Econ 203C: Systems Models

Problem Set 5

Michael Powell
Department of Economics, UCLA
May 15th, 2006

Question 1:
Consider the Generalized Method of Moments (GMM) estimator defined in Lecture Note 10:

$$\hat{\theta}_n = \underset{\theta \in \Theta}{\arg\min} m_n(\theta)' V_n^{-1} m_n(\theta),$$

where

$$V_n \xrightarrow{p} V, \text{ a non-stochastic non-singular matrix}$$

$$m_n(\theta) = \frac{1}{n}\sum_{i=1}^{n} \varphi(y_i, x_i; \theta), \text{ and}$$

$$E_0[\varphi(y_i, x_i; \theta_0)] = 0.$$

Note: The proofs should follow the same steps as for the proof of consistency for the MLE. In each step make sure to state the exact assumption(s) needed for the statement(s) made.

(a) Show that $\hat{\theta}_n$ is a consistent estimator for θ_0.

Solution: First, define $Q_n(\theta) \equiv m_n(\theta)' V_n^{-1} m_n(\theta)$. In order to prove the consistency of the GMM estimator, we need first to make the following three assumptions:

Assumption 1 (Uniform convergence): Define $m_0(\theta) \equiv E_0[\varphi(Y_i, X_i; \theta)]$. Then

$$\sup_{\theta \in \Theta} |m_n(\theta) - m_0(\theta)| \xrightarrow{p} 0.$$

An immediate implication of assumption 1 is that if we define

$$Q_0(\theta) \equiv E_0[\varphi(Y_i, X_i; \theta)]' V^{-1} E_0[\varphi(Y_i, X_i; \theta)].$$

Then

$$\sup_{\theta \in \Theta} |Q_n(\theta) - Q_0(\theta)| \xrightarrow{p} 0.$$

That is, assumption 1 implies that $Q_n \xrightarrow{p} Q_0$ uniformly in θ.

Assumption 2 (Identification): For all θ such that $||\theta - \theta_0|| \geq \varepsilon$, $Q_0(\theta) - Q_0(\theta_0) > 0$.

Assumption 3 (Continuity): For all n, $Q_n(\theta)$ is continuous in θ.

This assumption allows us to use the Mann-Wald theorem to establish that since $\hat{\theta}_n = \theta_0 + o_p(1)$, $Q_n\left(\hat{\theta}_n\right) = Q_n(\theta_0) + o_p(1)$.

Let $\varepsilon > 0$ be arbitrary. Suppose $||\theta - \theta_0|| \geq \varepsilon$. By Assumption 2, there exists a $\delta > 0$ such that $Q_0(\theta) - Q_0(\theta_0) \geq \delta$. This gives us that

$$\begin{aligned}
\Pr\left[\left|\left|\hat{\theta}_n - \theta_0\right|\right| \geq \varepsilon\right] &\leq \Pr\left[Q_0\left(\hat{\theta}_n\right) - Q_0(\theta_0) \geq \delta\right] \\
&= \Pr\left[Q_0\left(\hat{\theta}_n\right) - Q_n\left(\hat{\theta}_n\right) + Q_n\left(\hat{\theta}_n\right) - Q_0(\theta_0) \geq \delta\right] \\
&= \Pr\left[Q_0\left(\hat{\theta}_n\right) - Q_n\left(\hat{\theta}_n\right) + Q_n(\theta_0) - Q_0(\theta_0) + o_p(1) \geq \delta\right] \\
&\leq \Pr\left[\left\{\left|Q_0\left(\hat{\theta}_n\right) - Q_n\left(\hat{\theta}_n\right)\right| \geq \delta\right\} \cap \{|Q_n(\theta_0) - Q_0(\theta_0)| \geq \delta\}\right] \\
&\leq \Pr\left[2\sup_{\theta \in \Theta}|Q_n(\theta) - Q_0(\theta)| \geq \delta\right]
\end{aligned}$$

By assumption 2, the right hand side goes to zero as $n \to \infty$. Therefore,

$$\lim_{n\to\infty} \Pr\left[\left|\left|\hat{\theta}_n - \theta_0\right|\right| \geq \varepsilon\right] = 0$$

Or $\hat{\theta}_n \xrightarrow{p} \theta_0$. That is, $\hat{\theta}_n$ is consistent for θ_0.

(b) Show that $\hat{\theta}_n$ is asymptotically normal, that is $\sqrt{n}\left(\hat{\theta}_n - \theta_0\right) \xrightarrow{d} N(0, V)$.

Solution: Here, we have that

$$\begin{aligned} Q_n(\theta) &= \left(\frac{1}{n}\sum_{i=1}^{n} \varphi(Y_i, X_i; \theta)\right)' V_n^{-1} \left(\frac{1}{n}\sum_{i=1}^{n} \varphi(Y_i, X_i; \theta)\right) \\ &= m_n(\theta)' V_n^{-1} m_n(\theta) \end{aligned}$$

If we assume:

Assumption 4 (Interiority): $\theta_0 \in int(\Theta)$.

Assumption 5 (Differentiability): Q_n is twice continuously differentiable in θ.

This will guarantee that $\hat{\theta}_n$ will satisfy $\frac{\partial Q_n(\hat{\theta}_n)}{\partial \theta} = 0$. In addition,

$$\begin{aligned} \frac{\partial Q_n(\theta)}{\partial \theta} &= \left(\left[\frac{\partial Q_n(m_n(\theta))}{\partial m_n(\theta)'}\right]\left[\frac{\partial m_n(\theta)}{\partial \theta'}\right]\right)' \\ &= \frac{\partial m_n(\theta)}{\partial \theta}\left[2V_n^{-1} m_n(\theta)\right] \\ &= 2\frac{\partial m_n(\theta)}{\partial \theta}\left(V_n^{-1}\right)' m_n(\theta) \end{aligned}$$

Thus,

$$2\frac{\partial m_n\left(\hat{\theta}_n\right)}{\partial \theta}\left(V_n^{-1}\right)' m_n\left(\hat{\theta}_n\right) = 0$$

Or

$$\frac{\partial m_n\left(\hat{\theta}_n\right)}{\partial \theta}\left(V_n^{-1}\right)' m_n\left(\hat{\theta}_n\right) = 0 \tag{1}$$

Performing a first order Taylor expansion of $m_n(\theta)$ around θ_0, we have

$$m_n\left(\hat{\theta}_n\right) = m_n(\theta_0) + \frac{\partial m_n\left(\tilde{\theta}\right)}{\partial \theta'}\left(\hat{\theta}_n - \theta_0\right)$$

Where $\tilde{\theta} \in \left|\hat{\theta}_n, \theta_0\right|$. Plugging this result into (1), we have

$$\frac{\partial m_n\left(\hat{\theta}_n\right)}{\partial \theta}\left(V_n^{-1}\right)' m_n(\theta_0) + \frac{\partial m_n\left(\hat{\theta}_n\right)}{\partial \theta}\left(V_n^{-1}\right)' \frac{\partial m_n\left(\tilde{\theta}\right)}{\partial \theta'}\left(\hat{\theta}_n - \theta_0\right) = 0$$

Or, if we assume

Assumption 6 (Invertibility): The matrix $\left[\frac{\partial m_n(\theta)}{\partial \theta}\left(V_n^{-1}\right)' \frac{\partial m_n(\theta)}{\partial \theta'}\right]$ is nonsingular for all θ for all n.

Therefore,

$$\sqrt{n}\left(\hat{\theta}_n - \theta_0\right) = \left[-\frac{\partial m_n\left(\hat{\theta}_n\right)}{\partial \theta}\left(V_n^{-1}\right)' \frac{\partial m_n\left(\tilde{\theta}\right)}{\partial \theta'}\right]^{-1} \frac{\partial m_n\left(\hat{\theta}_n\right)}{\partial \theta}\left(V_n^{-1}\right)' \sqrt{n} m_n(\theta_0)$$

Next note that by the central limit theorem, assuming

Assumption 7 (i.i.d. data): (Y_i, X_i) are i.i.d.

Assumption 8 (Moment Conditions):

$$Var_0 (\varphi (Y_i, X_i; \theta_0)) = E_0 \left[\varphi (Y_i, X_i; \theta_0) \varphi (Y_i, X_i; \theta_0)' \right] < +\infty$$

Then,

$$\begin{aligned} \sqrt{n} m_n (\theta_0) &= \sqrt{n} \left(\frac{1}{n} \sum_{i=1}^{n} \varphi (Y_i, X_i; \theta_0) \right) \\ &= \sqrt{n} \left(\frac{1}{n} \sum_{i=1}^{n} \varphi (Y_i, X_i; \theta_0) - \underbrace{E_0 [\varphi (Y_i, X_i; \theta_0)]}_{=0} \right) \\ &\xrightarrow{d} N (0, \Lambda (\theta_0)) \end{aligned}$$

Where

$$\Lambda (\theta_0) = E_0 \left[\varphi (Y_i, X_i; \theta_0) \varphi (Y_i, X_i; \theta_0)' \right]$$

If we further assume that

Assumption 9 (Uniform Upper Bound):

$$E_0 \left[\sup_{\theta \in \Theta} \left\| \frac{\partial m_n (\theta)}{\partial \theta} \right\| \right] < +\infty$$

Then by the uniform law of large numbers, we will have

$$\frac{\partial m_n (\theta)}{\partial \theta} = \frac{1}{n} \sum_{i=1}^{n} \frac{\partial \varphi (Y_i, X_i; \theta)}{\partial \theta} \xrightarrow{p} E_0 \left[\frac{\partial \varphi (Y_i, X_i; \theta)}{\partial \theta} \right]$$

Uniformly in θ. Therefore, we have

$$\begin{aligned} \frac{\partial m_n \left(\hat{\theta}_n \right)}{\partial \theta} &\xrightarrow{p} E_0 \left[\frac{\partial \varphi (Y_i, X_i; \theta_0)}{\partial \theta} \right] \\ \frac{\partial m_n \left(\tilde{\theta} \right)}{\partial \theta'} &\xrightarrow{p} E_0 \left[\frac{\partial \varphi (Y_i, X_i; \theta_0)}{\partial \theta'} \right] \end{aligned}$$

Define $A (\theta_0) \equiv E_0 \left[\frac{\partial \varphi (Y_i, X_i; \theta_0)}{\partial \theta} \right]$. Since $V_n \xrightarrow{p} V$, by the Mann-Wald theorem, we have that $V_n^{-1} \xrightarrow{p} V^{-1}$. Putting this together with Slutsky's theorem and another application of the continuity theorem, we have

$$\left[-\frac{\partial m_n \left(\hat{\theta}_n \right)}{\partial \theta} \left(V_n^{-1} \right)' \frac{\partial m_n \left(\tilde{\theta} \right)}{\partial \theta'} \right]^{-1} \frac{\partial m_n \left(\hat{\theta}_n \right)}{\partial \theta} \left(V_n^{-1} \right)' \xrightarrow{p} \left[-A (\theta_0) \left(V^{-1} \right)' A (\theta_0)' \right]^{-1} A (\theta_0) \left(V^{-1} \right)'$$

Define this quantity to be $B (\theta_0)$.

Using Slutsky's theorem once again, we have

$$\sqrt{n} \left(\hat{\theta}_n - \theta_0 \right) \xrightarrow{d} N \left(0, B (\theta_0) \Lambda (\theta_0) B (\theta_0)' \right)$$

Where

$$\begin{aligned} \Lambda (\theta_0) &= E_0 \left[\varphi (Y_i, X_i; \theta_0) \varphi (Y_i, X_i; \theta_0)' \right] \\ B (\theta_0) &= \left[-A (\theta_0) \left(V^{-1} \right)' A (\theta_0)' \right]^{-1} A (\theta_0) \left(V^{-1} \right)' \\ A (\theta_0) &= E_0 \left[\frac{\partial \varphi (Y_i, X_i; \theta_0)}{\partial \theta} \right] \end{aligned}$$

Question 2:
Consider the following model:

$$y_i^* = z_i'\gamma + u_i,$$

for $i = 1, \ldots, n$ where u_i conditional on z_i has a normal distribution, that is,

$$u_i | z_i \sim N\left(0, \sigma_v^2\right).$$

Define

$$y_i = \begin{cases} 1 & \text{if } y_i^* > 0, \\ 0 & \text{otherwise.} \end{cases}$$

(a) Show that $\Pr(Y_i = 1 | Z_i) = \Phi\left(\frac{Z_i'\gamma}{\sigma_v}\right)$. Explain why γ and σ_v cannot be separately identified. Denote $\theta = \frac{\gamma}{\sigma_v}$, and let θ_0 be the population parameter.

Solution: First note that if $u_i | Z_i \sim N\left(0, \sigma_v^2\right)$, then $-u_i | Z_i \sim N\left(0, \sigma_v^2\right)$. Thus

$$\begin{aligned} \Pr[Y_i = 1 | Z_i] &= \Pr[Y_i^* > 0 | Z_i] \\ &= \Pr[Z_i'\gamma + u_i > 0 | Z_i] \\ &= \Pr[-u_i < Z_i'\gamma | Z_i] \\ &= \Pr\left[-\frac{u_i}{\sigma_v} < \frac{Z_i'\gamma}{\sigma_v} \middle| Z_i \right] \\ &= \Phi\left(\frac{Z_i'\gamma}{\sigma_v}\right) \end{aligned}$$

Suppose we have some alternative model in which

$$Y_i^* = Z_i'\gamma^* + \nu_i$$

Where

$$-\nu_i | Z_i \sim N\left(0, (\sigma_v^*)^2\right)$$

And

$$Y_i = \begin{cases} 1 & \text{if } Y_i^* > 0, \\ 0 & \text{otherwise.} \end{cases}$$

Suppose $\gamma^* \neq \gamma$ and $\sigma_v^* \neq \sigma_v$ but $\frac{\gamma^*}{\sigma_v^*} = \frac{\gamma}{\sigma_v}$. Then

$$\begin{aligned} \Pr[Y_i = 1 | Z_i] &= \Pr[Z_i'\gamma^* + \nu_i > 0 | Z_i] \\ &= \Pr[-\nu_0 < Z_i'\gamma^* | Z_i] \\ &= \Pr\left[-\frac{\nu_0}{\sigma_v^*} < \frac{Z_i'\gamma^*}{\sigma_v^*} \middle| Z_i \right] \\ &= \Phi\left(\frac{Z_i'\gamma^*}{\sigma_v^*}\right) \end{aligned}$$

Thus,

$$\Phi\left(\frac{Z_i'\gamma}{\sigma_v}\right) = \Phi\left(\frac{Z_i'\gamma^*}{\sigma_v^*}\right)$$

These two models are therefore observationally equivalent. That is, we cannot tell the difference between the model which has the parameters (γ, σ_v) and the one that has the parameters (γ^*, σ_v^*). Define $\theta = \frac{\gamma}{\sigma_v}$. Then

$$\Pr[Y_i = 1 | Z_i] = \Phi(Z_i'\theta)$$

(b) Provide the normalized log likelihood function for θ.

Solution: Since $Y_i|\, Z_i \sim Bernoulli\left(\Pr\left[Y_i = 1|\, Z_i\right] = \Phi\left(Z_i'\theta\right)\right)$, we have

$$f_{Y_i}\left(y_i|\, Z_i\right) = \Phi\left(Z_i'\theta\right)^{y_i}\left[1 - \Phi\left(Z_i'\theta\right)\right]^{1-y_i}$$

The likelihood function is therefore

$$L\left(\theta|\, Y_i, Z_i\right) = \prod_{i=1}^{n}\left\{\Phi\left(Z_i'\theta\right)^{Y_i}\left[1 - \Phi\left(Z_i'\theta\right)\right]^{1-Y_i}\right\}$$

Taking logs,

$$\log L\left(\theta|\, Y_i, Z_i\right) = \sum_{i=1}^{n}\left\{Y_i \log\Phi\left(Z_i'\theta\right) + \left(1 - Y_i\right)\log\left[1 - \Phi\left(Z_i'\theta\right)\right]\right\}$$

And the normalized log likelihood function is therefore,

$$\frac{1}{n}\log L\left(\theta|\, Y_i, Z_i\right) = \frac{1}{n}\sum_{i=1}^{n}\left\{Y_i \log\Phi\left(Z_i'\theta\right) + \left(1 - Y_i\right)\log\left[1 - \Phi\left(Z_i'\theta\right)\right]\right\}$$

(c) Provide the first order conditions for $\hat{\theta}_n$, the estimator for θ_0.
Solution: Taking first order conditions, we have

$$(\theta) : \frac{1}{n}\sum_{i=1}^{n}\left\{\frac{Y_i}{\Phi\left(Z_i'\hat{\theta}_n\right)}\phi\left(Z_i'\hat{\theta}_n\right)Z_i - \frac{1 - Y_i}{1 - \Phi\left(Z_i'\hat{\theta}_n\right)}\phi\left(Z_i'\hat{\theta}_n\right)Z_i\right\} = 0$$

Rearranging,

$$0 = \frac{1}{n}\sum_{i=1}^{n}\left\{\frac{Z_iY_i\phi\left(Z_i'\hat{\theta}_n\right) - Z_iY_i\phi\left(Z_i'\hat{\theta}_n\right)\Phi\left(Z_i'\hat{\theta}_n\right) + Z_iY_i\phi\left(Z_i'\hat{\theta}_n\right)\Phi\left(Z_i'\hat{\theta}_n\right) - Z_i\phi\left(Z_i'\hat{\theta}_n\right)\Phi\left(Z_i'\hat{\theta}_n\right)}{\Phi\left(Z_i'\hat{\theta}_n\right)\left[1 - \Phi\left(Z_i'\hat{\theta}_n\right)\right]}\right\}$$

Or

$$\frac{1}{n}\sum_{i=1}^{n}\left\{\frac{Y_i - \Phi\left(Z_i'\hat{\theta}_n\right)}{\Phi\left(Z_i'\hat{\theta}_n\right)\left[1 - \Phi\left(Z_i'\hat{\theta}_n\right)\right]}\phi\left(Z_i'\hat{\theta}_n\right)Z_i\right\} = 0$$

(d) Show that the estimator obtained by solving the first order conditions in (c) can be viewed as a Method of Moments estimator.
Solution: Define

$$\varphi\left(Y_i, Z_i; \theta\right) = \frac{Y_i - \Phi\left(Z_i'\theta\right)}{\Phi\left(Z_i'\theta\right)\left[1 - \Phi\left(Z_i'\theta\right)\right]}\phi\left(Z_i'\theta\right)Z_i$$

Here, we note that $\dim\left(Z_i\right) = \dim\left(\theta\right)$ and there is no need to use generalized method of moments. As shown in part (c), $\hat{\theta}_n$ solves

$$\frac{1}{n}\sum_{i=1}^{n}\varphi\left(Y_i, Z_i; \hat{\theta}_n\right) = 0$$

Assuming $\left(Y_i, Z_i\right)$ are i.i.d. and $E_0\left[\sup_{\theta\in\Theta}\left|\varphi\left(Y_i, Z_i; \theta\right)\right|\right] < +\infty$. Then by the uniform law of large numbers

$$\frac{1}{n}\sum_{i=1}^{n}\varphi\left(Y_i, Z_i; \theta\right) \xrightarrow{p} E_0\left[\varphi\left(Y_i, Z_i; \theta\right)\right]$$

Uniformly in θ. This gives us that $\hat{\theta}_n \xrightarrow{p} \theta_0$ where θ_0 is such that $E_0\left[\varphi\left(Y_i, Z_i; \theta\right)\right] = 0$ if and only if $\theta = \theta_0$. It thus only remains to show that $E_0\left[\varphi\left(Y_i, Z_i; \theta\right)\right] = 0$ if and only if $\theta = \theta_0$

$$\begin{aligned}
E_0\left[\varphi\left(Y_i, Z_i; \theta\right)\right] &= E_0\left[\frac{Y_i - \Phi\left(Z_i'\theta\right)}{\Phi\left(Z_i'\theta\right)\left[1-\Phi\left(Z_i'\theta\right)\right]}\phi\left(Z_i'\theta\right) Z_i\right] \\
&= E\left[E_0\left[\left.\frac{Y_i - \Phi\left(Z_i'\theta\right)}{\Phi\left(Z_i'\theta\right)\left[1-\Phi\left(Z_i'\theta\right)\right]}\phi\left(Z_i'\theta\right) Z_i\right| Z_i\right]\right] \\
&= E\left[\left\{E_0\left[\left.Y_i\right| Z_i\right] - \Phi\left(Z_i'\theta\right)\right\}\frac{\phi\left(Z_i'\theta\right) Z_i}{\Phi\left(Z_i'\theta\right)\left[1-\Phi\left(Z_i'\theta\right)\right]}\right]
\end{aligned}$$

Next, we have that $E_0\left[\left.Y_i\right| Z_i\right] = \Pr\left[Y_i = 1 \middle| Z_i, \theta_0\right] = \Phi\left(Z_i'\theta_0\right)$. Therefore,

$$E_0\left[\varphi\left(Y_i, Z_i; \theta\right)\right] = E\left[\left\{\Phi\left(Z_i'\theta_0\right) - \Phi\left(Z_i'\theta\right)\right\}\frac{\phi\left(Z_i'\theta\right) Z_i}{\Phi\left(Z_i'\theta\right)\left[1-\Phi\left(Z_i'\theta\right)\right]}\right] = 0$$

If and only if $\theta = \theta_0$. That is, $\hat{\theta}_n$ is a method of moments estimator.

Question 3:
Consider the linear model given by

$$\begin{aligned} y_i &= x_i' \beta_0 + \varepsilon_i \\ E_0 [\varepsilon_i | x_i] &= 0, \end{aligned}$$

for $i = 1, \ldots, n$, where x_i is a $K \times 1$ vector of regressors, that is, $x_i = \begin{bmatrix} x_{1i} \\ \vdots \\ x_{Ki} \end{bmatrix}$.

Define the vector

$$z_i = \left(x_i', x_{1i}^2, \ldots, x_{Ki}^2, x_{1i}x_{2i}, \ldots, x_{1i}x_{Ki}\right)',$$

and the following moment functions:

$$\begin{aligned} \varphi_1 (y_i, x_i; \beta) &= (y_i - x_i'\beta) x_i, \text{ and} \\ \varphi_2 (y_i, z_i; \beta) &= (y_i - x_i'\beta) z_i. \end{aligned}$$

(a) Define the population parameter vector β_0.
Solution: Since $E_0 [\varepsilon_i | X_i] = 0$, we have that β_0 solves

$$E_0 [Y_i | X_i] = X_i'\beta_0$$

Or

$$E_0 [Y_i - X_i'\beta_0 | X_i] = 0$$

(b) Show that when evaluated at the population value β_0,

$$\begin{aligned} E_0 [\varphi_1 (y_i, x_i; \beta_0)] &= 0, \text{ and} \\ E_0 [\varphi_2 (y_i, z_i; \beta_0)] &= 0. \end{aligned}$$

Solution: From part (a), we have that

$$E_0 [Y_i - X_i'\beta_0 | X_i] = 0$$

This gives us

$$\begin{aligned} 0 &= E_0 [Y_i - X_i'\beta_0 | X_i] X_i = E_0 [(Y_i - X_i'\beta_0) X_i | X_i] = E_0 [\varphi_1 (Y_i, X_i; \beta_0)] \\ 0 &= E_0 [Y_i - X_i'\beta_0 | X_i] Z_i = E_0 [(Y_i - X_i'\beta_0) Z_i | X_i] = E_0 [\varphi_2 (Y_i, Z_i; \beta_0)] \end{aligned}$$

Since Z_i is a function of X_i.

(c) Define the optimal GMM estimators for β_0 based on $\varphi_1 (y_i, x_i; \beta)$ and $\varphi_2 (y_i, z_i; \beta)$, say $\hat{\beta}_n^1$, and $\hat{\beta}_n^2$, respectively.
Solution: Based on $\varphi_1 (Y_i, X_i; \beta)$, we have that β_0 solves

$$E_0 [\varphi_1 (Y_i, X_i; \beta_0)] = 0$$

We define the GMM estimator by

$$\begin{aligned} \hat{\beta}_n^1 &= \arg\min_{\beta \in B} \left[\frac{1}{n} \sum_{i=1}^n \varphi_1 (Y_i, X_i; \beta)\right]' V_n^{-1} \left[\frac{1}{n} \sum_{i=1}^n \varphi_1 (Y_i, X_i; \beta)\right] \\ &= \arg\min_{\beta \in B} \left[\frac{1}{n} \sum_{i=1}^n (X_i Y_i - X_i X_i'\beta)\right]' V_n^{-1} \left[\frac{1}{n} \sum_{i=1}^n (X_i Y_i - X_i X_i'\beta)\right] \\ &= \arg\min_{\beta \in B} (X'Y - X'X\beta)' V_n^{-1} (X'Y - X'X\beta) \\ &= \arg\min_{\beta \in B} \left(Y'XV_n^{-1}X'Y - \beta'X'XV_n^{-1}X'Y - Y'XV_n^{-1}X'X\beta + \beta'X'XV_n^{-1}X'X\beta\right) \\ &= \arg\min_{\beta \in B} \left(Y'XV_n^{-1}X'Y - 2\beta'X'XV_n^{-1}X'Y + \beta'X'XV_n^{-1}X'X\beta\right) \end{aligned}$$

Taking first order conditions,

$$(\beta) : -2X'XV_n^{-1}X'Y + 2X'XV_n^{-1}X'X\hat{\beta}_n^1 = 0$$

Rearranging

$$X'XV_n^{-1}X'X\hat{\beta}_n^1 = X'XV_n^{-1}X'Y$$

If we assume that $X'X$ is invertible, then

$$\begin{aligned}
\hat{\beta}_n^1 &= \left(X'XV_n^{-1}X'X\right)^{-1} X'XV_n^{-1}X'Y \\
&= \left(X'X\right)^{-1} V_n \left(X'X\right)^{-1} X'XV_n^{-1}X'Y \\
&= \left(X'X\right)^{-1} V_nV_n^{-1}X'Y \\
&= \left(X'X\right)^{-1} X'Y
\end{aligned}$$

Which is the standard OLS estimator. This shows that OLS is a GMM estimator.
Based on $\varphi_2\left(Y_i, Z_i; \beta\right)$, we have that β_0 solves

$$E_0\left[\varphi_2\left(Y_i, Z_i; \beta_0\right)\right] = 0$$

Since in general, $\dim\left(Z_i\right) = \ell > k = \dim\left(\beta\right)$, the generalized method of moments estimator would be such that

$$\hat{\beta}_n^2 = \arg\min_{\beta\in B} \underbrace{\left[\frac{1}{n}\sum_{i=1}^n \varphi_2\left(Y_i, Z_i; \beta\right)\right]'}_{k\times\ell} \underbrace{V_n^{-1}}_{\ell\times\ell} \underbrace{\left[\frac{1}{n}\sum_{i=1}^n \varphi_2\left(Y_i, Z_i; \beta\right)\right]}_{\ell\times k}$$

Where $V_n^{-1} \xrightarrow{p} V^{-1}$ which is non-stochastic. Plugging in the values

$$\begin{aligned}
\hat{\beta}_n^2 &= \arg\min_{\beta\in B} \left[\frac{1}{n}\sum_{i=1}^n \left(Z_iY_i - Z_iX_i\beta\right)\right]' V_n^{-1} \left[\frac{1}{n}\sum_{i=1}^n \left(Z_iY_i - Z_iX_i\beta\right)\right] \\
&= \arg\min_{\beta\in B} \left(\frac{1}{n}Z'Y - \frac{1}{n}Z'X\beta\right)' V_n^{-1} \left(\frac{1}{n}Z'Y - Z'X\beta\right) \\
&= \arg\min_{\beta\in B} \left(Z'Y - Z'X\beta\right)' V_n^{-1} \left(Z'Y - Z'X\beta\right) \\
&= \arg\min_{\beta\in B} \left(Y'Z - \beta'X'Z\right) V_n^{-1} \left(Z'Y - Z'X\beta\right) \\
&= \arg\min_{\beta\in B} Y'ZV_n^{-1}Z'Y - \beta'X'ZV_n^{-1}Z'Y - Y'ZV_n^{-1}Z'X\beta + \beta'X'ZV_n^{-1}Z'X\beta \\
&= \arg\min_{\beta\in B} Y'ZV_n^{-1}Z'Y - 2\beta'X'ZV_n^{-1}Z'Y + \beta'X'ZV_n^{-1}Z'X\beta
\end{aligned}$$

Taking first order conditions

$$(\beta) : -2X'ZV_n^{-1}Z'Y + 2X'ZV_n^{-1}Z'X\hat{\beta}_n^2 = 0$$

Rearranging

$$X'ZV_n^{-1}Z'X\hat{\beta}_n^2 = X'ZV_n^{-1}Z'Y$$

If we assume that $\left(X'ZV_n^{-1}Z'X\right)^{-1}$ exists, we have

$$\hat{\beta}_n^2 = \left(X'ZV_n^{-1}Z'X\right)^{-1} X'ZV_n^{-1}Z'Y$$

This is nothing other than the weighted IV estimator with $\Pi = X'ZV_n^{-1}$. This shows that the weighted IV estimator is a GMM estimator. Similarly, if we let $V_n^{-1} = (Z'Z)^{-1}$, then

$$\begin{aligned}
\hat{\beta}_n^2 &= \left(X'Z\left(Z'Z\right)^{-1}X\right)^{-1} X'Z\left(Z'Z\right)^{-1}Z'Y \\
&= \left(X'P_Z X\right)^{-1} X'P_Z Y = \hat{\beta}_{2SLS}
\end{aligned}$$

That is, the $2SLS$ estimator is a GMM estimator.

(d) Provide the asymptotic covariance matrices for $\hat{\beta}_n^1$ and $\hat{\beta}_n^2$, from (c).
Solution: From question 1, part (b), we have that

$$\sqrt{n}\left(\hat{\beta}_n^1 - \beta_0\right) \xrightarrow{d} N\left(0, B\left(\beta_0\right)\Lambda\left(\beta_0\right)B\left(\beta_0\right)'\right)$$

Where

$$\begin{aligned}
\Lambda\left(\beta_0\right) &= E_0\left[\varphi_1\left(Y_i, X_i; \beta_0\right)\varphi_1\left(Y_i, X_i; \beta_0\right)'\right] \\
B\left(\beta_0\right) &= \left[-A\left(\beta_0\right)\left(V^{-1}\right)' A\left(\beta_0\right)'\right]^{-1} A\left(\beta_0\right)\left(V^{-1}\right)' \\
A\left(\beta_0\right) &= E_0\left[\frac{\partial\varphi_1\left(Y_i, X_i; \beta_0\right)}{\partial\beta}\right]
\end{aligned}$$

Plugging in the values, we have

$$\begin{aligned}
\Lambda\left(\beta_0\right) &= E_0\left[\left(Y_i - X_i'\beta_0\right)X_iX_i'\left(Y_i - X_i'\beta_0\right)'\right] \\
&= E_0\left[\varepsilon_i X_i X_i' \varepsilon_i'\right] \\
&= E_0\left[X_iX_i'\varepsilon_i^2\right]
\end{aligned}$$

$$A\left(\beta_0\right) = E_0\left[X_iX_i'\right] \equiv \Sigma_{XX}$$

$$\begin{aligned}
B\left(\beta_0\right) &= \left[-\Sigma_{XX}\left(V^{-1}\right)'\Sigma_{XX}'\right]^{-1}\Sigma_{XX}\left(V^{-1}\right)' \\
&= -\Sigma_{XX}^{-1}
\end{aligned}$$

This gives us

$$B\left(\beta_0\right)\Lambda\left(\beta_0\right)B\left(\beta_0\right) = \Sigma_{XX}^{-1}E_0\left[X_iX_i'\varepsilon_i^2\right]\Sigma_{XX}^{-1}$$

And therefore

$$\sqrt{n}\left(\hat{\beta}_n^1 - \beta_0\right) \xrightarrow{d} N\left(0, \Sigma_{XX}^{-1}E_0\left[X_iX_i'\varepsilon_i^2\right]\Sigma_{XX}^{-1}\right)$$

Which is the standard sandiwch form covariance matrix.
For $\hat{\beta}_n^2$, we have

$$\sqrt{n}\left(\hat{\beta}_n^2 - \beta_0\right) \xrightarrow{d} N\left(0, B\left(\beta_0\right)\Lambda\left(\beta_0\right)B\left(\beta_0\right)'\right)$$

Where

$$\begin{aligned}
\Lambda\left(\beta_0\right) &= E_0\left[\varphi_2\left(Y_i, X_i; \beta_0\right)\varphi_2\left(Y_i, X_i; \beta_0\right)'\right] \\
B\left(\beta_0\right) &= \left[-A\left(\beta_0\right)\left(V^{-1}\right)' A\left(\beta_0\right)'\right]^{-1} A\left(\beta_0\right)\left(V^{-1}\right)' \\
A\left(\beta_0\right) &= E_0\left[\frac{\partial\varphi_2\left(Y_i, X_i; \beta_0\right)}{\partial\beta}\right]
\end{aligned}$$

Plugging in the values,

$$\begin{aligned}\Lambda(\beta_0) &= E_0\left[(Y_i - X_i'\beta_0) Z_i Z_i' (Y_i - X_i'\beta_0)'\right] \\ &= E_0\left[\varepsilon_i Z_i Z_i' \varepsilon_i'\right] = E_0\left[Z_i Z_i' \varepsilon_i^2\right]\end{aligned}$$

$$A(\beta_0) = E_0\left[Z_i X_i'\right] \equiv \Sigma_{XZ}$$

$$B(\beta_0) = \left[-\Sigma_{XZ}\left(V^{-1}\right)'\Sigma_{XZ}'\right]^{-1}\Sigma_{XZ}\left(V^{-1}\right)'$$

Thus,

$$B(\beta_0)\Lambda(\beta_0)B(\beta_0)' = \left[\Sigma_{XZ}\left(V^{-1}\right)'\Sigma_{XZ}'\right]^{-1}\Sigma_{XZ}\left(V^{-1}\right)'E_0\left[Z_i Z_i'\varepsilon_i^2\right]V^{-1}\Sigma_{XZ}'\left[\Sigma_{XZ}V^{-1}\Sigma_{XZ}'\right]^{-1}$$

Which gives us

$$\sqrt{n}\left(\hat{\beta}_n^2 - \beta_0\right) \xrightarrow{d} N\left(0, \left[\Sigma_{XZ}\left(V^{-1}\right)'\Sigma_{XZ}'\right]^{-1}\Sigma_{XZ}\left(V^{-1}\right)'E_0\left[Z_i Z_i'\varepsilon_i^2\right]V^{-1}\Sigma_{XZ}'\left[\Sigma_{XZ}V^{-1}\Sigma_{XZ}'\right]^{-1}\right)$$

(*e*) Provide a consistent estimator for the asymptotic covariance matrix of $\hat{\beta}_n^2$ from (*d*). Show that the proposed estimator is, in fact, a consistent estimator for its population counterpart, using the assumptions put forward in Lecture Note 10.

Solution: Recall in part (*d*) that

$$\sqrt{n}\left(\hat{\beta}_n^2 - \beta_0\right) \xrightarrow{d} N\left(0, \Lambda_0(\beta_0)\right)$$

Where

$$\Lambda_0(\beta_0) = \left[\Sigma_{XZ}\left(V^{-1}\right)'\Sigma_{XZ}'\right]^{-1}\Sigma_{XZ}\left(V^{-1}\right)'E_0\left[Z_i Z_i'\varepsilon_i^2\right]V^{-1}\Sigma_{XZ}'\left[\Sigma_{XZ}V^{-1}\Sigma_{XZ}'\right]^{-1}$$

I claim that

$$\begin{aligned}\hat{\Lambda}_n\left(\hat{\beta}_n^2\right) \equiv & \left[\left(\frac{1}{n}X'Z\right)\left(V_n^{-1}\right)'\left(\frac{1}{n}Z'X\right)\right]^{-1}\left(\frac{1}{n}X'Z\right)\left(V_n^{-1}\right)'\left[\frac{1}{n}\sum_{i=1}^{n}Z_i Z_i'\hat{\varepsilon}_i^2\right]\cdot \\ & V_n^{-1}\left(\frac{1}{n}Z'X\right)\left[\left(\frac{1}{n}X'Z\right)V_n^{-1}\left(\frac{1}{n}Z'X\right)\right]^{-1}\end{aligned}$$

Where $\hat{\varepsilon}_i^2 \equiv \left(Y_i - X_i'\hat{\beta}_n^2\right)'\left(Y_i - X_i'\hat{\beta}_n^2\right)$. In order to establish this, first note that assuming (X_i, Z_i) are i.i.d. and that $E\left[|X_i Z_i'|\right] < +\infty$, by the weak law of large numbers,

$$\begin{aligned}\frac{1}{n}X'Z &= \frac{1}{n}\sum_{i=1}^{n}X_i Z_i' \xrightarrow{p} E\left[X_i Z_i'\right] \equiv \Sigma_{XZ} \\ \frac{1}{n}Z'X &= \frac{1}{n}\sum_{i=1}^{n}Z_i X_i' \xrightarrow{p} E\left[Z_i X_i'\right] \equiv \Sigma_{XZ}'\end{aligned}$$

Also, we know that $V_n \xrightarrow{p} V$, so by the continuity theorem, $V_n^{-1} \xrightarrow{p} V^{-1}$. Next note that if we assume (Z_i, ε_i) are i.i.d. and that $E\left[\left|Z_i Z_i'\varepsilon_i^2\right|\right] < +\infty$, then

$$\frac{1}{n}\sum_{i=1}^{n}Z_i Z_i'\varepsilon_i^2 \xrightarrow{p} E\left[Z_i Z_i'\varepsilon_i^2\right]$$

By the weak law of large numbers. Also, since

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n} Z_i Z_i' \hat{\varepsilon}_i^2 - \frac{1}{n}\sum_{i=1}^{n} Z_i Z_i' \varepsilon_i^2 &= \frac{1}{n}\sum_{i=1}^{n} Z_i Z_i' \left(\hat{\varepsilon}_i^2 - \varepsilon_i^2\right) \\
&= \frac{1}{n}\sum_{i=1}^{n} Z_i Z_i' \left[\left(Y_i - X_i'\hat{\beta}_n^2\right)' \left(Y_i - X_i'\hat{\beta}_n^2\right) - \left(Y_i - X_i'\beta_0\right)' \left(Y_i - X_i'\beta_0\right)\right]
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{n}\sum_{i=1}^{n} Z_i Z_i' \left[\left(Y_i'Y_i - 2\left(\hat{\beta}_n^2\right)' X_i Y_i + \left(\hat{\beta}_n^2\right)' X_i X_i' \hat{\beta}_n^2\right) - \left(Y_i'Y_i - 2\beta_0' X_i Y_i + \beta_0' X_i X_i' \beta_0\right)\right] \\
&= \frac{1}{n}\sum_{i=1}^{n} Z_i Z_i' \left[2\left(\beta_0 - \hat{\beta}_n^2\right)' X_i Y_i + \left(\hat{\beta}_n^2\right)' X_i X_i' \hat{\beta}_n^2 - \beta_0' X_i X_i' \beta_0\right] \\
&= \frac{2}{n}\sum_{i=1}^{n} Z_i Z_i' \left(\beta_0 - \hat{\beta}_n^2\right) X_i Y_i + \frac{1}{n}\sum_{i=1}^{n} \left(\hat{\beta}_n^2\right)' X_i X_i' \hat{\beta}_n^2 - \frac{1}{n}\sum_{i=1}^{n} \beta_0' X_i X_i' \beta_0 \\
&= \frac{2}{n}\sum_{i=1}^{n} Z_i Z_i' \left(\beta_0 - \hat{\beta}_n^2\right) X_i Y_i + \frac{1}{n}\sum_{i=1}^{n} \left(\hat{\beta}_n^2\right)' X_i X_i' \hat{\beta}_n^2 - \frac{1}{n}\sum_{i=1}^{n} \left(\hat{\beta}_n^2\right)' X_i X_i' \beta_0 \\
&\quad + \frac{1}{n}\sum_{i=1}^{n} \left(\hat{\beta}_n^2\right)' X_i X_i' \beta_0 - \frac{1}{n}\sum_{i=1}^{n} \beta_0' X_i X_i' \beta_0 \\
&= \frac{2}{n}\sum_{i=1}^{n} Z_i Z' \left(\beta_0 - \hat{\beta}_n^2\right) X_i Y_i + \frac{1}{n}\sum_{i=1}^{n} \left(\hat{\beta}_n^2\right)' X_i X_i' \left(\hat{\beta}_n^2 - \beta_0\right) + \frac{1}{n}\sum_{i=1}^{n} \left(\hat{\beta}_n^2 - \beta_0\right)' X_i X_i' \beta_0
\end{aligned}$$

Where, by the Cauchy-Schwarz inequality,

$$\left|\frac{2}{n}\sum_{i=1}^{n} Z_i Z_i' \left(\beta_0 - \hat{\beta}_n^2\right) X_i Y_i\right| \leq ||Z_i'Z_i||^{\frac{1}{2}} \left|\left|\beta_0 - \hat{\beta}_n^2\right|\right|^{\frac{1}{2}} ||X_i Y_i||^{\frac{1}{2}} \xrightarrow{p} 0$$

$$\left|\frac{1}{n}\sum_{i=1}^{n} \left(\hat{\beta}_n^2\right)' X_i X_i' \left(\hat{\beta}_n^2 - \beta_0\right)\right| \leq \left|\left|\hat{\beta}_n^2\right|\right|^{\frac{1}{2}} ||X_i X_i'||^{\frac{1}{2}} \left|\left|\hat{\beta}_n^2 - \beta_0\right|\right|^{\frac{1}{2}} \xrightarrow{p} 0$$

And

$$\left|\frac{1}{n}\sum_{i=1}^{n} \left(\hat{\beta}_n^2 - \beta_0\right)' X_i X_i' \beta_0\right| \leq \left|\left|\hat{\beta}_n^2 - \beta_0\right|\right|^{\frac{1}{2}} ||X_i X_i'||^{\frac{1}{2}} ||\beta_0||^{\frac{1}{2}} \xrightarrow{p} 0$$

Since $\beta_0 \xrightarrow{p} \hat{\beta}_n^2$. This gives us that

$$\frac{1}{n}\sum_{i=1}^{n} Z_i Z_i' \hat{\varepsilon}_i^2 - \frac{1}{n}\sum_{i=1}^{n} Z_i Z_i' \varepsilon_i^2 \xrightarrow{p} 0$$

And therefore,

$$\frac{1}{n}\sum_{i=1}^{n} Z_i Z_i' \hat{\varepsilon}_i^2 \xrightarrow{p} E\left[Z_i Z_i' \varepsilon_i^2\right]$$

By Slutsky's theorem and the continuity theorem, then, we have that

$$\hat{\Lambda}_n\left(\hat{\beta}_n^2\right) \xrightarrow{p} \Lambda_0\left(\beta_0\right)$$

That is, it is a consistent estimator.

(f) Provide the asymptotic distribution for the GMM estimator based on $\varphi_2(y_i, z_i; \beta)$ when the weight matrix is $V_n = I$.

Solution: When $V_n = I$ for all n, we have that $V = I$ and $V^{-1} = I$. Therefore,

$$\begin{aligned} & \left[\Sigma_{XZ}\left(V^{-1}\right)'\Sigma'_{XZ}\right]^{-1}\Sigma_{XZ}\left(V^{-1}\right)' E_0\left[Z_i Z_i' \varepsilon_i^2\right] V^{-1}\Sigma'_{XZ}\left[\Sigma_{XZ}V^{-1}\Sigma'_{XZ}\right]^{-1} \\ = \quad & \left[\Sigma_{XZ}\Sigma'_{XZ}\right]^{-1}\Sigma_{XZ}E_0\left[Z_i Z_i' \varepsilon_i^2\right]\Sigma'_{XZ}\left[\Sigma_{XZ}\Sigma'_{XZ}\right]^{-1} \end{aligned}$$

Question 4:

For this exercise we use the results of question 3 above. In the excel file **ps5q4.xls** you are provided with the data for this exercise. There are six variables in columns 1 through 6 of the file, corresponding to $y, x_1, \ldots, x_5$, respectively, where $x_{1i} = 1$ for all $i = 1, \ldots, n$.

(a) The estimator in this part is based on $\varphi_1(y_i, x_i; \beta)$ from Question 3.

(1) Provide an estimate for β_0 and the corresponding standard errors, based on $V_n = I$.

Solution: Using $\varphi_1(Y_i; X_i, \beta)$ and $V_n = I$

$$\hat{\beta}_n^1 = \arg\min_{\beta \in B} Q_n^1(\beta)$$

Where

$$Q_n^1(\beta) = m_n^1(\beta)' m_n^1(\beta)$$

The first order conditions are

$$(\beta): \quad 2\frac{\partial m_n^1\left(\hat{\beta}_n^1\right)}{\partial \beta'} m_n^1\left(\hat{\beta}_n^1\right) = 0$$

Or

$$\begin{aligned} \left(\frac{1}{n}\sum_{i=1}^{n} X_i X_i'\right)\frac{1}{n}\sum_{i=1}^{n} X_i\left(Y_i - X_i'\hat{\beta}_n^1\right) &= 0 \\ X'X\left[X'Y - X'X\hat{\beta}_n^1\right] &= 0 \\ \hat{\beta}_n^1 &= (X'X)^{-1} X'Y \end{aligned}$$

The estimator for asymptotic covariance matrix is given by

$$\begin{aligned} \widehat{Var}\left(\hat{\beta}_n^1\right) &= \frac{1}{n}\hat{\Lambda}^1\left(\hat{\beta}_n^1\right) \\ &= \frac{1}{n}\left[\hat{A}_1\left(\hat{\beta}_n^1\right)\hat{A}_1\left(\hat{\beta}_n^1\right)'\right]_1^{-1}\hat{A}\left(\hat{\beta}_n^1\right)\hat{W}_1\left(\hat{\beta}_n^1\right)\hat{A}_1\left(\hat{\beta}_n^1\right)'\left[\hat{A}_1\left(\hat{\beta}_n^1\right)\hat{A}_1\left(\hat{\beta}_n^1\right)'\right]^{-1} \\ &= n\left[(X'X)(X'X)'\right]^{-1}(X'X)\hat{W}_1\left(\hat{\beta}_n^1\right)(X'X)'\left[(X'X)(X'X)'\right]^{-1} \end{aligned}$$

Where the elements are described in Question 3, part (e).

$$\begin{aligned} \hat{W}_1\left(\hat{\beta}_n^1\right) &= \frac{1}{n}\sum_{i=1}^{n}\left[\varphi_1\left(Y_i; X_i; \hat{\beta}_n^1\right)\varphi_1\left(Y_i; X_i; \hat{\beta}_n^1\right)'\right] \\ &= \frac{1}{n}\sum_{i=1}^{n}\left[X_i\left(Y_i - X_i'\hat{\beta}_n^1\right)\left(Y_i - X_i\hat{\beta}_n^1\right)' X_i'\right] \end{aligned}$$

Estimates for the parameter vector β_0 and the standard errors based on $V_n = I$

	Coefficient	Standard Error
$\hat{\beta}_1^1$	1.0169	0.0978
$\hat{\beta}_2^1$	1.0004	0.0236
$\hat{\beta}_3^1$	-1.0194	0.0229
$\hat{\beta}_4^1$	0.4928	0.0220
$\hat{\beta}_5^1$	-0.4896	0.0253

(2) Provide the optimal estimate for β_0 and the corresponding standard errors, based on the optimal matrix V_n.

Solution: Using $\varphi_1(Y_i; X_i, \beta)$ and $V_n = \frac{1}{n}\sum_{i=1}^{n}\left\{\left[\varphi_1\left(Y_i; X_i, \hat{\beta}_*^1\right)\right]\left[\varphi_1\left(Y_i; X_i, \hat{\beta}_*^1\right)\right]'\right\}$

Since this is the just identified case, the weight matrix does not affect the numerical estimator: $\hat{\beta}_n^1 = (X'X)^{-1} X'Y$

Since it does not affect the numerical estimator, it also does not affect the estimated covariance of the estimator. However if we calculated again in Matlab using the formula for V_n we will realize that the results don't change.

Estimates for the parameter vector β_0 and the standard errors based on the optimal matrix V_n

	Coefficient	Standard Errors
$\hat{\beta}_1^1$	1.0169	0.0978
$\hat{\beta}_2^1$	1.0004	0.0236
$\hat{\beta}_3^1$	-1.0194	0.0229
$\hat{\beta}_4^1$	0.4928	0.0220
$\hat{\beta}_5^1$	-0.4896	0.0253

(3) Compare the two estimates obtained in $(a.1)$ and $(a.2)$.

Solution: Since the estimators in $(a.1)$ and $(a.2)$ are identical, so too will be the estimates and standard errors.

(b) Repeat the exercise in (a) for $\varphi_2(y_i, z_i; \beta)$.

(1) Provide an estimate for β_0 and the corresponding standard errors, based on $V_n = I$.

Solution: Using $\varphi_2(Y_i; Z_i, \beta)$ and $V_n = I$, we have

$$\hat{\beta}_n^2 = \underset{\beta \in B}{\arg\min} Q_n^2(\beta)$$

Where

$$Q_n^2(\beta) = m_n^2(\beta)' m_n^2(\beta)$$

The first order conditions are thus

$$(\beta) : \frac{\partial m_n^2\left(\hat{\beta}_n^2\right)}{\partial \beta'} m_n^2\left(\hat{\beta}_n^2\right) = 0$$

$$\begin{aligned}
\left(\frac{1}{n}\sum_{i=1}^{n} X_i Z_i'\right)\frac{1}{n}\sum_{i=1}^{n} Z_i\left(Y_i - X_i'\hat{\beta}_n^2\right) &= 0 \\
X'Z\left[Z'Y - Z'X\hat{\beta}_n^2\right] &= 0 \\
\hat{\beta}_n^2 &= (X'ZZ'X)^{-1} X'ZZ'Y
\end{aligned}$$

From above, we have that the estimator for asymptotic covariance matrix is given by

$$\begin{aligned}
\widehat{Var\left(\hat{\beta}_n^2\right)} &= \frac{1}{n}\hat{\Lambda}^1\left(\hat{\beta}_n^2\right) \\
&= \frac{1}{n}\left[\hat{A}_2\left(\hat{\beta}_n^2\right)\hat{A}_2\left(\hat{\beta}_n^2\right)'\right]_1^{-1}\hat{A}\left(\hat{\beta}_n^2\right)\hat{W}_2\left(\hat{\beta}_n^2\right)\hat{A}_2\left(\hat{\beta}_n^2\right)'\left[\hat{A}_2\left(\hat{\beta}_n^2\right)\hat{A}_2\left(\hat{\beta}_n^2\right)'\right]^{-1} \\
&= n(X'ZZ'X)^{-1} X'Z\hat{W}_2\left(\hat{\beta}_n^2\right) Z'X(X'ZZ'X)^{-1}
\end{aligned}$$

Where the elements are described in Question 3, part (e).

$$\begin{aligned}
\hat{W}_2\left(\hat{\beta}_n^2\right) &= \frac{1}{n}\sum_{i=1}^{n}\left[\varphi_2\left(y_i; z_i; \hat{\beta}_n^2\right)\varphi_1\left(y_i; z_i; \hat{\beta}_n^2\right)'\right] \\
&= \frac{1}{n}\sum_{i=1}^{n}\left[z_i\left(y_i - x_i'\hat{\beta}_n^2\right)\left(y_i - x_i\hat{\beta}_n^2\right)' z_i'\right]
\end{aligned}$$

Using MATLAB to compute the estimates for the parameter vector β_0 and the standard errors based on $V_n = I$,

	Coefficient	Standard Errors
$\hat{\beta}_1^2$	1.1686	0.1577
$\hat{\beta}_2^2$	0.9878	0.0309
$\hat{\beta}_3^2$	-1.0459	0.0298
$\hat{\beta}_4^2$	0.4873	0.0281
$\hat{\beta}_5^2$	-0.5135	0.0349

(2) Provide the optimal estimate for β_0 and the corresponding standard errors, based on the optimal matrix V_n.

Solution: Using $\varphi_2(Y_i; Z_i, \beta)$ and $V_n = \frac{1}{n}\sum_{i=1}^{n}\left\{\left[\varphi_2\left(Y_i; Z_i, \hat{\beta}_n^2\right)\right]\left[\varphi_2\left(Y_i; Z_i, \hat{\beta}_n^2\right)\right]'\right\}$, we have

$$Q_n^2(\beta) = m_n^2(\beta)' V_{n,2}^{-1}(\beta) m_n^2(\beta)$$

The first order conditions are given by

$$(\beta): \quad \frac{\partial m_n^2\left(\hat{\beta}_n^2\right)}{\partial \beta'} V_{n,2}^{-1}\left(\hat{\beta}_n^2\right) m_n^2\left(\hat{\beta}_n^2\right) = 0$$

$$\left(\frac{1}{n}\sum_{i=1}^{n} X_i Z_i'\right) V_{n,2}^{-1}\left(\hat{\beta}_n^2\right) \frac{1}{n}\sum_{i=1}^{n} Z_i\left(Y_i - X_i'\hat{\beta}_n^2\right) = 0$$

$$\hat{\beta}_n^2 = \left[X'ZV_{n,2}^{-1}\left(\hat{\beta}_*^2\right)Z'X\right]^{-1} X'ZV_{n,2}^{-1}\left(\hat{\beta}_*^2\right)Z'Y$$

The estimator for the asymptotic covariance is therefore,

$$\begin{aligned}
\widehat{Var\left(\hat{\beta}_n^2\right)} &= \frac{1}{n}\hat{\Lambda}^2\left(\hat{\beta}_n^2\right) \\
&= n\left[\hat{A}_2\left(\hat{\beta}_n^2\right) V_{n,2}^{-1}\left(\hat{\beta}_n^2\right) \hat{A}_2\left(\hat{\beta}_n^2\right)'\right]^{-1} \\
&= n\left[X'ZV_{n,2}^{-1}\left(\hat{\beta}_n^2\right)Z'X\right]^{-1}
\end{aligned}$$

Where

$$\begin{aligned}
V_{n,2}^{-1}\left(\hat{\beta}_n^2\right) &= \frac{1}{n}\sum_{i=1}^{n}\left[\varphi_2\left(Y_i; Z_i, \hat{\beta}_n^2\right)\varphi_1\left(Y_i; Z_i, \hat{\beta}_n^2\right)'\right] \\
&= \frac{1}{n}\sum_{i=1}^{n}\left[Z_i\left(Y_i - X_i'\hat{\beta}_n^2\right)\left(Y_i - X_i\hat{\beta}_n^2\right)' Z_i'\right]
\end{aligned}$$

Using MATLAB, we get

	Coefficient	Standard Errors
$\hat{\beta}_1^2$	1.0177	0.0978
$\hat{\beta}_2^2$	0.9961	0.0233
$\hat{\beta}_3^2$	-1.0168	0.0227
$\hat{\beta}_4^2$	0.4903	0.0219
$\hat{\beta}_5^2$	-0.4875	0.0246

(3) Compare the two estimates obtained in $(b.1)$ and $(b.2)$.

Solution: Comparing

	Coefficient	Standard Errors
$\hat{\beta}_1^2$	1.1686	0.1577
$\hat{\beta}_2^2$	0.9878	0.0309
$\hat{\beta}_3^2$	-1.0459	0.0298
$\hat{\beta}_4^2$	0.4873	0.0281
$\hat{\beta}_5^2$	-0.5135	0.0349

To

	Coefficient	Standard Errors
$\hat{\beta}_1^2$	1.0177	0.0978
$\hat{\beta}_2^2$	0.9961	0.0233
$\hat{\beta}_3^2$	-1.0168	0.0227
$\hat{\beta}_4^2$	0.4903	0.0219
$\hat{\beta}_5^2$	-0.4875	0.0246

We note that the coefficient estimates have changed a little and the standard errors for the GMM estimation with the optimal weight matrix are lower. This confirms the idea that the optimal weight matrix is indeed optimal.

(c) Compare and discuss the differences and similarities between the optimal estimates (and their corresponding vectors of standard errors) from (a) and (b).

Solution: For part (a), both estimates are identical. As noted in part (b.3), the coefficient estimates have changed a little and the standard errors for the GMM estimation with the optimal weight matrix are lower. This confirms the idea that the optimal weight matrix is indeed optimal.

(d) Construct a consistent estimate for the optimal weight matrix bawsed on the optimal estimate from (b) and re-estimate the optimal GMM. Discuss the results in comparison with the results obtained in (b).

Solution: A consistent estimator for the optimal weight matrix based on the optimal estimator from (b) will be

$$\begin{aligned}
V_{n,2}^{-1}\left(\hat{\beta}_n^2\right) &= \frac{1}{n}\sum_{i=1}^{n}\left[\varphi_2\left(Y_i; Z_i, \hat{\beta}_n^2\right)\varphi_1\left(Y_i; Z_i, \hat{\beta}_n^2\right)'\right] \\
&= \frac{1}{n}\sum_{i=1}^{n}\left[Z_i\left(Y_i - X_i'\hat{\beta}_n^2\right)\left(Y_i - X_i\hat{\beta}_n^2\right)' Z_i'\right]
\end{aligned}$$

Using MATLAB, we will have

$$\widehat{Var\left(\hat{\beta}_n^2\right)} = \begin{bmatrix} 0.0096 & -0.0011 & -0.0008 & -0.0009 & -0.0013 \\ -0.0011 & 0.0005 & 0.0000 & -0.0000 & 0.0000 \\ -0.0008 & 0.0000 & 0.0005 & -0.0000 & -0.0000 \\ -0.0009 & -0.0000 & -0.0000 & 0.0005 & 0.0001 \\ -0.0013 & 0.0000 & -0.0000 & 0.0001 & 0.0006 \end{bmatrix}$$

And therefore, given this consistent estimate, the optimal GMM estimator will be

$$\hat{\beta}_n^2 = \begin{bmatrix} \hat{\beta}_1^2 \\ \hat{\beta}_2^2 \\ \hat{\beta}_3^2 \\ \hat{\beta}_4^2 \\ \hat{\beta}_5^2 \end{bmatrix} = \begin{bmatrix} 1.0183 \\ 0.9960 \\ -1.0164 \\ 0.4900 \\ -0.4876 \end{bmatrix}$$

(e) Construct the Wald statistic for testing the hypothesis: $H_0 : \beta_2^2 + \beta_4^2 = \beta_3^2 + \beta_5^2$ based on the optimal estimates for β_0, from both (a) and (b). Discuss briefly your conclusions.

Solution: In order to construct the Wald statistic we first define define

$$W_0 = \left[\Gamma\hat{\beta}^2 - \gamma_0\right]' \left[\Psi\left(\hat{\beta}\right)\right]^{-1} \left[\Gamma\hat{\beta}^2 - \gamma_0\right] \sim \chi^2(1)$$

Let $\Gamma = \begin{bmatrix} 0 & 1 & -1 & 1 & -1 \end{bmatrix}, \gamma_0 = 0$ and $\hat{\beta}^2 = \begin{bmatrix} \hat{\beta}_1^2 \\ \hat{\beta}_2^2 \\ \hat{\beta}_3^2 \\ \hat{\beta}_4^2 \\ \hat{\beta}_4^2 \end{bmatrix}$. Define $h(\beta) = \beta_2^2 + \beta_4^2 - \beta_3^2 - \beta_5^2$. Then the null hypothesis is $H_0 : h(\beta) = 0$. Using the delta method we know that

$$\begin{aligned}
\Psi\left(\hat{\beta}\right) &= \Delta\hat{\Lambda}\left(\hat{\beta}\right)\Delta' = \left[\left.\frac{\partial h(\beta)}{\beta}\right|_{\beta=\hat{\beta}}\right] \hat{\Lambda}\left(\hat{\beta}\right) \left[\left.\frac{\partial h(\beta)}{\beta}\right|_{\beta=\hat{\beta}}\right]' \\
&= \begin{bmatrix} \frac{\partial h(\beta)}{\beta_1} & \frac{\partial h(\beta)}{\beta_2} & \frac{\partial h(\beta)}{\beta_3} & \frac{\partial h(\beta)}{\beta_4} & \frac{\partial h(\beta)}{\beta_5} \end{bmatrix}_{\beta=\hat{\beta}} \left[\hat{\Lambda}\left(\hat{\beta}\right)\right] \begin{bmatrix} \frac{\partial h(\beta)}{\beta_1} \\ \frac{\partial h(\beta)}{\beta_2} \\ \frac{\partial h(\beta)}{\beta_3} \\ \frac{\partial h(\beta)}{\beta_4} \\ \frac{\partial h(\beta)}{\beta_5} \end{bmatrix}_{\beta=\hat{\beta}} \\
&= \begin{bmatrix} 0 & 2\beta_2 & -2\beta_3 & 2\beta_4 & -2\beta_5 \end{bmatrix}_{\beta=\hat{\beta}} \left[\hat{\Lambda}\left(\hat{\beta}\right)\right] \begin{bmatrix} 0 \\ 2\beta_2 \\ -2\beta_3 \\ 2\beta_4 \\ -2\beta_5 \end{bmatrix}_{\beta=\hat{\beta}}
\end{aligned}$$

Therefore we obtain

$$\Gamma\widehat{\beta} - \gamma_0 \sim N\left(\Gamma\beta - \gamma_0, \Psi\left(\beta_0\right)\right)$$

Using $\varphi_1\left(Y_i; X_i, \beta\right)$ we can replace the know values of $\Gamma$and γ_0, we obtain

$$\begin{aligned}
W_1 &= \left(\Gamma\hat{\beta}_1 - \gamma_0\right)' \left[\Delta_1\hat{\Lambda}^1\left(\hat{\beta}_1\right)\Delta_1'\right]^{-1} \left(\Gamma\hat{\beta}_1 - \gamma_0\right) \\
W_1 &= \frac{\left(\hat{\beta}_{1,2} + \hat{\beta}_{1,4} - \hat{\beta}_{1,3} - \hat{\beta}_{1,5}\right)^2}{\Psi_1\left(\hat{\beta}_{1,2}, \hat{\beta}_{1,3}, \hat{\beta}_{1,4}, \hat{\beta}_{1,5}\right)} \\
W_1 &= 0.2293
\end{aligned}$$

Since W_1 critical value at 95% is 3.8415 , we fail to reject null hypothesis. Using $\varphi_2\left(y_i; z_i; \beta\right)$ we can replace the know values of $\Gamma$and γ_0, we obtain

$$\begin{aligned}
W_2 &= \left(\Gamma\hat{\beta}_2 - \gamma_0\right)' \left[\Delta_2\hat{\Lambda}^2\left(\hat{\beta}_2\right)\Delta_2'\right]^{-1} \left(\Gamma\hat{\beta}_2 - \gamma_0\right) \\
W_2 &= \frac{\left(\hat{\beta}_{2,2} + \hat{\beta}_{2,4} - \hat{\beta}_{2,3} - \hat{\beta}_{2,5}\right)^2}{\Psi_2\left(\hat{\beta}_{2,2}, \hat{\beta}_{2,3}, \hat{\beta}_{2,4}, \hat{\beta}_{2,5}\right)} \\
W_2 &= 0.2859
\end{aligned}$$

Since W_2 critical value at 95% is 3.8415 , we fail to reject null hypothesis.