AS.180.633: Econometrics **Spring 2020**

Homework 9: Suggested Solutions

Instructor: Yingyao Hu *By: Tong Zhou*

12.4

The reduced form between the regressors $\boldsymbol{x}_i$ and instruments $\boldsymbol{z}_i$ takes the form

$$\boldsymbol{x}_i = \boldsymbol{\Gamma}'\boldsymbol{z}_i + \boldsymbol{u}_i$$

or

$$\boldsymbol{X} = \boldsymbol{Z}\boldsymbol{\Gamma} + \boldsymbol{U}$$

where $\boldsymbol{x}_i$ is $k \times 1$, $\boldsymbol{z}_i$ is $l \times 1$, $\boldsymbol{X}$ is $n \times k$, $\boldsymbol{Z}$ is $n \times l$, $\boldsymbol{U}$ is $n \times k$, and $\boldsymbol{\Gamma}$ is $l \times k$. The parameter $\boldsymbol{\Gamma}$ is defined by the population moment condition

$$\mathbb{E}[\boldsymbol{z}_i\boldsymbol{u}_i'] = \boldsymbol{0}$$

Show that the method of moments estimator for $\boldsymbol{\Gamma}$ is $\widehat{\boldsymbol{\Gamma}} = (\boldsymbol{Z}'\boldsymbol{Z})^{-1}(\boldsymbol{Z}'\boldsymbol{X})$.

Proof. The moment condition $\mathbb{E}[\boldsymbol{z}_i\boldsymbol{u}_i'] = \boldsymbol{0}$ implies

$$\boldsymbol{\Gamma} = \left(\mathbb{E}[\boldsymbol{z}_i\boldsymbol{z}_i']\right)^{-1}\left(\mathbb{E}[\boldsymbol{z}_i\boldsymbol{x}_i']\right)$$

So a natural MoM estimator is

$$\widehat{\boldsymbol{\Gamma}} = (\boldsymbol{Z}'\boldsymbol{Z})^{-1}(\boldsymbol{Z}'\boldsymbol{X})$$

□

12.7

Take the linear model

$$y_i = x_i\beta + e_i$$
$$\mathbb{E}[e_i|x_i] = 0.$$

where x_i and β are 1×1.

(a) Show that $\mathbb{E}[x_i e_i] = 0$ and $\mathbb{E}[x_i^2 e_i] = 0$. Is $\boldsymbol{z}_i = \left(x_i \ x_i^2\right)'$ a valid instrument for estimation of β?

(b) Define the 2SLS estimator of β, using $\boldsymbol{z}_i$ as an instrument for x_i. How does this different from OLS?

Proof. **(a)**

By the *Law of Iterated Expectations*,

$$\begin{aligned}
\mathbb{E}[x_i e_i] &= \mathbb{E}[x_i \mathbb{E}[e_i | x_i]] = 0 \\
\mathbb{E}[x_i^2 e_i] &= \mathbb{E}[x_i^2 \mathbb{E}[e_i | x_i]] = 0
\end{aligned}$$

It implies $\mathbb{E}[\boldsymbol{z}_i e_i] = \boldsymbol{0}$, implying it is a valid instrument.

(b)

Consider the projection matrix $\boldsymbol{P_Z}$, by Exercise 3.7

$$\boldsymbol{P_Z X} = \boldsymbol{X}$$

therefore

$$\begin{aligned}
\widehat{\beta}_{\mathrm{2SLS}} &= \left(\boldsymbol{X}'\boldsymbol{P_Z X}\right)^{-1}\left(\boldsymbol{X}'\boldsymbol{P_Z y}\right) \\
&= \left(\left(\boldsymbol{P_Z X}\right)'\left(\boldsymbol{P_Z X}\right)\right)^{-1}\left(\left(\boldsymbol{P_Z X}\right)'\left(\boldsymbol{P_Z y}\right)\right) \\
&= \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{y} \\
&= \widehat{\beta}_{\mathrm{OLS}}
\end{aligned}$$

$\square$

12.8

Suppose that price and quantity are determined by the intersection of the linear demand and supply curves

$$\begin{aligned}
\text{Demand: } Q &= a_0 + a_1 P + a_2 Y + e_1 \\
\text{Supply: } Q &= b_0 + b_1 P + e_2
\end{aligned}$$

where income Y and wage W are determined outside the market. In this model, are the parameters identified?

Proof. Yes. All parameters can be identified.

We propose three numerically equivalent methods.

Method 1: 2SLS

Take the demand function as an example.

- 1st Stage: Regress P on instrument variables $\begin{bmatrix}1 & Y & W\end{bmatrix}$, and obtain the fitted value $\widehat{P}$.
- 2nd Stage: Regress Q on $\begin{bmatrix}1 & \widehat{P} & Y\end{bmatrix}$

The supply function can be consistently estimated in the same way.

Method 2: IV Methods

Like method 1, all estimators can be consistently estimated separately. For each regression, since the numbers of instrument variables and regressors are equal, IV method can be directly employed by formula (12.26) of Hansen's notes:

$$\begin{pmatrix}\hat{a}_0 & \hat{a}_1 & \hat{a}_2\end{pmatrix}' = (\boldsymbol{Z}'\boldsymbol{X}_1)^{-1}(\boldsymbol{Z}'\boldsymbol{Q})$$
$$\begin{pmatrix}\hat{b}_0 & \hat{b}_1 & \hat{b}_2\end{pmatrix}' = (\boldsymbol{Z}'\boldsymbol{X}_2)^{-1}(\boldsymbol{Z}'\boldsymbol{Q}),$$

where $\boldsymbol{X}_1$ includes $1, P$ and Y, $\boldsymbol{X}_2$ includes $1, P$ and W.

Method 3: Simultaneous Estimation

P and Q can be solved for:

$$\begin{cases} P &= \dfrac{b_0 - a_0}{a_1 - b_1} + \left(\dfrac{b_2}{a_1 - b_1}\right) W - \left(\dfrac{a_2}{a_1 - b_1}\right) Y + \dfrac{e_2 - e_1}{a_1 - b_1} \\ Q &= \dfrac{a_1 b_0 - a_0 b_1}{a_1 - b_1} + \left(\dfrac{a_1 b_2}{a_1 - b_1}\right) W - \left(\dfrac{a_2 b_1}{a_1 - b_1}\right) Y + \dfrac{a_1 e_2 - b_1 e_1}{a_1 - b_1}. \end{cases}$$

Observe that

$$
(1) \qquad \begin{cases} a_1 &= \dfrac{\mathsf{Cov}(Q,W)}{\mathsf{Cov}(P,W)} \\ \\ b_1 &= \dfrac{\mathsf{Cov}(Q,Y)}{\mathsf{Cov}(P,Y)} \end{cases}
$$

So natural estimators for a_1 and b_1 are:

$$
\begin{aligned}
\widehat{a}_1 &= \frac{\text{sample covariance btw Q and W}}{\text{sample covariance btw P and W}} \\
&= \frac{\sum_i (Q_i - \overline{Q})(W_i - \overline{W})}{\sum_i (P_i - \overline{P})(W_i - \overline{W})}.
\end{aligned}
$$

and

$$
\begin{aligned}
\widehat{b}_1 &= \frac{\text{sample covariance btw Q and Y}}{\text{sample covariance btw P and Y}} \\
&= \frac{\sum_i (Q_i - \overline{Q})(Y_i - \overline{Y})}{\sum_i (P_i - \overline{P})(Y_i - \overline{Y})}.
\end{aligned}
$$

Then a_0 and a_2 can be estimated by a simple regression of $(Q - \widehat{a}_1 P)$ on Y, while b_0 and b_2 can be estimated by a simple regression of $(Q - \widehat{b}_1 P)$ on W.

It is easy to show all estimators are consistent.

$\square$

12.10

Consider the model

$$
\begin{aligned}
y_i &= \boldsymbol{x}_i' \boldsymbol{\beta} + e_i \\
\boldsymbol{x}_i &= \boldsymbol{\Gamma}' \boldsymbol{z}_i + \boldsymbol{u}_i \\
\mathbb{E}[\boldsymbol{z}_i e_i] &= \boldsymbol{0} \\
\mathbb{E}[\boldsymbol{z}_i \boldsymbol{u}_i'] &= \boldsymbol{0}
\end{aligned}
$$

with y_i scalar and $\boldsymbol{x}_i$ and $\boldsymbol{z}_i$ each a k vector. You have a random sample $(y_i, \boldsymbol{x}_i, \boldsymbol{z}_i : i = 1, \ldots, n)$. Take the control function equation

$$e_i = \boldsymbol{u}_i' \boldsymbol{\gamma} + \varepsilon_i$$
$$\mathbb{E}[\boldsymbol{u}_i \varepsilon_i] = \mathbf{0}$$

and assume for simplicity that $\boldsymbol{u}_i$ is observed. Inserting into the structural equation we find

$$y_i = \boldsymbol{x}_i' \boldsymbol{\beta} + \boldsymbol{u}_i' \boldsymbol{\gamma} + \varepsilon_i.$$

The control function estimator $\left(\widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{\gamma}}\right)$ is OLS estimation of this equation.

(a) Show that $\mathbb{E}[\boldsymbol{x}_i \varepsilon_i] = \mathbf{0}$ (algebraially)

(b) Derive the asymptotic distribution of $\left(\widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{\Gamma}}\right)$.

Proof. **(a)**

By

$$\mathbb{E}[\boldsymbol{x}_i \varepsilon_i] = \boldsymbol{\Gamma}' \underbrace{\mathbb{E}[\boldsymbol{z}_i e_i]}_{=\mathbf{0}} - \boldsymbol{\Gamma}' \underbrace{\mathbb{E}[\boldsymbol{z}_i \boldsymbol{u}_i']}_{=\mathbf{0}} \boldsymbol{\gamma} + \underbrace{\mathbb{E}[\boldsymbol{u}_i \varepsilon_i]}_{=\mathbf{0}} = \mathbf{0}$$

(b)

Define $\boldsymbol{v}_i = \begin{pmatrix} \boldsymbol{x}_i \\ \boldsymbol{u}_i \end{pmatrix}$, then

$$\mathbb{E}[\boldsymbol{v}_i \varepsilon_i] = \mathbf{0}$$

and

$$y_i = \boldsymbol{v}_i' \begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\gamma} \end{pmatrix} + \varepsilon_i.$$

Then, the OLS estimator

$$\begin{pmatrix} \widehat{\boldsymbol{\beta}} \\ \widehat{\boldsymbol{\gamma}} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\gamma} \end{pmatrix} + \left(\sum_i \boldsymbol{v}_i \boldsymbol{v}_i' \right)^{-1} \left(\sum_i \boldsymbol{v}_i \varepsilon_i \right).$$

Its asymptotic distribution can be routinely established:

$$\sqrt{n} \left(\begin{pmatrix} \widehat{\boldsymbol{\beta}} \\ \widehat{\boldsymbol{\gamma}} \end{pmatrix} - \begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\gamma} \end{pmatrix} \right) \xrightarrow{\mathrm{d}} \mathcal{N} \left(\mathbf{0}, \boldsymbol{Q}_{vv}^{-1} \boldsymbol{\Omega} \boldsymbol{Q}_{vv}^{-1} \right)$$

where $\boldsymbol{Q}_{vv} = \mathbb{E}[\boldsymbol{v}_i \boldsymbol{v}_i']$ and $\boldsymbol{\Omega} = \mathbb{E}[\boldsymbol{v}_i \boldsymbol{v}_i' \varepsilon_i^2]$.

$\square$