A "Politically Robust" Experimental Design for Public Policy Evaluation, with Application to the Mexican Universal Health Insurance Program

Gary King
Institute for Quantitative Social Science
Harvard University

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

Project References

Project References

- A 'Politically Robust' Experimental Design for Public Policy Evaluation, with Application to the Mexican Universal Health Insurance Program Gary King, Emmanuela Gakidou, Nirmala Ravishankar, Ryan T. Moore, Jason Lakin, Manett Vargas, Martha María Téllez-Rojo, Juan Eugenio Hernández Ávila, Mauricio Hernández Ávila, Héctor Hernández Llamas. Forthcoming, JPAM.

Project References

- A ‘Politically Robust’ Experimental Design for Public Policy Evaluation, with Application to the Mexican Universal Health Insurance Program Gary King, Emmanuela Gakidou, Nirmala Ravishankar, Ryan T. Moore, Jason Lakin, Manett Vargas, Martha María Téllez-Rojo, Juan Eugenio Hernández Ávila, Mauricio Hernández Ávila, Héctor Hernández Llamas. Forthcoming, JPAM.

- The Essential Role of Pair Matching in Cluster-Randomized Experiments, with Application to the Mexican Universal Health Insurance Evaluation Kosuke Imai, Gary King, and Clayton Nall.

Project References

- A 'Politically Robust' Experimental Design for Public Policy Evaluation, with Application to the Mexican Universal Health Insurance Program Gary King, Emmanuela Gakidou, Nirmala Ravishankar, Ryan T. Moore, Jason Lakin, Manett Vargas, Martha María Téllez-Rojo, Juan Eugenio Hernández Ávila, Mauricio Hernández Ávila, Héctor Hernández Llamas. Forthcoming, JPAM.

- The Essential Role of Pair Matching in Cluster-Randomized Experiments, with Application to the Mexican Universal Health Insurance Evaluation Kosuke Imai, Gary King, and Clayton Nall.

- Public Policy for the Poor? A Randomized Evaluation of the Mexican Universal Health Insurance Program Gary King, Emmanuela Gakidou, Kosuke Imai, Jason Lakin, Ryan T. Moore, Nirmala Ravishankar, Manett Vargas, Martha María Téllez-Rojo, Juan Eugenio Hernández Ávila, Mauricio Hernández Ávila, Héctor Hernández Llamas.

Lessons from Experimental Failures

Lessons from Experimental Failures

- Many large scale public policy experiments fail

Lessons from Experimental Failures

- Many large scale public policy experiments fail
- The “problem”: Politicians pursuing short term goals

Lessons from Experimental Failures

- Many large scale public policy experiments fail
- The "problem": Politicians pursuing short term goals
- Citizens: you plan to randomly assign me?

Lessons from Experimental Failures

- Many large scale public policy experiments fail
- The "problem": Politicians pursuing short term goals
- Citizens: you plan to randomly assign me?
- E.g., Mexican anti-poverty program: Some governors "miraculously" found money so citizens in control groups could also participate

Lessons from Experimental Failures

- Many large scale public policy experiments fail
- The "problem": Politicians pursuing short term goals
- Citizens: you plan to randomly assign me?
- E.g., Mexican anti-poverty program: Some governors "miraculously" found money so citizens in control groups could also participate
- Numerous randomized evaluations torpedoed by politicians

Lessons from Experimental Failures

- Many large scale public policy experiments fail
- The "problem": Politicians pursuing short term goals
- Citizens: you plan to randomly assign me?
- E.g., Mexican anti-poverty program: Some governors "miraculously" found money so citizens in control groups could also participate
- Numerous randomized evaluations torpedoed by politicians
- All perfectly legitimate; a natural consequence in a democracy

Lessons from Experimental Failures

- Many large scale public policy experiments fail
- The "problem": Politicians pursuing short term goals
- Citizens: you plan to randomly assign me?
- E.g., Mexican anti-poverty program: Some governors "miraculously" found money so citizens in control groups could also participate
- Numerous randomized evaluations torpedoed by politicians
- All perfectly legitimate; a natural consequence in a democracy
- Our proposed research design:

Lessons from Experimental Failures

- Many large scale public policy experiments fail
- The "problem": Politicians pursuing short term goals
- Citizens: you plan to randomly assign me?
- E.g., Mexican anti-poverty program: Some governors "miraculously" found money so citizens in control groups could also participate
- Numerous randomized evaluations torpedoed by politicians
- All perfectly legitimate; a natural consequence in a democracy
- Our proposed research design:
  - uses data as efficiently as possible

Lessons from Experimental Failures

- Many large scale public policy experiments fail
- The "problem": Politicians pursuing short term goals
- Citizens: you plan to randomly assign me?
- E.g., Mexican anti-poverty program: Some governors "miraculously" found money so citizens in control groups could also participate
- Numerous randomized evaluations torpedoed by politicians
- All perfectly legitimate; a natural consequence in a democracy
- Our proposed research design:
  - uses data as efficiently as possible
  - includes key fail-safe components

Seguro Popular: A Massive Reform

Seguro Popular: A Massive Reform

- medical services, preventive care, pharmaceuticals, and financial health protection

Seguro Popular: A Massive Reform

- medical services, preventive care, pharmaceuticals, and financial health protection
- beneficiaries: 50M Mexicans (half of the population) with no regular access to health care, particularly those with low incomes.

Seguro Popular: A Massive Reform

- medical services, preventive care, pharmaceuticals, and financial health protection
- beneficiaries: 50M Mexicans (half of the population) with no regular access to health care, particularly those with low incomes.
- Cost in 2005: $795.5 million in new money

Seguro Popular: A Massive Reform

- medical services, preventive care, pharmaceuticals, and financial health protection
- beneficiaries: 50M Mexicans (half of the population) with no regular access to health care, particularly those with low incomes.
- Cost in 2005: $795.5 million in new money
- Cost when implemented: additional 1% of GDP

Seguro Popular: A Massive Reform

- medical services, preventive care, pharmaceuticals, and financial health protection
- beneficiaries: 50M Mexicans (half of the population) with no regular access to health care, particularly those with low incomes.
- Cost in 2005: $795.5 million in new money
- Cost when implemented: additional 1% of GDP
- Demand-based allocations, with stewardship

Seguro Popular: A Massive Reform

- medical services, preventive care, pharmaceuticals, and financial health protection
- beneficiaries: 50M Mexicans (half of the population) with no regular access to health care, particularly those with low incomes.
- Cost in 2005: $795.5 million in new money
- Cost when implemented: additional 1% of GDP
- Demand-based allocations, with stewardship
- One of the largest health reforms of any country in last 2 decades

Seguro Popular: A Massive Reform

- medical services, preventive care, pharmaceuticals, and financial health protection
- beneficiaries: 50M Mexicans (half of the population) with no regular access to health care, particularly those with low incomes.
- Cost in 2005: $795.5 million in new money
- Cost when implemented: additional 1% of GDP
- Demand-based allocations, with stewardship
- One of the largest health reforms of any country in last 2 decades
- Most visible accomplishment of the Fox administration

Seguro Popular: A Massive Reform

- medical services, preventive care, pharmaceuticals, and financial health protection
- beneficiaries: 50M Mexicans (half of the population) with no regular access to health care, particularly those with low incomes.
- Cost in 2005: $795.5 million in new money
- Cost when implemented: additional 1% of GDP
- Demand-based allocations, with stewardship
- 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

SPS Evaluation

SPS Evaluation

- Frenk and Fox asked: How can one democratically elected government "tie the hands" of their successors?

SPS Evaluation

- Frenk and Fox asked: How can one democratically elected government "tie the hands" of their successors?
- Their theory:

SPS Evaluation

- Frenk and Fox asked: How can one democratically elected government "tie the hands" of their successors?
- Their theory:
  - Commission an independent evaluation

SPS Evaluation

- Frenk and Fox asked: How can one democratically elected government "tie the hands" of their successors?
- Their theory:
  - Commission an independent evaluation
  - (They are true believers in SPS)

SPS Evaluation

- Frenk and Fox asked: How can one democratically elected government "tie the hands" of their successors?
- Their theory:
  - Commission an independent evaluation
  - (They are true believers in SPS)
  - Like in science: make themselves vulnerable to being proven wrong

SPS Evaluation

- Frenk and Fox asked: How can one democratically elected government "tie the hands" of their successors?
- Their theory:
  - Commission an independent evaluation
  - (They are true believers in SPS)
  - Like in science: make themselves vulnerable to being proven wrong
  - If we show SPS is a success: elimination would be difficult

SPS Evaluation

- Frenk and Fox asked: How can one democratically elected government "tie the hands" of their successors?
- Their theory:
  - Commission an independent evaluation
  - (They are true believers in SPS)
  - Like in science: make themselves vulnerable to being proven wrong
  - If we show SPS is a success: elimination would be difficult
  - If SPS is a failure: who cares about extending it

SPS Evaluation

- Frenk and Fox asked: How can one democratically elected government "tie the hands" of their successors?
- Their theory:
  - Commission an independent evaluation
  - (They are true believers in SPS)
  - Like in science: make themselves vulnerable to being proven wrong
  - If we show SPS is a success: elimination would be difficult
  - If SPS is a failure: who cares about extending it
- One of the largest policy experiments to date

SPS Evaluation

- Frenk and Fox asked: How can one democratically elected government "tie the hands" of their successors?
- Their theory:
  - Commission an independent evaluation
  - (They are true believers in SPS)
  - Like in science: make themselves vulnerable to being proven wrong
  - If we show SPS is a success: elimination would be difficult
  - If SPS is a failure: who cares about extending it
- One of the largest policy experiments to date
- Maybe the largest randomized health policy experiment ever

SPS Evaluation

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

Is Randomization Always Unethical in Public Policy?

Is Randomization Always Unethical in Public Policy?

- Not ethical to randomly assign health care to Mexicans!

Is Randomization Always Unethical in Public Policy?

- Not ethical to randomly assign health care to Mexicans!
- Is it ok to randomly assign the order in which people are informed about the program?

Is Randomization Always Unethical in Public Policy?

- Not ethical to randomly assign health care to Mexicans!
- Is it ok to randomly assign the order in which people are informed about the program?
- Program implementation always includes arbitrary decisions, made by low level officials

Is Randomization Always Unethical in Public Policy?

- Not ethical to randomly assign health care to Mexicans!
- Is it ok to randomly assign the order in which people are informed about the program?
- Program implementation always includes arbitrary decisions, made by low level officials
- If decisions are arbitrary, they can be randomized

Is Randomization Always Unethical in Public Policy?

- Not ethical to randomly assign health care to Mexicans!
- Is it ok to randomly assign the order in which people are informed about the program?
- Program implementation always includes arbitrary decisions, made by low level officials
- If decisions are arbitrary, they can be randomized
- Generalization: randomization is acceptable at one level below that at which politicians care

Is Randomization Always Unethical in Public Policy?

- Not ethical to randomly assign health care to Mexicans!
- Is it ok to randomly assign the order in which people are informed about the program?
- Program implementation always includes arbitrary decisions, made by low level officials
- If decisions are arbitrary, they can be randomized
- Generalization: randomization is acceptable at one level below that at which politicians care
- We were able to randomize at the "health cluster" level, the health clinic and catchment area around it — except in areas favored by politicians or presently infeasible to offer services

Matched-Pair, Cluster-Randomized Design

Matched-Pair, Cluster-Randomized Design

- Random assignment of clusters (rather than individuals)

Matched-Pair, Cluster-Randomized Design

- Random assignment of clusters (rather than individuals)
  - (control groups will get SPS later)

Matched-Pair, Cluster-Randomized Design

- Random assignment of clusters (rather than individuals)
  - (control groups will get SPS later)
  - Less efficient, but politically feasible

Matched-Pair, Cluster-Randomized Design

- Random assignment of clusters (rather than individuals)
  - (control groups will get SPS later)
  - Less efficient, but politically feasible
  - Reduced interference between units

Matched-Pair, Cluster-Randomized Design

- Random assignment of clusters (rather than individuals)
  - (control groups will get SPS later)
  - Less efficient, but politically feasible
  - Reduced interference between units
  - Frequently used: 68% of political science experiments;

Matched-Pair, Cluster-Randomized Design

- Random assignment of clusters (rather than individuals)
  - (control groups will get SPS later)
  - Less efficient, but politically feasible
  - Reduced interference between units
  - Frequently used: 68% of political science experiments; "exponential increase" in public health and medicine;

Matched-Pair, Cluster-Randomized Design

- Random assignment of clusters (rather than individuals)
  - (control groups will get SPS later)
  - Less efficient, but politically feasible
  - Reduced interference between units
  - Frequently used: 68% of political science experiments; "exponential increase" in public health and medicine; common in public policy, education, psychology, etc.

Matched-Pair, Cluster-Randomized Design

- Random assignment of clusters (rather than individuals)
  - (control groups will get SPS later)
  - Less efficient, but politically feasible
  - Reduced interference between units
  - Frequently used: 68% of political science experiments; "exponential increase" in public health and medicine; common in public policy, education, psychology, etc.
- Matched-Pair (rather than complete) randomization

Matched-Pair, Cluster-Randomized Design

- Random assignment of clusters (rather than individuals)
  - (control groups will get SPS later)
  - Less efficient, but politically feasible
  - Reduced interference between units
  - Frequently used: 68% of political science experiments; "exponential increase" in public health and medicine; common in public policy, education, psychology, etc.
- Matched-Pair (rather than complete) randomization
  - both are unbiased (i.e., on average)

Matched-Pair, Cluster-Randomized Design

- Random assignment of clusters (rather than individuals)
  - (control groups will get SPS later)
  - Less efficient, but politically feasible
  - Reduced interference between units
  - Frequently used: 68% of political science experiments; "exponential increase" in public health and medicine; common in public policy, education, psychology, etc.
- Matched-Pair (rather than complete) randomization
  - both are unbiased (i.e., on average)
  - More efficient: up to 38 times more in these data!

Matched-Pair, Cluster-Randomized Design

- Random assignment of clusters (rather than individuals)
  - (control groups will get SPS later)
  - Less efficient, but politically feasible
  - Reduced interference between units
  - Frequently used: 68% of political science experiments; "exponential increase" in public health and medicine; common in public policy, education, psychology, etc.
- Matched-Pair (rather than complete) randomization
  - both are unbiased (i.e., on average)
  - More efficient: up to 38 times more in these data!
  - Smaller standard errors: up to 6 times smaller

Detailed Design Summary

Detailed Design Summary

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

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)

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.

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)

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

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

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 ≈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)

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

Remaining in study: 148 clusters in 7 states

Complete vs. Matched-Pair Random Assignment

Complete vs. Matched-Pair Random Assignment

Goal: equivalence between treatment and control groups

Complete vs. Matched-Pair Random Assignment

Goal: equivalence between treatment and control groups

Complete Randomization

Complete vs. Matched-Pair Random Assignment

Goal: equivalence between treatment and control groups

Complete Randomization

- Equivalence on average (or with large n) and nothing goes wrong

Complete vs. Matched-Pair Random Assignment

Goal: equivalence between treatment and control groups

Complete Randomization

- Equivalence on average (or with large n) and nothing goes wrong
- But, if we lose even one unrepresentative cluster:

Complete vs. Matched-Pair Random Assignment

Goal: equivalence between treatment and control groups

Complete Randomization

- Equivalence on average (or with large n) and nothing goes wrong
- But, if we lose even one unrepresentative cluster:
  - Equivalence of treated and control clusters fails

Complete vs. Matched-Pair Random Assignment

Goal: equivalence between treatment and control groups

Complete Randomization

- Equivalence on average (or with large n) and nothing goes wrong
- But, if we lose even one unrepresentative cluster:
  - Equivalence of treated and control clusters fails
  - The key benefit of random assignment (unbiasedness) is lost

Complete vs. Matched-Pair Random Assignment

Goal: equivalence between treatment and control groups

Complete Randomization

- Equivalence on average (or with large n) and nothing goes wrong
- But, if we lose even one unrepresentative cluster:
  - Equivalence of treated and control clusters fails
  - The key benefit of random assignment (unbiasedness) is lost
  - (E.g., are poor, unhealthy clusters are more likely to drop out?)

Complete vs. Matched-Pair Random Assignment

Goal: equivalence between treatment and control groups

Complete Randomization

- Equivalence on average (or with large n) and nothing goes wrong
- But, if we lose even one unrepresentative cluster:
  - Equivalence of treated and control clusters fails
  - The key benefit of random assignment (unbiasedness) is lost
  - (E.g., are poor, unhealthy clusters are more likely to drop out?)

Matched Pair Randomization

Complete vs. Matched-Pair Random Assignment

Goal: equivalence between treatment and control groups

Complete Randomization

- Equivalence on average (or with large n) and nothing goes wrong
- But, if we lose even one unrepresentative cluster:
  - Equivalence of treated and control clusters fails
  - The key benefit of random assignment (unbiasedness) is lost
  - (E.g., are poor, unhealthy clusters are more likely to drop out?)

Matched Pair Randomization

1. Matching: in-sample control for observed confounders

Complete vs. Matched-Pair Random Assignment

Goal: equivalence between treatment and control groups

Complete Randomization

- Equivalence on average (or with large n) and nothing goes wrong
- But, if we lose even one unrepresentative cluster:
  - Equivalence of treated and control clusters fails
  - The key benefit of random assignment (unbiasedness) is lost
  - (E.g., are poor, unhealthy clusters are more likely to drop out?)

Matched Pair Randomization

1. Matching: in-sample control for observed confounders
2. Random assignment: control for all confounders on average

Complete vs. Matched-Pair Random Assignment

Goal: equivalence between treatment and control groups

Complete Randomization

- Equivalence on average (or with large n) and nothing goes wrong
- But, if we lose even one unrepresentative cluster:
  - Equivalence of treated and control clusters fails
  - The key benefit of random assignment (unbiasedness) is lost
  - (E.g., are poor, unhealthy clusters are more likely to drop out?)

Matched Pair Randomization

1. Matching: in-sample control for observed confounders
2. Random assignment: control for all confounders on average
3. Pairing: provides failure safeguard: drop entire pair, and remaining treatment and control groups remain equivalent (if pair was lost due to pre-treatment variables)

Complete vs. Matched-Pair Random Assignment

Goal: equivalence between treatment and control groups

Complete Randomization

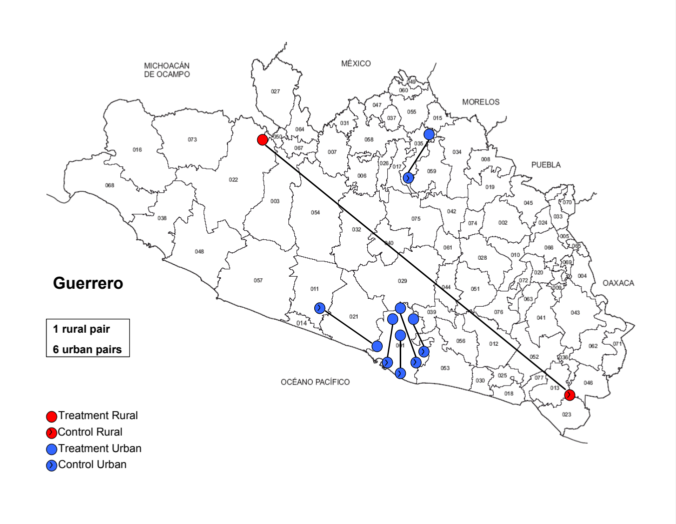
- Equivalence on average (or with large n) and nothing goes wrong
- But, if we lose even one unrepresentative cluster:
  - Equivalence of treated and control clusters fails
  - The key benefit of random assignment (unbiasedness) is lost
  - (E.g., are poor, unhealthy clusters are more likely to drop out?)

Matched Pair Randomization

1. Matching: in-sample control for observed confounders
2. Random assignment: control for all confounders on average
3. Pairing: provides failure safeguard: drop entire pair, and remaining treatment and control groups remain equivalent (if pair was lost due to pre-treatment variables)

All experiments should use matched pairs when feasible

Matched Pairs, Guerrero

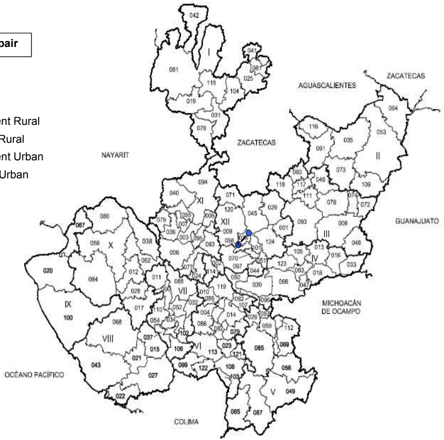


Matched Pairs, Jalisco

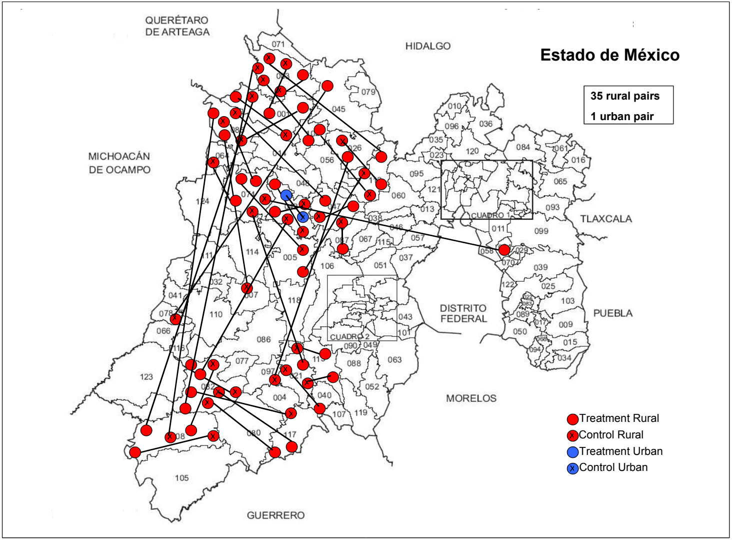
Jalisco

1 urban pair

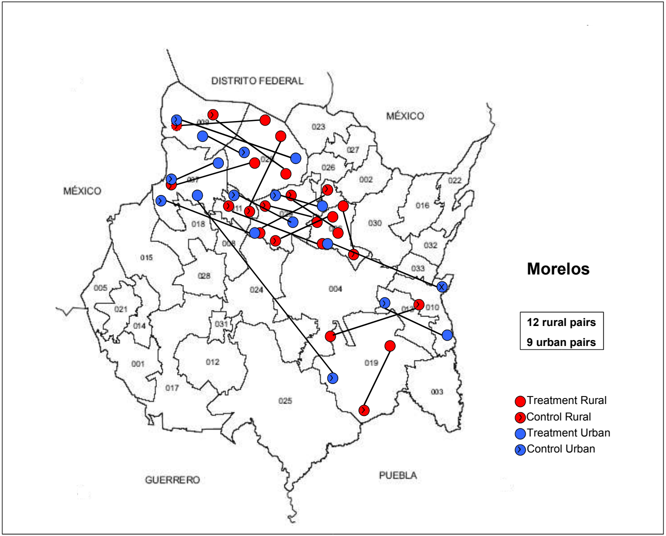
Treatment Rural
Control Rural
Treatment Urban
Control Urban



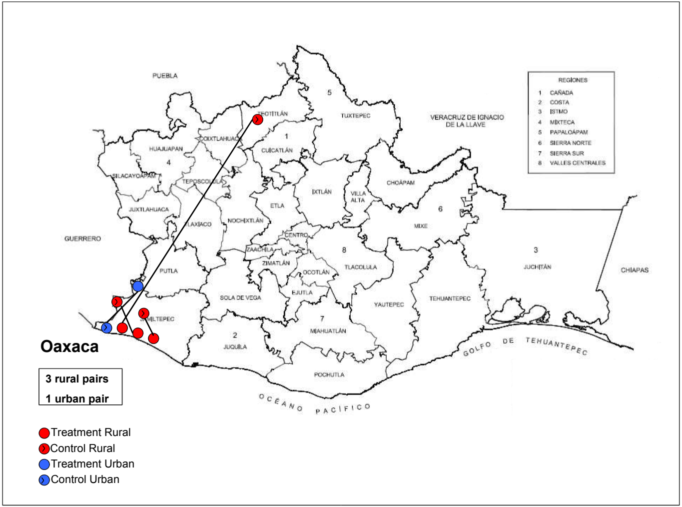
Matched Pairs, Estado de México



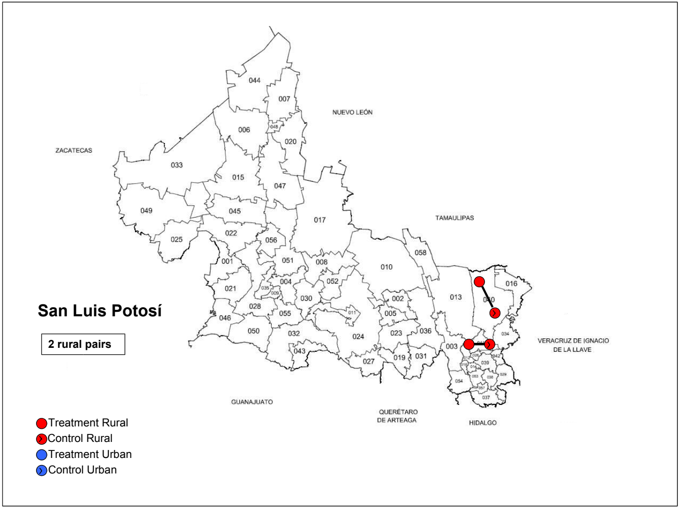
Matched Pairs, Morelos



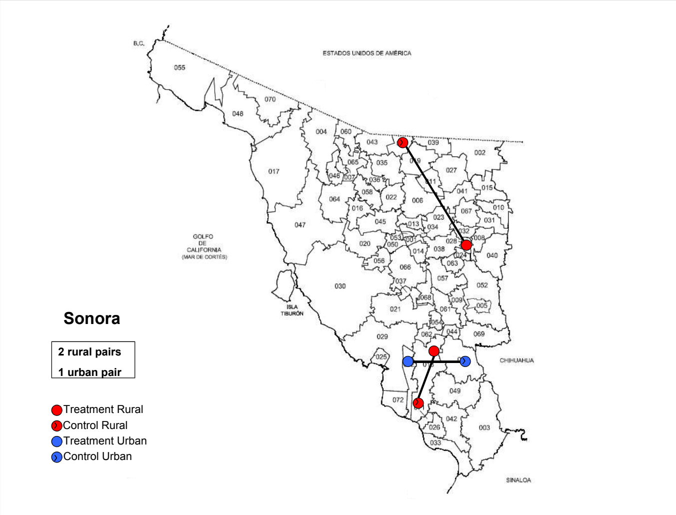
Matched Pairs, Oaxaca



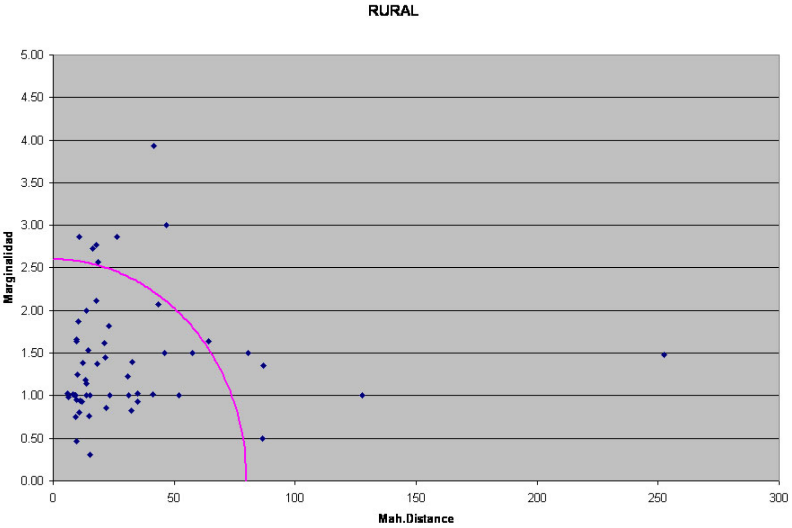
Matched Pairs, San Luis Potosí



Matched Pairs, Sonora



Choosing Pairs for the Survey



Design and Analysis Strategy is Triply Robust

Design and Analysis Strategy is Triply Robust

Design has three parts

Design and Analysis Strategy is Triply Robust

Design has three parts

1. Matching pairs on observed covariates

Design and Analysis Strategy is Triply Robust

Design has three parts

1. Matching pairs on observed covariates
2. Randomization of treatment within pairs

Design and Analysis Strategy is Triply Robust

Design has three parts

1. Matching pairs on observed covariates
2. Randomization of treatment within pairs
3. Parametric analysis adjusts for remaining covariate differences

Design and Analysis Strategy is Triply Robust

Design has three parts

1. Matching pairs on observed covariates
2. Randomization of treatment within pairs
3. Parametric analysis adjusts for remaining covariate differences

Triple Robustness

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

Design and Analysis Strategy is Triply Robust

Design has three parts

1. Matching pairs on observed covariates
2. Randomization of treatment within pairs
3. Parametric analysis adjusts for remaining covariate differences

Triple Robustness

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

Two Additional Checks if Triple Robustness Fails

Design and Analysis Strategy is Triply Robust

Design has three parts

1. Matching pairs on observed covariates
2. Randomization of treatment within pairs
3. Parametric analysis adjusts for remaining covariate differences

Triple Robustness

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

Two Additional Checks if Triple Robustness Fails

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

Design and Analysis Strategy is Triply Robust

Design has three parts

1. Matching pairs on observed covariates
2. Randomization of treatment within pairs
3. Parametric analysis adjusts for remaining covariate differences

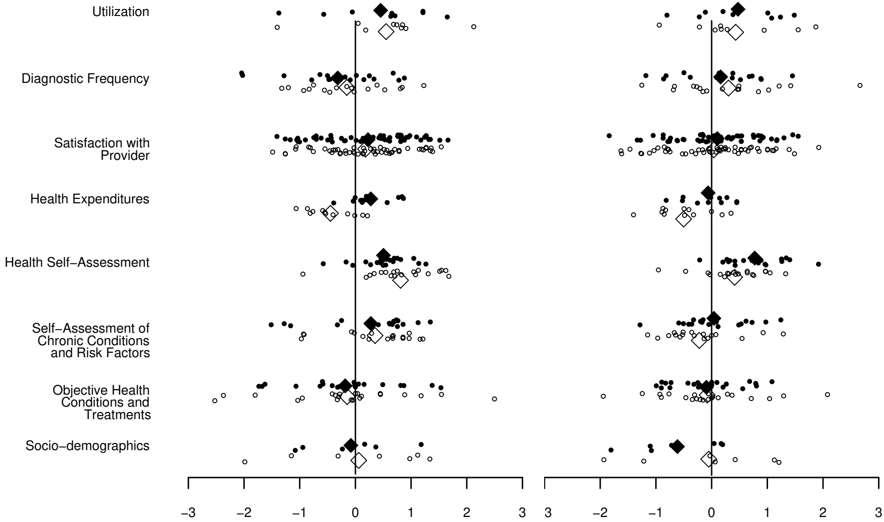
Triple Robustness

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

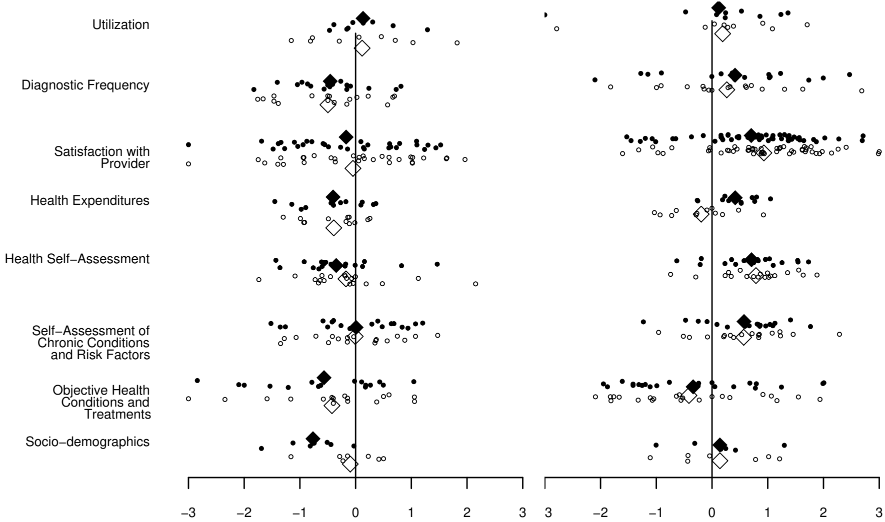
Two Additional Checks if Triple Robustness Fails

1. If one of the three works, then "effect of SPS" 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

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



ITT on Outcome Measures at Baseline, for wealthy families (left) and middle income families (right)



Matched-Pair Cluster-Randomized Designs in Polisci

Matched-Pair Cluster-Randomized Designs in Polisci

- Special research designs require special methods

Matched-Pair Cluster-Randomized Designs in Polisci

- Special research designs require special methods
- Prop. of polisci CREs which ignore the design:

Matched-Pair Cluster-Randomized Designs in Polisci

- Special research designs require special methods
- Prop. of polisci CREs which ignore the design: 100%

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:

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%

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.:

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,

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,

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,

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,

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

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 MPs:

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 MPs: 0%

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 MPs: 0%
- Conclusion: we're leaving a lot of information on the table!

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 MPs: 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

For more information

http://GKing.Harvard.edu

Effect of SP Rollout at Baseline: 1 of many

(Expected effects at 10 months: small, medium, large)

Dependent Variable [mean; SD]

Skilled birth attendance [0.9; 0.13]
Cholesterol cov. [0.07; 0.08]
Diarrhea children [0.86; 0.12]
Resp Infection children [0.64; 0.2]
Cervical exam [0.22; 0.11]
Papsmear [0.29; 0.12]
Flu vaccine [0.19; 0.1]
Diabetes [0.46; 0.18]
Hypertension cov. [0.33; 0.11]
Antenatal care [0.51; 0.22]
Mammography [0.05; 0.04]
Glasses [0.13; 0.07]