Econ 675 Assignment 1

Nathan Mather

September 27, 2018

Contents

List of Figures

List of Tables

1 Question 1: Simple Linear Regression with Measurement Error

1.1 OLS estimator

$\hat{\beta}_{ls} = (\tilde{x}'\tilde{x})^{-1}\tilde{x}'y$ and we want to show that $\hat{\beta}_{ls} \to_p \lambda\beta$

First note that

$$y = \beta(\tilde{x} - \mu) + \epsilon = \beta\tilde{x} + (\epsilon - \beta\mu)$$

So The measurement error in x becomes part of the error term in the regression. This means OLS will lead to a negative bias in $\hat{\beta}_{ls}$ if the true β is positive and a positive bias in $\hat{\beta}_{ls}$ if the true β is negative (an attenuation bias). In order to determine the magnitude of the bias consider the following.

$$\hat{\beta}_{ls} = \frac{\text{Cov}(\tilde{x}, y)}{\text{Var}(\tilde{x})} = \frac{\text{Cov}(x+\mu, \beta x+\epsilon)}{\text{Var}(x+\mu)} = \frac{\beta\text{Cov}(x,x) + \text{Cov}(x,\epsilon) + \text{Cov}(\mu, \beta x) + \text{Cov}(\mu, \epsilon)}{\text{Var}(x+\mu)}$$

$$= \frac{\beta\text{Var}(x)}{\text{Var}(x+\mu)} \rightarrow_p \frac{\beta\sigma_x^2}{\sigma_x^2+\sigma_\mu^2} = \lambda\beta$$

$$\text{This implies that } \lambda = \frac{\sigma_x^2}{\sigma_x^2+\sigma_\mu^2}$$

1.2 Standard Errors

Start with $\hat{\epsilon} = y - \hat{\beta}_{ls}(x+\mu)$

Now add and subtract the True error term $\epsilon = y - \beta x$ and collect terms to get $\hat{\epsilon} + \epsilon - \epsilon = \epsilon - (y - \beta x) + y - \hat{\beta}_{ls} x - \hat{\beta}_{ls}\mu = \epsilon + (\beta - \hat{\beta}_{ls})x - \hat{\beta}_{ls}\mu$

recall that $\hat{\beta}_{ls} \rightarrow_p \lambda\beta$ and that ϵ, x, μ are all uncorrelated. This implies that $\hat{\sigma_\epsilon^2} \rightarrow_p \sigma_\epsilon^2 + (1-\lambda)^2\beta^2\sigma_x^2 + \lambda^2\beta^2\sigma_\mu^2$

so this is biased upwards since we are adding positive terms to the true value

next to compute the probability limit of $\hat{\sigma}_\epsilon^2(\tilde{x}'\tilde{x}/n)^{-1}$

$$\hat{\sigma}_\epsilon^2(\tilde{x}'\tilde{x}/n)^{-1} = \frac{\hat{\sigma}_\epsilon^2}{\hat{\sigma}_{\tilde{x}}^2} \rightarrow_p \frac{\sigma_\epsilon^2 + (1-\lambda)^2\beta^2\sigma_x^2 + \lambda^2\beta^2\sigma_\mu^2}{\sigma_x^2+\sigma_\mu^2}$$

$$= \frac{\sigma_x^2}{\sigma_x^2+\sigma_\mu^2}\left(\frac{\sigma_\epsilon^2}{\sigma_x^2}\right) + \frac{\sigma_x^2}{\sigma_x^2+\sigma_\mu^2}(1-\lambda)^2\beta^2 + \frac{\sigma_\mu^2}{\sigma_x^2+\sigma_\mu^2}\lambda^2\beta^2 = \lambda\left(\frac{\sigma_\epsilon^2}{\sigma_x^2}\right) + \lambda(1-\lambda)^2\beta^2 + (1-\lambda)\lambda^2\beta^2$$

now note that $\lambda(1-\lambda)^2\beta^2 + (1-\lambda)\lambda^2\beta^2 = \beta^2\lambda(1-\lambda)[(1-\lambda)+\lambda] = \beta^2\lambda(1-\lambda)$
Combining these gives us that

$$\frac{\hat{\sigma}_\epsilon^2}{\hat{\sigma}_{\tilde{x}}^2} \rightarrow_p \frac{\lambda\sigma_\epsilon^2}{\sigma_x^2} + \lambda(1-\lambda)\beta^2$$

multiplying the first term by λ biases the result downwards but the second term is positive so it biases the result upwards. So the overall result of the bias cannot be signed in general

1.3 t-test

$$\frac{\hat{\beta}_{ls}}{\sqrt{\hat{\sigma}_\epsilon^2(\tilde{x}'\tilde{x}/n)^{-1}}} \rightarrow_p \frac{\lambda\beta}{\sqrt{\lambda\frac{\sigma_\epsilon^2}{\sigma_x^2} + \lambda(1-\lambda)\beta^2}} = \frac{\sqrt{\lambda}\beta}{\sqrt{\frac{\sigma_\epsilon^2}{\sigma_x^2} + (1-\lambda)\beta^2}}$$

which is smaller than

$$\frac{\beta}{\sqrt{\frac{\sigma_\epsilon^2}{\sigma_x^2}}}$$

So the t-test is downward biased

1.4 Second measurement, Consistency

$$y = x\beta + \epsilon$$

by assumption $\mathrm{E}[\breve{x}\epsilon] = 0$
Now multiply y by $\breve{x}'$ and take the expectation to get $\mathrm{E}[\breve{x}'y] = \mathrm{E}[\breve{x}'x]\beta$
Now assuming $\mathrm{E}[\breve{x}'x]$ is full rank we get $\beta = (\mathrm{E}[\breve{x}'x])^{-1}\mathrm{E}[\breve{x}'y]$
So $\hat{\beta}_{IV} = (\breve{x}'x)^{-1}\breve{x}'y$
Now to show it is consistent
$\hat{\beta}_{IV} = (\breve{x}'x)^{-1}\breve{x}'(x\beta + \epsilon) = \beta + (\frac{\breve{x}'x}{n})^{-1}(\frac{\breve{x}'\epsilon}{n}) \to_p \beta$
since $\mathrm{E}[\breve{x}'\epsilon] = 0$ so $\frac{\breve{x}'\epsilon}{n} \to_p 0$ by LLN

1.5 Second measurement, Distribution

$$\sqrt{n}(\hat{\beta}_{IV} - \beta) = (\breve{x}'x)^{-1}\breve{x}'\epsilon = \sqrt{n}\left(\frac{\breve{x}'x}{n}\right)^{-1}\left(\frac{\breve{x}'\epsilon}{n}\right)$$

Now using the CLT we get

$$\sqrt{n}\left(\frac{\breve{x}'\epsilon}{n}\right) \xrightarrow{d} N(0, \mathrm{E}[\breve{x}'\epsilon'\epsilon\breve{x}])$$

Now all together we get

$$\sqrt{n}(\hat{\beta}_{IV} - \beta) \xrightarrow{d} N(0, \mathrm{E}[\breve{x}'x]^{-1}\mathrm{E}[\breve{x}'\epsilon'\epsilon\breve{x}]E[x\breve{x}']^{-1})$$

1.6 Second measurement, Inference

To create a confidence interval robust to Standard errors we want to use the following, unsimplified, version of the asymptotic variance estimator.

$$\hat{V}_{IV} = Avar(\hat{\beta}_{IV}) = (\breve{x}'x)^{-1}\left(\sum_{i=1}^{n}\epsilon_i^2\breve{x}_i'\breve{x}_i\right)(\breve{x}'x)^{-1}$$

We also showed above that

$$\sqrt{n}(\frac{\hat{\beta}_{IV}}{\sqrt{\hat{V}_{IV}}}) \to_d \mathcal{N}(\beta, 1)$$

Inverting the standard normal distribution and the following confidence interval

$$\left[\hat{\beta}_{IV} - \Phi^{-1}\left(1 - \frac{(1-\alpha)}{2}\right)\left(\sqrt{\frac{\hat{V}_{IV}}{n}}\right), \hat{\beta}_{IV} + \Phi^{-1}\left(1 - \frac{(1-\alpha)}{2}\right)\left(\sqrt{\frac{\hat{V}_{IV}}{n}}\right)\right]$$

where $\alpha = 0.95$ in this case

1.7 Validation sample, Consistency

First note that $(\frac{1}{n}\tilde{x}'\tilde{x}) \to_p \sigma_x^2 + \sigma_u^2$ and as shown in part 1 $\hat{\beta}_{ls} \to_p \beta\frac{\sigma_x^2}{\sigma_x^2+\sigma_\mu^2}$

Now we define $\hat{\beta}_{VS} = \hat{\beta}_{ls}\left(\frac{1}{n}\frac{\tilde{x}'\tilde{x}}{\breve{\sigma}_x^2}\right)$

and by Slutsky's theorem we get that $\hat{\beta}_{VS} \to_p \beta$

1.8 Validation sample, Distribution

We know from section 1.7 that $\hat{\beta}_{VS} = \hat{\beta}_{ls}\left(\frac{1}{n}\frac{\tilde{x}'\tilde{x}}{\check{\sigma}_x^2}\right)$

We can break this into three pieces and define $\hat{\beta}_{VS}$ in the following way

$$\hat{\beta}_{VS} = g(a,b,c) = \frac{ab}{c}$$
$$a = \hat{\beta}_{ls}$$
$$b = \frac{1}{n}\tilde{x}'\tilde{x}$$
$$c = \check{\sigma}_x^2$$

g is a continuous function so we can apply the delta method.

$$\sqrt{n}\left(g\left(\hat{\beta}_{ls}, \frac{1}{n}\tilde{x}'\tilde{x}, \check{\sigma}_x^2\right) - g\left(\lambda\beta, \sigma_x^2 + \sigma_\mu^2, \sigma_x^2\right)\right) \to_d \mathcal{N}\left(\nabla g\left(\lambda\beta, \sigma_x^2 + \sigma_\mu^2, \sigma_x^2\right)' \mathbf{\Sigma} \nabla g\left(\lambda\beta, \sigma_x^2 + \sigma_\mu^2, \sigma_x^2\right)\right)$$

$$V_{vs} = \nabla g\left(\lambda\beta, \sigma_x^2 + \sigma_\mu^2, \sigma_x^2\right)' \mathbf{\Sigma} \nabla g\left(\lambda\beta, \sigma_x^2 + \sigma_\mu^2, \sigma_x^2\right)$$

1.9 Validation sample, Inference

Similar to problem 1.6 we have that

$$\sqrt{n}\left(\frac{\hat{\beta}_{VS}}{\sqrt{\hat{V}_{VS}}}\right) \to_d \mathcal{N}(\beta, 1)$$

Inverting the standard normal distribution and the following confidence interval

$$\left[\hat{\beta}_{VS} - \Phi^{-1}\left(1 - \frac{(1-\alpha)}{2}\right)\left(\sqrt{\frac{\hat{V}_{VS}}{n}}\right), \hat{\beta}_{VS} + \Phi^{-1}\left(1 - \frac{(1-\alpha)}{2}\right)\left(\sqrt{\frac{\hat{V}_{VS}}{n}}\right)\right]$$

where $\alpha = 0.95$ in this case

1.10 FE estimator, Consistency

First note that because we have $T = 2$, the FE estimator is equivalent to the first-difference (FD) estimator. That is

$$\hat{\beta}_{FE} = \hat{\beta}_{FD}$$
$$\left(\frac{1}{n}\sum_{i-1}^{n}(\tilde{x}_{i2} - \tilde{x}_{i1})^2\right)^{-1}\left(\frac{1}{n}\sum_{i-1}^{n}(\tilde{x}_{i2} - \tilde{x}_{i1})(y_{i2} - y_{i1})\right)$$

Not by using the WLLN:

$$\frac{1}{n}\sum_{i-1}^{n}(\tilde{x}_{i2} - \tilde{x}_{i1})^2 \to_p \mathrm{E}[(\tilde{x}_{i2} - \tilde{x}_{i1})^2] = \mathrm{E}[(x_{i2} - x_{i1} + u_{i2} - u_{i1})^2]$$

$$= \mathrm{E}[(x_{i2} - x_{i1})^2] + \mathrm{E}[(u_{i2} - u_{i1})^2] + 2\mathrm{E}[(x_{i2} - x_{i1})(u_{i2} - u_{i1})] = \sigma^2_{\Delta x} + \sigma^2_{\Delta u}$$

since $\mathrm{E}[x_{it}u_{it}] = 0 \ \forall \ t, s \in \{1, 2\}$ Next

$$\frac{1}{n}\sum_{i-1}^{n}(\tilde{x}_{i2} - \tilde{x}_{i1})(y_{i2} - y_{i1}) \to_p \mathrm{E}[(\tilde{x}_{i2} - \tilde{x}_{i1})(y_{i2} - y_{i1})]$$

$$= \mathrm{E}[(x_{i2} - x_{i1} + u_{i2} - u_{i1})(x_{i2}\beta - x_{i1}\beta + e_{i2} - e_{i1})]$$

$$= \mathrm{E}[(x_{i2} - x_{i1})^2]\beta + \mathrm{E}[(x_{i2} - x_{i1})(e_{i2} - e_{i1})] + \mathrm{E}[(x_{i2} - x_{i1})(u_{i2} - u_{i1})]\beta + \mathrm{E}[(u_{i2} - u_{i1}(e_{i2} - e_{i1}))]$$

$$= \mathrm{E}[(x_{i2} - x_{i1})^2]\beta = \sigma_{\Delta x}\beta$$

since $\mathrm{E}[x_{it}u_{it}] = \mathrm{E}[x_{it}e_{it}] = \mathrm{E}[u_{it}e_{it}] \ \forall t, s \in \{1, 2\}$. Finally we can put these together by the CMT to get.

$$\hat{\beta}_{FE} \to_p \frac{\sigma^2_{\Delta x}}{\sigma^2_{\Delta x} + \sigma^2_{\Delta u}}\beta$$

2 Question 2: Implementing Least-Squares Estimators

2.1 part 1

Start by adding and subtracting $x\tilde{\beta}$ to get

$$(y - x\tilde{\beta} + x\tilde{\beta} - x\beta)'W(y - x\tilde{\beta} + x\tilde{\beta} - x\beta)$$

$$= (y - x\tilde{\beta})'W(y - x\tilde{\beta}) + (y - x\tilde{\beta})'W(x\tilde{\beta} - x\beta) + (x\tilde{\beta} - x\beta)'W(y - x\tilde{\beta}) + (x\tilde{\beta} - x\beta)'W(x\tilde{\beta} - x\beta)$$
$$= (y - x\tilde{\beta})'W(y - x\tilde{\beta}) + 2(x\tilde{\beta} - x\beta)'W(y - x\tilde{\beta}) + (x\tilde{\beta} - x\beta)'W(x\tilde{\beta} - x\beta)$$

Now we need to find $\tilde{\beta}$ to minimize this equation. We want to set the middle term to zero so we need a $\tilde{\beta}$ such that $\tilde{\beta}'x'W(y - x\tilde{\beta}) = \beta'x'W(y - x\tilde{\beta})$

we pick $\tilde{\beta}$ such that $x'W(y - x\tilde{\beta}) = 0$ giving us

$$\tilde{\beta} = (x'W'x)^{-1}(x'Wy)$$

Now when we minimize over β the first term is irrelevent as it does not include a β. The middle term is 0 so it does not matter. The last term is positive semi definite and so it is minimized by setting $\beta = \tilde{\beta}$

2.2 Part 2

$$\sqrt{n}(\hat{\beta}(w) - \beta) = \sqrt{n}((x'Wx)^{-1}x'W(x\beta + \epsilon) - \beta) = \sqrt{n}((x'Wx)^{-1}x'W\epsilon)$$

$$= ((\frac{1}{n}x'Wx)^{-1}\sqrt{n}(\frac{1}{n}x'W\epsilon))$$

under appropriate assumptions we have by LLN that $(\frac{1}{n}x'Wx) \to_p A$
We also have that $\sqrt{n}(\frac{1}{n}x'W\epsilon) \to_d \mathcal{N}(0, B)$ by CLT
In this case we get $B = \frac{1}{n}\mathbb{V}[x'W\epsilon] = \frac{1}{n}\mathbb{E}[x'W\epsilon'\epsilon Wx]$
And we have that $V(W) = A^{-1}BA^{-1}$

2.3 Part 3

To estimate $V(W) = A^{-1}BA^{-1}$ we are mostly just putting hats on things

$$\hat{A} = \frac{1}{n}(x'\hat{W}x)$$

$$\hat{B} = \frac{1}{n}(x'\hat{W}\hat{\epsilon}'\hat{\epsilon}\hat{W}x)$$

so that gives us

$$\hat{V}(W) = \frac{1}{n}(x'\hat{W}x)^{-1}(x'\hat{W}\hat{\epsilon}'\hat{\epsilon}\hat{W}x)(x'\hat{W}x)^{-1}$$

3 Question 3: Analysis of Experiments

3.1 title

$$\begin{aligned}
\mathrm{E}[T_{DM}] = \mathrm{E}[\bar{Y}_1] - \mathrm{E}[\bar{Y}_0] &= \mathrm{E}\left[\frac{1}{N_1}\sum_{i=1}^{n} D_i(1)Y_i\right] - \mathrm{E}\left[\frac{1}{n-N_1}\sum_{i=1}^{n} D_i(0)Y_i\right] \\
&= \frac{1}{N_1}\sum_{i=1}^{n}(D_i(1)\mathrm{E}[Y_i]) - \frac{1}{n-N_1}\sum_{i=1}^{n}(D_i(0)\mathrm{E}[Y_i]) \\
&= \frac{1}{N_1}\sum_{i=1}^{n}(D_i(1))\,\mathrm{E}[Y_i(T_i)|T_i=1] - \frac{1}{n-N_1}\sum_{i=1}^{n}(D_i(0))\,\mathrm{E}[Y_i(T_i)|T_i=0]
\end{aligned}$$

Now note that since T_i is random:

$$\mathrm{E}[Y_i(T_i)|T_i=1] = \mathrm{E}[Y_i(1)]$$

$$\mathrm{E}[Y_i(T_i)|T_i=0] = \mathrm{E}[Y_i(0)]$$

Together this gives us:

$$\mathrm{E}[T_{DM}] = \mathrm{E}[Y_i(1)] - \mathrm{E}[Y_i(0)]$$

or

$$\tau_{ATE} = \frac{1}{n}\sum_{i=1}^{n} Y_i(1) - \frac{1}{n}\sum_{i=1}^{n} Y_i(0)$$