Labor Economics I

Problem Set 4 - Suggested Solutions

April 5, 2024

Question 1 - Search and Matching

1) Read article by Crépon et al. (2013) that we discussed in class (can be found in Lecture 8 folder). Answer the following questions:

a) Explain briefly what the treatment was and who was eligible for it?

- Treatment: Job-placement program that helped job seekers find a durable job(in addition to the standard assistance and counselling normally available to French job seekers). The program consisted of two parts: Phase I) counselling job seeker to find durable job; Phase II) advise job seeker in new job (to help them keep it)
- Eligible: young graduates (with at least a two-year degree) who have been unemployed for over 6 months

Question 1b

b) How should the treatment affect the hiring probabilities of treated workers?

- Treated workers should be more likely to have a fixed work contract (over 6 months long) than unassigned workers overall, and higher probability still than unassigned workers in their area.

Question 1c

c) According to the model, how does the share of workers receiving counselling affect hiring probabilities of workers that did not receive counselling?

- The treatment moves the Beveridge curve to the right and has no effect on the labour demand curve.
- This decreases the exit rate of the untreated workers, thus lowering their hiring probabilities.

Question 1d

d) Explain briefly how would you draw the labor demand curve in Figure 1 if the model would be similar to Pissarides (2000), where return to scale in production function is constant.

- The labour demand curve would be horizontal (which means the shift in the Beveridge curve does not lead to any displacement).

Question 1e

e) Table 4 presents the estimates of the unconstrained reduced form equation that pasted below. How should we interpret the coefficient β_{25}?

$$\begin{aligned} y_{ic} = &\beta_{25} Z_{ic} P_{25c} + \beta_{50} Z_{ic} P_{50c} + \beta_{75} Z_{ic} P_{75c} + \beta_{100} Z_{ic} P_{100_c} \\ &+ \delta_{25} P_{25c} + \delta_{50} P_{50c} + \delta_{75} P_{75c} \\ &+ X_{ic} \gamma_4 + u_{ic} \end{aligned}$$

- Coefficient β_{25} indicates the effect of being assigned to treatment in an area where 25% of eligible job seekers was assigned to treatment, compared to being unassigned in the same area (since the model includes main effects for treatment assignment intensity, we always compare assigned and unassigned workers within these areas) .
- According to the table 6, treated workers in areas with 25% of workers treated are 1.6 percentage points more likely to find a job contract.

Question 1f

f) How should one interpret δ_{75}? What is its value according to Table 4?

- Coefficient δ_{75} indicates the effect of not being invited to treatment in an area where 75% of workers are assigned to treatment, compared to being unassigned in the super control group where assignment rate was 0%.
- According to table IV workers in these areas are 1.6 percentage points more likely to find treatment, as compared to workers in areas where no one is assigned to treatment.

Question 2 - Fertility

We use Fertility Small Data Set which is subset of original data used in Angrist and Evans (1998)

Variable Description

1. morekids =1 if mom had more than 2 children (0, if just 2 children)
2. boy1st =1 if 1st child was a boy
3. boy2nd =1 if 2nd child was a boy
4. samesex =1 if 1st two children same sex
5. agem1 age of mom at census
6. black =1 if mom is black
7. hispan =1 if mom is Hispanic
8. othrace =1 if mom is not black, Hispanic or white
9. weeksm1 mom's weeks worked in 1979

Question 2a

a) Run an OLS regression that explains weeks worker for mother with having more than 2 kids. Do not add any another controls for your regression. Interpret the coefficient on "morekids".

```
. reg weeksm1 morekids
```

Source	SS	df	MS	Number of obs	=	30,000
				F(1, 29998)	=	538.16
Model	254515.369	1	254515.369	Prob > F	=	0.0000
Residual	14187250.9	29,998	472.939893	R-squared	=	0.0176
				Adj R-squared	=	0.0176
Total	14441766.3	29,999	481.408257	Root MSE	=	21.747

weeksm1	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
morekids	-6.008217	.2589951	-23.20	0.000	-6.515858	-5.500575
_cons	21.4782	.1591503	134.96	0.000	21.16626	21.79014

- On average, women with more than two children worked 6 weeks fewer than women with only two children.

Question 2b

b) Add controls that you want (motivate why you want them). How do you interpret the coefficient on more kids variable?

```
. reg weeksm1 morekids agem1 black hispan othrace
```

Source	SS	df	MS		
Model	688472.176	5	137694.435	Prob > F	0.0000
Residual	13753294.1	29,994	458.534844	R-squared	0.0477
				Adj R-squared	0.0475
Total	14441766.3	29,999	481.408257	Root MSE	21.413

Number of obs	=	30,000				
F(5, 29994)	=	300.29				
Prob > F	=	0.0000				
R-squared	=	0.0477				
Adj R-squared	=	0.0475				
Root MSE	=	21.413				

weeksm1	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
morekids	-6.896003	.2576912	-26.76	0.000	-7.401088	-6.390917
agem1	.8417492	.036849	22.84	0.000	.7695236	.9139748
black	11.53423	.5525237	20.88	0.000	10.45126	12.6172
hispan	-.2473107	.5214906	-0.47	0.635	-1.269455	.7748333
othrace	3.334692	.5892012	5.66	0.000	2.179833	4.489552
_cons	-4.52276	1.122571	-4.03	0.000	-6.723048	-2.322472

Question 2b

- Controls: age and race dummies. Effectively comparing weeks worked among mothers of the same age and race. Controlling for children's gender should be irrelevant.
- Coefficient: on average, women with more than two children worked 6.9 weeks fewer than women with only two children.

Question 2c

c) Pick a variable from the data that can be used to instrument having more children (same as used in Angrist and Evans(1998)). Discuss whether this instrument is valid or not.

- Angrist and Evans (1998) exploit the preference of parents to have children of both genders. In other words, parents who have two children of the same gender should be more likely to have a third child compared to parents having two children of different genders.

$\Rightarrow$ Potential instrument: samesex. Is it a valid instrument?

Question 2c

1. Random assignment: Gender of children should be orthogonal to observed parent characteristics.
2. Relevance: Probability of having more than two children: 41.1% (samesex = 1) vs. 34.4% (samesex = 0).
3. Exclusion restriction: samesex should not affect weeks worked through some other mechanism than the probability of having another child.
4. Endogeneity: Weeks worked should be orthogonal to the gender of children.

$\Rightarrow$ samesex seems to be a valid instrument for the probability of having more than two children. It would be invalid if one of these assumptions is violated.

Question 2d

d) Following Angirst and Evans (1998), estimate the effect of having more than two kids on female labor supply (weeksm1) using the instrumental variable method. Interpret the coefficient and discuss how it differs from the coefficient from the OLS model in part b).

```
. ivregress 2sls weeksm1 agem1 black hispan othrace (morekids = samesex), first

First-stage regressions
```

Number of obs	= 30,000
F(5, 29994)	= 157.10
Prob > F	= 0.0000
R-squared	= 0.0255
Adj R-squared	= 0.0254
Root MSE	= 0.4786

morekids	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
agem1	.016438	.0008182	20.09	0.000	.0148342	.0180417
black	.0962641	.0123369	7.80	0.000	.0720833	.1204449
hispan	.1484327	.0116251	12.77	0.000	.1256471	.1712183
othrace	.0235263	.0131685	1.79	0.074	-.0022846	.0493372
samesex	.0678624	.0055269	12.28	0.000	.0570294	.0786954
_cons	-.1730167	.0252522	-6.85	0.000	-.2225122	-.1235212

Question 2d

Instrumental variables 2SLS regression

Number of obs	= 30,000
Wald chi2(5)	= 787.51
Prob > chi2	= 0.0000
R-squared	= 0.0471
Root MSE	= 21.418

weeksm1	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
morekids	-5.780746	3.644609	-1.59	0.113	-12.92405	1.362555
agem1	.8234973	.0699881	11.77	0.000	.6863231	.9606714
black	11.42628	.6551618	17.44	0.000	10.14219	12.71037
hispan	-.4117677	.7479726	-0.55	0.582	-1.877767	1.054232
othrace	3.307789	.5958159	5.55	0.000	2.140011	4.475567
_cons	-4.370342	1.227826	-3.56	0.000	-6.776838	-1.963847

Instrumented: morekids
Instruments: agem1 black hispan othrace samesex

Question 2d

- The IV estimate is slightly smaller than the OLS estimate (-5.78 $<$ -6.90).
- The IV estimate is noisier (higher SE).
- The OLS estimate is likely to suffer from omitted variable bias. Mothers who choose to have more than two children might have different preferences for labor supply than those who choose otherwise.

Question 2e

e) Lundborg et al. 2017 uses alternative strategy to estimate the impact of children on women's labor supply. Why the estimates of fertility on female labor should is likely to differ between Angrist and Evans, 1998, and Lundborg et al. 2017?

- **Angrist & Evans (1998)**: exogenous variation in the probability of having an additional child after having two children.
- **Lundborg et al. (2017)**: exogenous variation in the probability of having a child at all.
- $\Rightarrow$ Intensive vs. Extensive fertility margin.
- The effect on labor supply and earnings of the first compared to the third child is likely to be different. The instruments identify effects for different groups, whose local average treatment effects are different. For example, IVF-takers are likely to be older and more educated.