Técnicas Econométricas para Estimar Funciones de Demanda

Dr. José Luis Arrufat

Resumen

El trabajo analiza dos temas: la transformación de Box-Cox y la estimación conjunta de ecuaciones de demanda, con ajuste parcial y estructura de errores AR(1). Con respecto a éste, se implementa un algoritmo basado en el uso de variables instrumentales que provee estimadores consistentes y eficientes. El trabajo consigna los detalles del programa de computación utilizado, escrito en lenguaje Gauss. Por último, se utiliza el programa para estimar un modelo de demanda de tráfico telefónico interurbano para las ciudades de Córdoba y Rosario, utilizando datos para el período 1991.02 a 1993.10.

Abstract

The paper addresses two main topics: the Box-Cox transformation and the estimation of a System of Seemingly Unrelated Regression Equations (SUR), with partial adjustment and AR(1) errors. In connection with the latter, an algorithm based on instrumental variables is implemented. The algorithm provides consistent and efficient estimates of the parameters of the model. The paper lists the code for the computer programme, which is written in the Gauss programming language. The programme is subsequently used to estimate the parameters of a long-distance telephone demand system for the cities of Córdoba and Rosario, for the period February 1991 to October 1993.

CODIGO JEL : C3

TÉCNICAS ECONOMÉTRICAS PARA ESTIMAR FUNCIONES DE DEMANDA

Dr. José Luis Arrufat*

Universidad Nacional de Córdoba

I. Introducción

La Econometría provee un marco conceptual adecuado para el estudio de funciones de demanda. En efecto, basándose en nociones provenientes de la teoría económica como de la matemática y de la estadística - según la célebre definición del premio Nobel de Economía Ragnar Frisch - se aboca a la tarea de formular y estimar modelos económicos, procediendo, asimismo a la contrastación de hipótesis. En el caso concreto de las funciones de demanda existen algunas propiedades importantes que resulta interesante contrastar, tales como la condición de agregación de Engel, la de simetría de efectos cruzados, la de homogeneidad de grado cero en precios e ingreso, entre otras. Estos contrastes son posibles cuando se estima un sistema completo. Un caso igualmente interesante es el que contempla la modelización en escala espacial, a través de distintas ecuaciones para distintas ciudades, por ejemplo. En este contexto, y para el estudio de la demanda de tráfico telefónico, algunas hipótesis interesantes para contrastar están dadas por la igualdad de elasticidades precio, ingreso o tamaño de la red telefónica para los diversos centros urbanos incluidos en el estudio.

Sin embargo, el marco conceptual genérico antes referido no resulta en modo alguno completo para emprender estimaciones empíricas, por las razones que a continuación se detallan. Primero, no es evidente cuál es la forma funcional adecuada para el estudio de la relaciones de interés. Según cuál sea la forma funcional elegida, las elasticidades precio estimadas serán constantes en caso de optar por una especificación lineal en logaritmos o, en caso de plantear que el logaritmo natural de la cantidad demandada depende linealmente del propio precio, se obtendrá, por el contraria, una elasticidad que resulta ser proporcional al valor de la variable precio. Este último caso, que se ha planteado recientemente en algunos debates vinculados a estimaciones de elasticidades de demanda de servicio telefónico básico, tanto para el caso de tráfico urbano como interurbano, se traduce en que a mayores precios del minuto de comunicación telefónica se deben esperar mayores elasticidades en valor absoluto.

No es posible establecer, *a priori*, cuál de las especificaciones alternativas antes mencionadas resulta más adecuada a los fines del análisis. Es por ello que muchas veces la práctica econométrica suele resolver esta cuestión de forma más o menos expeditiva postulando una determinada forma funcional. Existe, sin embargo, una técnica conocida como transformación de Box-Cox, que posibilita contrastar hipótesis referidas a la forma funcional adecuada y cuya implementación empírica se ha intentado en este trabajo.

Una cuestión adicional de gran importancia está dada por la posibilidad de realizar estimaciones de Regresiones aparentemente no relacionadas, conocidas en la literatura de habla inglesa como SUR o SURE ("*Seemingly Unrelated Regression Equations*"). En efecto, esta técnica resulta de gran utilidad por cuanto permite estimar conjuntamente un grupo de ecuaciones de demanda, así como contrastar la validez de cierto tipo de restricciones como la condición de agregación de Engel, homogeneidad, simetría, etc.

Si bien la utilización de los paquetes de programas más habituales en Econometría, tales como el TSP (*Time Series Processor*) permite la implementación de esta técnica, su empleo no resulta completamente satisfactorio por dos motivos:

Primero, cuando se postula que el sistema a estimar responde a un esquema de ajuste parcial, se incluye, obviamente, la propia variable endógena rezagada entre las variables explicativas. Resulta, por lo tanto, de fundamental importancia disponer de una técnica de estimación por variables instrumentales, dado que ello asegura la consistencia de los estimadores obtenidos en caso de que el término de error aleatorio de la ecuación sea autorregresivo de primer orden. De la lectura de los manuales del programa TSP no surge con claridad cuál es el algoritmo efectivamente empleado. Tampoco claro si es que la presencia de variables endógenas rezagadas supone un planteo especial adecuado a ese caso por parte del algoritmo.

En segundo lugar, la literatura considera habitualmente procesos autorregresivos más generales, que involucran los términos de error de más de una ecuación al mismo tiempo, tal como en el algoritmo de estimación propuesto por Spencer (1979).

Lo que resta del presente trabajo se organiza como sigue. En la segunda sección se presenta una brevísima síntesis de la principales ideas subyacentes para la transformación de Box y Cox y acerca de la utilidad potencial que esta técnica brinda para ayudar en la búsqueda de la forma funcional más adecuada en la estimación de funciones de demanda. En la tercera sección, se describe detalladamente la implementación del algoritmo de Spencer (1979), obtenido a partir de la aplicación del estimador en dos etapas de Hatanaka. En la cuarta sección se aplican los métodos desarrollados en la sección precedente para la estimación de funciones de demanda de servicio telefónico básico. La estimación efectuada tiene un carácter fundamentalmente ilustrativo del uso de las técnicas econométricas referidas en el presente trabajo. En la quinta sección se resumen las ideas principales del trabajo y se presentan las conclusiones obtenidas. Un anexo contiene un listado detallado de algunos programas escritos en lenguaje GAUSS para la implementación del método de estimación que se presenta en la tercera sección. Por último se consignan las referencias bibliográficas.

II. La transformación de Box y Cox

La transformación de Box y Cox[1]consiste en plantear la inclusión de una variable "y", en una ecuación a estimar, se haga a través de la siguiente forma funcional:

$$y^{(\lambda)} = (y^{\lambda} - 1) / \lambda$$

La ventaja que reporta el empleo de esta forma funcional es la siguiente. Si λ es igual a cero, la variable "y" entre en forma logarítmica[2]. Si por el contrario, λ es igual a 1, la variable "y" entra en forma lineal. La idea subyacente en la utilización de esta transformación consiste, por lo tanto, en valerse de la fórmula anterior para estimar una ecuación, incluyendo la propia λ como un parámetro adicional a estimar. Una vez realizada la estimación, y según sea el valor asumido por λ, se puede decidir si mejor forma funcional a adoptar para el problema proviene de la forma logarítmica, de la lineal, o inclusive de alguna otra diferente de estas dos. La técnica, tal como la plantearon Box y Cox es de tal naturaleza que posibilita el planteo de contrastes estadísticos apropiados para el valor de λ.[3]

En el marco de las tareas de investigación realizadas para el presente trabajo se desarrollaron los programas de cómputo adecuados para implementar esta transformación. Al efecto, se utilizaron resultados contenidos en Fomby y otros (1984), Green (1993), Judge y otros (1985 y 1988), Hill (1989), Spitzer (1978, 1982 y 1984) . Se procedió a verificar el correcto funcionamiento de los programas, escritos en lenguaje GAUSS, mediante la estimación de modelos generados artificialmente mediante el método de Monte Carlo.

Sin embargo, luego de experimentar con esta transformación se optó por no utilizarla en la estimación de modelos de demanda de servicio telefónico básico, para llamadas interurbanas del tipo desarrollado en Abdala, Arrufat, Colomé y Neder (1996), por cuanto la base de datos disponible resulta demasiado corta. En efecto, dichos autores contaron con series de datos mensuales correspondientes al período febrero de 1991 a octubre de 1993. Esto hace un total de 33 observaciones mensuales, referidas a cinco centros urbanos: Córdoba, Corrientes, Posadas, Resistencia y Rosario. La experiencia acumulada en las simulaciones reveló que las estimaciones de λ carecen de la precisión suficiente cuando se cuenta con un número tan reducido de observaciones.

III. Modelos SUR con ajuste parcial y errores aleatorios AR(1)

Se presenta a continuación una breve descripción del algoritmo de Spencer (1979). Considérese la siguiente ecuación, que es la i-ésima de un total de M ecuaciones que se estimarán conjuntamente:

$$y_i = Z_i b_i + u_i \qquad \mathrm{i} = 1,2,\ldots,\mathrm{M} \tag{1}$$

El significado de los símbolos que se emplean en (1) es el siguiente:

Z_i es la matriz de TxK_i de variables explicativas para la i-ésima ecuación, siendo T el número de observaciones y K_i el número de variables explicativas.

b_i es un vector de $K_i x1$ de parámetros a estimar.

y_i es un vector de Tx1 de observaciones de la i-ésima variable explicada del modelo de M ecuaciones.

u_i , es, por su parte, el vector de términos aleatorios correspondiente a la i-ésima ecuación, siendo sus dimensiones Tx1. Su esperanza matemática es el vector nulo, mientras que su matriz de varianzas y covarianzas se define más adelante.

El modelo de M ecuaciones de la fórmula (1), se puede reexpresar como sigue:

$$y = Zb + u \tag{2}$$

En la ecuación (2), y es un vector de TMx1, que se obtiene 'apilando" (*stacking)* los vectores columna y_i para las M ecuaciones que se estiman conjuntamente, en el mismo orden en que plantean en (1). La matriz Z, cuyas dimensiones son TMxK, siendo K la sumatoria de las K_i definidas anteriormente, es una matriz diagonal por bloques, que se obtiene a partir de la siguiente expresión: diag(Z_1, Z_2, ..., Z_M).

El vector b , por su parte, es un vector columna que tiene K filas, cuyos elementos son los parámetros del sistema. Por último, u es el vector de términos de error aleatorio, cuyas dimensiones son Tmx1, que se supone tiene la estructura autorregresiva de orden uno, de acuerdo con la expresión que se presenta a continuación:

$$u = (R' \otimes I_T)u_{-1} + e \qquad \textbf{(3)}$$

En la expresión (3) R es una matriz de MxM cuyas raíces características se ubican dentro del círculo unitario para satisfacer las condiciones de estacionariedad, u_{-1} simboliza el vector de errores aleatorios, rezagado en un período. Por último, e es un vector columna de TM filas, que simboliza términos de error aleatorios, cuya esperanza es el vector nulo de dimensione conformables y cuya matriz de varianzas y covarianzas se supone de la estructura que se presenta a continuación:

$$E[ee'] = \Sigma \otimes I_T \qquad \textbf{(4)}$$

En la fórmula (4), Σ simboliza una matriz de MxM, que se supone definida positiva y simétrica, por ser una matriz de varianzas y covarianzas.

La particularidad del modelo presentado por Spencer, constituido por las ecuaciones (1) a (4) es que no se imponen restricciones sobre la matriz R, salvo, obviamente, las que se refieren a sus raíces características para garantizar la estacionariedad y que se especifica expresamente que la matriz Z contiene la propia variable endógena rezagada en un período entre los regresores del modelo, para capturar una característica básica de un modelo de ajuste parcial. Es a partir de esta última característica del modelo que se plantea la necesidad de contar con un algoritmo de estimación adecuado, con el fin de obtener estimadores consistentes y asintóticamente eficientes.

Las ecuaciones que se presentan a continuación siguen a Hatanaka (1976) y Dhrymes and Taylor (1976) y plantean un estimador en dos etapas. La primera de esas etapas consiste en obtener estimadores consistentes para los b_i, simbolizados por el agregado de una tilde, vale decir, $\tilde{b}$. Dichos estimadores se obtienen a partir de cada una de las M ecuaciones, tomadas individualmente, a partir de la utilización del método de variables instrumentales, aunque podrían utilizarse otros alternativos. A partir de esos estimadores, se deben calcular los términos de la fórmula (5), que se presenta a continuación:

$$\tilde{u}_i = y_i - Z_i \tilde{b}_i \qquad \mathrm{i = 1,2,\ldots,M} \qquad \textbf{(5)}$$

Agrupando cada uno de los M términos calculados en (5), se obtiene:

$$\tilde{U} = [\tilde{u}_1 \ \tilde{u}_2 \ \dots \ \tilde{u}_M] \qquad \textbf{(6)}$$

La fórmula siguiente plantea la estimación de la matriz R, simbolizada por $\tilde{R}$, a partir de la técnica de regresión por mínimos cuadrados ordinarios para la M ecuaciones conjuntamente, tal como se expresa en la siguiente fórmula:

$$\tilde{R} = (\tilde{U}'_{-1}\tilde{U}_{-1})^{-1}\tilde{U}'_{-1}\tilde{U} \qquad \textbf{(7)}$$

Se procede ahora a calcular las expresiones (8) a (12), que se presentan a continuación:

$$\tilde{E} = \tilde{U} - \tilde{U}_{-1}\tilde{R} \qquad \textbf{(8)}$$

$$\tilde{\Sigma} = T^{-1}\tilde{E}'\tilde{E} \qquad \textbf{(9)}$$

$$H = [Z - (\tilde{R}' \otimes I_T)Z_{-1}, \ (I_M \otimes \tilde{U}_{-1})] \qquad \textbf{(10)}$$

$$\hat{q}^* = \begin{pmatrix} \hat{b} \\ \hat{r} - \tilde{r} \end{pmatrix} = [H'(\tilde{\Sigma}^{-1} \otimes I_T)H]^{-1}H'(\tilde{\Sigma}^{-1} \otimes I_T)[y - (\tilde{R}' \otimes I_T)y_{-1}] \qquad \textbf{(11)}$$

$$\hat{q} = \begin{pmatrix} \hat{b} \\ \hat{r} \end{pmatrix} = \hat{q}^* + \begin{pmatrix} 0 \\ \tilde{r} \end{pmatrix} \qquad \textbf{(12)}$$

Se presenta, seguidamente una brevísima caracterización de las expresiones anteriores. La fórmula (11) plantea la estimación por mínimos cuadrados generalizados factibles, es decir una vez obtenidos los estimadores apropiados para los elementos de la matriz $\tilde{\Sigma}$, obtenida en (9), y a su vez, proveniente de $\tilde{E}$, a partir de la fórmula (8).

Debe destacarse que el vector columna de estimadores de los parámetros en (11), tiene (K + M^2) filas. Ello resulta del número total de parámetros estimados conjuntamente que asciende a K para los parámetros b, mientras que los M^2 restantes corresponden a los estimadores de los elementos de la matriz R, no necesariamente diagonal, ni siquiera simétrica, cuyas raíces características deben estar comprendidas dentro del círculo unitario.

Por último, en la fórmula (12) se procede a corregir el estimador obtenido en la ecuación anterior sumando los valores obtenidos para la estimación de la matriz R, a través de $\tilde{R}$ en la fórmula (7). Se hace imprescindible, obviamente, vectorizar esta última matriz para poder realizar dicha operación. La vectorización adoptada es la vectorización por columnas.

Spencer considera, asimismo, la iteración del procedimiento aquí descripto. Para ello, es necesario repetir las secuencias desde la fórmula (5), deteniendo el algoritmo cuando se cumpla un número prefijado de iteraciones o bien cuando la diferencia obtenida entre dos estimaciones sucesivas sea menor que un cierto valor pequeño fijado como tolerancia.

Una cuestión adicional, de gran importancia práctica para la implementación del presente algoritmo, es que las matrices intervinientes en la fórmula (12) pueden ser de grandes dimensiones. Téngase presente que H es una matriz de $TMx(K+M^2)$. La implementación de este algoritmo puede requerir, cuando M, T o K son grandes, de una administración eficiente de la memoria disponible en la computadora.

Una vez concluida la fase de desarrollo de los programas y verificada la correcta estimación de modelos por medio de ejemplos sencillos basados en la metodología de Monte Carlo, utilizando los programas que se detallan en el anexo, se realizó una estimación con fines ilustrativos, referida a la demanda de servicio telefónico básico, que se detalla en la sección siguiente.

IV. Estimación de un modelo SUR de demanda de servicio telefónico

En tiempos recientes se han realizado dos estudios de demanda de servicio telefónico básico, tanto para llamadas urbanas como interurbanas. En el primero, NERA (1995), estima ecuaciones de demanda de servicio telefónico, tanto para llamadas urbanas como interurbanas, incluyendo además una ecuación estimada para explicar el tráfico total saliente para las llamadas internacionesl. En el segundo, Abdala, Arrufat, Colomé y Neder (1996), por el contrario, sólo se estiman ecuaciones de demanda para llamadas urbanas e iinterurbanas.

Una diferencia importante entre los enfoques utilizados en los dos estudios citados es que el primero no tiene en cuenta la modelización espacial, agregando, por consiguiente todo el tráfico en una sola variable. Tampoco considera, obviamente, el tamaño de la red de destino en forma desagregada. El trabajo de NERA, por lo que respecta a la modelización interurbana adopta un esquema interesante. En efecto, para paliar el problema planteado por series estadísticas de corta duración, incluye en forma separada el tráfico para cada una de las bandas horarias (normal, reducida y nocturna-feriados). Sin embargo, resulta difícil comprender el significado de un esquema de errores aleatorios autorregresivo, como el que plantean, en ese contexto.

En Abdala, Arrufat, Colomé y Neder (1996) se estimaron una serie de modelos por medio de la técnica SUR, para los centros urbanos de Córdoba, Corrientes, Posadas, Resistencia y Rosario, todos ellos ubicados en el área de exclusividad de Telecom.

Tal como se señala en ese trabajo, sus autores enfrentaron algunas limitaciones de disponibilidad de datos, con las consecuencias que se resumen a continuación:

1) Los modelos estimados se refieren sólo a la categoría de casas de familia. Ello sólo representa, aproximadamente, el 34% del tráfico total cursado, en minutos.

2) Se intentó incluir una variable explicativa de la calidad del servicio telefónico, probando para ello con el nivel de digitalización, que no resultó adecuada. En efecto, no pudieron contar con una variable que midiera el grado de congestión del sistema telefónico.

3) La variable correcta para la estimación de las ecuaciones econométricas debe expresarse en minutos y no en pulsos, lo que requirió de transformaciones adecuadas teniendo en cuenta los ritmos y los duraciones medias de las comunicaciones, así como también la composición del tráfico entre las distintas bandas horarias y claves tarifarias.

4) La estacionalidad presente en las series utilizadas resultó, en gran medida, una consecuencia no desead del uso de ponderadores muestrales.

5) No se pudieron aplicar métodos de corrección que resulten completamente adecuados para neutralizar los efectos de los ciclos de facturación (*billing cycle)*. Por lo tanto, el tratamiento que recibió este problema resultó bastante rudimentario.

6) Las estimaciones quedaron restringidas al ámbito de Telecom. Para el caso de las llamadas interurbanas, sólo se estudiaron las ciudades de Córdoba, Corrrientes, Posadas, Resistencia y Rosario. Para estas ciudades, adicionalmente, sólo pudieron estimarse ecuaciones para las claves interurbanas comprendidas entre la uno y la cuatro. Para facillitar la referencia se consigna que la primera clave se refiere a llamadas cursadas hasta 30 Km, la segunda desde 30 hasta 55Km, la tercera, desde 55 hasta 110 Km y la cuarta, incluye distancias comprendidas entre los 110 y los 170 Km.

A los fines de ilustrar el uso del estimador SUR para modelos de ajuste parcial con errores aleatorios AR(1), con matriz R no necesariamente diagonal, se presentan a continuación los resultados correspondientes a un modelo similar al estimado por Abdala, Arrufat, Colomé y Neder (1996) para las ciudades de Córdoba y Rosario y para llamadas efectuadas desde cada una de esas ciudades a distancias comprendidas entre 30 y 55

Km del origen. No se han incorporado las restantes ciudades del estudio referido por cuanto ello hubiera redundadoa en un gran aumento en el tamaño de las matrices para implementar la técnica de mínimos cuadrados generalizados factibles con variables instrumentales de Spencer descripta en la sección anterior. La matriz H, por ejemplo, de la fórmula (10), es de dimensiones MTx(K + M^2). De utilizarse los cinco centros urbanos, M asumiría el valor 5. El período utilizado para la estimación está comprendido entre febrero de 1991 y octubre de 1993, lo que da un total, T, de 33 observaciones mensuales. Como cada una de las ecuaciones contiene seis parámetros a estimar, el nómero total de parámetros a estimar, K, asciende a 30. Las dimensiones de H para este caso ascenderían a 165 filas por 55 columnas.

El modelo estimado en el presente trabajo considera que el logaritmo natural del tráfico saliente para cada ciudad (Córdoba y Rosario), medido en minutos, incluye las siguientes variables:

a) ordenada al origen.

b) la propia variable endógena rezagada en un período, por tratarse de modelos de ajuste parcial.

c) una variable *dummy* que asume el valor 1 para los meses del año 1993 y 0 en los restantes meses de la muestra, para captar algunas características de los datos utilizados. En efecto, debido al empleo de ponderadores muestrales en la elaboración de la base de datos se puede observar un salto bastante apreciable en algunas variables de tráfico[4].

d) el logaritmo natural del valor del minuto, ponderado según el tráfico cursado en las bandas horarias normal, reducida y nocturna-feriados, en términos reales.

e) el valor real del abono del servicio telefónico, expresado en logaritmos naturales.

f) el logaritmo natural de la red de destino. Se entiende por red de destino, el número total de teléfonos que pueden alcanzarse desde cualquiera de las ciudades consideradas en la muestra y comprendidas en las distancias de entre 30 y 55 Km. En la terminología de uso habitual en telefonía, las llamadas cursadas a estas distincias se denominan de Clave tarifaria 2.[5]

g) Se postula que el proceso generador de los errores aleatorios del modelo responde a procesos autorregresivos con matriz R no diagonal.

Tal como se desprende de las consideraciones anteriores, y por tratarse de un modelo lineal en logaritmos, el modelo a estimar postula elasticidades constantes del tráfico para las ciudades de Córdoba y Rosario, con respecto al precio (valor del minuto ponderado), del abono y del tamaño de la red de destino.

Se presenta a continuación una breve síntesis de los resultados obtenidos. La Tabla 1 consigna los resultados obtenidos para las elasticidades de corto plazo, usando un modelo SUR simple, mientras que la Tabla 2 contiene los resultados de la estimación a partir del modelo SUR con errores AR(1). En ambos casos, los números entre paréntesis debajo de los coeficientes son los correspondientes errores *standard* estimados.

Tabla 1

Elasticidades de corto plazo estimadas

Método SUR

Ciudad	Precio	Abono	Red
Córdoba	-0.344 (0.213)	-0.048 (0.031)	0.745 (0.316)
Rosario	-0.407 (0.205)	-0.078 (0.032)	0.581 (0.142)

Tabla 2

Elasticidades de corto plazo estimadas

Método SUR con errores AR(1)

Ciudad	Precio	Abono	Red
Córdoba	-0.410 (0.211)	-0.053 (0.030)	0.694 (0.345)

Rosario	-0.507 (0.202)	-0.094 (0.031)	0.623 (0.135)

A partir del análisis de los resultados contenidos en las Tablas 1 y 2, se pueden realizar las siguientes consideraciones:

1) Las elasticidades precio son mayores en valor absoluto en la Tabla 2 y sus errores *standard* ligeramente inferiores. Así, por ejemplo, la elasticidad precio de corto plazo para Córdoba se estima en -0.344 en la Tabla1, mientras que asciende a -0.410 en la Tabla 2. Los errores *standard* estimados son de 0.213 y 0.211, respectivamente.

2) Si se analizan las elasticidades de demanda con respecto al abono, se extraen conclusiones similares a las del punto anterior.

3) Las comparaciones que involucran las elasticidades con respecto al tamaño de la red de destino no presentan un perfil tan definido como el encontrado en los dos puntos anteriores. En efecto, para la ciudad de Córdoba, la elasticidad estimada en 0.745 (Tabla 1), cae a 0.694 (Tabla 2), observándose, al mismo tiempo, un pequeño aumento en el error *standard* asociado. Sin embargo, cuando se examinan los resultados pertinentes para la ciudad de Rosario se detecta un patrón diferente. Para esta última ciudad la comparación entre los resultados de las Tablas 1 y 2 revela que la elasticidad red aumente de 0.581 a 0.623. Los errores *standard* estimados ascienden a 0.142 y 0.135, respectivamente.

V. Resumen y conclusiones

En la segunda sección se presentó una muy breve caracterización de la transformación de Box y Cox. Si bien se escribieron los programas de cómputos apropiados para la implementación de esta técnica, se optó finalmente por no utilizarla en la modelización de un conjunto de ecuaciones de demanda de tráfico interurbano.

En la sección se presentó el algoritmo de Spencer (1979) cuya implementación se realizó en la sección 4 para el estudio de un modelo de demanda de tráfico interurbano para las ciudades de Córdoba y Rosario. Las elasticidades obtenidas en el presente trabajo tienen un error *standard* menor que las obtenidas a partir de un estimador alternativo. Sin embargo, es posible que el mayor logro obtenido en las nuevas estimaciones provenga del uso del estimador de variables instrumentales, más que de la estructura autorregresiva bastante más general que posibilita el algoritmo aquí implementado.

Anexo

Programas de cómputo utilizados

En el presente Anexo se transcriben diversos programas de cómputo escrito en lenguaje GAUSS. Dichos programas se han utilizado para la implementación de las distintas etapas de estimación del estimador por variables instrumentales y mínimos cuadrados generalizados factibles, tal como requiere el algoritmo propuesto por Spencer (1979).

En la medida de lo posible se ha usado una notación similar a la utilizada en el trabajo de referencia. La diferencia tal vez más importante radica en el uso de letras minúsculas en lugar de las mayúsculas. Así, por ejemplo, cuando se simboliza el número total de observaciones se utiliza la letra t, en lugar de T y m es el número de ecuaciones del sistema de regresiones aparentemente no relacionadas (SUR), en lugar de M. El producto de Kronecker se simboliza en el lenguaje GAUSS como *.*. Otros símbolos propios de dicho lenguaje son el concatenador vertical y el horizontal, representados por | y ~, respectivamente.

El procedimiento GENERA.G tiene por objeto generar datos similares a los utilizados por el propio Spencer para la verificación del funcionamiento adecuado del algoritmo. Téngase presente que no se ha intentado aquí realizar un ejercicio completo del tipo de Monte Carlo sino solamente apreciar, a partir de un número pequeño de replicaciones, si el programa de cómputo tenía el desempeño satisfactorio documentado por el Spencer a partir de las simulaciones exhaustivas a las que él mismo lo sometió.

Los insumos utilizados por el procedimiento son (x1,x2,m,t,r,sigma,beta1,beta2), que se describen brevemente a continuación. Tanto x1 como x2 se ha tomado del artículo de Spencer (1979), pág. 230, habiendo sido utilizados previamente por otros autores. Se trata, en ambos casos, de series de 10 valores cada una, que se repiten el número necesario de veces, simbolizada por nveces, de tal forma de lograr una muestra de t valores para las variables exógenas explicativas del modelo.

Por lo que respecta a m, se reitera que simboliza el número de ecuaciones del modelo, mientras que t representa el tamaño total de la muestra, para cada ecuación tomada individualmente. Los símbolos r y sigma, representan la matrices de igual nombre, mientras que beta1 y beta2 se refieren a los vectores del verdadero valor de los parámetros de igual nombre.

Los errores aleatorios para la simulación del modelo se simbolizan como eps, obteniéndose a partir del generador de número normales, transformados a través de la descomposición de Cholesky de la matriz sigma.

El procedimiento devuelve (y1, y2, z1, z2) al programa que lo invoca. Tal como se puede apreciar, los dos primeros símbolos se corresponden con los vectores de variables explicadas, para cada una de las dos ecuaciones que forman el conjunto a estimar, mientras que los dos últimos representan las matrices de variables explicativas correspondientes a esas ecuaiones. Debe tenerse presente que la correspondiente variable endógena rezada en un período, propia del tipo de modelo de ajuste parcial adoptado, ocupa la segunda columna en cada una de las matrices z. Con respecto a esta matriz, su primera columna está constituida por una variable que siempre asume el valor 1, asociada a la ordenada al origen de cada una de las ecuaciones y su tercera columna corresponde a la variable exógena apropiada para cada ecuación. Tal como está escrito, el procedimiento sólo contempla el caso m = 2, es decir, 2 ecuaciones. Sin embargo, resulta sencillo extenderlo al caso más general.

El código correspondiente al procedimiento GENERA.G es el siguiente.

```
proc (4) = genera(x1,x2,m,t,r,sigma,beta1,beta2);
                local x1,x2,m,t,r,sigma,beta1,beta2,nveces,nv;
                local x1v,x2v,eps,u10,u20,r11,r22,ik,u1v,u2v;
                local y1m1,y2m1,z1r,z2r,z1,z2,y1,y2;
                local u,um1;
                nveces = t/rows(x1);
                nv = 1;
do while nv le nveces;
        if(nv == 1);
        x1v = x1; x2v = x2;
        else;
        x1v = x1v | x1; x2v = x2v | x2;
        endif;
        nv = nv + 1;
endo;
eps = rndn(t,m)*chol(sigma);
```

```
u10 = 0; u20 = 0;
r11 = r[1,1]; r22 = r[2,2];
ik = 1;
do while ik le t;
        if(ik == 1);
            u1v = r11*u10 + eps[1,1];
            u2v = r22*u20 + eps[1,2];
        else;
            u1v = u1v | r11*u1v[ik-1,1] + eps[ik,1];
            u2v = u2v | r22*u2v[ik-1,1] + eps[ik,2];
        endif;
        ik = ik + 1;
endo;
y1m1 = 0; y2m1 = 0;
ik = 1;
do while ik le t;
        if(ik == 1);
           z1r = 1 ~ y1m1 ~ x1v[1,1];
           z2r = 1 ~ y2m1 ~ x2v[1,1];
           z1 = z1r; z2 = z2r;
           y1 = z1r*beta1 + u1v[1,1];
           y2 = z2r*beta2 + u2v[1,1];
        else;
           z1r = 1 ~ y1[ik-1,1] ~ x1v[ik,1];
           z1 = z1 | z1r;
           z2r = 1 ~ y2[ik-1,1] ~ x2v[ik,1];
           z2 = z2 | z2r;
           y1 = y1 | z1r*beta1 + u1v[ik,1];
           y2 = y2 | z2r*beta2 + u2v[ik,1];
        endif;
        ik = ik + 1;
endo;
" Matriz de varianzas y covarianzas estimada de eps ";
(1/t)*eps'*eps;
```

```
" Matriz de varianzas y covarianzas estimada de u   ";
 (1/t)*(u1v~u2v)'*(u1v~u2v);
u = u1v ~ u2v; um1 = (0|u1v[1:t-1,1])~(0|u2v[1:t-1]);
" Regresión de u sobre um1  ";
 u/um1;
 retp(y1,y2,z1,z2);
endp;
```

Seguidamente se transcribe el programa SPENCER.PRG.

```
/*  Programa spencer.prg  que implementa el procedimiento de estimación para sistemas de ecuaciones aparentamente no relacionadas (SUR), con ajuste parcial y errores autorregresivos AR(1), con matriz no necesariamente diagonal r, desarrollado por Spencer, David E. (1979), Estimation of a Dynamic System of Seemingly Unrelated Regressions with Autoregressive Disturbances, Journal of Econometrics, Vol. 10, pp. 227-241. m es el número de ecuaciones; t el número de observaciones para cada ecuación. El producto m*t arroja como resultado el número total de observaciones.
*/
cls; print; ? ; "   Trabajando... Espere por favor ";
  nrepl = 5;
m = 2;
  t = 60;
ts = hsec;
x1 = {2,0,1,3,9,1,4,9,4,3};
x2 = {3,1,4,5,7,2,3,5,6,8};
r = {0.40 0,0 0.80};
sigma = {1 0.50,0.50 1};
beta1 = {1,0.5,1};
beta2 = {1,0.50,1};
"  Beta1            " beta1';
"  Beta2            " beta2';
  irepl = 1;
do while irepl le nrepl;
        {y1,y2,z1,z2} = genera(x1,x2,m,t,r,sigma,beta1,beta2);
```

```
proc (1) = iv(y,z);
	local y,z,t,iz,btilde;
	t = rows(z);
	iz = ones(t,1) ~ (0|z[1:t-1,3]) ~ z[.,3];
	btilde = inv(iz'z)*iz'y;
	retp(btilde);
endp;

bt1 = iv(y1,z1);
bt2 = iv(y2,z2);
" *** Irepl  " irepl;
" bt1 igual a " bt1';
" bt2 igual a " bt2';
/*
	La siguiente sección se agrega para iterar el cálculo del procedimiento de
	estimación.
*/
nest = 10;
iest = 1;
do while iest le nest;
	if (iest == 1);
		u1t = y1 - z1*bt1;
		u2t = y2 - z2*bt2;
	else;
		bt1 = thetah[1:k1,.];
		bt2 = thetah[k1+1:k1+k2,.];
		u1t = y1 - z1*bt1;
		u2t = y2 - z2*bt2;
	endif;
	utilde = u1t ~ u2t;
	utm1 = zeros(1,m) | utilde[1:t-1,.];
	"Número de condición de utm1  " cond(utm1);
	rtilde = utilde/utm1;
	etilde = utilde - utm1*rtilde;
```

```
            sigtilde = (1/t)*(etilde'*etilde);
            k1 = rows(bt1); k2 = rows(bt2);
            z = (z1 ~ zeros(t,k2)) | (zeros(t,k1) ~ z2);
            y = y1 | y2;
            zm1 = zeros(1,k2+k1)|z[1:t-1,.]|zeros(1,k2+k1)|z[t+1:2*t-1,.];
            h = z - (rtilde'.*. eye(t))*zm1;
            h = h ~ (eye(m) .*. utm1);
            ym1 = zeros(1,1)|y1[1:t-1,.]|zeros(1,1)|y2[1:t-1,.];
            ynew = y - (rtilde' .*. eye(t))*ym1;
            sigi = invpd(sigtilde);
            thetah = invpd(h'(sigi.*.eye(t))*h)*(h'(sigi.*.eye(t)))*ynew;
            thetah = thetah + (zeros(k1+k2,1)|vec(rtilde));
"=============================================================";
            " Iest  " iest;;
            " thetah     " thetah';
            iest = iest + 1;
      endo;
      " Irepl " irepl;
      " thetah  " thetah';
      if (irepl == 1);
            result= thetah;
      else;
      result = result ~ thetah;
      endif;
      irepl = irepl + 1;
endo;
et = (hsec - ts)/100;
lprint " tiempo " et;
lprint " result " result;
lprint " medias " meanc(result')';
lprint " m ximos" maxc(result')';
lprint " m¡nimos" minc(result')';
"    ***  Final del trabajo  *** ";
```

Bibliografía

Abdala, Manuel Angel; Arrufat, José Luis; Colomé, Rinaldo Antonio y Neder, Angel Enrique (1996), Elasticidades de Demanda de Servicio Telefónico Básico en Argentina, *Cuadernos de Economía*, *Pontificia Universidad Católica de Chile*, Año 33, Nº 100, diciembre, págs. 397-424.

Buse, A. (1979), Goodness-of-Fit in the Seemingly Unrelated Regressions Model. A Generalization, *Journal of Econometrics*, Vol. 10, págs. 109-113.

Dhrymes, P.J. and Taylor, J.B. (1976), On an Efficient Two-Step Estimator for Dynamic Simultaneous Equations Models with Autorregresive Errors, *International Economic Review*, Vol 17, June, pp. 362-376.

Fomby, T.B., Hill, R.C. and Johnson, S.R. (1984), *Advanced Econometric Methods*, Springer, New York.

Greene, William H. (1993), *Econometric Analysis. Second Edition*, Macmillan, New York.

Harvey, A.C. (1990), *The Econometric Analysis of Time Series. Second Edition*, Philip Allan, New York.

Hatanaka, Michio (1974), An Efficient Two-Step Estimator for the Dynamic Adjustment Model with Autoregressive Errors, *Journal of Econometrics*, Vol. 2, págs. 199-200.

Hatanaka, Michio (1976), Several Efficient Two-Step Estimators for the Dynamic Simulteaneuos Equations Model with Autoregressive Disturbances, *Journal of Econometrics*, Vol. 4, págs. 189-204.

Hill, R. Carter (1989), *Learning Econometrics Using Gauss - A Computer Handbook to Accompany "Introduction to theTheory and Practice of Econometrics - Second Edition"*,Wiley, New York.

Johnston, J. (1984), *Econometric Methods. Third Edition*, McGraw-Hill, New York.

Judge, G.G, Hill, R. Carter, Griffiths, William E., Luetkepohl, Helmut and Lee,Tsoung-Chao (1988), *Introduction to the Theory and Practice of Econometrics-Second Edition*, Wiley, New York.

Judge, G.G, Griffiths, W.E., Hill, R.Carter, Luetkepohl, Helmut and Lee, Tsoung-Chao (1985), *The Theory and Practice of Econometrics - Second Edition*, Wiley, New York.

Kmenta, Jan (1986), *Elements of Econometrics- Second Edition*, Macmillan, New York.

NERA (1995), *Restructuring of Basic Telephone Service Tariffs*, Statistical Annex to: Final Report for the Ministry of Economy. Republic of Argentina, August, mimeo, London.

Spencer, David E. (1979), Estimation of a Dynamic System of Seemingly Unrelated Regressions with Autoregressive Disturbances, *Journal of Econometrics*, Vol. 10, págs. 227-241.

Spitzer, J.J. (1978), A Monte Carlo Investigation of the Box-Cox Transformation in Small Samples, *Journal of the American Statistical Association*, Vol. 73, págs. 488-495.

Spitzer, J.J (1982), A Primer on Box-Cox Estimation, *Review of Economics and Statistics*,Vol. 64, págs. 307-313.

Spitzer, J.J. (1984), Variance Estimates in Models with the Box-Cox Transformation: Implications for Estimation and Hypothesis Testing, *Review of Economics and Statistics*, Vol. 66, págs. 645-652.

Srivastava, V.K. and Dwivedi, T.D. (1979), Estimation of Seemingly Unrelated Regression Equations. A Brief Survey, *Journal of Econometrics*, Vol. 10, págs. 15-32.

Theil, Henri (1971), *Principles of Econometrics*, Wiley, New York.

* Instituto de Economía y Finanzas, Facultad de Ciencias Económicas, UNC. Av. Valparaíso sin nº, Agencia postal 4, Ciudad Universitaria, (5000) Córdoba. Tel.: (051) 334084, Fax: (051) 334092; *e-mail*: jarrufat@eco.uncor.edu.

Se agradece la valiosa colaboración del Lic. A. Enrique Neder por discusiones mantenidas a lo largo de las distintas etapas del presente trabajo.

[1]Véase, por ejemplo, Judge y otros (1988) o Fomby y otros (1985).

[2]Resulta inmediato comprobar, que si en la fórmula anterior se toma límite para λ que tiende a cero, se obtiene una indeterminación. La aplicación de la Regla de L'Hôpital conduce, en forma inmediata, a la transformación basada en el logaritmo natural de y.

[4] Para más detalles, ver Abdala, Arrufat, Colomé, Neder (1996).