PhD Econometrics 1: Study Questions Class 3
Imperial College London
Hormoz Ramian

Solutions

Question 1:

(1.1) Define $\boldsymbol{M}_1 = \boldsymbol{I} - \boldsymbol{x}_1(\boldsymbol{x}_1'\boldsymbol{x}_1)^{-1}\boldsymbol{x}_1'$ then,

$$\boldsymbol{M}_1\boldsymbol{y} = 0 + \boldsymbol{M}_1\boldsymbol{x}_2\boldsymbol{\beta}_2 + \boldsymbol{M}_1\boldsymbol{u} \tag{1}$$

Applying the least squares method gives $\widehat{\boldsymbol{\beta}}_2 = (\boldsymbol{x}_2'\boldsymbol{M}_1\boldsymbol{x}_2)^{-1}\boldsymbol{x}_2'\boldsymbol{M}_1\boldsymbol{y}$ which unbiased in the case of stochastic regressors because (using law of iterated expectations),

$$\begin{aligned}\mathbb{E}[\widehat{\boldsymbol{\beta}}_2] &= \mathbb{E}\left\{\mathbb{E}\left[(\boldsymbol{x}_2'\boldsymbol{M}_1\boldsymbol{x}_2)^{-1}\boldsymbol{x}_2'\boldsymbol{M}_1\boldsymbol{y}|\boldsymbol{x}\right]\right\} \\ &= \boldsymbol{\beta}_2 + \mathbb{E}\left\{(\boldsymbol{x}_2'\boldsymbol{M}_1\boldsymbol{x}_2)^{-1}\boldsymbol{x}_2'\boldsymbol{M}_1\mathbb{E}\left[\boldsymbol{u}|\boldsymbol{x}\right]\right\} \\ &= \boldsymbol{\beta}_2\end{aligned}$$

as the last term in the second equation is $\mathbb{E}\left[\boldsymbol{u}|\boldsymbol{x}\right] = 0$.

(1.2)

$$\begin{aligned}\text{var}[\widehat{\boldsymbol{\beta}}_2|\boldsymbol{x}] &= (\boldsymbol{x}_2'\boldsymbol{M}_1\boldsymbol{x}_2)^{-1}\boldsymbol{x}_2'\boldsymbol{M}_1\mathbb{E}\left[\boldsymbol{u}\boldsymbol{u}'|\boldsymbol{x}\right]\boldsymbol{M}_1\boldsymbol{x}_2(\boldsymbol{x}_2'\boldsymbol{M}_1\boldsymbol{x}_2)^{-1} \\ &= \sigma^2(\boldsymbol{x}_2'\boldsymbol{M}_1\boldsymbol{x}_2)^{-1}\boldsymbol{x}_2'\boldsymbol{M}_1\boldsymbol{x}_2(\boldsymbol{x}_2'\boldsymbol{M}_1\boldsymbol{x}_2)^{-1} \\ &= \sigma^2(\boldsymbol{x}_2'\boldsymbol{M}_1\boldsymbol{x}_2)^{-1} \\ &= \sigma^2([\boldsymbol{M}_1\boldsymbol{x}_2]'[\boldsymbol{M}_1\boldsymbol{x}_2])^{-1} \qquad (2)\end{aligned}$$

noting that $[\boldsymbol{M}_1\boldsymbol{x}_2]'[\boldsymbol{M}_1\boldsymbol{x}_2]$ is a scalar. Furthermore, setting up an auxiliary regression by regressing $\boldsymbol{x}_2$ on $\boldsymbol{x}_1$ yields the following results:

$$\begin{aligned}\boldsymbol{x}_2 &= \boldsymbol{x}_1\boldsymbol{\theta} + \boldsymbol{v} \qquad (3) \\ \widehat{\boldsymbol{\theta}} &= (\boldsymbol{x}_1'\boldsymbol{x}_1)^{-1}\boldsymbol{x}_1'\boldsymbol{x}_2 \qquad (4) \\ \widehat{\boldsymbol{x}}_2 &= \boldsymbol{x}_1\widehat{\boldsymbol{\theta}} \qquad (5) \\ \widehat{\boldsymbol{v}} &= \boldsymbol{x}_2 - \widehat{\boldsymbol{x}}_2 \\ &= \boldsymbol{x}_2 - \boldsymbol{x}_1(\boldsymbol{x}_1'\boldsymbol{x}_1)^{-1}\boldsymbol{x}_1'\boldsymbol{x}_2 \\ &= \boldsymbol{M}_1\boldsymbol{x}_2 \qquad (6)\end{aligned}$$

also that total sum of squares and R_2^2 from auxiliary regression are,

$$\begin{aligned}\text{TSS}_2 &= \text{RSS}_2 + \text{ESS}_2 \qquad (7) \\ R_2^2 &= 1 - \frac{\text{RSS}_2}{\text{TSS}_2} \qquad (8)\end{aligned}$$

implies that $\text{RSS}_2 \equiv [\boldsymbol{M}_1\boldsymbol{x}_2]'[\boldsymbol{M}_1\boldsymbol{x}_2] = (1 - R_2^2)\text{TSS}_2$, therefore from equation (2),

$$\text{var}[\widehat{\boldsymbol{\beta}}_2|\boldsymbol{x}] = \sigma^2([\boldsymbol{M}_1\boldsymbol{x}_2]'[\boldsymbol{M}_1\boldsymbol{x}_2])^{-1} = \sigma^2\left\{(1 - R_2^2)\text{TSS}_2\right\}^{-1}. \tag{9}$$

(1.3) When $\boldsymbol{x}_1$ is omitted we have,

$$\begin{aligned}\mathbb{E}[\widehat{\boldsymbol{\beta}}_2] &= \mathbb{E}\left\{\mathbb{E}\left[(\boldsymbol{x}_2'\boldsymbol{x}_2)^{-1}\boldsymbol{x}_2'\boldsymbol{y}|\boldsymbol{x}\right]\right\} \\ &= \mathbb{E}\left\{\mathbb{E}\left[(\boldsymbol{x}_2'\boldsymbol{x}_2)^{-1}\boldsymbol{x}_2'\left\{\boldsymbol{x}_1\boldsymbol{\beta}_1 + \boldsymbol{x}_2\boldsymbol{\beta}_2 + \boldsymbol{u}\right\}|\boldsymbol{x}\right]\right\} \\ &= \boldsymbol{\beta}_2 + \mathbb{E}\left\{\mathbb{E}\left[(\boldsymbol{x}_2'\boldsymbol{x}_2)^{-1}\boldsymbol{x}_2'\boldsymbol{x}_1\boldsymbol{\beta}_1|\boldsymbol{x}\right]\right\}\end{aligned}$$

which is unbiased[1] iff $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are orthogonal (or if $\boldsymbol{\beta}_1 = \boldsymbol{0}$). However, even in the case of

[1] Although this is in general biased, but knowledge about $\boldsymbol{x}_2'\boldsymbol{x}_1$ and $\boldsymbol{\beta}_1$ can be useful information to determine

orthogonality, efficiency is not guaranteed. Denote the vector of residuals from the regression with omitted variable with $\boldsymbol{\varepsilon}$,

$$\boldsymbol{\varepsilon} = \boldsymbol{y} - \boldsymbol{x}_2\widehat{\boldsymbol{\beta}}_2 = \boldsymbol{M}_2\boldsymbol{y} \tag{10}$$

Computing regression (omitted $\boldsymbol{x}_1$) variance term $\widetilde{s}^2$:

$$\begin{aligned}
\widetilde{s}^2 &= \frac{1}{N-k}[\boldsymbol{M}_2\boldsymbol{y}]'[\boldsymbol{M}_2\boldsymbol{y}] \\
&= \frac{1}{N-k}(\boldsymbol{M}_2\boldsymbol{x}_1\boldsymbol{\beta}_1 + \boldsymbol{M}_2\boldsymbol{u})'(\boldsymbol{M}_2\boldsymbol{x}_1\boldsymbol{\beta}_1 + \boldsymbol{M}_2\boldsymbol{u}) \\
&= \frac{1}{N-k}\left(\boldsymbol{u}'\boldsymbol{M}_2\boldsymbol{u} + \boldsymbol{\beta}_1'\boldsymbol{x}_1'\boldsymbol{M}_2\boldsymbol{x}_1\boldsymbol{\beta}_1\right)
\end{aligned}$$

noting that cross-term $2\boldsymbol{u}'\boldsymbol{M}_2\boldsymbol{x}_1\boldsymbol{\beta}_1$ is zero and that $\boldsymbol{u}'\boldsymbol{M}_2\boldsymbol{u}$ is the RSS from the correctly specified model:

$$\widetilde{s}^2 = \sigma^2 + \frac{1}{N-k}\boldsymbol{\beta}_1'\boldsymbol{x}_1'\boldsymbol{M}_2\boldsymbol{x}_1\boldsymbol{\beta}_1$$

when regressors are orthogonal $\boldsymbol{x}_1'\boldsymbol{M}_2\boldsymbol{x}_1 = \boldsymbol{x}_1'(\boldsymbol{I} - \boldsymbol{x}_2(\boldsymbol{x}_2'\boldsymbol{x}_2)^{-1}\boldsymbol{x}_2')\boldsymbol{x}_1 = \boldsymbol{x}_1'\boldsymbol{x}_1$:

$$\widetilde{s}^2 = \sigma^2 + \frac{1}{N-k}\boldsymbol{\beta}_1'\boldsymbol{x}_1'\boldsymbol{x}_1\boldsymbol{\beta}_1$$

where the second term is strictly positive if $\boldsymbol{\beta}_1 \neq 0$ requiring stronger conditions relative to unbiasedness.

(1.4) Denote the vector of residuals from the regression with one regressor with $\boldsymbol{e}_1$ and that of the full regression with $\boldsymbol{e}_2$, and respective variances with s_1^2 and s_2^2:

$$\begin{aligned}
\boldsymbol{e}_2 &= \boldsymbol{y} - \boldsymbol{x}_1\widehat{\boldsymbol{\beta}}_1 - \boldsymbol{x}_2\widehat{\boldsymbol{\beta}}_2 && (11)\\
\boldsymbol{M}_1\boldsymbol{e}_2 &= \boldsymbol{M}_1(\boldsymbol{y} - \boldsymbol{x}_2\widehat{\boldsymbol{\beta}}_2) \\
&= \boldsymbol{M}_1(\boldsymbol{I} - \boldsymbol{x}_2(\boldsymbol{x}_2'\boldsymbol{M}_1\boldsymbol{x}_2)^{-1}\boldsymbol{x}_2'\boldsymbol{M}_1)\boldsymbol{y} && (12)
\end{aligned}$$

then,

$$\begin{aligned}
\boldsymbol{e}_2'\boldsymbol{e}_2 &= \boldsymbol{y}'(\boldsymbol{I} - \boldsymbol{x}_2(\boldsymbol{x}_2'\boldsymbol{M}_1\boldsymbol{x}_2)^{-1}\boldsymbol{x}_2'\boldsymbol{M}_1)'\boldsymbol{M}_1(\boldsymbol{I} - \boldsymbol{x}_2(\boldsymbol{x}_2'\boldsymbol{M}_1\boldsymbol{x}_2)^{-1}\boldsymbol{x}_2'\boldsymbol{M}_1)\boldsymbol{y} \\
&= \boldsymbol{y}'\boldsymbol{M}_1\boldsymbol{y} - \boldsymbol{y}'\boldsymbol{M}_1\boldsymbol{x}_2(\boldsymbol{x}_2'\boldsymbol{M}_1\boldsymbol{x}_2)^{-1}\boldsymbol{x}_2'\boldsymbol{M}_1\boldsymbol{y} \\
&= \boldsymbol{e}_1'\boldsymbol{e}_1 - \boldsymbol{y}'\boldsymbol{M}_1\boldsymbol{x}_2(\boldsymbol{x}_2'\boldsymbol{M}_1\boldsymbol{x}_2)^{-1}\boldsymbol{x}_2'\boldsymbol{M}_1\boldsymbol{y} && (13)
\end{aligned}$$

Comparing variance requires the following conjecture. Suppose, $s_1^2 > s_2^2$ then:

$$\frac{1}{N-k}\boldsymbol{e}_1'\boldsymbol{e}_1 > \frac{1}{N-k-1}\boldsymbol{e}_2'\boldsymbol{e}_2 \tag{14}$$

substituting for $\boldsymbol{e}_1'\boldsymbol{e}_1$ from equation (13),

$$\begin{aligned}
\frac{N-k-1}{N-k}\frac{\boldsymbol{e}_2'\boldsymbol{e}_2 + \boldsymbol{y}'\boldsymbol{M}_1\boldsymbol{x}_2(\boldsymbol{x}_2'\boldsymbol{M}_1\boldsymbol{x}_2)^{-1}\boldsymbol{x}_2'\boldsymbol{M}_1\boldsymbol{y}}{\boldsymbol{e}_2'\boldsymbol{e}_2} &> 1 \\
\frac{N-k-1}{N-k}\frac{\boldsymbol{e}_2'\boldsymbol{e}_2 + \widehat{\boldsymbol{\beta}}_2'\boldsymbol{x}_2'\boldsymbol{M}_1\boldsymbol{x}_2\widehat{\boldsymbol{\beta}}_2}{\boldsymbol{e}_2'\boldsymbol{e}_2} &> 1 \\
\frac{\boldsymbol{e}_2'\boldsymbol{e}_2 + \widehat{\boldsymbol{\beta}}_2'\boldsymbol{x}_2'\boldsymbol{M}_1\boldsymbol{x}_2\widehat{\boldsymbol{\beta}}_2}{\boldsymbol{e}_2'\boldsymbol{e}_2 + \boldsymbol{e}_2'\boldsymbol{e}_2/(N-k-1)} &> 1 \\
\frac{\boldsymbol{e}_2'\boldsymbol{e}_2 + \widehat{\boldsymbol{\beta}}_2'\boldsymbol{x}_2'\boldsymbol{M}_1\boldsymbol{x}_2\widehat{\boldsymbol{\beta}}_2}{\boldsymbol{e}_2'\boldsymbol{e}_2 + s_2^2} &> 1
\end{aligned}$$

the direction of the bias, e.g. if unobservable $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are positively (negatively) correlated and $\boldsymbol{x}_1$ and $\boldsymbol{y}$ are positively (negatively) correlated then $\boldsymbol{\beta}_2$ is overestimated. Alternatively, if the relationships between regressors and the unobservable and the outcome variable are of the opposite directions, then the estimator for $\boldsymbol{\beta}_2$ is underestimated.

For simplicity suppose $\widehat{\boldsymbol{\beta}}_2$ is a scalar then,

$$\frac{\widehat{\beta}_2^2}{s_2^2/([\boldsymbol{M}_1\boldsymbol{x}_2]'[\boldsymbol{M}_1\boldsymbol{x}_2])} \quad > \quad 1$$

indicating that the associated (squared) t-ratio of the additional regressor has to be above one in order to have lower variance (of extended regression).

Question 2:

(2.1) Pre-multiplying the regression with $\boldsymbol{P}_z$, gives $\boldsymbol{P}_z\boldsymbol{y} = \boldsymbol{P}_z\boldsymbol{x} + \boldsymbol{P}_z\boldsymbol{u}$, then applying the least squares methods yields,

$$\begin{aligned} \widehat{\boldsymbol{\beta}}_{IV} &= \left(\boldsymbol{x}'\boldsymbol{z}(\boldsymbol{z}'\boldsymbol{z})^{-1}\boldsymbol{z}'\boldsymbol{x}\right)^{-1}\boldsymbol{x}'\boldsymbol{z}(\boldsymbol{z}'\boldsymbol{z})^{-1}\boldsymbol{z}'\boldsymbol{y} \qquad (15)\\ &= \left[\frac{\boldsymbol{x}'\boldsymbol{z}}{n}\left(\frac{\boldsymbol{z}'\boldsymbol{z}}{n}\right)^{-1}\frac{\boldsymbol{z}'\boldsymbol{x}}{n}\right]^{-1}\frac{\boldsymbol{x}'\boldsymbol{z}}{n}\left(\frac{\boldsymbol{z}'\boldsymbol{z}}{n}\right)^{-1}\frac{\boldsymbol{z}'\boldsymbol{y}}{n}\\ &= \boldsymbol{\beta} + \left[\frac{\boldsymbol{x}'\boldsymbol{z}}{n}\left(\frac{\boldsymbol{z}'\boldsymbol{z}}{n}\right)^{-1}\frac{\boldsymbol{z}'\boldsymbol{x}}{n}\right]^{-1}\frac{\boldsymbol{x}'\boldsymbol{z}}{n}\left(\frac{\boldsymbol{z}'\boldsymbol{z}}{n}\right)^{-1}\frac{\boldsymbol{z}'\boldsymbol{u}}{n}\end{aligned}$$

then,

$$\sqrt{n}(\widehat{\boldsymbol{\beta}}_{IV} - \boldsymbol{\beta}) = \left[\frac{\boldsymbol{x}'\boldsymbol{z}}{n}\left(\frac{\boldsymbol{z}'\boldsymbol{z}}{n}\right)^{-1}\frac{\boldsymbol{z}'\boldsymbol{x}}{n}\right]^{-1}\frac{\boldsymbol{x}'\boldsymbol{z}}{n}\left(\frac{\boldsymbol{z}'\boldsymbol{z}}{n}\right)^{-1}\frac{\boldsymbol{z}'\boldsymbol{u}}{\sqrt{n}} \qquad (16)$$

Denote $\boldsymbol{\Sigma}_{xz} = \text{plim } n^{-1}\boldsymbol{x}'\boldsymbol{z}$, $\boldsymbol{\Sigma}_{zz} = \text{plim } n^{-1}\boldsymbol{z}'\boldsymbol{z}$ and[2] that by CLT,

$$\frac{\boldsymbol{z}'\boldsymbol{u}}{\sqrt{n}} \xrightarrow{d} N(\boldsymbol{0}, \sigma^2\boldsymbol{\Sigma}_{zz}) \qquad (17)$$

$$\sqrt{n}(\widehat{\boldsymbol{\beta}}_{IV} - \boldsymbol{\beta}) \xrightarrow{d} N(\boldsymbol{0}, \boldsymbol{V}) \qquad (18)$$

where $\boldsymbol{V} = \sigma^2(\boldsymbol{\Sigma}_{xz}\boldsymbol{\Sigma}_{zz}^{-1}\boldsymbol{\Sigma}_{xz})^{-1}$, confirming the consistency and limiting distribution of the IV estimator. Moreover, notice that when $l = k$, then, $\boldsymbol{x}'\boldsymbol{z}$ is full rank and invertible such that the (inverse) power can be distributed over individual pairs leading to:

$$\begin{aligned} \widehat{\boldsymbol{\beta}}_{IV} &= (\boldsymbol{x}'\boldsymbol{z})^{-1}(\boldsymbol{z}'\boldsymbol{z})(\boldsymbol{z}'\boldsymbol{x})^{-1}(\boldsymbol{x}'\boldsymbol{z})(\boldsymbol{z}'\boldsymbol{z})^{-1}\boldsymbol{z}'\boldsymbol{y}\\ &= (\boldsymbol{x}'\boldsymbol{z})^{-1}\boldsymbol{z}'\boldsymbol{y}\\ &= \boldsymbol{\beta} + (\boldsymbol{x}'\boldsymbol{z})^{-1}\boldsymbol{z}'\boldsymbol{u}\end{aligned}$$

re-arranging yields:

$$\sqrt{n}(\widehat{\boldsymbol{\beta}}_{IV} - \boldsymbol{\beta}) = \left(\frac{\boldsymbol{x}'\boldsymbol{z}}{n}\right)^{-1}\frac{\boldsymbol{z}'\boldsymbol{u}}{\sqrt{n}} \qquad (19)$$

Question 3:

(3.1) Vectorizing the structural and reduced form equations, respectively, gives:

$$\begin{bmatrix} 1 & -\alpha_1 \\ -\alpha_2 & 1 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \beta_1 & 0 \\ 0 & \beta_2 \end{bmatrix}\begin{bmatrix} x \\ w \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \qquad (20)$$

$$\boldsymbol{A}\boldsymbol{y} = \boldsymbol{B}\boldsymbol{z} + \boldsymbol{u} \qquad (21)$$

and

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \pi_{11} & \pi_{21} \\ \pi_{12} & \pi_{22} \end{bmatrix}\begin{bmatrix} x \\ w \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \end{bmatrix} \qquad (22)$$

$$\boldsymbol{y} = \boldsymbol{\Pi}\boldsymbol{z} + \boldsymbol{\epsilon} \qquad (23)$$

[2]Since the off diagonal element of $\boldsymbol{\Sigma}_{xz}$ is $\text{cov}(\boldsymbol{x}, \boldsymbol{z}) = \text{cov}(\boldsymbol{z}, \boldsymbol{x})$ then $\boldsymbol{\Sigma}_{xz} = \boldsymbol{\Sigma}'_{xz} = \boldsymbol{\Sigma}_{zx}$.

re-arranging equations (21) and (23) requires $\mathbf{\Pi}' = \boldsymbol{B}'(\boldsymbol{A}^{-1})'$:

$$\mathbf{\Pi}' \quad = \quad \begin{bmatrix} \beta_1 & 0 \\ 0 & \beta_2 \end{bmatrix} \times \frac{1}{1-\alpha_1\alpha_2} \begin{bmatrix} 1 & \alpha_1 \\ \alpha_2 & 1 \end{bmatrix}' \tag{24}$$

thus

$$\begin{aligned} \pi_{11} &= (1-\alpha_1\alpha_2)^{-1}\beta_1 \\ \pi_{12} &= (1-\alpha_1\alpha_2)^{-1}\alpha_2\beta_1 \\ \pi_{21} &= (1-\alpha_1\alpha_2)^{-1}\alpha_1\beta_2 \\ \pi_{22} &= (1-\alpha_1\alpha_2)^{-1}\beta_2 \end{aligned}$$

(3.2) Solving for the structural parameters gives:

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} \frac{\pi_{21}}{\pi_{22}} \\ \frac{\pi_{12}}{\pi_{11}} \end{bmatrix}, \quad \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = (\pi_{11}\pi_{22} - \pi_{12}\pi_{21}) \begin{bmatrix} \frac{1}{\pi_{22}} \\ \frac{1}{\pi_{11}} \end{bmatrix} \tag{25}$$

(3.3) The relevance condition (w vs. y) depends on $\pi_{22} \neq 0$ that is to say α_1 and β_1 are not identified if $\pi_{22} = 0$.