Matching Methods

Matching: Overview

- The ideal comparison group is selected such that matches the treatment group using either a comprehensive baseline survey or time invariant characteristics
- The matches are selected on the basis of similarities in observed characteristics
- This assumes no selection bias based on unobserved characteristics
  - Take the ITN Example from Yesterday: Households who were more concerned about malaria also took other preventative actions
  - All such differences must be in the data in order for the match to produce a valid estimate of project impacts

Propensity-Score Matching (PSM)

Propensity score matching: match treated and untreated observations on the estimated probability of being treated (propensity score). **Most commonly used.**

- Match on the basis of the **propensity score**

$$P(X) = Pr(d=1|X)$$

  - D indicates participation in project
  - Instead of attempting to create a match for each participant with exactly the same value of X, we can instead match on the probability of participation.

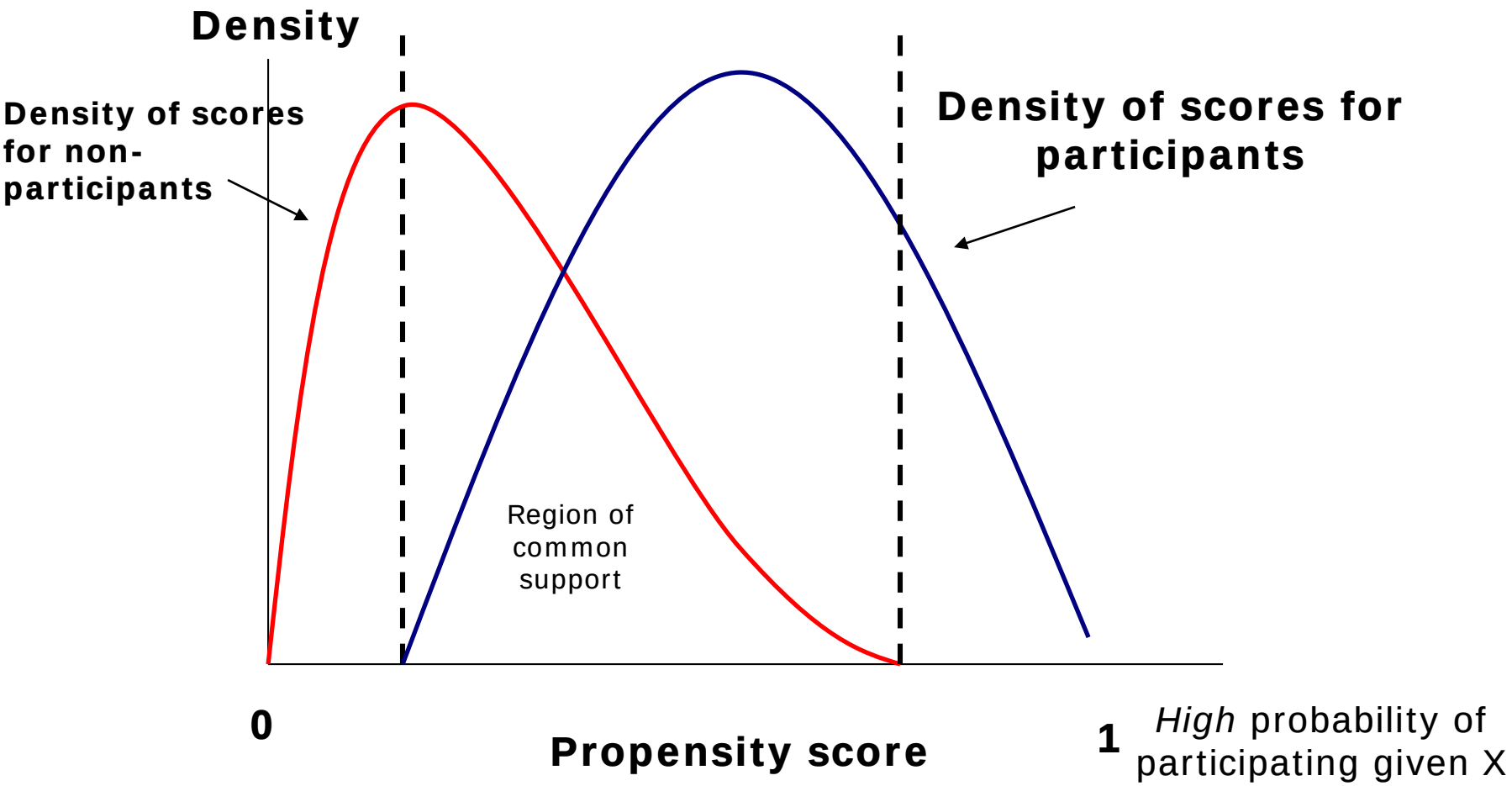
PSM: Key Assumptions

- Key assumption: participation is independent of outcomes conditional on X_i

$$E[y_0 \mid X, d = 1] = E[y_0 \mid X, d = 0]$$

  - This is false if there are unobserved outcomes affecting participation

- Enables matching not just at the mean but balances the distribution of observed characteristics across treatment and control


Density
Density of scores for non-participants
Density of scores for participants
Region of common support
0
Propensity score
1
High probability of participating given X

Steps in Score Matching

1. Need representative and comparable data for both treatment and comparison groups
2. Use a logit (or other discrete choice model) to estimate program participations as a function of observable characteristics
3. Use predicted values from logit to generate propensity score $p(x_i)$ for all treatment and comparison group members

Calculating Impact using PSM

4. Match Pairs:
   - Restrict sample to common support (as in Figure)
   - Need to determine a *tolerance limit*: how different can control individuals or villages be and still be a match?
     - Nearest neighbors, nonlinear matching, multiple matches

5. Once matches are made, we can calculate impact by comparing the means of outcomes across participants and their matched pairs

PSM vs Randomization

- Randomization does not require the *untestable* assumption of independence conditional on observables
- PSM requires large samples and good data:
  1. Ideally, the same data source is used for participants and non-participants
  2. Participants and non-participants have access to similar institutions and markets, and
  3. The data include *X* variables capable of identifying program participation and outcomes.

Lessons on Matching Methods

- Typically used when neither randomization, RD or other quasi experimental options are not possible
  - Case 1: no baseline. Can do ex-post matching
  - Dangers of ex-post matching:
    - Matching on variables that change due to participation (i.e., endogenous)
    - What are some variables that won't change?
- Matching helps control only for OBSERVABLE differences, not unobservable differences

More Lessons on Matching Methods

- Matching becomes much better in combination with other techniques, such as:
  - Exploiting baseline data for matching and using difference-in-difference strategy
  - If an assignment rule exists for project, can match on this rule
- Need good quality data
  - Common support can be a problem if two groups are very different

Case Study: Piped Water in India

- Jalan and Ravaillion (2003): Impact of piped water for children's health in rural India

- Research questions of interest include:

1. Is a child less vulnerable to diarrhoeal disease if he/she lives in a HH with access to piped water?
2. Do children in poor, or poorly educated, HH have smaller health gains from piped water?
3. Does income matter independently of parental education?

Piped Water: the IE Design

- Classic problem for infrastructure programs: randomization is generally not an option (although randomization in timing may be possible in other contexts)
- The challenge: observable and unobservable differences across households with piped water and those without
  - What are differences for such households in Nigeria?
- Jalan and Ravallion use cross-sectional data
  - 1993-1994 nationally representative survey on 33,000 rural HH from 1765 villages

PSM in Practice

- To estimate the propensity score, authors used:
- Village level characteristics
  - Including: Village size, amount of irrigated land, schools, infrastructure (bus stop, railway station)
- Household variables
  - Including: Ethnicity / caste / religion, asset ownership (bicycle, radio, thresher), educational background of HH members
- Are there variables which can not be included?
  - Only using cross-section, so no variables influenced by project

Piped Water: Behavioral Considerations

- IE is designed to estimate not only impact of piped water but to look at how benefits vary across group
- There is therefore a behavioral component: poor households may be less able to benefit from piped water b/c they do not properly store water
- With this in mind, Are there any key variables missing?

Potential Unobserved Factors

- The behavioral factors – importance put on sanitation and behavioral inputs – are also likely correlated with whether a HH has piped water

- However, there are no behavioral variables in data: water storage, soap usage, latrines
  - These are unobserved factors NOT included in propensity score

Table 3
Impacts of piped water on diarrhea prevalence and duration for children under five

	Prevalence of diarrhea		Duration of illness	
	Mean for those with piped water (st. dev.)	Impact of piped water (st. error)	Mean for those with piped water (st. dev.)	Impact of piped water (st. error)
Full sample	0.0108	−0.0023*	0.3254	−0.0957*
	(0.046)	(0.001)	(1.650)	(0.021)
Stratified by household income per capita (quintiles)				
1 (poorest)	0.0155	0.0032*	0.4805	0.0713
	(0.055)	(0.001)	(2.030)	(0.053)
2	0.0136	0.0007	0.4170	0.0312
	(0.051)	(0.001)	(1.805)	(0.051)
3	0.0083	−0.0039*	0.2636	−0.1258*
	(0.038)	(0.001)	(1.418)	(0.042)
4	0.0100	−0.0036*	0.3195	−0.1392*
	(0.044)	(0.001)	(1.703)	(0.048)
5	0.0076	−0.0068*	0.1848	−0.2682*
	(0.042)	(0.001)	(1.254)	(0.036)

Piped Water: Impacts

- Disease prevalence among those with piped water would be 21% higher without it
- Gains from piped water exploited more by wealthier households and households with more educated mothers
- Even find counterintuitive result for low income, illiterate HH: piped water is associated with higher diarrhea prevalence

Design	When to use	Advantages	Disadvantages
Randomization	▫Whenever feasible ▫When there is variation at the individual or community level	▫Gold standard ▫Most powerful	▫Not always feasible ▫Not always ethical
Randomized Encouragement Design	▫When an intervention is universally implemented	▫ Provides exogenous variation for a subset of beneficiaries	▫Only looks at sub-group of sample ▫Power of encouragement design only known ex post
Regression Discontinuity	▫If an intervention has a clear, sharp assignment rule	▫ Project beneficiaries often must qualify through established criteria	▫Only look at sub-group of sample ▫Assignment rule in practice often not implemented strictly
Difference-in-Differences	▫If two groups are growing at similar rates ▫ Baseline and follow-up data are available	▫Eliminates fixed differences not related to treatment	▫Can be biased if trends change ▫Ideally have 2 pre-intervention periods of data
Matching	▫ When other methods are not possible	▫Overcomes observed differences between treatment and comparison	▫Assumes no unobserved differences (often implausible)