MODELOS TIPO BOX-COX: PROPOSICIÓN DE UNA ESTRATEGIA DE ESTIMACION*

DIETER WUNDER M.

Departamento de Economía, Universidad de Chile

HECTOR GUTIERREZ L.

Departamento de Economía, Universidad de Chile

Abstract:

This paper proposes a strategy for evaluating and estimating Box-Cox models which are widely used for example in estimation of hedonic price equations in urban economics. The proposition has in mind some difficulties arising in practical implementation of Spitzer's [1982] conclusion about the equivalence between results obtained with the maximum likelihood technique, by one hand, and the ones obtained with the iterative and non linear least squared technique, by the other. The strategy is shown estimating a sites and services' s hedonic price equation.

1. Aspectos metodológicos introductorios

Desafortunadamente en economía, tal como en otras ciencias, los fundamentos teóricos por sí solos no son a menudo suficientes para proveernos de una forma funcional única a ser utilizada en el proceso de estimación de modelos, no obstante en algunos casos permite discernir entre tipos de relaciones funcionales que se deben utilizar. La estimación de una ecuación de precios hedónicos provee un claro ejemplo de lo anterior: el modelo desarrollado por Sherwin Rosen [1974] no permite inferir cuál es la

* Se agradecen los comentarios del editor de este Journal y de un árbitro anónimo, que nos sugirieron reorganizar la presentación de este artículo a su presente forma. También se agradecen los comentarios de Martín Paldam a una versión preliminar y la colaboración del Ministerio de la Vivienda chileno. Finalmente se agradece la lealtad y eficiencia de nuestro ayudante Nelson Fuenzalida y la ayuda de Claudio Bonacic en las estimaciones por Máxima Verosimilitud. Mayores detalles se presentan en un Documento de la Serie Investigación a publicar en 1992 en el Departamento de Economía de la Universidad de Chile.

relación funcional exacta a utilizar en el proceso de estimación, aunque entrega argumentos que permiten "descartar" funciones lineales a priori¹. Y si la forma funcional es no lineal, la pregunta es: ¿Cómo puede estimarse dicha relación?

Al respecto, se pueden examinar tres variantes o alternativas que van desde menos a mayores grados de sofisticación econométrica, los que a continuación se describen brevemente: a) Aproximación de serie de potencias; b) Transformaciones linealizantes; y c) Transformaciones de Box-Cox.

a) *Aproximación por serie de potencias*: La primera variante consiste en estimar, en vez de la especificación no lineal directamente, una aproximación de ésta, que puede ser de primer orden (lineal) o de orden mayor (cuadrática, cúbica, etc.). El primer caso puede implementarse vía estimación de múltiples segmentos lineales para distintos rangos de valores de las variables independientes. El segundo caso trata de ajustar a la función no lineal un polinomio de segundo, tercer o mayor orden, dependiendo del grado de aproximación de la expansión de Serie de Taylor que se utilice; este método requiere estimar la variable dependiente en función de las variables independientes en sus valores originales, de las variables independientes elevadas al cuadrado, elevadas al cubo, etc.

Las principales desventajas de estos métodos pueden resumirse en las siguientes: el primero de ellos requiere un importante esfuerzo y además sus resultados son de escasa utilización práctica, sobre todo si la función original es "muy" no lineal, por cuanto requeriría reportar distintas especificaciones estimadas, una para cada rango de valores muestrales. En relación al segundo método, se producen dos problemas: uno de pérdida de grados de libertad, derivado del gran número de variables independientes que aparecen en el lado derecho de la ecuación, y los muy probables problemas de multicolinealidad a que la estimación de dicha especificación puede dar origen; esto último generalmente no constituye un problema si el propósito de la estimación es realizar predicciones (en tanto se considere que el problema de colinealidad está también presente para el rango de valores de las variables independientes que se utilizarán al momento de predecir), aunque sí lo es cuando el propósito básico de la estimación es realizar análisis estructural. Esta variante no se explora intensivamente en este estudio tanto por las razones antes comentadas como por el hecho de que las dos variantes siguientes pueden ser consideradas superiores a ésta.

b) *Transformaciones linealizantes*: Con la segunda variante, de transformaciones linealizantes ("linearizing transformations"), se parte de un modelo que es no lineal en las variables, pero lineal en los parámetros, y mediante algún tipo de transformación algebraica se obtiene un modelo que aparece como lineal, pudiéndose estimar por el simple método de Mínimos Cuadrados Ordinarios, MCO: ejemplos típicos son las transformaciones log-log (logaritmo-logaritmo), las semi-log (logaritmo-lineal) y la logaritmo-recíproca. El problema básico con este tipo de métodos es que la elección de la forma no lineal específica es *ad hoc* o arbitraria; pareciera que la elección preferida para precios hedónicos es la semi-log, pero no existe mucha justificación para esta elección particular². El problema a que conlleva ajustar distintas transformaciones, sobre todo cuando los modelos son no anidados ("nonnested"), es que no existe un criterio fácil de implementar para elegir entre ellos³.

c) *Transformaciones de Box-Cox*: La tercera variante de modelos a estimar cae en el contexto de estimaciones no lineales directas; bajo esta modalidad no es necesaria la

preselección de transformaciones *a priori*, por cuanto es posible estimar sus parámetros simultáneamente con los de las variables independientes presentes en la ecuación. En este contexto, uno de las técnicas más frecuentemente utilizadas en la estimación de ecuaciones hedónicas es la denominada transformación de Box-Cox [1964] siguiente:

$$y_i^{(\lambda_i)} = \beta_0 + \beta_1 X_{1i}^{(\lambda_1)} + \beta_2 X_{2i}^{(\lambda_2)} + \dots + \beta_k X_{ki}^{(\lambda_k)} + \epsilon_i \quad (1)$$

para todo $i = 1, 2, \dots, n$

$$\text{en donde: } z^{(\lambda)} = \begin{cases} \frac{(z^\lambda - 1)}{\lambda} & \text{si } \lambda \neq 0 \\ \ln z & \text{si } \lambda = 0 \end{cases} \quad \text{para } z \geq 0; \quad z = X_i, y \quad (2)$$

con:

y = variable dependiente o a explicar.

$X_1, X_2, X_3, \dots, X_k$ = variables independientes (que pueden por supuesto incluir variables dummies o dicotómicas, en cuyo caso la transformación no está definida y , por ende, no se aplica).

ϵ = término de error aleatorio supuesto normalmente distribuido, con media igual a cero y varianza constante.

λ_y = parámetro de transformación de la variable dependiente, a ser estimado.

$\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_k$ = parámetros de transformación de las variables independientes o características, a ser estimados.

$\beta_0, \beta_1, \beta_2, \dots, \beta_k$ = parámetros que acompañan a las variables independientes, a ser estimados también.

(El subíndice i representa las n distintas observaciones disponibles en la muestra.)

El método de estimación a utilizar puede ser el de Máxima Verosimilitud (MV), el de Mínimos Cuadrados no Lineales (MCNL) o el de Mínimos Cuadrados Iterativos (MCI); al respecto Spitzer [1982] provee una excelente discusión respecto de estas tres opciones. En este artículo se examina esta variante de estimación no lineal.

2. Proposición de una estrategia de estimación

La ecuación (1) recién presentada se puede denominar como transformación de Box-Cox generalizada por cuanto las transformaciones presentadas en (2) no sólo aplican a la variable dependiente (y), sino que también a las distintas variables independientes ($X_1, X_2, X_3, \dots, X_k$) continuas⁴. Versiones más sencillas sobre la especificación (1) dan origen a una variedad de transformaciones de Box-Cox; dichas versiones pueden denominarse restringidas por cuanto presuponen valores específicos (normalmente 1, 0, $\delta - 1$) para los distintos λ 's (del lado derecho e izquierdo). Una de esas versiones sencillas corresponde a aquella que se origina al suponer que todos los parámetros de transformación de las variables independientes (es decir, $\lambda_1, \lambda_2, \dots, \lambda_k$) son iguales a uno

o, lo que es lo mismo, suponer que las variables independientes aparecen en el lado derecho de la ecuación en forma lineal. A modo de ejemplo, es precisamente este tipo de especificación la que utiliza el BID en su trabajo en Bogotá, Colombia, y que a continuación se ilustra (ver BID [1988]):

$$\frac{y^i - 1}{\lambda} = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + \epsilon_i \quad (3)$$

para todo $i = 1, 2, \dots, n$

Sin duda el grado de complejidad envuelto en las estimaciones de formas funcionales como las (1) y (3) no es comparable; lo anterior no sólo porque el número de coeficientes a estimar difiere sustancialmente, sino que también porque la forma funcional a optimizar es muy distinta, todo lo cual hace más prolongado el proceso de estimación de los parámetros. Asimismo, la significancia estadística de los parámetros puede diferir sustancialmente y también la dificultad en su interpretación, sobre todo cuando se agregan efectos interactivos⁵.

Independiente de lo anterior y con el propósito de enfrentar el problema de la estimación desde una perspectiva lo más amplia posible (o, si se quiere, lo menos restringido), en esta sección se propone una estrategia supuestamente general para estimaciones no lineales, partiendo de la más sencilla denotada como especificación (3), a la más compleja o especificación (1); asimismo se proponen otras especificaciones denominadas "intermedias" con el propósito de poder llegar en definitiva a la "mejor" forma funcional vía la realización de tests de razón de verosimilitudes para modelos anidados⁶.

Pero antes de presentar la estrategia de estimación, conviene tener presente un aspecto adicional y que es inherente a la estimación de especificaciones no lineales en los parámetros y que en gran medida justifica la estimación de transformaciones específicas más sencillas, previo a la estimación de modelos más generales (o menos restringidos); nos referimos a la selección de los denominados valores de partida de los parámetros ("starting values") para iniciar el proceso de iteración camino a la convergencia.

La estrategia de estimación propuesta contempla las siguientes etapas:

Etapla 1: En la primera etapa se regresionan por MCO transformaciones lineales específicas del tipo lineal-lineal, logaritmo-lineal y logaritmo-logaritmo. Los coeficientes estimados por estos métodos serán sesgados y sus varianzas estarán fuertemente subestimadas, invalidando la realización de test de hipótesis y, en particular, la comparación entre los R² que se obtengan⁷, tal como prueba Spitzer, *op. cit.*; todo esto ocurre porque se adoptan exógenamente valores supuestos para los parámetros lambda. No obstante lo anterior, el propósito inicial es sólo auscultar valores de partida para las estimaciones de la forma no lineal más general, además del deseado "precalentamiento" de partida.

Etapla 2: Si bien la etapa anterior es de utilidad en la determinación de valores de partida para la aplicación de métodos de estimación no lineales, dicha utilidad es en cierta medida restrictiva; esto por cuanto las transformaciones de la etapa anterior obligan a que los λ 's tomen valores puntuales, generalmente 0 ó 1 (y, por ende, condi-

cionan los β 's que de allí se derivan). Lo anterior puede considerarse que restringe demasiado el rango de valores de los λ 's, pudiendo existir por ejemplo "mejores ajustes" para valores como por ejemplo de 0,3 (valores entre -2,0 y +2,0 no son extraños de encontrar en estimaciones empíricas). Una alternativa para solucionar el problema anterior consiste en aplicar el denominado método de Mínimos Cuadrados Iterativos, MCI. Si bien este método pertenece a los que estiman la forma no lineal indirectamente, pues descansa en estimaciones por MCO, es más sencilla que las no lineales y puede demostrarse que sus resultados son equivalentes a los de ellos, previo ciertas transformaciones, tal como lo prueba Spitzer. El método consiste simplemente en darse valores para λ_i en el rango desde -2,0 a +2,0, en un "grid search", por ejemplo considerando valores -2,0, -1,9, ..., +1,9, +2,0, y regresionar por MCO para cada uno de ellos.

No obstante la sencillez de esta variante iterativa, conviene tener presente que las desviaciones estándar que se obtienen para los parámetros estimados nuevamente aparecen apreciablemente subestimadas, invalidando por ende los test de hipótesis (al igual que en el caso anterior), a menos que se esté dispuesto a aceptar que el valor identificado (o elegido) para el parámetro de transformación de la variable dependiente (λ_0) sea el verdadero y, por ende, no esté sujeto a errores en su estimación, lo cual obviamente es inaceptable en un contexto estadístico.

Etapla 3: Aquí se propone la estimación del modelo (1) con la técnica MV a partir de los resultados obtenidos en la etapa anterior. Se adopta el supuesto que los errores ϵ_i se distribuyen "aproximadamente" Normal; esto último, citando a Armenta y Powell, *op. cit.*, de página 352, porque "el término de error en el modelo (1) estrictamente no puede estar distribuido Normal, excepto cuando λ es igual a cero... y... el modelo que se intenta estimar no es uno estadístico". Esta situación es típica con precios hedónicos y se presenta porque éstos son estrictamente positivos, lo que conduce a que $y_i^{(\lambda)}$ también lo sean, con lo que el término estocástico ϵ_i a su vez no puede ser negativo, lo que contradice el supuesto de normalidad y conduce a lo más a una Normal truncada; la excepción es cuando λ_i es cero porque $y_i^{(0)}$ tiende al logaritmo natural de y_i que si podría adoptar teóricamente valores negativos, tal como lo señalan Armenta y Powell.

La técnica de estimación a utilizar consiste en identificar la función de verosimilitud de la observaciones de la muestra, la cual es maximizada con respecto a los parámetros de interés; Spitzer deduce dicha función de verosimilitud. Aquí es necesario recurrir a algoritmos computacionales de estimación que a su vez requieren de valores de partida para los parámetros a estimar, los que se obtienen de la segunda etapa.

Un aspecto que no debe desestimarse es la dificultad para implementar la técnica, además que conviene estar preparado para la contingencia de que el algoritmo no converja. Ante esto, puede optarse por utilizar los resultados de Spitzer: repetir el grid search de la etapa anterior pero después de dividir cada una de las observaciones y_i por su media geométrica muestral y escoger aquella regresión de mínimo SSR; de esa manera se identifica el λ_i de MV⁸.

Etapla 4: El artículo de Spitzer muestra que existen cuatro maneras de obtener los mismos resultados: 1) MV; 2) MV concentrada; 3) MCI; y 4) MCNL. Aquí se explora la cuarta variante de mínimos cuadrados no lineales.

El punto de partida de esta etapa es el reconocimiento de que la implementación del algoritmo MCNL en paquetes estadísticos (hasta donde conocemos) requieren reescribir la ecuación del modelo, dejando "Y" de la siguiente forma:

$$Y_i = \left\{ \left(\beta_1 + \beta_2 X_{2i}^{(\lambda_2)} + \dots + \beta_k X_{ki}^{(\lambda_k)} \right) \lambda_i + 1 \right\}^{1/\lambda_i} + H_i \quad (4)$$

para todo $i = 1, 2, \dots, n$

donde: $Z^{\lambda} = \frac{Z^{\lambda} - 1}{\lambda}$ si $\lambda \neq 0$ (en cuyo caso $Z(\lambda)$ converge a $\ln(Z)$), para $Z =$ cualquier X_k continuo, $y \{Z\lambda_i + 1\}^{1/\lambda_i}$ converge a $\text{EXP}\{Z\}$ para $\lambda_i = 0$.

Aquí es crucial tener presente que el modelo (4) es *distinto* del modelo (1), por cuanto difieren en el supuesto acerca del comportamiento del término de error, una sutil pero reveladora diferencia. Más aún, como no es posible implementar en los paquetes estadísticos el modelo (1) por MCNL, sino el (4) con la misma técnica, se debe estar preparado para obtener resultados distintos a los de las tres etapas anteriores, aparte de que en este mismo instante nos salimos del contexto estudiado por Spitzer.

El supuesto que se adopta aquí es que el término de error de la ecuación (4) se distribuye Normal con media cero y varianza constante y , por ende, y_i se distribuye de similiar forma⁹. En cambio, en el modelo (1), el supuesto que prevalece es que el término de error de esa ecuación, ϵ_i , se distribuye Normal y consecuentemente $y_i^{(\lambda_i)}$ se distribuye Normal; pero no así y_i , en general, cuya función de densidad de probabilidad no si quiera existe.

En definitiva, con esta etapa se obtienen estimaciones que no son directamente comparables con las anteriores; en particular, los modelos (1) y (4) no son anidados por lo que su comparación requerirá de algún trabajo adicional.

Si bien el lector puede preguntarse acerca de la razón de plantear un modelo alternativo y distinto al modelo (1), la dificultad para saber *a priori* acerca de cuál es el verdadero modelo hace recomendable su planteamiento.

Etapla 5: En esta etapa se comparan los resultados de estimar los modelos (1) y (4). Ante la imposibilidad de realizar un test de especificación anidado que permita discernir por uno de ellos, éstos se comparan por medio de un indicador de bondad de ajuste ("goodness of fit").

El cálculo propuesto parte reconociendo que los residuos que se estimen para el modelo (1) son en realidad desviaciones en torno a $y_i^{(\lambda_i)}$ y no de y_i , tal como lo son los residuos que se estimen del modelo (4). Luego, se propone recalcular los residuos del modelo (1), transformándolos de modo que correspondan a desviaciones en torno a y_i ; ello pone a ambos modelos sobre iguales bases para comparar.

La transformación implícita para el modelo (1) es una simplificación del método de transformación propuesto por Miller, *op. cit.*, y utilizado en el reporte BID [1988], *op. cit.*, para funciones del tipo $y_i^{(\lambda_i)} = f(X_i, \beta_i, \lambda_i) + \epsilon_i$, donde N es un número entero positivo y corresponde al valor aproximado más cercano de λ_i que produce un número entero. Miller, *op. cit.*, demuestra que el factor de corrección es un término que debe sumarse a $f(X_i, \beta_i, \lambda_i)^{\lambda_i}$, sin embargo, en nuestro caso optamos por aproximar el factor de corrección a un valor igual a la media muestral de los errores estimados del modelo (1), equivalentes a $y - f(X_i, \beta_i, \lambda_i) \cdot \lambda_i + 1$ ¹⁰; ello tiene un sentido obvio: que los residuos transformados del modelo (1) tengan media muestral nula, una de las ecuaciones normales de cualquier estimación por mínimos cuadrados. Por otra parte, el caso de $\lambda_i = 0$, que no puede escribirse de la forma $1/N$, es el conocido caso en que y_i se distri-

buye Lognormal si $\ln(Y_i)$ se distribuye Normal; para la transformación que corresponde en este caso, ver Miller, *op. cit.* 10

La retransformación de residuos del modelo (1) que se ha propuesto permite, por una parte, calcular SSR ("sum of squared residual") que son comparables con el SSR que se obtiene directamente del modelo (4) y , por otra, calcular R^2 también comparables; con ello ya se tiene la medida de bondad de ajuste para elegir. (Como "corolario" se destaca que esa manera de comparar es aplicable si se quisiera elegir entre modelos no anidados en otras etapas, tales como los de las etapas 1 y 2.)

Etapla 6: Un aspecto que resulta ser de interés, una vez elegido un modelo en la etapa 5, consiste en testear si la especificación general puede ser bien representada por una especificación más sencilla, como por ejemplo la Lineal-Lineal o la Log-Lineal en vez de Box-Cox "a ambos lados", por la facilidad para interpretarse dicha especificación funcional. Para poder discernir entre la especificación general y la restringida, se puede aplicar un test de razón de verosimilitudes¹¹.

3. Aplicación de la estrategia de estimación: el caso de precios hedónicos de lotes con servicios

Para ilustrar la aplicación de la estrategia de estimación sugerida, en esta sección se presentan los resultados de seguirla en un caso particular. Pero antes de pasar a las seis etapas de estimación, se presentan algunas referencias que permitirán familiarizarse con el ejemplo.

a) **Elementos del modelo de formación de precios hedónicos:** Como se estima una ecuación hedónica, a continuación se hace una breve referencia al modelo. La potencia de la teoría de precios hedónicos para hacer estudios de economía aplicada es apreciable, aun cuando ello no sea trivial. Desde el punto de vista de las aplicaciones empíricas existen al menos dos versiones o etapas posibles de elejear: en la primera, o etapa más simple, se estima una relación funcional entre el valor de mercado y sus atributos, con el propósito de determinar la contribución de cada uno de ellos al precio del bien; con la segunda etapa se identifican los parámetros de cada una de las curvas de demanda y oferta por cada uno de los atributos que posee el bien. La primera etapa es la *forma reducida*, porque los parámetros que se estiman reflejan la influencia combinada de oferta y de demanda en equilibrio, en tanto la segunda etapa trata de separar dichas influencias y se denomina la *forma estructural*.

En un contexto en donde el interés es realizar análisis de bienestar, la estimación de la forma estructural es la adecuada pues permite identificar qué cambios hacer a un bien tal como un lote con servicios para lograr mayor bienestar social; pero en un contexto en donde el interés es únicamente estimar precios de mercado y no disponibilidad a pagar, sólo basta estimar la forma reducida de la ecuación hedónica de precios. Por otro lado, la correcta identificación de la forma estructural puede ser una tarea poco extensa, tal como lo prueba la experiencia del BID [1987]¹², por lo que muchas veces no queda más que conformarse con estimaciones de la forma reducida, adecuadamente complementadas con la utilización de otros métodos tal como el de valuación contin-

Si a lo anterior se agrega la experiencia chilena de estimar ecuaciones hedónicas, que nos advierte sobre las dificultades incluso de hacer estimaciones aceptables de la forma reducida al tratarse de viviendas en barrios pobres¹³, podrá entenderse la cautela que debe tenerse. En definitiva, se tiene la meta de estimar la forma reducida de una ecuación hedónica de precios de mercado para lotes con servicios.

b) *Identificación de la variable precio*: Para la obtención de la información básica con la que identificar la ecuación hedónica se consideró la toma de una encuesta especialmente diseñada. Se eligió a Santiago de Chile por facilidades de acceso a los hogares a encuestar y se optó por partir considerando a 1.200 propiedades estadísticamente representativas; ello permitió tener información confiable de 1.095 casos, por 105 donde no se ubicó personas idóneas para responder, ya que se muestró sin reemplazo para evitar sesgos; la encuesta se tomó en julio de 1991. De varias alternativas de precio (valor del lote con servicios sin vivienda vs. con vivienda, valor del activo vs. de arriendo, y precio de demanda vs. de oferta vs. de mercado) se eligió como variable dependiente al precio de mercado estimado por los encuestados para un lote con servicios sin la vivienda. (Ver Anexo 1)

c) *Identificación de variables independientes*: La elección de variables independientes consideró tanto características del lote con servicios como otras del entorno. Ello condujo a su vez a la definición de algunas variables dummy, tanto dicotómicas como polidíctomas; alternativamente, en algunos casos se optó por indizar variables. (Ver Anexo 2)

d) *Resultados de aplicar la estrategia de estimación*: A continuación se muestran los resultados obtenidos al aplicar la estrategia de estimación al caso de lotes con servicios.

A. Etapa 1

En la primera etapa se regresionan transformaciones lineales específicas del tipo lin-lin, log-lin, log-log y log-recíproco: los principales resultados de estas regresiones se reportan en el Cuadro 1. Recuérdese que los R² que se obtienen para las distintas especificaciones no son comparables entre sí por cuanto la variable dependiente sufre distintas transformaciones; sin embargo, en el cuadro se presentan R² corregidos que sí son comparables. La diferencia, y la no comparabilidad, se origina en que R² es la fracción de la varianza de y^(a) que es explicada y que reportan automáticamente los paquetes estadísticos; en cambio, los R² corregidos son la fracción explicada de la varianza de y, lo que los hace comparables; de acuerdo al procedimiento de transformación que se describe en la etapa 5 de la sección con la estrategia de estimación (y que es una aplicación del "corolario" descrito ahí); la inclusión de los R² corregidos es sólo con fines ilustrativos y no para testeo.

CUADRO 1
TRANSFORMACIONES LINEALIZANTES DEL MODELO (1)

Variable	Lin-Lin $\lambda_1=1 \lambda_2=1$	Log-Lin $\lambda_1=0 \lambda_2=1$	Log-Log $\lambda_1=0 \lambda_2=0$	Log-Recip. $\lambda_1=0 \lambda_2=1$
C	124*E+4 (225*E+3)	139*E-1 (416*E-3)	204*E-1 (177*E-2)	113*E-1 (371*E-3)
ALUMPUBL	268*E+2 (234*E+2)	727*E-4 (425*E-4)	111*E-3 (763*E-4)	-121*E-3 (125*E-3)
ANTGCA1	745*E 0 (749*E 0)	210*E-5 (136*E-5)	563*E-4 (531*E-4)	354*E-3 (125*E-2)
CALCAS2	509*E+2 (292*E+2)	134*E-3 (532*E-4)	282*E-3 (113*E-3)	-534*E-3 (220*E-3)
CGVALIME	357*E 0 (322*E 0)	112*E-5 (586*E-6)	390*E-4 (237*E-4)	-536*E-4 (217*E-3)
CGVCENTR	-691*E+1 (103*E+1)	-140*E-4 (187*E-5)	-228*E-2 (320*E-3)	400*E 0 (513*E-1)
CGVTRAB	-176*E-1 (165*E 0)	-614*E-7 (300*E-6)	-406*E-4 (345*E-3)	171*E-2 (858*E-3)
DBASURA	-579*E+2 (327*E+2)	-364*E-4 (595*E-4)	-317*E-4 (594*E-4)	-283*E-4 (583*E-4)
DCALLE	423*E+2 (385*E+2)	661*E-4 (701*E-4)	862*E-4 (718*E-4)	155*E-3 (705*E-4)
DITTIPO	920*E+2 (344*E+2)	127*E-3 (625*E-4)	137*E-3 (624*E-4)	150*E-3 (628*E-4)
DVERBDA	-687*E+1 (42*E+2)	119*E-4 (621*E-4)	108*E-4 (620*E-4)	159*E-4 (623*E-4)
SUSTEST	169*E+1 (509*E 0)	495*E-5 (925*E-6)	716*E-3 (129*E-3)	-990*E-1 (161*E-1)
TPOCCOM11	-233*E+1 (200*E+1)	-754*E-5 (364*E-5)	-563*E-4 (838*E-4)	-398*E-3 (918*E-3)
TPOBSCUE	195*E+1 (120*E+1)	436*E-5 (217*E-5)	931*E-4 (431*E-4)	-455*E-3 (332*E-3)
TPOPARAD	-857*E+1 (476*E+1)	-945*E-5 (865*E-5)	-540*E-4 (443*E-4)	164*E-3 (117*E-3)
TPOPOSTA	128*E+1 (903*E 0)	264*E-5 (164*E-5)	833*E-4 (508*E-4)	-133*E-3 (609*E-3)
TPOSALUID	-823*E 0 (117*E+1)	120*E-6 (213*E-5)	-549*E-5 (448*E-4)	-927*E-4 (395*E-3)
TPOTELDI	-208*E 0 (229*E+1)	-687*E-5 (416*E-5)	-935*E-4 (368*E-4)	235*E-3 (102*E-3)
R squared corregido	0,188861	0,150420	0,147810	0,144340
R squared	0,188861	0,237175	0,273661	0,260841
S.E. of Regression	431684,1	0,785043	0,784780	0,791675
Log likelihood	-12442,05	-1,008,942	-1,008,652	-1,016,219
Mean of dep. var.	509153,8	12,76276	12,76276	12,76276
N° observ.	865	865	865	865

a/ Variable independiente en logaritmo.

b/ Variable independiente como recíproco.

Nota: Ver en Anexo 2 la definición y descripción de las variables. Los valores entre paréntesis corresponden a desviaciones estándares de los coeficientes. Las que están subscritas de acuerdo a lo indicado en el texto. El número de observaciones corresponde a aquellas en que existía información de todas las variables independientes de las 1.095 disponibles, por no respuesta de algunos encuestados.

B. Etapa 2

Siguiendo la costumbre de la literatura, los resultados de esta etapa que se reportan en Cuadro 2 corresponden al valor de λ_y que maximiza el R^2 , aunque conviene recordar que estos R^2 no son comparables entre sí; esto es especialmente relevante en este contexto por cuanto algunos autores reportan el modelo para el λ_y que maximiza el R^2 como el "mejor" de entre todos los valores alternativos para λ_y utilizados en el "grid search". Lo anterior presume que el criterio de máximo R^2 es conveniente, lo cual es incorrecto. Con fines ilustrativos, en el cuadro también se reportan los resultados del "grid search" para un valor de $\lambda_y=0,979$ que más adelante, en la etapa 4, se obtiene como coeficiente de regresión al estimar el modelo (4) por MCNL.

Para apreciar lo erróneo que puede ser el comparar modelos con distintas transformaciones para la variable dependiente (distintos λ_y) en base a los R^2 , se calcularon coeficientes de bondad de ajuste para ambos casos a partir de predicciones para "y" de cada modelo (en vez de para $y^{(\lambda_y)}$), obteniendo que la SSR (suma de residuos al cuadrado) para el caso de $\lambda_y=0,1$ (SSR=1,6243*10⁴) arroja valores superiores a los obtenidos para el modelo con $\lambda_y=0,9785943$ (SSR=1,578413*10⁴).

Derivado de lo anterior es posible también calcular los R^2 corregidos (que sí son comparables) obteniendo R^2 corregido = 0,16526 para $\lambda_y=0,1$ y un R^2 corregido de 0,13885 para $\lambda_y=0,98$; esto es, se obtiene un mayor R^2 corregido para $\lambda_y=0,98$, a pesar que el mayor R^2 (sin corregir) se obtuvo para $\lambda_y=0,1$.

Finalmente se recuerda que las desviaciones estándares de los coeficientes que se obtienen y reportan en el Cuadro 2 están apreciablemente subestimadas al adoptarse valores particulares para λ_y .

C. Etapa 3

Se reporta que las estimaciones por MV fueron exitosas en aproximar valores para los diversos parámetros envueltos en la especificación funcional (1); pero no fue posible obtener la matriz de varianzas y covarianzas de dichos coeficientes por cuanto la batería de algoritmos utilizados no fue capaz de entregar el inverso del Hessiano involucrado o, en otros términos, la matriz de varianzas y covarianzas de los parámetros, lo que imposibilita la inferencia sobre la significancia estadística de cualquier parámetro individual; no obstante, aún es posible efectuar algún tipo de inferencias a nivel global utilizando el valor reportado para el logaritmo de la función de verosimilitud evaluada en el máximo.

La matriz de varianzas y covarianzas puede ser aproximada por vías "indirectas", como por ejemplo el procedimiento sugerido por Berndt, Hall, Hall y Hausmanm descrito por R.E. Quandt [1984]: pág. 734, pero ello requiere de un esfuerzo considerable. Se adoptó la estrategia de realizar dichos cálculos sólo si se concluyese que se había identificado un modelo superior a otro, pues en caso contrario sería inofensivo. Sin embargo, y a pesar que no fue necesario hacerlo, tal como se mostrará en etapas sucesivas, la matriz de varianzas y covarianzas fue estimada para los dos casos mostrados en el Cuadro 3 que reporta los resultados de esta etapa (tal vez con la idea que uno debe hacerlo una vez en la vida y ésta era nuestra oportunidad).

En el Cuadro 3 no sólo se reporta el modelo general (1), sino que también sus dos versiones restringidas equivalentes que involucran suponer que los λ_x son iguales a uno y a cero. Destaca que el valor obtenido para λ_y retona valores comparables con los re-

CUADRO 2

MINIMOS CUADRADOS ITERATIVOS DEL MODELO (1)

Variable	Box-Cox-Lin $\lambda_x=1$	Variable	Box-Cox-Log $\lambda_x=0$	Variable	Box-Cox-Lin $\lambda_x=1$
LAMBDA	0.10 (n.d.)	LAMBDA	0.10 (n.d.)	LAMBDA	0.979 (n.d.)
C	303*E-1 (148*E-2)	C	541*E-1 (629*E-2)	C	935*E+3 (170*E+3)
ANTIGCA1	755*E-5 (484*E-5)	LNANTIG	200*E-3 (189*E-3)	ANTIGCA1	568*E0 (58*E0)
CALCAS2	463*E-3 (189*E-3)	LNALCAS	965*E-3 (403*E-3)	CALCAS2	383*E+2 (218*E+2)
SUSTEST	175*E-4 (329*E-5)	LNUSPST	253*E-2 (457*E-3)	SUSTEST	128*E-1 (379*E0)
DBASURA	-145*E-3 (212*E-3)	DBASURA	-126*E-3 (211*E-3)	DBASURA	-281*E+2 (244*E+2)
DITTPPO	470*E-3 (222*E-3)	DITTPPO	508*E-3 (222*E-3)	DITTPPO	683*E+2 (256*E+2)
TPOPARRAD	-357*E-4 (308*E-4)	LNTPPAR	-213*E-3 (158*E-3)	TPOPARRAD	-634*E+1 (354*E+1)
TPOTELDI	-226*E-4 (148*E-4)	LNTPOTEL	-307*E-3 (131*E-3)	TPOTELDI	-212*E0 (171*E+1)
CGVTRAB	-238*E-6 (107*E-5)	LNCGVTRA	-151*E-3 (123*E-3)	CGVTRAB	-140*E-1 (123*E0)
CGVALIME	386*E-5 (208*E-5)	LNCGVALI	136*E-3 (843*E-4)	CGVALIME	270*E0 (240*E0)
TPOESCUE	155*E-4 (773*E-5)	LNTPOESC	327*E-3 (153*E-3)	TPOESCUE	146*E+1 (890*E0)
TPOSALUD	113*E-6 (758*E-5)	LNTPOSAL	-251*E-4 (159*E-3)	TPOSALUD	-595*E0 (873*E0)
CGVCENTR	-509*E-4 (666*E-5)	LNCGVCTR	-829*E-2 (114*E-2)	CGVCENTR	-519*E-1 (767*E0)
TPOPOSTA	933*E-5 (584*E-5)	LNTPOPOS	292*E-3 (181*E-3)	TPOPOSTA	957*E0 (672*E0)
TPPOCOMI	-267*E-4 (129*E-4)	LNTPOCOM	-192*E-3 (298*E-3)	TPPOCOMI	-178*E-1 (149*E-1)
DCALLE	236*E-3 (249*E-3)	DCALLE	310*E-3 (255*E-3)	DCALLE	314*E+2 (287*E+2)
DVEREDA	291*E-4 (221*E-3)	DVEREDA	255*E-4 (221*E-3)	DVEREDA	-523*E-1 (255*E-2)
ALUMPUBL	254*E-3 (151*E-3)	LNALUDM	384*E-3 (271*E-3)	ALUMPUBL	203*E+2 (170*E+2)
R^2 corregido	0.165260	R^2 corregido	0.1647048	R^2 corregido	0.18885
R squared	0.274838	R squared	0.274978	R squared	0.19206
S.E. of Regress.	2.790734	S.E. of Regress.	2.790734	S.E. of Regress.	321620.8
Log likelihood	-2106.124	Log likelihood	-2106.041	Log likelihood	-12187.5
Mean of dep. var.	25.9805	Mean of dep. var.	25.9805	Mean of dep. var.	389926.7
Nº observ.	865	Nº observ.	865	Nº observ.	865

Nota: Ver en Anexo 2 la definición y descripción de las variables. Los valores entre paréntesis corresponden a las desviaciones estándares de los coeficientes, las que están subestimadas de acuerdo a lo indicado en el texto; n.d. = no disponible, puesto que se supone λ_x es exógeno al modelo. El número de observaciones corresponde a aquellas en que exista información de todas las variables independientes de las 1.095 disponibles, por no respuesta de algunos encuestados.

CUADRO 3

ESTIMACION DEL MODELO (1) POR MÁXIMA VEROSIMILITUD

Variable	Box-Cox General	Variable	Box-Cox-Lin λ_{λ}^* s=1	Variable	Box-Cox-Log λ_{λ}^* s=0
VALESST LAMBDA	115*E-3	VALESST LAMBDA	112*E-3 (299*E-4) 337*E-1 (914*E-2)	VALESST LAMBDA	111*E-3 (306*E-4) 606*E-1 (215*E-1)
C	-347*E-4	C		C	
ANTGCAL1 LAMBDA	878*E-5 121*E-2	ANTGCAL1 LAMBDA	880*E-5 (669*E-5) 540*E-3	LNANTG	(241*E-3) (210*E-2)
CALCAS2 LAMBDA	575*E-4 396*E-2	CALCAS2 LAMBDA	(283*E-3) 204*E-4 (858*E-5)	LNCALCAS	110*E-2 (385*E-3)
SUSITEST LAMBDA	347*E-4 896*E-3	SUSITEST LAMBDA	(273*E-3) 553*E-3 (343*E-3)	LNSUPST	290*E-2 (123*E-2)
DBASURA LAMBDA	-590*E-4	DBASURA LAMBDA	(173*E-3) 275*E-3 (314*E-3)	DBASURA	-147*E-3 (259*E-3)
DITTPPO	612*E-3	DITTPPO	553*E-3 (343*E-3)	DITTPPO	(355*E-2) 353*E-3
DCALLE	368*E-3	DCALLE	275*E-3 (314*E-3)	DCALLE	353*E-3 (325*E-3)
DVEREDA	848*E-8	DVEREDA	327*E-4 (260*E-3) -423*E-4	DVEREDA	265*E-4 (255*E-3) -248*E-3
TPOPARAD LAMBDA	-459*E-3 -475*E-3	TPOPARAD LAMBDA	(444*E-4) (260*E-3) -202*E-4	LNTPOPAR	(234*E-3) (348*E-3) -179*E-3
TPOTELDI LAMBDA	-544*E-3 -204*E-3	TPOTELDI LAMBDA	(202*E-4) -384*E-6 (128*E-5)	LNTPOTEL	(188*E-3) (153*E-3) 154*E-3
CGVTRAB LAMBDA	-102*E-1 142*E-2	CGVTRAB LAMBDA	(862*E-5) 448*E-5 (118*E-5)	LNGCVTR	-179*E-3 (153*E-3) 154*E-3
CGVALIME LAMBDA	741*E-4 264*E-3	CGVALIME LAMBDA	184*E-4 (118*E-4) 692*E-7	LNGCVALLI	383*E-3 (233*E-3) -298*E-4
TPOESCUE LAMBDA	112*E-3 -116*E-2	TPOESCUE LAMBDA	(118*E-4) (862*E-5) -597*E-4	LNTPOESC	383*E-3 (233*E-3) -298*E-4
TPOSALUD LAMBDA	221*E-3 -112*E-2	TPOSALUD LAMBDA	(862*E-5) -597*E-4 (248*E-2)	LNTPOSAL	(185*E-3) (408*E-2) 333*E-3
CGVCENTR LAMBDA	929*E-3 533*E-3	CGVCENTR LAMBDA	(794*E-5) 109*E-4 (794*E-5)	LNGVCEN	-953*E-2 (408*E-2) 333*E-3
TPOPOSTA LAMBDA	884*E-3 196*E-1	TPOPOSTA LAMBDA	(215*E-4) -314*E-4 (215*E-4)	LNTPOPOS	(251*E-3) -222*E-3 (339*E-3)
TPOCOMI LAMBDA	150*E-2	TPOCOMI LAMBDA		LNTPOCOM	
ALUMPUBL LAMBDA	453*E-10 175*E-1	ALUMPUBL LAMBDA	298*E-3 (199*E-3)	LNALUM	439*E-3 (328*E-3)
R squared correctido	0.171470	R squared correctido	0.165820	R squared correctido	0.16594
Log likelihood	-12035.5	Log likelihood	10.44	Log likelihood	10.0
Mean of dep. var.	509153.8	Mean of dep. var.	509153.8	Mean of dep. var.	509153.8
Nº observ.	865	Nº observ.	865	Nº observ.	865

Nota: Ver en Anexo 2 la definición y descripción de las variables. Los valores entre paréntesis corresponden a las desviaciones estándares de los coeficientes, aproximados de acuerdo al método sugerido por Berndt, Hall, Hall and Hausman. Ver R. Quandt, *op. cit.*; éstas están subestimadas de acuerdo a lo indicado en el texto. Las desviaciones estándares correspondientes a los parámetros del modelo general no se pudieron obtener por problemas de convergencia. El número de observaciones corresponde a aquellas en que existió información de todas las variables independientes de las 1,095 disponibles, por no respuesta de algunos encuestados.

portados en la segunda y cuarta columna del Cuadro 2: esto no debiera causar sorpresa por cuanto los resultados reportados allí son versiones restringidas del modelo que nos preocupa en esta etapa; sin embargo, los valores no tienen por qué coincidir y su cercanía tiene algo de casual porque lo que se reporta en Cuadro 2 no corresponde al λ_{λ} que minimiza el SSR, sino el que maximiza el R². Note que la comparación de los resultados reportados en ambos cuadros es posible en forma directa para los parámetros β 's, λ 's y σ^2 ; pero no lo es de la misma forma para la función de verosimilitud para lo cual se requirieron cálculos adicionales pues con la técnica de MV se obtiene la función de verosimilitud para los γ , y en el "grid search" para $\gamma_1^{(0)}$.

Una mirada a los resultados anteriores podría llevar al lector a inferir que la calidad de las estimaciones presentadas aquí es muy inferior a los resultados obtenidos mediante la técnica de grid search reportados en el Cuadro 2, especialmente en lo referido a los valores para las desviaciones estándares de los parámetros y los valores para los correctivos R². En este sentido, y como ya ha sido mencionado en etapas anteriores, la comparación directa entre R² de ambos resultados no es correcta. En relación a la importante alza mostrada en las desviaciones estándar de los parámetros (o, equivalentemente, la abrupción caía en los test t de dichos coeficientes), tal como había sido anunciado, era de esperar; lo anterior se explica fundamentalmente porque al estimar el parámetro λ_{λ} en forma simultánea (o conjunta) con los parámetros β 's, se tiene (y debe tenerse) en cuenta tanto la varianza propia de λ_{λ} , como la covarianza entre dicho parámetro y el resto de los coeficientes presentes en la ecuación. Para una prueba formal de este punto, véase Spitzer, *op. cit.* Sin embargo, conviene tener presente que las desviaciones estándares de parámetros que se reportan en el Cuadro 3 aún están subestimadas, pues corresponden a las que se calcularon para los modelos restringidos; en realidad, las únicas que no estarían subestimadas serían las del modelo general que no se calcularon.

Aplicando el test de Razón de Verosimilitud entre el modelo no restringido y cada uno de los modelos restringidos del Cuadro 3 (añadidos del primero), no es posible inferir si el modelo lineal en las X_{λ} (es decir $\lambda_{\lambda}=1$), o el modelo con logaritmos en las X_{λ} ($\lambda_{\lambda}=0$) son estadísticamente distintos del modelo general (con λ_{λ} endógenos)¹⁴. De lo anterior se deduce que el grado de variabilidad que permite el método de MV en la estimación de los λ_{λ} es de una magnitud no despreciable, lo cual despertaría dudas sobre su calidad; sin embargo, esto no puede corroborarse en esta oportunidad por la ya mencionada carencia de la matriz de varianzas y covarianzas de los parámetros estimados para el modelo general. Por otra parte, y reconociendo el método de comparación entre modelos no anidados propuesto en la sección con la estrategia de estimación para los modelos (1) y (4), para la etapa 5, es posible la elección entre el modelo con $\lambda_{\lambda}=1$ y el con $\lambda_{\lambda}=0$. Los resultados que se obtienen al comparar la suma de los residuos al cuadrado (SSR = "sum of squared residuals") de ambos modelos restringidos permiten inferir que el modelo con logaritmos en las X_{λ} es marginalmente inferior al modelo lineal en las X_{λ} , en cuanto a su capacidad para predecir el comportamiento de la variable dependiente observado¹⁵.

En resumen, en esta tercera etapa de estimación del modelo (1) por el método de MV hemos encontrado que el modelo lineal en las variables independientes ($\lambda_{\lambda}=1$) puede ser considerado "levemente superior" al modelo con logaritmos en las X_{λ} ($\lambda_{\lambda}=0$) y, a la vez, prácticamente indistinguible del modelo más general.

D. Etapa 4

Aplicando la técnica de MCNL se estiman tres modelos distintos. El primero de ellos corresponde al modelo general (4) y los otros dos a versiones restringidas de éste; la primera versión restringida corresponde a una en donde se supone que las variables independientes aparecen en forma lineal y la otra en forma logarítmica. El origen de estos modelos restringidos obedece a la misma lógica seguida en la etapa anterior, de identificar eventualmente especificaciones funcionales más simples que la general y, consecuentemente, más fáciles de interpretar y de utilizar para propósitos de predicción. Asimismo, como las versiones restringidas son modelos anidados de la especificación más general, es relativamente fácil inferir sobre la pertinencia estadística de adoptar la especificación más restringida. En el Cuadro 4 se presentan los resultados obtenidos.

CUADRO 4
MINIMOS CUADRADOS NO LINEALES DEL MODELO (4)

Variable	Box-Cox General	Variable	Box-Cox-Lin $\lambda_{\chi^2, s=1}$	Variable	Box-Cox-Log $\lambda_{\chi^2, s=0}$
VALESST	823*E-3	VALESST	979*E-3	VALESST	961*E-3
LAMBDA	(391*E-3)	LAMBDA	(324*E-3)	LAMBDA	(175*E-2)
C	132*E+3	C	945*E+3	C	714*E+3
ANTGCCAI	(103*E+4)	ANTGCCAI	(390*E+4)	LNANTG	(157*E+5)
LAMBDA	358*E-1	LAMBDA	554*E 0	LAMBDA	169*E-1
LAMBDA	(503*E 0)	LAMBDA	(242*E-1)	LAMBDA	(425*E+2)
LAMBDA	121*E-2	LAMBDA	(343*E-2)	LAMBDA	(141*E-2)
CALCAS2	160*E+1	CALCAS2	386*E+2	LNCALCAS	(322*E+3)
LAMBDA	(109*E+2)	LAMBDA	(164*E+3)	LAMBDA	(526*E-2)
LAMBDA	259*E-2	LAMBDA	(526*E-2)	LAMBDA	(526*E-2)
SUSTEST	225*E+2	SUSTEST	127*E+1	LNSURPST	361*E-2
LAMBDA	(185*E+3)	LAMBDA	(541*E+1)	LAMBDA	(820*E+3)
LAMBDA	285*E-4	LAMBDA	(148*E-2)	LAMBDA	(132*E 0)
DBASURA	-397*E-1	DBASURA	-288*E-2	DBASURA	-451*E+1
DBASURA	(205*E+2)	DBASURA	(123*E+3)	DBASURA	(107*E+3)
DTTPO	981*E+1	DTTPO	694*E+2	DTTPO	142*E-2
DCALLE	(510*E+2)	DCALLE	319*E+2	DCALLE	809*E-1
DVEREDA	187*E+2	DVEREDA	(142*E+3)	DVEREDA	(189*E+3)
DVEREDA	-168*E-1	DVEREDA	-517*E+1	DVEREDA	-12457.14
TPOPARAD	(954*E+1)	TPOPARAD	(346*E+2)	LNTPOPAR	(383*E+2)
LAMBDA	-913*E-1	LAMBDA	-657*E+1	LNTPOPAR	-842*E+1
LAMBDA	(487*E+2)	LAMBDA	(278*E+2)	LNTPOPAR	(193*E+3)
LAMBDA	-436*E-3	LAMBDA	(996*E-3)	LNTPOPAR	(996*E-3)
TPOTELDI	131*E 0	TPOTELDI	-166*E 0	LNTPOTEL	-286*E 0
LAMBDA	(308*E+1)	LAMBDA	(177*E+1)	LAMBDA	(144*E+2)
LAMBDA	451*E-3	LAMBDA	(110*E-1)	LAMBDA	(110*E-1)

CUADRO N.º 4

MINIMOS CUADRADOS NO LINEALES DEL MODELO (4)

Variable	Box-Cox General	Variable	Box-Cox-Lin $\lambda_{\chi^2, s=1}$	Variable	Box-Cox-Log $\lambda_{\chi^2, s=0}$
CGVTRAB	-647*E+2	CGVTRAB	-130*E-1	LNGCVTR	441*E+2
LAMBDA	(346*E+3)	LAMBDA	(134*E 0)	LAMBDA	(995*E+3)
CGVALLME	-871*E-3	CGVALLME	271*E 0	LNGVALL	209*E+1
LAMBDA	(556*E-3)	LAMBDA	(117*E+1)	LAMBDA	(483*E+2)
LAMBDA	(185*E+1)	LAMBDA	558*E-3	LAMBDA	(170*E-2)
TPOESCUE	253*E 0	TPOESCUE	148*E+1	LNTPOESC	521*E+1
LAMBDA	(180*E+1)	LAMBDA	(628*E+1)	LAMBDA	(120*E+3)
LAMBDA	(145*E-2)	LAMBDA	929*E-3	LAMBDA	(145*E-2)
TPOSALUD	-135*E+1	TPOSALUD	-628*E 0	LNTPOSAL	-310*E+1
LAMBDA	(102*E+2)	LAMBDA	(291*E+1)	LAMBDA	(727*E+2)
LAMBDA	221*E-3	LAMBDA	(198*E-2)	LAMBDA	(198*E-2)
CGVCENTR	-101*E+1	CGVCENTR	-523*E-1	LNGVCENT	-163*E+3
LAMBDA	(172*E+2)	LAMBDA	(222*E+2)	LAMBDA	(372*E+4)
LAMBDA	929*E-3	LAMBDA	219*E 0	LAMBDA	(276*E-2)
TPOPOSTA	(276*E-2)	TPOPOSTA	969*E 0	LNTPOPOST	483*E-1
LAMBDA	(168*E+1)	LAMBDA	(410*E+1)	LAMBDA	(110*E+3)
LAMBDA	884*E-3	LAMBDA	(159*E-2)	LAMBDA	(159*E-2)
TPOCOMII	-183*E-2	TPOCOMIII	-173*E-1	LNTPOCOM	188*E+1
LAMBDA	(244*E-1)	LAMBDA	(743*E+1)	LAMBDA	(330*E+2)
LAMBDA	241*E-2	LAMBDA	(355*E-2)	LAMBDA	(355*E-2)
ALUMPUBL	199*E+2	ALUMPUBL	199*E+2	LNALUM	-108*E-1
LAMBDA	(528*E+3)	LAMBDA	(871*E+2)	LAMBDA	(274*E+2)
LAMBDA	-482*E-2	LAMBDA	(132*E 0)	LAMBDA	(132*E 0)

R squared

correctido

R squared

correctido

R squared

S.E. of Regress.

Log likelihood

Mean of dep. var.

Mean of dep. var.

Mean of dep. var.

Mean of dep. var.

Nº observ.

Nº observ.

Nº observ.

Nota: Ver en Anexo 2 la definición y descripción de las variables. Los valores entre paréntesis corresponden a las desviaciones estándares de los coeficientes; las correspondientes al modelo general no están subíndicadas aunque sí lo están las de los modelos restringidos, de acuerdo a lo que se indica en el texto. El número de observaciones corresponde a aquellas en que existía información de todas las variables independientes de las 1.095 disponibles, por no respuesta de algunos encuestados.

Aplicando ahora el test de Razón de Verosimilitud entre el modelo no restringido y cada uno de los modelos restringidos (anidados del primero), es posible inferir que el modelo lineal en los X_k (es decir $\lambda_k=1$) no es estadísticamente distinto del modelo general (con λ_k endógenos); sin embargo no es posible arribar a similar inferencia para el modelo con logaritmos en los X_k ($\lambda_k=0$)¹⁶. Esta conclusión permite simplificar apreciablemente la especificación general del modelo (4) con ventajas evidentes para fines de interpretación y de predicción. Y aquí vale recordar que similar conclusión se alcanzó cuando las comparaciones fueron hechas para análogas versiones en el contexto de estimación por MV (pero, en rigor, de otro modelo).

E. Etapa 5

Antes de proceder a comparar los modelos (1) y (4), conviene detenerse a resumir lo que se ha encontrado en las cuatro etapas anteriores. En las etapas 1 a 3, que corresponden al modelo (1), se concluyó que el modelo general ("Box-Cox a ambos lados") puede restringirse a uno Box-Cox Lineal o a otro Box-Cox Log, y que la restricción lineal al modelo es levemente superior. Por otra parte, en la etapa 4 se llega a la conclusión que el modelo general (4) puede restringirse al modelo Box-Cox Lineal también. Lo que corresponde ahora es elegir uno de los modelos Box-Cox Lineal comparando los SSR o R² corregidos, donde el modelo (4) estimado por MCNL presenta ventajas claras sobre el modelo (1) estimado por MV; además, lo mismo ocurre al comparar las variantes restringidas *vis-à-vis* ¹⁷. Ello en definitiva conduce a dejar de lado el modelo (1) porque el (4) parece mejor.

Tal vez los resultados anteriores son consecuencia de los problemas asociados a estimar el modelo (1) con la técnica de MV y que el modelo (4) no posee; uno de ellos, ya mencionado, de Amemiya y Powell, *op. cit.*, se refiere a que (1) no es estrictamente un modelo estadístico. El segundo problema se relaciona con la falta de robustez del método MV frente a violaciones al supuesto de normalidad de los errores del método MV, tal como describen también Amemiya y Powell en pág. 380 y Amemiya, *op. cit.*, en pág. 380.

F. Etapa 6

Esta etapa se desarrolla a partir de los resultados de la etapa 5, en que se eligió el modelo (4) en su versión restringida Box-Cox Lineal; ahora corresponde estudiar si es posible restringir más aún dicha especificación: a una Lineal-Lineal. Al respecto, se reporta que al testear se concluye que la especificación lineal-lineal no es estadísticamente distinta de la especificación Box-Cox-Lineal¹⁸. Esto completa la ejemplificación de la estrategia de estimación sugerida en este artículo¹⁹.

4. Conclusiones

En este artículo se ha propuesto una estrategia de estimación de ecuaciones no lineales del tipo Box-Cox, incluyendo un procedimiento para elegir entre los modelos siguientes (referidos antes como modelos (1) y (4), respectivamente):

$$Y_i^{(\lambda)} = \beta_1 + \beta_2 X_{2i}^{(\lambda)} + \dots + \beta_k X_{ki}^{(\lambda)} + \varepsilon_i; \quad \varepsilon_i \text{ iid } N(0, \sigma^2)$$

$$Y_i = \left(\beta_1 + \beta_2 X_{2i}^{(\lambda)} + \dots + \beta_k X_{ki}^{(\lambda)} \right)^{1/\lambda} + \mu_i; \quad \mu_i \text{ iid } N(0, \sigma_\mu^2)$$

donde: $Z^{(\lambda)} = \frac{Z^\lambda - 1}{\lambda}$ si $\lambda \neq 0$ (en cuyo caso $Z(\lambda)$ converge a $\ln(Z)$), para $Z = Y$ o cualquier X_k continuo, y $\{Z\lambda_y + 1\}^{1/\lambda_y}$ converge a $\text{EXP}[Z]$ para $\lambda_y = 0$. Las conclusiones obtenidas son:

1. Los dos modelos anteriores aparentan ser iguales; pero el supuesto de comportamiento del término de error los hace distintos. El primer modelo puede estimarse con la técnica de máxima verosimilitud (MV) y el segundo con la técnica de mínimos cuadrados no lineales (MCNL).
2. La estimación de la ecuación hedónica de precios en la segunda forma nace, por un lado, de la demostración de John Spitzer [1982] de que al estimar el primer modelo por MV se obtendría idéntico resultado que al estimar una transformación trivial del mismo modelo por el método de MCNL y, por otro, del hecho de que la estimación por MCNL (hasta donde nosotros sabemos) requiere de que la variable dependiente aparezca *despejada* o libre de la influencia de cualquier parámetro endógeno, tal como está escrito el segundo modelo. Si la restricción de software recién mencionada es universal (o efectiva), se concluye:
 - i) Que estimar el primer modelo directamente por MCNL es imposible, lo que invalida ese método como alternativa de estimación a la de MV;
 - ii) Que la utilización de la técnica MCNL para estimar el modelo (1) requiere vías indirectas con un grid search, más la aplicación de las fórmulas de Spitzer, que no son fáciles de implementar; y
 - iii) Que es innecesario dividir cada observación de la variable dependiente por su media geométrica muestral ya que ello sólo es necesario para que los resultados de estimar el modelo (1) por MCNL coincidan con los de estimar el mismo modelo por MV, que es uno de los resultados de Spitzer; es innecesario por cuanto la técnica MCNL que se puede implementar no conduce al modelo (1) sino al modelo (4).
3. A primera vista, el poder predictivo del modelo (4) estimado por MCNL puede aparecer considerablemente inferior al del modelo (1) estimado por MV, cuando se comparan coeficientes de determinación al cuadrado (R²), tal como ocurrió en ejemplo ilustrativo para precios hedónicos de hoteles con servicios. Dicha conclusión puede ser errónea cuando la variable dependiente sufre transformaciones, tal como en el modelo (1), pues los coeficientes R² no son directamente comparables. Utilizando una simplificación de la retransformación propuesta por D.M. Miller [1984], para obtener una medida de bondad de ajuste se puede llegar a inferir que el modelo (4) estimado por MCNL sea superior al modelo (1) estimado por MV, si es que los R² corregidos (que se derivan de sumas de residuos al cuadrado comparables) aparecen superiores para el primer caso. Y ello es precisamente lo que sucedió en el caso práctico estudiado en este trabajo.

Lo anterior puede ayudar de paso a explicar dos interrogantes que a menudo se plantean en la literatura aplicada de estimaciones de ecuaciones hedónicas de precios:

- i) Por qué la transformación semilogarítmica es la preferida entre los econométristas (de entre las transformaciones linealizantes típicas); y
 - ii) Por qué, en el contexto de Mínimos Cuadrados Iterativos, la elección del parámetro de transformación de la variable dependiente (λ_y) es aquel que minimiza la suma de residuos al cuadrado (SSR) del modelo (1) pero previamente escalado (i.e., dividiendo previamente cada observación de la variable dependiente por su media geométrica muestral), y por qué no lo es el λ_y de mayor R^2 (y que será el mismo, ya sea que el modelo haya sido previamente escalado o no).
4. La superioridad del método de estimación por MCNL, por sobre el de MV puede o no deberse a que la última no es robusta ante la violación del supuesto de normalidad de los errores, lo cual ocurre en el contexto de este estudio y se expone claramente en Anemíya y Powell [1981]. Ahora, en el contexto más práctico de la estimación hedónica de esta investigación, se rescata una interrogante más motivar más investigación en el contexto de economía urbana.

Independientemente de los argumentos que se esgrimen en la literatura, y de los aquí esgrimidos, la elección entre el valor de arrendo y el valor de mercado del activo como variable dependiente a utilizar en la ecuación hedónica de precios es una cuestión no bien resuelta y amerita más investigación. Al respecto, queda por explicar el marcado patrón de heteroscedasticidad encontrado al regresar entre dichas variables.

Notas:

- 1 Rosen [1974] argumenta que no existe razón para presuponer una especificación funcional lineal, lo cual es posible sólo en el contexto que los bienes puedan ser reemplazados en sus características o en partes de ellas; supone que lo anterior no es posible o no es económico hacerlo. Como ejemplo aclaratorio cita: i) "los autos de seis pies de largo no son equivalentes a uno de doce pies de largo pues ellos no pueden ser conducidos simultáneamente"; y ii) "un auto de 12 pies de largo por medio año más otro de 6 pies por el medio año restante no es igual que uno de 9 pies por todo el año".
- 2 Nuestra impresión es que esta especificación es usualmente escogida, pues tiende a integrar los mejores R^2 comparadas con el resto. La elección de este tipo de especificación (o para estos efectos de cualquier otra de las otras) en base al criterio del mayor R^2 es esencialmente erróneo por cuanto, como se verá en detalle más adelante, las transformaciones que aplican a la variable dependiente en cada caso son distintas y por ende los R^2 no son comparables.
- 3 No obstante, en este estudio se realizan estimaciones de este estilo para identificar valores de partida a la iteración de los algoritmos a usar en la estimación.
- 4 Note que la especificación descrita no incluye efectos interactivos entre distintas variables (sean éstas cuantitativas o cualitativas); la inclusión de este tipo de efectos haría de la especificación una más general aún, aunque complicando de paso la estimación y la interpretación de los coeficientes. Respecto de esto último, véase Casel y Mendelshon [1985].
- 5 Al respecto, Casel y Mendelshon, *op. cit.*, refiriéndose a especificaciones del tipo (1) para estimar ecuaciones hedónicas presentadas arriba en el texto pero además incluyendo efectos interactivos, menciona cuatro críticas al uso de las transformaciones de Box-Cox para estimaciones de ecuaciones de precios hedónicos: 1) que el criterio de escoger la ecuación que mejor se ajusta a los datos (más alto

R^2) no necesariamente conlleva a estimadores más precisos de los precios implícitos; 2) que las transformaciones no lineales entregan coeficientes estimados para las pendientes y para los precios implícitos a menudo demasiado complejos como para ser utilizados en la práctica, sobre todo cuando aparecen efectos interactivos entre distintas variables independientes; 3) que las transformaciones en cuestión no permiten manejar valores negativos para las variables en juego, tal como para $\lambda_y = 1/2$ y que origina un número imaginario si $X_i < 0$; y 4) que pueden ser inapropiadas para fines de predicción, por cuanto si lo que interesa es predecir valores de la variable dependiente en sus valores originales, dado que las estimaciones se realizan en base a variables transformadas, dichas predicciones serían sesgadas.

En nuestro parecer, sólo 3) puede ser considerado de importancia práctica en el contexto de este trabajo; para afrontar esa dificultad se podría optar por simplificar algebraicamente el modelo (1), por ejemplo forzando $\lambda_s = 1$, que vendría a ser una simplificación cuya validez debe ponderarse en cada caso particular (y que en caso de ser aceptable podría conducir a estar en el mismo contexto de este trabajo, tal vez con un modelo anidado de (1)). Respecto a 1), se obvia el uso de R^2 como criterio de selección entre modelos (andidos); 2) es una cuestión más de estilo y que conduce por ejemplo a descartar variables interactivas (que el modelo (1) no tiene); y respecto a 4), se utilizarán métodos para corregir dicho problema.

6 Es conveniente tener en mente que en el contexto de estimación no lineal que precupara, los intervalos de confianza usuales (y por ende los tests de significancia) son válidos sólo asintóticamente. En el caso del test F estándar, su uso es válido sólo como una aproximación y su calidad variará de caso en caso. En el contexto de máximo verosímil, el test de razón de verosimilitud puede ser utilizado invocando la normalidad asintótica del estimador de máxima verosimilitud. Para mayores detalles puede consultarse un clásico en esta materia: Goldfeld y Quandt [1972]; asimismo, puede consultarse a Anemíya [1984].

7 Habitualmente, el criterio que se utiliza para decidir sobre el "mejor ajuste" es el máximo R^2 (o R^2 ajustado), tal como se indica en la nota de pie de página 4; al respecto, conviene tener presente que los R^2 no son directamente comparables si la variable dependiente no es la misma, aun cuando se tenga la impresión de que son comparables porque la medida de R^2 es adimensional. La alternativa es obtener valores estimados de la variable dependiente para construir R^2 's comparables o, alternativamente, transformar los datos para asegurar comparabilidad. Para una medida de calidad de ajuste en modelos no lineales, véase Haessel [1978]. Para un ejemplo del tipo de transformaciones, véase Rao y Miller [1971]. Más adelante se utilizará un criterio denominado de bondad de ajuste para realizar comparaciones de este estilo al comparar modelos no anidados.

8 Aquí conviene tener en mente, además, que un grid search simultáneo para λ_y y para los λ_x 's de las variables independientes puede ser innamajable. Ante ello, puede ser práctico seguir el método iterativo de Box-Tidwell, en que para cada λ_x del grid search se identifican los λ_y 's con un algoritmo que converge al de MV; ver, por ejemplo, Maddala [1977]: p. 315. Otra alternativa es aplicar MCNL para cada λ_x del grid search.

9 Note que el supuesto de normalidad de los errores no es requerido para la derivación de los estimadores por Mínimos Cuadrados; asimismo, dichos estimadores serán consistentes incluso en el caso de que se viole el supuesto de normalidad de los errores ya que sólo es necesario para la realización de test de hipótesis; más aún, recordando lo puntualizado por Anemíya y Powell, *op. cit.*, será "aproximadamente Normal".

10 Aquí conviene tener presente que la generalización del resultado de Miller es esencialmente una simplificación cuya validez teórica no ha sido probada. Aquí se agradece el comentario del referee que lo hizo notar y convidaría seguir su sugerencia de utilizar el método de corrección no paramétrico de Duan [1983].

11 Normalmente, debiera considerarse una etapa final donde se resuelven problemas como qué variables interactivas agregar y qué hacer con coeficientes de regresión que aparecen con signos cambiados. Aquí no se propone ninguna estrategia, más que recomendar que ello se haga después de las etapas 1 a la 6; lo único que se menciona es que cada econométrista intentará la suya.

12 Ver BID [1987], sección b, página 21: "Las alternativas abiertas para futuros análisis de proyectos usando este método parecen ser, ya sea trazar sólo la estimación de la primera etapa hedónica y a la vez descartar paralelamente en otras técnicas no hedónicas con el propósito de generar medidas de beneficios que puedan ser utilizadas como chequeos de plausibilidad, o insistir en consultores excepcionalmente bien calificados que generen medidas de beneficio de corto plazo vía variación compensatoria junto con las usuales estimaciones de precios de la forma reducida para establecer límites superiores e inferiores para los beneficios".

13 Ver P.U. Católica de Chile [1987], Tomo II, págs. N.º102-103: "Para analizar los antecedentes se comenzó dividiendo el parque de viviendas en dos grupos: las permanentes y las mediegas. Sin embargo, las estimaciones mejoraron notablemente en términos de R^2 y del número de coeficientes significativos al juntar ambas poblaciones y estudiar viviendas simplemente. Este hecho no es difícil de

- explicar toda vez que al aumentar la diversidad de tipos de vivienda, aumenta la heterogeneidad y, por lo tanto, la variabilidad de las variables independientes, lo que facilita la posibilidad de discriminación entre ellas... "Se estudiaron dos divisiones más. La primera entre los hogares pertenecientes a los 6 deciles más pobres versus el resto. El resultado mostró que el comportamiento estadístico de las regresiones en el caso de los deciles más pobres fue muy bajo en términos de ajuste. El valor de R² oscilaba entre 0,11 y 0,35, en tanto que el de los deciles superiores fue sustancialmente mejor, con R² entre 0,28 y 0,50".
- 14 Para el caso del modelo lineal en las X's, el test de razón de verosimilitud arroja un valor de 12,6, que comparado con el valor de tabla para una Chi-Cuadrado con 13 grados de libertad al 95% de confianza (22,4) no permite rechazar la hipótesis de que todos los λ_i sean iguales a uno. Para el caso del modelo logarítmico en las X's, el mismo test arroja un valor de 12,2 que comparado con el valor de tabla tampoco permite rechazar la hipótesis de que todos los λ_i sean iguales de cero o, en otras palabras, que dichos valores sean estadísticamente distintos de cero.
- 15 El valor de SSR para el modelo con $\lambda_i = 1$ asciende a 1,623239*10⁴, mientras que el correspondiente al modelo con $\lambda_i = 0$ es 1,624369*10⁴. En términos de R² corregidos, los valores corresponden a 0,165582 y 0,16524, respectivamente.
- 16 Para el caso del modelo lineal en las X's el test de razón de verosimilitud arroja un valor de 9,82, que comparado con el valor de tabla para una Chi-Cuadrado con 13 grados de libertad al 95% de confianza (22,4) no permite rechazar la hipótesis de que todos los λ_i sean iguales a uno. Para el caso del modelo logarítmico en las X's, el test de razón de verosimilitud arroja un valor de 147,98, que comparado con el valor de tabla permite rechazar la hipótesis de que todos los λ_i sean iguales de cero o, en otras palabras, que dichos valores sean estadísticamente distintos de cero.
- 17 Los valores de SSR corregidos para el modelo general son (donde los R² corregidos se escriben entre paréntesis): SSR_{gen} = 1,560538*10⁴ (0,1980122); SSR_{uv} = 1,612234*10⁴ (0,17147). Los valores de SSR para el modelo restringido (lineal en X's) son: SSR_{merc} = 1,578384*10⁴ (0,188867); SSR = 1,622239*10⁴ (0,16582).
- 18 El test de razón de verosimilitud para la comparación entre el modelo Box-Cox-lineal y el lineal-lineal arroja un valor de 0, lo que inmediatamente conduce a no rechazar la hipótesis nula de que ambos modelos son similares. En cualquier caso el valor crítico de tabla para una Chi-cuadrado con un grado de libertad al 95% de confianza es igual a 3,84. Por otra parte, el test de razón de verosimilitud para la comparación entre el modelo Box-Cox general y el lineal-lineal arroja un valor de 9,82, lo que comparado con el valor crítico de una Chi-cuadrado con 14 grados de libertad al 95% de confianza de 23,7 no permite rechazar la hipótesis nula de que ambos modelos son similares. En términos de poder predictivo o de R² corregido, el modelo lineal-lineal se comporta muy parecido al modelo Box-Cox-lineal y al modelo Box-Cox general, siendo los valores 0,1888661, 0,188867 y 0,1980122, respectivamente.
- 19 Al lector interesado en aspectos no tratados en este artículo, y referidos a este ejemplo en particular, tales como la incorporación de variables interactivas y la interpretación de resultados, se le remite a nuestra publicación como Documento de Investigación citado al principio del artículo.

Referencias

1. AMEMITYA, TAKESHI and JAMES POWELL. [1981], "A Comparison of the Box-Cox Maximum Likelihood Estimator and the Non-Linear Two-Stage Least Squares Estimator", *Journal of Econometrics* V. 17, pp. 351-381.
2. AMEMITYA, TAKESHI [1984], Capítulo 6: "Non-Linear Regression Models", *Handbook of Econometrics* V. 1. Editores: Zvi Griliches y Michael Intriligator, North Holland Publishing Company, pp. 333-389.
3. BID [1987], "Hedonic Rent (or Property Value) Functions and Project Benefit Measurement: Some Methodological and Empirical Issues", *Reporte OEO/WP-03/87* (Washington).
4. BID [1988], "Empirical Issues in the Estimation of Hedonic Rent or Property Value Equations and their Use in Prediction", *Reporte OEO/WP-02/88* (Washington).
5. BOX, G.E.P. and D.R. COX [1964], "An Analysis of Transformations", *Journal of the Royal Statistical Society*, serie B, pp. 211-252.
6. CASSELL, ERIC y ROBERT MENDELSSOHN [1985], "The Choice of Functional Forms for Hedonic Price Equations: Comment", *Journal of Urban Economics* V. 18, pp. 135-42.
7. DUMAN, N. [1983], "Smearing Estimate: A Nonparametric Transformation Method", *Journal of the American Statistical Association*, V. 78, N° 383, pp. 605-610.

8. GOLDFELD, S.M. and R.E. QUANDT [1972], *Nonlinear Methods in Econometrics*, Contributions to Economic Analysis, North Holland Publishing Company.
9. HAESSEL [1978], "Measuring Goodness of Fit in Linear and Non Linear Models", *Southern Economic Journal*, V. 44, pp. 648-52.
10. MADDALA [1977], *Econometrics* McGraw-Hill.
11. MILLER, DON M. [1984], "Reducing Transformation Bias in Curve Fitting", *The American Statistician*, V. 38, N° 2 (May), pp. 124-126.
12. PONTFICIA U. CATOLICA DE CHILE [1987], *Diagnóstico de Alternativas de Sistemas Habitacionales* (Santiago).
13. QUANDT, R.E. [1984], Capítulo 12: "Computational Problems and Methods", *Handbook of Econometrics*, V. 1, op.cit., pp. 699-764.
14. RAO Y MILLER [1971], *Applied Econometrics*, Belmont California: Wadsworth.
15. ROSEN, SHERWIN [1974], "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition", *Journal of Political Economy*, V. 82, pp. 34-55.
16. SPITZER, JOHN I. [1982], "A Primer on Box-Cox Estimation", *Review of Economics and Statistics*, V. LXIV (May), pp. 307-313.
17. WUNDER, DIETER y HECTOR GUTIERREZ [1992], "Precios de Mercado de Soluciones de Vivienda Progressiva en el Gran Santiago", por aparecer como Documento Serie Investigación, Departamento de Economía Universidad de Chile.

Anexo I

ELECCION DE LA VARIABLE DEPENDIENTE

Algunos de los problemas básicos a resolver en la elección de la variable dependiente usada se refieren a: 1) Estimar la ecuación hedónica directamente a partir de precios de lotes con servicios o indirectamente con valores de propiedades que además incluyen una vivienda; 2) Estimar directamente a partir de precios del activo o indirectamente con valores de arriendo; y 3) Diferencias en respuestas según se trate de disposición a vender o a comprar o a valor de mercado.

1) *Valor del lote con servicios sin vivienda vs. con vivienda*: El problema básico se origina en que los lotes con servicios no existen como propiedad separada de la vivienda que sus moradores le adosan, por lo que, al momento de encuestar, se tiene la opción de preguntar por el valor comercial que los moradores atribuyen al lote con servicios o, alternativamente, preguntar por el valor de la propiedad completa y deducir de alguna manera el del lote con servicios sin la vivienda. Esto condujo a incluir en la encuesta preguntas tanto del valor del lote con servicios mismo, como del valor que incluye a la vivienda.

Si bien puede sospecharse del valor de los lotes con servicios sin vivienda que reportan los encuestados, pues no existen en realidad, la tarea de estimar una ecuación hedónica para precios de propiedades con vivienda y obtener de ahí la ecuación sin ella fue desestimada. Ello se debe esencialmente a la extrapolación que involucra, pues, por ejemplo, se requiere obtener el valor de un lote con servicios sin vivienda a partir del de lotes con viviendas con cierto número de piezas (en vez de sin piezas), y con ciertos m²

* Si bien estos problemas no afectan la estrategia de estimación que se sugiere seguir en este artículo, se ha considerado que su mención es interesante al menos para motivar mayor investigación, aparte del interés que puedan tener los resultados.

construidos (en vez de los 6 u 8 del lote); esa extrapolación no es trivial por posible cambio estructural en la ecuación hedónica. Además, los encuestadores sugirieron que, a pesar de todo, la estimación del valor del lote sin la vivienda es estimada con mayor precisión por los encuestados que con la vivienda, tal vez por ser más homogénea.

2) *Valor de activo vs. de arriendo*: Partiendo de la base de que el mercado de compra-venta de propiedades con lotes con servicios es poco activo, puede postularse que la información de precios que tengan las personas es imprecisa, lo que conduce a postular la alternativa de inferir el precio del activo a partir de valores de arriendo. Esto condujo a preguntar a los encuestados tanto por valores de activo como de arriendo.

La consistencia entre valores de arriendo y de activo se examinó mediante una regresión entre ellos, sin intercepto. Ello condujo a un valor medio de arriendo igual a 1,24% del valor de lotes con servicios con vivienda, y de 1,12% para el caso de lotes con servicios sin vivienda; ambos valores medios están de acuerdo a los esperados por corresponder a un valor no muy distinto y algo mayor al 1% que utilizan los corredores de propiedades como regla aproximada para valorar arriendos. Pero, además, en la regresión se encontró un patrón de comportamiento heterocedástico de los residuos muy marcado; ello sugiere que uno de los dos valores, arriendo o valor de activo, entrega mejor información que el otro, aunque sin referenciar a cuál.

Finalmente se optó por elegir a la variable precio del activo en vez de valor de arriendo. Esta elección se fundamentó en los objetivos de la investigación, en que se desea obtener precios de activo, ya que elegir el valor de arriendo requiere un paso adicional al tener que convertirlos después en valor de activo, y no se puede estar seguro que esos errores de conversión, más los propios de la estimación de una ecuación hedónica para arriendos, sean inferiores a los de estimar directamente una ecuación hedónica para precios de activo. Pero para decidir también se tomó en cuenta la experiencia de los encuestadores y de profesionales del Ministerio de la Vivienda, que sugirieron que el valor de activo era estimado con mayor precisión (tal vez porque se tiene una amplia mayoría de propietarios que viven en su propiedad, en que su propiedad es su activo máspreciado, la cual no están dispuestos a arrendar según revela el hecho que no lo hacen). (En el documento BID [1987], se llega a la misma conclusión en su sección F.5 página 31: Additional Sampling Issues: Property Value Or Rentals?)

3) *Precio de demanda vs. de oferta vs. de mercado*: Como las respuestas de precio que se obtienen no corresponden a transacciones efectivas, excepto en el valor de arriendos, es de rigor preguntarse qué responden los encuestados. Por una parte, los propietarios pueden responder por el valor mínimo al cual venderían la propiedad, que es un precio de oferta, y los arrendatarios, por otra, podrían responder por el valor máximo al cual comprarían la propiedad, que es un precio de demanda; alternativamente, propietarios y arrendatarios podrían responder sobre el precio de mercado, que teóricamente se le puede ubicar entre precio de demanda y de oferta. Esto condujo a incluir en la encuesta preguntas tanto de valor de mercado estimado como disposición a vender (a propietarios) o a comprar (a arrendatarios), según correspondiera.

Los análisis estadísticos practicados condujeron a concluir que el valor de oferta reportado por propietarios era significativamente mayor que el de mercado, y que el de demanda reportado por los arrendatarios era significativamente inferior; ambas conclusiones se obtuvieron aplicando dos test no paramétricos para diferencia entre medias:

MODELOS TIPO BOX-COX: PROPOSICION DE UNA ESTRATEGIA DE ESTIMACION 101

test de signos y de Wilcoxon. Esto da razonable seguridad para usar la información de precios de mercado estimados por los encuestados.

Anexo 2

VARIABLES INDEPENDIENTES O EXPLICATIVAS

El listado de variables que se utiliza en las regresiones es el siguiente:

C: Define la constante en la regresión.

ALUMPUBL: Representa a la variable alumbrado público. Toma el valor 1 si el encuestado tiene alumbrado público, 2 si no tiene y 3 si lo tiene, pero no funciona.

ANTIGCAl: Corresponde a la antigüedad de la caseta (obtenida a partir de la fecha en que el Estado entregó la solución). Originalmente esta variable estaba medida en años, de cero a 8 años por ejemplo; lo anterior tuvo que ser modificado, puesto que algunas especificaciones requerían tomar logaritmos de las variables, y dicha transformación no está definida para valores cero. En definitiva la estrategia utilizada fue multiplicar la variable por 12 para dejarla expresada en meses, asignando "arbitrariamente" un valor de 6 para los valores que antes correspondían a cero.

CALCAS2: Esta variable es un índice que representa la calidad de la caseta (1=mal, 2=aceptable y 3=bueno). La información proveniente de la encuesta entrega información para la calidad de los muros, del piso y del techo de la solución; el índice pondera la calidad de muros, techo y pisos en razón de 3:2:1, es decir calcas 2=0.5 calidad muros + 0.33 calidad de techo + 0.17 por la calidad del piso.

DBASURA: Es una variable dummy que representa recolección de la basura en el barrio al cual pertenece el entrevistado. Toma el valor 1 si la recolección es cada dos días o menos, y el valor 0 si es cada tres días o más.

DCALLE: Es una variable dummy que adopta el valor 1 si la calle de enfrente a la propiedad del entrevistado está pavimentada y 0 en cualquier otro caso.

DTTPO: Variable dummy que adopta el valor 1 si el barrio es tranquilo y seguro todo el tiempo, y 0 en cualquier otro caso (tranquilo sólo de día, tranquilo sólo de día y en días de trabajo y, finalmente, intranquilo todo el tiempo).

DVEREDA: Es una variable dummy que adopta el valor 1 si la vereda adyacente a la propiedad del entrevistado está pavimentada y 0 en cualquier otro caso.

SUSJTEST: Corresponde a la superficie del sitio de la solución, estimada por el encuestador; es una variable cuantitativa, continua y está medida en m².

TPOCOMII: Corresponde al tiempo hasta la comisaría más cercana, en minutos.

TPOESCUE: Es el tiempo a la escuela que asisten los niños del hogar, en minutos.

TPOPARAD: Corresponde al tiempo hasta la parada de buses más cercana, en minutos.

TPPOSTA: Es el tiempo a la "posta" (centro de salud de urgencia) más cercana, en minutos.

TPPOSALUD: Corresponde a la variable tiempo hasta el centro de salud, en minutos.

TPOTELDI: Corresponde al tiempo hasta el teléfono más cercano disponible en el día (expresado en minutos).

CGVALIME: Corresponde al Costo Generalizado de Viaje hasta el lugar de compra de los alimentos, en pesos por viaje de ida. Este tipo de índice, ampliamente utilizado en la evaluación de proyectos de transporte, agrega (es decir, suma) al gasto monetario efectivo en que se incurre para viajar hasta el lugar de destino el tiempo que toma hacerlo, ponderado por un factor, de manera tal que el tiempo quede expresado en unidades monetarias (en el caso chileno, pesos). La ponderación corresponde al equivalente en pesos de una hora de transporte, determinado a partir del ingreso por hora promedio de los individuos o personas de las que se obtienen ambas variables. En evaluación de proyectos se suele utilizar como valor de una hora de transporte un tercio del ingreso por hora promedio de las personas, criterio que fue adoptado en esta investigación. Los datos de ingreso promedio fueron obtenidos de la encuesta ocupacional realizada en marzo de 1991, por la Unidad de Encuestas-Departamento de Economía de la Universidad de Chile. En síntesis, se determinó que para las personas del nivel socio-económico de los encuestados, el valor promedio de una hora de transporte corresponde aproximadamente a 86,3 pesos (moneda de julio de 1991).

CGVCCENTR: Corresponde al Costo Generalizado de Viaje hasta el centro de la ciudad, en pesos por viaje de ida.

CGVTRAB: Esta variable representa el Costo Generalizado de Viaje hasta el lugar de trabajo del jefe de hogar, en pesos por viaje de ida.

UNA DEMANDA POR DINERO MENSUAL PARA CHILE: 1983:1-1992:8*

JAIMÉ Z. APT**

ILADES/Georgetown University

JORGE A. QUIROZ**

ILADES/Georgetown University

Abstract:

Conventional wisdom in Chile today, adopted by critical observers and monetary authorities alike, points to the hypothesis that money demand is an unstable and erratic function, even at quarterly frequency. As a consequence of this, so the argument goes, the Central Bank should attempt to achieve price stabilization by focusing primarily on interest rates, without paying much attention to nominal monetary aggregates. This paper presents an econometric estimation of a monthly money demand function for Chile during the period 1983:1-1992:8. The estimated function satisfies strong stability tests and passes a number of other statistical criteria. The demand displays the necessary flexibility to accommodate important short run fluctuations which results in a fit to the data superior to all money demand functions estimated so far in Chile. The finding presented in this paper strongly suggests to rethink the potential role that nominal monetary objectives may have on price stabilization policies.

1. Introducción

El problema de la existencia de una demanda agregada de dinero estable ha sido recurrente en la literatura económica mundial así como en la chilena. Si bien las conclusiones que podrían deducirse de la existencia de dicha demanda no son compartidas

* Una versión preliminar de este artículo fue publicada como documento de trabajo de ILADES/Georgetown University.

** Se agradecen los comentarios y sugerencias de Patricio Arrau y Francisco Rosende. Errores y omisiones son de exclusiva responsabilidad de los autores.