

## **Incorporación de atributos intangibles en modelos de elección discreta**

**Gustavo García Melero**

Profesor Asociado, Pontificia Universidad Católica de Valparaíso, Chile

**Pablo Coto Millán**

Catedrático de Universidad, Universidad de Cantabria, España

**Rubén Sainz González**

Profesor Contratado Doctor, Universidad de Cantabria, España

### **RESUMEN**

La principal herramienta en la modelación de la demanda de transporte durante muchos años ha sido el Logit Multinomial (MNL), pero sus limitaciones, fundamentalmente el supuesto de independencia de alternativas irrelevantes, han motivado diferentes mejoras, destacando entre ellas el Logit Mixto (ML).

El ML permite mayor flexibilidad en las preferencias individuales, aunque no por ello ha colmado las expectativas de los modeladores, ya que el objetivo es representar la realidad de la forma más veraz posible.

De este modo, se ha planteado la inclusión de elementos subjetivos, a través de modelos de utilidad híbridos que consideran atributos tangibles, fácilmente identificables para cada alternativa, así como elementos intangibles asociados a las percepciones y actitudes de los individuos, expresados a través de variables latentes.

El principal objetivo de este trabajo es identificar las variables latentes relevantes en la modelación de la elección modal en el ámbito urbano, así como la estimación de MNL y ML con y sin variables latentes, con el fin de determinar si la inclusión de variables latentes mejora los resultados en la estimación de los modelos.

La aplicación empírica se efectúa con datos obtenidos a partir de una encuesta de Preferencias Reveladas realizada en Santander (España), a la que se incorporó un cuestionario con indicadores de percepción, que permiten captar las variables latentes.

La principal conclusión es que los modelos ML se muestran superiores a los MNL y, en ambos casos, la incorporación de variables latentes permite obtener mejores resultados.

## 1. INTRODUCCIÓN

Durante muchos años, el Modelo de Elección Discreta (MED) más empleado en la modelación de la demanda de transporte ha sido el MNL. El supuesto de independencia de alternativas irrelevantes ha supuesto una limitación, que ha generado la propuesta de diferentes alternativas en este campo, destacando actualmente entre ellas el ML.

El ML ha proporcionado a los modeladores una forma más flexible de representar la realidad a través de las preferencias individuales, aunque no es capaz de recoger todos los elementos determinantes de la demanda de transporte que se observan en la práctica.

De este modo, se están desarrollando otras mejoras en la modelación de la demanda de transporte como es la inclusión de elementos subjetivos en los MED. Este asunto se ha convertido en uno de los principales temas de análisis y discusión en la literatura actual a nivel mundial, una vez que se ha podido avanzar tanto en su marco teórico como en las limitaciones para su aplicación empírica.

En la última década se han planteado modelos de utilidad híbridos (Ben-Akiva et al., 2002; Walker y Ben-Akiva, 2002) que consideran atributos tangibles, fácilmente identificables para cada alternativa (variables explicativas clásicas, por ejemplo el tiempo y el coste), así como elementos intangibles asociados a las percepciones y actitudes de los individuos, que son expresados a través de variables latentes.

Cabe destacar la formulación de un nuevo marco de utilidad aleatoria generalizado (Ben-Akiva et al, 1999), en el que se intenta superar las limitaciones del modelo estándar. Los avances en este sentido han sido muy importantes en los últimos años, siendo posible la delimitación de las ampliaciones introducidas:

1. Estructuras de error flexibles (*logit kernel o mixed logit*)
2. Combinación de datos de Preferencias Reveladas y Declaradas
3. Clases latentes
4. Variables latentes

El modelo estándar de comportamiento de los individuos viene sufriendo serias críticas por parte de los principales investigadores en este campo, siendo una de las principales la no consideración de aquellos aspectos ligados a las percepciones y actitudes de los individuos (heterogeneidad).

De este modo, la inclusión de elementos subjetivos en MED, mediante variables latentes, ha resurgido con fuerza en la literatura internacional (Ashok et al., 2002; Bolduc et al., 2008; Bollen et al., 2007; Daziano y Rizzi, 2015; Maldonado, 2014; Paulsen et al., 2014; Temme et al., 2008).

Este tipo de modelos han sido utilizados en mayor medida en el campo de la ingeniería, donde se han desarrollado mediante su aplicación a los modelos de localización residencial (Walker y Li, 2007) y seguridad vial (Moen, 2007), entre otros. En menor medida se han utilizado en el campo de la economía o el marketing (O'Neill, V. et al., 2014), derivándose actualmente un gran potencial de utilización en estas áreas.

Estos modelos híbridos son una herramienta fundamental para comprender el comportamiento de los individuos, ya que consideran aspectos psicológicos asociados a las percepciones de la población ante su realidad.

En este trabajo, se prestará especial atención a los MED híbridos con variables latentes, y se identificarán las variables latentes más relevantes (y su grado de incidencia), en la modelación de situaciones de elección modal en el ámbito urbano.

Por tanto, en el siguiente apartado se presentará la metodología empleada. En el tercer apartado, se presentarán los datos utilizados. Los resultados obtenidos de las estimaciones econométricas se presentarán en el siguiente. Finalmente, se entregarán las principales conclusiones y se realizará una discusión de los resultados.

## 2. METODOLOGÍA

Con el fin de modelar la demanda de transporte en el ámbito urbano, se utilizarán varios tipos de MED, específicamente el MNL y el ML, lo cual permitirá establecer cuál de ellos representa en mayor medida la realidad.

Se plantearán distintas especificaciones econométricas, entre las cuales deben destacarse las que integren dentro de las variables explicativas a las variables latentes relevantes.

Los MED se basan en que la probabilidad de que los individuos elijan una determinada alternativa, viene dada por sus características socioeconómicas y por la utilidad relativa de la alternativa.

La utilidad medible se define generalmente como una combinación lineal de variables, donde cada variable representa un atributo de la alternativa o del viajero, en tanto que los coeficientes representan la influencia relativa de cada atributo.

La base teórica de los MED es la Teoría de la Utilidad Aleatoria, la cual asume:

1. Los individuos ( $Q$ ) actúan racionalmente y eligen siempre la alternativa que maximiza su utilidad neta personal, sujeta a sus restricciones.
2. Existe un conjunto de alternativas disponibles  $A = \{A1, \dots, A2, \dots, An\}$  y un conjunto de atributos medibles de los individuos y de las alternativas ( $X$ ). Cada individuo dispone de un conjunto  $x \in X$  de atributos, el cual se enfrenta a un conjunto de alternativas

disponibles  $A(q) \in A$ .

3. Cada alternativa  $A_j \in A$  tiene asociada una utilidad  $U_{jq}$  para cada individuo  $q$ . El modelador desconoce algunos elementos considerados por el individuo cuando hace una elección, por lo tanto se considera que la utilidad  $U_{jq}$  tiene dos componentes:
  - Componente sistemática ( $V_{jq}$ ), que es función de los atributos medibles  $x$
  - Componente aleatoria ( $\varepsilon_{jq}$ ), que representa las preferencias individuales, y además incluye los errores de medición del modelador.

A partir de los supuestos anteriores, el modelador puede proponer que la utilidad vendrá dada por:

$$U_{jq} = V_{jq} + \varepsilon_{jq} \quad (1)$$

$$V_{jq} = \sum_k \theta_{kj} X_{jkq} \quad (2)$$

El individuo  $q$  elige la alternativa que le proporciona su máxima utilidad, es decir elige  $A_j$  si y sólo si:

$$U_{jq} \geq U_{iq} \forall A_i \in A(q) \quad (3)$$

Por lo tanto, la probabilidad de elegir la alternativa  $A_j$  viene dada por:

$$P_{jq} = \text{Prob}\{\varepsilon_{iq} \leq \varepsilon_{jq} + (V_{jq} - V_{iq}), \forall A_i \in A(q)\} \quad (4)$$

donde la distribución de los residuos  $\varepsilon$  es desconocida.

El MNL es el MED más sencillo y el más popularmente utilizado; dicho modelo puede ser generado aceptando que los residuos aleatorios se distribuyen Gumbel IID.

Por otro lado, el ML es un modelo muy flexible que puede aproximar cualquier MED ya que permite considerar patrones muy generales de correlación, variaciones en las preferencias y heterocedasticidad mediante una correcta especificación de parámetros y variables. Presenta una función de utilidad del siguiente tipo:

$$U_{jqt} = \theta_q X_{jqt} + \varepsilon_{jqt} \quad (5)$$

La posterior inclusión de variables latentes en los MED permite considerar aspectos psicológicos (percepciones y actitudes) en la modelación del comportamiento de los individuos, aumentando la capacidad explicativa del modelo. Por otro lado, las variables latentes permiten incorporar implícitamente las características socioeconómicas a nivel

individual, así como sus apreciaciones respecto de las alternativas disponibles.

Al incluir variables latentes se obtiene una función de utilidad expresada por la ecuación 8, donde  $\beta_{il}$  son los parámetros a calibrar asociados a dichas variables. Las variables latentes no son conocidas, por lo que el MED debe acompañarse por un modelo MIMIC (Multiple Indicators and Multiple Causes), que está integrado por las ecuaciones estructurales

$$\eta_{ilq} = \sum_r \alpha_{ilr} s_{iqr} + v_{ilq} \quad (6)$$

y de medición

$$y_{ipq} = \sum_l \gamma_{ilp} \eta_{ilq} + \zeta_{ipq} \quad (7)$$

, abordándose en conjunto:

$$U_{iq} = \sum_k \theta_{ik} X_{ikq} + \sum_l \beta_{il} \eta_{ilq} + \varepsilon_{iq} \quad (8)$$

donde  $\eta_{ilq}$  son las variables latentes,  $s_{iqr}$  son las características de los individuos y de las alternativas,  $y_{ipq}$  son los indicadores de percepción,  $i$  indica una alternativa,  $l$  indica a una variable latente,  $q$  indica a un individuo,  $r$  indica una variable explicativa y  $p$  indica a un indicador.

### 3. OBTENCIÓN Y ANÁLISIS DE DATOS

La obtención de los datos necesarios para el desarrollo de este trabajo de investigación ha requerido de una estrategia mixta de investigación social, mediante la combinación de tres métodos cualitativos: Mega Grupos Focales (MGF), Talleres Participativos y Grupos Focales (GF). (Figura 1)

En su implementación debe estar presente, al menos, un moderador y un observador por cada sesión, con el fin de que los objetivos establecidos previamente en la pauta-guía se cumplan.

Para llevar a cabo el proceso de participación ciudadana, se ha considerado que la forma de contar con individuos representativos era a través de la convocatoria de las Asociaciones de Vecinos (AAVV), al ser representantes de la ciudadanía.

Los MGF se han realizado con el fin de atraer a la población objetivo al proceso de participación, desarrollándose dos ejercicios de este tipo en días diferentes, con el fin de extraer la información del primero, que sirviese de input en el segundo de ellos.

En los MGF se deben tratar temas genéricos y de interés común, con el fin de facilitar la interacción entre los asistentes. Además, se debe tener un criterio de selección de los

participantes a cada uno de los MGF, que estará compuesto entre 20 y 30 personas, con el fin de que las problemáticas sean comunes y se fomente el debate entre ellos.

Por otro lado, los Talleres Participativos se deben realizar dentro de los propios MGF, seleccionando problemáticas ya declaradas para tratarlas en varios sub-grupos de 3 o 4 miembros. De esta forma, es posible generar un debate más acotado, que permita trasladar al resto de miembros del MGF las principales conclusiones obtenidas para su puesta en común, así como plantear soluciones.

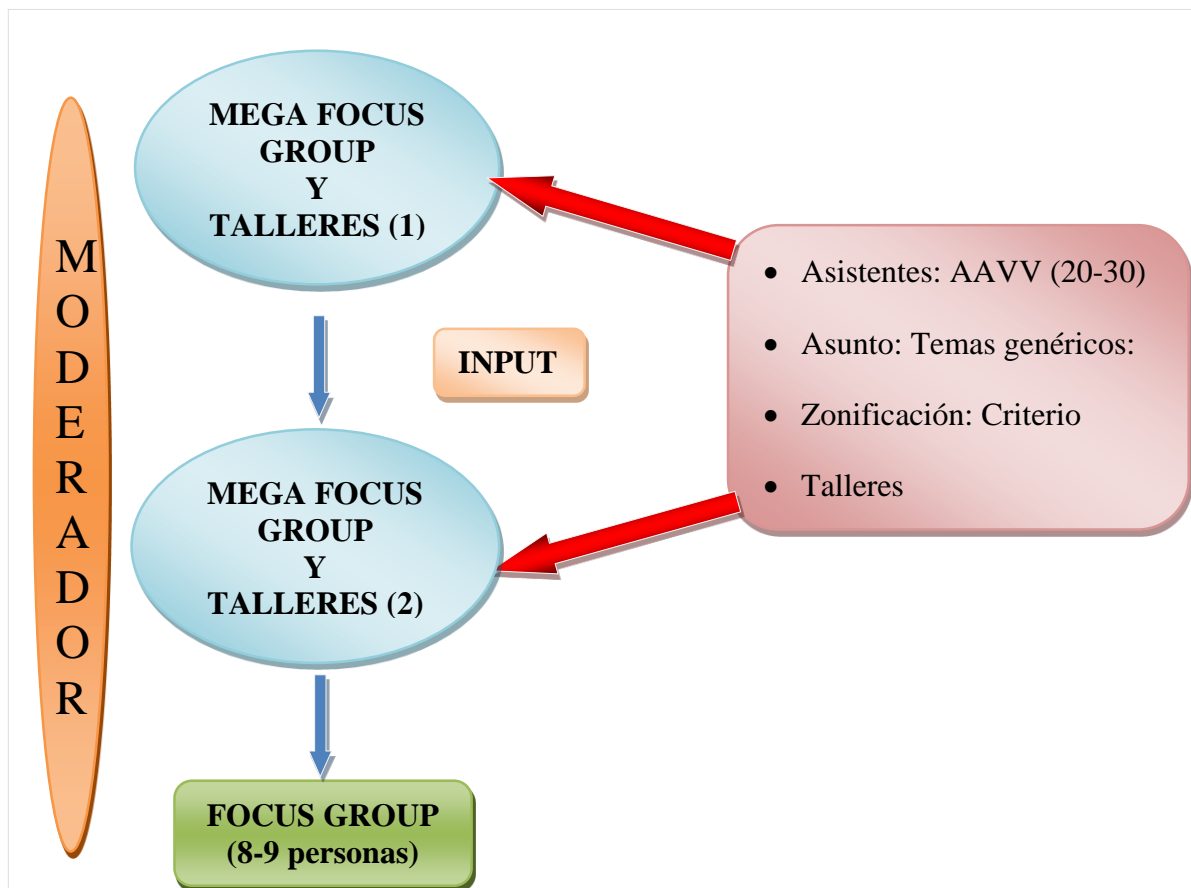
Una vez que se han determinado las principales problemáticas que afectan a la movilidad sostenible, es posible establecer las variables latentes que son consideradas por los individuos y los indicadores de percepción más relevantes para la población.

Posteriormente, se debe realizar un GF unos días después de los MGF, en el que el número de participantes se reducirá a menos de 10 asistentes, los cuales también serán representantes de AAVV. En este caso, no se discriminará entre ellas en base a ningún criterio, ya que se pretende conocer si ante un mismo cuestionario de indicadores de percepción se responde de forma similar.

Además, es interesante valorar el hecho de contrastar las problemáticas que afectan a cada uno, en un marco de puesta en común, una vez que se dispone de la información extraída de los MGF.

El funcionamiento de los MGF y de los GF es similar, siendo la única diferencia entre ambos el hecho de que el número de asistentes es mayor en el MGF. Las sesiones son orientadas en todo momento por el moderador, con el fin de no desviar el debate de los objetivos fijados previamente.

Una vez que se ha obtenido toda la información necesaria de la estrategia de participación ciudadana, es necesario realizar el diseño de la versión inicial de la encuesta. Se realizará un proceso de mejora, hasta que se llegue finalmente al diseño de la encuesta piloto, que será sometida a un proceso de evaluación y rediseño, el cual permitirá definir la encuesta definitiva (Ortúzar y Willumsem, 2001).



**Figura 1. Esquema metodológico**

Se realizaron 1182 encuestas domiciliarias de movilidad en la ciudad de Santander, en las que se cuestionaba por las características personales, por los viajes habituales y por la valoración que entregaban a los indicadores de percepción para cada modo de transporte.

De este modo, respecto a la información recabada se puede destacar lo siguiente:

1. Características de los individuos

En primer lugar, la muestra de encuestados se compone en un 54% de mujeres mientras que el 46% son varones.

Teniendo en cuenta los rangos de edad de los encuestados, el mayor porcentaje (22%) se encuentran representados por personas de entre 45 y 54 años de edad. A estos le siguen con un 20% las personas de más de 65 años, con un 16% el grupo de entre 55 y 64 años, con un 16% las personas con edades comprendidas entre los 35 y 44 años, y por último con un 12% las personas jóvenes de menos de 24 años de edad.

El 72% de los encuestados afirman tener licencia para conducir en contraposición a un 28% que no disponen de ella.

El 32% de los encuestados disponen de estudios superiores universitarios y el 29%

tiene estudios de EGB completados. Tan solo el 5% no disponen de estudios y un 17% tienen bachillerato o una formación profesional de grado medio o superior.

La mayor parte de los encuestados (48%) disponen de ingreso mensual por familia de entre 1000 y 2500 €. Por otro lado, el 24% de los hogares encuestados poseen un ingreso familiar de menos de 1000 € mensuales. Cabe destacar que, el 1% de las familias encuestadas ganan entre 2500 y 5000 €/mes y tan solo el 4% ganan más de 5000€. Por último, el 7% de los encuestados no han contestado a esta pregunta.

Por otro lado, el 61% afirman disponer de vehículo para sus desplazamientos, y un 39% de los encuestados no disponen de vehículo privado.

## 2. Información sobre los viajes de las personas del hogar

El 47% de los encuestados originan sus viajes para ir a su domicilio, el 8% para ir de compras y el 16% por motivos laborales. Por otro lado el 10% de los viajes se motivan por ocio y diversión y un 5% por estudios. Por último, el 8% estaría en la categoría de otros motivos, y el 2% de los viajes están originados por sanidad.

El 46% de los encuestados tienen el destino de sus viajes con el motivo de regreso a su domicilio, el 8% para ir de compras y el 16% por motivos laborales. Por otro lado, el 11% se motivan por ocio y diversión y un 5% por estudios. Por último, el 8% está en la categoría de otros motivos, y el 2% están originados por motivos de sanidad.

En cuanto al coste del billete en caso de utilización del transporte público de pasajeros, el 69% de los encuestados utilizan la tarjeta frente al 11% que utilizan el bono y el billete sencillo. Tan solo el 2% dicen utilizar billete sencillo de ida y vuelta.

Respecto a la frecuencia de realización de los viajes, el 56% de los viajes realizados se hacen de manera diaria frente a un 28% de los mismos que se efectúan de manera semanal. El 9% de los viajes son realizados esporádicamente y el 4% se realizan puntualmente. Sólo el 3% de los viajes se realiza con frecuencia mensual.

El 36% de los viajes han sido realizados a pie, el 34% en automóvil conduciendo (6% como acompañante), y el 19% en transporte público. El resto suponen en bicicleta (1%), tren (4%), taxi (1%) y motocicleta (3%).

Finalmente, por localidad de origen de los viajes, cabe destacar que el 96 % de los viajes realizados por los encuestados tienen su origen en el municipio de Santander, y tan solo el 4% de los mismos tienen su origen fuera.



#### 4. ESTIMACIÓN Y RESULTADOS

Las variables latentes son factores que, aun cuando influyen en el comportamiento y percepciones individuales (elecciones de los individuos), su influencia no puede ser cuantificada en la práctica (por ejemplo, confort, seguridad, comodidad, confiabilidad).

La inclusión de variables latentes en los MED permite considerar aspectos psicológicos, como percepciones y actitudes, en la modelación del comportamiento de los individuos. De esta manera, la capacidad explicativa del modelo aumenta, pudiendo ajustarse a la realidad de mejor manera.

Por otro lado, las variables latentes permiten incorporar implícitamente las características socioeconómicas a nivel individual, así como sus apreciaciones respecto de las alternativas disponibles o las políticas asociadas.

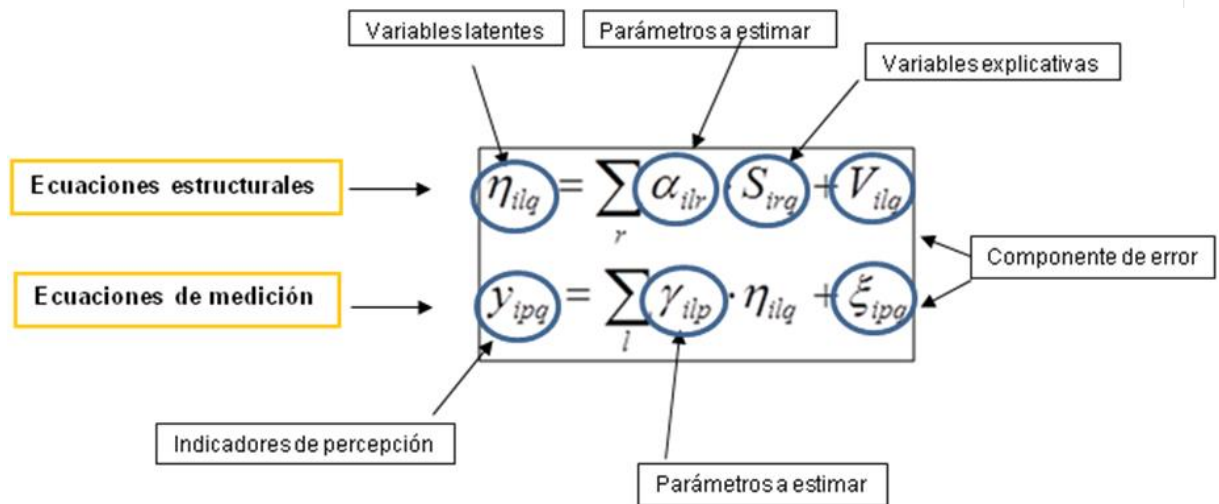
La dificultad para su medición radica tanto en su intangibilidad (no existe escala de medición) como en su intrínseca subjetividad (diferentes personas pueden percibir las de modo totalmente diferente). En ocasiones, ni tan siquiera sabemos ni de su existencia.

La identificación de variables latentes requiere suplementar una encuesta estándar de preferencias reveladas, con preguntas destinadas a capturar la percepción de los usuarios respecto a aspectos que no se pueden medir de otra forma, generando indicadores de percepción (medidas objetivas de las variables latentes para cada individuo). (Figura 2)

Califique entre 1 (poco satisfactorio) y 10
Características
1.- Adecuación del tipo de transporte a sus necesidades horarias
2.- Comodidad durante el viaje (asientos, espacio,...)
3.- Facilidad de acceso al modo de transporte
4.- Disponibilidad de información
5.- Fiabilidad del tiempo de espera
6.- Fiabilidad del tiempo de viaje
7.- Protección personal en el vehículo en caso de accidente
8.- Seguridad al viajar

**Figura 2. Indicadores de percepción**

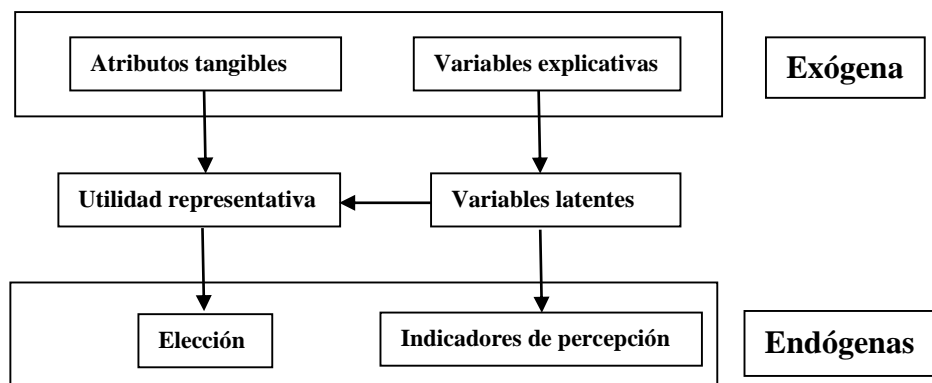
Para poder estimar las variables latentes se utilizan modelos MIMIC, que presentan la siguiente estructura:



**Figura 3. Esquema de los modelos MIMIC**

Posteriormente, en los MED híbridos la función de utilidad depende de las variables clásicas (atributos de individuos y alternativas) más las variables latentes (ecuación 8)

Por tanto, el marco de actuación de los MED con variables latentes se puede presentar del siguiente modo:



**Figura 4. Marco de actuación de los modelos de variables latentes**

En la práctica, se utiliza un método de estimación secuencial debido a la limitación en el software comercial, en el que primero se estima el valor de las variables latentes para posteriormente incluir en los MED:

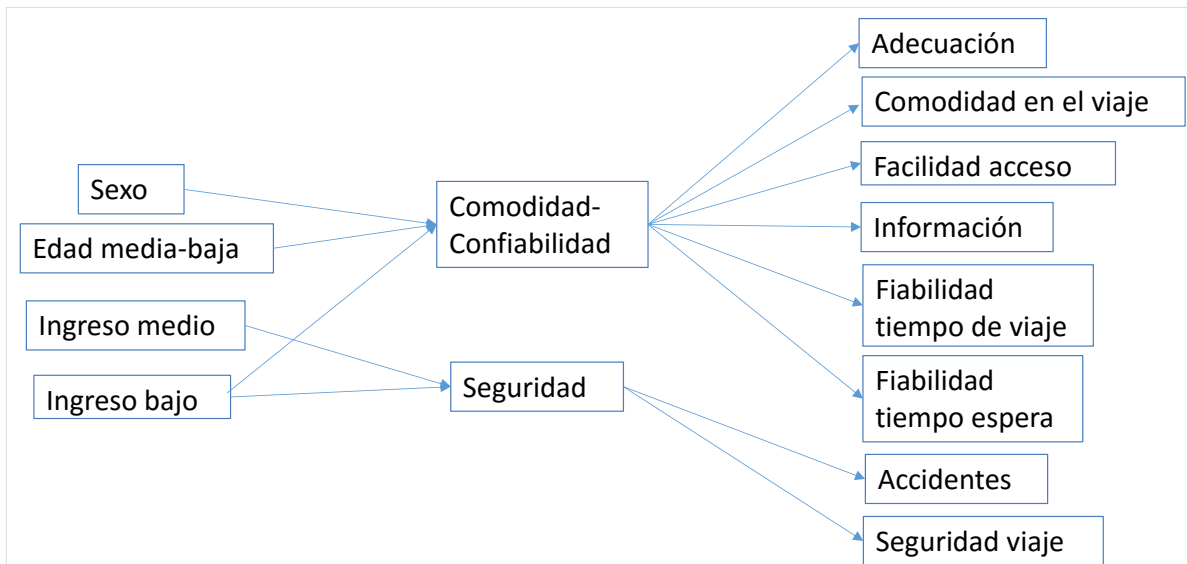
$$\begin{aligned}
 \eta_{ilq} &= \sum_r \alpha_{ilr} \cdot S_{irq} + V_{ilq} \\
 y_{ipq} &= \sum_l \gamma_{ilp} \cdot \eta_{ilq} + \xi_{ipq}
 \end{aligned}$$
$$U_{iq} = \sum_k \theta_{ik} \cdot X_{ikq} + \sum_l \beta_{il} \cdot \eta_{lq} + \varepsilon_{iq}$$

$$diq = \begin{cases} 1 & \text{si } U_{iq} \geq U_{jq}, \quad \forall j \in A(q) \\ 0 & \text{en otro caso} \end{cases}$$

El enfoque simultáneo (Raveau et al., 2010) debiera proporcionar estimadores eficientes y consistentes pero ha sido menos utilizado debido a su complejidad y a las limitaciones computacionales. Este es el principal motivo por el cual prevalece en la literatura empírica el enfoque secuencial.

En este caso, se ha procedido a estimar un modelo MIMIC para la variable latente Comodidad-Confiable, que viene explicada por tres variables socioeconómicas: sexo, pertenencia al tramo de edad media-baja (35-44 años) y pertenencia al nivel de ingreso bajo (<1000 €/mes). Por otro lado, esta variable latente explica los siguientes indicadores de percepción: adecuación del tipo de transporte a sus necesidades horarias, comodidad durante el viaje, facilidad de acceso al modo de transporte, disponibilidad de información, fiabilidad del tiempo de espera, fiabilidad del tiempo de viaje. (Figura 5)

Del mismo modo, se incluye en el modelo la variable latente Seguridad, que viene explicada por dos variables socioeconómicas: pertenencia al nivel de ingreso bajo (<1000 €/mes) y pertenencia al nivel de ingreso medio (1000-2500 €/mes). Esta variable latente explica los siguientes indicadores de percepción: Protección personal en el vehículo en caso de accidente y Seguridad al viajar. (Figura 5)



**Figura 5. Estructura del modelo MIMIC**

Una vez estimados los parámetros asociados al modelo MIMIC (Tabla 1), es posible obtener el valor de la variable latente para cada uno de los individuos, y así introducirlas como variables explicativas en el MED.

	$\alpha$ Sexo/Com	$\alpha$ INGBAJO-Com	$\alpha$ INGMED-BAJO-Seg	$\alpha$ EMED-BAJA/Com	$\alpha$ INGBAJO-Seg
BUS	0,159 (2,493)	0,162 (2,841)	-0,191 (-2,651)	-0,194 (-2,129)	0,172 (2,354)
COCHE CONDUCTOR	-0,193 (-2,863)	-0,372 (-4,717)	-0,225 (-3,206)	0,226 (2,403)	-0,37 (-4,512)

**Tabla 1. Resultados del modelo MIMIC**

En este caso, se han estimado dos MED: MNL y ML. En ambos casos, se han estimado modelos con y sin variables latentes (Comodidad-Confiability/Seguridad), que se han incluido en las funciones de utilidad del bus y del coche. Se han considerado sólo tres modos de transporte (coche-conductor, bus y caminando), ya que los viajes en los otros modos son prácticamente marginales.

Cabe destacar que la introducción del coste se ha realizado utilizando la expresión del modelo de tasa de gasto (Jara-Díaz y Farah, 1987), que propone introducir el coste dividiéndolo entre la tasa de gasto, la cual se calcula como el cociente entre el ingreso y el tiempo disponible para ello. Es decir, suponen que los individuos no pueden elegir libremente el número de horas dedicadas al trabajo, siendo relevante el tiempo que los individuos tienen para gastar y no el propio ingreso. Este enfoque ha sido muy utilizado, ya que proporciona excelentes resultados cuando se trabaja con individuos que tienen ingresos fijos (Ortúzar et al., 2000; Cherchi y Ortúzar, 2002).

De esta manera, la especificación econométrica de las funciones de utilidad para el modelo con variables latentes adopta la siguiente forma (en el otro caso sólo se elimina la parte correspondiente a la variable latente):

$$U(pie) = K_{pie} + T_{pie} * T_{viaje\ pie}$$

$$U(coche) = T_{coche} * T_{viaje\ coche} + C_{coche} * C_{viaje\ coche} + Com * Comodidad + Seg * Seguridad$$

$$U(bus) = T_{bus} * T_{viaje\ bus} + C_{bus} * C_{viaje\ bus} + T_{esp} * T_{espera\ bus} + Com * Comodidad + Seg * Seguridad$$

A continuación se presentan los principales resultados obtenidos en la estimación de los distintos modelos MNL y ML con y sin variables latentes:

Parámetro	MNL		MNL+LATENTES		ML		ML+LATENTES	
	Valor	Test-t	Valor	Test-t	Valor	Test-t	Valor	Test-t
Kpie	-0,360	-2,304	-5,693	-6,388	-0,120	-0,599	-7,250	-6,011
Tpie	-0,042	-11,199	-0,044	-11,149	-0,051	-8,196	-0,057	-8,034
Tcochc	-0,264	-5,042	-0,250	-4,466	-0,283	-5,073	-0,319	-4,311
Tbus	-0,013	-2,331	-0,023	-3,613	-0,013	-2,268	-0,026	-3,318
Tesp	-0,091	-3,346	-0,131	-4,149	-0,094	-3,233	-0,175	-4,225
Ccochc	-2,031	-3,212	-2,216	-3,192	-1,990	-3,007	-2,895	-3,310
Cbus	-3,091	-12,777	-4,082	-14,649	-3,228	-12,047	-5,189	-10,608
Com	-	-	1,865	10,420	-	-	2,992	5,480
Seg	-	-	2,692	5,649	-	-	3,437	5,471
Tpie (St.Dev.)	-	-	-	-	0,018	3,040	0,013	1,899
Como (St. Dev.)	-	-	-	-	-	-	1,790	2,823
Seg (St.Dev.)	-	-	-	-	-	-	0,743	4,008
l(θ)	-926,18		-823,82		-922,79		-814,74	
vst	7,80		6,77		8,55		6,61	

**Tabla 2. Resultados de los modelos**

En todos los modelos, todas las variables relevantes presentan los signos correctos y son estadísticamente significativas al 95% de confianza.

Al observar los resultados, se puede constatar la superioridad del modelo ML híbrido con variables latentes respecto al modelo más tradicional, que sólo incluye variables clásicas (tiempo y coste). Del mismo modo, el MNL híbrido se muestra superior al MNL que sólo considera las variables clásicas.

En primer lugar, se observa que la log-verosimilitud es menor en los modelos con variables latentes, lo que indica que son modelos más explicativos.

Se debe destacar el hecho de que el Valor Subjetivo del Tiempo es inferior en el modelo con variables latentes. Este resultado es coherente con la literatura (Raveau et al. 2010) existente y demuestra que los modelos más tradicionales tienden a sobreestimar este valor.

Finalmente, las variables latentes presentan el signo correcto y son estadísticamente

significativo, por lo que su inclusión mejora el modelo, al hacerlo más explicativo.

Por otro lado, si analizamos las elasticidades directas de cada uno de los modelos se puede observar que en todos los casos los signos son correctos ya que antes variaciones en el precio o el tiempo de viaje de los modos de transporte disminuye la demanda de los mismos.

Además, se observa que la mayor sensibilidad se presenta en la relación entre el precio y la demanda del transporte público. En cualquier caso, las elasticidades-precio del modelo ML con variables latentes son mayores que las del resto de los modelos, al igual que las asociadas a los tiempos del transporte público. Por otro lado, disminuye la sensibilidad en el caso del tiempo de viaje en coche.

	MNL	MNL+LATENTES	ML	ML+LATENTES
	Valor	Valor	Valor	Valor
Precio;demanda coche	-0,13	-0,14	-0,12	-0,15
Precio;demanda bus	-1,09	-1,44	-1,11	-1,67
Tiempo de viaje; demanda coche	-0,20	-0,19	-0,20	-0,18
Tiempo de viaje; demanda bus	-0,14	-0,25	-0,15	-0,27
Tiempo de espera; demanda bus	-0,21	-0,30	-0,21	-0,37

**Tabla 3. Elasticidades directas**

## 5. CONCLUSIONES

Este estudio muestra una aplicación empírica desarrollada en la ciudad de Santander (España), donde la realización de una encuesta de movilidad urbana, ha permitido obtener los datos requeridos para la estimación de los modelos econométricos que incorporen variables latentes.

De los datos analizados se desprende que la mayor parte de los viajes se realizan en vehículo privado o en bus, si bien las propias características de la ciudad (tamaño medio, estructura urbana centralizada), hacen que los desplazamientos caminando aún tengan un peso importante sobre el total.

En este trabajo se han podido identificar las variables latentes relevantes en la elección modal urbana existente en la ciudad de Santander, identificándose como tales la Comodidad-confiabilidad y la Seguridad. Estas variables latentes son explicadas por variables socioeconómicas (sexo, edad, nivel de ingresos). A su vez, la variable latente explica a distintos indicadores de percepción, cuyos valores conocemos mediante los datos obtenidos en la encuesta de movilidad.

Estas variables se muestran relevantes a la hora de elegir el modo de transporte, al ser estadísticamente significativas en la estimación del MED, y al presentar un alto valor en sus parámetros asociados.

Además, los modelos híbridos se muestran superiores a los MED más convencionales (aquellos que sólo consideran las variables explicativas clásicas), lo cual se puede apreciar a través de los valores alcanzados por el Log Likelihood. Del mismo modo, se ha podido comprobar que los modelos ML se muestran superiores a los modelos MNL ya que su capacidad explicativa es superior.

Por otro lado, se debe señalar que los errores de especificación debidos a la omisión de variables explicativas relevantes, como son las variables latentes, generan soluciones no óptimas. En este sentido, los valores subjetivos del tiempo en los modelos híbridos son inferiores, siendo este hecho de vital importancia para la valoración de los ahorros de tiempo, ya que los ahorros de tiempo son los mayores beneficios que se obtienen de las actuaciones en materia de transporte.

Finalmente, se debe señalar que las elasticidades también se ven afectadas tanto por el tipo de MED considerados como por la no inclusión de las variables latentes en las funciones de utilidad, lo cual implica que las decisiones acerca de las políticas de transporte basadas en modelos inadecuados pueden estar sesgadas

## REFERENCIAS

BEN-AKIVA, M., WALKER, J., BERNARDINO, A.T., GOPINATH, D.A., MORIKAWA, T. & POLYDOROPOULOU, A. (2002): Integration of choice and latent variable models. *In Perpetual Motion: Travel Behaviour Research Opportunities and Challenges*, 431-470. Pergamon, Amsterdam.

ASHOK, K., DILLON, W. R. & YUAN, S. (2002): Extending discrete choice models to incorporate attitudinal and other latent variables. *Journal of Marketing Research* 39, 31-46.

BEN-AKIVA, M., MCFADDEN, D., TRAIN, K., WALKER, J., BHAT, C., BIERAIRE, M., BOLDUC, D., BOERSCH-SUPAN, A., BROWNSTONE, D., BUNCH, D.S, DALY, A., DE PALMA, A., GOPINATH, D., KARLSTROM, A. & MUNIZAGA, M.A. (2002): Hybrid choice models: Progress and challenges. *Marketing Letters* 13-3, 163-175.

BEN-AKIVA, M., MCFADDEN, D., GARLING, T., GOPINATH, D., WALKER, J., BOLDUC, D., BORSCH-SUPAN, A., DELQUIE, P., LARICHEV, O., MORIKAWA, T., POLYDOROPOULOU & A., RAO, V. (1999): Extended framework for modeling choice behavior. *Marketing Letters* 10-3, 187-203.

BOLDUC, D., BOUCHNER, N. y ÁLVAREZ-DAZIANO, R. (2008): Hybrid choice modeling of new technologies for car choice in Canada. *Transportation Research Record: Journal of the Transportation Research Board* 2082, 63-71.

BOLLEN, K. A., KIRBY, J.B., CURRAN, P.J., PAXTON, P.M. & CHEN, F. (2007): Latent variable models under misspecification: two-stage least squares (2SLS) and maximum likelihood (ML) estimators. *Sociological Methods & Research* 36, 48-86.

CHERCHI, E., & ORTÚZAR, J. de D. (2002): Mixed RP/SP models incorporating interaction effects: modelling new suburban train services in Cagliari. *Transportation* 29(4), 371–395.

DAZIANO, R. & RIZZI, L. (2015): Analyzing the impact of a fatality index on a discrete, interurban mode choice model with latent safety, security and comfort. *Safety Science* 78, 11-19.

JARA-DÍAZ, S.R., & FARAH, M. (1987): Transport Demand and User's Benefits with Fixed Income: the Goods/Leisure Trade off Revisited. *Transportation Research*, 21B, 165-170.

MALDONADO-HINAREJOS, R., SIVAKUMAR, A. & POLAK, J.W. (2014): Exploring the role of individual attitudes and perceptions in predicting the demand for cycling: a hybrid



choice modelling approach. *Transportation*, 41, 1287-1304.

MOEN, B. E. (2007): Determinants of safety priorities in transport – the effect of personality, worry, optimism, attitudes and willingness to pay. *Safety Science* 45, 848-863.

O'NEILL, V., HESS, S. & CAMPBELL, D. (2014): A question of taste: Recognising the role of latent preferences and attitudes in analysing food choices. *Food quality and preferences*, 32, 299-310.

ORTÚZAR, J. DE D., RONCAGLIO, D.A. & VELARDE, U.C. (2000): Interactions and independence in stated preference modeling. In J. de D. Ortúzar (ed.), *Stated Preference Modelling Techniques*. Perspectives 4, PTRC, London

ORTÚZAR, J. DE D. & WILLUMSEM, L.G. (2001): *Modelling Transport*. 3<sup>o</sup> edición, John Wiley & Sons, Chichester.

PAULSEN, M., TEMME, D., VIJ, A. & WALKER, J. (2014): Values, attitudes and travel behavior: a hierarchical latent variable mixed logit model of travel mode choice. *Transportation*, 41, 873-888.

RAVEAU, S., ÁLVAREZ-DAZIANO, R., YÁÑEZ, M.F., BOLDUC, D. & ORTÚZAR, J. DE D. (2010): Sequential and simultaneous estimation of hybrid discrete choice models, *Transportation Research Record*, 2156, 131-139.

TEMME, D., PAULSEN, M. & DANNEWALD, T. (2008): Incorporating latent variables into discrete choice models – a simultaneous estimation approach using SEM software. *BuR - Business Research* 1, 220-237.

WALKER, J. L. & LI, J. (2007): Latent lifestyle preferences and household location decisions. *Journal of Geographical Systems* 9, 77-101.

WALKER, J. L., & BEN-AKIVA, M. (2002): Generalized random utility model. *Mathematical Social Sciences* 43, 303-343.