Econometrics 522A
Spring, 2019
Hidehiko Ichimura, TA: Zejin Shi

PS #9 (Due 3 April, 2019)

1. Consider estimating the linear regression model $y_i = x_i'\beta_0 + u_i$, where x_i is a $K \times 1$ vector, using J instrumental variables, z_i, which is a $J \times 1$ vector, $J \geq K$. Sampling is i.i.d. Consider the GMM estimation of β_0 using

$$\hat{g}(b)' W \hat{g}(b),$$

where $\hat{g}(b) = N^{-1} \sum_{i=1}^{N} g_i(b)$, where $g_i(b) = z_i(y_i - x_i'b)$, as the objective function. The first order condition is

$$0 = \frac{\partial \hat{g}(b)'}{\partial b} W \hat{g}(b) = \hat{G}' W \hat{g}(b),$$

where $\hat{G} = N^{-1} \sum_{i=1}^{N} z_i x_i'$.

   (a) Using the probability limit of $\hat{G}$, denoted G and W, explain which K linear combinations of z_i vector is used as the K IV in the GMM estimation.

2. We claimed that using $W = \Omega^{-1}$, where $\Omega = E[g_i(\beta) g_i'(\beta)]$ yields the optimal GMM estimator. The following sequence of questions meant to demonstrate this.

   (a) Show that the asymptotic variance-covariance matrix of the GMM estimator using general weight matrix W is

$$(G'WG)^{-1} G'W\Omega WG(G'WG)^{-1},$$

   where $G = E(z_i x_i')$.

   (b) Show that the variance-covariance matrix of the optimal GMM estimator is

$$(G'\Omega^{-1}G)^{-1}.$$

   (c) Show that

$$\begin{aligned}
&(G'WG)^{-1}G'W\Omega WG(G'WG)^{-1} - (G'\Omega^{-1}G)^{-1} \\
&\quad = (G'WG)^{-1}[G'W\Omega WG - (G'WG)(G'\Omega^{-1}G)^{-1}(G'WG)](G'WG)^{-1} \\
&\quad\quad = (G'WG)^{-1}G'W[\Omega - G(G'\Omega^{-1}G)^{-1}G')]WG(G'WG)^{-1}.
\end{aligned}$$

   (d) Let T be such that $\Omega = T'T$. Recall that $\Omega^{-1} = T^{-1}(T')^{-1}$ and that $(T')^{-1} = (T^{-1})'$. Using these results from earlier problem sets and denoting Let $C = (T^{-1})'G$, show that the last expression equals

$$\begin{aligned}
&(G'WG)^{-1}G'WT'[I_J - (T')^{-1}G(G'T^{-1}(T')^{-1}G)^{-1}G'T^{-1}]TWG(G'WG)^{-1} \\
&\quad = (G'WG)^{-1}G'WT'[I_J - C(C'C)^{-1}C']TWG(G'WG)^{-1}.
\end{aligned}$$

   (e) Show that the last expression is positive semi-definite.

3. Consider the following simultaneous equations model

$$\begin{aligned}
y_i^S &= \beta_1^S + \beta_2^S p_i + \beta_3^S x_i + u_i^S \\
y_i^D &= \beta_1^D + \beta_2^D p_i + \beta_3^D z_i + u_i^D
\end{aligned}$$

where p_i is determined by $y_i^S = y_i^D$ and that we observe, for each i, the equilibrium quantity y_i that satisfies the equality. The first equation is meant to be a supply function and the second equation is meant to be a demand function. In this model we assume that (u_i^S, u_i^D) given (x_i, z_i) has conditional mean zero.

(a) Write down the reduced form equations for p_i and y_i.

(b) What is the order condition for the supply parameters?

(c) What is the rank condition for the demand parameters?

(d) Discuss how you will estimate the demand parameters efficiently by GMM under heteroskedasticity.

4. Use the data set at `http://economics.mit.edu/faculty/angrist/data1/data/anglavy99` to

   (a) reproduce Angrist and Lavy full sample and discontinuity sample in Table III using R.

   (b) Construct a 95% confidence interval for the class-size effect for both samples using heterskedasticity robust standard error.