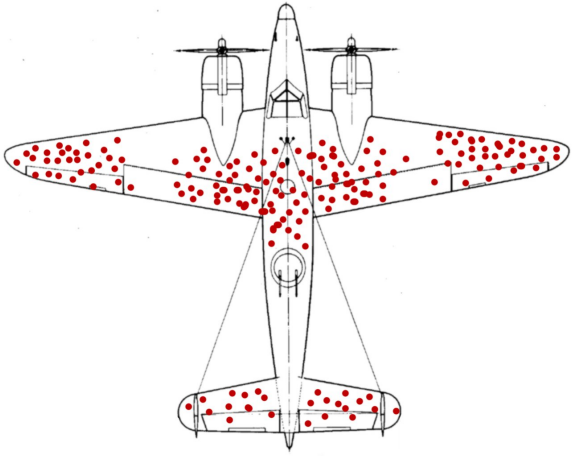
Research Design

The Advantages and Challenges of Randomized Controlled Trials

Miles Corak
Department of Economics
The Graduate Center, City University of New York
MilesCorak.com @MilesCorak

Lecture 1
Labor Economics I
ECON 87100



Three ingredients of a successful research program

1. Academic rigour
2. Policy relevance
3. Broadly communicated

Three ingredients of a successful research program

1. Academic rigour

- Understand and encompass the existing literature
- Innovative, yet appropriate, use of data
- Appropriate causal inference

Three ingredients of a successful research program

1. Academic rigour
2. Policy relevance
   - Tied to new facts or trends
   - Framed in terms of policy levers
   - Timely

Three ingredients of a successful research program

1. Academic rigour
2. Policy relevance
3. Broadly communicated

- Accessible to a wide range of audiences
- High potential for media coverage
- Partnered with policy makers

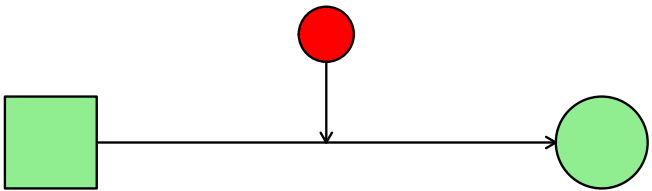
Our challenge in this course:

Academic rigour in making appropriate causal inference as it applies to important issues in the economics of work and pay

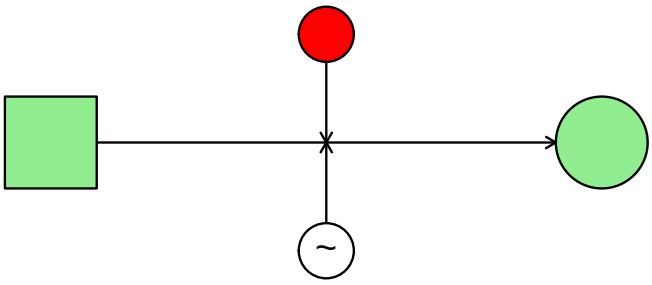
Develop an understanding of theory and empirical practice for making causal statements

- Just what is the causal problem?
- How is the Randomized Controlled Trial an "ideal" solution?
- What are the challenges (disadvantages) of Randomized Controlled Trials?

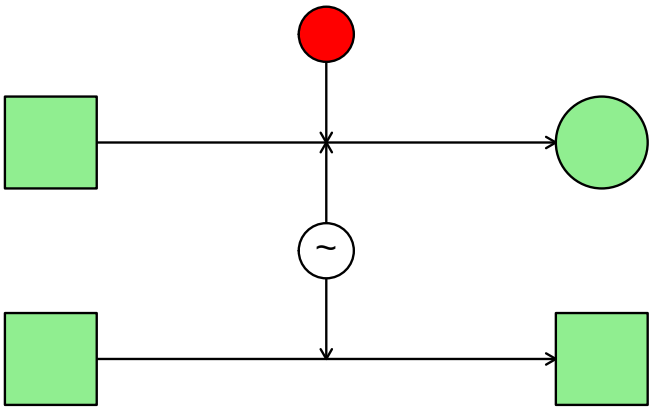
The causal problem, and the “ideal” solution



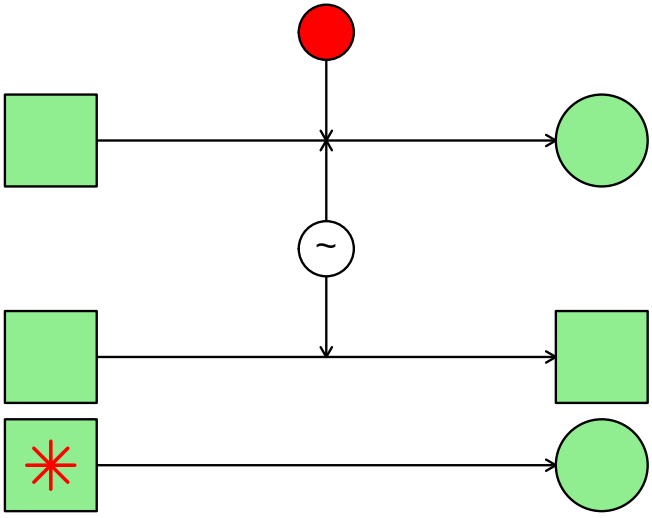
The causal problem, and the “ideal” solution



The causal problem, and the “ideal” solution



The causal problem, and the “ideal” solution



An example: the Polio Vaccination Trial of 1954

rates of cases per 100,000

	Size	Rate
Treatment	200,000	28
Control	200,000	71
No consent	350,000	46

1. Polio is a disease of hygiene
   - implying that low income poeple are less likely to contract the disease, poorer hygiene during the early years implying the development of antibodies among children who came into contact with a weak form of the virus.
2. Participation in the trial depended upon income
   - low income families were less likely to participate in the experiment, and consent to their children receving the vaccine
   - participants were more likely to be from higher income families, and therefore more prone to the disease

Potential outcomes framework and the self-selection problem

1. some notation
   - let i index individuals, with $i = 1...N$
   - let D index program participation
     - $D = 1$ means a program participant
     - $D = 0$ means a non participant
   - let $Y_{i,j}$ indicate an outcome of interest for individual i
     - where $j = 1$ and $j = 0$ indicates receipt or not of the treatment
2. we wish to estimate the impact on the outcome for individual i when treated
   - $\Delta = Y_{i,1} - Y_{i,0}$
   - we cannot observe the same individual in two different states

Potential outcomes framework and the self-selection problem

3. we focus on the average over a number of individuals
   - $E(\Delta) = E(Y_{i,1} - Y_{i,0})$
   - but in designing an experiment we actually estimate
     - $E(\Delta|D_i = 1) = E(Y_{i,1} - Y_{i,0}|D_i = 1)$
     - this is the impact of treatment on the treated
     - not all individuals in the treatment group need be "compliers"
   - to estimate $E(Y_{i,1} - Y_{i,0}|D_i = 1)$ we need to estimate $E(Y_{i,0}|D_i = 1)$
     - but this can't be estimated from $E(Y_{i,0}|D_i = 0)$ without raising the threat of a selection bias

Potential outcomes framework and the self-selection problem

the selection bias

$$\Delta^* = E(Y_{i,1}|D=1) - E(Y_{i,0}|D=0)$$

$$= E(Y_{i,1}|D=1) + E(Y_{i,0}|D=1)$$
$$-E(Y_{i,0}|D=1) - E(Y_{i,0}|D=0)$$

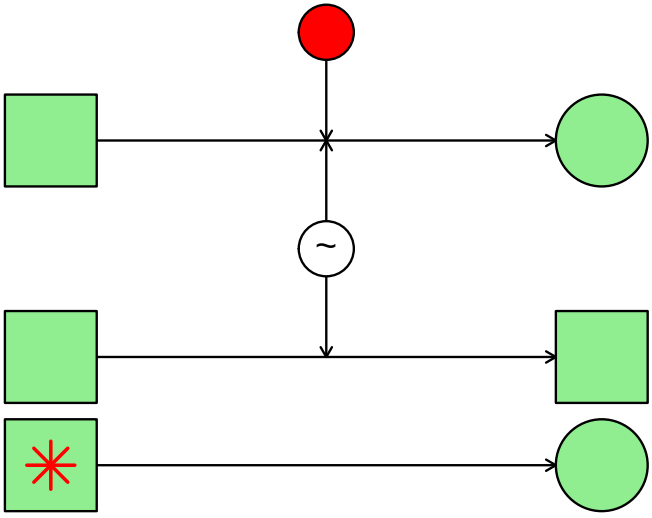
$$= E(Y_{i,1}|D=1) - E(Y_{i,0}|D=1)$$
$$+E(Y_{i,0}|D=1) - E(Y_{i,0}|D=0)$$

the true effect plus another term reflecting a difference in the expected outcome in the absense of treatment between participants and non participants

The advantage of randomized controlled trials

1. Randomized controlled trials address the bias that potentially results from the selection problem
   - selection arises from the fact that there is missing information on factors that influence both the participation in the program and the outcome
   - by randomly denying treatment a properly run trial creates variation in exposure to the treatment that is independent of the decision to participate
   - observational studies rely on collecting more information, or using more sophisticated statistical methods to model the selection process

The causal problem, and the “ideal” solution



The advantage of randomized controlled trials

1. Randomized controlled trials address the bias that potentially results from the selection problem
2. They offer the possibility of measuring treatments not previously observed
3. They are easy to communicate

Questions to discuss next class

1. What are the disadvantages of randomized controlled trials?
2. What are the challenges of implementing a randomized controlled trial in the social sciences?
3. In what ways does the "Self-Sufficiency Project" reflect these advantages, and address the challenges?