
MODELACION CONJUNTA DE LA ASIGNACION DE TIEMPO AL TRABAJO Y ELECCION MODAL CONSIDERANDO CORRELACION DE ERRORES

Marcela A. Munizaga
Departamento de Ingeniería Civil, Universidad de Chile
Casilla 228-3, Santiago, Chile
Fono: (56-2)-6784380; Fax: (56-2)-6894206
e-mail: mamuniza@cec.uchile.cl

Rodrigo Correia
Departamento de Ingeniería de Transporte
Pontificia Universidad Católica de Chile
Casilla 306, Cod. 105, Santiago 22, Chile
e-mail: rcorreia@puc.cl

Sergio Jara-Díaz
Departamento de Ingeniería Civil
Universidad de Chile
Casilla 228-3, Santiago, Chile
e-mail: jaradiaz@cec.uchile.cl

Juan de Dios Ortúzar
Departamento de Ingeniería de Transporte
Pontificia Universidad Católica de Chile
Casilla 306, Cod. 105, Santiago 22, Chile
e-mail: jos@ing.puc.cl

RESUMEN

La modelación conjunta de actividades y viajes es un problema complejo que sólo recientemente ha recibido una adecuada atención en la literatura. Recientemente se ha reportado una modelación microeconómica de la asignación de tiempo a actividades, que permite derivar dos ecuaciones principales: una relativa a la elección de modo de transporte y otra concerniente a la asignación de tiempo al trabajo. Desde el punto de vista de la estimación, parece natural esperar alguna interrelación entre los términos de error de ambas ecuaciones, pero a la fecha sólo se ha reportado resultados preliminares asumiendo independencia. A fin de avanzar en esta línea, se adaptó un modelo econométrico discreto/continuo que permite postular una estructura de error relativamente general en la especificación estocástica del problema. En este trabajo se muestra que al hacer esto, se obtienen modelos más verosímiles y resultados más consistentes que los obtenidos asumiendo errores independientes. El método es eficiente y relativamente simple frente al problema y puede ser fácilmente extendido a casos más complejos a futuro.

1. INTRODUCCION

La modelación de demanda de transporte ha mostrado avances notables en el último tiempo. Al reconocer formalmente que ésta es una demanda derivada (surge a partir de los deseos y necesidades de las personas por realizar actividades), parece particularmente relevante modelar conjuntamente las decisiones de asignación de tiempo a actividades y las decisiones de viajes. En este contexto, a partir de la microeconomía, Jara-Díaz y Guevara (2002) formulan un modelo basado en el enfoque de De Serpa (1971) en el cual las personas maximizan su utilidad decidiendo qué bienes consumir y de qué forma gastar su tiempo, sujeto a restricciones técnicas y ambientales. A partir de este modelo se derivan dos ecuaciones fundamentales: una relativa a la elección de modo de transporte y otra a la asignación de tiempo al trabajo. Jara-Díaz y Guevara (2002) estiman el modelo asumiendo independencia entre los términos de error de ambas funciones, pero reconocen que éste es claramente un supuesto cuestionable.

En este trabajo se revisa las diversas formas de liberar el supuesto de independencia incorporando una estructura de correlación, se elige una de ellas y se aplica experimentalmente. Para ello se describe brevemente el modelo de Jara-Díaz y Guevara (2002) y se formula el desafío econométrico planteado. Luego se presenta una breve revisión de la literatura relacionada con la especificación de errores en modelos discretos/continuos, optando por un enfoque apropiado al contexto específico. Posteriormente el método se aplica al mismo caso de modelación utilizado por Jara-Díaz y Guevara (2002) y se realiza una comparación que permite evaluar los beneficios de utilizar un marco econométrico más flexible. Finalmente se destacan las principales conclusiones y posibles extensiones de este trabajo.

2. FORMULACION DEL PROBLEMA

La literatura sobre comportamiento del consumidor ha recibido un aporte decisivo desde las perspectivas de los modelos de producción en el hogar, oferta de trabajo y transporte. Esta es la consideración del tiempo como variable central. En estos modelos, la utilidad depende de la asignación de tiempo a actividades y del consumo, considerando restricciones de ingreso, tiempo total y tecnológicas. A partir de la formulación de De Serpa (1971), bajo un enfoque de elección discreta de modo (Train y McFadden, 1978), Jara-Díaz y Guevara (2002) derivan dos ecuaciones fundamentales: una relativa a la elección de modo de transporte (1) y otra para la asignación de tiempo al trabajo (2).

$$U_q^j \approx \gamma^j + \gamma^t t_q^j + \gamma^c c_q^j + \varepsilon_q^j \quad (1)$$

$$T_{wq}^j = \beta(\tau - T_q^j) + \alpha \frac{c_q^j}{w_q} + \sqrt{\left[\beta(\tau - T_q^j) + \alpha \left(\frac{c_q^j}{w_q} \right) \right]^2 - [2(\alpha + \beta) - 1] \frac{c_q^j}{w_q} (\tau - T_q^j) + \eta_q^j} \quad (2)$$

En la ecuación (1) U_q^j es una aproximación de la utilidad indirecta condicional a la elección de modo j para el individuo q , es decir, el valor máximo que puede alcanzar la función de utilidad para un modo dado. Por su parte γ^j representa una constante modal para la alternativa j , γ^t un parámetro para el tiempo (agregado), γ^c un parámetro del costo y ε_q^j el error asociado al modo j .

En lo relativo a las variables, t_q^j es el tiempo (agregado) del modo j para el individuo q , c_q^j equivalentemente es el costo del modo j para el individuo q .

En la segunda ecuación (2), T_{wq}^j es el tiempo asignado al trabajo por el individuo q condicional a la elección del modo j . De la misma manera, τ es el tiempo total disponible, T_q^j representa el tiempo de viaje del modo j para la persona q , w_q la tasa salarial, y η_q^j el error asociado al trabajo. Por último α y β , que no tienen una interpretación directa, son sólo manipulaciones algebraicas definidas en el trabajo de Jara-Díaz y Guevara (2002). Su uso, en conjunto con el modelo discreto, permite obtener todas las componentes del valor subjetivo del tiempo (VST). Esta ecuación, no-lineal, se obtiene de las condiciones de primer orden del problema, resolviendo para un modo dado. Esto es, se deriva una expresión que dé cuenta del tiempo de trabajo, condicional al modo elegido, que maximiza la utilidad (tiempo de trabajo óptimo).

En definitiva el modelo plantea que la persona elige en forma simultánea el modo de transporte y el tiempo que asigna al trabajo. Desde el punto de vista teórico el individuo calcula tiempos de trabajo óptimos para cada modo pero aquellos tiempos no asignados están ocultos para el analista.

Para estimar los modelos planteados en (1) y (2) es preciso suponer una cierta estructura de los errores asociados. Ambas ecuaciones están sujetas a errores que provienen de diversas fuentes: la observación empírica de las variables está sujeta a errores típicos de la recolección de datos, los supuestos del modelo están sujetos a error de modelación (todos los usuarios se comportan de la misma forma, forma funcional, entre otros). Si se analiza ambas ecuaciones por separado, parece natural plantear que el modelo de elección discreta asociado a la ecuación (1) tendrá un término de error aditivo Gumbel, que corresponde a la distribución más razonable si se busca encontrar un máximo. El caso de la ecuación (2) es menos claro, dado que se trata de una ecuación no lineal. Se decidió que lo más conservador en este caso era asumir la existencia de un término aleatorio que distribuye Normal dada las múltiples fuentes de error con distribución desconocida que contribuyen en forma independiente.

Ahora bien, dado que existen variables que aparecen en ambas ecuaciones y que están sujetas a error de medición, es natural postular que exista correlación entre los términos aleatorios de cada uno de los modelos. Se puede pensar, además, en la existencia de variables no observadas (latentes) que pueden afectar ambas decisiones sin que exista una causalidad de por medio. Más importante aún, existe una razón analítica para postular que exista esta correlación; la función de utilidad indirecta es condicional al modo elegido, e implica el cálculo de la máxima utilidad que el individuo puede conseguir dado que eligió un modo determinado. Esto se traduce en que la función de utilidad indirecta, aunque no aparezca explícitamente debido a la aproximación lineal, sea función del tiempo asignado al trabajo.

3. MODELACION ECONOMETRICA DE SISTEMAS DISCRETOS/CONTINUOS

Los problemas de elección discretos/continuos han recibido alguna atención en la literatura. Los primeros enfoques para tratar el problema eran de tipo secuencial. Típicamente en dos etapas, en la primera se obtienen los parámetros de la elección discreta mediante máxima verosimilitud;

para la segunda se pueden adoptar dos enfoques. Si se cree que hay sesgo de endogeneidad, esto es, si se considera que alguna de las variables de la ecuación continua es endógena a la elección discreta, ella se puede reemplazar por una función de las probabilidades obtenidas en el primer paso. Por otra parte, si se cree que existe un sesgo de selección, es decir que las observaciones de la ecuación continua dependen de la elección discreta, se aumenta la regresión añadiendo un término adicional que es función de las probabilidades de la primera etapa. La mayoría de los casos puede ser tratado mediante cualquiera de los dos métodos y la idea general es tomar en cuenta la simultaneidad de las decisiones introduciendo correlación entre los errores aleatorios de la parte discreta y la parte continua.

Los referentes clásicos del método secuencial utilizando el enfoque de corrección por selección, son los trabajos de Heckman (1979) y de Dubin y McFadden (1984). El primero de estos trabajos asume que ambos errores distribuyen Normal y es el método más popular debido a lo sencilla que resulta su aplicación. El segundo amplía los términos de corrección por selección cuando el error de la parte discreta es Gumbel (dando paso a un modelo Logit, ver Ortúzar y Willumsen, 2001), pero sigue suponiendo que el error asociado a la parte continua es Normal.

El siguiente paso en el desarrollo de este problema econométrico consistió en buscar modelos que maximizaran una función de verosimilitud única, que contuviera toda la información que se posee sobre el fenómeno a fin de estimar parámetros eficientes, consistentes e insesgados. Esto debido a que el procedimiento de calibración de dos etapas no es tan eficiente como el de máxima verosimilitud con información completa, y presenta grandes problemas en la construcción de la matriz de covarianza asintóticamente correcta¹.

Hanemann (1984) postula que los dos pilares fundamentales de un modelo discreto/continuo son las funciones de utilidad indirecta y la función de densidad conjunta de los errores de la parte discreta y continua. El primer aspecto está relacionado con la microeconomía del problema y se encuentra plenamente desarrollado en el modelo de Jara-Díaz y Guevara (2002). Por otra parte, la densidad conjunta de los errores es ciertamente la clave en la especificación estocástica. La definición adecuada de una densidad conjunta permite tratar en forma irrestricta correlación entre errores. En el caso dicotómico, dicha especificación es relativamente sencilla. Por ejemplo, si se asume que ambos distribuyen conjuntamente Normal y, por ende las distribuciones marginales son también Normales se obtiene como resultado que el modelo de elección discreta es un Probit (Ortúzar y Willumsen, 2001). El caso dicotómico más común en transporte es la tenencia y uso de automóvil; De Jong (1989) modela esta situación presentando un buen ejemplo de modelo discreto/continuo estimado por máxima verosimilitud con información completa.

Cuando la situación que se desea modelar incluye una elección discreta policotómica, el problema se torna más complejo. Se sabe que el modelo Probit es más difícil de tratar en casos de elecciones múltiples, y que se hace necesario recurrir a simulación. Lee (1983) propone una metodología para el caso en que existan razones *a priori* para asumir ciertas distribuciones marginales específicas, que de alguna forma da vuelta el problema; esto es, la atención no se centra en especificar una función de densidad conjunta para los errores sino que, a partir de

¹ Si bien Lee *et al* (1980) proponen un procedimiento con este fin, éste se complica enormemente en casos en que el conjunto de elección discreta presenta más de dos alternativas.

funciones marginales específicas, en proponer funciones conjuntas interesantes. Se entiende por función conjunta interesante a cualquiera que permita correlación irrestricta entre los errores. El método consiste básicamente en generar una transformación que permite convertir la distribución de los errores especificada *a priori*, en errores Normal estándar. Así, la distribución conjunta de estos nuevos errores (transformados adecuadamente) es también Normal estándar. En el apéndice de este trabajo se encuentra un resumen de la derivación del modelo.

Krishnamurthi y Raj (1988) utilizan esta técnica para modelar la elección de marca y su consumo. Barnard y Hensher (1992) lo hacen para modelar la elección de lugar de compra y la función de gasto. Ambos trabajos asumen un error Gumbel en la elección discreta y un error Normal en la parte continua. Los trabajos de Bhat (1998; 2001) ofrecen una extensión del modelo de Lee (1983) cuando existe una elección discreta policotómica y dos decisiones continuas e introducen modelos jerárquicos en la elección discreta.

Siguiendo otras líneas de investigación Sanga (1999) utiliza un enfoque Probit por simulación cuando existe una elección discreta policotómica. Aún así su trabajo resulta difícil de ser aplicado cuando el conjunto de elección es muy grande. Bolduc *et al* (2002) extienden el trabajo de Heckman (1979) y Dubin y McFadden (1984) desarrollando términos de corrección de selección cuando el modelo de elección discreta es un Logit Mixto.

Sobre la base de esta revisión, se llegó a la conclusión que el enfoque de Lee constituía una estrategia interesante para el caso en estudio, ya que no presenta problemas para considerar elecciones discretas múltiples y también permite incorporar correlación entre la alternativa elegida en el conjunto discreto y el nivel de la variable continua.

4. MODELO DE ACTIVIDADES Y VIAJES CONSIDERANDO CORRELACION

Para desarrollar el método de Lee (1983) al sistema de ecuaciones (1) y (2) asumiremos que el error asociado a la elección de modo de transporte distribuye Gumbel (0,1) y el error en la ecuación de asignación de tiempo al trabajo distribuye Normal (0, σ^2). Esta elección permite mantener el problema dentro de los límites tratables y posibilita incorporar, eventualmente, especificaciones jerárquicas que permitan correlación entre alternativas. La condición de elección de la alternativa *i* está dado por:

$$U_i^* > \underset{j \neq i}{\text{Max}} U_j^* \tag{3}$$

Como ya se mencionó, esta utilidad puede descomponerse en una componente observable más un término de error que, distribuye Gumbel (0,1). Con esto, la condición presentada en la ecuación (3) puede ser reescrita de la siguiente forma:

$$\tag{4}$$

La parte del lado derecho de esta última ecuación representa un término aleatorio que podemos llamar ω_i . La distribución de probabilidad asociada a este término depende de la distribución de

las variables aleatorias U y ε . Ambas variables están estrechamente relacionadas ya que U distribuye Gumbel a causa de ε . Sabiendo que el máximo de términos Gumbel distribuye también Gumbel y que la resta de dos funciones Gumbel independientes e idénticas (iid) distribuye Logística, se da paso al conocido modelo Logit.

Utilizando el método de Lee (1983) el error ω_i se puede transformar en un error Normal estándar aplicando la función Normal inversa. Una propiedad interesante que posee esta transformación es que por ser monótona creciente tiene la capacidad de mantener la condición de elección del modo i , $V_i \geq \omega_i$, de la siguiente manera:

$$J(V_i) = \Phi^{-1}[F(V_i)] \geq \Phi^{-1}[F(\omega_i)] = \omega_i^* \quad (5)$$

en que $F(\cdot)$ representa la función de distribución Logística característica del modelo Logit, en donde:

$$F(V_i) = \frac{\exp(V_i)}{\exp(V_i) + \sum_{\substack{j=1 \\ i \neq j}} \exp(V_j)} \quad (6)$$

Con estas definiciones se ha logrado formular un problema de elección policotómico con errores Gumbel como uno de elección dicotómica con errores Normales; el término de error ω_i^* reúne inobservables relativos a la alternativa i en relación a todas las demás alternativas j . Así podemos presentar la parte discreta del sistema de ecuaciones propuesto por Jara-Díaz y Guevara (2002) como:

$$R_{qi}^* = J(V_i) - \omega_i^* \quad (7)$$

$$R_{qi} = 1 \quad \text{si} \quad R_{qi}^* > 0 \quad (8)$$

$$R_{qi} = 0 \quad \text{si} \quad R_{qi}^* < 0 \quad (9)$$

Acá R_{qi} toma el valor uno si la persona q elige la alternativa i y cero en caso contrario. R_{qi}^* representa la propensión latente, no observada, del individuo q por elegir el modo i . El nuevo término de error ω_i^* entra con signo negativo en esta ecuación de propensión, lo que es relevante al analizar los posibles signos de la correlación entre los términos de error ω_i^* y η_i . Concretamente si el signo de esta correlación es negativo, ello indicará que inobservables que aumentan la propensión a elegir el modo i aumentan igualmente el tiempo asignado al trabajo.

Ya que ω_i^* y η_i distribuyen marginalmente Normal, su función de distribución conjunta y distribución condicional son igualmente Normal. Aprovechando la descomposición de la probabilidad conjunta y el hecho que T_{wi} se observa sólo si la alternativa i es elegida, o sea si se cumple (5), podemos escribir la función de verosimilitud para la alternativa i , escogida por N individuos, de la siguiente forma:

$$(10)$$

Con ello, usando notación estándar, la función de verosimilitud total queda:

$$L = \prod_{i=1}^T \left\{ \prod_{i=1}^N \left[\frac{1}{\sigma} \phi(t_i) \cdot \Phi(b_i) \right]^{R_{qi}} \right\} \quad (11)$$

En donde T representa el número total de modos de transporte considerados y σ es la desviación estándar de η . Por su parte las cantidades t_i y b_i se definen de la siguiente manera:

$$t_i = \frac{T_{wi} - \left\{ \beta(\tau - T_i^j) + \alpha \frac{c_i^j}{w_i} + \sqrt{\left[\beta(\tau - T_i^j) + \alpha \left(\frac{c_i^j}{w_i} \right) \right]^2 - [2(\alpha + \beta) - 1] \frac{c_i^j}{w_i} (\tau - T_i^j)} \right\}}{\sigma} \quad (12)$$

$$b_i = \frac{\Phi^{-1} [F(V_i^*)] - \rho_i t_i}{\sqrt{1 - \rho_i^2}} \quad (13)$$

Es decir, t_i representa el término $g_2(\eta_{ii})$, mientras que b_i representa a $g_1(\omega_{ii}^*/\eta_{ii})$. Por otra parte, el factor ρ_i representa la correlación entre inobservables de la elección del modo i e inobservables de la asignación de tiempo al trabajo. Este parámetro permite acomodar factores subyacentes, no observados, que influyen en la elección de modo y la asignación de tiempo al trabajo y debe ser incluido en la función de verosimilitud. No obstante esto, la idea del modelo no es encontrar un valor específico para esta correlación, sino obtener estimadores para los parámetros de interés más eficientes e insesgados.

5. APLICACION NUMERICA

Un objetivo interesante de este modelo econométrico es dar cuenta de la correlación entre los términos de error para el modelo de Jara-Díaz y Guevara (2002). Por esto se decidió utilizar la misma base de datos a fin de poder comparar las virtudes del modelo propuesto. La muestra utilizada se construyó a partir de información sobre 377 trabajadores reportada en la encuesta origen destino de 1991 de Santiago (Ortuzar *et al*, 1993). Estos trabajadores fueron elegidos en forma aleatoria dentro de un conjunto que presentaba un esquema de actividades sencillo: casa-viaje-trabajo-viaje-casa. Los datos contienen información respecto de tiempos asignados a cuatro actividades agregadas durante un día de trabajo normal: trabajar, estaba en el hogar, viaje al trabajo y viaje de vuelta al hogar.

Para segmentar la muestra por ingreso, los individuos se agruparon inicialmente en dos estratos definidos en base al ingreso familiar neto: entre \$110.000 y \$405.000 (\$ de 1991) y con un ingreso superior. Para algunos efectos, el estrato medio se separó en dos sub-estratos: medio-bajo y medio-alto. La base de datos utilizada también incluye niveles de servicio, es decir tiempos de caminata, espera y viaje propiamente tal. Además contiene información relativa a los costos de cada modo para cada individuo. Respecto a esto último es posible afirmar que el costo de viaje promedio es de \$170 pesos para el estrato medio y de \$226 para el estrato alto.

En la Tabla 1 se reporta los resultados del mejor modelo obtenido utilizando el método propuesto por Lee (1983) y los resultados para el modelo original de Jara-Díaz y Guevara (2002); en ambos casos se fijó como alternativa de referencia el bus. Es interesante mencionar que a pesar de

intentar especificaciones de tipo jerárquico, que permitieran postular modos más similares entre sí, el test de razón de verosimilitud (LR, ver Ortúzar y Willumsen, 2001) arrojó preferencia hacia el modelo Logit simple.

Tabla 1
Resultados de modelación. Base de datos EOD 1991

Variables	Independiente		Con Correlación	
	Estimador	t-est.	Estimador	t-est
Constantes modales				
Auto chofer	0,4505	1,7	0,606	2,3
Auto acompañante	2,2286	2,0	2,1852	2,0
Taxi colectivo	-2,4617	-7,0	-2,4256	-6,9
Metro	-0,4155	-1,3	-0,5393	-1,6
Caminata	0,5343	1,4	0,0293	0,1
Taxi	-0,7729	-1,5	-0,3527	-0,7
Bus-Metro	-2,2807	-4,3	-2,0965	-3,9
Taxi colectivo-Metro	-2,4407	-4,7	-2,3807	-4,6
Parámetros				
γ tiempo estrato medio-bajo	-0,0708	-3,6	-0,0721	-3,8
γ costo estrato medio-bajo	-0,0058	-3,3	-0,0084	-4,2
γ tiempo estrato medio-alto	-0,0576	-4,2	-0,0476	-3,6
γ costo estrato medio-alto	-0,0032	-3,6	-0,0043	-4,1
γ tiempo estrato alto	-0,091	-3,9	-0,075	-3,4
γ costo estrato alto	-0,0019	-2,2	-0,0023	-2,3
β estrato medio	0,0961	13,3	0,0957	11,2
β estrato alto	0,1208	13,6	0,1316	13,9
α	-2,2729	-6,3	-3,2607	-6,7
σ	160,4409	27,0	164,1584	24,7
Correlaciones				
ρ Achof	-		0,6531	6,3
ρ Bus	-		0,3243	2,2
ρ Cami	-		-0,4512	-3,0
ρ Tax	-		0,509	4,0
Log Verosimilitud	-2644,9		-2627,1	
N° Observaciones	366		366	
Test LR	-		35,6	
χ^2 (4, 99,9%)	-		18,5	

Como se puede ver, el modelo que incorpora correlación es muy superior al que asume que todas las correlaciones son cero. Un test de LR rechaza la hipótesis nula incluso al 99,9% de confianza. De la misma forma, es posible apreciar que la única correlación con signo negativo corresponde a la del modo caminata. Esto sugiere que inobservables que aumentan la propensión a elegir la caminata, aumentan igualmente el tiempo asignado al trabajo. Se puede especular que dicho signo puede ocultar la cercanía al lugar de trabajo.

Otro punto de comparación de los resultados es mediante los valores de las distintas componentes del tiempo, a saber: valor subjetivo de ahorrar tiempo de viaje (VSATV), valor del ocio, valor de asignar tiempo al trabajo (VATT) y valor de asignar tiempo a viajes (VATV). Las expresiones para calcularlos se definen en Jara-Díaz y Guevara (2002). Los valores del estadístico t presentados en la Tabla 2, fueron calculados utilizando una aproximación en series de Taylor de

primer orden (ver Jara-Díaz *et al*, 1998; Ortúzar y Willumsen, 2001). Para examinar la bondad de esta aproximación también se calculó intervalos de confianza para el VSATV de acuerdo a la metodología propuesta por Armstrong *et al* (2001).

Tabla 2
Comparación de Valores del Tiempo [\$/min]

Resultados Estrato Medio (Tasa salarial: 22,8)							
	Valor Subjetivo de Ahorrar Tiempo de Viaje			Valor del Ocio		VATT	VATV
	Valor	t-est	IC	Valor	t-est		
Independiente	16,1	4,3	9,9 – 27,3	0,9	8,5	-21,9	-15,2
Con Correlación	10,5	4,1	6,1 – 17,3	0,6	8,7	-22,1	-9,9
Variación porcentual	-35		-35	-26		+1	+35
Resultados Estrato Alto (Tasa salarial: 50,4)							
	Valor Subjetivo de Ahorrar Tiempo de Viaje			Valor del Ocio		VATT	VATV
	Valor	t-est	IC	Valor	t-est		
Independiente	45,9	2,3	21,4 – 370,8	2,4	8,1	-48,1	-43,6
Con Correlación	33,7	2,2	14,3 – 198,0	1,7	8,0	-48,7	-32,0
Variación porcentual	-27		-47	-28		+1	-27

VATT: valor de asignar tiempo al trabajo. VATV: valor de asignar tiempo a viajes.

Se puede ver que tanto el VSATV como el Valor del Ocio disminuyen en porcentajes importantes (cerca al 30%) al incorporar el efecto de correlación a la modelación. Esto se considera un resultado positivo, ya que generalmente se le critica a las estimaciones de VSATV en nuestro país el hecho de ser demasiado elevadas en relación al ingreso (Gaudry *et al*, 1989). Por otra parte, el tamaño de los intervalos de confianza se reduce en magnitudes importantes en ambos estratos (casi a la mitad en el estrato alto), indicándonos que se ha ganado en precisión de las estimaciones.

6. CONCLUSIONES

Una primera conclusión de este trabajo es que al incorporar la posibilidad de capturar correlación a la modelación econométrica del modelo de Jara-Díaz y Guevara (2002), se obtiene un resultado significativamente mejor desde el punto de vista estadístico (test de razón de verosimilitud), y mejora la estimación de las distintas componentes del Valor Subjetivo del Tiempo, entregando estimaciones más razonables, y con mayor precisión.

Con respecto a la interpretación de los coeficientes de correlación que resultaron significativos, se debe destacar el caso de la correlación entre caminata y tiempo de trabajo. Dada la formulación del modelo, el hecho de que el valor sea negativo implica que factores no observados que aumentan la propensión a elegir el modo caminata, aumentan igualmente el tiempo asignado al trabajo. Es posible especular que una variable no observada (en este caso, no incluida en la modelación), como la cercanía al lugar de trabajo, puede causar este efecto. Por otra parte, los modos que tienen correlación de signo positivo, indicarían que factores no observados que aumentan la propensión a elegir estos modos, disminuyen el tiempo asignado al trabajo (se puede especular que se esté intentando evitar horas punta). Es importante insistir en que la correlación modelada se da entre inobservables relegados al término de error.

En cuanto a las líneas futuras de extensión de esta investigación, se puede mencionar la posibilidad de probar otras estructuras de error, utilizando por ejemplo el marco que provee un modelo tipo Mixed Logit (Train, 2002). Por otra parte, dado que la formulación planteada por Jara-Díaz y Guevara (2002) permite incorporar otras actividades, resulta un hecho positivo que el modelo econométrico permita extensiones a casos con dos asignaciones continuas. Si bien la función de verosimilitud se vuelve más complicada, sigue siendo tratable (Bhat, 1998).

AGRADECIMIENTOS

Esta investigación fue financiada parcialmente por FONDECYT y el programa Milenium.

REFERENCIAS

Armstrong, P.M, R.A. Garrido y J.deD. Ortúzar (2001) Confidence intervals to bound the value of time. **Transportation Research 37E**, 143-161.

Barnard, P.O. y D.A. Hensher (1992) Joint estimation of a polychotomous discrete-continuous choice system: an analysis of the spatial distribution of retail expenditures. **Journal of Transport Economics and Policy XXVI**, 299-312.

Bhat, C.R. (1998) A model of post-home arrival activity participation behavior. **Transportation Research 32B**, 387-400.

Bhat, C.R. (2001) Modeling the commute activity-travel pattern of workers: formulation and empirical analysis. **Transportation Science 35**, 61-79.

Bolduc, D., L. Khalaf, y E. Moyneur (2002) Joint discrete/continuous models with possibly weak identification. Working Paper, Department of Economics, Université Laval, Québec.

De Jong, G.C. (1989) Simulating car costs changes using an indirect utility model of car ownership and annual mileage. **Proceedings PTRC's 17th Summer Annual Meeting**, Brighton.

De Serpa (1971) A theory of the economics of time. **The Economic Journal 81**, 828-846.

Dubin, J. y D. McFadden (1984) An econometric analysis of residential electric appliance holdings and consumption. **Econometrica 52**, 345-362.

Gaudry, M.J.I, S.R. Jara-Díaz y J.deD. Ortúzar (1989) Value of time sensitivity to model specification. **Transportation Research 23B**, 151-158.

Hanemann, W.M. (1984) Discrete/Continuous models of consumer demand. **Econometrica 52**, 541-561.

Heckman, J.J. (1979) Sample selection bias as a specification error. **Econometrica 47**, 153-161.

Jara-Díaz, S. y C.A. Guevara (2002) Behind the subjective value of travel time savings: the perception of work, leisure and travel from a joint mode choice-activity model. **Journal of Transport Economics and Policy** (en imprenta).

Jara-Díaz, S.R., J.de D. Ortúzar y R. Parra (1988) Valor Subjetivo del Tiempo considerando efecto ingreso en la partición modal. **Actas del V Congreso Panamericano de Ingeniería de Tránsito y Transporte**, Universidad de Mayaguez, Puerto Rico.

Krishnamurthi, L., y S.P. Raj (1988) A model of brand choice and purchase quantity price sensitivities. **Marketing Science** 7, 1-20.

Lee, L.F. (1983) Generalized econometric models with selectivity. **Econometrica** 51, 507-512.

Ortúzar, J. de D., A.M. Ivelic, H. Malbrán y A. Thomas (1993) The 1991 Great Santiago origin-destination survey: methodological design and main results. **Traffic Engineering and Control** 34, 362-368.

Ortúzar, J. de D. y L.G. Willumsen (2001) **Modelling Transport**. Tercera Edición, John Wiley and Sons, Chichester.

Sanga, D. (1999) **Estimation des Modèles Économétriques de Choix Discrets/Continus avec Choix Polytomiques Interdépendants: Une Approche par Simulation**. PhD Thesis, Université Laval, Québec

Train, K. (2002) **Discrete Choice Methods with Simulation**. University Press, Cambridge.

Train, K. y D. McFadden (1978) The goods/leisure trade off and disaggregate work trip mode choice models. **Transportation Research** 12, 349-353.

ANEXO

TRANSFORMACION DE LEE

Este apéndice sintetiza las ideas centrales de la transformación de Lee, que permite expresar la función de verosimilitud de un modelo discreto/continuo en términos de la distribución Normal, permitiendo su calibración econométrica incluyendo ciertos términos de correlación.

En particular supóngase una situación en que la persona i tiene M posibles elecciones discretas (en nuestro caso M modos de transporte) con una utilidad latente (no observada) denominada U_{si}^* . Asimismo su elección continua, condicional a la elección de la alternativa s es Y_{si} . Con estas definiciones se puede formular el siguiente sistema de ecuaciones:

$$(A.1)$$

$$(A.2)$$

En la ecuación (A.1) z_{si} es un vector de variables exógenas para la persona i , como por ejemplo el tiempo del modo s . Consecuentemente γ_s es un vector de parámetros a ser estimado y η_{si} es un error aleatorio con una distribución

marginal definida *a priori*. En la ecuación (A.2), x_{si} es un vector de variables exógenas como nivel educacional del individuo i , β_s un vector de parámetros a ser estimados, μ_{si} una variable aleatoria con distribución marginal definida *a priori* y σ_s su varianza. Adicionalmente se asume que $E(\mu_{si}/x_1, \dots, x_M, z_1, \dots, z_M) = 0$ y $E(\eta_{si}/x_1, \dots, x_M, z_1, \dots, z_M) = 0$.

La condición necesaria para que se produzca la elección de la alternativa s está dada por la ecuación (A.3) en que para ahorrar notación se omite el subíndice del individuo i .

$$U_s^* > \underset{\substack{j=1, \dots, M \\ j \neq s}}{\text{Max}} U_j^* \quad (\text{A.3})$$

Esto significa que la alternativa s es escogida cuando su utilidad es mayor que el máximo de las utilidades de las otras alternativas (la utilidad de s es mayor que la de la segunda mejor alternativa). Esta utilidad U_s^* puede descomponerse en una parte observable más un término de error. Así es posible aislar la parte aleatoria de la elección y escribir la condición (A.4).

$$V_s > \underset{\substack{j=1, \dots, M \\ j \neq s}}{\text{Max}} U_j^* - \eta_s \quad (\text{A.4})$$

Sea entonces I una variable múltiple con valores de 1 a M y sea $I=s$ si la alternativa s es elegida. Utilizando las ecuaciones (A.1) y (A.4) la condición de elección de la alternativa s se puede escribir como:

$$I = S \quad \text{ssi} \quad z_s \gamma_s > \eta_s^* \quad \text{en donde} \quad \eta_s^* \equiv \underset{\substack{j=1, \dots, M \\ j \neq s}}{\text{Max}} U_j^* - \eta_s \quad (\text{A.5})$$

Sean $G(\mu)$ y $F(\eta_s^*)$ las distribuciones marginales especificadas *a priori* para μ y η_s^* . Lee (1983) establece que los errores μ y η_s^* , que distribuyen de acuerdo a estas funciones G y F , pueden transformarse en variables aleatorias Normal estándar, utilizando la función Normal inversa. Si Φ es la función distribución Normal estándar, es posible crear nuevos términos de error:

$$\eta_s^{**} = J_1(\eta_s^*) \equiv \Phi^{-1}(F(\eta_s^*)) \quad (\text{A.6})$$

$$\mu_* = J_2(\mu) \equiv \Phi^{-1}(G(\mu)) \quad (\text{A.7})$$

La propiedad deseable que tiene esta transformación es que permite ahora que los nuevos términos de error η_s^{**} y μ_* distribuyan conjuntamente Normal, con media cero y coeficiente de correlación ρ . Esto permite especificar correlación irrestricta entre ellos. En particular interesa poder definir la probabilidad de observar ambos errores en forma conjunta. Haciendo uso de la propiedad Bayesiana clásica de probabilidades es posible escribir:

$$f_p(\eta_s^{**}, \mu_*) = g_1(\mu_*) \cdot g_2(\eta_s^{**} / \mu_*) \quad (\text{A.8})$$

Tomando $g_s(\cdot)$ como la función de densidad de $G_s(\cdot)$ y las variables dummy $D_s = 1$ si, y sólo si, $I=s$, la log-verosimilitud del modelo de elección múltiple con muestra aleatoria de tamaño N y con individuos i es:

$$\ln L(\beta, \gamma, \rho, \sigma) = \sum_{s=1}^M \sum_{i=1}^N \left\{ D_{si} \cdot \ln g_s \left(\frac{Y_{si} - x_{si} \beta_s}{\sigma_s} \right) + D_{si} \cdot \left(\ln \Phi \left[J_{1s}(z_{si} \gamma_s) - \frac{\rho_s \cdot J_{2s}((Y_{si} - x_{si} \beta_s) / \sigma_s)}{\sqrt{1 - \rho_s^2}} \right] \right) \right\} \quad (\text{A.9})$$

Como se puede apreciar, el primer término corresponde a la realización continua y el segundo a la realización discreta condicional a la realización continua. ρ y σ son parámetros a ser estimados dentro del modelo. Ese coeficiente de correlación ρ representa la correlación entre la alternativa elegida en el problema discreto modificado (cuyo error aparece con signo negativo) y la variable continua. Por lo tanto, un valor negativo de ρ indica la presencia de correlación positiva entre el error de una cierta alternativa discreta, si esta es elegida, y la variable continua.