Econometrics Practice

Date: 27/02/2007

OUTLINE

1.Efficient GMM (Proposition 3.5)

2.Variable Addition Test

3. Weak Instruments (Q10)

1. **Efficient GMM Proposition 3.5: Optimal Choice of the weighting matrix**

We will prove the proposition 3.5.11 in Hayashi Book p. 213, namely

A lower bound for the asymtotic variance of the GMM estimators indexed by $\hat{W}$ is given by $(\Sigma_{xz}^{'} W \Sigma_{xz})^{-1}$. This lower bound is achieved if $\hat{W}$ is such that plim $\hat{W}$ =W=S^{-1}

$$\begin{aligned}
\hat{\delta}(\hat{W}) &= (S_{xz}^{'} \hat{W} S_{xz})^{-1} S_{xz}^{'} \hat{W} s_{xy} \\
Avar(\hat{\delta}(\hat{W})) &= (\Sigma_{xz}^{'} W \Sigma_{xz})^{-1} \Sigma_{xz}^{'} W S W \Sigma_{xz} (\Sigma_{xz}^{'} W \Sigma_{xz})^{-1} \\
(\Sigma_{xz}^{'} W \Sigma_{xz})^{-1} \Sigma_{xz}^{'} W S W \Sigma_{xz} (\Sigma_{xz}^{'} W \Sigma_{xz})^{-1} &\geq (\Sigma_{xz}^{'} S^{-1} \Sigma_{xz})^{-1}
\end{aligned}$$

for any symmetric positive definite matrix W, where

$$\begin{aligned}
S &= E(g_i g_i^{'}), g_i = x_i \varepsilon_i \\
\Sigma_{xz} &: = \textit{population matrix of covariances} \text{ with full column rank} \\
S_{xz_{(K \times L)}} &= \frac{1}{n} \sum_{i=1}^{n} x_i z_i^{'}.
\end{aligned}$$

Proof. There are two ways to show the inequality

a) Using the property that if we have a symmetric and idempotent matrix A$\Rightarrow$positive semidefinite

$$\begin{aligned}
A^{'} &= A \\
AA &= A
\end{aligned}$$

then we will show that $\forall x$, $x^{'} A x \geq 0$, since $x^{'} AAx = x^{'} A^{'} Ax = z^{'} z \geq 0$, $z = Ax$. The fact we will exploit is the following

A $-$ B is positive semidefinite iff B^{-1} $-$ A^{-1} is positive semidefinite

We define $Q = \Sigma_{xz}^{'}S^{-1}\Sigma_{xz} - \Sigma_{xz}^{'}W\Sigma_{xz}(\Sigma_{xz}^{'}WSW\Sigma_{xz})^{-1}\Sigma_{xz}^{'}W\Sigma_{xz}$, where A=$(\Sigma_{xz}^{'}W\Sigma_{xz})^{-1}\Sigma_{xz}^{'}WSW\Sigma_{xz}(\Sigma_{xz}^{'}W\Sigma_{xz})^{-1}$,B=$(\Sigma_{xz}^{'}S^{-1}\Sigma_{xz})^{-1}$, so left to show that Q is positive semidefinite. Now we define

$$\begin{aligned} H &= C\Sigma_{xz} \\ S^{-1} &= C^{'}C \Longleftrightarrow C^{-1}(C^{'})^{-1} = S \\ G &= (C^{'})^{-1}W\Sigma_{xz} \end{aligned}$$

then we will write the Q in terms of the new variables defined, i.e.

$$\begin{aligned} Q &= \Sigma_{xz}^{'}S^{-1}\Sigma_{xz} - \Sigma_{xz}^{'}W\Sigma_{xz}(\Sigma_{xz}^{'}WSW\Sigma_{xz})^{-1}\Sigma_{xz}^{'}W\Sigma_{xz} \\ Q &= H^{'}H - \Sigma_{xz}^{'}C^{'}(C^{'})^{-1}W\Sigma_{xz}(\Sigma_{xz}^{'}WC^{-1}(C^{'})^{-1}W\Sigma_{xz})^{-1}\Sigma_{xz}^{'}WC^{-1}C\Sigma_{xz} \\ Q &= H^{'}H - H^{'}G(G^{'}G)^{-1}G^{'}H = H^{'}(I - G(G^{'}G)^{-1}G^{'})H = H^{'}M_GH \\ M_G &= (I - G(G^{'}G)^{-1}G^{'}) \text{ (annihilator)} \end{aligned}$$

since M_G is the annihilator matrix we have seen in the first chapter, it has the desired properties, namely it is symmetric and idempotent, therefore Q is positive semidefinite which completes the proof.

b) The second way exploits the fact that

$$\begin{aligned} a &\geq b \Leftrightarrow b^{-1} \geq a^{-1} \\ (\Sigma_{xz}^{'}W\Sigma_{xz})^{-1}\Sigma_{xz}^{'}WSW\Sigma_{xz}(\Sigma_{xz}^{'}W\Sigma_{xz})^{-1} &\geq (\Sigma_{xz}^{'}S^{-1}\Sigma_{xz})^{-1} \\ & if \text{ and only if} \\ \Sigma_{xz}^{'}S^{-1}\Sigma_{xz} &\geq \Sigma_{xz}^{'}W\Sigma_{xz}(\Sigma_{xz}^{'}WSW\Sigma_{xz})^{-1}\Sigma_{xz}^{'}W\Sigma_{xz} \end{aligned}$$

We will use the following three properties in the proof

i. A is positive semidefinite ⇔ Blocks on diagonal are positive semidefinite

ii. A is positive semidefinite ⇔ A^{-1} is positive semidefinite

iii. A is positive semidefinite ⇔ $B^{'}AB$ is positive semidefinite. (A symmetric)

We will use iii.⇒ii.⇒i. Let's define

$$\begin{aligned} B &= \begin{bmatrix} I & SW\Sigma_{xz} \\ 0 & 0 \end{bmatrix} \\ A &= \begin{bmatrix} S^{-1} & 0 \\ 0 & S^{-1} \end{bmatrix} \\ B'AB &= \begin{bmatrix} I & 0 \\ \Sigma_{xz}^{'}WS & 0 \end{bmatrix} \begin{bmatrix} S^{-1} & 0 \\ 0 & S^{-1} \end{bmatrix} \begin{bmatrix} I & SW\Sigma_{xz} \\ 0 & 0 \end{bmatrix} = \\ B'AB &= \begin{bmatrix} S^{-1} & W\Sigma_{xz} \\ \Sigma_{xz}^{'}W & \Sigma_{xz}^{'}WS^{-1}W\Sigma_{xz} \end{bmatrix} \end{aligned}$$

then we will use the following formula

$$D = \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix} \Rightarrow D^{-1} = \begin{bmatrix} (D_{11} - D_{12}D_{22}^{-1}D_{21})^{-1} & ... \\ ... & ... \end{bmatrix}$$

so

$$\begin{aligned} D &= \begin{bmatrix} S^{-1} & W\Sigma_{xz} \\ \Sigma_{xz}^{'}W & \Sigma_{xz}^{'}WS^{-1}W\Sigma_{xz} \end{bmatrix} \\ D^{-1} &= \begin{bmatrix} (S^{-1} - W\Sigma_{xz}(\Sigma_{xz}^{'}WSW\Sigma_{xz})^{-1}\Sigma_{xz}^{'}W)^{-1} & ... \\ ... & ... \end{bmatrix} \end{aligned}$$

following the logical order iii.⇒ii.⇒i. , we see that the positive semidefineteness of the diagonal element implies that

$$\begin{aligned} S^{-1} &\geq W\Sigma_{xz}(\Sigma_{xz}^{'}WSW\Sigma_{xz})^{-1}\Sigma_{xz}^{'}W \Rightarrow \\ \Sigma_{xz}^{'}S^{-1}\Sigma_{xz} &\geq \Sigma_{xz}^{'}W\Sigma_{xz}(\Sigma_{xz}^{'}WSW\Sigma_{xz})^{-1}\Sigma_{xz}^{'}W\Sigma_{xz} \end{aligned}$$

this observation completes the proof.

2. **Variable Addition Test**: Testing a subset of the orthogonality conditions.

Assume we have the following model in our mind

$$y_i = z_i^{'}\delta + \varepsilon_i$$

we split the the instruments into two subgroups

$$x_i = \begin{bmatrix} x_{i1} \\ x_{i2} \end{bmatrix}$$

for simplification we assume that all the regressors are predetermined and equal to the first subgroup, i.e.

$$\begin{aligned} z_i &= x_{i1} \\ y_i &= x_{i1}^{'}\delta + \varepsilon_i \end{aligned}$$

using the whole set of the instruments we have the following *augmented model*

$$y_i = x_{i1}^{'}\delta + x_{i2}^{'}\alpha + \varepsilon_i = x_i^{'}\gamma + \varepsilon_i$$

We have the following null hypothesis

$$H_0 : x_{i2} \text{ predetermined} \equiv H_0 : \alpha = 0$$

the two hypothesis are equivalent, because if x_{i2} predetermined, then their influence on the dependent variable should be not direct, but through x_{i1}

$$x_{i1} = x_{i2}\beta + \nu_i$$

example 3.3 on page 221 describes such a situation, whether schooling is predetermined in the wage equation. From proposition 3.7, we know that we have to calculate the J statistics. We first run the OLS regression on the augmented model

$$\hat{\gamma}_{OLS} = S_{xx}^{-1} s_{xy}$$

then J statistic becomes

$$\begin{aligned} J_1 &= n g_n(\hat{\gamma}_{OLS})' \hat{S}^{-1} g_n(\hat{\gamma}_{OLS}) \\ g_n(\hat{\gamma}) &= (s_{xy} - S_{xx}\hat{\gamma}_{OLS}) \\ J_1 &= n(s_{xy} - S_{xx}\hat{\gamma}_{OLS})' \hat{S}^{-1}(s_{xy} - S_{xx}\hat{\gamma}_{OLS}) \end{aligned}$$

substituting $\hat{\gamma}_{OLS} = S_{xx}^{-1} s_{xy}$

$$J_1 = n(s_{xy} - S_{xx}S_{xx}^{-1}s_{xy})' \hat{S}^{-1}(s_{xy} - S_{xx}S_{xx}^{-1}s_{xy}) = 0$$

The second statistic is obtained by $\widetilde{\tilde{\gamma}}$ that minimizes $J(\tilde{\gamma}, \hat{S}^{-1}) = n(s_{xy} - S_{xx}\tilde{\gamma})' \hat{S}^{-1}(s_{xy} - S_{xx}\tilde{\gamma})$ s.t $\alpha = 0$, where

$$\widetilde{\tilde{\gamma}} = \begin{bmatrix} \widetilde{\delta}(\hat{S}^{-1}) \\ 0 \end{bmatrix}$$

Since J1=0, the likelihood ratio test LR becomes

$$\begin{aligned} LR &= n(s_{xy} - S_{xx}\widetilde{\tilde{\gamma}})' \hat{S}^{-1}(s_{xy} - S_{xx}\widetilde{\tilde{\gamma}}) \\ LR &= n(s_{xy} - S_{xx_1}\widetilde{\delta}(\hat{S}^{-1}))' \hat{S}^{-1}(s_{xy} - S_{xx_1}\widetilde{\delta}(\hat{S}^{-1})) \end{aligned}$$

where

$$\begin{aligned} S_{xx} &= \begin{bmatrix} S_{x_1x_1} & S_{x_1x_2} \\ S_{x_2x_1} & S_{x_2x_2} \end{bmatrix} \\ S_{xx}\widetilde{\tilde{\gamma}} &= \begin{bmatrix} S_{x_1x_1}\widetilde{\delta}(\hat{S}^{-1}) \\ S_{x_2x_1}\widetilde{\delta}(\hat{S}^{-1}) \end{bmatrix} = S_{xx_1}\widetilde{\delta}(\hat{S}^{-1}) \\ S_{xx} &= \begin{bmatrix} S_{xx_1} & S_{xx_2} \end{bmatrix} \end{aligned}$$

We have shown that the LR test becomes the *Hansen's (p.217) J statistics* for restricted regression.

$$\begin{aligned} y_i &= x_{i1}'\delta + \varepsilon_i \\ \text{with } x_i &= \begin{bmatrix} x_{i1} \\ x_{i2} \end{bmatrix} \text{ Instruments} \end{aligned}$$

3. **Weak Instruments:** We have the following simple model

$$y_i = z_i\delta + \varepsilon_i \quad \text{i=1....n}$$

where y_i, z_i and ε_i are scalar random variables and we have one single instrument x_i s.t. $z_i = x_i\beta + \nu_i$. We also define

$$\begin{aligned} \eta_i &= \begin{bmatrix} \varepsilon_i \\ \nu_i \end{bmatrix} \\ g_i &= \eta_i x_i = \begin{bmatrix} g_{1i} \\ g_{2i} \end{bmatrix} = \begin{bmatrix} x_i\varepsilon_i \\ x_i\nu_i \end{bmatrix} \end{aligned}$$

We have the following *assumptions*

i. $\{x_i, \eta_i\}$ is ergodic stationary

ii. g_i is martingale difference sequence, a fortiori $E(g_i) = 0$, with $E(g_i g_i^{'}) = S$ positive definite

iii. $E(x_i^2) = \sigma_x^2 > 0$

iv. $\beta \neq 0$

Then

$$\begin{aligned} \hat{\delta} &= S_{xz}^{-1} s_{xy} \\ S_{xz} &= \frac{1}{n}\sum_{i=1}^{n} x_i z_i \\ s_{xy} &= \frac{1}{n}\sum_{i=1}^{n} x_i y_i \end{aligned}$$

Given our assumptions we have the following observations

1. $\sigma_{xz} = E(x_i z_i) \neq 0$

$$E(x_i z_i) = E(x_i(x_i\beta + \nu_i)) = E(x_i^2)\beta + E(x_i\nu_i)$$

by assumption iii

$$E(x_i^2)\beta = \sigma_x^2\beta \neq 0$$

by assumption ii (m.d.s) since $E(g_i) = 0$

$$E(x_i\nu_i) = 0$$

thus

$$E(x_i z_i) = \sigma_x^2\beta \neq 0$$

2. $\hat{\delta} \to_p \delta, where\ \hat{\delta} = S_{xz}^{-1} s_{xy}$

   Notice that z_i is ergodic stationary, since by assumption i x_i is ergodic stationary. So, assumption i guarantees that

$$S_{xz} \to_p \sigma_{xz} = \sigma_x^2 \beta$$

   recalling the sampling error formula

$$\hat{\delta} - \delta = S_{xz}^{-1} s_{x\varepsilon} = (\frac{1}{n}\sum_{i=1}^{n} x_i z_i)^{-1} (\frac{1}{n}\sum_{i=1}^{n} x_i \varepsilon_i) \to_p 0$$

   since the first term convergences in probability $(\sigma_x^2\beta)^{-1}$(by Lemma 2.3b) and the second term convergences in probability to 0, hence the sampling error as well converges in probability to 0. (by Lemma 2.4b)

   to derive the following results we replace the assumption iii with

   iii$'$ $\beta = \frac{1}{\sqrt{n}}$

3. $S_{xz} \to_p 0$

$$\begin{aligned}
S_{xz} &= \frac{1}{n}\sum_{i=1}^{n} x_i z_i =^{iii'} \frac{1}{n}\sum_{i=1}^{n} x_i (x_i \frac{1}{\sqrt{n}} + \nu_i) = \\
&= \frac{1}{\sqrt{n}} (\frac{1}{n}\sum_{i=1}^{n} x_i^2) + \frac{1}{n}\sum_{i=1}^{n} x_i \nu_i \to_p 0 \\
(\frac{1}{n}\sum_{i=1}^{n} x_i^2) &\to {}_p \frac{\sigma_x^2}{\sqrt{n}} \to_p 0 \text{ as } n \to \infty \\
\frac{1}{n}\sum_{i=1}^{n} x_i \nu_i &\to {}_p 0 \\
S_{xz} &\to {}_p 0 + 0 = 0
\end{aligned}$$

4. $\sqrt{n} S_{xz} \to_d \sigma_x^2 + a, \quad a \sim N(0, s_{22})$ where s_{22} (2,2) element of S

$$\begin{aligned}
\sqrt{n}(\frac{1}{n}\sum_{i=1}^{n} x_i z_i) &= \sqrt{n}(\frac{1}{n}\frac{1}{\sqrt{n}}\sum_{i=1}^{n} x_i^2) + \sqrt{n}(\frac{1}{n}\sum_{i=1}^{n} x_i \nu_i) = \\
&= (\frac{1}{n}\sum_{i=1}^{n} x_i^2) + \sqrt{n}\bar{g}_2 \\
\begin{bmatrix} \bar{g}_1 \\ \bar{g}_2 \end{bmatrix} &= \begin{bmatrix} \frac{1}{n}\sum_{i=1}^{n} x_i \varepsilon_i \\ \frac{1}{n}\sum_{i=1}^{n} x_i \nu_i \end{bmatrix} \\
(\frac{1}{n}\sum_{i=1}^{n} x_i^2) &\to {}_p \sigma_x^2 \text{ by ergodic stationarity} \\
\sqrt{n}\bar{g}_2 &\to {}_d N(0, s_{22}) \text{ by m.d.s + CLT}
\end{aligned}$$

so we have shown that

$$\sqrt{n}S_{xz} \to_d \sigma_x^2 + N(0, s_{22})$$

Question: Is $\hat{\delta} = S_{xz}^{-1} s_{xy}$ *consistent* for δ?
Answer: No.

$$\begin{aligned}
\hat{\delta} - \delta &= (\sqrt{n}S_{xz})^{-1}(\sqrt{n}\bar{g}_1) \to_d (\sigma_x^2 + a)^{-1} b \\
(\sqrt{n}S_{xz})^{-1} &\to {}_d(\sigma_x^2 + a)^{-1} \\
(\sqrt{n}\bar{g}_1) &\to {}_d N(0, s_{11}) = b
\end{aligned}$$