Public Policy for the Poor? A Randomized Evaluation of the Mexican Universal Health Insurance Program

Gary King Institute for Quantitative Social Science Harvard University

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

(Talk at USAID Conference, "Mind the Gap: Research & Evaluation Methods for Scaling Up Evidenced-Based Interventions" 6/1/10)

Gary King (Harvard)

æ

イロト イヨト イヨト イヨト

• New Evaluation Design: Gary King et al., A 'Politically Robust' Experimental Design for Public Policy Evaluation, with Application to the Mexican Universal Health Insurance Program *Journal of Policy Analysis and Management*, 26, 3 (2007): 479-506.

- New Evaluation Design: Gary King et al., A 'Politically Robust' Experimental Design for Public Policy Evaluation, with Application to the Mexican Universal Health Insurance Program *Journal of Policy Analysis and Management*, 26, 3 (2007): 479-506.
- New Statistical Methods: Kosuke Imai, Gary King, and Clayton Nall. The Essential Role of Pair Matching in Cluster-Randomized Experiments, with Application to the Mexican Universal Health Insurance Evaluation *Statistical Science*, 24, 1 (2009): pp. 29–53.

- New Evaluation Design: Gary King et al., A 'Politically Robust' Experimental Design for Public Policy Evaluation, with Application to the Mexican Universal Health Insurance Program *Journal of Policy Analysis and Management*, 26, 3 (2007): 479-506.
- New Statistical Methods: Kosuke Imai, Gary King, and Clayton Nall. The Essential Role of Pair Matching in Cluster-Randomized Experiments, with Application to the Mexican Universal Health Insurance Evaluation *Statistical Science*, 24, 1 (2009): pp. 29–53.
- Results: Gary King et al., Public Policy for the Poor? A Randomized 10-Month Evaluation of the Mexican Universal Health Insurance Program *The Lancet*, 373 (25 April 2009): 1447-1454.

- New Evaluation Design: Gary King et al., A 'Politically Robust' Experimental Design for Public Policy Evaluation, with Application to the Mexican Universal Health Insurance Program *Journal of Policy Analysis and Management*, 26, 3 (2007): 479-506.
- New Statistical Methods: Kosuke Imai, Gary King, and Clayton Nall. The Essential Role of Pair Matching in Cluster-Randomized Experiments, with Application to the Mexican Universal Health Insurance Evaluation *Statistical Science*, 24, 1 (2009): pp. 29–53.
- Results: Gary King et al., Public Policy for the Poor? A Randomized 10-Month Evaluation of the Mexican Universal Health Insurance Program *The Lancet*, 373 (25 April 2009): 1447-1454.
- Copies at ~ http://gking.harvard.edu

Seguro Popular Evaluation

Gary King (Harvard)

æ

イロト イヨト イヨト イヨト

• Program: medical services, preventive care, pharmaceuticals, and financial health protection

B ▶ < B ▶

- Program: medical services, preventive care, pharmaceuticals, and financial health protection
- Beneficiaries: 50M Mexicans with no regular access to health care, particularly those with low incomes.

- Program: medical services, preventive care, pharmaceuticals, and financial health protection
- Beneficiaries: 50M Mexicans with no regular access to health care, particularly those with low incomes.
- Result: Among world's largest health reforms in 2 decades

- Program: medical services, preventive care, pharmaceuticals, and financial health protection
- Beneficiaries: 50M Mexicans with no regular access to health care, particularly those with low incomes.
- Result: Among world's largest health reforms in 2 decades
- Politics: Most visible accomplishment of Fox administration; Major issue in the 2006 presidential campaign

- Program: medical services, preventive care, pharmaceuticals, and financial health protection
- Beneficiaries: 50M Mexicans with no regular access to health care, particularly those with low incomes.
- Result: Among world's largest health reforms in 2 decades
- Politics: Most visible accomplishment of Fox administration; Major issue in the 2006 presidential campaign
- Evaluation: Financial protection (money for the poor rarely makes it there), utilization, & health

- Program: medical services, preventive care, pharmaceuticals, and financial health protection
- Beneficiaries: 50M Mexicans with no regular access to health care, particularly those with low incomes.
- Result: Among world's largest health reforms in 2 decades
- Politics: Most visible accomplishment of Fox administration; Major issue in the 2006 presidential campaign
- Evaluation: Financial protection (money for the poor rarely makes it there), utilization, & health
- Size: The largest randomized health policy experiment in history

- Program: medical services, preventive care, pharmaceuticals, and financial health protection
- Beneficiaries: 50M Mexicans with no regular access to health care, particularly those with low incomes.
- Result: Among world's largest health reforms in 2 decades
- Politics: Most visible accomplishment of Fox administration; Major issue in the 2006 presidential campaign
- Evaluation: Financial protection (money for the poor rarely makes it there), utilization, & health
- Size: The largest randomized health policy experiment in history
- Problem: Most large scale public policy experiments fail!

- Program: medical services, preventive care, pharmaceuticals, and financial health protection
- Beneficiaries: 50M Mexicans with no regular access to health care, particularly those with low incomes.
- Result: Among world's largest health reforms in 2 decades
- Politics: Most visible accomplishment of Fox administration; Major issue in the 2006 presidential campaign
- Evaluation: Financial protection (money for the poor rarely makes it there), utilization, & health
- Size: The largest randomized health policy experiment in history
- Problem: Most large scale public policy experiments fail!
- Solution: New evaluation design with fail-safe components

イロト イポト イヨト イヨト

• Complete Randomization (used in Oportunidades evaluation)

± ►.

Complete Randomization (used in Oportunidades evaluation)

• Flip coin to assign program to each area

Complete Randomization (used in Oportunidades evaluation)

- Flip coin to assign program to each area
- If one area is lost:

Complete Randomization (used in Oportunidades evaluation)

- Flip coin to assign program to each area
- If one area is lost:
 - treated and control groups are incomparable

Complete Randomization (used in Oportunidades evaluation)

- Flip coin to assign program to each area
- If one area is lost:
 - treated and control groups are incomparable
 - all advantages of randomization are gone

O Complete Randomization (used in Oportunidades evaluation)

- Flip coin to assign program to each area
- If one area is lost:
 - treated and control groups are incomparable
 - all advantages of randomization are gone

O Complete Randomization (used in Oportunidades evaluation)

- Flip coin to assign program to each area
- If one area is lost:
 - treated and control groups are incomparable
 - all advantages of randomization are gone

Matched-Pair Randomization (used in Seguro Popular evaluation)

• Match areas in pairs on background characteristics

Complete Randomization (used in Oportunidades evaluation)

- Flip coin to assign program to each area
- If one area is lost:
 - treated and control groups are incomparable
 - all advantages of randomization are gone

- Match areas in pairs on background characteristics
- Flip coin once for each pair: one area within each pair gets the program

O Complete Randomization (used in Oportunidades evaluation)

- Flip coin to assign program to each area
- If one area is lost:
 - treated and control groups are incomparable
 - all advantages of randomization are gone

- 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:

Complete Randomization (used in Oportunidades evaluation)

- Flip coin to assign program to each area
- If one area is lost:
 - treated and control groups are incomparable
 - all advantages of randomization are gone

- 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

Complete Randomization (used in Oportunidades evaluation)

- Flip coin to assign program to each area
- If one area is lost:
 - treated and control groups are incomparable
 - all advantages of randomization are gone

- 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

Complete Randomization (used in Oportunidades evaluation)

- Flip coin to assign program to each area
- If one area is lost:
 - treated and control groups are incomparable
 - all advantages of randomization are gone

- 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

Complete Randomization (used in Oportunidades evaluation)

- Flip coin to assign program to each area
- If one area is lost:
 - treated and control groups are incomparable
 - all advantages of randomization are gone

- 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:

Complete Randomization (used in Oportunidades evaluation)

- Flip coin to assign program to each area
- If one area is lost:
 - treated and control groups are incomparable
 - all advantages of randomization are gone

- 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:
 - More efficient: up to 38 times!

Complete Randomization (used in Oportunidades evaluation)

- Flip coin to assign program to each area
- If one area is lost:
 - treated and control groups are incomparable
 - all advantages of randomization are gone

- 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:
 - More efficient: up to 38 times!
 - Smaller standard errors: up to 6 times smaller

Complete Randomization (used in Oportunidades evaluation)

- Flip coin to assign program to each area
- If one area is lost:
 - treated and control groups are incomparable
 - all advantages of randomization are gone

- 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:
 - More efficient: up to 38 times!
 - Smaller standard errors: up to 6 times smaller
 - We can find effects where complete randomization cannot

Complete Randomization (used in Oportunidades evaluation)

- Flip coin to assign program to each area
- If one area is lost:
 - treated and control groups are incomparable
 - all advantages of randomization are gone

- 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:
 - More efficient: up to 38 times!
 - 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

æ

イロト イヨト イヨト イヨト

Define 12,284 "health clusters" that tile Mexico's 31 states; each includes a health clinic and catchment area

- Define 12,284 "health clusters" that tile Mexico's 31 states; each includes a health clinic and catchment area
- Persuaded 13 of 31 states to participate (7,078 clusters)
- Define 12,284 "health clusters" that tile Mexico's 31 states; each includes a health clinic and catchment area
- Persuaded 13 of 31 states to participate (7,078 clusters)
- Match clusters in pairs on background characteristics.

- Define 12,284 "health clusters" that tile Mexico's 31 states; each includes a health clinic and catchment area
- Persuaded 13 of 31 states to participate (7,078 clusters)
- Match clusters in pairs on background characteristics.
- Select 74 pairs (based on necessary political criteria, closeness of the match, likelihood of compliance)

- Define 12,284 "health clusters" that tile Mexico's 31 states; each includes a health clinic and catchment area
- Persuaded 13 of 31 states to participate (7,078 clusters)
- Match clusters in pairs on background characteristics.
- Select 74 pairs (based on necessary political criteria, closeness of the match, likelihood of compliance)
- Sandomly assign one in each pair to receive encouragement to affiliate, better health facilities, drugs, and doctors

- Define 12,284 "health clusters" that tile Mexico's 31 states; each includes a health clinic and catchment area
- Persuaded 13 of 31 states to participate (7,078 clusters)
- Match clusters in pairs on background characteristics.
- Select 74 pairs (based on necessary political criteria, closeness of the match, likelihood of compliance)
- Sandomly assign one in each pair to receive encouragement to affiliate, better health facilities, drugs, and doctors
- \bigcirc Survey $\approx\!32,000$ random households in 50 of the 74 treated and control unit pairs

- Define 12,284 "health clusters" that tile Mexico's 31 states; each includes a health clinic and catchment area
- Persuaded 13 of 31 states to participate (7,078 clusters)
- Match clusters in pairs on background characteristics.
- Select 74 pairs (based on necessary political criteria, closeness of the match, likelihood of compliance)
- Sandomly assign one in each pair to receive encouragement to affiliate, better health facilities, drugs, and doctors
- \bigcirc Survey $\approx\!32,000$ random households in 50 of the 74 treated and control unit pairs
- Repeat surveys in 10 months and subsequently to see effects

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



Matched Pairs, Estado de México



3

イロト 不得下 イヨト イヨト

э

イロト イヨト イヨト イヨト

Design has three parts

∃ ▶ ∢

Design has three parts

Matching pairs on observed covariates

Design has three parts

- Matching pairs on observed covariates
- 2 Randomization of treatment within pairs

Design has three parts

- Matching pairs on observed covariates
- 2 Randomization of treatment within pairs
- If necessary statistically adjust for differences

Design has three parts

- Matching pairs on observed covariates
- 2 Randomization of treatment within pairs
- If necessary statistically adjust for differences

Triple Robustness

If matching or randomization or statistical analysis is right, but the other two are wrong, results are still unbiased

Design has three parts

- Matching pairs on observed covariates
- 2 Randomization of treatment within pairs
- If necessary statistically adjust for differences

Triple Robustness

If matching or randomization or statistical analysis is right, but the other two are wrong, results are still unbiased

Two Additional Checks if Triple Robustness Fails

Design has three parts

- Matching pairs on observed covariates
- 2 Randomization of treatment within pairs
- If necessary statistically adjust for differences

Triple Robustness

If matching or randomization or statistical analysis is right, but the other two are wrong, results are still unbiased

Two Additional Checks if Triple Robustness Fails

If one of the three works, then "effect of SP" on time 0 outcomes (measured in baseline survey) must be zero

Design has three parts

- Matching pairs on observed covariates
- ② Randomization of treatment within pairs
- If necessary statistically adjust for differences

Triple Robustness

If matching or randomization or statistical analysis is right, but the other two are wrong, results are still unbiased

Two Additional Checks if Triple Robustness Fails

- If one of the three works, then "effect of SP" on time 0 outcomes (measured in baseline survey) must be zero
- 2 If we lose pairs, we check for selection bias by rerunning this check

Gary King (Harvard)

2

メロト メポト メヨト メヨト

• Positive effects detected now:

æ

イロト イヨト イヨト イヨト

- Positive effects detected now:
 - Catastrophic expenditures slashed

э

- ∢ ∃ ▶

- Positive effects detected now:
 - Catastrophic expenditures slashed
 - Out-of-pocket expenditures drastically reduced

• Positive effects detected now:

- Catastrophic expenditures slashed
- Out-of-pocket expenditures drastically reduced
- High citizen satisfaction

.∋...>

• Positive effects detected now:

- Catastrophic expenditures slashed
- Out-of-pocket expenditures drastically reduced
- High citizen satisfaction
- Positive effects not (yet?) seen:

• Positive effects detected now:

- Catastrophic expenditures slashed
- Out-of-pocket expenditures drastically reduced
- High citizen satisfaction
- Positive effects not (yet?) seen:
 - Expenditures on medicines

• Positive effects detected now:

- Catastrophic expenditures slashed
- Out-of-pocket expenditures drastically reduced
- High citizen satisfaction

• Positive effects not (yet?) seen:

- Expenditures on medicines
- Utilization (preventative and procedures)

• Positive effects detected now:

- Catastrophic expenditures slashed
- Out-of-pocket expenditures drastically reduced
- High citizen satisfaction

• Positive effects not (yet?) seen:

- Expenditures on medicines
- Utilization (preventative and procedures)
- Risk factors

• Positive effects detected now:

- Catastrophic expenditures slashed
- Out-of-pocket expenditures drastically reduced
- High citizen satisfaction

• Positive effects not (yet?) seen:

- Expenditures on medicines
- Utilization (preventative and procedures)
- Risk factors

• Positive effects detected now:

- Catastrophic expenditures slashed
- Out-of-pocket expenditures drastically reduced
- High citizen satisfaction

• 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

• Positive effects detected now:

- Catastrophic expenditures slashed
- Out-of-pocket expenditures drastically reduced
- High citizen satisfaction

• Positive effects not (yet?) seen:

- Expenditures on medicines
- Utilization (preventative and procedures)
- Risk factors

- Only 66% of automatically affiliated Oportunidades respondents were aware of this fact
- Low affiliation rates for the poor submerged within wealthier areas

• Positive effects detected now:

- Catastrophic expenditures slashed
- Out-of-pocket expenditures drastically reduced
- High citizen satisfaction

• Positive effects not (yet?) seen:

- Expenditures on medicines
- Utilization (preventative and procedures)
- Risk factors

- Only 66% of automatically affiliated Oportunidades respondents were aware of this fact
- Low affiliation rates for the poor submerged within wealthier areas
- Developed new and more powerful evaluation design and statistical methods, tuned to the needs of Mexico

• Positive effects detected now:

- Catastrophic expenditures slashed
- Out-of-pocket expenditures drastically reduced
- High citizen satisfaction

• Positive effects not (yet?) seen:

- Expenditures on medicines
- Utilization (preventative and procedures)
- Risk factors

- Only 66% of automatically affiliated Oportunidades respondents were aware of this fact
- Low affiliation rates for the poor submerged within wealthier areas
- Developed new and more powerful evaluation design and statistical methods, tuned to the needs of Mexico
- Evaluation design: being adopted around the world

http://GKing.Harvard.edu

æ

(日) (周) (三) (三)

ITT on Outcome Measures at Baseline, for all families (left) and poor families, in Oportunidades (right)



Gary King (Harvard)

Public Policy for the Poor?

11 / 17

Effect of Encouragement on Seguro Popular Affiliation



Horizontal axes: per-capita asset ownership deciles of areas (poorer to the left). Vertical axes: percentage point causal effect of encouragement to affiliate on Seguro Popular affiliation.

Poor areas, not poor households, are affiliated the most

Effect on % of Households with Catastrophic Health Expenditures

	All Study Participants			Experimental Compliers		
	Average	ITT	SE	Average	CACE	SE
	(Control)			(Control)		
All	8.4	1.9*	(.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

(日) (周) (三) (三)

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

	All Study Participants			Experimental Compliers		
	Average	ITT	SE	Average	CACE	SE
	(Control)			(Control)		
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)

Image: Image:

B ▶ < B ▶

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

	All Study Participants			Experimental Compliers		
	Average	ITT	SE	Average	CACE	SE
	(Control)			(Control)		
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)

Image: Image:

∃ ▶ ∢ ∃ ▶
Utilization: Overall

	All Study Participants			Experimental Compliers		
	Average	ITT	SE	Average	CACE	SE
	(Control)			(Control)		
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)

2

イロト イヨト イヨト イヨト

Clusters are Representative On Measured Variables



Public Policy for the Poor?