Econometrics Practice

Date: 28/02/2007

OUTLINE

1.Hausman Principle in GMM

2.Taylor Expansion of the Sampling Error

1. **Hausman Principle in GMM: Testing Subset of Orthogonality Conditions under conditional Homoscedasticity**

First we have solved the 9th analytical question at the end of chapter 3 of Hayashi Book, the answer can be found in Hayashi's website. Using that result, we derive the *Hausman Statistic* introduced in page 233 (3.8.22). We have the same notation as before, the matrix $Z_{(n\times p)}$ of regressors and the matrix $X_{(n\times K)}$ of full set of instruments. We can split the the subset of instruments into two subsets $X = [X_1\ X_2]$, where $X_{1(n\times K_1)}$ is the subset of the instruments that we trust, i.e we think that they are predetermined. Since under conditional homoscedastcity assumption the two-step GMM estimator is the 2SLS estimator, we obtain two GMM estimators, one using the full set of instruments and one with the first subset of the instruments, respectively

$$\begin{aligned}\hat{\delta} &= (Z^{'}PZ)^{-1}Z^{'}Py, \quad P = X(X^{'}X)^{-1}X^{'} \\ \bar{\delta} &= (Z^{'}P_1Z)^{-1}Z^{'}P_1y, \quad P_1 = X_1(X_1^{'}X_1)^{-1}X_1^{'}\end{aligned}$$

Since the GMM estimator $\hat{\delta}$ that uses the whole set of instrumets exploits more orthogonality conditions, it is more asymtotically more efficient than the one obtained using only the first subset, i.e.

$$Avar(\bar{\delta}) \geq Avar(\hat{\delta})$$

Exploiting the result we have derived in the analytical question we have

$$Avar(\bar{\delta} - \hat{\delta}) = Avar(\bar{\delta}) - Avar(\hat{\delta})$$

Under conditional homoscedasticity assumption $Avar(\hat{\delta})$ can be consistently estimated by

$$\widehat{Avar(\hat{\delta})} = n.\hat{\sigma}^2(Z^{'}PZ)^{-1}$$

also under the same assumption $Avar(\bar{\delta})$ can be consistently estimated by

$$\widehat{Avar(\bar{\delta})} = n.\hat{\sigma}^2 (Z^{'} P_1 Z)^{-1}$$

notice that the same estimator $\hat{\sigma}^2$,s used in both consistent estimators. Then the resulting estimator of $Avar(\bar{\delta} - \hat{\delta})$ becomes

$$\widehat{Avar(\bar{\delta} - \hat{\delta})} = n.\hat{\sigma}^2 ((Z^{'} P_1 Z)^{-1} - (Z^{'} P Z)^{-1})$$

Then the Hausman Statistics becomes

$$\begin{aligned}
H &= \sqrt{n}(\bar{\delta} - \hat{\delta})^{'} \{\widehat{Avar(\bar{\delta} - \hat{\delta})}\} \sqrt{n}(\bar{\delta} - \hat{\delta}) \\
H &= \sqrt{n}(\bar{\delta} - \hat{\delta})^{'} \{n.\hat{\sigma}^2 ((Z^{'} P_1 Z)^{-1} - (Z^{'} P Z)^{-1})\}^{-1} \sqrt{n}(\bar{\delta} - \hat{\delta}) \\
H &= \frac{(\bar{\delta} - \hat{\delta})^{'} \{((Z^{'} P_1 Z)^{-1} - (Z^{'} P Z)^{-1})\}^{-1} (\bar{\delta} - \hat{\delta})}{\hat{\sigma}^2}
\end{aligned}$$

this statistic is asymtotically chi-squared distributed with min(K-K_1,L-s) degrees of freedom, where

$$\begin{aligned}
K &= \# \text{ of instruments} \\
K_1 &= \# \text{ of instruments in } X_1 \\
L &= \# \text{ of regressors} \\
s &= \# \text{ of regressors that are retained in } X_1.
\end{aligned}$$

2. **Taylor Expansion of the Sampling Error for GMM**: First we recall that the objective function for GMM is

$$\begin{aligned}
Q_n(\theta) &= -\frac{1}{2} g_n(\theta)^{'} \hat{W} g_n(\theta) \\
g_n(\theta)_{K \times 1} &= \frac{1}{n} \sum_{t=1}^{n} g(w_t; \theta)
\end{aligned}$$

to derive the asymtotic normality of GMM estimator, we will exploit the Mean Value Theorem to $g_n(\theta)$ as opposed to the case in M-Estimation where the theorem was applied to first order condition of the maximization. The first order condition for maximization of the objective function in GMM is the following expression

$$0_{p \times 1} = \frac{\partial Q_n(\hat{\theta})}{\partial \theta} = -G_n(\hat{\theta})^{'}_{(p \times K)}.\hat{W}_{(K \times K)}.g_n(\hat{\theta})_{(K \times 1)}$$

where $G_n(\theta)$ is the Jacobian of $g_n(\theta)$:

$$G_n(\theta)_{(K \times p)} = \frac{g_n(\theta)}{\partial \theta^{'}}$$

If we apply the Mean value theorem to $g_n(\theta)$, we obtain the mean-value expansion

$$g_n(\hat{\theta}) = g_n(\theta_0) + G_n(\bar{\theta})(\hat{\theta} - \theta_0)$$

where θ_0 is the true value, $\hat{\theta}$ is the estimator and $\bar{\theta}$ is a value between the former two, i.e. $\bar{\theta} \in [\ \theta_0, \hat{\theta}]$. Then, we can plug in this expression into the first order condition and solve for the sampling error and we obtain

$$\begin{aligned}
\sqrt{n}(\hat{\theta} - \theta_0) &= -B^{-1}c. \quad (7.3.32) \\
B_{(p\times p)} &= -G_n(\hat{\theta})'_{(p\times K)}\hat{W}_{(K\times K)}G_n(\bar{\theta})_{(K\times p)} \\
c_{(p\times 1)} &= -G_n(\hat{\theta})'_{(p\times K)}\hat{W}_{(K\times K)}\frac{1}{\sqrt{n}}\sum_{t=1}^{n} g(w_t;\theta_0)_{(K\times 1)} \\
G_n(\hat{\theta})_{(K\times p)} &= \frac{g_n(\hat{\theta})}{\partial\theta'}_{|\theta=\hat{\theta}}
\end{aligned}$$

Our task is to show the same result (7.3.32) using the Taylor expansion intead of the Mean Value Theorem under the 5 assumptions of proposition 7.10 that states the asymtotic normality of GMM. (For assumptions, see page 480) In other words, we aim to show that the asymtotic distribution of the nonlinear GMM estimation can also be obtained by taling the linear (Taylor) approximation of $g_n(\theta)$ around the true parameter value.

Formally we need to show

$$\begin{aligned}
\sqrt{n}(\hat{\theta} - \theta_0) &= -\psi^{-1}\sqrt{n}\frac{\partial Q_n(\theta_0)}{\partial\theta'} + o_p \\
\psi &= -G'WG \\
\text{plim } G_n(\theta_0 &= G) \\
\text{plim } \hat{W} &= W
\end{aligned}$$

Proof.

a) Let $B^{-1} = \psi^{-1} + Y_n$ where $Y_n \to_p 0$. We apply the Taylor expansion to B

$$B = -G_n(\hat{\theta})'\hat{W}G_n(\bar{\theta}) = -G_n(\theta_0)'\hat{W}G_n(\bar{\theta}) - \frac{\partial G_n(\bar{\bar{\bar{\theta}}})}{\partial\theta'}\hat{W}G_n(\bar{\theta})(\hat{\theta} - \theta_0)$$

where $\bar{\bar{\bar{\theta}}} \in [\theta_0, \hat{\theta}]$. We know that the GMM estimator is consistent, i.e. $\hat{\theta} \to_p \theta_0$, since both $\bar{\theta}, \bar{\bar{\bar{\theta}}} \in [\ \theta_0, \hat{\theta}]$, it also holds that $\bar{\theta} \to_p \theta_0, \bar{\bar{\bar{\theta}}} \to_p \theta_0$, we further assume that we have a consistent estimate $\hat{W}$, i.e. $\hat{W} \to_p W$, also since we have ergodic stationarity assumption and have consistent estimators

$$\begin{aligned}
G_n(\hat{\theta}) &= \frac{1}{n}\sum_{t=1}^{n}\frac{\partial g(w_t;\hat{\theta})}{\partial\theta'} \to_p G_n(\theta_0) = \frac{1}{n}\sum_{t=1}^{n}\frac{\partial g(w_t;\theta_0)}{\partial\theta'} \to_p G \\
G_n(\bar{\theta}) &= \frac{1}{n}\sum_{t=1}^{n}\frac{\partial g(w_t;\bar{\theta})}{\partial\theta'} \to_p G_n(\theta_0) = \frac{1}{n}\sum_{t=1}^{n}\frac{\partial g(w_t;\theta_0)}{\partial\theta'} \to_p G
\end{aligned}$$

where G is a finite value, a number, e.g. $E(\frac{\partial g(w_t;\theta_0)}{\partial \theta^{'}})$. Finally, by the second assumption of proposition 7.10, we have

$$\frac{\partial G_n(\overline{\overline{\overline{\theta}}})}{\partial \theta^{'}} = \frac{1}{n}\sum_{t=1}^{n}\frac{\partial^2 g(w_t;\overline{\overline{\overline{\theta}}})}{\partial \theta \partial \theta^{'}} \rightarrow_p E(\frac{\partial g(w_t;\theta_0)}{\partial \theta \partial \theta^{'}}) = H^{'} < \infty$$

since $\overline{\overline{\overline{\theta}}} \rightarrow_p \theta_0$ and $g(w_t;\theta)$ is continuously differentiable in θ for any w_t. Hence, we obtain

$$\begin{aligned} \text{plim B} &= -G^{'}WG - H^{'}WG\text{plim}(\hat{\theta}-\theta_0) = -G^{'}WG + o_p \\ B^{-1} &= \psi^{-1} + o_p = \psi^{-1} + Y_n \\ \text{plim } B^{-1} &= \psi^{-1} + 0. \end{aligned}$$

b) We also need to show that

$$\begin{aligned} c &= -G_n(\hat{\theta})^{'}\hat{W}\frac{1}{\sqrt{n}}\sum_{t=1}^{n} g(w_t;\theta_0) = \sqrt{n}\frac{\partial Q_n(\theta_0)}{\partial \theta^{'}} + y_n \\ y_n &\rightarrow {}_p 0 \end{aligned}$$

From the FOC of GMM objective function maximization we know that

$$\sqrt{n}\frac{\partial Q_n(\theta_0)}{\partial \theta^{'}} = -G_n(\theta_0)^{'}\hat{W}\frac{1}{\sqrt{n}}\sum_{t=1}^{n} g(w_t;\theta_0)$$

so we can write

$$c - \sqrt{n}\frac{\partial Q_n(\theta_0)}{\partial \theta^{'}} = -(G_n(\hat{\theta}) - G_n(\theta_0))^{'}\hat{W}\frac{1}{\sqrt{n}}\sum_{t=1}^{n} g(w_t;\theta_0)$$

we know from assumption 3 of proposition 7.10 that

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n} g(w_t;\theta_0) \rightarrow_d N(0, S_{(K\times K)})$$

so we also know from the properties of normal distribution

$$\hat{W}\frac{1}{\sqrt{n}}\sum_{t=1}^{n} g(w_t;\theta_0) \rightarrow_d N(0, WSW)$$

If we focus on $-(G_n(\hat{\theta}) - G_n(\theta_0))^{'}$, using the add/substract trick we can show that

$$(G_n(\hat{\theta}) - G_n(\theta_0) - G + G = [(G_n(\hat{\theta}) - G) - (G_n(\theta_0) - G)]$$

given that plim $G_n(\hat{\theta}) = G_n(\theta_0) = G$, this term converges in probability to 0. Hence we can write,

$$c = \sqrt{n}\frac{\partial Q_n(\theta_0)}{\partial \theta^{'}} + o_p \Rightarrow \sqrt{n}\frac{\partial Q_n(\theta_0)}{\partial \theta^{'}} + y_n$$

Combining the results in section (a) and (b),

$$\begin{aligned} B^{-1} &= \psi^{-1} + Y_n, \quad Y_n \rightarrow_p 0 \\ c &= \sqrt{n}\frac{\partial Q_n(\theta_0)}{\partial \theta^{'}} + y_n, \; y_n \rightarrow_p 0 \end{aligned}$$

we complete the proof by showing

$$\begin{aligned} \sqrt{n}(\hat{\theta} - \theta_0) &= -B^{-1}c = -(\psi^{-1} + Y_n)(\sqrt{n}\frac{\partial Q_n(\theta_0)}{\partial \theta^{'}} + y_n) \\ &= -\psi^{-1}\sqrt{n}\frac{\partial Q_n(\theta_0)}{\partial \theta^{'}} + o_p. \end{aligned}$$