ECON 710A - Problem Set 3

Alex von Hafften*

2/16/2020

1. Let $(Y, X', Z')'$ be a random vector such that $Y = X'\beta + U$, $E[U|Z] = 0$ where $E[ZX']$ is invertible and $E[Y^4 + ||X||^4 + ||Z||^4] < \infty$. Also, let $\{(Y_i, X_i', Z_i')'\}_{i=1}^n$ be a random sample from the distribution of $(Y, X', Z')'$. We showed in lecture 4 that $\sqrt{n}(\hat{\beta}^{IV} - \beta) \to_d N(0, \Omega)$, $\Omega = E[ZX']^{-1}E[ZZ'U^2]E[XZ']^{-1}$. Now suppose that $X = (X_1, X_2')'$, $Z = (Z_1, X_2')'$, and $E[U^2|Z] = \sigma_U^2$ and let Ω_{11} be the upper left entry in Ω.

(i) Show that $E[ZX']$ is invertible iff $E[ZZ']$ is invertible and $\pi_1 \neq 0$ where $(\pi_1, \pi_2')' = E[ZZ']^{-1}E[ZX_1]$.

Observe that

$$E[ZX'] = E[\begin{pmatrix} Z_1 \\ X_2 \end{pmatrix} \begin{pmatrix} X_1 & X_2' \end{pmatrix}] = \begin{pmatrix} E[Z_1X_1] & E[Z_1X_2'] \\ E[X_2X_1] & E[X_2X_2'] \end{pmatrix}$$

$$E[ZZ'] = E[\begin{pmatrix} Z_1 \\ X_2 \end{pmatrix} \begin{pmatrix} Z_1 & X_2' \end{pmatrix}] = \begin{pmatrix} E[Z_1^2] & E[Z_1X_2'] \\ E[X_2Z_1] & E[X_2X_2'] \end{pmatrix}$$

First, if either $E[ZZ']$ or $E[ZX']$ are invertible then, $E[X_2X_2']$ is invertible. Assume for sake of a contradiction that $E[X_2X_2']$ is not invertible. Then there exists some nonzero t such that $E[X_2X_2']t = 0 \implies t'E[X_2X_2']t = t'0 = 0 \implies E[X_2't] = 0 \implies E[Z_1X_2']t = 0$. This implies that $E[ZX']$ and $E[ZZ']$ are not invertible, which is a contradiction:

$$E[ZX'](0, t')' = \begin{pmatrix} E[Z_1X_1] & E[Z_1X_2'] \\ E[X_2X_1] & E[X_2X_2'] \end{pmatrix} \begin{pmatrix} 0 \\ t' \end{pmatrix} = \begin{pmatrix} E[Z_1X_1](0) + E[Z_1X_2']t' \\ E[X_2X_1](0) + E[X_2X_2']t' \end{pmatrix} = 0$$

$$E[ZZ'](0, t')' = \begin{pmatrix} E[Z_1^2] & E[Z_1X_2'] \\ E[X_2Z_1] & E[X_2X_2'] \end{pmatrix} \begin{pmatrix} 0 \\ t' \end{pmatrix} = \begin{pmatrix} E[Z_1^2](0) + E[Z_1X_2']t \\ E[X_2Z_1](0) + E[X_2X_2']t \end{pmatrix} = 0$$

By the block inversion formula, if $E[ZX']$ is invertible iff

$$E[Z_1X_1] - E[Z_1X_2']E[X_2X_2']^{-1}E[X_2X_1] \neq 0 \iff E[\tilde{Z}_1X_1] \neq 0$$

where $\tilde{Z}_1 := Z_1 - X_2'E[X_2X_2']^{-1}E[X_2Z_1]$. Denote $\pi_1 := \frac{E[\tilde{Z}_1X_1]}{E[\tilde{Z}_1^2]}$,

$$E[\tilde{Z}_1X_1] \neq 0 \iff \pi_1 E[\tilde{Z}_1^2] \neq 0 \iff \pi_1 \neq 0 \text{ and } E[\tilde{Z}_1^2] = E[Z_1^2] - E[Z_1X_2']E[X_2X_2']^{-1}E[X_2Z_1] \neq 0$$

By the block inversion formula, $E[Z_1^2] - E[Z_1X_2']E[X_2X_2']^{-1}E[X_2Z_1] \neq 0 \iff E[ZZ']$ is invertible. $\square$

*I worked on this problem set with a study group of Michael Nattinger, Andrew Smith, and Ryan Mather. I also discussed problems with Sarah Bass, Emily Case, Danny Edgel, and Katherine Kwok.

(ii) Show that $\Omega_{11} = \frac{\sigma_U^2}{E[\tilde{Z}_1^2]\pi_1^2}$ where $\tilde{Z}_1 = Z_1 - X_2' E[X_2 X_2']^{-1} E[X_2 Z_1]$.

Using the block inversion formula, define A and B:

$$\begin{aligned}
A &= E[ZX']^{-1} \\
&= \begin{pmatrix} E[Z_1 X_1] & E[Z_1 X_2'] \\ E[X_2 X_1] & E[X_2 X_2'] \end{pmatrix}^{-1} \\
B &= E[XZ']^{-1} \\
A_{11} &= B_{11} \\
&= (E[Z_1 X_1] - E[X_2 X_1] E[X_2 X_2']^{-1} E[X_2 X_1])^{-1} \\
&= E[\tilde{Z}_1 X_1]^{-1} \\
A_{12} &= B_{21} \\
&= -E[Z_1 X_1]^{-1} E[Z_1 X_2'] E[\tilde{Z}_1 X_1]^{-1} \\
A_{21} &= B_{12} \\
&= -E[\tilde{Z}_1 X_1]^{-1} E[X_2 X_1] E[Z_1 X_1]^{-1} \\
A_{22} &= B_{22} \\
&= E[X_2 X_2']^{-1} + E[X_2 X_2']^{-1} E[X_2 X_1] E[\tilde{Z}_1 X_1]^{-1} E[Z_1 X_2'] E[X_2 X_2']^{-1}
\end{aligned}$$

Homoskedastic variance-covariance matrix:

$$\begin{aligned}
\Omega &= \sigma_U^2 E[ZX']^{-1} E[ZZ'] E[XZ']^{-1} \\
&= \sigma_U^2 A E[ZZ'] B \\
&= \sigma_U^2 \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} E[Z_1^2] & E[Z_1 X_2'] \\ E[X_2 Z_1] & E[X_2 X_2'] \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \\
&= \sigma_U^2 \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} B_{11} E[Z_1^2] + B_{21} E[Z_1 X_2'] & B_{12} E[Z_1^2] + B_{22} E[Z_1 X_2'] \\ B_{11} E[X_2 Z_1] + B_{21} E[X_2 X_2'] & B_{12} E[X_2 Z_1] + B_{22} E[X_2 X_2'] \end{pmatrix} \\
&= \sigma_U^2 \begin{pmatrix} A_{11} B_{11} E[Z_1^2] + A_{11} B_{21} E[Z_1 X_2'] + A_{12} B_{11} E[X_2 Z_1] + A_{12} B_{21} E[X_2 X_2'] & ... \\ ... & ... \end{pmatrix}
\end{aligned}$$

$$\begin{aligned}
\Omega_{11} &= \sigma_U^2 [A_{11} B_{11} E[Z_1^2] + A_{11} B_{21} E[Z_1 X_2'] + A_{12} B_{11} E[X_2 Z_1] + A_{12} B_{21} E[X_2 X_2']] \\
&= \sigma_U^2 [E[\tilde{Z}_1 X_1]^{-1} E[\tilde{Z}_1 X_1]^{-1} E[Z_1^2] \\
&+ E[\tilde{Z}_1 X_1]^{-1} (-E[Z_1 X_1]^{-1} E[Z_1 X_2'] E[\tilde{Z}_1 X_1]^{-1}) E[Z_1 X_2'] \\
&+ (-E[Z_1 X_1]^{-1} E[Z_1 X_2'] E[\tilde{Z}_1 X_1]^{-1}) E[\tilde{Z}_1 X_1]^{-1} E[X_2 Z_1] \\
&+ (-E[Z_1 X_1]^{-1} E[Z_1 X_2'] E[\tilde{Z}_1 X_1]^{-1})(-E[Z_1 X_1]^{-1} E[Z_1 X_2'] E[\tilde{Z}_1 X_1]^{-1}) E[X_2 X_2']] \\
&= \frac{\sigma_U^2}{E[\tilde{Z}_1^2]} \frac{E[\tilde{Z}_1^2]^2}{E[\tilde{Z}_1 X_1]^2} \\
&= \frac{\sigma_U^2}{E[\tilde{Z}_1^2]\pi_1^2}
\end{aligned}$$

(iii) Explain, using "regression language", what $(\pi_1, \pi_2')'$ and $\tilde{Z}_1$ is.

$(\pi_1, \pi_2')'$ are the coefficients of the projection of X_1 onto Z. $\tilde{Z}_1$ is the regression error from the projection of Z_1 on X_2.

(iv) Show that $\Omega_{11} \geq \frac{\sigma_U^2}{E[Z_*^2]}$ where $Z_* = E[X_1|Z] - X_2' E[X_2 X_2']^{-1} E[X_2 E[X_1|Z]]$ and provide restrictions on $E[X_1|Z]$ such that $\Omega_{11} = \frac{\sigma_U^2}{E[Z_*^2]}$.

First, note that

$$\begin{aligned}
E[\tilde{Z}_1 X_2] &= E[X_2 \tilde{Z}_1] \\
&= E[X_2(Z_1 - X_2' E[X_2 X_2']^{-1} E[X_2 Z_1])] \\
&= E[X_2 Z_1 - X_2 X_2' E[X_2 X_2']^{-1} E[X_2 Z_1]] \\
&= E[X_2 Z_1] - E[X_2 X_2'] E[X_2 X_2']^{-1} E[X_2 Z_1] \\
&= E[X_2 Z_1] - E[X_2 Z_1] \\
&= 0
\end{aligned}$$

This implies:

$$\begin{aligned}
E[\tilde{Z}_1 X_2] &= 0 \\
\implies E[\tilde{Z}_1 X_2] E[X_2 X_2']^{-1} E[X_2 X_2'] &= 0 \\
\implies E[\tilde{Z}_1 X_2 E[X_2 X_2']^{-1} E[X_2 E[X_1|Z]]] &= 0
\end{aligned}$$

$$\begin{aligned}
E[\tilde{Z}_1 X_1] &= E[\tilde{Z}_1 E[X_1|Z]] \\
&= E[\tilde{Z}_1 E[X_1|Z]] + E[\tilde{Z}_1 X_2 E[X_2 X_2']^{-1} E[X_2 E[X_1|Z]]] \\
&= E[\tilde{Z}_1 (E[X_1|Z] + X_2 E[X_2 X_2']^{-1} E[X_2 E[X_1|Z]])] \\
&= E[\tilde{Z}_1 Z_*]
\end{aligned}$$

By Cauchy-Schwarz,

$$\begin{aligned}
\Omega_{11} &= \frac{\sigma_U^2}{E[\tilde{Z}_1^2] \pi_1^2} \\
&= \frac{\sigma_U^2 E[\tilde{Z}_1^2]}{E[\tilde{Z}_1 X_1]^2} \\
&= \frac{\sigma_U^2 E[\tilde{Z}_1^2]}{E[\tilde{Z}_1 Z_*]^2} \\
&\geq \frac{\sigma_U^2 E[\tilde{Z}_1^2]}{E[\tilde{Z}_1^2] E[Z_*^2]} \\
&= \frac{\sigma_U^2}{E[Z_*^2]}
\end{aligned}$$

The restriction on $E[X_1|Z]$:

$$
\begin{aligned}
E[X_1|Z] &= Z_1\pi_1 + X_2'\pi_2 \\
\implies Z_* &= Z_1\pi_1 + X_2'\pi_2 - X_2'E[X_2X_2']^{-1}E[X_2(Z_1\pi_1 + X_2'\pi_2)] \\
&= Z_1\pi_1 + X_2'\pi_2 - X_2'E[X_2X_2']^{-1}E[X_2Z_1]\pi_1 - X_2'E[X_2X_2']^{-1}E[X_2X_2']\pi_2 \\
&= Z_1\pi_1 + X_2'\pi_2 - X_2'E[X_2X_2']^{-1}E[X_2Z_1]\pi_1 - X_2'\pi_2 \\
&= (Z_1 - X_2'E[X_2X_2']^{-1}E[X_2Z_1])\pi_1 \\
&= \tilde{Z}_1\pi_1 \\
\implies \frac{\sigma_U^2}{E[Z_*^2]} &= \frac{\sigma_U^2}{E[(\tilde{Z}_1\pi_1)^2]} \\
&= \frac{\sigma_U^2}{E[\tilde{Z}_1^2\pi_1^2]} \\
&= \frac{\sigma_U^2}{E[\tilde{Z}_1^2]\pi_1^2} \\
&= \Omega_{11}
\end{aligned}
$$

(v) Suppose that $X_2 = 1$. Write Ω_{11}/σ_U^2 as a function of variances and covariances involving Z_1 and X_1.

$$
\begin{aligned}
X_2 &= 1 \\
\implies Z_1 - (1)E[(1)(1)]^{-1}E[(1)Z_1] &= Z_1 - E[Z_1] \\
\implies \Omega_{11} &= \frac{\sigma_U^2 E[\tilde{Z}_1^2]}{E[\tilde{Z}_1X_1]^2} \\
&= \frac{\sigma_U^2 E[(Z_1 - E[Z_1])^2]}{E[(Z_1 - E[Z_1])X_1]^2} \\
&= \frac{\sigma_U^2 Var[Z_1]}{Cov[Z_1, X_1]^2}
\end{aligned}
$$

2. Let $(Y, X, Z)'$ be a random vector such that $Y = X\beta_1 + U, E[U|Z] = 0$ where $E[Y_4 + X_4 + Z_4] < \infty$. Also, let $\{(Y_i, X_i, Z_i)'\}_{i=1}^n$ be a random sample from the distribution of $(Y, X', Z')'$. Let h be a function of Z such that $E[h(Z)^4] < \infty$.

(i) Provide conditions such that $E[h(Z)(Y - X\beta)] = 0$ iff $\beta = \beta_1$.

On the condition that $E[h(Z)X] \neq 0$:

$$\begin{aligned}
E[h(Z)(Y - X\beta)] &= E[h(Z)(X\beta_1 + U - X\beta)] \\
&= E[h(Z)U] + E[h(Z)X(\beta_1 - \beta)] \\
&= E[h(Z)E[U|Z]] + E[h(Z)X](\beta_1 - \beta) \\
&= E[h(Z)X](\beta_1 - \beta) \\
\iff \beta = \beta_1
\end{aligned}$$

(ii) Derive a method of moments estimator of β_1, say $\hat{\beta}_1^h$, using the IV moment in (i).

The method of moments estimator is:

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^n h(Z_i)(Y_i - X_i\hat{\beta}_1^h) = 0 \\
\implies \frac{1}{n}\sum_{i=1}^n h(Z_i)Y_i - \frac{1}{n}\sum_{i=1}^n h(Z_i)X_i\hat{\beta}_1^h = 0 \\
\implies \hat{\beta}_1^h &= \frac{\frac{1}{n}\sum_{i=1}^n h(Z_i)Y_i}{\frac{1}{n}\sum_{i=1}^n h(Z_i)X_i} \\
&= \frac{\frac{1}{n}\sum_{i=1}^n h(Z_i)(X_i\beta_1 + U_i)}{\frac{1}{n}\sum_{i=1}^n h(Z_i)X_i} \\
&= \frac{\frac{1}{n}\sum_{i=1}^n h(Z_i)X_i\beta_1}{\frac{1}{n}\sum_{i=1}^n h(Z_i)X_i} + \frac{\frac{1}{n}\sum_{i=1}^n h(Z_i)U_i}{\frac{1}{n}\sum_{i=1}^n h(Z_i)X_i} \\
&= \beta_1 + \frac{\frac{1}{n}\sum_{i=1}^n h(Z_i)U_i}{\frac{1}{n}\sum_{i=1}^n h(Z_i)X_i}
\end{aligned}$$

(iii) Under the conditions provided in (i), show that $\sqrt{n}(\hat{\beta}_1^h - \beta_1) \to_d N(0, \Omega^h)$ for some asymptotic variance $\Omega^h \geq 0$.

$$\sqrt{n}(\hat{\beta}_1^h - \beta_1) = \frac{\frac{1}{\sqrt{n}}\sum_{i=1}^n h(Z_i)U_i}{\frac{1}{n}\sum_{i=1}^n h(Z_i)X_i}$$

Because $E[h(Z)X]^2 < E[h(Z)^2]E[X^2] < \infty$, the law of large numbers implies that

$$\frac{1}{n}\sum_{i=1}^n h(Z_i)X_i \to_p E[h(Z)X]$$

By the condition from (i), $E[h(Z)X] \neq 0$. Because $E[h(Z)^2U^2]^2 \leq E[h(Z)^4]E[U^4] < \infty$, the central limit theorem implies that

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n h(Z_i)U_i \to_d N(0, E[h(Z)^2U^2])$$

Finally, by the continuous mapping theorem,

$$\sqrt{n}(\hat{\beta}_1^h - \beta_1) = \frac{\frac{1}{\sqrt{n}}\sum_{i=1}^n h(Z_i)U_i}{\frac{1}{n}\sum_{i=1}^n h(Z_i)X_i} \to_d N(0, \frac{E[h(Z)^2U^2]}{E[h(Z)X]^2})$$

(iv) Show that $\Omega^h \geq E[\frac{E[X|Z]^2}{E[U^2|Z]}]^{-1}$ and find a function h, such that Ω^h achieves this lower bound.

From (iii),

$$\begin{aligned}
\Omega^h &= \frac{E[h(Z)^2U^2]}{E[h(Z)X]^2} \\
&= \frac{E[h(Z)^2E[U^2|Z]]}{E[h(Z)E[X|Z]]^2} \\
&= \frac{E[h(Z)^2E[U^2|Z]]}{E[h(Z)\sqrt{E[U^2|Z]}\frac{E[X|Z]}{\sqrt{E[U^2|Z]}}]^2} \\
&\geq \frac{E[h(Z)^2E[U^2|Z]]}{E[h(Z)^2E[U^2|Z]]E[\frac{E[X|Z]^2}{E[U^2|Z]}]} \\
&= [E[\frac{E[X|Z]^2}{E[U^2|Z]}]]^{-1}
\end{aligned}$$

If $h(Z) = \frac{E[X|Z]}{E[U^2|Z]}$:

$$\begin{aligned}
\Omega_h &= \frac{E[[\frac{E[X|Z]}{E[U^2|Z]}]^2U^2]}{E[\frac{E[X|Z]}{E[U^2|Z]}X]^2} \\
&= [E[\frac{E[X|Z]^2}{E[U^2|Z]}]]^{-1}
\end{aligned}$$

3. Consider the data from Angrist and Krueger (1991) provided on the course website and the following linear model for $log(wage)$ as a function of educationand additional control variables: $log(wage) = \beta_0 + educ \cdot \beta_1 + \sum_{t=31}^{39} 1\{yob = t\}\beta_t + \sum_{s=1}^{50} 1\{sob = s\}\gamma_s + U$, where yob is year of birth and sob is state of birth. As instruments for $educ$ consider three instruments: $1\{qob = 2\}, 1\{qob = 3\}, 1\{qob = 4\}$ where qob is quarter of birth. Using matrix algebra and your preferred statistical software, write code that loads the data and computes the 2SLS estimate of β_1 and the heteroskedasticity robust standard error stemming from the variance estimator formula (12.40) on page 354 of Bruce Hansen's textbook. The solution to this exercise should include code and the two numbers produced by it.

```r
data <- read_csv("AK91.csv", col_types = "ddddd")

n <- nrow(data)

# prep variables
y <- data$lwage
x_1 <- data$educ
controls <- cbind(rep(1, n),
                  to.dummy(data$yob, prefix = "yob")[,2:10],
                  to.dummy(data$sob, prefix = "sob")[,2:51])
instruments <- to.dummy(data$qob, prefix = "qob")[,2:4]

z <- cbind(instruments, controls)
x <- cbind(x_1, controls)

k <- ncol(x)
l <- ncol(z)

# Estimating 2sls beta using 12.29 in Hansen
pi_hat <- solve(t(z) %*% z) %*% t(z) %*% x_1
z_2 <- cbind(z %*% pi_hat, controls)
beta_2sls <- solve(t(z_2) %*% x) %*% t(z_2) %*% y

# Estimating heteroskedastic robust standard errors using 12.40 in Hansen
e_hat <- as.numeric(y - x %*% beta_2sls)

omega <- crossprod(z_2 * e_hat)

varcov <- solve(t(z_2) %*% x) %*% omega %*% solve(t(x) %*% z_2)

print(beta_2sls[1])
```

```
## [1] 0.1083936
```

```r
print(sqrt(varcov[1,1]))
```

```
## [1] 0.01954637
```