PUBLICACIONES DE 4er CURSO

Grado: Economía

Asignatura: ECONOMETRÍA III

Tema 3: Evaluación
Apartado 3.4: Discriminación de Modelos Econométricos

Grupos:

Profesores: Antonio Aznar y M. Teresa Aparicio

Departamento de ANÁLISIS ECONÓMICO

Curso Académico 2014/15

Facultad de
Economía y Empresa
Universidad Zaragoza

Apartado 3.4
Discriminación de Modelos Econométricos

Los resultados que vamos a presentar en este apartado se han derivado dentro de un marco con dos modelos lineales anidados lo cual, creemos, no resta generalidad al planteamiento.

Los dos modelos lineales anidados pueden escribirse como:

$$M1: y = X_1\beta_1 + u_1 \tag{3.4.1}$$

$$M2: y = X_2\beta_2 + u_2 \tag{3.4.2}$$

en donde y es el vector Tx1 de observaciones de la variable dependiente; X_2 es la matriz Txk_2 de observaciones de las k_2 variables que aparecen en M_2 y X_1 es la matriz Txk_1 de observaciones de las k_1 variables que aparecen en M1; por el carácter anidado de ambos modelos se tiene que: $X_2 = (X_1, X^*)$ en donde X^* es una matriz Tx $(k_2- k_1)$ de las observaciones de las (k_2-k_1) variables que aparecen en M2 pero no en M1; u_1 y u_2 son, cada uno de ellos, vectores de T perturbaciones aleatorias. Cuando asumimos que el Proceso Generador de Datos (PGD) es M_1 entonces se cumple que:

$$u_1 \sim N(0, \sigma_1^2 I_T)$$

y cuando suponemos que el PGD es M2 entonces se tiene:

$$u_2 \sim N(0, \sigma_2^2 I_T)$$

El modelo M2 puede escribirse como:

$$y = X_1\beta_1 + X^*\beta^* + u_2 \tag{3.4.3}$$

Asumiremos que se dispone de T_1 observaciones extramuestrales y que, para cada una de ellas, podemos escribir:

$$M1: y_p = x'_{p1}\beta_1 + u_{p1} \qquad p = 1, ..., T_1 \tag{3.4.4}$$

M2: $y_p = x_{p2}'\beta_2 + u_{p2} = x_{p1}'\beta_1 + x_p^{*'}\beta^* + u_{p2}$ $\quad p = 1, ..., T_1$

(3.4.5)

Si el PGD es M_1 entonces $u_{p1} \sim N\left(0, \sigma_1^2 I_{T_1}\right)$ y se distribuye independientemente del resto de las perturbaciones de los periodos extramuestrales de M1 y de los elementos del vector u_1. Si el PGD es M2 entonces $u_{p2} \sim N\left(0, \sigma_2^2 I_{T_1}\right)$ y se distribuye independientemente del resto de las perturbaciones de los periodos extramuestrales de M2 y de los elementos de u_2.

Para cada modelo se define el vector de residuos MCO como:

$$\hat{u}_i = y - X_i\hat{\beta}_i \tag{3.4.6}$$

con
$$\hat{\beta}_i = (X_i'X_i)^{-1}X_i'y \qquad i = 1, 2 \tag{3.4.7}$$

Los estimadores MCO de las varianzas de las perturbaciones de los modelos vienen dados por:

$$\hat{\sigma}_i^2 = \frac{\hat{u}_i'\hat{u}_i}{T - k_i} \qquad i = 1, 2 \tag{3.4.8}$$

y los estimadores Máximo-Verosímiles (MV) por:

$$\tilde{\sigma}_i^2 = \frac{\hat{u}_i'\hat{u}_i}{T} \qquad i = 1, 2 \tag{3.4.9}$$

Para cada modelo y periodo extramuestral obtenemos el predictor MCO y el correspondiente error de predicción que escribiremos como:

$$e_{pi} = y_p - \hat{y}_{pi} = y_p - x_{pi}'\hat{\beta}_i \qquad i = 1, 2 \tag{3.4.10}$$

en donde suponemos que para cada período extramuestral que hemos denotado con el subíndice p, cuando se calculan los estimadores MCO de β_i, i = 1, 2, utilizando la expresión escrita en

(3.4.7)

Nuestro punto de partida es que los dos modelos son esféricos, es decir, que las perturbaciones de los dos modelos no tienen correlación serial y son homoscedásticas. Una vez que se garantiza el cumplimiento de esta característica, en la literatura econométrica se ha propuesto el uso de determinados criterios para comparar modelos alternativos.

En el marco de este trabajo sería imposible analizar las características de todos los criterios desarrollados en dicha literatura. Nosotros vamos a limitar nuestro análisis a los criterios que aparecen en el Cuadro 3.4.1.

En este cuadro aparece la denominación del criterio, el estadístico en el que se basa y la regla de decisión que se utiliza. Todos los términos que aparecen en el Cuadro han sido ya definidos; simplemente, indicar que F_{ε} $[(k_2-k_1), T-k_2]$ y $\chi^2_{\varepsilon}(k_2 - k_1)$ hacen referencia a los puntos críticos correspondientes al nivel de significación y grados de libertad indicados.

La justificación de la mayor parte de los criterios que aparecen en el Cuadro 3.4.1 puede encontrarse en Aznar (1989) y Aznar y Trívez (1993). La justificación de los criterios PIC y PICF dentro de un marco bayesiano puede verse en Phillips (1996) y en Phillips y Ploberger (1994, 1996). Estos autores derivan el criterio PIC como una aproximación asintótica a la densidad predictiva de la razón de verosimilitud de las dos hipótesis que se contrastan. El criterio PICF se obtiene condicionando la aproximación anterior a la información correspondiente a un primer periodo muestral.

CUADRO 3.4.1. Criterios y Regla de Decisión.

Criterio	**Estadístico**	**Regla de decisión: se rechaza M1 frente a M2 cuando:**
Coeficiente de Determinación	$R_i^2 = 1 - \frac{\hat{u}_i'\hat{u}_i}{\sum(y_t - \bar{y})^2}$	$R_1^2 < R_2^2$
Cef.Det.Corregido	$\bar{R}_i^2 = 1 - \frac{T-1}{T-k_i}\left(\frac{\hat{u}_i'\hat{u}_i}{\sum(y_t - \bar{y})^2}\right)$	$\bar{R}_1^2 < \bar{R}_2^2$
Contrastes t y F	$F = \frac{(\hat{u}_1'\hat{u}_1 - \hat{u}_2'\hat{u}_2)/(k_2 - k_1)}{\hat{u}_2'\hat{u}_2/(T-k_2)}$	F> $F_\varepsilon[(k_2 - k_1), T - k_2]$
Wald	$W = T\frac{\tilde{\sigma}_1^2 - \tilde{\sigma}_2^2}{\tilde{\sigma}_2^2}$	$W > \chi_\varepsilon^2(k_2 - k_1)$
Multiplicadores de Lagrange	$LM = T\frac{\tilde{\sigma}_1^2 - \tilde{\sigma}_2^2}{\tilde{\sigma}_1^2}$	$LM > \chi_\varepsilon^2(k_2 - k_1)$
Razón de Verosimilitud	$LR = T\ln\left(\frac{\tilde{\sigma}_1^2}{\tilde{\sigma}_2^2}\right)$	$LR > \chi_\varepsilon^2(k_2 - k_1)$
C_p de Mallows	$C_{pi} = \frac{\tilde{\sigma}_1^2}{\tilde{\sigma}_2^2} + \frac{2k_i}{T-k_2}$	$C_{p1} > C_{p2}$
Akaike	$AIC_i = \ln\tilde{\sigma}_i^2 + \frac{2k_i}{T}$	$AIC_1 > AIC_2$
Densidad Predictiva	$PIC_i = \hat{u}_i'\hat{u}_i + \frac{(k_2 - k_1)\hat{u}_2'\hat{u}_2}{T-k_2}\ln\tilde{\sigma}_2^2 + $ $+ \frac{\hat{u}_2'\hat{u}_2}{T-k_2}\ln(X_i'X_i)$	$PIC_1 > PIC_2$
Información	$BEC_i = \tilde{\sigma}_i^2 + \frac{k_i \ln T}{T - K_2}\tilde{\sigma}_2^2$	$BEC_1 > BEc_2$
Schwarz	$SBIC_i = \ln\tilde{\sigma}_i^2 + \frac{k_i \ln T}{T}$	$SBIC_1 > SBIC_2$
Error Cuadrático Medio de Predicción	$ECMP_i = \frac{1}{T_1}\sum_{p=1}^{T_1} e_{pi}^2$	$ECMP_1 > ECMP_2$

Densidad Predictiva Condicional	$\mathrm{PICF_i} = \sum_{P=1}^{T_i} \ln\left(\tilde{\sigma}^2_{1p}\left(1+C_{pi}\right)\right) + \sum_{p=1}^{T_i} \frac{e^2_{pi}}{\tilde{\sigma}^2_{pi}\left(1+C_{pi}\right)}$	$PICF_1 > PICF_2$

A continuación, vamos a formular, de forma alternativa, la regla de decisión expresando cada criterio como una combinación de un factor de ajuste y de un factor de parsimonia, destacando cómo es este factor de parsimonia el que marca la diferencia entre los criterios.

La expresión genérica que nos va a servir para formular la regla de decisión de todos los criterios es la siguiente: se rechaza el modelo M1 frente al modelo M2 cuando:

$$\hat{u}_1^{'}\hat{u}_1 > \hat{u}_2^{'}\hat{u}_2 \,.h(\ ) \tag{3.4.12}$$

h( ) es lo que hemos llamado factor de parsimonia. Es una función cuyos argumentos dependen del criterio que se considere pero que siempre toma un valor superior a 1 y es creciente de la medida que se toma del tamaño del modelo.

La forma que adopta h( ) para todos los criterios formulados en términos de las sumas de cuadrados de residuos puede verse en la columna 2 del Cuadro 3.4.2. Obviando los criterios $\bar{R}^2$ y PIC se pueden distinguir tres grupos: el primero está formado por los criterios F, C_p y BEC; el segundo está formado por los criterios W y LM y el tercero abarca a los criterios LR, AIC, SBIC. Derivemos ahora h( ) para el $\bar{R}^2$ y uno de cada uno de los tres grupos.

Utilizando el criterio $\bar{R}^2$ se rechaza M1 frente a M2 cuando:

$$1 - \frac{T-1}{T-k_1} \cdot \frac{\hat{u}_1^{'}\hat{u}_1}{\Sigma(y_t - \bar{y})^2} < 1 - \frac{T-1}{T-k_2} \cdot \frac{\hat{u}_2^{'}\hat{u}_2}{\Sigma(y_t - \bar{y})^2}$$

Eliminando términos comunes y agrupando se obtiene:

$$\hat{u}_1'\hat{u}_1 > \hat{u}_2'\hat{u}_2 \cdot \frac{T-k_1}{T-k_2}$$

Utilizando el criterio W, se rechaza M1 frente a M2 cuando:

$$T\,\frac{\hat{u}_1'\hat{u}_1 - \hat{u}_2'\hat{u}_2}{\hat{u}_2'\hat{u}_2} > \chi_\varepsilon^2(k_2-k_1)$$

que es equivalente a:

$$\hat{u}_1'\hat{u}_1 > \hat{u}_2'\hat{u}_2 \cdot \left(1+\frac{\chi_\varepsilon^2}{T}\right)$$

Para el criterio F, se rechaza M1 cuando:

$$\frac{(\hat{u}_1'\hat{u}_1 - \hat{u}_2'\hat{u}_2)/(k_2-k_1)}{\hat{u}_2'\hat{u}_2/(T-k_2)} > F_\varepsilon[(k_2-k_1),T-k_2]$$

que es equivalente a:

$$\frac{\hat{u}_1'\hat{u}_1}{\hat{u}_2'\hat{u}_2} > 1+\frac{k_2-k_1}{T-k_2}F_\varepsilon$$

o bien

$$\hat{u}_1'\hat{u}_1 > \hat{u}_2'\hat{u}_2\left(1+\frac{k_2-k_1}{T-k_2}F_\varepsilon\right)$$

Por último, con el criterio LR se rechaza M1 frente a M2 cuando:

$$T\ln\left(\frac{\hat{u}_1'\hat{u}_1}{\hat{u}_2'\hat{u}_2}\right) > \chi_\varepsilon^2\,(k_2-k_1)$$

Esta expresión es equivalente a:

$$\left(\frac{\hat{u}_1'\hat{u}_1}{\hat{u}_2'\hat{u}_2}\right) > \exp\left\{\frac{\chi_\varepsilon^2}{T}\right\}$$

de donde se obtiene la expresión para h( ).

Vamos a terminar esta sección reescribiendo todas las expresiones h( ) en la forma que corresponde al criterio F redefiniendo la forma que toma el punto crítico F_ε. Los resultados aparecen en la última columna del Cuadro 3.4.2.

Escribimos h( ) en la forma siguiente:

$$1 + \frac{(k_2 - k_1)}{T - k_2} F \qquad (3.4.13)$$

Para el criterio $\bar{R}^2$ podemos escribir:

$$\frac{T - k_1}{T - k_2} = 1 + \frac{k_2 - k_1}{T - k_2} F_{\bar{R}^2}$$

con $$F_{\bar{R}^2} = 1$$

A partir de estos resultados se pueden concretar los valores que toman puntos críticos implícitos correspondientes a los diferentes criterios.

CUADRO 3.4.2. Formas Alternativas de La Regla de Decisión

Criterio	FORMA 1: $\hat{u}_1'\hat{u}_1 < \hat{u}_2'\hat{u}_2\, h(\ )$ $h\ (\ )$	$h(F) = 1 + \frac{k_2 - k_1}{T - k_2}F$ F
$\bar{R}^2$	$(T - k_1)/(T - k_2)$	$F_{\bar{R}^2} = 1$
F	$1 + \frac{(k_2 - k_1)}{T - k_2}F_\varepsilon$	$F_F = F_\varepsilon$
W	$1 + \frac{\chi^2_\varepsilon}{T}$	$F_W = \frac{T - k_2}{T}\ \frac{\chi^2_\varepsilon}{k_2 - k_1}$
LM	$\frac{T}{T - \chi^2_\varepsilon}$	$F_{LM} = \frac{T - k_2}{k_2 - k_1}\ \frac{\chi^2_\varepsilon}{T - \chi^2_\varepsilon}$
LR	$\exp\left(\frac{\chi^2_\varepsilon}{T}\right)$	$F_{LR} = \left(\exp\left(\frac{\chi^2_\varepsilon}{T}\right) - 1\right)\frac{T - k_2}{k_2 - k_1}$
C_p	$1 + \frac{2(k_2 - k_1)}{T - k_2}$	$F_{Cp} = 2$
AIC	$\exp\left(\frac{2(k_2 - k_1)}{T}\right)$	$F_{AIC} = \left(\exp\left(\frac{2(k_2 - k_1)}{T}\right) - 1\right)\frac{T - k_2}{k_2 - k_1}$
PIC	$1 + \frac{k_2 - k_1}{T - k_2}\ln\tilde{\sigma}^2_2 + \frac{1}{T - k_2}$ $\ln\frac{\lvert X_2'X_2\rvert}{\lvert X_1'X_1\rvert}$	$F_{PIC} = -\ln\tilde{\sigma}^2_2 + \frac{1}{k_2 - k_1}\ln\frac{\lvert X_2'X_2\rvert}{\lvert X_1'X_1\rvert}$
BEC	$1 + \frac{\ln T(k_2 - k_1)}{T - k_2}$	$F_{BEC} = \ln T$
$SBIC$	$\exp\left((k_2 - k_1)\frac{\ln T}{T}\right)$	$F_{SBIC} = \left(\exp\left(\frac{\ln T\,(k_2 - k_1)}{T}\right) - 1\right)\frac{T - k_2}{k_2 - k_1}$

Podemos fijar un nivel de significación, tal como $\varepsilon = 0{,}05$, y teniendo en cuenta los grados de libertad obtener el correspondiente punto crítico $F_{0,05}$. A partir de este valor, podemos decir que cualquier criterio que tenga un valor de la F –tal como aparecen en la última columna del Cuadro 3.4.2- superior a $F_{0,05}$ dicho criterio garantizará un nivel del ε inferior al 5%. Por el contrario, para el criterio para el que el valor de la F sea inferior a $F_{0,05}$ podemos decir que la garantía no llegará a ese nivel del 5%.

Una ilustración numérica de estos resultados puede verse en el Cuadro 3.4.3. En este cuadro se calculan los valores de la F que corresponden a los diferentes criterios, suponiendo 2 tamaños muestrales y dos diferencias de tamaños entre los dos modelos.

CUADRO 3.4.3. Punto Crítico Implícito T = 45(T = 125) ε = 0,05

PUNTO CRÍTICO	$k_2 = 2,\ k_1 = 1$	$k_2 = 5,\ k_1 = 1$
$F_{\bar{R}^2}$	1	1
$F_{0,05}$	4,08(3,92)	2,61(2,45)
F_W	3,67(3,78)	2,11(2,27)
F_{LM}	4,01(3,89)	2,67(2,46)
F_{LR}	3,83(3,84)	2,34(2,36)
F_{Cp}	2	2
F_{AIC}	1,95(1,98)	1,94(1,98)
F_{BEC}	3,89(4,82)	3,80(4,82)
F_{SBIC}	3,79(4,84)	4,01(5,01)

A partir del contenido del Cuadro 3.4.3 se observa que los resultados son diferentes según sean los valores que toman (k_2 - k_1) y el tamaño muestral.

En general, podemos decir que los contrastes W, LM y LR se comportan de forma similar a como lo hace el contraste F. Los criterios $\bar{R}^2$, C_p y AIC son menos parsimoniosos que el contraste F aunque la diferencia tiende a decrecer conforme la diferencia

$(k_2 - k_1)$ tiende a ser mayor. Por último, los criterios BEC y SBIC tienden a ser más parsimoniosos que el criterio F, haciéndose la distancia mayor conforme la diferencia $(k_2 - k_1)$ se hace mayor.

Cualquiera que sea la aproximación que se siga, la conclusión a la que se llega es que la diferencia entre los diferentes criterios radica en el diferente peso que asignan al factor de parsimonia; según sea la ponderación asignada, la combinación que resulta de los dos tamaños de error difiere.

Referencias

Aznar, A. (2012): "Curso de Econometría" Copy Center Digital. Zaragoza.