



THE WORLD BANK



Randomized Experiments

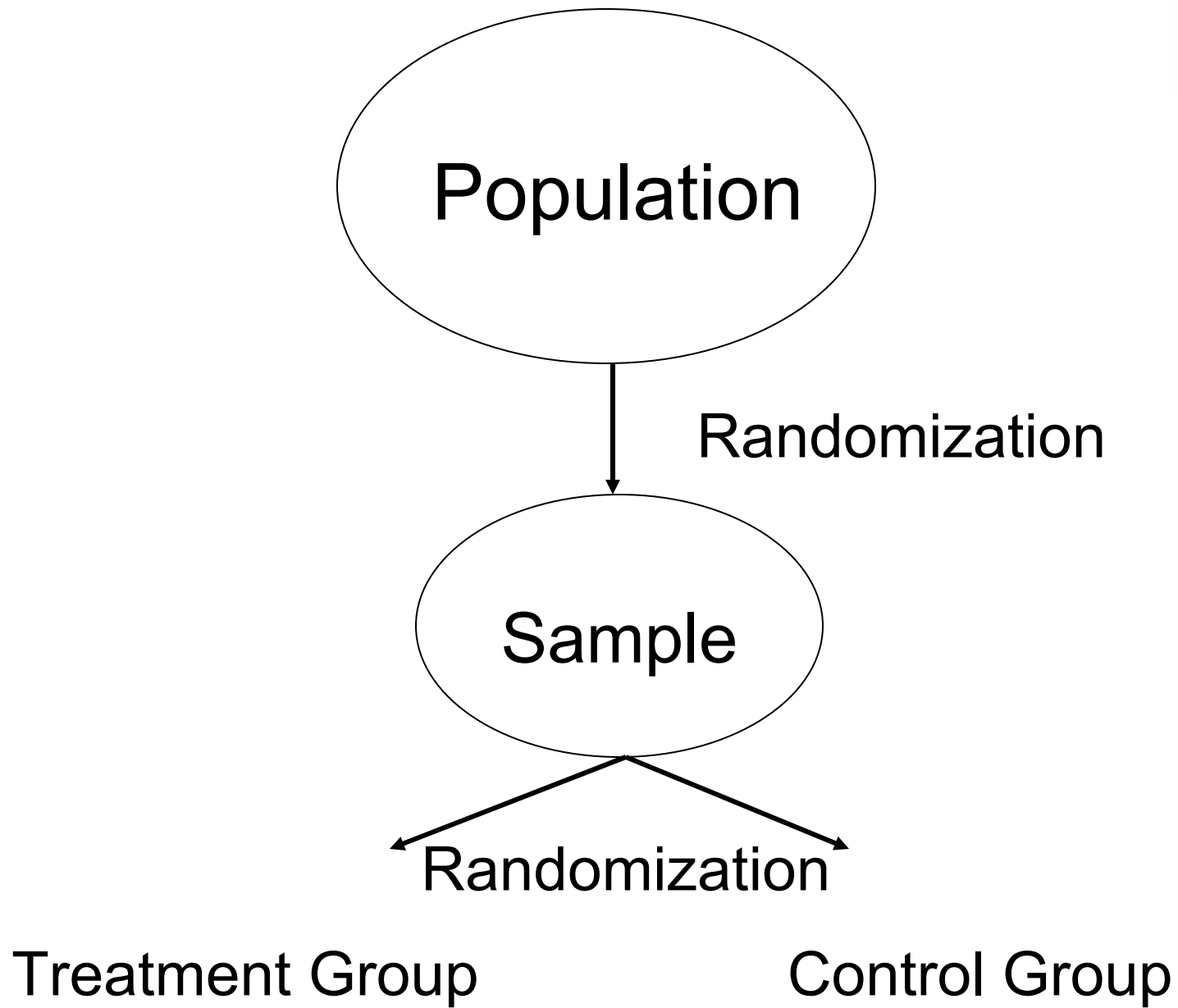


Randomized Trials

- 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.
- In principle, randomized trials ensure that outcomes in the control group really do capture the counterfactual for a treatment group.

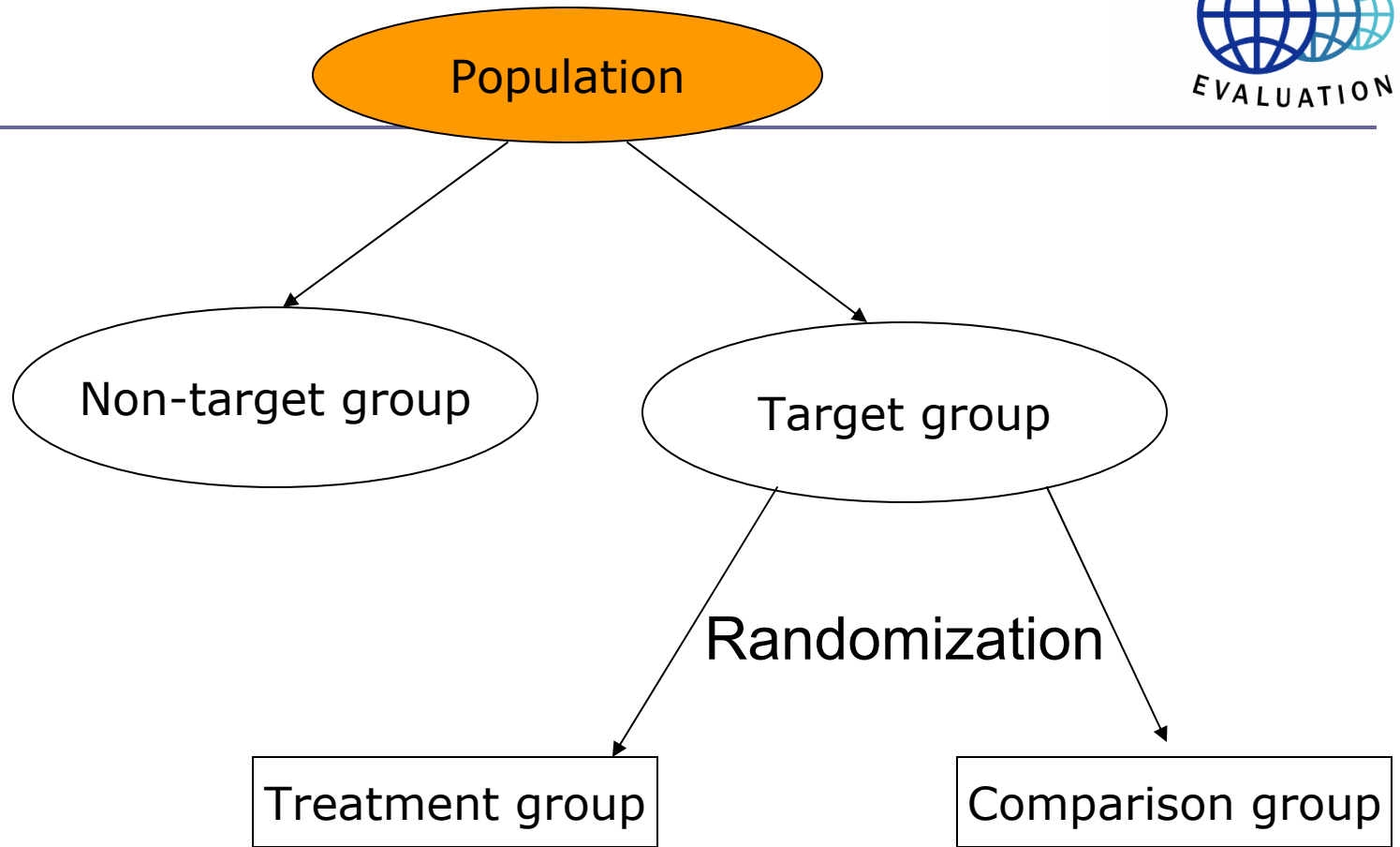
Randomization

- ❑ To answer causal questions, statisticians recommend a formal two-stage statistical model.
- ❑ In the first stage, a random sample of participants is selected from a defined population.
- ❑ In the second stage, this sample of participants is randomly assigned to treatment and comparison (control) conditions.



External & Internal Validity

- The purpose of the first-stage is to ensure that the results in the sample will represent the results in the population within a defined level of sampling error (**external validity**).
- The purpose of the second-stage is to ensure that the observed effect on the dependent variable is due to some aspect of the treatment rather than other confounding factors (**internal validity**).



Two-Stage Randomized Trials

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

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

- Thus, the estimator

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

- Consistently estimates ATE

One-Stage Randomized Trials

- ▣ Instead, if randomization takes place on a selected subpopulation –e.g., list of volunteers–, it only ensures:

$$[\bar{Y}_0 | D = 1] = [\bar{Y}_0 | D = 0]$$

- And hence, the estimator

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

- Only estimates TOT Consistently

Randomized Trials

- Furthermore, even in idealized randomized designs,
 1. Not all participants will receive the intended treatment (non-compliance)
 2. Not all participants will complete the dependent measure (attrition)

- Even the best-designed randomized experiment will at best only approximate the ideal of the two-stage statistical model.

Randomized Trials

- Nevertheless, the important point is that, in the absence of difficulties such as noncompliance or loss to follow up, assumptions play a minor role in randomized experiments, and no role at all in randomized tests of the hypothesis of no treatment effect.
- In contrast, inference in a nonrandomized experiment requires assumptions that are not at all innocuous.

Randomized Trial of Computers for Education, Colombia



- ▣ The program Computers for Education (Colombia) re-furbish computers donated by private firms and install them in public schools. It trains teachers in the pedagogic uses of computers with the help of a local university
- ▣ In 2006, 100 schools were subject to a randomization: 50 of them received computers, 50 did not

Randomized Trial of Computers for Education, Colombia



Appendix B. BASE LINE, COLOMBIA: SOME SELECTED STATISTICS

	Mean		Difference	Mean	Difference
	Urban	Rural	Urban/Rural	Sample	Treatment/Control
SCHOOL INFORMATION					
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)

Randomized Trial of Computers for Education, Colombia



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)

Does reducing class size improve elementary school education?

- ❑ Project STAR (Student-Teacher Achievement Ratio): 4-year experiment designed to evaluate the effect on learning of small class sizes.
- ❑ Focus of the experiment: 3 different class arrangements for kindergarten through third 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.

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.

$$Y_i = \beta_0 + \beta_1 \text{SmallClass}_i + \beta_2 \text{RegAide}_i + u_i$$

Project STAR: Differences Estimates of Effect on Standardized Test Scores of Class Size Treatment Group

Regressor	Grade			
	K	1	2	3
Small Class	13.90*** (2.45)	29.78*** (2.83)	19.39*** (2.71)	15.59*** (2.40)
Regular Size with aide	0.31 (2.27)	11.96*** (2.65)	3.48 (2.54)	-0.29 (2.27)
Intercept	918.04*** (1.63)	1,039.39*** (1.78)	1,157.81*** (1.82)	1,228.51*** (1.68)
Number of Observations	5,786	6,379	6,049	5,967

Does reducing class size improve elementary school education?

- The estimates presented suggest that:
 1. Reducing class size has an effect on test performance,
 2. But adding an aide to a regular sized class has a much smaller effect, possible zero.

However, the estimates presented ignore both attrition and non-compliance. These two nuisances were high, and hence, the results might be biased.

Possible Solution to Attrition and Non-compliance

- Non-compliance:
 1. Intention to Treat Analysis
 2. Instrumental Variables Analysis (Local Average Treatment Effect)

- Attrition (Hidden bias)
 1. Search for ignorability among sub-groups.
 2. Instrumental Variables & Matching Methods.

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

Angrist et al. (2002)
AER

Motivation

- ❑ This paper presents evidence on the impact of one of the largest school voucher programs to date: Programa de APLICACION de Cobertura de la Educacion Secundaria (PACES).
- ❑ Treatment: 125,000 pupils with vouchers covering somewhat more than half the cost of private secondary school.
- ❑ Vouchers were renewable as long as students maintained satisfactory academic performance.

Design

- ❑ The authors interviewed 1,600 PACES applicants in 1998, stratifying to obtain approximately equal numbers of winners and losers.
- ❑ For practical reasons, interviewing was limited to the 1995 and 1997 applicant cohorts from Bogota and the 1993 applicant cohort from Jamundi, a suburb of Cali.
- ❑ Telephones were used for the majority of interviews. Approximately 60% response rate. Response is independent of treatment assignment.

Personal Characteristics and Voucher Status

Dependent variable	Bogotá 1995		Bogotá 1997		Jamundi 1993	
	Loser means	Won voucher	Loser means	Won voucher	Loser means	Won voucher
Age at time of survey	15.0 (1.4)	-0.013 (0.078)	13.2 (1.4)	-0.259 (0.171)	17.2 (1.4)	-0.375 (0.217)
Male	0.501	0.004 (0.029)	0.527	-0.047 (0.061)	0.365	0.110 (0.077)
Mother's highest grade completed	5.9 (2.7)	-0.079 (0.166)	5.9 (2.7)	0.654 (0.371)	4.4 (2.7)	1.46 (0.494)
Father's highest grade completed	5.9 (2.9)	-0.431 (0.199)	5.5 (2.5)	0.929 (0.388)	5.2 (2.9)	0.737 (0.640)
Mother's age	40.7 (7.3)	-0.027 (0.426)	38.7 (6.6)	-0.146 (0.808)	43.6 (8.8)	-0.736 (1.42)
Father's wage	44.4 (8.1)	0.567 (0.533)	41.9 (7.3)	0.265 (0.973)	45.5 (9.1)	1.92 (1.61)
Father's wage (>2 min wage)	0.100	0.005 (0.021)	0.088	-0.008 (0.043)	0.133	-0.092 (0.056)
N	583	1,176	131	277	74	165

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 (I)

Dependent variable	Bogotá 1995				Combined sample	
	Loser means (1)	No controls (2)	Basic controls (3)	Basic +19 barrio controls (4)	Basic controls (5)	Basic +19 barrio controls (6)
Using any scholarship in survey year	0.057 (0.232)	0.509*** (0.023)	0.504*** (0.023)	0.505*** (0.023)	0.526*** (0.019)	0.521*** (0.019)
Ever used a scholarship	0.243 (0.430)	0.672*** (0.021)	0.663*** (0.022)	0.662*** (0.022)	0.636*** (0.019)	0.635*** (0.019)
Started 6th grade in private	0.877 (0.328)	0.063*** (0.017)	0.057*** (0.017)	0.058*** (0.017)	0.066*** (0.016)	0.067*** (0.016)
Started 7th grade in private	0.673 (0.470)	0.174*** (0.025)	0.168*** (0.025)	0.171*** (0.024)	0.170*** (0.021)	0.173*** (0.021)
Currently in private school	0.539 (0.499)	0.160*** (0.028)	0.153*** (0.027)	0.156*** (0.027)	0.152*** (0.023)	0.154*** (0.023)
Highest grade completed	7.5 (0.960)	0.164*** (0.053)	0.130*** (0.051)	0.120*** (0.051)	0.085** (0.041)	0.078** (0.041)
Currently in school	0.831 (0.375)	0.019 (0.022)	0.007 (0.020)	0.007 (0.020)	-0.002 (0.016)	-0.002 (0.016)
Sample size	562		1,147		1,577	

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.

*** significant at 1% ** significant at 5% * significant at 10%

Educational Outcomes and Voucher Status (II)

Dependent variable	Bogotá 1995				Combined sample	
	Loser means (1)	No controls (2)	Basic controls (3)	Basic +19 barrio controls (4)	Basic controls (5)	Basic +19 barrio controls (6)
Finished 6th grade	0.943 (0.232)	0.026** (0.012)	0.023* (0.012)	0.021* (0.011)	0.014 (0.011)	0.012 (0.010)
Finished 7th grade (excludes Bogotá 97)	0.847 (0.360)	0.040** (0.020)	0.031 (0.019)	0.029 (0.019)	0.027 (0.018)	0.025 (0.018)
Finished 8th grade (excludes Bogotá 97)	0.632 (0.483)	0.112*** (0.027)	0.100*** (0.027)	0.094*** (0.027)	0.077*** (0.024)	0.074*** (0.024)
Repetitions of 6th grade	0.194 (0.454)	-0.066*** (0.024)	-0.059** (0.024)	-0.059** (0.024)	-0.049*** (0.019)	-0.049*** (0.019)
Ever repeated after lottery	0.224 (0.417)	-0.060*** (0.023)	-0.055** (0.023)	-0.051** (0.023)	-0.055*** (0.019)	-0.053*** (0.019)
Total repetitions since lottery	0.254 (0.508)	-0.073*** (0.028)	-0.067** (0.027)	-0.064** (0.027)	-0.058*** (0.022)	-0.057*** (0.022)
Years in school since lottery	3.7 (0.951)	0.058 (0.052)	0.034 (0.050)	0.031 (0.050)	0.015 (0.044)	0.012 (0.043)
Sample size	562		1,147			1,577

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.

*** significant at 1% ** significant at 5% * significant at 10%

Test Results

Variable	OLS results (1)	OLS results with covariates (2)	RE (3)	RE with covariates (4)	Sample size (5)
Total Points	0.217* (0.116)	0.205* (0.108)			282
Math scores	0.178 (0.120)	0.153 (0.114)			282
Reading scores	0.204* (0.115)	0.203* (0.114)			283
Writing scores	0.126 (0.116)	0.128 (0.105)			283
Pooled test scores			0.170* (0.095)	0.148* (0.088)	846
Math and reading scores			0.192* (0.101)	0.162* (0.096)	568

Robust standard errors are reported in parentheses.

*** significant at 1% ** significant at 5% * significant at 10%

Observational Studies

- ❑ Economists rarely have the opportunity to randomize variables like educational attainment, immigration or minimum wages. Normally, we must rely on observational studies.
- ❑ An *observational study* is an empiric investigation of treatments, policies, or exposures and the effects they cause.
- ❑ They differ from an experiment in that the investigator cannot control the assignment of treatment to subjects.

- Cochran (1965) defined an observational study as an empiric investigation in which:

«...The objective is to elucidate cause-and-effect relationships ... [in which] it is not feasible to use controlled experimentation, in the sense of being able to impose the procedures or treatments whose effects it is desired to discover, or to assign subjects at random to different procedures.»

References

- ❑ Rosenbaum, Paul (2002): Observational Studies, Springer. Chapter 2.
- ❑ W. G. Cochran (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.
- ❑ Angrsit, J. and V. Lavy (2002): “The Effect of High School Matriculation Awards: Evidence from Randomized Trials”, NBER Working Paper.