



**UNIVERSIDAD DE CONCEPCIÓN
FACULTAD DE INGENIERÍA
DEPARTAMENTO DE INGENIERÍA
CIVIL**



**EVALUACIÓN DE ENDOGENEIDAD PRODUCIDA POR OMISIÓN DE
VARIABLES PSICOSOCIALES EN MODELOS DE ELECCIÓN DISCRETA**

POR

Lorenzo Manuel Muñoz Torrejón

Tesis presentada a la Facultad de Ingeniería de la Universidad de Concepción para optar al grado de Magíster en Ciencias de la Ingeniería con Mención en Ingeniería Civil

Profesor Guía
Sebastián Astroza Tagle

Profesores Co-guía
Alejandro Tudela Román
Maximiliano Lizana Maldonado

Mayo 2024
Concepción (Chile)
© 2023 Lorenzo Manuel Muñoz Torrejón

© 2024 Lorenzo Manuel Muñoz Torrejón

Se autoriza la reproducción total o parcial, con fines académicos, por cualquier medio o procedimiento, incluyendo la cita bibliográfica del documento

RESUMEN

La endogeneidad es un problema inevitable en modelos de elección discreta que causa estimación de parámetros inconsistentes. Dicha inconsistencia se puede ver reflejada en indicadores económicos que se utilizan para evaluar proyectos de transporte, como el valor subjetivo del tiempo y elasticidades. Por otro lado, se tiene evidencia del impacto que tiene la inclusión de variables psicosociales, tales como el hábito, la actitud y el afecto, en modelos de partición modal, cuya omisión podría causar endogeneidad. Estos antecedentes motivan estudiar la endogeneidad por omisión de variables psicosociales en modelos de partición modal.

Se proponen dos metodologías para evaluar endogeneidad, el Método de los Indicadores Múltiples (MIS) y el Método de Variables Latentes Integrado (ICLV). Para el primero, no se han encontrado aplicaciones con respecto a variables psicosociales. Más aún, se tienen muy pocos estudios que incluyen varias de estas variables de manera simultánea. Además, se proponen ambas metodologías interactuando multiplicativamente con el tiempo de viaje. Los datos provienen de encuestas realizadas a funcionarios, académicos y alumnos de la Universidad de Concepción y de la Universidad de la Frontera. Los modos estudiados corresponden al automóvil, la bicicleta y el bus.

Se demuestra que el afecto, la actitud y el hábito tienen un impacto significativo en la elección de modo de transporte, para todos los modelos estimados. Esto genera cambios sustanciales en los indicadores mencionados. Se tiene que el valor subjetivo del tiempo aumenta si se consideran las variables mencionadas aditivamente, pero disminuye si se consideran interactuando con el tiempo de viaje. Las elasticidades del tiempo de viaje se ven disminuidas para la bicicleta y para el automóvil al usar las metodologías propuestas, mientras que para el bus se ve aumentada. Por último, la elasticidad del costo también se ve disminuida para el automóvil y aumentada para el bus. Esto podría tener una repercusión directa en la evaluación de políticas públicas, dado que estas dependen de la partición modal, la cual podría estar sesgada.

ABSTRACT

Endogeneity is an unavoidable problem in discrete choice models that causes inconsistent estimation of parameters. This can be seen reflected in economic indicators that are used to evaluate infrastructure projects, such as the subjective value of time and elasticities. On the other hand, there is evidence of the impact of the inclusion of psychosocial variables in discrete choice models, such as affection, attitude and habit. The omission of these variables could cause endogeneity. This motivates studying endogeneity due to omission of psychosocial variables in discrete choice models.

Two methodologies are proposed to evaluate endogeneity, the Multiple Indicator Solution (MIS) and the Integrated Choice Latent Variable Method (ICLV). For the first method, no applications have been found regarding psychosocial variables. Furthermore, there are very few studies that include multiple psychosocial variables simultaneously. Moreover, both methodologies are proposed interacting as a product with travel time. The data comes from surveys carried out among officials, academics and students of Universidad de Concepción and the Universidad de la Frontera. The modes studied correspond to car, bicycle, and bus.

It is shown that affect, attitude, and habit have a significant impact on the choice of transport mode, for all the estimated models. This generates substantial changes in the indicators mentioned. The subjective value of time increases if the psychosocial variables are considered in an additive way but decreases if they are considered interacting with travel time. The elasticities of travel time decrease for the bicycle and for the car when using the proposed methodologies, while for the bus it increases. Finally, cost elasticity is also decreases for the car and increases for the bus. This could have a direct impact in evaluating public policies, since modal share is required for their assessment, and modal share may be biased.

ÍNDICE DE CONTENIDOS

CAPÍTULO 1: INTRODUCCIÓN	2
1.1. Motivación	2
1.2. Hipótesis de trabajo	5
1.3. Objetivo general	5
1.4. Objetivos específicos	5
1.5. Metodología de trabajo	5
1.6. Alcance	6
1.7. Principales Resultados y Conclusiones	6
1.8. Organización del texto	7
CAPÍTULO 2: ENDOGENEIDAD EN MODELOS DE ELECCIÓN DISCRETA	6
2.1. Introducción	6
2.2. Endogeneidad en Modelos con Interacción	11
2.3. Endogeneidad por Omisión de Variables Psicosociales	11
2.4. Conclusión	15
CAPÍTULO 3: DATOS UTILIZADOS PARA LA MODELACIÓN	16
3.1. Introducción	16
3.2. Medición de Atributos de los Viajes y de Variables Psicosociales	16
3.3. Análisis de Consistencia Estadística	20
3.4. Características Sociodemográficas de la Muestra	22
3.5. Variables Consideradas en la Modelación	25
3.6. Conclusión	27
CAPÍTULO 4: METODOS DE MODELACIÓN	28
4.1. Introducción	28
4.2. Modelos de Elección Discreta Aplicados a Partición Modal	29
4.3. Método de los Indicadores Múltiples	31
4.4. Análisis Factorial y Modelos de Ecuaciones Estructurales	36
4.5. Modelos de Variables Latentes Integradas	40
4.6. Modelos con Interacción	44
4.7. Valor Subjetivo del Tiempo y Elasticidades	46
4.8. Conclusión	48

CAPÍTULO 5: ANÁLISIS DE RESULTADOS	50
5.1. Introducción	50
5.2. Modelo Básico	51
5.2. Modelo con el Método de los Indicadores Múltiples	53
5.3. Método de Variables Latentes Integrado	56
5.4. Modelos con Interacción	62
5.5. Valor Subjetivo del Tiempo y Elasticidades	66
5.6. Conclusión	72
CAPÍTULO 6: CONCLUSIONES	76
6.1. Conclusiones	76
6.2. Líneas de Investigación Futura	77
CAPÍTULO 7: REFERENCIAS	79
CAPÍTULO 8: ANEXOS	83
ANEXO 8.1. Contribución a los Objetivos de Desarrollo Sostenible	83

ÍNDICE DE TABLAS

Tabla 2.1: Metodologías para la evaluación de la endogeneidad	11
Tabla 3.1: Métodos de medición para variables psicosociales	21
Tabla 3.2: Valores del α de Cronbach desagregados según dimensión del afecto	22
Tabla 3.3: Características sociodemográficas de la muestra	24
Tabla 3.4: Promedios y desviaciones estándar de las variables psicosociales	25
Tabla 3.5: Variables consideradas en la modelación	27
Tabla 5.1: Resultados para el modelo básico	52
Tabla 5.2: Resultados para el modelo corregido con el método MIS	54
Tabla 5.3: Resultados para el modelo corregido con el método ICLV	58
Tabla 5.4: Resultados para los modelos de ecuaciones estructurales	60
Tabla 5.5: Resultados para modelos MIMIC	62
Tabla 5.6: Resultados para el modelo corregido con el método MIS con interacción	65
Tabla 5.7: Resultados para el modelo corregido con el método ICLV con interacción	66
Tabla 5.8: Porcentajes de aporte de las variables en funciones de utilidad	68
Tabla 5.9: Intervalos de confianza para el VST en métodos MIS y ICLV	70
Tabla 5.10: VST desagregado por modo para modelos con interacción con el tiempo de viaje	70
Tabla 8.1: Análisis factorial para el afecto hacia la bicicleta	88
Tabla 8.2: Análisis factorial para el afecto hacia el automóvil	88
Tabla 8.3: Análisis factorial para el afecto hacia el bus	89
Tabla 8.4: Análisis factorial para la actitud hacia la bicicleta	89
Tabla 8.5: Análisis factorial para la actitud hacia el automóvil	90
Tabla 8.6: Análisis factorial para la actitud hacia el transporte público	90
Tabla 8.7: Resultados de los SEM para modelo con interacción	91
Tabla 8.8: Resultados de los modelos MIMIC para modelo con interacción	92

ÍNDICE DE FIGURAS

Figura 2.1: Esquema de la Teoría del Comportamiento Interpersonal de Triandis (1977)	15
Figura 4.1: Modelo de elección incorporando variables latentes	42
Figura 5.1: Elasticidad del tiempo de viaje para la bicicleta	72
Figura 5.2: Elasticidad del tiempo de viaje para el automóvil	72
Figura 5.3: Elasticidad del tiempo de viaje para el transporte público	73
Figura 5.4: Elasticidad del costo para el automóvil	75
Figura 5.5: Elasticidad del costo para el transporte público	75

CAPÍTULO 1: INTRODUCCIÓN

1.1. Motivación

En el campo de la modelación de demanda de viajes, la elección de modo de transporte es la decisión más importante que afecta la congestión vehicular (Abdel-Aal, 2017). En general, dicha modelación se realiza a partir de modelos de elección discreta, que cuantifican los pesos de las variables involucradas en dicha decisión. Estos modelos no están exentos de irregularidades en la estimación, ya que, a priori, no se conoce la totalidad de las variables que serán significativas para explicar la elección. Si bien se pueden generar aproximaciones, a partir de la intuición y la experiencia, no es trivial el representar adecuadamente un fenómeno observado. Por lo tanto, se corre el riesgo de que la modelación sufra de problemas, como es la presencia de endogeneidad.

La endogeneidad es una anomalía inevitable para los modelos de elección discreta, y se presenta causando estimaciones inconsistentes de los coeficientes calculados. Dicha inconsistencia se presenta fundamentalmente en las magnitudes, signos y pruebas estadísticas asociadas a los coeficientes del modelo, cuya interpretación es la utilidad marginal de la variable utilizada para modelar la decisión (Guerrero *et al.*, 2020). Por lo tanto, este fenómeno puede llevar a que la toma de decisiones, en términos de análisis de políticas públicas, predicciones, conclusiones o evaluaciones conductuales, sea defectuosa (Guerrero *et al.*, 2021). Los modelos sin corrección por endogeneidad han sido los únicos realmente utilizados en la práctica para realizar predicciones. Además, los efectos adversos de dicha anomalía se incrementan con el paso de los años, generando impactos severos en la predicción de escenarios futuros (Guevara *et al.*, 2020).

En particular, para los modelos que no sean corregidos por endogeneidad, se tiene que los valores obtenidos en la estimación de la partición modal, como el valor subjetivo del tiempo (VST), las elasticidades y las utilidades marginales de las variables consideradas, son inconsistentes. Dichos indicadores son cruciales para la planificación de transporte y para la evaluación social de proyectos. Más aún, existe evidencia de que estos indicadores cambian

sustancialmente en magnitud al omitir variables significativas para la elección de modo de transporte (Fernández-Antolín *et al.*, 2016; Guerrero *et al.*, 2021). Por lo tanto, existe la necesidad de desarrollar métodos de corrección para el problema de la endogeneidad y cuantificar sus efectos. En otras palabras, las conclusiones económicas derivadas de la estimación podrían diferir mucho de los valores que hayan corregido esta anomalía, lo que podría significar una reasignación de recursos importante.

Por otro lado, se tiene evidencia del impacto que tienen las variables psicosociales en la elección de modo de transporte. En efecto, la literatura ha mostrado que los modelos de elección discreta aplicados a partición modal son sensibles a las condiciones psicosociales de los usuarios, como el afecto (Ababio-Donkor *et al.*, 2020), la actitud y el hábito (Lizana *et al.*, 2020). También, se ha estudiado el efecto de incorporar múltiples variables psicosociales en la elección de modo de transporte. Galdames *et al.* (2011) estudian el rol del afecto, la actitud y los factores sociales, Devika *et al.* (2020) estudian la influencia de la actitud, la norma subjetiva, la intención y el control conductual percibido. Para todos los estudios mencionados las variables psicosociales se mostraron como variables relevantes para explicar la elección de modo de transporte. Además, Soza-Parra y Cats (2023) realizaron una revisión bibliográfica, concluyendo que los motivos instrumentales, la autonomía, los factores afectivos, motivos simbólicos, la norma social, y motivos medioambientales son factores influyentes en la decisión de viajar en automóvil.

Por ejemplo, la actitud hacia un modo de transporte podría estar correlacionada con el tiempo de viaje, o deberse principalmente al ahorro de tiempo que significaría el escoger tal modo, para una interpretación práctica. En dicho ejemplo, si se omite la variable actitud, es probable que el modelo sufra de endogeneidad.

Una fuente común de endogeneidad es la omisión de atributos significativos para representar el suceso observado, como el hacinamiento en el transporte público (Tirachini y Guevara, 2020), la comodidad en elección de modo de transporte (Guevara y Polanco, 2016), atributos correlacionados con el precio en elección de vivienda (Guevara y Ben-Akiva, 2006) y

atributos del servicio de televisión en elecciones de compañías televisivas (Petrin y Train, 2010).

Existen diversas metodologías para evaluar este tipo de endogeneidad. En este caso, debido a las características de los datos para la modelación, pues se requiere información bastante específica, se utilizarán el Método de Variables Latentes Integrado (Ben-Akiva *et al.*, 2002) y el Método de los Indicadores Múltiples (Guevara y Polanco, 2016). Ambos métodos cuentan con ventajas y desventajas, comparándolos a partir de indicadores económicos, como es el valor subjetivo del tiempo y la elasticidad.

El Método de Variables Latentes Integrado (ICLV, por sus siglas en inglés) ha sido ampliamente utilizado en la literatura para incluir variables cualitativas. Por ejemplo, Temme *et al.* (2008) consideran la flexibilidad al viajar, Daly *et al.* (2012) incluyen la seguridad y Hess *et al.* (2018) incorporan la actitud hacia el automóvil. Por otro lado, algunas de las aplicaciones del Método de Indicadores Múltiples (MIS) son los trabajos de Guevara y Polanco (2016) en elecciones residenciales, Fernández-Antolín *et al.* (2016) y Guevara *et al.* (2020) en elecciones de modo de transporte, Mariel *et al.* (2018) en evaluación medioambiental y Palma *et al.* (2018) en elecciones de vinos.

En específico, respecto del MIS, se necesita más trabajo empírico para determinar el impacto real de su contribución a la práctica (Guevara y Polanco, 2016). Además, una brecha encontrada en la literatura es que existen pocas aplicaciones del Método de los Indicadores Múltiples con datos reales (Fernández-Antolín *et al.*, 2016). De esta manera, se hipotetiza que este método puede ser una buena aproximación para la representación más sofisticada del comportamiento, que se realiza a partir del método ICLV, pero evitando los enormes costos de tiempo y modelación que esto conlleva. Otra aplicación potencial de este método podría ser la detección de la endogeneidad, para una posterior modelación más compleja.

En relación con las variables psicosociales, no se han encontrado correcciones de endogeneidad por su omisión en la modelación de la partición modal, es decir, cuantificar el impacto efectivo de la inclusión (o exclusión) de la actitud, el afecto y el hábito hacia un

Introducción

modo de transporte, de manera simultánea, donde la actitud es la evaluación personal hacia el objeto, el afecto se relaciona con la emocionalidad hacia un modo y el hábito se ve condicionado principalmente por la conducta previa. Incluso, la mayoría de los estudios realizados para incluir variables cualitativas suelen considerar una o dos de estas variables. En el presente estudio se incluyen seis variables latentes y se evalúa el impacto de su incorporación, análisis que no ha sido reportado previamente.

Más aún, no se han encontrado aplicaciones prácticas del método MIS para modelos con interacción entre variables con datos reales, es decir, a partir de una función de utilidad no lineal, donde se obtiene la utilidad marginal de una variable sobre otra. Este análisis de interacción sí se ha realizado para el método ICLV, resultando relevante un análisis comparativo (Cherchi y Ortúzar, 2002). En específico, interacción es de una variable psicosocial con otra variable del nivel de servicio presente en la utilidad; para este caso se estudiará la interacción con el tiempo de viaje, generando una comparación entre ambas metodologías.

Por lo tanto, se evaluará la endogeneidad causada por la omisión de variables psicosociales en los modelos de elección discreta. La contribución principal de este trabajo es (i) determinar la existencia de endogeneidad causada por la omisión de variables psicosociales para la muestra en cuestión, (ii) cuantificar las implicancias de omitir dichas variables en los modelos de elección discreta, y (iii) comparar metodologías existentes en la literatura para solucionar el problema de la endogeneidad. De esta manera, los resultados de este trabajo permiten generar sugerencias para encontrar estimaciones consistentes, que representen de mejor manera el comportamiento de los usuarios. Esto genera una solución a la problemática planteada por Guerrero *et al.* (2021), con respecto a la estimación sesgada de indicadores económicos, aportando a una mejor planificación de transporte y una evaluación más realista de proyectos. A la fecha, este trabajo presenta la primera aplicación del método MIS para corregir la endogeneidad por omisión de variables psicosociales.

1.2. Hipótesis de trabajo

La endogeneidad producida por la omisión de variables psicosociales en modelos de partición modal causa una estimación inconsistente de coeficientes, que se puede ver reflejada en indicadores económicos; y la corrección de está generaría parámetros consistentes.

1.3. Objetivo general

Corregir la endogeneidad producida por la omisión de variables psicosociales en la elección de modo de la muestra, correspondiente a estudiantes, funcionarios y académicos de la Universidad de Concepción y de la Universidad de la Frontera.

1.4. Objetivos específicos

- Comparar distintas metodologías para corregir la endogeneidad presente en modelos de elección discreta, en particular los métodos MIS e ICLV.
- Analizar el rol de los coeficientes obtenidos a partir de la estimación, obteniendo las variables significativas para la muestra.
- Evaluar indicadores a partir de las distintas metodologías, como el valor subjetivo del tiempo y la elasticidad de la demanda respecto del tiempo de viaje y del costo.
- Validar una aplicación del modelo MIS con datos reales para corregir dicha endogeneidad.

1.5. Metodología de trabajo

Para la realización de la presente tesis se comenzó por una revisión bibliográfica con respecto a los modelos de elección discreta, la incorporación de variables psicosociales, la endogeneidad presente en éstos y las distintas metodologías para evaluar la endogeneidad. Se realizó un orden y limpieza de los datos disponibles, junto con un posterior análisis descriptivo y de consistencia estadística. Posteriormente, se estimaron los modelos de interés,

probando distintas especificaciones para las funciones de utilidad. Luego, se analizaron los resultados obtenidos y sus implicancias. Por último, se compararon las distintas metodologías en función de indicadores obtenidos a partir de las estimaciones, como lo son el valor subjetivo del tiempo y la elasticidad de la demanda respecto al tiempo de viaje y al costo.

1.6. Alcance

Los datos para la modelación fueron obtenidos a partir de encuestas de preferencias reveladas realizadas a personas de la Universidad de Concepción y de la Universidad de la Frontera, Chile. Esto constituye un muestreo por conveniencia e implica que las conclusiones realizadas sólo aplicarán para esta población. El marco teórico, los supuestos y la estimación matemática corresponde a los modelos MIS y ICLV, basándose fundamentalmente en los trabajos de Ben-Akiva *et al.* (2002), Guevara y Polanco (2016), y Fernández-Antolín *et al.* (2016).

1.7. Principales Resultados y Conclusiones

En la presente Tesis se estimaron cinco modelos: un modelo que no presenta correcciones de endogeneidad, un modelo corregido con la metodología MIS, un modelo corregido con la metodología ICLV y ambas correcciones con una versión que presenta la interacción de las variables psicosociales con el tiempo de viaje. Se obtuvo que, para ambas metodologías, el afecto, la actitud y el hábito corresponden a variables significativas para explicar la elección de modo de transporte de los usuarios. Además, la omisión de dichas variables produce una estimación errónea de los coeficientes. Esto se puede observar directamente en términos de índices económicos: el valor subjetivo del tiempo de viaje y las elasticidades del tiempo de viaje y del costo, pues, se aprecian cambios sustanciales en los valores con la inclusión de las variables mencionadas. Un aporte concreto al campo de los modelos de elección discreta es que se propone una aplicación del modelo MIS para la inclusión de las variables psicosociales, y la demostración de que los resultados de este modelo son más consistentes que el omitir dichas variables en la modelación.

1.8. Organización del texto

Este documento se organiza de la siguiente manera: en la Sección 2 se expone la revisión bibliográfica del fenómeno de la endogeneidad, mientras que en la Sección 3 se describen los datos utilizados para la modelación. La Sección 4 presenta los métodos de modelación y de corrección de endogeneidad. En la Sección 5 se muestran los resultados obtenidos y en la Sección 6 se postulan las conclusiones y comentarios finales.

CAPÍTULO 2: ENDOGENEIDAD

2.1. Aspectos Generales

La endogeneidad es una anomalía inevitable para modelos econométricos, dentro de los cuales se encuentran los modelos de elección discreta utilizados para calcular la utilidad que le reporta a un usuario n la elección de un modo i , a partir de una serie de variables. Este problema se presenta causando estimaciones de coeficientes inconsistentes, esto es, distintos de los valores que se obtienen corrigiendo este problema, independiente del tamaño de la muestra. Esta inconsistencia se presenta principalmente en términos de magnitud, pero incluso podrían darse casos donde el signo de los coeficientes no sea el correcto. Por ejemplo, obtener que más tiempo de viaje signifique una utilidad positiva para el usuario o que el viaje sea causa de placer.

Los coeficientes mencionados se interpretan como las utilidades marginales de la variable asociada, esto es, la contribución que tiene cada unidad de la variable en cuestión en la utilidad que reporta la elección del modo estudiado. En función de estos se calculan indicadores económicos que frecuentemente se utilizan para evaluaciones de proyectos de infraestructura. Luego, la presencia de endogeneidad puede llevar a que la toma de decisiones, en términos de análisis de políticas públicas, predicciones, conclusiones o evaluaciones conductuales, sea defectuosa (Guerrero *et al.*, 2021). Incluso, la endogeneidad es el problema más severo al que un modelo de elección discreta puede enfrentarse en su objetivo de realizar inferencias causales o predicciones (Guevara, 2015), pues, como los coeficientes están erróneamente calculados (incluso en términos de signos), dichas predicciones estarán sesgadas.

2.2. Endogeneidad en Modelos de Elección Discreta

En los modelos de elección discreta la utilidad se representa como la suma entre una parte medible y una parte aleatoria. En palabras técnicas, la endogeneidad ocurre cuando algunas variables de la parte determinísticas de la utilidad no son independientes de la componente

Endogeneidad

de error de ésta. Si una variable es endógena, se malinterpretan los cambios en la componente de error, pues podrían ser cambios en la variable endógena, siendo imposible la estimación correcta de los coeficientes (Guevara y Ben-Akiva, 2012).

En la aplicación a los modelos de partición modal, es común la omisión de variables como la comodidad, seguridad, e incluso variables psicosociales que pueden ser significativas para el usuario. La razón es que generalmente estas variables son difíciles de medir y de recolectar, por lo que, la gran mayoría de estos modelos sufren de endogeneidad. Es más, según Guerrero *et al.* (2020), no se han encontrado modelos de elección en un modelo de transporte estratégico clásico que hayan sido corregidos por endogeneidad. Además, es muy difícil saber, preliminarmente, si un modelo estimado sufre o no de endogeneidad, precisamente porque no se conoce el modelo “real”. Una representación adecuada del fenómeno debería idealmente considerar todas las variables significativas para éste, lo que es prácticamente imposible. Por lo tanto, se debe tener un conocimiento adecuado del procesamiento de los datos y de las limitaciones de las herramientas de análisis.

Existen diversas fuentes de endogeneidad, así como distintas maneras de evaluarla y cuantificarla. En general se pueden describir cinco causas:

1. Atributos omitidos, donde se tiene una representación parcial del fenómeno que se está intentando reproducir, como lo que se mencionó anteriormente. Se debe principalmente a la imposibilidad del investigador para identificar y/o medir dichos atributos.
2. Errores de especificación en el modelo, donde se tienen descuidos de heterogeneidad o usar una forma funcional errónea. Por ejemplo, realizar supuestos de relaciones lineales cuando en realidad éstas serían cuadráticas o logarítmicas.
3. Errores de medición en los atributos, donde el error medido estará correlacionado con la componente de error aditiva de la utilidad. Esto se puede dar midiendo *in situ*, por ejemplo, realizando conteos vehiculares, o en respuestas asociadas a encuestas.
4. Determinaciones simultáneas, para casi cualquier modelo donde se utilicen datos reales, pues los atributos son resultados de un equilibrio. Por ejemplo, el tiempo de

viaje depende de la elección de modo del usuario, la elección de modo incide en el flujo vehicular, y el flujo vehicular incide en el tiempo de viaje.

5. Autoselección, donde los individuos se auto-seleccionan dentro de la muestra de una manera independiente a la del modelador. Por ejemplo, si se están intentando modelar accidentes y sus consecuencias, los usuarios que usen casco naturalmente tendrían consecuencias menos fatales.

En este caso se estudiará la endogeneidad producida por la omisión de variables significativas para explicar el fenómeno, y que estén correlacionadas con otra variable económica incluida, tal como el tiempo de viaje o el costo de viaje. Es decir, el primer caso mencionado en el listado anterior. En definitiva, interesa evaluar cómo la omisión de variables psicosociales incide en la estimación de coeficientes.

Además, hay dos métodos formales para detectar la endogeneidad:

1. Hausman (1978): propuso comparar los coeficientes de dos modelos que serían consistentes, bajo la hipótesis nula de que no habrá endogeneidad, pero podrían diferir en su eficiencia.
2. Aplicar una metodología para su corrección y verificar si los cambios son significativos para un nivel de confianza dado.

Existen diversas metodologías para la evaluación de la endogeneidad, todas con virtudes y desventajas. Se tiene el método BLP (Berry *et al.*, 1995), el método de máxima verosimilitud (Park y Gupta, 2009), el *Control Function Method* (Petrin y Train, 2010), el método *proxy* (Guevara, 2015), el método de variables latentes integrado (Ben-Akiva *et al.*, 2002) y el método de indicadores múltiples (Guevara y Polanco, 2016). La decisión de usar un método u otro se basa en qué información está disponible, los supuestos que el modelador está dispuesto a hacer y los costos computacionales asociados (Guerrero *et al.*, 2022). Luego, en la presente Tesis se optó por los dos últimos. El detalle y la derivación matemática se explicitan en el Capítulo 4; en esta sección serán descritos brevemente. En la Tabla 2.1 se resumen las metodologías para la evaluación de endogeneidad.

Tabla 2.1: Metodologías para la evaluación de la endogeneidad

Metodología	Autor(es)
BLP	Berry et al. (1995)
Máxima Verosimilitud	Park y Gupta (2009)
Control Function Method	Petrin y Train (2010)
Proxy	Guevara (2015)
Variables Latentes	Ben-Akiva et al. (2002)
MIS	Guevara y Polanco (2016)

Es importante mencionar que la corrección de endogeneidad no es gratuita, pues en regresiones lineales todas las variables tienen un error asociado. Luego, se deben realizar bastantes supuestos. Es criterio del modelador saber escoger bien estos supuestos, para llegar a estimaciones consistentes, sin comprometer el modelo.

En un principio, se corregía la endogeneidad con el método BLP, que consiste en abordar la endogeneidad a partir de regresiones lineales. Esto se realiza con la inclusión de una constante específica (*ASC*, por sus siglas en inglés) para cada alternativa, que debería medir la parte no observable en la elección. Generalmente, esta constante específica se considera en los modelos de elección discreta, pero no entrega información sobre las variables omitidas, señalando que sólo existirían.

El *Control Function Method* (Petrin y Train, 2010) utiliza, como mínimo, una variable instrumental por cada variable endógena. La variable instrumental corresponde a un instrumento de medición para la variable endógena, que debe estar correlacionado con ésta y ser exógeno con respecto a la función de utilidad. Este método era un buen candidato para utilizar en la presente evaluación, pero, precisamente debido a la dificultad de contar con los requerimientos del instrumento, fue descartado. Su aplicación podría ser homologada con el Método de los Indicadores Múltiples (MIS, por sus siglas en inglés), pero este utiliza indicadores como variables instrumentales.

El MIS fue propuesto originalmente por Wooldridge (2010) para modelos lineales. Su aplicación a modelos de elección discreta fue realizada por Guevara y Polanco (2016) y

requiere al menos dos indicadores adecuados para medir la variable omitida. Primero, uno de los indicadores se incluye como una variable explicativa en la utilidad del modelo de elección. A partir de esta inclusión, se elimina la endogeneidad de las otras variables, y el indicador pasa a ser la única variable endógena, pues arrastra consigo errores de medición. Luego, se utiliza el segundo (o los demás) indicadores como un instrumento para lidiar con la endogeneidad inducida. Este método ha tenido diversas aplicaciones a modelos de elección discreta aplicados a partición modal, Fernández-Antolín *et al.* (2016) estudian el *amor* hacia el automóvil, Guevara *et al.* (2020) estudian el hacinamiento en transporte público y Mariel *et al.* (2018) estudian la valoración ambiental.

Otro método para lidiar con los problemas inducidos por la endogeneidad es el método de variables latentes (ICLV, por sus siglas en inglés), donde, a partir de modelos de ecuaciones estructurales, se cuantifica el peso efectivo de la variable omitida en el modelo de elección. Además, entrega información sobre cada uno de los indicadores utilizados para medir la variable y sobre las características sociodemográficas utilizadas para definirla. Por lo tanto, se cuenta con un método que es más sofisticado, pero computacionalmente mucho más complejo y costoso en términos de tiempo. Sin embargo, no está del todo claro si las componentes de error de los métodos de ecuaciones estructurales están (o no) correlacionados con las otras variables explicativas en la utilidad, lo que sí se logra con el método MIS.

Ambos modelos buscan solucionar el problema de la endogeneidad con enfoques distintos. En resumen, el método MIS se utiliza para considerar la heterogeneidad de los factores omitidos entre modos e individuos, mientras que el método ICLV se utiliza para considerar la heterogeneidad en percepciones entre individuos (Fernández-Antolín *et al.*, 2016). Una ventaja importante del MIS sobre el ICLV es que es mucho más barato computacionalmente. Para tener una referencia, el orden de magnitud de tiempo de optimización del MIS para la muestra y las variables estudiadas en esta Tesis es de menos de un minuto, mientras que el método ICLV demora varias horas. Un planteamiento interesante es utilizar secuencialmente ambos métodos: el MIS para la *detección* de la endogeneidad y el ICLV para *analizarla*.

Endogeneidad

En definitiva, el método ICLV debería ser más eficiente porque considera la verosimilitud de las ecuaciones estructurales, ocupa más información, pero requiere supuestos distribucionales más fuertes y la disponibilidad de ecuaciones estructurales adecuadas para la variable latente.

En el presente caso, esto es, preferencias reveladas a partir de las respuestas a una encuesta, la fuente de endogeneidad es principalmente la imposibilidad del investigador de cuantificar atributos omitidos. En particular, la comodidad es comúnmente fuente de endogeneidad, porque es frecuente que dicho atributo esté correlacionado con el tiempo de viaje. Esto se traduce en una sobreestimación del valor subjetivo del tiempo, un indicador clave para la toma de decisiones en términos de políticas públicas (Guevara y Polanco, 2016).

2.3. Endogeneidad en Modelos con Interacción

Por otro lado, es posible que la sensibilidad del usuario hacia alguna variable de servicio tenga directa relación con las variables psicosociales en cuestión. Por ejemplo, podría darse que el hábito que un usuario tiene de viajar en automóvil se deba al tiempo de viaje que este modo le significa, o que el afecto hacia el transporte público se vea condicionado por la tarifa.

Guevara (2010) explica en su tesis doctoral por qué es más difícil lidiar con la endogeneidad producida en modelos no lineales, en términos de utilidad, que en modelos lineales. En particular, se tiene que las correcciones por endogeneidad implicarían un cambio de escala en modelos de elección. Sin embargo, el supuesto adicional que se debe hacer es que la interacción del indicador con la variable de servicio es un indicador de la interacción de la variable omitida con la variable económica, lo que no necesariamente es trivial.

2.4. Endogeneidad por Omisión de Variables Psicosociales

Una fuente común de endogeneidad es la omisión de variables latentes, es decir, variables que no se pueden medir directamente, sino que se infieren a partir de otras variables

cuantificables. Además, se sospecha que indicadores como el valor subjetivo del tiempo se vean condicionados por este tipo de variables, como se ha visto en la literatura (Fernández-Antolín *et al.*, 2016; Atasoy *et al.*, 2013). Dentro del universo de las variables latentes, se pueden encontrar las variables psicosociales.

Para analizar políticas públicas se pueden realizar distintos escenarios, prediciendo cambios en atributos observables. Dichos escenarios podrían estar incompletos sin la inclusión de las variables latentes en los modelos de predicción (Vij y Walker, 2016). Otra aplicación práctica de esta inclusión es en publicidad, para modificar la conducta de los usuarios, por ejemplo, políticas de no manejar si se ha bebido alcohol, como probaron Scaglonari *et al.* (2015) en Lugano, Suiza. Además, esta modelación podría reducir la varianza de los coeficientes estimados, y consecuentemente, incrementar la precisión de indicadores como las tasas marginales de sustitución.

En particular, se han estudiado las implicancias de considerar (u omitir) variables como el afecto (Ababio-Donkor *et al.*, 2020), la actitud y el hábito (Lizana *et al.*, 2020), el afecto, la actitud y los factores sociales (Galdames *et al.*, 2011), la actitud, la norma subjetiva, la intención y el control conductual percibido (Devika *et al.*, 2020). Más aún, Soza-Parra y Cats (2023) realizaron una revisión bibliográfica, concluyendo que los motivos instrumentales, la autonomía, los factores afectivos, motivos simbólicos, la norma social, y motivos medioambientales son factores influyentes en la decisión de viajar en automóvil.

Se tienen distintas variables que serán tentativas para incluir en la modelación. En particular, se evaluará cómo la omisión de variables psicosociales, que pueden ser significativas para la elección de modo del usuario, incide en los resultados de los modelos de elección discreta. Para la especificación de las variables psicosociales, y su rol en la conducta de los usuarios, se utilizó la Teoría del Comportamiento Interpersonal de Triandis (1977), abreviada como TCI; la Figura 2.1 muestra un esquema de la TCI.

Endogeneidad

Este modelo de comportamiento incorpora de manera explícita las variables psicosociales que se estudiarán. A grandes rasgos, se tienen tres tipos de factores que inciden en la intención de un individuo.

- Actitud: compuesta por la expectativa y el valor hacia un objeto actitudinal.
- Factores sociales: constituidos por norma, rol y autoconcepto.
- Factores afectivos: fundamentalmente condicionados por las emociones.

Además, se tiene la influencia del comportamiento no razonado repetitivo en la conducta de un individuo, que resulta en el factor del hábito. Finalmente, tanto el hábito como la intención son mediados por las condiciones facilitadoras para llevar a cabo cierta conducta. Para el caso de este estudio, la conducta mencionada será el modo de transporte que eligen los usuarios en sus viajes hacia la universidad.

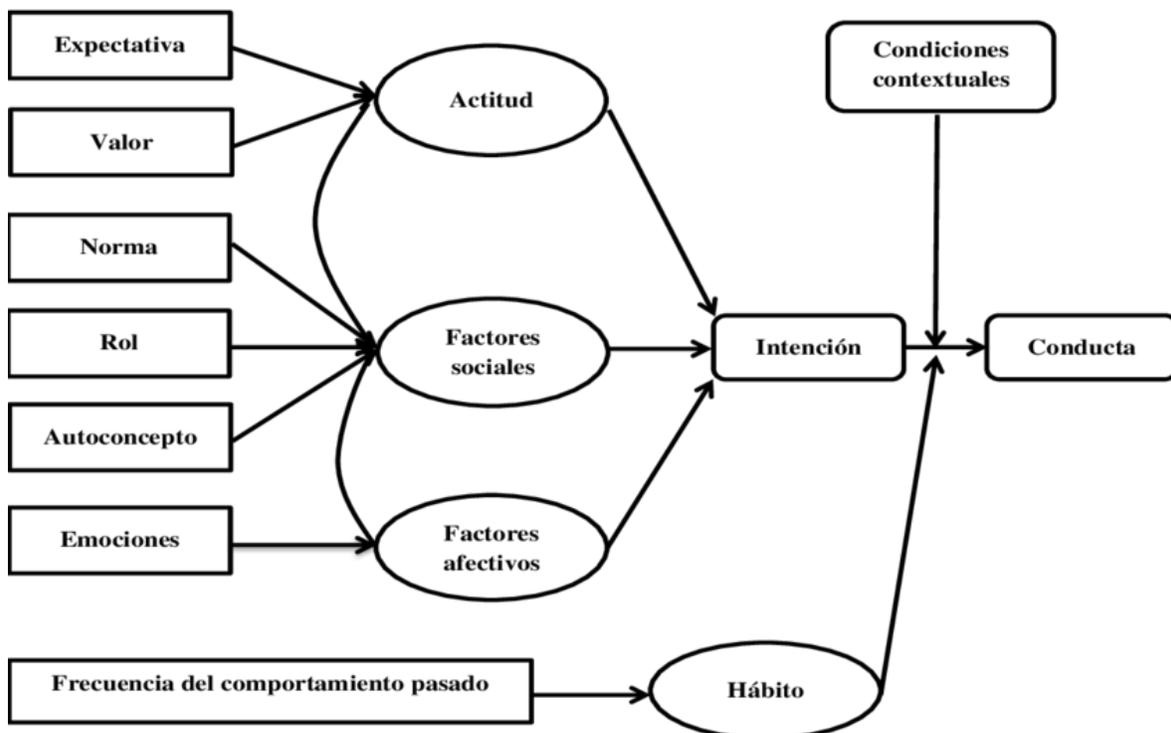


Figura 2.1: Esquema de la Teoría del Comportamiento Interpersonal de Triandis (1977)

Alternativamente, existen teorías análogas a la TCI de Triandis, que buscan explicar la conducta de las personas: la Teoría de la Acción Razonada (Fishbein y Azjen, 1975) y la Teoría de la Conducta Planificada (Azjen, 1999). Por lo tanto, se podrían tener distintos

Endogeneidad

enfoques para la incorporación de variables psicosociales en la conducta, pero se considera la teoría de Triandis, pues los fundamentos entre teorías son similares. Más aún, se tiene que las teorías mencionadas corresponden a subconjuntos de la TCI de Triandis, por lo que se reafirma el enfoque elegido.

La actitud está definida como una evaluación personal y duradera de algún aspecto del objeto actitudinal, y tiene una dirección y una intensidad específica (Domarchi *et al.*, 2010). Para su interpretación se siguió la Teoría de Expectativa-Valor (Reeve, 2005), donde se postula que esta viene dada por el producto entre la expectativa y el valor hacia un objeto actitudinal, obteniéndose una respuesta adimensional. La expectativa es el resultado esperado de una conducta por parte del usuario, y el valor es que tan importante se perciba ese resultado.

Los factores sociales corresponden a promotores de la intención para elegir un modo de transporte (Tudela *et al.*, 2011) y corresponden a la influencia que tiene la sociedad sobre el comportamiento del individuo. Este factor tiene tres dimensiones, ya mencionadas:

- Rol: comportamiento esperado dada una cierta ubicación en una escala social.
- Norma social: peso relativo de la conducta de los otros sobre la conducta de un individuo.
- Autoconcepto: visión de cada cual sobre sí mismo. Relacionado directamente con la auto valoración afectiva.

Los factores afectivos están relacionados con el valor emocional asignado a un objeto o a las acciones que se realizan con dicho objeto, tal como un modo de transporte, donde este aspecto es fundamentalmente instintivo (Tapia, 2014). Se identifican cuatro dimensiones para el afecto, donde Osgood *et al.* (1976) definen las tres primeras:

- Evaluación: se define como la respuesta actitudinal de la persona sobre un concepto.
- Potencia: se define como la percepción del individuo sobre la fortaleza del objeto actitudinal.
- Activación: indica el nivel de alerta que significa el objeto evaluado en el usuario.

Corraliza (1987) agrega una cuarta dimensión para el afecto:

- Control: el nivel de manejo que percibe el individuo frente al objeto actitudinal.

El hábito está postulado explícitamente como mediador de la conducta, y lo define como una predisposición a actuar según el ejercicio previo, sin un razonamiento, no estando condicionado por una intención. Surge a partir de una condición de estabilidad y es muy difícil de cambiar.

Las condiciones contextuales (o facilitadoras) se refieren a la facilidad o dificultad de llevar a cabo lo que se ha identificado como intención, correspondiendo a varios atributos que representan las características socioeconómicas y demográficas de quien toma la decisión, y la atractividad relativa de las otras alternativas (Idris *et al.*, 2015). Los signos de las utilidades marginales que se obtengan explicarán si efectivamente la condición facilita el comportamiento (signo positivo), o si lo desincentiva (signo negativo).

2.5. Conclusión

En el presente capítulo se describe detalladamente el fenómeno de la endogeneidad, las causas que puede tener, las implicaciones que tiene para los modelos de elección discreta y con las consecuencias que significaría para la evaluación de políticas públicas sin su corrección. Además, se describen las metodologías más utilizadas para su corrección y evaluación, y se concluye que, por sus características, se evaluarán el método MIS y el método ICLV. También se presentaron las variables psicosociales que inciden en el comportamiento de las personas, con la estructura postulada por Triandis, que explica el comportamiento de los usuarios a partir de dichas variables.

CAPÍTULO 3: DATOS USADOS PARA LA MODELACIÓN

3.1. Introducción

En este capítulo se analizan los datos obtenidos de la encuesta realizada a los estudiantes, académicos, y personal de apoyo y administrativo de la Universidad de Concepción (UdeC) y de la Universidad de la Frontera (UFRO). Se comienza con una descripción general de la información, para posteriormente, se realizar los análisis estadísticos pertinentes, con la finalidad de evaluar la consistencia de los datos, pues datos que no cumplan con esta condición pueden inducir anomalías en los modelos. Como se tiene un muestreo por conveniencia, se debe realizar una estadística descriptiva apropiada, para comprender de mejor manera el comportamiento de los usuarios estudiados.

3.2. Medición de Atributos de los Viajes y de Variables Psicosociales

Los datos utilizados provienen de dos encuestas realizadas en contextos universitarios en Chile durante 2017 y 2019. En particular, la primera fue realizada en la UdeC entre la primera mitad de 2017 y la segunda mitad de 2018, recibiendo 4.643 respuestas. De éstas, se tiene 700 respuestas válidamente emitidas, correspondientes a estudiantes, y trabajadores de la UdeC. De las respuestas válidas, se deben excluir los usuarios cautivos, esto es, los usuarios que tienen un solo modo de transporte disponible, ya que no son útiles para estimar los modelos. Luego de este filtro resulta un total de 525 usuarios válidos para el análisis. En términos de infraestructura, se tiene que Concepción posee una red segregada de ciclovías que alcanzó los 14 km el año 2018, generando una buena conectividad entre el campus de la UdeC y diferentes sectores (Lizana *et al.*, 2021). Por otro lado, en la UFRO la encuesta se llevó a cabo entre noviembre de 2018 y mayo de 2019. Se recibieron 608 respuestas válidamente emitidas, y, al aplicar los mismos criterios que para la UdeC, se obtuvieron 557 usuarios factibles de análisis. La red de ciclovías de Temuco era de 28 kilómetros aproximadamente al año 2019, pero la mayoría de las ciclovías están ubicadas en caminos rurales, y no se tiene una conexión directa con el campus de la UFRO. Con respecto al transporte público de Concepción, se tienen distintos modos: bus, taxi colectivo y tren, pero

en este caso sólo se analizó el bus, pues resultó ser el más utilizado. Generalmente, los buses transitan por las calles troncales del centro de Concepción y llegan hasta las diez comunas que constituyen el Gran Concepción. Para el caso de Temuco, sólo se tiene bus y taxi colectivo, y se ofrecen trazados que llegan directamente a la UFRO. La operación de los buses en ambas ciudades es similar en términos de frecuencias y tarifas.

Es importante mencionar que para los datos de la bicicleta se consultó por la distancia, en vez del tiempo, por lo tanto, se debió generar la conversión para evaluar el valor subjetivo del tiempo de viaje. Se utilizó una velocidad promedio de 15 km/h para los usuarios de la bicicleta, lo que se considera un valor adecuado según la literatura (Jensen *et al.*, 2010).

La encuesta constó de 104 preguntas y fue enviada por correo electrónico. La literatura respalda que la recolección de datos en universidades ha demostrado ser eficiente en estudios de viajes en bicicleta y transporte público, principalmente porque se tiene consideración geográfica (heterogeneidad en la ubicación de los hogares de los usuarios) y sencillez de toma y procesamiento de datos vía *e-mail* (Thigpen *et al.*, 2015).

Si bien el número de preguntas es alto, y podría resultar tedioso para responder, la mayoría de las preguntas son de valoración rápida. Dicha encuesta comenzó por consultar características sociodemográficas de los usuarios. Luego, se realizaron preguntas sobre el viaje que se realizó a la universidad el mismo día de aplicación del instrumento, y sobre características del modo e infraestructura que se utilizó; detalles sobre esto se pueden encontrar en Lizana *et al.* (2021). Finalmente, preguntó por las características psicosociales de los usuarios y con respecto a modos específicos (automóvil, bicicleta y transporte público), como el hábito, la actitud, el afecto y la norma social. Se preguntó por diversas características de movilidad para el viaje hacia la universidad el día en que se realizó la encuesta. Se tiene el modo elegido, el tiempo de viaje, el costo asociado y la disponibilidad para los modos encuestados. Para las condiciones contextuales, generalmente se tienen variables binarias (para la presencia o ausencia de una variable, como si el usuario es de género masculino) o porcentuales (por ejemplo, para la cantidad del tramo que posee ciclovías); más adelante se mencionan las condiciones contextuales utilizadas en este trabajo.

Respecto de las variables psicosociales, se tienen tres métodos de medición: la escala de Likert (Likert, 1934), el diferencial semántico de Osgood (Osgood *et al.*, 1976) y el cuestionario de respuesta-frecuencia de Verplanken (Verplanken *et al.*, 1994). Cabe destacar que todas las variables medidas son adimensionales.

Para la actitud, se tienen de la encuesta la expectativa y el valor que el usuario le asigna a cada modo de transporte. Ambos factores se midieron usando la misma escala de Likert de 5 puntos, que representa el grado de conformidad que presenta el encuestado con respecto a una afirmación presentada, teniendo un número entre 1 y 25 para la actitud. La encuesta consiste en una serie de declaraciones que el usuario debe valorar según le parezca, por ejemplo: “Es importante que mi viaje a la universidad sea amigable con el medio ambiente”. Para dicha afirmación se puede responder una de las siguientes alternativas: muy en desacuerdo, en desacuerdo, indiferente, de acuerdo y muy de acuerdo. Es importante ofrecer un número impar de respuestas posibles para poder capturar si el usuario es indiferente ante tal declaración. Para esta variable se contó con seis dimensiones directamente obtenidas de la encuesta, por lo tanto, se le preguntó a cada individuo qué tanto relaciona cada uno de los siguientes aspectos con el modo de transporte en cuestión: diversión, económicamente barato, saludable, seguro, ambientalmente amigable y cómodo.

El diferencial semántico de Osgood, que corresponde a una escala de valoración orientada a capturar la connotación del significado de un concepto (Osgood *et al.*, 1975), se usó para medir la componente afectiva. En este instrumento se plantean dos antónimos en cada extremo de una escala, calificando un concepto según los adjetivos ofrecidos; por ejemplo: “Relaciona la bicicleta con los siguientes conceptos”, donde las alternativas son incómodo y cómodo. Este método dependerá directamente de las emociones del usuario sobre los adjetivos planteados, pudiendo surgir anomalías. Un ejemplo concreto de esta irregularidad observado en la literatura es la discrepancia entre la población chilena y la población europea al mencionar la palabra “desértico”. Se ha notado que el usuario chileno tiende a asociarlo como un sinónimo de “árido”, lo que se atribuye principalmente a su cercanía con el desierto de Atacama, mientras que el usuario europeo lo tiende a asociar con “solitario”.

El diferencial usado tenía cuatro pares de adjetivos antónimos para cada una de las cuatro dimensiones, para cada modo. Para la dimensión de evaluación, se tenían los siguientes pares de antónimos: Bueno-Malo, Incómodo-Cómodo, Agradable-Desagradable y Limpio-Sucio. Mientras que, para la dimensión de potencia, se tenían los siguientes pares de antónimos: Fuerte-Débil, Grandioso-Insignificante, Flexible-Inflexible y Grande-Pequeño. Por otro lado, para la dimensión de activación se tenían: Poblado-Despoblado, Rápido-Lento, Activo-Pasivo y Tranquilo-Intranquilo. Por último, para la dimensión de control se tuvo: Sencillo-Complejo, Seguro-Inseguro, Claro-Oscuro y Conocido-Desconocido. Para cada par de antónimos se consideró que el primero de los dos nombrados correspondía al adjetivo positivo, siendo el de mayor puntuación (7 puntos).

Por último, se tiene el cuestionario de respuesta-frecuencia de Verplanken *et al.*, (1994), que consiste en una lista de actividades hipotéticas donde el usuario debe responder en qué modo de transporte las realizaría, donde las alternativas disponibles son automóvil particular, transporte público y bicicleta. Además, se debe enfatizar en que los usuarios respondan rápidamente y sin meditar mucho su respuesta, pues la idea es capturar el hábito que tienen los usuarios para realizar esas actividades, sin mayor raciocinio y que sean sinceros con su elección. Luego, el valor del hábito será un número entero del 0 al 10, desagregado por modo. Se tienen tres principios en dicho método de medición: no se recopila información sobre la frecuencia del comportamiento pasado, considera más de un viaje específico y es de fácil aplicación para la medición de variables psicosociales (Lizana *et al.*, 2021). El listado de actividades propuesto fue: actividades al aire libre, visitar a un amigo, visitar a un familiar, ir de compras, hacer deporte, salir de noche, ir de compras al supermercado, ir a un restaurant, ir al cine e ir a la estación de bus.

En la Tabla 3.1, se resumen los métodos de medición utilizados, desagregados por variable psicosocial.

Tabla 3.1: Métodos de medición para variables psicosociales

Variable psicosocial	Método de medición
Actitud	Escala de Likert de 5 puntos
Valoración afectiva	Diferencial semántico de Osgood
Hábito	Cuestionario de respuesta-frecuencia de Verplanken

Se generó una base de datos, correspondiente a las respuestas a la encuesta realizada. En particular, interesan de dicha base de datos los factores psicosociales de los usuarios, y como inciden éstos en la elección de su modo de transporte.

3.3. Análisis de Consistencia Estadística

Es importante revisar si hay consistencia en las respuestas de los usuarios, pues respuestas inconsistentes inducirían anomalías en los resultados. Para lo anterior se utilizará el α de Cronbach (Tavakol y Dennick, 2011). Dicho indicador adoptará valores entre 0 y 1, y busca explicar la consistencia entre las respuestas de los usuarios, al medir un mismo objeto. Cuanto más próximo este el valor a 1, significará que las preguntas son más consistentes entre sí. Valores aceptados se consideran sobre 0,5. La Ecuación 3.1 presenta el cálculo para el indicador descrito, donde k es la cantidad de preguntas incluidas en la escala, σ_j^2 es la varianza en las respuestas a la pregunta x y σ_T^2 es la varianza total de todas las preguntas.

$$\alpha = \frac{k}{k-1} \left[1 - \frac{(\sum_{x=1}^k \sigma_x^2)}{\sigma_T^2} \right], \quad (3.1)$$

Es importante recalcar que este análisis sólo aplica para preguntas que están midiendo lo mismo, pero de distinta manera. Como solo se tiene este caso para las dimensiones del afecto, únicamente se analizarán estas variables. La disponibilidad de distintas preguntas o *indicadores* para evaluar una misma variable será de suma importancia para un análisis posterior.

Tabla 3.2: Valores del α de Cronbach desagregados según dimensión del afecto

Dimensión	α de Cronbach para el afecto		
	Bicicleta	Automóvil	Transporte público
Evaluación	0,58	0,53	0,48
Potencia	0,46	0,69	0,52
Activación	0,26	0,48	0,57
Control	0,68	0,67	0,68

La mayoría de los resultados arroja respuestas consistentes entre sí, a excepción de la dimensión de activación para la bicicleta y para el automóvil, la dimensión de potencia para la bicicleta y la dimensión de evaluación para el transporte público. De esta manera, se tiene una propuesta tentativa de las dimensiones a incluir en la modelación. Esto deberá complementarse con un análisis factorial, que se presenta en los anexos.

Por último, el software *RStudio* entrega como resultado el valor que tendría el α de Cronbach si se omitiese un indicador, teniendo una suerte de análisis de sensibilidad asociado a una dimensión del afecto. Esta información será relevante, pues se utilizará como criterio para el uso de indicadores a futuro.

3.4. Características Sociodemográficas de la Muestra

En la Tabla 3.3 se muestra el análisis de frecuencias de los datos obtenidos a partir de la encuesta. Los porcentajes desagregados a nivel de modo se refieren a la fracción de usuarios del modo en cuestión, mientras que los porcentajes del total se refieren a la porción de usuarios de los tres modos en conjunto.

Se tiene una muestra diversa con respecto al ingreso de los usuarios. En cuanto a la ocupación, la mayoría de los encuestados eran estudiantes de pregrado, seguido de los trabajadores de la universidad y, en menor medida, estudiantes de posgrado. En términos etarios, se tiene un 48% de usuarios menores a 25 años, un 21% entre 25 y 34 años, un 13% entre 35 y 44 años, un 13% entre 45 y 59 años y un 4% sobre 59 años. Se tiene una ligera mayoría en cuanto al género por parte del género femenino (53%).

De alguna manera, y preliminarmente, se explican algunas elecciones de los usuarios a partir de las condiciones contextuales. Por ejemplo, pocos usuarios de bicicleta y transporte público viajaron acompañados, mientras que la mitad de los usuarios de automóvil viajó con acompañante. Además, gran cantidad de los usuarios de bicicleta viajó menos de 2,5 km para llegar a la universidad (68%). Cabe destacar que para los viajes en automóvil se consideraron

Datos Usados Para la Modelación

viajeros como pasajeros y como conductores, por lo que se explica que no todos los individuos tengan licencia de conducir (89%).

Tabla 3.3: Características sociodemográficas de la muestra

Variable	Bicicleta		Automóvil		Bus		Total	
Tamaño de la muestra	122	(11%)	423	(39%)	540	(50%)	1085	
Ocupación								
Pregrado	87	(71%)	95	(22%)	428	(79%)	610	(56%)
Posgrado	12	(10%)	19	(4%)	29	(5%)	60	(6%)
Académico	10	(8%)	178	(42%)	17	(3%)	205	(19%)
No académico	13	(11%)	131	(31%)	66	(12%)	210	(19%)
Edad (años)								
≤24	65	(53%)	79	(19%)	374	(69%)	518	(48%)
25-34	43	(35%)	83	(20%)	105	(19%)	231	(21%)
35-44	8	(7%)	106	(25%)	30	(6%)	144	(13%)
45-59	6	(5%)	113	(27%)	27	(5%)	146	(13%)
≥60	0	(0%)	42	(10%)	4	(1%)	46	(4%)
Ingreso (CLP·1000)								
≤700	50	(41%)	69	(16%)	335	(62%)	454	(42%)
700-1250	33	(27%)	83	(20%)	120	(22%)	236	(22%)
1250-2250	29	(24%)	109	(26%)	63	(12%)	201	(19%)
>2250	10	(8%)	162	(38%)	22	(4%)	194	(18%)
Condiciones facilitadoras								
Género masculino	88	(72%)	199	(47%)	226	(42%)	513	(47%)
Niños en casa	24	(20%)	161	(38%)	131	(24%)	316	(29%)
Licencia de conductor	68	(56%)	375	(89%)	188	(35%)	631	(58%)
Viaje acompañado	5	(4%)	212	(50%)	64	(12%)	281	(26%)
Viaje con objetos livianos	15	(12%)	135	(32%)	118	(22%)	268	(25%)
Distancia <2,5 km	83	(68%)	124	(29%)	147	(27%)	354	(33%)

El mayor porcentaje de los usuarios del automóvil son de ingreso alto y académicos de las universidades. Por otro lado, generalmente los usuarios del transporte público y de la bicicleta son estudiantes de pregrado, de nivel socioeconómico bajo o medio-bajo y de una edad menos a 25 años. Son pocos los casos donde un usuario de ingreso alto viajó en bus o bicicleta.

En la Tabla 3.4 se muestran las evaluaciones de los usuarios para las variables psicosociales consultadas, en términos de promedios y desviaciones estándar. Para el hábito, esto se

Datos Usados Para la Modelación

muestra tanto para los usuarios que escogieron el modo en cuestión, como para la totalidad de la muestra. Se tienen hábitos bastante bajos para la bicicleta y para el transporte público, y relativamente altos para el automóvil, considerando que el mínimo es 0 y el máximo es 10 para desarrollar las actividades potencialmente realizadas en dicho modo.

Las actitudes (que pueden tomar valores entre 1 y 25, pues es el producto entre la expectativa y el valor) hacia la bicicleta se notan cercanas o sobre el valor medio. Para el automóvil sucede lo mismo, a excepción de la dimensión ambiental y de diversión. En cuanto al transporte público, sólo dos dimensiones superan la media: la económica y la de seguridad.

Para el afecto (que varía entre 1 y 7) se tiene que el valor medio es 4. Por lo tanto, todas las dimensiones para la bicicleta y el automóvil están sobre la media, mientras que para el transporte público ninguna dimensión supera dicho valor.

Tabla 3.4: Promedios y desviaciones estándar de las variables psicosociales

Variable	Promedios			Desviaciones estándar		
	Bicicleta	Auto	Bus	Bicicleta	Auto	Bus
Hábito						
Todos los usuarios	1,2	4,6	2,6	1,9	3,5	2,7
Usuarios respectivos	4,1	7,4	4,1	2,6	2,6	2,8
Actitud						
Comodidad	13,4	20	12,8	5	5,9	5,3
Diversión	11,8	11	8,2	5,1	5,5	4,2
Económico	21,3	12	16,8	5,3	5	5,4
Salud	18,2	14	11,1	6	5,5	4,9
Ambiental	19,6	9,9	11,8	5,7	4,7	5,7
Seguridad	12,1	20,2	15,4	4,8	5,6	5,3
Afecto						
Evaluación	4,4	4,9	3,5	1,3	1,1	1,2
Control	4,4	5,3	3,8	1,1	1,1	1,3
Activación	4,4	4,1	4	1	0,8	1,3
Potencia	4,1	4,3	4	1,2	1,2	1,1

Como se busca estudiar el comportamiento de un grupo en específico, se tiene un muestreo por conveniencia. Dicho muestreo no es aleatorio y, generalmente, se tienen características similares en la población objetivo, por lo que no necesariamente se asegura representatividad.

Sin embargo, se asegura que la encuesta se realiza de una manera fácil, rápida y económica, y se relaciona directamente con los objetivos planteados.

3.5. Variables Consideradas en la Modelación

En la Tabla 3.5 se presenta la totalidad de las variables usadas para modelar junto con los valores que estas pueden tomar. Esta lista corresponde a las variables originalmente utilizadas, pues se espera que haya coeficientes que no sean estadísticamente significativos. Una vez realizado el filtro de dichos coeficientes, se tendrá el modelo definitivo. Se tiene información de la bicicleta, el automóvil y el transporte público, por lo que se estudiarán estos tres modos de transporte. Hay algunas variables que se escalaron, lo que se realiza para comparar ordenes de magnitud parecidos, y facilitar su comprensión al lector y al modelador.

Tabla 3.5: Variables consideradas en la modelación

Categoría	Variable	Descripción
Atributos tradicionales		
Atributos del modo	Tiempo de viaje	Tiempo total de viaje · 100, en minutos.
	Costo de viaje	Costo monetario de viaje · 1000, en \$CLP.
Condiciones facilitadoras		
Características	Género	Binaria. 1 si el encuestado es varón.
Sociodemográficas	Licencia de conducir	Binaria. 1 si el encuestado tiene licencia de conducir.
	Edad 25-35	Binaria. 1 si el encuestado tiene entre 25 y 35 años.
	Estudiante	Binaria. 1 para estudiantes de pregrado o posgrado.
	Niños en casa	Ordinal. Número de niños en casa menores a 15 años.
	Ingreso alto	Binaria. 1 si el ingreso familiar supera los \$CLP 2.250.000.
Infraestructura	Ciclovías	Proporción de la ruta utilizada con presencia de ciclovías.
	Estacionamientos bicicleta	Binaria. 1 si hay estacionamientos para bicicletas disponibles.
Familiaridad con la bicicleta	Habilidad bicicleta	Ordinal. De 1 a 5 según la habilidad declarada.
	Ciclismo de fin de semana	Ordinal. De 1 a 4 según la frecuencia de ciclismo los fines de semana.
Temas prácticos	Acompañante	Binaria. 1 si el encuestado suele viajar acompañado.
	Distancia	Binaria. 1 si la distancia del viaje es menor a 2,5 km.
	Objetos livianos	Binaria. 1 si el usuario suele viajar sólo con objetos livianos.
Variables psicosociales		
Afecto	Evaluación modo <i>i</i>	Indicador para la evaluación del modo <i>i</i> (1-7).
	Potencia modo <i>i</i>	Indicador para la potencia del modo <i>i</i> (1-7).
	Activación modo <i>i</i>	Indicador para la activación del modo <i>i</i> (1-7).
	Control modo <i>i</i>	Indicador para el control del modo <i>i</i> (1-7).
Actitud	Comodidad modo <i>i</i>	Indicador para la comodidad del modo <i>i</i> (1-25).
	Diversión modo <i>i</i>	Indicador para la diversión del modo <i>i</i> (1-25).
	Ambientalismo modo <i>i</i>	Indicador para los factores ambientales del modo <i>i</i> (1-25).
	Economía modo <i>i</i>	Indicador para los factores económicos del modo <i>i</i> (1-25).
	Salud modo <i>i</i>	Indicador para los factores de salud del modo <i>i</i> (1-25).
Hábito	Seguridad modo <i>i</i>	Indicador para la seguridad del modo <i>i</i> (1-25).
	Hábito modo <i>i</i>	Cantidad de actividades realizadas en el modo <i>i</i> (1-10).

3.6. Conclusión

En este capítulo se describieron los datos que se utilizarán para los modelos que se presentan en el Capítulo 4. Se presentan las características de movilidad y psicosociales de los usuarios. También se estudió por modo de transporte, según correspondiese. Esto es de particular interés, dado el muestreo por conveniencia que se tiene. Para cada variable psicosocial de la TCI se explican las respectivas metodologías de medición.

Además, se analizaron las consistencias entre las respuestas de los encuestados, obteniendo las variables, y las dimensiones de las variables, que serán apropiadas para analizar. De esta manera, se comienzan a definir las variables preliminares a incluir en la modelación.

CAPÍTULO 4: METODOS DE MODELACIÓN

4.1. Introducción

Los modelos de demanda se basan en relaciones observadas para grupos de individuos, donde se busca representar la atractividad de una alternativa i para un individuo n como una función de utilidad (U_{in}), donde se pueden identificar las variables que serán relevantes para la elección de los usuarios (Ortúzar y Willumsen, 2011). Dichos modelos se han utilizado en diversas disciplinas a partir de la década de los setenta, generando un particular campo de investigación en la ingeniería de transporte.

Los modelos arrojan los coeficientes θ_i que replican las elecciones observadas, los coeficientes máximo-verosímiles, que son los que maximizan la probabilidad de que un hecho modelado haya sucedido. Dichos coeficientes se interpretan como la utilidad marginal de la variable asociada θ_i , junto con un intercepto θ_i^0 , denotado usualmente en la literatura como ASC (constante específica de la alternativa). La notación del superíndice del intercepto se mantendrá en todo el documento para ecuaciones análogas.

Estos modelos suelen utilizar variables de movilidad, correspondientes a características del modo de transporte y del viaje en sí (x_{in}). Ejemplos de esto son el tiempo de viaje, el tiempo de espera, el costo, la disponibilidad de estacionamientos, entre otros. Además, se tiene una componente de error asociada (e_{in}) como complemento a la parte determinística de la utilidad. En la Ecuación 4.1 se presenta un ejemplo genérico de lo mencionado.

$$U_{in} = \theta_i^0 + \sum_n \theta_i \cdot x_{in} + e_{in}, \quad (4.1)$$

Si bien esto ha servido a lo largo de una disciplina que es relativamente nueva, se considera que este enfoque tiene margen de crecimiento y que tiene muchos supuestos que en la práctica no necesariamente se cumplen. Por ejemplo, se supone que los individuos son seres racionales, es decir, eligen el modo que les reporta la mayor utilidad, y si bien muchos lo son, no necesariamente todos. Otro ejemplo es que se puede dar el caso de que una persona esté

dispuesta a sacrificar parte de una variable de servicio (como el costo monetario o el tiempo de viaje) para aportar al medioambiente, eligiendo un modo de transporte más sustentable, pero que demore más. Estos modelos no están exentos de endogeneidad, por lo tanto, se debe generar un análisis pertinente para evitar que las conclusiones realizadas a partir de los resultados estén sesgadas. En el presente capítulo se describirán los métodos mencionados anteriormente para corregir este problema: el método MIS y el método ICLV.

En este trabajo se incluirán variables que no se pueden medir directamente ni observar, sino que deben ser medidas a partir de otros factores observados o indicadores, dichas variables se denominan variables latentes. La inclusión de variables latentes en los modelos de elección discreta, en conjunto con las mencionadas variables de servicio, se denomina comúnmente en la literatura como modelo híbrido de elección (Bolduc y Daziano, 2010).

4.2. Modelos de Elección Discreta Aplicados a Partición Modal

Los modelos de elección discreta permiten determinar la importancia relativa de las variables que inciden en la elección de los usuarios. Para el caso de estudio, dicha elección corresponde al modo de transporte i frente a un conjunto finito de alternativas. Se busca explicar el comportamiento un usuario n a partir de variables explicativas que se obtienen de la base de datos. Las relaciones que se establecerán *a priori* nacen fundamentalmente de la experiencia, del sentido físico y de la estadística descriptiva realizada. A medida que se obtengan resultados, se podrá saber si efectivamente las variables propuestas explican el fenómeno en cuestión, o no.

En la encuesta realizada se consulta por el modo escogido y los atributos de dicho modo; por lo tanto, lo que se realiza es estimar la utilidad que reporta el modo que escogió el usuario. A partir de esto, lo que se obtiene son los coeficientes mencionados anteriormente: θ_i . Es decir, se obtendrá la valorización marginal que le asigna la muestra a cada atributo del modo de transporte. Es importante mencionar que el peso efectivo de cada variable en la utilidad es el producto entre el coeficiente y la variable, no sólo el coeficiente. Esto es relevante para

evitar confusiones al evaluar únicamente las magnitudes de los coeficientes y justifica el hecho de que algunas variables se escalen.

La función de utilidad del modelo de referencia se muestra en la Ecuación 4.2 y se denomina de esta manera porque no incluye corrección por endogeneidad. La ecuación arrojará la utilidad (U_{in}), a partir de atributos (x_{in}), coeficientes que representan la valorización marginal del usuario por cada atributo (θ_i) y un factor de error asociado (e_{in}) que es independiente, exógeno e idénticamente distribuido (IID) con distribución Gumbel. El supuesto de la distribución de la componente de error se mantiene para todos los modelos posteriores. Además, se incluye una serie de condiciones contextuales (W_{inl}), con su coeficiente respectivo (μ_{il}).

$$U_{in} = \theta_i^0 + \sum_n \theta_i \cdot x_{in} + \sum_l \mu_{il} \cdot W_{inl} + e_{in}, \quad (4.2)$$

En particular, se generó un *modelo logit multinomial (MNL)*, según lo descrito anteriormente, donde la probabilidad de que un usuario elija un determinado modo viene dada por la Ecuación 4.3, en función de la utilidad reportada por el modo escogido y de la utilidad de los demás modos disponibles.

$$P_{in} = \frac{\exp(U_{in})}{\sum_j \exp(U_{jn})}, \quad (4.3)$$

Se propuso un modelo a partir del *modelo de utilidad aleatoria*, que sirve como marco teórico para el presente trabajo. Se supuso que los usuarios serán seres racionales y escogen el modo que les reporte la mayor utilidad. Otro supuesto importante es que los usuarios poseen información perfecta, luego son capaces de discriminar razonablemente entre alternativas.

Por otro lado, la estimación de los coeficientes de los modelos se realizó en *Python*, usando como referencia modelos propuestos en *Pandas Biogeme* (Bierlaire, 2003). Este es un paquete de código abierto diseñado para la estimación de máxima verosimilitud de modelos paramétricos. Se debieron hacer modificaciones pertinentes para el caso de estudio,

principalmente en las variables utilizadas y la metodología de estimación. Cabe destacar que se realizaron modificaciones sustanciales a la plantilla ofrecida, pues no se encontraron códigos de libre uso que incorporaran más de una variable latente. El hecho de incorporar y estimar múltiples variables latentes fue uno de los principales desafíos de la presente Tesis. Otro desafío importante fue ajustar la plantilla proporcionada a una escala de 7 puntos, en vez de la plantilla disponible que tiene sólo 5 puntos.

4.3. Método de los Indicadores Múltiples

Como se presentó en el Capítulo 2, la endogeneidad que se evaluará corresponde a la que se presenta cuando se omiten variables significativas para la modelación, en particular variables psicosociales relevantes para la elección de modo de los usuarios. El primer método para corregir dicha endogeneidad corresponde al Método de los Indicadores Múltiples (MIS, por sus siglas en inglés). Fue originalmente presentado por Wooldridge (2010) para modelos lineales genéricos, y su aplicación a los modelos de elección discreta fue realizada por Guevara y Polanco (2016), quienes argumentan que el método MIS implica una mejora en términos de ajuste, pero, más importante aún, una estimación más plausible de tasas marginales de sustitución. A grandes rasgos, se introduce un factor de corrección en el modelo de elección para obtener resultados adecuados.

Algunas de las aplicaciones de este método en modelos de elección discreta son los trabajos de Guevara y Polanco (2016) en elecciones residenciales, Fernández-Antolín *et al.* (2016) y Guevara *et al.* (2020) en elecciones de modo, Mariel *et al.* (2018) en evaluación medioambiental, y Palma *et al.* (2018) en elecciones de vinos.

Considerando la Ecuación 4.2 como modelo de comparación, se asume que se tendrá una variable omitida q_{in} para representar adecuadamente el fenómeno de la elección de modo, ponderado por su respectiva utilidad marginal (γ_i). Además, se considera que la utilidad de los usuarios dependerá de un factor económico (t_{in}), que a su vez está correlacionado con la variable omitida mencionada anteriormente. Un buen ejemplo de este factor económico es el

tiempo de viaje asociado a la alternativa i . Por lo tanto, la función de utilidad adecuada para dicha alternativa viene dada por la Ecuación 4.4.

$$U_{in} = \theta_i^0 + \sum_n \theta_i \cdot x_{in} + \theta_t \cdot t_{in} + \gamma_i \cdot q_{in} + e_{in}, \quad (4.4)$$

donde θ_i^0 , θ_i , θ_t y γ_i son coeficientes por estimar y e_{in} es la componente de error asociada. Si el término $\gamma_i \cdot q_{in}$ fuera omitido, estaría incluido en la componente de error. Luego, dicha componente estaría correlacionada con t_{in} , causando endogeneidad por la omisión de variables significativas.

Para la corrección de esta anomalía, se asumen como mínimo dos indicadores que tienen relación directa con la variable omitida (I_{1in} y I_{2in}), definidos en las Ecuaciones 4.5 y 4.6. Se debe suponer que dichos indicadores están correlacionados entre sí, y que la única fuente de correlación es la variable endógena en el modelo. Si bien estos supuestos son fuertes, son no son más fuertes que los que proponen las demás metodologías para corregir por endogeneidad, pues, como se mencionó antes, esta corrección no es gratuita.

Se asume I_{2in} como un instrumento para I_{1in} , y que la relación entre la variable omitida y el indicador debe ser lineal. Posteriormente, se realizó un reordenamiento de variables, para la inclusión de q_{in} en la función de utilidad, que se presenta en la Ecuación 4.7.

$$I_{1in} = \alpha^0 + \alpha_q \cdot q_{in} + e_{I_{1in}}, \quad (4.5)$$

$$I_{2in} = \tau^0 + \tau_q \cdot q_{in} + e_{I_{2in}}, \quad (4.6)$$

$$q_{in} = \frac{I_{1in} - \alpha^0 - e_{I_{1in}}}{\alpha_q}, \quad (4.7)$$

Se debe suponer además que los pares de variables $(q_{in}, e_{I_{1in}})$, $(x_{in}, e_{I_{1in}})$, $(q_{in}, e_{I_{2in}})$, $(x_{in}, e_{I_{2in}})$ y $(e_{I_{1in}}, e_{I_{2in}})$ son independientes, y α_q y τ_q son distintos de 0. Al combinar las Ecuaciones 4.4, 4.5, 4.6 y 4.7, se obtiene la Ecuación 4.8 como función de utilidad. Esto aplica solamente de manera ilustrativa, pues sigue siendo endógena por la correlación entre

I_{1in} y $e_{I_{1in}}$, que resulta de la Ecuación 4.5. Además, se debe definir $\omega_q = \frac{\gamma_i}{\alpha_q}$, para facilitar la legibilidad.

$$U_{in} = \theta_i^0 + \sum_n \theta_{in} \cdot x_{in} + \theta_t \cdot t_{in} + \omega_q \cdot (I_{1in} - \alpha_0 - e_{I_{1in}}) + e_{in}, \quad (4.8)$$

La metodología consiste en formular una regresión lineal de un indicador en función de otro, para corregir los errores de medición inducidos por el indicador independiente, considerando las mismas variables en la regresión y en la función de utilidad. Así, se tiene una aplicación del Método de Control de Función (CFM, por sus siglas en inglés), propuesto por Guevara (2010), donde el instrumento es un indicador. Algebraicamente, esto se representa en la Ecuación 4.9, junto con la componente de error del indicador, en la Ecuación 4.10.

$$I_{1in} = \rho_0 + \rho_1 \cdot I_{2in} + \rho_q \cdot q_{in} + \rho_x \cdot x_{in} + \hat{\zeta}_{in}, \quad (4.9)$$

$$e_{I_{1in}} = \theta_\zeta \cdot \zeta_{in} + v_{in}, \quad (4.10)$$

El término $\hat{\zeta}_{in}$, que corresponde al residual de la regresión, captura la parte endógena de I_{1in} , y v_{in} es una componente de error que es exógena bajo los supuestos considerados. Por último, sustituyendo la Ecuación 4.10 en la Ecuación 4.8, se tiene:

$$U_{in} = \theta_i^0 + \sum_n \theta_i \cdot x_{in} + \theta_t \cdot t_{in} + \omega_q \cdot I_{1in} - \omega_q \cdot \alpha_0 - \omega_q \cdot \theta_\zeta \cdot \hat{\zeta}_{in} - \omega_q \cdot v_{in} + e_{in}, \quad (4.11)$$

Esto puede ser reordenado, y, denotando $\widetilde{\theta}_i^0 = \theta_i^0 - \omega_q \cdot \alpha_0$, $\omega_\zeta = -\omega_q \cdot \theta_\zeta$ y $\widetilde{e}_{in} = -\omega_q \cdot v_{in} + e_{in}$, resulta la Ecuación 4.12, donde ya no hay endogeneidad producida por la omisión de la variable estudiada (q_{in}).

$$U_{in} = \widetilde{\theta}_i^0 + \sum_n \theta_i \cdot x_{in} + \theta_t \cdot t_{in} + \omega_q \cdot I_{1in} + \omega_\zeta \cdot \hat{\zeta}_{in} + \widetilde{e}_{in}, \quad (4.12)$$

Si el coeficiente del indicador incluido directamente en la función de utilidad (ω_q) y el coeficiente que captura los errores de medición correspondientes a dicho indicador (ω_ζ) resultan significativos en la estimación, entonces se tenía endogeneidad en el modelo original, producto de la omisión de variables psicosociales. El coeficiente ω_ζ no tiene interpretación conductual directa, sino matemática. En palabras simples, es una corrección asociado a la función de utilidad.

Una diferencia fundamental entre los instrumentos propuestos por el CFM (mencionado en el Capítulo 2) y los indicadores planteados por el MIS es que, en el primero, se requiere al menos un instrumento por cada variable endógena, lo que condiciona su calidad y poder explicativo; por otro lado, en el MIS se requieren al menos dos indicadores por cada variable endógena, que deben satisfacer las condiciones estipuladas anteriormente. En general, los indicadores son de más fácil acceso que los instrumentos (Guevara y Polanco, 2016), pues estos últimos deben cumplir las condiciones estipuladas por Hausman (1978), esto es, estar correlacionados con la variable endógena e independientes de la componente de error de la utilidad, de manera simultánea.

Una potencial pregunta es si el aumento en el poder explicativo (o de verosimilitud) del modelo se pudiera deber exclusivamente a la inclusión un indicador válido para la variable omitida y si será necesario incluir los residuales de la regresión, que conllevan esfuerzo algebraico y de programación. Aquello no estaría solucionando el problema de fondo, que es estimar adecuadamente los coeficientes; por consiguiente, no basta con sólo incluir en indicador directamente en la función de utilidad, pues si se hiciera, aún habría endogeneidad.

La estimación de dos pasos, es decir, realizar primero la regresión y después estimar los coeficientes asociados a la función de utilidad, es más fácil de estimar y puede lidiar con un conjunto más amplio de distribuciones de las componentes de error. Mientras que la estimación *Full Information Maximum Likelihood* (en un solo paso) es usualmente más eficiente en términos de tiempo y permite el cálculo directo de errores estándar (Guevara y Ben-Akiva, 2012). Además, la estimación secuencial es más fácil de implementar, pero podría comprometer la eficiencia y requiere la estimación de los errores estándar a partir de

bootstrapping. Sin embargo, es más robusta frente a los supuestos de modelos alternativos (Guevara *et al.*, 2020). Por lo tanto, en este caso se optó por una estimación secuencial.

En estudios anteriores se planteaba que los indicadores debían ser continuos (Guevara, 2015), lo que constituía un problema debido a que, en la práctica, los indicadores se obtienen a partir de encuestas, que suelen usar escalas discretas, como las mencionadas en el Capítulo 3. Sin embargo, existe evidencia de que los indicadores discretos pueden ser tan buenos como los continuos con este método (Fernández-Antolín *et al.*, 2016), pero se debe ser claro en que esto corresponde a una aproximación. Más aún, Guevara *et al.*, (2020) validan los indicadores continuos para corregir con el MIS.

Este método ofrece la ventaja de capturar la endogeneidad asociada a una variable omitida de manera rápida, en términos de programación y cálculo; y directa, en comparación con otras metodologías que utilizan aproximaciones para calcular la verosimilitud. Además, implica supuestos que no conllevan mayor riesgo en comprometer los resultados. También ofrece la ventaja de incluir múltiples indicadores para distintas variables en la misma función de utilidad (realizando la metodología pertinente), lo que acentúa la ventaja de la rapidez de uso.

Una desventaja importante es que este modelo induce un cambio de escala en los coeficientes estimados para la utilidad, debido a la inclusión directa del indicador y de los residuales de la regresión en la Ecuación 4.12. Por lo tanto, no se pueden comparar directamente los coeficientes en términos de magnitudes, por lo tanto, se deben analizar indicadores que sean funciones de estos coeficientes, como cociente o como tasa de cambio. En este caso se analizará el valor subjetivo del tiempo de viaje y la elasticidad del tiempo de viaje y del costo. Otra limitación, también producida por la inclusión directa del indicador I_{in} y los residuales de la regresión δ_{in} en la función de utilidad, es que se pueden imposibilitar potenciales predicciones. Esto se podría solucionar con la inclusión de características sociodemográficas en las regresiones asociadas y, por consiguiente, en las funciones de utilidad (Fernández-Antolín *et al.*, 2016).

4.4. Análisis Factorial y Modelos de Ecuaciones Estructurales

Para la aplicación del método ICLV, que es mucho más costoso computacionalmente que el método MIS, es pertinente realizar un análisis factorial para reducir la cantidad de indicadores por incorporar en la modelación. El análisis factorial se utiliza para medir correlaciones entre variables a incluir en el modelo y se busca describir la correlación entre varias variables en términos de un número reducido de factores no observables (Washington *et al.*, 2011). En otras palabras, corresponde a una técnica de reducción de dimensionalidad, generando una estructura de factores comunes. Por otro lado, si no existe correlación entre las variables originales el análisis factorial no es útil para el modelo, pues se deberán incluir en este todas las variables utilizadas en un principio.

En definitiva, se busca explicar las variables latentes en función de indicadores, que aportan una carga determinada a cada variable, desagregadas por modo de transporte. Para el caso del afecto, corresponderán a los pares de antónimos desagregados según dimensión. Por otro lado, para la actitud, corresponderán a las valoraciones actitudinales según la expectativa y el valor mencionados.

Se tiene el análisis factorial confirmatorio (CFA) y el exploratorio (EFA) para relacionar las variables medidas con el aspecto que se busca medir. El primero se conduce principalmente por supuestos y expectativas, pues la relación entre variable y aspecto se plantea por el modelador. Por otro lado, el segundo busca descubrir la estructura subyacente que poseen los datos, es decir, encontrar las relaciones entre las variables y el aspecto a medir.

En este caso, se presentará el análisis factorial confirmatorio, pues se tienen los componentes de la encuesta desagregados por variable psicosocial. Dicha encuesta fue validada previamente por Domarchi (2007), y de esta manera se respeta la consistencia entre dimensiones por modo. De otra manera, es decir, realizando un análisis factorial exploratorio, se encontraron preguntas de la encuesta que generan dimensiones distintas entre modos, agrupando indicadores diferentes para estas. Por lo tanto, se optó por respetar la validación

realizada anteriormente e incorporar las preguntas de la encuesta a la dimensión definida previamente.

Los indicadores que serán relevantes para modelar la variable latente correspondiente deberán presentar una carga normalizada mayor a 0,5 para representar dicha variable. De lo contrario, serán descartados, pues se tendría una modelación adecuada para la variable latente con los indicadores que cumplan aquel requisito. Esta regla no es taxativa, sino que dependerá directamente del modelador, sin embargo, existe evidencia de que un valor de 0,5 es adecuado (Yong y Pearce, 2013).

Con respecto a los Modelos de Ecuaciones Estructurales (SEM, por sus siglas en inglés) son una herramienta ampliamente utilizada por las ciencias sociales. En particular, se usarán para la inclusión de las variables latentes en el modelo. Los SEM tienen dos componentes, el modelo de medición y el modelo estructural. El modelo de medición se preocupa de qué tan bien los indicadores miden las variables latentes y el modelo estructural analiza el qué tan bien las variables del modelo están relacionadas entre sí (Washington *et al.*, 2011). Tienen la ventaja de que pueden lidiar con variables endógenas y exógenas simultáneamente y capturar constructos latentes a partir de indicadores específicos. Estos indicadores son análogos a los utilizados en el método MIS.

En resumen, se mide cuál es el peso relativo de cada respuesta de la encuesta en la variable psicosocial que se está midiendo, estableciendo relaciones de causalidad, al mismo tiempo que se mide el peso relativo del factor psicosocial en el resultado del modelo, que en este caso es la utilidad que le reporta el modo de transporte al individuo. Los SEM incluyen un análisis factorial dentro del modelo de elección discreta, pero es crucial la reducción de dimensionalidad para optimizar tiempos de cálculo.

Las variables de personalidad no se pueden medir directamente, ni presentan una unidad de medida específica. Por lo tanto, se debió realizar una encuesta pertinente al caso para capturarlas e incorporarlas en modelos de elección discreta. Para esta incorporación se destacan dos estimaciones: secuencial o simultánea. La estimación secuencial separa las

estimaciones de modo y de variables latentes, tratando el problema en etapas distintas. Por otro lado, la estimación simultánea estima ambos coeficientes en conjunto. Ambos enfoques entregan coeficientes que no presentan diferencias en cuanto a significancia estadística, pero la verosimilitud mejora en el caso simultáneo, por lo tanto, se optó por este enfoque. Si bien las diferencias entre valores de coeficientes entre ambas estimaciones no son tan grandes como cuando se omiten las variables latentes, se puede incurrir en serios errores al momento de evaluar políticas de transporte (Raveau *et al.*, 2010).

Por otro lado, las variables latentes tienen una componente de error que está simbolizada por la parte no observable de la variable. Por lo tanto, se tiene el riesgo de que las incógnitas sean más que la cantidad de ecuaciones disponibles y se debe tener particular cuidado con aquello. Además, la estimación simultánea no impone ninguna restricción sobre el número de variables explicativas, mientras que para la estimación secuencial se requieren al menos tantas variables explicativas como variables latentes se consideren.

La definición de las variables latentes (A_{in}) se realizó a través de una serie de condiciones facilitadoras l , que corresponden a atributos sociodemográficos del usuario, familiaridad de uso con el modo en cuestión, infraestructura y temas prácticos. Este enfoque se denomina *Multiple Indicators Multiple Causes* (MIMIC). Dicho método prueba las determinantes que pueden estar afectando la evaluación de los indicadores que se usan para explicar cada variable latente (Ortúzar y Willumsen, 2011). A partir de lo anterior se define el peso de la condición facilitadora en la variable latente (W_{inl}).

De esta manera, se plantea una relación lineal estructural, en función del intercepto correspondiente al modo i (λ_i^0), la contribución marginal de cada variable latente (λ_{il}) y una componente de error aleatoria para el modo respectivo (ψ_{in}). Lo mencionado anteriormente se presenta en la Ecuación 4.13.

$$A_{in} = \lambda_i^0 + \sum_l \lambda_{il} \cdot W_{inl} + \psi_{in}, \quad (4.13)$$

Asimismo, la variable latente del afecto al modo i , explica las respuestas de los encuestados sobre los indicadores del afecto a través de ecuaciones de medición. En este caso se tienen las respuestas (I_{inr}) del usuario n , para un indicador psicosocial r como una función de la constante específica asociada a cada factor (α_{ir}^0), de la variable latente A_{in} ponderada por la carga asociada a cada factor (α_{ir}) y una componente de error aleatoria con distribución normal (z_{inr}). Esto se presenta en la Ecuación 4.14, donde el lector podrá notar que la ecuación es análoga a los indicadores utilizados en el MIS.

$$I_{inr} = \alpha_{ir}^0 + \alpha_{ir} \cdot A_{in} + z_{inr}, \quad (4.14)$$

Los resultados de las SEM se incluyen directamente en los resultados de la estimación del modelo, debido a que su estimación y la de los coeficientes asociados fue simultánea. Es importante destacar que, para cada ecuación estructural, hay una dimensión que no se calculará directamente, sino que se establecerá como intercepto de la variable latente. Esta es una práctica común en variables explicativas y se explica porque lo que interesa es las diferencias relativas entre las dimensiones, es por esto que en los resultados se incluyen valores fijados para la dimensión de evaluación y para la dimensión de comodidad. Lo anterior implica que, para las dimensiones mencionadas, el intercepto en los modelos estimados tenga un valor fijo de 0 y la carga respectiva un valor establecido de 1.

En cuanto a las herramientas computacionales utilizadas: para el análisis factorial se utilizó *RStudio*, mientras que, para los SEM, como se integran directamente en el modelo de elección, se utilizó *Pandas Biogeme* (Bierlaire, 2003).

4.5. Modelos de Variables Latentes Integradas

Otra manera de solucionar el problema de la endogeneidad es a partir del método de variables latentes integradas (ICLV, por sus siglas en inglés). En este caso, un factor latente captura un atributo cualitativo no observado (Fernández-Antolín *et al.*, 2016). Estos modelos permiten al analista cuantificar el impacto de constructos latentes en el comportamiento y

constituyen un análisis más profundo de la variable no observada. Sin embargo, se deben realizar más supuestos y conlleva un costo computacional más alto.

Los modelos ICLV han demostrado ser un enfoque más flexible, ser más explicativos y mejorar la especificación de los modelos de elección discreta. Además, ofrecen una mejor bondad de ajuste que los modelos tradicionales y, por lo tanto, una representación del comportamiento más satisfactoria (Ben Akiva *et al.*, 1997). La aplicación de estos métodos requiere una consideración cuidadosa del marco teórico conductual y del diseño de recolección de datos para generar buenos indicadores y variables causales que apoyen el marco teórico.

Estos modelos han sido usados extensivamente en la última década para entender mejor las actitudes y factores psicológicos que dirigen las preferencias de los individuos hacia bienes y políticas. La popularidad de estos modelos se basa en que utilizar indicadores directamente en el modelo de elección (sin utilizar los SEM, ni corregir los errores de medición con el método MIS); sin embargo, es considerado defectuoso metodológicamente, debido a la falta de causalidad y la inducción de errores de medición y endogeneidad (Budziński y Czajkowski, 2022).

Además, son métodos poderosos para mejorar las representaciones de la toma de decisiones, pues se espera que los modelos más complejos se ajusten mejor a los datos que modelos más simples. En general, se cree que la información adicional de los ICLV, a partir de la incorporación de indicadores de medición, lleva a una reducción en la varianza de los coeficientes estimados; sin embargo, la evidencia para respaldar esto es limitada.

Dichos modelos se han utilizado en diversos contextos, incorporando variables explicativas que anteriormente se omitían. En la literatura se han identificado interacciones entre creencias, valores y emociones cuando se elige un modo de transporte (Idris *et al.*, 2015). También se han identificado factores como la actitud, el hábito y el afecto como determinantes para explicar el comportamiento de los usuarios (Lizana *et al.*, 2020; Ababio-Donkor *et al.*, 2020).

Se componen de dos “sub-modelos”: el modelo de variables latentes y el modelo de elección discreta. El primero estima la relación estructural entre características sociodemográficas y variables latentes y el segundo estima la utilidad de cada alternativa. Luego, la verosimilitud asociada a este tipo de modelos incluye la verosimilitud del modelo de elección y de los SEM (Kim y Lee, 2023).

Si bien la TCI de Triandis postula más variables psicosociales como mediadoras del comportamiento, los factores sociales ofrecen cierta problemática, pues los usuarios no tienden a mostrar transparencia al momento de autoevaluarse. Además, se tiene un sesgo de opinión al momento de responder por la opinión que tienen terceros sobre el usuario. Por ende, no se consideraron la norma, el rol ni el autoconcepto. Sí se consideraron como variables latentes la actitud y el afecto. También utilizó el hábito como tercera variable psicosocial, pero no corresponde a una variable latente, pues se mide directamente. En la Figura 4.1 se simboliza el modelo de elección incorporando las variables latentes (x_{in}^*).

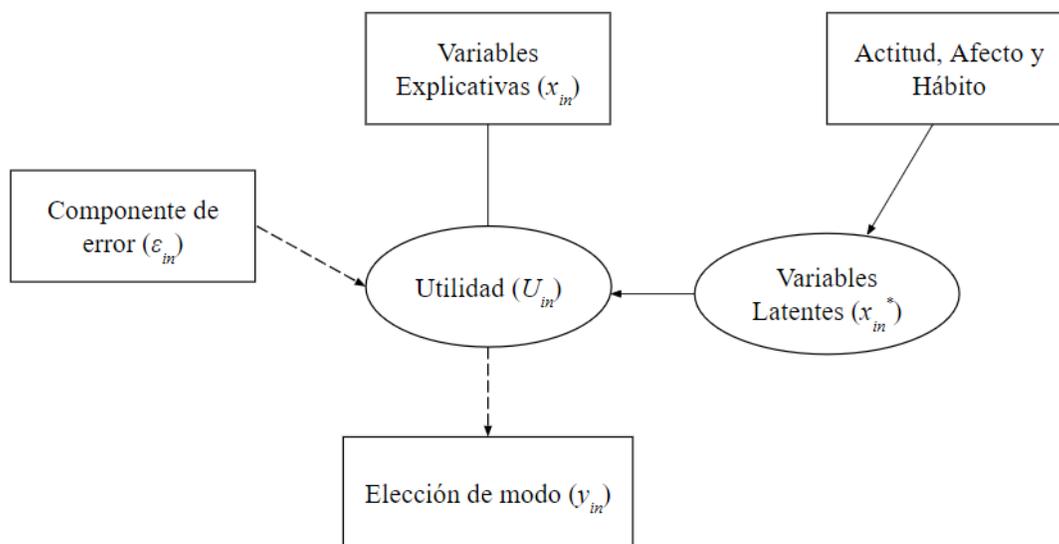


Figura 4.1: Modelo de elección incorporando variables latentes

Para la segunda metodología, se usó la notación de x_{in}^* para simbolizar al vector de dichas variables. Análogamente, a los coeficientes de las variables latentes se les denominó δ_i . Luego, se tiene la Ecuación 4.15 para el modelo ICLV. También se incorporó el hábito, que es una variable que se mide directamente. Se usará la notación H_{in} para esta variable y κ_i para su respectiva utilidad marginal.

$$U_{in} = \theta_i^0 + \sum_n \theta_i \cdot x_{in} + \sum_n \delta_i \cdot x_{in}^* + \kappa_i \cdot H_{in} + \sum_i \mu_{il} \cdot W_{nil} + e_{in}, \quad (4.15)$$

Además, pueden sufrir de ecuaciones estructurales débiles, donde las variables observables son malos predictores de las variables latentes, y de ecuaciones de medición débiles, donde las respuestas para indicadores apropiados no están disponibles o no son consultados (Vij y Walker, 2016). Otra desventaja asociada a estos modelos es que, si los datos son insuficientes, en términos de cantidad o calidad, entonces los beneficios de usar dichos modelos son limitados, y los costos asociados a su implementación serán en vano.

Sin embargo, ofrecen diversas ventajas. La primera es la habilidad de modelar explícitamente heterogeneidad no observada. Además, se tienen ganancias en eficiencia estadística de los coeficientes estimados, debido a la información adicional entregada por la inclusión de las variables latentes. Luego, se tiene mayor realismo conductual, es decir, los modelos ICLV representan de manera más transparente la toma de decisiones de las personas, en comparación con un modelo de *caja negra*. Por último, se tiene que, en términos de políticas públicas, se espera que estos modelos puedan realizar predicciones más sensibles, pudiendo generar políticas mejor dirigidas (Abou-Zeid y Ben-Akiva, 2014).

Para la estimación de más de una variable latente es necesario usar el método de simulación de MonteCarlo, en vez de integración numérica (Raychaudhuri, 2008). Esto podría ser considerado una desventaja, dado que la solución es aproximada. Este método no es determinista, y se utiliza para aproximar integrales muy complejas o costosas de calcular. Además, es más apropiado para integrales multidimensionales y ofrece la facilidad de que se puede estimar modificando el término de la probabilidad condicional en el código.

Este método en aproximar la integral de la función $g(x)$ como con la suma de un gran número de términos. El procedimiento consiste en elegir una variable aleatoria X , definida en un intervalo $[a, b]$, tal que se puedan generar iteraciones de X , y asumiendo que se conoce la función de densidad de la probabilidad (f_X). Luego, se deben generar R sorteos x_r desde X . Por último, se calcula con la Ecuación 4.16 (genérica), donde I será el símil a la ecuación de función de la verosimilitud.

$$I \approx \frac{1}{R} \sum_{r=1}^R \frac{g(x_r)}{f_X(x_r)}, \quad (4.16)$$

Se observó que la convergencia, determinada fundamentalmente por la razón de verosimilitud (*test LR*) y por la variación ínfima (o nula) de los coeficientes se logra a partir de las 5000 iteraciones, por lo que se mantuvo ese número para la estimación de MonteCarlo. Los cambios en términos de la razón mencionada son marginales a partir de esta cantidad de simulaciones.

4.6. Modelos con Interacción

Más aún, se pueden generar factores de corrección para las variables de servicio, pues se hipotetiza que las variables psicosociales estudiadas tendrán un impacto sobre, por ejemplo, el tiempo de viaje. Así, el factor que pondera a esta variable, es decir, la utilidad que se tiene por cada unidad de dicha variable estará compuesto por una suma entre el factor propio de la variable de servicio y un factor de corrección, que está condicionado por las variables psicosociales mencionadas. De esta manera, se plantea un ejemplo para el tiempo de viaje en la Ecuación 4.17, donde se tendrá el peso efectivo de la variable latente en la utilidad del usuario (x_{in}^*), multiplicado por la valoración que le atribuye el usuario a la interacción entre dicha variable y el tiempo de viaje (β_{t_v}).

$$t_v \cdot (\theta_{t_v} + x_{in}^* \cdot \beta_{t_v}), \quad (4.17)$$

Estos modelos se denominan modelos con interacción, precisamente porque la variable latente interactúa con la variable de servicio y con su coeficiente respectivo. Evidencias en la literatura han demostrado que los modelos con interacción son significativamente superiores a los modelos que incluyen las variables latentes de forma aditiva, fundamentalmente en la verosimilitud del modelo. Además, si los términos de interacción no son incluidos (y son relevantes) la distribución del término de error puede no satisfacer las hipótesis del *modelo logit multinomial*, induciendo anomalías en el modelo (Cherchi y Ortúzar, 2002). Por otro lado, los índices económicos en estudio se podrían ver modificados por los términos con interacción, generando cambios importantes en sus magnitudes.

Luego, se tiene la función de utilidad para el modelo con interacción, donde se incorporan las variables latentes interactuando con el tiempo de viaje. Dicha interacción viene dada por el producto entre la variable psicosocial en cuestión, el tiempo de viaje (t_v) y el coeficiente que se obtendrá a partir de la modelación (β_{in}). El significado de los coeficientes de interacción se interpreta como el efecto marginal de la variable psicosocial sobre la utilidad marginal del tiempo. Esto se presenta en la Ecuación 4.18.

$$U_{in} = \theta_i^0 + \sum_n \theta_i \cdot x_{in} + \sum_n \beta_{in} \cdot x_{in}^* \cdot t_v + \kappa_i \cdot H_{in} + \sum_l \mu_{il} \cdot W_{inl} + e_{in}. \quad (4.18)$$

La aplicación principal de los modelos con interacción es explorar la sensibilidad de cambios en variables para los usuarios o para realizar predicciones. Además, implican una especificación de la utilidad que no es lineal, pues se ponderan dos variables por un coeficiente. Se espera que los coeficientes que interactúan con las variables psicosociales y el tiempo de viaje (β_{in}) tengan signo positivo, porque el efecto de tener afecto, actitud o hábito positivos hacia un modo de transporte implicaría una reducción en la valorización del tiempo de viaje. Eventualmente, se podrían tener valores negativos del valor subjetivo del tiempo de viaje, lo que se interpretaría como un goce por viajar. Esto es común en usuarios que son muy apasionados por el automóvil o por personas que utilizan el transporte público para realizar tareas productivas. En ese caso, les es indiferente pasar más tiempo viajando.

Para realizar lo mencionado, en el modelo ICLV basta con definir de esta manera la función de utilidad. Para el modelo MIS, también se debe realizar lo anterior, pero se debe añadir el supuesto de que la interacción del indicador con el tiempo de viaje ($I_{in} \cdot t_{in}$) es un indicador de la interacción de la variable omitida con el tiempo de viaje ($q_{in} \cdot t_{in}$), lo que no necesariamente es trivial. En palabras simples, el supuesto es análogo al que se debe realizar para contar con indicadores válidos para el MIS, pero se extrapola a la interacción de las variables. La validación de lo anterior se puede revisar en Fernández-Antolín *et al.* (2016).

Anteriormente se mencionó que las metodologías para corregir la endogeneidad no pueden ser directamente comparadas entre sí, pero el modelo con interacción puede ser aplicado a ambas metodologías (por separado, naturalmente). Lo que sí se podría hacer es comparar, por ejemplo, un modelo MIS con un modelo MIS con interacción, pues la metodología es idéntica. Esto también aplica para los modelos ICLV. De esta manera, se espera realizar dicha comparación, además de los indicadores económicos mencionados.

4.7. Valor Subjetivo del Tiempo y Elasticidades

Como se mencionó anteriormente, es crucial para los objetivos de la presente Tesis poder comparar las metodologías para evaluar la endogeneidad. Esto, sin embargo, no se puede realizar directamente, pues el modelo MIS, al incluir directamente un indicador en la función de utilidad, introduce un cambio de escala en la especificación del modelo. Por lo tanto, una práctica común es comparar indicadores económicos, que se obtienen a partir de las utilidades marginales que resultan de los modelos. En este caso, se comparará el valor subjetivo del tiempo (VST) de viaje y la elasticidad del tiempo de viaje.

El VST (en \$/min) de un la muestrase calcula dividiendo el coeficiente del tiempo (1/min) por el coeficiente del costo monetario (1/\$), para un modelo de utilidad lineal en variables. Es importante mencionar que el coeficiente del costo es independiente del modo de transporte, al contrario del coeficiente del tiempo de viaje. Esto se debe a que el primero corresponde a la utilidad marginal del ingreso, que no depende del modo, mientras que el tiempo de viaje sí depende del modo.

Este índice corresponde a una tasa marginal de sustitución, donde se mide el *trade-off* que un usuario está dispuesto a realizar entre dos variables. En otras palabras, es la cantidad de dinero que el usuario está dispuesto a pagar para ahorrar una unidad de tiempo de viaje. Depende directamente del usuario, estando sujeto a características sociodemográficas, psicosociales y del viaje (Abou-Zeid *et al.*, 2010). Se puede interpretar de dos maneras. Primero, la voluntad de sustituir el tiempo de viaje por una actividad más placentera o útil. Segundo, la percepción directa de la reducción del tiempo de viaje en sí (Jara-Díaz y Guevara, 2003). La Ecuación 4.19 presenta el cálculo del valor subjetivo del tiempo de manera tradicional.

$$VST_{tradicional} \left(\frac{\$}{min} \right) = \frac{\theta_{tiempo}}{\theta_{costo}}, \quad (4.19)$$

Más aún, a partir de los valores de los coeficientes, de los respectivos *test t*, del nivel de significancia estadística establecido (95%) y de la correlación entre ambos coeficientes, se pueden obtener los límites inferior y superior, y la mediana del intervalo de confianza del valor del tiempo (Armstrong *et al.*, 2001). Estos valores se utilizan para realizar un análisis más profundo, pues, en la Ecuación 4.19 se obtiene sólo el promedio del valor del tiempo.

Para un modelo con interacción el coeficiente del tiempo de viaje es toda la expresión que pondera al tiempo de viaje mencionada en la Ecuación 4.20, es decir, se incluye la interacción con la variable psicosocial y se genera un ajuste para el coeficiente del tiempo, modificando su valor total. El nuevo valor subjetivo del tiempo depende directamente de los dos coeficientes mencionados anteriormente, de los coeficientes de interacción que sean significativos estadísticamente (β_{in}) y de los valores de las variables latentes obtenidos según los modelos de ecuaciones estructurales (x_{in}^*). Lo explicado se presenta en la Ecuación 4.20.

$$VST_{interacción} \left(\frac{\$}{min} \right) = \frac{\theta_{tiempo} + \sum_n \beta_{in} \cdot x_{in}^*}{\theta_{costo}}, \quad (4.20)$$

Por otro lado, la elasticidad es un concepto económico utilizado para cuantificar la variación relativa experimentada por una variable al cambiar otra, también relativamente. Para la modelación de la demanda representa el porcentaje de variación de la probabilidad de escoger una alternativa i ($P_n(i)$) a partir del incremento de un 1% en algún factor económico, típicamente el tiempo de viaje o el costo monetario de viajar. A diferencia del VST, que es genérico para el modelo, la elasticidad depende de cada individuo.

Se espera que tome valores negativos, pues un aumento en estas variables es un desincentivo para utilizar la alternativa en cuestión, y que esté entre 0 y -1 (0% y -100% en porcentajes). Se ejemplificará en función del mismo atributo económico mencionado en la Sección 4.3 (t_{in}). Luego, la elasticidad de t_{in} ($E_{t_{in}}^{P_n(i)}$) viene dada por la Ecuación 4.21.

$$E_{t_{in}}^{P_n(i)} = \frac{\partial P_n(i)}{\partial t_{in}} \frac{t_{in}}{P_n(i)}, \quad (4.21)$$

El cálculo de este índice se realiza *a posteriori*, es decir, después de la estimación del modelo. Como las variables económicas son independientes de la especificación de la función de utilidad, el cálculo de la elasticidad es análogo para todos los modelos propuestos.

La elasticidad del tiempo de viaje se interpreta como el porcentaje de variación en la probabilidad de escoger una alternativa n a partir de un incremento de un por ciento en el tiempo de viaje de dicha alternativa. En particular, se calculó la elasticidad punto desagregada directa (Bierlaire, 2018). Corresponde a la elasticidad punto, pues se considera un cambio infinitesimal en la variable; es desagregada, ya que se refiere al modelo de elección de un individuo específico; y es directa, porque mide el impacto que genera el cambio de un atributo en la alternativa i en la probabilidad de elección de dicha alternativa.

4.8. Conclusión

En este capítulo se presentó el marco teórico que gobierna las metodologías para solucionar el problema de la endogeneidad por atributos omitidos, junto con la modelación matemática

respectiva. Además, se presentan los análisis preliminares de las dimensiones de las variables psicosociales, para comenzar la estimación de los modelos.

Por otro lado, se presentan los criterios para comparar las metodologías mencionadas, a partir de indicadores prácticos que se obtienen de los resultados de las estimaciones: el valor subjetivo del tiempo de viaje y la elasticidad del tiempo de viaje y del costo. Finalmente, se mencionan los *software* utilizados para el manejo de la base de datos y la estimación de los modelos, en caso de que algún lector busque replicar o profundizar los modelos.

CAPÍTULO 5: ANÁLISIS DE RESULTADOS

5.1. Introducción

En este capítulo se presentan todos los resultados obtenidos que sean relevantes para cumplir con los objetivos estipulados, habiendo resultados que, por diversos motivos, se incluyen en los Anexos. De esta manera, se incluyen los valores de los coeficientes estimados, junto con las respectivas pruebas de significancia estadística. Se presentan cinco modelos, desagregados según las metodologías estipuladas en el Capítulo 4. Por último, se muestran los índices económicos utilizados para la comparación de modelos, y sus implicancias en la toma de decisiones en términos de políticas públicas.

Los coeficientes fueron interpretados según condiciones que deberán cumplir: tener sentido físico y económico, y ser estadísticamente significativos. Para lo último se usó como criterio el *test t de student*, que se obtuvo directamente de *Biogeme* y se comparará con valores estandarizados. Usualmente, se usa un 95% de confianza para este tipo de modelos, lo que implica que el valor crítico del *test t* es 1,96, dado el tamaño de muestra usada en las estimaciones. Basta con que el valor absoluto del *test t* de la variable obtenida sea mayor al valor crítico para que la variable sea estadísticamente significativa, y se deba incluir en el modelo. Se tiene, como complemento, el error estándar, que corresponde a la desviación estándar del coeficiente y se buscan valores reducidos, cercanos a 0. Como última medida de evaluación se tiene el valor-*p*, que indica la probabilidad de que el coeficiente calculado sea nulo estadísticamente, lo ideal es que esta probabilidad sea lo más cercano a 0 posible. En cualquier caso, debe ser menor a 0,05 por el nivel de confianza establecido.

Para unir ambas bases de datos, esto es, la correspondiente a la UdeC y la correspondiente a la UFRO, se debió verificar la homogeneidad de las componentes de error entre ambas bases de datos. De esta manera, se utilizó una variable auxiliar para comprobar que no existiesen diferencias entre los coeficientes por el mero hecho de pertenecer a universidades diferentes.

Además, se presentan las verosimilitudes obtenidas para cada modelación. Esto no necesariamente será revelador para comparar entre modelos, pues las funciones de verosimilitud incluyen argumentos distintos, pero sí para comparar modelos aditivos con modelos con interacción.

5.2. Modelo Básico

Este modelo no incluye ninguna metodología para corregir la endogeneidad. La nomenclatura para este modelo se indicó en el capítulo anterior e incluye el tiempo de viaje, el costo monetario asociado a la alternativa i , y las condiciones facilitadoras mencionadas previamente, incluyendo inicialmente en la modelación propuesta todas las condiciones disponibles. Luego, según la significancia estadística obtenida, se fue acotando el modelo. Las condiciones facilitadoras disponibles en primera instancia corresponden a las siguientes: usuario de sexo masculino, presencia de acompañantes en su viaje, presencia de equipaje pesado, si el viaje correspondía a un fin de semana, presencia de ciclovías, presencia de estacionamiento para bicicletas, usuario de nivel socioeconómico alto, posesión de licencia de conducir, posesión de licencia de conducir clase A, usuario estudiante, usuario entre 25 y 35 años de edad, viaje a una distancia menor a 2,5 km y usuario con habilidad para andar en bicicleta. En particular, la edad mencionada se podría considerar arbitraria, pero corresponde a una parte representativa de la muestra (69%).

Por otro lado, se probó una especificación donde se dividiera el costo por el ingreso, pues es una práctica común en economía para considerar la renta del individuo y relativizar el efecto del costo. Este enfoque arrojó resultados inconsistentes en términos de significancia estadística, pues algunas variables relevantes, incluso el mismo costo monetario, no resultaron ser significativas para la elección de modo de transporte, lo que no tiene sentido en la práctica. Además, a partir de la Prueba de Razón de Verosimilitud (Horowitz, 1983) se concluyó que el modelo propuesto es equivalente en términos de la información que entrega, por lo tanto, se optó por el enfoque de utilizar directamente el costo como variable.

Los resultados obtenidos para el modelo básico se presentan en la Tabla 5.1. La nomenclatura para el modo de transporte será bicicleta ($n=1$), automóvil ($n=2$) y transporte público ($n=3$), manteniéndola para todo el documento.

Tabla 5.1: Resultados para el modelo básico

Variable (Coeficiente)	Valor	Error std.	Test-t	Valor-p
Nivel de servicio				
Intercepto automóvil (θ_2^0)	4,88	0,68	7,18	0,00
Intercepto transporte público (θ_3^0)	6,63	0,65	10,30	0,00
Costo (θ_1)	-0,42	0,15	-2,71	0,01
Tiempo de viaje (θ_2)	-1,84	0,40	-4,56	0,00
Condiciones facilitadoras bicicleta				
Viaje con objetos livianos ($\mu_{1,1}$)	1,05	0,38	2,79	0,01
Viaje con acompañante ($\mu_{1,2}$)	-1,06	0,55	-1,97	0,05
Género ($\mu_{1,3}$)	1,15	0,27	4,28	0,00
Ciclismo fines de semana ($\mu_{1,4}$)	1,30	0,13	10,00	0,00
Condiciones facilitadoras automóvil				
Viaje con objetos livianos ($\mu_{2,1}$)	-0,48	0,24	-1,98	0,05
Viaje con acompañante ($\mu_{2,2}$)	2,37	0,29	8,08	0,00
Género ($\mu_{2,3}$)	0,47	0,22	2,12	0,03
Ingreso alto ($\mu_{2,4}$)	1,88	0,28	6,74	0,00
Licencia de conductor ($\mu_{2,5}$)	2,46	0,26	9,51	0,00
Estudiante ($\mu_{2,6}$)	-2,15	0,22	-9,66	0,00
Verosimilitud				
LR	-497,99			

Como se mencionó anteriormente, lo que importa en los modelos de partición modal es la diferencia entre las utilidades reportadas para diferentes alternativas, pues una utilidad sin alternativas para comparar no contribuye al análisis deseado, fijando algunas variables como puntos de comparación. Lo mencionado anteriormente aplica para las funciones de utilidad, donde se utilizó el transporte público como modo de referencia, mientras que, para los modelos de ecuaciones estructurales, la dimensión de evaluación quedó como variable fija del afecto, y la dimensión de comodidad como variable fija de la actitud.

Se tienen utilidades marginales negativas para el costo y el tiempo de viaje, lo que se condice con lo esperado, pues se espera que, mientras más caro o más se demore un modo de transporte, menos utilidad le reporte al usuario. También se obtuvieron valores relativamente altos para las constantes específicas. En general, esto no es deseable, porque podría deberse a que las variables que no fueron consideradas en la modelación estén siendo representadas como una constante que no entrega ningún tipo de información.

En términos de condiciones facilitadoras, o inhabilitadoras si se tiene signo negativo, se tiene que incentivan el viajar en bicicleta el viajar con objetos livianos, el ser un usuario masculino y el realizar ciclismo los fines de semana. En cambio, desincentiva el viaje en bicicleta el viajar con acompañante. Las magnitudes para las condiciones contextuales de la bicicleta son similares entre sí. Para el automóvil, el viajar con acompañante, el poseer licencia de conducir, el ser usuario masculino y de ingreso alto reportan utilidades positivas para los viajes en este modo, mientras que el viajar con objetos livianos y el ser estudiante reportan utilidades negativas. Ciertamente los resultados se condicen con lo que se observa frecuentemente en la práctica.

En particular, llama la atención que la presencia de ciclovías, la habilidad para andar en bicicleta, y la disponibilidad de estacionamientos para bicicletas no resultaron relevante para la utilidad de la bicicleta. Se espera que las condiciones mencionadas se puedan incluir en la modelación a partir de los modelos MIMIC mencionados en el Capítulo 4.

5.3. Modelo con el Método de los Indicadores Múltiples

En esta sección se presentan los resultados para el modelo corregido por endogeneidad utilizando la metodología de los indicadores múltiples. De esta manera, los coeficientes obtenidos son análogos al modelo básico, pero se añaden los coeficientes propios de los indicadores y de los errores de medición inducidos por los indicadores (o residuales de la regresión). Lo anterior se presenta en la Tabla 5.2.

Análisis de Resultados

Es importante destacar que se debe comprobar si existe heterogeneidad en la muestra, para saber si las potenciales diferencias observables son producto de endogeneidad o heterogeneidad. Esto se controla indirectamente en el método ICLV. Para ello, se estimó un *modelo logit mixto*, que entrega la media y la desviación estándar de los coeficientes de interés. Se probó heterogeneidad con las variables tiempo y costo de viaje, tanto por separado, como de manera conjunta. La heterogeneidad para el tiempo de viaje no resultó significativa, mientras que para el costo monetario se observaron resultados anómalos, es decir, utilidades marginales de costo positivas y desviaciones estándar negativas, lo que no tiene sentido físico. De esta manera, se atribuye que las diferencias en la muestra son producto de la endogeneidad presente, y no de la heterogeneidad estudiada.

Tabla 5.2: Resultados para el modelo corregido con el método MIS

Variable (Coeficiente)	Valor	Error std.	Test-t	Valor-p
Nivel de servicio				
Intercepto automóvil (θ_2^0)	11,40	3,93	2,90	0,00
Intercepto transporte público (θ_3^0)	13,40	3,15	4,25	0,00
Costo (θ_1)	-0,47	0,19	-2,45	0,00
Tiempo de viaje (θ_2)	-2,49	0,72	-3,45	0,00
Condiciones facilitadoras bicicleta				
Viaje con objetos livianos ($\mu_{1,1}$)	1,08	0,47	2,31	0,02
Viaje con acompañante ($\mu_{1,2}$)	-1,41	0,64	-2,22	0,03
Género ($\mu_{1,3}$)	0,86	0,35	2,48	0,01
Ciclismo fines de semana ($\mu_{1,4}$)	0,56	0,16	3,43	0,00
Condiciones facilitadoras automóvil				
Viaje con objetos livianos ($\mu_{2,1}$)	-0,42	0,27	-1,55	0,12
Viaje con acompañante ($\mu_{2,2}$)	2,41	0,34	7,17	0,00
Género ($\mu_{2,3}$)	0,58	0,26	2,22	0,03
Ingreso alto ($\mu_{2,4}$)	1,55	0,34	4,59	0,00
Licencia de conductor ($\mu_{2,5}$)	1,75	0,31	5,59	0,00
Estudiante ($\mu_{2,6}$)	-1,84	0,27	-6,79	0,00
Coeficientes indicadores método MIS				
Afecto bicicleta ($\omega_{1,1}$)	1,61	0,68	2,38	0,02
Afecto automóvil ($\omega_{1,2}$)	0,55	0,24	2,29	0,01
Afecto transporte público ($\omega_{1,3}$)	0,43	0,21	2,06	0,03
Actitud bicicleta ($\omega_{2,1}$)	0,28	0,06	4,58	0,00
Actitud automóvil ($\omega_{2,2}$)	0,14	0,06	2,51	0,01
Actitud transporte público ($\omega_{2,3}$)	0,18	0,04	4,48	0,00

Análisis de Resultados

Variable (Coeficiente)	Valor	Error std.	Test-t	Valor-p
Coeficientes residuales método MIS				
Afecto bicicleta	-1,35	0,67	-2,02	0,04
Afecto automóvil	-0,30	0,05	-5,63	0,00
Afecto transporte público	-0,24	0,06	-3,83	0,00
Actitud bicicleta	-0,19	0,06	-3,08	0,00
Actitud automóvil	-0,12	0,05	-2,52	0,01
Actitud transporte público	-0,47	0,21	-2,22	0,01
Coeficientes hábito				
Hábito bicicleta (κ_1)	3,17	0,83	3,80	0,00
Hábito automóvil (κ_2)	2,35	0,50	4,70	0,00
Hábito transporte público (κ_3)	0,88	0,41	2,17	0,02
Verosimilitud				
LR	-386,35			

Lo primero que se puede observar es el aumento de magnitud de las constantes específicas, si bien esto podría implicar una peor representación del fenómeno, se debe recordar que este método induce un cambio de escala en la utilidad. Precisamente por esta razón es que se analizan tasas marginales de sustitución para la comparación entre modelos. Por lo tanto, se debe realizar una indagación completa antes de precipitarse a generar conclusiones.

Para las condiciones contextuales, se mantienen las relaciones de incentivo o desincentivo para los modos estudiados, en términos de signos. En general, estas variables cambian ligeramente su impacto en las utilidades de los usuarios, a excepción del género y la realización de ciclismo los fines de semana para la bicicleta, y de la posesión de licencia de conducir y el ser estudiante para el automóvil.

Se tiene que el modelo básico efectivamente sufría de endogeneidad producto de la omisión de variables psicosociales, pues todos los coeficientes introducidos en la Sección 4.3 resultaron relevantes para explicar el fenómeno. Esto significa que los resultados actualizados de las utilidades marginales deberían ser más consistentes que el modelo anterior.

La significancia estadística de los coeficientes de los indicadores implica que la variable medida con dicho indicador será relevante para explicar el fenómeno; por lo tanto, se tenía originalmente endogeneidad producto de la omisión de dicha variable. En cambio, la significancia estadística de los coeficientes asociados a los residuales de la regresión quiere decir que, producto de la inclusión directa de los indicadores en la función de utilidad, se indujeron errores de medición asociados, que se corrigen a partir de la regresión. En términos de signos, se tiene que el afecto y la actitud tienen un impacto positivo la utilidad de viajar en todos los modos disponibles, siendo notoriamente el afecto hacia la bicicleta la variable de mayor magnitud. Con respecto a los residuales, no se tiene una interpretación práctica, sino que expresan una corrección para la función de utilidad, que, por lo demás, fueron negativos en todos los casos.

Por otro lado, se incluye la variable psicosocial del hábito, pero sin ninguna corrección, pues no se utilizan indicadores para su medición, sino que se mide directamente. El hábito para todos los modos de transporte resultó en una utilidad marginal positiva y de mayor magnitud que el afecto y la actitud. Se tiene que el impacto del hábito de viajar bicicleta es mayor que el de viajar en automóvil y mucho mayor que el de viajar en transporte público.

Si bien el lector podría considerar que, como la inclusión de las nuevas variables tiene menos peso marginal que las condiciones contextuales, podría ser más beneficioso considerar nuevas o distintas condiciones contextuales para mejorar el poder explicativo del modelo; se debe recordar que la mayoría de estas variables son *booleanas*. De esta manera, se multiplicarían por 0 o 1, mientras que las variables psicosociales se multiplican por un indicador entero (entre 1 y 5, o 1 y 7). Luego, el peso efectivo de la variable psicosocial podría ser mayor que el de las potenciales nuevas condiciones contextuales.

Lo mismo sucede para las variables psicosociales hacia el automóvil. En el Capítulo 3 se mencionó que este modo poseía las mayores valoraciones por parte de los usuarios, por lo tanto, se podría esperar que las utilidades marginales fueran mayores que las de los otros modos. Sin embargo, es precisamente por la alta valoración que posee este modo, que la magnitud de estos coeficientes no se elevó particularmente, porque el peso efectivo de la

variable es su ponderación por el coeficiente respectivo. Con el objetivo de complementar las justificaciones anteriores, se presenta un análisis de los pesos efectivos de las variables en la función de utilidad, que se encuentra en la Sección 5.7.

Es importante mencionar que, en el caso del afecto, se encontraron anomalías en las dimensiones de potencia y de activación, pues algunos pares de antónimos contribuían de manera negativa para un modo y de manera positiva para otro. Esto sucedió, por ejemplo, con el par Pasivo-Activo y Grande-Pequeño. Por lo tanto, se optó por modelar el afecto en función de la dimensión de la Evaluación y del Control. Para la dimensión de Evaluación, las cargas de los antónimos Incómodo-Cómodo y Sucio-Limpio, no cumplieron con el criterio de carga para considerarse indicadores válidos. Por otro lado, para la dimensión de Control todos los pares de antónimos resultaron ser válidos como indicadores. Lo anterior aplica para todos los modos de transporte disponibles. Además, como se mencionó en el Capítulo 3, al no considerar un indicador se podría ver beneficiada la consistencia entre las respuestas. Esto se cumplió para los indicadores mencionados. En cambio, para la actitud de la bicicleta, el automóvil y el transporte público, todas las dimensiones resultaron relevantes. Por ende, todas se mantuvieron en la modelación. En el Anexo 8.2 se presentan los resultados del análisis factorial para las dimensiones del afecto y de la actitud, desagregadas por modo.

5.4. Método de Variables Latentes Integrado

En esta sección se presentan los resultados para el modelo corregido por endogeneidad utilizando el método ICLV. Para facilitar la comprensión y legibilidad, se mostrarán los resultados del modelo de elección discreta con los coeficientes obtenidos a partir de la metodología mencionada en la Tabla 5.3. Los resultados de los modelos de ecuaciones estructurales, es decir, el peso de cada indicador sobre la variable latente, y los resultados de los modelos MIMIC, esto es, el peso de las variables sociodemográficas del usuario sobre la variable latente, se presentan en las Tablas 5.4 y 5.5, respectivamente. Cabe destacar que se realizó una estimación secuencial antes de realizar la estimación simultánea, con la finalidad de obtener valores iniciales de iteración y optimizar el tiempo de cálculo en el programa.

La magnitud de las constantes específicas disminuyó considerablemente con esta especificación del modelo, incluso la probabilidad de que el intercepto del automóvil sea 0 aumentó de un 0% a un 80%. En cuanto a las condiciones contextuales, dejaron de ser significativos el ciclismo los fines de semana para facilitar los viajes en bicicleta y el género, el viaje con objetos livianos y el ser usuario de ingreso auto para el automóvil. Esto se podría explicar por la inclusión de las nuevas variables modeladas en los modelos MIMIC, que reflejan el peso de la variable sociodemográfica sobre la variable psicosocial, en vez de en la utilidad del modo de transporte.

Con respecto a las variables psicosociales, tanto las calculadas con los SEM como el hábito (que se incluye de manera análoga al método MIS), todas reportan utilidades positivas, como es esperado. Eso implica que por cada unidad que se tenga de una variable psicosocial, se aumentará la utilidad del usuario en viajar en el modo respectivo, lo que tiene sentido intuitivo. Para este caso, independiente del modo de transporte, el hábito pesa más marginalmente que el afecto. Asimismo, el afecto tiene un mayor efecto marginal que la actitud.

Es importante mencionar que, en la etapa de estadística descriptiva, las valoraciones del automóvil fueron más altas que las de la bicicleta y mucho más altas que las del transporte público. Esto podría suponer que los coeficientes del automóvil deberían tener mayor magnitud, pero se debe recordar que el peso efectivo de la variable es el producto de la variable (o valoración) por el coeficiente. Por lo tanto, se tendrá un peso mayor en la utilidad para los modos mejor valorados.

Tabla 5.3: Resultados para el modelo corregido con el método ICLV

Variable (Coeficiente)	Valor	Error std.	Test-t	Valor-p
Nivel de servicio				
Intercepto automóvil (θ_2^0)	-0,38	1,53	-0,25	0,80
Intercepto transporte público (θ_3^0)	3,68	1,19	3,09	0,00
Costo (θ_1)	-0,44	0,13	-3,46	0,02
Tiempo de viaje (θ_2)	-2,51	0,68	-3,71	0,00
Condiciones facilitadoras bicicleta				
Viaje con objetos livianos ($\mu_{1,1}$)	1,03	0,48	2,16	0,03

Análisis de Resultados

Variable (Coeficiente)	Valor	Error std.	Test-t	Valor-p
Viaje con acompañante ($\mu_{1,2}$)	-1,74	0,71	-2,44	0,01
Género ($\mu_{1,3}$)	0,81	0,36	2,25	0,02
Condiciones facilitadoras automóvil				
Viaje con acompañante ($\mu_{2,2}$)	2,64	0,72	3,67	0,00
Licencia de conductor ($\mu_{2,5}$)	1,83	0,92	1,99	0,04
Estudiante ($\mu_{2,6}$)	-1,62	0,32	-5,02	0,00
Coeficientes método ICLV				
Afecto bicicleta	2,11	0,92	2,28	0,02
Afecto automóvil	1,65	0,31	5,32	0,00
Afecto transporte público	0,59	0,29	2,06	0,03
Actitud bicicleta	0,36	0,13	2,72	0,01
Actitud automóvil	0,20	0,08	2,41	0,02
Actitud transporte público	0,19	0,08	2,23	0,03
Coeficientes hábito				
Hábito bicicleta	2,76	0,91	3,02	0,00
Hábito automóvil	2,33	0,62	3,79	0,00
Hábito transporte público	1,11	0,54	2,07	0,04
Verosimilitud				
LR	-34092,12			

En la Tabla 5.4 se muestran los resultados de los modelos de ecuaciones estructurales para la estimación de las variables latentes, que corresponden a la actitud y al afecto. Como se mencionó anteriormente, las dimensiones utilizadas para medir el afecto fueron la evaluación (fijada) y el control, mientras que para la actitud se utilizaron todas las dimensiones disponibles. Estos resultados reflejan claramente la ventaja que se tiene por el uso del método ICLV, pues se tiene información detallada del peso efectivo de cada una de las dimensiones sobre la variable psicosocial. Se debe recordar que los valores establecidos para las dimensiones de evaluación y comodidad fueron un valor nulo en intercepto y una pendiente unitaria en carga. Para el afecto, se tiene que el control tiene un peso mayor que la evaluación, tanto en carga como en intercepto. Por otro lado, se tiene una baja actitud con respecto al ambientalismo asociado al viajar en automóvil, y alta con respecto a la bicicleta. Los factores económicos mostraron una magnitud relativamente alta para la bicicleta y para el transporte público, lo que se puede deber a que viajar en bicicleta es prácticamente gratis y la tarifa máxima del bus no supera los \$600 CLP. La dimensión de salud mostró mayor

peso en la bicicleta que en los demás modos, lo que se podría atribuir al esfuerzo físico que conlleva viajar en dicho modo. Lo mismo sucedió con la dimensión de diversión, donde se podría hipotetizar que dicha actividad física resulta agradable para la muestra. En cuanto a la seguridad, el automóvil y el transporte público muestran valores similares entre sí, mientras que la bicicleta reporta valores bastante más bajos. Esto se podría explicar por la baja presencia de infraestructura adecuada para viajar en bicicleta hacia las universidades estudiadas. Por último, las cargas asociadas a todas las dimensiones, para ambas variables, resultaron positivas, lo que tiene sentido al momento de la interpretación práctica: por cada unidad de las dimensiones utilizadas aumenta el afecto o la actitud hacia el modo de transporte.

Tabla 5.4: Resultados para los modelos de ecuaciones estructurales

Variable (Coeficiente)	Valor	Error std.	Test-t	Valor-p
Afecto				
Intercepto control automóvil	1,13	0,56	2,02	0,04
Intercepto control bicicleta	0,44	0,17	2,52	0,02
Intercepto control bus	2,54	0,81	3,12	0,00
Carga control automóvil	0,98	0,46	2,12	0,03
Carga control bicicleta	1,99	0,68	2,93	0,00
Carga control bus	1,65	0,36	4,63	0,00
Actitud				
Intercepto ambientalismo automóvil	-7,08	0,50	-14,13	0,00
Intercepto diversión automóvil	-3,80	0,30	-12,50	0,00
Intercepto economía automóvil	-5,12	0,52	-9,88	0,00
Intercepto salud automóvil	-4,88	0,45	-10,80	0,00
Intercepto seguridad automóvil	0,73	0,34	2,15	0,03
Intercepto ambientalismo bicicleta	4,16	0,26	15,82	0,00
Intercepto diversión bicicleta	-0,57	0,12	-4,79	0,00
Intercepto economía bicicleta	5,75	0,37	15,37	0,00
Intercepto salud bicicleta	3,49	0,21	17,02	0,00
Intercepto seguridad bicicleta	-1,43	0,16	-9,05	0,00
Intercepto ambientalismo bus	-0,68	0,27	-2,58	0,01
Intercepto diversión bus	-3,02	0,39	-7,70	0,00
Intercepto economía bus	2,71	0,27	10,15	0,00
Intercepto salud bus	-1,03	0,26	-3,95	0,00
Intercepto seguridad bus	1,35	0,16	8,33	0,00
Carga ambientalismo automóvil	1,12	0,12	9,11	0,00

Análisis de Resultados

Variable (Coeficiente)	Valor	Error std.	Test-t	Valor-p
Carga diversión automóvil	0,60	0,07	9,16	0,00
Carga economía automóvil	0,92	0,13	7,33	0,00
Carga salud automóvil	1,27	0,11	11,76	0,00
Carga seguridad automóvil	0,82	0,08	10,26	0,00
Carga ambientalismo bicicleta	0,60	0,09	6,41	0,00
Carga diversión bicicleta	0,84	0,05	15,68	0,00
Carga economía bicicleta	0,59	0,11	5,51	0,00
Carga salud bicicleta	0,77	0,12	6,29	0,00
Carga seguridad bicicleta	0,65	0,05	12,31	0,00
Carga ambientalismo bus	0,77	0,18	4,18	0,00
Carga diversión bus	0,62	0,06	11,14	0,00
Carga economía bus	0,66	0,12	5,47	0,00
Carga salud bus	1,00	0,15	6,71	0,00
Carga seguridad bus	0,81	0,09	9,19	0,00

En la Tabla 5.5 se muestran los resultados de los modelos MIMIC que resultaron significativos para definir las variables psicosociales estudiadas. Estos resultados constituyen otra de las ventajas de utilizar los modelos ICLV, dado que permiten cuantificar el impacto de las características de los usuarios sobre las variables latentes, lo que puede ser útil para realizar predicciones o estudios de políticas públicas.

Para el afecto hacia la bicicleta, aportan positivamente el realizar ciclismo los fines de semana, el que el viaje sea de una distancia inferior a 2,5 km y que el usuario considere que tiene habilidad para andar en bicicleta, mientras que la posesión de licencia de conducir aporta negativamente. Con respecto al afecto hacia el automóvil, el realizar ciclismo los fines de semana y que el viaje sea de una distancia mayor a 2,5 km presentan signos negativos, y el ser un usuario de ingreso alto y poseer licencia de conductor presentan signos positivos. El afecto hacia el transporte público se ve aumentado únicamente por la ocupación del usuario, pero se ve disminuido si se pertenece a un nivel socioeconómico alto y se tiene licencia de conductor.

La actitud hacia la bicicleta mostró la mayor cantidad de variables explicativas, donde casi todas aportan de manera positiva: el viaje con objetos livianos, el realizar ciclismo los fines

de semana, la presencia de ciclovías, el ser de rango etario entre 25 y 35 años, el ser usuario masculino, el que la distancia del viaje sea inferior a 2,5 km y la habilidad para andar en bicicleta. Se debe tener en cuenta que el intercepto tiene una alta magnitud negativa, por lo que no necesariamente la actitud hacia la bicicleta adoptará valores tan altos. La única variable que mostró aportar negativamente fue la posesión de licencia de conducir. Para la actitud hacia el automóvil se tiene un coeficiente base positivo y de alta magnitud con respecto a las demás variables. Además, la posesión de licencia de conductor contribuye hacia esta variable. Sin embargo, aportan negativamente el rango etario entre 25 y 35 años, la habilidad para andar en bicicleta y la ocupación de estudiante. Por último, la actitud hacia el transporte público se ve condicionada positivamente por el viajar con objetos livianos y la ocupación de estudiante, y negativamente por el ser un usuario de ingreso alto, el poseer licencia de conducir, y el ser un usuario con habilidad para andar en bicicleta.

Tabla 5.5: Resultados para modelos MIMIC

Variable (Coeficiente)	Valor	Error std.	Test-t	Valor-p
Afecto bicicleta				
Intercepto	-1,64	0,53	-3,11	0,00
Ciclismo los fines de semana	0,18	0,07	2,49	0,01
Licencia de conducir	-0,24	0,09	-2,67	0,01
Distancia inferior a 2,5 km	0,36	0,17	2,16	0,03
Habilidad para andar en bicicleta	0,39	0,12	3,25	0,00
Afecto automóvil				
Intercepto	2,41	1,04	2,32	0,02
Ciclismo los fines de semana	-0,37	0,15	-2,55	0,01
Ingreso alto	0,61	0,25	2,49	0,03
Licencia de conducir	0,80	0,31	2,60	0,01
Distancia inferior a 2,5 km	-0,76	0,34	-2,24	0,03
Afecto bus				
Intercepto	-2,70	0,38	-7,16	0,00
Ingreso alto	-0,65	0,25	-2,59	0,02
Licencia de conducir	-0,46	0,16	-2,97	0,00
Estudiante	0,59	0,22	2,71	0,01
Actitud bicicleta				
Intercepto	-5,93	1,22	-4,86	0,00
Viaje con objetos livianos	0,59	0,21	2,80	0,01
Ciclismo los fines de semana	0,68	0,14	4,84	0,00

Análisis de Resultados

Variable (Coeficiente)	Valor	Error std.	Test-t	Valor-p
Presencia de ciclovías	1,89	0,48	3,93	0,00
Edad entre 25 y 35 años	0,60	0,27	2,23	0,03
Licencia de conducir	-0,70	0,25	-2,76	0,01
Género	0,83	0,24	3,47	0,00
Distancia inferior a 2,5 km	1,48	0,52	2,86	0,00
Habilidad para andar en bicicleta	0,73	0,15	4,83	0,00
Actitud automóvil				
Intercepto	6,18	0,94	6,55	0,00
Edad entre 25 y 35 años	-0,81	0,21	-3,79	0,00
Licencia de conducir	0,56	0,24	2,36	0,01
Habilidad para andar en bicicleta	-0,52	0,21	-2,43	0,01
Estudiante	-0,74	0,13	-5,66	0,00
Actitud bus				
Intercepto	0,31	0,69	0,45	0,96
Viaje con objetos livianos	0,99	0,32	3,15	0,00
Ingreso alto	-1,11	0,39	-2,88	0,00
Licencia de conducir	-0,65	0,30	-2,17	0,02
Habilidad para andar en bicicleta	-0,50	0,24	-2,07	0,04
Estudiante	1,44	0,63	2,27	0,04

Si bien, a partir de la cantidad y calidad de resultados obtenidos con esta metodología, se presentan diversas ventajas, este tipo de modelos no está exento de críticas. Por ejemplo, Chorus y Kroensen (2014) plantean que los indicadores de las variables latentes son medidos, casi sin excepción, en un momento particular en un tiempo determinado. En general, la evidencia es aún insuficiente para usar percepciones o indicadores como bases para apoyar políticas de demanda de viajes que apunten a modificar los comportamientos de los viajes, únicamente a partir de cambios en los niveles de la variable latente. También postulan que las valoraciones de las variables latentes pueden tener una gran correlación con la alternativa escogida y que no se tienen maneras de cuantificar los cambios de valoraciones para un mismo usuario, con el paso del tiempo. Se reconoce que se tiene mucho por mejorar aún en todo tipo de modelación, pero sin duda se provee un marco teórico robusto para idear nuevas soluciones y campos de investigación.

5.5. Modelos con Interacción

En la presente sección se presentan los resultados asociados a los modelos con interacción introducidos en el Capítulo 4, para el método MIS y el método ICLV, en las Tablas 5.6 y 5.7, respectivamente. Para los SEM y los MIMIC del modelo ICLV con interacción se tuvieron resultados análogos en cuanto a las variables que resultaron explicativas y sus signos. Sin embargo, las magnitudes cambian. Lo anterior se reporta en los Anexos 8.3 y 8.4 para facilitar la comprensión del documento.

Se tienen condiciones similares a los modelos anteriores para las interpretaciones de las constantes específicas, de las condiciones facilitadoras y de la utilidad marginal del costo, para ambas metodologías. Además, en ambos casos se puede observar que la magnitud del coeficiente del tiempo de viaje aumenta considerablemente en comparación con los modelos sin interacción. Sin embargo, se debe considerar el efecto del aporte de las variables psicosociales, que generan una especie de “ajuste” para este coeficiente.

El signo de todas las utilidades marginales de las variables psicosociales interactuando con el tiempo de viaje resultó positivo, lo que disminuiría el peso efectivo del tiempo de viaje en las decisiones de los usuarios. Lo anterior quiere decir que, mientras mayor sea la valoración de la variable en cuestión, menor desutilidad reporta el tiempo que toma viajar. Por ejemplo, mientras mayor afecto se tiene hacia la bicicleta, menos relevante será el tiempo de viaje en la decisión de los usuarios.

Otro aspecto relevante por mencionar, que aplicó para ambas metodologías, es que se intentó la inclusión de las variables psicosociales de manera aditiva en conjunto con las variables que interactúan con el tiempo de viaje, pero no resultaron ser significativas estadísticamente. Se podría considerar, de esta manera, que la relevancia de las variables interactivas es superior a la de las variables aditivas. Además, en ambos casos se tienen aumentos en términos de verosimilitud. Por lo tanto, se postula que la representación del fenómeno es más adecuada.

Tabla 5.6: Resultados para el modelo corregido con el método MIS con interacción

Variable	Valor	Error std.	Test-t	Valor-p
Nivel de servicio				
Intercepto automóvil	2,35	1,09	2,16	0,03
Intercepto transporte público	4,36	0,81	5,40	0,00
Costo	-0,44	0,21	-2,11	0,03
Tiempo de viaje	-6,55	2,15	-3,05	0,00
Condiciones facilitadoras bicicleta				
Viaje con objetos livianos	1,07	0,43	2,47	0,01
Viaje con acompañante	-1,46	0,66	-2,23	0,03
Género	1,18	0,32	3,65	0,00
Ciclismo fines de semana	1,23	0,16	7,94	0,00
Condiciones facilitadoras automóvil				
Viaje con objetos livianos	-0,63	0,26	-2,40	0,02
Viaje con acompañante	2,40	0,32	7,45	0,00
Género	0,63	0,24	2,59	0,02
Ingreso alto	1,78	0,29	6,10	0,00
Licencia de conductor	1,97	0,29	6,75	0,00
Estudiante	-1,93	0,25	-7,60	0,00
Coefficientes método MIS interactuando con tiempo de viaje				
Afecto bicicleta	1,66	0,89	1,86	0,00
Afecto automóvil	1,06	0,55	1,93	0,03
Afecto transporte público	0,94	0,48	1,97	0,05
Actitud bicicleta	0,52	0,18	2,97	0,03
Actitud automóvil	0,28	0,11	2,55	0,02
Actitud transporte público	0,19	0,07	2,83	0,00
Coefficientes hábito interactuando con tiempo de viaje				
Hábito bicicleta	4,30	0,71	6,06	0,00
Hábito automóvil	0,75	0,16	4,69	0,00
Hábito transporte público	2,08	0,27	7,70	0,00
Coefficientes residuales método MIS				
Afecto bicicleta	0,55	0,14	4,01	0,00
Afecto automóvil	0,09	0,04	2,16	0,01
Afecto transporte público	-0,41	0,19	-2,13	0,03
Actitud bicicleta	0,21	0,04	5,95	0,01
Actitud automóvil	0,09	0,04	2,38	0,01
Actitud transporte público	0,07	0,03	1,94	0,06
Verosimilitud				
LR	-377,06			

Análisis de Resultados

Para la metodología MIS se tiene que el hábito tiene un mayor peso que el afecto, excepto para el automóvil, y que el afecto tiene un mayor peso que la actitud sobre el tiempo de viaje, para todas las alternativas. Por otro lado, para la metodología ICLV, el hábito corresponde al coeficiente con mayor magnitud para todas las alternativas, y la actitud pesó más que el afecto, para todos los modos disponibles. Es importante recordar que no se pueden realizar comparaciones directas entre las metodologías en términos de magnitudes de coeficientes, la comparación adecuada se presenta en la Sección 5.5.

Tabla 5.7: Resultados para el modelo corregido con el método ICLV con interacción

Variable	Valor	Error std.	Test-t	Valor-p
Nivel de servicio				
Intercepto automóvil	1,01	0,99	1,02	0,31
Intercepto transporte público	1,77	1,01	1,75	0,08
Costo	-0,65	0,13	-5,00	0,00
Tiempo de viaje	-11,30	0,91	-12,42	0,00
Condiciones facilitadoras bicicleta				
Viaje con objetos livianos	1,08	0,49	2,22	0,03
Viaje con acompañante	-1,62	0,74	-2,20	0,03
Género	1,07	0,34	3,18	0,00
Condiciones facilitadoras automóvil				
Viaje con acompañante	2,46	0,41	5,97	0,00
Licencia de conductor	2,05	0,37	5,56	0,00
Estudiante	-2,32	0,59	-3,96	0,00
Coefficientes método ICLV interactuando con tiempo de viaje				
Afecto bicicleta	1,32	0,20	6,60	0,00
Afecto automóvil	0,71	0,26	2,73	0,01
Afecto transporte público	1,22	0,74	1,65	0,00
Actitud bicicleta	2,85	1,12	2,54	0,02
Actitud automóvil	1,10	0,49	2,24	0,03
Actitud transporte público	0,45	0,18	2,56	0,00
Coefficientes hábito interactuando con tiempo de viaje				
Hábito bicicleta	7,51	0,48	15,65	0,00
Hábito automóvil	1,20	0,32	3,75	0,00
Hábito transporte público	3,65	0,25	14,60	0,00
Verosimilitud				
LR	-34039,86			

Para el caso del ICLV, se podría pensar que el aumento en verosimilitud es marginal, pero se debe considerar que la función de verosimilitud de este modelo incluye la verosimilitud de los modelos de ecuaciones estructurales, que en este caso corresponden a 24 regresiones lineales, y aportan la mayor parte de la verosimilitud.

5.6. Peso Relativo de Variables en la Utilidad

Con el objetivo de conocer el peso que tendría cada variable en cada modelo propuesto, se realizó un análisis cuantitativo. Como se tiene cada variable es única para cada usuario, pero se tiene un coeficiente que es constante para toda la muestra, se utilizó el promedio entre las variables de la muestra y se ponderó por el coeficiente respectivo. Así, se pudieron obtener las contribuciones porcentuales de las variables modeladas en las utilidades estudiadas. En la Tabla 5.8 se presentan estos resultados y se utilizó la siguiente nomenclatura para optimizar espacio en el documento: 1 para el modelo básico, 2 para el modelo MIS aditivo, 3 para el modelo MIS con interacción, 4 para el modelo ICLV aditivo, 5 para el modelo ICLV con interacción, *B* para la bicicleta, *A* para el automóvil y *TP* para el transporte público.

Tabla 5.8: Porcentajes de aporte de las variables en funciones de utilidad

Variable/Función	U _{1B}	U _{1A}	U _{1TP}	U _{2B}	U _{2A}	U _{2TP}	U _{3B}	U _{3A}	U _{3TP}	U _{4B}	U _{4A}	U _{4TP}	U _{5B}	U _{5A}	U _{5TP}
ASC	-	21%	25%	-	16%	18%	-	12%	14%	-	5%	12%	-	14%	21%
Tiempo de viaje	27%	26%	46%	9%	16%	34%	16%	10%	28%	33%	15%	25%	68%	26%	45%
Costo	-	17%	29%	-	17%	11%	-	13%	18%	-	10%	10%	-	6%	13%
Viaje con objetos livianos	11%	1%	-	1%	0%	-	2%	1%	-	10%	-	-	5%	-	-
Viaje acompañado	12%	7%	-	2%	3%	-	3%	6%	-	18%	8%	-	8%	9%	-
Género	24%	3%	-	1%	1%	-	4%	3%	-	15%	-	-	9%	-	-
Ciclismo los fines de semana	26%	-	-	1%	-	-	4%	-	-	-	-	-	-	-	-
Ingreso alto	-	4%	-	-	1%	-	-	3%	-	-	-	-	-	-	-
Licencia de conductor	-	6%	-	-	3%	-	-	8%	-	-	9%	-	-	13%	-
Estudiante	-	15%	-	-	5%	-	-	11%	-	-	12%	-	-	19%	-
Afecto	-	-	-	31%	11%	6%	-	-	-	6%	19%	23%	-	-	-
Actitud	-	-	-	17%	11%	10%	-	-	-	4%	9%	22%	-	-	-
Hábito	-	-	-	2%	4%	1%	-	-	-	13%	13%	8%	-	-	-
Residuales afecto	-	-	-	27%	6%	4%	18%	5%	13%	-	-	-	-	-	-
Residuales actitud	-	-	-	10%	5%	17%	18%	9%	5%	-	-	-	-	-	-
Interacción afecto	-	-	-	-	-	-	17%	8%	11%	-	-	-	1%	3%	16%
Interacción actitud	-	-	-	-	-	-	17%	9%	8%	-	-	-	5%	9%	1%
Interacción hábito	-	-	-	-	-	-	1%	1%	2%	-	-	-	5%	1%	5%

De lo anterior se puede desprender que, en general, las variables más significativas en la toma de decisiones son el tiempo de viaje y el costo monetario. En cuanto a las constantes específicas de la alternativa, se obtuvieron valores altos, lo que significa que gran parte del fenómeno se está representando con una constante, lo que no es deseable, pues no permite generar ningún tipo de conclusión. Esto se acentúa en las funciones de utilidad del transporte público, pues no se incluyeron condiciones facilitadoras en el análisis de ese modo, como se justificó anteriormente. Las condiciones facilitadoras tienen un gran impacto en el modelo básico, pero cuando se consideran las variables psicosociales, estas empiezan a perder relevancia. Más aún, con la inclusión de variables psicosociales algunas condiciones facilitadoras dejan de ser significativas para un nivel de confianza de 95%.

El peso de las variables psicosociales depende del modo de transporte y de su inclusión matemática (aditiva o interactiva). Generalmente, el afecto fue la variable que más aportó en la toma de decisiones, y se notó que la inclusión del hábito tuvo mayor impacto en el modelo

ICLV. Los residuales asociados a la metodología MIS no tienen interpretación física, como se explicó anteriormente, sin embargo, mostraron tener un gran peso en las utilidades, lo que se interpreta como que existe un gran desfase entre un modelo que no considera variables psicosociales y un modelo que sí las considera. Por último, en el modelo MIS con interacción se notaron aportes de mayor magnitud en las variables interactivas que para la metodología ICLV con interacción.

Naturalmente, este análisis es general, pues se basa en promedios, tanto para las utilidades como para la población y sus variables respectivas, lo que constituye un supuesto potente. Sin embargo, se presentan observaciones que contribuyen a la interpretación de los resultados.

5.7. Valor Subjetivo del Tiempo y Elasticidades

En esta sección se presentan los resultados para los índices económicos utilizados para comparar las metodologías propuestas: el valor subjetivo del tiempo de viaje (VST) y las elasticidades para el tiempo de viaje y del costo monetario de viajar. En la Tabla 5.9 se revelan los intervalos de confianza para el VST comparados con el modelo que no corrige endogeneidad. En ésta se puede apreciar que la metodología MIS ofrece un intervalo de confianza más grande que el modelo básico y que para el método ICLV este intervalo se reduce. Sin embargo, tanto el promedio como la mediana aumentan con ambos enfoques. Esto tiene una interpretación directa: las variables psicosociales del individuo tienen incidencia en la elección de modo de los usuarios y producen cambios sustanciales en la valoración del tiempo de viaje en términos monetarios. De esta manera, se concluye que, para aproximarse a la representación más sofisticada del suceso observado (ICLV), el MIS es una herramienta válida y ofrece mejores resultados que no incluir los atributos omitidos en la modelación.

Tabla 5.9: Intervalos de confianza para el VST en métodos MIS y ICLV

Valor del tiempo (CLP/min)	Base	MIS	ICLV
Promedio	44,3	53,0	57,4
Límite inferior	24,2	22,7	29,3
Mediana	26,3	37,6	53,4
Límite superior	138,7	153,3	119,8

Tabla 5.10: VST desagregado por modo para modelos con interacción con el tiempo de viaje

Valor del tiempo (CLP/min)	MIS interacción	ICLV interacción
Bicicleta	33,8	27,2
Automóvil	7,2	8,5
Transporte público	34,5	39,4

Por otro lado, en la Tabla 5.10 se presentan la valoración del tiempo de viaje considerando los efectos de interacción con variables psicosociales para la muestra. El análisis de la influencia de cada variable se discutió en la Sección 5.4. A partir de la interacción del hábito, del afecto y de la actitud hacia las alternativas estudiadas se puede ver una disminución en la valoración del tiempo de viaje, en comparación con el modelo básico (44,4 CLP/min). Para la bicicleta y el transporte público se tuvieron disminuciones similares de alrededor de un 23%, mientras que para el automóvil una disminución de gran magnitud (83,7%). A partir de esto, se mencionó que la interacción de las variables psicosociales con el tiempo de viaje tiene un poder explicativo mayor, aunque podría ser marginalmente mayor, en términos de verosimilitud. Pero, más importante aún, se puede concluir que se tienen valores muy distintos para la valorización del tiempo de viaje si se consideran estos efectos.

En las Figuras 5.1, 5.2, 5.3, 5.4 y 5.5 se muestran los diagramas de caja de las elasticidades del tiempo de viaje y del costo de viaje. Todos los valores obtenidos fueron valores entre 0 y -1, y representan la probabilidad de dejar de elegir ese modo si se aumenta un porcentaje de la variable respectiva. Como se tiene que las utilidades marginales para obtener las elasticidades son derivadas de la función de utilidad, no se presenta la elasticidad del costo de viaje para la bicicleta, pues no se incluye la variable del costo en su función de utilidad. En conjunto, se muestran los cinco modelos estimados: el modelo básico, el MIS, el ICLV y

estos dos últimos considerando efectos de interacción (representados en la gráfica con el sufijo “Inter”).

En el transporte público las personas son más sensibles a cambios. Por otro lado, el automóvil resulta ser el modo donde los usuarios muestran mayor resistencia a cambiar de modo de transporte, tanto para el costo como el tiempo de viaje. La bicicleta muestra resultados intermedios entre ambos modos, mostrando una mayor disposición a cambiar de modo si se eleva en un 1% el tiempo de viaje que el automóvil, pero menor que el bus.

Nuevamente se puede apreciar que los mayores cambios para este índice se producen con el modelo ICLV, el cual debería arrojar los resultados más próximos a la realidad, debido a que es un modelo más sofisticado. En cuanto al modelo MIS, se pueden observar cambios significativos en comparación con el modelo base, tanto en los límites del intervalo de confianza, como en la mediana de los resultados. Más aún, para todos los resultados, se tiene que la incorporación de variables psicosociales con la metodología MIS ofrece resultados consistentes en cuanto al impacto en los índices económicos estudiados, en comparación con la metodología ICLV. Esto se puede apreciar en los gráficos como una suerte de “escalera”, siendo la metodología MIS el peldaño de al medio.

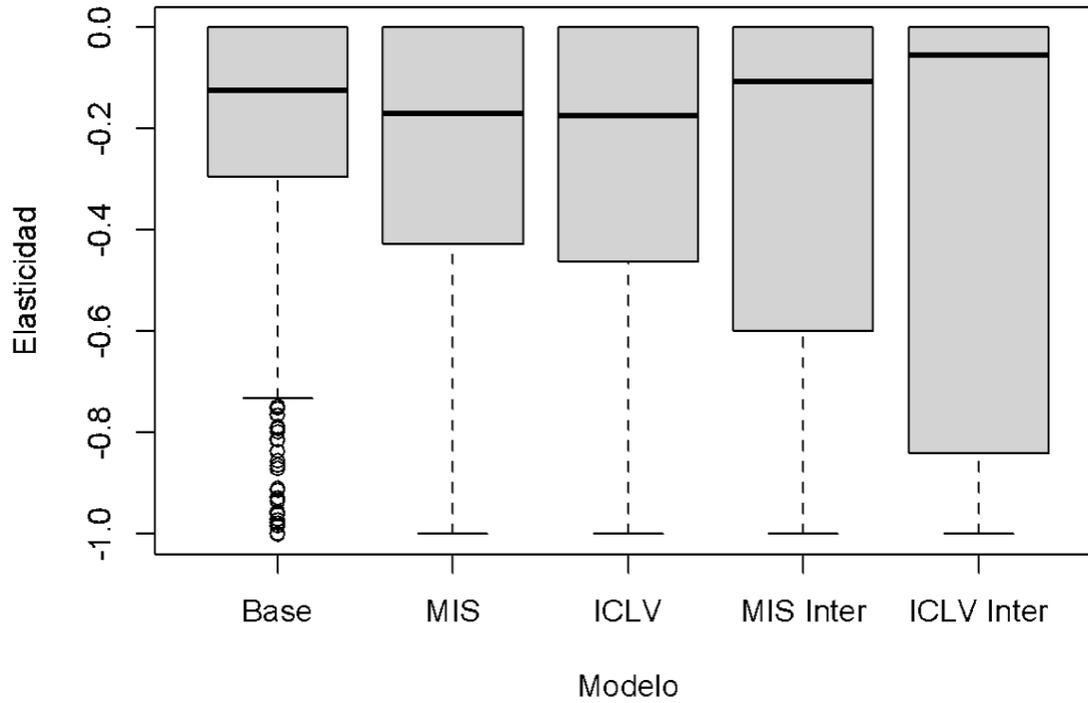


Figura 5.1: Elasticidad del tiempo de viaje para la bicicleta

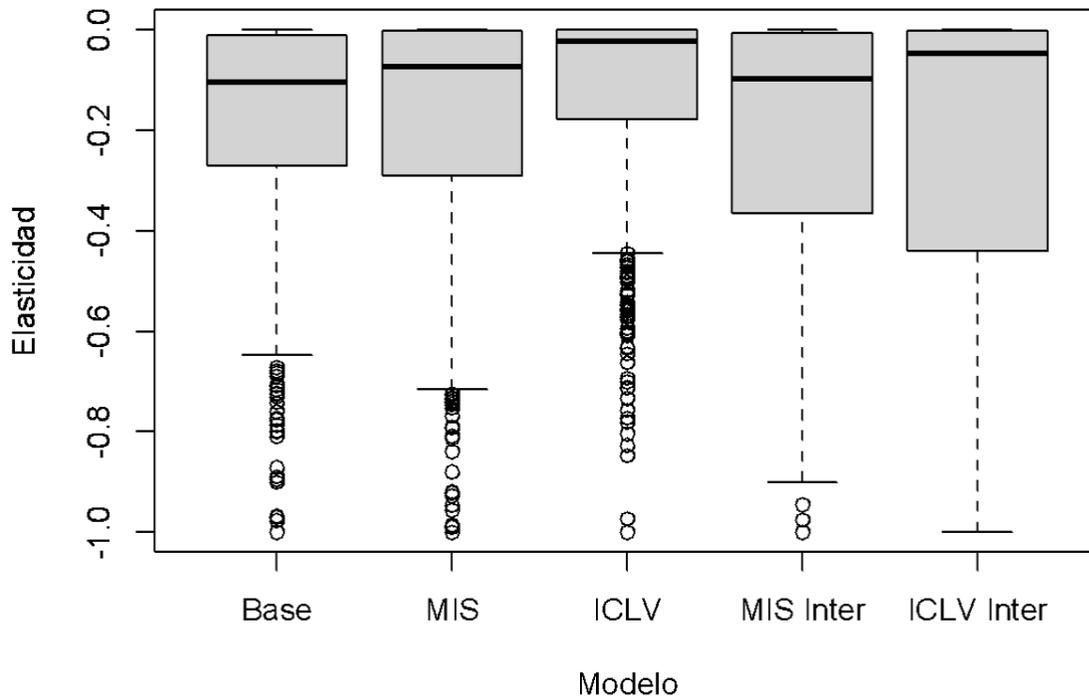


Figura 5.2: Elasticidad del tiempo de viaje para el automóvil

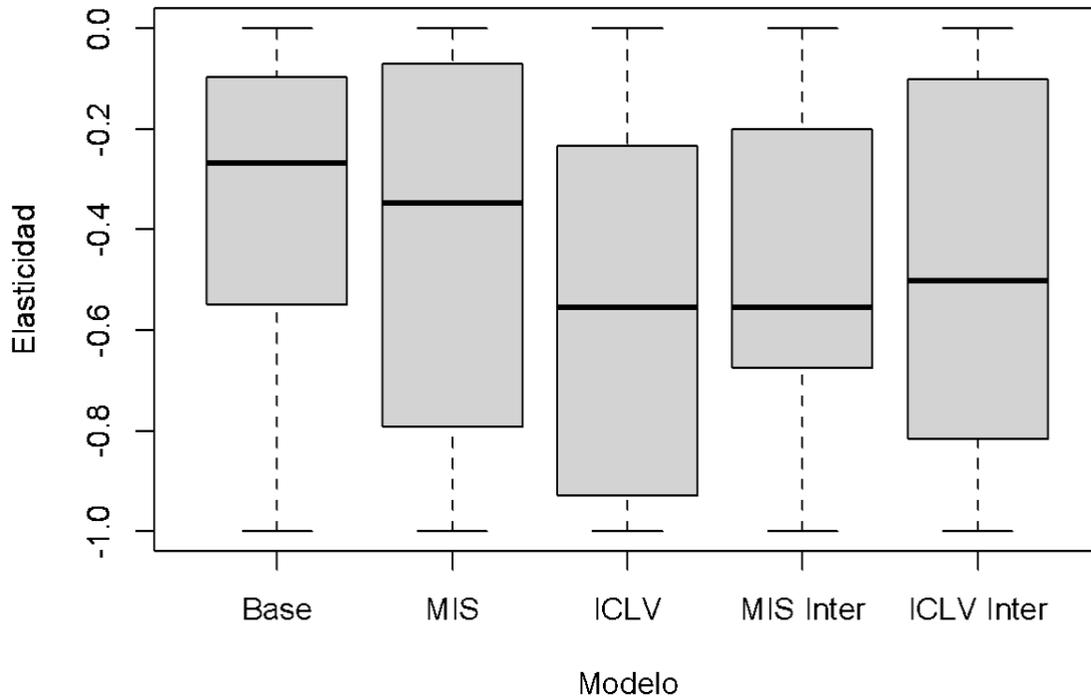


Figura 5.3: Elasticidad del tiempo de viaje para el transporte público

La muestra mostró mayor disposición de cambiarse de modo de transporte cuando se aumenta el tiempo de viaje del transporte público, lo que se condice con las valoraciones hacia este modo en la estadística descriptiva. En cambio, para el automóvil sucedió lo opuesto: los usuarios se mostraron menos dispuestos a cambiar. La elasticidad del tiempo de viaje para la bicicleta mostró valores intermedios; lo anterior aplica para todos los modelos estimados.

Se prueba que la incorporación de las variables psicosociales en los modelos de partición modal produce cambios significativos en la elasticidad del tiempo de viaje. El primer análisis se para los modelos aditivos MIS y ICLV. Se puede observar que ambas metodologías ofrecen resultados casi idénticos entre sí para la elasticidad del tiempo de viaje de la bicicleta, y muestran diferencias significativas con respecto al modelo base. En cuanto a la elasticidad del tiempo de viaje del automóvil y del transporte público, se muestra que, para el primero, la incorporación de dichas variables produce una disminución en este índice, y para el bus la aumenta. En ambos gráficos se nota que el modelo ICLV acentúa los resultados del MIS. A partir de las valoraciones realizadas en la Sección 3.4 se tiene que, generalmente, las valoraciones hacia el automóvil son positivas, mientras que para el autobús son

fundamentalmente negativas. Por lo tanto, se tienen resultados adecuados en términos de elasticidad. La semejanza o diferencia entre ambas metodologías pareciera tener directa relación con la cantidad de usuarios que eligió algún modo de transporte: para la bicicleta se tienen 122 usuarios, para el automóvil 423 y para el transporte público 540. Este estudio se deja para una investigación futura.

Los efectos de la interacción muestran resultados similares con los modelos aditivos en cuanto a la disminución o aumento de la elasticidad del tiempo de viaje para todas las alternativas, pero el tamaño cambia. En particular, para la bicicleta se tiene una mayor disminución de dicha elasticidad para ambas metodologías. Para el auto, los modelos aditivos muestran resultados similares a los modelos con interacción, si se comparan ambos métodos con el método con interacción respectiva, pero los modelos con interacción presentan intervalos de confianza más amplios. Con respecto al transporte público, el modelo MIS con interacción muestra una elasticidad de mayor tamaño que el MIS aditivo, mientras que el intervalo de confianza del ICLV con interacción es menor en magnitud que el del ICLV aditivo. La contradicción entre modelos aditivos y con interacción no necesariamente es preocupante, pues las interpretaciones de ambos efectos son distintas.

En las Figuras 5.4 y 5.5 se presentan las elasticidades del costo monetario, que en todos los casos muestra ser menor que la del tiempo de viaje. Esto quiere decir que, independiente de la incorporación las variables psicosociales y de la metodología de evaluación, la muestra se mostró más sensible a un cambio porcentual en el tiempo de viaje, que en el costo monetario asociado a viajar. Al incorporar dichas variables en la modelación, se aprecia una disminución de este índice para el automóvil y un aumento para el bus, análogo a lo que sucede con el tiempo de viaje.

Las comparaciones de ambas metodologías con su modelo interactivo respectivo muestran resultados muy similares para la elasticidad del costo del automóvil, en términos de intervalos de confianza y de la mediana. Por otro lado, para la elasticidad del costo del transporte público, los modelos con interacción muestran mayores valoraciones que los modelos aditivos.

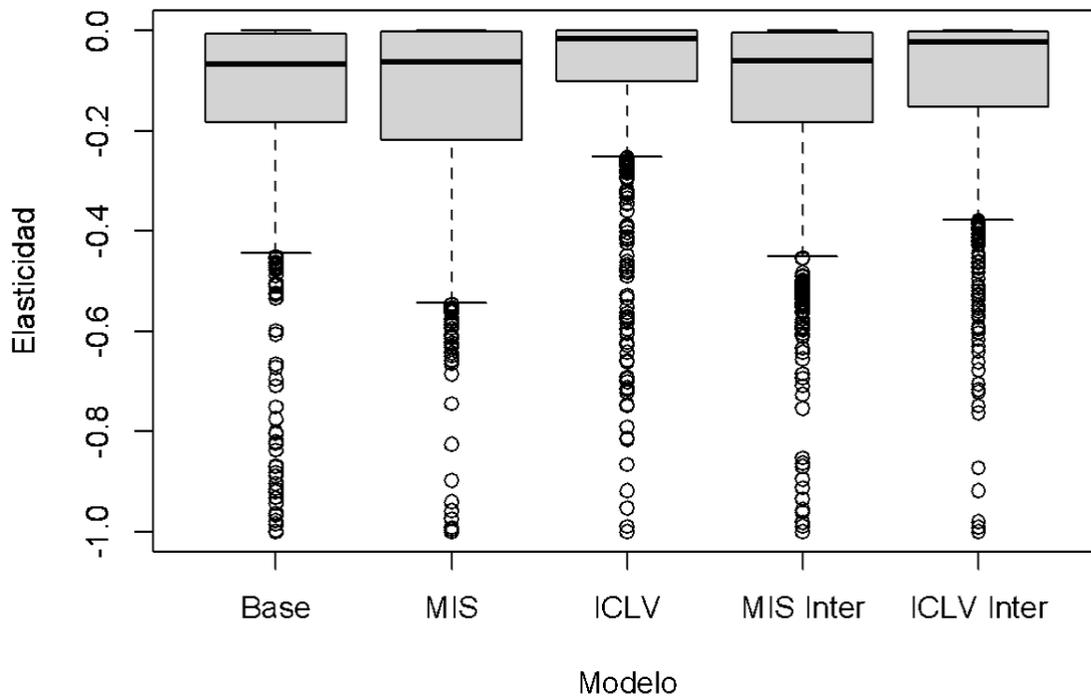


Figura 5.4: Elasticidad del costo para el automóvil

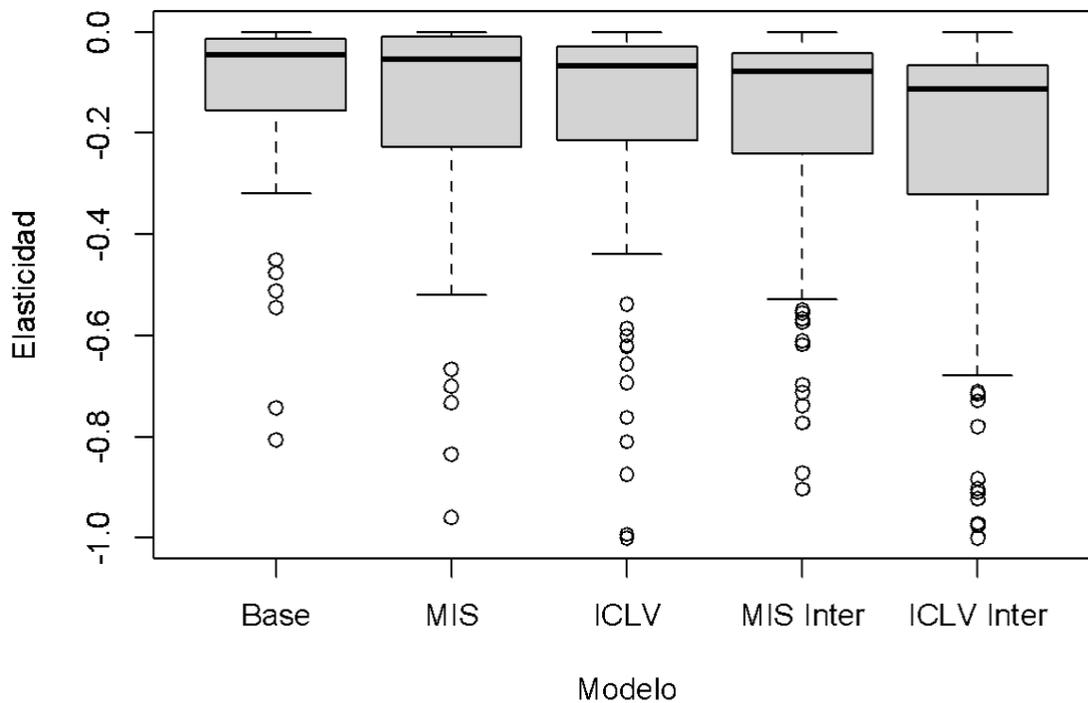


Figura 5.5: Elasticidad del costo para el transporte público

5.8. Conclusión

En el presente capítulo se presentaron los resultados de las metodologías para corregir la endogeneidad explicadas en el Capítulo 4, junto con un modelo que no corrige esta anomalía. Se concluye que la incorporación de variables psicosociales tiene un impacto significativo en la estimación de coeficientes, pues todos, independiente del método, resultaron relevante estadísticamente, y se presentan cambios de magnitud en las demás utilidades marginales utilizadas en la modelación.

Además, se evalúan ambas metodologías en función de indicadores económicos, donde también se obtienen cambios sustanciales en comparación con un modelo básico. Por lo tanto, se sugiere evaluar la potencial presencia de endogeneidad en modelos de partición modal utilizados para evaluar políticas públicas, dado que se podrían estar asumiendo comportamientos que pueden diferir mucho de lo que sucede en realidad.

Cabe destacar que el modelo ICLV ofrece la mayor cantidad de información, por lo tanto, debería ser el modelo que mejor represente el fenómeno que se está modelando. Sin embargo, ofrece desventajas significativas en términos de costos asociados a su estimación y cantidad de información que se debe recolectar para implementarlo. Para este caso en particular, el método MIS demoró menos de un minuto en estimar los coeficientes, mientras que para el método ICLV más exigente en términos de ecuaciones estructurales, el tiempo de cálculo fue de varios días.

Por otro lado, se aplica por primera vez la metodología MIS para la inclusión de variables psicosociales postuladas en la TCI de Triandis, y se concluye que es una herramienta adecuada para la corrección de endogeneidad producida por la omisión de dichas variables. Se ofrecen diversas ventajas con respecto al método ICLV: los costos son mucho menores, requiere supuestos menos potentes y es fácil de implementar. De igual manera, se postula que ambas metodologías ofrecen ventajas y desventajas con respecto a la otra alternativa, y se pueden aplicar a distintos escenarios.

Por último, es importante mencionar que las condiciones facilitadoras relevantes para explicar el comportamiento resultaron las mismas para todos los modelos, a excepción del ICLV, donde algunas dejan de ser significativas. Esto se explica porque las nuevas especificaciones buscan generar modelos más sofisticados, es decir, incluir variables explicativas nuevas, en vez de considerar nuevamente las anteriores.

Un supuesto importante que se realizó es que la utilidad marginal del indicador utilizado será constante para toda la muestra, lo que podría ser cuestionable, dado que se están buscando estudiar características propias del individuo. Dicho efecto se podría capturar con clases latentes, necesitando distintos indicadores para cada clase, o con estimando un *modelo logit mixto*, asumiendo una distribución para los coeficientes estimados, que, como se mencionó anteriormente, se intentó para la presente muestra.

CAPÍTULO 6: CONCLUSIONES

6.1. Conclusiones

En la presente Tesis se evalúa la endogeneidad producida por omitir variables psicosociales en modelos de partición modal con dos enfoques distintos: el método ICLV y el método MIS. A partir de esto, se demuestra que se presentan diferencias relevantes entre un modelo que presenta dicha corrección y un modelo que no evalúa estos impactos, para ambas metodologías estudiadas. Esto se puede apreciar fundamentalmente en los índices económicos estudiados: tasas marginales de sustitución y elasticidades. De esta manera, se confirma la hipótesis planteada: el afecto, la actitud y el hábito hacia un modo de transporte tienen impactos significativos en la toma de decisiones de los usuarios con respecto a qué modo utilizar para viajar.

Además, esta investigación muestra que la aplicación del método MIS para corregir la endogeneidad producida por la omisión de variables psicosociales es una herramienta válida para ser utilizada en la práctica debido a su simplicidad y parsimonia. Otra ventaja importante es que puede ser aplicado en cualquier *software* econométrico. A partir de la brecha encontrada en la literatura, se presenta una aplicación del MIS con datos reales, y se demuestra que sus resultados son más consistentes que un modelo que no corrige endogeneidad.

También se presenta una validación con datos reales de la metodología MIS para modelos con interacción, en conjunto con el método ICLV con interacción, que ya había sido estudiado en la literatura. Esto podría ser valioso para otros investigadores, pues se requieren supuestos y técnicas menos difíciles de aplicar en algunos casos. De esta manera, se estudiaron los efectos de las variables psicosociales sobre el tiempo de viaje, generando modelos que muestran una mayor verosimilitud que una función de utilidad lineal. Más aún, estas variables se desagregan según modo de transporte, y se demuestra que el impacto de dichas variables es más significativo para el automóvil que para el transporte público y la bicicleta (para los cuales no deja de ser relevante). Es importante mencionar que ambos métodos dependen directamente de la calidad (y eventualmente cantidad) de los indicadores.

Como mostró Guevara (2015), los resultados pueden estar gravemente sesgados si los indicadores no satisfacen los supuestos mencionados en el Capítulo 4.

Una de las principales ventajas del método ICLV sobre el MIS es la posibilidad de modelar las variables latentes con los modelos MIMIC. De esta manera, se cuantifica el efecto marginal de las características sociodemográficas de la muestra sobre la variable latente. Eso puede resultar muy útil al momento de implementar políticas públicas que busquen incidir en la elección de modo de los usuarios, pues se puede apuntar hacia características personales en vez de enfoques tradicionales que se basan fundamentalmente en la infraestructura.

6.2. Líneas de Investigación Futura

Una línea de investigación futura interesante es evaluar la endogeneidad producida por correlaciones entre las variables psicosociales. Por ejemplo, es posible que el afecto y el hábito hacia un modo de transporte estén correlacionados, precisamente porque se retroalimentan entre sí. La metodología adecuada para esta evaluación se denomina ecuaciones simultáneas y no se han encontrado artículos en la literatura que se refieran a este método en ingeniería de transporte, sólo se encontraron aplicaciones en materias económicas. La razón por la que no se evaluó en el presente trabajo es la falta de disponibilidad de variables instrumentales para plantear dichas ecuaciones, que naturalmente no es trivial. Se sospecha que, precisamente por la dificultad para conseguir esos instrumentos, es porque no se ha realizado una evaluación pertinente en la literatura.

Como se dijo en el Capítulo 1, nunca se sabe por completo si el modelo está incorporando todas las variables que son significativas para representar el fenómeno. El alcance de la presente Tesis es evaluar la endogeneidad producida por la omisión de variables psicosociales, pero es posible que haya otra variable relevante que no se haya considerado; un buen ejemplo es el hacinamiento, considerado por Guevara *et al.* (2020). Uno de los principales problemas asociado a la incorporación de nuevas (y vanguardistas) variables es la dificultad de su medición. Por lo tanto, se debe generar una encuesta que, además, incorpore las variables utilizadas en la presente modelación.

CAPÍTULO 7: REFERENCIAS

Abdel-Aal, M. M. M. (2017). Value of time determination for the city of Alexandria based on a disaggregate binary mode choice model. *Alexandria Engineering Journal*, 56(4), 567-578.

Abou-Zeid, M., & Ben-Akiva, M. (2014). Hybrid choice models. In *Handbook of choice modelling* (pp. 383-412). Edward Elgar Publishing.

Abou-Zeid, Maya, Ben-Akiva, Moshe, Bierlaire, Michel, Choudhury, Charisma, Hess, Stephane, 2010. Attitudes and value of time heterogeneity. In: de Voorde, Eddy Van, Vanellander, Thierry (Eds.), *Applied Transport Economics: A Management and Policy Perspective*, pp. 523–545.

Ajzen, I. (1991) The theory of planned behavior. *Organizational Behavior and Human Decision Processes*. 50. 179-211.

Aldarete C. (2011) El Rol de los Factores Sociales en la Elección del Modo de Transporte. Memoria de Título Ingeniero Civil. Departamento de Ingeniería Civil. Universidad de Concepción. Concepción.

Armstrong, P., Garrido, R., Ortúzar, J. Confidence intervals to bound the value of time (2001). *Transportation Research Part E: Logistics and Transportation Review*, Volume 37, Issues 2 -3, Pages 143-161, ISSN 1366-5545, [https://doi.org/10.1016/S1366-5545\(00\)00019-3](https://doi.org/10.1016/S1366-5545(00)00019-3).

Atasoy, B., Glerum, A., Bierlaire, M., (2013) Attitudes towards mode choice in Switzerland, *disP - The Planning Review*, 49:2, 101-117. DOI: 10.1080/02513625.2013.827518.

Ababio-Donkor, A., Saleh, W., Fonzone, A., (2020) Understanding transport mode choice for commuting: the role of affect, *Transportation Planning and Technology*, 43:4, 385-403, DOI: 10.1080/03081060.2020.1747203.

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*, 431–470. <https://doi.org/10.1016/B978-008044044-6/50022-X>.

Berry, S., Levinsohn, J., Pakes, A., 1995. Automobile prices in market equilibrium. *Econometrica* 63, 841–890.

Bierlaire, M. (2003) BIOGEME: A free package for the estimation of discrete choice models. Proceedings of the 3rd Swiss Transportation Research Conference, Ascona, Switzerland.

Bierlaire, M. (2018). Estimating choice models with latent variables with PandasBiogeme. Report TRANSP-OR, 181227.

Bierlaire, M. (2019). Monte-Carlo integration with PandasBiogeme. Report TRANSP-OR, 191231.

Bierlaire, M. (2020). A short introduction to PandasBiogeme. A short introduction to PandasBiogeme.

Bolduc, D., Boucher, N., 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. DOI: 10.3141/2082-08.

Budziński, W., Czajkowski, M. Endogeneity and Measurement Bias of the Indicator Variables in Hybrid Choice Models: A Monte Carlo Investigation. *Environ Resource Econ* 83, 605–629 (2022). <https://doi.org/10.1007/s10640-022-00702-0>.

Cherchi, E., & Ortúzar, J. D. D. (2002). Mixed RP/SP models incorporating interaction effects. *Transportation*, 29(4), 371-395.

Cherchi, E., Ortúzar, J. (2006). Income, Time Effects and Direct Preferences in a Multimodal Choice Context: Application of Mixed RP/SP Models with Non-Linear Utilities. *Networks and Spatial Economics* 6(1):7-23. DOI: 10.1007/s11067-006-7682-7.

Referencias

Chorus, C.G., Kroesen, M., 2014. On the (im-)possibility of deriving transport policy implications from hybrid choice models. *Transport Policy* 36, 217–222.

Collins, C., Chambers, S. (2005). Psychological and Situational Influences on Commuter-Transport-Mode Choice. *Environment and Behavior* 37(5):640-661. DOI: 10.1177/0013916504265440.

Corraliza, J. (1987) *La Experiencia del Ambiente. Percepción y Significado del Medio Construido*. Tecnos. Madrid.

Cortés, A. (2017). Factores que Explican el Uso de la Bicicleta en los Viajes al Campus de la Universidad de Concepción. Memoria de Título Ingeniero Civil. Departamento de Ingeniería Civil. Universidad de Concepción. Concepción.

Daly, A., Hess, S., Patrui, B., Potoglou, D., & Rohr, C. (2012). Using ordered attitudinal indicators in a latent variable choice model: a study of the impact of security on rail travel behaviour. *Transportation*, 39, 267-297.

De Maghalaes, D., Rivera-Gonzalez, C. (2021). Car users' attitudes towards an enhanced bus system to mitigate urban congestion in a developing country. *Transport Policy* 110(5). DOI: 10.1016/j.tranpol.2021.06.013.

Devika, R., Harikrishna, M., & Anjaneyulu, M. V. L. R. (2020). Influence of psychological factors in mode choice decision making: A structural equation modeling approach. *Transportation research procedia*, 48, 2821-2830.

Domarchi, C., Tudela, A., González, A. (2008). Effect of attitudes, habit and affective appraisal on mode choice: an application to university workers. *Transportation*, Springer, vol. 35(5), pages 585-599, August. DOI: 10.1007/s11116-008-9168-6.

Fernández-Antolín, A., Guevara, C. A., De Lapparent, M., & Bierlaire, M. (2016). Correcting for endogeneity due to omitted attitudes: Empirical assessment of a modified MIS method using RP mode choice data. *Journal of choice modelling*, 20, 1-15.

Fishbein, M. & Ajzen, I. (1975) *Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research*. Addison-Wesley. Reading.

Guerrero, T. E., Guevara, C. A., Cherchi, E., & de Dios Ortúzar, J. (2022). Characterizing the impact of discrete indicators to correct for endogeneity in discrete choice models. *Journal of choice modelling*, 42, 100342.

Guerrero, T.E., Guevara, C.A., Cherchi, E. Addressing endogeneity in strategic urban mode choice models. *Transportation* 48, 2081–2102 (2021). <https://doi.org/10.1007/s11116-020-10122-y>.

Guevara, C. A., & Polanco, D. (2016). Correcting for endogeneity due to omitted attributes in discrete-choice models: the multiple indicator solution. *Transportmetrica A: Transport Science*, 12(5), 458-478.

Guevara, C. A., & Ben-Akiva M. (2012). Change of Scale and Forecasting with the Control-Function Method in Logit Models. *Transportation Science* 46.3: 425–437.

Guevara, C. A., & Ben-Akiva, M. (2006). Endogeneity in Residential Location Choice Models. *Transportation Research Record*, 1977(1), 60-66. <https://doi.org/10.1177/0361198106197700108>.

Guevara, C. A., 2015. Critical assessment of five methods to correct for endogeneity in discrete-choice models. Volume 82, (2015), Pages 240-254. ISSN 0965-8564. <https://doi.org/10.1016/j.tra.2015.10.005>.

Guevara, C. A., Tirachini, A., Hurtubia, R., Dekker, T., (2020). Correcting for endogeneity due to omitted crowding in public transport choice using the Multiple Indicator Solution (MIS) method. *Transportation Research Part A: Policy and Practice*. Volume 137, Pages 472-484. ISSN 0965-8564. <https://doi.org/10.1016/j.tra.2018.10.030>.

Hausman, J. A. (1978). Specification Tests in Econometrics. *Econometrica*, 46(6), 1251–1271. <https://doi.org/10.2307/1913827>.

Hess, S., Spitz, G., Bradley, M., & Coogan, M. (2018). Analysis of mode choice for intercity travel: Application of a hybrid choice model to two distinct US corridors. *Transportation Research Part A: Policy and Practice*, 116, 547-567.

Horowitz, J. L., (1983) Statistical Comparison of Non-Nested Probabilistic Discrete Choice Models. *Transportation Science* 17(3):319-350. <http://dx.doi.org/10.1287/trsc.17.3.319>.

Huan, N., Hess, S., Yao, E. (2021). Understanding the effects of travel demand management on metro commuters' behavioural loyalty: a hybrid choice modelling approach. *Transportation* 49(1). DOI: 10.1007/s11116-021-10179-3.

Idris, A., Habib K. M. N., Tudela, A., & Shalaby, A., (2015) Investigating the effects of psychological factors on commuting mode choice behaviour. *Transportation Planning and Technology*, 38:3, 265-276, DOI: 10.1080/03081060.2014.997451

Jensen, P., Rouquier, J. B., Ovtracht, N., & Robardet, C. (2010). Characterizing the speed and paths of shared bicycle use in Lyon. *Transportation research part D: transport and environment*, 15(8), 522-524.

Johansson, M.V., Heldt, T., Johansson, P. (2005). The effects of attitudes and personality traits on mode choice. *Transportation Research Part A Policy and Practice* 40(6):507-525. DOI: 10.1016/j.tra.2005.09.001.

Kim, J., & Lee, B. (2023). Campus commute mode choice in a college town: An application of the integrated choice and latent variable (ICLV) model. *Travel Behaviour and Society*, 30, 249-261.

Likert, R. (1934) A technique for the measurement of attitudes. *Archives of Psychology*. 140. 1- 55.

Referencias

Lizana, M., Tudela, A., Tapia, A. (2021). Analysing the influence of attitude and habit on bicycle commuting. *Transportation Research Part F: Traffic Psychology and Behaviour*. Volume 82. Pages 70-83. ISSN 1369-8478. <https://doi.org/10.1016/j.trf.2021.07.015>.

Mariel, P., Hoyos, D., Artabe, A., Guevara, C. A. (2018). A multiple indicator solution approach to endogeneity in discrete-choice models for environmental valuation. *Science of The Total Environment*, Volume 633. Pages 967-980, ISSN 0048-9697, <https://doi.org/10.1016/j.scitotenv.2018.03.254>.

Muñoz, B., Monzon, A. & Daziano R. (2016). The Increasing Role of Latent Variables in Modelling Bicycle Mode Choice, *Transport Reviews*, 36:6, 737- 771, DOI: 10.1080/01441647.2016.1162874.

Ortúzar, J. & Willumsen L. (2011) *Modelling Transport*. 4th edition. Wiley and Sons, Chichester.

Park, S. & Gupta, S., 2009. Simulated maximum likelihood estimator for the random coefficient logit model using aggregate data. *J. Market. Res.* 46, 531–542.

Paulssen, M., Temme, D., Vij, A., & Walker, J. L. (2014). Values, attitudes and travel behavior: a hierarchical latent variable mixed logit model of travel mode choice. *Transportation*, 41, 873-888.

Petrin, A., and K. Train. (2002). Omitted Product Attributes in Discrete-Choice Models. Working Paper, Department of Economics, University of California, Berkeley, CA.

Petrin, A., Train, K., (2010). A control function approach to endogeneity in consumer choice models. *J. Market. Res.* 47, 3–13.

Raveau, S., Álvarez-Daziano, R., Yáñez, M. F., Bolduc, D., & Ortúzar, J. D. D. (2010). Sequential and simultaneous estimation of hybrid discrete choice models: Some new findings. *Transportation Research Record*, 2156(1), 131-139. <https://doi.org/10.3141/2156-15>.

Raychaudhuri, S. (2008). Introduction to monte carlo simulation. In 2008 Winter simulation conference (pp. 91-100). IEEE.

Reeve, J. (1994) Motivación y Emoción. McGraw-Hill/Interamericana de España. Madrid.

Román, C., Martín J. C., Espino, R., Cherchi, E., Ortúzar, J., Rizzi., L. I., González R. M., Amador, F. J. (2014). Valuation of travel time savings for intercity travel: The Madrid-Barcelona corridor. *Transport Policy* 36, 105-117. <https://doi.org/10.1016/j.tranpol.2014.07.007>.

Scagnolari S, Walker J, Maggi R. (2015). Young drivers' night-time mobility preferences and attitude toward alcohol consumption: A Hybrid Choice Model. *Accid Anal Prev*; 83:74-89. doi: 10.1016/j.aap.2015.07.002. Epub 2015 Jul 26. PMID: 26221792.

Soza-Parra, J. & Cats, O., (2023). The Role of Personal Motives in Determining Car Ownership and Use: A Literature Review. SSRN: <https://ssrn.com/abstract=4576694> or <http://dx.doi.org/10.2139/ssrn.4576694>

Stock, J., & Yogo, M. (2005). Testing for Weak Instruments in Linear IV Regression. In D. Andrews & J. Stock (Eds.), *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg* (pp. 80-108). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511614491.006.

Tapia, A. (2016). El Impacto de las Variables de Personalidad en la Decisión de Inversión de un Proyecto de Transporte. Un Estudio de Caso en Concepción. Memoria de Título Ingeniero Civil. Departamento de Ingeniería Civil. Universidad de Concepción. Concepción.

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International journal of medical education*, 2, 53.

Thigpen, C. G., Driller, B. K., & Handy, S. L. (2015). Using a stages of change approach to explore opportunities for increasing bicycle commuting. *Transportation Research Part D: Transport and Environment*, 39, 44-55. <https://doi.org/10.1016/j.trd.2015.05.005>.

Referencias

Guerrero, T. E., Guevara, C. A., Cherchi, E., Ortúzar, J. D. D., (2022) Forecasting with strategic transport models corrected for endogeneity, *Transportmetrica A: Transport Science*, 18:3, 708-735, DOI: 10.1080/23249935.2021.1891154.

Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge university press. *Transportation Planning and Technology*, 38:3, 265-276, DOI: 10.1080/03081060.2014.997451.

Triandis, H. (1977) *Interpersonal Behavior*. Brooks and Cole, Monterrey.

Vij, A., Walker, J. (2016). How, when and why integrated choice and latent variable models are latently useful. *Transportation Research Part B: Methodological*, Volume 90, Pages 192-217, ISSN 0191-2615, <https://doi.org/10.1016/j.trb.2016.04.021>.

Walker, J. (2001). *Extended Discrete Choice Models: Integrated Framework, Flexible Error Structures and Latent Variables*. Ph.D. Thesis. Massachusetts Institute of Technology, Dept. of Civil and Environmental Engineering.

Washington, S., Karlaftis, M.G., Mannering, F., & Anastasopoulos, P. (2020). *Statistical and Econometric Methods for Transportation Data Analysis* (3rd ed.). Chapman and Hall/CRC. <https://doi.org/10.1201/9780429244018>.

Wooldridge, Jeffrey M., (2010). *Econometric Analysis of Cross Section and Panel Data*. MIT Pres.

Yong, A.G. and Pearce, S. (2013) A Beginner's Guide to Factor Analysis: Focusing on Exploratory Factor Analysis. *Tutorials in Quantitative Methods for Psychology*, 9, 79-94. <https://doi.org/10.20982/tqmp.09.2.p079>

CAPÍTULO 8: ANEXOS

ANEXO 8.1. Contribución a los Objetivos de Desarrollo Sostenible

Los Objetivos de Desarrollo Sostenible (ODS) constituyen un llamamiento universal a la acción para poner fin a la pobreza, proteger el planeta y mejorar las vidas y las perspectivas de las personas en todo el mundo. Seleccione a cuál de los 17 ODS contribuye su trabajo de Tesis:

- ODS-1 : Fin de la pobreza.
- ODS-2 : Hambre cero.
- ODS-3 : Salud y bienestar.
- ODS-4 : Educación de calidad.
- ODS-5 : Igualdad de género.
- ODS-6 : Agua limpia y saneamiento.
- ODS-7 : Energía asequible y no contaminante.
- ODS-8 : Trabajo decente y crecimiento económico.
- ODS-9 : Industria, innovación e infraestructura.
- ODS-10 : Reducción de las desigualdades.
- ODS-11 : Ciudades y comunidades sostenibles.
- ODS-12 : Producción y consumo responsables.
- ODS-13 : Acción por el clima.
- ODS-14 : Vida Submarina.
- ODS-15 : Vida de ecosistemas terrestres.
- ODS-16 : Paz, justicia e instituciones sólidas.
- ODS-17 : Alianzas para lograr los objetivos.

Vinculación

La presente Tesis se vincula con el ODS-9, particularmente en la sección de innovación. Muchos de los modelos de partición modal en Chile incorporan solamente variables de servicio, mientras que en el presente trabajo se incorporan variables psicosociales, esto es, un modelo vanguardista y poco utilizado en la práctica.

Además, se considera que se vincula con el ODS-11. Se aporta a las ciudades sostenibles a partir de la incorporación de factores medioambientales en el modelo propuesto. En particular, se tendrá un factor de corrección para las variables de servicio, a partir de factores actitudinales.

ANEXO 8.2. Análisis Factorial

Tabla 8.1: Análisis factorial para el afecto hacia la bicicleta

Indicadores		Valor	Error std.	Valor-z	P(> z)
Control bicicleta					
Bueno	Malo	2,41	0,058	41,4	0,00
Incómodo	Cómodo	-0,08	0,061	-1,3	0,19
Agradable	Desagradable	2,09	0,054	38,5	0,00
Sucio	Limpio	-0,01	0,061	-0,1	0,89
Potencia bicicleta					
Débil	Fuerte	-0,01	0,065	-0,1	0,93
Grandioso	Insignificante	1,62	0,055	29,4	0,00
Flexible	Inflexible	1,71	0,058	29,3	0,00
Grande	Pequeño	-0,15	0,064	-2,4	0,02
Activación bicicleta					
Poblado	Despoblado	0,02	0,071	0,3	0,74
Rápido	Lento	-1,08	0,073	-14,9	0,00
Pasivo	Activo	-0,07	0,059	-1,2	0,24
Tranquilo	Intranquilo	-1,12	0,075	-15,1	0,00
Control bicicleta					
Complejo	Sencillo	1,21	0,07	17,3	0,00
Inseguro	Seguro	0,82	0,061	13,5	0,00
Oscuro	Claro	1,17	0,067	17,4	0,00
Desconocido	Conocido	1,30	0,068	19,3	0,00

Tabla 8.2: Análisis factorial para el afecto hacia el automóvil

Indicadores		Valor	Error std.	Valor-z	P(> z)
Control automóvil					
Bueno	Malo	1,90	0,054	35,0	0,00
Incómodo	Cómodo	0,15	0,038	3,8	0,00
Agradable	Desagradable	2,29	0,054	42,2	0,00
Sucio	Limpio	0,18	0,058	3,1	0,00
Potencia automóvil					
Débil	Fuerte	0,38	0,058	6,5	0,00
Grandioso	Insignificante	1,78	0,052	34,1	0,00
Flexible	Inflexible	2,01	0,058	34,9	0,00
Grande	Pequeño	1,43	0,055	26,2	0,00
Activación automóvil					
Poblado	Despoblado	0,80	0,066	12,0	0,00
Rápido	Lento	2,31	0,056	40,9	0,00
Pasivo	Activo	0,31	0,066	4,7	0,00
Tranquilo	Intranquilo	1,64	0,057	28,7	0,00
Control automóvil					
Complejo	Sencillo	1,11	0,067	16,5	0,00
Inseguro	Seguro	1,15	0,05	22,9	0,00
Oscuro	Claro	1,36	0,063	21,7	0,00
Desconocido	Conocido	1,08	0,05	21,7	0,00

Tabla 8.3: Análisis factorial para el afecto hacia el bus

Indicadores		Valor	Error std.	Valor-z	P(> z)
Control transporte público					
Bueno	Malo	1,15	0,068	16,8	0,00
Incómodo	Cómodo	0,01	0,048	0,1	0,89
Agradable	Desagradable	2,07	0,082	25,2	0,00
Sucio	Limpio	0,00	0,039	0,0	0,97
Potencia transporte público					
Débil	Fuerte	0,64	0,06	10,5	0,00
Grandioso	Insignificante	1,11	0,052	21,5	0,00
Flexible	Inflexible	1,41	0,056	25,3	0,00
Grande	Pequeño	1,41	0,06	23,4	0,00
Activación transporte público					
Poblado	Despoblado	2,21	0,06	37,0	0,00
Rápido	Lento	-0,40	0,058	-6,8	0,00
Pasivo	Activo	0,34	0,061	5,6	0,00
Tranquilo	Intranquilo	-1,34	0,056	-23,9	0,00
Control transporte público					
Complejo	Sencillo	1,11	0,066	16,9	0,00
Inseguro	Seguro	1,11	0,058	19,3	0,00
Oscuro	Claro	1,24	0,055	22,7	0,00
Desconocido	Conocido	0,92	0,066	13,9	0,00

Tabla 8.4: Análisis factorial para la actitud hacia la bicicleta

Indicadores	Valor	Error std.	Valor-z	P(> z)
Comodidad	3,76	0,183	20,6	0,00
Diversión	3,81	0,169	22,5	0,00
Ambientalismo	2,83	0,158	17,9	0,00
Salud	3,77	0,168	22,4	0,00
Economía	2,12	0,155	13,7	0,00
Seguridad	2,49	0,186	13,4	0,00

Tabla 8.5: Análisis factorial para la actitud hacia el automóvil

Indicadores	Valor	Error std.	Valor-z	P(> z)
Comodidad	2,68	0,165	16,2	0,00
Diversión	2,64	0,167	15,8	0,00
Ambientalismo	3,24	0,17	19,1	0,00
Salud	4,21	0,187	22,5	0,00
Economía	2,58	0,191	13,5	0,00
Seguridad	2,25	0,161	13,9	0,00

Tabla 8.6: Análisis factorial para la actitud hacia el transporte público

Indicadores	Valor	Error std.	Valor-z	P(> z)
Comodidad	3,21	0,168	19,1	0,00
Diversión	2,40	0,132	18,1	0,00
Ambientalismo	2,77	0,176	15,7	0,00
Salud	3,63	0,147	24,6	0,00
Economía	2,18	0,189	11,5	0,00
Seguridad	2,82	0,171	16,5	0,00

ANEXO 8.3. Resultados SEM para modelo con interacción**Tabla 8.7: Resultados de los SEM para modelo con interacción**

Coficiente	Valor	Error std.	Test-t	Valor-p
Afecto				
Intercepto control automóvil	1,06	0,5	2,12	0,01
Intercepto control bicicleta	1,4	0,2	7,16	0,03
Intercepto control bus	2,6	0,2	13	0
Carga control automóvil	0,92	0,37	2,49	0
Carga control bicicleta	2,08	0,67	3,11	0
Carga control bus	5,79	0,64	9,05	0
Actitud				
Intercepto ambientalismo automóvil	-6,93	0,45	-15,43	0
Intercepto diversión automóvil	-3,79	0,31	-12,39	0
Intercepto economía automóvil	-5,16	0,49	-10,62	0
Intercepto salud automóvil	-4,74	0,48	-9,85	0
Intercepto seguridad automóvil	0,71	0,25	2,89	0
Intercepto ambientalismo bicicleta	4,17	0,19	22,18	0
Intercepto diversión bicicleta	-0,57	0,12	-4,9	0
Intercepto economía bicicleta	5,77	0,26	22,45	0
Intercepto salud bicicleta	3,5	0,17	21,21	0
Intercepto seguridad bicicleta	-1,41	0,14	-10,07	0
Intercepto ambientalismo bus	-0,66	0,17	-3,98	0
Intercepto diversión bus	-2,95	0,16	-18,44	0
Intercepto economía bus	2,65	0,18	15,14	0
Intercepto salud bus	-0,98	0,15	-6,55	0
Intercepto seguridad bus	1,35	0,14	9,64	0
Carga ambientalismo automóvil	1,08	0,12	9,08	0
Carga diversión automóvil	0,6	0,08	7,91	0
Carga economía automóvil	0,93	0,14	6,84	0
Carga salud automóvil	1,23	0,13	9,39	0
Carga seguridad automóvil	0,82	0,13	6,41	0
Carga ambientalismo bicicleta	0,59	0,08	7,74	0
Carga diversión bicicleta	0,82	0,06	14,38	0
Carga economía bicicleta	0,59	0,11	5,57	0
Carga salud bicicleta	0,76	0,07	11,41	0
Carga seguridad bicicleta	0,67	0,06	12,11	0
Carga ambientalismo bus	0,78	0,09	8,44	0
Carga diversión bus	0,64	0,06	10,89	0
Carga economía bus	0,65	0,09	6,98	0
Carga salud bus	1,03	0,08	12,29	0
Carga seguridad bus	0,84	0,09	9,17	0

ANEXO 8.4. Resultados MIMIC para modelo con interacción**Tabla 8.8: Resultados de los modelos MIMIC para modelo con interacción**

Coefficiente	Valor	Error std.	Test-t	Valor-p
Afecto bicicleta				
Intercepto	-1,54	0,49	-3,15	0
Ciclismo los fines de semana	0,18	0,06	2,94	0
Licencia de conducir	-0,23	0,11	-2,09	0,04
Distancia inferior a 2,5 km	0,24	0,11	2,23	0,03
Habilidad para andar en bicicleta	0,4	0,13	2,96	0
Afecto automóvil				
Intercepto	2,41	0,69	3,49	0
Ciclismo los fines de semana	-0,27	0,07	-3,56	0
Ingreso alto	0,44	0,22	2,05	0,04
Licencia de conducir	0,78	0,23	3,36	0
Distancia inferior a 2,5 km	-0,62	0,24	-2,57	0,01
Afecto bus				
Intercepto	-2,51	0,15	-16,62	0
Ingreso alto	-0,33	0,12	-2,83	0,02
Licencia de conducir	-0,21	0,09	-2,33	0,02
Estudiante	0,89	0,14	6,59	0
Actitud bicicleta				
Intercepto	-6,05	0,7	-8,62	0
Viaje con objetos livianos	0,59	0,29	2,03	0,04
Ciclismo los fines de semana	0,77	0,24	3,23	0
Presencia de ciclovías	1,79	0,88	2,03	0,04
Edad entre 25 y 35 años	0,64	0,31	2,05	0,04
Licencia de conducir	-0,76	0,27	-2,82	0
Género	0,71	0,35	2,06	0,02
Distancia inferior a 2,5 km	1,38	0,58	2,36	0,02
Habilidad para andar en bicicleta	0,76	0,24	3,24	0
Actitud automóvil				
Intercepto	5,87	0,52	11,33	0
Edad entre 25 y 35 años	-0,65	0,28	-2,31	0,02
Licencia de conducir	0,78	0,23	3,39	0
Habilidad para andar en bicicleta	-0,49	0,16	-3,08	0
Estudiante	-0,67	0,14	-4,78	0
Actitud bus				
Intercepto	-0,58	0,79	-0,73	0,47
Viaje con objetos livianos	0,75	0,35	2,13	0,02
Ingreso alto	-0,66	0,11	-6	0
Licencia de conducir	-1,03	0,21	-4,89	0,01
Habilidad para andar en bicicleta	-0,48	0,17	-2,76	0,01
Estudiante	1,35	0,31	4,39	0