Econometrics 710
Final Exam, Spring 2012
Sample Answers

1. The estimator $\widehat{\sigma}^2$ is not appropriate, as the LS residuals $\widehat{e} = Y - \widehat{X}\widehat{\beta}$ are not appropriate estimates of the errors $e = Y - X\beta$. Instead, we want to use the residuals $\widetilde{e} = Y - X\widehat{\beta}$ and the estimator

$$\widetilde{\sigma}^2 = \frac{1}{n}\widetilde{e}'\widetilde{e}$$

One way to see this is to think of the just-identified case, and write the moment conditions

$$\begin{aligned}\mathbb{E}\left(z_i\left(y_i - x_i'\beta\right)\right) &= 0\\ \mathbb{E}\left(\left(y_i - x_i'\beta\right)^2 - \sigma^2\right) &= 0\end{aligned}$$

Then the moment estimator of the parameters $\left(\beta, \sigma^2\right)$ are the solutions to the two empirical analogs. The solution to the first equation is $\widehat{\beta}$, and the solution to the second equation is $\widetilde{\sigma}^2$, not $\widehat{\sigma}^2$.

Another way to see this is write out the estimators. Suppose for extreme simplicity there was no estimation error, so that $\widehat{\beta} = \beta$ and $\widehat{\Gamma} = \Gamma$. Then if we write the reduced form as $X = Z\Gamma + u$, then $\widehat{X} = Z\Gamma = X - u$, and

$$\widehat{e} = Y - \widehat{X}\widehat{\beta} = e + X\beta - (X - u)\beta = e + u\beta \neq e$$

and $\widehat{\sigma}^2 = \frac{1}{n}\widehat{e}'\widehat{e}$ is clearly estimating $\operatorname{var}\left(e_i + u_i'\beta\right) \neq \operatorname{var}\left(e_i\right).$

2.

(a) The GMM/IV/2SLS estimators for β_1 and β_2 are

$$\begin{aligned}\widehat{\beta}_1 &= \left(Z_1'X_1\right)^{-1}\left(Z_1'Y_1\right)\\ \widehat{\beta}_2 &= \left(Z_2'X_2\right)^{-1}\left(Z_2'Y_2\right)\end{aligned}$$

with asymptotic distributions

$$\begin{aligned}\sqrt{n}\left(\widehat{\beta}_1 - \beta_1\right) &\xrightarrow{d} N(0, V_1)\\ \sqrt{n}\left(\widehat{\beta}_2 - \beta_2\right) &\xrightarrow{d} N(0, V_2)\end{aligned}$$

where the two normal distributions are independent,

$$\begin{aligned}V_1 &= Q_1^{-1}\Omega_1 Q_1^{-1\prime}\\ V_2 &= Q_2^{-1}\Omega_2 Q_2^{-1\prime}\\ Q_1 &= \mathbb{E}\left(z_{1i}x_{1i}'\right)\\ Q_2 &= \mathbb{E}\left(z_{2i}x_{2i}'\right)\\ \Omega_1 &= \mathbb{E}\left(z_{1i}z_{1i}'e_{1i}^2\right)\\ \Omega_2 &= \mathbb{E}\left(z_{2i}z_{2i}'e_{2i}^2\right)\end{aligned}$$

Thus

$$\sqrt{n}\left(\left(\widehat{\beta}_1 - \beta_1\right) - \left(\widehat{\beta}_2 - \beta_2\right)\right) \xrightarrow{d} N(0, V_1 + V_2)$$

A Wald-type test for $H_0 : \beta_1 = \beta_2$ is

$$W_n = n\left(\widehat{\beta}_1 - \widehat{\beta}_2\right)'\left(\widehat{V}_1 + \widehat{V}_2\right)^{-1}\left(\widehat{\beta}_1 - \widehat{\beta}_2\right)$$

where

$$
\begin{aligned}
\widehat{V}_1 &= \widehat{Q}_1^{-1}\widehat{\Omega}_1\widehat{Q}_1^{-1\prime} \\
\widehat{V}_2 &= \widehat{Q}_2^{-1}\widehat{\Omega}_2\widehat{Q}_2^{-1\prime} \\
\widehat{Q}_1 &= \frac{1}{n}\sum_{i=1}^{n} z_{1i}x_{1i}' \\
\widehat{Q}_2 &= \frac{1}{n}\sum_{i=1}^{n} z_{2i}x_{2i}' \\
\widehat{\Omega}_1 &= \frac{1}{n}\sum_{i=1}^{n} z_{1i}z_{1i}'\hat{e}_{1i}^2 \\
\widehat{\Omega}_2 &= \frac{1}{n}\sum_{i=1}^{n} z_{2i}z_{2i}'\hat{e}_{2i}^2 \\
\hat{e}_{1i} &= y_{1i} - x_{1i}'\widehat{\beta}_1 \\
\hat{e}_{2i} &= y_{2i} - x_{2i}'\widehat{\beta}_2
\end{aligned}
$$

(b) Since $\widehat{V}_1 \xrightarrow{p} V_1$ and $\widehat{V}_2 \xrightarrow{p} V_2$, under H_0, $W_n \xrightarrow{d} \chi_k^2$, the chi-square distribution with k degrees of freedom.

(c) We select a significance level α (typically 5%) and set the critical value c to be the $1-\alpha$ quantile of the χ_k^2 distribution. We compute W_n, and reject H_0 if $W_n > c$ and do not reject H_0 if $W_n < c$. Equivalently, we compute the p-value $p_n = 1 - F_k(W_n)$, where $F_k(u)$ is the χ_k^2 distribution function, and reject if $p_n < \alpha$.

3.

(a) The model is a linear projection. Thus the appropriate estimator for (β_1, β_2) is least-squares

$$y_i = x_{1i}\widehat{\beta}_1 + x_{2i}\widehat{\beta}_2 + \widehat{e}_i$$

The plug-in estimator for θ is

$$\widehat{\theta} = \widehat{\beta}_1\widehat{\beta}_2$$

(b) By standard asymptotic theory

$$\sqrt{n}\left(\widehat{\beta} - \beta\right) \xrightarrow{d} N(0, V)$$

where $V = Q^{-1}\Omega Q^{-1}$ in the standard notation. By the delta method

$$\sqrt{n}\left(\widehat{\theta} - \theta\right) \xrightarrow{d} N(0, V_\theta)$$

where $V_\theta = h'Vh$ and

$$h = \begin{pmatrix} \dfrac{\partial}{\partial \beta_1}\beta_1\beta_2 \\ \\ \dfrac{\partial}{\partial \beta_2}\beta_1\beta_2 \end{pmatrix} = \begin{pmatrix} \beta_2 \\ \\ \beta_1 \end{pmatrix}$$

It is important that you calculate what h is! You can also write $V_\theta = \beta_2^2 V_{11} + 2\beta_1\beta_2 V_{21} + \beta_1^2 V_{22}$ after partitioning V.

(c) $C_{asy} = \widehat{\theta} \pm 2s(\widehat{\theta})$ where $s(\widehat{\theta}) = \sqrt{n^{-1}\widehat{h}'\widehat{V}\widehat{h}}$, with $\widehat{V}$ the standard estimator for V (write it out) and $\widehat{h}' = (\ \widehat{\beta}_2 \quad \widehat{\beta}_1\ )'$.

(d) The percentile bootstrap method:

i. Draw n iid observations $(y_i^*, x_{1i}^*, x_{2i}^*)$ randomly with replacement from the original sample $\{y_i, x_{1i}, x_{2i}\}$
ii. Compute $\widehat{\beta}_1^*$ and $\widehat{\beta}_2^*$ by least-squares regression of y_i^* on (x_{1i}^*, x_{2i}^*). Set $\widehat{\theta}^* = \widehat{\beta}_1^* \widehat{\beta}_2^*$
iii. Repeat this B times, where B is a large number ($B \geq 1000$). Let $\widehat{\theta}_b^*$, $b = 1, ..., B$ be the results of the B simulations
iv. Compute the $\alpha/2$ and $1 - \alpha/2$ empirical quantiles of the $\widehat{\theta}_b^*$, say $\widehat{q}_{\alpha/2}$ and $\widehat{q}_{1-\alpha/2}$
v. The bootstrap percentile interval is $C_{boot} = [\widehat{q}_{\alpha/2}, \widehat{q}_{1-\alpha/2}]$

4. There are a number of approaches. You can estimate the reduced form by nonparametric series method, either a power series or a spline. The number of terms can be determined by cross-validation on the reduced form. Then the predicted value $\hat{x}_i$ can be used as an instrument to yield $\hat{\beta} = \left(\sum_i \hat{x}_i x_i\right)^{-1} \sum_i \hat{x}_i y_i$. Or, once the series terms for the reduced form have been determined, β can be estimated by GMM. Alternatively, you could use a kernel regression estimator for g. The bandwidth could be selected by cross-validation on the reduced form. Then the predicted value $\hat{x}_i$ could be computed from the kernel estimator and used as an instrument. For either IV approach (series or kernel) the predicted value could be computed using a leave-one-out estimator, yielding a JIVE (jacknife instrumental variables estimator).

   An answer should carefully explain the estimator you recommend, and all steps.