Exercises Week 4

Econometrics

1. **Exercise 8.2 in ETM**: Consider the simple IV estimator computed first with an $n \times K$ matrix W of instrumental variables, and then with another $n \times K$ matrix WJ, where J is a $K \times K$ nonsingular matrix. Show that the two estimators coincide. Why does this fact show that $\hat{\beta}_{IV}$ depends on W only through the orthogonal projection matrix P_W?

2. **Exercise 8.3 in ETM**: Show that, if the matrix of instrumental variables W is $n \times K$, with the same dimensions as the matrix X of explanatory variables, then the generalized IV estimator is identical to the simple IV estimator.

3. **Exercise 8.8 in ETM**: Suppose that W_1 and W_2 are, respectively, $n \times l_1$ and $n \times l_2$ matrices of instruments, and that W_2 consists of W_1 plus $l_2 - l_1$ additional columns. Prove that the generalized IV estimator using W_2 is asymptotically more efficient than the generalized IV estimator using W_1. To do this, you need to show that the matrix

$$(X^T P_{W_1} X)^{-1} - (X^T P_{W_2} X)^{-1}$$

   is positive semidefinite.

Hint: See Excerise 3 on the Exercise Set for Week 2.

4. **Measurement errors in OLS**: In this exercise you are going to show the effect that measurement errors have on the OLS estimator.

Do the following:

i. Generate a sample of size 50 from the model

$$y_t = \beta_1 + \beta_2 x_t + u_t,$$

   with $\beta_1 = 1$ and $\beta_2 = 0.8$.

For simplicity, assume that x_t are $NID(2, 2)$ and that the u_t are $NID(0, 1)$.

ii. Now generate noisy versions of the regressors and regressand $y_t^* = y_t + u_{y,t}$, $x^* = x_t + u_{x,t}$, with $u_{y,t} \sim N(0, \sigma_y^2)$, $u_{x,t} \sim N(0, \sigma_x^2)$ for some values $\sigma_y^2$and σ_x^2 of your choosing.

iii. Estimate by OLS the equations

$$y_t^* = \beta_1 + \beta_2^y x_t + u_t,$$

   and

$$y_t = \beta_1 + \beta_2^x x_t^* + u_t,$$

   and compare the value of the estimator for β_2 and its standard deviation between equations.

iv. Repeat at least 100 times and find the averages and variance of $\hat{\beta}_2^x$ and $\hat{\beta}_2^y$. Use these averages to estimate the bias of the OLS estimators of β_2. What happens to the bias and variance if you increase σ_y^2 and/or σ_x^2? Explain.

5. **Exercise 8.28 in ETM**: The file *demand-supply.data* (available here) contains 120 artificial observations on a demand-supply model. The demand equation is

$$q_t = \beta_1 + \beta_2 X_{t2} + \beta_3 X_{t3} + \gamma p_t + u_t,$$

where q_t is the log of quantity, p_t is the log of price, X_{t2} is the log of income, and X_{t3} is a dummy variable that accounts for regular demand shifts.

Estimate this equation by OLS and 2SLS, using the variables X_{t4} and X_{t5} as additional instruments.

Does OLS estimation appear to be valid here?

Does 2SLS estimation appear to be valid here? Perform whatever tests are appropriate to answer these questions.