TECNIA, Vol.11 N° 1, págs.9 - 17, 2001
Universidad Nacional de Ingeniería
Lima - Perú

EVIDENCIA EMPÍRICA DE LOS MÍNIMOS CUADRADOS EN EL MODELO DE REGRESIÓN DE COEFICIENTES ALEATORIOS

Luis Huamanchumo de la Cuba
Escuela Profesional de Ingeniería Estadística-Universidad Nacional de Ingeniería

RESUMEN

El presente artículo estudia algunas propiedades en muestras finitas de varios estimadores del Coeficiente de Respuesta Medio en un Modelo de Regresión Lineal de Coeficientes Aleatorios. Para este fin, fue necesario diseñar un experimento muestral. Así, se obtuvo evidencias acerca del Sesgo, Consistencia y Eficiencia a partir de 15,120 estimaciones. De acuerdo a esto, no sólo el estimador Mínimos Cuadrados Generalizados en Dos Etapas (MCG2E) produjo los mejores resultados sino también el estimador Mínimos Cuadrados Ordinarios (MCOII) al obtener ganancias significativas en Eficiencia cuando fue estimado a partir de un modelo sin Error de Especificación. Es necesario ampliar dicha investigación incluyendo el estimador Máximo Verosímil.
PALABRAS CLAVE: Coeficientes Aleatorios, Mínimos Cuadrados Generalizados, Sesgo, Consistencia y Eficiencia.

ABSTRACT

This paper studies some Finite Sample Properties of several estimators of the Mean Response Coefficient in a Linear Regression Model with Random Coefficients. An Experiment was designed to explore that. In this context, results about Bias, Consistency and Efficiency were obtained from 15,120 estimates. According to this, not only Two Steps Generalized Least Square (MCG2E) performed better than the other alternative estimators but also Ordinary Least Square had gotten gains in Efficiency when the Linear Regression Model without Constant Term was used. Further investigations about Maximun Likelihood Models are needed.
KEYWORDS: Random Coefficient Models, Generalized Least Square, Bias, Consistency, Efficiency.

INTRODUCCION

Los Modelos de Regresión de Coeficientes Aleatorios constituyen una herramienta poderosa en el análisis de la información contenida en datos de tipo panel. Los diversos factores que influyen en los mecanismos de la toma de decisiones, tales como: diversidad cultural, diferencias en la asignación de recursos, distribución geográfica, etc.; son incorporados en el modelo bajo el supuesto de aleatoriedad de los coeficientes. En este sentido, la combinación de los datos que conforman el panel, permitirá ganar grados de libertad al aumentar la muestra de corte transversal, manteniendo fija la información temporal correspondiente, puesto que sólo nos interesará estimar los parámetros de dicha distribución los cuales son independientes al tamaño de la muestra.

Los resultados empíricos producto del experimento muestral nos otorgará ciertas ventajas que no podríamos obtener con datos reales debido al carácter no-experimental de las variables socioeconómicas. La simulación de dichas variables permitirá aislar ciertos factores perturbadores, así como también, la replicación bajo las mismas condiciones experimentales, el bloqueo de resultados de acuerdo a las variables de interés y mantener las mismas condiciones iniciales de aleatoriedad las cuales no serán influenciadas por el efecto temporal que incorpora un panel de datos.

Los resultados permitirán disponer de una base empírica que contribuya a decidir sobre qué técnica de estimación es más conveniente a las necesidades de investigación.

Podemos esperar que las propiedades muestrales del estimador por Mínimos Cuadrados Generalizados en Dos Etapas y por Mínimos Cuadrados Ordinarios sean influenciados por los valores de los parámetros de la distribución de los coeficientes. Cuando la muestra de corte transversal es mucho mayor que la de la serie temporal, la información respecto a la variabilidad de la distribución de los coeficientes es mucho más relevante que la correspondiente a la estructura del término errático de la serie temporal.

Primero, se discutirá la formulación fundamental del modelo de coeficientes aleatorios así como los métodos de estimación alternativos. En segundo lugar, se desarrolla la metodología que incluye las definiciones operativas y el diseño del experimento. Finalmente, se analizan los resultados experimentales obtenidos de los cuales se desprenderán las conclusiones y recomendaciones para futuras investigaciones.

EL MODELO

La ecuación 1.1 muestra la relación lineal para el individuo 'i' de la muestra de corte transversal, la cual resume el impacto aleatorio y proporcional a 'β_i' sobre la variable endógena 'y_i' debido a una variación en la variable exógena X_i.

$$\underline{y}_i = x_i \underline{\beta}_i + u_i \quad , \quad i=1,...n \qquad (1.1)$$

donde, 'yi' es un vector Tx1 de las observaciones sobre la variable endógena, 'Xi' es una matriz TxK de observaciones de la variable exógena, 'β_i' es el vector de coeficientes aleatorios no observables y 'u_i' el vector Tx1 de perturbaciones.

Supuestos:

(a) Los tamaños de las muestras 'n' y 'T' son tales que n>k y T>k.

(b)Las variables independientes son no estocásticas en el sentido de que las Xi son fijas en repetidas muestras sobre Yi. El rango de Xi es K.

(c) El vector de perturbaciones Ui están independientemente distribuidos con vector de media cero y matriz varianza-covarianza de Ui igual a $\sigma_{ii} I_T$.

(d) Los vectores de coeficientes βi (i=1,...,n) están independientemente e idénticamente distribuidos como una Normal Multivariada con media E(βi)=β y varianza Var(βi)=Δ, el cual es no singular.

(e) Los vectores Ui y βj son independientes para todo i,j=1,...,n.

Expresando el coeficiente de respuesta como la suma de un componente fijo, el coeficiente de respuesta medio β, y un componente aleatorio δi, es decir, βi = β+δi en la ecuación (1.1) tenemos el modelo transformado:

$$\mathbf{yi = \beta Xi + \theta i} \quad i=1,.,n \qquad (1.2)$$

El término θi=δi Xi+ui $\sim_{iid} N_2(0,H_i(\theta))$ es el nuevo término de perturbación para el modelo transformado cuya matriz varianza-covarianza es:

$$\mathbf{H_i(\theta)=X_i \Delta X_i + \sigma_{ii} \Omega_{ii}} \ , \quad i=1,.,n \qquad (1.3)$$

El primer término resume la variabilidad de los coeficientes y, el segundo, la variabilidad del término errático de la serie temporal.

Métodos de Estimación

Se plantean dos modelos: el Modelo I que será estimado con el método de los Mínimos Cuadrados Ordinarios, y el Modelo II el cual será estimado con todas las técnicas consideradas. Así tenemos que:

Modelo I $\quad y_i = \beta_{0i} + \beta_1 X_{1i}$

Modelo II $\quad y_i = \beta_1 X_{1i} + \beta_{21} X_{2i}$

a) Mínimos Cuadrados Ordinarios con Término Constante (MCOI)

Este método utiliza el Modelo I para estimar el coeficiente de respuesta medio β. De acuerdo a esto, el estimador sería:

$$b = (X'X)^{-1}X'Y \qquad (1.4)$$

donde X= $(X_1, X_2, ..., X_n)$ es una matriz (nTx2) cuya primera columna son unos e Y es un vector (nTx1). La matriz varianza-covarianza para el estimador (1.4) está dado por $V(b)=\sigma^2 (X'X)^{-1}$

b) Mínimos Cuadrados Ordinarios sin Término Constante (MCOII)

Bajo la técnica MCOII estimaremos por Mínimos Cuadrados Ordinarios el Modelo II, sin término constante. La formulación para el estimador del coeficiente de respuesta medio es como (1.4) donde sólo cambia la matriz X debido a que en este caso no hay término constante.

c) Mínimos Cuadrados Generalizados (MCG)

Este método de estimación incorpora los supuestos de heteroscedasticidad y correlación de errores incorporados en la muestra de corte transversal y la serie temporal respectivamente. La formulación del estimador de respuesta medio sería:

$$b(\theta)=[\sum_j X_j'(X_j'\Delta X_j+\sigma_{jj}I_T)^{-1}X_j]^{-1}[\sum_i X_{i'}'(X_i'\Delta X_i+\sigma_{ii}I_T)^{-1}Y_i] \quad (1.5)$$

Aplicando la identidad de Rao [1], el estimador queda transformado

$$b(\theta)=\sum_{i=1}^{n} W_i(\theta)b_i \quad (1.6)$$

La ecuación (1.6) expresa al estimador del coeficiente de respuesta medio como la suma ponderada de los coeficientes estimados de los parámetros correspondientes a cada unidad de la muestra de corte transversal bi = $(X_i'X_i)^{-1}X_i'Y_i$. El factor de ponderación Wi es calculado por:

$$W_i(\theta)=[\sum_{j=1}^{n}\{\Delta+\sigma_{jj}(X_j'X_j)^{-1}\}^{-1}]\{\Delta+\sigma_{ii}(X_i'X_i)^{-1}\}^{-1} \quad (1.7)$$

d) Mínimos Cuadrados Generalizados en Dos Eta pas (MCG2E) [2]

Este método presenta la misma formulación básica del estimador tal como se presenta en la ecuación (1.5) sólo que el método MCG realiza las estimaciones conociendo los parámetros poblacionales. En este método que estimar 'Ω_{ii} ' ($\hat{\Omega}_{ii}$),'ρ_i'($\hat{\rho}_i$) y 'σ' en una primera etapa. En la segunda, estimaremos el coeficiente de respuesta medio β.

El estimador de Ω_{ii} ($\hat{\Omega}_{ii}$) se obtiene reemplazando 'ρ_i' por $\hat{\rho}_i$ en Ωii. Más aún,

$$\hat{\rho}_i=\frac{\sum_{t=2}^{T}\hat{u}_{it}\hat{u}_{i,t-1}}{\sum_{t=2}^{T}\hat{u}_{i,t-1}^{2}} \quad (1.8)$$

donde $\hat{u}_{it}$ es el t-ésimo elemento de

$$\hat{u}_i=y_i-X_i\underline{b}_i \quad (1.9)$$

Sabiendo que $\hat{R}_i'\hat{R}_i=\hat{\Omega}_{ii}^{-1}$ tenemos que:

$$\hat{R}_i=\begin{bmatrix}\sqrt{1-\hat{\rho}_i^2} & 0 & 0 & \cdots & 0 & 0\\ -\hat{\rho}_i & 1 & 0 & \cdots & 0 & 0\\ 0 & -\hat{\rho}_i & 1 & \cdots & 0 & 0\\ 0 & 0 & -\hat{\rho}_i & \cdots & 0 & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & 0 & \cdots & -\hat{\rho}_i & 1\end{bmatrix} \quad (1.10)$$

El estimador de Δ lo obtenemos por [3]:

$$\hat{\Delta}(\hat{\rho})=\frac{S_b(\hat{\rho})}{n-1}-\frac{1}{n}\sum_{i=1}^{n}s_{ii}(X_i'\hat{\Omega}^{-1}{}_{ii}X_i)^{-1} \quad (1.11)$$

donde:

$$S_b(\hat{\rho})=\sum_{i=1}^{n}\underline{b}_i(\hat{\rho}_i)\underline{b}_i'(\hat{\rho}_i)-\frac{1}{n}\sum_{i=1}^{n}\underline{b}_i(\hat{\rho}_i)\sum_{i=1}^{n}\underline{b}_i'(\hat{\rho}_i) \quad (1.12)$$

En (1.12) los estimadores de los coeficientes para cada unidad 'i' de la muestra de corte transversal son tratados como una variable aleatoria.

e) Mínimos Cuadrados Generalizados de Hildreth&Houck (MCGHH)

El estimador de Hildreth & Houck [4] es como la ecuación (1.5), donde el estimador de la matriz varianza-covarianza de los Coeficientes β se calcula bajo el enfoque de estos autores. Así, el modelo sería:

$$b_{hh}(\hat{\Delta}_{hh})=\sum_{i=1}^{n}W_i(\hat{\Delta}_{hh})b_i \quad (1.13)$$

y,

$$W_i(\hat{\Delta}_{hh})=\left[\sum_j X_j'H^{-1}(\Delta_{hh})X_j\right]^{-1}X_i'H^{-1}(\Delta_{hh})X_i \quad (1.14)$$

donde Δ_{hh} se estima a partir de la ecuación (1.15), según la formulación de Hildreth & Houck:

$$\dot{u}=\dot{M}\dot{X}\underline{d}+\varepsilon=\dot{G}\underline{d}+\varepsilon \quad (1.15)$$

donde $E(\varepsilon)=0$ y $E(\varepsilon\varepsilon')=\varphi$ y $M=I-X'(X'X)^{-1}X'$ es la matriz TxT idempotente. Así, estimamos el vector de varianzas **d** que luego será reemplazado en la diagonal principal de $H_i(\hat{\Delta}_{hh}) = \dot{X}_i\hat{\Delta}_{hh}$.

f) Mínimos Cuadrados por Etapas de Aitken (MCEA) [5]

El estimador del coeficiente de respuesta medio es la suma ponderada de las estimaciones de los coeficientes para cada unidad de la muestra de corte transversal. Este método consta de dos etapas. En una primera etapa, el estimador de φ se obtiene utilizando el estimador de Hildreth & Houck especificado en la ecuación (1.15), es decir, $\varphi=MH(\Delta_{hh})M'$. La segunda etapa, consiste en utilizar el estimador de φ en la fórmula de Aitken para **d**:

$$d_{AITKEN} = (\dot{G}\varphi^{-1}\acute{G})^{-1}\dot{G}\varphi\dot{u} \qquad (1.16)$$

el cual será finalmente incorporado en $H(\Delta_{AITKEN})$.

g) Mínimos Cuadrados Ponderados de Theil (MCPT)

Theil propone un procedimiento mediante el cual los elementos fuera de la diagonal principal de φ en (1.15) son reemplazados por ceros obteniéndose $2\varphi_{MCPT}$, el cual es utilizado para estimar σ_{MCPT} por Mínimos Cuadrados Ponderados en (1.17):

$$\hat{\sigma}_{MCPT} = \dot{Z}'(2\hat{\varphi}_{MCPT})^{-1}\dot{Z})^{-1}\dot{Z}(2\hat{\varphi}_{MCPT})^{-1}\hat{\theta} \qquad (1.17)$$

Los estimados obtenidos a partir de la ecuación anterior corresponderán a las varianzas estimadas de los parámetros aleatorios a partir del cual tendremos el estimador de Theil para Δ: $\hat{\Delta}(\sigma_{MCPT})$. El estimador del coeficiente de respuesta medio es como el indicado en la ecuación (1.13) reemplazamdo Δ_{hh} por $\Delta(\sigma_{MCPT})$.

Propiedades de los Estimadores y su Operacionalización

Un estimador puntual se dice insesgado si el valor esperado de la diferencia entre el estimador y el verdadero valor del parámetro (precisión), es igual a cero. Para efectos de nuestro trabajo de investigación la definición de dicha propiedad estaría dada por el ***sesgo porcentual*** definido éste como el sesgo dividido por el valor del parámetro por cien por ciento. El valor del parámetro está dado por los valores que definen los 24 escenarios en el espacio parametral.

La consistencia de un estimador está dada por su capacidad de converger al verdadero valor del parámetro poblacional conforme la muestra aumenta. Sin embargo; a pesar de la evidencia, en el contexto del experimento nos interesará además las características de esta convergencia. En este sentido, la definición para nuestros fines es el de ***comportamiento errático*** del sesgo el cual podemos definirlo como el esquema presentado en las realizaciones observadas de éste a medida que aumenta el tamaño de la muestra de corte transversal. Es decir, el comportamiento errático es evidente en términos experimentales cuando el sesgo cambia de signo de positivo a negativo o viceversa o cuando no converge a cero monotónicamente al aumentar el tamaño de la muestra. El indicador de comportamiento errático lo definimos como el número de casos de convergencia no monotónica dividido por el número de casos totales por cien por ciento.

Un estimador insesgado de mínima varianza será eficiente. El presente trabajo de investigación toma como referencia al estimador Mínimos Cuadrados Generalizados (MCG) porque conocemos sus propiedades exactas dado que es estimado a partir del conocimiento *apriori* de los parámetros poblacionales. Por otro lado, debido a la existencia de distribuciones asimétricas, o a curtosis que se alejan de esquemas mesocúrticos y distribuciones bimodales que caracterizan a las muestras pequeñas utilizamos el rango intercuartílico como medida de variabilidad. En consecuencia, el ***indicador RI*** lo definimos como la razón entre el Rango Intercuartílico de una Técnica de Estimación X entre el Rango Intercuartílico de los Mínimos Cuadrados Generalizados (MCG) por cien por ciento.

DISEÑO DEL EXPERIMENTO

Para el estudio de las propiedades muestrales de los diferentes estimadores es necesario la generación de datos por simulación a partir de un conjunto de parámetros conocidos [6].

Se seleccionaron cuatro valores para el coeficiente de respuesta medio de tal modo que formaran un cuadrilátero en el espacio parametral. En cuanto a la matriz varianza-covarianza de los coeficientes aleatorios fue necesario considerar dos casos: primero, cuando la variabilidad de los parámetros es mínima, es decir, $\det(\Delta)\longrightarrow 0$ y, un segundo caso cuando la variabilidad es significativamente mayor que cero, $\det(\Delta)>0$.
Con respecto a la estructura autoregresiva del término errático se seleccionó tres valores para el coeficiente ρ (0.1, 0.5 y 0.9).

Al considerar siete técnicas de estimación para tres tamaños de muestra (10, 20 y 50) y 30 replicaciones en los 24 escenarios definidos en el espacio parametral fue necesario realizar 15,120 estimaciones.

Generación de las Variables Exógenas y Endógenas.

La generación de números aleatorios provenientes de una distribución normal se realizó mediante el generador de números pseudoaleatorios del *SPSS v7.5* [7]. Se optó por no elaborar programas específicos porque no es de nuestro interés los métodos y algoritmos empleados en la generación de dichos números. En segundo lugar, creemos que dicho trabajo hubiera implicado una asignación no óptima de recursos, principalmente de tiempo, en un tema que no es de primordial interés para esta investigación. En consecuencia, consideramos que una sucesión es aleatoria cuando los números pseudoaleatorios satisfagan un cierto conjunto de pruebas de aleatoriedad. Desde este punto de vista, el método para generar una sucesión es totalmente independiente.

La variable exógena será generada de tal modo que presente las mismas características de una variable económica: correlación serial, tendencia, etc. Para ello planteamos el siguiente esquema generador:

$$\begin{aligned} x_{1it} &= 0.5t + 1.2x_{1,i,t-1} + \omega_{it} \\ x_{2it} &= 2.5t + 7.6x_{1,i,t-1} + \omega_{it} \end{aligned} \qquad (2.1)$$

donde , $\omega \sim N_2(0,\Lambda)$ para $\Lambda = \begin{bmatrix} 7.0 & 0 \\ 0 & 30.0 \end{bmatrix}$

La observación inicial $X_{i,o}$ es generada, bajo las mismas condiciones, para la muestra de corte transversal mediante una corrida de números pseudoaleatorios en cada una de las 30 replicaciones.
En cuanto a las variables endógenas, las observaciones se generaron siguiendo un modelo lineal de parámetros aleatorios sin intercepto (Modelo II). Así tenemos:

$$y_{i,t} = \beta_{1,i}x_{1,i,t} + \beta_{2,i}x_{2,i,t} + u_{it} \qquad (2.2)$$

donde $u_{it} \sim U(0,5)$.

Finalmente, en cuanto al esquema autorregresivo de los errores $\mathbf{u}_i$ tenemos que:

$$u_{it} = \rho u_{i,t-1} + \varepsilon_{it} \quad ; \quad \varepsilon \sim N(0,5)\ , \quad 0<|\rho|<1 \qquad (2.3)$$

donde se analizará todos los casos para los cuales 'ρ ' toma los siguientes valores: 0.1, 0.5, 0.9.

RESULTADOS EXPERIMENTALES

El análisis comparativo de las propiedades muestrales de los diferentes estimadores nos lleva a organizar la exposición en dos partes. En la primera, se centra la atención en lo que es Insesgamiento y Consistencia; en la segunda parte, se analiza la Eficiencia.

Sesgo y Consistencia

El método de estimación Mínimos Cuadrados Ordinarios ha sido utilizado bajo los dos modelos especificados anteriormente, en este sentido, dado que el mecanismo generador de la variable endógena se basó en el Modelo II podemos considerar que la técnica MCOI incurre en un error de especificación del modelo, el caso de variables omitidas. El cuadro 1; muestra los casos, en términos porcentuales, donde el sesgo relativo ha sido mayor al 10% del verdadero valor del coeficiente de respuesta medio.

Los Mínimos Cuadrados Ordinarios presentan una significativa pérdida de precisión al hacerse una incorrecta especificación del modelo. En efecto, como puede observarse el modelo MCOI presenta pérdidas de precisión, si lo comparamos con el modelo MCOII, evidenciado esto último por los niveles más altos de sesgo relativo mayor al 10% con respecto al conjunto de técnicas de estimación para los tamaños de muestra especificados. Así, a pesar de que la precisión mejora conforme el tamaño de muestra aumenta, parael caso Mínimos Cuadrados Ordinarios con el Modelo I los casos de pérdida de precisión siempre se mantienen por encima del promedio conjunto.

La pérdida de precisión es más importante aún, cuando la muestra de corte transversal se reduce comparado a la muestra de series temporales T/N→1 debido a que la información relativa al mecanismo generador de los errores, resumido en ρ, es más relevante. Por esa razón, el método MCOI presenta en el 50% de los casos pérdidas de precisión cuando el tamaño de muestra es 10. Inverso es el resultado mostrado

Cuadro 1. *Casos (%) con Sesgo Relativo Mayor al 10% según Tamaño de Muestra y Técnica de Estimación*

Técnica de Estimación	Tamaño de Muestra 10	20	50	TOTAL
MCOI	50	25	13	**29**
MCOII	19	6	2	**9**
MCG	17	15	2	**11**
MCG2E	21	17	4	**14**
MCGHH	19	13	23	**18**
MCEA	44	38	25	**35**
MCPT	44	38	25	**35**
TOTAL	**29**	**21**	**13**	**21**

FUENTE: (Huamanchumo, 2000)

Cuadro2. *Casos (%) con Sesgo Relativo Mayor al 10% según Tamaño de Muestra y Coeficiente ρ por Técnicas de Estimación*

	Tamaño de Muestra								
	10			20			50		
Coeficiente ρ	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9
Técnicas de Estimación									
MCG	19	19	13	19	19	6	-	-	6
MCG2E	19	25	19	19	19	13	-	13	-
MCGHH	19	19	19	13	13	13	25	25	19
MCEA	44	44	44	38	38	38	25	25	25
MCPT	44	44	44	38	38	38	25	25	25
TOTAL	**29**	**30**	**28**	**22**	**22**	**19**	**13**	**14**	**13**

FUENTE: (Huamanchumo, 2000)

por MCOII, puede verse que cuando (T/N→0) la producción de éste se manifiesta en un 2% de casos de pérdida de precisión, esto último refleja la pérdida de importancia de la información de ρ cuando la muestra de corte transversal aumenta en la combinación de datos.

Como se podía esperar el Método de Mínimos Cuadrados Generalizados (MCG) y Mínimos Cuadrados Generalizados en Dos Etapas (MCG2E) muestran los niveles de precisión más altos -las pérdidas de precisión se mantienen significativamente bajas comparado al promedio general por tamaño de muestra- con respecto al conjunto de técnicas de estimación.

En el cuadro 2, puede verse que para los Mínimos Cuadrados Generalizados, a diferencia de lo visto para los Mínimos Cuadrados Ordinarios, la información respecto a 'ρ' sólo es relevante, en muestras pequeñas, cuando sus valores son cercanos a 1, es decir, en este caso las ganancias son significativas en precisión.

Este resultado era previsibe en la medida que el efecto de 'ρ' pueda hacerse evidente en muestras pequeñas sólo cuando toma valores altos. Por otro lado, cuando la estimación por Minimos Cuadrados Generalizados se hace en dos etapas (MCG2E) las ganancias de precisión se dan para valores extremos de 'ρ' inclusive en muestras de corte transversal grandes. En otras palabras, para 'ρ' cercano a 0.5 las pérdidas de precisión para MCG2E son evidentes. La estabilidad observada en los métodos MCEA y MCPT en todo el rango de variación de 'ρ' se debe a la incorporación de mayor información respecto al término errático en el modelo transformado (1.2).

Una primera aproximación al estudio de la Consistencia se presenta en el cuadro 3. Cuando la variabilidad de los coeficientes aleatorios es significativamente diferente de cero (det(Δ)> 0) los casos de comportamiento errático se reducen del 42% al 29%.

Cuadro 3. *Comportamiento Errático (%) del Sesgo según Coeficiente ρ y Δ [a] por Técnica de Estimación*

	Parámetro Poblacional Δ								TOTAL
	Δ_0				Δ_1				
Coeficiente ρ	**0.1**	**0.5**	**0.9**	TOTAL	**0.1**	**0.5**	**0.9**	TOTAL	
Técnica de Estimación									
MCOI	50	50	50	**50**	25	25	25	**25**	**38**
MCOII	38	38	50	**42**	13	13	13	**13**	**27**
MCG	25	38	38	**33**	25	13	25	**21**	**27**
MCG2E	25	25	50	**33**	25	13	13	**17**	**25**
MCGHH	38	38	38	**38**	25	13	38	**25**	**31**
MCEA	63	63	25	**50**	50	50	50	**50**	**50**
MCPT	63	63	25	**50**	50	50	50	**50**	**50**
TOTAL	**42**	**44**	**38**	42	**31**	**25**	**31**	29	**35**

FUENTE: (Huamanchumo, 2000)

a Matriz Varianza-Covarianza de los Parámetros Aleatorios

El caso MCG es importante analizar no sólo a raíz de los resultados obtenidos sino porque constituye el caso ideal entre las técnicas estudiadas debido a que utiliza en el proceso de estimación los verdaderos valores de los parámetros poblacionales. En el cuadro 3 podemos ver que, en términos promedio, ambos métodos tienen niveles más bajos de comportamiento errático. Sin embargo, dentro de cada uno de los escenarios se observa que el efecto de ρ es significativo para MCG2E, lo que no ocurre con MCG donde los casos de comportamiento errático se mantienen para los diferentes valores de ρ. Así, en escenarios donde la variabilidad de los coeficientes aleatorios es mayor ($\Delta 1$) la estructura de correlación del término errático caracterizado por un ρ cercano a 0.5 mejora significativamente la consistencia en el método de MCG2E al igual que MCG quien presenta un esquema más estable en escenarios donde la variabilidad de los parámetros aleatorios es menor ($\Delta 0$) inclusive.

Eficiencia

Sobre la base del ratio RI podemos observar la evidente superioridad del estimador MCG2E en un rango considerable del espacio parametral y tamaños de muestra. Del mismo modo, la evidencia empírica muestra las significativas ganancias en eficiencia del estimador Mínimos Cuadrado Ordinarios cuando el modelo utilizado es el correcto. En efecto, puede verse que el estimador Mínimos Cuadrados Ordinarios en el Modelo I (con error de especificación) califica en último lugar teniendo una significativa mejora cuando $\det(\Delta)\rightarrow+\infty$ aunque no hay ganancias asintóticas para este caso. Sin embargo, cuando se utiliza el Modelo II, las ganancias en eficiencia son importantes: los Mínimos Cuadrados Ordinarios (MCOII) ocupan el segundo lugar en todo el rango considerado del espacio parametral y tamaños de muestra.

Cuadro 4. Calificación del RI a por Parámetro Estimado según Tamaño de Muestra N, Método de Estimación y Escenarios definido por Δ [b]

		Parámetro PoblacionalΔ					
		Δ_0			Δ_1		
N		β_1	β_2	TOTAL	β_1	β_2	TOTAL
10	MCOI	6	n.a	**n.a**	4	n.a	**n.a**
	MCOII	2	2	**2**	2	2	**2**
	MCG2E	1	1	**1**	1	1	**1**
	MCGHH	3	3	**3**	3	3	**3**
	MCEA	4	4	**4**	6	5	**5**
	MCPT	5	5	**5**	5	4	**4**
20	MCOI	6	n.a	**n.a**	4	n.a	**n.a**
	MCOII	2	2	**2**	2	2	**2**
	MCG2E	1	1	**1**	1	1	**1**
	MCGHH	3	3	**3**	3	3	**3**
	MCEA	5	4	**4**	5	4	**4**
	MCPT	4	5	**4**	6	4	**5**
50	MCOI	6	n.a	**n.a**	4	n.a	**n.a**
	MCOII	2	2	**2**	2	2	**2**
	MCG2E	1	1	**1**	1	1	**1**
	MCGHH	3	3	**3**	3	3	**3**
	MCEA	4	4	**4**	6	5	**5**
	MCPT	5	4	**5**	5	4	**4**

FUENTE: (Huamanchumo, 2000)
a Ratio Rango Intercuartílico vs. Rango Intercuartílico de los MCG
b Matriz Varianza-Covarianza de los Parámetros Aleatorios

Para los métodos basados en el enfoque de Hildreth & Houck podemos ver que MCEA y MCPT presentan ventajas relativas dependiendo del contexto lo cual responde satisfactoriamente a nuestras expectativas sustentadas en las características que posée la matriz varianza-covarianza del término errático en cada uno de los casos.

Así, por ejemplo, mientras MCEA tiene mejor calificación respecto a eficiencia asintótica que MCPT en escenarios caracterizados por una menor variabilidad en los parámetros aleatorios (Δ_0), éste lo supera a aquél en escenarios cuya variabilidad de los parámetros aleatorios es mayor (Δ_1). La razón radica, como ya lo mencionamos, en que la matriz varianza-covarianza del término errático para MCEA posée mayor información respecto a los errores que el método MCPT ya que este último sólo le da importancia a las varianzas con las cuales pondera la estimación y no a la correlación entre ellos, en consecuencia, cuando las varianzas de los parámetros aleatorios son más significativas MCPT mejora en la estimación. Los resultados sobre eficiencia en Error Cuadrático Medio son similares a los obtenidos según Rango Intercuartílico [8].

CONCLUSIONES Y RECOMENDACIONES

1. Los métodos de estimación para el modelo de Regresión de Coeficientes Aleatorios producen estimadores Insesgados, Consistentes y Eficientes que son afectados por los valores de los Parámetros Poblacionales.

2. El estimador MCG2E obtiene la mejor calificación en Eficiencia y Eficiencia en Error Cuadrático Medio, además, la información respecto al término errático '$\mathbf{u}_i$' que incorporan los métodos MCG y MCG2E son más relevantes para la Consistencia cuando T/N tiende a 1.

3. El método Mínimo Cuadrados Ordinarios mostró ganancias en Eficiencia cuando se estimó sobre un modelo sin error de especificación.

4. Recomiendo desarrollar estudios similares que incluyan el Método de Máxima Verosimilitud. Al mismo tiempo, es importante desarrollar pruebas estadísticas de aleatoriedad de los coeficientes de modo que facilite la aplicación de dicho modelo.

Agradecimiento

Tengo que agradecer las facilidades brindadas y el apoyo incondicional a la tarea científica por parte del Ing. Sergio Cuentas, Director del Instituto de Investigaciones Económico Sociales-IECOS.

REFERENCIAS

1. Rao, C.R. Linear Statistical Models and Its Applications. New York: John Wiley & Sons. 1971.
2. Johnson, Richard Applied Multivariate Statistical Analysis. Third Edition. Prentice Hall International Inc. 1992.
3. Swamy, P. A. V. B. Statistical Inference in Random Coefficient Regression Models. Springer Verlag. 1971.
4. Hildreth, Clifford & Houck, James (junio, 1968) "Some Estimators for a Linear Model with Random Coefficients". Journal of the American Statistical Asociation, pp. 584-595.
5. Raj, Baldev (marzo, 1971) "Linear Regression with Random Coefficients: The Finite Sample and Convergence Properties". Journal of the American Statistical Association. Vol 70, No 349, pp. 127-137.
6. Haussman, Jerry y Taylor, William (marzo, 1979) "Attrition Bias in Experimental and Panel Data: The Gary Income Maintenance Experiment". Econometrica. Vol No 2, pp. 455-472.
7. Marija J., Norousis. SPSS. Advanced Statistics 6+.1 SPSS Inc. 1994.
8. Huamanchumo de la Cuba, Luis "Eficiencia en Muestras Finitas en el Modelo de Parámetros Aleatorios de Hildreth & Houck: Un Experimento Muestral". Tesis de Licenciatura. Escuela Profesional de Ingeniería Estadística.Universidad Nacional de Ingeniería, Octubre 2000.