Estudio de Montecarlo sobre métodos alternativos de estimación en un modelo con perturbaciones autocorrelacionadas

por EZEQUIEL URIEL JIMENEZ

Estadístico facultativo

0. INTRODUCCION

Griliches y Rao (1969) realizaron un estudio de Montecarlo sobre las propiedades de muestras pequeñas de varios métodos de estimación en modelos con perturbaciones autocorrelacionadas. El fin del presente trabajo es precisamente el investigar algunos aspectos sobre este tema que no han sido tratados por Griliches y Rao (como es la aplicación del método de «tanteo» [*scanning*]), o que no han recibido la suficiente atención en el estudio de estos autores (como es la influencia de la corrección de la primera observación según distintos tamaños de muestra).

En el epígrafe 1 plantearemos el problema, considerando distintos métodos alternativos de estimación. En el epígrafe 2 nos ocuparemos del diseño del experimento de simulación que vamos a realizar, dejando para el último epígrafe el análisis de resultados obtenidos.

1. PLANTEAMIENTO DEL PROBLEMA

En el modelo de regresión lineal clásico (1), las hipótesis básicas que se hacen sobre el término de perturbación aleatoria u_t son las siguientes:

1) $E(u_t) = 0.$
2) $E(u_t)^2 = \sigma^2.$
3) $E(u_t u_{t'}) = 0 \quad$ para $t \neq t'.$
4) u_t tiene distribución normal.

(1) Algunos autores, como A. S. Goldberger (1964), no incluyen la hipótesis de normalidad de las perturbaciones dentro del modelo lineal clásico. Cuando introducen dicha hipótesis, añaden el calificativo de normal. En nuestra exposición, el modelo lineal clásico (sin calificativos) incluirá la hipótesis de normalidad de las perturbaciones.

Las hipótesis anteriores pueden formularse conjuntamente en una sola diciendo que el vector de perturbaciones tiene una distribución normal multivariante con media 0 y matriz de covarianzas escalar. También se puede expresar de esta forma equivalente: el vector de perturbaciones es un vector normal esférico. Analíticamente,

$$u \sim\!\!> N(0, \sigma^2 I) \tag{1-1}$$

Si la hipótesis 3) no resulta admisible, como consecuencia de aplicar alguno de los contrastes pertinentes, es preciso reemplazarla por una hipótesis alternativa.

La hipótesis alternativa más utilizada es la de suponer que las perturbaciones siguen un esquema autorregresivo de primer orden:

$$\begin{aligned} &u_t = \rho\, u_{t-1} + \epsilon_t \\ &E(\epsilon) = 0 \\ &E(\epsilon\epsilon') = \sigma^2_\epsilon I \\ &\epsilon_t \text{ tiene distribución normal.} \end{aligned} \tag{1-2}$$

Por supuesto, se pueden adoptar esquemas autorregresivos de orden superior (es aconsejable adoptar un esquema autorregresivo de cuarto orden cuando se trabaja con series trimestrales sin desestacionalizar). Por nuestra parte, nos limitaremos a exponer el caso de esquemas de primer orden, ya que las soluciones que se obtengan son generalizables para esquemas de orden superior. También, con objeto de no complicar la notación, vamos a considerar un modelo con una sola variable exógena

$$Y_t = b_0 + b_1 X_t + u_t \tag{1-3}$$

que en notación matricial lo expresaremos por

$$y = Xb + u$$

La hipótesis, o mejor conjunto de hipótesis [1-2], reemplazan a las hipótesis del modelo lineal clásico formulado. En realidad, con este cambio, la única que queda alterada es la 3).

Bajo los anteriores supuestos, se puede deducir que la nueva matriz de covarianzas de las perturbaciones es la siguiente:

$$E(uu') = \sigma^2 W = \begin{bmatrix} 1 & \rho & \rho^2 & \cdots\cdots & \rho^{T-1} \\ \rho & 1 & \rho & \cdots\cdots & \rho^{T-2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \rho^{T-1} & \rho^{T-2} & \rho^{T-3} & \cdots\cdots & 1 \end{bmatrix} \tag{1-4}$$

En el caso de que se conozca ρ, la aplicación de MCG sería inmediata y aseguraría la obtención de estimadores lineales insesgados óptimos.

La aplicación de mínimos cuadrados generalizados (MCG) puede hacerse por cualquiera de estos dos procedimientos alternativos, pero que conducen a los mismos resultados (2):

1) Aplicación directa de la fórmula

$$\hat{b} = (X' W^{-1} X)^{-1} X' W^{-1} y \qquad [1\text{-}5]$$

2) Aplicación de un procedimiento en dos etapas. En la primera, se multiplican ambos miembros de la ecuación por la matriz no singular S, tal que se verifique que $S^{-1} S^{-1'} = W$. El modelo transformado es

$$Sy = SXb + Su \qquad [1\text{-}6]$$

En la segunda etapa se aplica mínimos cuadrados ordinarios (MCO) a [1-6]. El término de perturbación de [1-6] tiene matriz de covarianzas escalar y, por tanto, cumple las hipótesis del modelo lineal clásico.

En este caso, la matriz S es la siguiente:

$$S = \begin{bmatrix} \sqrt{1-\rho^2} & 0 \; 0 \; \ldots\ldots & 0 \; 0 \\ -\rho & 1 \; 0 \; \ldots\ldots & 0 \; 0 \\ 0 & -\rho \; 1 \; \ldots\ldots & 0 \; 0 \\ \ldots\ldots & \ldots\ldots\ldots\ldots & \ldots\ldots \\ 0 & 0 \; 0 \; \ldots\ldots & -\rho \; 1 \end{bmatrix} \qquad [1\text{-}7]$$

Aplicando la matriz [1-7] al modelo [1-3] éste quedaría transformado así, utilizando notación algebraica (3):

$$\left.\begin{array}{ll} t = 1 & Y_1 \sqrt{1-\rho^2} = b_0 \sqrt{1-\rho^2} + b_1 \sqrt{1-\rho^2}\, X_1 + \sqrt{1-\rho^2}\, u_1 \\ t = 2 \ldots T & Y_t - \rho\, Y_{t-1} = b_0 (1-\rho) + b_1 (X_t - \rho\, X_{t-1}) + \epsilon_t \end{array}\right\} \qquad [1\text{-}8]$$

En la práctica, muchas veces se prescinde de la transformación de la primera observación. Bajo ciertas condiciones, como demuestran Wonnacott y Wonnacott (1970), el prescindir de la transformación de la primera observación puede afectar a la eficiencia de los estimadores. Así, los citados autores muestran un ejemplo en el que los MCO arrojan unos estimadores más eficientes que los que resultan de la transformación indicada, pero sin la corrección de la primera observación.

Pero el problema más grave con que se enfrenta el económetra es que generalmente no conoce ρ, por lo que debe estimarlo conjuntamente con los restantes parámetros del modelo. A continuación examinamos distintos procedimientos de estimación, bajo el supuesto de que ρ tiene un valor desconocido.

---

(2) La exposición que realizamos coincide en gran parte con la de J. JOHNSTON (1972).

(3) El primer trabajo publicado en que se propone la transformación de la primera observación fue el de KADIYALA (1968). Sin embargo, parece que PRAISS y WINSTEN habían ya planteado en 1954 este problema en un documento interno de la Cowles Foundation, según manifiestan GRILICHES y RAO (1969).

a) Método simple

Consiste en los siguientes pasos:

1.º Estimación por MCO de [1-3], y a partir de los residuos que resultan estimar ρ por alguno de estos tres procedimientos alternativos:

— Coeficiente de autocorrelación empírico de los residuos.
— Estimación de la relación $\hat{u}_t = \rho\, \hat{u}_{t-1} + \epsilon_t$.
— Mediante la fórmula $\hat{\rho} = (2 - d)/2$, donde d es el estadístico de Durbin y Watson.

2.º Sustituir el valor obtenido de $\hat{\rho}$ en [1-8], y aplicar MCO.

b) Estimación no lineal

Consiste en la aplicación del criterio de mínimos cuadrados por procedimientos iterativos a la relación siguiente, puesto que las fórmulas analíticas de MCO no son aplicables por no ser el modelo lineal en los parámetros:

$$Y_t = b_0 (1 - \rho) + \rho\, Y_{t-1} + b_1 X_t + \rho\, b_1 X_{t-1} \qquad [1\text{-}9]$$

con la restricción de que $\widehat{b_1 \rho} = \hat{b}_1\, \hat{\rho}$.

c) Método iterativo

En este procedimiento, propuesto por Cochrane y Orcutt (1949), se toma como base la segunda relación de [1-8]. Se asigna un valor inicial arbitrario (4) a ρ, digamos $\hat{\rho}_1$, y se aplica MCO, ya que la ecuación es entonces lineal en los parámetros. Después, las estimaciones $\hat{b}_0$ y $\hat{b}_1$ obtenidas, se sustituyen en [1-8] y se minimiza la suma de los cuadrados con respecto a ρ, obteniéndose un nuevo valor $\hat{\rho}_2$, que a su vez se sustituiría en la ecuación, volviéndose a minimizar con respecto a b_0 y b_1, y así sucesivamente, hasta conseguir que la diferencia entre dos iteraciones sucesivas del valor de ρ fuera menor que un número prefijado. En este procedimiento iterativo se plantean dos problemas: la posibilidad de convergencia y la existencia de soluciones múltiples. Sargan (1964) demuestra que siempre se converge a un valor estacionario. Por otra parte, en un gran número de estudios que ha realizado este autor con sus alumnos no ha encontrado ningún caso de mínimos múltiples. Sobre este último punto señalaremos que mi experiencia particular al estimar el modelo PREFICO, patrocinado por el Instituto de Estudios Fiscales, no confirma el resultado obtenido por Sargan. El método

(4) Para conseguir lo antes posible la convergencia es conveniente utilizar un valor «razonable» para ρ. A estos efectos se puede utilizar cualquiera de los tres que se obtienen en el método simple.

que hemos aplicado ha sido el de «tanteo», que se examinará después. De todas formas, los resultados intermedios del método de «tanteo» son adecuados para analizar el problema de los mínimos locales, y hemos de decir que ha sido relativamente frecuente encontrarnos con ecuaciones que tuvieran varios mínimos locales.

d) MÉTODO DE DURBIN EN DOS ETAPAS

Durbin (1960) formula así la segunda ecuación de [1-8]:

$$Y_t = b_0(1-\rho) + \rho\, Y_{t-1} + b_1 X_t - b_1\, \rho\, X_{t-1} + \epsilon_t \qquad [1\text{-}10]$$

En la primera etapa se estima [1-10] por MCO, lo que permite obtener una estimación consistente de ρ (el coeficiente de Y_{t-1}). En la segunda etapa con la estimación de ρ se transforman las variables del modelo original. Este método ha sido aplicado con resultados satisfactorios por Lagares (1973), en su estimación de la función de inversión del sector industrial de nuestro país.

e) MÉTODO DE LA MÁXIMA VEROSIMILITUD

La función de verosimilitud de las perturbaciones ϵ_t es la siguiente:

$$P[\epsilon_1, \ldots, \epsilon_t] = \left[\frac{1}{2\pi\,\sigma_\epsilon^2}\right]^{-\frac{T}{2}} \cdot e^{-\frac{1}{2\sigma_\epsilon^2}\epsilon'\epsilon} \qquad [1\text{-}11]$$

El determinante del jacobiano de la transformación para pasar a la función de verosimilitud de las u_t (o de las Y_t) es

$$\left|\frac{\partial\,(\epsilon_1, \ldots, \epsilon_T)}{\partial\,(u_1, \ldots, u_T)}\right| = |\, S \,| = (1-\rho^2)^{1/2} \qquad [1\text{-}12]$$

(En [1-12] se ha tenido en cuenta la transformación de la primera perturbación: $u_1 = \epsilon_1/\sqrt{1-\rho^2}$).

Finalmente, la función de verosimilitud de las Y_t sería la siguiente:

$$P(Y_1, \ldots, Y_T/\sigma_\epsilon^2, b, \rho) = \left[\frac{1}{2\pi\,\sigma_\epsilon^2}\right]^{-\frac{S}{2}} (1-\rho^2)^{\frac{1}{2}} \cdot e^{-\frac{1}{2\sigma_\epsilon^2}(y-Xb)'W^{-1}(y-Xb)\,(1-\rho^2)} \qquad [1\text{-}13]$$

Si se conociera ρ, maximizar [1-13] equivaldría a minimizar su exponente, es decir, a aplicar MCG. Sin embargo, si se estima en el mismo proceso ρ obtendríamos en general resultados distintos debido a la presencia del factor $(1-\rho^2)^{1/2}$, según puede verse en Goldfeld y Quandt (1972). Para maximizar [1-13] es preciso utilizar, como en el método *b)*, procedimientos de cálculo numérico.

f) MÉTODO DE «TANTEO»

Este método fue propuesto inicialmente por Hildreth y Lu (1960) y popularizado por Dhrymes (1966). El principio en que se basa es muy simple:

Si en la segunda ecuación de [1-8] sustituimos ρ por un estimador $\hat{\rho}$, se puede expresar así:

$$Y_t - \hat{\rho} Y_{t-1} = b_0 (1 - \hat{\rho}) + b_1 (X_t - \hat{\rho} X_{t-1}) + \eta_t \qquad [1\text{-}14]$$

donde

$$\eta_t = \epsilon_t + (\rho - \hat{\rho})\, u_{t-1}$$

El método de «tanteo» (o *scanning*) consiste en ir dando distintos valores a ρ en el intervalo $(-1,1)$ —o en otro que se fije de antemano—, y estimar por MCO b_0 y b_1, en [1-14], para cada uno de ellos. Se retendría como estimación definitiva aquella para la cual la suma de los cuadrados de los residuos fuera mínima. Se trata, pues, de elegir el ρ que minimice la expresión

$$\Sigma\, \eta_t^2 = \Sigma\, [\epsilon_t + (\rho - \hat{\rho})\, u_{t-1}]^2 \qquad [1\text{-}15]$$

Dicho valor será

$$\hat{\rho} = \rho + \frac{\Sigma\, \epsilon_t\, u_{t-1}}{\Sigma\, u_{t-1}^2} \qquad [1\text{-}16]$$

Es fácil probar que éste es un estimador consistente. Además es eficiente, puesto que en definitiva lo que se hace es aplicar MCG —si se corrige también la primera observación— después de determinar ρ mediante una aproximación mínimo cuadrática. Frente al método de Cochrane-Orcutt señalaremos que no se corre el peligro de converger a un mínimo local.

Todos los métodos examinados son consistentes y asintóticamente más eficientes que el método de MCO. Se presenta el problema, sin embargo, de conocer su comportamiento en muestras pequeñas. A tal fin va dirigido el estudio de Montecarlo realizado por Griliches y Rao (1969).

Estos autores comparan la eficiencia para muestras de veinte observaciones de los siguientes métodos:

1) MCO.

2) Método simple con corrección de la primera observación (a este método lo denominan Praiss-Winsten, debido a que fueron los que pusieron de manifiesto la conveniencia de corregir la primera observación).

3) Método de Cochrane-Orcutt sin corregir la primera observación.

4) Método de Durbin en dos etapas, con y sin la corrección de la primera obsevación.

5) Método no lineal.

Las conclusiones a que llegan Griliches y Rao son las siguientes: *a)* MCO es menos eficiente que los otros métodos cuando $|\rho| > 0{,}3$; *b)* el método de Durbin, con la corrección de la primera observación, es el que arroja los mejores resultados; *c)* el método no lineal no es más eficiente que los métodos 2), 3) y 4), y presenta el inconveniente de que los cálculos son mucho más complicados.

En nuestro trabajo pretendemos completar el estudio de Griliches y Rao en los siguientes puntos:

1.° El método simple lo aplicaremos con y sin la corrección de la primera observación. Para estimar ρ utilizaremos el coeficiente de autocorrelación empírico de los residuos.

2.° En lugar del procedimiento Cochrane-Orcutt, aplicaremos el procedimiento del *scanning*, con y sin la corrección de la primera observación. De esta forma se evita el problema de los mínimos locales. Griliches y Rao (1969, p. 258) afirman, sin embargo, que el método del «tanteo» no se adapta bien a los experimentos de Montecarlo. Esto es cierto en lo que se refiere a la estimación de ρ, pero en cambio no se presenta ninguna dificultad si se centra la atención exclusivamente, como es nuestro caso, en la estimación del coeficiente de la variable exógena.

El «tanteo» se hace en dos etapas. En la primera se hace el recorrido para saltos de longitud 0,198, en el intervalo $(-0{,}99;\ 0{,}99)$. En la segunda etapa, si ha resultado seleccionado ρ^*, la búsqueda se lleva a cabo en el intervalo $(\rho^* - 0{,}198;\ \rho^* + 0{,}198)$ para saltos de longitud 0,0198, aunque en ningún caso se supera el intervalo $(-0{,}99;\ 0{,}99)$. La aplicación de este método implica realizar treinta regresiones para cada estimación. Conviene señalar que este método se ha aplicado con y sin la corrección de la primera observación, separando completamente ambos casos, ya que pueden resultar —y de hecho resulta— una selección de ρ^* en la primera etapa distinto.

3.° Griliches y Rao utilizan solamente muestras de veinte observaciones. Como uno de nuestros propósitos es el de estudiar la influencia de la corrección de la primera observación para distinto número de observaciones, vamos a utilizar muestras de tamaño diez y veinte.

El método de Durbin en dos etapas lo aplicaremos con y sin la corrección de la primera observación, como en el trabajo de Griliches y Rao, pero en cambio no utilizaremos el método no lineal.

En resumen, los métodos que vamos a aplicar son los siguientes:

1) MCO (I).

2) Método simple, sin la corrección (II) y con la corrección de la primera observación (V).

3) Método de Durbin en dos etapas, sin la corrección (III) y con la corrección de la primera observación (VI).

4) Método de «tanteo» (o «Hilu»), sin la corrección (IV) y con la corrección de la primera observación (VII).

Con objeto de poder efectuar comparaciones se ha aplicado también el método de MCG (VIII), en que se utiliza el verdadero valor de ρ.

En el epígrafe siguiente pueden verse los detalles sobre el modelo que ha servido de base para la simulación.

2. DISEÑO DEL EXPERIMENTO DE SIMULACION

El modelo que hemos tomado como base para la simulación es el siguiente:

$$y_t = 1x_t + u_t \quad [2\text{-}1]$$

$$x_t = \lambda x_{t-1} + w_t \quad [2\text{-}2]$$

$$u_t = \rho u_{t-1} + \epsilon_t \quad [2\text{-}3]$$

$$\epsilon_t \longrightarrow N(0, \sigma^2_\epsilon) \quad [2\text{-}4]$$

$$w_t \longrightarrow N(0, \sigma^2_w) \quad [2\text{-}5]$$

$$E(\epsilon_t w_t) = 0 \quad [2\text{-}6]$$

Como puede observarse, coincide en su planteamiento con el utilizado por Griliches y Rao. Sin embargo, hemos adoptado distintos supuestos en cuanto al valor de λ, al carácter de X y al tamaño de la muestra.

En los experimentos diseñados por los citados autores se ha tomado siempre $|\lambda| < 1$. Este supuesto nos parece poco realista para reflejar la evolución de las series económicas —al menos las españolas— en los períodos recientes. Por ello nos ha parecido más razonable tomar un valor de $\lambda = 1{,}04$ que constituye una cierta aproximación de lo que ha sido el crecimiento medio de las series económicas, en términos reales, de la economía española en los últimos veinte años. Para el objetivo que perseguimos no hemos creído conveniente experimentar con diversos valores de λ.

Hemos tomado 1 como valor inicial de X (5) y asignado 0,012 a la desviación típica de w. Esto quiere decir que la importancia de w en la generación de los sucesivos valores de X va atenuándose a lo largo del tiempo. Por el valor que se ha asignado a la desviación típica de w, la probabilidad de que $X_{t+1} < X_t$ para cualquier t es muy baja, y a medida que crece t, esta probabilidad va disminuyendo. La serie X tendrá, pues, un crecimiento acumulativo a una tasa del 4 por 100 con la posibilidad de que para un período concreto existan ciertas variaciones en torno a este promedio.

Se ha calculado la varianza de u de forma que el coeficiente de determinación en [2-1] sea de 0,8. El valor inicial de u se ha generado a partir de la distribución resultante.

(5) Por X indicamos que la variable está expresada en valores originales, y por x que lo está en desviaciones respecto a la media.

Hemos supuesto que u está generada por un proceso autorregresivo estacionario, es decir, que $-1 < \rho < 1$, y hemos realizado experimentos para valores de ρ desde $-0{,}9$ a $0{,}9$, con pasos de $0{,}2$. También hemos considerado el caso de $\rho = 0$.

Para cada valor de ρ y la correspondiente varianza de σ^2_u, se ha determinado la varianza de ϵ teniendo en cuenta que

$$\sigma^2_\epsilon = (1 - \rho^2)\, \sigma^2_u \qquad [2\text{-}7]$$

Para cada uno de los dos tamaños de muestra (10 y 20) y para cada valor de ρ, se ha extraído 1 muestra de X, y 50 del término de perturbación. Es decir, en cada experimento hemos mantenido fija la muestra de X. Griliches y Rao no han encontrado diferencias sensibles en los resultados, derivados de mantener o no fija la muestra de X.

Los números aleatorios N (0, 1), a los que denominaremos Z_t, los hemos generado a partir de la siguiente expresión:

$$Z_t = \sum_{i=1}^{12} \zeta_i - 6 \qquad [2\text{-}8]$$

donde ζ_i es un número aleatorio uniforme con variación en el intervalo (0,1) y que son generados directamente por el miniordenador Hewlett-Packard 9830, que hemos utilizado en los cálculos.

Para cada experimento se han calculado los siguientes estadísticos:

— Sesgo: $\dfrac{1}{50} \sum_{i=1}^{50} \hat{b}_i - 1 = \bar{\hat{b}} - 1$

— Varianza: $\dfrac{1}{50} \sum_{i=1}^{50} [\hat{b} - \bar{\hat{b}}]^2$

— Error cuadrático medio (ECM): $\dfrac{1}{50} \sum_{i=1}^{50} [\hat{b}_i - 1]^2$

Como se sabe, los tres estadísticos están relacionados de la siguiente forma:

$$\text{ECM} = [\text{Sesgo}]^2 + \text{Varianza}$$

Los programas para la aplicación de estos experimentos han sido elaborados por el autor en lenguaje BASIC y ejecutados en el miniordenador ya citado. Dicha ejecución ha requerido noventa y cuatro horas de funcionamiento ininterrumpido. Téngase en cuenta que ha supuesto el cálculo de 72.600 regresiones, lo que permite estimar en unos cuatro segundos el tiempo que ha empleado en realizar cada regresión.

3. ANALISIS DE RESULTADOS

Los resultados de nuestro experimento de simulación se han recogido en los cuadros 1.1-1.11, 2,1, 2.2 y 3.

En los cuadros de la serie 1, uno para cada valor de ρ, aparecen reflejados los estadísticos sesgo, varianza y error cuadrático medio (ECM) obtenidos para cada procedimiento de estimación y cada tamaño de muestra. Todos los resultados están multiplicados por 100. En los cuadros 2.1 —para muestras de tamaño 10— y 2.2 —para muestras de tamaño 20—, se ha calculado, tomando como base el ECM, la eficiencia relativa de los métodos I-VII, en comparación con el método de MCG (VIII). Finalmente, en el cuadro 3, tomando también como base el ECM, se efectúa la comparación de eficiencia dentro de cada procedimiento de estimación, según se haga o no la corrección de la primera observación.

CUADRO 1.1

SESGO, VARIANZA, ERROR CUADRATICO MEDIO (ECM)

$\rho = -0,9$

METODOS DE ESTIMACION	TAMAÑO DE LA MUESTRA = 10			TAMAÑO DE LA MUESTRA = 20		
	Sesgo	Varianza	ECM	Sesgo	Varianza	ECM
MCO (I)	1,2444	0,6391	0,6546	0,1830	0,1647	0,1650
Sin corrección primera observación:						
Simple (II)	0,2598	0,1889	0,1896	— 0,2686	0,0907	0,0914
Durbin (III)	0,2442	0,1940	0,1946	— 0,2558	0,0905	0,0912
Hilu (IV)	0,2600	0,1945	0,1952	— 0,2776	0,0912	0,0920
Con corrección primera observación:						
Simple (V)	0,2694	0,1895	0,1902	— 0,1866	0,0923	0,0926
Durbin (VI)	0,3220	0,1834	0,1844	— 0,1980	0,0904	0,0908
Hilu (VII)	0,2426	0,1870	0,1876	— 0,2238	0,0917	0,0922
MCG (VIII)	0,2554	0,1859	0,1866	— 0,2340	0,0929	0,0934

CUADRO 1.2

SESGO, VARIANZA, ERROR CUADRATICO MEDIO (ECM)

$\rho = -0,7$

METODOS DE ESTIMACION	TAMAÑO DE LA MUESTRA = 10			TAMAÑO DE LA MUESTRA = 20		
	Sesgo	Varianza	ECM	Sesgo	Varianza	ECM
MCO (I)	— 1,1256	0,7051	0,7178	0,6548	0,2963	0,3006
Sin corrección primera observación:						
Simple (II)	— 2,0414	0,5427	0,5844	0,2158	0,2347	0,2352
Durbin (III)	— 1,9990	0,5604	0,6004	0,1818	0,2341	0,2344
Hilu (IV)	— 2,0414	0,5615	0,6032	0,2094	0,2334	0,2338
Con corrección primera observación:						
Simple (V)	— 1,7340	0,5279	0,5580	0,1566	0,2330	0,2332
Durbin (VI)	— 1,6996	0,5583	0,5872	0,1366	0,2306	0,2308
Hilu (VII)	— 1,7088	0,5442	0,5734	0,1370	0,2296	0,2298
MCG (VIII)	— 1,5962	0,5465	0,5720	0,1894	0,2314	0,2318

CUADRO 1.3

SESGO, VARIANZA, ERROR CUADRATICO MEDIO (ECM)

$\rho = -0,5$

METODOS DE ESTIMACION	TAMAÑO DE LA MUESTRA = 10			TAMAÑO DE LA MUESTRA = 20		
	Sesgo	Varianza	ECM	Sesgo	Varianza	ECM
MCO (I)	0,8398	0,7141	0,7212	1,9484	0,4298	0,4678
Sin corrección primera observación:						
Simple (II)	— 0,9398	0,9166	0,9254	1,1968	0,4429	0,4572
Durbin (III)	— 0,9988	0,9298	0,9398	1,1798	0,4441	0,4580
Hilu (IV)	— 1,0640	0,9419	0,9532	1,1920	0,4464	0,4606
Con corrección primera observación:						
Simple (V)	— 0,3558	0,6311	0,6324	1,4932	0,4162	0,4370
Durbin (VI)	— 0,4344	0,6521	0,6540	1,4576	0,4147	0,4374
Hilu (VII)	— 0,6566	0,6793	0,6836	1,4440	0,4211	0,4420
MCG (VIII)	— 0,1536	0,6906	0,6908	1,4772	0,4060	0,4278

CUADRO 1.4

SESGO, VARIANZA, ERROR CUADRATICO MEDIO (ECM)

$\rho = -0,3$

METODOS DE ESTIMACION	TAMAÑO DE LA MUESTRA = 10			TAMAÑO DE LA MUESTRA = 20		
	Sesgo	Varianza	ECM	Sesgo	Varianza	ECM
MCO (I)	— 0,3944	1,3978	1,3994	— 0,3784	0,5854	0,5868
Sin corrección primera observación:						
Simple (II)	— 0,2884	1,9252	1,9260	— 0,9312	0,6441	0,6528
Durbin (III)	— 0,3100	2,1444	2,1454	— 0,9722	0,6461	0,6556
Hilu (IV)	— 0,1776	2,1163	2,1166	— 0,9448	0,6465	0,6554
Con corrección primera observación:						
Simple (V)	— 0,1344	1,4362	1,4364	— 0,4714	0,5886	0,5908
Durbin (VI)	0,1796	1,4309	1,4312	— 0,4990	0,5977	0,6002
Hilu (VII)	0,0120	1,5172	1,5162	— 0,4668	0,5918	0,5940
MCG (VIII)	— 0,5710	1,3343	1,3376	— 0,6034	0,5944	0,5980

CUADRO 1.5

SESGO, VARIANZA, ERROR CUADRATICO MEDIO (ECM)

$\rho = -0,1$

METODOS DE ESTIMACION	TAMAÑO DE LA MUESTRA = 10			TAMAÑO DE LA MUESTRA = 20		
	Sesgo	Varianza	ECM	Sesgo	Varianza	ECM
MCO (I)	— 0,6894	3,1062	3,1110	— 0,4028	1,0124	1,0140
Sin corrección primera observación:						
Simple (II)	— 2,2956	3,9913	4,0440	— 0,6904	1,3478	1,3526
Durbin (III)	— 2,1142	3,7669	3,8116	— 0,7148	1,3483	1,3534
Hilu (IV)	— 2,4000	3,9360	3,9936	— 0,7016	1,3517	1,3566
Con corrección primera observación:						
Simple (V)	— 0,5454	3,0796	3,0826	— 0,5158	1,0843	1,0870
Durbin (VI)	— 0,7342	3,0160	3,0214	— 0,5546	1,1027	1,1056
Hilu (VII)	— 0,5052	3,0452	3,0478	— 0,3586	1,0761	1,0774
MCG (VIII)	— 0,7158	3,0917	3,0968	— 0,2854	1,0136	1,0138

CUADRO 1.6

SESGO, VARIANZA, ERROR CUADRATICO MEDIO (ECM)

$\rho = 0{,}1$

METODOS DE ESTIMACION	TAMAÑO DE LA MUESTRA = 10			TAMAÑO DE LA MUESTRA = 20		
	Sesgo	Varianza	ECM	Sesgo	Varianza	ECM
MCO (I)	1,3964	3,4301	3,4496	3,4296	1,4022	1,5198
Sin corrección primera observación:						
Simple (II)	3,0292	4,6998	4,7916	2,5588	1,5857	1,6512
Durbin (III)	3,4088	4,7482	4,8644	2,6234	1,5602	1,6290
Hilu (IV)	3,7840	5,3010	5,4442	2,5110	1,5885	1,6516
Con corrección primera observación:						
Simple (V)	1,7406	3,5861	3,6164	3,3538	1,4231	1,5356
Durbin (VI)	1,8888	3,6545	3,6902	3,2872	1,4015	1,5096
Hilu (VII)	1,9810	3,6998	3,7390	3,3844	1,4251	1,5396
MCG (VIII)	1,4580	3,5095	3,5308	3,4078	1,4077	1,5238

CUADRO 1.7

SESGO, VARIANZA, ERROR CUADRATICO MEDIO (ECM)

$\rho = 0{,}3$

METODOS DE ESTIMACION	TAMAÑO DE LA MUESTRA = 10			TAMAÑO DE LA MUESTRA = 20		
	Sesgo	Varianza	ECM	Sesgo	Varianza	ECM
MCO (I)	0,6976	4,7565	4,7614	— 1,7624	2,3075	2,3386
Sin corrección primera observación:						
Simple (II)	3,6024	6,1702	6,3000	— 2,1066	2,7286	2,7730
Durbin (III)	2,8308	5,8233	5,9034	— 1,8514	2,6247	2,6590
Hilu (IV)	4,0882	6,1049	6,2720	— 2,3178	2,7993	2,8530
Con corrección primera observación:						
Simple (V)	1,8668	4,2342	4,2690	— 2,3228	2,2412	2,2952
Durbin (VI)	1,6736	4,3656	4,3936	— 2,2112	2,2793	2,3282
Hilu (VII)	2,1360	4,0856	4,1312	— 2,4558	2,2779	2,3382
MCG (VIII)	1,7262	4,4818	4,5116	— 2,3090	2,3179	2,3712

CUADRO 1.8

SESGO, VARIANZA, ERROR CUADRATICO MEDIO (ECM)

$\rho = 0,5$

METODOS DE ESTIMACION	TAMAÑO DE LA MUESTRA = 10			TAMAÑO DE LA MUESTRA = 20		
	Sesgo	Varianza	ECM	Sesgo	Varianza	ECM
MCO (I)	6,4662	3,9403	4,3584	— 0,4650	3,3942	3,3964
Sin corrección primera observación:						
Simple (II)	4,4890	4,4459	4,6474	0,6446	3,8410	3,8452
Durbin (III)	5,2244	4,8265	5,0994	0,6256	3,7135	3,7174
Hilu (IV)	4,3056	4,5930	4,7784	0,9386	3,8572	3,8660
Con corrección primera observación:						
Simple (V)	6,2714	3,3323	3,7256	0,2176	3,6165	3,6170
Durbin (VI)	6,6342	3,6201	4,0602	0,1448	3,5300	3,5302
Hilu (VII)	6,2820	3,2778	3,6724	0,4746	3,6805	3,6828
MCG (VIII)	5,7790	3,4604	3,7944	0,2930	3,5701	3,5710

CUADRO 1.9

SESGO, VARIANZA, ERROR CUADRATICO MEDIO (ECM)

$\rho = 0,7$

METODOS DE ESTIMACION	TAMAÑO DE LA MUESTRA = 10			TAMAÑO DE LA MUESTRA = 20		
	Sesgo	Varianza	ECM	Sesgo	Varianza	ECM
MCO (I)	— 1,6676	3,4720	3,4998	— 0,9818	4,5802	4,5898
Sin corrección primera observación:						
Simple (II)	— 2,1362	5,2390	5,2846	— 1,4292	5,3890	5,4094
Durbin (III)	— 0,4800	6,8075	6,8098	— 2,2532	5,1266	5,1774
Hilu (IV)	— 1,6664	7,6370	7,6648	— 1,0924	5,4375	5,4494
Con corrección primera observación:						
Simple (V)	— 2,9520	2,6877	2,7748	— 0,9276	3,6286	3,6372
Durbin (VI)	— 2,2774	3,1809	3,2328	— 1,1986	3,6952	3,7096
Hilu (VII)	— 2,4872	2,6845	2,7464	— 0,7910	3,6369	3,6432
MCG (VIII)	— 1,4894	2,6468	2,6690	— 0,0506	3,5158	3,5158

CUADRO 1.10

SESGO, VARIANZA, ERROR CUADRATICO MEDIO (ECM)

$\rho = 0,9$

METODOS DE ESTIMACION	TAMAÑO DE LA MUESTRA = 10			TAMAÑO DE LA MUESTRA = 20		
	Sesgo	Varianza	ECM	Sesgo	Varianza	ECM
MCO (I)	— 0,5108	2,8912	2,8938	— 3,1400	5,5932	5,6918
Sin corrección primera observación:						
Simple (II)	0,4554	4,1339	4,1360	— 3,3332	6,1555	6,2666
Durbin (III)	1,7520	5,2481	5,2788	— 4,4602	6,4181	6,6170
Hilu (IV)	0,2100	6,2020	6,2024	— 3,2018	6,8097	6,9122
Con corrección primera observación:						
Simple (V)	— 0,5062	2,6768	2,6794	— 2,0058	4,2230	4,2632
Durbin (VI)	— 0,3028	2,7373	2,7382	— 2,7104	4,5223	4,5958
Hilu (VII)	— 0,4754	2,7037	2,7060	— 2,0240	4,0544	4,0954
MCG (VIII)	0,0168	2,7598	2,7598	— 2,6920	3,6057	3,6782

CUADRO 1.11

SESGO, VARIANZA, ERROR CUADRATICO MEDIO (ECM)

$\rho = 0$

METODOS DE ESTIMACION	TAMAÑO DE LA MUESTRA = 10			TAMAÑO DE LA MUESTRA = 20		
	Sesgo	Varianza	ECM	Sesgo	Varianza	ECM
MCO (I)	— 0,4032	3,0660	3,0676	1,6288	1,5763	1,6028
Sin corrección primera observación:						
Simple (II)	0,2388	5,0014	5,0020	2,0394	1,8016	1,8432
Durbin (III)	0,1122	4,6075	4,6076	2,0886	1,7916	1,8352
Hilu (IV)	0,2428	5,1420	5,1426	2,0710	1,8043	1,8472
Con corrección primera observación:						
Simple (V)	— 0,0586	3,1202	3,1202	1,9034	1,6284	1,6646
Durbin (VI)	0,1586	3,1471	3,1474	1,8538	1,6344	1,6688
Hilu (VII)	— 0,0636	3,1346	3,1346	1,9178	1,6346	1,6714
MCG (VIII)	— 0,4032	3,0660	3,0676	1,6288	1,5763	1,6028

CUADRO 2.1

EFICIENCIA RELATIVA DE LOS METODOS I-VII EN COMPARACION AL METODO DE MCG (VIII): ECM (VIII) / ECM (α)

α = I, II, ..., VII

(Tamaño de la muestra: 10 observaciones)

Método / Valor de ρ	I	II	III	IV	V	VI	VII
— 0,9	0,29	0,98	0,96	0,96	0,98	1,01	0,99
— 0,7	0,80	0,98	0,95	0,95	1,03	0,97	1,00
— 0,5	0,96	0,75	0,74	0,72	1,09	1,06	1,01
— 0,3	0,96	0,69	0,62	0,63	0,93	0,93	0,88
— 0,1	1,00	0,77	0,81	0,78	1,00	1,02	1,02
0,1	1,02	0,74	0,73	0,65	0,98	0,96	0,94
0,3	0,95	0,72	0,76	0,72	1,06	1,03	1,09
0,5	0,87	0,82	0,74	0,79	1,02	0,93	1,03
0,7	0,76	0,51	0,39	0,35	0,96	0,83	0,97
0,9	0,95	0,67	0,52	0,44	1,03	1,01	1,02
0	1,00	0,61	0,67	0,60	0,98	0,97	0,98

CUADRO 2.2

EFICIENCIA RELATIVA DE LOS METODOS I-VII EN COMPARACION AL METODO DE MCG (VIII): ECM (VIII) / ECM (α)

α = I, II, ..., VII

(Tamaño de la muestra: 20 observaciones)

Método / Valor de ρ	I	II	III	IV	V	VI	VII
— 0,9	0,57	1,02	1,02	1,02	1,01	1,03	1,01
— 0,7	0,77	0,99	0,99	0,99	0,99	1,00	1,01
— 0,5	0,91	0,94	0,93	0,93	0,98	0,98	0,97
— 0,3	1,02	0,92	0,91	0,91	1,01	1,00	1,01
— 0,1	1,00	0,75	0,75	0,75	0,93	0,92	0,94
0,1	1,00	0,92	0,94	0,92	0,99	1,01	0,99
0,3	1,01	0,86	0,89	0,83	1,03	1,02	1,01
0,5	1,05	0,93	0,96	0,92	0,99	1,01	0,97
0,7	0,77	0,65	0,68	0,65	0,97	0,95	0,97
0,9	0,65	0,59	0,56	0,53	0,86	0,80	0,90
0	1,00	0,87	0,87	0,87	0,96	0,96	0,96

CUADRO 3

EFICIENCIA RELATIVA DE LOS METODOS II, III Y IV, EN COMPARACION A LOS METODOS V, VI Y VII

Valor de ρ	ECM (V) / ECM (II) — Tamaño de la muestra 10	ECM (V) / ECM (II) — Tamaño de la muestra 20	ECM (VI) / ECM (III) — Tamaño de la muestra 10	ECM (VI) / ECM (III) — Tamaño de la muestra 20	ECM (VII) / ECM (IV) — Tamaño de la muestra 10	ECM (VII) / ECM (IV) — Tamaño de la muestra 20
— 0,9	1,00	1,01	0,95	1,00	0,96	1,00
— 0,7	0,95	0,99	0,98	0,98	0,95	0,98
— 0,5	0,68	0,96	0,70	0,96	0,72	0,96
— 0,3	0,75	0,91	0,67	0,92	0,72	0,91
— 0,1	0,76	0,80	0,79	0,82	0,76	0,79
0,1	0,75	0,93	0,76	0,93	0,69	0,93
0,3	0,68	0,83	0,74	0,88	0,66	0,82
0,5	0,80	0,94	0,80	0,95	0,77	0,95
0,7	0,53	0,67	0,47	0,72	0,36	0,67
0,9	0,65	0,68	0,52	0,69	0,44	0,59
0	0,62	0,90	0,68	0,91	0,61	0,90

Las conclusiones que se pueden deducir de nuestros experimentos de simulación, no olvidando el alcance limitado de este tipo de investigaciones, son las siguientes:

1.ª De los dos componentes de ECM, en todos los casos tiene mucha mayor importancia la varianza.

2.ª Para valores de ρ comprendidos entre —0,5 y 0,5, la eficiencia de los MCO no es inferior a la de los otros métodos de estimación.

3.ª Se incrementa de forma apreciable la eficiencia de las estimaciones cuando se procede a corregir la primera observación. Este incremento de eficiencia es mucho más patente en caso de muestra de tamaño 10. En cambio, los tres métodos en dos etapas que se han utilizado (simple, Durbin, Hilu), arrojan unos resultados similares. Por tanto, para muestras pequeñas parece mucho más discriminante —desde el punto de la eficiencia de las estimaciones—, la corrección o no de la primera observación que la elección de uno u otro procedimiento de estimación. En todo caso, cuando el coeficiente es pequeño en valor absoluto, parece lo más aconsejable utilizar MCO.

BIBLIOGRAFIA

COCHRANE, D., y ORCUTT, G. H. (1949): «Application of least squares to relations hips containing autocorrelated error terms», *Journal of the American Statistical Association*, vol. 44, pp. 32-61.

DHRYMES, P. J. (1966): «On the treatment of certain recurrent non-linearities in regression analysis», *Southern Economic Journal*, vol. 33, pp. 187-196.

DURBIN, J. (1960): «Estimation of parameters in time-series regression models», *Journal of the Royal Statistical Society*, Series B, vol. 22, pp. 139-153.

GOLDBERGER, A. S. (1964): *Econometric theory*, John Wiley (Nueva York).

GOLDFELD, S. M., y QUANDT, R. E. (1972): *Non linear methods in econometrics* (North-Holland).

GRILICHES, Z., y RAO, P. (1969): «Small-sample properties of several two-stage regression methods in the context of autocorrelated errors», *Journal of American Statistical Association*, vol. 64, pp. 253-272.

HILDRETH, C., y LU, J. Y. (1960): «Demand relations with autocorrelated disturbances», *Technical Bulletin* 276, Michigan State University, Agricultural Experiment Station.

JOHNSTON, J. (1972): *Econometric Methods*, 2.ª edición, Mc Graw-Hill.

KADIYALA, K. R. (1968): «A transformation used to circunvent the problem of autocorrelation», *Econometrica*, vol. 36, pp. 93-96.

SARGAN, J. D. (1964): «Wages and prices in the United Kingdom: A Study in Econometric Methodology», pp. 25-63, de la obra *Econometric Analysys for National Planning*, editada por P. E. Hart Butterwort (Londres).

WONNACOTT, R. J., y WONNACOTT, T. H. (1970): *Econometrics*, John Wiley.