

Randomized Evaluation of the Mexican Universal Health Insurance Program: Substantive and Methodological Findings

Kosuke Imai
Princeton University

Joint work with Gary King, Emmanuela Gakidou, Jason Lakin, Ryan T. Moore, Clayton Nall, Nirmala Ravishankar, Manett Vargas, Martha María Téllez-Rojo, Juan Eugenio Hernández Ávila, Mauricio Hernández Ávila, Héctor Hernández Llamas

April 27, 2010

Project References

- “A ‘Politically Robust’ Experimental Design for Public Policy Evaluation, with Application to the Mexican Universal Health Insurance Program” *Journal of Policy Analysis and Management*, January 2007.
- “The Essential Role of Pair Matching in Cluster-Randomized Experiments, with Application to the Mexican Universal Health Insurance Evaluation” (with discussions) *Statistical Science*, February 2009.
- “Public Policy for the Poor? A Randomised assessment of the Mexican Universal Health Insurance Programme” (with comments) *The Lancet*, April 2009.

Before Seguro Popular (2003)

- Two public healthcare providers:
 - ① IMSS (Mexican Institute of Social Security)
 - ② ISSTE (Institute of Security and Social Services for State Workers)
- All workers have a right to affiliate with IMSS but employers have to pay a minimum fee for each worker
- State workers affiliate with ISSTE
- Contract workers, unemployed etc. must purchase private insurance
- 50 million uninsured Mexicans (more than half of the population!)
- No regular access to health care, particularly those with low income
- Negative consequences:
 - ① large out-of-pocket healthcare expenditures
 - ② less access to and reduced quality in health services
 - ③ inefficient use of resources

Seguro Popular: A Massive Reform

- Article 4 of the Mexican constitution:

all persons have a right to the protection of their health

- Medical services, preventive care, pharmaceuticals, and financial health protection
- Voluntary and available for everyone but free to the poor
- Beneficiaries: intended to cover (by 2012) all 50M Mexicans who otherwise have no access to the healthcare system
- Cost in 2005: \$800 million in new money
- One of the largest health reforms of any country in last 2 decades
- Most visible accomplishment of the Fox administration
- Major issue in the 2006 presidential campaign
- Initial result: 20M beneficiaries by the end of 2007
- Still going: The World Bank just approved a \$1.25 billion loan to cover additional 10M Mexicans

Goals of SP & Evaluation Outcome Measures

- Financial Protection (money for the poor rarely makes it there)
 - Out-of-pocket expenditure
 - Catastrophic expenditure (8.4% of households, spend > 30% of annual disposable income on health)
 - Impoverishment due to health care payments
- Health System Effective Coverage
 - Percent of population receiving appropriate treatment by disease
 - Satisfaction of affiliates with Seguro Popular
- Health Care Facilities
 - Operations, office visits, emergencies, personnel, infrastructure and equipment, drug inventory.
- Health
 - Health status
 - All-cause mortality
 - Cause-specific mortality

- Frenk and Fox asked: How can one democratically elected government “tie the hands” of their successors?
 - Commission an independent evaluation
 - (They are true believers in SP)
 - Like in science: make themselves vulnerable to being proven wrong
 - If we show SP is a success: elimination would be difficult
 - If SP is a failure: who cares about extending it
- The largest randomized health policy experiment in history
- One of the largest policy experiments to date
- First cohort: 148 geographic areas, 1,380 localities, approximately 118,569 households, and about 534,457 individuals

Lessons from Previous Public Policy Experiments

- Most large scale public policy experiments fail
- Many failures are political
 - politicians: need to pursue short term goals
 - citizens: you plan to *randomly* assign *me*?
 - all perfectly legitimate; a natural consequence in a democracy
- E.g., Oportunidades program: Some governors “miraculously” found money for control groups to participate too (numerous similar examples worldwide)
- Previous evaluation designs ignored democratic politics
- We developed a new research design & new methods for Mexico:
 - includes **fail-safe components** for when politics intervenes
 - uses data far more efficiently to **find effects** and **save money**

Example of Fail-Safe Design Procedure (CR vs. MPR)

- 1 **Complete Randomization** (used in **Oportunidades** evaluation)
 - Flip coin to assign program to each area
- 2 **Matched-Pair Randomization** (used in **Seguro Popular** evaluation)
 - Match areas in pairs on background characteristics
 - Flip coin once for each pair: one area within each pair gets the program
 - If one area is lost:
 - Drop the other member of the pair
 - Remaining pairs are kept
 - Treated and control groups are still protected by randomization: advantages of the experiment survives
 - With our new statistical methods, the design:
 - **Smaller standard errors**: up to 6 times smaller!
 - **We can find effects where complete randomization cannot**
 - **Far less expensive** for the same impact

Detailed Design Summary

- 1 Define 12,284 “health clusters” that tile Mexico’s 31 states; each includes a health clinic and catchment area
- 2 Persuaded 13 of 31 states to participate (7,078 clusters)
- 3 Match clusters in pairs on background characteristics.
- 4 Select 74 pairs (based on necessary political criteria, closeness of the match, likelihood of compliance)
- 5 Randomly assign one in each pair to receive encouragement to affiliate, better health facilities, drugs, and doctors
- 6 Conduct baseline survey of each cluster’s health facility
- 7 Survey $\approx 32,000$ random households in 50 of the 74 treated and control unit pairs (chosen based on likelihood of compliance with encouragement and similarity of the clusters within pair)
- 8 Repeat surveys in 10 months and subsequently to see effects

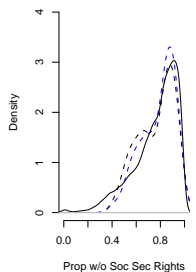
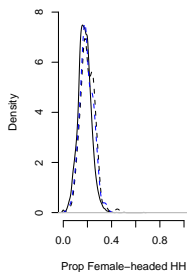
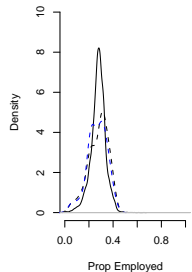
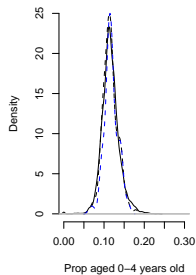
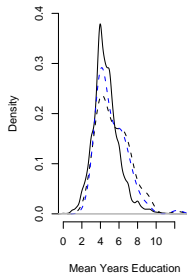
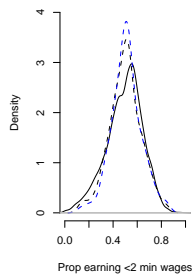
Matched-Pair Cluster-Randomized Designs in Polisci

- Special research designs require special methods
- Prop. of polisci CREs which ignore the design: 100%
- Prop. of polisci CREs making more assumptions than necessary: 100%
- MPDs \geq Complete Randomization w.r.t.: efficiency, bias, power, estimator simplicity, and robustness to political intervention
- Proportion of previous CREs in polisci that use MPDs: 0%
- Conclusion: we're leaving a lot of information on the table!
- Imai-King-Nall: prove above results and offer simple estimators for MPDs making minimal assumptions for both intent to treat and complier average treatment effects

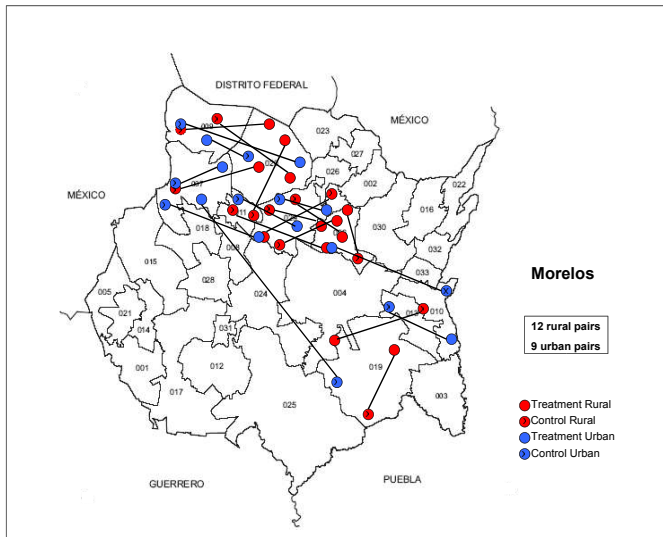
Remaining in study: 148 clusters (74 pairs) in 7 states



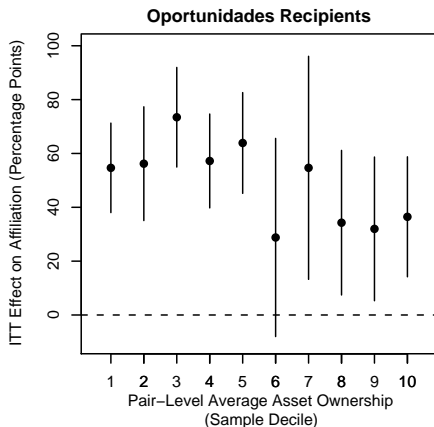
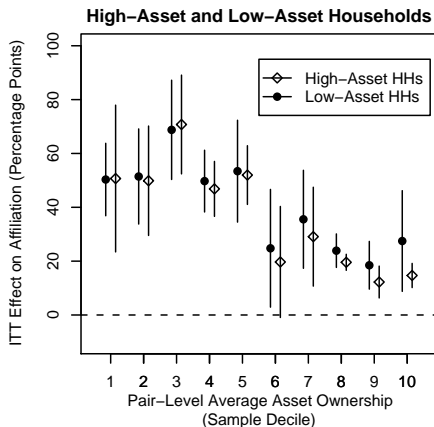
Clusters are Representative On Measured Variables



Matched Pairs, Morelos



Effect of Encouragement on Seguro Popular Affiliation



- Households in poorer areas have higher estimated average causal effects on their affiliation rate

Effect on % of Households with Catastrophic Expenditures

	All Study Participants			Experimental Compliers		
	Average (Control)	ITT	SE	Average (Control)	CACE	SE
All	8.4	1.9*	(0.9)	9.5	5.2*	(2.3)
Low Asset	9.9	3.0*	(1.3)	11.0	6.5*	(2.5)
High Asset	7.1	0.9	(0.8)	7.9	3.0	(2.7)
Female-Headed	8.5	1.4	(1.1)	10.6	3.8	(3.0)

“Catastrophic expenditures”: out-of-pocket health expenses $>$ 30% of post-subsistence income

Effect on Out-of-pocket Health Expenditures, I (in pesos)

	All Study Participants			Experimental Compliers		
	Average (Control)	ITT	SE	Average (Control)	CACE	SE
Overall:						
All	\$1631.3	\$258.0	(\$175)	\$1712.7	\$689.7	(\$453)
Low Asset	1360.2	425.6*	(197)	1502.6	915.3*	(392)
High Asset	1867.9	128.4	(201)	1933.2	428.2	(669)
Female-Headed	1509.1	156.5	(207)	1689.9	428.6	(566)
Inpatient Care:						
All	532.5	96.9*	(44)	557.1	259.1*	(112)
Low Asset	527.1	188.2*	(73)	579.0	404.8*	(142)
High Asset	537.2	31.1	(52)	536.2	103.6	(173)
Female-Headed	452.5	115.1*	(68)	510.0	315.2*	(182)
Outpatient Care:						
All	448.3	116.7*	(63)	499.1	312.0*	(161)
Low Asset	412.3	176.7*	(73)	466.3	380.0*	(147)
High Asset	479.7	81.9	(69)	533.0	272.9	(230)
Female-Headed	416.3	110.4	(75)	496.8	302.4	(202)

Effect on Out-of-pocket Health Expenditures, II (in pesos)

	All Study Participants			Experimental Compliers		
	Average (Control)	ITT	SE	Average (Control)	CACE	SE
Medicine:						
All	521.1	20.0	(41)	534.5	53.3	(109)
Low Asset	427.3	17.8	(46)	444.7	38.3	(100)
High Asset	603.0	29.4	(47)	627.5	98.1	(157)
Female-Headed	625.6	53.6	(55)	738.9	146.8	(151)
Medical Devices:						
All	139.7	-8.8	(23)	117.8	-23.4	(62)
Low Asset	72.0	-0.2	(20)	72.8	-0.5	(43)
High Asset	198.8	-16.5	(29)	165.6	-55.1	(98)
Female-Headed	155.5	10.9	(34)	162.8	30.0	(94)

Utilization: Overall

	All Study Participants			Experimental Compliers		
	Average (Control)	ITT	SE	Average (Control)	CACE	SE
Utilization (Procedures):						
Used Outpatient Services (%)	62.6	-1.5	(1.9)	64.8	-4.0	(5.2)
Outpatient Visits (count)	1.6	-0.03	(0.09)	1.7	-0.08	(0.23)
Hospitalized (%)	7.6	-0.2	(0.5)	7.9	-0.5	(1.5)
Hospitalizations (count)	0.1	-0.003	(0.006)	0.1	-0.01	(0.02)
Satisfaction with Provider (%)	68.0	-1.0	(1.6)	69.8	-2.6	(4.5)
Utilization (Preventative) (%):						
Eye Exam Last Yr.	10.0	-0.7	(0.7)	9.8	-1.8	(1.9)
Flu Vaccine	25.7	-1.8	(1.4)	27.2	-4.9	(3.7)
Mammogram Last Yr.	5.1	-0.9	(0.6)	5.2	-2.3	(1.6)
Cervical Last Yr.	21.8	-1.3	(2.0)	22.2	-3.2	(4.8)
Pap Test Last Yr.	31.9	-2.3	(2.1)	33.2	-5.8	(5.0)

Self-Assessment, Controlling for Baseline Levels

	ITT		CACE	
Overall Health	0.6	(2.2)	1.7	(6.0)
Mobility	0.2	(0.9)	0.6	(2.5)
Vigorous Activity	3.3	(2.4)	8.9	(6.4)
Self-Care	-0.2	(0.6)	-0.5	(1.6)
Soreness	1.0	(1.4)	2.6	(3.8)
Pain	1.1	(1.2)	3.0	(3.3)
Sleeping	1.0	(1.0)	2.6	(2.5)
Depression	0.6	(3.0)	1.5	(7.9)
Anxiety	0.8	(1.8)	2.1	(4.8)

- A **difference-in-difference test**: The causal effect of Seguro Popular on the change from baseline to followup in the difference between treated and control groups on health self-assessment variables

Risk Factors: Overall

	All Study Participants Average (Control)	ITT	SE	Experimental Average (Control)	Compliers CACE	SE
Doctor's Diagnosis (%):						
Diabetes	6.5	0.4	(0.4)	6.2	1.0	(1.2)
Hypertension	14.7	-1.1	(0.8)	15.0	-2.9	(2.1)
Cholesterol	5.6	-0.2	(0.4)	5.3	-0.6	(1.0)
Diet or Exercise Program (%):						
Hypertension	27.8	-0.6	(1.8)	28.4	-1.6	(5.0)
Cholesterol	11.4	-0.8	(1.1)	11.2	-2.1	(3.0)
Treated with Medication (%):						
Hypertension	35.2	0.8	(1.5)	34.5	2.2	(4.1)
Cholesterol	4.8	-0.1	(0.5)	4.5	-0.4	(1.5)
Risk Factors (%):						
Smoking	10.7	1.6*	(0.6)	10.9	4.3*	(1.7)
Seat Belt	28.2	1.0	(1.7)	25.4	2.6	(4.6)

Conclusions

- **Positive effects detected now:**
 - Catastrophic expenditures slashed
 - In-patient out-of-pocket expenditures drastically reduced
 - Out-patient out-of-pocket expenditures drastically reduced
 - Citizen satisfaction is high
- **Positive effects not yet seen:**
 - Expenditures on medicines
 - Utilization (preventative and procedures)
 - Risk factors
- **Other findings:**
 - Only 66% of automatically affiliated Oportunidades respondents were aware of this fact
 - More encouragement to affiliate might be devoted to finding the poor hidden within relatively “wealthier” clusters
 - Developed new and more powerful evaluation design and statistical methods, tuned to the needs of Mexico
 - These design and statistical methods are widely applicable

- Continued evaluation of long-term effects
- Political and economic consequences of Seguro Popular
- Merging municipality-level electoral data with evaluation data
- Does the receipt of health insurance change voting behavior?