ECMA 31000: Problem Set 7

Joe Hardwick

Question 1 (You must include exogenous regressors in the first stage) Let $x \in \mathbb{R}^{k+1}$be a random vector with first component $x_0 = 1$. Suppose you observe an iid sample of $\{y_i, x_i, z_i\}_{i=1}^n$. Consider the model

$$y = x'\beta + u; \qquad \mathrm{E}(ux) \neq 0, \mathrm{E}(zu) = 0.$$

Assume $\mathrm{E}(zx') < \infty$ and is invertible. Suppose there is a single endogenous regressor x_k.

a) Suppose you estimate this model using Two Stage Least Squares. Describe the two stage procedure. Show that the estimator is equal to the Instrumental Variables estimator. Is this a consistent estimate of β?

b) Now suppose you have more instruments than regressors. Show that the 2SLS estimator can be considered as an IV estimator after running the first stage stage regression. What are the instruments in this case? Is the 2SLS estimator consistent given the modelling assumptions?

c(i) Return to the exactly identified case. Let $x_{-k} = (x_0, \ldots, x_{k-1})$ and $\gamma_{-k} = (\gamma_0, \ldots, \gamma_{-k})$. Show that the coefficient on x_k in the second stage regression can be found by first running OLS on

$$y = x'_{-k}\gamma_{-k} + \gamma_k z + \epsilon, \tag{1}$$

to yield the OLS estimate $\hat{\gamma}_k$, and then dividing this estimate by $\hat{\pi}$, where $\hat{\pi}$ is the OLS coefficient on z in the regression

$$x_k = x'_{-k}\delta + \pi z + v.$$

Hint: When considering the second stage regression, substitute $\hat{x}_k = x'_{-k}\hat{\delta} + \hat{\pi}z$. Now compare the OLS minimization problems of the second stage regression and of model (1).

c(ii) Is $\hat{\gamma}_k$ a consistent estimate of β_k?

d) Suppose $k = 2$, the included instruments are $(1, x_1)$ and there is a single endogenous variable x_2. There is a single excluded instrument z_1, so the vector of instrumental variables is $z = (1, x_1, z_1)$. Explain, first mathematically, and then intuitively, why the following is incorrect and actually yields an inconsistent estimate of β when using 2 stage least squares:

"In the first stage regression, it's fine to omit exogenous regressors. Just regress the endogenous variables on the excluded instruments since this produces exogenous variation anyway." (The term "exogenous variation" means x_k is exogenous after projection because it is a linear combination of excluded $z's$).

Question 2 Let $x \in \mathbb{R}^{k+1}$be a random vector with first component $x_0 = 1$. Suppose you observe an iid sample of $\{y_i, x_i\}_{i=1}^n$. Consider the model

$$y = x'\beta + u; \qquad \mathrm{E}\left(u|x\right) = 0, Var\left(u|x\right) = \sigma^2.$$

Suppose you stack the observations and obtain

$$Y = X\beta + U; \mathrm{E}\left(U|X\right) = 0; \quad Var\left(U|X\right) = \sigma^2 I_n.$$

a) Justify each step in the following proof that $\hat{\sigma}^2 = \frac{SSR}{n-k-1}$ is an unbiased estimator of σ^2 conditional on X. <u>Hint</u>: Look up the properties of the trace operator.

First, $SSR = U'M_X U$. Using properties of the trace operator, we have

$$\begin{aligned}
\mathrm{E}\left(U'M_X U|X\right) &= \mathrm{E}\left(tr\left(U'M_X U\right)|X\right) \\
&= \mathrm{E}\left(tr\left(M_X UU'|X\right)\right) \\
&= tr\left(\mathrm{E}\left(M_X UU'|X\right)\right) \\
&= tr\left(M_X \mathrm{E}\left(UU'|X\right)\right) \\
&= \sigma^2 tr\left(M_X\right) \\
&= \sigma^2\left(tr\left(I_n\right) - tr\left(X\left(X'X\right)^{-1}X'\right)\right) \\
&= \sigma^2 n - \sigma^2 tr\left(\left(X'X\right)^{-1}X'X\right) \\
&= \sigma^2\left(n-k-1\right).
\end{aligned}$$

For parts (b)-(c) assume additionally that $y|x \sim \mathcal{N}\left(x'\beta, \sigma^2\right)$. Use the following facts to help answer the questions.

- Fact 1: Suppose $z \sim \mathcal{N}\left(0,1\right)$ and $Q \sim \chi^2_{n-k-1}$, and z is independent of Q. Then

$$t = \frac{z}{\sqrt{Q/\left(n-k-1\right)}} \sim t_{n-k-1}.$$

- Fact 2: If $x \sim \mathcal{N}\left(\mu, \Sigma\right)$ and if A and B are conformable constant matrices, Ax is independent of Bx iff $Cov\left(Ax, Bx\right) = 0$.

- Fact 3: $\frac{(n-k-1)\hat{\sigma}^2}{\sigma^2}|X \sim \chi^2_{n-k-1}$.

b) Show that the residual vector $\hat{U}$ and $\hat{\beta} - \beta$ are independent conditional on X. Argue that $\hat{\sigma}^2$ is independent of $\hat{\beta}$ conditional on X. <u>Hint</u>: First show that $\hat{U} = M_X U$, and argue that $U|X \sim \mathcal{N}\left(0, \sigma^2 I_n\right)$.

c) Suppose you wish to test the null hypothesis $H_0 : \beta_k = \beta_k^0$. Show that under H_0:

$$\frac{\hat{\beta}_k - \beta_k^0}{se\left(\hat{\beta}_k\right)} \sim t_{n-k-1}.$$

Now consider the model with weaker assumptions

$$y = x'\beta + u; \qquad \mathrm{E}\left(xu\right) = 0, Var\left(xu\right) = \mathrm{E}\left(u^2 xx'\right),$$

where $\mathrm{E}\left(u^2 xx'\right)$ is invertible and $\beta = (\beta_0, \beta_1, \beta_2) \in \mathbb{R}^3$.

d) Suppose we want to test $\beta_1 = \beta_2 = 1$. Describe in detail how to conduct this test.

e) Suppose you want to test the null hypothesis that $(\beta_1 - \beta_2)^2 = 0$. Can you use the test for non-linear restrictions? If not, how would you conduct a test of this hypothesis?

Question 3 Suppose $y = \beta x + u$, where y, x, u are scalar random variables, but $\mathrm{E}\left(xu\right) \neq 0$. Suppose there exists a scalar random variable z such that $\mathrm{E}\left(u|z\right) = 0$. Show by example that the IV estimate of β is generally biased.

Question 4 Consider the following regression model of the returns to schooling:

$$\ln\left(w\right) = \beta_0 + \beta_1 y + \beta_2 x + u,$$

where w denotes annual salary, y denotes years of schooling and x is a measure of previous work experience. Suppose y is endogenous x is exogenous and that z_1 is an instrument for y. Let $x = (1, y, x)$, $z = (1, z_1, x)$ and $\beta = (\beta_0, \beta_1, \beta_2)$. Suppose $\mathrm{E}\left(zx'\right), \mathrm{E}\left(zz'\right)$ and $Var\left(zu\right)$ exist, and that both $\mathrm{E}\left(zz'\right)$ and $Var\left(zu\right)$ are invertible. Suppose z_1 is relevant and valid. Suppose we observe an iid sample $\{w_i, y_i, x_i\}_{i=1}^n$.

a) What does it mean for x to be exogenous and y endogenous? Explain why we might expect y to be an endogenous variable.

b) Can we assume $\mathrm{E}\left(u\right) = 0$ without loss of generality?

c) What does it mean for z_1 to be relevant and valid?

d) Describe how to estimate β using the instrumental variables estimator.

e) Suppose you wish to test the null hypothesis that z_1 is irrelevant vs. the alternative that it is relevant at asymptotic level α. State the null and alternative hypotheses in terms of features of the joint distribution of z, x, and describe how you would perform the test.

f) Suppose $Cov\left(y, x\right) \neq 0$. Can we use x as an instrument for y?

Question 5 (Hausman Test) Suppose you observe an iid sample of $\{y_i, x_i, z_i\}_{i=1}^n$, where y is a scalar random variable, $x \in \mathbb{R}$ and $z \in \mathbb{R}$ are random variables, and consider the model

$$y = \beta x + u.$$

In this question we will use the instrumental variable z to test the assumption that $\mathrm{E}(xu) = 0$ holds in the linear model.

a) Write down the OLS estimator and the IV estimator of β using the variable z as the instrument. Write your answer using sums and also using matrices.

b) Under what conditions is z a relevant and valid instrumental variable? Show that $\hat{\beta}_{IV}$ is a consistent estimator of β under these assumptions.

c) Suppose that $\mathrm{E}(xu) = 0$ and $\mathrm{E}(zu) = 0$. Compute the joint asymptotic distribution of $\hat{\beta}_{OLS}$ and $\hat{\beta}_{IV}$. Hint: First, show that

$$\begin{aligned}
\begin{pmatrix} \hat{\beta}_{OLS} - \beta \\ \hat{\beta}_{IV} - \beta \end{pmatrix} &= \begin{pmatrix} (X'X)^{-1} & 0 \\ 0 & (Z'X)^{-1} \end{pmatrix} \begin{pmatrix} X'U \\ Z'U \end{pmatrix} \\
&= \begin{pmatrix} (X'X)^{-1} & 0 \\ 0 & (Z'X)^{-1} \end{pmatrix} \sum_{i=1}^{n} \begin{pmatrix} x_i \\ z_i \end{pmatrix} u_i.
\end{aligned}$$

Now use the CLT.

d) Explain how to use your result to conduct a test of asymptotic size α of $H_0 : \mathrm{E}(xu) = 0$ vs. $H_1 : \mathrm{E}(xu) \neq 0$. Assume that all relevant matrix inverses exist.

Question 6 Suppose you observe an iid sample of $\{y_i, x_i\}_{i=1}^n$, where y and x are scalar random variables, and you assume

$$y = \beta x + u; \quad \mathrm{E}(u|x) = 0.$$

a) Is the assumption $\mathrm{E}(u|x) = 0$ without loss of generality?

b) Show that $\mathrm{E}(xu) = 0$ and $\mathrm{E}(x^2 u) = 0$. Explain why assuming that $\mathrm{E}(u|x) = 0$ yields many different representations of β.

c) Suppose now that you wish to estimate β using a combination of the moments $\mathrm{E}(xu) = 0$ and $\mathrm{E}(x^2 u) = 0$ using GMM. Explain in detail how to calculate an asymptotically efficient GMM estimator of β based on these moments. Consider 2 cases – conditional homoskedasticity and heteroskedasticity of u.