ECON 120C: Review

Pietro Emilio Spini

February 7, 2020

Exercise 1. Here the measurement error concerns the outcome variable.

a) We hyave the following system:

$$\tilde{Y}_i = Y_i + w_i \tag{1}$$
$$Y = \beta_0 + \beta_1 X_i + u_i \tag{2}$$

Substituting (2) into (1) the equation substituting one obtains $\tilde{Y}_i = \beta_0 + \beta_1 X_i + u_i + w_i$ where both u_i and w_i are unobserved. Hence the unobserved term in the regression of $\tilde{Y}$ on X given by equation $\tilde{Y}_i = \beta_0 + \beta_1 X_i + v_i$ must have $v_i = u_i + w_i$.

b) We check 3 assumptions:

i) $\mathbb{E}[v|X] = \mathbb{E}[u + w|X] = \mathbb{E}[u|X] + \mathbb{E}[w|X] = 0 + 0 = 0$
The first equality follows from a). The second from linearity of the conditional expectation function. Finally $w \perp (u, X)$ implies $\mathbb{E}[w|X] = \mathbb{E}[w] = 0$ while $\mathbb{E}[u|X] = 0$ is already assumed about the correctly measured model.

ii) $(\tilde{Y}_i, X_i)$ are i.i.d. To see this, recall that (Y_i, X_i) are i.i.d and (w_i) are i.i.d and $w_i \perp (Y_j, X_j)$ for any i and j. Hence the full block (Y_i, X_i, w_i) is i.i.d. Lastly, by construction $\tilde{Y}_i = Y_i + w_i$, is a function solely of the ith observations of Y and w. Hence $(x_i, \tilde{Y}_i)$ is also i.i.d.

iii) $\mathbb{E}[\tilde{Y}]^4 = \mathbb{E}[(Y + w)^4] = \mathbb{E}[Y^4] + 4\mathbb{E}[Y^3 w] + 6\mathbb{E}[Y^2 w^2] + 4\mathbb{E}[Y w^3] + \mathbb{E}[w^4]$. Now each of these terms is finite since $\mathbb{E}[Y^4] < \infty$ and $\mathbb{E}[w^4] < \infty$. Hence $\mathbb{E}[\tilde{Y}^4]$ is finite.

c) Yes OLS estimates of the regression $\tilde{Y}$ on a constant and X are consistent since, by part b), they satisfy the classical regression hypothesis. It can be shown directly by noting $\beta_1 \xrightarrow{p} \frac{Cov(\tilde{Y},X)}{Var(X)}$ and decompose the covariance term. Exercise 2 of this set of notes provides a blueprint on how to prove the details.

d) We can characterize the variance of the $\hat{\beta}_1$ estimator. We have:

$$\sqrt{n}(\hat{\beta}_1 - \beta_1) = \sqrt{n}\left(\frac{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2} - \beta_1\right)$$

Using the usual decomposition we can separately study the limiting behavior of numerator and denominator. By LLN we have $\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 \xrightarrow{p} \sigma_X^2$. By the CLT (under homoskedasticity) we have:

$$\frac{\sqrt{n}}{n}\sum_{i=1}^{n}(X_i - \hat{X})(u_i + w_i) \xrightarrow{d} \mathcal{N}(0, \sigma_X^2(\sigma_U^2 + \sigma_W^2)) \tag{3}$$

hence we have, by Slutsky theorem:

$$\sqrt{n}(\hat{\beta}_1 - \beta_1) \xrightarrow{d} N\left(0, \frac{\sigma_U^2 + \sigma_W^2}{\sigma_X^2}\right)$$

Because the asymptotic distribution of the $\hat{\beta}$ estimator has a larger variance, confidence intervals constructed with the asymptotic distribution will be larger. The intuition for this is the fact that now, in addition to all other unobserved components u that are featured in the correctly measured model we have to take into account the additional noise coming from w.

e) Measurement error in X is a serious problem for both estimation and inference (i.e. hypothesis testing). On the other hand measurement error on Y is not problematic if all we want to do is obtain point estimates. For inference, the answer to d) says that estimates may be more noisy and tests may be under-powered since the tests statistic constructed with the asymptotic distribution is smaller that it would be without measurement error.

Exercise 2. The OLS estimator for β_1 computed from the regression for Y on $\hat{X}$ and a constant can be written as:

$$\hat{\beta}_1 = \frac{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \bar{Y})(\tilde{X}_i - \bar{\tilde{X}})}{\frac{1}{n}\sum_{i=1}^{n}(\tilde{X}_i - \bar{\tilde{X}})^2} \tag{4}$$

Hence from the Law of Large numbers and the continuous mapping theorem we have:

$$\hat{\beta}_1 \xrightarrow{p} \frac{Cov(Y, \tilde{X})}{Var(\tilde{X})}$$

For the numerator we have:

$$\begin{aligned}
Cov(Y, \tilde{X}) &= Cov(\beta_0 + \beta_1 X + u, \tilde{X}) \\
&= Cov(\beta_1 X + u, X + w) \\
&= \beta Cov(X, X) + \beta_1 Cov(X, w) + Cov(u, X) + Cov(u, w) \\
&= \beta_1 Var(X) \\
&= \beta_1 \sigma_X^2
\end{aligned}$$

For the denominator we have:

$$\begin{aligned}
Var(\tilde{X}) &= Var(X + w) \\
&= Var(X) + Var(w) + 2Cov(X, w) \\
&= \sigma_X^2 + \sigma_W^2
\end{aligned}$$

Hence putting them together we obtain:

$$\beta_1 \xrightarrow{p} \frac{\sigma_X^2}{\sigma_X^2 + \sigma_W^2} \beta_1 \tag{5}$$