Estimating Interpretable and Accurate Experimental Impacts

March 12, 2012

Presentation to OAH and ACYF
Teenage Pregnancy Prevention Conference

John Deke

MATHEMATICA
Policy Research, Inc.

Overview

- **Presenting Interpretable Impacts on Dichotomous Outcomes**
- **Estimating impacts of “treatment on the treated”**
- **Avoiding estimation pitfalls**

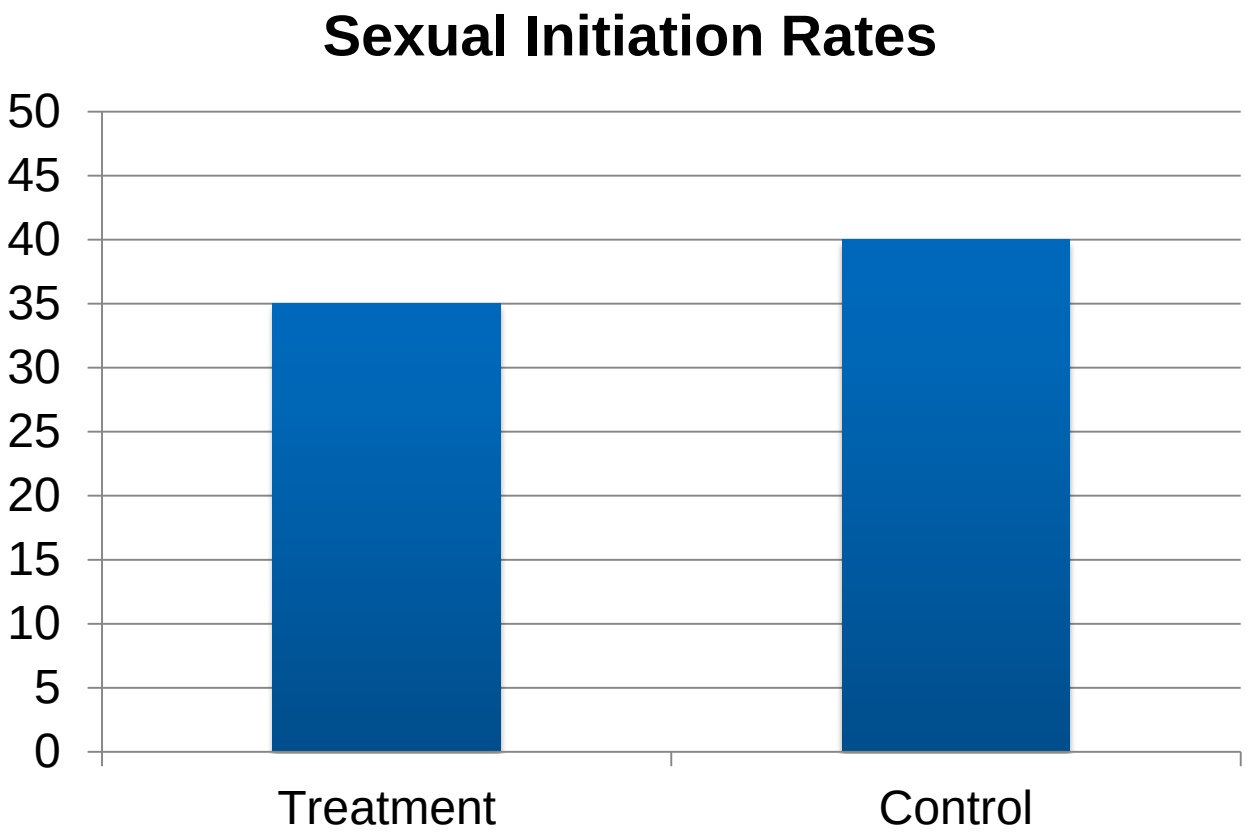
Presenting Interpretable Impact Findings

The Impact of an Offer

- A randomized controlled trial (RCT) directly supports estimation of the intent to treat (ITT) impact – the impact of offering a program
- Policy makers cannot force people to participate in a program – they can only offer a program
- The ITT impact is the difference in mean outcomes between the treatment and control groups:

$$I_{ITT} = \bar{Y}_T - \bar{Y}_C$$

Everyone Can Understand This



"35 percent of students assigned to the program group reported initiating sex after the first year compared to 40 percent of students in the control group."

But These Are Less Widely Understood

Odds ratio = 0.81

"Youth who participated in this intervention were 0.81 times less likely to initiate sex than those in the control group. In other words, for every intervention group youth, there are 1.23 youth in the control group who initiated sex at follow-up."

Log odds ratio = -0.21

Effect Size (Cox Index) = -0.13

Estimation

- Binary outcome and binary treatment indicator
- Simple t-test of equality of proportions between two groups
- Linear probability model
- Mean marginal effects from logistic regression

Two Strategies to Consider

- **Strategy 1**
  - **Estimate a logit, calculate and report mean marginal effects and statistical significance based on mean marginal effects**
- **Strategy 2**
  - **Estimate a logit, report odds ratio and statistical significance**
  - **If the impact is *not* statistically significant:**
    - Stop, because there's nothing to interpret
  - **If the impact *is* statistically significant:**
    - Report the average treatment effect estimated using a linear probability model to aid in interpretation

Estimating Impacts of Treatment on the Treated

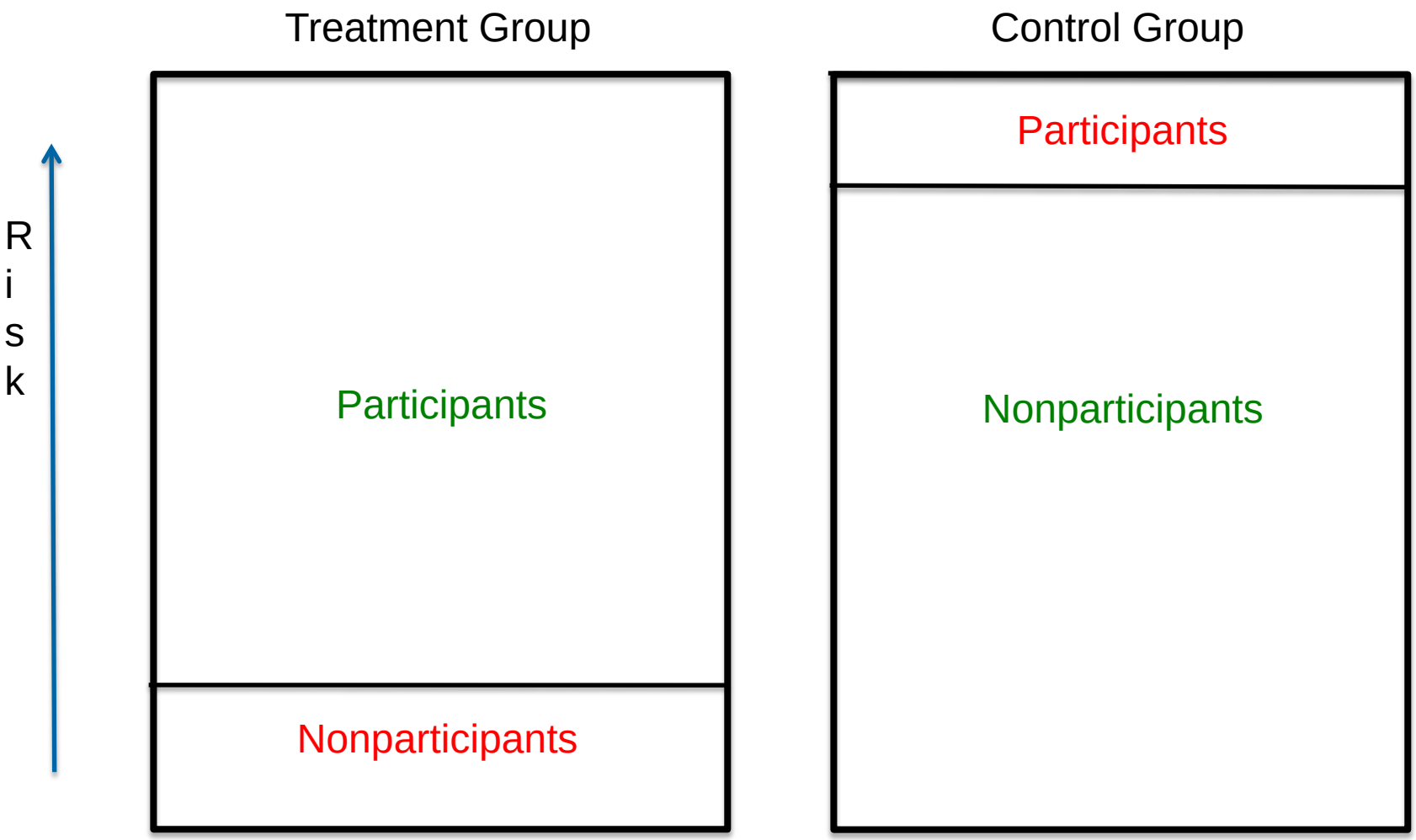
Impact of Participating in a Program

- **The ITT impact gives us the effect of *offering* the program to students**
- **ITT does not tell us the effect of *participating* in the program**
- **The complier average causal effect (CACE), also known as the impact of treatment on the treated (TOT), provides an unbiased estimate of the effect of participating in a program**

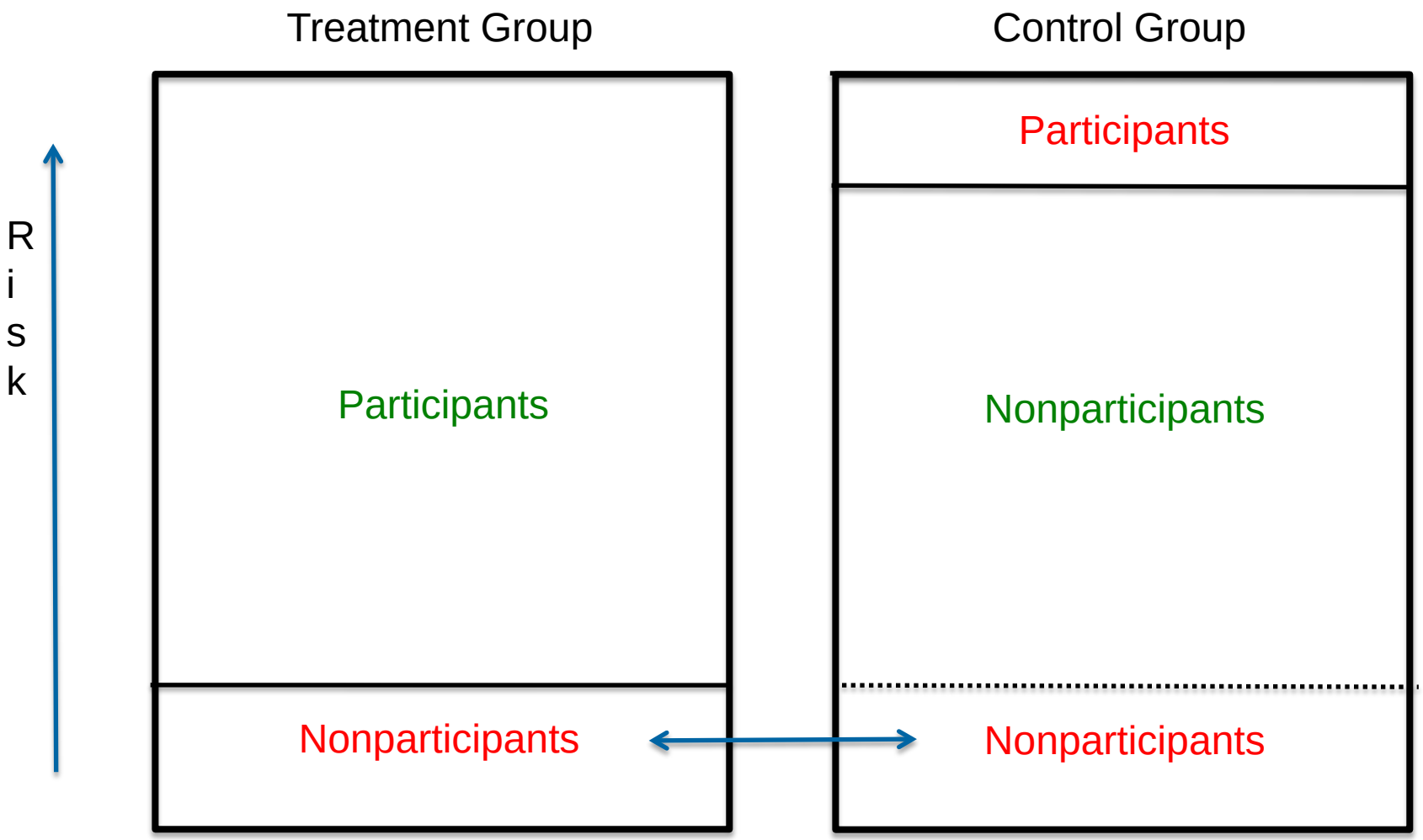
Main Challenge in Estimating CACE

- **Participation is not randomly assigned – those who choose to participate may be different from those who don't**
- **We need to find a way to make a valid comparison – to compare treatment group participants to a comparable group of non-participants in the control group**

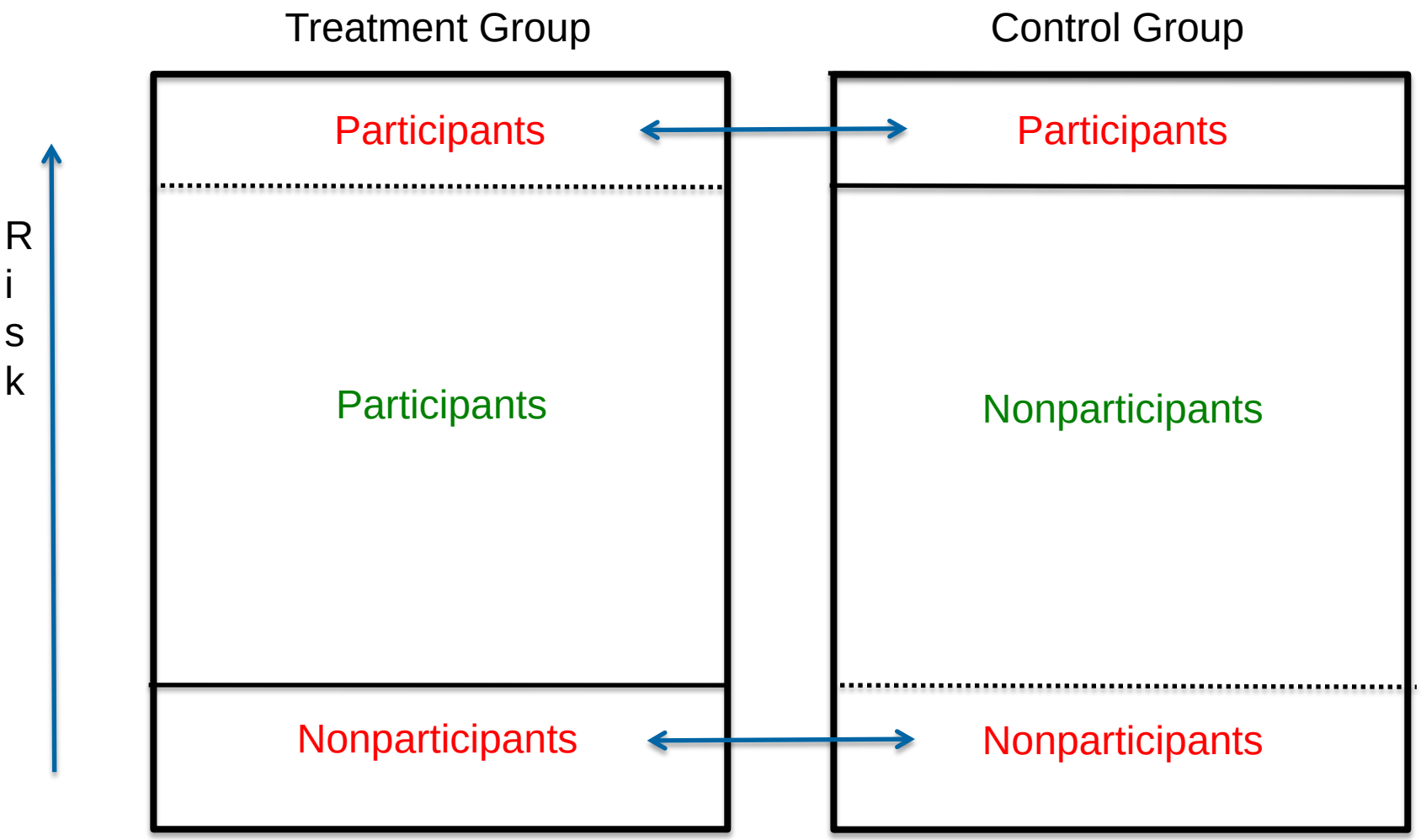
Looking For a Valid Comparison



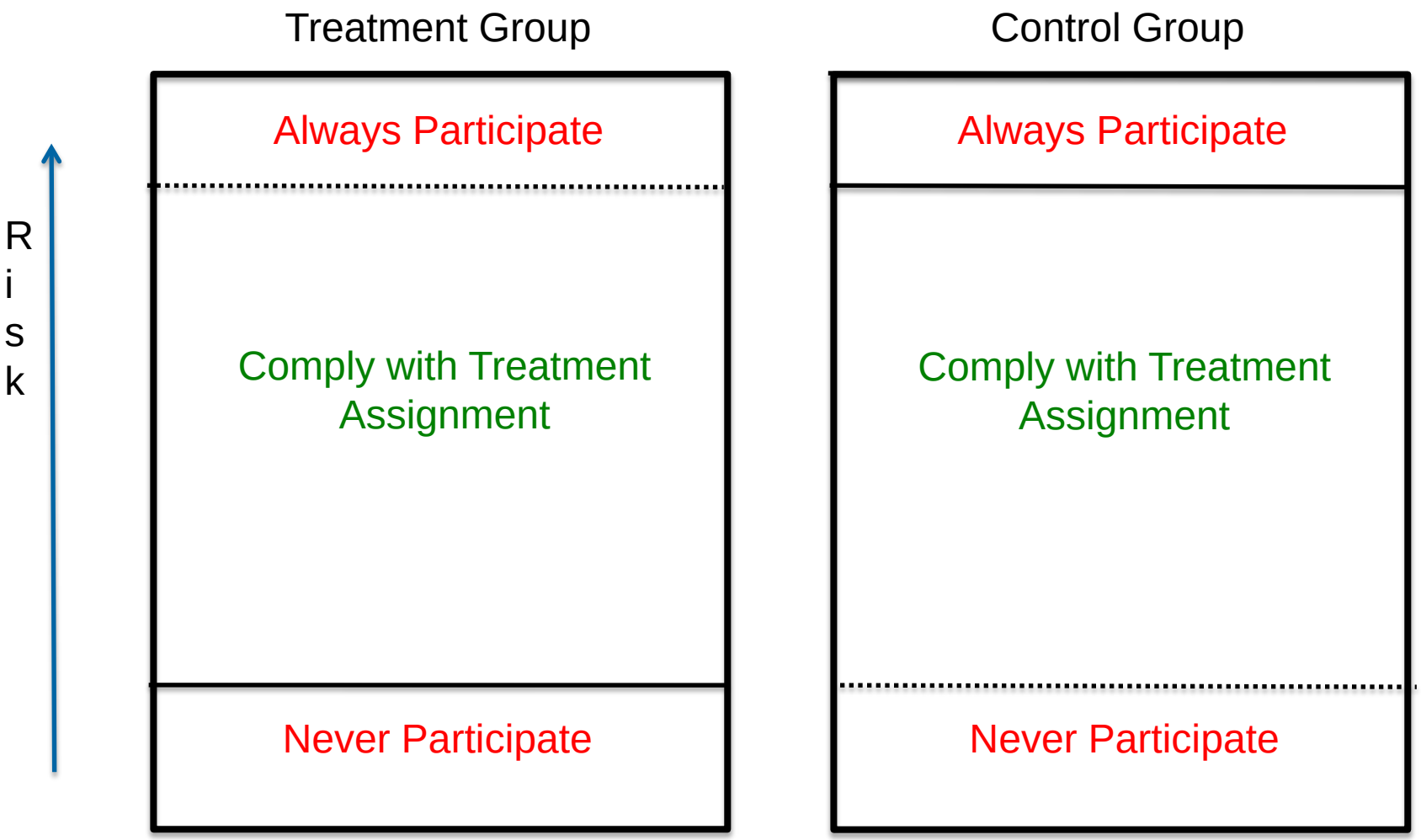
Looking For a Valid Comparison



Looking For a Valid Comparison



Looking For a Valid Comparison



Complier Average Causal Effect

- **Goal: calculate the impact for Compliers**
- **Impact on Compliers cannot be directly calculated but it can be inferred:**
  - **If we assume impacts on Always Participate and Never Participate groups are zero**
  - **Because we know by random assignment that all three groups are proportionately equal in size between the treatment and control groups**
  - **Because we directly observe the sizes of the Always Participate group and the Never Participate group**
- **Multiple estimation methods**
  - **Instrumental Variables (Angrist, Imbens, and Rubin 1996)**
  - **Bayesian approach (Imbens and Rubin 1997)**

Instrumental Variables Approach

- Two step method
- Estimate impact on program participation (I_P)
- Estimate ITT impact on outcome of interest (I_{ITT})
- Divide ITT impact by the impact on program participation ($I_{CACE} = I_{ITT} / I_P$)
- Correct standard errors take into account $Var(I_{ITT})$, $Var(I_P)$, and $Cov(I_{ITT}, I_P)$

A Strategy to Consider

- **Implement either strategy 1 or 2 from before**
- **If the ITT impact is statistically significant:**
  - **Report the CACE using two stage least squares (IV with linear probability models) to aid in interpretation**

Avoiding Estimation Pitfalls

Comparing Impacts Over Time

- Several studies will examine impacts at multiple points in time
- Different students may respond to surveys at different points in time, and response rates may vary by treatment status
- Don’t confound changes in the analysis sample with changes in impacts
- One solution: estimate impacts at multiple points in time for students present at all points in time

Once randomized always analyzed

- **Remember that the attrition calculation is the number of students with missing outcome data divided by the number of students <u>involved in random assignment</u>**
- **Survey entire consented sample at all data collection waves**
- **Don't stop collecting data just because:**
  - **Someone didn't respond to an earlier round**
  - **Someone has stopped attending the program**

Blocked Random Assignment

- **Blocked random assignment means that units were randomly assigned within "blocks"**
  - **Matched pairs of schools**
  - **School districts**
- **Impact regression must include block "dummy" variables. If not, then:**
  - **degrees of freedom will be wrong**
  - **impact could be biased if the a priori probability of assignment to the treatment group varies by block**

Adjustments For Clustering

- Adjust for clustering at the unit of random assignment
- Example: randomly assign schools, cluster standard errors at the school level
- Accounting for clustering at levels other than the unit of assignment is fine but not necessary
- Some methods to account for clustering:
  - Random effects / mixed effects / hierarchical linear modeling (HLM)
  - Generalized estimating equations (GEE) / Huber-White / the "sandwich" estimator

Be Careful When Mean-Centering

- **Mean Centering of covariates in HLM is a common practice**
- **Remember that covariates are most valuable when they explain variation at the level of random assignment**
- **Be careful centering covariates at the unit of assignment**
  - **This can destroy the explanatory power of a covariate**
  - **Only do this if you include a cluster-level version of the covariate**
  - **Alternative is grand-mean centering**

For more information

- You can reach us through:
  - Your TA liaison
  - TPPevalTA@mathematica-mpr.com
  - 1-866-336-3880
  - Your project officer
  - SharePoint website (coming soon)