Econ 741 Masao Ogaki

Spring 2002 5/30/02

ANSWERS TO THE SAMPLE FINAL EXAM PROBLEMS

The number in square brackets indicate after each question indicates the points allocated to the question and the numbers in the square brackets in the bold face indicate partial credit.

$$\mathbf{1}.\mathbf{a} - F[\mathbf{2}]; \mathbf{b} - F[\mathbf{2}]; \mathbf{c} - F[\mathbf{2}]; \mathbf{d} - T[\mathbf{2}]; \mathbf{e} - F[\mathbf{2}]; \mathbf{f} - F[\mathbf{2}]; \mathbf{g} - T[\mathbf{2}]$$

2.[**3**] A test is said to over-reject in small samples when the probability of rejecting the null hypothesis when it is true is larger than the nominal size when the nominal critical value is used in small samples. [**2**] In this case, the true critical value is greater than the nominal critical value. [**1**]

3.a.[**4**]

$$t_k = \frac{\sqrt{n}(b_k - \overline{\beta}_k)}{\sqrt{s^2[(\frac{1}{n}\mathbf{X}'\mathbf{X})^{-1}]_{kk}}}$$

The denominator converges in probability to the square root of $\sigma^2[E(\mathbf{x}_t\mathbf{x}_t')]_{kk}$, and the numerator converges almost surely to a normal random variable with mean zero and variance $\sigma^2[E(\mathbf{x}_t\mathbf{x}_t')]_{kk}$. Hence t_k converges in distribution to a standard normal random variable.

3.b. [**4**] The t_k statistic has the exact t distribution with $n - K$ degrees of freedom. Because the critical value based on the t distribution is always greater than 1.96 in finite samples, the test overrejects. [**2**] The actual size can be obtained from a t distribution table. The actual size is larger than 10 percent if the cutoff point for Student's t distribution for the 0.05 right-hand tail probability is greater than 1.96. This is true for $df = 1,2,3,4,5$, and not true if df is greater than 5. When $n = 5, df = 1$, and therefore the actual size is larger than 10 percent. For $n = 8$ and $9, df = 4$ and 5, respectively; so this is still true. For $n = 10, 11$, the df is greater than 5. Therefore, the actual size is smaller than 10 percent. [**2**]

4. [**15**]

4.a.[**7**] Using Proposition 3.1, we obtain

$$\sqrt{n}(\hat{\delta}_\ell(\hat{\mathbf{W}}) - (\bar{\delta}_\ell + \frac{\gamma}{\sqrt{n}})) = \sqrt{n}(\hat{\delta}_\ell(\hat{\mathbf{W}}) - \bar{\delta}_\ell) - \gamma \xrightarrow{d} N(\mathbf{0}, \sigma^2(\hat{\mathbf{W}})).$$

Hence

$$\sqrt{n}\ t_\ell \xrightarrow{d} N(\mu, 1)$$

where $\mu = \gamma/\sqrt{\sigma^2(\hat{\mathbf{W}})}$.

4.b. [**8**] From $3.a, Z = t_\ell - \mu$ converges in distribution to $N(0,1)$. The asymptotic power of t test based on $\hat{\delta}(\hat{\mathbf{W}}^*)$ is $Prob(t_\ell > 1.65) = Prob(t_\ell - 9/6 > 1.65 - 9/6) = Prob(Z > 0.15) = 0.4404$. The asymptotic power of the t test based on $\hat{\delta}(\hat{\mathbf{S}}^{-1})$ is $Prob(t_\ell > 1.65) = Prob(t_\ell - 9/4 > 1.65 - 9/4) = Prob(Z > -0.6) = 1 - 0.2743 = 0.7257$.

5 [**34**].

5.a. [**3**] The IV estimator is

$$\mathbf{b}_{IV} = (\sum_{t=1}^{T} \mathbf{z}_t \mathbf{x}_t')^{-1} \sum_{t=1}^{T} \mathbf{z}_t y_t$$

5.b. [**7**] From (A.2), Part (a) of Gordin's condition is satisfied. To see Part (b) of Gordin's condition, we note that $\mathbf{g}_t$ is in I_{t+3} because $\mathbf{g}_t = \mathbf{z}_t e_t = \mathbf{z}_t(y_t - \mathbf{x}_t'\beta)$. We also note that $E(\mathbf{g}_t|I_t) = E(\mathbf{z}_t e_t|I_t) = \mathbf{z}_t E(e_t|I_t) = \mathbf{0}$. Hence $E(\mathbf{g}_t|\mathbf{g}_{t-j}, \mathbf{g}_{t-j-1}, \ldots) =$

$E(E(\mathbf{g}_t|I_t)|\mathbf{g}_{t-j}, \mathbf{g}_{t-j-1}, \ldots) = 0$ for $j \geq 3$. Hence Part (b) of Gordin's condition is satisfied. Similarly, $\mathbf{r}_{tj} = E(\mathbf{g}_t|\mathbf{g}_{t-j}, \mathbf{g}_{t-j-1}, \ldots) - E(\mathbf{g}_t|\mathbf{g}_{t-j-1}, \mathbf{g}_{t-j-2}, \ldots) = 0$ for $j \geq 3$. Hence Part (c) of Gordin's condition is satisfied.

5.c. [**7**] Applying the Ergodic theorem to $\mathbf{z}_t\mathbf{x}_t$ and $\mathbf{z}_t e_t$, we have

$$\frac{1}{T}\sum_{t=1}^{T} \mathbf{z}_t\mathbf{x}_t' \xrightarrow{a.s.} E(\mathbf{z}_t\mathbf{x}_t')$$

$$\frac{1}{T}\sum_{t=1}^{T} \mathbf{z}_t e_t' \xrightarrow{a.s.} E(\mathbf{z}_t e_t)$$

Because of Assumption $(A5)$, $\frac{1}{T}\sum_{t=1}^{T} \mathbf{z}_t\mathbf{x}_t'$ is nonsingular for large enough T with probability one, and therefore,

$$\begin{aligned}
\mathbf{b}_{IV} - \mathbf{b}_0 &= (\frac{1}{T}\sum_{t=1}^{T} \mathbf{z}_t\mathbf{x}_t')^{-1}\frac{1}{T}\sum_{t=1}^{T} \mathbf{z}_t y_t - \mathbf{b}_0 \\
&= (\frac{1}{T}\sum_{t=1}^{T} \mathbf{z}_t\mathbf{x}_t')^{-1}\frac{1}{T}\sum_{t=1}^{T} \mathbf{z}_t(\mathbf{x}_t\mathbf{b}_0 + e_t) - \mathbf{b}_0 \\
&= (\frac{1}{T}\sum_{t=1}^{T} \mathbf{z}_t\mathbf{x}_t')^{-1}\frac{1}{T}\sum_{t=1}^{T} \mathbf{z}_t e_t \\
&\xrightarrow{a.s.} E(\mathbf{z}_t\mathbf{x}_t')^{-1}E(\mathbf{z}_t e_t) = 0.
\end{aligned}$$

5.d. [**9**] Applying the Gordin and Hansen's central limit theorem to $\mathbf{z}_t e_t$, we obtain

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T} \mathbf{z}_t e_t' \xrightarrow{d} N(0, \mathbf{S}).$$

where $\mathbf{S} = \mathbf{\Gamma}_{-2} + \mathbf{\Gamma}_{-1} + \mathbf{\Gamma}_0 + \mathbf{\Gamma}_1 + \mathbf{\Gamma}_2$, where $\mathbf{\Gamma}_j = E(e_t^2\mathbf{z}_t\mathbf{z}_{t-j}')$.[**6**] Hence

$$\sqrt{T}(\mathbf{b}_{IV} - \boldsymbol{\beta}) = (\frac{1}{T}\sum_{t=1}^{T} \mathbf{z}_t\mathbf{x}_t')^{-1}\frac{1}{\sqrt{T}}\sum_{t=1}^{T} \mathbf{z}_t e_t' \xrightarrow{d} E(\mathbf{z}_t\mathbf{x}_t')^{-1}N(0, \mathbf{S}).$$

Hence $\sqrt{T}(\mathbf{b}_T - \boldsymbol{\beta}) \xrightarrow{d} N(0, E(\mathbf{z}_t\mathbf{x}_t')^{-1}\mathbf{S}E(\mathbf{x}_t\mathbf{z}_t')^{-1})$.[**3**]

5.e. Hansen's J statistic has an asymptotic χ^2 distribution with 2 degrees of freedom. The critical value is 5.99. Hence the overidentifying restrictions are not rejected at the 5% level.

6.[**14**] Because $\mathbf{S}^{-1}$ is positive definite, there exists a nonsingular $K \times K$ matrix Λ such that $\mathbf{S}^{-1} = \Lambda'\Lambda$. Then $\mathbf{S} = \Lambda^{-1}(\Lambda')^{-1}$. Let $A_1 = \Lambda\Sigma_{xz}$ and $A_2 = (\Lambda')^{-1}\mathbf{W}\Sigma_{xz}$. Because $I - A_2(A_2'A_2)^{-1}A_2'$ is positive semidefinite by Fact 1, we have

$$A_1'A_1 \geq A_1'A_2(A_2'A_2)^{-1}A_2'A_1.$$

From Fact 2, we obtain

$$(A_1'A_1)^{-1} \leq (A_2'A_1)^{-1}A_2'A_2(A_1'A_2)^{-1}.$$

Because the asymptotic variance of $\hat{\delta}(\mathbf{S}^{-1})$ is $(\Sigma_{xz}'\Omega^{-1}\Sigma_{xz})^{-1} = (A_1'A_1)^{-1}$ and the asymptotic variance of $\hat{\delta}(\mathbf{W})$ is $(A_1'A_2)^{-1}A_2'A_2(A_2'A_1)^{-1}$, the conclusion follows from this inequality.

7.a. [**5**] The equations in the multiple equation GMM are identified if and only if each equation is identified. The order condition for an equation states that the number of instrumental variables (K_m) is greater than or equal to the number of parameters estimated (L_m). When $(1, h_i, \mathbf{q}_i')'$ is used as instruments, $K_1 = 3$, and $L_1 = 2$ for (1.1) and $K_2 = 3$ and $L_2 = 4$ for (1.2). Hence the order condition is satisfied by (1.1), but not by (1.2). Because the order condition is not satisfied for (1.2), the system is not identified.

However, the order condition is satisfied for (1.1). Therefore, it is possible that (1.1) is identified.

7.b. [**5**] When $k = 2, K_1 = 4$, and $L_1 = 2$ for (1.1) and $K_2 = 4$ and $L_2 = 4$ for (1.2). The order condition is satisfied by each equation. Therefore, it is possible for both (1.1) and (1.2) to be identified in the multiple equation GMM.

7.c. Because

$$E(\epsilon_i e_i \mathbf{x}_i \mathbf{x}_i') = E(E(\epsilon_i e_i \mathbf{x}_i \mathbf{x}_i' | \mathbf{x}_i)) = E(E(\epsilon_i e_i)\mathbf{x}_i \mathbf{x}_i') = \mathbf{0}$$

Equation (4.4.3) in the Summary is satisfied. Therefore, the efficient multiple equation GMM estimator and the efficient equation-by-equation GMM are asymptotically equivalent by Proposition 4.3. Because the 3SLS estimator is a multiple equation GMM estimator and the 2SLS estimator is an efficient equation-by-equation GMM estimator, the 3SLS estimator is asymptotically equivalent to the 2SLS estimator. Hence the 3SLS estimator is not more efficient asymptotically.