

Lecture IV: Experiments

Stanislao Maldonado

Universidad del Rosario
stanislao.maldonado@urosario.edu.co

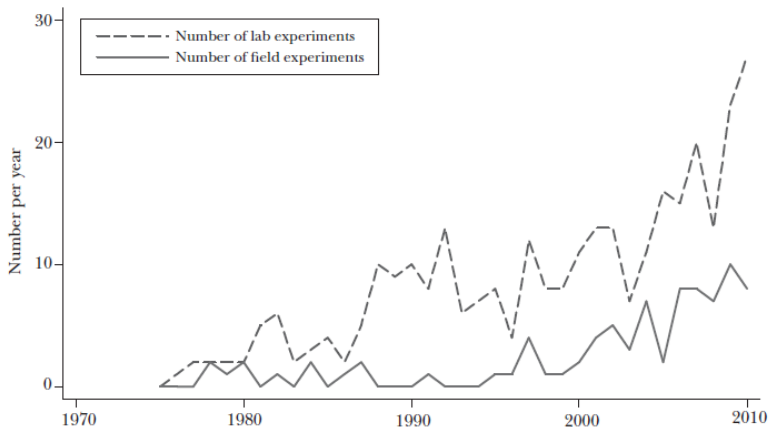
Impact Evaluation
Universidad del Rosario
February 14th, 2017

1. Randomized experiments in economics

- Experiments are increasingly used in several fields in economics (labor, economics of education, health economics, development, behavior, political economy, industrial organization, public economics, etc)
- Examples:
 - Effect of school inputs on learning (Glewwe and Kremer 2002)
 - Adoption of new technologies in agriculture (Duflo et al 2010)
 - Corruption in licenses (Bertrand et al 2006)
 - Moral hazard and adverse selection in consumer markets (Karlán et al 2005)
- Economics is becoming more experimental!

Figure 1

Number of Laboratory and Field Experiments Published in Five Top Economics Journals from 1975 to 2010



Note: The journals surveyed were the *American Economic Review*, *Econometrica*, the *Journal of Political Economy*, the *Quarterly Journal of Economics*, and the *Review of Economic Studies*.

- Other names:
 - Randomized assignment studies
 - Randomized controlled trials (RCT)
 - Randomized controlled experiments
 - Social experiments

Definition 1: Random assignment (Shadish et al 2002)

Any procedure that assigns units to treatment/control status based only on chance, in which each unit has a nonzero probability of being assigned to a treatment/control status.

- Random assignment is the same as random sampling?

- The answer is no!
 - Random sampling ensures that the selected sample is similar to a population
 - Random assignment makes samples of treatment and control units similar to each other
- Why randomization works?
 - Ensures alternative causes are not cofounded with treatment
 - Reduces plausibility of validity issues by distributing them randomly
 - Equates groups on the expected value of all pre-treatment characteristics
 - Allows the researcher to know and model the selection process
 - Allows the computation of a valid estimate of the error variance that it is also orthogonal to treatment

- Despite its power, random assignment is only one part of an experimental design
- A typical experiment involves (JPAL-MIT):
 - Design of the study
 - Random assign of units to treatment and control status
 - Collect baseline data
 - Verify randomization
 - Monitor the process to make sure that original design is not affected during the implementation
 - Collect follow-up data
 - Estimate the impacts of the treatment, assessing whether the impacts are statistically and practically significant

Typology of Randomized Experiments

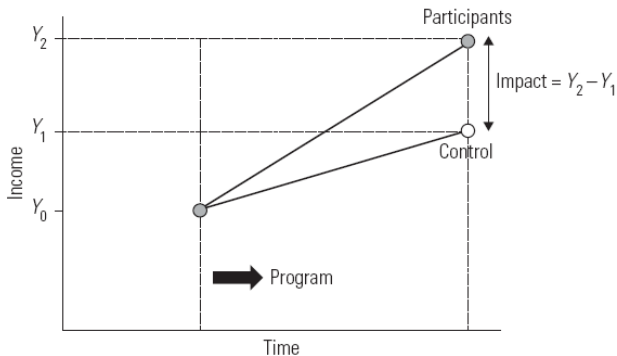
- Let's introduce some definitions:

Classical randomized experiment (Imbens and Rubin 2015)

A classical randomized experiment is a randomized experiment with an assignment mechanism that is individualistic, and unconfounded

- Following the definition of classical randomized experiment, 4 designs can be identified:
 - Bernoulli Trials
 - Completely Randomized Experiments
 - Stratified Randomized Experiments
 - Paired Randomized Experiments

Figure 3.1 The Ideal Experiment with an Equivalent Control Group



What can be randomized? (Glennester et al 2013)

- **Access:** We can choose which people will be offered the access to the program
- **Timing of access:** We can choose when to provide access to the program
- **Encouragement:** We can choose which people will be given encouragement to participate in the program

When randomize? (Shadish et al 2002 and Glennester et al 2013)

- **New program design:** When a problem has been identified but there is no agreement about what solution to implement
- **New programs:** When a program is new and being pilot-tested
- **New services:** When an existing program offers a new services
- **New people:** When a program is expanded to new areas
- **Oversubscription:** When there are more interested people than the program can serve
- **Undersubscription:** When not everyone who is eligible for the program takes it up

When randomize? (Shadish et al 2002 and Glennester et al 2013)

- **Rotation:** When the program's benefits or burden are to be shared by rotation
- **Admission cutoffs:** When the program has a merit cutoff and those just below the cutoff can be randomly admitted
- **Admission in phases:** When logistical and resource constraints means that not all the potential beneficiaries can be enrolled at one time and people can be randomly admitted in phases over time

When no randomize?

- When quick answers are needed
- When a great precision in estimating an effect is not needed
- When the treatment of interest cannot be manipulated
- When the contribution of the experiment to scientific/policy knowledge is expected to be low compared to its costs

2. Methods of Randomization (Duflo et al 2008)

- 1 Basic lottery (classic randomized designs)
- 2 Oversubscription
- 3 Randomized order of phase-in
- 4 Within-group randomization
- 5 Encouragement designs

A. Basic Lottery

- Simplest strategy to allocate treatment
- Typically involves the inclusion of a randomized strategy along the beginning of an intervention in which the sample was randomly allocated to treatment(s) and a comparison group
- Example: Duflo, Kremer and Robison (2006) on determinants of the adoption of a fertilizer in Western Kenya

B. Oversubscription

- Ideal when resources are limited and demand for a program exceeds supply
- Fair method to allocate resources
- Example: Angrist et al (2002)
 - Voucher program in Colombia (Programa de Ampliacion de Cobertura de la Educacion Secundaria-PACES)
 - Lottery for vouchers (50% cost) for private schools

C. Randomized order of phase-in

- Financial constraints forces to phase-in programs over time.
- Randomization is a fair way to decide which group will be treated first
- Facilitates cooperation of beneficiaries (reduces attrition)
- Example: Miguel and Kremer (2004) on deworming (75 schools in 3 phases, 25 each from 1998 to 2000)
- Problems: Prevent the estimation of long-run effects and expectation for future treatment may affect behavior

D. Within-group randomization

- Sometimes cooperation is compromised in randomized phase-in designs if control group units are not benefited in some way before being treated
- Characteristics of the experimental unit can be exploited to provide benefits to all participants without compromising the research design
- Example: Banerjee et al (2007)
 - Remedial education program using tutors
 - Each school was provided with a tutor but they were randomized to serve only one grade (3th or 4th).

E. Encouragement designs

- Allow evaluation of a program available in the entire study area but whose take up is not universal
- Randomizes encouragement to receive the treatment rather than the treatment itself
- Encouragement may serve as an IV for the treatment of interest

3. Experiments and Causal Effects

- Recall the key idea: Causal effects can be measured by randomly selecting individuals from a population and the randomly giving some of the individuals the treatment
- The effect of random assignment:

$$Y_i = \beta_0 + \beta_1 D_i + \mu_i \quad (1)$$

- We know:
 - If D is randomly assigned: $\mathbb{E}[\mu_i | D_i] = 0$
 - D is distributed independently of the omitted factor μ_i
 - Random assignment of D implies that the orthogonality condition holds

- Causal effect on Y on treatment level D (assuming a binary treatment variable):

$$\beta_1 = \mathbb{E}[Y_i/D = 1] - \mathbb{E}[Y_i/D = 0] \quad (2)$$

- We call this the **differences estimator** (DE):
 - Causal effect can be estimated by the difference in the sample average outcomes between the treatment and control groups
 - Equivalently: β can be estimated by OLS estimator b if treatment is randomly assigned

3.1 Potential problems with experiments

- There is no free lunch in economic research!
- Experiments have many advantages:
 - Less subject to methodological debates
 - Easier to convey
 - More convincing to policy-makers
- However, experiments may be subject to internal and external validity threats

Threats to internal validity

- Failure to randomize
- Failure to follow treatment protocol:
 - People don't do what they are asked to do (partial compliance)
- Attrition
 - Subjects dropping out of the study after being randomly assigned to treatment
- Experimental effects
 - Being in an experiment change behavior: treatment (Hawthorne effect), control (John Henry effect)
- Small sample

Threats to external validity

- Non representative sample

Population studied and the population of interest must be similar to justify generalizing results

- Non representative program or policy

Small-scale experiments can be quite different than the program/policy to be implemented

- General equilibrium effects

Turning a small and temporary small experimental program into a widespread and permanent program might change the economic environment

- Treatments vs. eligibility effects

Participation in an actual program is voluntary. A different effect should be expected

3.2 Regression estimators of causal effects using experimental data

- If treatment is randomly received:
 - Differences estimator is unbiased
 - But is this efficient?
- When experiment have some issues of internal validity, then the differences estimator is biased
- Controlling for observables can help to improve internal validity
- Differences Estimator with additional regressors (DER):

$$Y_i = \beta_0 + \beta_1 D_i + \beta_2 W_{1i} + \beta_3 W_{2i} + \dots + \beta_{r+1} W_{ri} + \mu_i \quad (3)$$

- What is the difference between a “treatment” and “control” variable?

- Conditional mean-zero assumption:

$$\mathbb{E}[\mu_i | D_i] = 0 \quad (4)$$

- Conditional mean independence assumption:

$$\mathbb{E}[\mu_i | D_i, W_{1i}, \dots, W_{ri}] = \alpha_0 + \alpha_1 W_{1i} + \dots + \alpha_r W_{ri} \quad (5)$$

- Conditional mean independence implies:

- μ can be correlated with W
 - Given W , μ does not depend on D

- When this assumption is true ?
 - When $\mathbb{E}[\mu_i|D_i] = 0$
 - D is randomly assigned
 - D is assigned randomly conditional on W
- Evaluating at $D = 1$ and $D = 0$:

$$\beta_1 = \mathbb{E}[Y_i/D = 1, W_{1i}, \dots, W_{ri}] - \mathbb{E}[Y_i/D = 0, W_{1i}, \dots, W_{ri}]$$

- W must reflect non-experimental or predetermined outcomes
- Reasons for using the DER:
 - Efficiency
 - Check for randomization
 - Adjust for “conditional” randomization

- Testing for randomization:

- **Testing for random receipt of treatment:** F-test for null hypothesis that treatment was received randomly

$$D_i = \alpha_0 + \alpha_1 W_{1i} + \dots + \alpha_r W_{ri} + v_i \quad (6)$$

- **Testing for random assignment:** F-test for null hypothesis that all the slope coefficients are zero

$$Z_i = \delta_0 + \delta_1 W_{1i} + \dots + \delta_r W_{ri} + v_i \quad (7)$$

3.3 Randomization does not justify regression

- What does randomization guarantee?
- Regression models assume a set of assumptions that are not linked to random assignment
- Freedman (2008):
 - DER may be biased but this bias tends to 0 as sample size increases
 - Asymptotically, DER may perform worse than DE
 - Nominal standard errors are severely biased
 - Although ATE estimator coincides with DE estimator, standard errors are different

4. Using experiments as a benchmark for evaluating non-experimental methods

- Experimental data can be exploited to assess the bias using non-experimental techniques
- Lalonde (1986) showed that many econometric procedures and comparison groups used in the literature provide estimates that are often far from experimental results
- Other studies:
 - Propensity score matching (Heckman et al 1997, Heckman et al 1998, Dehejia and Wahba 1999, Smith and Todd 2005, Diaz and Handa 2006, among others)
 - RDD (Budlemeyer and Skoufias 2003)
 - Matching (Abadie and Imbens 2006, McKenzie et al 2010, Arcenoux et al 2000)
 - Difference in differences (Glewwe et al 2004)
 - Instrumental variables (McKenzie et al 2010)

- Analyzes data from a randomized experiment evaluating a job training program, the National Supported Work Demonstration (NSW), to assess whether standard econometric procedures can reproduce experimental results
- To do so:
 - Construct alternative control groups from household surveys
 - Test standard methods: difference in differences, Heckman sample selection model, and IV
- He shows that experimental results cannot be replicable by using non-experimental techniques and control groups
- Experimental effect: 800-900 US\$

TABLE 4—EARNINGS COMPARISONS AND ESTIMATED TRAINING EFFECTS FOR THE NSW AFDC PARTICIPANTS USING COMPARISON GROUPS FROM THE *PSID* AND THE *CPS-SSA*^{a,b}

Name of Comparison Group ^d	Comparison Group Earnings Growth 1975–79 (1)	NSW Treatment Earnings Less Comparison Group Earnings				Difference in Differences: Difference in Earnings Growth 1975–79 Treatments Less Comparisons		Unrestricted Difference in Differences: Quasi Difference in Earnings Growth 1975–79		Controlling for All Observed Variables and Pre-Training Earnings	
		Pre-Training Year, 1975		Post-Training Year, 1979		Without Age (6)	With Age (7)	Unad-justed (8)	Ad-justed ^c (9)	Without AFDC (10)	With AFDC (11)
		Unad-justed (2)	Ad-justed ^c (3)	Unad-justed (4)	Ad-justed ^c (5)						
Controls	2,942 (220)	–17 (122)	–22 (122)	851 (307)	861 (306)	833 (323)	883 (323)	843 (308)	864 (306)	854 (312)	–
<i>PSID</i> -1	713 (210)	–6,443 (326)	–4,882 (336)	–3,357 (403)	–2,143 (425)	3,097 (317)	2,657 (333)	1746 (357)	1,354 (380)	1664 (409)	2,097 (491)
<i>PSID</i> -2	1,242 (314)	–1,467 (216)	–1,515 (224)	1,090 (468)	870 (484)	2,568 (473)	2,392 (481)	1,764 (472)	1,535 (487)	1,826 (537)	–
<i>PSID</i> -3	665 (351)	–77 (202)	–100 (208)	3,057 (532)	2,915 (543)	3,145 (557)	3,020 (563)	3,070 (531)	2,930 (543)	2,919 (592)	–
<i>PSID</i> -4	928 (311)	–5,694 (306)	–4,976 (323)	–2,822 (460)	–2,268 (491)	2,883 (417)	2,655 (434)	1,184 (483)	950 (503)	1,406 (542)	2,146 (652)
<i>CPS-SSA</i> -1	233 (64)	–6,928 (272)	–5,813 (309)	–3,363 (320)	–2,650 (365)	3,578 (280)	3,501 (282)	1,214 (272)	1,127 (309)	536 (349)	1,041 (503)
<i>CPS-SSA</i> -2	1,595 (360)	–2,888 (204)	–2,332 (256)	–683 (428)	–240 (536)	2,215 (438)	2,068 (446)	447 (468)	620 (554)	665 (651)	–
<i>CPS-SSA</i> -3	1,207 (166)	–3,715 (226)	–3,150 (325)	–1,122 (311)	–812 (452)	2,603 (307)	2,615 (328)	814 (305)	784 (429)	–99 (481)	1,246 (720)
<i>CPS-SSA</i> -4	1,684 (524)	–1,189 (249)	–780 (283)	926 (630)	756 (716)	2,126 (654)	1,833 (663)	1,222 (637)	952 (717)	827 (814)	–

^aThe columns above present the estimated training effect for each econometric model and comparison group. The dependent variable is earnings in 1979. Based on the experimental data, an unbiased estimate of the impact of training presented in col. 4 is \$851. The first three columns present the difference between each comparison group's 1975 and 1979 earnings and the difference between the pre-training earnings of each comparison group and the NSW treatments.

^bEstimates are in 1982 dollars. The numbers in parentheses are the standard errors.

^cThe exogenous variables used in the regression adjusted equations are age, age squared, years of schooling, high school dropout status, and race.

^dSee Table 2 for definitions of the comparison groups.

TABLE 4—EARNINGS COMPARISONS AND ESTIMATED TRAINING EFFECTS FOR THE NSW AFDC PARTICIPANTS USING COMPARISON GROUPS FROM THE *PSID* AND THE *CPS-SSA*^{a,b}

Name of Comparison Group ^d	Comparison Group Earnings Growth 1975–79 (1)	NSW Treatment Earnings Less Comparison Group Earnings				Differences: Difference in Earnings Growth 1975–79 Treatments Less Comparisons		Unrestricted Difference in Earnings Growth 1975–79 Quasi Difference in Earnings Growth 1975–79		Controlling for All Observed Variables and Pre-Training Earnings	
		Pre-Training Year, 1975		Post-Training Year, 1979		Without Age (6)	With Age (7)	Unadjusted (8)	Adjusted ^c (9)	Without AFDC (10)	With AFDC (11)
		Unadjusted (2)	Adjusted ^c (3)	Unadjusted (4)	Adjusted ^c (5)						
Controls	2,942 (220)	–17 (122)	–22 (122)	851 (307)	861 (306)	833 (323)	883 (323)	843 (308)	864 (306)	854 (312)	–
<i>PSID</i> -1	713 (210)	–6,443 (326)	–4,882 (336)	–3,357 (403)	–2,143 (425)	3,097 (317)	2,657 (333)	1746 (357)	1,354 (380)	1664 (409)	2,097 (491)
<i>PSID</i> -2	1,242 (314)	–1,467 (216)	–1,515 (224)	1,090 (468)	870 (484)	2,568 (473)	2,392 (481)	1,764 (472)	1,535 (487)	1,826 (537)	–
<i>PSID</i> -3	665 (351)	–77 (202)	–100 (208)	3,057 (532)	2,915 (543)	3,145 (557)	3,020 (563)	3,070 (531)	2,930 (543)	2,919 (592)	–
<i>PSID</i> -4	928 (311)	–5,694 (306)	–4,976 (323)	–2,822 (460)	–2,268 (491)	2,883 (417)	2,655 (434)	1,184 (483)	950 (503)	1,406 (542)	2,146 (652)
<i>CPS-SSA</i> -1	233 (64)	–6,928 (272)	–5,813 (309)	–3,363 (320)	–2,650 (365)	3,578 (280)	3,501 (282)	1,214 (272)	1,127 (309)	536 (349)	1,041 (503)
<i>CPS-SSA</i> -2	1,595 (360)	–2,888 (204)	–2,332 (256)	–683 (428)	–240 (536)	2,215 (438)	2,068 (446)	447 (468)	620 (554)	665 (651)	–
<i>CPS-SSA</i> -3	1,207 (166)	–3,715 (226)	–3,150 (325)	–1,122 (311)	–812 (452)	2,603 (307)	2,615 (328)	814 (305)	784 (429)	–99 (481)	1,246 (720)
<i>CPS-SSA</i> -4	1,684 (524)	–1,189 (249)	–780 (283)	926 (630)	756 (716)	2,126 (654)	1,833 (663)	1,222 (637)	952 (717)	827 (814)	–

^aThe columns above present the estimated training effect for each econometric model and comparison group. The dependent variable is earnings in 1979. Based on the experimental data, an unbiased estimate of the impact of training presented in col. 4 is \$851. The first three columns present the difference between each comparison group's 1975 and 1979 earnings and the difference between the pre-training earnings of each comparison group and the NSW treatments.

^bEstimates are in 1982 dollars. The numbers in parentheses are the standard errors.

^cThe exogenous variables used in the regression adjusted equations are age, age squared, years of schooling, high school dropout status, and race.

^dSee Table 2 for definitions of the comparison groups.

- How much do migrants stand to gain in income from moving across borders?
- Selection: income differences may be due to unobserved differences in ability, skills, motivation, etc.
- This paper uses experimental data (random selection of immigrants) from the Pacific Access Category (PAC):
 - PAC allows Tongans to participate in a visa lottery to migrate permanently to New Zealand
 - Survey to winners and losers + data about non-aplicants
- They use experimental data to study performance of non-experimental methods: first differences, OLS, DD, matching and IV

TABLE 1. Test for randomization.

	Sample Means Applicants		<i>t</i> -test of Equality of Means <i>p</i> -value
	Successful Ballots	Unsuccessful Ballots	
Age	33.6	33.9	0.80
Years of schooling	11.9	11.6	0.57
Proportion male	0.55	0.51	0.58
Proportion born on Tongatapu	0.75	0.79	0.51
Proportion who are married	0.60	0.65	0.52
Height of males	175.8	173.8	0.22
Height of females	166.4	167.3	0.57
Past income (NZ dollars per week)	80.9	76.6	0.74
Proportion with the following family members living in NZ at time of last application:			
Father/Father-in-law	0.38	0.44	0.45
Mother/Mother-in-law	0.40	0.35	0.46
Brother/Brother-in-law	0.72	0.71	0.78
Sister/Sister-in-law	0.64	0.60	0.63
Aunt or Uncle	0.65	0.55	0.17
Total Sample Size	120	78	

Note: Comparison of ex ante characteristics of principal applicants in successful and unsuccessful ballots.

TABLE 2. Experimental estimates of the income gain from migration.

	(1) OLS	(2) OLS	(3) OLS	(4) IV	(5) IV	(6) IV
Ballot success dummy	90.63*** (24.6)	89.74*** (24.2)	87.39*** (22.5)			
Male dummy		-29.07 (24.4)	-23.86 (22.0)		-33.10 (23.1)	-27.77 (20.9)
Married dummy		-4.493 (27.4)	24.54 (23.3)		-10.69 (26.6)	18.38 (22.4)
Age dummy		0.558 (1.64)	-0.886 (1.26)		0.987 (1.54)	-0.462 (1.13)
Years of education		13.43** (6.60)	4.605 (3.89)		12.03** (6.05)	3.274 (3.58)
Born on Tongatapu dummy		29.17 (18.9)	27.60* (14.7)		29.59 (18.0)	28.01** (13.7)
Height		1.281* (0.65)	0.381 (0.42)		1.249** (0.61)	0.353 (0.38)
Past income			0.662*** (0.095)			0.660*** (0.090)
Migration dummy				274.0*** (61.5)	281.1*** (61.6)	273.7*** (54.9)
Constant	104.1*** (11.8)	-297.9** (121)	-60.42 (81.5)	104.1*** (11.7)	-285.0** (116)	-48.59 (73.8)
First stage <i>F</i> -statistic on instrument				66.53	61.88	61.51
Observations	197	191	190	197	191	190
<i>R</i> -squared	0.04	0.14	0.27			

Notes: Robust standard errors in parentheses. *Significance at 10%; **Significance at 5%; *** Significance at 1%.
Dependent variable: Weekly income from work in New Zealand dollars.

TABLE 4. Non-experimental estimates.

Method	Sample Size	Estimate	s.e.	Percent Difference Compared to Experiment	Percent of Bootstrap Replications Where This is Greater than Experimental Estimate	Bootstrap 90% Confidence Interval for Difference (Non-experiment-experiment)
1) Using pre-migration income as the counterfactual	63	341.3	46.4	24.6	92.7	[−11.0, 150.2]
2) Selection on Observables: OLS regression						
PINZMS regressions including height						
Base specification	230	360.0	41.2	31.4	97.6	[17.9, 150.7]
Polynomials	230	347.5	42.5	26.8	95.2	[1.9, 148.5]
PINZMS regressions excluding height						
Base specification	244	362.6	40.9	32.3	98.3	[19.0, 153.7]
Polynomials	244	357.2	42.7	30.4	96.8	[11.2, 158.3]
TLFS regressions						
Base specification	4,043	368.9	43.9	34.6	99.5	[28.4, 158.6]
Polynomials	4,043	358.2	44.0	30.7	98.6	[15.9, 148.0]
3) Difference-in-Difference Regression						
Base specification	219	330.5	42.2	20.6	90.8	[−12.9, 120.8]
Polynomials	219	333.9	43.6	21.9	86.9	[−22.5, 130.8]
4) Instrumental variables using migrant network	219	497.5	233.7	81.6	86.6	[−154.4, 750.0]
5) Instrumental variables using log distance to NZIS office						
All islands	219	309.0	89.5	12.8	67.9	[−121.6, 201.9]
Tongatapu only	159	277.1	90.0	1.1	54.3	[−165.9, 147.8]

TABLE 5. Matching estimates.

	Sample Size	Estimate	s.e.	Percent Difference Compared to Experiment	Percent of Bootstrap Replications Where This is Greater Than Experimental Estimate	Bootstrap 90% Confidence Interval for Difference (Non-experiment-experiment)
<i>Abadie and Imbens (2006) nearest-neighbor matching on multiple covariates</i>						
<i>Single Nearest Neighbor Matching</i>						
A: Matching without using past income						
PINZMS sample using height as a control						
ATT	230 ^a	364.0	52.8	32.9	98.7	[21.0, 155.9]
bias-adjusted ATT	230 ^a	349.6	54.0	27.6	96.5	[5.4, 139.6]
Bias-adjusted ATT with interactions						
PINZMS excluding height	244 ^a	334.1	54.9	21.9	90.0	[-18.1, 136.7]
TLFS	4,043 ^a	359.2	46.7	31.1	98.6	[15.5, 147.2]
B: Matching using past income						
ATT	219 ^a	345.9	56.7	26.2	96.6	[4.8, 138.9]
bias-adjusted ATT	219 ^a	334.8	58.4	22.2	91.5	[-8.8, 127.0]
C: Matching using past income and interactions						
ATT	219 ^a	342.2	59.5	24.9	96.3	[3.5, 135.8]
bias-adjusted ATT	219 ^a	330.1	59.0	20.5	85.7	[-29.2, 141.6]
<i>Multiple Nearest Neighbor Matching</i>						
D: Bias-adjusted ATT, using past income and interactions						
Nearest 2 neighbors	219 ^a	330.6	56.0	20.7	90.4	[-14.1, 123.2]
Nearest 5 neighbors	219 ^a	328.6	54.5	19.9	90.9	[-15.5, 123.0]
Nearest 10 neighbors	219 ^a	330.4	53.3	20.6	90.7	[-13.9, 122.3]

Propensity Score Matching

Radius Matching on the Common Support

Radius 0.10	156	352.3	49.5	28.6	n.a.	n.a.
Radius 0.05	150	371.5	52.7	35.6	n.a.	n.a.

Kernel Matching on the Common Support

Gaussian kernel	163	329.6	47.9	20.3	92.1	[−11.2, 124.2]
Epanechnikov kernel	163	344.8	58.7	25.8	91.7	[−12.2, 142.1]

Kernel Matching on Trimmed Samples

ATT trimming 0.01 and 0.99

PINZMS	217	336.5	41.2	22.8	93.9	[−4.4, 132.0]
TLFS	1,385	367.7	42.1	34.2	98.7	[21.9, 154.9]

ATT trimming 0.05 and 0.95

PINZMS	200	329.2	49.5	20.1	93.8	[−3.2, 139.9]
TLFS	354	344.5	61.9	25.7	91.2	[−16.5, 186.3]

Note: All matching includes gender, age, marital status, years of education, and place of birth. Height is also included in all of the PINZMS matches except where indicated otherwise. Interactions are interactions of sex with each covariate, quartics in age and years schooling, and an interaction between age and schooling. n.a. = not available, since the Becker and Ichimura radius matching program (attr.do) would not find matches on certain bootstrap replications. ^aSample sizes for Nearest Neighbor Matching show the sample size over which nearest neighbors were sought. When only a single nearest neighbor is used for each migrant, the actual number of controls used is approximately equal to the number of migrants (less due to the same control being used as a match for multiple migrants, and more when there are ties in the closest match).