ARE 212 - Problem Set 5

Due May 2^{nd}

Part I: Theory (Optional)

1. Show that the parameter estimates for $\mathbf{b}_{2SLS}$ and $\mathbf{b}_{IV}$ are equivalent if we have a model with one endogenous variable and one instrumental variable.

2. Prove that for $E[Z'X]$ to be of full column rank, at least one of the θ_j in the linear projection $\mathbf{x}_k = \delta_o + \delta_1 \mathbf{x}_1 + \delta_2 \mathbf{x}_2 + \ldots + \delta_{k-1}\mathbf{x}_{k-1} + \theta_1 \mathbf{z}_1 + \theta_2 \mathbf{z}_2 + \ldots + \theta_M \mathbf{z}_M + \boldsymbol{\eta}$ has to be different from zero.

Part II (Applied): Instrumental Variables

1. Let's revisit the model and data from problem set 4. We would like to estimate the model:

$$\log(wage) = \beta_o + exper \cdot \beta_1 + tenure \cdot \beta_2 + married \cdot \beta_3 + south \cdot \beta_4 + urban \cdot \beta_5 + black \cdot \beta_6 + educ \cdot \beta_7 + abil \cdot \gamma + \varepsilon \quad (1)$$

   One of the big problems in the labor literature is that we do not (as econometricians) observe ability. If ability is not correlated with any of the right hand side variables, we can include it in the disturbance and nothing is lost by not observing it. If, however, it is correlated with one or more of the right hand side variables, OLS is no longer unbiased or consistent. Assume that ability is correlated with education and none of the other right hand side variables.

   (a) Derive the bias of β_7 and show what direction the bias goes in depending on whether the correlation between ability and education is positive or negative.

   (b) You showed in the first part that we can derive the sign/direction of the bias. One approach that has been take in the literature is using a "proxy" variable for the unobservable ability. We will use IQ here to proxy for ability. Estimate the model above excluding ability, record your parameter estimates, standard errors and R^2.

   (c) Estimate the model including IQ as a proxy, record your parameter estimates, standard errors and R^2.

   (d) What happens to returns to schooling? Does this result confirm your suspicion of how ability and schooling are expected to be correlated?

2. This problem from Wooldridge asks you to try and recreate some of the results in Card(1995), which is on reserve and on the website. Use the dataset card.raw on the website.

   (a) Read the data in your favorite statistical program. Plot the series make sure your data are read in correctly.

   (b) Estimate a $\log(wage)$ regression via Least Squares with *educ*, *exper*, $exper^2$, *black*, *south*, *smsa*, *reg*661 through *reg*668 and *smsa*66 on the right hand side. Check your results against Table2, column 5.

   (c) Estimate a reduced form equation for *educ* containing all of the explanatory variables and the dummy variable *nearc*4. Is the partial correlation between *nearc*4 and *educ* statistically significant?

   (d) Estimate the $\log(wage)$ equation by instrumental variables, using *nearc*4 as an instrument for *educ*. Compare the 95% confidence interval for the return to educutioan to that obtained from the Least Squares regression above.

   (e) Now use multiple instruments. Use *nearc*2 and *nearc*4 as instruments for *educ*. Comment on the significance of the partial correlations of both instruments in the reduced form. Show your standard errors from the second stage and compare them to the correct standard errors.

   (f) Conduct a Hausman test for endogeneity of educ. Report your test statistic, critical value and p-value.