

Strategies for Selecting Comparison Groups

Thomas D. Cook

OPRE Methodological Advancement Meeting

Washington, D.C.

September 6, 2012

Two Quite Distinct Reasons for Selecting Groups for Comparison

- To identify the causal contrast tested and thus how the causal agent should be labeled – Part 1 of this presentation.
- To identify whether a statistical association between treatment and outcome is likely to be causal when an RCT or RDD is not feasible – Part 2.

Part 1: Labeling the Causal Agent More Accurately

We know, but mostly forget that:

The causal agent is not an entity labeled, say, Head Start (HS); it is always a contrast between this and:

A no-treatment control where “nothing” happens – very rare since we cannot suspend children’s development for a period!

A no-formal services control where HS is contrasted with children receiving no formal child care services at all. They stay at home.

A business-as-usual comparison experiencing the “usual” mix of services available – HS relative to whatever services non-HS children get or not

An alternative treatment “control” —HS compared to programs with similar or at least overlapping aims -- say, State Pre-K programs

These Contrasts entail different Causal Questions

1. What is absolute merit of HS, hypothetically compared to children in cold storage for a year?
2. What is the merit of HS relative to children getting no center-based pre-K services?
3. What is merit of Head Start relative to whatever other (heterogeneous) mix of services parents cobble together, or not?
4. What is merit of HS relative to some clearly labeled alternative pre-K policy?

Framings differ in Size of Hurdle for Demonstrating Positive Effects

- Relative to children with no formal pre-K exposure – often lowest hurdle
- Relative to the mix of center-based services for HS-eligible children not in HS – higher hurdle and probably most prevalent question now
- Relative to some other center-based program with partially overlapping goals – head-to-head comparison reflecting how choices are made in much political life – emergent framing today

Each Comparison is important, but for different Questions and Audiences

- We could answer all these differently framed causal questions in a single research project, but this is practically difficult and studies with multiple contrast groups are not common.
- Different stakeholders tend to be more interested in some contrasts than others,
- Since there is no reason all these different comparison groups will give the same causal answer, the comparison group you select is vital

Moving to Opportunity Example

For public housing residents, what are the effects of moving to a Census tract with <10% poverty relative to two different experimental contrasts:

- a. Staying in 50% poverty neighborhood and moving from there as choices and opportunities arise (business as usual)
- b. Getting a Section 8 Housing Voucher public housing residents can only use to move (alternative treatment)

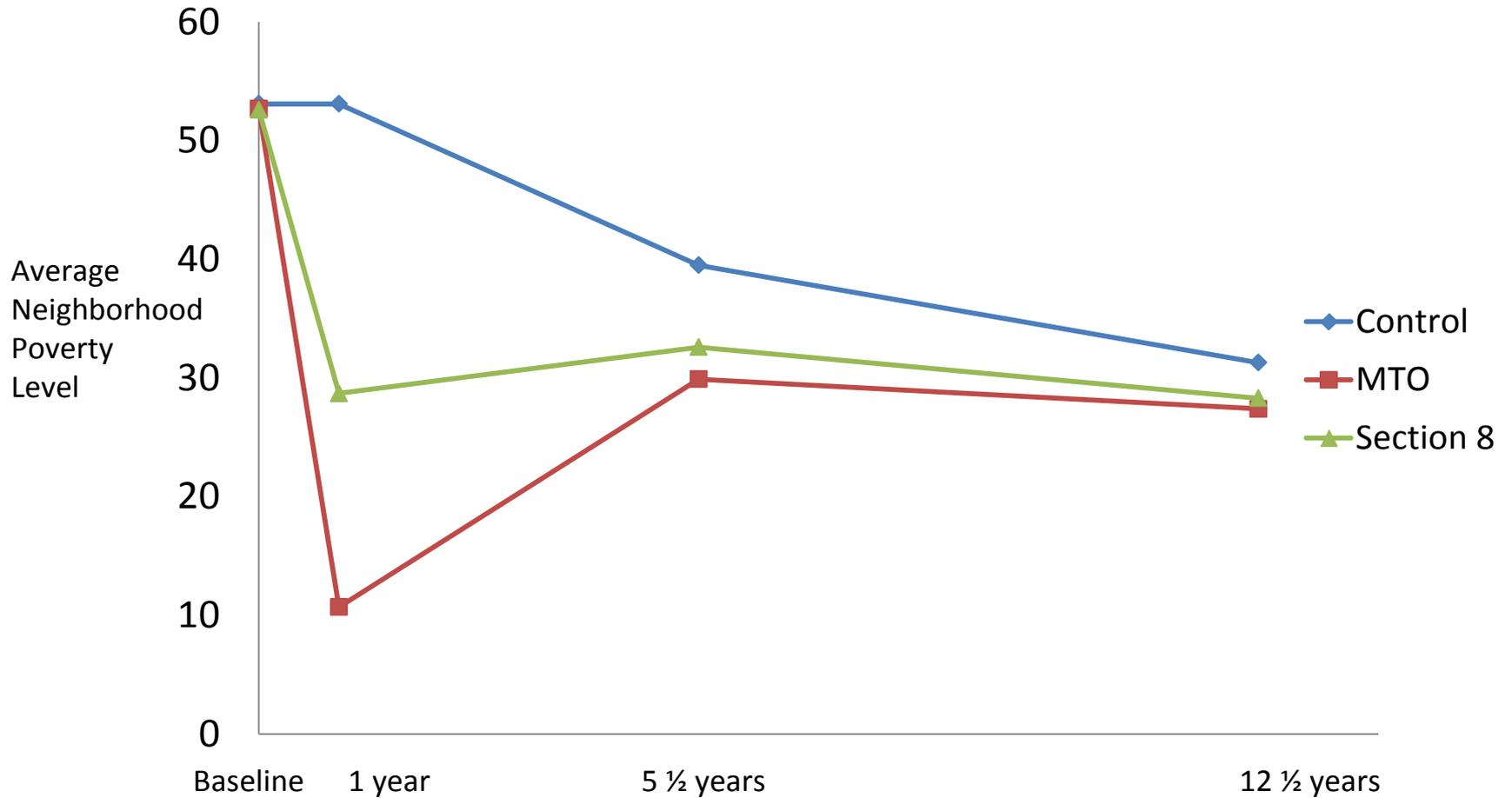
MTO was implemented as 10-15 year field experiment. It included measures of tract poverty to describe the planned treatment contrasts.

Census Tract Poverty Level Means after 1 year:

Business as Usual Controls	=	53%
Section 8 Group	=	29%
MTO Group	=	11%

Dramatic initial success in implementing
the desired treatment contrasts

Tract Poverty Means over Time:



What was the Causal Agent?

1. *If we contrast MTO and controls only.* We get health effects and could be due to initially moving to a more affluent neighborhood, at least initially.
2. *If we contrast MTO and Section 8 alternative treatment.* Difference in initial tract poverty rate but not most health outcomes. Does reduced poverty rate matter, or only a “large” reduction in this rate?
3. *If we contrast Section 8 and controls.* Difference in initial poverty rate but not most health outcomes. Some kinds of move do not seem to make a reliable difference to outcomes with estimands and power of tests used.
4. *If we contrast all three groups.* Clear differences in initial tract poverty rate, and health outcomes vary with this rate. Should MTO be analyzed as a parametric RCT? Is causal agent an as yet unidentified quite high level of tract poverty reduction? And then how is this reduction related to time? Is it due to just the initial change or subsequent also?

MTO Comparison Groups: Conclusion

- Including the alternative treatment group enriched (and also muddied) the findings in ways impossible with just a 2 group design
- Specifically, it prevented the overly general conclusion that reducing tract poverty improves health outcomes linked to glucose regulation
- But the alternative treatment group also suggests that we probably need to invoke issues related to the timing and dosage of any tract poverty changes if we are to understand how housing mobility improves health.

Part 2: Comparison Group
Choice and Causal
Identification: The Value of
Intact Group Matching when
no RCT or RDD is possible

Second Purpose

- Later at this conference, both Stuart and Thoemmes will discuss matching individual cases, usually persons
- This addresses the dominant current framing: Given the different populations receiving treatment and comparison status, how do we match cases to make them comparable on the most important observables?
- Question that I address here is different: Given the treatment group, how can the study sampling design help select a maximally similar control group and so reduce the role individual case matching is required to play later? (n.b., “minimize” NOT “eliminate”)

How can we know the Intact Group Matching Reduces Bias?

- Strategy here is to test to what extent intact group matching attains causal estimates similar to those from an RCT sharing the same treatment group
- I do not have time here to go into technical details about such “within study comparisons”, but it may become clearer as I briefly walk you through 3 examples that also exemplify what “intact group matching” means operationally.

Aiken, West et al. (1998)

- The Setting
- The RCT
- The RCT results for the multiple choice and essay-writing outcomes
- The Non-Equivalent Comparison Population
- Why it is Intact?
- What it Matches on
- The RCT vs Intact Group Matched Results

Bloom, Michalopoulos et al., (2002)

- The Setting
- The Within-City RCTs
- The Within-City Non-Equivalent Comparisons from other Job Training Centers
- Why it is Intact?
- What it Matches on
- The RCT vs Intact Group Matched Results similar in 3 of 4 cities and across all 4.

Diaz & Handa (2006)

- The Setting
- The RCT
- The non-Equivalent but Eligible Villages with Eligible Families
- What is the Intact comparison?
- What is matched
- Results comparing the RCT and intact group matched design.

What We are not Claiming

- In these 3 of 3 known cases, intact group matching led to equivalence on observables *and unobservables*
- This cannot be guaranteed as a general result
- However, such matching will reduce the difference between the treatment and control populations
- This reduction means more common support and thus better propensity score or OLS analyses of causal hypotheses
- This last is the most defensible and important rationale for intact group matching. It is a prelude to case matching and not an alternative

Conclusions

- This presentation only begins this conference's discussion of comparison group choice and matching. It asserts 2 main things:
 1. That strong causal claims depend on matching processes BOTH in a study's sampling design AND in its analysis of individual cases
 2. Selecting comparison groups is also (and perhaps especially) important for labeling a study's causal agent, for understanding how high the hurdle is a program has to jump, and how policy-relevant its findings will likely be
- For what it is worth, my view is that the causal question framing and the contrast groups it requires come first; comparison groups to facilitate causal inference, while extremely important, are secondary

How comparison groups are selected to facilitate cause will sometimes depend on whether we compare a program to a no-treatment control, a business as usual control, or an alternative treatment. That's for another day.