Problem Set I

Econometrics III

Nurfatima Jandarova

February 5, 2017

Exercise 1

(a) The LS estimate of α_3:

$$\hat{\alpha}_3 = \frac{\sum\limits_{i=1}^{n} y_1 y_2}{\sum\limits_{i=1}^{n} y_1^2} = \frac{5}{110} = \frac{1}{22}$$

The 2SLS estimate of α_3:

First, estimate

$$y_1 = \gamma x + \varepsilon = \frac{\alpha_2}{1-\alpha_1\alpha_3}x + \frac{u_1 + \alpha_1 u_2}{1-\alpha_1\alpha_3}$$

$$\hat{\gamma} = \frac{\sum\limits_{i=1}^{n} x_i y_{1i}}{\sum\limits_{i=1}^{n} x_i^2} = \frac{120}{360} = \frac{1}{3}$$

Now, regress y_2 on the fitted value of y_1

$$\hat{\alpha}_3 = \frac{\sum\limits_{i=1}^{n} \hat{\gamma} x_i y_{2i}}{\sum\limits_{i=1}^{n} (\hat{\gamma} x_i)^2} = \frac{1}{\hat{\gamma}} \frac{\sum\limits_{i=1}^{n} x_i y_{2i}}{\sum\limits_{i=1}^{n} x_i^2} = 3\frac{120}{360} = 1$$

(b) No. If one tries to apply 2SLS in the other direction, i.e., first estimate $\hat{\beta}$ from $y_2 = \underbrace{\frac{\alpha_3\alpha_2}{1-\alpha_1\alpha_3}}_{\beta} x + \frac{\alpha_3 u_1 + u_2}{1-\alpha_1\alpha_3}$ and then regress y_1 on $\hat{y}_2$ and x, there is an issue of multicollinearity since $\hat{y}_2 = \hat{\beta}x$. If, on the other hand, one estimates the following equations

$$y_1 = \frac{\alpha_2}{1-\alpha_1\alpha_3}x + \frac{u_1 + \alpha_1 u_2}{1-\alpha_1\alpha_3}$$
$$y_2 = \frac{\alpha_2\alpha_3}{1-\alpha_1\alpha_3}x + \frac{u_2 + \alpha_3 u_1}{1-\alpha_1\alpha_3}$$

the estimated coefficients in front of x could only help to recover α_3, but not α_1 or α_2.

(c) We should use the estimate of α_3 obtained by 2SLS as it removes the endogeneity issue. So, using the result above, where we found $\hat{\alpha}_3 = 1$, the predicted value is $\hat{y}_2 = \hat{\alpha}_3 55 = 55$.

Exercise 2

The presence of measurement errors creates attenuation bias as is evidenced from Table 1; and the larger the measurement error, the larger is the magnitude of the bias. However, running an IV regression of y on x_3, which was instrumented by x_1 and x_2, vastly improved the estimate of

Regressors		OLS		IV	2SLS (manual)
	(1)	(2)	(3)	(4)	(5)
x_1	0.4342				
	(0.0240)				
x_2		0.2206			
		(0.0167)			
x_3			0.1496	1.0095	1.0095
			(0.0136)	(0.0613)	(0.0511)
Constant	0.2756	0.3826	0.4192	0.0001	0.0001
	(0.0158)	(0.0133)	(0.0123)	(0.0323)	(0.0269)

Note: Dependent variable is y. Standard errors reported in parentheses. In an IV regression x_3 was instrumented by x_1 and x_2.

Table 1: Regression results

the coefficient, bringing it much closer to the true value. The p-value of the Wald test that an IV coefficient is equal to 1 was 0.8771. Hence, we fail to reject the null hypothesis and may conclude that an IV coefficient is indeed equal to the true coefficient.
Performing the two-stage least squares manually (5th column of Table 1) results in the same point estimate; however, with lower standard errors as anticipated. Regressing y on $\hat{x}_3$ is not the same as regressing y on x_3 because by fitting the value of x_3 we are throwing away part of the variation in the regressor. Hence, the correct standard deviation should be adjusted upwards as in 4th column (done automatically by the command *ivregress*).
It's better to use both x_2 and x_3 because their combination allows to eliminate the attenuation bias in the asymptotics, i.e., yields consistent estimator for β. This is due to the fact that both variables only share common information on the true value of x, while the measurement errors are independent of each other. Whereas using x_1 alone, one gets inconsistent estimator, even though, the associated measurement error is the smallest of the three.
Consider first the regression of y on x_1 alone. The resulting estimator is inconsistent:

$$y_i = \phi x_{1i} + \nu_i$$

$$\hat{\phi} = \frac{\frac{1}{n}\sum_{i=1}^{n} x_{1i} y_i}{\frac{1}{n}\sum_{i=1}^{n} x_{1i}^2} \xrightarrow{p} \frac{\mathbb{E}(x_i y_i)}{\mathbb{E}(x_i^2) + \mathbb{E}(v_{1i}^2)} \neq \beta$$

However, if we use 2SLS using x_2 as an instrument for x_3 (or the other way around), we obtain a consistent estimator for β, as is shown below.

First-stage regression

$$x_{3i} = \gamma x_{2i} + \varepsilon_{1i}$$

$$\hat{\gamma} = \frac{\sum_{i=1}^{n} x_{2i} x_{3i}}{\sum_{i=1}^{n} x_{2i}^2} = \frac{\sum_{i=1}^{n} (x_i + v_{2i})(x_i + v_{3i})}{\sum_{i=1}^{n} (x_i + v_{2i})^2} \xrightarrow{p} \frac{\mathbb{E}(x_i^2)}{\mathbb{E}(x_i^2) + \mathbb{E}(v_{2i}^2)}$$

Second-stage regression

$$y_i = \alpha \hat{x}_{3i} + \varepsilon_{2i} = \alpha \hat{\gamma} x_{2i} + \varepsilon_{2i}$$

$$\hat{\alpha} = \frac{\sum_{i=1}^{n} \hat{x}_{3i} y_i}{\sum_{i=1}^{n} \hat{x}_{3i}^2} = \frac{\sum_{i=1}^{n} \hat{\gamma} x_{2i} y_i}{\sum_{i=1}^{n} (\hat{\gamma} x_{2i})^2} = \frac{1}{\hat{\gamma}} \frac{\sum_{i=1}^{n} x_{2i} y_i}{\sum_{i=1}^{n} x_{2i}^2} \xrightarrow{p} \frac{\cancel{\mathbb{E}(x_i^2) + \mathbb{E}(v_{2i}^2)}}{\mathbb{E}(x_i^2)} \frac{\mathbb{E}(x_i y_i)}{\cancel{\mathbb{E}(x_i^2) + \mathbb{E}(v_{2i}^2)}} = \frac{\mathbb{E}(x_i y_i)}{\mathbb{E}(x_i^2)} = \beta$$

However, if the two measurement errors, v_{2i} and v_{3i} were correlated, then the optimal choice depends on the magnitude of correlation relative to the magnitude of the first measurement error. Probability limits of first- and second-stage estimators under the assumption of correlated measurement errors are illustrated below.

$$\hat{\gamma} = \frac{\sum_{i=1}^{n}(x_i + v_{2i})(x_i + v_{3i})}{\sum_{i=1}^{n}(x_i + v_{2i})^2} \xrightarrow{p} \frac{\mathbb{E}(x_i^2) + \mathbb{E}(v_{2i}v_{3i})}{\mathbb{E}(x_i^2) + \mathbb{E}(v_{2i}^2)}$$

$$\hat{\alpha} = \frac{1}{\hat{\gamma}} \frac{\sum_{i=1}^{n}(x_i + v_{2i})y_i}{\sum_{i=1}^{n}(x_i + v_{2i})^2} \xrightarrow{p} \frac{\cancel{\mathbb{E}(x_i^2) + \mathbb{E}(v_{2i}^2)}}{\mathbb{E}(x_i^2) + \mathbb{E}(v_{2i}v_{3i})} \frac{\mathbb{E}(x_i y_i)}{\cancel{\mathbb{E}(x_i^2) + \mathbb{E}(v_{2i}^2)}} = \frac{\mathbb{E}(x_i y_i)}{\mathbb{E}(x_i^2) + \mathbb{E}(v_{2i}v_{3i})}$$

Hence, whenever $\mathbb{E}(v_{2i}v_{3i}) \leq \mathbb{E}(v_{1i}^2)$, the combination of the two measurement errors would still be preferable than to just using x_{1i} alone. On the other hand, if the errors v_{2i} and v_{3i} are highly correlated, using x_{1i} alone would be better.