CCT's Long-Term Effects

Econ 8784

More Schooling and More Learning? Effects of a Three-Year Conditional Cash Transfer Program in Nicaragua After 10 years Barham, Macours, Maluccio 2013

- Objective: Effect of Nicaraguan CCT, Red de Proteccion Social (RPS), on education and learning, and labor market outcomes
- What is the motivation for the paper?
  - Little known about longer-term effects of CCTs
    - But important since long-term poverty reduction through human capital accumulation is a goal of CCTs
    - Mexican CCT:15-21 year old - increase in education, no effect on achievement, decreased labor market participation (Behrman, Parker, Todd 2009, 2011)
  - Does increased schooling lead to increased learning and better jobs?
    - Millennium Development Goals: led to increase enrollment but concern as to whether students are learning (Beatty & Pritchett, 2012)

Nicaraguan CCT program

- Fixed 3-year transfer period (no re-certification)
- Two types of transfers to household:
  - Education: children aged 7-13 and < grade 4 completed
    - enrolled in school and attend 85% of the time
  - Food security transfer: all households receive
    - Children <5: regular preventative health visits
    - Mothers: attend health education workshops
- Social marketing on nutrition and education
- Transfer given to mothers every other month
- Size of transfer: ~18% pre-program household expenditures
- Conditions were enforced, take-up ~85%
- WHAT MAKES THIS SET UP DIFFERENT FROM PROGRESA ?

Randomized Design

- Stratified randomized intervention at locality level
  - 42 localities in 6 rural municipalities
  - **WHAT DOES THIS MEAN FOR HETEROGENEITY ANALYSIS?**
- Randomized phase-in of 42 localities (21 each group)
  - Early treatment group 2000-2003, then transfers stop
  - Late treatment group 2003-2005, then program ends

Data: 2010 Follow-up Survey

- Sample:
  - Original baseline household survey (1764 households)
  - Drawn from 2000 census of program areas (source of control variables)
  - Oversample of children approximately age 11 at the start of the program from baseline census (502 households)
- Survey:
  - Household and individual-level surveys
  - Cognitive and achievement tests done in the household
  - Only for <13 at the start of the program (< 23 in 2010)
- Extensive tracking:
  - Nicaragua and Costa Rica
  - Attrition: Household < 7%; Education 12%, Tests 19%
  - No significant difference in attrition rate between groups
    - For those with tests: Coef. 0.009, P-value 0.77

Other data

- Survey data for randomized evaluation
  - 2000, 2001, 2002, 2004
  - Issues of attrition
- In 2004, a non-random comparison group was added so could try to evaluate absolute effects of the program

Identification Strategy

- Select cohort of children more likely to benefit from transfers in early treatment group than the late treatment group
- Exploit program/schooling features to determine a cohort with large potential differences in years the education transfers received
  1. Randomized phase-in (timing)
  2. Age-specific conditionalities (7-13 years old)
  3. Fixed 3-year duration of transfers
  4. Age at which children are at risk of dropout.

What age groups are focused on and why?

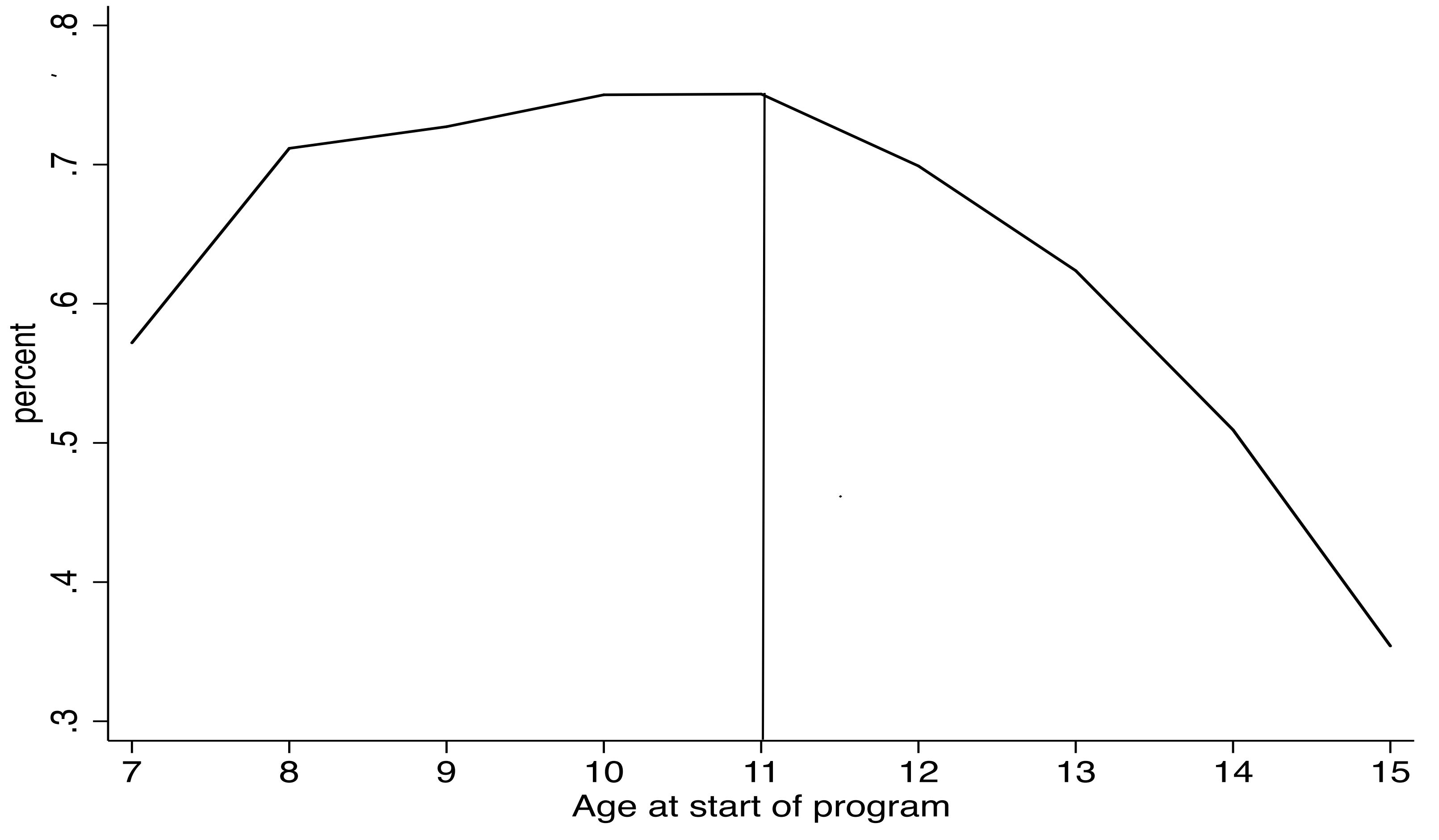
Years Eligible for Education Transfer

Age at Start of Program	Years of Eligible for Educ. Transfer		Difference
	Early Treat.	Late Treat.	
9	3	2	1
10	3	1	2
11	**3**	**0**	**3**
12	2	0	2
13	1	0	1
14	0	0	0

- Difference based on age eligibility, ignoring fact may not have completed grade 4.

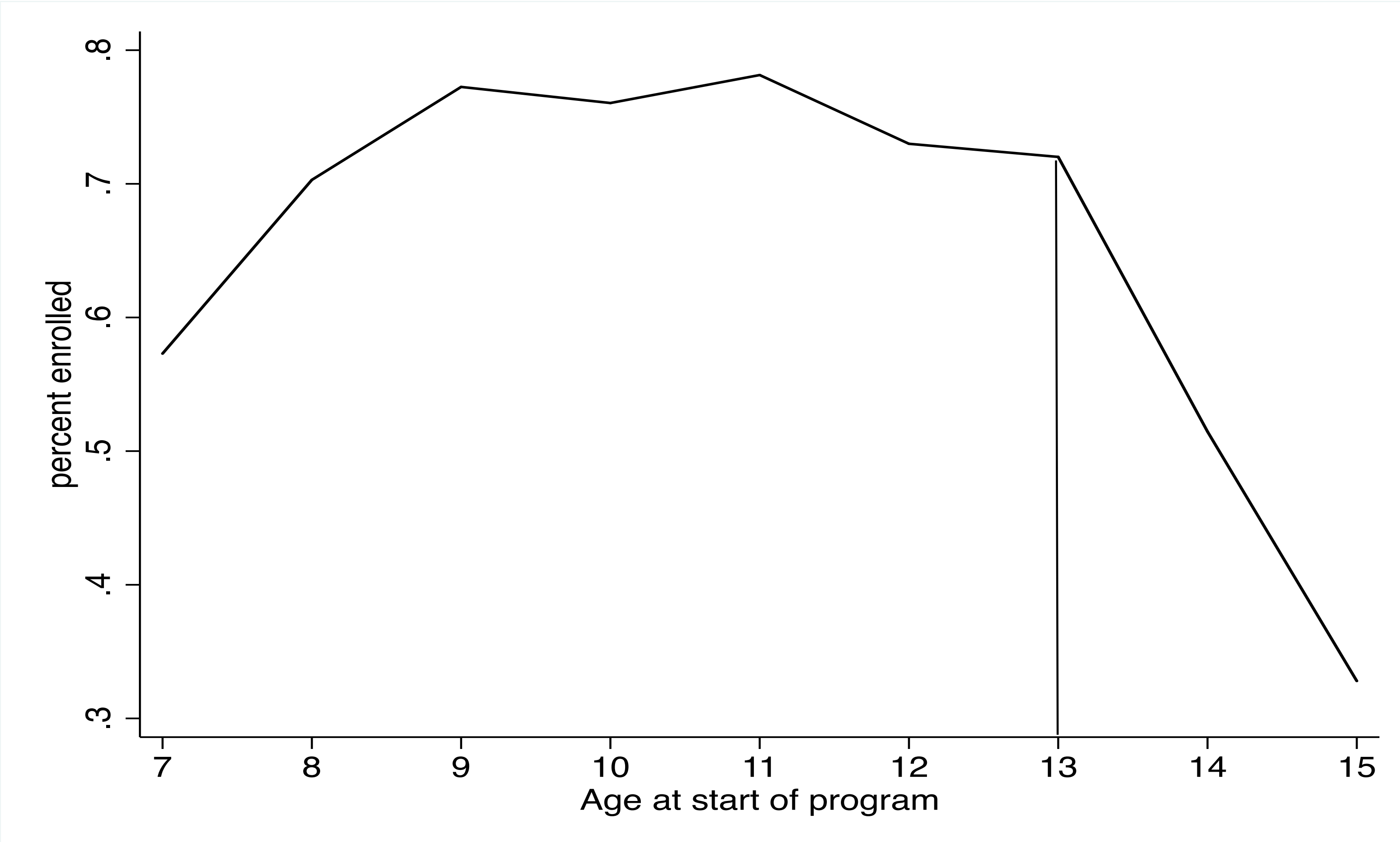
Mean Enrollment Rate in 2000 Boys

Focus 9-12 year olds



Mean Enrollment Rate in 2000 Girls

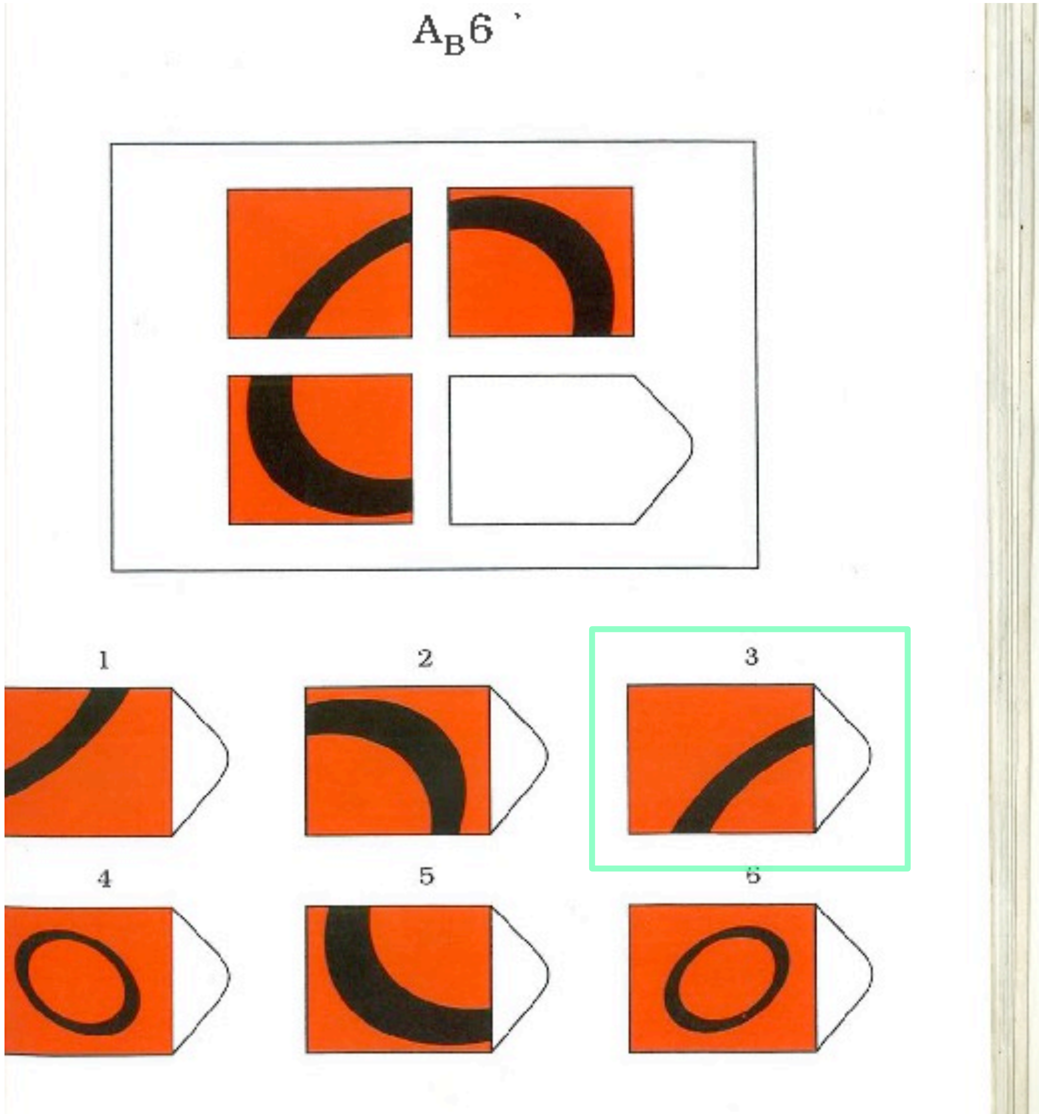
Focus 13-14 year olds – tests not available



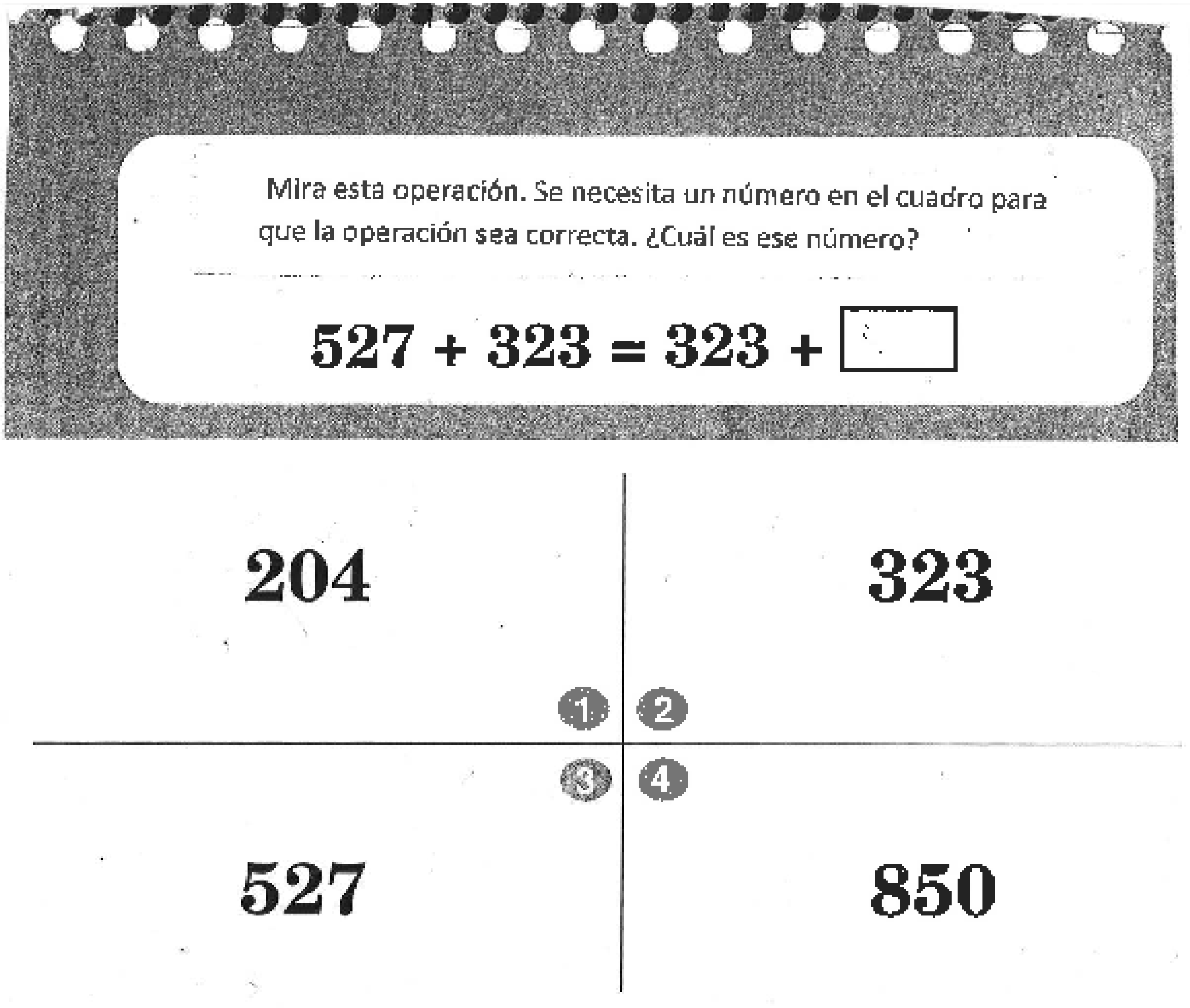
Tests in 2010 Survey

- Achievement tests
  - Woodcock-Muñoz: **spelling, reading fluency, word identification**
  - Woodcock-Muñoz: 3-minute math – **math fluency**
  - Peabody Individual Achievement Test (PIAT) – **math problems**
- Mixed Cognitive and Achievement Tests
  - TVIP (PPVT-Peabody Picture Vocabulary Test) – **receptive vocabulary**
  - Digit Span (backwards & forwards) – **memory math**
- Cognitive Tests: Raven Colored Progressive Matrices
- All converted to z-scores (internally standardized on late treatment for across test comparison)

Raven Colored Progressive Matrices (A, AB, B)



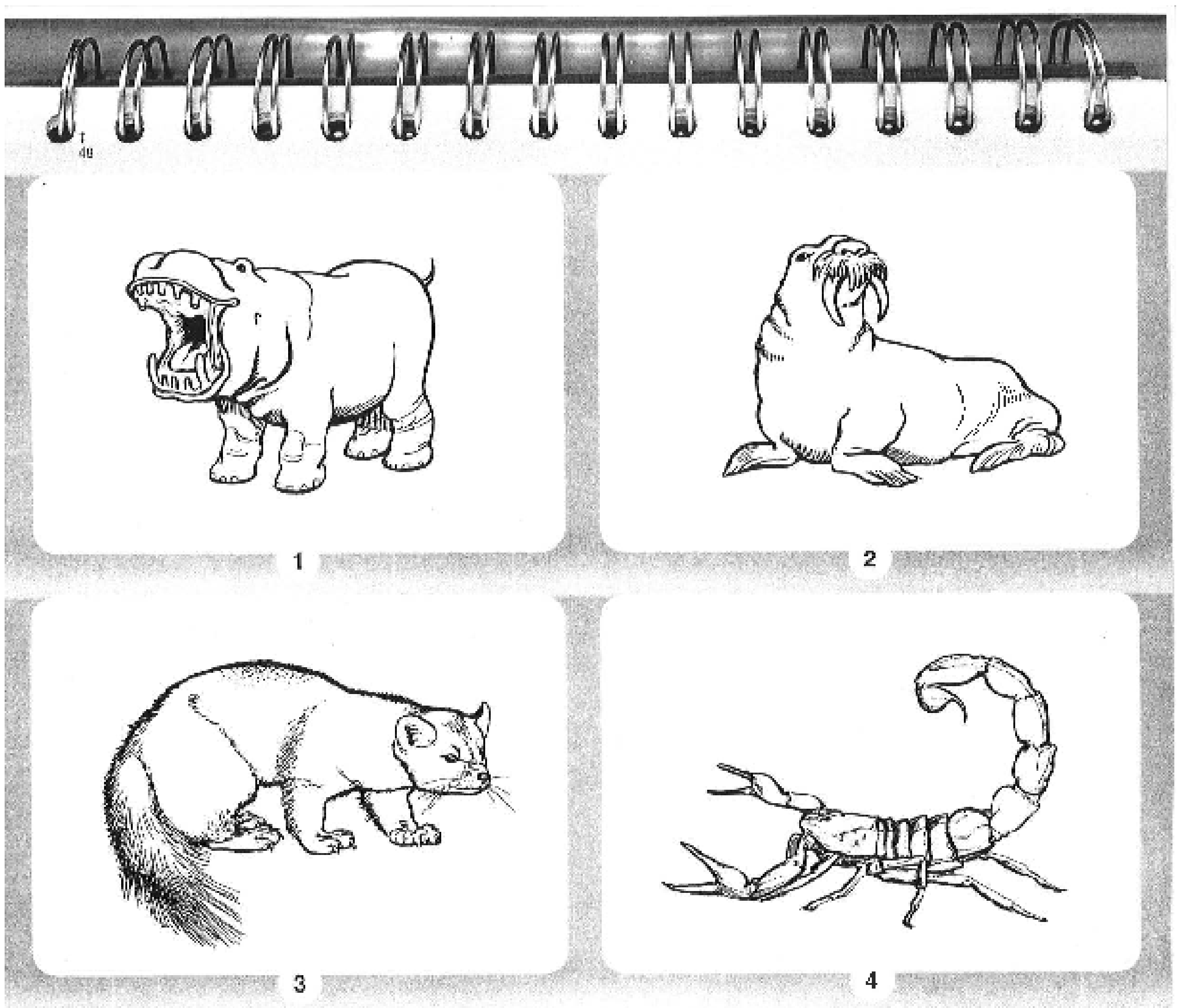
Peabody Individual Achievement Test-Revised (PIAT-R vol. III-Math)


Mira esta operación. Se necesita un número en el cuadro para que la operación sea correcta. ¿Cuál es ese número?

$$527 + 323 = 323 + \boxed{}$$

204 ①	② 323
527 ③	④ 850

Peabody Picture Vocabulary Test (Spanish version – TVIP)



ITT Empirical specification

Individual outcomes

$$Y_k = \alpha_k T + \beta_k X + \varepsilon_k$$

Y : Internally standardized outcomes – what does this mean?

T : Intent-to-treat indicator=1 if early treatment/0 control

X : Strata and 3 month date of birth dummies, individual and household level baseline controls (e.g. grades attained)

Standard errors clustered at the locality level

Family of outcomes (SURE)

$$\bar{\alpha} = \frac{1}{K} \sum_{k=1}^{K} \hat{\alpha}_k$$

Kling, Leibman and Katz (2007)

Absolute vs Differential effects

- Why do we get absolute effects in 2002, but differential effects in 2004, and 2010?
  - 2002, the control group did not phase-in. So could see total effect of the program.
  - 2004, the control has phased in, so it is the differential effect between early and late treated groups.

Short-Term Educational Attainment

TABLE 1: 2002 AND 2004 EXPERIMENTAL IMPACTS ON EDUCATION

	Grades Attained	Completed Grade 4 =1	Enrolled =1	Attended School More Than 85% of Time =1	Read and Write =1
	(1)	(2)	(3)	(4)	(5)
Panel A: 2002 — Absolute Effects					
ITT	0.361***	0.053*	0.182***	0.360***	0.150***
	(0.094)	(0.031)	(0.042)	(0.055)	(0.034)
N	475	475	475	475	475
R^2	0.828	0.747	0.191	0.271	0.324
Mean late treatment	2.396	0.277	0.733	0.544	0.735
Panel B: 2004 — Differential Effects					
ITT	0.487***	0.086*	-0.049	-0.100	0.124***
	(0.155)	(0.045)	(0.063)	(0.066)	(0.029)
N	458	458	458	458	458
R^2	0.598	0.467	0.262	0.239	0.241
Mean late treatment	3.585	0.536	0.626	0.564	0.815

- Why does it go negative in Panel B for enrolled?
- Notice the mean for the late treatment group what story does it tell?

Table 1 Questions

What are the absolute effects on Education in 2002:

- Grades attainted?
  - ITT effect on education of 0.361 grades
- What is the percent increase?
  - Divide 0.361 by the late treatment mean of 2.396 *100=15%
- Completing grade 4?
  - Completed education increase of 5.3 percentage points?
  - Percent increase is .053. /.277 = 19%
- Can read and write?
  - Increase of 15 percentage points or 20 percent

2010 Educational Results

TABLE 2: 2010 DIFFERENTIAL EXPERIMENTAL IMPACTS FOR EDUCATION FAMILY AND LITERACY

	Education Family Z-Score	Education Family Components			Read and Write =1
		Grades Attained	Completed Grade 4 =1	Enrolled =1	
	(1)	(2)	(3)	(4)	(5)
ITT	0.098**	0.288*	0.035	0.045**	0.052**
	(0.043)	(0.167)	(0.024)	(0.021)	(0.021)
N	1,007	1,006	1,006	1,005	1007
R^2	0.346	0.425	0.356	0.096	0.175
Mean late treatment	-0.026	5.498	0.747	0.181	0.874

Notes: *** $p<0.01$, ** $p<0.05$, * $p<0.10$. Standard errors are clustered at the locality level and given in parentheses. Regressions are weighted to account for sampling and attrition providing population estimates. Differential ITT results compare early to late treatment groups in 2010. The family z-score averages the z-scores for the individual components and is calculated if any component is available. Z-scores are calculated using the mean and standard deviation of the late treatment group. Controls include three-month age group indicators and indicator variables for grades attained at baseline, stratification, and region. All variables measured using the 2010 household instrument.

- Why do we want to look at the results for the family?

Table 2 Questions

- Interpret the program effect on the education family?
  - The program led to a .098 SD increase in education outcomes as a whole.
- What unit are the components of the family in?
  - Not in SD, in their actual units. They are put into SD for the family measure.
- What is the ITT effect on grades attained in 2010
  - Program led to an incease in education of 0.288 years. This represents a 0.288/5.498= 5 percent increase over the late treatment group mean.

2010 Test Results

Boys 9-12 year olds

TABLE 3: 2010 DIFFERENTIAL EXPERIMENTAL IMPACTS FOR LEARNING AND COGNITION FAMILIES (Z-SCORES)

	Learning			Mixed Cognition and Learning	Cognition (Raven)
	Math and Spanish	Math	Spanish		
	(1)	(2)	(3)	(4)	(5)
ITT	0.183**	0.160**	0.204**	0.113	-0.016
	(0.070)	(0.069)	(0.081)	(0.082)	(0.095)
N	907	905	907	906	906
R^2	0.448	0.395	0.437	0.380	0.215

Notes: *** p<0.01, ** p<0.05, * p<0.10. Standard errors are clustered at the locality level and given in parentheses. Regressions are weighted to account for sampling and attrition providing population estimates. Differential ITT results compare early to late treatment groups in 2010. Z-scores are calculated using the mean and standard deviation of the late treatment group. Controls include three-month age group indicators and indicator variables for grades attained at baseline, stratification, and region. All variables measured using the 2010 individual instrument.

Learning and Cognition Family

TABLE 4: 2010 DIFFERENTIAL EXPERIMENTAL IMPACTS FOR LEARNING AND COGNITION BY TEST (Z-SCORES)

	Learning Family Components					Mixed Cognition and Learning Family Components	
	Math Fluency	Math Problems	Reading Fluency	Spelling	Word Identification	Receptive Vocabulary	Memory Math
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ITT	0.183**	0.179**	0.137**	0.252***	0.206**	0.128	0.094
	(0.070)	(0.082)	(0.067)	(0.078)	(0.087)	(0.102)	(0.074)
N	907	904	904	898	905	906	902
R^2	0.448	0.349	0.346	0.425	0.358	0.353	0.291

Notes: *** p<0.01, ** p<0.05, * p<0.10. Standard errors are clustered at the locality level and given in parentheses. Regressions are weighted to account for sampling and attrition providing population estimates. Differential ITT results compare early to late treatment groups in 2010. The family z-score is calculated by averaging the z-score for the individual components where the average is determined even if one component is missing. Z-scores are calculated using the mean and standard deviation of the late treatment group. Controls include three-month age group indicators and indicator variables for grades attained at baseline, stratification, and region. All variables measured

Participation in Economic Activities

Past 12 months - Boys 9-12 year olds

TABLE 5: 2010 DIFFERENTIAL EXPERIMENTAL IMPACTS FOR LABOR MARKET PARTICIPATION, AND MIGRATION

	Labor Market Participation Family Z-Score	Labor Market Participation Family Components				Permanent Migration Out of Municipality =1
		Worked Off-Farm =1 (last 12 months)	Migrated for Work =1 (last 12 months)	Ever Had a Salaried Non-Agricultural Job =1	Ever Worked in Urban Area =1	
	(1)	(2)	(3)	(4)	(5)	(6)
ITT	0.272***	0.062***	0.093***	0.084**	0.065*	-0.019
	(0.075)	(0.022)	(0.032)	(0.036)	(0.034)	(0.028)
N	1,006	1,006	1,006	998	998	1,007
R^2	0.146	0.074	0.169	0.132	0.115	0.102
Mean late treatment	-0.018	0.828	0.312	0.226	0.127	0.150

Notes: *** p<0.01, ** p<0.05, * p<0.10. Standard errors are clustered at the locality level and given in parentheses. Regressions are weighted to account for sampling and attrition providing population estimates. Differential ITT results compare early to late treatment groups in 2010. The family z-score is calculated by averaging the z-score for the individual components where the average is determined even if one component is missing. Z-scores are calculated using the mean and standard deviation of the late treatment group. Controls include three-month age group indicators and indicator variables for grades attained at baseline, stratification, and region. All variables measured using the 2010 household instrument.

Impact on Earnings

TABLE 6: 2010 DIFFERENTIAL EXPERIMENTAL IMPACTS FOR EARNINGS FAMILY AND COMPONENTS

	Family Z-Score	Earnings Family Components (C$)			
		Earnings Per Month Worked (last 12 months)	Annual Earning (last 12 months)	Maximum Earnings (last 12 months)	Maximum Salary Ever
	(1)	(2)	(3)	(4)	(5)
Panel A: Rank of Earnings					
ITT	0.194***	41.780**	25.568	49.313**	43.899**
	(0.057)	(19.471)	(18.493)	(19.684)	(19.290)
N	1,006	1,006	1,006	1,006	998
R^2	0.097	0.082	0.095	0.094	0.094
Mean late treatment		497	503	498	486.9
Panel B: Earnings — Five Percent Trim					
ITT	0.192***	201.152***	595.013	211.421***	142.260*
	(0.067)	(63.624)	(619.322)	(69.318)	(71.919)
N	997	956	956	956	955
R^2	0.097	0.085	0.084	0.107	0.071
Mean late treatment		1436	8222	1619	228

Impact on Social-Emotional Outcomes

TABLE 8: 2010 DIFFERENTIAL EXPERIMENTAL IMPACTS FOR SOCIO-EMOTIONAL FAMILY OUTCOMES

	Family Z-Score	Socio-emotional Family Components			
		Positive Self Evaluation	Optimism	Stress	Negative Self Evaluation
	(1)	(2)	(3)	(4)	(5)
ITT	0.053	0.249**	0.287***	0.170**	0.155*
	(0.039)	(0.093)	(0.078)	(0.071)	(0.086)
N	900	900	900	900	900
R^2	0.152	0.194	0.180	0.097	0.106

Notes: *** p<0.01, ** p<0.05, * p<0.10. Standard errors are clustered at the locality level and given in parentheses. Regressions are weighted to account for sampling and attrition providing population estimates. Differential ITT results compare early to late treatment groups in 2010. Socio-emotional components are the first four factors resulting from exploratory factor analysis of all socio-emotional questions. The family z-score is calculated by averaging the z-score for the individual components. Z-scores are calculated using the mean and standard deviation of the late treatment group. Controls include three-month age group indicators and indicator variables for grades attained at baseline, stratification, and region. All variables measured using the 2010 individual instrument.

- Why is the family not big or statistically significant?

Heterogeneity by Strata

TABLE 9: 2010 DIFFERENTIAL EXPERIMENTAL IMPACTS FOR ALL FAMILIES BY STRATA

	Education		Learning Family Z-Score	Labor Market Participation Family Z-Score	Earnings Family Z-Score		Socio-emotional Family Z-Score
	Grades Attained	Family Z-Score			Absolute (5 % Trim)	Rank	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ITT	0.667**	0.170**	0.360***	0.201**	0.081	0.223**	0.185***
	(0.250)	(0.072)	(0.127)	(0.088)	(0.094)	(0.085)	(0.055)
Four Highest Strata (=1) * ITT	-0.629*	-0.119	-0.294**	0.119	0.185	-0.049	-0.220***
	(0.336)	(0.093)	(0.145)	(0.138)	(0.119)	(0.106)	(0.079)
Test: ITT + Highest Strata * ITT = 0							
P-value	0.861	0.350	0.342	0.005	0.002	0.018	0.504
N	1,006	1,007	907	1,006	997	1,006	900
R^2	0.427	0.347	0.453	0.146	0.099	0.097	0.158

Notes: *** p<0.01, ** p<0.05, * p<0.10. Standard errors are clustered at the locality level and given in parentheses. The strata are ordered from 1 to 7 with one being the poorest. Highest strata includes strata 4–7, hence coefficient on ITT indicates estimates on 3 poorest strata. Regressions are weighted to account for sampling and attrition providing population estimates. Differential ITT results compare early to late treatment groups in 2010. Z-scores are calculated using the mean and standard deviation of the late treatment group. Controls include three-month age group indicators and indicator variables for grades attained at baseline, stratification, and region. Variables in columns 1-2 and 4-6 measured using the 2010 household survey; variables in columns 3 and 7 measured using the 2010 individual survey.

What other analysis was done and why?

- Trying to determine absolute effects of the program
  - Helps determine if zeros, are no program effect or if the early and late treatment group have the same positive effect
  - If the differential effects are under-estimates
- Matching
  - Compares the treated to a comparison group Comparison groups, not randomized, never received program
    - Include data on this group starting in 2004 survey
  - To try to determine what any zeros are and the absolute effects rather than the differential effects.
- Double-difference
  - Uses 1995 and 2005 census data (before and after)
  - Uses treatment and comparison group municipalities
    - Program covered about 90% of people in treated munis

TABLE B1.1: 2010 DIFFERENTIAL EXPERIMENTAL IMPACTS, ALTERNATIVE SPECIFICATIONS

	Education		Learning Family Z-Score	Labor Market Participation Family Z-Score	Earnings Family Z-Score		Socio-emotional Family Z-Score
	Grades Attained	Family Z-Score			Rank	Absolute (5% Trim)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Age 9–12 Main results							
ITT	0.288*	0.098**	0.183**	0.272***	0.194***	0.192***	0.053
	(0.167)	(0.043)	(0.070)	(0.075)	(0.057)	(0.067)	(0.039)
N	1,006	1,007	907	1,006	1,006	997	900
Panel B: Age 9-12 Excluding Over-Sample Children							
ITT	0.375*	0.124**	0.172**	0.266**	0.224***	0.214**	0.027
	(0.197)	(0.061)	(0.069)	(0.118)	(0.081)	(0.104)	(0.056)
N	496	497	444	497	497	491	441
Panel C: Age 11 Only							
ITT	0.681*	0.177*	0.283***	0.283**	0.183	0.223	-0.046
	(0.362)	(0.099)	(0.079)	(0.127)	(0.129)	(0.135)	(0.070)
N	352	353	321	353	353	349	318
Panel D: Age 9–12 Strata Controls Only							
ITT	0.361	0.109	0.326**	0.165	0.133**	0.165**	0.128*
	(0.324)	(0.078)	(0.141)	(0.099)	(0.059)	(0.063)	(0.070)
N	1,006	1,007	907	1,006	1,006	997	900
Panel E: Age 9–12 Extended Controls							
ITT	0.289*	0.105**	0.165**	0.240***	0.179***	0.182***	0.045
	(0.167)	(0.044)	(0.067)	(0.069)	(0.057)	(0.064)	(0.039)
N	1,006	1,007	907	1,006	1,006	997	900
Panel F: Age 9–12 Sampling Weights (No attrition correction)							
ITT	0.351**	0.105**	0.175***	0.266***	0.195***	0.188***	0.033
	(0.153)	(0.040)	(0.063)	(0.070)	(0.057)	(0.069)	(0.041)
N	1,006	1,007	907	1,006	1,006	997	900
Panel G: Age 9-12 Attrition Correction using Observations From Intensive & Regular tracking							
ITT	0.296	0.089	0.164***	0.282***	0.195***	0.181**	0.0450
	(0.188)	(0.054)	(0.062)	(0.068)	(0.073)	(0.08)	(0.047)
N	1,006	1,007	907	1,006	1,006	997	900
Panel H: Restrict Household Survey Variables to Individual survey Sample and Weight							
ITT	0.351*	0.138***		0.245***	0.100	0.121	
	(0.176)	(0.048)		(0.080)	(0.071)	(0.077)	-
N	906	907		907	907	898	
Panel I: Lee bounds							
Lower Bound	0.456*	0.018	0.237***	-0.084	0.152**	0.149**	0.043
	(0.237)	(0.060)	(0.077)	(0.070)	(0.072)	(0.072)	(0.053)

Other Robustness

- Age 11 only
- Age 9-12 Strata controls only
- Age 9-12 Extended controls
- Sample weights no attrition correction
- Restrict HH survey variables to individual sample
- Attrition: use different measure Lee bounds
- Look at the use of proxy variables
- Look at different age groups included
  - 7-13 – why look at those?

What goes in the appendix and why? Would you put them somewhere different?

- Baseline balance
- Robustness
- Description of the program
- Description of variables
- Details on how tracking was done
- Spill-over analysis
- Matching analysis details
- Double difference analysis details

TABLE B1.2: 2010 DIFFERENTIAL EXPERIMENTAL IMPACTS, ROBUSTNESS INFERENCE

	Education Family Z-Score	Learning Family Z-Score	Labor Market Participation Family Z-Score	Earnings Family Z-Score		Socio-emotional Family Z-Score
				Rank	Absolute (5% Trim)	
	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Multiple Hypothesis Testing – Familywise Error Rate Adjusted p-values						
p-value	0.034	0.020	0.005	0.005	0.013	0.186
Panel B: Randomization Inference						
Exact p-value	0.036	0.030	0.002	0.001	0.008	0.256

Notes: Regressions include boys ages 9-12 at the start of the program in November 2000 and are weighted to account for sampling and attrition providing population estimates. Differential ITT results compare early to late treatment groups in 2010. The family z-score averages the z-scores for the individual components and is calculated if any component is available. Z-scores are calculated using the mean and standard deviation of the late treatment group. Controls include three-month age group indicators and indicator variables for grades attained at baseline, stratification, and region. Variables in columns 1-2 and 4-6 are measured using the 2010 household survey; variables in column 3 and 7 are measured using the 2010 individual survey. Panel A adjusts the p-values for multiple hypothesis testing using the familywise error rate following Anderson (2008) based on the variables included in the table. Panel B shows Fisher exact p-values obtained through randomization inference using Young (2017)'s randomization-t.

Robustness for Inference Appendix B Multiple Hypothesis Testing

Multiple hypothesis testing across equations

- The more outcomes you look at the more likely you are to find something that is statistically different by random chance.
- Adjusts the standard errors for the fact that looked at multiple outcomes.
  - This paper does it across families.
- For false discovery rates download from Michael Anderson website
- https://are.berkeley.edu/~mlanderson/ARE_Website/Research.html

Randomization Inference

- Assesses whether an observed realization of statistics, like a treatment effect, is unlikely to be observed by chance so is statically significant.
  - Could the outcome have been observed by chance even if the treatment had no effect.
- Most test statistics rely on large sample sizes and asymptotics properties.
  - Good in the case of small samples, stratified or clustered treatment assignments, or non-standard randomization techniques.
- Uncertainty is over which units within your sample are assigned to treatment.
  - What we have been doing is uncertainty over the specific sample of the population you drew.
- Now in Stata ritest

Randomization Inference

- Randomization inference for *p*-values, based on the idea that the specific units in our sample **that are treated** are random.
- Chance of a treatment-control difference in outcomes of any given magnitude simply based on which units are assigned to the treatment group – even if the treatment has no effect.
- Process re-assigns “treatment” at random, to compute the probability of differences of various magnitudes under the null hypothesis that the treatment does nothing.

References: Susan Athey and Guido Imbens. 2016 “The Econometrics of Randomized Experiments”

https://jasonkerwin.com/nonparibus/2017/09/25/randomization-inference-vs-bootstrapping-p-values/

Correcting for attrition

- Estimates on learning are large and precise
- Bounding Methods
  - Horowitz and Manski 2000: uninformative as are too wide.
  - Lee (2009): similar results – not much differential attrition between group.
    - Deals with differential attrition due to treatment.
- Weighting by inverse of probability to attrite
  - Make probability to attrite based on baseline characteristics.
  - See this paper:

Steps to deal with attrition

- Determine attrition rate
- Check for differential attrition
  - Are rates different for treatment & control or some other relevant groups
    - Even if the rates are the same, different types of people may have attrited in the T & C areas
      - Program itself may have caused differential attrition.
  - Is attrition correlated with observables
    - Regression of attrition on baseline observables interacted with the treatment variable.
      - Hard to do if no baseline data
  - Use some kind of attrition correction
    - Put on treatment effect bounds.
    - Problem is bounds are large when attrition is high
  - JHR Spring 1998 – special issue on attrition

Manski-Horwitz (2000) Bounds

- 2000 - “Nonparametric analysis of randomized experiments with missing covariate and outcome data,” *Journal of the American Statistical Association,* 95:77-84.
- Non-parametric
- Construct upper bound
  - Assign best value of outcome to treatment attritors, and worse value of outcome to control attritors
    - E.g. if outcome is 1 if sick last week. Then assign 0 to treatment and 1 to control attritors.
- Construct lower bound: do the opposite of upper
- Variable most be bounded (have a max and min).
- Can yield wide uniformative bounds (variable has wide value range).
- Good for binary outcomes.

Lee (2009) bounds

- "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects," *Review of Economic and Statistics*, 76:1071-1102
- Bounds average treatment effect for never-attriters
- Intuition: trim so that share of observed individuals is same for T&C
  - Throw out fraction of obs from the group that has less attrition.
  - Lower bound: drop fraction of obs that contribute most to result
  - Upper bound: drop fraction of obs that contribute least to result.
- Monotonicity Assumption
  - Assignment to treatment can only affect attrition in one direction
  - Doesn't cause some people to stay but other people to leave

Do Conditional Cash Transfers Improve Economic Outcomes in the Next Generation? Evidence from Mexico

Susan Parker and Tom Vogl

Introduction

- **Goal:** Examine the Effects of a the Mexican CCT 13 years after program start on education, labor market and economic outcomes
- Why topic is important
  - Key policy tool: in > 80 mainly developing countries
  - A Goal: to reduce long-term poverty, like to know if this happens
  - Early CCT introduced late 90s early 2000
    - Can now look at effects of program in adulthood
- Why Mexican CCT important
    - One of first large scale CCTs to use randomized roll out
    - Short-term effects on rural areas heavily studied

Introduction

- Longer-term effects of CCTs difficult to study
  - Early CCT study design use a randomized phase-in
    - Long-term analysis: early and late phase-in groups
    - Mexican CCT: phased-in control group after 1.5 months
  - Differential effects often zero in long-run using randomization
    - Doesn’t mean there is no absolute impact of program
      - Program effect on early and late phase-in groups the same
    - Time too short between phase-in groups, late treat catch-up
  - Get creative to estimate something closer to absolute effects
- Kuddos for trying to estimate something close to absolute effect
  - Use roll out of program over time in Mexico
  - Use 2010 census

Data

- Census 2010: 10 percent sample of household
  - Program starts 1997: so 13 years after program start.
  - Kid who was starting/finishing 13/17 and starting Highschool in 1997:
    - 1997 is 26/30
    - 1999 is 24/28
    - 2001 (when urban areas come in) is 22/36
- Sample includes 20-39 year olds
  - Kids in the 30s are unlikely to have received the program or at least for long.
- Merge on:
  - Administrative program data on number of households enrolled in Progress by year and municipality
- Restrict to municipalities with high or very high marginality index in 1990
  - Good this is prior to the program
- Aggregate to municipality level (same as Barham 2011)
  - boarders change over time so aggregate to have a consistent sent of municipalities over time

Outcomes

- Education: grades completed
- Labor Market:
  - Labor force participation
  - Wage work
  - Agricultural work
  - Weekly labor hours
  - Monthly labor income
- Household level outcomes
  - Housing condition and durable ownership – in an index
  - Use principal components to make index.
- Household composition
- Geographic mobility: moved between 2005 and 2010

Mechanisms

- How does the program translate into better labor market outcomes?
  - Don’t go into it too much in the paper
  - Talk about education and greater parental income
  - Don’t try to separate out effects.

Treatment Indicator - ENROLL

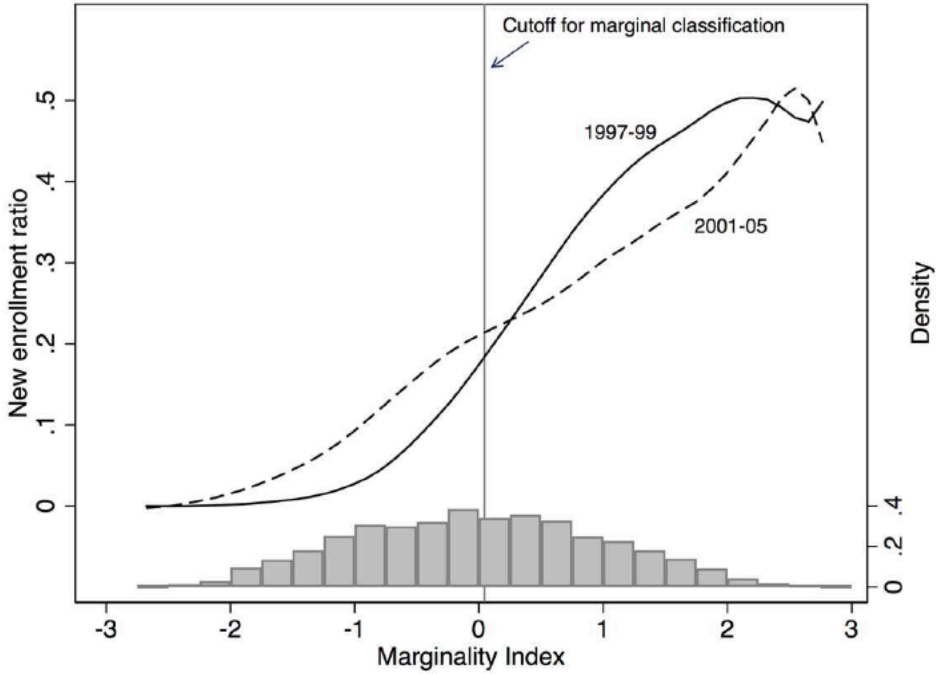
- Same as Barham (2011) - at municipality level
  - Number of Households / Total number of household in municipality in 1997
- Enroll: is a confusing name
  - Not educational enrollment
  - Enrollment into the program

Research Design

- Two sources of variation
  - Role out of the program over time in across municipalities
    - Same as Barham (2011)
  - Variation in exposure of different age cohorts
- Fully Exposed Cohort: age 7-11 in 1997, 20-34 in 2010
  - Most intensely treated because received treatment prior to high school when kids begin to dropout.
- Not Exposed or Pre-Program Cohort: 16+ in 1997, 29+ in 2010
  - Likely to have dropped out because Progresa came too late
- Partially Exposed: Age 12-14
  - Some may have already dropped out prior to high school but those who stayed on helped by program to make through high school.
- Focus on two roll out phases
  - 1997-1999: rural roll out phase
  - 2001-2005: urban roll out phase
  - Skip 2000 which was an election year

Program Intensity Against Marginality Index
-Program highest in poorest municipalities

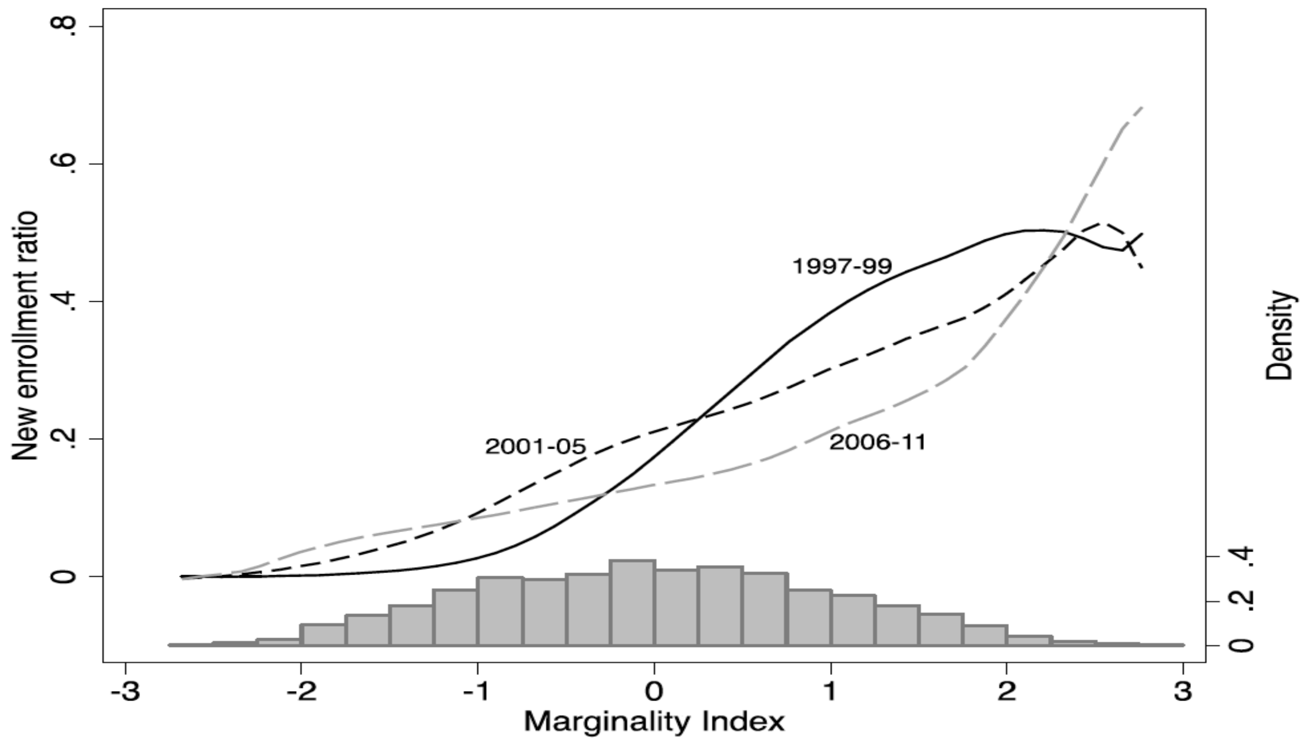
Figure 2: Municipal Economic Conditions and New Enrollment Intensity, 1997-99 vs. 2001-05



Note: Curves are local linear regressions with bandwidths of 0.25; bars are a histogram of the marginality index. The new enrollment ratio is the count of new households enrolled divided by the estimated number of households in the municipality in 1997. The marginalization index is defined (using 1990 census data) by the Mexican government as the normalized first principal component of nine municipal population shares: the share living in communities with less than 5000 inhabitants, the share earning less than twice the minimum wage, the share illiterate, and the shares with less than primary school, without a toilet, without electricity, without running water, with crowding, and with a dirt floor.

New Enrollment and Marginality Index
Drop after 2005: not in poorest municipalities

Figure A2: Economic Conditions and Enrollment Intensity, 1997-99, 2001-05, and 2006-11



Note: The new enrollment ratio is the count of new households enrolled divided by the estimated number of households in 1997. The marginality index is defined (using 1990 census data) by the Mexican government as the normalized first principal component of nine municipal population shares: the share illiterate, the share with less than primary school education, the share without a toilet, the share without electricity, the share without running water, the share with crowding as measured by number of rooms divided by household size, the share with a dirt floor, the share living in communities with less than 5000 inhabitants, and the share earning less than twice the minimum wage. Curves are local linear regressions with bandwidths of 0.25.

Compare Penetration and Poverty by Marginality Index

Table 2: Municipal Marginality and Program Rollout

	Very low/low/medium marginality	High/very high marginality
	(1)	(2)
A. Program rollout		
Cumulative enrollment ratio, 1999	0.06	0.34
	(0.07)	(0.16)
Cumulative enrollment ratio, 2005	0.19	0.63
	(0.15)	(0.22)
B. Marginality index and components		
Index (1st principal component)	-0.79	0.86
	(0.56)	(0.57)
Share illiterate	0.13	0.34
	(0.07)	(0.13)
Share without toilet	0.26	0.60
	(0.17)	(0.19)
Share without electricity	0.12	0.37
	(0.10)	(0.25)
Share without running water	0.19	0.51
	(0.15)	(0.24)
Share with dirt floor	0.22	0.62
	(0.14)	(0.22)
Share earning <2x minimum wage	0.69	0.86
	(0.12)	(0.08)
Share with primary education or less	0.46	0.70
	(0.12)	(0.10)
Share with crowding	0.60	0.74
	(0.10)	(0.08)
Share in localities with pop. < 5000	0.60	0.95
	(0.36)	(0.14)
Number of municipalities	1239	1143
Number of states	32	24

Note: The cumulative enrollment ratio in year *t* is the number of households enrolled in Progresa up to and including *t* divided by the estimated number of households in 1997. The remainder of the analysis focuses on the high or very high marginality municipalities in column (2), where program activities were concentrated.

Research Design

- DD design using special and cohort variation
  - Assume 2 cohorts Young (treated) and Old (nontreated)
- Typical
  - $Y_{mt} = \delta_m + \eta_t + Enroll_m * CohortYoung_t + \varepsilon_{mt}$
  - They will call the CohortYoung - Post
- What is the assuming of this model for DD?
  - Compare Young and Old Cohort
  - Early versus late entry to Progress (or High versus lower enrollment)
- Trends outcomes between cohorts, would be the same in higher versus low intensity municipalities, had the program not happened.

Basic Research Design in Picture Double Difference Strategy

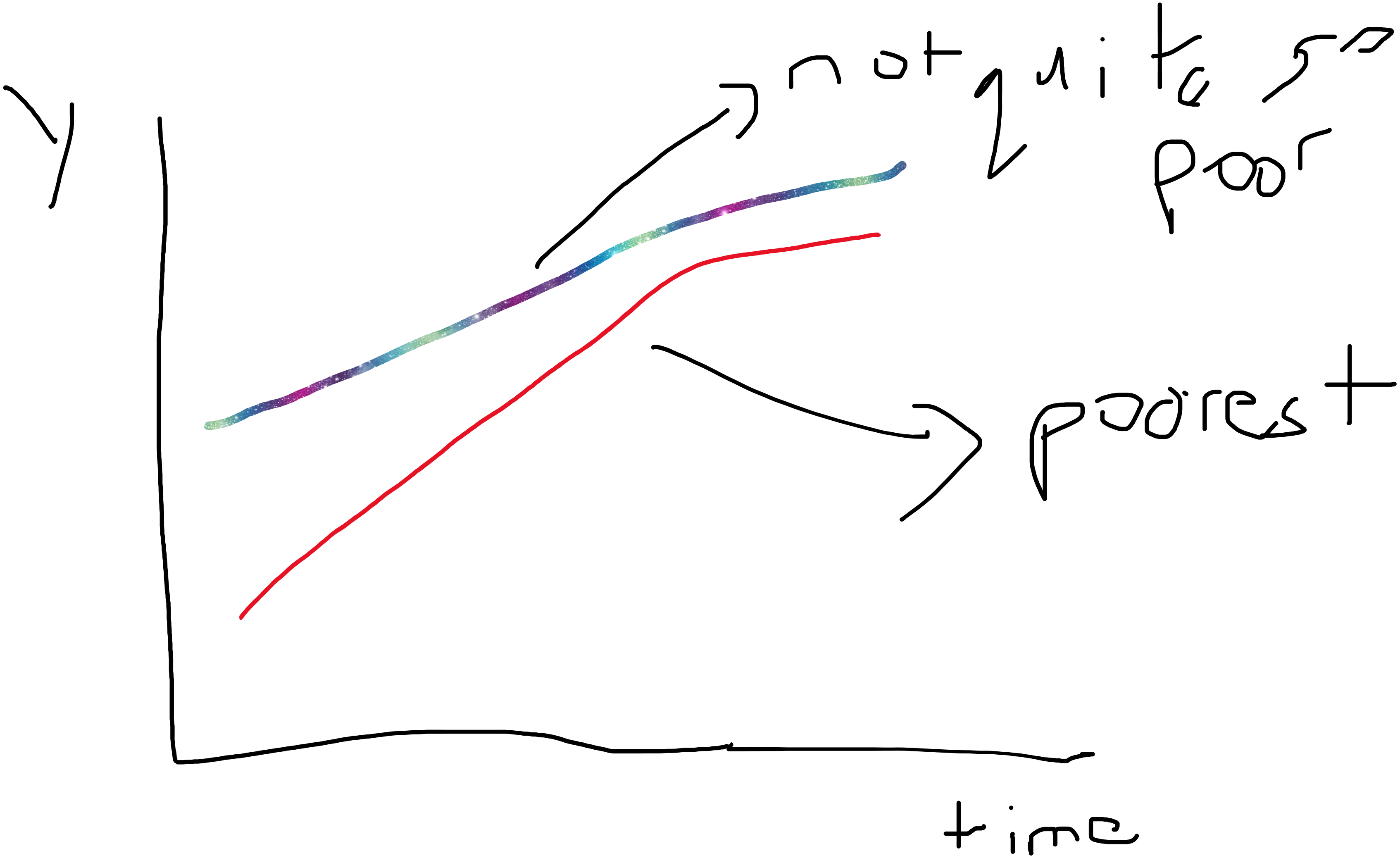
Cohorts (post) Age in 1997	High Enroll in 1999 (roll out 97-2000)	Low Enroll in 1999 (roll out 2001-2005)
7-11 (High Exposure)		
15-19 (Low Exposure)		

- If just use the , **double-difference strategy**
- Includes municipalities were program rolled out in 1997-2000 and 2001-2005

Research Design

- Concern: cross-cohort trends not the same
  - Initial poverty differ between higher and lower enrollment municipalities
  - Poorer municipalities may have steeper trends in labor outcomes with out the program because they start at a lower level of employment so more room to improve.

Assume no program



Research Design

- Want to compare
  - Among municipalities with the same level of enrollment in 2005, did those who had more enrollment prior to 2000 experience larger gains in the young cohort compared to the hold cohort.
  - Problem, the types of localities that were enrolled prior to 2000 are different than those after 2000
    - Prior to 2000 were rural localities after 2000 urban localities.
- To do this include: $enroll_m^{2005} * post_t$

Estimation Regression

Regression Equation: includes time and municipality fixed-effect

$$Y_{imt} = \delta_m + \eta_t + \beta(enroll_m^{1999} * post_t) + \gamma(enroll_m^{2005} * post_t) + \varepsilon_{imt}$$

- Post: 1 if in the younger cohort (<12 in 1997)/0 age 15-20 in 1997
- Program rolled out at locality level over more than 10 years
  - Locality level lower than from muni: Enroll_m varies over time as new localities enter program
- γ: argue this absorbs differences in cross-cohort trends?
- What is the problem with this?
  - $enroll_m^{2005}$: This is not preprogram, so it is controlling for program effect in the younger cohorts (treated cohort).

Basic Research Design in Picture

Cohorts (post) Age in 1997	High Enroll in 1999 (roll out 97-2000)	Low Enroll in 1999 (roll out 2001-2005)	Age in 2001/2005
7-11 (High Exposure)			2001: 10-14 2005: 14-18
15-19 (Low Exposure)			

- Low Enroll Group partially exposed to program 2001-2005
- Controls for partial exposure in municipalities rolled out between 2001-2005

Estimation Regression

Regression Equation: includes time and municipality fixed-effect

$$Y_{imt} = \delta_m + \eta_t + \beta(enroll_m^{1999} * post_t) + \gamma(enroll_m^{2005} * post_t) + \varepsilon_{imt}$$

Robustness:

Let baseline characteristics differ between cohorts

-X is have a dirt floor, have electricity

-This tried to control for any differences at baseline between cohorts.

$$= \delta_m + \eta_t + \beta(enroll_m^{1999} * post_t) + \gamma(enroll_m^{2005} * post_t) + (post_t * X_m^{1990})'\Psi$$

Treats to Validity

- Endogenous program intensity
  - Argue lower bound if worse of areas come in first.
  - Do you agree?
- Endogenous migration
  - Assign program exposure to people using their pre-program municipality
  - Census includes information on municipality in 2005 and state of birth
  - Use municipality in 2005 as a proxy for location prior to program in 1997

Table 3: Program Impacts on Educational Attainment

	Men		Women	
	(1)	(2)	(3)	(4)
A. Grades completed				
Enrollment ratio, 1999	1.310	1.373	1.370	1.339
× post cohort	[0.405]	[0.225]	[0.392]	[0.441]
Mean (SD)	7.90 (3.96)		7.65 (4.07)	
Observations	319,714		375,892	
B. Some secondary				
Enrollment ratio, 1999	0.165	0.176	0.288	0.290
× post cohort	[0.054]	[0.045]	[0.057]	[0.081]
Mean (SD)	0.58 (0.49)		0.54 (0.50)	
Observations	320,423		376,753	
C. Some high				
Enrollment ratio, 1999	0.145	0.098	0.161	0.105
× post cohort	[0.058]	[0.032]	[0.042]	[0.036]
Mean (SD)	0.26 (0.44)		0.25 (0.43)	
Observations	320,423		376,753	
D. Some college				
Enrollment ratio, 1999	0.023	0.025	-0.015	-0.014
× post cohort	[0.017]	[0.018]	[0.023]	[0.019]
Mean (SD)	0.09 (0.28)		0.08 (0.28)	
Observations	320,423		376,753	
Municipality and cohort FE	X	X	X	X
1990 marginality interactions		X		X

Note: Brackets contain standard errors clustered at the state level. All regressions additionally control for the interaction of the post indicator with the cumulative enrollment ratio in 2005. Sample includes individuals from high and very high marginality municipalities who were aged 7-11 and 15-19 in 1997.

Reading Results

- Given with and without the interaction of baseline vars with cohorts
- Point estimate is the program effect if enroll goes from 0 to 100 percent
  - Full roll out, all households in a municipality covered
- Grades completed estimate is ~ 1.37 men and women
  - This means if every household in the municipality were treated the program effect on education is 1.37 grades for men and women.
  - To make into a percent divide by the mean
  - 1.37/7.9 for men and 1.37/7.7
  - So about 17-18 percent increase
- Issue: they do not do below
  - all household in a municipality are not covered
  - Need to multiple this by average number of households covered
- Some Secondary: how to do read these point estimates?
  - Men 17.6 percentage points, women 29 percentage points

Table 4: Program Impacts on Labor Market Outcomes

	Men		Women	
	(1)	(2)	(4)	(5)
A. Working				
Enrollment ratio, 1999	-0.025	-0.020	0.063	0.105
× post cohort	[0.037]	[0.030]	[0.036]	[0.032]
Mean (SD)	0.81 (0.40)		0.26 (0.44)	
Observations	320,133		377,236	
B. Working for wage				
Enrollment ratio, 1999	0.037	0.039	0.067	0.107
× post cohort	[0.043]	[0.037]	[0.032]	[0.034]
Mean (SD)	0.48 (0.50)		0.17 (0.36)	
Observations	313,459		374,600	
C. Working in agriculture				
Enrollment ratio, 1999	-0.114	-0.087	-0.003	0.004
× post cohort	[0.065]	[0.038]	[0.009]	[0.010]
Mean (SD)	0.39 (0.49)		0.03 (0.17)	
Observations	317,865		376,067	
D. Health insurance from job				
Enrollment ratio, 1999	0.089	0.078	0.031	0.023
× post cohort	[0.049]	[0.034]	[0.028]	[0.018]
Mean (SD)	0.14 (0.35)		0.13 (0.34)	
Observations	319,695		376,409	
E. Weekly labor hours				
Enrollment ratio, 1999	3.448	3.032	3.459	5.547
× post cohort	[1.681]	[1.457]	[1.753]	[1.675]
Mean (SD)	36.18 (24.24)		10.30 (20.54)	
Observations	316,879		376,081	
F. Monthly labor earnings				
Enrollment ratio, 1999	725	281	335	438
× post cohort	[541]	[395]	[185]	[172]
Mean (SD)	2081 (4839)		670 (2168)	
Observations	307,326		373,088	
Municipality and cohort FE	X	X	X	X
1990 marginality interactions		X		X

Note: Brackets contain standard errors clustered at the state level. All regressions additionally control for the interaction of the post indicator with the cumulative enrollment ratio in 2005. All labor market outcomes are unconditional on labor force participation. Sample includes individuals from high and very high marginality municipalities who were aged 7-11 and 15-19 in 1997.

Table 5: Program Impacts on Compensation, Conditional on Participation, Men

	Working for wage	Weekly labor hours	Monthly labor earnings	Log earnings	Sectoral earnings score	Log earnings score	Log wage	Sectoral wage score	Log wage score
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A. Without 1990 marginality interactions									
1999 ratio	0.105	5.550	972	0.249	646	0.126	0.120	11.474	0.110
× post	[0.048]	[1.926]	[677]	[0.176]	[309]	[0.069]	[0.154]	[6.417]	[0.066]
B. With 1990 marginality interactions									
1999 ratio	0.110	4.703	448	0.115	498	0.098	0.026	8.248	0.083
× post	[0.047]	[1.865]	[510]	[0.132]	[197]	[0.044]	[0.102]	[4.100]	[0.042]
Mean	0.49	43	2006	7.81	40,711	8.15	2.60	88	4.32
(SD)	(0.50)	(18)	(5091)	(0.79)	(2717)	(0.56)	(0.79)	(58)	(0.54)
Observations	248,428	251,848	242,295	147,674	254,358	252,581	145,930	252,581	252,581

Note: Estimates are conditional on working. Brackets contain standard errors clustered at the state level. All regressions additionally control for the interaction of the post indicator with the cumulative enrollment ratio in 2005. Sample includes male workers from high and very high marginality municipalities who were aged 7-11 and 15-19 in 1997. The wage is calculated as monthly earnings divided by monthly hours. The sectoral earnings and wage scores are obtained by regressing the earnings and wages on indicators for 1-digit occupation and industry codes (from the International Standard Classification of Occupations and the International Standard Industrial Classification) in the nationwide sample of prime-age male workers, ages 35-44.

Table 6: Program Impacts on Household and Demographic Outcomes

	Men		Women	
	(1)	(2)	(4)	(5)
A. HH monthly labor earnings				
Enrollment ratio, 1999	735	1211	1092	1284
× post cohort	[701]	[650]	[502]	[461]
Mean (SD)	3856 (8146)		1863 (7429)	
Observations	312,730		376,384	
B. Index of housing conditions				
Enrollment ratio, 1999	0.263	0.177	0.278	0.179
× post cohort	[0.116]	[0.070]	[0.098]	[0.081]
Mean (SD)	-0.02 (1.00)		0.02 (0.99)	
Observations	315,190		370,997	
C. Index of durable goods				
Enrollment ratio, 1999	0.162	0.121	0.216	0.137
× post cohort	[0.082]	[0.063]	[0.070]	[0.065]
Mean (SD)	-0.02 (1.00)		0.02 (1.01)	
Observations	316,412		372,371	
D. New municipality last 5 years				
Enrollment ratio, 1999	0.099	0.074	0.119	0.090
× post cohort	[0.079]	[0.049]	[0.054]	[0.043]
Mean (SD)	0.07 (0.25)		0.07 (0.26)	
Observations	321,910		378,674	
E. Urban residence				
Enrollment ratio, 1999	0.083	0.061	0.117	0.061
× post cohort	[0.061]	[0.037]	[0.050]	[0.031]
Mean (SD)	0.35 (0.48)		0.36 (0.48)	
Observations	321,910		378,674	
Municipality and cohort FE	X	X	X	X
1990 marginality interactions		X		X

Note: Brackets contain standard errors clustered at the state level. All regressions additionally control for the interaction of the post indicator with the cumulative enrollment ratio in 2005. Sample includes individuals from high and very high marginality municipalities who were aged 7-11 and 15-19 in 1997. The housing index is the standardized first principal component of indicators for having a dirt floor, a modern roof, sewage, a flush toilet, piped water, and electricity; the durables index is the same for indicators for having a car, a mobile phone, a computer, a washer, a refrigerator, a television, and a hot water heater.

Robustness

- Examine results without $(enroll_m^{2005} * post_t)$
- Add controls: all interacted by post
  - Politics
    - Municipality PRI vote share in 1994 pres election.
  - School construction
    - New schools per capita in 1995-2000 and 2000-2005
    - Interacted by post
  - Violence
    - Change in municipality homicide rate between 2010 and 2006.
- Falsification Test:
  - Use data set where should be no program effect
  - Census 1990, instead of 2010, so lag 20 years.
  - Municipality of residence based on age in 1977

Quibbles

2. **Limit to municipalities that look more similar and more details in paper on the roll out**

- Roll out not random and at <u>locality</u> level
  - 1997-2000: poor rural areas 1997-2000
    - 1997: high marginality, permanent health clinic, not geographically isolates, rural
    - 1998: drop condition for permanent health clinic, big incorporation
    - 1999-2001: add geographically isolate
  - 2001-2005: urban areas & ???
  - Aggregate to municipality level – tricky as treated composition changes
- Table A2: show municipality incorporated between 97-99 and 2000-2005 are different
  - Show statistical significant of the difference in means on table
- Trends in outcomes may not be similar in absence of program when municipalities are different to start

Quibbles

3. Enroll variables are endogenous
   - Only the poor in a locality got the treatment
   - % treated in a municipality varies with poverty
   - Use a 1/0 treatment variable as a robustness check
4. Mechanisms: look at others
   - Learning: literacy rate
   - Fertility for women: have a child
5. Earning – results noisy for men, trim or winsorize
   5. Winsorize means to top code potential outliers
6. Analysis at locality level?
7. Add 1990 census variables not in marginality index
8. Standard stuff
   - Heterogeneity with baseline census characteristics
   - Adjust for multiple hypothesis testing: families of outcomes
   - Attrition Bounds