Econ 241B Econometrics
Final Examination

March 20, 2018

Please answer all questions. Show your work. You are allow to use one sheet (both sides) for formulas.

1. Let $\{(y_t, \boldsymbol{x}_t)\}_{t=1}^T$ be a sample of length T. Define

$$S_T^2 = (T-k)^{-1} \sum_{t=1}^{T} (y_t - \boldsymbol{x}_t' \boldsymbol{b}_T)^2$$

as the OLS error variance estimator where $\boldsymbol{x}_t$ is a $k \times 1$ vector, $\boldsymbol{b}_T$ is the OLS estimator of $\boldsymbol{\beta}$ from a sample of length T.

a. Assume that $\mathbb{E}[u_t^2 | \boldsymbol{X}] = \sigma^2$ $(t = 1, \ldots, T)$ and $\boldsymbol{X} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_T)$. Show the limit in probability of S_T^2 and clearly state any assumptions needed.

Solution: Multiplying by T both in the numerator and denominator

$$S_T^2 = \frac{T}{T-k} \left(\frac{1}{T} \sum_{t=1}^{T} \hat{u}_t^2 \right)$$

we know that $\frac{T}{T-k}$ converges to 1 as T grows large and, therefore, in probability. What is the limit in probability of the following term?

$$\frac{1}{T} \sum_{t=1}^{T} \hat{u}_t^2$$

We have that

$$\frac{1}{T} \sum_{t=1}^{T} \hat{u}_t^2 = \frac{1}{T} \sum_{t=1}^{T} (y_t - \boldsymbol{x}_t' \boldsymbol{b}_T)^2$$

adding and substracting $\boldsymbol{x}_t' \boldsymbol{\beta}$,

$$\frac{1}{T} \sum_{t=1}^{T} \hat{u}_t^2 = \frac{1}{T} \sum_{t=1}^{T} (y_t - \boldsymbol{x}_t' \boldsymbol{\beta} - \boldsymbol{x}_t'(\boldsymbol{b}_T - \boldsymbol{\beta}))^2$$

By the definition of u_t,

$$\frac{1}{T} \sum_{t=1}^{T} \hat{u}_t^2 = \frac{1}{T} \sum_{t=1}^{T} (u_t - \boldsymbol{x}_t'(\boldsymbol{b}_T - \boldsymbol{\beta}))^2$$

Expanding,

$$\frac{1}{T} \sum_{t=1}^{T} \hat{u}_t^2 = \frac{1}{T} \sum_{t=1}^{T} u_t^2 - 2\frac{1}{T} \sum_{t=1}^{T} (\boldsymbol{b}_T - \boldsymbol{\beta})' x_t u_t + (\boldsymbol{b}_T - \boldsymbol{\beta})' \left(\frac{1}{T} \sum_{t=1}^{T} \boldsymbol{x}_t \boldsymbol{x}_t' \right) (\boldsymbol{b}_T - \boldsymbol{\beta}) \tag{1}$$

Let's show the convergence in probability of each of the terms of Equation 1

$$\begin{aligned}
\operatorname{plim} \frac{1}{T} \sum_{t=1}^{T} (\boldsymbol{b}_T - \boldsymbol{\beta})' \boldsymbol{x}_t u_t &= \operatorname{plim} (\boldsymbol{b}_T - \boldsymbol{\beta})' \operatorname{plim} \frac{1}{T} \sum_{t=1}^{T} \boldsymbol{x}_t u_t \\
&= \boldsymbol{0}' \cdot \mathbb{E}[\boldsymbol{x}_t' u_t] = 0 \\
&= \boldsymbol{0}' \cdot \boldsymbol{0} = 0
\end{aligned}$$

Given that $\boldsymbol{b}_T$ is a consistent estimator of $\boldsymbol{\beta}$ and that by assumption $\{\boldsymbol{x}_t u_t\}$ is a sequence of stationary and ergodic martingale differences. Additionally, we know that the limit in probability of a product is equal to the product of the limits in probability provided the limits exist.

$$\begin{aligned}\operatorname{plim}(\boldsymbol{b}_T-\boldsymbol{\beta})'\left(\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{x}_t\boldsymbol{x}_t'\right)(\boldsymbol{b}_T-\boldsymbol{\beta}) &= \operatorname{plim}(\boldsymbol{b}_T-\boldsymbol{\beta})'\operatorname{plim}\left(\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{x}_t\boldsymbol{x}_t'\right)\operatorname{plim}(\boldsymbol{b}_T-\boldsymbol{\beta})\\ &= \mathbf{0}'\cdot\boldsymbol{Q}\cdot\mathbf{0}\\ &= 0\end{aligned}$$

Again, given that $\boldsymbol{b}_T$ is a consistent estimator of $\boldsymbol{\beta}$, $\{\boldsymbol{x}_t\}$ is stationary and ergodic, and $\mathbb{E}[\boldsymbol{x}_t\boldsymbol{x}_t'] = \boldsymbol{Q}$ exists. In addition, assume that there is an intercept in $\boldsymbol{X}$

Taking the limit in probability of Equation 1 and plugging the above limits in probability, we have the following

$$\begin{aligned}\operatorname{plim}\frac{1}{T}\sum_{t=1}^{T}\hat{u}_t^2 &= \operatorname{plim}\frac{1}{T}\sum_{t=1}^{T}u_t^2\\ \operatorname{plim}\frac{1}{T}\sum_{t=1}^{T}\hat{u}_t^2 &= \mathbb{E}[u_t^2|\boldsymbol{X}]\\ \operatorname{plim}S_T^2 = \mathbb{E}[u_t^2|\boldsymbol{X}] &= \operatorname{Var}(u_t|\boldsymbol{X}) = \sigma^2\end{aligned}$$

b. Assume that $\mathbb{E}[u_t^2|\boldsymbol{X}] = \sigma_t^2$ $(t = 1,\dots,T)$ and $\boldsymbol{X} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_T)$. Show the limit in probability of S_T^2 and clearly state any assumptions needed.

Solution: Following a procedure similar to part a, we get that

$$\operatorname{plim}\frac{1}{T}\sum_{t=1}^{T}\hat{u}_t^2 = \operatorname{plim}\frac{1}{T}\sum_{t=1}^{T}u_t^2$$

Since $\mathbb{E}[u_t^2|\boldsymbol{X}] = \sigma_t^2$ is not constant across $t = 1,\dots,T$, the number of parameters grows as the sample length grows $(T\to\infty)$ and there is not feasible consistent estimation of the error variance matrix.

Nevertheless, it is still possible to get a consistent estimator of the OLS estimator asymptotic variance using the correction suggested by Eicker, Huber and White. In particular,

$$\left(\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{x}_t\boldsymbol{x}_t'\right)^{-1}\frac{1}{T}\sum_{t=1}^{T}\hat{u}_t^2\boldsymbol{x}_t\boldsymbol{x}_t'\left(\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{x}_t\boldsymbol{x}_t'\right)^{-1}\xrightarrow{p} AVar(\boldsymbol{\beta}_T)$$

where $AVar(\boldsymbol{\beta}_T)$ is the asymptotic variance of $\boldsymbol{\beta}_T$.

2. Suppose $\sqrt{n}\left(\widehat{\mu}-\mu\right) \rightsquigarrow \mathcal{N}\left(0,\nu^2\right)$, where $\mu \neq 0$. Set $\beta = \frac{1}{\mu}$ and $\widehat{\beta} = \frac{1}{\widehat{\mu}}$.

a. Determine the distribution for $\sqrt{n}\left(\widehat{\beta}-\beta\right)$.

Solution: By the Delta Method,

$$\sqrt{n}\left(g\left(\widehat{\mu}\right)-g\left(\mu\right)\right) \rightsquigarrow \mathcal{N}\left(0, g'\left(\mu\right)^2\nu^2\right).$$

Here $g\left(\mu\right) = \frac{1}{\mu}$ and $g'\left(\mu\right) = \frac{-1}{\mu^2}$ so

$$\sqrt{n}\left(\widehat{\beta}-\beta\right) \rightsquigarrow \mathcal{N}\left(0, \frac{1}{\mu^4}\nu^2\right).$$

b. Use a Taylor expansion for $\widehat{\beta} = g\left(\widehat{\mu}\right)$ to justify the asymptotic distribution you proposed in part a.

Solution: The Taylor expansion is

$$\frac{1}{\widehat{\mu}} = \frac{1}{\mu} + \left(\frac{-1}{\mu^2}\right)(\widehat{\mu} - \mu) + \frac{1}{2}\left(\frac{2}{\mu^3}\right)(\widehat{\mu} - \mu)^2 + R.$$

Note $\sqrt{n}\,(\widehat{\mu} - \mu) = O_P(1)$ and $\sqrt{n}\,(\widehat{\mu} - \mu)^2 = o_P(1)$ (as are all higher order terms) thus

$$\sqrt{n}\left(\frac{1}{\widehat{\mu}} - \frac{1}{\mu}\right) = \left(\frac{-1}{\mu^2}\right)\sqrt{n}\,(\widehat{\mu} - \mu) + o_P(1).$$

This is the first-order approximation that yields the Delta Method

$$\sqrt{n}\left(\widehat{\beta} - \beta\right) \rightsquigarrow \mathcal{N}\left(0, \frac{1}{\mu^4}\nu^2\right).$$

c. From a sample of data with $n = 20$, the values $\widehat{\mu} = 5$ and $\widehat{\nu}^2 = 500$ are obtained. Find the associated estimate for β and the estimated standard errors for $\widehat{\mu}$ and $\widehat{\beta}$.

Solution: From the reported values, $\widehat{\beta} = \frac{1}{5}$, the estimated standard errors are

$$\text{for } \widehat{\mu} : \widehat{\sigma}_{\widehat{\mu}} = \sqrt{\frac{\widehat{\nu}^2}{n}} = 5 \quad \text{for } \widehat{\beta} : \widehat{\sigma}_{\widehat{\beta}} = \sqrt{\frac{1}{\widehat{\mu}^4}\frac{\widehat{\nu}^2}{n}} = \frac{1}{5}.$$

d. With the values from part c, test $H_0 : \mu = 1$ against $H_1 : \mu \neq 1$. Recast the hypothesis test in terms of β and use the values from part c to test. Explain why these tests should, or should not, reach the same conclusion.

Solution: The value of the test statistic for $H_0 : \mu = 1$ is

$$\frac{5-1}{5} = .8 \text{ which is not large enough to reject at even the 10\% level.}$$

For β the null hypothesis is $H_0 : \beta = 1$ with alternative $H_1 : \beta \neq 1$. The value of the test statistic is

$$\frac{\frac{1}{5} - 1}{\frac{1}{5}} = 1 - 5 = -4 \text{ which is large enough to reject at even the 1\% level.}$$

These tests should reach the same conclusion. What has gone wrong? As the Taylor expansion of part b makes clear, the approximation in part a is only accurate in a local neighborhood for $\widehat{\mu}$ about μ. Yet the estimated variance is 500, indicating that $\widehat{\mu}$ is not contained in a local neighborhood of μ with high probability, so the reported variance for $\frac{1}{\widehat{\mu}}$ is too small. Indeed, with such a large variance, there is a substantial probability that $\widehat{\mu}$ lies close to 0, causing $\frac{1}{\widehat{\mu}}$ to blow up.