Causal Attribution in Environmental Program Evaluation

Gary T. Henry
MacRae Professor of Public Policy
University of North Carolina at Chapel Hill

2008 Environmental Evaluator's Networking Forum
June 12, 2008

Eating Your Vegetables: Revisiting Causal Attribution

- Not just RCT vs. everything else
- The Received Theory of Causality has Changed
  - Campbell & Stanley (1966); Cook and Campbell (1979); Shadish, Cook and Campbell (2002)
  - Design-based logic of inquiry approach
  - Objective: to establish that a causal relationship exists and can reasonably be generalized
  - Method: Making alternative explanations implausible

Potential Outcomes

Units, individuals or plots of land, have alternative potential outcomes, for example, recycling/not recycling or deforested/forested, respectively.

Objective of Establishing a Causal Impact

Each unit has alternative outcomes

Evaluation Question: Does an environmental program alter the potential outcomes in the desired direction for a unit?

For a particular unit, we would like to observe the outcome after the intervention occurred for two conditions:

1. If the unit was included in the intervention;
2. If the unit was not included in the intervention.

The objective is an unbiased estimate of the effect of the program

Rubin's Causal Model (RCM)

Fortunately, someone (Donald Rubin) has done the math for us…

(but unfortunately the fine print says we have to collect the data)

The Objective: Unbiased estimate of the effect of treatment

Possible assignments (X, treatment or control)

Potential outcomes (Y)

- Y_{iT} = outcome for individual i after exposure to treatment
- Y_{iC} = outcome for individual i after exposure to control

Back to Potential Outcomes

Unit	Potential Outcome without Program	Potential Outcome without Program (Y_{iC})	Potential Outcome with Program	Potential Outcome with Program (Y_{iT})	Label
1	deforested	0	forested	1	Program success
2	forested	1	forested	1	No difference
3	deforested	0	deforested	0	No difference
4	forested	1	deforested	0	Program failure

These four units exhaust all of the logical possibilities

Treatment Effect for Unit *i*

$$\tau_i = (Y_{Ti} - Y_{Ci})$$

The fundamental problem with causal inference:

It is ***impossible*** to observe the ***ideal*** comparison

All designs including RCTs are ***approximations*** of the ideal

Causal inference requires assumptions: RCTs require the fewest

Extrapolation of treatment effects to target population requires additional assumptions

Potential Outcomes

	Strata	Units in Study Population	*Potential Outcomes*		*Label*
			Y_T	Y_C	
	1	40	1	0	Program success
	2	20	1	1	No Difference
	3	20	0	0	No difference
	4	20	0	1	Program failure

Program produced 60 forested parcels
No program produces 40 forested parcels
The program effect was 20 forested parcels or a 1.5 increase in forested parcels

Transformation from "scientific" issue to statistical issue

The average treatment effect (ATE)

$$\tau = E(Y_T - Y_C)$$

$$\tau = E(Y_T) - E(Y_C)$$

$$\tau = (\overline{Y}_T) - (\overline{Y}_C)$$

Possible Outcomes (Unobservable)

Independence = Equivalence of the Study Population Percentages for Each Strata in Control and Treatment

	Strata	Percentage of Study Population	*Possible Outcomes*	
			Y_T	Y_C
Control Group	**1**	20	1	0
	2	10	1	1
	3	10	0	0
	4	10	0	1
Treatment Group	**1**	20	1	0
	2	10	1	1
	3	10	0	0
	4	10	0	1

Possible Outcomes (Observable)

Independence = Equivalence of the Study Population Percentages for Each Strata in Control and Treatment

	Strata	Percentage of Study Population	*Possible Outcomes*	
			Y_T	Y_C
Control Group	**1**	20	?	0
	2	10	?	1
	3	10	?	0
	4	10	?	1
Treatment Group	**1**	20	1	?
	2	10	1	?
	3	10	0	?
	4	10	0	?

The Independence Assumption

To complete the ingredients needed for causal attribution (unbiased effect size estimate) we need a switch to assign units to treatment and control

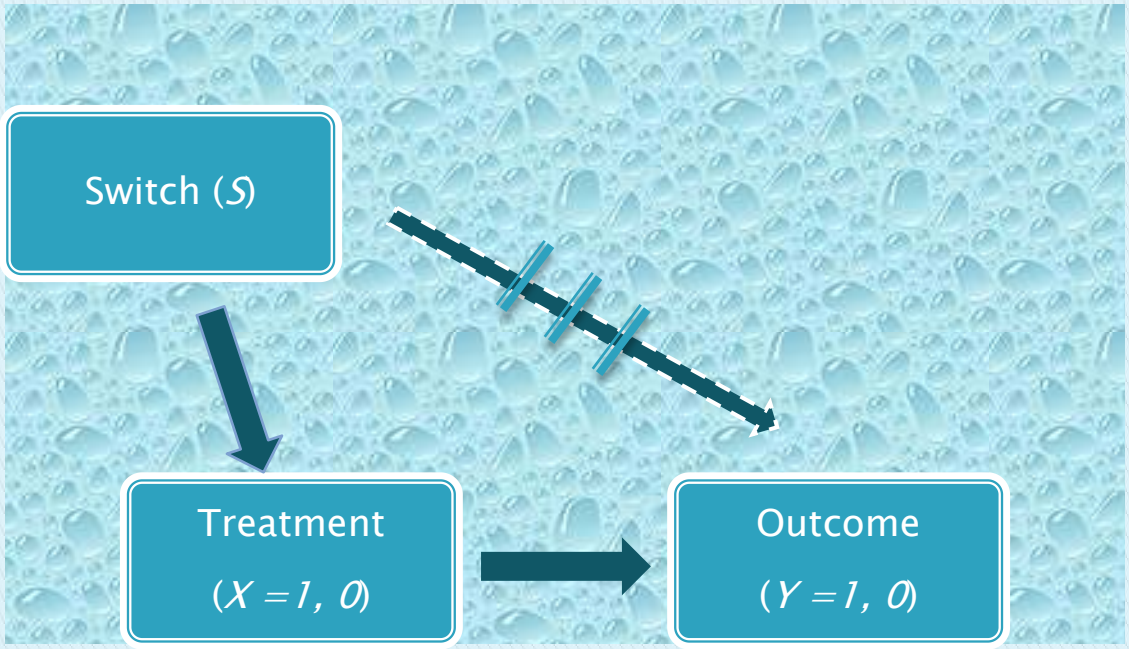
We need a switch that meets the independence assumption: creates equivalent groups…

The Switch *(S)*

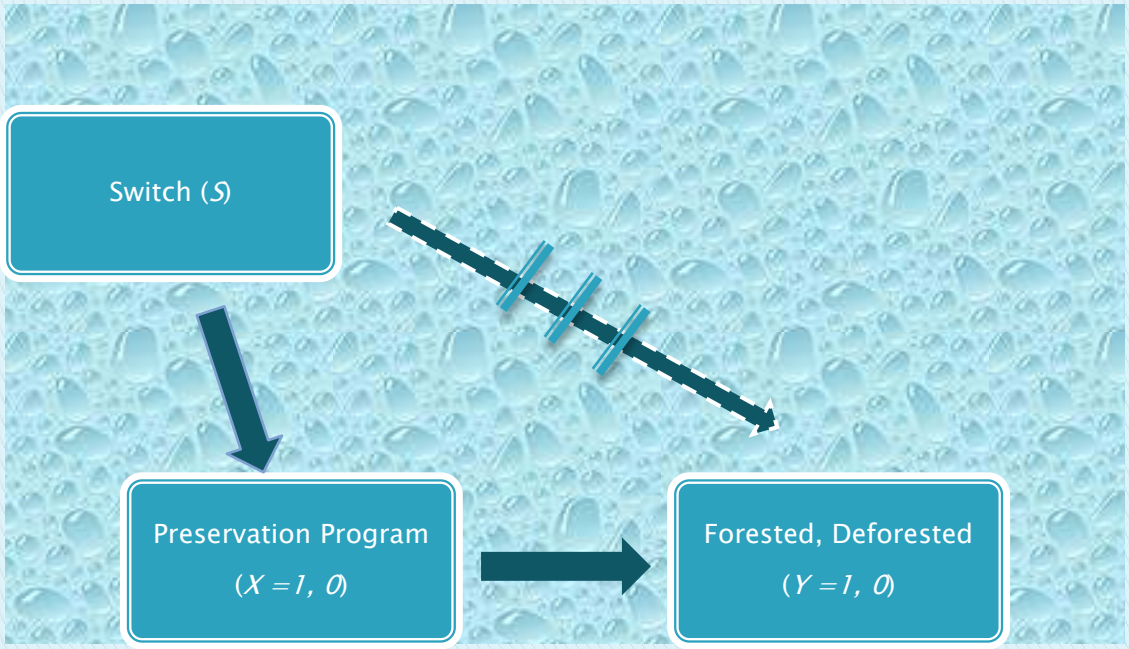
The switch assigns each individual to treatment (S = 1) or control (S = 0)

$$E(Y_i \mid S_i = 1) = E(Y_i \mid S_i = 0)$$

The Process



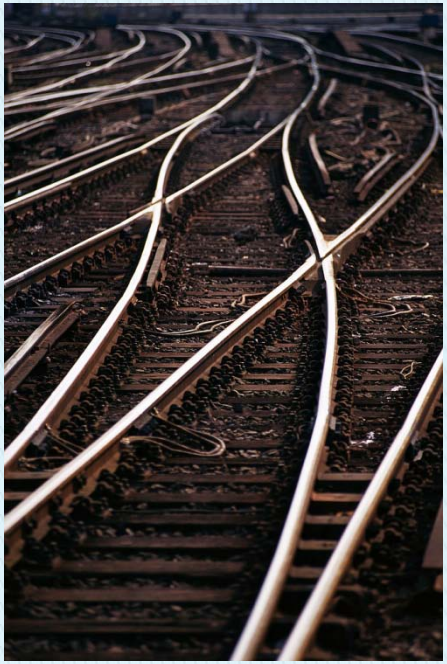
The Process


Switch (*S*)
Preservation Program
(*X* = 1, 0)
Forested, Deforested
(*Y* = 1, 0)

Observed outcomes: X = Treatment (1) or Control (0)

	Strata	Percentage of Study Population	*Potential Outcomes*		*Observed outcomes*		
			Y_T	Y_C	X	Y_T	Y_C
Control Group	**1**	20	1	0	0	*	0
	2	10	1	1	0	*	1
	3	10	0	0	0	*	0
	4	10	0	1	0	*	1
Treatment Group	**1**	20	1	0	1	1	*
	2	10	1	1	1	1	*
	3	10	0	0	1	0	*
	4	10	0	1	1	0	*

Choosing your switch?



Choosing your switch?

1. Random assignment to treatment & control
   - If independence produces equivalence, “extraneous” sources of variation (aka influence of disturbing variables) are equally distributed across treatment and control
   - Simplifies analysis
2. Matched sampling
3. Matched sampling using propensity scores
   - Propensity scores are each individual’s probability of being assigned to treatment
   - Matches based on finding individual in control similar to each individual in treatment based on propensity scores
4. Cutoff on assignment variable assigns individuals to treatment and control (regression discontinuity)
   - If model correctly specified, produces unbiased estimate of average treatment effect
5. Instrumental variable
6. Fixed effects (within individual estimates for panel data)

Or using regression to adjust estimates…

Evidence about Bias Reduction

Several important studies about differences in effect sizes between experimental and observational studies

Lipsey and Wilson (1992)

Weisburd, Lum & Petrosino (2001)

Glazerman, Levy & Myer (2003) matched sample labor force interventions; assumed randomized experiment unbiased

1. Matching works well (better w/ one-on-one matching extensive covariates;
2. Regression works well (better with specification tests, numerous controls, especially pretests);
3. Large sample studies less biased
4. Controls selected from “similar” sites

Large consensus that regression discontinuity is second best switch after randomized control trials (van der Klaauw 2003; Trochim, Cappelleri, Reichhardt 1991)

Things to think about…

1. What kind of evidence is needed to influence environmental policy and program decisions?
2. Is there a program on the horizon for which it would be helpful to have this information?
   … likely to have large benefits?
   … highly controversial?
3. Can you find the resources to invest in obtaining trustworthy information about program effects?
4. Consider the extrapolation problem – how to estimate effects on target population based on study population.