Econometrics 2

PROBLEM SET 3   18-2-2020

Exercise - Theory

Exercise - 1

Potential outcomes

$$\begin{cases} Y_1 = X\beta^1 + U^1 \\ Y_0 = X\beta^0 + U^0 \end{cases}$$

Assignment to treatment $D(Z) \in \{0,1\}$

Based in an unobservable rule

$D^*(Z) = Z\theta + U^D$ $D(Z) \begin{cases} \text{if } Z\theta + u^D \geq 0, \ D(Z) = 1 \\ \text{if } Z\theta + u^D < 0, \ D(Z) = 0 \end{cases}$

- we can say $Z\theta$ gains from assignment to treatment u^D costs $\Longrightarrow$ ① if $D(Z) = 1$ $Z\theta \geq u^D$ through theory of revealed preference
- if $D(Z) = 0$, $Z\theta < u^D$ through theory of revealed preference

It holds that $(U^1\ U^0\ U^D)$ independent of Z, X:

$cov(U^1, Z) = 0$ & $E(U^1) =$
$cov(U^0, Z) = 0$ $E(U^0) =$
$cov(U^D, Z) = 0$ $E(U^D) = 0$
$cov(U_1, X) = 0$
$cov(U^0, X) = 0$
$cov(U^D, X) = 0$

(U^1, U^0, U^D) are jointly normally distributed

with $\begin{bmatrix} U^D \\ U^1 \\ U^D \end{bmatrix} \sim N\left(0, \begin{bmatrix} 1 & \sigma_{1D} & \sigma_{0D} \\ \sigma_{1D} & \sigma_1^2 & \sigma_{1D} \\ \sigma_{0D} & \sigma_{1D} & \sigma_0^2 \end{bmatrix}\right)$ ③

$U^D \sim N(0,1)$, standardiz. normal distrib

- For $X=x$ & $Z=z$ the ATT assuming noselection bias

$ATT(x, z, D(z)=1) =$

$E[Y^1 - Y^0 / x, z, D(z)=1] \overset{\text{by observed outcomes}}{=}$

$E[(x b^1 + U^1) - (x b^0 + U^0) / x, z (D(z)=1]$

$= E(x(b^1 - b^0) / x, z, D(z)=1] +$

$+ E(U^1 - U^0 / x, z, D(z)=1] =$

because for give $X=x$ & b_1, b_0 determ. comp all comp is obs.

$= \underbrace{x(b^1 - b^0)}_{\text{observed component}} + \underbrace{E(U^1 - U^0 / x, z, D(z)=1)}_{\text{unobserved}}$ (1)

We know that

$D(z)=1$ if $z\theta + U^D \geq 0$

So

unobs component can be written

2

$E(U^1 - U^0 / x, z, D(z)=1) =$

$E(U^1 - U^0 / x, z, z\theta + U^D \geq 0) =$

$E(U^1 - U^0 / x, z, U^D \geq -z\theta) =$

$E[U^1 / x, z, U^D \geq -z\theta] - E[U^0 / x, z, U^D \geq -z\theta]$

expectation of each part of the unobserved component given that the benefits of taking treatment are higher than the costs so for indiv. that took up the treatment.

So (I) becomes

$(I) = x(\beta^1 - \beta^0) + E[U^1 / x, z, U^D \geq -z\theta] -$
$- E[U^0 / x, z, U^D \geq -z\theta]$

As X, Z are independent with respect ~~to~~ U^1, U^2, U^D we can drop $z=Z, x=X$ from the conditional expectection

$ATT(x, z, D(z)=1) = x(\beta^1 - \beta^0) + E(U^1 / x, z, U^D \geq -z\theta) - E(U^0 / x, z, U^D \geq z\theta)$)*

$x(\beta^1 - \beta^0) + E(U^1 / U^D \geq -z\theta) - E(U^0 / U^D \geq -z\theta)$

we can standarize the distribution of U^1, U^0, U^D by multiply & divide by $\sigma_1, \sigma_0, \sigma_D$ respectively

$x(\beta^1 - \beta^0) + \sigma_1 E\left[\frac{U_1}{\sigma_1} / \frac{U_D}{\sigma_D} \geq -\frac{z\theta}{\sigma^D}\right] - \sigma_0 E\left[\frac{U_0}{\sigma_0} / \frac{U_D}{\sigma_D} \geq -\frac{z\theta}{\sigma_D}\right]$

3 x

· we know that $E[U^0/U^P] = \underbrace{\frac{\sigma_{0D}}{\sigma^2_D}}_{\frac{\operatorname{Cov}(U_0,U_D)}{\operatorname{Var}(U_D)}} U_D$

regressor coefficient

· we know that correl

$\rho_{0,1} = \operatorname{Cor}(U^{1,0}, U^D) = \frac{\sigma_{0,1D}}{\sigma_{0,1}, \sigma_D}$ → covar(0/1, D)

(ρ_0 or ρ_1 respectively)

$\sigma_{0,1}$ → Var(0/1), σ_D → var(D)

· So $E\left[\frac{U_0}{\sigma_0} / \frac{U_D}{\sigma_D}\right] = \frac{\sigma_{0D}}{\sigma_D^2} \frac{\cancel{\sigma_D}}{\sigma_0} U_D =$

$= \frac{\sigma_{0D}}{\sigma_D \sigma_0} U_D = \rho_0 \frac{U_D}{\sigma_D}$

& $E\left[\frac{U_1}{\sigma_1} / \frac{U_D}{\sigma_D}\right] = \rho_1 \frac{U_D}{\sigma_D}$

· We know that U_1, U_0 are distrib normally (jointly & individually)

Left-truncated moments of standard random variable(x) holds

$E[x / x > \alpha] = \frac{\phi(a)}{1-\Phi(a)}$, $\phi(a)$ PDF, $\Phi(a)$ CDF

Comes from $E(X)$: mean of truncated (between a & b) normal distribution

$E(X) = \mu + \sigma \frac{\phi(a) - \phi(b)}{\Phi(b) - \Phi(a)}$

- Since in our case distribution is left truncated, $b \to \infty \Rightarrow \phi(b)=0, \Phi(b)=1$
- Since standardize normal distribution $\gamma = 0, \sigma = 1$

$$\gamma + \sigma \frac{\phi(a) - \phi(b)}{\Phi(b) - \Phi(a)} = \frac{\phi(a)}{1 - \Phi(a)}$$

→ Based on the above points:

$$ATT(x, z, D(z)=1) = x(\beta^1 - \beta^0) + \sigma_1 E\left[\frac{U_1}{\sigma_1} \middle| \frac{U_D}{\sigma_D} \geq -\frac{z\vartheta}{\sigma_D}\right] - \sigma_0 E\left[\frac{U_0}{\sigma_0} \middle| \frac{U_D}{\sigma_D} \geq -\frac{z\vartheta}{\sigma_D}\right] =$$

$$= x(\beta^1 - \beta^0) + \rho_1 \sigma_1 E\left[\frac{U^D}{\sigma^D} \middle| \frac{U^D}{\sigma^D} \geq -\frac{z\vartheta}{\sigma_D}\right] - \rho_0 \sigma_0 E\left[\frac{U^D}{\sigma^D} \middle| \frac{U^D}{\sigma^D} \geq -\frac{z\vartheta}{\sigma^D}\right] =$$

$$x(\beta^1 - \beta^0) + \rho_1 \sigma_1 \frac{\phi\left(-\frac{z\vartheta}{\sigma_D}\right)}{1 - \Phi\left(-\frac{z\vartheta}{\sigma_D}\right)} - \rho_0 \sigma_0 \frac{\phi\left(-\frac{z\vartheta}{\sigma_D}\right)}{1 - \Phi\left(-\frac{z\vartheta}{\sigma_D}\right)} =$$

$$= X(\beta^1-\beta^0) + \rho_1\sigma_1 \frac{\phi\left(\frac{Z\theta}{\sigma_D}\right)}{1-\left(1-\Phi\left(-\frac{Z\theta}{\sigma_D}\right)\right)}$$

$$-\rho_0\sigma_0 \frac{\phi\left(\frac{Z\theta}{\sigma_D}\right)}{1-\left(1-\Phi\left(\frac{Z\theta}{\sigma_D}\right)\right)}$$ by symmetry of the normal distribution

$$= X(\beta^1-\beta^0) + \rho_1\sigma_1 \frac{\phi\left(\frac{Z\theta}{\sigma_D}\right)}{\Phi\left(\frac{Z\theta}{\sigma_D}\right)} -$$

$$-\rho_0\sigma_0 \frac{\phi(Z\theta)}{\Phi\left(\frac{Z\theta}{\sigma_D}\right)}$$ we know that $\sigma_D = 1$

$$= X(\beta^1-\beta^0) + \rho_1\sigma_1 \frac{\phi(Z\theta)}{\Phi(Z\theta)} - \rho_0\sigma_0 \frac{\phi(Z\theta)}{\Phi(Z\theta)}$$

$$\boxed{= X(\beta^1-\beta^0) + (\rho_1\sigma_1-\rho_0\sigma_0) \frac{\phi(Z\theta)}{\Phi(Z\theta)}}$$

So in our case there is selection into treatment based on potential gains from treatment that are differential amongst treatment and control group.

We estimate consistently the ATT by making statistical assumptions about the distributions of the unobs. components (following Heckman 1978)

Exercise 2

Average Treatment Effect

$ATE(x,z) = E[Y^1 - Y^0 / x,z] = E[Y^1/x,z] -$

$- E[Y^0/x,z] =$

$E[xB^1 + U^1/x,z] - E[xB^0 + U^0/x,z] =$

$= xB^1 + \underbrace{E(U^1/x,z)}_{\text{by exogeneity assumption: } 0} - xB^0 - \underbrace{E(U^0/x,z)}_{\text{by exogeneity assumption } 0}$

$= x(B^1 - B^0)$

From Exercise 1

Average Treatment Effect on the Treatment

$ATT(x,z,D(z)=1) = x(B^1 - B^0) + \underbrace{(\rho_{1\sigma_1} - \rho_{0\sigma_0})}_{*} *$

$\underbrace{\cdot \frac{\phi(z\theta)}{\Phi(z\theta)}}_{*}$

component * captures self-selection into treatment.

- For them to be equal there should be no self selection into treatment and component * should be zero

7α

For no self selection to treatment gains from treated for those treated should be the same as the average individual

$cov(U^D, U^1 - U^0) = 0$ (dif gains from treatment should be uncorrelated with factors that determine participation into treatment choice)

$$\circledast \; [\rho_1\sigma_1 - \rho_0\sigma_0] \frac{\phi(Z\theta)}{\Phi(Z\theta)} = 0 \iff$$

$$\rho_1\sigma_1 - \rho_0\sigma_0 = 0 \iff \rho_1\sigma_1 = \rho_0\sigma_0$$

But if $cov(U^D, U^1 - U^0) > 0$ people self-select into treatment based on their gains from treatment. In that case gains from treatment for those observed to have taken the treatment exceed those on average

so

$$\rho_1\sigma_1 > \rho_0\sigma_0 \Longrightarrow ATT > ATE$$

7b

Ex.3

a) LATE Effect of treatment on compliers

Based on $D^*(Z) = Z\vartheta + u_D$ for someone to be a complier if $Z = z' \Rightarrow D^* = z'\vartheta + u_D$ & $z'\vartheta + u_D \geq 0$ (benefits outrank the costs, $D(z') = 1$

and if $Z = z'$, $z' \neq z$ then $D^* = z\vartheta + u_D$ & $z\vartheta + u_D < 0$ (cost outrank benefits), $D(z) = 0$

If i was an always-taker $D^*(Z) = 1$ for both z & z'

If i was never-taker $D^*(Z) = 0$ for both z & z'

So $LATE(x, D(z)=0, D(z')=1) =$

$E[Y^1 - Y^0 / x, D(z)=0, D(z')=1] =$

$= E[Y^1 / x, D(z)=0, D(z')=1] - E[Y^0 / x, D(z)=0, D(z')=1] =$

$= E[x\beta^1 + U^1 / x, D(z)=0, D(z')=1] -$

$- E[x\beta^0 + U^0 / x, D(z)=0, D(z')=1] =$

$= x\beta^1 + E[U^1 / x, D(z)=0, D(z')=1] -$

$(x\beta^0 - E[U^0 / x, D(z)=0, D(z')=1] =$

$= x(\beta^1 - \beta^0) + E[U^1 / x, D(z)=0, D(z')=1] -$

$- E[U^0 / x, D(z)=0, D(z')=1]$ (I)

• For compliers

$D(z)=0 \Leftrightarrow z\vartheta + U^D < 0 \Leftrightarrow U^D < -z\vartheta$

$\& D(z')=1 \Leftrightarrow z'\vartheta + U^D \geq 0 \Leftrightarrow U^D \geq -z'\vartheta$

$(I) = x(\beta^1-\beta^0) + E[U^1 / x, U^D < -z\vartheta, U^D \geq -z'\vartheta]$

$- E[U^0 / x, U^D < -z\vartheta, U^D \geq -z'\vartheta] =$

$= x(\beta^1-\beta^0) + E[U^1 / x, -z'\vartheta \leq U^D < -z\vartheta]$

$- E[U^0 / x, -z'\vartheta \leq U^D < z\vartheta] =$

per independence of U^1, U^D, U^0 and x, z we can drop the x

So: $LATE(x, D(z)=0, D(z')=1) =$

$x(\beta^1-\beta^0) + E[U^1 / -z'\vartheta \leq U^D < -z\vartheta]$

$- E[U^0 / -z'\vartheta \leq U^D < -z\vartheta] =$ (by normalization)

$= x(\beta^1-\beta^0) + \sigma_1 E\left[\frac{U_1}{\sigma_1} / -\frac{z'\vartheta}{\sigma_D} \leq \frac{U^D}{\sigma_D} < -\frac{z\vartheta}{\sigma_D}\right]$

$- \sigma_0 E\left[\frac{U^0}{\sigma_0} / -\frac{z'\vartheta}{\sigma_D} \leq \frac{U^D}{\sigma_D} < -\frac{z\vartheta}{\sigma_D}\right] =$

We know from Exercise 1

$E\left[\frac{U_0}{\sigma_0} / \frac{U_D}{\sigma_D}\right] = \rho_0 \frac{U_D}{\sigma_D}, \quad E\left[\frac{U_1}{\sigma_1} / \frac{U_D}{\sigma_D}\right] = \rho_1 \frac{U^D}{\sigma_D}$

We also know that a both side truncated moment

$$E[x \mid a \le x \le b] = \frac{\phi(a') - \phi(\beta)}{\Phi(\beta) - \Phi(a')},$$

$$\alpha' = \frac{\alpha - \mu_x}{\sigma_x}$$

$$\beta = \frac{b - \mu_x}{\sigma_x}$$

$$LATE\,(x, D(z)=0, D(z')=1) = x(\beta^1 - \beta^0) + \rho_1\sigma_1 \frac{\phi\left(-\frac{z'\theta}{\sigma_D}\right) - \phi\left(-\frac{z\theta}{\sigma_D}\right)}{\Phi\left(-\frac{z\theta}{\sigma_D}\right) - \Phi\left(-\frac{z'\theta}{\sigma_D}\right)} - \rho_0\sigma_0 \frac{\phi\left(-\frac{z'\theta}{\sigma_D}\right) - \phi\left(-\frac{z\theta}{\sigma_D}\right)}{\Phi\left(-\frac{z\theta}{\sigma_D}\right) - \Phi\left(-\frac{z'\theta}{\sigma_D}\right)}$$

by symmetry of the normal distribution

$$= x(\beta^1 - \beta^0) + \rho_1\sigma_1 \frac{\phi\left(\frac{z'\theta}{\sigma_D}\right) - \phi\left(\frac{z\theta}{\sigma_D}\right)}{\left(1 - \Phi\left(\frac{z\theta}{\sigma_D}\right)\right) - \left(1 - \Phi\left(\frac{z'\theta}{\sigma_D}\right)\right)} - \rho_0\sigma_0 \frac{\phi\left(\frac{z'\theta}{\sigma_D}\right) - \phi\left(\frac{z\theta}{\sigma_D}\right)}{\left(1 - \Phi\left(\frac{z\theta}{\sigma_D}\right)\right) - \left(1 - \Phi\left(\frac{z'\theta}{\sigma_D}\right)\right)} =$$

$$= x(\beta^1 - \beta^0) + \rho_1\sigma_1 \frac{\phi(z'\theta) - \phi(z\theta)}{\Phi(z'\theta) - \Phi(z\theta)} - \rho_0\sigma_0 \frac{\phi(z'\theta) - \phi(z\theta)}{\Phi(z'\theta) - \Phi(z\theta)} =$$

$$\boxed{= x(\beta^1-\beta^0) + (\rho_1\sigma_1 - \rho_0\sigma_0)\,\frac{\phi(z'\theta)-\phi(z\theta)}{\Phi(z'\theta)-\Phi(z\theta)}}$$

b) LATE: effect of treatment on compliers

$$\underbrace{\frac{E[Y_i / Z_i=1]-E[Y_i/Z_i=0]}{E[D_i/Z_i=1]-E[D_i/Z_i=0]}}_{\text{wald estimator}} =$$

$$= E[Y_{1i}-Y_{0i} \;/ D_{1i} > D_{0i}]$$

ATU: weighted avrge of the effects of the treatment on compliers (who in that case didn't receive treatment) and on never-taken

$$E[Y_{1i}-Y_{0i}/D_i=0] = \underbrace{E[Y_{1i}-Y_{0i}/D_{1i}=0]}_{\text{effect on never taker}}\;\underbrace{P(D_{1i}=0/D_i=0)}_{\text{probability/portion of never-takers that someone is never-taker}}$$

$$+\underbrace{E[Y_{1i}-Y_{0i}/D_{1i}>D_{0i}]}_{\text{Effect on compliers LATE}}\;\underbrace{P(D_{1i}>D_{0i}, Z_i=0/ D_i=0)}_{\text{probability that someone in control group is complier / portion of compliers}}$$

From the above we see that if there is perfect compliance (no never-takers) then LATE and ATU is the same

In case where instrument is a compulsory schooling reform that forces everyone to take-up the treatment (eg Black Devereux & Salvanes (2005); Devereux & Hart (2010)) and everyone conforms with the law/reform in place then the LATE and ATU is the same

Econometrics 2

Problem Set 3 1

Exercise 1 - Theory

Question 4

$\sqrt{}$ Marginal Treatment Effects at $x = X^0$, $u^D = U^D$ (define MTE at a specif point)

We know $MTE(v, x) = E[Y_1 - Y_0 / U^D = u^D, X = x]$

$= E[Y_1 / U^D = u^D, X = x] - E[Y_0 / U^D = u^D, X = x]$

(we know $Y_1 = X\beta^1 + U_1$ & $Y_0 = X\beta^0 + U_0$)

$= E[X\beta^1 + U_1 / U^D = u^D, X = x] - E[X\beta^0 + U^0 / U^D = u^D, X = x]$

$x\beta_1 + E[U^1 / u^D, x] - x\beta^0 + E[U^0 / u^D, x]$

$= x(\beta_1 - \beta_0) + E[U^1 / U^D, x] - E[U^0 / U^D, x]$

by independence from x we can drop out x from the conditioning

$$= X(\beta^1-\beta^0) + E[U^1/u^D] -$$

$$- E[U^0/u^D]$$

(standarize conditional distribution as in previous questions)

$$= X(\beta^1-\beta^0) + \sigma_1 E\Big[\frac{U^1}{\sigma_1} \Big/ \underbrace{\frac{u_D}{\sigma_D} = \frac{U_D}{\sigma_D}}_{\substack{\text{we know} \\ u^D = U^D \\ (\cdot U^D \text{ evaluated} \\ \text{at } u_D)}}\Big] -$$

$$- \sigma_0 E\Big[\frac{U^0}{\sigma_0} \Big| \frac{u^D}{\sigma_D} = \frac{U^D}{\sigma_D}\Big] =$$

(We know $E\Big[\frac{U_0}{\sigma_0} \Big/ \frac{U^D}{\sigma^D}\Big] = \frac{\sigma_{0D}}{\sigma_D^2 \sigma_0} U_D$

$$\overset{*}{=} \rho_0 \frac{U^D}{\sigma_D}$$

& $E\Big[\frac{U_1}{\sigma_1} \Big/ \frac{U^D}{\sigma^D}\Big] \overset{*}{=} \frac{\sigma_{1D}}{\sigma_1} U_D$)

$$= X(\beta^1 - \beta^0) + \sigma_1 \frac{\sigma_{1D}}{\sigma_D \sigma_1} u_D$$

$$- \sigma_0 \frac{\sigma_{0D}}{\sigma_D \sigma_0} u_D =$$

(We know that $\rho_1 = \frac{\sigma_{1D}}{\sigma_D \sigma_1}$

& $\rho_0 = \frac{\sigma_{0D}}{\sigma_D \sigma_0}$)

$$= X(\beta^1 - \beta^0) + \sigma_1 \rho_1 u^D - \sigma_0 \rho_0 u^D$$

$$\Leftrightarrow X(\beta^1 - \beta^0) + \rho_1 \sigma_1$$

$$\boxed{MTE(x, u^D) = X(\beta^1 - \beta^0) + (\rho_1 \sigma_1 - \rho_0 \sigma_0) u^D}$$