Are Observational Studies Valid?

Hamsa S. Bastani

5/11/15

LaLonde (1986)

- “Evaluating the Econometric Evaluations of Training Programs with Experimental Data”
  - Field experiment, random assignment of participants
  - NSW (National Supported Work) for ex drug addicts, ex convicts, etc.
  - Followed for 5 years

LaLonde (1986)

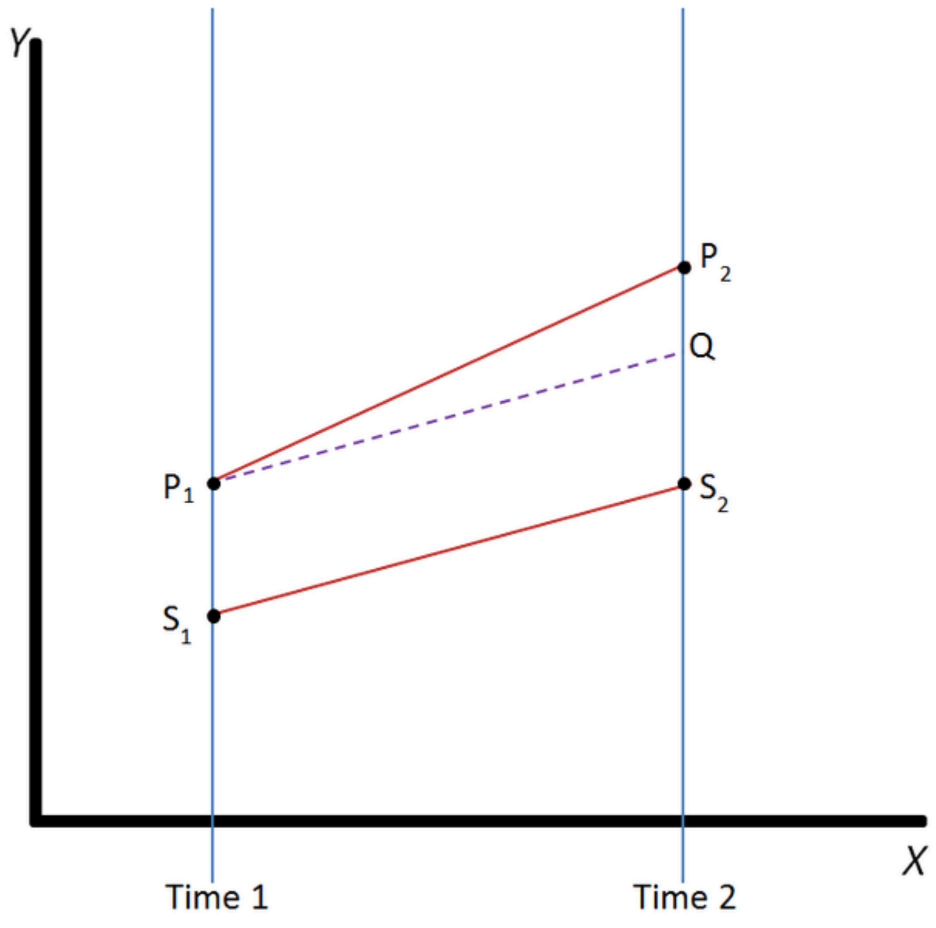
- Non-experimental group
  - PSID (Panel Study of Income Dynamics)
    - longitudinal survey of 9000 families about economic, social, and health behavior
  - CPS (Current Population Survey)
    - monthly survey of households about labor, employment, and unemployment

Results

- Non-experimental methods fail to recover true estimate

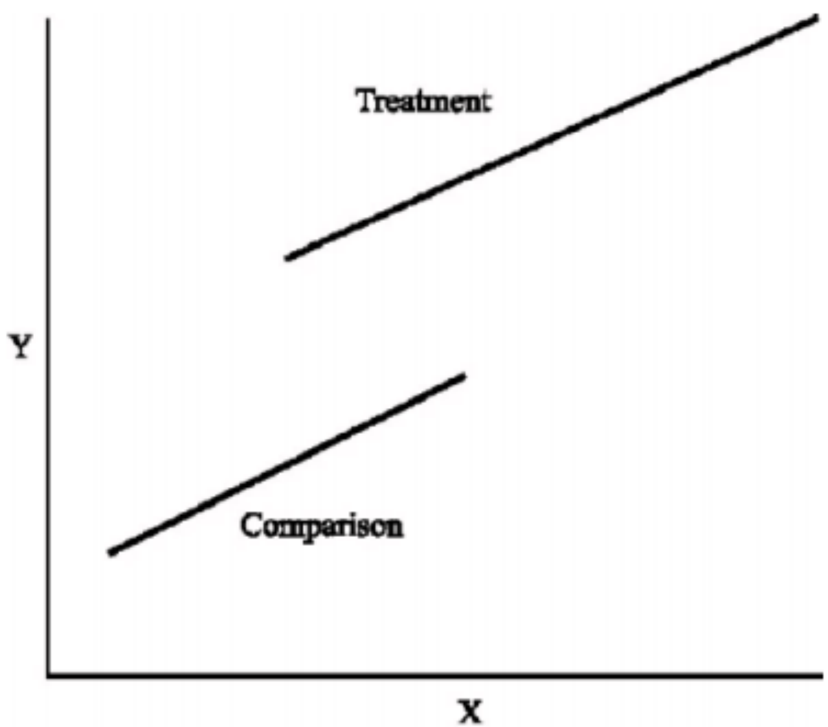
Control group	Regression	DID
Experimental group		
NSW	798	856
Non-experimental groups		
PSID1-PSID3	[-8,067 ; -509]	[-1,325 ; -650]
CPS1-CPS3	[-4,416 ; 224]	[-1,388 ; 195]

Diff-in-Diff



What's the problem?

- Lack of overlap of match variables
- Inappropriate weighting



Dehejia-Wahba (1999)

- "Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation in Training Programs"
  - Used propensity score matching
  - Restricted sample to individuals whose 1974 income was known (~60%)

Dehejia-Wahba (1999)

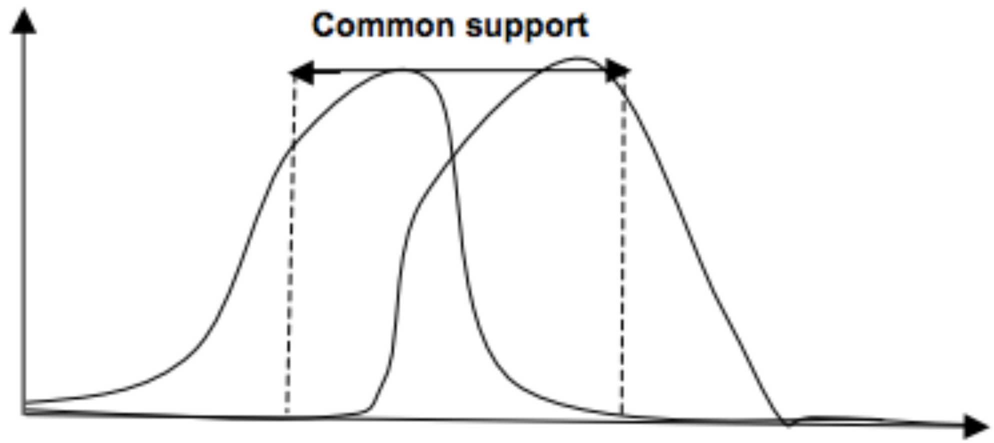
Control group	Regression	Stratification	PSM
Experimental group			
NSW	1,672	1,672	1,672
Non-experimental groups			
PSID1	731	1,494	1,473
PSID2	683	2,220	1,480
PSID3	825	2,235	1,549
CPS1	972	1,774	1,616
CPS2	790	1,622	1,563
CPS3	1,326	2,219	662

Smith-Todd (2005)

- Dispute sample restriction in Dehejia-Wahba
- Estimates without restriction are much more biased
- Need to adapt estimator to context / support

Key Assumptions

- Common support: $\forall X, 0 < E(T|X = x) < 1$



Key Assumptions

- In a natural experiment, we have strong ignorability: $(Y_0, Y_1) \perp\!\!\!\perp T$
- In matching, we need the Conditional Independence Assumption (CIA):

$$Y_{i0}, Y_{i1} \perp\!\!\!\perp T|X$$

- For propensity score matching,

$$p(X) = P(T = 1|X) = E(T|X)$$

$$Y_{i0}, Y_{i1} \perp\!\!\!\perp T|X \Rightarrow Y_{i0}, Y_{i1} \perp\!\!\!\perp T|p(X)$$

Abadie and Imbens (2006)

- "Large Sample Properties of Matching Estimators for Average Treatment Effects"
- Matching estimators with fixed # matches
  - Highly non-smooth function of data
  - Not sqrt-N consistent in general
  - Even when consistent, does not achieve semiparametric efficiency bound (exception: PSM)
  - Bootstrap fails, Abadie-Imbens give a consistent variance estimator

Types of matching…

- Propensity score matching
  - assumes model is known
  - uni-variate CIA
- Multiple covariates
  - integer linear program to get best balance
  - not sqrt-N statistically consistent
  - unsure which covariates to balance more, especially in higher dimensions
  - Bias-variance tradeoff