

# Using Random Assignment to Evaluate Social Programs: Examples from the U.S.

Charles Michalopoulos  
MDRC

Presentation for annual conference of the Dansk EvalueringsSelskabs

11 September, 2009

# Origins of random assignment evaluations in U.S.

- Medicine: Need for evidence in regulating drugs
- Social policy:
  - 1960s: No hard evidence for shaping new policies
  - War on Poverty: Mandated and funded evaluations
- Evaluation of CETA training program
  - Non-experimental methods
  - Widely divergent results

# Results from CETA evaluation (non-experimental)

Group	Smallest estimated effect	Largest estimated effect
White women	408	778
White men	-4	500
Minority women	336	762
Minority men	-104	658

Results depend on statistical assumptions

Policymakers unsure what to conclude

# One reaction: More use of random assignment

- 1960-1970s
  - Negative Income Tax experiments
  - RAND health insurance study
  - National Supported Work Demonstration
- Challenge grants:
  - Foundation-funded random assignment evaluations of state policies in 1980s
  - Showed it could be done in real-world settings
- Large programs in some federal agencies require randomized evaluations
  - More than 30 major studies involving 300,000 people
- Training for disadvantaged workers
  - 1970-1985: 2 random assignment evaluations, 20 others
  - 1985-2000: 9 random assignment evaluations, 6 others

# Random assignment in many areas

- Training and education
- Welfare-to-work, employment retention and advancement
- Increasing employment among hard-to-employ groups
- Financial work incentives
- Youth programs
- Child development
- Health and disability
- Housing assistance
- Criminal justice
- Unemployment insurance

# Why use random assignment?

- Guaranteed unbiased estimate for the chosen sample
  - Nonexperimental methods subject to greater debate
- Statistical properties well understood
  - Less argument about statistical assumptions
- Easy to analyze and present
  - Simple comparisons of treatment and control groups

# Why use random assignment: qualitative evaluations are not enough\*

- Supported Work Demonstration
  - Staff ranked programs at 14 sites based on qualitative assessments
  - Result: no correlation between staff rankings and program effects
- New Chance Demonstration
  - Staff systematically assessed fidelity of program in 16 sites
  - More faithfully implemented programs had better outcomes, but not better impacts
- Job Corps
  - Performance measurement calculations were made in each site
  - No relationship between performance measurement and effects

\* As described in Hollister (2009), *Journal of Policy Analysis and Management*, 27(3): 611-615

# Why use random assignment: outcomes are not program impacts

Site	Outcome: 4 <sup>th</sup> Quarter Employment		Impact*
	Program Group	Control Group	
Grand Rapids MI	67% (#1)	61%	6% points, or 10% (#2)
Columbus OH	65% (#2)	63%	2% points, or 3% (#4)
Atlanta GA	55% (#3)	50%	5% points, or 10% (#3)
Riverside CA	45% (#4)	37%	8% points, or 22% (#1)

\*All impacts are statistically significant

Conclusion: Rank on outcomes not the same as rank on program impacts

# When random assignment might be ethically used

- Testing addition to a program
  - Example: health benefits for disabled
- Too many applicants for a program
  - Example: National Guard residential youth program
- When cutting a program, hold a group harmless
  - Example: testing welfare time limit
- Not clear which policy is superior
  - Example: domestic violence

# Example: when a federal agency controls the program

## ■ Background

- Disability beneficiaries not eligible for public health insurance for 2 years

## ■ Policy issue

- Does providing health care benefits to new disability beneficiaries improve health and increase employment?
- Do employment services enhance impact of health care benefits?

## ■ Steps in carrying out study

- Expert panel helped design health benefits package
- Decided to include only people without insurance, not close to retirement
- Obtained approval of institutional review board
- Individuals recruited by telephone interviews
- Interviewers get assignment from computerized system
- Data collected through surveys and administrative records

# Example: working with a local agency

## ■ Background

- Families receiving child care subsidies pay part of cost of care
- Amount increases with income, which might discourage employment

## ■ Policy issue

- Does reducing copayment encourage increased work?

## ■ Steps in carrying out study

- Convinced state to participate in the study, pay for reduced copayments
- Chose experimental copayment schedule
- At state request, included all applicants statewide
- Obtained approval from state review board
- Trained local staff to know about new copayment amounts
- Random assignment done by state using algorithm provided by us
- Outcomes from administrative data (survey failed)

# Example: a program created for purposes of an evaluation

## ■ Background

- Low-income couples divorce at higher rates than other couples
- Children of divorced couples do worse than others

## ■ Policy issue

- Does helping couples have better relationships keep them together and benefit their children?

## ■ Steps in carrying out study

- Experts in marriage helped design program
- Federal government funded programs
- Eight sites chosen for evaluation after extensive search
- Pilot programs run before sites approved for study
- Individuals recruited by local program staff
- Randomization via website (usually not staff providing the program)
- Data collected through surveys and observations

# Example: using excess applications to carry out random assignment

## ■ Background

- Troubled youth more likely to drop out of school and fare worse in life
- National Guard runs residential program for these youth

## ■ Policy issue

- Does the National Guard program help youth graduate from high school, find work, and stay out of trouble with the criminal justice system?

## ■ Steps in carrying out study

- Program already existed and funded by federal government
- Had to convince high-level staff the study was important
- Sites chosen for the evaluation only if they were over-enrolled
- Studying youth required consent from parents
- Randomization to program or control group happened whether or not youth agreed to be in the study
- Data collected through surveys

# Three cautions

- Randomization does not produce certainty
  - Importance of testing in several places
  - Theory still important
- Some questions require other methods
  - Example: can a policy be well implemented
  - Example: effects of a program run in the past
  - Example: difficult to measure entry effects
  - Example: cannot artificially reduce benefits for a study
- Random assignment not always ethical
  - Example: enough funding to provide services to everyone who is eligible

# For more information

[www.mdrc.org](http://www.mdrc.org)

[charles.michalopoulos@mdrc.org](mailto:charles.michalopoulos@mdrc.org)