

**Tutorial para la estimación de Ecuaciones Simultáneas por el método de Mínimos Cuadrados en dos Etapas en Easyreg**

**Julio César Alonso C.**

**No. 18**

**Marzo de 2009**

## APUNTES DE ECONOMÍA

**ISSN 1794-029X**

No. 18, Marzo de 2009

Editor

Julio César Alonso C.

[jcalonso@icesi.edu.co](mailto:jcalonso@icesi.edu.co)

Vanessa Ospina López

Asistente de Edición

Gestión Editorial

Departamento de Economía - Universidad Icesi

[www.icesi.edu.co](http://www.icesi.edu.co)

Tel: 5552334 ext: 8398. Fax: 5551441

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

## TUTORIAL PARA LA ESTIMACIÓN DE ECUACIONES SIMULTÁNEAS POR EL MÉTODO DE MÍNIMOS CUADRADOS EN DOS ETAPAS EN EASYREG

Julio Cesar Alonso C<sup>1</sup>.

Marzo de 2009

### Resumen

*Este documento presenta una breve introducción a cómo estimar un modelo por el método de mínimos cuadrados en dos etapas con el paquete econométrico gratuito 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, Mínimos Cuadrados en dos etapas, Modelo de Klein, Variables instrumento.

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.

---

<sup>1</sup> Profesor del Departamento de Economía y Director del Centro de Investigación en Economía y Finanzas (CIENFI) de la Universidad Icesi, [icalonso@icesi.edu.co](mailto:icalonso@icesi.edu.co).

Al terminar este documento usted estará en capacidad de estimar una ecuación por el método de mínimos cuadrados en dos etapas (MC2E) en EasyReg.

Para este tutorial nos basaremos en el Ejemplo 18.6 y 20.2 de Gujarati. (Página 627 y 677). El ejemplo considera el siguiente modelo macroeconómico conocido como el modelo I de Klein para la economía de Estados Unidos en el período 1920-1941:

$$\text{Función de Consumo} \quad C_t = \beta_0 + \beta_1 P_t + \beta_2 (W + W')_t + \beta_3 P_{t-1} + u_{1t} \quad (1)$$

$$\text{Función de Inversión} \quad I_t = \beta_4 + \beta_5 P_t + \beta_6 P_{t-1} + \beta_7 K_{t-1} + u_{2t} \quad (2)$$

$$\text{Demanda de Trabajo} \quad W_t = \beta_8 + \beta_9 X_t + \beta_{10} X_{t-1} + \beta_{11} t + u_{3t} \quad (3)$$

$$\text{Identidad} \quad Y_t + T_t = C_t + I_t + G_t \quad (4)$$

$$\text{Identidad} \quad Y_t = W'_t + W_t + P_t \quad (5)$$

$$\text{Identidad} \quad K_t = K_{t-1} + I_t \quad (6)$$

$$\text{Identidad} \quad X_t = (Y + T - W')_t \quad (7)$$

donde  $C_t$  = Consumo  $I_t$  = Inversión

$G_t$  = Gasto del gobierno

$P_t$  = Utilidades

$K_t$  = Stock de Capital

$W_t$  = Nómina del sector Privado

$T_t$  = Impuestos

$W'_t$  = Nómina del sector público

$t$  = tiempo

$Y_t$  = Ingresos después de Impuestos

Los datos para este ejercicio serán encontrados en la página Web de Apuntes de Economía.

Note que existen siete variables endógenas ( $C_t$ ,  $I_t$ ,  $P_t$ ,  $X_t$ ,  $W_t$ ,  $K_t$  y  $Y_t$ ) y siete exógenas ( $P_{t-1}$ ,  $W'_t$ ,  $K_{t-1}$ ,  $T_t$ ,  $t$ ,  $G_t$  y  $X_{t-1}$ ). El lector puede darse cuenta fácilmente que las primeras tres ecuaciones de comportamiento están sobre-identificadas.

Es bien sabido, que estas ecuaciones estructurales de (1) a (3) no se deben estimar directamente por el método de MCO, pues existe una gran probabilidad que los estimadores MCO sean sesgados gracias a la posible correlación del término de error,

de cada una de las tres primeras ecuaciones, y las variables endógenas que aparecen como variables explicativas. Así, no podemos emplear el método de Mínimos Cuadrados Ordinarios para estimar los coeficientes de estas ecuaciones estructurales, pero podemos emplear el método de MC2E.

A continuación, encontrará los pasos a seguir para estimar la ecuación (1) por MC2E. Además se discutirán brevemente los pasos necesarios para efectuar el test de simultaneidad de Hausman.

Antes de iniciar, cargue los datos en EasyReg International y cree una nueva variable:  $(W + W')_t$ .

### 1 Estimación por el Método de Mínimos Cuadrados en Dos Etapas.

Noten que el sistema de la forma estructural se puede describir de la siguiente forma reducida:

$$P_t = \pi_{11} + \pi_{12}P_{t-1} + \pi_{13}K_{t-1} + \pi_{14}X_{t-1} + \pi_{15}t + \pi_{16}T + \pi_{17}G + \mu_{1t} \quad (8)$$

$$(W + W')_t = \pi_{21} + \pi_{22}P_{t-1} + \pi_{23}K_{t-1} + \pi_{24}X_{t-1} + \pi_{25}t + \pi_{26}T + \pi_{27}G + \mu_{2t} \quad (9)$$

$$X_t = \pi_{31} + \pi_{32}P_{t-1} + \pi_{33}K_{t-1} + \pi_{34}X_{t-1} + \pi_{35}t + \pi_{36}T + \pi_{37}G + \mu_{3t} \quad (10)$$

Es importante darse cuenta que el resto de variables endógenas se pueden determinar por medio de identidades.

Los MC2E implican dos pasos. Si queremos estimar la ecuación (1), primero se deben estimar  $\hat{P}_t$  y  $(\overline{W} + W')_t$  a partir de las ecuaciones de la forma reducida (8) y (9)<sup>2</sup>. El segundo paso es estimar  $C_t = \beta_0 + \beta_1\hat{P}_t + \beta_2(\overline{W} + W')_t + \beta_3P_{t-1} + u_{1t}$ . Estos dos pasos son efectuados simultáneamente por EasyReg International. A continuación, veremos

<sup>2</sup> Esta forma reducida puede ser estimada por el método de MCO. Para ver como hacerlo en EasyReg, puede ver Alonso(2007).

cómo efectuar una estimación por el método de mínimos cuadrados en dos etapas (también conocido como el método de variables instrumento).

Antes de ver cómo efectuar la estimación de **¡Error! No se encuentra el origen de la referencia.** por el método de MC2E, es importante diferenciar las variables en cuatro grupos:

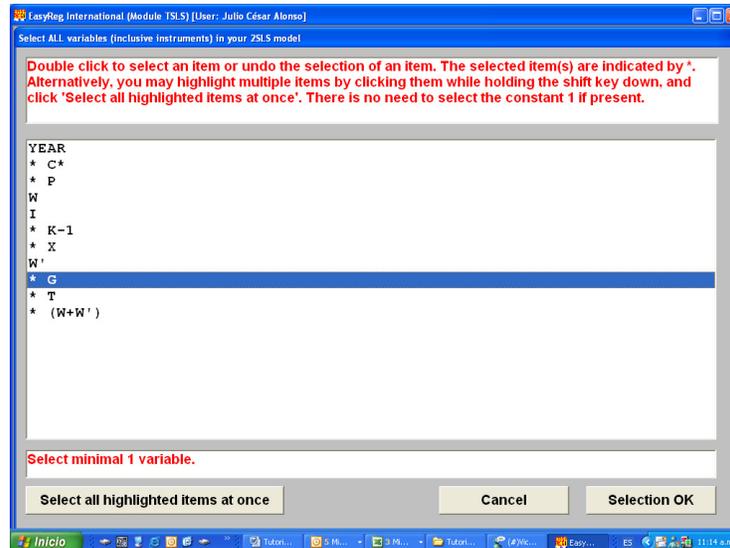
- variable endógena dependiente para la ecuación (1) ( $C_t$ ),
- variables endógenas explicativas para la ecuación **¡Error! No se encuentra el origen de la referencia.** ( $P_t$  y  $(W + W')$ ),
- variables exógenas o predeterminadas que se emplean como variables explicativas en la ecuación **¡Error! No se encuentra el origen de la referencia.** ( $P_{t-1}$ ), y
- otras variables exógenas ( $K_{t-1}$ ,  $T_t$ ,  $t$ ,  $G_t$  y  $X_{t-1}$ ).

Ahora, haga clic en “Menu/Single equation models/Two-stage least squares models”, observará la siguiente ventana:

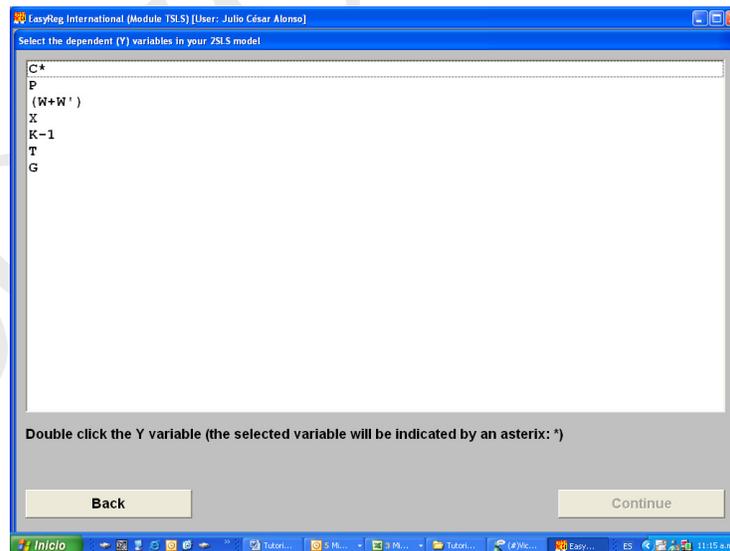


Posteriormente, observará una ventana muy familiar para usted. En esta ventana escoja, haciendo doble clic sobre ellas, todas las variables que mencionamos

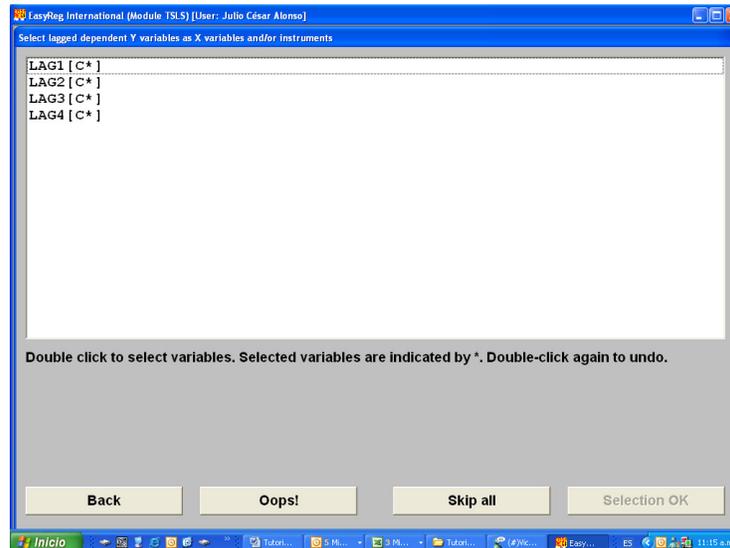
anteriormente en la clasificación de variables en cuatro grupos, i.e., “C\*”, “P”, “W+W’”, “X”, “K-1”, “T”, y “G”.



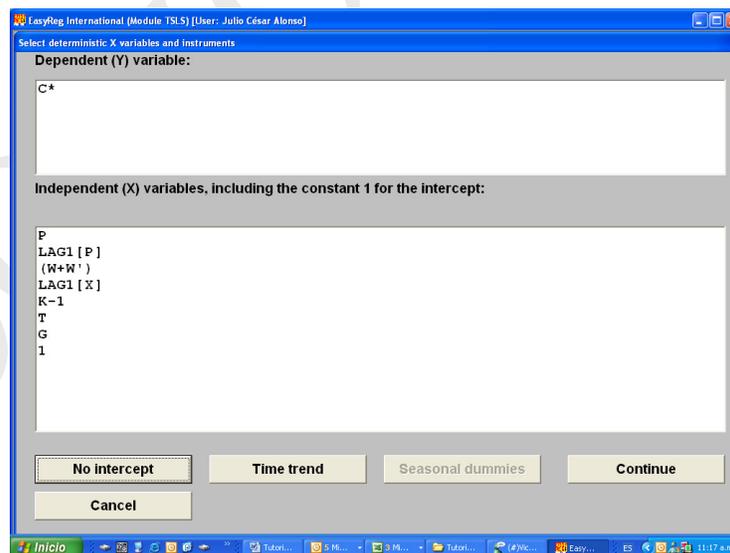
Haga clic en el botón “*Selection OK*”. A continuación haga clic en el botón “*No*” y posteriormente en el botón “*Continue*”. Observará la siguiente ventana:



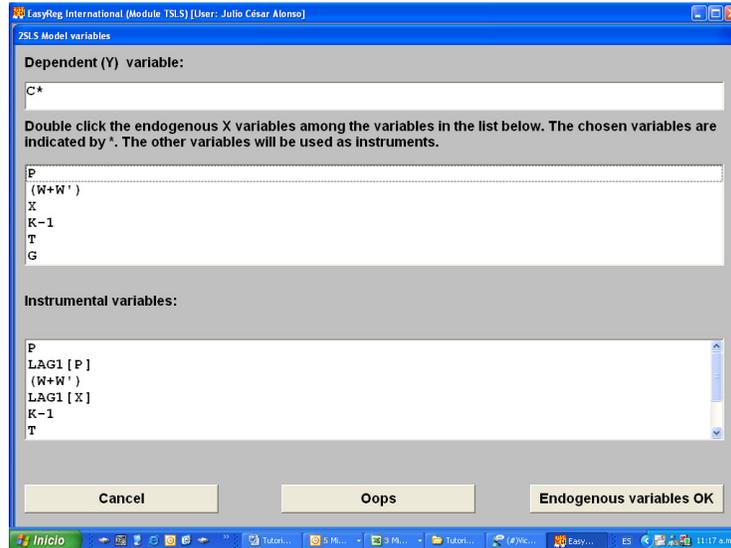
Escoja la variable “C\*” y haga clic en el botón “*Continue*” dos veces. En la siguiente ventana, haga clic en el botón “*Selection OK*”. Observará la siguiente ventana:



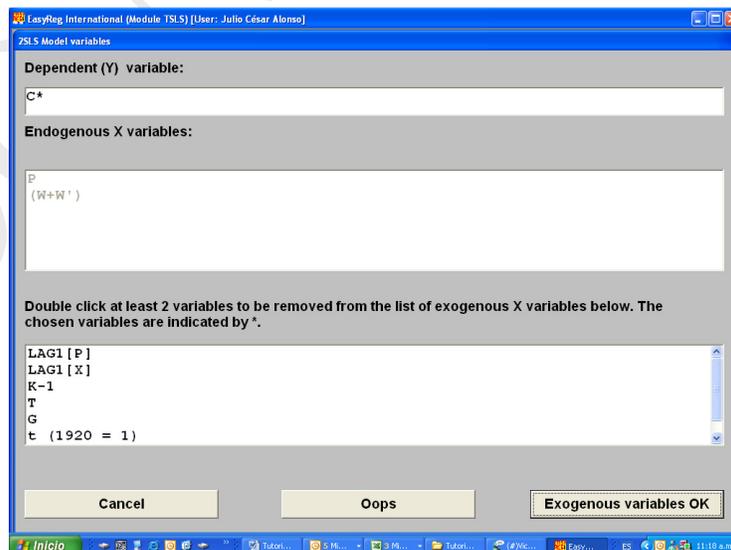
Haga clic en el botón “*Skip all*”. Tenga mucho cuidado en las siguientes ventanas, asegúrese de escoger todas las variables necesarias. Debe escoger las siguientes variables: “*P*”, “*LAG1[P]*”, “*K-1*”, “*T*”, “*W+W*”, “*G*” y “*LAG1[X]*” (¡¡¡Atención!!! No incluya la variable “*X*”, sólo incluya “*LAG1[X]*”). Una vez seleccionadas estas variables observará la siguiente ventana:



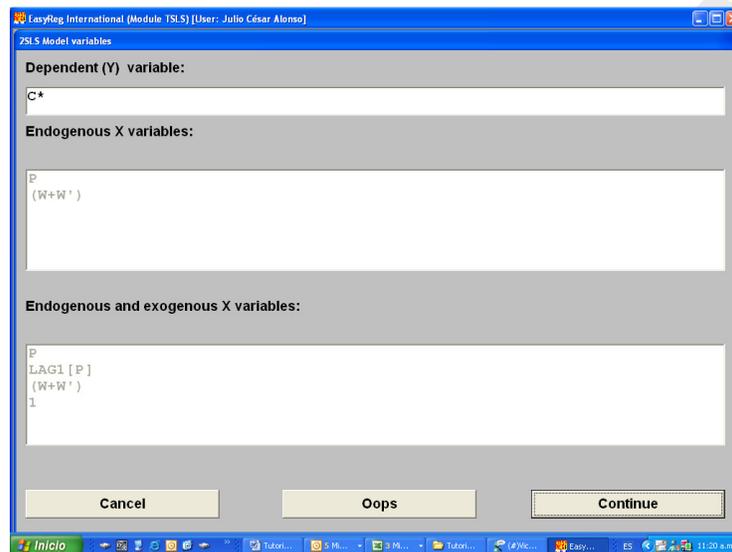
Ahora, se debe incluir una tendencia, para esto haga clic en el botón “*Time trend*” y posteriormente en el botón “*Continue*” dos veces. Observará la siguiente ventana.



Recuerde que se desea estimar la siguiente ecuación:  $C_t = \beta_0 + \beta_1 P_t + \beta_2 (W + W')_t + \beta_3 P_{t-1} + u_{1t}$ . En la primera ventana, “*Dependent (Y) variable*”, encontrará la variable dependiente de la ecuación estimada. En la segunda ventana, usted debe escoger las variables endógenas incluidas como variables explicativas; en este caso escoja “*P*” y “*W+W*” y haga clic en el botón “*Endogenous variables OK*”. Observará la siguiente ventana.



En la tercera ventana, panel inferior, se deben dejar las variables exógenas empleadas como variables explicativas en la ecuación (1). Para esto, tenemos que hacer clic en aquellas variables que no se emplean como variables explicativas y son exógenas en la ecuación  $C_t = \beta_0 + \beta_1 P_t + \beta_2 (W + W')_t + \beta_3 P_{t-1} + u_{1t}$ . Escoja las variables “K-1”, “T”, “G”, “LAG1[X]” “t(1920=1)” haciendo doble clic sobre ellas y posteriormente haga clic en el botón “Exogenous variables OK”. Observará la siguiente ventana:



Haga clic en el botón “Continue” dos veces. Observará los resultados reportados en la Tabla 1.

Note que la estructura de estos resultados es mas o menos parecida a los que se obtienen cuando se estima un modelo por el método de MCO. Así, en este caso tendremos que la ecuación estructural estimada está dada por (t-calculado entre paréntesis)

$$C_t = 16.551662 + 0.016946P_t + 0.810369(W + W')_t + 0.216319P_{t-1}$$

(11.271)
(0.129)
(18.105)
(1.814)

Además, note que todas las opciones de crear gráficos y de test de Wald están también disponibles para las estimaciones de MC2E.

Antes de continuar estimando las otras ecuaciones, es importante resaltar que es necesario incluir en nuestra estimación todas las variables exógenas del modelo, así aparentemente en el último paso algunas de estas variables exógenas sean descartadas.

Recuerden que los MC2E implican la estimación de los instrumentos  $\hat{P}_t$  y  $(\overline{W} + W')_t$ , a partir de todas las variables exógenas del modelo ( $P_t = \pi_{11} + \pi_{12}P_{t-1} + \pi_{13}K_{t-1} + \pi_{14}X_{t-1} + \pi_{15}t + \pi_{16}T + \pi_{17}G + \mu_{1t}$ , y  $(W + W')_t = \pi_{21} + \pi_{22}P_{t-1} + \pi_{23}K_{t-1} + \pi_{24}X_{t-1} + \pi_{25}t + \pi_{26}T + \pi_{27}G + \mu_{2t}$ ). Posteriormente, estos valores estimados son empleados en la estimación de la ecuación estructural deseada:  $C_t = \beta_0 + \beta_1\hat{P}_t + \beta_2(\overline{W} + W')_t + \beta_3P_{t-1} + u_{1t}$ . Así, es necesario seleccionar todas las variables endógenas del sistema pues esto le permite a EasyReg Internacional reconocer cuáles son todas las variables exógenas del sistema que serán empleadas en la estimación de los instrumentos  $\hat{P}_t$  y  $(\overline{W} + W')_t$ . Noten que el último paso de la estimación de los MC2E, implica la exclusión de las variables exógenas que están en el sistema, pero que no se encuentran en la ecuación estructural particular a estimar. Así, debe ser claro que no es lo mismo no incluir desde un inicio todas las variables endógenas del sistema que incluirlas.

**Tabla 1. Resultados de la estimación de la ecuación (1) por el método de MC2E.**

Two-stage least squares:
Dependent variable:
$Y = C^*$
Characteristics:
$C^*$
First observation = 1(=1920)
Last observation = 22(=1941)
Number of usable observations: 22

**Tabla 1. Resultados de la estimación de la ecuación (1) por el método de MC2E.  
(Cont.)**

Minimum value: 3.9800000E+001		
Maximum value: 6.9700000E+001		
Sample mean: 5.3350000E+001		
X variables, including instrumental variables:		
X(1) = P		
X(2) = LAG1[P]		
X(3) = (W+W')		
X(4) = LAG1[X]		
X(5) = K-1		
X(6) = T		
X(7) = G		
X(8) = t (1920 = 1)		
X(9) = 1		
WARNING: The effective degrees of freedom is only 12. Therefore, the estimation results may be unreliable!		
Endogenous X variables:		
Y*(1)=P		
Y*(2)=(W+W')		
Exogenous X variables:		
X*(1)=LAG1[P]		
X*(2)=1		
2SLS estimation results for Y = C*		
Variables	2SLS estimate	t-value [p-value]
P	0.016946	0.129

		[0,89728]
LAG1[P]	0.216319	1.814
		[0,06970]
(W+W')	0.810369	18.105
		[0,00000]
1	16.551662	11.271
		[0,00000]

[The p-values are two-sided and based on the normal approximation]

Standard error of the residuals = 11.359987E-001

Residual sum of squares (RSS) = 21.938381E+000

Total sum of squares (TSS) = 94.142950E+001

R-square = 0.976697

Adjusted R-square = 0.972584

Effective sample size (n) = 21

If the model is correctly specified, in the sense that the conditional expectation of the model error  $u$  relative to the instrumental variables and all lagged dependent ( $Y$ ) variables and lagged instrumental variables equals zero, then the 2SLS parameter estimators  $b(1), \dots, b(4)$ , minus their true values, times the square root of the sample size  $n (=21)$ , are (asymptotically) jointly normally distributed with zero mean vector and variance matrix:

$$\begin{bmatrix} 0.3617949 & -0.2484492 & -0.0396465 & -0.3976598 \\ -0.2484492 & 0.2986728 & -0.0137703 & -0.1234936 \\ -0.0396465 & -0.0137703 & 0.0420702 & -0.8499592 \\ -0.3976598 & -0.1234936 & -0.8499592 & 45.2866280 \end{bmatrix}$$

provided that the conditional variance of the model error  $u$  is constant ( $u$  is homoskedastic)

Noten que los resultados difieren en algunos decimales de la Tabla 20.6 reportada en el libro de Gujarati, esta discrepancia se debe a que no estamos empleando todos los decimales de los datos originales.

Ahora, se puede proceder a estimar las otras dos ecuaciones de comportamiento del sistema de ecuaciones de la forma estructural.

Noten que para el caso de la segunda ecuación ( $I_t = \beta_4 + \beta_5 P_t + \beta_6 P_{t-1} + \beta_7 K_{t-1} + u_{2t}$ ) tenemos los siguientes grupos de variables:

- variable endógena dependiente para la ecuación (2) ( $I_t$ ),
- variables endógenas explicativas para la ecuación (2) ( $P_t$ ),
- variables exógenas o predeterminadas que se emplean como variables explicativas en la ecuación (2) ( $P_{t-1}$  y  $K_{t-1}$ ), y,
- otras variables exógenas ( $T_t$ ,  $t$ ,  $G_t$  y  $X_{t-1}$ ).

Si usted replica los pasos efectuados anteriormente haciendo las modificaciones necesarias, encontrará los resultados reportados en la Tabla 2. (Asegúrese que sí puede replicar estos resultados).

**Tabla 2. Resultados de la estimación de la ecuación (2) por el método de MC2E.**

Two-stage least squares: Dependent variable: Y = I  Characteristics: I First observation = 1(=1920) Last observation = 22(=1941) Number of usable observations: 22 Minimum value: -6.2000000E+000 Maximum value: 5.6000000E+000 Sample mean: 1.3318182E+000  X variables, including instrumental variables: X(1) = P X(2) = LAG1[P] X(3) = K-1 X(4) = T X(5) = G X(6) = LAG1[X] X(7) = t (1920 = 1) X(8) = 1  WARNING: The effective degrees of freedom is only 13. Therefore, the estimation results may be unreliable!  Endogenous X variable: Y*=P
----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Exogenous X variables:

$$X^*(1)=LAG1[P]$$

$$X^*(2)=K-1$$

$$X^*(3)=1$$

2SLS estimation results for  $Y = I$

Variables	2SLS estimate	t-value [p-value]
P	0.149318	0.775 [0,43855]
LAG1[P]	0.616719	3.405 [0,00066]
K-1	-0.157914	-3.928 [0,00009]
1	20.306052	2.419 [0,01554]

[The p-values are two-sided and based on the normal approximation]

Standard error of the residuals = 13.085972E-001

Residual sum of squares (RSS) = 29.111256E+000

Total sum of squares (TSS) = 25.232666E+001

R-square = 0.884629

Adjusted R-square = 0.864269

Effective sample size (n) = 21

If the model is correctly specified, in the sense that the conditional expectation of the model error  $u$  relative to the instrumental variables and all lagged dependent ( $Y$ ) variables and lagged instrumental variables equals zero, then the 2SLS parameter estimators  $b(1), \dots, b(4)$ , minus their true values, times the square root of the sample size  $n$  ( $=21$ ), are (asymptotically) jointly normally distributed with zero mean vector and variance matrix:

0.7802583	-0.6700960	0.1089401	-24.0472928
-0.6700960	0.6890001	-0.1006685	20.2186011
0.1089401	-0.1006685	0.0339326	-6.9948031
-24.0472928	20.2186011	-6.9948031	1479.2036935

provided that the conditional variance of the model error  $u$  is constant  
( $u$  is homoskedastic)

Finalmente, se puede proceder a estimar la última ecuación de comportamiento del sistema de ecuaciones de la forma estructural. Note que para el caso de la tercera ecuación,  $(W_t = \beta_8 + \beta_9 X_t + \beta_{10} X_{t-1} + \beta_{11} t + u_{3t})$ , tenemos los siguientes grupos de variables:

- variable endógena dependiente para la ecuación (3) ( $W_t$ ),
- variables endógenas explicativas para la ecuación (3) ( $X_t$ ),
- variables exógenas o predeterminadas que se emplean como variables explicativas en la ecuación (3) ( $X_{t-1}$  y  $t$ ), y
- otras variables exógenas ( $T_t$ ,  $G_t$ ,  $P_{t-1}$  y  $K_{t-1}$ ).

Si ustedes replican los pasos efectuados anteriormente haciendo las modificaciones necesarias, encontrarán los resultados reportados en la Tabla 3. (Asegúrese que si puede replicar estos resultados).

**Tabla 3. Resultados de la estimación de la ecuación (3) por el método de MC2E.**

Two-stage least squares: Dependent variable: $Y = W$  Characteristics: W First observation = 1(=1920) Last observation = 22(=1941) Number of usable observations: 22 Minimum value: 2.5500000E+001 Maximum value: 5.3300000E+001 Sample mean: 3.6018182E+001  X variables, including instrumental variables: $X(1) = X$ $X(2) = \text{LAG1}[X]$ $X(3) = T$ $X(4) = G$ $X(5) = \text{LAG1}[P]$ $X(6) = K-1$ $X(7) = t (1920 = 1)$ $X(8) = 1$  WARNING: The effective degrees of freedom is only 13. Therefore, the estimation results may be unreliable!  Endogenous X variable: $Y^*=X$  Exogenous X variables:
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

$$X^*(1)=LAG1[X]$$

$$X^*(2)=t(1920 = 1)$$

$$X^*(3)=1$$

2SLS estimation results for  $Y = W$

Variables	2SLS estimate	t-value	[p-value]
X	0.438637	11.073	[0,00000]
LAG1[X]	0.146883	3.402	[0,00067]
t(1920 = 1)	0.130450	4.028	[0,00006]
1	-0.063932	-0.056	[0,95555]

[The p-values are two-sided and based on the normal approximation]

Standard error of the residuals = 76.716227E-002

Residual sum of squares (RSS) = 10.005145E+000

Total sum of squares (TSS) = 79.490955E+001

R-square = 0.987413

Adjusted R-square = 0.985192

Effective sample size (n) = 21

If the model is correctly specified, in the sense that the conditional expectation of the model error  $u$  relative to the instrumental variables and all lagged dependent ( $Y$ ) variables and lagged instrumental variables equals zero, then the 2SLS parameter estimators  $b(1), \dots, b(4)$ , minus their true values, times the square root of the sample size  $n (=21)$ , are (asymptotically) jointly normally distributed with zero mean vector and variance matrix:

**Tabla 3. Resultados de la Estimación de la ecuación (3) por el método de MC2E.  
(Cont.).**

0.0329542	-0.0311395	-0.0080242	-0.0772017
-0.0311395	0.0391423	0.0013277	-0.4154792
-0.0080242	0.0013277	0.0220306	0.1405569
-0.0772017	-0.4154792	0.1405569	27.6302283
provided that the conditional variance of the model error $u$ is constant			

## 2 Referencias

Alonso, Julio C. (2007) **“Tutorial para la estimación de un modelo de regresión múltiple e inferencia con EasyReg”**, Apuntes de economía, No. 13. Septiembre.

Bierens, H. J. (2009), **"EasyReg International"**, Department of Economics, Pennsylvania State University.  
(<http://econ.la.psu.edu/~hbierens/EASYREG.HTM> )

Gujarati D. (1997) *Econometría Básica* 3ª Edición Mc.Graw Hill