

# Inclusión de Variables Latentes en Modelos de Elección Discreta para Usuarios de Transporte Público Interurbano

Javier Egaña Ugrinovic, Juan de Dios Ortúzar, Luis Ignacio Rizzi  
 Departamento de Ingeniería de Transporte, Pontificia Universidad Católica de Chile  
 Casilla 306, Código 105, Santiago 22, Chile.  
 Fax: (562) 5530281  
 E-mail: jnegana@uc.cl; jos@ing.puc.cl; lrizzi@ing.puc.cl

## RESUMEN

En este trabajo se combina un modelo de elección discreta con un modelo de variables latentes a fin de estudiar la elección modal entre bus y tren para viajes interurbanos, utilizando un enfoque secuencial. La principal innovación de este estudio es la inclusión de atributos modales en el modelo de variables latentes. Otra contribución de este estudio es mostrar como los resultados pueden ser afectados arbitrariamente por el modelador según la normalización elegida del modelo de variables latentes. Este último resultado arroja ciertas dudas sobre la utilidad de estos modelos para el análisis económico, en especial si estamos interesados en obtener disposiciones al pago y / o elasticidades.

*Palabras claves:* variables latentes, ecuaciones estructurales, modelos de elección discreta.

## ABSTRACT

We combine a structural equation model and a discrete choice model to study mode choice for interurban public transport users using a two-stage estimation approach. The principal innovation of this case study is the inclusion in the structural model of a set of objective values that relate to trip characteristics rather than travellers' socioeconomic attributes. Another contribution of our analysis is to show how the results can be arbitrarily affected by the modeller's choice of normalization in estimating the structural equation model. This results cast doubt on the usefulness of this type of models for economic analysis, especially if we are interested in some key values as subjective valuations or elasticities.

*Keywords:* latent variables, structural equations, discrete choice models.

## 1. INTRODUCCIÓN

Este estudio entrega una nueva aplicación de modelos integrados de variables latentes y elección discreta en un contexto de elección modal entre bus y tren para viajes interurbanos desde la ciudad capital de Santiago de Chile a ciudades ubicadas al sur del país (Curicó, Talca, Linares y Chillán). Este caso de estudio presenta como novedad la inclusión en el modelo estructural de un conjunto de variables objetivas relacionadas con las características del viaje más que con las características socioeconómicas de los viajeros. Otro aporte del estudio es mostrar como los resultados de la modelación pueden ser afectados de manera arbitraria por el modelador dependiendo de las normalizaciones realizadas al estimar el modelo híbrido. Por último, también obtendremos la disposición a pagar por la seguridad vial de los usuarios de buses, aunque, como veremos, este resultado es sensible a la normalización adoptada del modelo de variables latentes.

El resto de este artículo se organiza de la siguiente manera. En la segunda sección se describe el modelo integrado de variables latentes y elección discreta. Esta misma sección contiene una revisión bibliográfica de los estudios realizados a la fecha según nuestro conocimiento. La tercera sección presente el caso de estudio descrito en el párrafo anterior. Esta sección describe el tipo de datos utilizados, los modelos estimados y sus resultados. La cuarta sección contiene las conclusiones.

## 2. EL MODELO INTEGRADO DE VARIABLES LATENTES Y ELECCIÓN DISCRETA

Las ecuaciones estructurales son técnicas estadísticas desarrolladas para determinar relaciones entre grupos de variables observables y grupos de variables latentes (conceptos no observables). Los modelos de medición son aquellas formas funcionales que vinculan estos conceptos no medibles, con variables e indicadores<sup>3</sup> observables y/o medibles. La integración de variables latentes en modelos de elección discreta se debe primordialmente al impulso dado por el trabajo de McFadden (1986) en donde presenta prototipos de modelos que incluyen indicadores psicométricos en las funciones de utilidad. Ben-Akiva et al. (1999) y McFadden (2000) sugieren integrar los modelos MIMIC (*multiple indicator and multiple cause*) con los modelos de elección discreta. Los modelos MIMIC son un caso particular de modelos de medición, donde variables objetivas alimentan las variables latentes que a su vez alimentan las variables indicadoras. A fin de integrar ambos modelos, estas variables latentes son incorporadas dentro de la función de utilidad de cada alternativa modal.

<sup>3</sup> Típicamente una variable indicadora corresponde a la evaluación que hace el encuestado sobre alguna pregunta del tipo "en una escala de 1 a 10, califique la calidad del servicio al cliente".

A continuación se presenta la estructura de un modelo MIMIC típico (Bollen, 1989, pp. 18):

$$\begin{aligned}\eta &= \Gamma x + \zeta \\ y &= \Lambda_y \eta + \varepsilon\end{aligned}\quad (1)$$

donde  $\eta_i$  representa un vector de variables latentes (seguridad operacional, comodidad);  $x$  (vector de dimensión  $q$ ) es un conjunto de variables observadas (variables socioeconómicas, características del modo, etc.) medidas sin error y que afectan la percepción de las variables latentes;  $\varepsilon$  y  $y$  (vector de dimensión  $p$ ), un conjunto de variables indicadoras que caracterizan la percepción de las variables latentes;  $\zeta_i$  y  $\varepsilon$  son variables aleatorias con media cero y matriz de covarianza  $\Psi$  y  $E$  respectivamente, y los vectores  $\Gamma$  y  $\Lambda_y$  denotan  $q + p$  parámetros a ser estimados. La ecuación 1 se denomina “modelo estructural” y la ecuación 2 se denomina “modelo de medición”.

Para que el modelo sea identificable basta que  $p \geq 2$  y  $q \geq 1$ ; también hay que fijar un coeficiente del vector  $\Lambda_y$  a fin de determinar la escala de  $\eta_i$ . Para que el modelo MIMIC sea identificable (además de que no hayan más parámetros que ecuaciones en (1) y (2)), se debe normalizar la escala de las variables latentes. Stapleton (1978) recomienda fijar en 1 cualquier elemento distinto de cero de cada columna de  $\Lambda_y$ .

La función indirecta de utilidad de la alternativa  $i$  ( $i = \text{bus, tren}$ ) para el individuo  $n$  considerando las variables latentes  $\eta$ , viene dada por:

$$U_i = z_i \alpha + \eta_i \beta + \mu_i \quad (3)$$

donde  $z$  es un vector de atributos como tiempo de viaje, costo de viaje, etc.;  $\alpha$  y  $\beta$  son parámetros a estimar asociados a las variables  $z$  and  $\eta$ ,  $U$  es utilidad aleatoria y  $\mu$ , un término de error aleatorio. Si se supone que  $\mu$  es un error aditivo que distribuye IID Gumbel con media cero, se obtiene la típica expresión del modelo Logit Binario (MNL):

$$P_{bus} = \frac{e^{V_{bus}}}{e^{V_{bus}} + e^{V_{tren}}} \quad (4)$$

$$P_{tren} = 1 - P_{bus} \quad (5)$$

Finalmente se debe integrar en una ecuación la probabilidad conjunta de observar una determinada elección y un conjunto de valores particulares para las variables indicadoras:

$$\begin{aligned}P(d_i = 1, y | z, x, \alpha, \beta, \Delta_y, \Gamma) = \\ \int_{\eta} P(d_i = 1 | z, \eta, \alpha, \beta) g(y | \eta, \Lambda_y) h(\eta | x, \Gamma) d\eta\end{aligned}\quad (6)$$

donde la variable binaria  $d_i = 1$  significa que la alternativa  $i$  fue elegida.

La ecuación (6) es una integral compleja que tiene que ser calculada numéricamente. Existen dos maneras de estimar la función de log-verosimilitud que corresponde a estos modelos híbridos, ya sea mediante técnicas de estimación simultánea o mediante técnicas de estimación secuencial. En ambos casos suele suponerse que la función de densidad  $h(\eta | x, \Gamma)$  distribuye Normal.

Si se adopta el segundo camino de estimación (estimación secuencial), el proceso es el siguiente. En primer lugar, se estima el modelo MIMIC que permite calcular el valor de las variables latentes. En la segunda etapa, el estimador puntual de cada variable latente es introducido en el modelo de elección discreta y se estima éste. Dado que las variables latentes son en rigor variables aleatorias, el cálculo de los errores estándar de los

coeficientes del modelo de elección discreta es subestimado; este se debe a que el modelo de elección discreta se estima con el valor esperado de la variable latente desconociendo la variabilidad en torno de este valor medio. Murphy y Topel (1985) proponen un método para corregir la estimación de los errores estándar, que aplica en este caso. Con dicha corrección, se obtienen errores estándar asintóticamente correctos. En este estudio, adoptaremos el enfoque de estimación secuencial y aplicaremos la corrección de Murphy y Topel (1985).

A los efectos de calcular la probabilidad de elegir la alternativa  $i$  antes cambios en cierta variable de control en fase predictiva o si se quiere calcular elasticidades, la ecuación correspondiente es:

$$\begin{aligned}P(d_i = 1 | z, x, \alpha, \beta, \Gamma) = \\ \int_{\eta} P(d_i = 1 | z, \eta, \alpha, \beta) h(\eta | x, \Gamma) d\eta\end{aligned}\quad (7)$$

Si  $P(d_i = 1 | z, \eta, \alpha, \beta)$  fuese una probabilidad logit, el cálculo de (7) supone integrar dicha probabilidad logit sobre la región de variación de  $\eta$ . Si bien esta integral no puede ser calculada mediante una fórmula cerrada, puede ser fácilmente resuelta utilizando simulación de Montecarlo.

En esta revisión, no profundizaremos en los aspectos teóricos de la formulación de este tipo de modelos híbrido ni en los aspectos econométricos de la estimación de los modelos. Excelentes referencias al respecto son Mc Fadden (1986), Ben Akiva et al. (1999), Walker (2001), Ben Akiva et al. 2002 y Golob (2003).

### 3. CASO DE ESTUDIO

Se encuestó a 600 usuarios de transporte público en terminales de buses y trenes, 300 usuarios de cada modo. Todas estas personas estaban a punto de realizar su viaje interurbano. En la encuesta se registraba el servicio de bus o tren que la persona estaba a punto de abordar y se le hacían varias preguntas relacionadas con el viaje. Se les pedía que evalúe en una escala del 1 (uno) al 7 (siete) a cada una de las siguientes variables indicadoras ( $y$ ) tanto para el modo bus como para el modo tren:

- La *conveniencia* de los horarios de viaje (mayor disponibilidad de viajes diarios)
- La *facilidad* para viajar con niños o bultos pesados
- El *aprovechamiento del tiempo* en el viaje para leer, escribir, trabajar o realizar otras actividades
- La *comodidad* durante el viaje (calidad de los asientos, espacios disponible, entre otras)
- La *puntualidad* en el cumplimiento del horario del viaje
- La *relajación* durante el viaje (capacidad para disfrutar el viaje)
- La *seguridad* al viajar (sensación de que no ocurra un accidente en el transcurso del viaje)
- La *seguridad* en el vehículo/coche tren (nivel de protección brindado por el vehículo en caso de accidente)
- La *seguridad* en cuanto a robo/hurto al interior del vehículo/coche tren.
- La *seguridad* en cuanto a robo/hurto en Terminales.

También se recabó información socioeconómica y se hicieron consultas que inquirían sobre los siguientes aspectos:

- Experiencias relativas a accidentes o robos sufridos en bus y/o tren. Se esperaba que las vivencias previas influyan las percepciones de seguridad de los usuarios con respecto al modo.

- Si el usuario viajaba o no con computador portátil. Se observó que las personas que viajan con computador personal, por lo general trabajan durante el viaje y además tienen una preocupación extra con respecto a la seguridad. Esta condición, podría resultar causa de las variables latentes consideradas.
- Siguiendo las recomendaciones de Lyons y Urry (2005) y de Lyons et al. (2007) se incluyeron preguntas relacionadas con la actividad que la persona realiza durante el viaje (aquella actividad que le dedicará la mayor cantidad de tiempo). Se espera que las actividades realizadas afecten la percepción de la comodidad.

### 3.1 Modelo MIMIC

El enfoque de estimación adoptado es el secuencial. En una primera etapa se estima el modelo MIMIC y en una segunda el modelo de elección discreta. Mediante un análisis de componentes principales, fue posible determinar tres factores que explicaban mayormente la varianza de la información contenida en las diez variables indicadoras ( $y$ ). Estos tres factores son la comodidad, la seguridad y la conveniencia horaria. En cuanto a la variable seguridad, la misma contiene aspectos de seguridad ciudadana y de seguridad operacional. En cuanto a la conveniencia horaria, esta variable indicadora quedó aislada, por lo que es considerada entonces una variable latente. Sin embargo, al estimar los modelos de ecuaciones estructurales, no se obtuvieron resultados sensatos en torno a ella y fue descartada del análisis de aquí en más.

Una vez realizado el análisis de componentes principales, se procedió con la estimación del modelo estructural, utilizando el software AMOS. Para definir qué elementos de  $A_y$  (indicadores) normalizar en el Modelo 1, se realizó lo siguiente:

- En primer lugar se estimó el modelo MIMIC con una normalización cualquiera de los elementos  $A_y$  (teniendo en cuenta las condiciones de Stapleton, (1978)).
- En segundo lugar se analizaron las columnas de la matriz  $A_y$  recientemente estimada; cada una contiene los coeficientes  $\lambda$  estimados y exactamente un “uno” producto de la normalización utilizada.
- Se escoge el mayor  $\lambda$  en valor absoluto de cada columna.
- Finalmente, se realiza una nueva estimación del modelo MIMIC normalizando en “uno” aquellos  $\lambda$  escogidos en iii).

Siguiendo este procedimiento, todos los elementos de la matriz  $A_y$  tendrán como valor máximo “uno”. En la Tabla 2 se puede observar el detalle de los indicadores que fueron normalizados siguiendo este procedimiento (aquellos indicadores que toman valor “1” son los normalizados). La idea de utilizar esta normalización es permitir una interpretación un poco más clara de la matriz  $A_y$  resultante.

La Tabla 1 presenta la estimación de la ecuación (1) para el Modelo 1, donde se observa el impacto que el conjunto de variables medibles ( $x$ ) tiene sobre las variables latentes ( $\eta$ ) comodidad y seguridad.

TABLA 1: Modelo Estructural de Variables Latentes – Modelo 1

Variable Objetiva	Comodidad Bus	Seguridad Bus	Comodidad Tren	Seguridad Tren
¿Ha visto o sufrido algún robo en bus o terminal de bus?		-0,1794 (-1,79)		
¿Ha visto o sufrido algún accidente en bus?		-0,4438 (-4,28)		
Desarrollo de Actividades Intelectuales	-0,1985 (-2,03)		0,1833 (2,78)	
Edad	0,0064 (2,3)	0,0087 (3,26)	0,0034 (1,78)	0,0059 (2,23)
¿Viaja en Grupo?	0,3216 (2,84)		0,2992 (3,86)	
Ingreso Alto	-0,4103 (-4,25)			
¿Viaja con notebook?		-0,2542 (-2,03)		0,2421 (2,91)
¿Ha visto o sufrido algún accidente en tren?				-0,1875 (-1,33)
Índice de Fatalidades		-0,2088 (-1,35)		0,4738 (2,99)
Desarrollo de Actividades Sociales	-0,2315 (-1,45)		-0,1072 (-0,99)	
¿Ha visto o sufrido algún robo en tren o terminal de tren?				-0,4135 (-2,23)

Notas: Test-t en paréntesis. Si la persona lee, estudia o trabaja durante el viaje, se considera que realiza actividades intelectuales. Si la persona gana sobre los \$400.000 mensuales, se considera de ingreso alto. Índice de fatalidades, medido en fatalidades por pasajero transportado de buses interurbanos. Si la persona conversa con otro pasajero o juega con sus hijos/nietos/sobrinos durante el viaje, se considera que realiza actividades sociales. El haber sido víctima o testigo de un accidente o robo, afecta negativamente la percepción de seguridad de los usuarios. El viajar con notebook afecta negativamente la sensación de seguridad del bus, ya que las personas temen ser víctimas de un robo. Este efecto es contrario en el tren. El realizar algún tipo de actividad intelectual que requiera concentración durante el viaje, impacta negativamente en la sensación de comodidad del bus y positivamente en el tren. Esto se debe a que el tren presenta mejores condiciones para el desarrollo de estas actividades: asientos más amplios, movimiento suave y constante del carro, electricidad para poder conectar el notebook (disponible en primera clase), entre otras. Por otra parte, el desarrollo de actividades sociales influye negativamente en la sensación de comodidad de ambos modos; sin embargo, este impacto es mayor en el caso del bus.

En relación al tema de la seguridad operacional, el tipo de servicios ferroviarios que operan hoy día no registra víctimas fatales entre los viajeros del ferrocarril en su historia reciente<sup>4</sup>. Por el contrario, si hay usuarios de servicios de buses que fallecen producto de un accidente de bus. El índice de accidentes entrega la probabilidad que un usuario de bus fallezca en un accidente fatal y se expresa como el número de buses-kilómetros que deben recorrerse por la ruta en que se circula para que un pasajero de bus muera en un accidente. Este índice afecta a la percepción de la seguridad de los dos modos de transporte: a mayor valor del índice empeora la percepción de la seguridad operacional del modo bus y mejora la percepción de la misma para el modo tren.

Por otro lado, el índice de fatalidades tuvo el efecto esperado: a medida que aumentan los accidentes de buses interurbanos con fallecidos para las rutas en cuestión, disminuye la sensación de seguridad del bus y simultáneamente, el tren se percibe como un modo comparativamente más seguro. La variable Viaja en Grupo influye positivamente en la sensación de comodidad de los usuarios, tanto de tren como de bus. Es de esperar que la sensación de comodidad de los usuarios sea mayor si la persona que se sienta al lado es alguien conocido. El “Ingreso Alto” influye negativamente en la sensación de comodidad del bus (las personas de ingresos más altos son “usuarios más exigentes”). Por último, al igual que los resultados reportados por Vredin Johansson et al. (2005) y Temme et al. (2007), la edad influye positivamente en las sensaciones de comodidad y de seguridad del bus y del tren.

La Tabla 2 muestra la estimación de la ecuación 2 para el Modelo 1, que muestra la vinculación entre las variables latentes y las variables indicadoras ( $y$ ). Todos los efectos son consistentes, una mejor percepción de la comodidad y/o seguridad produce un aumento en la nota que las personas colocan a las variables indicadoras. En particular, destacamos la percepción de la seguridad. La variable latente seguridad se manifiesta a través de indicadores de seguridad ciudadana y de seguridad operacional.

<sup>4</sup> Esta información fue comunicada por la Gerencia de la Empresa de Ferrocarriles del Estado.

En otras palabras, ambos tipos de seguridad se consolidan en una variable latente seguridad. Esta variable también se manifiesta a través de la relajación que los pasajeros pueden lograr durante el viaje.

Probablemente, a mayor percepción de seguridad mayor nivel de relajación se alcance. La relajación es la única variable indicadora que se ve afectada tanto por la seguridad percibida como por la comodidad del viaje: un pasajero manifiesta sentirse más relajado no sólo si está más cómodo sino también si se siente más seguro.

**TABLA 2: Modelo de Medición de Variables Latentes – Modelo 1**

	Comodidad Bus	Seguridad Bus	Comodidad Tren	Seguridad Tren
La facilidad para viajar con niños o bultos pesados	0,7429 (12,75)		0,9877 (11,91)	
El aprovechamiento del tiempo en el viaje para leer, escribir, trabajar o realizar otras actividades	0,8100 (12,91)		0,8427 (10,86)	
La comodidad durante el viaje (calidad de los asientos, espacios disponible, entre otras)	1 (n.a.)		1 (n.a.)	
La puntualidad en el cumplimiento del horario del viaje	0,7736 (13,51)		0,7723 (10,93)	
La relajación durante el viaje (capacidad para disfrutar el viaje)	0,7384 (12,09)	0,8315 (9,78)	0,7750 (11,52)	0,5024 (8,06)
La seguridad al viajar (sensación de que ocurra un accidente en el transcurso del viaje)		0,9986 (10,79)		0,7872 (9,12)
La seguridad en el vehículo/coche tren (nivel de protección brindado por el vehículo en caso de accidente)		1 (n.a.)		0,9934 (11,15)
La seguridad en cuanto a robo/hurto al interior del vehículo/coche tren		0,9621 (11,33)		1 (n.a.)
La seguridad en cuanto a robo/hurto en Terminales		0,2976 (4,53)		0,1834 (4,09)

Nota: test t entre paréntesis. Para los indicadores normalizados se presenta (n.a.), es decir, "no aplica".

Similarmente, en las Tablas 3 y 4 se presentan los resultados de las ecuaciones (1) y (2) para el Modelo 2, que normaliza de manera diferente el modelo MIMIC. El proceso para definir qué elementos de  $\Lambda_y$  (indicadores) normalizar en el Modelo 2 fue más sencillo. Simplemente se escogieron 4 indicadores al azar (uno por cada columna de la matriz  $\Lambda_y$ ) para ser normalizados, tomando en cuenta las recomendaciones de Stapleton (1978), es decir, fijar en 1 cualquier elemento distinto de cero de cada columna de  $\Lambda_y$ . El detalle de los indicadores normalizados puede observarse en la Tabla 10 (aquellos indicadores que toman valor "1" son los normalizados). Producto de esta normalización se observan algunas diferencias en los resultados del modelo MIMIC en relación variable latente "seguridad". Para el Modelo 1, el indicador que afecta mayormente a la variable latente "Seguridad Bus" es "La seguridad en el vehículo/coche tren (nivel de protección brindado por el vehículo en caso de accidente)", sin embargo, para el Modelo 2, tal indicador pasa a ser "La seguridad en cuanto a robo/hurto al interior del vehículo/coche tren". Por otra parte, para los indicadores asociados a la comodidad, es decir, "Comodidad Bus" y "Comodidad Tren", se mantiene el orden de importancia de los efectos de los indicadores sobre dichas variables latentes. Algo similar se observa en los resultados del modelo estructural para el Modelo 1 y el Modelo 2. Las variables objetivas asociadas a la comodidad, no presentan grandes diferencias entre ambos modelos (en términos de magnitud), sin embargo, para el Modelo 2, las variables asociadas a robos y accidentes en bus o tren, pierden importancia en términos de magnitud con respecto al Modelo 1.

**TABLA 3: Modelo Estructural de Variables Latentes - Modelo 2**

Variable Objetiva	Comodidad Bus	Seguridad Bus	Comodidad Tren	Seguridad Tren
¿Ha visto o sufrido algún robo en bus o terminal de bus?		-0,0534 (-1,74)		
¿Ha visto o sufrido algún accidente en bus?		-0,1321 (-3,27)		
Desarrollo de Actividades Intelectuales	-0,1475 (-2,02)		0,1810 (2,74)	
Edad	0,0048 (2,3)	0,0026 (2,75)	0,0033 (1,77)	0,0011 (2,00)
¿Viaja en Grupo?	0,2389 (2,81)		0,2956 (3,78)	
Ingreso Alto	-0,3048 (-4,13)			
¿Viaja con notebook?		-0,1888 (-2,02)		0,2391 (2,84)
¿Ha visto o sufrido algún accidente en tren?				-0,0344 (-1,30)
Índice de Fatalidades		-0,0622 (-1,27)		0,0869 (2,53)
Desarrollo de Actividades Sociales	-0,1720 (-1,44)		-0,1059 (-0,98)	
¿Ha visto o sufrido algún robo en tren o terminal de tren?				-0,0758 (-1,98)

Nota: test t entre paréntesis.

**TABLA 4: Modelo de Medición de Variables Latentes – Modelo 2**

Indicador	Comodidad Bus	Seguridad Bus	Comodidad Tren	Seguridad Tren
La facilidad para viajar con niños o bultos pesados	1 (n.a.)		1 (n.a.)	
El aprovechamiento del tiempo en el viaje para leer, escribir, trabajar o realizar otras actividades	1,09 (11,14)		0,8532 (11,00)	
La comodidad durante el viaje (calidad de los asientos, espacios disponible, entre otras)	1,3460 (12,75)		1,0125 (11,91)	
La puntualidad en el cumplimiento del horario del viaje	1,0413 (11,13)		0,7819 (9,92)	
La relajación durante el viaje (capacidad para disfrutar el viaje)	0,9939 (10,03)	1 (n.a.)	0,7846 (10,14)	1 (n.a.)
La seguridad al viajar (sensación de que ocurra un accidente en el transcurso del viaje)		2,7935 (4,93)		2,7397 (4,18)
La seguridad en el vehículo/coche tren (nivel de protección brindado por el vehículo en caso de accidente)		3,3549 (4,75)		4,2930 (4,11)
La seguridad en cuanto a robo/hurto al interior del vehículo/coche tren		3,3597 (4,52)		5,4174 (4,01)
La seguridad en cuanto a robo/hurto en Terminales		3,2322 (4,47)		5,4536 (4,09)

Nota: test t entre paréntesis. Para los indicadores normalizados se presenta (n.a.), es decir, "no aplica".

### 3.2 Modelo de Elección Discreta

En la Tabla 5, se muestran los resultados de un modelo logit (Modelo Base) sin variables latentes. En este modelo, todas las variables explicativas son estadísticamente significativas y presentan el signo correcto. Los coeficientes del costo de viaje y tiempo de viaje son negativos e implican un valor subjetivo del tiempo de viaje<sup>5</sup> de \$44 el minuto, un valor acorde a los ingresos observados en la muestra<sup>6</sup>. Se estimaron dos interacciones entre la variable costo y las variables binaria clase de viaje (turista o primera) y si el viaje lo paga la empresa o el viajero. En ambos casos se obtienen coeficientes positivos; es decir, si se viaja en clase primera o si paga la empresa la utilidad marginal del ingreso disminuye. Se estimó también una interacción entre el tiempo de viaje y el género de las personas: la desutilidad provocada por un minuto más de viaje es mayor para los hombres que para las mujeres.

La Tabla 5 muestra los resultados del Modelo 1 que contiene las dos variables latentes comodidad y seguridad. Este modelo de elección discreta fue estimado como un modelo logit binario en

<sup>5</sup> Todos los valores subjetivos del tiempo fueron calculados por enumeración muestral.

<sup>6</sup> El valor subjetivo del tiempo de viaje para viajes interurbanos en vehículos suele ser bastante mayor. De todas maneras, se trata de individuos con ingresos personales inferiores a los conductores de automóvil interurbano.

una segunda etapa en el marco de la estimación secuencial y los test-t fueron calculados siguiendo la corrección de Murphy y Topel (1985). Cuatro de los cinco coeficientes mencionados en el párrafo anterior siguen siendo significativos: costo de viaje, tiempo de viaje y las interacciones entre el costo y la variable binaria clase del viaje y entre costo y la variable binaria si el viaje lo paga la empresa o el viajero. El valor subjetivo del tiempo de viaje apenas sube a \$43 el minuto. Las dos variables latentes, comodidad y seguridad, son significativas y presentan el signo correcto: a mayor percepción de comodidad y/o seguridad mayor es la utilidad modal. También hay una sensible mejora según el índice de ajuste  $Rho^2$ . Más definitivo es el test de razón de verosimilitud entre el Modelo 1 y el Modelo Base como versión restringida indica que el Modelo 1 es estadísticamente superior.

La última columna de la Tabla 5 muestra los resultados del Modelo 2. Producto de la diferente normalización del modelo MIMIC, en esta ocasión, la variable latente seguridad deja de ser significativa para un valor de  $\alpha = 0.05$ . De esta manera, vemos que la normalización del modelo MIMIC no es indiferente: con dos normalizaciones distintas, en un caso concluimos que la variable latente seguridad es estadísticamente significativa y en otro caso, no. Como se aprecia en la tabla, los estimadores puntuales del vector de variables  $z$  no cambian<sup>7</sup>.

Desde el punto de vista económico, un resultado que es de alta relevancia es la elasticidad de la probabilidad de elección del modo  $i$  ante cambios en alguna variable de control  $(x, z)$ . La Tabla 6 muestra las elasticidades de la probabilidad de elegir la alternativa bus o tren ante cambios en los costos de viaje y/o tiempos de viaje del mismo modo de transporte. Las elasticidades mostradas son la elasticidad promedio; es decir, se calculó la elasticidad para cada individuo de la muestra y luego se promedió (método de enumeración muestral (sample enumeration)). Estas elasticidades fueron calculadas para el Modelo Base, el Modelo 1 y el Modelo 2<sup>8</sup>. La Tabla 7 muestra las elasticidades de la probabilidad de elegir la alternativa bus o tren ante un cambio porcentual en el índice de fatalidades<sup>9</sup>. Se observa que esta elasticidad es bastante menor que las elasticidades del costo y tiempo del viaje. Además la elasticidad cambia de manera importante según sea el modelo 1 o el modelo 2. En particular, para el Modelo 2, la elasticidad es casi cuatro veces menor.

TABLA 5: Modelo Logit Binario

Parámetro $\theta$	Variable	Modelo Base	Modelo 1	Modelo 2
Cvia (miles de pesos)	costo de viaje	-0.8030 (-9.12)	-0.7410 (-8.28)	-0.7408 (-8.35)
Cvia_cat	Interacción entre costo y categoría de viaje (1 si es categoría alta, 0 e.o.c)	0.4820 (9.83)	0.4286 (8.58)	0.4291 (8.51)
cvia_pas_e	Interacción entre costo y si la empresa paga el pasaje (1 si la empresa paga, 0 e.o.c)	0.3018 (3.32)	0.2183 (2.36)	0.2171 (2.31)
Tvia (minutos)	tiempo de viaje	-0.02104 (-3.90)	-0.0196 (-3.524)	-0.0202 (-3.61)
Tvia_sex	interacción entre tiempo de viaje y sexo (1 si la es hombre, 0 e.o.c)	-0.003897 (-1.14)	0,0001 (-0,04)	0,0005 (-0,12)
seg	variable latente seguridad	n.a.	1,1150 (2,55)	1,6830* (1,13)
com	variable latente comodidad	n.a.	1,6120 (5,1)	2,2240 (5,53)
Kbus	constante modal del bus (corregida)	0.9499 (-0.75)	1,1440 (1,17)	1,0580 (0,97)
Ktren	constante modal del tren (corregida)	-3,9087	-3,9090	-3,9090
Log-Likelihood	Log verosimilitud	-287,02	-267,26	-267,57
	$Rho^2$	0.310	0.357	0.357

Nota: Test t entre paréntesis. Además para el Modelo 1 y el Modelo 2, los test t se presentan corregidos según Murphy y Topel (1985).

\*No es significativo para un valor de  $\alpha=0.05$ . Las constantes modales fueron corregidas siguiendo la fórmula propuesta señalada por Cosslett (1981, ecuación 2.81) cuando se trabaja con muestras basadas en la elección.

TABLA 6: Elasticidades Costo y Tiempo de Viaje

Variable	Elasticidades Modelo Base		Elasticidades Modelo 1		Elasticidades Modelo 2	
	Bus	Tren	Bus	Tren	Bus	Tren
	Costo	-1,263	-1,578	-1,424	-2,238	-1,562
Tiempo de viaje	-1,934	-1,233	-1,801	-1,414	-2,018	-1,586

Nota: Las elasticidades agregadas fueron calculadas vía enumeración muestral.

TABLA 7: Elasticidades con Respecto al Índice de Fatalidades

Variable	Elasticidades Modelo 1	Elasticidades Modelo 2
	Bus	Bus
Índice de fatalidades	-0,086	-0,024

Nota: Las elasticidades agregadas fueron calculadas vía enumeración muestral.

#### 4. CONCLUSIONES

Hemos desarrollado una nueva aplicación de un modelo integrado de variables latentes y elecciones discretas para estudiar el impacto que la percepción de seguridad y comodidad tienen sobre la elección de modo de transporte público para viajes interurbanos. Hemos expandido el conjunto de variables que explican la percepción de las variables latentes, prestando especial importancia a ciertas características del viaje.

A partir de los modelos estimados, se concluye que la inclusión de variables latentes contribuye a una mejor modelación. Así, se obtienen modelos con mejor ajuste estadístico y se logra también una mejor comprensión del fenómeno bajo análisis. Sin embargo, hemos visto que ciertos resultados obtenidos dependen significativamente de la normalización adoptada para el modelo de variables latentes (modelo MIMIC). En particular, al momento de estimar la disposición al pago por reducción de víctimas fatales para usuarios de transporte público y/o la elasticidad de la probabilidad de elección del bus con respecto al índice de seguridad vial, los resultados difieren de manera importante según la normalización adoptada. Se espera que futuras investigaciones puedan arrojar más luces en relación a este tema, a fin de que los resultados de estos modelos híbridos no solo permitan obtener mejores ajustes estadísticos sino también proveer información económica relevante para la toma de decisiones y/o la evaluación social de proyectos.

<sup>7</sup> El Modelo 2 presenta un valor subjetivo del tiempo de viaje de \$44 el minuto.

<sup>8</sup> Para el modelo base, la elasticidad se calcula utilizando la típica fórmula correspondiente a los modelos logit (Ortúzar y Willumsen, 2001). Para los modelos 1 y 2, el cálculo de la elasticidad es más complejo. Las probabilidades de elección se calculan aplicando la ecuación (7) mediante simulación de Montecarlo, considerando mil réplicas por individuo. La elasticidad de la probabilidad de elegir la alternativa  $i$  ante cambios en determinada variable  $z$  (o  $x$ ) es igual a

$$\frac{\partial P(d_i = |z, x, \alpha, \beta, \Gamma)}{\partial z} \frac{z}{P(d_i = |z, x, \alpha, \beta, \Gamma)} \left( 0 \frac{\partial P(d_i = |z, x, \alpha, \beta, \Gamma)}{\partial \alpha} \frac{x}{P(d_i = |z, x, \alpha, \beta, \Gamma)} \right)$$

Estas derivadas fueron calculadas numéricamente.

<sup>9</sup> En la nota a pie 8, se explica como se calculan las elasticidades.

## AGRADECIMIENTOS

Agradecemos a los proyectos FONDECYT 1050672 y 1085287 y al Instituto Milenio Sistemas Complejos en Ingeniería (Proyecto 05 004F) por el apoyo financiero provisto.

## REFERENCIAS

Ben-Akiva, M., D. McFadden, T. Gärling, D. Gopinath, J. Walker, D. Bolduc, A. Börsch-Supan, P. Delquié, O. Larichev, T. Morikawa, A. Polydoropoulou, and V. Rao (1999) "Extended Framework for Modeling Choice Behavior", **Marketing Letters** 10(3), 187-203.

Bollen, K. A. (1989). **Structural Equations with Latent Variables**. Nueva York, EE.UU.: John Wiley and Sons, Inc.

Golob, T. (2003). Structural equation modeling for travel behavior research. **Transportation Research** 37B, 1-25.

Lyons G., Jain J. & Holley D. (2007). The use of travel time by rail passengers in Great Britain. **Transportation Research Part A**, 41, 1, 107-120.

Lyons G., Urry J. (2005). Travel time use in the information age. **Transportation Research Part A**, 39, 2-3, 257-276.

McFadden, D. (1986). The Choice Theory Approach to Marketing Research. **Marketing Science**, 5 (4), 275-97.

McFadden, D. (2000). "Disaggregate Behavioral Travel Demand's RUM Side: A 30-Year Retrospective", documento presentado en la IATBR (**International Association of Travel Behavior Research**). Gold Coast, Queensland, Australia.

Murphy K. & Topel H. (1985). Estimation and Inference in Two-Step Econometric Models. **Journal of Business and Economic Statistics**, 1985, 3(4), 370-379.

Stapleton David (1978). Analyzing Political Participation Data with a Mimic Model. **Sociological Methodology**, 9, 52-74.

Temme Dirk, Paulssen Marcel & Dannewald Till (2007). Integrating latent variables in discrete choice models - How higher-order values and attitudes determine consumer choice. SFB 649 Discussion Papers SFB649DP2007-065, Sonderforschungsbereich 649, Humboldt University, Berlin, Germany.

Vredin Johansson, M., Heldt, T., & Johansson, P. (2005). **Latent variables in a travel mode choice model: Attitudinal and behavioural indicator variables**. Working Paper 2005:5, Uppsala University.

Walker, J. (2001). **Extended Discrete Choice Models: Integrated Framework, Flexible Error Structures and Latent Variables**. PhD Thesis, MIT.