17.802 – Recitation 6

Selection on Observables and Match-Makers

Yiqing Xu

MIT

March 14, 2014

The Training Data

- Training is the mother of policy evaluation and an area where new ideas of causal inference are tested.
  - Straightforward counterfactual
  - Classic identification problem (selection bias)
  - Clear experimental benchmark

- Debate over evaluation methods and the debate of model-based methods against randomized trials
  - All start with Lalonde (1986)
  - Changes the trend of social science
    (1360 citations so far, a really high bar for a Ph.D. thesis)

TABLE 5—EARNINGS COMPARISONS AND ESTIMATED TRAINING EFFECTS FOR THE NSW MALE PARTICIPANTS USING COMPARISON GROUPS FROM THE *PSID* AND THE *CPS-SSA*[a,b]

Name of Comparison Group[d]	Comparison Group Earnings Growth 1975–78	NSW Treatment Earnings Less Comparison Group Earnings				Difference in Differences: Difference in Earnings Growth 1975–78 Treatments Less Comparisons		Unrestricted Difference in Differences: Quasi Difference in Earnings Growth 1975–78		Controlling for All Observed Variables and Pre-Training Earnings
		Pre-Training Year, 1975		Post-Training Year, 1978						
		Unadjusted	Adjusted[c]	Unadjusted	Adjusted[c]	Without Age	With Age	Unadjusted	Adjusted[c]	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Controls	$2,063	$39	$ − 21	$886	$798	$847	$856	$897	$802	$662
	(325)	(383)	(378)	(476)	(472)	(560)	(558)	(467)	(467)	(506)
PSID-1	$2,043	− $15,997	− $7,624	− $15,578	− $8,067	$425	− $749	− $2,380	− $2,119	− $1,228
	(237)	(795)	(851)	(913)	(990)	(650)	(692)	(680)	(746)	(896)
PSID-2	$6,071	− $4,503	− $3,669	− $4,020	− $3,482	$484	− $650	− $1,364	− $1,694	− $792
	(637)	(608)	(757)	(781)	(935)	(738)	(850)	(729)	(878)	(1024)
PSID-3	($3,322	($455	$455	$697	− $509	$242	− $1,325	$629	− $552	$397
	(780)	(539)	(704)	(760)	(967)	(884)	(1078)	(757)	(967)	(1103)
CPS-SSA-1	$1,196	− $10,585	− $4,654	− $8,870	− $4,416	$1,714	$195	− $1,543	− $1,102	− $805
	(61)	(539)	(509)	(562)	(557)	(452)	(441)	(426)	(450)	(484)
CPS-SSA-2	$2,684	− $4,321	− $1,824	− $4,095	− $1,675	$226	− $488	− $1,850	− $782	− $319
	(229)	(450)	(535)	(537)	(672)	(539)	(530)	(497)	(621)	(761)
CPS-SSA-3	$4,548	$337	$878	− $1,300	$224	− $1,637	− $1,388	− $1,396	$17	$1,466
	(409)	(343)	(447)	(590)	(766)	(631)	(655)	(582)	(761)	(984)

[a] The columns above present the estimated training effect for each econometric model and comparison group. The dependent variable is earnings in 1978. Based on the experimental data an unbiased estimate of the impact of training presented in col. 4 is $886. The first three columns present the difference between each comparison group's 1975 and 1978 earnings and the difference between the pre-training earnings of each comparison group and the NSW treatments.

[b] Estimates are in 1982 dollars. The numbers in parentheses are the standard errors.

[c] The exogenous variables used in the regression adjusted equations are age, age squared, years of schooling, high school dropout status, and race.

[d] See Table 3 for definitions of the comparison groups.

Reactions to Lalonde (1986)

According to Josh Angrist...

- Initial reaction:
  - Heckman and Hotz (1985): specification testing gets it right
  - Heckman et al. (1997): Better data helps
- Dehejia and Wahba (1999): the propensity scores solves the problem
  - Simth and Todd (2001,2005): No, it doesn't; DW sample is special
  - DW (2005): Yes, it does
  - MHE and Klines: don't need the score; get the controls right
- Cook and Wong (2005, 2007): "well designed" observational studies (RD) come close to an RCT benchmark

Selection on Observables

- How to understand this assumption?
- Selection... on... observables, literally!
  - What to control for? (usually more important)
  - How to control for them?
  - Dale and Krueger (2002) for an example
- If you don't have it, you feel very bad about a cross-sectional observational study

Table 2.2A. Private School Effects Estimated Using Exact Applicant Matches.

	No selection controls			Selection controls		
	(1)	(2)	(3)	(4)	(5)	(6)
Private school	0.296***	0.211**	0.176*	-0.115	-0.123	-0.083
	(0.074)	(0.071)	(0.068)	(0.103)	(0.105)	(0.096)
Own SAT score/100		0.053***	0.017		0.022	-0.004
		(0.011)	(0.011)		(0.012)	(0.014)
Predicted log(parental income)			0.217***			0.186*
			(0.048)			(0.072)
Female			-0.425***			-0.444***
			(0.023)			(0.044)
Black			-0.026			-0.091
			(0.074)			(0.084)
Hispanic			0.031			-0.093
			(0.096)			(0.174)
Asian			0.355***			0.301***
			(0.068)			(0.075)
Other/missing race			-0.405			-0.088
			(0.367)			(0.608)
High school top 10 percent			0.100**			0.051
			(0.032)			(0.049)
High school rank missing			0.030			-0.027
			(0.048)			(0.062)
Athlete			0.139**			0.112
			(0.041)			(0.063)
Selection controls	N	N	N	Y	Y	Y
N	2,330	2,330	2,330	2,330	2,330	2,330

Notes: Models (1)-(3) include no selection controls. Models (4)-(5) include fixed effects for groups formed by matching students according to the average SAT score of each school at which they were accepted or rejected. Each model is estimated using only observations with exact school-SAT matches. The models are estimated by WLS and weighted to make the sample representative of the population of students at the C&B institutions. Each model also includes a constant term. Robust standard errors, clustered on the school attended, are shown in parentheses. * significant at 5%; ** significant at 1%; ***significant at 0.1%

			First Application		Second Application		Third Application		
Applicant Group	Applied to	Student	School	Admission Decision	School	Admission Decision	School	Admission Decision	1996 Earnings
A	Ivy, Leafy, All State	1	Ivy (Private)	Enroll	Leafy (Private)	Reject	All State (Public)	Admit	120,000
		2		Admit		Reject		Enroll	90,000
B	Ball State, Altered State, Smart	3	Ball State (Public)	Admit	Altered State (Public)	Admit	Smart (Private)	Enroll	100,000
		4		Admit		Admit		Enroll	30,000
		5		Enroll		Admit		Admit	70,000
C	Leafy	6	Leafy (Private)	Enroll					90,000
		7		Enroll					75,000
D	Ivy, All State, Ball State	8	Ivy (Private)	Reject	All State (Public)	Enroll	Ball State (Public)	Admit	110,000
		9		Reject		Admit		Enroll	60,000

Notes: Enroll indicates school attended; enrollment decisions are highlighted in grey. Private schools: Leafy, Ivy, and Smart. Public Schools: All State, Ball State, and Altered State

Table 2.1: The College Matching Matrix

General Guidance for Cross-sectional Observational Data

- Make assumptions
  - Selection on observables
  - Common support
- Conditioning
  - Sub-classification
  - Matching
  - Regression
  - Or a combination
- Sensitivity analysis
- Instrument variables (which are really hard to find)

Objective

- Assumptions
  - Selection on observables: $D_i \perp\!\!\!\perp (Y_i(0), Y_i(1))|X_i$
  - Common support: $0 < e(x) < 1$
    in which $e(x) = Pr(D_i = 1|X_i = x)$ is the propensity score
- Data:
  $$\{Y_i^{obs}, D_i, X_i\}, \text{ in which } Y_i^{obs} = \begin{cases} Y_i(0) & \text{if } D_i = 0 \\ Y_i(1) & \text{if } D_i = 0 \end{cases}$$
- Estimand:

$$\begin{aligned} \tau &= \mathbb{E}[Y_i(1) - Y_i(0)] \\ &= \mathbb{E}[\mathbb{E}[Y_i(1) - Y_i(0)|X_i]] \\ &= \mathbb{E}[\mathbb{E}[Y_i(1)|X_i]] - \mathbb{E}[\mathbb{E}[Y_i(0)|X_i]] \\ &= \mathbb{E}[\mathbb{E}[Y_i^{obs}|D_i = 1, X_i]] - \mathbb{E}[\mathbb{E}[Y_i^{obs}|D_i = 0, X_i]] \end{aligned}$$

$$\tau_{treat} = \mathbb{E}[Y_i^{obs}|D_i = 1] - \mathbb{E}[\mathbb{E}[Y_i^{obs}|D_i = 0, X_i]|D_i = 1]$$

Matching and Regression (when it works)

- Matching gives you:

$$\tau = \sum_x \tau_x P(X_i = x)$$

$$\tau_{treat} = \sum_x \tau_x P(X_i = x | D_i = 1)$$

- Regression gives you:

$$\tau_R = \frac{\mathbb{E}[\sigma_D^2(X_i)\tau_X]}{\mathbb{E}[\sigma_D^2(X_i)]}, \quad \sigma_D^2(X_i) = var(D|X)$$

$$= \frac{\sum_x \tau_x P(X_i = x)[e(x)(1-e(x))]}{\sum_x P(X_i = x)[e(x)(1-e(x))]}$$

- ATT weights covariate cells in proportion to the probability of treatment; while regression weights in proportion to the conditional variance of treatment.

Gosh, so many estimators!

- Sub-classfication
- Oaxaca-Blinder (Kline 2011)
- Matching of different kinds
- Matching on PS
- PS reweighting
- PS trimming
- Regression on mean-balanced sample

- The heart of the problem is to make (some moments of) the two joint distributions similar

When Regression Doesn't Work?

- Lack of common support

  Regression replies on extrapolation, which requires right specification
- Weights

  Regression takes the linearity assumption very seriously, thus observations with extreme values are influential

- Instead, matching changes estimand in the lack of common support
- Selection on observables is always an issue

Practical Advice – Imbens (2014)

I. Design
   - Determining your estimand
   - Assessing overlap
   - Trimming the data based on propensity score (specification searches are allowed)

II. Assessing selection on observables:
   - placebo tests for pre-treatment covariates, esp. lagged outcome

III. Analysis based on trimmed data:
   - 1.a Sub-classification based on PS
   - 1.b Matching with replacement on PS and bias-adjustment
   - 2 Conditional variance based on Abadie and Imbens (2006) (slide 100)

In Summary

- Be careful, the SOO assumption literally means "selection on observables"
- Both regression and matching are reweighting schemes, though the latter is often more transparent
- Bootstrap is not allowed for matching

- What to control for is often more important than how to control for them
- Practically, trimming (to ensure common support) is important
- If you do it right, all methods should give you roughly the same result

Midterm Review

- Potential outcome framework
- Randomized experiments
- Inference
- Selection on observables

Potential Outcome Framework

- Notations and simple derivation
- SUTVA
- Estimands: ATE, ATT, ATC ...
- Bias equation

Randomized Experiments

- Random assignment assumption
- Point estimation (difference-in-means, blocking)
- Variance estimation
- Covariate adjustment
- Using regression to estimate ATE

Inference

- SATE and PATE
- Clustering
- Hypothesis test and power calculations
- Bootstrap
- Permutation inference
- Sharp Null and Weak Null

Selection on Observables

- The SOO and overlap assumptions
- "Bad control" and post-treatment bias
- Sub-classification
- Matching (mechanics and bias-correction)
- PS methods
- Regression anatomy