Answer Keys to Problem Set 2

I. The model is

$$Y = X_1'\beta_1 + X_2'\beta_2 + \epsilon$$
$$E(\epsilon|X) = 0$$
$$E(\epsilon^2|X) = \sigma^2$$

where $X = (X_1, X_2)$, with the dimensions of X_1 and X_2 be $k_1 \times 1$ and $k_2 \times 1$. Consider a mis-specified regression $Y = X_1'\hat{\beta}_1 + \hat{\epsilon}$ and define the error variance estimator $s^2 = \frac{1}{n-k_1}\sum_{i=1}^n \hat{\epsilon}_i^2$. Find $E(s^2|X)$.

$$\hat{\beta}_{OLS} = (X_1'X_1)^{-1}X_1'y$$

$$\hat{\varepsilon} = M_1 y = M_1(X_1\beta_1 + X_2\beta_2 + \varepsilon) = M_1(X_2\beta_2 + \varepsilon)$$

$$M_1 X_1 \beta = (I - X_1(X_1'X_1)^{-1}X_1')X_1\beta = 0$$

$$\hat{\varepsilon}'\hat{\varepsilon} = (M_1 X_2'\beta_2 + M_1\varepsilon)'(M_1 X_2'\beta_2 + M_1\varepsilon)$$
$$= \beta_2' X_2 M_1' M_1 X_2'\beta_2 + \beta_2' X_2 M_1' M_1 \varepsilon + \varepsilon' M_1 M_1 X_2'\beta_2 + \varepsilon' M_1' M_1 \varepsilon$$
$$= \beta_2' X_2 M_1 X_2'\beta_2 + \beta_2' X_2 M_1 \varepsilon + \varepsilon' M_1 X_2'\beta_2 + \varepsilon' M_1 \varepsilon$$

$$E[s^2|X] = \frac{\hat{\varepsilon}'\hat{\varepsilon}}{n-k_1} = E[\beta_2' X_2 M_1 X_2'\beta_2 + \varepsilon' M_1 \varepsilon | X]$$
$$= \frac{\beta_2' X_2 M_1 X_2'\beta_2}{n-k_1} + \sigma^2$$

$$E(\varepsilon' M_1 \varepsilon \mid x) = E(tr(\varepsilon' M_1 \varepsilon) \mid X) = E[tr(M_1 \varepsilon' \varepsilon) \mid x]$$

$$= tr(E(M_1 \varepsilon' \varepsilon) \mid X) = tr(M)\sigma^2 = n-k_1 \sigma^2$$

$E[\beta_2' x_2 M_1 \varepsilon \mid x] = 0$ and $E[\varepsilon' M_1 X_2' \beta_2 \mid x] = 0$. because $E[\varepsilon \mid X] = 0$

II. Consider the model,

$$
\begin{aligned}
\mathbf{y} &= \mathbf{X}\beta + \epsilon \\
E(\epsilon|\mathbf{X}) &= 0 \\
E(\epsilon\epsilon'|\mathbf{X}) &= \mathbf{\Omega}
\end{aligned}
$$

Assume for simplicity that $\mathbf{\Omega}$ is known. Consider the OLS and GLS estimators $\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ and $\tilde{\beta} = (\mathbf{X}'\mathbf{\Omega}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{\Omega}^{-1}\mathbf{y}$. Please compute the (conditional) covariance between $\hat{\beta}$ and $\tilde{\beta}$:

$$
E\left(\left(\hat{\beta}-\beta\right)\left(\tilde{\beta}-\beta\right)'\middle|\mathbf{X}\right)
$$

and the (conditional) covariance matrix for $\hat{\beta}-\tilde{\beta}$:

$$
E\left(\left(\hat{\beta}-\tilde{\beta}\right)\left(\hat{\beta}-\tilde{\beta}\right)'\middle|\mathbf{X}\right).
$$

These two covariance matrices play a pivotal role in the development of specification tests in Hausman (1978).

$$E[(\hat{\beta}-\beta)(\tilde{\beta}-\beta)|x] = E[(x'x)^{-1}x'\varepsilon\varepsilon'\Omega^{-1}x(x'\Omega^{-1}x)^{-1}|x] = (x'\Omega x)^{-1}$$

$$E[(\hat{\beta}-\tilde{\beta})(\hat{\beta}-\tilde{\beta})'|x] = E[((\hat{\beta}-\beta)-(\tilde{\beta}-\beta))((\hat{\beta}-\beta)-(\tilde{\beta}-\beta))'|x]$$

$$= (x'x)^{-1}x'\Omega x(x'x)^{-1} - 2(x'\Omega^{-1}x)^{-1} + (x'\Omega^{-1}x)^{-1}$$

$$= (x'x)^{-1}x'\Omega x(x'x)^{-1} - (x'\Omega^{-1}x)^{-1}$$

III. In a regression model $y = \mathbf{X}\beta + \epsilon$, some columns of X represent endogenous independent variables.

1. Explain the problem of endogenous independent variables, what is its consequence in terms of finite and asymptotic properties of OLS estimator?
2. If you have a matrix of exogenous variables $\mathbf{Z}$ where the dimension of $\mathbf{Z}$ is identical to the dimension of $\mathbf{X}$. How do you use $\mathbf{Z}$ as the instrumental variables (IV) for $\mathbf{X}$ to estimate the parameter β. Provide the formula of IV estimator, $\hat{\beta}_{IV}$.
3. If you have a matrix of exogenous variables Z where the dimension of $\mathbf{Z}$ is larger than the dimension of $\mathbf{X}$. How do you construct the best instrumental variables (IV) for $\mathbf{X}$ to estimate the parameter β. Provide the formula of IV estimator, $\hat{\beta}_{IV}$.
4. Given a real-world example of a situation where IV estimation is needed because of inconsistency of OLS, and specify suitable instruments.

#1

if endogenous, $E[u|x] \neq 0$, $E[x'u] \neq 0$.

$E(\hat{\beta}_{OLS}|x) = \beta + (x'x)^{-1}x'E[u|x] \neq \beta$ (finite sample, biased)

$\hat{\beta}_{OLS} = \beta + (x'x)^{-1}x'u \xrightarrow{P} \beta + E[x'x]^{-1}E[x'u] \neq \beta$ (Asymptotically, inconsistent)

#2. if just-identified. (detail note p48)

$E[z'u] = E[z'(y - x\beta)] = E[z'y] - E[z'x]\beta = 0$

$\beta = E[z'x]^{-1}E[z'y]$

$\hat{\beta} = (z'x)^{-1}z'y$

#3 over identified (detail P49)

Firsely. we project f to span Z by $P_Z \triangleq Z(Z'Z)^{-1}Z$ and denote

$$\hat{X} \triangleq P_Z X = Z(Z'Z)^{-1}Z'X$$

Second, obtain $\hat{\beta}_{IV}$ from the regression of y on $\hat{X}$

$$\hat{\beta}_{IV} = (\hat{X}'\hat{X})^{-1}\hat{X}y$$

IV. Consider the three-equations model, $y = \beta x + u$; $x = \lambda u + \epsilon$; $z = \gamma\epsilon + v$, where independent errors u, ϵ, and v are i.i.d. normal with mean 0 and variances, respectively, σ_u^2, σ_ϵ^2, and σ_v^2.

1. Show that $\text{plim}(\hat{\beta}_{\text{ols}} - \beta) = \lambda\sigma_u^2/(\lambda^2\sigma_u^2 + \sigma_\epsilon^2)$.
2. Show that the square correlation $\rho_{xz}^2 = (\gamma\sigma_\epsilon^2)^2/[(\lambda^2\sigma_u^2 + \sigma_\epsilon^2)(\gamma^2\sigma_\epsilon^2 + \sigma_v^2)]$.
3. Show that $\hat{\beta}_{\text{IV}} = m_{zy}/m_{zx} = \beta + m_{zu}/(\lambda m_{zu} + m_{z\epsilon})$, where, for example, $m_{zy} = \frac{1}{n}\sum_{i=1}^{n} z_i y_i$.
4. Show that $\hat{\beta}_{\text{IV}} - \beta \approx 1/\lambda$ if γ (or ρ_{xz}) ≈ 0.
5. Show that $\hat{\beta}_{\text{IV}} - \beta \approx \infty$ if $m_{zu} = -\gamma\sigma_\epsilon^2/\lambda$.
6. What do the last two results imply regarding the finite-sample biases and the moment of $\hat{\beta}_{\text{IV}} - \beta$ when the instruments are poor?

#1

$$\text{plim}(\hat{\beta}_{ols} - \beta) = \frac{E[xu]}{E[x^2]} = \frac{E[\lambda u^2]}{E[\lambda^2 u^2 + \varepsilon^2]} = \frac{\lambda\sigma^2}{\lambda^2\sigma_u^2 + \sigma_\varepsilon^2}$$

#2

$$\rho^2 = \frac{E(xz)^2}{var(x)\,var(z)} = \frac{\gamma^2 E[\varepsilon^2]^2}{E[\lambda^2 u^2 + \varepsilon^2]\,E[\gamma^2\varepsilon^2 + v^2]} = \frac{(\gamma\sigma_\varepsilon^2)^2}{(\lambda^2\sigma_u^2 + \sigma_\varepsilon^2)(\gamma^2\sigma_\varepsilon^2 + \sigma_v^2)}$$

#3.

$$\hat{\beta}_{IV} = \left(\sum_{i=1}^{n} x_i z_i\right)^{-1}\left(\sum_{i=1}^{n} z_i y_i\right) = \beta + \frac{\sum_{i=1}^{n} z_i u_i}{\sum_{i=1}^{n} z_i x_i}$$

$$= \beta + \frac{\sum_{i=1}^{n} z_i u_i}{\sum_{i=1}^{n} z_i(\lambda u_i + \varepsilon_i)} = \beta + \frac{m_{zu}}{\lambda m_{zu} + m_{z\varepsilon}}$$

#4.

$$\hat{\beta}_{IV} - \beta = \frac{m_{zu}}{\lambda m_{zu} + m_{z\varepsilon}} = \frac{\sum_{i=1}^{n} z_i u_i}{\lambda \sum_{i=1}^{n} z_i u_i + \gamma \sum_{i=1}^{n} \varepsilon_i^2 + \sum_{i=1}^{n} v_i \varepsilon_i} \xrightarrow{p} \frac{E[z \cdot u]}{\lambda E[z\ u] + \gamma \sigma_\varepsilon^2 + E[v\varepsilon]}$$

$$m_{z\varepsilon} = \frac{1}{n} \sum_{i=1}^{n} z_i \varepsilon_i$$

$$= \frac{1}{n} \sum_{i=1}^{n} \gamma \varepsilon_i^2 + \frac{1}{n} \sum_{i=1}^{n} v_i \varepsilon_i$$

$$= \frac{1}{\lambda}$$, because $\gamma \approx 0$ and $cov(v\varepsilon) = 0$

#5

$$\frac{\sum_{i=1}^{n} z_i u_i}{\lambda \sum_{i=1}^{n} z_i u_i + \gamma \sum_{i=1}^{n} \varepsilon_i^2 + \sum_{i=1}^{n} v_i u_i} = \frac{-\gamma \frac{\sigma_\varepsilon^2}{\lambda}}{-\gamma \sigma_\varepsilon^2 + \gamma \sum_{i=1}^{n} \varepsilon_i^2 + \sum_{i=1}^{n} v_i \varepsilon_i} \xrightarrow{p} \frac{-\gamma \frac{\sigma_\varepsilon^2}{\lambda}}{\cancel{-\gamma \sigma_\varepsilon^2} + \cancel{\gamma \sigma_\varepsilon^2} + E[v\varepsilon]}$$

$$= \frac{-\gamma \frac{\sigma_\varepsilon^2}{\lambda}}{0}$$

#6.

less correlation between x and z, the IV will cause biases

IV not exogeneous, $\hat{\beta}_{IV}$ will not be unbiased.

V.2

Notice that we will obtain more efficient estimator,

as variance of data getting large. More information. more. efficient.