Econ 704 Econometrics I - Discussion 10

Professor: Jack Porter

Exercise 1. (WLS, Econ 710 Midterm, 2015) Let $(y_i, x_i)_{i=1}^n$ be i.i.d. sample, where y_i is $\mathbb{R}$-valued and x_i is $\mathbb{R}^k$-valued. The parameter of interest β is given by

$$\beta = \underset{b\in\mathbb{R}^k}{\arg\min} E\left[(y_i - x_i'b)^2\tau(x_i)\right],$$

where τ is a known, scalar, non-negative, bounded function. Define $e_i = y_i - x_i'\beta$.

(1) Give an explicit formula for β.

(2) Show that $E[\tau(x_i)x_ie_i] = 0$.

(3) Under what condition (other than $\tau(x) = 1$) is β equal to the population OLS coefficient, i.e. $\beta = \underset{b}{\arg\min} E\left[(y_i - x_i'b)^2\right]$?

(4) Write down an estimator $\hat{\beta}$ for β.

(5) Under what condition is $\hat{\beta}$ you suggest unbiased for β?

(6) Show that $\hat{\beta}$ is consistent for β.

(7) Find the asymptotic distribution of $\sqrt{n}(\hat{\beta} - \beta)$?

(8) Show how this asymptotic distribution simplifies when $E[e_i^2|x_i] = \sigma^2$.

(9) Suggest an estimator for the asymptotic variance-covariance matrix of $\sqrt{n}(\hat{\beta} - \beta)$.

Proof. (1) Taking FOC yields $0 = E[x_i(y_i - x_i'\beta)\tau(x_i)]$. Assuming $E[\tau(x_i)x_ix_i']$ is invertibale and reordering yields $\beta = E[\tau(x_i)x_ix_i']^{-1}E[\tau(x_i)x_iy_i]$.

(2) $E[\tau(x_i)x_ie_i] = E[\tau(x_i)x_i(y_i - x_i'\beta)] = E[\tau(x_i)x_iy_i] - E[\tau(x_i)x_ix_i']\beta = E[\tau(x_i)x_iy_i] - E[\tau(x_i)x_iy_i] = 0$.

(3) Assume linearity $E[y_i|x_i] = x_i'\gamma$. Clearly population OLS coefficient is γ. Also,

$$\begin{aligned}
\beta &= E[\tau(x_i)x_ix_i']^{-1}E[\tau(x_i)x_iy_i] \\
&= E[\tau(x_i)x_ix_i']^{-1}E\left[E[\tau(x_i)x_iy_i|x_i]\right] \\
&= E[\tau(x_i)x_ix_i']^{-1}E\left[\tau(x_i)x_iE[y_i|x_i]\right] \\
&= E[\tau(x_i)x_ix_i']^{-1}E\left[\tau(x_i)x_ix_i'\gamma\right] \\
&= \gamma.
\end{aligned}$$

So β is equal to population OLS coefficient under linearity assumption.

(4) Use the sample analogue of $\beta = E[\tau(x_i)x_ix_i']^{-1}E[\tau(x_i)x_iy_i]$. In particular, an estimator for β is $\hat{\beta} = \left(\frac{1}{n}\sum_{i=1}^n \tau(x_i)x_ix_i'\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^n \tau(x_i)x_iy_i\right)$.

(5) Assume linearity $E[y_i|x_i] = x_i'\beta$. Then $E[e_i|x_i] = 0$. Conditional mean of $\hat{\beta}$ is

$$\begin{aligned} E[\hat{\beta}|x_1,\ldots,x_n] &= E\left[\left(\frac{1}{n}\sum_{i=1}^{n}\tau(x_i)x_ix_i'\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}\tau(x_i)x_i(x_i'\beta+e_i)\right)|x_1,\ldots,x_n\right] \\ &= \beta + \left(\frac{1}{n}\sum_{i=1}^{n}\tau(x_i)x_ix_i'\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}\tau(x_i)x_iE[e_i|x_1,\ldots,x_n]\right) \\ &= \beta. \end{aligned}$$

We also have $E[\hat{\beta}] = E\left[E[\hat{\beta}|x_1,\ldots,x_n]\right] = 0$.

(6) WLLN and Continuous Mapping Theorem imply

$$\hat{\beta} = \left(\frac{1}{n}\sum_{i=1}^{n}\tau(x_i)x_ix_i'\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}\tau(x_i)x_iy_i\right) \xrightarrow{p} E[\tau(x_i)x_ix_i']^{-1}E[\tau(x_i)x_iy_i] = \beta.$$

(7) Using CLT, Continuous Mapping Theorem, and Slutsly's Theorem,

$$\begin{aligned} \sqrt{n}(\hat{\beta}-\beta) &= \left(\frac{1}{n}\sum_{i=1}^{n}\tau(x_i)x_ix_i'\right)^{-1}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\tau(x_i)x_ie_i\right) \\ &\xrightarrow{d} E[\tau(x_i)x_ix_i']^{-1}N\left(0, Var(\tau(x_i)x_ie_i)\right) \\ &= N(0, E[\tau(x_i)x_ix_i']^{-1}E[\tau(x_i)^2e_i^2x_ix_i']E[\tau(x_i)x_ix_i']^{-1}). \end{aligned}$$

(8) Variance expression above can be rewritten as

$$\begin{aligned} &E[\tau(x_i)x_ix_i']^{-1}E[\tau(x_i)^2e_i^2x_ix_i']E[\tau(x_i)x_ix_i']^{-1} \\ =&E[\tau(x_i)x_ix_i']^{-1}E\left[E[\tau(x_i)^2e_i^2x_ix_i'|x_i]\right]E[\tau(x_i)x_ix_i']^{-1} \\ =&E[\tau(x_i)x_ix_i']^{-1}E\left[E[e_i^2|x_i]\tau(x_i)^2x_ix_i'\right]E[\tau(x_i)x_ix_i']^{-1} \\ =&\sigma^2E[\tau(x_i)x_ix_i']^{-1}E\left[\tau(x_i)^2x_ix_i'\right]E[\tau(x_i)x_ix_i']^{-1} \end{aligned}$$

(9) Use $\widehat{V} = \left(\sum_{i=1}^{n}\tau(x_i)x_ix_i'\right)^{-1}\left(\sum_{i=1}^{n}\tau(x_i)^2\hat{e}_i^2x_ix_i'\right)\left(\sum_{i=1}^{n}\tau(x_i)x_ix_i'\right)^{-1}$, where $\hat{e}_i = y_i - x_i\hat{\beta}$. □

Exercise 2. (Bruce Hansen, Econ 710, Midterm 2001) The model is

$$y_i = x_i'\beta + \epsilon_i, \quad E(\epsilon_i|x_i) = 0.$$

An econometrician is worried about the impact of some unusually large values of the regressors. The model is thus estimated on the subsample for which $|x_i| \leq c$, for some fixed c. Let $\tilde{\beta}$ denote the OLS estimator on this subsample. It equals

$$\tilde{\beta} = \left(\sum_{i=1}^{n}x_ix_i'1(|x_i|\leq c)\right)^{-1}\left(\sum_{i=1}^{n}x_iy_i1(|x_i|\leq c)\right)$$

where $1(\cdot)$ denotes the indicator function.

(a) Show that $\tilde{\beta} \to_p \beta$.

(b) Find the asymptotic distribution of $\sqrt{n}(\tilde{\beta} - \beta)$.

(c) Suppose instead the model is

$$y_i = x_i'\beta + \epsilon_i, \quad E(x_i\epsilon_i) = 0.$$

Does result (a) change?

Proof. (a) Since

$$\tilde{\beta} = \beta + \left(\frac{1}{n} \sum_{i=1}^{n} x_i x_i' 1(|x_i| \le c) \right)^{-1} \left(\frac{1}{n} \sum_{i=1}^{n} x_i \epsilon_i 1(|x_i| \le c) \right)$$

so that we check the converging value of the second term. By WLLN,

$$\left(\frac{1}{n} \sum_{i=1}^{n} x_i x_i' 1(|x_i| \le c) \right)^{-1} \to_p E(x_i x_i' 1(|x_i| \le c))^{-1}$$

and the next term is

$$\left(\frac{1}{n} \sum_{i=1}^{n} x_i \epsilon_i 1(|x_i| \le c) \right) \to_p E(x_i \epsilon_i 1(|x_i| \le c)) = E\left[x_i 1(|x_i| \le c) E\left[\epsilon_i | x_i\right]\right] = 0$$

since $E\left[\epsilon_i | x_i\right] = 0$ by assumption. Therefore $\tilde{\beta} \to_p \beta$.

(b) By above we know

$$\sqrt{n}(\tilde{\beta} - \beta) = \left(\frac{1}{n} \sum_{i=1}^{n} x_i x_i' 1(|x_i| \le c) \right)^{-1} \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} x_i \epsilon_i 1(|x_i| \le c) \right).$$

The first term is the same as above, so it converges to $E(x_i x_i' 1(|x_i| \le c))^{-1} =: \Gamma^{-1}$. For the second term, we can use CLT, so

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} x_i \epsilon_i 1(|x_i| \le c) \to_d N(0, Var(x_i \epsilon_i 1(|x_i| \le c)).$$

Then what is $Var(x_i \epsilon_i 1(|x_i| \le c)$? It is by definition $E(x_i x_i' \epsilon_i^2 1(|x_i| \le c)) - E(x_i \epsilon_i 1(|x_i| \le c)) E(x_i \epsilon_i 1(|x_i| \le c))' = E(x_i x_i' \epsilon_i^2 1(|x_i| \le c)) =: \Omega$. So

$$\sqrt{n}(\tilde{\beta} - \beta) = \left(\frac{1}{n} \sum_{i=1}^{n} x_i x_i' 1(|x_i| \le c) \right)^{-1} \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} x_i \epsilon_i 1(|x_i| \le c) \right) \to_d \Gamma^{-1} N(0, \Omega) \sim N(0, \Gamma^{-1} \Omega \Gamma^{-1'}).$$

(c) The result of (a) is changed. In (a), our result came from showing $E(x_i \epsilon_i 1(|x_i| \le c)) = 0$. Unfortunately, $E(x_i \epsilon_i) = 0$ cannot be a sufficient condition for that. Let's take a counter-example. Let $c = 1$, and x_i and ϵ_i have the following distribution.

$x_i \backslash \epsilon_i$	-1	0	1
-1	$\frac{1}{12}$	0	0
0	0	$\frac{5}{8}$	0
2	$\frac{1}{6}$	0	$\frac{1}{8}$

Then $E(x_i\epsilon_i) = 1 \cdot \frac{1}{12} + (-2) \cdot \frac{1}{6} + 2 \cdot \frac{1}{8} = 0$ but $E(x_i\epsilon_i 1(|x_i| \leq 1)) = 1 \cdot \frac{1}{12} + 0 + 0 = \frac{1}{12} \neq 0$. This is also a counter-example to show that $E(x_i\epsilon_i) = 0$ may not imply $E(\epsilon_i|x_i) = 0$. Conditioning on $x_i = -1$, the probability distribution of ϵ_i degenerates, $E(\epsilon_i|x_i = -1) = -1 \neq 0$. □

Exercise 3. (2SLS, ECON 710 Final 2013) Take a linear equation with endogeneity and a just-identified linear reduced form

$$y_i = x_i\beta + \epsilon_i, x_i = \gamma z_i + u_i$$

where both x_i and z_i are scalar. Assume that $\mathbb{E}(z_i\epsilon_i) = 0$ and $\mathbb{E}(z_i u_i) = 0$.

(1) Write down the standard 2SLS estimator $\hat{\beta}_{2SLS}$ for β using z_i as an instrument for x_i.

(2) Find the asymptotic distribution for $\hat{\beta}_{2SLS}$. Write the asymptotic variance as a function of $\Omega = \mathbb{E}(z_i^2\epsilon^2), Q = \mathbb{E}(z_i^2)$, and γ.

Proof. (1) Using the formula for 2SLS estimator, we have

$$\begin{aligned}\hat{\beta}_{2SLS} &= \left[\left(\frac{1}{n}\sum_{i=1}^{n} x_i z_i'\right)\left(\frac{1}{n}\sum_{i=1}^{n} z_i z_i'\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} z_i x_i'\right)\right]^{-1} \\ &\quad \cdot \left(\frac{1}{n}\sum_{i=1}^{n} x_i z_i'\right)\left(\frac{1}{n}\sum_{i=1}^{n} z_i z_i'\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} z_i y_i\right).\end{aligned}$$

Since z_i and x_i are scalar, each element in the first square bracket is invertible, and we can simplify it as

$$\hat{\beta}_{2SLS} = \left(\frac{1}{n}\sum_{i=1}^{n} x_i z_i\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} z_i y_i\right).$$

(2) The 2SLS estimator can be rewritten as

$$\begin{aligned}\hat{\beta}_{2SLS} &= \left(\frac{1}{n}\sum_{i=1}^{n} x_i z_i\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} z_i y_i\right) \\ &= \left(\frac{1}{n}\sum_{i=1}^{n} x_i z_i\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} z_i(x_i\beta + \epsilon_i)\right) \\ &= \beta + \left(\frac{1}{n}\sum_{i=1}^{n} x_i z_i\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} z_i\epsilon_i\right).\end{aligned}$$

You can easily prove that this converges in probability to β (Check this!). For asymptotic normality,

$$\begin{aligned}\sqrt{n}(\hat{\beta}_{2SLS} - \beta) &= \left(\frac{1}{n}\sum_{i=1}^{n} x_i z_i\right)^{-1}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} z_i\epsilon_i\right) \\ &= \left(\gamma\frac{1}{n}\sum_{i=1}^{n} z_i^2 + \frac{1}{n}\sum_{i=1}^{n} z_i u_i\right)^{-1}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} z_i\epsilon_i\right) \\ &\to_d N(0, Q^{-1}\Omega Q^{-1}/\gamma^2) = N(0, \Omega/Q^2\gamma^2)\end{aligned}$$

The convergence holds by WLLN, CLT, $\mathbb{E}(z_i u_i) = 0$ and $\mathbb{E}(z_i\epsilon_i) = 0$ and the last equality holds since Q is scalar. (Check this!) □