

Guía de estudio del 1º Parcial. Econometría 4º Curso, Licenciatura de Economía. USC. 22-2-2013.

Profesoras María del Carmen Guisán Seijas y Pilar Expósito Díaz.

Bibliografía básica: Guisán, M.C. (1997) Econometría. Editorial McGraw-Hill. Guisán, M.C. (2012) Documento nº 121 en la serie *Economic Development*: <http://www.usc.es/economet/econometria.htm>

TEMA 1: AMPLIACIONES DEL MODELO CLÁSICO

Recordamos las hipótesis del Modelos de Regresión Lineal Clásico (MRLC) y añadimos una referencia al caso de que esas hipótesis se incumplan. La estimación del modelo consiste en calcular un *vector b de estimadores* de los elementos del *vector de parámetros* β , a partir del conocimiento de los datos de las variables y y X y de supuestos acerca del comportamiento de la variable desconocida ε . En general utilizamos letras latinas para elementos que son conocidos o se pueden calcular (y , X , b) y letras griegas para elementos desconocidos que tratamos de estimar (β , ε). Las hipótesis del modelo clásico son:

1) Hipótesis de linealidad: Existe una relación lineal entre la *variable explicada* por la ecuación del modelo y_t (que es el *regresando* o *variable endógena* en el momento t) y un conjunto de variables: x_0 , x_1 , x_2 , ..., x_k (que son los *regresores*: el término constante $x_0=1$, también llamado “*regresor ficticio*”, y las *variables explicativas* o *variables predeterminadas*: x_1 , x_2 , ..., x_k) que forman el vector fila $x'(t)$, que se expresa de la siguiente forma:

$$y_t = x'(t) \beta + \varepsilon; \quad y = X \beta + \varepsilon; \quad (1)$$

y es un vector columna de orden $T \times 1$ de observaciones de la variable endógena y_t (siendo T el tamaño de la muestra o número de observaciones, temporales o atemporales de la variable y_t),

$x'(t)$ es un vector fila de *regresores*, de orden $1 \times k^*$ ($k^*=k+1$ si el modelo incluye x_0 , mientras que $k^*=k$ si el modelo no incluye x_0), siendo k el número de *variables predeterminadas* o *variables explicativas*. X es la matriz de orden $T \times k^*$ cuyas filas son los vectores $x'(t)$ desde $t=1, 2, \dots, T$. Las variables predeterminadas pueden ser *exógenas* (no explicadas por el modelo) o *endógenas retardadas* (explicadas por el modelo en un momento anterior, por ejemplo $x_{1t} = y_{t-1}$).

El vector de *parámetros* β es un vector columna, de orden $k^* \times 1$ de parámetros desconocidos, siendo su traspuesto β' un vector fila de k^* elementos: $\beta' = (\beta_0 \ \beta_1 \ \beta_2 \ \beta_k)$ mientras que ε es un vector columna de orden $T \times 1$ de la *perturbación aleatoria*, que es una variable desconocida que recoge pequeños efectos de variables omitidas que explican la parte de y_t que no está relacionada con $x'(t)$.

La *ordenada en el origen* β_0 se incluye habitualmente en los modelos en los que las variables están medidas en niveles, es decir no son incrementos, mientras que en general se excluye en los modelos en incrementos o en el modelo dinámico mixto (modelo en el que y_t y y_{t-1} están medidos en niveles y las variables exógenas en incrementos).

Estimadores Mínimo-Cuadráticos Ordinarios (MCO) en el MRLC:

$b = (X'X)^{-1} X'y$ posee muy buenas propiedades y minimiza la Suma de Cuadrados de los Errores (SCE). (2)

Matriz de covarianzas de b es $V(b) = \sigma_\varepsilon^2 (X'X)^{-1}$, se estima mediante $\hat{V}(b) = s^2 (X'X)^{-1}$ (3)

El cálculo de b , de sus varianzas estimadas y t -de Student puede realizarse con el programa E-Views:

LS y $c \ x_1 \ x_2 \ \dots \ x_k$ (donde c es el término constante que se incluye si el modelo tiene regresor x_0).

INCUMPLIMIENTO DE LA HIPÓTESIS DE LINEALIDAD: La hipótesis se incumple si el modelo es no lineal, por ejemplo en una función potencial como es la función de producción Cobb-Douglas:

$$Q_t = \gamma_0 \ KD_t^{\beta} \ L_t^{\alpha} \ e^{\lambda t} \quad (4)$$

Donde la Producción real (Q) se expresa en función del capital disponible (KD), el empleo (L) y de cambios en eficiencia que se producen a los largo del tiempo. En el exponente del número e figura un parámetro λ multiplicado por la variable t que representa el tiempo.

En el modelo no lineal podemos optar por dos opciones: a) linealizarlo, especialmente si es fácilmente linealizable como ocurre en el caso de la función potencial, b) estimar los parámetros $(\gamma_0, \beta, \alpha, \lambda)$ mediante Mínimos-Cuadrados No Lineales (NLS). En este ejemplo:

a) LS $\log(Q)$ c $\log(KD)$ $\log(L)$ TI (donde $\log$ se refiere a logaritmo neperiano y TI es la variable tiempo)

b) NLS $Q=c(0)*KD^{c(1)}*L^{c(2)}*\exp(c(3)*TI$, siendo $c(0)$, $c(1)$, $c(2)$ y $c(3)$ estimadores de los parámetros.

2. Hipótesis de que la perturbación es un ruido blanco. En el Modelo Clásico se asume que el vector de perturbaciones aleatorias, ϵ , tiene esperanza matemática nula y *matriz de covarianzas escalar* con un término constante en la diagonal principal (σ_ϵ^2) y elementos nulos fuera de la diagonal principal. Este tipo de variable se denomina *ruido blanco*:

Esperanza matemática del vector ϵ : $E(\epsilon) = 0$, siendo 0 un vector columna nulo de orden $T \times 1$

Matriz de varianzas-covarianzas, o matriz de covarianzas, del vector ϵ : $V = E\epsilon\epsilon' = \sigma_\epsilon^2 I$

siendo σ_ϵ^2 la varianza de ϵ_t (para todo $t=1,2,\dots,T$) e I una matriz identidad de orden $T \times T$.

En general en el Modelo Clásico se asume que el vector ϵ sigue una distribución Normal, lo que permite realizar diversos contrastes de hipótesis. La hipótesis de normalidad y la de esperanza nula se cumplen habitualmente, pero la hipótesis de matriz V escalar se incumple en muchos casos.

INCUMPLIMIENTO DE LA HIPÓTESIS DE MATRIZ V ESCALAR: Esta hipótesis puede incumplirse debido a que los elementos no diagonales, de la matriz V , son no nulos (autocorrelación) o a que los elementos diagonales no son iguales a una constante (heterocedasticidad). La autocorrelación se presenta con cierta frecuencia en series temporales. La heterocedasticidad se presenta con cierta frecuencia en series atemporales. También pueden presentarse los dos problemas en el caso de una combinación de series temporales y atemporales. El Modelo de Regresión Lineal Generalizado (MRLG) es un modelo similar al clásico en el que se incumple la hipótesis de V escalar. En ese caso tenemos dos opciones para convertir el modelo en clásico: a) mejorar la especificación del modelo (puede incluirse una variable importante que estaba omitida o revisar la forma funcional) para que se cumplan la hipótesis de V escalar en el modelo de y_t . b) transformar las variables, como se indica en la bibliografía, de forma que la relación entre el vector y^* y la matriz X^* de variables transformadas se adapta a las hipótesis del modelo clásico. En general la primera de estas opciones es la preferible pero no siempre es posible realizarla.

3. Hipótesis de X no estocástica y $r(X)=k^*$. En el modelo clásico se asume, a efectos de simplificación de la demostración de las buenas propiedades del vector de estimadores b , que la matriz X es no estocástica. Esas propiedades se mantienen en alto grado (especialmente si la muestra es grande), aunque X sea estocástica si los regresores no tienen alta correlación contemporánea con la perturbación aleatoria. También se asume que el rango de X es igual al número de regresores (k^*) y no hay multicolinealidad perfecta.

INCUMPLIMIENTO DE LA HIPÓTESIS X NO ESTOCÁSTICA: En el Modelo de Regresión Lineal Estocástico (MRLE) el vector b de estimadores MCO conserva buenas propiedades si no hay correlación contemporánea entre cada regresor x_{it} y ϵ_t , es decir si el vector $\Sigma_{x\epsilon}$ es nulo.

En el Modelo de Regresión Lineal Estocástico sin correlación contemporánea el vector b de estimadores MCO es un estimador consistente de β , y tiene también otras buenas propiedades (en general tiene insesgadez asintótica), por lo que es correcto utilizarlo para la estimación de los parámetros, ya que el límite probabilístico de b tiende a β :

$$\text{plim } b = \beta + \Sigma_{xx}^{-1} \Sigma_{x\epsilon} = \beta + \Sigma_{xx}^{-1} 0 = \beta; \quad \text{siendo } \Sigma_{x\epsilon} = E x(t) \epsilon_t = 0 \quad (5)$$

En el Modelo de Regresión Lineal Estocástico con correlación contemporánea entre algún regresor y la perturbación, el vector b de estimadores MCO es un estimador inconsistente de β , ya que $\text{plim } b$ es distinto de β , pues $\Sigma_{x\epsilon}$ no es nulo. Si la inconsistencia es débil el vector b puede también utilizarse. El método general de los momentos (MGM o GMM) o su caso particular el método de variables instrumentales (VI) se pueden utilizar para obtener estimadores consistentes. Las principales causas de este problema son: a) modelos autorregresivos con autocorrelación de la perturbación, b) modelos multiecuacionales con interdependencia, c) modelo con errores aleatorios de cierta importancia en un regresor.

MULTICOLINEALIDAD: La matriz X no puede tener relaciones lineales exactas entre sus columnas ya que entonces se presentaría el problema de multicolinealidad perfecta, el rango de X sería menor que k^* y el determinante de $(X'X)$ sería nulo, y en consecuencia no podría calcularse el vector b . La multicolinealidad perfecta puede evitarse pero es frecuente que entre los regresores de un modelo econométrico haya correlaciones lineales elevadas aunque no sean exactas. En ese caso el determinante de $(X'X)$ será próximo a cero y los elementos diagonales de la inversa de $(X'X)$ tendrán valores muy elevados lo que causa imprecisión en los estimadores de los parámetros teniendo en cuenta la fórmula de $V(b)$. Aumentar el tamaño de la muestra, si es posible, suele ser una solución efectiva para este problema.

TEMA 2: EL MODELO DE REGRESION LINEAL GENERALIZADO (MRLG)

El modelo generalizado (MRLG) es un modelo similar al clásico (MRLC), en el que cambia la hipótesis relativa a la matriz V que ya no es una matriz escalar, y por lo tanto la perturbación no es un *ruido blanco*. En ese caso, dada la matriz X , el vector b de estimadores mínimo-cuadráticos ordinarios (MCO) deja de ser óptimo y el vector de estimadores mínimo-cuadrático generalizados (MCG) es preferible porque tiene varianzas menores y por lo tanto proporciona estimadores más precisos:

$$\text{Estimadores MCG del modelo de } y: \quad \hat{b} = (X'V^{-1}X)^{-1} X'V^{-1}y \quad V(\hat{b}) = (X'V^{-1}X)^{-1} \quad (6)$$

$$\text{Estimadores MCO del modelo de } y: \quad b = (X'X)^{-1} X'y \quad V(b) = (X'X)^{-1} X'V X (X'X)^{-1} \quad (7)$$

$$\text{Estimadores MCO del modelo de } y^*: \quad b^* = (X^{*'}X^*)^{-1} X^{*'}y^* \quad V(b^*) = s^{*2} (X^{*'}X^*)^{-1} \quad (8)$$

donde $s^{*2} = \text{SCE}^*/(T-k^*)$, siendo $\text{SCE}^* = y^{*'}y^* - b^{*'}X^{*'}y^*$ la suma de los cuadrados de los errores de e^*_t .

Los estimadores MCO del modelo del regresando transformado y^* coinciden con los estimadores MCG del modelo de y . Este método se utiliza con frecuencia para obtener los estimadores MCG.

En el MRLG los estimadores MCO aunque no son óptimos pueden utilizarse calculando sus varianzas estimadas mediante la fórmula (7) lo que se conoce como *corrección de White*.

Goldberger(1962) demostró que en el MRLG el *predictor óptimo* de y en el momento τ es:

$$\hat{y}_\tau = x'(\tau) \hat{b} + w'V^{-1}e; \quad w' = (e_1e_\tau \quad e_2e_\tau \dots e_\tau e_\tau); \quad e = y - Xb \quad (9)$$

TEMA 3: AUTOCORRELACIÓN

La autocorrelación surge con frecuencia, en series temporales, por la omisión de alguna variable explicativa relevante y también puede surgir por una inadecuada forma funcional o por otros motivos. Con frecuencia la inclusión de una variable explicativa importante, que estaba omitida, es la mejor solución a este problema, como vemos en el modelo de empleo de España de la sección 1 del Documento 121. Si esa solución no es posible es preferible utilizar los estimadores MCG, en vez de los estimadores MCO, para realizar contrastes de hipótesis y para predecir, porque tienen mejores propiedades.

En caso de utilizar los estimadores MCO en el MRLG hay que tener en cuenta que ya no son óptimos pero que aún tienen buenas propiedades, y que la fórmula de $V(b)$ es (7) y no la fórmula (3) del MRLC.

El estadístico Durbin-Watson (dw) es el método habitual de contrastar la hipótesis H_0 de ausencia de autocorrelación de la perturbación aleatoria, frente a la hipótesis alternativa de presencia de autocorrelación. En series temporales es un resultado incluido en el programa de estimación. Un valor de dw próximo a 2 indica que el coeficiente de autocorrelación ρ es nulo. Las tablas del estadístico dw nos proporcionan los límites inferior y superior (d_L y d_U), y distinguimos 5 zonas en las que puede situarse el valor de dw :

$$\begin{aligned} \text{Dos zonas de rechazo de } H_0: & \quad dw < d_L \text{ en caso de } \rho > 0; \quad dw > d_U \text{ en caso de } \rho < 0 \\ \text{Una zona de aceptación de } H_0: & \quad d_U < dw < 4-d_U \text{ en caso de } \rho = 0 \\ \text{Dos zonas de incertidumbre:} & \quad d_L < dw < d_U; \quad 4-d_U < dw < 4-d_L \end{aligned} \quad (10)$$

En el caso de modelos con datos anuales y autocorrelación del tipo $AR(1)$. En caso de un modelo autorregresivos con autocorrelación es preferible utilizar el estadístico h de Durbin. Con datos de periodicidad inferior al año (mensual, trimestral, etc.) conviene aplicar métodos de contraste más generales, como el estadístico de Box-Pierce para un $ARMA(p,q)$ (tema que se estudia en el 2º parcial).

La estimación por MCG en caso de autocorrelación es sencilla mediante el programa E-Views, indicando que hay autocorrelación mediante $AR(1)$ u otro tipo de $ARMA(p,q)$, y no es necesario programar el método de transformación de las variables:

LS Y C X1 X2 AR(1)

La fórmula (9), para predecir, en el caso de autocorrelación $AR(1)$ y un período de antelación es:

$$\hat{y}_\tau = x'(\tau) \hat{b} + \rho \hat{e}_{\tau-1}; \quad \hat{e}_{\tau-1} = y_{\tau-1} - x'(\tau-1) \hat{b}$$

TEMA 4. HETEROCEDASTICIDAD

La heterocedasticidad está presente, a veces, en series atemporales y en la combinación de series temporales y atemporales. Una causa es el aumento de la dispersión de ε_t cuando aumenta algún regresor.

Se contrasta mediante el test de la razón de verosimilitud (LR), el test de White, u otros tests que se indican en la bibliografía. En la sección 2 del Documento 121 se realizan estos contrastes en un modelo de la tasa de empleo de las regiones españolas.

El test de la razón de verosimilitud es útil cuando dividimos la muestra en p grupos (por ejemplo en 4 grupos según el nivel de renta en una ecuación de consumo), y estimamos la varianza de la perturbación en cada uno de los grupos a partir de los residuos de la ecuación estimada por MCO. El estadístico que se utiliza es μ , el cual bajo la hipótesis H_0 de homogeneidad de varianzas (homocedasticidad) se distribuye como una ji-cuadrado de $(p-1)$ grados de libertad, siendo p el número de grupos considerados:

Si $\mu > \chi^2_{\alpha}$ se rechaza H_0 , $\mu = -2 \ln(\lambda)$, $\ln(\lambda) = (T_1 \ln s_1 + T_2 \ln s_2 + \dots + T_p \ln s_p) - T \ln s$ donde T_i ($i=1, \dots, p$) es el tamaño muestral de cada grupo, s_i la raíz cuadrada de la varianza de la perturbación estimada en cada grupo, mientras que T es el tamaño muestral total y s es la raíz cuadrada de la varianza de la perturbación estimada en el conjunto total. El valor crítico χ^2_{α} es aquel punto de la ji-cuadrado que deja a su derecha una cola de probabilidad α y a su izquierda una probabilidad $(1-\alpha)$. Si fijamos el valor habitual $\alpha=0.05$, obtenemos el valor crítico en las tablas de la distribución ji-cuadrado para $(p-1)$ grados de libertad.

El test de White realiza la regresión de los cuadrados de los residuos de la estimación MCO respecto a x_0 , a los valores de cada variable predeterminada, a los valores del cuadrado de dichas variables, y a los productos cruzados entre las variables explicativas. Si todos los parámetros son nulos, excepto la ordenada en el origen, se confirma que hay varianza constante y por lo tanto homocedasticidad. Si se rechaza la nulidad conjunta, del subconjunto de parámetros sin la ordenada en el origen, se afirma que hay heterocedasticidad.

Si existe heterocedasticidad la estimación MCG es mejor que la estimación MCO, y el predictor óptimo de la relación (9) se reduce al primer sumando porque el segundo sumando es nulo. La transformación de las variables puede realizarse multiplicando por el cociente (s_1/s_i) para $i=1, 2, \dots, p$ siendo p el número de grupos.

A efectos de simplificación se pueden utilizar los estimadores MCO para el contraste de nulidad y para la predicción, pues aunque no son óptimos todavía tienen buenas propiedades. En ese caso es importante recordar que la estimación de la matriz $V(b)$ debe realizarse con la fórmula (7) y no con la fórmula (3).

TEMA 5: MULTICOLINEALIDAD Y SELECCIÓN DE REGRESORES

Un alto grado de relación lineal entre los regresores es frecuente, ya que ocurre entre cualquier par de variables que tengan tendencias lineales en el tiempo aunque no haya entre ellas ninguna relación causal entre ellas. La multicolinealidad ocasiona valores grandes para los elementos de $V(b)$ y dificulta la selección de regresores, con situaciones de incertidumbre cuando se contrasta la nulidad de un parámetro no nulo. El aumento del tamaño y de la diversidad de la muestra en general reduce la multicolinealidad.

Es importante destacar que las situaciones posibles, en un contraste H_0 de nulidad de un parámetro, no son sólo *aceptación* o *rechazo*. *No aceptación* muchas veces no significa *rechazo* sino *incertidumbre*. Por ejemplo si el *intervalo de confianza* de un parámetro, β_i dado por $(b_i - s_{bi} t_{\alpha/2}, b_i + s_{bi} t_{\alpha/2})$, (donde $t_{\alpha/2}$ es un valor próximo a 2 dado por las tablas correspondientes), es $(-5, 3)$ tenemos gran incertidumbre porque todos los valores del intervalo no podrían rechazarse y no hay evidencia clara de que el parámetro sea nulo.

La selección de regresores no debe ser sólo mecánica sino basarse en un conocimiento de la teoría y de la realidad económica. Los métodos mecánicos que se utilizan habitualmente son: a) todas las regresiones posibles, b) selección hacia delante (*forward*), c) selección hacia atrás (*backward*), que en general es el método más sencillo, d) selección paso a paso (*stepwise*).

Variables irrelevantes incluidas: No es conveniente la inclusión de variables explicativas irrelevantes porque aumenta la imprecisión de los estimadores y las situaciones de incertidumbre especialmente si las variables irrelevantes incluidas ocasionan algo grado de multicolinealidad.

Variables relevantes excluidas: Para contrastar la nulidad de los parámetros y efectuar predicciones el modelo econométrico funciona bien a veces incluso si existen variables relevantes excluidas, siempre que exista un relación lineal fuerte y estable entre las variables excluidas y las incluidas, como puede verse en el capítulo correspondiente de la bibliografía. Hay que tener en cuenta que el efecto de las variables excluidas afecta al valor de los parámetros y de sus estimadores, ya que su valor depende del efecto de las variables incluidas, del efecto de las variables excluidas y de la relación lineal entre las excluidas y las incluidas.

Anexo 1. Ejercicios.

Ejercicio 1. Autocorrelación: Empleo y salarios en España, 1995-2012

Año	LTE	PIB00E	W00E	PAE	Año	LTE	PIB00E	W00E	PAE
1995	12313	474.85	22.24	16027	2004	17481	658.09	22.00	19695
1996	12634	486.33	22.27	16290	2005	18431	681.88	22.15	20343
1997	13036	505.14	21.96	16306	2006	19162	709.28	22.22	20999
1998	13530	527.72	22.00	16707	2007	20361	734.63	22.38	22195
1999	14234	552.76	21.87	16956	2008	20124	740.95	22.60	22851
2000	15222	580.67	21.71	17706	2009	18888	713.36	23.71	23037
2001	15752	601.86	21.77	17619	2010	18394	711.22	23.14	23026
2002	16160	618.13	21.87	18315	2011	18105	708.38	22.57	23149
2003	16657	637.27	22.00	18899	2012	17282	701.30	22.32	23261

Notas: LTE y PAE son el Empleo Total y la Población Activa (miles de personas); PIB00E= PIB real en Dólares del año 2000, W00E=coste salarial real medio en miles de Dólares del año 2000. $LTE_{94}=12012$. El punto indica separación decimal. Fuente: Datos de España obtenidos, o elaborados, a partir de estadísticas de la OCDE y del INE.

Ecuación 1. Dependent Variable: LTE. Method: Least Squares. Sample: 1995 2012				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
LTE(-1)	0.975779	0.006696	145.7325	0.0000
D(PIB00E/W00E)	467.2689	75.84314	6.160991	0.0000
D(PAE)	0.954677	0.214183	4.457302	0.0005
R-squared	0.989072	Mean dependent var		16542.51
Adjusted R-squared	0.987614	S.D. dependent var		2578.528
S.E. of regression	286.9653	Akaike info criterion		14.30761
Sum squared resid	1235237.	Schwarz criterion		14.45601
Log likelihood	-125.7685	Hannan-Quinn criter.		14.32807
Durbin-Watson stat	2.012697			

Ecuación 2. Dependent Variable: LTE. Method: Least Squares. Sample: 1995 2012				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
LTE(-1)	0.981043	0.012082	81.19934	0.0000
D(PAE)	1.331202	0.373464	3.564474	0.0026
R-squared	0.961417	Mean dependent var		16542.51
Adjusted R-squared	0.959006	S.D. dependent var		2578.528
S.E. of regression	522.0769	Akaike info criterion		15.45795
Sum squared resid	4361028.	Schwarz criterion		15.55688
Log likelihood	-137.1215	Hannan-Quinn criter.		15.47159
Durbin-Watson stat	1.056538			

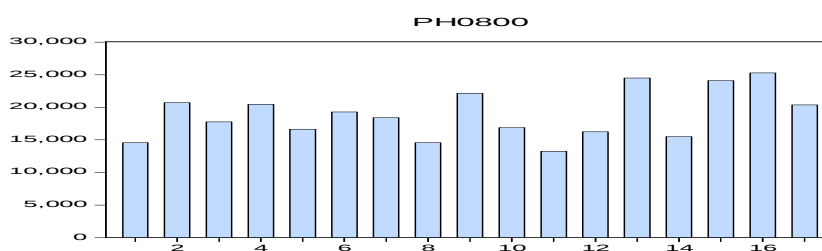
Ecuación 3. Dependent Variable: LTE. Method: Least Squares. Sample: 1995 2012				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
LTE(-1)	0.988217	0.021216	46.57893	0.0000
D(PAE)	0.760491	0.343537	2.213707	0.0428
AR(1)	0.671304	0.249185	2.693993	0.0167
R-squared	0.972533	Mean dependent var		16542.51
Adjusted R-squared	0.968871	S.D. dependent var		2578.528
S.E. of regression	454.9399	Akaike info criterion		15.22922
Sum squared resid	3104555.	Schwarz criterion		15.37761
Log likelihood	-134.0630	Hannan-Quinn criter.		15.24968
Durbin-Watson stat	1.776720			

Comentarios: En la ecuación 2, dw es bajo y menor que $d_L=1.13$, lo que indica presencia de autocorrelación positiva. Por tratarse de un modelo autorregresivos con autocorrelación es mejor utilizar el h de Durbin, el cual toma el valor $1.87 < 1.96$ para $\alpha=0.05$, >1.64 para $\alpha=0.10$. Indica que hay indicios de autocorrelación, pero que no es excesiva y no es imprescindible la estimación por mínimos cuadrados generalizados. Observamos que el mejor resultado es el de la ecuación 1, en el que hemos corregido la autocorrelación al incluir una variable explicativa relevante que estaba excluida en la ecuación 2, y observamos que el Error Estándar (S.E.) baja de 522 en la ecuación 2 a 286 en la ecuación 1, mientras que la corrección mecánica de la autocorrelación al pasar de la ecuación 2 a la 3 daría un peor resultado con un S.E. de 454.

Ejercicio 2. Test de heterocedasticidad: Tasa de empleo y PIB por habitante (PH) en 17 CCAA, año 2008

Dependent Variable: LHT08. Method: Least Squares. Sample: 1 17				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	296.8910	22.88962	12.97055	0.0000
PH0800/1000	7.556014	1.192088	6.338469	0.0000
R-squared	0.728144	Mean dependent var		439.4435
Adjusted R-squared	0.710020	S.D. dependent var		32.60335
S.E. of regression	17.55684	Akaike info criterion		8.678895
Sum squared resid	4623.640	Schwarz criterion		8.776920
Log likelihood	-71.77061	Hannan-Quinn criter.		8.688639
F-statistic	40.17619	Durbin-Watson stat		2.500290
Prob(F-statistic)	0.000013			

Heteroskedasticity Test: White				
F-statistic	1.750060	Prob. F(2,14)		0.2097
Obs*R-squared	3.400093	Prob. Chi-Square(2)		0.1827
Scaled explained SS	1.019989	Prob. Chi-Square(2)		0.6005
Test Equation: Dependent Variable: RESID^2				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3459.578	1730.387	1.999309	0.0654
PH0800/1000	-330.3066	182.6939	-1.807979	0.0921
(PH0800/1000)^2	8.256277	4.693771	1.758986	0.1004
R-squared	0.200005	Mean dependent var		271.9788
Adjusted R-squared	0.085721	S.D. dependent var		246.1074
S.E. of regression	235.3228	Akaike info criterion		13.91858
F-statistic	1.750060	Durbin-Watson stat		2.739711



obs	LHT08	PH00	PH0800	POB00	POB08
1.An	384	11538	14575	7340	8202
2.Ar	461	16365	20730	1190	1327
3.As	418	13081	17766	1077	1080
4.Bl	475	19282	20450	846	1073
5.Cn	415	14845	16621	1716	2076
6.Cb	447	14634	19301	531	582
7.Cl	418	14164	18397	2479	2557
8.Cm	413	12307	14546	1734	2043
9.Ct	475	19072	22126	6262	7364
10.Cv	443	15102	16907	4121	5030
11.Ex	373	9965	13252	1069	1098
12.Ga	431	12163	16238	2732	2784
13.Ma	489	21281	24500	5205	6272
14.Mu	440	13132	15508	1149	1426
15.Na	466	19927	24109	544	620
16.Pv	461	19182	25306	2099	2157
17.Ri	462	17826	20393	264	318

Comentarios: Se acepta la hipótesis de homocedasticidad porque los estadísticos del test de White toman valores bajos y sus *probabilities* son superiores a 0.05 y a 0.10. La hipótesis se rechaza si el estadístico toma valores elevados y la *probability* es igual o inferior a 0.05.

Anexo 2: Conceptos básicos de Econometría I: Estimadores MCO y contrastes en el modelo clásico.

Cuadro 2.1
Resultados de la estimación MCO

Vector de estimadores de β :	$b = (X'X)^{-1}X'y$
Suma de cuadrados de errores:	$SCE = y'y - b'X'y$
Estimador de σ^2 :	$S^2 = \frac{SCE}{T - k - 1}$
Matriz estimada de covarianzas de b :	$\hat{V}(b) = S^2(X'X)^{-1}$
Suma de cuadrados totales:	$SCT = \sum_i y_i^2 - T\bar{y}^2$; $\bar{y} = T^{-1} \sum_i y_i$
Coefficiente de determinación:	$R^2 = 1 - \frac{SCE}{SCT}$
Error estándar:	$E.S. = \sqrt{S^2}$
Raíz del error cuadrático medio:	$RECM = \sqrt{\frac{SCE}{T}}$; $\%RECM = \frac{RECM}{\bar{y}} \times 100$

Fuente: Guisán(1997): Econometría

Cuadro 3.1
Resumen de los contrastes de significatividad examinados en el capítulo

H_0	H_1	Estadístico a calcular	Grados de libertad	Criterio para rechazar H_0
$\beta_i = \beta_i^*$ ($i = 0, 1, \dots, k$)	$\beta_i \neq \beta_i^*$	$t_i = \frac{b_i - \beta_i^*}{S_{b_i}}$	$(T - k - 1)$	$ t_i > t_{\alpha/2}$
$\beta = 0$	$\beta \neq 0$ ($k + 1$) $\times 1$	$F = \frac{b'X'y/n}{SCE/d}$	$n = k + 1$ $d = T - k - 1$	$F > F_\alpha$
$\underline{\beta} = 0$	$\underline{\beta} \neq 0$ ($k \times 1$)	$F = \frac{(SCT - SCE)/n}{SCE/d}$	$n = k$ $d = T - k - 1$	$F > F_\alpha$
$\beta_2 = 0$	$\beta_2 \neq 0$ ($k - H$) $\times 1$	$F = \frac{(SCE_1 - SCE)/n}{SCE/d}$	$n = k - H$ $d = T - k - 1$	$F > F_\alpha$
$\beta = \beta^*$	$\beta \neq \beta^*$ ($k + 1$) $\times 1$	$F = \frac{(b - \beta^*)'(X'X)(b - \beta^*)/n}{SCE/d}$	$n = k + 1$ $d = T - k - 1$	$F > F_\alpha$
$\pi = \pi^*$ ($\pi = \delta'\beta$)	$\pi \neq \pi^*$	$t = \frac{p - \pi^*}{S_p}$	$T - k - 1$	$ t > t_{\alpha/2}$

Símbolos utilizados:

Fuente: Guisán(1997): Econometría.

Símbolos utilizados: H_0 =Hipótesis a contrastar; H_1 = Hipótesis alternativa; $t_{\alpha/2}$ valor crítico de la distribución t-de Student para un contraste dos colas; F_α valor crítico de la distribución F de Snedecor para un contraste de una sola cola; β vector de parámetros de la ecuación (generalmente de orden $(k+1) \times 1$) mientras que $\underline{\beta}$ con un guión debajo indica que es un vector de orden $k \times 1$ que excluye a la ordenada en el origen β_0 ; β_2 es un subsector de orden $(k-H)$; n indica grados de libertad del numerador; d indica grados de libertad del denominador, SCE es la suma de los cuadrados de errores de la regresión respecto a todos los regresores, mientras que SCE_1 es la suma de cuadrados de errores respecto a x_0 y al subconjunto de H variables explicativas, excluyendo $k-H$ regresores. Nota: los contrastes de significatividad pueden referirse a un solo parámetro, en el caso de la t-de Student, o a un conjunto de parámetros como en el caso de la F de Snedecor. En el caso de un solo parámetro la hipótesis más frecuente es la de nulidad (H_0 es $\beta_i = 0$).