Measuring the Cost Effects of SNAP Demonstrations

Possible Methodologies

Secretary’s Innovation Group
Milwaukee, WI
August 2, 2013

Douglas J. Besharov
School of Public Policy
University of Maryland

and
The Atlantic Council
with the assistance of
Jacob Klerman
Abt Associates
The Policy Change

Imposition of work-related requirements on non-disabled, working age adult SNAP recipients with children over 1.

Including

> job search,
> job preparation, and, possibly,
> job training.
Objective

Measure effects on costs and caseloads for up to five years

And, to the extent possible, employment and earnings.

SNAP Expenditures 1996-2012

Billions of dollars

© Douglas Besharov, 2013
Possible Methodologies

**Individual level (statewide sample)**
- Random assignment of individuals
- Matched individuals
- Propensity score matching

**County or office level**
- Random assignment of counties or offices
- Matched counties
- Difference-in-differences
- Stratified random assignment of counties
- Randomized phase-in

**Statewide**
- Interrupted time series
Randomized Experiments

Create a “control” group (not just a “comparison” group) by assigning individuals or other units of analysis to program and nonprogram groups based on chance (“at random” or the functional equivalent of a coin toss).
Random Assignment of Individuals
(Across state)

**Pros**

- Strong causal validity for *exit* effects only
  - *If* limited spillover and contamination
- Only modest holdback required (under 10 percent)

**Cons**

- Possible spillover/contamination
- Does not capture *entry* effects
- Imprecise estimates, if small holdback
- Can be difficult to administer (continuous randomization)
- Can be expensive because difficult to implement
- Holdback must continue for subsequent cost savings estimates
Matching Studies

Match individuals or other units of analysis in the program group with a similar nonprogram group, based on selected individual or aggregate observed variables that are presumed to influence participation in the program and/or the outcomes and impacts of interest.

Comparison group
Program group

SNAP expenditures at baseline
SNAP expenditures after one year

Program effect
Matched Individuals
(In other states)

**Pros**
- Avoids holdback

**Cons**
- Especially weak causal validity
- Does not capture entry effects
- Difficult coordination with other states

---

© Douglas Besharov, 2013
Create a comparison group from nonparticipants who have the characteristics that seem to have led those in the program to participate. Statistical techniques (often logit or probit) are used to estimate the probability of individuals or other units of analysis being in the program, and these “propensity scores” are then used to select the members of a comparison group.

Total Effect = Average of individual program effects
Propensity Score Matching

Showing more common support

Amount of annual SNAP benefits

Propensity Score

© Douglas Besharov, 2013
Propensity Score Matching  
(In other states)

**Pros**
- Avoids holdback

**Cons**
- Weak causal validity
- Does not capture entry effects
- Possible weak generalizability
- Difficult coordination with other states
Randomized Experiments

Create a “control” group (not just a “comparison” group) by assigning individuals or other units of analysis to program and nonprogram groups based on chance (“at random” or the functional equivalent of a coin toss).
Random Assignment of Counties (or Offices) (Within state)

**Pros**

- Strong causal validity
- Precise estimates, if sufficient holdback (10-20 percent)
- Captures entry effects
- Relatively inexpensive (about $200K+ per state)

**Cons**

- Possible spillover/contamination
- Imprecise estimates, if small holdback
- Requires substantial holdback (10-20 percent)
- Holdback must continue for subsequent cost savings estimates
Matching Studies

Match individuals or other units of analysis in the program group with a similar nonprogram group, based on selected individual or aggregate observed variables that are presumed to influence participation in the program and/or the outcomes and impacts of interest.

Comparison group
Program group

SNAP expenditures at baseline

SNAP expenditures after one year

Program effect
Matched Counties
(Within state or in other states)

Pros
• Captures entry effects
• Relatively inexpensive

Cons
• Weak causal validity (weaker if other states)
• Possible spillover/contamination
• Imprecise estimates, if small holdback
• Requires substantial holdback (10-20 percent)
• Holdback must continue for subsequent cost savings estimates
Difference-in-Differences Studies

Assume that there are unobserved differences between the program and comparison groups. They seek to use the preintervention differences in the outcome or impact measures to control for those unobserved differences.

**SNAP expenditures at baseline**

**SNAP expenditures after one year**

Comparison group

Program group

Program effect

Predicted difference if no program

© Douglas Besharov, 2013
Difference-in-Differences (Within state or in other states)

**Pros**

- Moderate-to-strong causal validity, if sufficient holdback within state (10-20 percent)
- Captures entry effects
- Relatively inexpensive

**Cons**

- Weak causal validity, if small holdback
- Weaker causal validity if in other states
- Possible spillover/contamination
- Imprecise estimates, if small holdback
- Requires substantial holdback (10-20 percent) unless in other states
- Holdback must continue for subsequent cost savings estimates
Stratified Random Assignment Studies

Divide the population into subgroups and then randomly draws a sample from each subgroup, proportional to the subgroup’s percentage of the total population. The sample is then randomly assigned.
Stratified Random Assignment of Counties (or Offices) (Within state)

**Pros**
- Very strong causal validity
- Precise estimates, if sufficient holdback (10-20 percent)
- Captures entry effects
- Relatively inexpensive

**Cons**
- Possible spillover/contamination
- Imprecise estimates, if small holdback
- Requires substantial holdback (10-20 percent)
- Holdback must continue for subsequent cost savings estimates

© Douglas Besharov, 2013
Randomized Phase-in Designs

Generate a control group based on variations in the timing of the intervention on the target population.

Control groups 1+2+3  Control groups 2+3  Control group 3

Program groups

Control groups

Baseline expenditures  +6 months expenditures  +12 months expenditures  +18 months expenditures

Program group 1

Program effect 1

Program group 2

Program effect 2

Program group 3

Program effect 3

Total Program Effect

© Douglas Besharov, 2013
Randomized Phase-in of Counties (or Offices) (Within state)

**Pros**
- Strong causal validity, if sufficient sample size
- Captures entry effects
- Requires only a temporary holdback
- Relatively inexpensive

**Cons**
- Possible spillover/contamination
- Requires timely and efficient implementation
- Imprecise estimates
- No estimates of cost savings after full implementation
- Effect takes more time to document
Interrupted Time Series Studies

Compare individuals, a changing population of individuals in the same program, or other units of analysis to themselves *over an extended period of time* before and after the intervention (the “interruption”)

![Graph showing SNAP Expenditures over time with an intervention effect highlighted.](image-url)
Interrupted Time Series Studies
With a comparison time series

Comparison State Time Series

SNAP Expenditures

Time

© Douglas Besharov, 2013
Interrupted Time Series (Statewide)

**Pros**
- Reasonable causal validity, if change is large and relatively immediate
- Captures entry effects
- No holdback required
- Relatively inexpensive

**Cons**
- Weak causal validity, if change is small and not immediate
- Requires timely and efficient implementation
- No estimates of cost savings after full implementation
<table>
<thead>
<tr>
<th>Methodology</th>
<th>Causal Validity</th>
<th>Entry Effects</th>
<th>Precision</th>
<th>Small or No Holdback</th>
<th>Low Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random assignment of individuals</td>
<td>?</td>
<td>N</td>
<td>?</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Matched individuals</td>
<td>N</td>
<td>N</td>
<td>?</td>
<td>Y</td>
<td>?</td>
</tr>
<tr>
<td>Propensity score matching</td>
<td>N</td>
<td>N</td>
<td>?</td>
<td>Y</td>
<td>?</td>
</tr>
<tr>
<td>Random assignment of counties or offices</td>
<td>Y</td>
<td>Y</td>
<td>?</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>Matched counties</td>
<td>N</td>
<td>Y</td>
<td>?</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>Difference-in-differences</td>
<td>?</td>
<td>Y</td>
<td>?</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>Stratified random assignment of counties</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>Randomized phase-in</td>
<td>?</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Interrupted time series</td>
<td>?</td>
<td>Y</td>
<td>?</td>
<td>Y</td>
<td>Y</td>
</tr>
</tbody>
</table>