

Measuring the Cost Effects of SNAP Demonstrations

Possible Methodologies

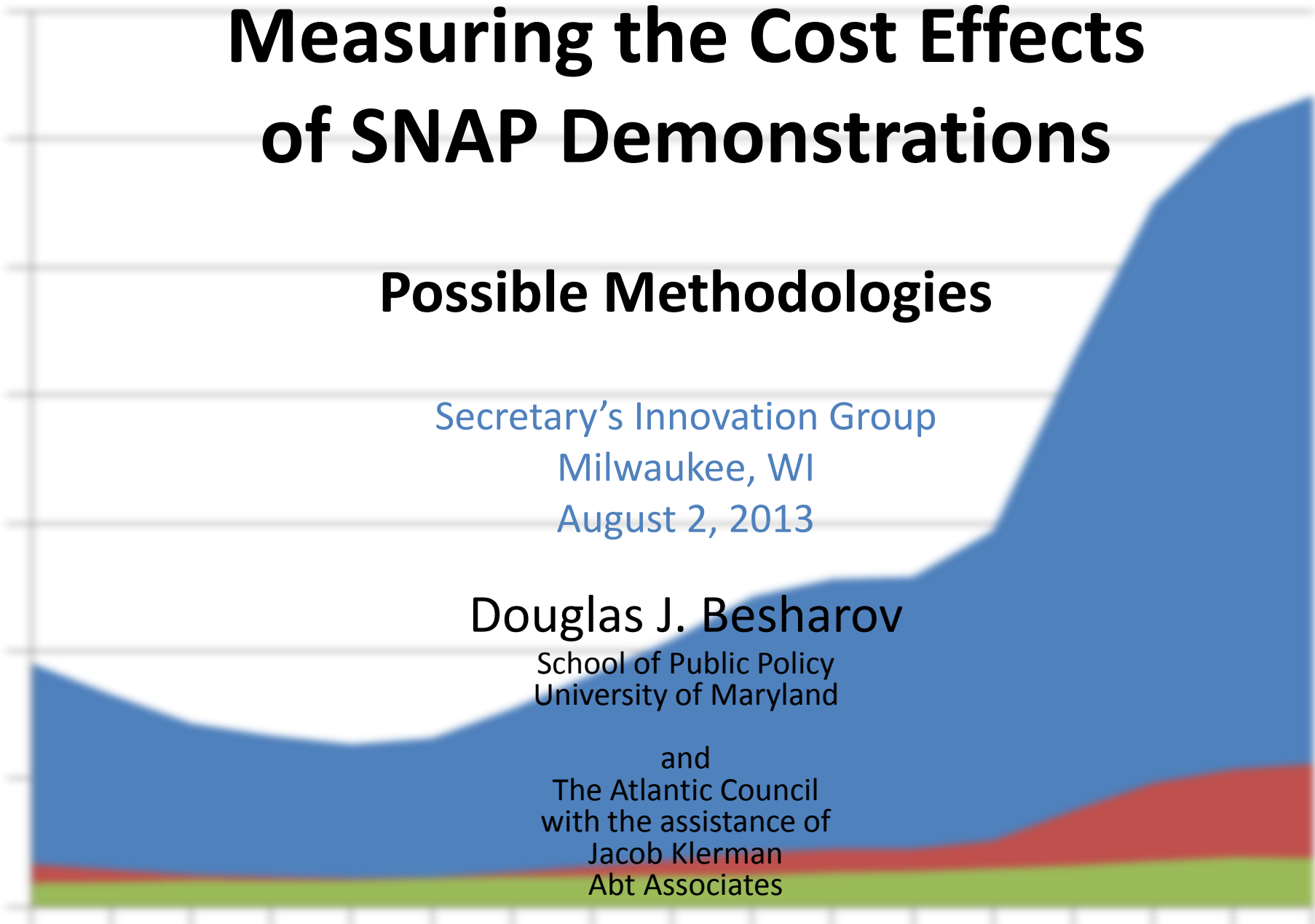
Secretary's Innovation Group
Milwaukee, WI
August 2, 2013

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.

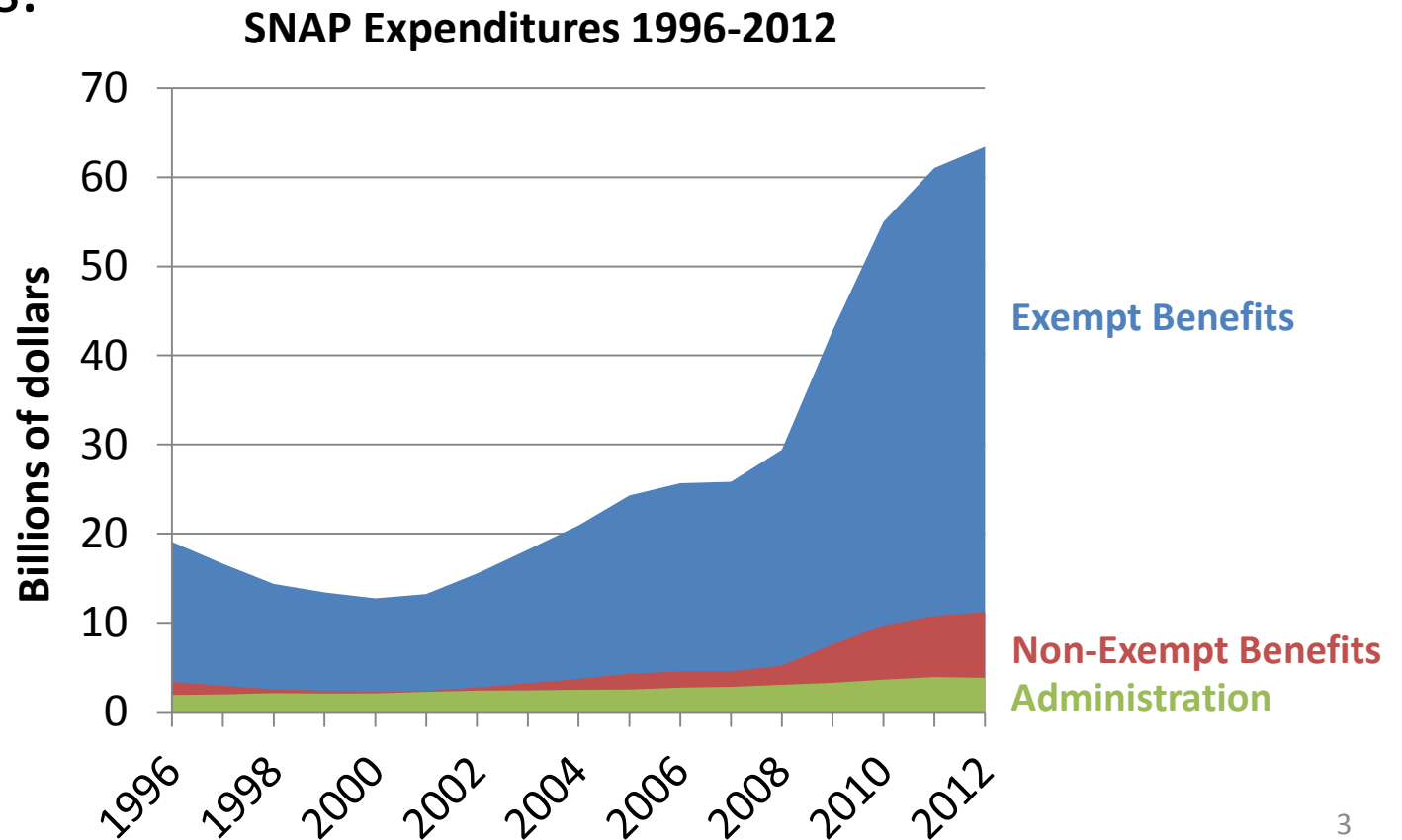
Including

- > 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.



Possible Methodologies

Individual level (statewide sample)

- Random assignment of individuals
- Matched individuals
- Propensity score matching

County or office level

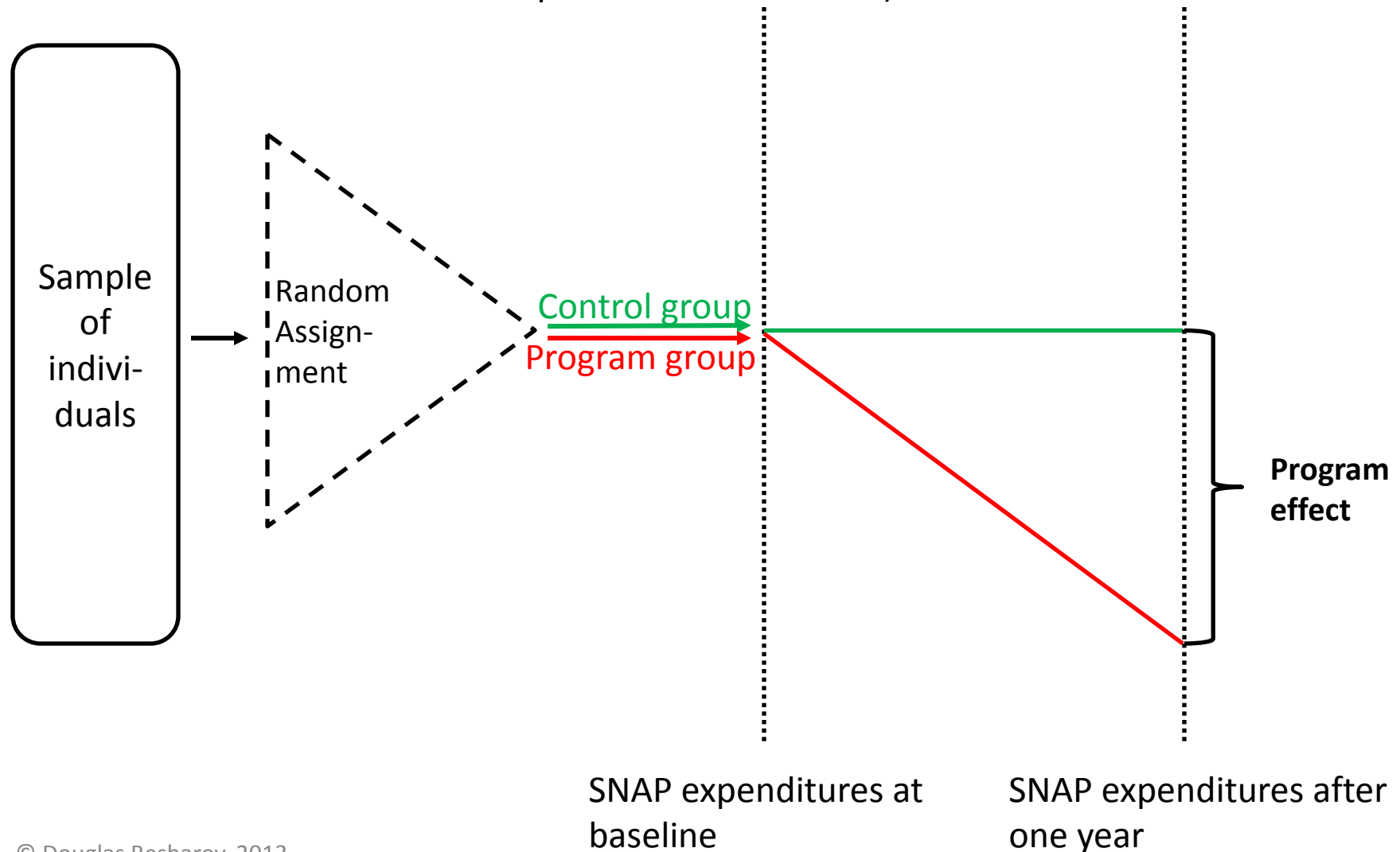
- Random assignment of counties or offices
- Matched counties
- Difference-in-differences
- Stratified random assignment of counties
- Randomized phase-in

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

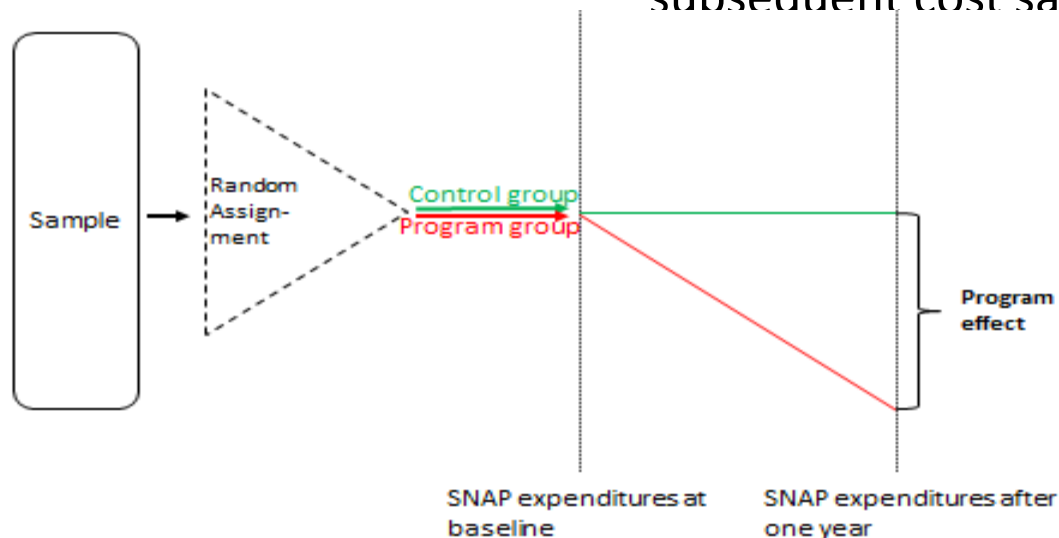
(Across state)

Pros

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

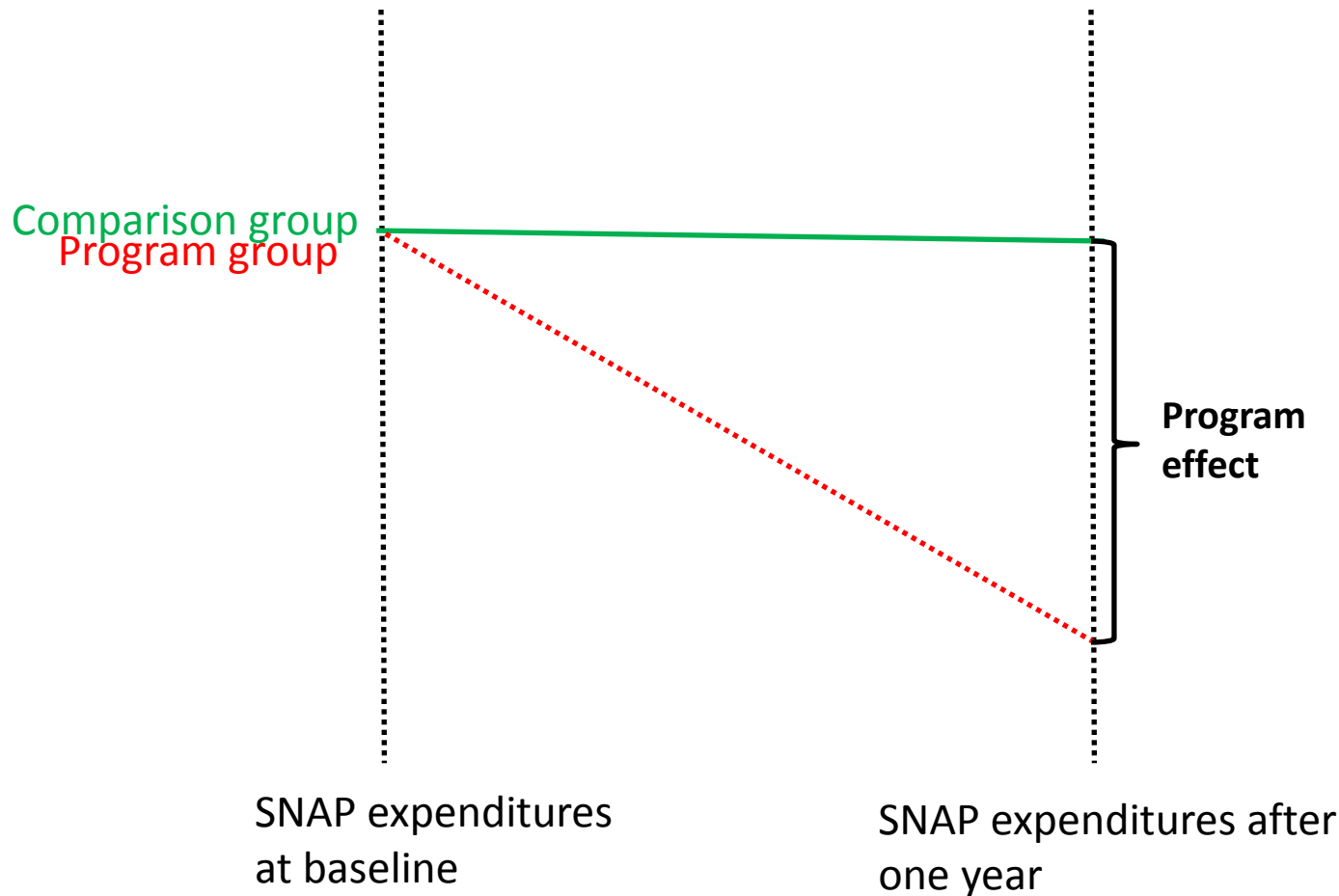
Cons

- 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.



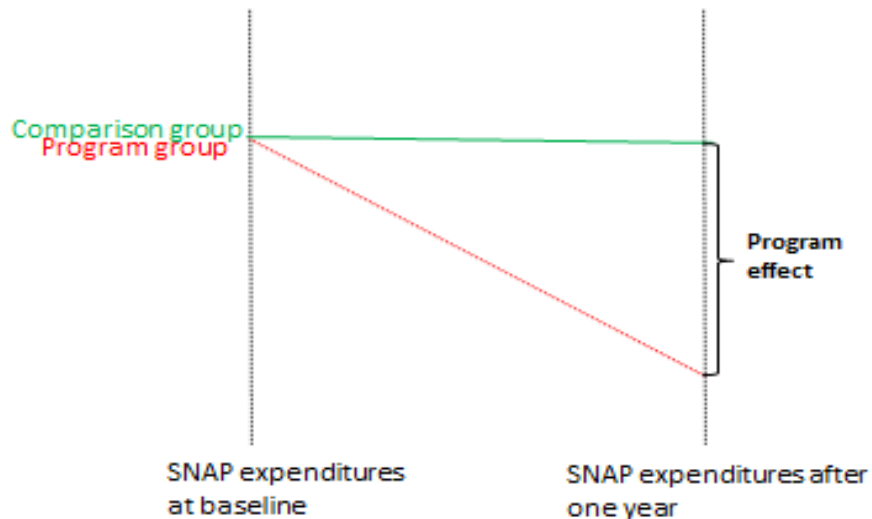
Matched Individuals (In other states)

Pros

- Avoids holdback

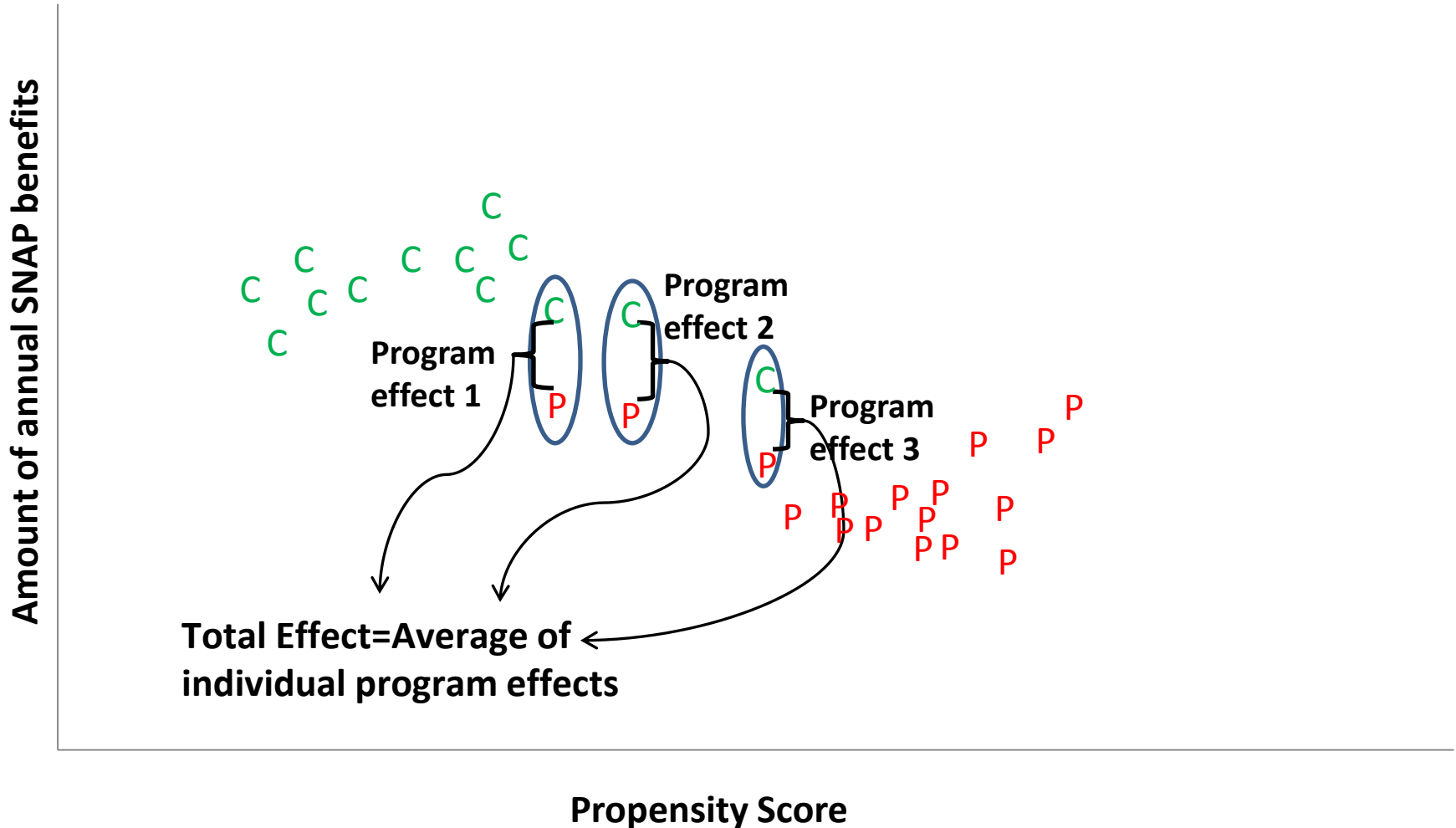
Cons

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



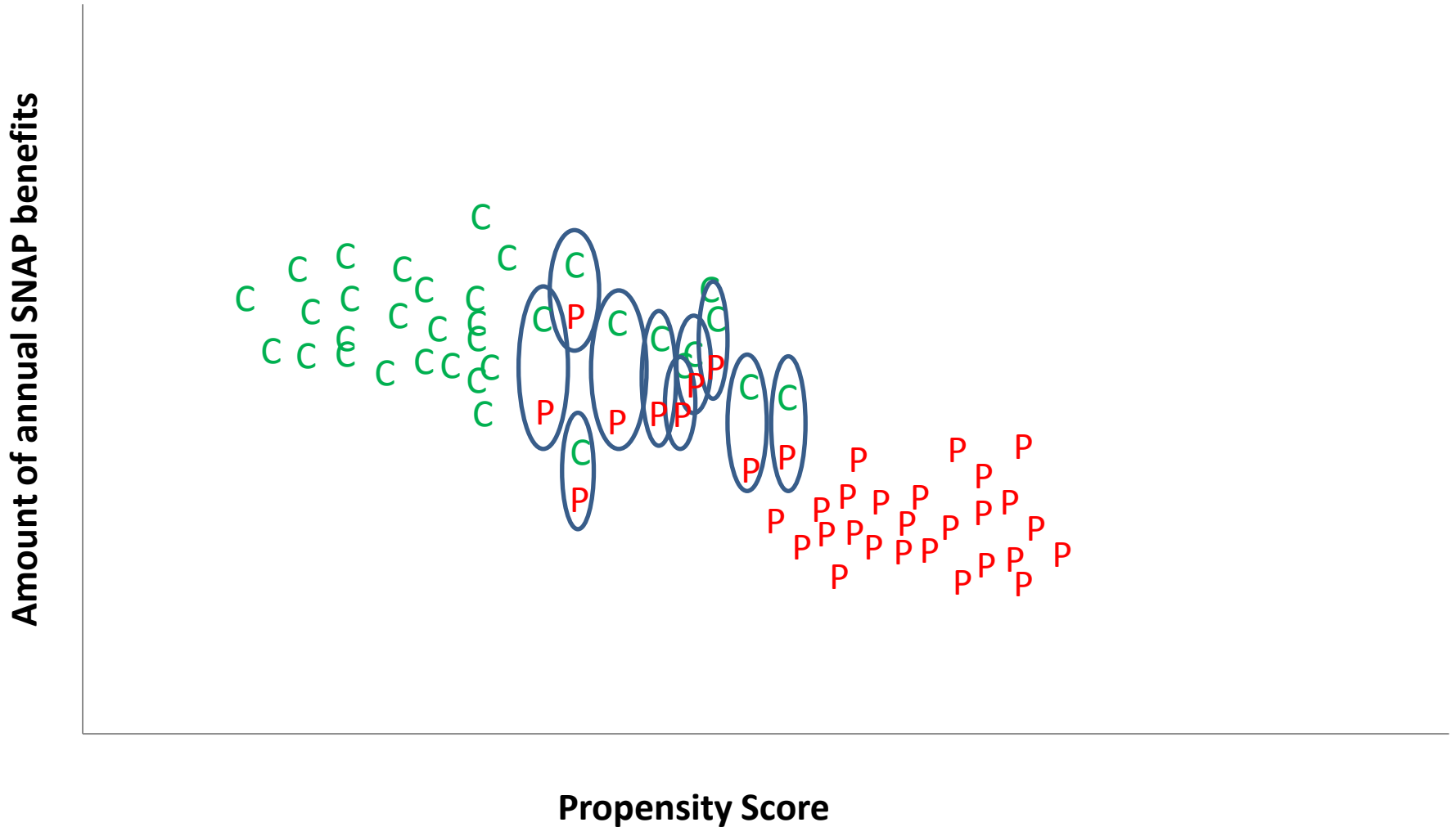
Propensity Score Matching

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.



Propensity Score Matching

Showing more common support



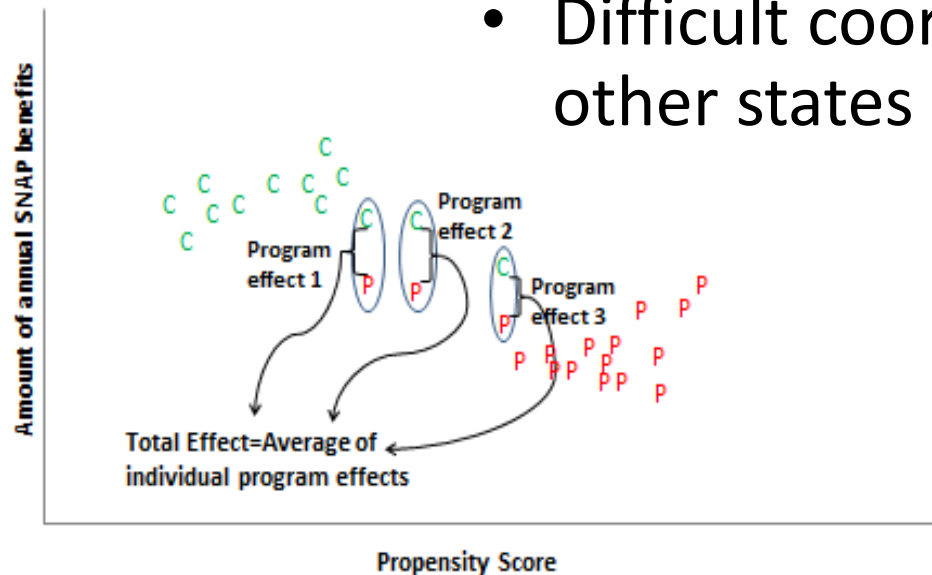
Propensity Score Matching (In other states)

Pros

- Avoids holdback

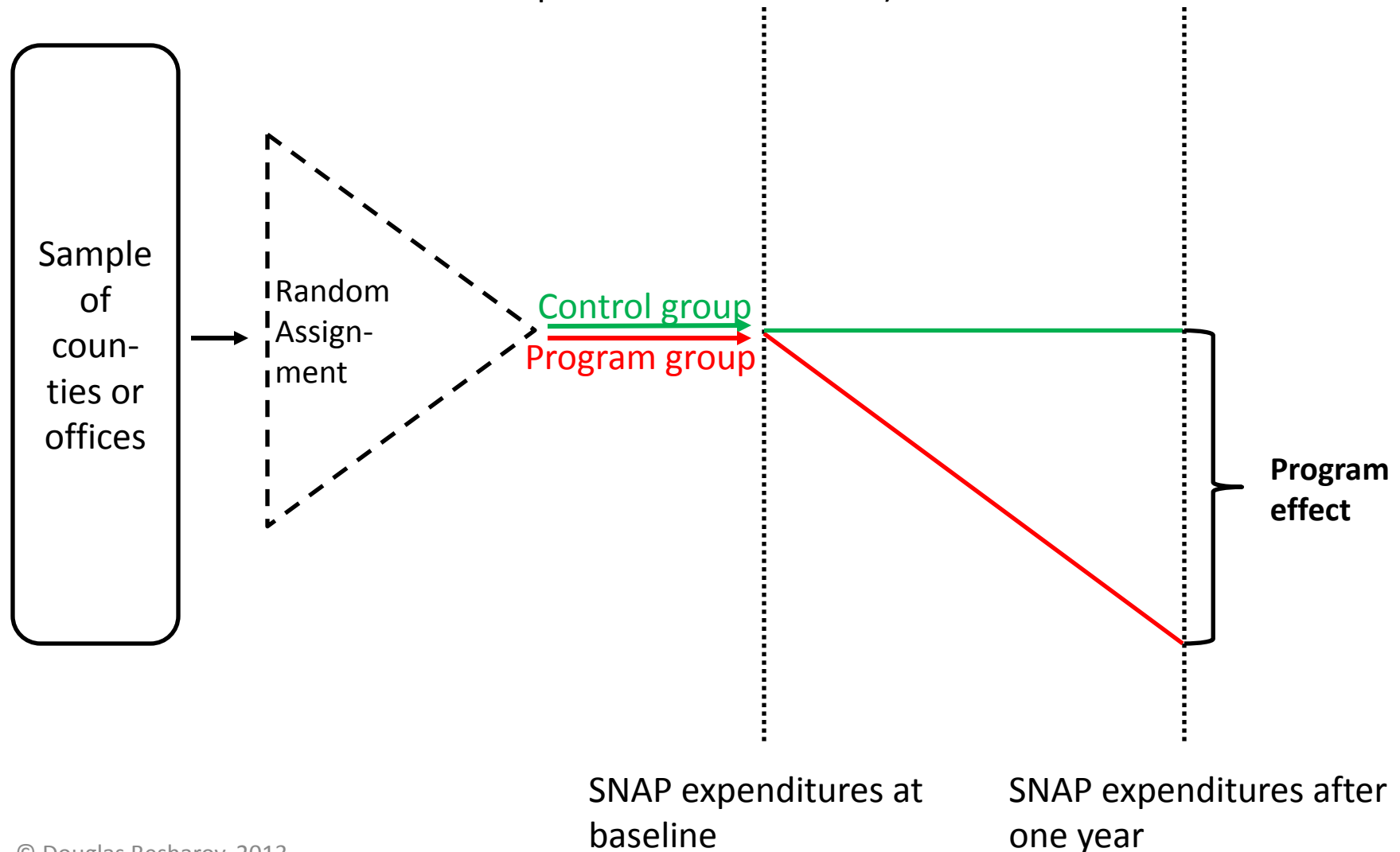
Cons

- 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).



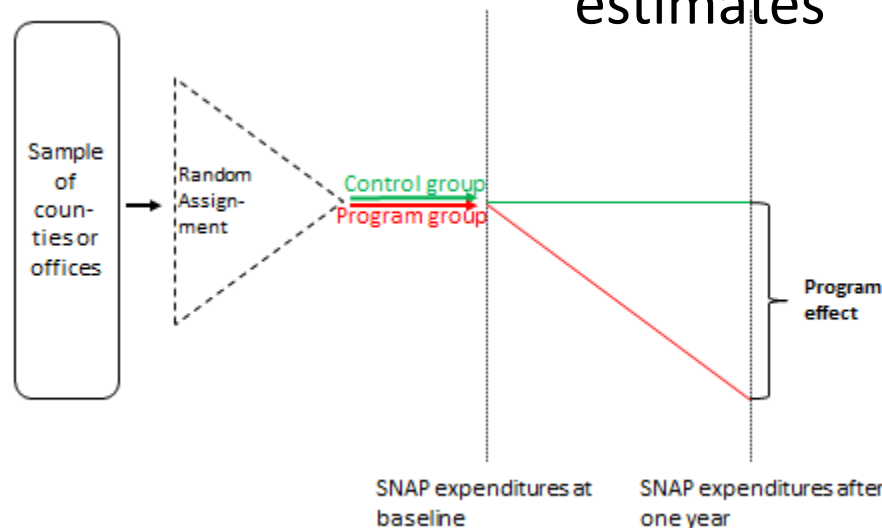
Random Assignment of Counties (or Offices) (Within state)

Pros

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

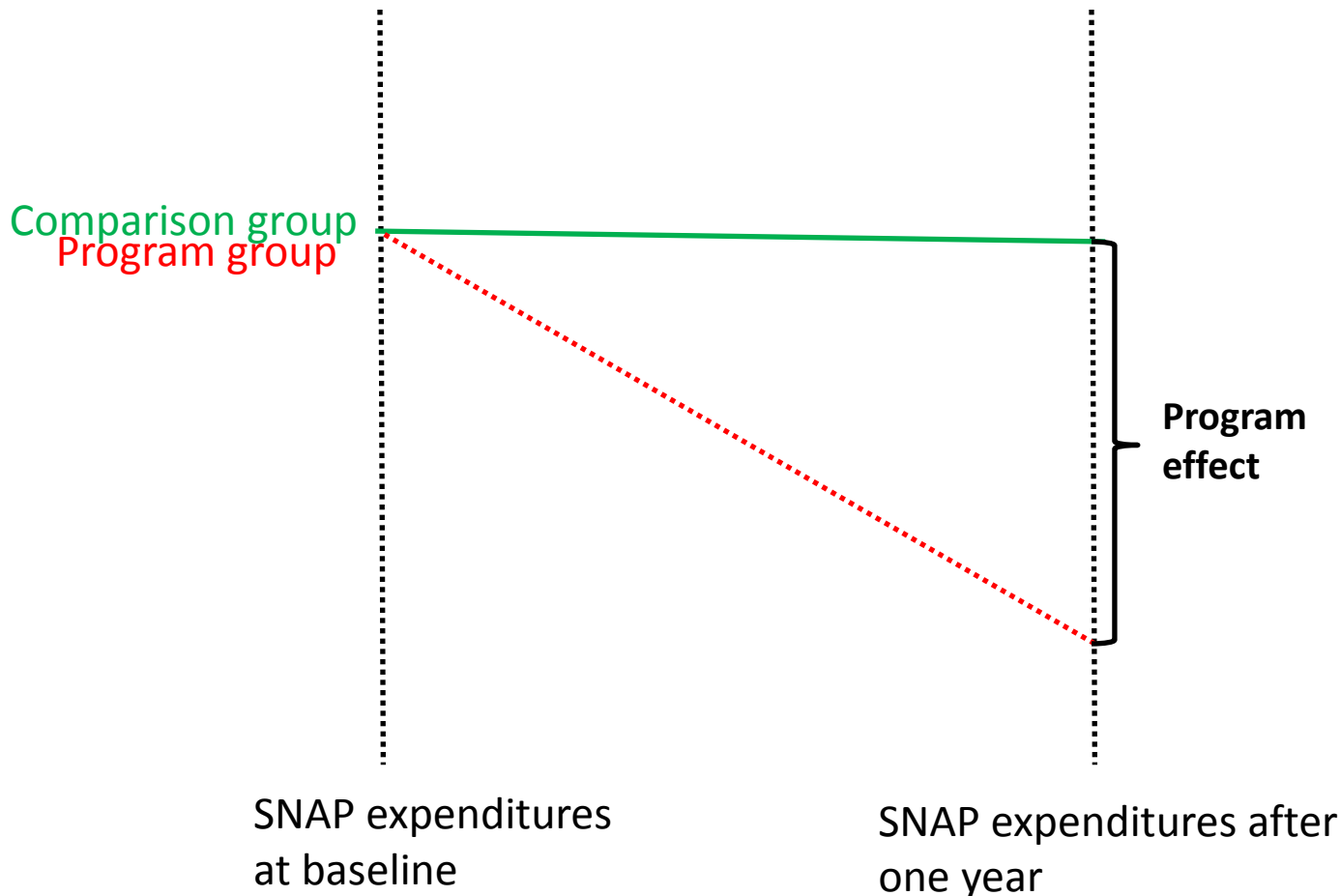
Cons

- 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.



Matched Counties

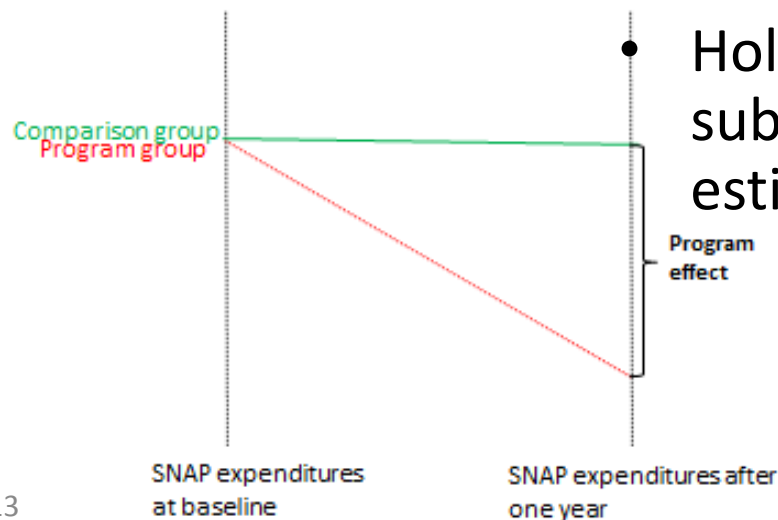
(Within state or in other states)

Pros

- Captures entry effects
- Relatively inexpensive

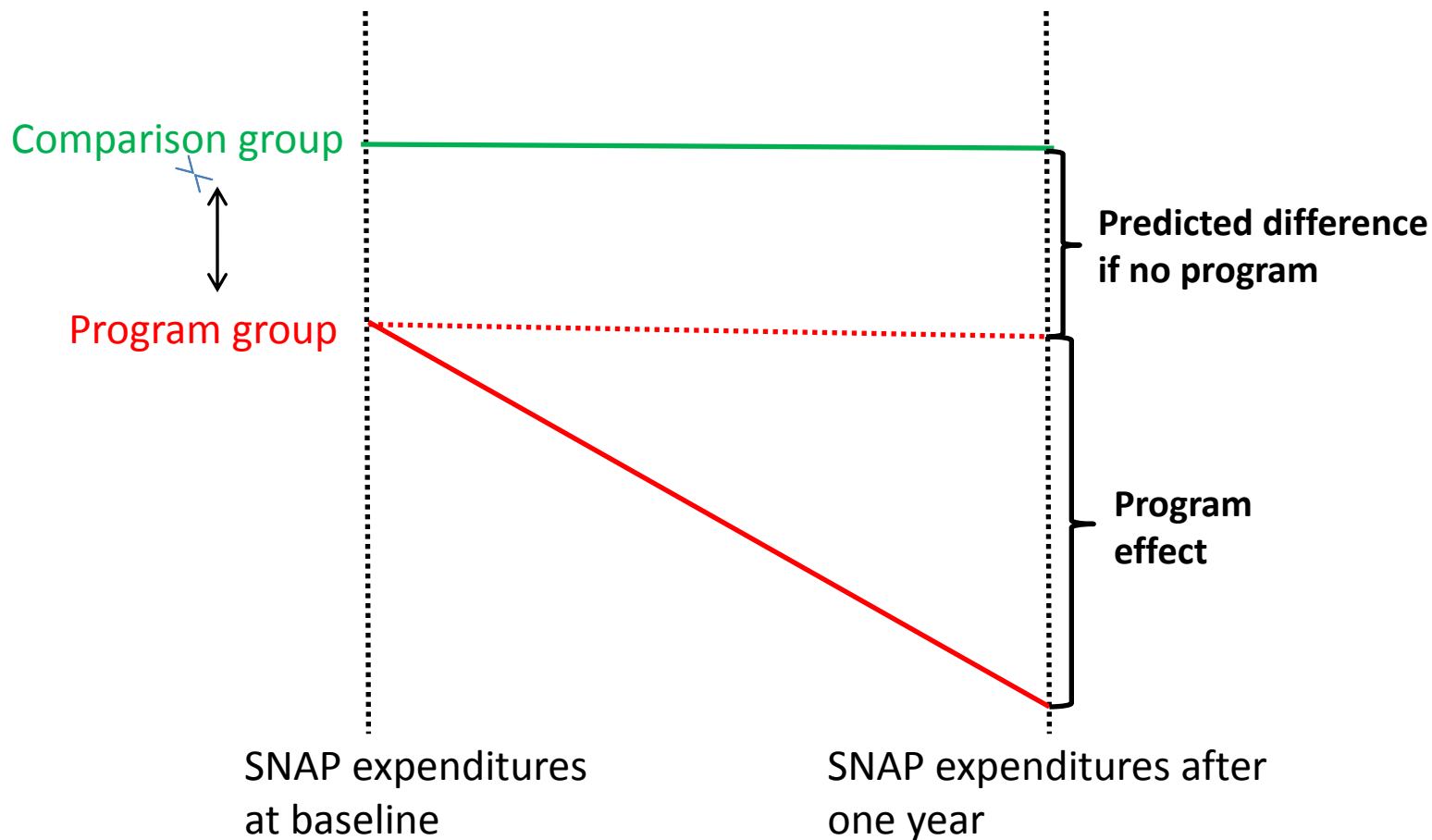
Cons

- 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.



Difference-in-Differences

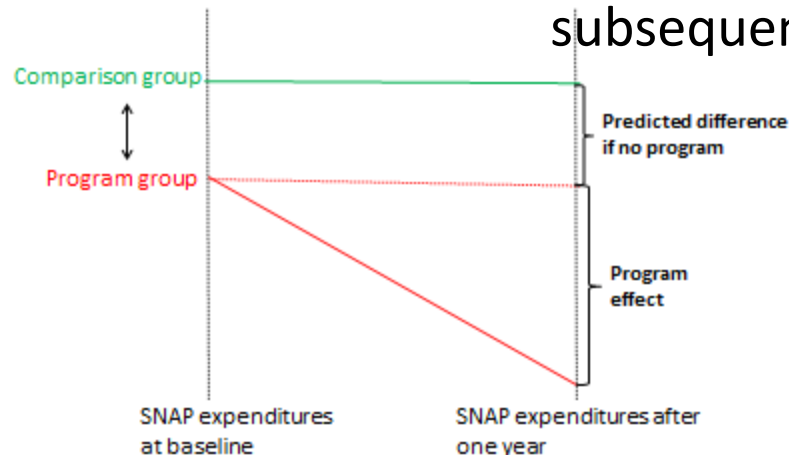
(Within state or in other states)

Pros

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

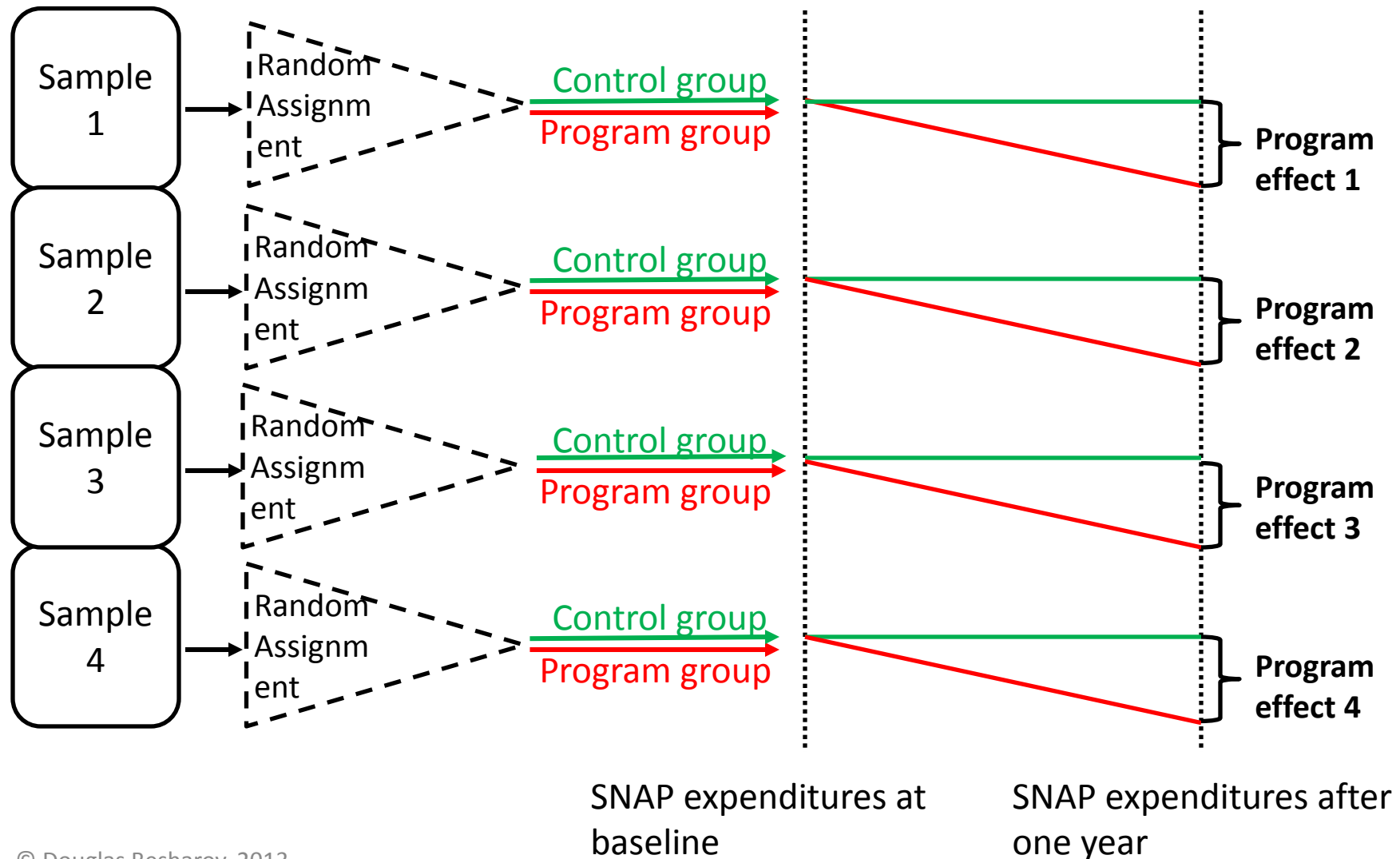
Cons

- 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.



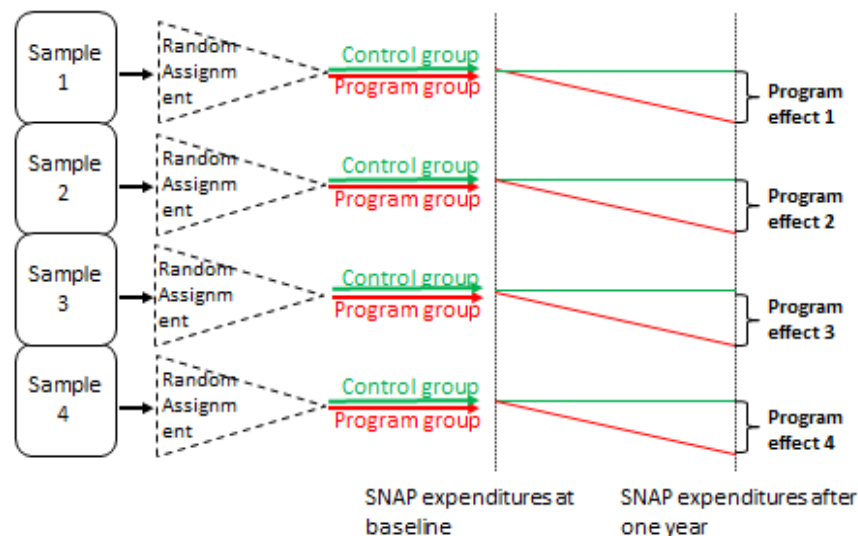
Stratified Random Assignment of Counties (or Offices) (Within state)

Pros

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

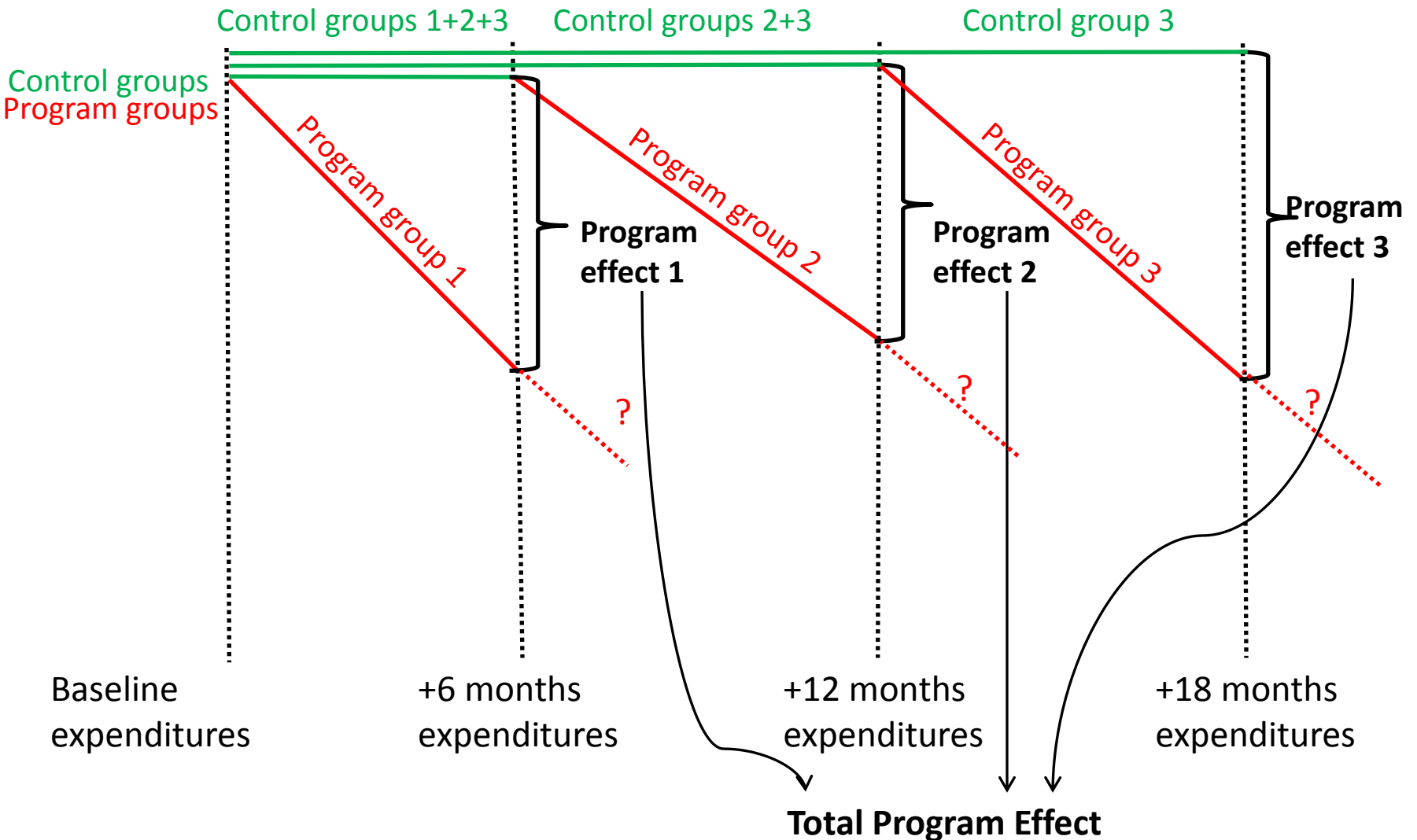
Cons

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



Randomized Phase-in Designs

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

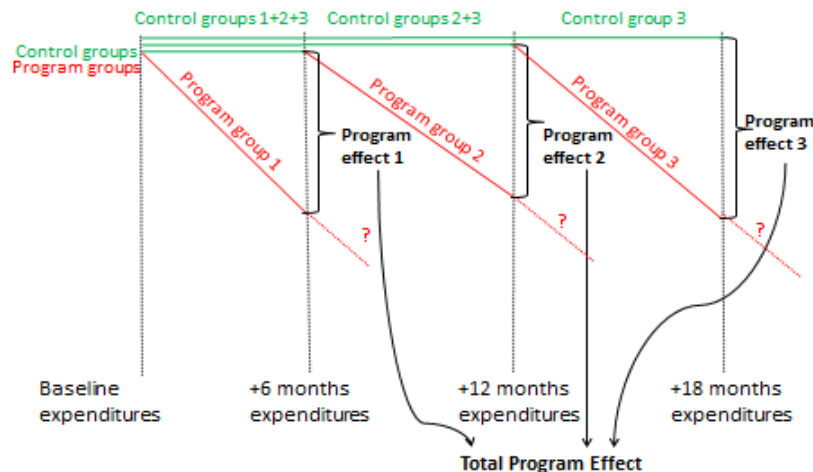


Randomized Phase-in of Counties (or Offices) (Within state)

Pros

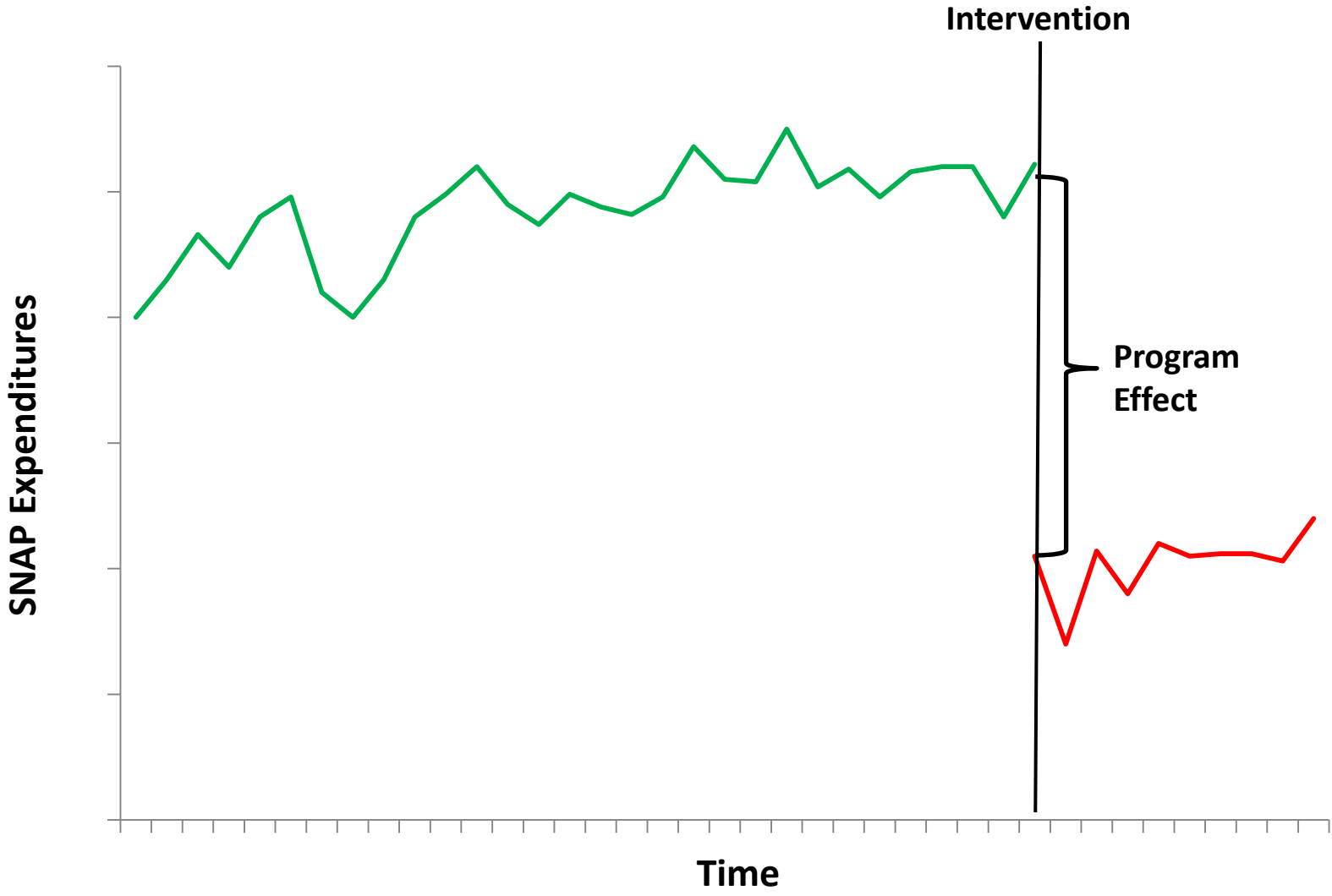
Cons

- Strong causal validity, if sufficient sample size
- Captures entry effects
- Requires only a temporary holdback
- Relatively inexpensive
- Possible spillover/contamination
- Requires timely and efficient implementation
- Imprecise estimates
- No estimates of cost savings after full implementation
- Effect takes more time to document



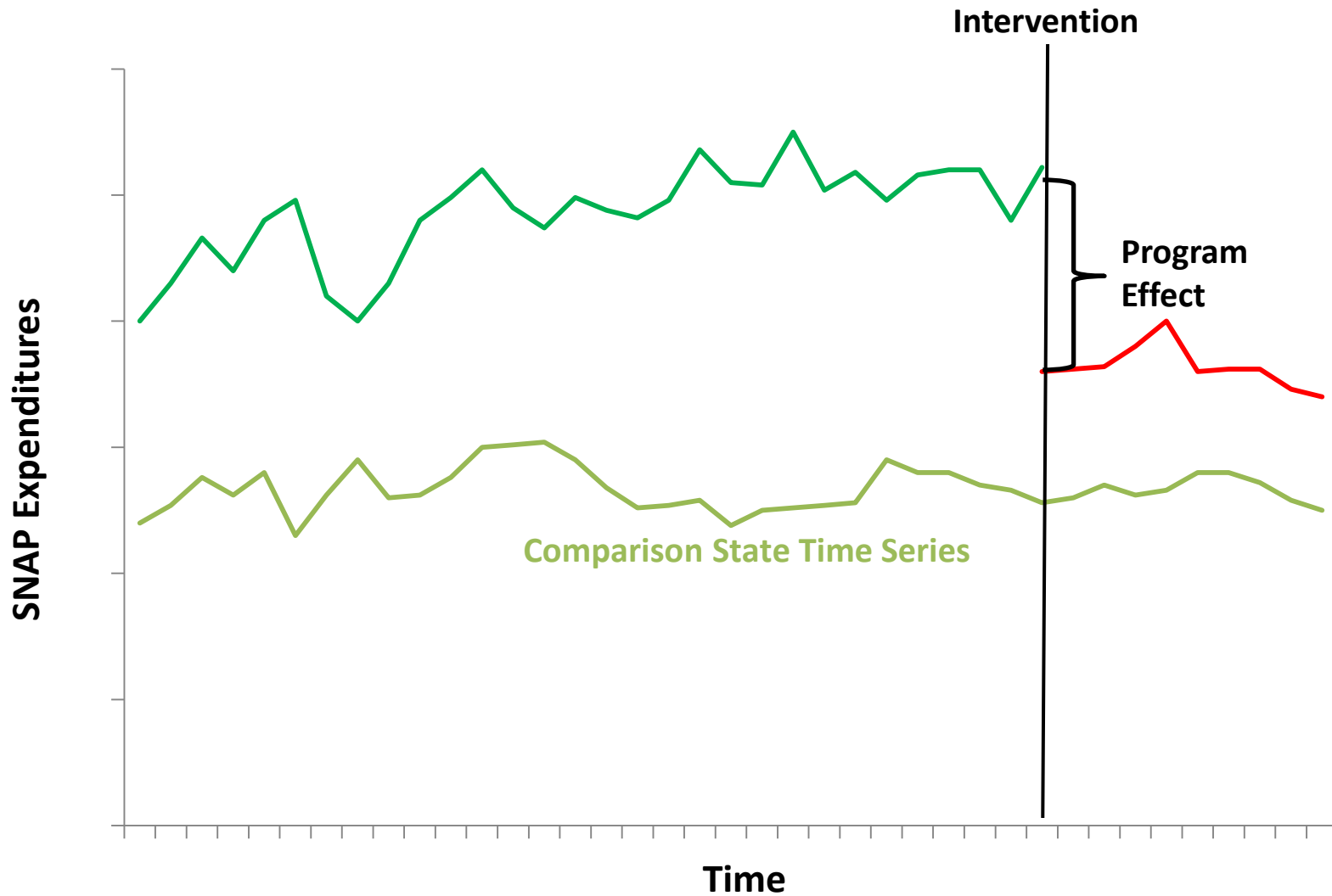
Interrupted Time Series Studies

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”)



Interrupted Time Series Studies

With a comparison time series



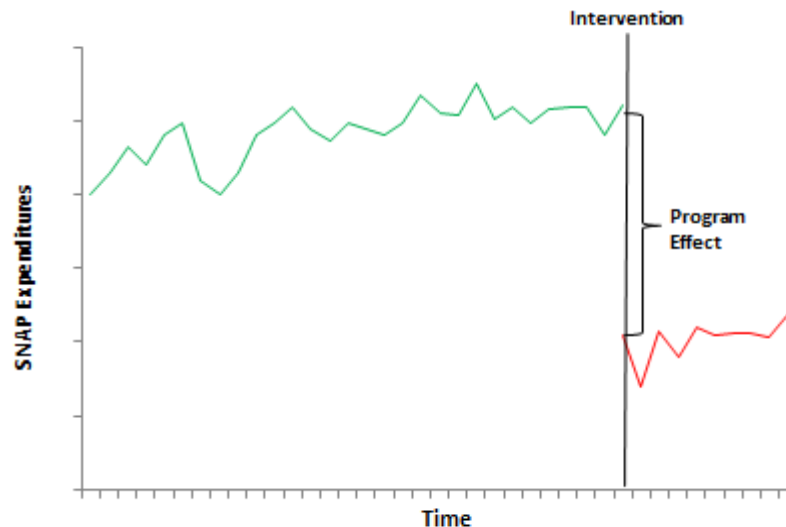
Interrupted Time Series (Statewide)

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



Methodology	Causal Validity	Entry Effects	Precision	Small or No Holdback	Low Cost
Random assignment of individuals	?	N	?	Y	N
Matched individuals	N	N	?	Y	?
Propensity score matching	N	N	?	Y	?
Random assignment of counties or offices	Y	Y	?	N	Y
Matched counties	N	Y	?	N	Y
Difference-in-differences	?	Y	?	N	Y
Stratified random assignment of counties	Y	Y	Y	N	Y
Randomized phase-in	?	Y	N	Y	Y
Interrupted time series	?	Y	?	Y	Y