The Essential Role of Pair Matching in Cluster-Randomized Experiments, with Application to the Mexican Universal Health Insurance Evaluation

Kosuke Imai
Princeton University

Joint work with Gary King (Harvard) & Clayton Nall (Stanford)

September 6, 2012

Published in *Statistical Science* (2009) with discussions

Cluster-Randomized Experiments (CREs)

- Problem of many field experiments:
  - unit of randomization = clusters of individuals
  - unit of interest = individuals

- Public health and medicine:
  CREs have "risen exponentially since 1997" (Campbell, 2004)

- Political science:
  About 2/3 of field experiments are CREs

- Education:
  Randomization of classrooms and schools

Advantages of CRE

- Feasibility
- Cluster-level treatment
- Interference between units
  - Standard potential outcomes framework: $Y_i(T_i = 1)$ and $Y_i(T_i = 0)$
  - Potential outcomes of one unit may depend on treatment status of other units: many potential outcomes for each unit
  - Examples: peer effects, contagion, spill-over effects
  - Causal inference with such interference is notoriously difficult
  - Cluster randomization limits the number of potential outcomes: all units in the same cluster receives the treatment vs. no unit does
  - Avoids the interference problem rather than "solving" it

Main Disadvantage of CREs and Possible Solution

- Problem: Loss of efficiency

$$\text{CRE variance} \;=\; \text{usual variance} \times \{1 + (n-1)\rho\}$$

  where n is the cluster size and ρ is the intracluster correlation coefficient
- Number of clusters is often small

- Matched-Pair Designs (MPDs) to improve efficiency:
  1. Pair clusters based on background characteristics
  2. Within each pair, randomly assign one cluster to the treatment group and the other to the control group

- Idea: Eliminate as much difference between treated and control groups as possible before randomization of treatment assignment

Common Arguments Against MPDs

- "Analytical limitations" of MPDs (Klar and Donner, 1997):
  1. inability to test for homogeneity of causal effects across clusters
  2. difficulties in estimating the intracluster correlation coefficient
  3. Concerns about losing both clusters in a pair in event of randomization failure (Donner and Klar, 2000)
- In 10 or fewer pairs, MPDs can lose power (Martin *et al.* 1993)
- Our paper shows that these concerns are unfounded

Contributions of Our Paper

- Conclusion: **pair-matching should be used whenever feasible**
  - MPDs improve bias, efficiency, and power
  - Not pairing $=$ throwing away data!
- Existing estimator is based on a highly restrictive model
- Propose new simple design-based estimators and s.e.'s
- Demonstrate advantages using data from the Mexico study
- Present quantities of interest for CREs

Design-based Analysis of CREs under MPDs

- Existing Model-based approach: assume DGP for observed data
- The standard estimator assumes homogeneity across clusters
  $\Longrightarrow$ no point of matching to begin with!

- Our Design-based approach avoids modeling assumptions (Neyman, 1923)
- Randomness comes from:
  1. *randomization* of treatment assignment
  2. *random sampling* of clusters and units within clusters

- Recommendation: match on cluster sizes and prognostic covariates

Motivating Study: Seguro Popular de Salud (SPS)

- Article 4 of the Mexican constitution:

  all persons have a right to the protection of their health

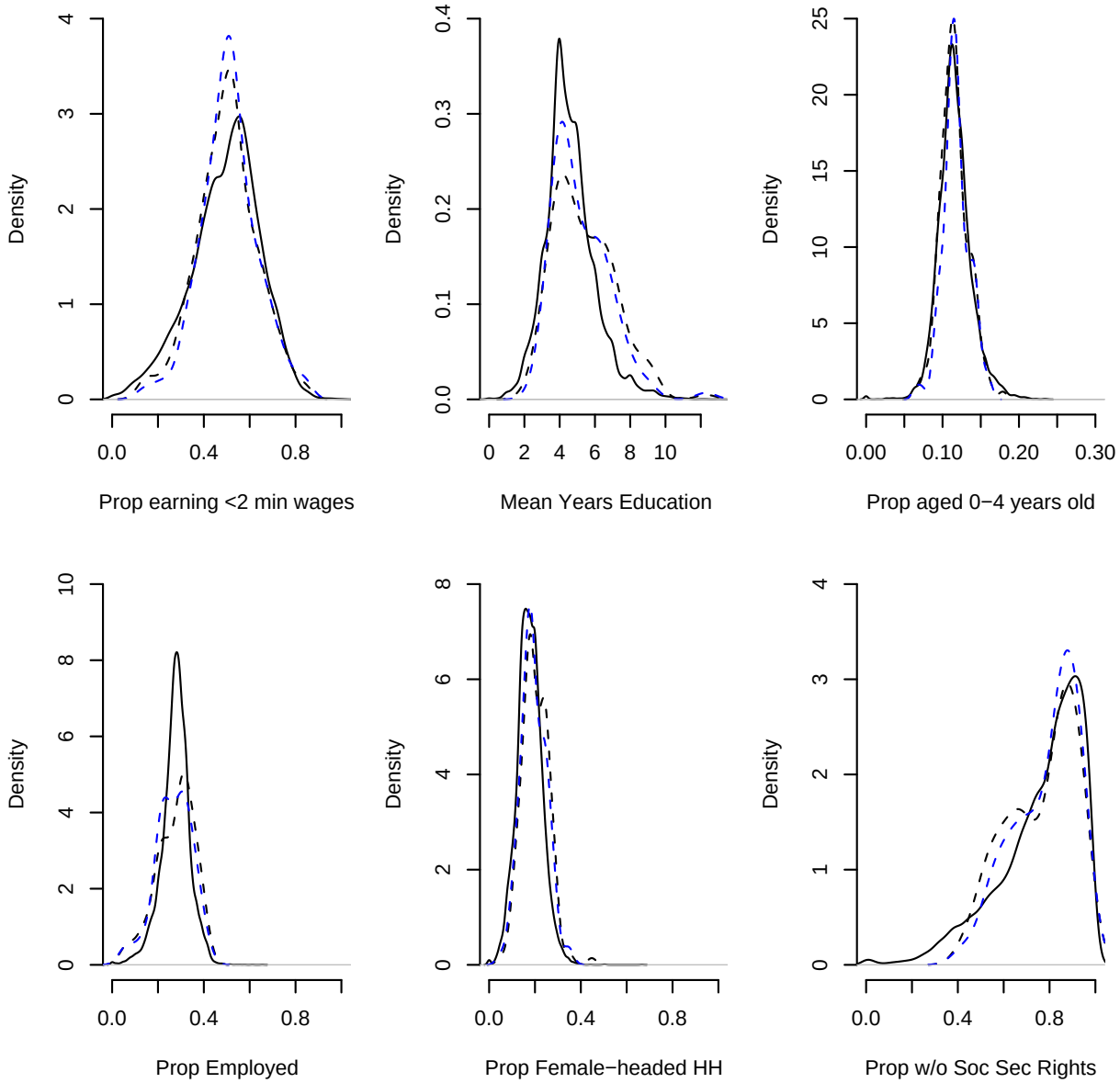
- SPS provides medical services, preventive care, pharmaceuticals, and financial health protection
- Voluntary and available for everyone but free to the poor
- Beneficiaries: intended to cover (by 2012) all 50M Mexicans who otherwise have no access to the healthcare system
- A key goal: reduce out-of-pocket health expenditures

- Randomized evaluation commissioned by the Fox administration
- One of the largest policy experiments to date

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

Clusters are Representative On Measured Variables



Quantities of Interest Depend on Sampling

Quantities	Clusters	Units within Clusters	Inferential Target
SATE	Observed	Observed	Observed sample
CATE	Observed	Sampled	Population within observed clusters
UATE	Sampled	Observed	Observable units within pop. of clusters
PATE	Sampled	Sampled	Population

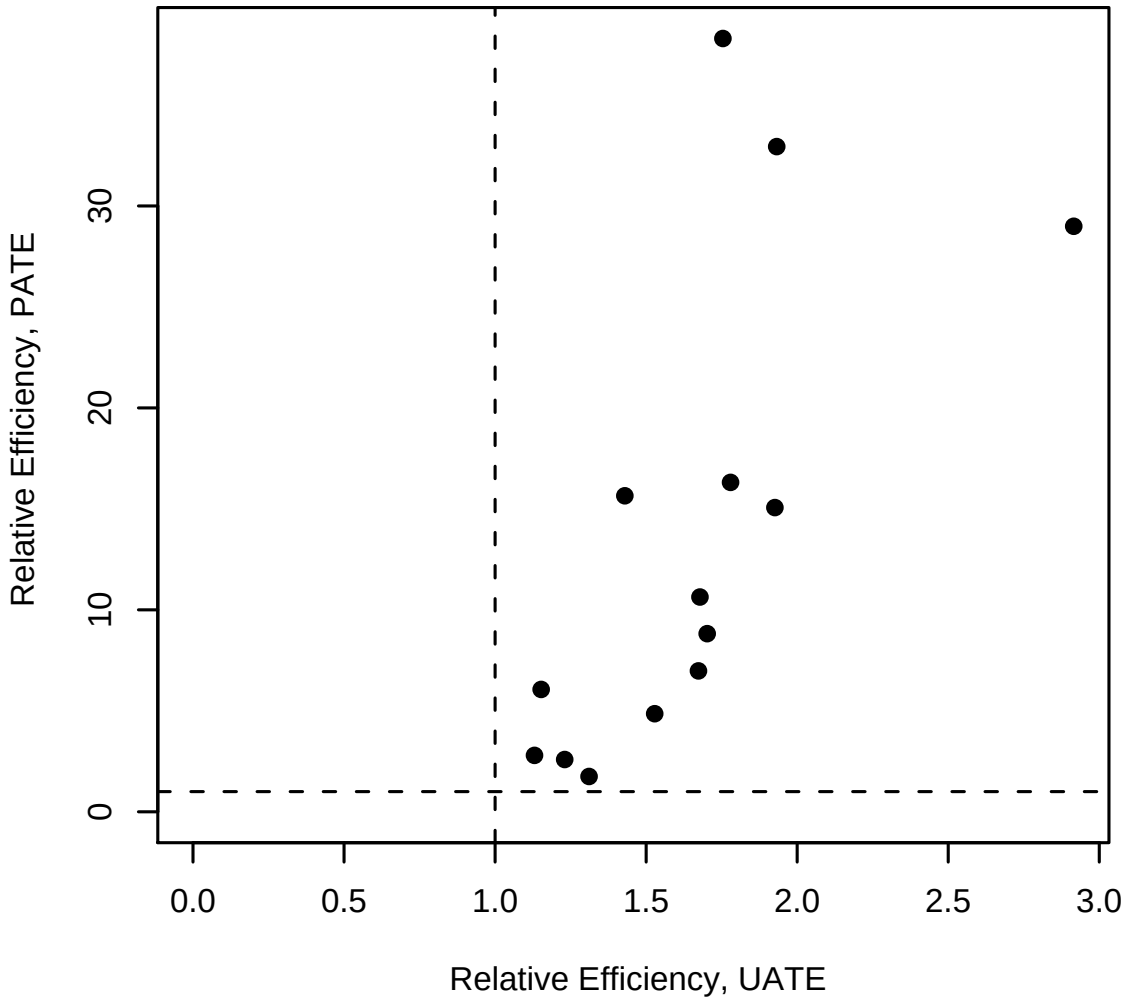
Main Finding: Effect of SPS on % of Households with Catastrophic Expenditures

	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

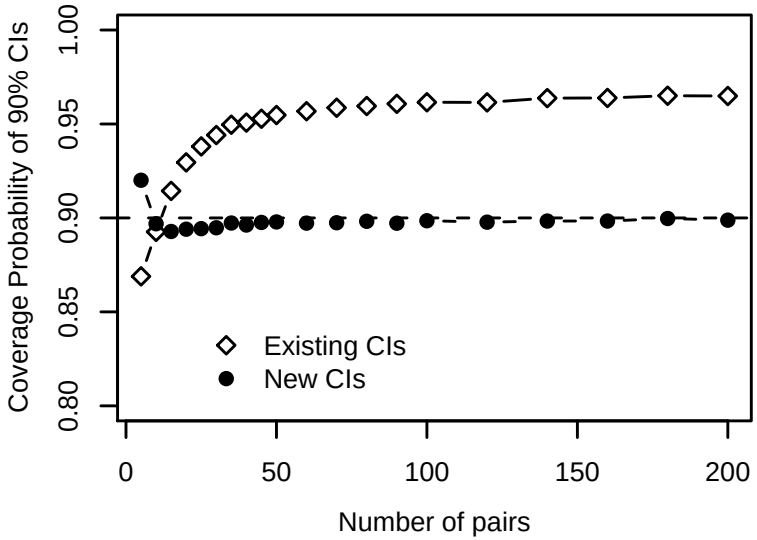
Efficiency Gains: MPD vs. Complete Randomization

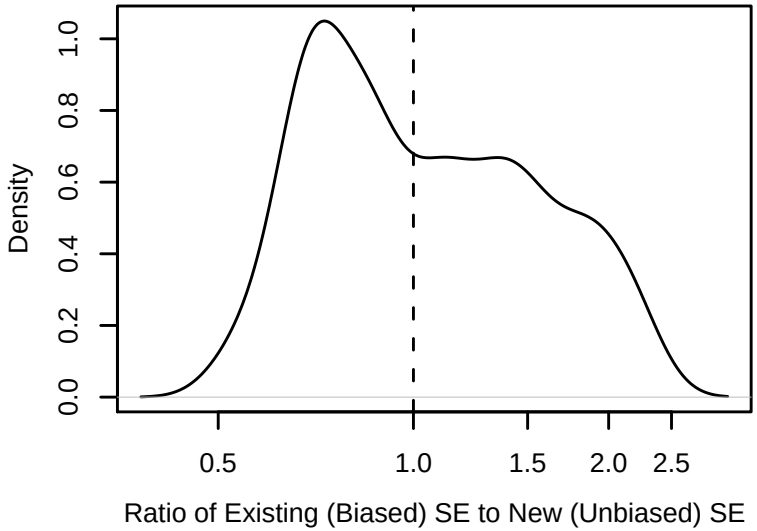
- Unit ATE: MPDs 1.1 to 2.9 times more efficient
- Population ATE: MPDs 1.8 to 38.3 times more efficient!

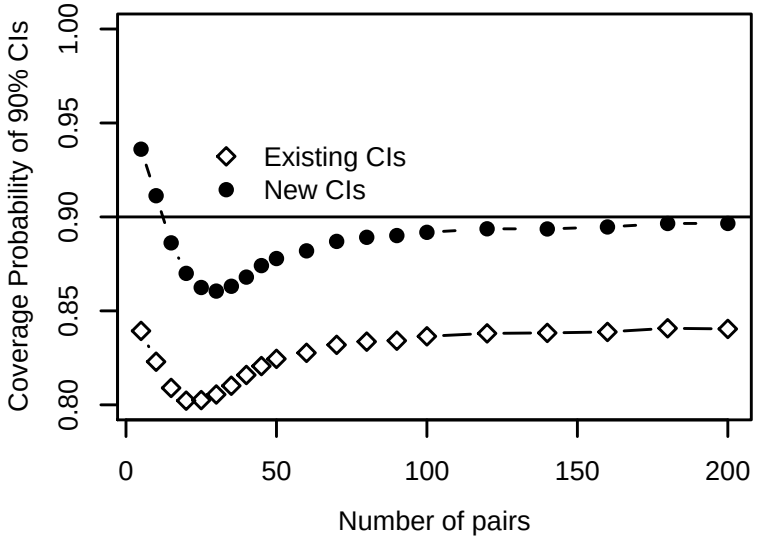


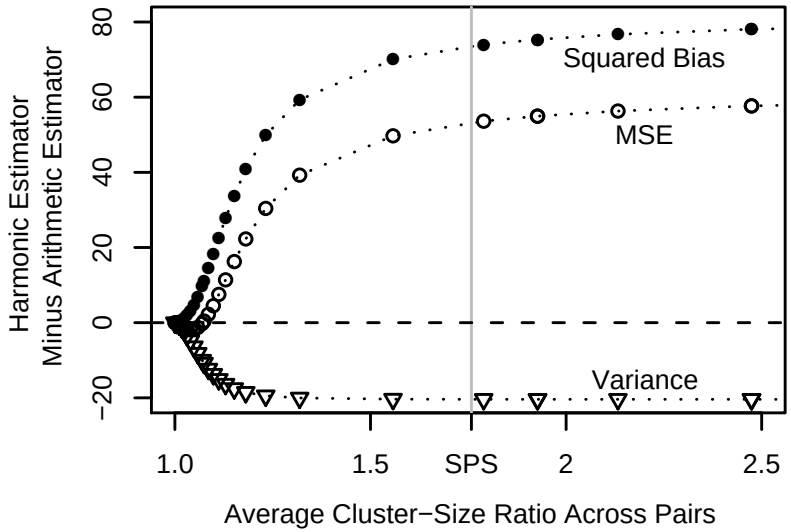
Bias and Inefficiency of Existing Approach

- Simulations Based on Mexico Data









Other Findings of SPS Evaluation

- Positive effects detected:
  - 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

Concluding Remarks

- Field experiments often require cluster randomization
- Problem: Loss of statistical efficiency

- Our recommendation: MPDs for CREs
  1. Select quantities of interest
  2. Identify pre-treatment covariates for matching
  3. Pair clusters based on the covariates and cluster sizes
  4. Randomize treatment within each pair
  5. Use design-based methods to analyze the data
- Our design-based estimators avoid modeling assumptions

- MPDs are preferred from perspectives of bias, efficiency, & power
- May affect CONSORT, Cochrane Collaboration, Council guidelines, etc.