ECON 710 QUARTER 1: PROBLEM SET 2

EMILY CASE

Exercise 1

Suppose $(Y, X, Z)'$ is a vector of random variables such that

$$Y = \beta_0 + X\beta_1 + U, \qquad \mathbb{E}[U|Z] = 2$$

where $Cov(Z, X) \neq 0$ and $\mathbb{E}[Y^2 + X^2 + Z^2] < \infty$. Additionally, let $\{(Y_i, X_i, Z_i)'\}_{i=1}^n$ be a random sample from the model with $\widehat{Cov}(Z, X) \neq 0$.

$$\hat{\beta}_1^{IV} = \frac{\widehat{Cov}(Z,Y)}{\widehat{Cov}(Z,X)} = \frac{\frac{1}{n}\sum_{i=1}^{n}(Z_i - \bar{Z}_n)(Y_i - \bar{Y}_n)}{\frac{1}{n}\sum_{i=1}^{n}(Z_i - \bar{Z}_n)(X_i - \bar{X}_n)}$$

$$\hat{\beta}_0^{IV} = \bar{Y} - \bar{X}\hat{\beta}_1^{IV}$$

(i) *Does $\hat{\beta}_1^{IV} \to^p \beta_1$?*
Note that by LLN and CMT,

$$\widehat{Cov}(Z,Y) = \frac{1}{n}\sum_{i=1}^{n}(Z_i - \bar{Z}_n)(Y_i - \bar{Y}_n) \to^p \mathbb{E}[(Z_i - \bar{Z}_n)(Y_i - \bar{Y}_n) = Cov(Z,Y)$$

and by similar logic $\widehat{Cov}(Z, X) \to^p Cov(Z, X)$. Then, by CMT,

$$\begin{aligned}
\hat{\beta}_1^{IV} &\to^p \frac{Cov(Z,Y)}{Cov(Z,X)} \\
&= \frac{Cov(Z, \beta_0 + X\beta_1 + U)}{Cov(Z,X)} \\
&= \frac{\beta_1 Cov(Z,X) + Cov(Z,U)}{Cov(Z,X)} \\
&= \beta_1
\end{aligned}$$

I worked on this Problem set with Sarah Bass, Michael Nattinger, Alex von Hafften, and Danny Edgel.

Since $Cov(Z,U) = 0$. [1]

(ii) *Does* $\hat{\beta}_0^{IV} \to^p \beta_0$*?*

$$\begin{aligned}\hat{\beta}_0^{IV} &= \bar{Y} - \bar{X}\hat{\beta}_1^{IV} \\ &\to^2 \bar{Y} - \bar{X}\beta_1\end{aligned}$$

EXERCISE 2

Consider the simultaneous model

$$\begin{aligned} Y &= \beta_0 + X\beta_1 + U, & \mathbb{E}[U|Z] = 0 \\ X &= \pi_0 + Z\pi_1 + V, & \mathbb{E}[V|Z] = 0 \end{aligned}$$

where $\mathbb{E}[Y^2 + X^2 + Z^2] < \infty$. Additionally, let $\{(Y_i, X_i, Z_i)'\}_{i=1}^n$ be a random sample from the model with $\frac{1}{n}\sum_{i=1}^{n}(Z_i - \bar{Z}_n)^2 > 0$.

(i) *Under what conditions (on β_0, β_1, π_0, π_1) is Z a valid instrument for X?*
(ii) *Show that*

$$Y = \gamma_0 + Z\gamma_1 + \epsilon, \qquad \mathbb{E}[\epsilon|Z] = 0$$

where γ_0, γ_1 and ϵ are some functions of β_0, β_1, π_0, π_1, U, and V. In particular show that $\gamma_1 = \pi_1\beta_1$.

EXERCISE 3

The paper "Children and Their Parents' Labor Supply: Evidence fromExogenous Variation in Family Size" by J. Angrist and W. Evans (AE98) considerslabor supply responses to the number of children in the household.
They consider models of the form

$$Y = \beta_0 + X_1\beta_1 + X_2'\beta_2 + U$$

where Y is some measure of the parents' labor supply, X_1 is a binary variable indicating "more than 2 children in the household", and X_2 is a vector of (assumed) exogenous variables that control for race, age, and whether any of the children is a

[1] **check this**

$$\begin{aligned} Cov(Z,U) &= \mathbb{E}[(Z - \mathbb{E}[Z])(U - \mathbb{E}[U])] \\ &= \mathbb{E}[ZU] - \mathbb{E}[\mathbb{E}[Z]U] - \mathbb{E}[Z\mathbb{E}[U]] + \mathbb{E}[\mathbb{E}[Z]\mathbb{E}[U]] \\ &= \mathbb{E}[ZU] - \mathbb{E}[Z]\mathbb{E}[U] \\ &= \mathbb{E}[Z\mathbb{E}[U|Z]] - E[Z]\mathbb{E}[\mathbb{E}[U|Z]] \\ &= 2\mathbb{E}[Z] - 2E[Z] \end{aligned}$$

boy. For the next two questions we will focus on the case where Y is a binary variable indicating whether the mother worked during the year.

(i) *Provide a causal interpretation of* β_1.

(ii) *Discuss why or why not you think that* X_1 *could be endogenous. If you think it is, discuss the direction of the (conditional) bias in OLS relative to the causal parameter.*

(iii) *Repeat the previous two questions when* Y *is a binary variable indicating whether the husband worked during the year.*

(iv) *Discuss why or why not you think that the binary variable* Z_1 *which indicates whether the two first children are of the same sex is a valid instrument for* X_1.

(v) *Estimate the reduced form regression of* X_1 *on* Z_1 *and* X_2*, do the results suggest that* Z_1 *is relevant?*

(vi) *(Attempt to) replicate the first three rows of Table 7 columns 1, 2, 5, 7, and 8 in AE98. Interpret the empirical results in relation to your discussion of the previous questions.*