TUTORIAL PARA LA ESTIMACIÓN DE UN MODELO CON PRESENCIA DE AUTOCORRELACIÓN EN EASYREG

Julio César Alonso C.

No. 16

Septiembre de 2008

UNIVERSIDAD ICESI
DEPARTAMENTO DE ECONOMÍA

APUNTES DE ECONOMÍA

ISSN 1794-029X

No. 16, Septiembre de 2008

Editor

Julio César Alonso C.

jcalonso@icesi.edu.co

Vanessa Ospina López

Asistente de Edición

Gestión Editorial

Departamento de Economía - Universidad Icesi

www.icesi.edu.co

Tel: 5552334 ext: 8398. Fax: 5551441

Calle 18 # 122-135 Cali, Valle del Cauca, Colombia

TUTORIAL PARA LA ESTIMACIÓN DE UN MODELO CON PRESENCIA DE AUTOCORRELACIÓN EN EASYREG

Julio Cesar Alonso C[1].

Septiembre de 2008

Resumen

Este documento presenta una breve introducción a tres pruebas de autocorrelación heteroscedasticidad (Durbin-Watson, Box-Pierce y la prueba de rachas) empleando el paquete econométrico gratuito EasyReg. Así mismo, se presenta la manera de corregir un modelo con autocorrelación de primer orden por medio de diferencias generalizadas empleando EasyReg. Este documento está dirigido principalmente a estudiantes de pregrado de un curso de econometría o cualquier lector con conocimientos básicos del modelo de regresión múltiple.

Palabras Clave: EasyReg, Regresión múltiple, Prueba de *Durbin-Watson, prueba de Box-Pierce, prueba de rachas, corrección por autocorrelación, diferencias generalizadas.*

Apuntes de Economía es una publicación del Departamento de Economía de la Universidad Icesi, cuya finalidad es divulgar las notas de clase de los docentes y brindar material didáctico para la instrucción en el área económica a diferentes niveles. El contenido de esta publicación es responsabilidad absoluta del autor.

[1] Profesor del Departamento de Economía y Director del Centro de Investigación en Economía y Finanzas (CIENFI) de la Universidad Icesi, jcalonso@icesi.edu.co.

Al terminar este tutorial usted estará en capacidad de:

- Analizar gráficamente los residuos de una regresión para detectar la presencia de autocorrelación.
- Efectuar el test de Durbin-Watson.
- Efectuar el test de Rachas (Runs Test).
- Efectuar el test de Box-Pierce.
- Efectuar el test de Ljung-Box
- Corregir un modelo con autocorrelación por el método de Durbin.

Para este tutorial emplearemos datos generados del siguiente modelo que corresponde a un modelo de una sola ecuación que pretende explicar los movimientos de la tasa de interés de los bonos de tesorería de una economía ficticia:

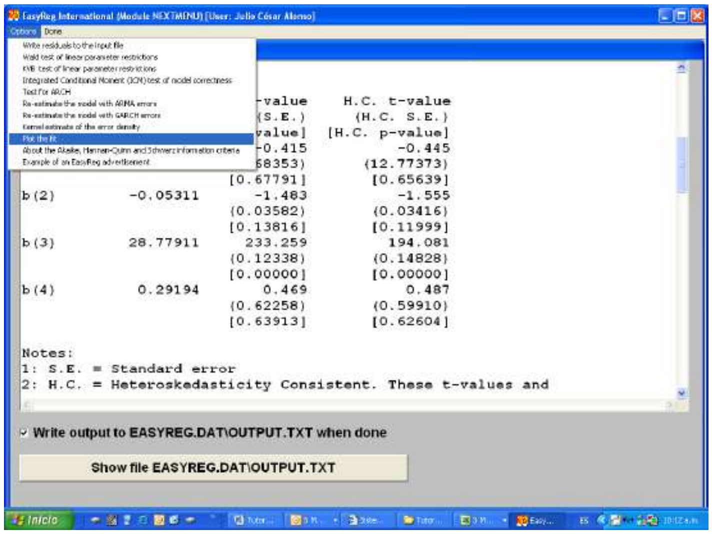
$$R_t = 1.41 + 0.097IP_t - 0.0769M_t + 29.2P_t + \varepsilon_t \tag{1}$$

$$\varepsilon_t = \rho\varepsilon_{t-1} + \nu_t \tag{2}$$

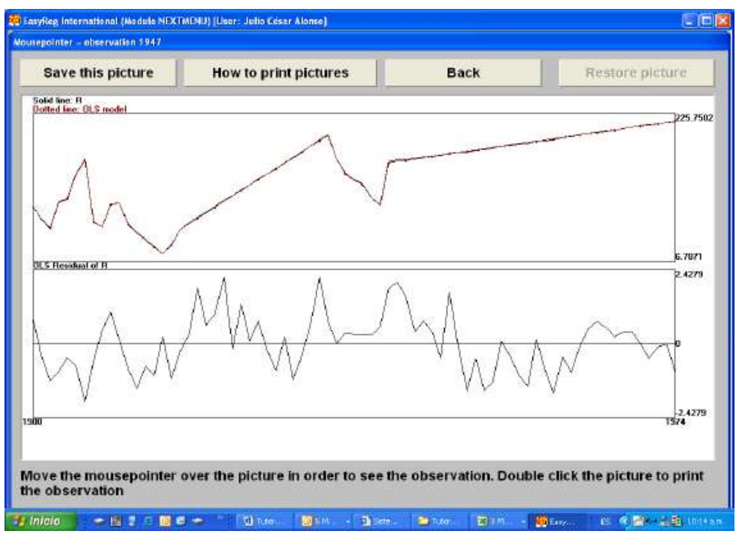
donde $\rho = 0.6$, $\nu_t \sim N(0,1)$, R_t representa la tasa de interés de los bonos de tesorería, IP_t corresponde a un índice de producción industrial, M_t corresponde a la demanda de dinero, y P_t corresponde al cambio en el nivel de precios. Se generaron 75 datos del anterior modelo correspondientes al período 1900-1974. Estos datos se encuentran en la página Web de Apuntes de Economía.

1 Análisis Gráfico de los residuos y test de Durbin-Watson

Cargue los datos y corra la regresión correspondiente al modelo (1) empleando el método de mínimos cuadrados ordinarios (Para ver detalles de cómo hacer esto revise Alonso (2007)). Haga clic en el botón *"Options"*. Note que cuando trabaja con series de tiempo, existe una posibilidad más en el anterior menú que no existe cuando se trabaja con datos de corte transversal. Para graficas los errores, haga clic en la opción *"Plot the fit"*.



Verá el siguiente gráfico:



En el panel superior del gráfico, EasyReg grafica los valores estimados de la variable dependiente ($\hat{R}_t$) y sus correspondientes valores observados (R_t). En el panel inferior, se encuentra el gráfico de los residuos estimados ($\hat{\varepsilon}_t$) versus el tiempo (Recuerde que usted puede grabar este gráfico haciendo clic en la opción "*Save this picture*".) En este gráfico, se ven claros síntomas de autocorrelación positiva, pues los residuos parecen tener una fuerte persistencia. En otras palabras, cuando la serie alcanza valores positivos estos tienden a persistir, lo mismo ocurre con valores negativos.

Ahora haga clic en el botón "Back", esto lo llevará de retorno a los resultados del modelo estimado. En los resultados, inmediatamente debajo del estadístico F global, observará el estadístico Durbin-Watson. En este caso, tenemos que DW es igual a 1.029736. Como lo discutimos, un DW menor que dos puede implicar la presencia de autocorrelación. Para efectuar una prueba formal de la presencia o no de autocorrelación, necesitamos consultar la tabla de valores críticos para el estadístico Durbin-Watson. En este caso, con un nivel de significancia del 5% (3 variables explicatorias y un intercepto y 75 observaciones) tenemos que $d_l = 1.543$ y $d_u = 1.709$.

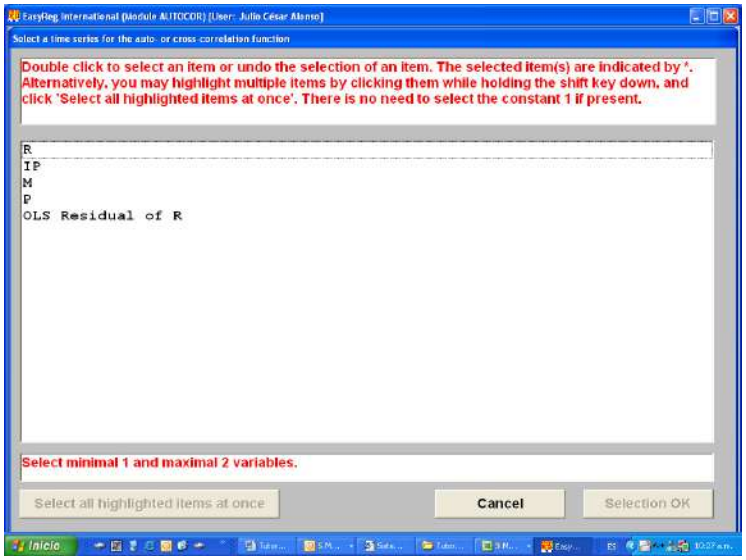
Entonces tenemos que $DW = 1.029 < d_u = 1.709$, y por tanto podemos rechazar la hipótesi nula de no autocorrelación, a favor de la hipótesis nula de que existe autocorrelación. Es más, dado que $DW < d_l$, podemos rechazar la hipótesis nula de no autocorrelación positiva a favor de la hipótesis alterna (autocorrelación positiva). Además, como es bien sabido, $DW \approx 2(1-\hat{\rho})$, de tal forma que $\hat{\rho} \approx 1 - \frac{DW}{2} = 0.4851$, sin embargo, este estimador es sesgado y no lo deberíamos emplear para cualquier corrección.

2 Cálculo de la autocorrelación para los residuos, test de Box-Pierce y de Ljung-Box para los residuos y gráfico de las autocorrelaciones

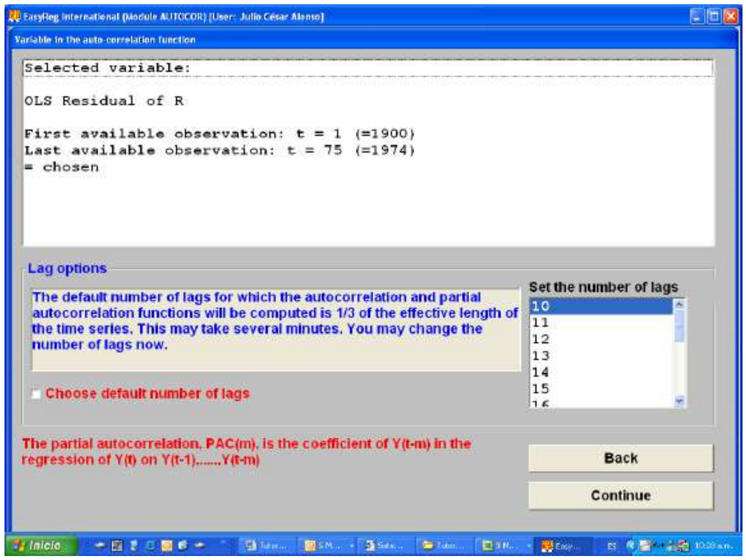
Haga clic en el menú "*"Options*" *y* seleccione la opción "*Write residuals to the input file*", esto creará una nueva variable con el nombre "*OLS Residuals of R*". Ahora regrese a la ventana principal de EasyReg (Haga clic en la opción "*Done*").

Ahora, haga clic en el botón "*Menu / Data analysis / Auto/Cross correlations*".

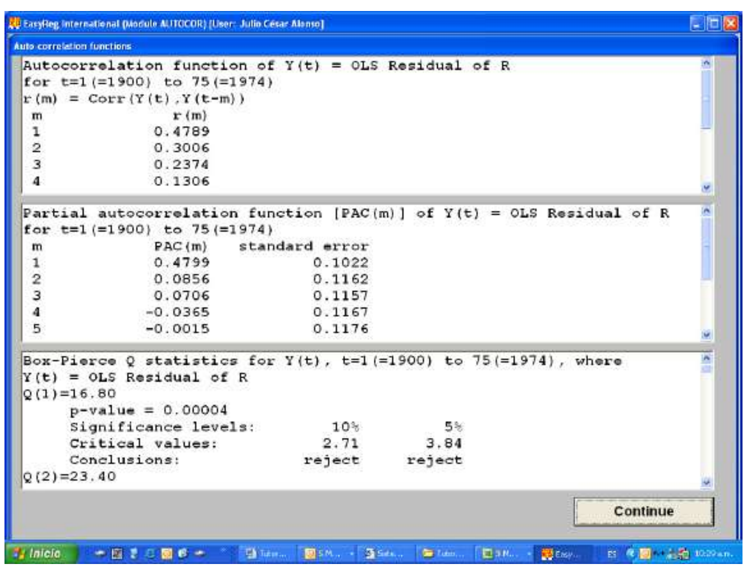
Ahora, verá la siguiente ventana en la que debe escoger la variable para la cual desea conocer las correlaciones y correlaciones parciales, así como los correspondientes test de Box-Pierce y de Ljung-Box.



Escoja la variable "*OLS Residual of R*" y haga clic en el botón "*Selection OK*". Haga clic en el botón "*No*" de la siguiente ventana y posteriormente en "*Continue*". Observará la siguiente ventana:



En esta ventana usted puede escoger el número de rezagos para los cuales quiere calcular la autocorrelación y la correlación parcial. El número de rezagos pre-establecido es un tercio de las observaciones. En nuestro caso, esto implicaría calcular hasta la autocorrelación 25, esto es excesivo y además demorado. En la ventana "*Set the number of* lags", escoja el mínimo número de rezagos, es decir, 10 en este caso. Haga clic en el botón "*Continue*". Observará la siguiente ventana:



En el primer panel, observará la autocorrelación para la serie de los residuos. En el segundo panel, encontrará la autocorrelación parcial y su respectiva desviación estándar. En el tercer panel, encontrarán los estadísticos de Box-Pierce y de Ljung-Box. Los resultados de ambas pruebas están reportados en las Tabla 1 y Tabla 2 .

Tabla 1. Resultados del Test de estadísticos de Box-Pierce sobre los residuos estimados

```
Box-Pierce Q statistics for Y(t), t=1(=1900) to 75(=1974), where
Y(t) = OLS Residual of R
Q(1)=16.80
    p-value = 0.00004
    Significance levels:     10%      5%
    Critical values:        2.71     3.84
    Conclusions:          reject   reject
Q(2)=23.40
    p-value = 0.00001
    Significance levels:     10%      5%
    Critical values:        4.61     5.99
    Conclusions:          reject   reject
Q(3)=27.42
    p-value = 0.00000
    Significance levels:     10%      5%
    Critical values:        6.25     7.81
    Conclusions:          reject   reject
Q(4)=28.62
    p-value = 0.00001
    Significance levels:     10%      5%
    Critical values:        7.78     9.49
    Conclusions:          reject   reject
Q(5)=29.10
    p-value = 0.00002
    Significance levels:     10%      5%
```

Critical values:	9.24	11.07
Conclusions:	reject	reject

Note que para probar la H_0:no autocorrelación versus H_A:por lo menos una autocorrelación es diferente de cero, podemos considerar diferentes rezagos. Siempre existirá el problema de determinar cuántos rezagos emplear para el test. La prueba puede no detectar la autocorrelación con rezagos de orden superior, si se consideran pocos rezagos. Sin embargo, si se emplean muchos rezagos, la prueba puede tener bajo poder. Para más detalles, consulte Ljung y Box (1979) y Harvey (1990, 1993).

Como se puede observar en la Tabla 1 y Tabla 2, los estadísticos de Box-Pierce y los de Ljung-Box son iguales a 16.8 si sólo se considera un rezago. En este caso, estos dos estadísticos son iguales dado que el número de observaciones es relativamente grande y el número de regresores es pequeño[2]. Este estadístico se compara con el valor de la distribución Chi-cuadrado con un grado de libertad (2.71 y 3.84 para niveles de significancia de 10% y 5%, respectivamente). Entonces se puede rechazar la hipótesis nula a favor de la presencia de un error que sigue un proceso autoregresivo. Noten que si se consideran más rezagos, la decisión sigue siendo la misma.

Tabla 2. Resultados del Test de estadísticos de Ljung-Box sobre los residuos estimados

Ljung-Box Q statistics for Y(t), t=1(=1900) to 75(=1974), where		
Y(t) = OLS Residual of R		
Q(1)=16.80		
p-value = 0.00004		
Significance levels:	10%	5%

---

[2] Note que los estadísticos de Box-Pierce y Ljung-Box son $Q = n\sum_{k=1}^{p} r_k^2$ y $Q' = n(n+2)\sum_{k=1}^{p} \frac{r_k^2}{n+k}$, respectivamente. Así cuando n es lo suficientemente grande y el número de regresores es pequeño, entonces $Q' = n\sum_{k=1}^{p} \frac{(n+2)r_k^2}{n+k} \approx n\sum_{k=1}^{p} r_k^2 = Q$.

Critical values: 2.71 3.84
Conclusions: reject reject

Q(2)=23.40
p-value = 0.00001
Significance levels: 10% 5%
Critical values: 4.61 5.99
Conclusions: reject reject

Q(3)=27.42
p-value = 0.00000
Significance levels: 10% 5%
Critical values: 6.25 7.81
Conclusions: reject reject

Q(4)=28.62
p-value = 0.00001
Significance levels: 10% 5%
Critical values: 7.78 9.49
Conclusions: reject reject

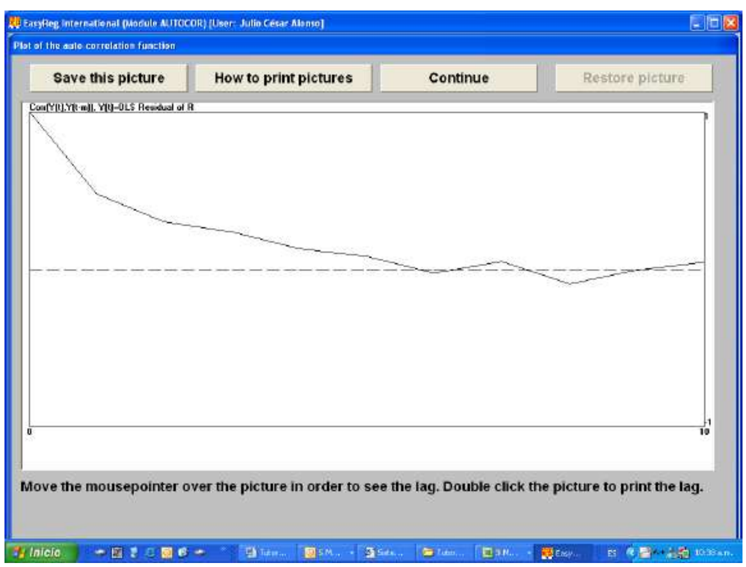
Q(5)=29.10
p-value = 0.00002
Significance levels: 10% 5%
Critical values: 9.24 11.07
Conclusions: reject reject

Q(6)=29.13
p-value = 0.00006
Significance levels: 10% 5%
Critical values: 10.64 12.59
Conclusions: reject reject

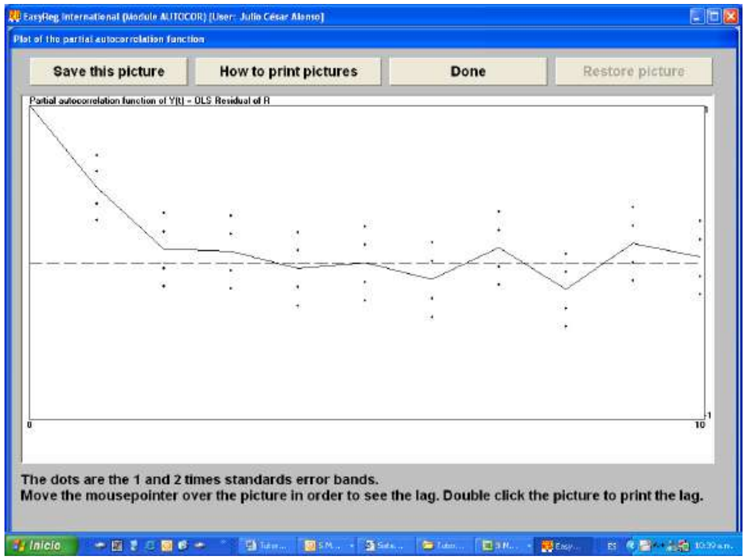
Q(7)=29.29
p-value = 0.00013
Significance levels: 10% 5%
Critical values: 12.02 14.07
Conclusions: reject reject

Q(8)=29.86		
p-value = 0.00022		
Significance levels:	10%	5%
Critical values:	13.36	15.51
Conclusions:	reject	reject
Q(9)=29.86		
p-value = 0.00046		
Significance levels:	10%	5%
Critical values:	14.68	16.92
Conclusions:	reject	reject
Q(10)=29.98		
p-value = 0.00086		
Significance levels:	10%	5%
Critical values:	15.99	18.31
Conclusions:	reject	reject

Haga clic en el botón "*Continue*". A continuación observará el gráfico de las correlaciones estimadas versus el número de rezagos.



Si hace clic en el botón "*Continue*" observará el gráfico de la función de autocorrelación parcial.



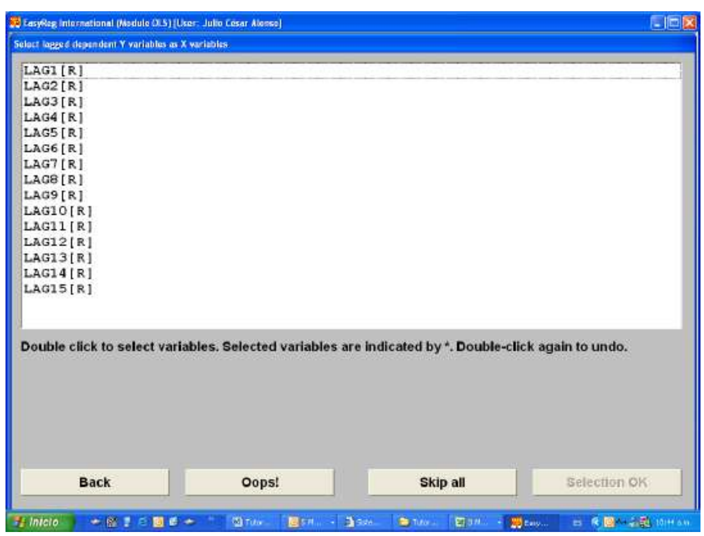
3 Método de corrección de Durbin

Los pasos para efectuar esta corrección son los siguientes:

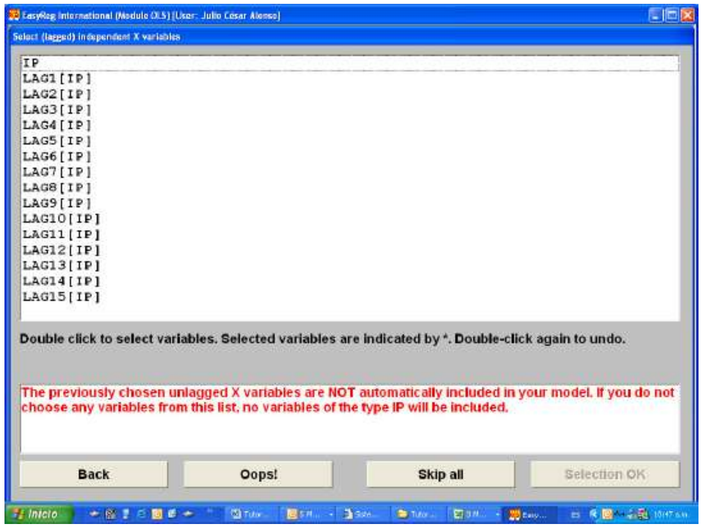
1. Estime el siguiente modelo:
$$R_t = \beta_1 + \beta_2 IP_t + \beta_3 M_t + \beta_4 P_t + \beta_5 IP_{t-1} + \beta_6 M_{t-1} + \beta_7 P_{t-1} + \rho R_{t-1} + \nu_t$$
2. Realice las siguientes transformaciones $R_t^* = R_t - \hat{\rho} R_{t-1}$, $IP_t^* = IP_t - \hat{\rho} IP_{t-1}$, etc.
3. Estime el siguiente modelo $R_t^* = \beta_1^* + \beta_2 IP_t^* + \beta_3 M_t^* + \beta_4 P_t^* + \nu_t$, donde $\beta_1^* = \beta_1 (1 - \hat{\rho})$.

Para correr la regresión del Paso 1 haga clic en "*Menu / Single equation Models/Linear regression models*".

A continuación, haga doble clic en las variables "*R*", "*M*", "*IP*", "*P*". Haga clic en el botón "*Selection OK*", después en "*No*" y posteriormente en "*Continue*". Escoja la variable dependiente "R" y posteriormente haga clic en "*Continue*" dos veces. En la ventana de selección de variables independientes, haga clic en *"Selection OK",* verá la siguiente ventana:



Haga nuevamente clic en "*LAG1[R]*" y posteriormente en el botón "*Selection OK*". Observará la siguiente ventana:



Ahora, seleccione "*IP*", luego "*LAG1[IP]*" y posteriormente haga clic en el botón "*Selection OK*". Repita este procedimiento con las otras variables independientes. Encontrará que $\hat{\rho} = 0.47307$.

Ya podemos efectuar el segundo paso, es decir, necesitamos crear las siguientes variables: $R_t^* = R_t - \hat{\rho} R_{t-1}$, $IP_t^* = IP_t - \hat{\rho} IP_{t-1}$, etc. Para ello, es necesario crear las variables rezagadas.

Para esto vaya al menú "*Transform variables*" y haga clic en el botón "*Time series transformations*". Haga clic en el botón "*Lag: x(t-m)*" y nuevamente en el botón "*Lag: x(t-m)*". Haga doble clic en las variables "*R*", "*M*", "*IP*", "*P*" y haga clic en el botón "*Selection OK*", posteriormente en el botón "*OK*". Las variables han sido creadas bajo el nombre "*LAG1[R]*", "*LAG1[M]*", etc. Cree las variables $R_t^* = R_t - \hat{\rho} R_{t-1}$, $IP_t^* = IP_t - \hat{\rho} IP_{t-1}$, etc. (Emplee la opción "*linear combination of variables*").

Ahora, estime el modelo $R_t^* = \beta_1^* + \beta_2 IP_t^* + \beta_3 M_t^* + \beta_4 P_t^* + \nu_t$; obtendrá el siguiente resultado (Note que este modelo ya no tiene ningún problema de autocorrelación).

Tabla 3. Modelo corregido por el método de Durbin.

```
Dependent variable:
Y = R-.47307xLAG1[R]

Characteristics:
R-.47307xLAG1[R]
  First observation = 2(=1971)
  Last observation  = 75(=2044)
  Number of usable observations: 74
  Minimum value: -1.7484537E+001
  Maximum value: 1.1890411E+002
  Sample mean:   7.6080162E+001

X variables:
X(1) = IP-.47307xLAG1[IP]
X(2) = M-.47307xLAG1[M]
X(3) = P-.47307xLAG1[P]
X(4) = 1

Model:
Y = b(1)X(1) +.....+ b(4)X(4)  + U,
where U is the error term, satisfying
E[U|X(1),...,X(4)] = 0.

OLS estimation results
Parameters  Estimate      t-value    H.C. t-value
                          (S.E.)     (H.C. S.E.)
                          [p-value]  [H.C. p-value]
      b(1)    -1.47407    -0.093       -0.082
                        (15.80407)   (18.02629)
                         [0.92569]    [0.93483]
      b(2)    -0.06384    -1.563       -1.318
```

		(0.04084)	(0.04844)
		[0.11801]	[0.18754]
b(3)	28.78570	189.377	156.295
		(0.15200)	(0.18418)
		[0.00000]	[0.00000]
b(4)	0.18549	0.448	0.385
		(0.41370)	(0.48235)
		[0.65389]	[0.70057]

Notes:

1: S.E. = Standard error

2: H.C. = Heteroskedasticity Consistent. These t-values and standard errors are based on White's heteroskedasticity consistent variance matrix.

3: The two-sided p-values are based on the normal approximation.

Effective sample size (n) = 74

Variance of the residuals = 0.887503

Standard error of the residuals = 0.942074

Residual sum of squares (RSS)= 62.125182

Total sum of squares (TSS) = 88644.817900

R-square = 0.999299

Adjusted R-square = 0.99926

Overall F test: F(3,70) = 33270.40

p-value = 0.00000

Significance levels:	10%	5%
Critical values:	2.16	2.74
Conclusions:	reject	reject

4 Referencias.

- Alonso (2007) "Tutorial para la estimación de un modelo de regresión múltiple e inferencia con EasyReg" Apunte de Economía, No.13, Septiembre.

- Bierens, H. J. (2010), "EasyReg International", Department of Economics, Pennsylvania State University(http://econ.la.psu.edu/~hbierens/EASYREG.HTM)

- Harvey, Andrew C. (1990) The Econometric Analysis of Time Series, 2nd edition, MIT Press.

- Harvey, Andrew C. (1993) Time Series Models, 2nd edition, MIT Press.

- Ljung, G. and G. Box (1979) "On a Measure of Lack of Fit in Time Series Models," Biometrika, 66, 265–270.