (0.0330)	12.54 (0.0311)	-1.54*** (0.54)	11.91 (0.0237)	0.17 (0.51)
Number of siblings	3.2 (0.0337)	4.16 (0.0322)	-0.96*** (0.25)	3.77 (0.0239)	-0.12 (0.28)
Work	0.16 (0.0050)	0.23 (0.0047)	-0.07*** (0.02)	0.2 (0.0035)	-0.01 (0.02)
Attend school last year	0.97 (0.1757)	0.98 (0.1544)	-0.0075 (0.0049)	0.97 (0.1635)	-0.0013 (0.0048)
Repeated grade last year	0.29 (0.4542)	0.38 (0.4861)	-0.0920*** (0.0251)	0.35 (0.4755)	0.0011 (0.0276)
Did not attend school last week	0.21 (0.4104)	0.24 (0.4250)	-0.0222 (0.0231)	0.23 (0.4193)	-0.0097 (0.0264)
How many days	2.15 (2.9484)	1.9 (1.6404)	0.2571* (0.1500)	2 (2.2373)	0.1003 (0.1523)
Like the school	0.98 (0.1313)	0.98 (0.1427)	0.0033 (0.0038)	0.98 (0.1382)	-0.0022 (0.0039)
Know internet	0.48 (0.4996)	0.35 (0.4758)	0.1314** (0.0517)	0.4 (0.4900)	-0.0009 (0.0592)
Uses internet (if yes)	0.8 (0.3993)	0.65 (0.4756)	0.1464*** (0.0422)	0.73 (0.4459)	-0.0077 (0.0514)
Hours of study outside school	1.47 (1.0064)	1.31 (0.9610)	0.1565*** (0.0559)	1.38 (0.9826)	0.0812 (0.0542)
Test scores: Language pool	0.45 (0.2667)	0.4 (0.2579)	0.0520** (0.0207)	0.42 (0.2627)	0.0072 (0.0229)
Test scores: Mathematics pool	0.33 (0.2574)	0.31 (0.2437)	0.0239 (0.0235)	0.31 (0.2495)	-0.0077 (0.0231)



## Example 2: Does reducing class size improve elementary school education?

---

- ❑ Project STAR (Student-Teacher Achievement Ratio)
- ❑ 4-year experiment to evaluate the effect of small class sizes on learning, in kindergarten through 3rd grade.
- ❑ Treatment levels:
  1. Regular class size: 22-25 students and a single teacher.
  2. Small class: 13-17 students and a single teacher.
  3. Teacher's aide: regular-sized class plus a teacher's aide.

## Example 2: Does reducing class size improve elementary school education?

---

- ❑ Each school had at least one class of each type.
- ❑ Students entering kindergarten in a participating school were randomly assigned to one of these three groups.
- ❑ Teachers were also assigned randomly.

## Estimates of Effect of Treatments on Standardized Test Scores

$$Y_i = \beta_0 + \beta_1 I_i^{SmallClass} + \beta_2 I_i^{Aide} + u_i$$

	<b>Grade</b>			
Regressor	<b>Kinder</b>	<b>1st</b>	<b>2nd</b>	<b>3rd</b>
Intercept	918.0*** (1.6)	1,039.4*** (1.8)	1,157.8*** (1.8)	1,228.5*** (1.7)
Small Class	13.9*** (2.5)	29.8*** (2.8)	19.4** (2.7)	15.6*** (2.4)
Regular Size with aide	0.3 (2.3)	12.0*** (2.7)	3.5 (2.5)	-0.3 (2.3)
<i>Number of Observations</i>	5,786	6,379	6,049	5,967

# Example 2: Does reducing class size improve elementary school education?

---

## □ Findings:

- Reducing class size has an effect on test performance,
- Adding a teacher's aide to a regular sized class has a much smaller effect, possible zero.

## □ Caveat:

- These estimates ignore both attrition and non-compliance.
- These two nuisances were high => the results might be biased.

# Non-Compliance and Attrition: Solutions?

---

## □ **Non-compliance:**

1. Intention to Treat Analysis
  - The main reason for advocating ITT analysis is that it maintains the baseline comparability achieved by the randomizing process.
2. Instrumental Variables Analysis (Local Average Treatment Effect)

## □ **Attrition (Hidden bias)**

1. Make sure that there is no difference in attrition between treatment and comparison groups
2. Use Instrumental Variables & Matching Methods

---

## **Example 3:**

# **Vouchers for Private Schooling in Colombia: Evidence from a Randomized Natural Experiment**

**Angrist et al. (2002)**

**AER**

# Programa de Aplicación de Cobertura de la Educación Secundaria (PACES)

---

## □ Program benefits

- School voucher programs for attendance of private secondary schools
- Vouchers covering somewhat more than half the cost of private secondary school
- Vouchers were renewable if students maintained satisfactory academic performance

## □ Beneficiaries

- 125,000 “treated” pupils
- Beneficiaries were chosen by lottery from a pool of eligible applicants

# Design of the PACES evaluation

---

- Interviewed 1,600 PACES applicants in 1998
  - Similar numbers of lottery winners and losers
  - 1995 and 1997 applicant cohorts from Bogota
  - 1993 applicant cohort from Jamundi
  
- Interview method: through telephone
  - Response rate: +/- 60%
  - Response is independent of treatment assignment (i.e. same for lottery winners and losers)

## Personal Characteristics and Voucher Status for Bogotá 1995

Dependent variable	Loser means	Won voucher
Age at time of survey	15.0 (1.4)	-0.013 (0.078)
Male	0.501	0.004 (0.029)
Mother's highest grade completed	5.9 (2.7)	-0.079 (0.166)
Mother's age	40.7 (7.3)	-0.027 (0.426)
Father's age	44.4 (8.1)	0.567 (0.533)
Father's wage (>2 min wage)	0.100	0.005 (0.021)
N	583	1,176

Notes: The table reports voucher losers' means and the estimated effect of winning a voucher. Numbers in parentheses are standard deviations in columns of means and standard errors in columns of estimated voucher effects.

## Educational Outcomes and Voucher Status, Bogotá 1995

Dependent variable	Loser means (1)	No controls (2)	Basic controls (3)	Full controls (4)
Finished 6th grade	0.943 (0.232)	0.026** (0.012)	0.023* (0.012)	0.021* (0.011)
Finished 8th grade	0.632 (0.483)	0.112*** (0.027)	0.100*** (0.027)	0.094*** (0.027)
Repetitions of 6th grade	0.194 (0.454)	-0.066*** (0.024)	-0.059** (0.024)	-0.059** (0.024)
Ever repeated after lottery	0.224 (0.417)	-0.060*** (0.023)	-0.055** (0.023)	-0.051** (0.023)
Years in school since lottery	3.7 (0.951)	0.058 (0.052)	0.034 (0.050)	0.031 (0.050)
Sample size	562		1,147	

Notes: The table reports voucher losers' means and the estimated effect of winning a voucher.

Numbers in parentheses are standard deviations in columns of means and standard errors in 28 columns of estimated voucher effects. \*\*\* significant at 1% \*\* significant at 5% \* significant at 10%

# References

---

- ❑ Rosenbaum, Paul (2002): Observational Studies, Springer. Chapter 2.
- ❑ Cochran, W. G. (1965): “The planning of observational studies of human populations”, *Journal of the Royal Statistics Association Series A 128*, pp. 134-155, with discussion.
- ❑ Angrist, J., E. Bettinger, E. Bloom, E. King and M. Kremer (2002): “Vouchers for Private Schooling in Colombia: Evidence from a Randomized Natural Experiment”, *American Economic Review*, 92, pp. 1535-58.
- ❑ Angrist, J. and V. Lavy (2002): “The Effect of High School Matriculation Awards: Evidence from Randomized Trials”, NBER Working Paper.