

Barranquilla, Noviembre de 2011

Señores

Comité de Proyectos de Grados e Investigaciones
División de Ingenierías
Departamento de Ingeniería Civil y Ambiental
Programa de Maestría en Ingeniería Civil
Fundación Universidad de Norte
Ciudad.

Cordial Saludo:

Por medio de la presente certifico que he asesorado en la elaboración del informe final del trabajo de grado que se titula “ESTIMACIÓN ECONOMETRICA DEL VALOR SUBJETIVO DEL TIEMPO. ANÁLISIS PARA EL PROYECTO METRO DE BOGOTÁ”. Con toda la estructura metodología, teórica y contenido en el proceso de investigación, elaborado por la estudiante Margareth Giovanna Gutiérrez Torres, para optar el título de Magister en Ingeniería Civil.

Atentamente

Víctor Manuel Cantillo Maza
Director de Tesis

Barranquilla, Noviembre de 2011

Señores

Comité de Proyectos de Grado e Investigaciones
División de Ingenierías
Departamento de Ingeniería Civil y Ambiental
Programa de Magister en Ingeniería Civil
Fundación Universidad del Norte
Ciudad.

Cordial Saludo:

Por medio de la presente me dirijo a ustedes para presentar el informe final del trabajo de grado para optar el título de Magister en Ingeniería Civil, que tiene por nombre “ESTIMACIÓN ECONOMETRICA DEL VALOR SUBJETIVO DEL TIEMPO. ANÁLISIS PARA EL PROYECTO METRO DE BOGOTÁ”.

Atentamente:

MARGARETH GUTIERREZ TORRES

Barranquilla, Noviembre de 2011

Nota de Aceptación

Firma del evaluador

Firma del evaluador

Barranquilla, Noviembre de 2011

Nota de Aceptación

Firma del evaluador

Firma del evaluador

ESTIMACIÓN ECONOMETRICA DEL VALOR SUBJETIVO DEL TIEMPO. ANÁLISIS PARA EL PROYECTO METRO DE BOGOTÁ

Por

MARGARETH GUTIERREZ TORRES

Presentado al Departamento de Ingeniería Civil y Ambiental
Como requisito parcial para optar al título de
MAESTRÍA EN INGENIERÍA CIVIL

En la

Fundación Universidad del Norte



Barranquilla, 22 noviembre 2011

AGRADECIMIENTOS

Mi primer y principal agradecimiento a Dios, por haberme dado la oportunidad de hacer parte de este gran proceso, por la capacidad, las ganas y la fortaleza de seguir adelante hasta el final. A mis padres, por acompañarme, apoyarme y estar presentes en cada uno de los momentos de mi vida. A mi esposo Nelson Rojas por darme aliento en los momentos en que creía no poder continuar, confiar en mí y transmitirme esa confianza para seguir. A Víctor Cantillo por su invaluable apoyo, conocimiento, paciencia, confianza y guía desde el primer momento en que decidí avanzar en este camino de crecimiento.

TABLA DE CONTENIDO

1	INTRODUCCIÓN.....	1
1.1	OBJETIVOS DE LA TESIS.....	1
1.2	HIPÓTESIS	1
1.3	CONTRIBUCIONES.....	2
1.4	ESTRUCTURA DE LA TESIS	2
2	MARCO CONCEPTUAL ECONOMETRICO.....	3
2.1	EL COMPORTAMIENTO DEL CONSUMIDOR RACIONAL	3
2.2	LA TEORIA DE LA UTILIDAD ALEATORIA	4
2.3	LOS MODELOS DE ELECCIÓN DISCRETA.....	5
2.3.1	Modelo Logit Multinomial.....	7
2.3.2	Modelos con Variaciones Sistemáticas.....	10
2.4	MODELO LOGIT MIXTO.....	10
2.4.1	Formulación del Modelo	11
2.4.2	Estimación Clásica	13
2.4.3	Estimación Bayesiana	17
2.5	VALOR SUBJETIVO DEL TIEMPO	19
2.5.1	Antecedentes	19
2.5.2	Estimación del Valor Subjetivo del Tiempo	21
3	DISEÑO DEL EXPERIMENTO	24
3.1	DATOS DE PREFERENCIA REVELADA Y PREFERENCIA DECLARADA	24
3.2	DISEÑO EXPERIMENTAL	26
3.2.1	Definición de Atributos y Niveles	27
3.3	INFORMACIÓN SECUNDARIA.....	34
3.4	METODOLOGIA DE LA TOMA DE INFORMACION	35
3.5	ANALISIS ESTADISTICO DE LA MUESTRA	37
4	FORMULACION Y EVALUACION DE MODELOS.....	46
4.1	ESPECIFICACIÓN DE MODELOS.....	47
4.1.1	Estimación Clásica	48
4.1.2	Estimación Bayesiana	60
4.1.3	Comparación de modelos.....	67
5	CONCLUSIONES.....	70
6	BIBLIOGRAFÍA.....	73

LISTA DE FIGURAS

<i>Figura 2.1 Dimensión 1 y 2 de secuencia pseudo-aleatorias, 300 puntos</i>	16
<i>Figura 2.2 Dimensión 1 y 2 de secuencia de Halton, 300 puntos</i>	16
<i>Figura 2.3 Dimensión 14 y 15 de secuencia de Halton</i>	16
<i>Figura 2.4 Primer ciclo de puntos de secuencia de Halton</i>	17
<i>Figura 3.1 Opciones de elección para el caso de viajes cortos</i>	32
<i>Figura 3.2 Cuestionario para encuestas PR y PD</i>	33
<i>Figura 3.3 Metodología del trabajo de campo</i>	37
<i>Figura 3.4 Zonificación del estudio</i>	37
<i>Figura 3.5 Entrevistados según estrato socioeconómico</i>	38
<i>Figura 3.6 Entrevistados según ocupación</i>	38
<i>Figura 3.7 Encuestados según cargo desempeñado</i>	38
<i>Figura 3.8 Modo alternativo al modo considerado</i>	39
<i>Figura 3.9 Motorización de usuarios de transporte privado</i>	39
<i>Figura 3.10 Motorización de todos los encuestados</i>	40
<i>Figura 3.11 Encuestas según modo de viaje</i>	40
<i>Figura 3.12 Encuestados según motivo de viaje</i>	41
<i>Figura 3.13 Frecuencia de viaje</i>	41
<i>Figura 3.14 Encuestas según localidades de destino</i>	42
<i>Figura 3.15 Encuestas según localidades de origen</i>	42
<i>Figura 3.16 Encuestas según tiempo de viaje en auto</i>	43
<i>Figura 3.17 Encuestas según tiempo de viaje en Transmilenio</i>	44
<i>Figura 3.18 Encuestas según tiempo de viaje en Bus</i>	44
<i>Figura 3.19 Encuestados según tiempo de viaje en Taxi</i>	45
<i>Figura 4.1 Distribución del valor subjetivo del tiempo de viaje, para estratos 1-2-3</i>	54
<i>Figura 4.2 Distribución del valor subjetivo del tiempo de acceso, para estratos 1-2-3</i>	54
<i>Figura 4.3 Distribución del valor subjetivo del tiempo de viaje, para estratos 4-5-6</i>	55
<i>Figura 4.4 Distribución del valor subjetivo del tiempo de acceso, para estratos 4-5-6</i>	56
<i>Figura 4.5 Distribución del valor subjetivo del tiempo de viaje, para estratos 1-2-3, transporte público</i>	59
<i>Figura 4.6 Distribución del valor subjetivo del tiempo de acceso, para estratos 1-2-3, transporte público</i>	59
<i>Figura 4.7 Distribución del valor subjetivo del tiempo de viaje, para estratos 4-5-6, transporte público</i>	60
<i>Figura 4.8 Distribución del valor subjetivo del tiempo de acceso, para estratos 4-5-6, transporte público</i>	60
<i>Figura 4.9 Valor subjetivo del tiempo de viaje para usuarios de estrato 2-3</i>	62
<i>Figura 4.10 Valor subjetivo del tiempo de acceso para usuarios de estrato 2-3</i>	63
<i>Figura 4.11 Valor subjetivo del tiempo de viaje para usuarios de estratos 4-5-6</i>	63
<i>Figura 4.12 Valor subjetivo del tiempo de acceso para usuarios de estratos 4-5-6</i>	64
<i>Figura 4.13 Valor subjetivo del tiempo de viaje para usuarios de transporte público de estratos 1, 2 y 3</i>	66

<i>Figura 4.14 Valor subjetivo del tiempo de acceso para usuarios de transporte público de estratos 1,2y3.</i>	<i>66</i>
<i>Figura 4.15 Valor subjetivo del tiempo de viaje para usuarios de estratos 456.</i>	<i>67</i>
<i>Figura 4.16 Valor subjetivo del tiempo de acceso para usuarios de estratos 456.</i>	<i>67</i>

LISTA DE TABLAS

<i>Tabla 3.1 Escenarios del Experimento.....</i>	<i>30</i>
<i>Tabla 3.2 Niveles de atributos para viajes largos.....</i>	<i>30</i>
<i>Tabla 3.3 Niveles de atributos para viajes cortos.....</i>	<i>30</i>
<i>Tabla 3.4 Diseño encuesta de preferencias declaradas PD1</i>	<i>31</i>
<i>Tabla 3.5 Tiempo de viaje en auto.....</i>	<i>43</i>
<i>Tabla 3.6 Tiempos de viaje en Transmilenio.....</i>	<i>43</i>
<i>Tabla 3.7 Tiempos de viaje en Bus.....</i>	<i>44</i>
<i>Tabla 3.8 Tiempos de viaje en Taxi</i>	<i>45</i>
<i>Tabla 4.1 Datos disponibles para estimación de modelos.....</i>	<i>47</i>
<i>Tabla 4.2 Especificación de modelos.....</i>	<i>50</i>
<i>Tabla 4.3 Modelos estimados para usuarios de transporte privado.....</i>	<i>51</i>
<i>Tabla 4.4 Tasas marginales de sustitución</i>	<i>52</i>
<i>Tabla 4.5 Especificación de modelos para usuarios de transporte público.....</i>	<i>56</i>
<i>Tabla 4.6 Modelos estimados para usuarios de transporte público.....</i>	<i>57</i>
<i>Tabla 4.7 Tasas marginales de sustitución usuarios de transporte público</i>	<i>58</i>
<i>Tabla 4.8 Especificación de modelos mixed logit ML</i>	<i>61</i>
<i>Tabla 4.9 Modelos estimados para usuarios de transporte privado, estimación bayesiana</i>	<i>61</i>
<i>Tabla 4.10 Especificación de modelos mixed logit ML para usuarios de transporte público</i>	<i>64</i>
<i>Tabla 4.11 Modelos estimados para usuarios de transporte público, estimación bayesiana</i>	<i>65</i>
<i>Tabla 4.12 Elasticidades de los diferentes modos, respecto al tiempo y el costo.....</i>	<i>69</i>

1 INTRODUCCIÓN

La presente investigación tiene su marco de referencia en el estudio de la percepción de los individuos sobre el valor subjetivo del tiempo y algunos atributos adicionales al momento de elegir el modo de transporte. Esta clase de estudio se ha venido desarrollando en los últimos años por medio de técnicas econométricas como son los modelos de elección discreta. A través de este documento se tendrá un acercamiento a las percepciones de los individuos en diferentes contextos de análisis en la ciudad de Bogotá, de los usuarios de transporte masivo y de aquellos cautivos de un nuevo medio de transporte como es el Metro y el Tren de Cercanías, proyectos de transporte de carácter estratégico que tendrán un gran impacto sobre dicha ciudad. Esto con el fin de obtener información valiosa para la evaluación social de proyectos y evaluar el comportamiento de los individuos al elegir modo de transporte. A su vez, se compararán los valores encontrados con referentes en uso y diferentes enfoques de modelación.

1.1 OBJETIVOS DE LA TESIS

En la presente investigación se buscó desarrollar un proceso metodológico para estimar el valor subjetivo del tiempo en el contexto del proyecto Metro de Bogotá utilizando modelos de elección discreta.

Entre los objetivos de este trabajo se encuentran:

- Conceptualizar el proceso de elección considerando la percepción de cada uno de los individuos al momento de elegir el modo de transporte.
- Realizar una revisión bibliográfica de las técnicas usadas para la estimación del valor subjetivo del tiempo con modelos econométricos de elección discreta.
- Aplicar los modelos propuestos a casos de elección entre alternativas de transporte, utilizando datos reales mediante los cuales se busca evaluar el valor subjetivo del tiempo en el contexto del proyecto metro de Bogotá.
- Considerar variaciones individuales, sistemáticas y aleatorias en la estimación del valor subjetivo del tiempo.
- Diseñar los instrumentos para la toma de información necesaria que permita estimar modelos de elección discreta en el contexto evaluado.

1.2 HIPÓTESIS

La hipótesis básica en la cual se desarrolla esta investigación considera que la valoración del tiempo no es homogénea en la población, y que está fuertemente influenciada por las variables socioeconómicas. Adicionalmente, se desea probar que es diferente la percepción que se tienen de las distintas componentes del tiempo (viaje en vehículo, acceso y transbordos).

1.3 CONTRIBUCIONES

Este trabajo de investigación describe un desarrollo metodológico para la estimación del valor subjetivo del tiempo en el contexto del transporte privado y público, que es aplicado al proyecto metro de la ciudad de Bogotá. Teniendo en cuenta los antecedentes que se tenían sobre el tema, el presente trabajo aporta un desarrollo metodológico para la estimación de las tasas marginales de sustitución a nivel del individuo, por medio de técnicas econométricas como son los modelos de elección discreta.

En el caso de los modelos tipo logit multinomial o logit simple se presentan análisis de variaciones sistemáticas, aleatorias y combinaciones de las dos, para la estimación del valor subjetivo del tiempo de viaje y sus componentes.

Para la evaluación de los parámetros en los modelos se realiza una estimación en dos enfoques diferentes a saber, la estimación clásica o por máxima verosimilitud y las técnicas bayesianas para estimación de parámetros individuales.

Adicionalmente, en los modelos se involucra una variable latente como es el confort o comodidad, la cual es muy útil para la estimación de la demanda en el sistema Metro, como un nuevo medio de transporte en la ciudad.

Esta podría considerarse como una de las primeras investigaciones que se desarrolla sobre este tema en la ciudad de Bogotá, buscando determinar cuáles son las preocupaciones que inciden de manera significativa en la transferencia de un individuo cautivo de un cierto modo hacia otro.

1.4 ESTRUCTURA DE LA TESIS

Este documento está organizado como se indica a continuación:

- Capítulo 2. Aquí se presenta el marco conceptual, y la teoría que sustenta el desarrollo metodológico de los capítulos siguientes, incluyendo la teoría de utilidad aleatoria y los modelos de elección discreta. Adicionalmente se desarrollan las estimaciones del modelo logit mixto encontrados en la literatura, los cuales se aplicarán en el presente trabajo.
- Capítulo 3. Se enfoca en el diseño del instrumento y su aplicación, y la metodología utilizada para la toma de información.
- Capítulo 4. Se estudian los diferentes modelos logit obtenidos luego de procesar la información y los valores subjetivos del tiempo según las metodologías desarrolladas en el marco conceptual.
- Capítulo 5. Finalmente, este capítulo resume las conclusiones y aportes más importantes derivadas del presente trabajo de investigación.

2 MARCO CONCEPTUAL ECONÓMICO

A continuación se presenta un marco conceptual sobre los términos y definiciones que se estarán empleando a medida que avanza este documento.

2.1 EL COMPORTAMIENTO DEL CONSUMIDOR RACIONAL

En este contexto de análisis se parte de la teoría del consumidor racional (*homo economicus*), donde se postula que un individuo actúa de forma racional siempre escogiendo el camino de acción que maximice su utilidad neta personal, sujeto a ciertas restricciones ambientales, las cuales pueden ser legales, sociales, físicas o presupuestarias.

De esta manera el problema que enfrenta el consumidor puede resumirse como un modelo de optimización donde el objetivo es maximizar la función de utilidad sujeto a ciertas restricciones.

El grupo de alternativas factibles corresponde al conjunto de todas las combinaciones que el consumidor puede elegir, así si $p = (p_1, p_2, \dots, p_k)$ es el vector de precios del conjunto de bienes X , e I el ingreso disponible por el consumidor q , el conjunto de las combinaciones posibles está dado por:

$$A(q) = \{x \in X : p \cdot x \leq I\} \quad (1)$$

así el problema que enfrenta el consumidor puede expresarse como:

$$\begin{aligned} & \text{Max} U(x) \\ & \text{s.a. } p \cdot x \leq I \\ & x \in X \end{aligned} \quad (2)$$

Cuando las alternativas son modos de transporte o en su defecto rutas, las restricciones pueden ser el tiempo, el costo y el ingreso, por lo cual el problema de maximización vendría dado por la siguiente expresión:

$$\begin{aligned} & \max_m \left\langle \max_x U(X, q_m, tv_m, t_x) / m \in M \right\rangle \\ & \text{S.A. } P_i X_i + C_m = I \\ & tv_m + \sum_{\tau} t_{\tau m} = T \\ & m * n = 0, \forall m \neq n, m \in M, n \in M \end{aligned} \quad (3)$$

donde P_i y X_i son el precio y la cantidad del bien i , C_m es el costo de usar el modo o alternativa m , I representa el ingreso y M es el conjunto de alternativas. A su vez, t_v es el tiempo de viaje y t_{τ} son los distintos usos τ que los individuos le dan al tiempo si eligen la alternativa m . Aquí se tienen dos restricciones: la primera es de presupuesto y la segunda de tiempo, en el sentido que el tiempo disponible para realizar las actividades y para viajar

no puede exceder cierto valor (por ejemplo, en un día las actividades no pueden exceder las 24 horas disponibles).

Ahora bien, la noción de utilidad es un artificio teórico muy conveniente, consistente en asociar un índice al nivel de satisfacción relativo correspondiente a consumir un bien particular, por unidad de tiempo. Los bienes *per se* no producen utilidad, ésta se deriva de los servicios asociados a ellos. El nivel de utilidad obtenido es una combinación adecuadamente definida de los atributos, ponderados por la importancia relativa de cada uno en cuanto a contribuir a la utilidad total del bien en particular.

2.2 LA TEORIA DE LA UTILIDAD ALEATORIA

La teoría de la utilidad aleatoria (Domencich & McFadden, 1975) (Williams, 1977) postula que:

- Los individuos pertenecen a cierta población homogénea Q , actúan en forma racional y tienen información perfecta.
- Existe un conjunto $A = \{A_1, \dots, A_i, \dots, A_J\}$ de alternativas disponibles y un conjunto X de vectores, que incluyen las características de los individuos y de las alternativas. Así el conjunto de alternativas disponibles para un individuo q en particular es $A(q)$ y va a tener asociado un conjunto de atributos.
- Cada alternativa $A_i \in A$ tiene asociada una utilidad U_{iq} para el individuo q . El modelador, al ser un observador, no posee información completa de todos los factores considerados por los individuos al realizar su elección; por tanto, supone que esta utilidad se puede representar por dos componentes:
 - Un componente determinístico V_{iq} que es función de los atributos medidos u observados X , llamado utilidad sistemática o representativa. Es una práctica común utilizar una función aditiva y lineal en los parámetros $V_{iq} = \sum_{k=1}^K \theta_{ikq} X_{ikq}$ donde X_{ikq} son funciones numéricas que representan el valor del atributo k de la alternativa A_i para el individuo q . Se supone que los parámetros θ son constantes para todos los individuos pudiendo variar entre alternativas; éstos se obtienen mediante un proceso de estimación, donde las observaciones de las elecciones realizadas por una muestra de individuos se ajusta al modelo, típicamente mediante el método de máxima verosimilitud.
 - Una parte aleatoria ε_{iq} , que refleja la idiosincrasia y gustos particulares de cada individuo, además de errores de medición y observación por parte del modelador. En general se supone que los residuos ε son variables aleatorias con media cero y una distribución de probabilidad a especificar.

De esta manera se tiene:

$$U_{iq} = V_{iq} + \varepsilon_{iq} \quad (4)$$

El individuo q escoge la alternativa que maximiza su utilidad, esto es, escoge A_i si y sólo si se cumple (Ortúzar & Willumsen, 2001):

$$U_{iq} \geq U_{jq} \quad \forall A_j \in A(q) \quad (5)$$

expresado en las componentes,

$$V_{iq} - V_{jq} \geq \varepsilon_{jq} - \varepsilon_{iq} \quad \forall A_j \in A(q) \quad (6)$$

como no se conoce $\varepsilon_{jq} - \varepsilon_{iq}$, no es posible determinar si se cumple la relación anterior, por lo tanto se asignan probabilidades. Así la probabilidad de que el individuo q escoja la alternativa i es:

$$P_{iq} = \text{Prob}\{\varepsilon_{jq} \leq \varepsilon_{iq} + (V_{iq} - V_{jq}), \forall A_j \in A(q)\} \quad (7)$$

si $f(\varepsilon) = f(\varepsilon_1, \dots, \varepsilon_N)$, es la función distribución de las variables aleatorias:

$$P_{iq} = \int_{\varepsilon_{iq}=-\infty}^{\infty} \int_{\varepsilon_{1q}=-\infty}^{V_{iq}-V_{1q}+\varepsilon_{iq}} \dots \int_{\varepsilon_{Nq}=-\infty}^{V_{iq}-V_{Nq}+\varepsilon_{iq}} f(\varepsilon_{1q}, \varepsilon_{2q}, \dots, \varepsilon_{Nq}) d\varepsilon_{1q} \dots d\varepsilon_{Nq} \quad (8)$$

por lo tanto la probabilidad de elección es una integral multidimensional sobre la densidad de la porción no observada de la utilidad, diferentes modelos son obtenidos dependiendo de los supuestos sobre la distribución de los ε .

2.3 LOS MODELOS DE ELECCIÓN DISCRETA

Los modelos de elección discreta han sido desarrollados para describir el comportamiento de un consumidor o usuario ante la escogencia entre un grupo de opciones mutuamente excluyentes. En general postulan que la probabilidad de que un individuo escoja una cierta opción, es una función de sus características socioeconómicas y de lo atractiva que resulte la alternativa en cuestión en comparación a las demás (Ortúzar & Willumsen, 2001).

En general postulan que:

- La probabilidad de que un individuo escoja una cierta opción, es una función de sus características socioeconómicas y de lo atractiva que resulte la alternativa en cuestión en comparación a las demás (Ortúzar & Willumsen, 2001).

Usualmente se supone una decisión donde existe un número de alternativas finitas y sólo es posible elegir una de ellas. El consumidor, basándose en los principios de la teoría de utilidad aleatoria, elige la alternativa que le brinde mayor utilidad. En modelos de elección discreta, ha llegado a ser la norma basarse en este enfoque.

Algunas de las propiedades de estos modelos son:

- A diferencia de los modelos agregados, que generalmente se calibran con datos que se han agrupado de alguna forma, los modelos de elección discreta se calibran usando observaciones de la conducta de los individuos como datos de entrada. Dado esto, esta clase de modelos pueden ser más eficientes en el uso de información y poseen menor probabilidad de sesgos debidos a correlaciones entre unidades agregadas.
- Los modelos desagregados, como también se les conoce en la literatura, son probabilísticos.
- A las variables explicativas que se incluyen en los modelos, se les calibran coeficientes en forma explícita. Estos coeficientes permiten una representación más flexible de las variables de política que sean consideradas relevantes para el estudio. A su vez, los coeficientes tiene una interpretación directa en términos de determinar la importancia relativa a cada atributo en la elección.

Los axiomas de la modelación desagregada del comportamiento de la demanda son:

- Los individuos representan la unidad básica en la toma de decisiones
- Los individuos seleccionan la alternativa que más desee o le sea más útil entre todas las disponibles. Esta elección depende de los atributos de la alternativa y las características socioeconómicas del individuo.

La metodología del análisis desagregado de demanda se basa en especificar una forma funcional y paramétrica para la parte representativa V de la función de utilidad y a su vez una función de distribución del término aleatorio. Para estimar los modelos, el procedimiento estadístico utilizado regularmente es máxima verosimilitud.

Dependiendo de la distribución del término error, mostrado en la expresión de la teoría mencionada y basándose en ella misma, se dará origen a distintos modelos probabilísticos de elección discreta. Los residuos ε_{iq} son variables aleatorias con media cero, que determinarán distintos modelos probabilísticos dependiendo de la distribución estadística que se considere. En la mayoría de las aplicaciones empíricas, las distribuciones usadas han sido las de Valor Extremo Generalizados, que dan lugar a los modelos de tipo logit, y la Normal, que da lugar a los modelos de tipo Probit. En la aplicación empírica de este trabajo se han utilizado únicamente modelos tipo logit. Cuando se considera un modelo con variación sistemática de los gustos donde el término de error ε_{iq} se distribuye idéntica e independientemente (iid) Gumbel se obtiene el denominado logit multinomial (MNL) o logit simple (McFadden D. , 1974). Por el contrario, cuando se supone variación aleatoria de los gustos y la misma distribución para los ε_{iq} , se alcanza una especificación de tipo logit mixto (ML) o de parámetros aleatorios donde la utilidad de la alternativa i viene dada por la siguiente expresión:

$$U_{iq} = (\theta + \mu_q) x_{iq} + \varepsilon_{iq} = \theta x_{iq} + \mu_q x_{iq} + \varepsilon_{iq} \quad (9)$$

que es un caso particular de la expresión más general de un modelo logit mixto:

$$U_{iq} = V_{iq} + \eta_{iq} + \varepsilon_{iq} \quad (10)$$

Donde $\varepsilon_{iq} \sim \text{Gumbel}(0, \sigma^2)$ y $\eta_{iq} \sim f(\eta/\theta)$, siendo f una función de densidad general y η los parámetros fijos que caracterizan su distribución en la población.

Por otro lado, se tiene el modelo Logit Jerárquico o Anidado (NL) (Williams, 1977), el cual fue desarrollado como una extensión del MNL, en el que se considera una partición disjunta de las alternativas en grupos de tal forma que las alternativas incluidas en cada grupo presentan correlación en sus términos de error; siguiendo estos también una distribución Gumbel. Este modelo es apropiado cuando las opciones a las que se enfrenta el individuo pueden ser agrupadas en nidos de tal manera que la propiedad de independencia de alternativas irrelevantes (IIA) del MNL (Train K., Discrete Choice Methods with Simulation, 2009), se mantenga para las alternativas dentro del mismo nido y no se cumpla para las opciones que pertenecen a diferentes grupos. Los modelos MNL y NL no pueden manejar el problema de variedad en los gustos, al trabajar con parámetros fijos para toda la población. No son los modelos apropiados cuando existe heterocedasticidad entre individuos o atributos.

2.3.1 Modelo Logit Multinomial

Es el modelo de elección discreta más ampliamente utilizado; su popularidad se debe al hecho de que la integral de la fórmula de probabilidad de elección, ecuación (8), es cerrada. Se obtiene a partir de suponer que los términos de error siguen una distribución iid Gumbel¹, con media cero y varianza σ^2 . Esto quiere decir que los términos no están correlacionados y poseen la misma varianza tanto a nivel de alternativas como individuos.

Así, la probabilidad de que un individuo q escoja la alternativa i , está dada por:

$$P_{iq} = \frac{\exp(\beta \cdot V_{iq})}{\sum_{A_j \in A(q)} \exp(\beta \cdot V_{jq})} \quad (11)$$

donde β es un factor de escala relacionado con la varianza del término de error,

$\beta = \frac{\pi}{\sqrt{6}\sigma}$ El factor de escala generalmente no es identificable, por lo que es necesario fijar

este valor (escalamiento de la matriz de varianza covarianza), usualmente se asume

$\beta = 1$ lo que implica que $\sigma^2 = \frac{\pi^2}{6}$ (Ortúzar & Willumsen, 2001).

Por otro lado, se puede decir que el modelo MNL exhibe la propiedad de independencia de alternativas irrelevantes y puede no ser adecuado cuando las alternativas están correlacionadas, tal como se mencionó previamente. Esto implica que la proporción entre

¹ También conocida como Valor Extremo, Valor Extremo Tipo I o Weibull.

dos alternativas no cambia por la inclusión o exclusión de otra alternativa. La propiedad IIA (independencia de alternativas irrelevantes) implica que la probabilidad de escoger una alternativa A en vez de otra, debe ser independiente de la presencia o ausencia de una tercera alternativa no escogida.

Algunas desventajas del MNL pueden resumirse así (Train K. , Discrete Choice Methods with Simulation, 2009)

- Los modelos logit pueden representar variaciones sistemáticas en los gustos pero variaciones aleatorias.
- Estos modelos no pueden representar situaciones donde los factores no observados se correlacionan en el tiempo.

Dado que los modelos MNL no permiten analizar los problemas cuando existe heterogeneidad en las preferencias o gustos de los individuos, las aproximaciones que se utilizan para captar la heterogeneidad de las preferencias de los individuos son dos. La primera consiste en usar una especificación donde el parámetro de cada atributo es una función de las características socioeconómicas observadas de los individuos (edad, sexo, ingreso, posesión de vehículo,...). Esto permite recoger la variación sistemática de los gustos e identificar cuáles son las fuentes de variación de las distintas disposiciones a pagar. A partir de este enfoque se construyen los modelos con variaciones sistemáticas, propuestos por (Fowkes & Wardman, 1988), en los que se incluye la variación heterogénea en los gustos, asumiendo que los parámetros de los diferentes atributos dependen de las características de los individuos de una manera determinística.

La segunda aproximación recoge la heterogeneidad aleatoria a través de la especificación de un modelo logit mixto o de parámetros aleatorios, planteada por Train (2003). Estas dos aproximaciones pueden emplearse simultáneamente en un mismo modelo que hace posible captar, además de la variación sistemática de las preferencias, la heterogeneidad no observada. Estos modelos se presentan posteriormente bajo los títulos de Modelos con Variaciones Sistemáticas (MVS), Modelos Mixed Logit (ML) y Modelos con Variaciones Combinadas (MVC).

Ahora bien, para estimar este tipo de modelos, se emplea el método de máxima verosimilitud para determinar los coeficientes θ que aparecen en la definición de V (Ortúzar & Willumsen, 2001). En el caso de que V sea de forma lineal en los parámetros, la idea detrás del método de máxima verosimilitud es la siguiente: aunque una muestra pueda provenir (o ser generada) de distintas poblaciones, existe una para la cual hay mayor probabilidad que esto ocurra. Así, los estimadores máximo-verosímiles son el conjunto de parámetros que generarían más a menudo la muestra observada. Para ilustrar esta idea, consideremos una muestra de n observaciones de alguna variable z , provenientes de una población caracterizada por un parámetro desconocido θ (pudiendo ser éste la media, varianza, etc.) y denotémoslas por $(z_1, z_2, \dots, z_n)$. Ya que z es una variable aleatoria, tiene una función densidad asociada $f(z / \theta)$, puesto que depende de

los valores de θ . Si todos los valores de z en la muestra son independientes, podemos escribir la función de densidad como:

$$f(z_1, z_2, \dots, z_n / \theta) = f(z_1 / \theta) \cdot f(z_2 / \theta) \cdot \dots \cdot f(z_n / \theta) \quad (12)$$

La interpretación de esta función conjunta es que los z son variables y θ es fijo. Si ahora invertimos el proceso y suponemos que los z son conocidos y fijos, y θ una variable, se puede interpretar a la expresión anterior como una función de verosimilitud y no de densidad conjunta. Si se maximiza esa función respecto a θ , el resultado se llama estimador de máxima verosimilitud, ya que es el valor de θ que tiene mayor probabilidad de haber generado la muestra observada; esta idea puede ser extendida a varios parámetros.

Para maximizar la función de verosimilitud, se procede como es usual, esto es, derivando parcialmente respecto de los θ e igualando a cero. Normalmente se maximiza la función logaritmo de la verosimilitud puesto que es más manejable y tiene el mismo máximo. Para el MNL la función de verosimilitud se comporta bien y tiene un máximo único si la utilidad es lineal en los parámetros. Si todas las observaciones de la muestra son independientes, la expresión para la función de verosimilitud es sencillamente la multiplicación de las probabilidades de escoger, de acuerdo al modelo, la alternativa efectivamente seleccionada por cada individuo.

La función de verosimilitud puede escribirse, en general, como:

$$L = L(\theta) = \prod_{q=1}^Q \prod_{A_j \in \underline{A}(q)} P_{jq}^{\epsilon_{jq}} \quad (13)$$

La función a maximizar es: $\ln L(\theta) = l(\theta) = \sum_{q=1}^Q \sum_{A_j \in \underline{A}(q)} g_{jq} \ln P_{jq}$ donde:

$$g_{jq} = \begin{cases} 1 & \text{si } q \text{ escoge } A_j \\ 0 & \text{en otros casos} \end{cases}$$

Maximizando $l(\theta)$ es posible encontrar un conjunto de estimadores máximo verosímiles $\hat{\underline{\theta}}$ que se distribuyen asintóticamente Normal $(\underline{\theta}; \underline{V})$ en que :

$$\underline{V} = - \left\{ E \left(\frac{\partial^2 l(\theta)}{\partial \underline{\theta}^2} \right) \right\}^{-1} \quad (14)$$

Además, $-2l(\theta)$ distribuye asintóticamente χ^2 con Q grados de libertad.

2.3.2 Modelos con Variaciones Sistemáticas

La manera tradicional de captar la heterogeneidad de las preferencias ha consistido en la introducción de interacciones entre los atributos de las alternativas (tiempo de viaje, costo, frecuencia, etc.) y las características socioeconómicas de los individuos tales como sexo, edad, nivel de ingreso, entre otros (Train K. , Recreation Demand Models with Taste Differences Over People, 1998); (Morey & Rossman, 2002). De esta manera, el parámetro de cada atributo θ_k es una función de las características socioeconómicas observadas de los individuos. Este método permite detectar únicamente la variación sistemática de los gustos.

La forma de la función de utilidad en este caso sería:

$$V_{iq} = \left(\alpha_0 + \sum_l \alpha_l s_{lq} \right) t_{iq} + \left(\theta_0 + \sum_l \theta_l s_{lq} \right) c_{iq} + \left(\gamma_0 + \sum_l \gamma_l s_{lq} \right) f_{iq} \quad (i=1,2,\dots) \quad (15)$$

Donde t es tiempo, c es costo y f es frecuencia y las variables s_{lq} son variables dummy que representan las características socioeconómicas l del individuo q , tal como lo explican en (Ortúzar & Willumsen, 2001). Esta ecuación establece que dadas unas características de los diferentes individuos, es posible obtener diferentes coeficientes para un mismo atributo. Es claro que esta formulación de la utilidad sistemática no implica que los parámetros están distribuidos aleatoriamente en la población, por el contrario, asume que los parámetros dependen, en una manera determinística, de las características de los individuos.

2.4 MODELO LOGIT MIXTO

El modelo logit mixto (Train K. , Discrete Choice Methods with Simulation, 2009) es en la actualidad una de las herramientas más poderosa en la estimación de modelos de elección discreta, entre sus características principales está la capacidad de aproximar cualquier modelo de utilidad aleatoria (McFadden & Train, 2000) y que además no tiene las limitaciones de un modelo logit simple.

Estos modelos tienen una amplia trayectoria, ya que los primeros estudios han sido desarrollados desde los años 70's con la investigación realizada por Cardell y Reddy (1977) en la cual introducen el concepto de variación en los gustos de los individuos en un modelo logit simple, sin embargo, la poca capacidad computacional de la época para estimar modelos con especificaciones tan complejas como ésta no permitió su utilización tan común como lo es ahora.

Unos años después el trabajo paralelo realizado por Boyd y Mellman (1980) y Cardell y Dunbar (1980) sobre la reducción del consumo de gasolina en automóviles presenta la primera aplicación de un modelo logit mixto, donde el gran proceso computacional que es inherente a un logit mixto se lleva a cabo una sola vez, para el mercado como un todo y no para cada individuo, facilitando la estimación.

Una década después las investigaciones realizadas contaban con mejores herramientas computacionales y se fueron presentando trabajos como el realizado por Ben Akiva y

Bolduc (1996) donde encuentran que el modelo mixed logit se comporta bien, obviando el ruido y dejando intacto el término Gumbel, por otra parte, el trabajo realizado por McFadden y Train (2000) el cual ha sido una gran contribución para este tipo de modelos prueba que cualquier modelo de comportamiento de utilidad aleatoria puede ser representando por un logit mixto y presenta pruebas de especificación fáciles de implementar, algunas de las cuales se proponen en una investigación realizada por Chesher y Santos (2003). Adicionalmente a estos trabajos representativos, se han desarrollado un gran número de investigaciones sobre este tema donde se destaca algunas como Mehndiratta and Hansen (1997), Train, McFadden, y Goett (1987), Revelt y Train (1998), Srinivasan and Mahmassani (2005), Greene (2000), Bhat (2001), entre otros.

El logit mixto se define en la base de la forma funcional de su probabilidad de elección (Train K. , Discrete Choice Methods with Simulation, 2009), de esta forma, las probabilidades de elección de un logit mixto son las integrales de probabilidades estándar de un logit multinomial pero agregando parámetros de densidad, que pueden expresarse de la siguiente manera:

$$P_{ni} = \int L_{ni}(\beta) f(\beta) d\beta \quad (16)$$

Donde $L_{ni}(\beta)$ es la probabilidad logit evaluada en el parámetro β :

$$L_{ni}(\beta) = \frac{e^{V_{ni}(\beta)}}{\sum_{j=1}^J e^{V_{nj}(\beta)}} \quad (17)$$

Y $f(\beta)$ es la función de densidad. $V_{ni}(\beta)$ es la porción observada de la función de utilidad que depende de los parámetros β .

En este orden de ideas, un modelo logit mixto se puede definir como una mezcla de la función logit evaluada para diferentes valores de β 's donde $f(\beta)$ es la distribución mixta. La formula más utilizada para este modelo es donde $f(\beta)$ es especificada como continua, y puede tomar distintas distribuciones de densidad como son la normal, lognormal, entre otras.

2.4.1 Formulación del Modelo

La probabilidad de un modelo logit mixto se deriva generalmente con la aplicación de coeficientes aleatorios (Train, 2009). El individuo n enfrenta la elección de J alternativas, en una situación de elección t , donde la utilidad del individuo n de la alternativa j se especifica como:

$$U_{njt} = \beta_n' X_{njt} + \varepsilon_{njt} \quad (18)$$

Donde X_{nj} son las variables observadas que se relacionan con las alternativas y el individuo, β_n es un vector de coeficientes de estas variables por cada individuo n que representa el gusto individual, el cual tiene media b y una matriz de covarianza W , y ε_{njt} es

un término aleatorio que es iid valor extremo. Los coeficientes varían por individuo en la población con densidad $f(\beta)$. Esta densidad es una función de los parámetros β , esta especificación es la misma que para un logit simple siendo que los parámetros β cambian con el individuo. En este caso el modelador puede observar los x_{nj} 's pero no los β_n ni las condiciones o los ε_{nj} .

La probabilidad condicional viene dada por:

$$L_{ni}(\beta_n) = \frac{e^{\beta'x_{ni}}}{\sum_j e^{\beta'x_{nj}}} \quad (19)$$

Por lo tanto la probabilidad de elección no condicional es entonces la integral de $L_{ni}(\beta_n)$ sobre todas las posibles variables de β_n :

$$P_{ni} = \int \frac{e^{\beta'x_{ni}}}{\sum_j e^{\beta'x_{nj}}} f(\beta) d\beta \quad (20)$$

La cual es la probabilidad logit mixta.

La formulación del modelo logit multinomial tiene importantes limitaciones (Sillano & Ortúzar, 2005) principalmente por la rigidez de su estructura de error (una matriz de covarianza diagonal con varianzas iguales). Para superar estas limitaciones la expresión de utilidad aleatoria para U_{njt} se puede reescribir en una forma más general como:

$$U_{njt} = \beta_n X_{njt} + \Omega_{njt} Y_{njt} + \varepsilon_{njt} \quad (21)$$

Donde β_n es ahora el vector de parámetros de utilidad marginal para el individuo n ; Y_{njt} es un vector de cargas que describe los componentes del error de acuerdo a la estructura del modelo deseado, y Ω_{njt} es un vector de componentes estocásticos los cuales siguen una distribución dada por el modelador, con media cero y varianza desconocida. Los demás términos son los mismos que en la ecuación (18).

Partiendo de que la utilidad que un individuo n puede obtener de la opción A_j en t situaciones de elección, viene dada por la expresión mostrada en la ecuación (18), donde β_n tiene media b y una matriz de covarianza W . La varianza para cada periodo t es igual a $\sigma_t^2 = \pi^2 / (6\mu_t^2)$, $t = 1, \dots, T$, donde μ_t es un factor de escala para ese periodo, pero puede ser igual para todos los periodos (Godoy & Ortúzar, 2004).

Para cada uno de los individuos el β_n puede expresarse en dos componentes de la siguiente manera:

$$\beta_n = b + \eta_n \quad (22)$$

Donde b es la media poblacional y η_n son las desviaciones individuales que representan las preferencias de cada individuo en relación con los gustos medios en la población.

Al ser variables estos gustos en la población de acuerdo con una función de densidad $f(\beta/b, W)$ (Train K. , Discrete Choice Methods with Simulation, 2009), como ε_{int} es independiente de cada situación de elección t , la probabilidad que un individuo n haga una secuencia de elecciones, condicionado a β viene dada por:

$$P(y_n / X_n, \beta) = \prod_{t=1}^T \left(\frac{\exp(\mu_t(\beta' \cdot X_{y_{nt}}))}{\sum_{j=1}^J \exp(\mu_t(\beta' \cdot X_{jnt}))} \right) \quad (23)$$

Donde y_{nt} identifica la opción tomada por el individuo n en el experimento t , y y_n es el vector que identifica la secuencia de elecciones para cada individuo. Si la utilidad es lineal en β , entonces $V_{njt}(\beta) = \beta_n' X_{njt}$. Finalmente, la probabilidad de observar una secuencia de elecciones se obtiene integrando la probabilidad condicional sobre todos los posibles valores de β usando su distribución en la población:

$$P(y_n / X_n, b, W) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} P(y_n / X_n, \beta) f(\beta/b, W) d\beta \quad (24)$$

Para la estimación de esta clase de modelos se han desarrollado varios procedimientos, siendo los más conocidos el método de máxima verosimilitud o llamado “clásico”, y el método bayesiano.

2.4.2 Estimación Clásica

Para la estimación de este tipo de modelos se ha usado comúnmente el procedimiento de máxima verosimilitud (Train K. , Discrete Choice Methods with Simulation, 2009). La utilidad está definida por la ecuación $U_{njt} = \beta_n' X_{njt} + \varepsilon_{njt}$ (18), donde los coeficientes β_n distribuyen con una densidad $f(\beta|\theta)$, donde θ se refieren colectivamente a los parámetros de esta distribución (como su media y su covarianza). El modelador especifica la forma funcional de $f(\cdot)$ y desea estimar los parámetros θ . Teniendo las probabilidades de elección:

$$P_{ni} = \int L_{ni}(\beta) f(\beta|\theta) d\beta \quad (25)$$

Donde según lo mostrado en la ecuación (19) la probabilidad condicional viene dada por:

$$L_{ni}(\beta_n) = \frac{e^{\beta_n' x_{ni}}}{\sum_{j=1}^J e^{\beta_n' x_{nj}}}$$

Las probabilidades se aproximan por medio de simulación (Train K. , Discrete Choice Methods with Simulation, 2009) para cualquier valor de θ dado:

1. Dar un valor a β de la función $f(\beta | \theta)$ y denotarlo como β^r con el superíndice $r = 1$ refiriéndose a la primera iteración.
2. Calcular la probabilidad logit $L_{ni}(\beta^r)$ con esta primera iteración.
3. Repetir los pasos 1 y 2 varias veces y promediar los resultados. Este promedio es la probabilidad simulada:

$$\check{P}_{ni} = \frac{1}{R} \sum_{r=1}^R L_{ni}(\beta^r) \quad (26)$$

Donde R es el número de iteraciones. $\check{P}_{ni}$ es un estimador insesgado de P_{ni} por construcción, estrictamente positivo por lo tanto es posible definir el logaritmo natural de este estimador con el cual se aproxima la función de verosimilitud.

La probabilidad simulada se inserta en la función de verosimilitud para obtener la verosimilitud simulada:

$$SLL = \sum_{n=1}^N \sum_{j=1}^J d_{nj} \ln \check{P}_{nj} \quad (27)$$

Donde d_{nj} es igual a uno si n escoge j y cero en otro caso. El estimador de máxima verosimilitud simulada es el valor de θ que maximiza SLL . Usualmente, diferentes iteraciones se toman para cada observación.

Esta especificación simple trata los coeficientes que entran a la utilidad como variantes en la población pero constante en alguna situación de elección para cada individuo. Por lo tanto en el caso de datos en panel la simulación varía de la siguiente manera:

La utilidad de la alternativa j en un caso de elección t por una persona n viene dada por $U_{njt} = \beta'_n X_{njt} + \varepsilon_{njt}$ y ε_{njt} siendo un término aleatorio que es iid valor extremo en el tiempo, en los individuos y en las alternativas.

Considere una secuencia de alternativas, $i = \{i_1, \dots, i_T\}$. La probabilidad condicional de observar que el individuo n realice esa secuencia de elección, dado valores fijos de los parámetros del modelo β_n , es el producto de las funciones logit:

$$L_{ni}(\beta) = \prod_{t=1}^T \left[\frac{e^{\beta'_n x_{nit}}}{\sum_j e^{\beta'_n x_{njt}}} \right] \quad (28)$$

Como β_n es desconocido la probabilidad de elección no condicional es dada por la integral de la ecuación (31) sobre todos los valores de β . $P_{ni} = \int L_{ni}(\beta) f(\beta) d\beta$

La probabilidad se simula similar que la probabilidad con un solo periodo. Se tiene una iteración de β de su distribución. Se calcula la función logit de cada periodo, y el producto de estos logit se multiplica. Este proceso se repite para un número de iteraciones, y se promedia el resultado.

Para incrementar la precisión en la estimación se pueden usar secuencias de baja discrepancia, las más utilizadas en este campo son la secuencia de Sobol o la secuencia de Halton (Garrido & Silva, 2004) (Bliemer, Rose, & Hess, Approximation of Bayesian Efficiency in Experimental Choice Designs, 2006).

Secuencias de baja discrepancia

El concepto de secuencias de baja discrepancia está asociado a que números sucesivos son agregados en una posición tan lejos como sea posible de los puntos ya generados, evitando de esta forma la agrupación de puntos (clustering). Las secuencias son construidas en base a la idea de que cada nuevo punto sea repelido por los otros. Bajos valores de discrepancia implican alta uniformidad. En la Figura 2.1 y la Figura 2.2 se presentan 300 puntos pseudo-aleatorios y 300 puntos de secuencias de Halton respectivamente, claramente los primeros entregan una mejor cobertura del área de integración.

La discrepancia está directamente relacionada con el error de integración. Una baja discrepancia de las secuencias tiene como resultado un bajo error de integración, con errores del orden de N^{-1} en el mejor caso y tasas de convergencia más rápidas.

El problema de las secuencias de baja discrepancia en altas dimensiones se debe a la pérdida de uniformidad, ya que se obtienen secuencias altamente dependientes. En la Figura 2.3 se puede ver las secuencias de Halton para los números primos 43 y 47, que corresponden a la dimensión de integración 14 y 15 respectivamente. Es claro que las secuencias son altamente dependientes y ordenadas en un patrón fijo.

Pese a esto, investigaciones como las de Paskov (1994), Paskov y Traub (1995) y Paskov (1997), han mostrado que los métodos de Monte Carlo basado en secuencias de baja discrepancia (LDMC) pueden ser muy efectivos para problemas de integrales multidimensionales (integrales de hasta 360 dimensiones) en el área de finanzas, obteniendo resultados más cercanos a la mejor tasa de convergencia que a la convergencia del peor caso. Esto ocurre principalmente debido a que los integrandos típicamente utilizados en finanzas son funciones suaves.

La secuencia de Halton (Halton, 1960) en S dimensiones se obtiene pareando S secuencias de Van der Corput unidimensionales, basadas en S enteros primos, $r_1, r_2, \dots, r_S$ (usualmente los primeros S primos).

Figura 2.1 Dimensión 1 y 2 de secuencia pseudo-aleatorias, 300 puntos

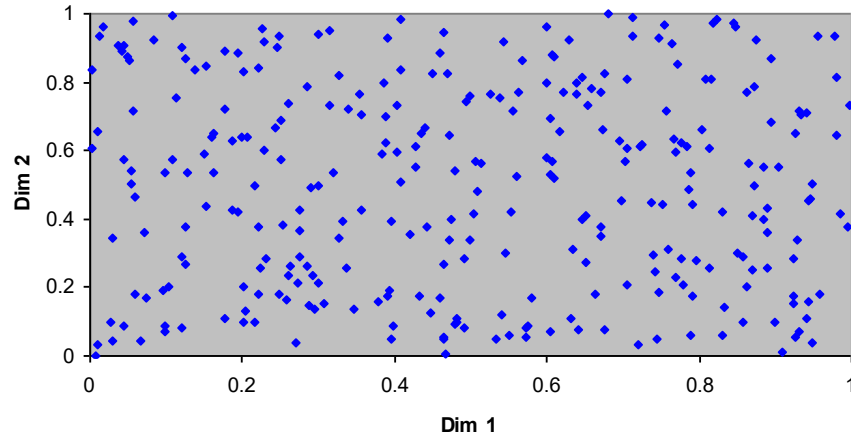


Figura 2.2 Dimensión 1 y 2 de secuencia de Halton, 300 puntos

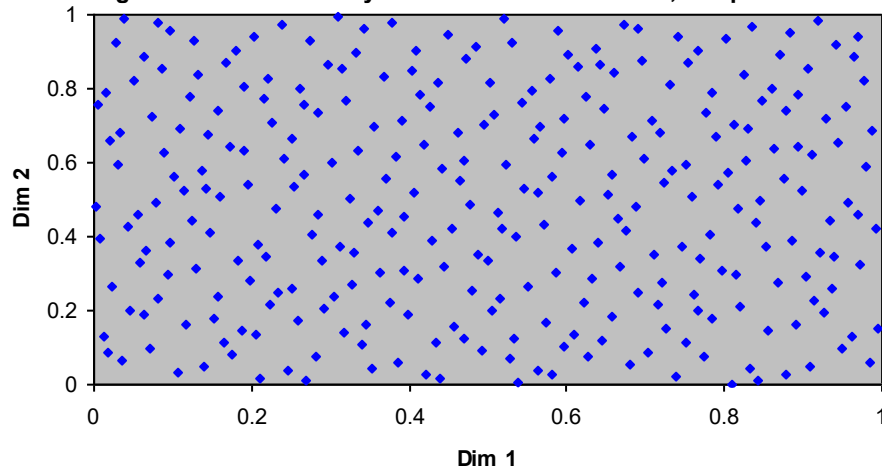
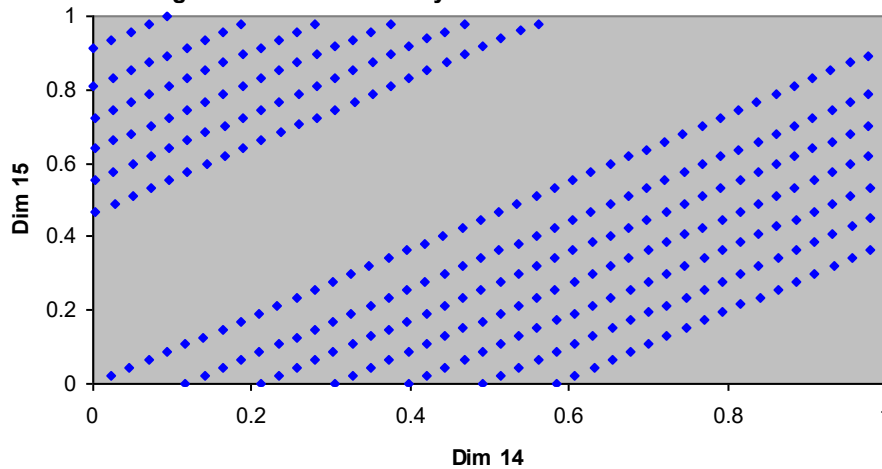


Figura 2.3 Dimensión 14 y 15 de secuencia de Halton



La secuencia correspondiente al primo r abarca ciclos de largo r de números monótonamente crecientes. Los primeros nueve números de la secuencia correspondiente a la base tres son $0, 1/3, 2/3, 1/9, 4/9, 7/9, 2/9, 5/9, 8/9$. Como se puede ver, los nuevos números que se agregan a la secuencia tienden a llenar los espacios

vacíos a la izquierda de los números anteriores, cubriendo mejor el intervalo (0,1). También se puede ver que la secuencia correspondiente al primo tres abarca ciclos de largo tres, de números monótonamente crecientes. Por esta característica, los términos iniciales de dos secuencias son altamente dependientes, al menos hasta el primer ciclo de cada secuencia; por ejemplo:

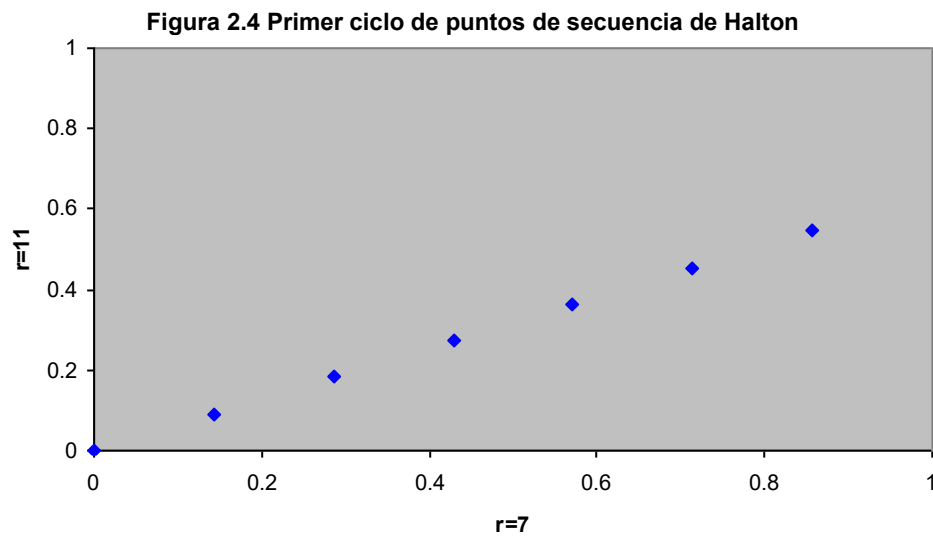
Secuencia para $r=7$

$$\left\{0, \frac{1}{7}, \frac{2}{7}, \frac{3}{7}, \frac{4}{7}, \frac{5}{7}, \frac{6}{7}, \frac{1}{49}, \frac{8}{49}, \frac{15}{49}, \frac{22}{49}, \dots\right\}$$

Secuencia para $r=11$

$$\left\{0, \frac{1}{11}, \frac{2}{11}, \frac{3}{11}, \frac{4}{11}, \frac{5}{11}, \frac{6}{11}, \frac{7}{11}, \frac{9}{11}, \frac{10}{11}, \frac{1}{121}, \frac{12}{121}, \dots\right\}$$

En la Figura 2.4 se presentan los siete primeros puntos que resultan de parear las secuencias para $r=7$ y $r=11$.



Para evitar esta dependencia se eliminan los primeros términos tomándose como base el largo del ciclo mayor.

En la presente investigación se estará utilizando secuencias Halton para la estimación de parámetros aleatorios a través de técnicas bayesianas, como se describe en el capítulo siguiente.

2.4.3 Estimación Bayesiana

La utilización de técnicas Bayesianas en la estimación de modelos logit mixtos está siendo ampliamente utilizada desde la introducción de este concepto por Allenby y Lenk (1994) y Allenby (1997) para el caso de coeficientes que distribuyen normal, estos autores demuestran de qué manera pueden estimarse los parámetros de estos modelos sin

necesidad de calcular sus probabilidades de elección (Train K. , Discrete Choice Methods with Simulation, 2009), siendo esta técnica de estimación una alternativa a la estimación por medio de máxima verosimilitud. Un aporte adicional que entregan las técnicas Bayesianas es su utilización para estimaciones de parámetros individuales, que en los casos de cálculo de valores subjetivos del tiempo (VST) son de gran ayuda para evaluación de proyectos en infraestructura, ya que se cuenta con valores desagregados de la población y no con la media poblacional como generalmente se ha trabajado.

En un problema de estimación, el procedimiento de máxima verosimilitud clásico presenta dos dificultades principales, que con las técnicas Bayesianas pueden evitarse. En primer lugar, los procedimientos Bayesianos no requieren maximización de ninguna función, ya que en casos de modelos logit mixtos o probit especialmente aquellos con distribuciones lognormales, pueden ser numéricamente difíciles de estimar. En segundo lugar, propiedades de estimación esperadas, como consistencia y eficiencia del modelo se pueden obtener de una manera más relajada con técnicas Bayesianas que con estimaciones clásicas (Train K. , Discrete Choice Methods with Simulation, 2009).

Para la estimación de los modelos logit mixtos, las técnicas bayesianas han empezado a utilizarse ampliamente (Train, 2001; Sillano and Ortúzar, 2004; Godoy and Ortúzar, 2004). En este caso se mostrará la aproximación desarrollada por Allenby (1997), implementado por Sawtooth Software (1999), y generalizado por Train (2001).

Para una muestra de N individuos, donde β_n distribuye normal en la población con media b y covarianza W , $\beta \sim N(b, W)$ la probabilidad de que el individuo n observe las elecciones, condicionadas a β , está dada por:

$$L(y_n / \beta) = \prod_t \left(\frac{e^{(\mu_t(\beta' X_{nynt}))}}{\sum_{j=1}^J e^{(\mu_t(\beta' X_{nynt^j}))}} \right) \quad (29)$$

La probabilidad no condicional es la integral $L(y_n / \beta)$ de sobre todos los valores de β :

$$L(y_n / b, W) = \int L(y_n / \beta) f(\beta / b, W) d\beta \quad (30)$$

La distribución posterior de b y W es, por definición:

$$K(b, W | Y) \propto \prod_n L(y_n | b, W) k(b, W) \quad (31)$$

donde $k(b, W)$ es un valor a priori de b y W .

Es posible iterar directamente con el algoritmo de Metropolis-Hastings (MH), el cual sería computacionalmente muy lento. Sin embargo, la probabilidad de elección $L(y_n / b, W)$ es una integral no cerrada y puede ser aproximada por simulación, cada iteración del

algoritmo MH requerirá simular $L(y_n / b, W)$ para cada n . Las iteraciones pueden ser más fácil y rápidas si cada β_n se considera un parámetro a través de b y W .

En la investigación realizada por Train (2001) se discute como las medias posteriores de la estimación Bayesiana puede ser analizada desde una perspectiva clásica. Esto gracias al teorema de Bernstein-von Mises, el cual dice que, asintóticamente, la distribución posterior de un estimador bayesiano converge a una distribución normal la cual es la misma distribución asintótica del estimador de máxima verosimilitud. Esto significa que el análisis estadístico clásico (por ejemplo, la construcción de los test estadísticos t para analizar la significancia de los parámetros) puede ser desarrollado por estimadores bayesianos sin comprometer la interpretación de los resultados.

2.5 VALOR SUBJETIVO DEL TIEMPO

2.5.1 Antecedentes

La respuesta a la pregunta ¿el tiempo de viaje tiene un valor? ha sido tema de discusión a través de los años, y principalmente el hecho de saber qué valor se le puede dar a este tiempo. Diferentes autores de la literatura en el mundo han debatido sobre este tema simplemente porque los ahorros en el tiempo son el beneficio más importante de proyectos de mejora en el transporte (Fernandez, 1997).

La mayoría de los estudios distinguen entre evaluación de valores del tiempo y valores del tiempo subjetivo. El último corresponde a, por ejemplo, el valor del parámetro asociado con el tiempo de viaje en vehículo en las funciones generalizadas de costo y el cual debe haber sido derivado estimando un modelo de demanda con datos empíricos. El valor de evaluación es el usado, como el nombre lo indica (Becker, 1965), para comparar esquemas alternativos los cuales producen diferentes niveles de tiempo y otras fuentes de ahorro. Se discute, por lo tanto, que el valor subjetivo del tiempo refleja mayormente la habilidad del viajero a pagar y no el valor intrínseco del ahorro del tiempo particular. Por esto muy a menudo el valor del tiempo usado para propósitos de evaluación es un valor *equitativo*, tomado como el mismo para todos los viajeros, independientemente de su edad o grupo socioeconómico (Ortúzar & Willumsen, 2001).

Algunos modelos desarrollados en diversos contextos tratan el tema y permiten un entendimiento global y profundo en cuanto a la valoración de distintas actividades, como el tiempo de viaje y de trabajo, mediante la solución de problemas microeconómicos de maximización de utilidad que además de considerar las variables precio y cantidad que determinan el consumo, incluyen también la variable tiempo e implícitamente el espacio.

Entre los primeros desarrollos se destaca el trabajo realizado por (Becker, 1965), quien introdujo el concepto de valor del tiempo en el modelo clásico. Este trabajo propone un modelo basado en la incorporación del tiempo en la función de utilidad, en la cual ésta depende del consumo de bienes finales Z_i que no pueden comprarse y cuya producción requiere la adquisición de otros bienes, tales como el tiempo, T_i y ciertos bienes del mercado X_i .

Por otra parte el modelo de (De Serpa, 1971) propone una modelación analítica del valor del tiempo de viaje, el cual similarmente al modelo clásico del consumidor se plantea como un problema de optimización cuya función objetivo consiste en maximizar la utilidad directa individual. Sin embargo, a diferencia del modelo clásico, usa una formulación de la función utilidad en la que se incluyen variables que representan las cantidades de tiempo que el individuo racional dedica a diferentes actividades (trabajo, ocio, viajes, recreación, entre otros).

En un desarrollo posterior, Train y McFadden (1978) trataron por primera vez en forma rigurosa el problema de especificación de las funciones de utilidad para el caso de modelos de demanda de elección discreta. Ellos plantean un problema de consumo individual para el caso de la elección de modos de transporte en los viajes de trabajo, en el que la utilidad directa individual tiene la expresión $U(G,L)$, donde G representa el gasto en el consumo de un bien generalizado x , y L representa la cantidad de tiempo dedicado al ocio. Este trabajo proporciona un soporte microeconómico para la utilización de modelos de elección discreta en la evaluación del tiempo de viaje, ya que proponen al tiempo como un bien discreto y una resolución del problema de una manera secuencial.

Bates y Roberts (1986) siguiendo los mismos planteamientos anteriores desarrollan un modelo en el que se supone que el tiempo de trabajo no es una variable que el individuo pueda elegir libremente, sino que es una constante externa (una variable exógena). En tal caso el tiempo de trabajo t_w desaparece del planteamiento del problema del consumidor y se supone que el ingreso total monetario disponible por el individuo (Y) es un dato fijo. Consecuentemente, el tiempo total disponible por el individuo (T^*) es igual a $T^* = T - t_w$, donde t_w el tiempo dedicado al trabajo y T es 24 horas menos el tiempo requerido para las necesidades biológicas básicas. Un enfoque similar es propuesto por Jara-Díaz et al. (1988), quienes usando el mismo supuesto que Bates y Roberts en cuanto a considerar como exógena la variable tiempo de trabajo t_w , y modificando en consecuencia el modelo usado por Train y McFadden (1987), Jara-Díaz et al. (1988), plantean que el verdadero tiempo disponible que el individuo puede asignar libremente es $(T-t_w)$.

Sin embargo, conviene plantear un nuevo enfoque que consista en realizar segmentaciones del modelo, de acuerdo con las características de los individuos considerados y las condiciones en que se realizan los viajes analizados. Además analizar empíricamente si los valores de los coeficientes de la función utilidad estimados para cada segmento planteado, son estadísticamente diferentes. Un aspecto que reviste gran interés es la asignación de un valor del tiempo a las personas sin empleo, que podrían tener una disposición nula a pagar por reducir su tiempo de viaje.

Un caso se puede analizar en el trabajo realizado por Pérez, Martínez y Ortúzar (2003), en el cual realizan un modelo usando segmentación por diferentes aspectos como la renta, entre otros. Obteniendo así medidas de la disposición a pagar para disminuir el tiempo de viaje.

Varias investigaciones han permitido el planteamiento de modelos en contextos concretos. Vilain y Bhandari (2002), desarrollaron modelos de elección discreta en los

cuales analizando las posibles diferencias entre medidas subjetivas y sociales, partiendo de datos directos de preferencias reveladas y declaradas en Croacia (Vilain & Bhandari, 2002) realizan estimaciones del valor del tiempo. A su vez, Mackie, Jara-Diaz y Fowkes (2001) revisan la elasticidad del valor del tiempo, los ahorros en el tiempo de trabajo, el tamaño de los ahorros del tiempo y los modos de transporte. Prasetyo, Fukuda, Yoshino y Yai (2003), en un estudio realizado en Tokio y Jakarta presentan un método de análisis del valor del tiempo de actividad, discutiendo que el beneficio en el ahorro del tiempo de viaje debe ser evaluado en más detalle los fines de semana. En otro trabajo relacionado, Jiang y Morikawa (2004), examinan cómo y por qué los ahorros en el valor del tiempo de viaje cambian con el viaje y el ambiente socio-económico por medio de un análisis estático comparativo de cambios en el valor del tiempo de viaje, costo de viaje y tiempo de trabajo; también se discuten diferencias entre las variaciones en los ahorros del tiempo de viaje y el valor del tiempo. En un artículo reciente Cantillo y Ortúzar (2005) muestran la influencia de la utilización de modelos no compensatorios en la estimación del valor del tiempo.

Otro autor con un tema importante para esta investigación, es Metz (2004), con su artículo "Valor del tiempo - ¿Variable o constante?", en el cual hace un análisis de la manera empírica en la que se concluye que el valor del tiempo es más o menos una variable, pero que sin embargo en algunos análisis de transporte se llega a la conclusión que el valor del tiempo de viaje es una variable.

Un punto también importante de tratar es el hecho de considerar que el valor del tiempo de viaje puede ser nulo en algunos casos, tema tratado por Richardson (2003), en un artículo en el cual reúne evidencia por medio de encuestas de preferencia revelada, de que existe un grupo de persona con un valor del tiempo nulo, haciendo parte de aquellos que usan los medios de transporte público, los usuarios privados si tienen un valor del tiempo.

Por otra parte, en los últimos años las investigaciones se han enfocado en encontrar valores subjetivos del tiempo para cada individuo, haciendo uso de técnicas bayesianas para la modelación (Train, 2009) modelos logit mixtos con parámetros individuales como los usados por (Sillano & Ortúzar, 2005) donde pudieron encontrar casos donde los individuos no responden seriamente a una encuesta de preferencias declaradas, pudieron hallar muestras representativas con un número pequeño de individuos, finalmente, presentando una comparación de las técnicas bayesianas y clásicas para la estimación del VST en la cual concluyen que no presentan diferencias significativas entre una técnica y otra, sin embargo en el caso de las bayesianas arrojan mejores resultados con muestras más pequeñas.

2.5.2 Estimación del Valor Subjetivo del Tiempo

Una de las aplicaciones de los modelos de elección discreta que se derivan de la teoría de la utilidad aleatoria es el cálculo de esta valoración subjetiva del atributo k-ésimo de una alternativa i . El atributo habitualmente más estimado es el Valor Subjetivo del Tiempo (VST) de viaje, aunque en este trabajo también se presenta los resultados con respecto al tiempo de acceso y la penalidad del transbordo.

El valor del tiempo se deriva de los estudios basados en la teoría macroeconómica del comportamiento del consumidor. En el área de transporte, este valor se obtiene del análisis de la elección hecha por el individuo (elección discreta). El método más aceptado para estimar el VST consiste en hallar la tasa marginal de sustitución entre el tiempo de viaje y el costo de viaje, usualmente a partir de las funciones de utilidad calibradas a partir de modelos desagregados de elección discreta basados en la teoría de utilidad aleatoria, como los planteados en este trabajo. La interpretación de dicho valor es cuánto está dispuesto a pagar una persona por reducir el tiempo de viaje en una unidad.

La valoración del tiempo se basa en la teoría de que el tiempo es un recurso económico, del cual disponen todos los individuos en la misma cantidad. Cada individuo puede distribuir este recurso a cada actividad en diferentes cantidades de tal manera que la distribución tendrá un impacto diferente en su presupuesto y su utilidad. Una suposición básica de esta teoría es que los individuos distribuyen su tiempo de tal manera que se maximice su utilidad, teniendo en cuenta que el tiempo no se puede almacenar como el dinero, sólo se puede transferir a otra actividad. Otra suposición es que la distribución del tiempo entre actividades para un individuo puede tener diferentes valores, los cuales se pueden medir en términos monetarios. Por ejemplo, una persona puede estar dispuesta a aumentar su tiempo de viaje si el costo de este disminuye.

El valor del tiempo es, en términos microeconómicos, un mecanismo a través del cual se intenta captar el equivalente monetario de la disminución, en una unidad, del tiempo necesario para realizar una actividad. Cuando esta equivalencia captura la percepción individual, se habla del valor subjetivo del tiempo o disponibilidad a pagar. “Si se nos pidiese una explicación intuitiva acerca de por qué preferiríamos, como parte de un esquema normal de actividades, demorarnos menos en el viaje diario a o desde el trabajo, probablemente llegaríamos a una conclusión que no por trivial deja de ser relevante: un viaje más breve podría permitirnos dormir más, desayunar más relajados, leer el diario, conversar con la familia, o trabajar más, entre otras posibilidades. En todos los casos se trata de sustituir parte de una actividad como el viaje por otra más placentera o útil”. (Jara – Díaz, 1999).

Beesley (1974) sugirió por primera vez una metodología para medir el valor del tiempo de viaje, la cual se enfoca en las condiciones para una medida exitosa (Hensher, 2001). Hensher (2001) explica que existen 3 aspectos de valoración del tiempo que pueden tener gran impacto en el VST, a saber: la heterogeneidad y no linealidad del tiempo, la estrategia de diseño para experimentos de preferencia declarada y la exploración de estructuras de covarianza del error de las alternativas en los modelos de elección discreta. Con respecto al tema del diseño de experimentos, afirma que este estimado del VST es sensible al número de alternativas por escenario, al número de escenarios presentados y al rango y niveles de los atributos tiempo y costo a analizar / estimar / determinar.

Tradicionalmente, los valores subjetivos de los ahorros de tiempo de viaje (VSTV) se han obtenido a partir de modelos que imponen la homogeneidad de los gustos y que permiten derivar un valor único de las disposiciones a pagar para un individuo medio ficticio. Este

es el caso de los modelos en los que se especifica una función indirecta de utilidad lineal y con parámetros genéricos (MNL), la expresión del VSTV para un individuo q vendría dada por el cociente entre el parámetro del tiempo y el del costo, esto es,

$$VST = \frac{\partial V_{iq} / \partial T_{iq}}{\partial V_{iq} / \partial C_{iq}} = \frac{\beta_{T,q}}{\beta_{C,q}} \quad (32)$$

Sin embargo, éste puede ser un supuesto demasiado restrictivo y la disposición a pagar puede diferir entre los individuos de acuerdo a sus preferencias y a determinadas características socioeconómicas. El desarrollo reciente de las técnicas de estimación por simulación ha facilitado la utilización de modelos más flexibles como el logit mixto o de parámetros aleatorios que permite captar la existencia de variación en los gustos en la población. Sin embargo, a pesar de la relativa sencillez con que estos modelos pueden ser estimados, las implicaciones de su utilización a la hora de derivar VSTV no están del todo claras, siendo la evidencia empírica al respecto poco concluyente. En este sentido, los estudios realizados hasta el momento muestran que los modelos que imponen la homogeneidad de las preferencias en una población que presenta variación en los gustos pueden conducir tanto a sobrestimar como a subestimar los VSTV.

Aunque el tema de calcular distribuciones de los parámetros individuales para cada atributo presenta un gran reto para el modelador, se hace más interesante cuando este análisis se centra en el cálculo de VST o WTP (Willingness To Pay). Cuando se tienen parámetros fijos, este cálculo se resume al presentado en la expresión anterior. Cuando se tratan factores aleatorios, como es el caso de los hallados en los modelos ML, esto es un reto. Existen dos enfoques básicos para abordar este tema: usar toda la información que se tiene de la distribución o simplemente emplear la media y desviación estándar de la misma. El primer enfoque es el más empleado y el más complejo de los dos.

3 DISEÑO DEL EXPERIMENTO

Un experimento de preferencias reveladas (PR) y preferencias declaradas (PD) fue realizado en la ciudad de Bogotá con el fin de analizar la transferencia modal del transporte privado al público y de un modo a otro, que pudiese ocurrir ante la introducción de nuevos modos de transporte como el metro y el tren de cercanías. La población objetivo está constituida por los potenciales usuarios del sistema metro en el marco de un Sistema Integrado de Transporte Público, enfatizando en los poseedores de vehículos, dada la cautividad hacia el transporte público de quienes no tienen acceso al transporte privado. En el análisis se ha incluido otra opción como es el taxi, de gran aceptación entre los poseedores de vehículo ya que su operación es muy similar. Como complemento a este primer experimento se incluye un adicional donde se analiza la transferencia modal para usuarios de transporte público, quienes no tienen disponible el transporte privado ni el taxi.

3.1 DATOS DE PREFERENCIA REVELADA Y PREFERENCIA DECLARADA

La información empleada para estimar modelos de elección discreta puede provenir principalmente de tres fuentes de datos: preferencias reveladas (PR), preferencias declaradas (PD) o una combinación de ambos, datos mixtos (PR/PD).

Los primeros, más tradicionales, corresponden al comportamiento de consumo observado directamente de los individuos. Así, los datos de PR generalmente representan un corte transversal y se obtienen a partir de encuestas que intentan medir el valor de los atributos de las alternativas disponibles por los individuos muestreados. La principal ventaja es su realismo, aunque también tienen limitaciones entre las que se destacan la posible existencia de correlación entre las variables, la poca variabilidad de los datos, la existencia de errores de medición, la dificultad de evaluar variables cualitativas, su alto costo y la dificultad de estudiar la demanda por alternativas no existentes en el mercado.

En los últimos años han ganado popularidad las técnicas de PD, que se basan en declaraciones de los individuos acerca de cuáles son sus preferencias ante situaciones que describen escenarios de elección hipotéticos, contruidos por el investigador según un diseño experimental. Estas técnicas tienen la ventaja de sortear varias de las limitaciones de las PR, pero no están exentas de problemas; entre éstos destaca la posible falta de realismo asociada a la potencial existencia de diferencias entre lo que la gente declara que haría y lo que realmente haría si llega la ocasión.

Al hablar de datos PD se hace referencia a un conjunto de técnicas que se basan en declaraciones de individuos acerca de cuáles son sus preferencias cuando se les presentan opciones que describen una serie de situaciones o escenarios hipotéticos contruidos por el investigador en un diseño experimental.

En los últimos años numerosas investigaciones han apuntado a describir el proceso a seguir para diseñar un experimento con preferencias declaradas. Desde los años 80's a través de las investigaciones desarrolladas por Louviere y Woodworth (1983) y por

Louviere, Hensher and Swait (2000) se proponen este tipo de experimentos en diferentes campos como el transporte, la salud, el marketing, entre otros. Posteriormente, diferentes autores han sugerido métodos para la construcción de este tipo de experimentos (Anderson y Wiley 1992, Bunch, Louviere, y Anderson 1996) basándose en su mayoría en métodos que son extensiones de diseños experimentales lineales estándar (Zwerina, Huber y Kuhfeld, 1996) los cuales presentan limitaciones en dos características principales: los modelos de elección de probabilidad son no lineales en los parámetros, implicando que la eficiencia estadística del diseño depende de un vector de parámetros (desconocido normalmente). Por otra parte, la eficiencia de los diseños de elección depende de la creación de un perfil apropiado y colocarlos apropiadamente en un set de elección. Con el fin de sortear estas limitaciones, nacen los llamados diseños eficientes.

Los cuatro principios de los diseños eficientes descritos por Huber y Zwerina (1996) son:

- Ortogonalidad. Esta condición se satisface cuando los niveles de cada atributo varía independientemente de los demás.
- Balance de niveles. Cuando los niveles de cada atributo aparecen en la misma proporción.
- Traslapo mínimo. Se cumple cuando las alternativas junto con su set de elección no tienen niveles de atributos traslapados.
- Balance de utilidad. Se satisface cuando las utilidades de las alternativas dentro del conjunto de elección son las mismas.

En una investigación desarrollada por Zwerina, Huber y Kuhfeld (1996) desarrollan un algoritmo para minimizar el D-error, aproximando los cuatro principios descritos anteriormente para satisfacer un diseño eficiente. Usando el algoritmo de Fedorov con la posibilidad de usar computadores sencillos de fácil acceso.

En el campo del transporte autores como Rose y Bliemer (2006) presentan un estado del arte para el diseño de este tipo de experimentos, en este trabajo describen diferentes tipos de diseños como son los diseños ortogonales y diferentes tipos de diseños eficientes. Algunos autores desarrollan métodos para construir diseños con PD para modelos logit específicos como son, logit mixto (Sándor y Wedel, 2002, Bliemer y Rose, 2008) y logit jerárquico (Bliemer et al, 2007) entre otros, ampliamente utilizados en el campo de la modelación en transporte.

Finalmente, la modelación con datos mixtos de PR/PD aparece como una alternativa que permite combinar las bondades de ambas técnicas. Sin embargo, son muchos los desafíos que involucra el uso de datos mixtos (Hensher et al, 1999), siendo relevante el originado por la distinta naturaleza de los datos, que implica que los errores asociados a ambos tipos de información pueden ser diferentes. Al respecto Ben Akiva y Morikawa (1990) propusieron la introducción de un factor de escala, que permite igualar las varianzas de sus errores y que está dado por:

$$\mu^2 = \frac{\sigma_{PR}^2}{\sigma_{PD}^2} \quad (33)$$

donde σ_{PR}^2 y σ_{PD}^2 son las varianzas de los errores asociados a los datos de PR y PD, respectivamente. El factor de escala μ puede ser estimado conjuntamente con los demás parámetros del modelo (Bradley & Daily, Estimation of logit choice models using mixed stated preference and revealed preference information, 1997). Esta aproximación puede generalizarse a la aplicación conjunta de bases de datos (aún pudiendo ser ambas de PD o de PR), cada una de ellas con distinta varianza.

En el presente trabajo se desarrollaron modelos de elección modal sólo con los datos de preferencia declarada, a pesar, de que el experimento contaba con datos de PR. La razón principal de trabajar sólo con las PD se debió a la falta de información veraz en cuanto a tiempos y distancias proporcionados por los encuestados, arrojando así modelos con resultados poco esperados.

3.2 DISEÑO EXPERIMENTAL

El instrumento aplicado incluye varias partes. La primera hace referencia a información socioeconómica del encuestado, destacando ocupación, estrato socioeconómico, vehículos en el hogar, género, edad y estructura del hogar. La segunda, es una encuesta de preferencias reveladas en la cual se pregunta sobre el último viaje realizado, incluyendo información sobre el motivo, hora, frecuencia, tanto para el modo escogido como para los modos alternativos, costos, tiempos de viaje, tiempos de acceso y transbordos. Otra parte de la encuesta son las preguntas de preferencias declaradas, cuyo diseño tuvo en cuenta las siguientes etapas:

- a. Definición del ámbito de elección. Dado el interés en estimar la demanda del Metro se generaron escenarios hipotéticos de elección contrastando el Sistema Integrado de Transporte (SITP), considerando como columna vertebral el Metro, con modos alternativos, siendo ellos el auto particular y el taxi. Es importante resaltar que en el SITP, al aplicar la encuesta se aclaró que incluye además el sistema BRT (Transmilenio), y el sistema alimentador, con integración operacional y tarifaria. En el juego, el individuo debía escoger la alternativa de su conveniencia en atención a los atributos allí indicados, siendo una encuesta del tipo “elección”. Puede notarse que en un conjunto de encuestas se confrontó el SITM tomando como eje el Metro con el auto particular y el taxi, y en otro se confrontó el SITM tomando como eje el sistema troncal sobre buses (Transmilenio) con el auto particular y con el taxi.
- b. Diseño preliminar de la encuesta. Para este diseño inicial, fueron escogidas inicialmente las variables relevantes, y el número de niveles de cada una. El análisis llevó a considerar las variables tiempo de viaje en vehículo, tiempo de acceso (incluye caminata y espera), costo y número de transbordos. Siguiendo las últimas tendencias, se procuró diseños eficientes, que permiten mejor

aprovechamiento de los datos que los tradicionales diseños puramente ortogonales (Rose y Bliemer, 2006).

- c. Realización de grupo focal. Se utilizó la técnica de grupos focales, orientado por un experto para obtener información útil. El diseño inicial fue presentado a un conjunto de 15 personas, representativo de la población a estudiar, para obtener retroalimentación que permitiese mejorar el instrumento. Aspectos como el lenguaje, las variables y los niveles en que se encontraban, el ámbito de elección y otros fueron evaluados por el grupo, obteniendo información útil para revisar el diseño preliminar. Los encuestados proporcionaron información de gran valor sobre las tarifas y costos y el proceso por el cual toman las decisiones de viajes. Un aspecto importante fue el procurar diseños adaptivos a las características del viaje del individuo; por ello, se preparó una encuesta para viajes cortos y otra para viajes largos.
- d. Aplicación de encuesta piloto. El diseño mejorado fue probado mediante la realización de una prueba piloto, aplicada a un conjunto de 70 individuos. La prueba piloto permitió verificar la pertinencia del instrumento y evaluar variables como los tiempos de aplicación y la recepción de la encuesta. Los resultados de la encuesta permitieron estimar modelos preliminares y verificar su consistencia microeconómica. Se encontró que los atributos y niveles definidos permitían recuperar los parámetros correspondientes de la función de utilidad. En síntesis, la prueba piloto permitió afinar el diseño del instrumento definitivo.
- e. Aplicación de la encuesta. En los sitios escogidos se llevó a cabo el trabajo de campo, aplicando un muestreo aleatorio dentro de cada estrato. En total fueron entrevistados 1912 individuos, de los cuales a 1125 se les confrontó el SITM considerando como eje troncal el Metro, en tanto que al resto de ellos se les confrontó el SITM considerando como eje troncal el sistema sobre buses (Transmilenio). Posteriormente, con la información obtenida se configuró una base de datos que permitiera su tratamiento para los cálculos.
- f. Estimación de modelos. A partir de los datos de la encuesta se procedió a estimar los modelos de elección involucrando las diferentes alternativas. En cada uno de ellos se realizó la correspondiente evaluación a la luz de criterios estadísticos y de consistencia microeconómica.

3.2.1 Definición de Atributos y Niveles

Para el presente estudio se estableció el enfoque de la encuesta, la cual estaría dirigida a usuarios de vehículo particular y taxi, y por otra parte usuarios de transporte público cautivos del mismo, todos ellos en sus viajes más frecuentes a su sitio de actividades (trabajo y estudio principalmente). Después de haber analizado la situación del transporte urbano en la región metropolitana de Bogotá y teniendo en cuenta los antecedentes que se tienen sobre el tema, se llegó a la conclusión que los aspectos más relevantes al momento de decidir qué modo de transporte elegir son: el costo, el tiempo de viaje y el tiempo de acceso. Estas fueron las variables consideradas en el estudio, para el caso del

automóvil y del taxi. En el caso del transporte público que para efectos de este estudio será el SITP se agregan otras dos variables como son el número de transbordos y el confort. Para el caso del SITP en unas encuestas se asoció con la operación de un metro, con posibilidad de transbordo integrado con Transmilenio y rutas alimentadoras; en otro conjunto de datos se asoció a un SITP en el cual el recorrido troncal se realiza en un sistema sobre buses (Transmilenio), con posibilidad de transbordo. Hubo otras variables consideradas en el diseño original que fueron la seguridad y el confort; sobre esta última en el grupo focal se notó la diversidad de interpretaciones: algunos lo asociaban con la ocupación del vehículo (lleno o vacío), otros lo asociaban a la cola en paraderos, otros lo asociaban al estado de los vehículos o a la tecnología. Como resultado del grupo focal y de la prueba piloto se optó por explicar parcialmente el efecto de la seguridad e incluir solo el confort y hacer referencia a percepciones de la seguridad, a través de la constante modal; ello además, permite un diseño menos complejo, dado que incluir una quinta en el diseño incrementa sustancialmente el diseño experimental lo cual puede redundar en los resultados (Caussade et al, 2005).

El costo o tarifa está definido en caso del automóvil como el costo del estacionamiento más un costo de operación, dado principalmente por el consumo de combustible; a su vez, el costo del taxi y del transporte público está representado en la tarifa del pasaje. El tiempo de viaje se definió como el tiempo en movimiento, separado del tiempo de acceso que incluye los tiempos de caminata antes y después de abordar el vehículo y el tiempo de espera en paradero o en espera del taxi. Un transbordo es cada intercambio vehicular entre modos motorizados (sin incluir la caminata ni la bicicleta), y se supuso que éstos se realizan sin que medie un nuevo pago, dado que se trata de un sistema con integración tarifaria. En caso de que haya un pago por transbordo, este se deberá considerar en la tarifa.

La definición del confort se toma como un análisis técnico en donde está determinado por el estrato socioeconómico de la persona, dado que existen diferentes apreciaciones de la siguiente manera:

- Estratos 1, 2 y 3
Confort Bajo: Vehículo muy lleno, poca posibilidad de ir sentado, frecuencia irregular.
Confort Medio: Vehículo lleno, poca posibilidad de ir sentado, frecuencia constante y confiable.
Confort Alto: Vehículo semilleno, alta posibilidad de ir sentado, frecuencia constante y confiable.
- Estrato 4, 5 y 6
Confort Bajo: Frecuencia irregular, muy variable y poco confiable. Vehículo lleno.
Confort Medio: Frecuencia constante y confiable. Vehículo lleno.
Confort Alto: Frecuencia constante y confiable. Vehículo semilleno y amplitud de espacios.

Para el diseño, los atributos tarifa y tiempo de viaje fueron tratados en tres niveles tomando un valor nominal referencial de acuerdo con los valores actuales. Para el atributo tiempo de acceso se consideró dos niveles. En el caso de transbordo se trabajó con dos niveles (0 y 1 transbordo) para viajes cortos, definidos como aquellos que son menores de 30 minutos; y tres niveles (0, 1 y 2 transbordos) para viajes largos entendidos como aquellos mayores de 30 minutos). Por último, para el confort se definieron tres niveles (bajo, medio y alto) según se precisó anteriormente.

A objeto de recolectar esta información de los valores actuales, fue necesario realizar una investigación de mercado, complementada con la información tomada del grupo focal, en el cual se mostraron las alternativas, los atributos y los escenarios, con el fin de que los individuos evaluaran si los valores presentados correspondían a la realidad. En el diseño experimental, el conjunto de elecciones estaba constituido por dos alternativas.

La combinación de niveles define distintos escenarios que constituyen el diseño experimental. La práctica más común ha sido la de usar diseños ortogonales, lo cual implica asegurar que cada atributo varía independientemente de los demás, de forma que se pueda aislar su efecto en la función de utilidad indirecta. La ortogonalidad está asegurada cuando se consideran todas las combinaciones posibles de los niveles de las variables; esto se constituye en un diseño factorial completo, que crece exponencialmente con el número de atributos y niveles.

Un diseño factorial completo permite estimar no sólo los efectos aislados de todos los atributos sino también todas las posibles interacciones entre ellos. El problema es que en estos casos el número de escenarios puede ser muy elevado y el experimento puede producir fatiga en el entrevistado. En el caso de este estudio, en los viajes largos el número de escenarios es $54 = 27 \times 2$, mientras que para los viajes cortos es $36 = 9 \times 4$. Por tanto, lo que se suele hacer en la práctica, y es de hecho lo realizado en este estudio, es utilizar un diseño factorial fraccionado, que consiste en seleccionar solo un subconjunto de escenarios del diseño factorial completo. Sin embargo, en la gran mayoría de las aplicaciones empíricas, se observa que los efectos lineales explican alrededor de un 70 - 90% de la varianza observada – condición de monotonicidad en los efectos lineales (Hensher, Louviere, & Swait, 1999); esto sugiere, en diseño experimental, considerar principalmente efectos lineales, y, si hubiese motivos, alguna(s) interacción(es) dobles y/o algún(os) efecto(s) cuadrático(s).

Si bien los diseños ortogonales son un excelente punto de partida, en este caso se utilizó el criterio de “diseños eficientes” (Huber y Zwerina, 1996). Un experimento se considerará eficiente si los parámetros estimados con los datos recolectados de su aplicación son confiables (baja magnitud de sus errores estándar). Por lo tanto, es interesante comparar diseños mediante la matriz de covarianza asociada a los parámetros; ésta matriz es la inversa de la matriz de información obtenida de la aplicación del experimento, a partir de la función de log-verosimilitud asociada al modelo. Para lograr este diseño, se utilizó como información preliminar los parámetros de modelos de selección modal previamente calibrados en Bogotá.

Atendiendo a la recomendación de presentarle al usuario 9 o 10 escenarios, el diseño del experimento presenta únicamente 9 escenarios. El diseño ortogonal base fue tomado de Kocur et al (1982) y resulta ser el siguiente:

Tabla 3.1 Escenarios del Experimento

1111	2122	3113
1212	2213	3221
1323	2311	3312

Fuente: Elaboración propia

Este diseño es una fracción del original, en el cual no se tienen en cuenta las interacciones de segundo orden o de orden superior y se puede probar que es ortogonal. Ahora bien, esta fracción representa el diseño en caso de una alternativa; es decir, un modo que viaje desde un origen a un destino determinado. Dado que en los escenarios se presentan 2 alternativas para cada par origen – destino, es necesario determinar el orden de estos 9 escenarios para cada una de ellas. Para definir los conjuntos de elección se utilizó teniendo en cuenta los balances de utilidades sugeridos por Huber y Zwerina (1996). Se evaluaron distintas alternativas de grupos de elección y combinación de niveles en los atributos para generar las diferentes encuestas, una para viajes cortos, otra para viajes largos y, dependiendo de si tienen o no disponible su vehículo particular en ese día, se confronta con el auto o con el taxi. La construcción de los diseños implicó utilizar las técnicas de intercambio de niveles (swapping) y reetiquetamiento (relabelling) (Huber y Zwerina, 1996; Hensher, Louviere y Swait 1999; Sándor y Wedel, 2001; Sándor y Wedel, 2002).

En el juego se contrastó el SITP con el auto y con el taxi, considerando en ambos casos escenarios de viajes cortos y viajes largos. Los niveles empleados en los diseños se presentan en la Tabla 3.2 para los viajes largos y en la Tabla 3.3 para los viajes cortos. Los diseños obtenidos se presentan en la Tabla 3.4.

Tabla 3.2 Niveles de atributos para viajes largos

	SITP (Metro – Transmilenio)			Transporte Privado (Auto – Taxi)		
	1	2	3	1	2	3
Costo	1500	1800	2100	4000	5500	7000
Tiempo de Viaje	20	25	30	30	40	50
Tiempo de Acceso	10	20		5	10	
Transbordo	0	1	2			

Tabla 3.3 Niveles de atributos para viajes cortos

	SITP (Metro – Transmilenio)			Transporte Privado (Auto – Taxi)		
	1	2	3	1	2	3
Costo	1500	1800	2100	3000	4500	6000
Tiempo de Viaje	10	15	20	10	20	30
Tiempo de Acceso	10	20		5	10	
Transbordo	0	1				

Tabla 3.4 Diseño encuesta de preferencias declaradas PD1

1	Alternativa 1		Alternativa 2	
	Tiempo de viaje	0	Tiempo de viaje	0
	Costo	0	Costo	0
	Tiempo de Acceso	0	Tiempo de Acceso	0
	Transbordo	0		
2	Alternativa 1		Alternativa 2	
	Tiempo de viaje	0	Tiempo de viaje	1
	Costo	1	Costo	1
	Tiempo de Acceso	0	Tiempo de Acceso	0
	Transbordo	1		
3	Alternativa 1		Alternativa 2	
	Tiempo de viaje	0	Tiempo de viaje	2
	Costo	2	Costo	0
	Tiempo de Acceso	1	Tiempo de Acceso	0
	Transbordo	0		
4	Alternativa 1		Alternativa 2	
	Tiempo de viaje	1	Tiempo de viaje	2
	Costo	0	Costo	1
	Tiempo de Acceso	1	Tiempo de Acceso	1
	Transbordo	1		
5	Alternativa 1		Alternativa 2	
	Tiempo de viaje	1	Tiempo de viaje	0
	Costo	1	Costo	2
	Tiempo de Acceso	0	Tiempo de Acceso	1
	Transbordo	1		
6	Alternativa 1		Alternativa 2	
	Tiempo de viaje	1	Tiempo de viaje	1
	Costo	2	Costo	0
	Tiempo de Acceso	0	Tiempo de Acceso	1
	Transbordo	0		
7	Alternativa 1		Alternativa 2	
	Tiempo de viaje	2	Tiempo de viaje	0
	Costo	0	Costo	1
	Tiempo de Acceso	0	Tiempo de Acceso	0
	Transbordo	0		
8	Alternativa 1		Alternativa 2	
	Tiempo de viaje	2	Tiempo de viaje	1
	Costo	1	Costo	2
	Tiempo de Acceso	1	Tiempo de Acceso	0
	Transbordo	0		
9	Alternativa 1		Alternativa 2	
	Tiempo de viaje	2	Tiempo de viaje	2
	Costo	2	Costo	2
	Tiempo de Acceso	0	Tiempo de Acceso	0
	Transbordo	1		

Como se ha indicado, la encuesta de preferencias declarada se complementa con una de datos de preferencia revelada, que está dividida en tres partes: la primera está relacionada con la disponibilidad de auto particular para la realización del viaje, bien sea en calidad de conductor o de acompañante. Si la respuesta es afirmativa, se continua adelante con la encuesta, en caso contrario se le agradecerá al encuestado y en ese momento se termina la encuesta.

La segunda parte de la encuesta PR está relacionada con la identificación del encuestado, en cuanto su ocupación, edad y sexo. Estas variables aunque que pudieran ser incluidas en el modelo, son además importantes en el momento de revisar que las encuestas no estén sesgadas a cierto tipo de población, y que, por el contrario, se garantice tienen una representación razonable del mercado.

En la tercer parte se indaga sobre el viaje más reciente en modo particular, para el periodo de máxima demanda en la mañana (06:00 - 09:00 horas), detallando los atributos del viaje. Es importante conocer de alguna manera, los atributos del modo alterno del usuario, por ejemplo cuando tiene restricción vehicular, además de información general del encuestado: ingresos, ocupación, motivo del viaje, edad, sexo entre otros conceptos de información socioeconómica.

Como se indicó anteriormente, el muestreo para el tema de encuestas de preferencias declaradas es un tema bastante complejo y se puede resumir en las siguientes 2 preguntas: ¿Quién debe ser encuestado? y ¿Cuántas encuestas deben ser realizadas? Atendiendo a la primera pregunta, se puede decir que la población objetivo de esta investigación son personas que realizan su viaje cotidiano en las horas pico de la mañana y que tienen acceso al vehículo particular para realizar ese viaje, bien sea hombre o mujer mayores de 18 años. A los usuarios se les aplicó la encuesta en su lugar de trabajo o de estudio.

En referencia al muestreo, para la construcción de los modelos, un total de 1912 individuos encuestados constituyen la muestra, información que puede ser utilizada para estimar modelos de PR. En el caso de los datos de PD, dado que se tienen nueve observaciones por individuo en el primer caso y ocho para las segundas encuestas realizadas, hay disponibles 16.421 pseudo-individuos.

Los cuadernillos utilizados en desarrollo del experimento pueden consultarse en el anexo, como ilustración, en la Figura 3.2 y en la Figura 3.1 se presentan ejemplos de las opciones de elección de los individuos.

Figura 3.1 Opciones de elección para el caso de viajes cortos
PD1 VIAJE CORTO

1		TARIFA	TIEMPO	ACCESO	TRANSBORDOS	CONFORT
		\$ 1.500,00	10	10	0	MEDIO
		TARIFA	TIEMPO	ACCESO	TRANSBORDOS	CONFORT
2		\$ 1.200,00	25	10	0	MEDIO
		TARIFA	TIEMPO	ACCESO	TRANSBORDOS	CONFORT
		\$ 1.800,00	15	10	1	ALTO
2		TARIFA	TIEMPO	ACCESO	TRANSBORDOS	CONFORT
		\$ 1.800,00	10	10	1	ALTO
		TARIFA	TIEMPO	ACCESO	TRANSBORDOS	CONFORT
2		\$ 1.400,00	25	15	0	BAJO
		TARIFA	TIEMPO	ACCESO	TRANSBORDOS	CONFORT
		\$ 1.500,00	15	10	0	MEDIO

Fuente: Proyecto Metro Bogotá

Figura 3.2 Cuestionario para encuestas PR y PD



PROYECTO METRO BOGOTÁ
Encuesta de Preferencias a individuos



1. IDENTIFICACIÓN DE LA ENCUESTA

Lugar Fecha / / Folio

Encuestador Digitador Hora

2. IDENTIFICACIÓN DEL USUARIO

¿Dispone de vehículo para realizar en él su viaje cotidiano?

☐ Sí Conductor ☐ Acompañante ☐
☐ No

Ocupación Género Edad

Estudiante ☐ Directivo ☐ Femenino ☐
 Empleado ☐ Profesional ☐ Masculino ☐
 Independiente ☐ Técnico ☐
 Otro ☐ Auxiliar ☐

18 - 25 AÑOS ☐
 26 - 35 AÑOS ☐
 36 - 45 AÑOS ☐
 > 45 AÑOS ☐

3. INFORMACIÓN DEL VIAJE

¿Utilizó vehículo particular para realizar viaje cotidiano en esta mañana?

☐ Sí ☐ No → Pico y Placa ☐ Otro ☐

¿A qué hora inició su viaje? :

Modo en que realizó el viaje ¿Cuál fue el motivo de su viaje?

☐ B ☐ TM ☐ TX Trabajo ☐
 Estudio ☐
 Otro ☐

¿A qué hora terminó su viaje? :

¿Con qué frecuencia realiza este viaje a la semana?

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

VIAJE EN AUTO

Tiempo de Viaje :

Costo del viaje (Incluye parqueadero)

¿Cuántas cuadras camina al parqueadero? Cuadras

VIAJE EN TRANSMILENIO

Tiempo de Viaje :

Costo del viaje

¿Cuántas cuadras camina? Cuadras

¿Cuánto tiempo espera en el paradero? Minutos

Transbordos 0 1 2 Confort A H B

VIAJE EN BUS

Tiempo de Viaje :

Costo del viaje

¿Cuántas cuadras camina? Cuadras

¿Cuánto tiempo espera en el paradero? Minutos

Transbordos 0 1 2 Confort A H B

VIAJE EN TAXI

Tiempo de Viaje :

Costo del viaje

¿Cuántas cuadras camina para tomar el taxi? Cuadras

¿Cuánto tiempo espera el taxi? Minutos

4. PREFERENCIAS DECLARADAS

Considere el viaje que usted ha realizado. A continuación le mostraremos ocho situaciones en las que usted deberá escoger entre Metro, Transmilenio, Bus. Las tres alternativas cambiarán de acuerdo a cada diseño presentado

PD N°

☐ 1	Metro ☐	Transmilenio ☐	Bus ☐
☐ 2	Metro ☐	Transmilenio ☐	Bus ☐
☐ 3	Metro ☐	Transmilenio ☐	Bus ☐
☐ 4	Metro ☐	Transmilenio ☐	Bus ☐
☐ 5	Metro ☐	Transmilenio ☐	Bus ☐
☐ 6	Metro ☐	Transmilenio ☐	Bus ☐
☐ 7	Metro ☐	Transmilenio ☐	Bus ☐
☐ 8	Metro ☐	Transmilenio ☐	Bus ☐

5. INFORMACIÓN SOCIOECONÓMICA

Número de autos en hogar 1 2 M Número de personas en hogar

¿En qué barrio vive? ¿En qué estrato vive?

Es posible realizar su viaje en Transmilenio SI NO

OBSERVACIONES

Fuente: Proyecto Metro Bogotá

3.3 INFORMACIÓN SECUNDARIA

En la ciudad de Bogotá luego de la implementación del modo BRT conocido como, Transmilenio se han realizado una serie de estudios con el fin de analizar la movilidad y otros aspectos del transporte en la ciudad. Entre esos estudios se destacan principalmente tres de ellos:

- La recopilación de información adelantada por la empresa Logitrans - Logit en el año 2002. De este estudio no se cuenta con información detallada, excepto algunos resultados obtenidos del mismo como son la hora de máxima demanda para el transporte público en la ciudad es 6:30 a 7:30 am, y el número total de viajes en transporte público es de 567,886 en esa hora.
- La encuesta de movilidad realizada en el 2005.
- La encuesta para la actualización y validación de las matrices de transporte público y revisión de la base de datos de oferta física y operativa del modelo de transporte de la ciudad de Bogotá del año 2007, realizada por Cal y Mayor y Asociados. Este trabajo se divide en cuatro actividades principales: la toma de información primaria, revisión de los modelos de transporte, conformación de banco de datos para la modelación y aplicación de encuestas de preferencia declaradas para elección modal.

Como toma de información primaria se realizaron dos estudios: frecuencias y ocupación visual y encuestas origen – destino en unidades de transporte público en ciertos tramos definidos en corredores de la ciudad. En los estudios de frecuencia y ocupación se definieron 59 corredores, 10 corredores más de los presentados en el estudio de Logitrans – Logit en el 2002 con el fin de garantizar la cobertura adecuada del sistema de transporte de la ciudad, para las tres horas de aforo (6:00 a 9:00, considerada la hora pico en la ciudad) se tiene un volumen total de pasajeros de 1,651,364. Por otra parte, las encuestas origen – destino se dividen en dos metodologías, las encuestas para usuarios de transporte público colectivo y las encuestas para usuarios de transporte público masivo, el tamaño de la muestra se obtuvo luego de la aplicación del procedimiento de muestreo probabilístico con un error de estimación del 1.5% y un nivel de confianza del 99%, teniendo en cuenta el volumen total de pasajeros del estudio de Frecuencia y Ocupación, el número de encuestas a aplicar es de 7,339.

En la siguiente actividad que consistió en los modelos de transporte se encuentra la zonificación del modelo la cual luego de corregirla y adaptarla se dividieron en 764 zonas urbanas de Bogotá y 60 zonas externas, considerándose la UPZ como macrozona de referencia. El modelo de transporte público y de transporte privado se revisó con base en el entregado por el banco en el escenario 9003, concluyendo que la hora pico en la ciudad es de 6:30 a 7:30 de la mañana, dato que concuerda con el entregado por el estudio de Logitrans – Logit en el 2002.

Con respecto al inventario de tránsito se tomó información entregada por la Secretaria de Movilidad sobre volúmenes vehiculares de los años 2005 y 2006 en cinco estaciones maestras distribuidas en intersecciones de corredores importantes de la ciudad

desagregados según tipo de vehículos y en periodos de cada 15 minutos. Adicionalmente, la información de cartillas de volúmenes vehiculares para Bogotá de la STT de los años 2000 – 2004. Luego de depuración y procesamiento de esta información se concluye que al ser las tasas de crecimiento de tráfico muy altas (en algunos casos superiores del 60%) se tomarán las tasas según el Producto Interno Bruto.

Por otra parte las velocidades de recorrido se toman para 14 corredores principales de la ciudad para el periodo pico de mañana de 6:00 a 9:00 a.m., información entregada por la Secretaría de Movilidad de la ciudad.

Finalmente, para las encuestas de preferencia declaradas se tomaron en cuenta combinaciones de alternativas de elección: sistema de Transmilenio solo, sistema de Transmilenio más alimentador, sistema de Transmilenio más alimentador más bus (2 transbordos), sistemas de solo buses y sistemas de solo buses con transbordo. Los atributos elegidos para el modelo fueron: tiempo de viaje en vehículo, tiempo de caminata, tiempo de espera en paradero, tarifa y numero de transbordos, cada uno con dos niveles de variación. El modelo seleccionado fue el más simple de ellos, donde se desagregaba por estrato socioeconómico, teniendo un valor subjetivo del tiempo según estratos socioeconómicos así: estratos 1y2 \$/min 18,4, 3y4 \$/min 25,2 y 5 y 6 \$/min 45,5.

Como conclusión del estudio se observa que el número de viajes en transporte público en la ciudad de Bogotá pasa de 567,886 viajes (Logitrans – Logit, 2003) a 658.739 viajes al 2007.

Teniendo en cuenta las sugerencias y recomendaciones de estas fuentes secundarias, se seleccionan los atributos y niveles de las alternativas como preliminares para el grupo focal y la encuesta piloto. Estos atributos a saber son: tiempo de viaje, tiempo de espera en el paradero, tiempo de caminata, tarifa, número de transbordos y una variable latente como es el confort. Adicionalmente, para seleccionar la muestra se tomaron resultados obtenidos de estos estudios previos, como preliminares del nuestro.

3.4 METODOLOGIA DE LA TOMA DE INFORMACION

Partiendo de los datos disponibles (los resultados de la encuesta realizada en el 2007 y de la Encuesta de Movilidad 2005), para la selección de la muestra del trabajo de campo se ha seguido el siguiente procedimiento:

- Segmentación de la demanda en grupos de comportamiento homogéneo respecto a la selección modal. Las variables de segmentación más relevantes son el ingreso, la posesión de vehículo y la ocupación. Se definieron las categorías de ingreso según la estratificación socioeconómica utilizada para tarificación de servicios públicos (estratos 1 a 6). A su vez, según la posesión de vehículo se definieron tres categorías: no poseedores de vehículos y por tanto cautivos del transporte público, usuarios con un vehículo en el hogar y usuarios con dos o más vehículos en el hogar. Finalmente, según la ocupación se consideró dos categorías; estudio y trabajo. Es importante resaltar que se encontró una fuerte correlación entre la posesión de vehículo y el número de autos.

- Determinación de la muestra para cada segmento de la población. Diseños eficientes de encuestas de PD permiten lograr buenos modelos con menos datos que los requeridos cuando se trabaja con datos de PR. Aunque hay indicios respecto de que una muestra de 30 entrevistas por segmento puede ser suficiente, investigaciones sugieren al menos entre 75 y 100 entrevistas por segmento (Bradley and Kroes, 1990; Swanson et al, 1992; Nerella, S. and Bhat, C., 2004). En el trabajo de campo realizado, se encuestó un total de 1912 individuos. Para el primer grupo de encuestados (usuarios de transporte privado) un total de 1.125 individuos se le presentó un juego de nueve situaciones de elección, lo que resulta en un total de 10.125 pseudo-individuos. El segundo grupo de encuestados, usuarios de transporte público cautivos del mismo, un total de 787 con un juego de ocho situaciones de elección para un número de 6.296 pseudo-individuos. Este tamaño muestral permitió cubrir de manera suficiente cada uno de los segmentos en que fue dividida la población.
- Selección del ámbito geográfico para la muestra. La zona que se trabajó fue específicamente el centro expandido de la ciudad (calle 6 “centro” y calle 100, entre los cerros orientales y la Avenida Caracas extendida a la calle 100 con Autopista Norte). La elección se hizo considerando el área de la ciudad que más viajes demanda, dada la presencia en ella de oficinas, instituciones educativas, de manera que allí hay una importante concentración de la población objeto de estudio. La encuesta fue aplicada en oficinas, instituciones educativas, centros comerciales, entre otros.
- Reparto de la muestra de cada segmento entre los diversos ámbitos geográficos. Por su condición, la muestra fue dividida convenientemente según la ocupación, de manera que el 65% trabaja, el 25% estudia y el 10% tenía otra ocupación. Por su parte, según estratos, el 4% de la muestra corresponde a ingresos bajos, el 70% a ingresos medios y el 26% a ingresos altos. Esta distribución se explica porque la encuesta fue aplicada a poseedores de vehículos en su mayoría.

La Figura 3.3 muestra esquemáticamente la metodología seguida durante el trabajo de campo necesario para la estimación de modelos.

Figura 3.3 Metodología del trabajo de campo

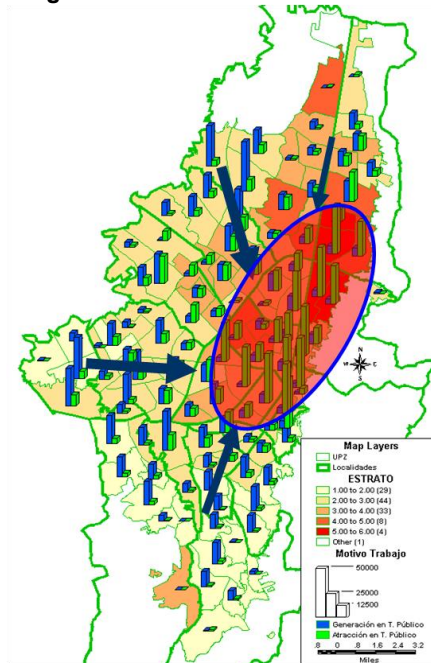


Fuente: Proyecto Metro de Bogotá

3.5 ANALISIS ESTADISTICO DE LA MUESTRA

La zonificación tiene un rol importante al momento de programar las labores de campo, por lo tanto se realizó una apropiada zonificación como se muestra en la Figura 3.4.

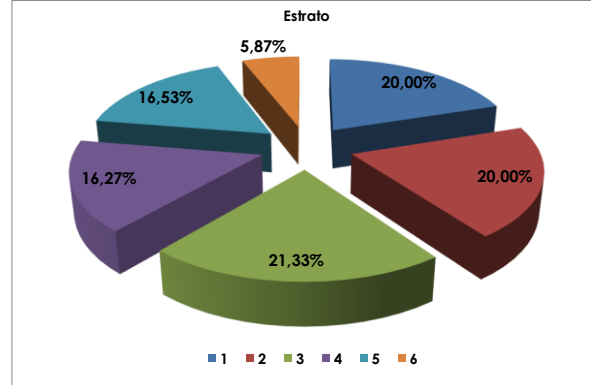
Figura 3.4 Zonificación del estudio



Fuente: Proyecto Metro de Bogotá

La distribución de la muestra según estrato se puede apreciar en la Figura 3.5. La población se encuentra distribuida de la siguiente manera los estratos 1 y 2 representan el 20% de la muestra, mientras que el estrato 3 el 21,33%, el estrato 4 el 16,27%, el estrato 5 el 16,53% y el estrato 6 el 5,87%.

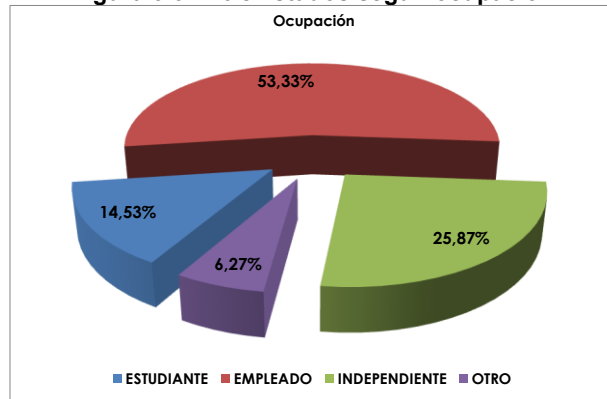
Figura 3.5 Entrevistados según estrato socioeconómico



Fuente: Proyecto Metro de Bogotá

Las personas entrevistadas según ocupación se distribuyen un 53,33% se encuentra desempeñando una labor en una empresa bien sea de carácter público o privado, un 14,53% es estudiante, el 25,87% independiente y el 6,27% realiza otra actividad, se muestra en la Figura 3.6.

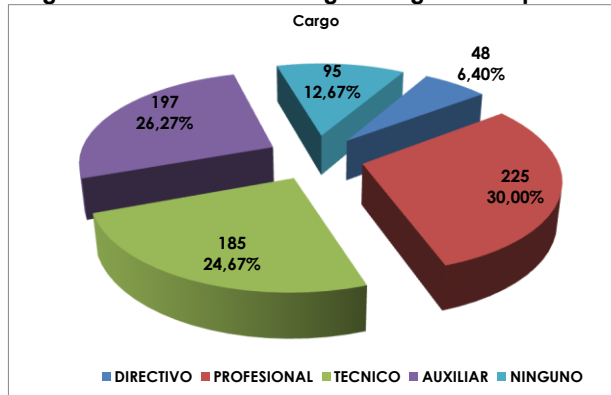
Figura 3.6 Entrevistados según ocupación



Fuente: Proyecto Metro de Bogotá

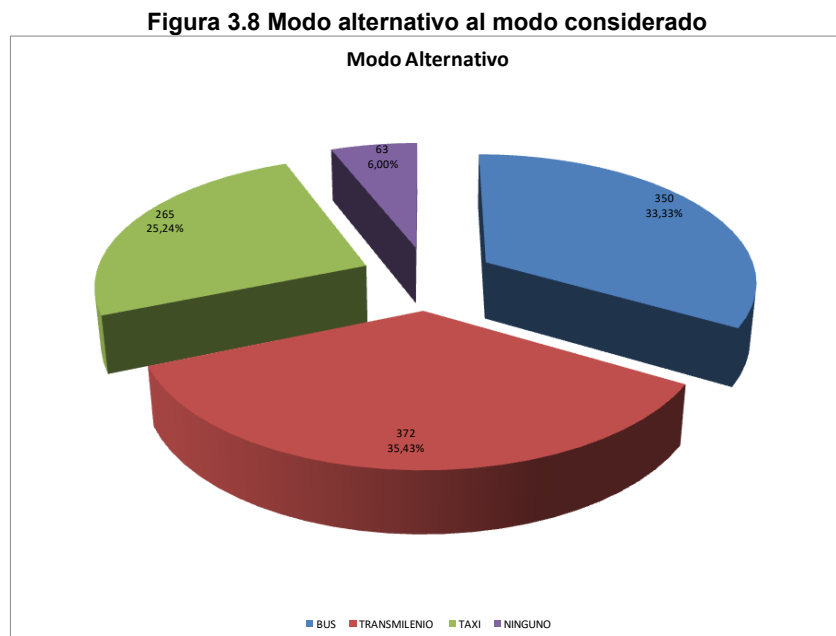
Referente al nivel educativo, se observa que un 30,00% de los encuestados tienen una educación profesional, contrastando con 12,67% que no tienen educación alguna, se muestra en la Figura 3.7.

Figura 3.7 Encuestados según cargo desempeñado



Fuente: Proyecto Metro de Bogotá

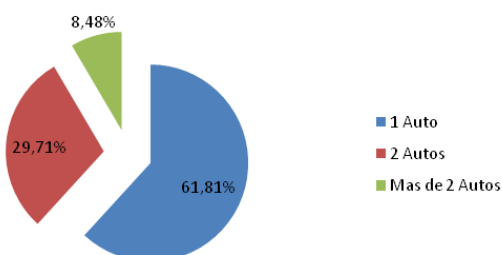
En el caso de la encuesta realizada a usuarios de transporte privado, se les consulta el modo alternativo considerado cuando no tienen acceso al auto, las respuestas son las que se presentan en la Figura 3.8. Allí se observa que en primera instancia se considera Transmilenio, seguido muy de cerca por el bus y luego el taxi.



Fuente: Proyecto Metro de Bogotá

En referencia a la motorización, la Figura 3.9 indica que el 62% de los entrevistados posee un vehículo, el 30% dos vehículos y un 8% posee tres o más autos en su hogar. Como era de esperar, la mayor parte de estos últimos corresponde a los estratos 5 y 6; esto es, los de mayores ingresos.

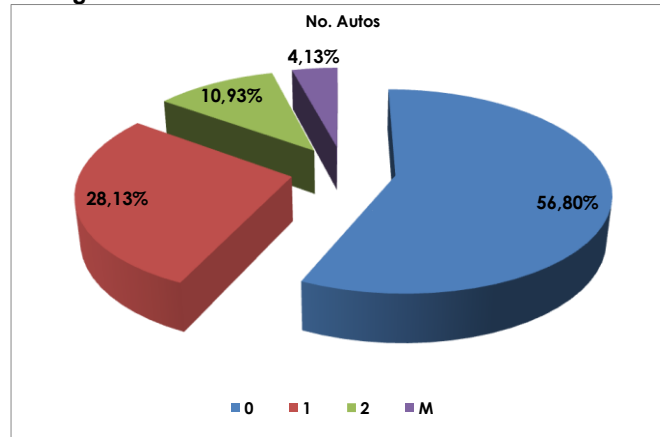
Figura 3.9 Motorización de usuarios de transporte privado
Número de Autos



Fuente: Proyecto Metro de Bogotá

Ahora bien, tomando toda la muestra el porcentaje descrito varia considerablemente, en la Figura 3.10 se aprecia que el porcentaje mayor no tiene vehículos más de la mitad, seguido por los poseedores de un vehículo con el 28,13%.

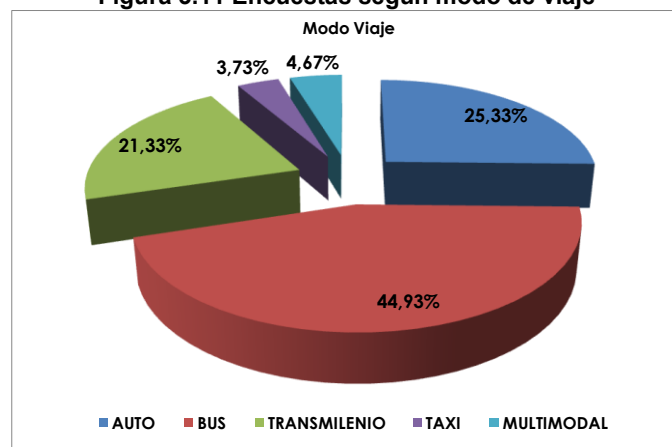
Figura 3.10 Motorización de todos los encuestados



Fuente: Proyecto Metro de Bogotá

Por otra parte para la muestra total se puede apreciar que el 21.33% de los encuestados elige el Transmilenio como alternativa de transporte por encima de otras opciones, ver Figura 3.11.

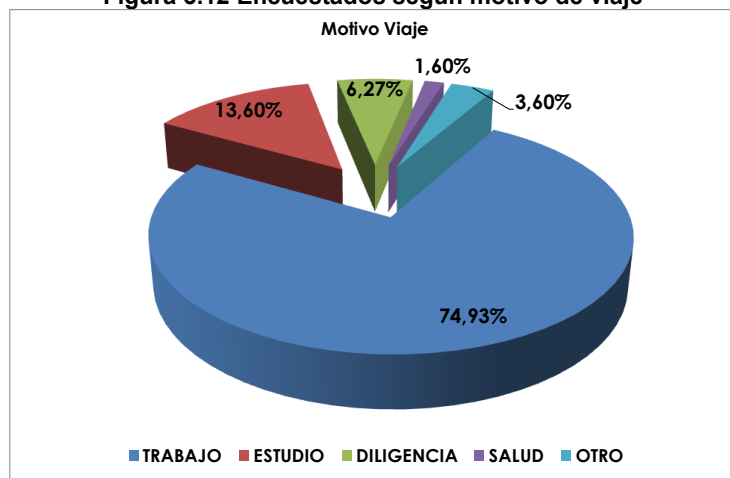
Figura 3.11 Encuestas según modo de viaje



Fuente: Proyecto Metro de Bogotá

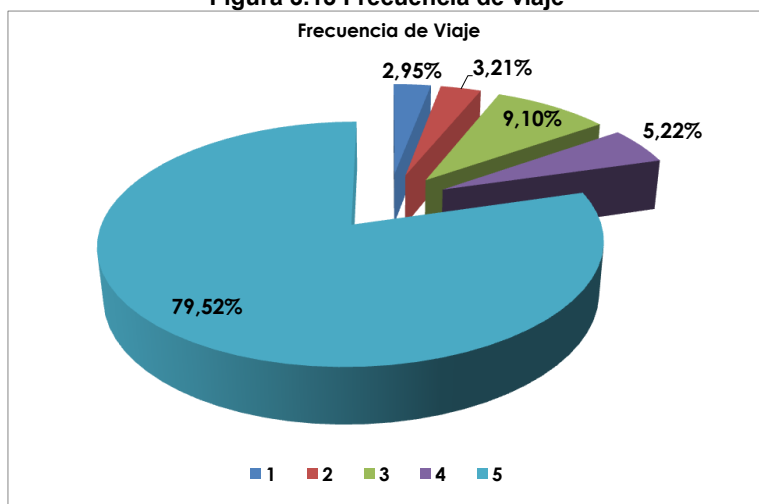
En la Figura 3.12 se puede apreciar la distribución en la muestra según el motivo de viaje el cual en su mayoría es motivo trabajo en un 75%. Según la frecuencia de viajes de este mismo tipo que hace en la semana se aprecia que al ser la mayor parte de los viajes por motivo trabajo la frecuencia en un porcentaje similar a cinco veces a la semana, se muestra en la Figura 3.13.

Figura 3.12 Encuestados según motivo de viaje



Fuente: Proyecto Metro de Bogotá

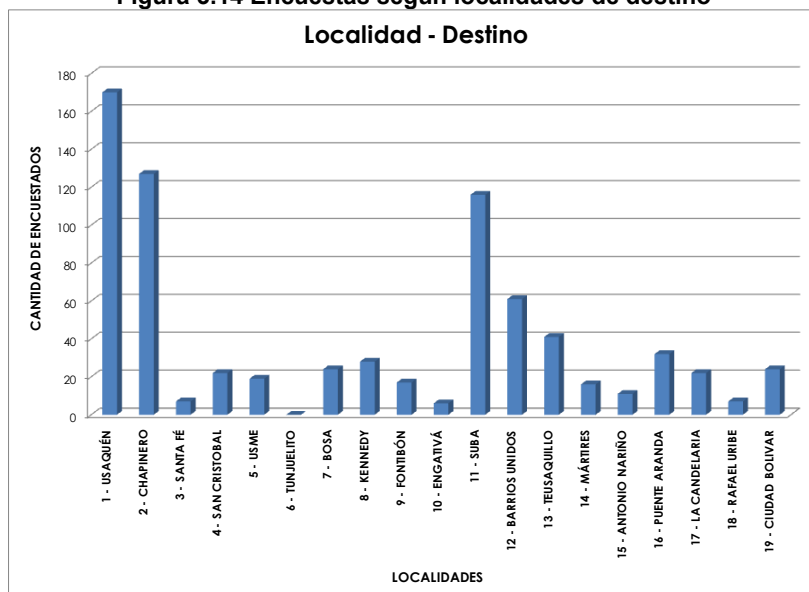
Figura 3.13 Frecuencia de viaje



Fuente: Proyecto Metro de Bogotá

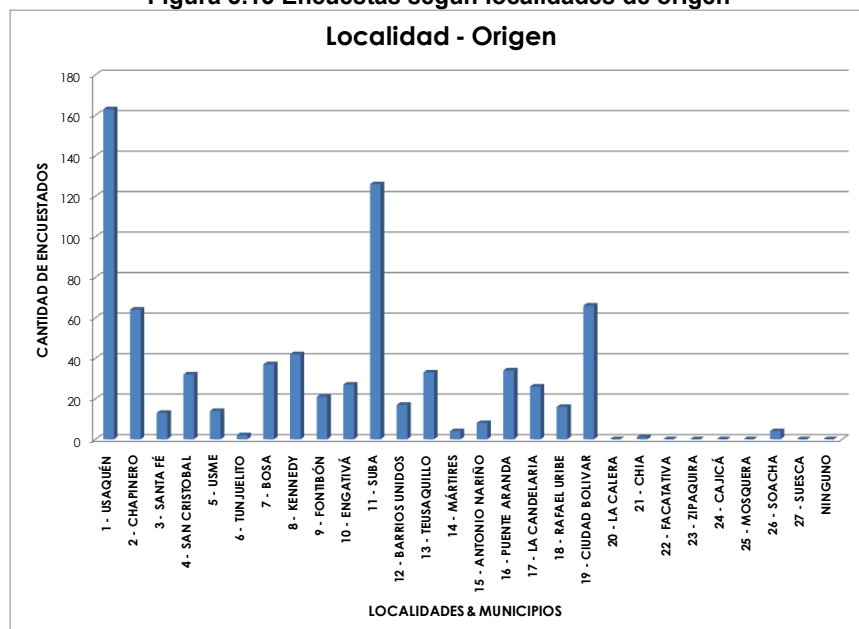
Las encuestas realizadas según origen o destino se analizaron según las localidades UPZ como se puede ver en la Figura 3.14 y Figura 3.15.

Figura 3.14 Encuestas según localidades de destino



Fuente: Proyecto Metro de Bogotá

Figura 3.15 Encuestas según localidades de origen



Fuente: Proyecto Metro de Bogotá

Puede notarse que las tres (3) principales zonas de atracción de viajes para nuestros encuestados son Usaqué, Chapinero y Suba respectivamente. De la misma forma son las zonas con mayor generación de viaje, incluyendo en esta categoría a Ciudad Bolívar.

En cuanto a tiempo de viaje en vehículo se analizó la muestra de tal forma que se tuviera desagregado por modo.

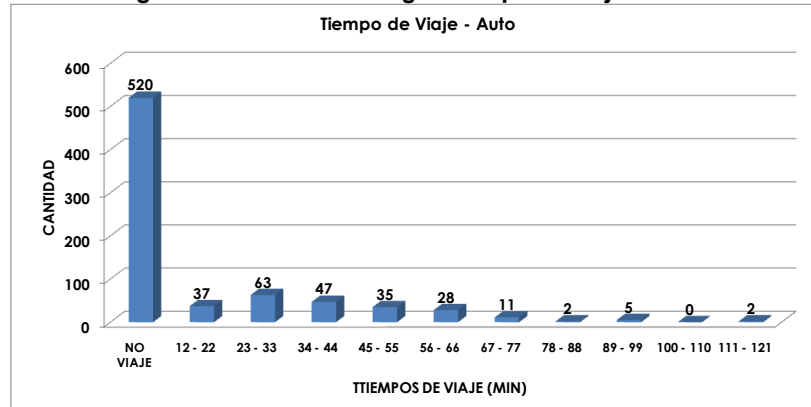
Tabla 3.5 Tiempo de viaje en auto

Tiempo de Viaje - Auto							Longitud Intervalo
No.	Intervalos Tiempo (min)	Intervalos	Xi	ni	fi		
1	0 - 11	NO VIAJE	6	520	0,69		
2	12 - 22	12 - 22	17	37	0,05	11	
3	23 - 33	23 - 33	28	63	0,08		
4	34 - 44	34 - 44	39	47	0,06		
5	45 - 55	45 - 55	50	35	0,05		
6	56 - 66	56 - 66	61	28	0,04		
7	67 - 77	67 - 77	72	11	0,01		
8	78 - 88	78 - 88	83	2	0,00		
9	89 - 99	89 - 99	94	5	0,01		
10	100 - 110	100 - 110	105	0	0,00		
11	111 - 121	111 - 121	116	2	0,00		
			Σ	750	1,00		

Fuente: Proyecto Metro de Bogotá

Los tiempos de viaje en auto son en su mayoría trayectorias cortas, de menos de 15 minutos, con un 70% aproximadamente del total de los viajes en vehículo particular.

Figura 3.16 Encuestas según tiempo de viaje en auto



Fuente: Proyecto Metro de Bogotá

De la misma forma los viajes en Transmilenio son cortos, aunque hay una cantidad considerable de viajes largos (de más de 30 minutos).

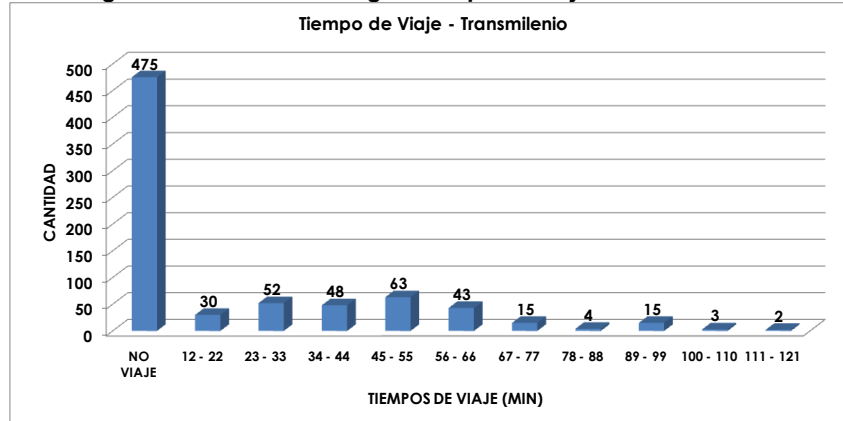
Tabla 3.6 Tiempos de viaje en Transmilenio

Tiempo de Viaje - Transmilenio						
No.	Intervalos Tiempo (min)		Intervalos	Xi	ni	fi
1	0	- 11	NO VIAJE	6	475	0,63
2	12	- 22	12 - 22	17	30	0,04
3	23	- 33	23 - 33	28	52	0,07
4	34	- 44	34 - 44	39	48	0,06
5	45	- 55	45 - 55	50	63	0,08
6	56	- 66	56 - 66	61	43	0,06
7	67	- 77	67 - 77	72	15	0,02
8	78	- 88	78 - 88	83	4	0,01
9	89	- 99	89 - 99	94	15	0,02
10	100	- 110	100 - 110	105	3	0,00
11	111	- 121	111 - 121	116	2	0,00
				Σ	750	1.00

Longitud Intervalo
11

Fuente: Proyecto Metro de Bogotá

Figura 3.17 Encuestas según tiempo de viaje en Transmilenio



Fuente: Proyecto Metro de Bogotá

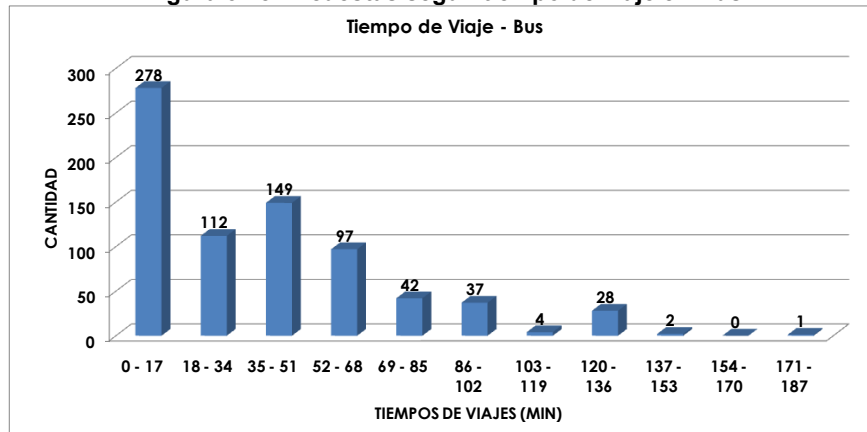
Retomando el concepto aplicado en la encuesta de viajes cortos y largos, menores de 30 minutos y mayores de 30 minutos, respectivamente. En el caso de tiempos de viaje en bus, la mayoría de estos son trayectos largos en aproximadamente el 60% de la muestra.

Tabla 3.7 Tiempos de viaje en Bus

Tiempo de Viaje - Bus						Longitud Intervalo
No.	Intervalos Tiempo (min)	Intervalos	Xi	ni	fi	
1	0 - 17	0 - 17	9	278	0,37	17
2	18 - 34	18 - 34	26	112	0,15	
3	35 - 51	35 - 51	43	149	0,20	
4	52 - 68	52 - 68	60	97	0,13	
5	69 - 85	69 - 85	77	42	0,06	
6	86 - 102	86 - 102	94	37	0,05	
7	103 - 119	103 - 119	111	4	0,01	
8	120 - 136	120 - 136	128	28	0,04	
9	137 - 153	137 - 153	145	2	0,00	
10	154 - 170	154 - 170	162	0	0,00	
11	171 - 187	171 - 187	179	1	0,00	
Σ				750	1,00	

Fuente: Proyecto Metro de Bogotá

Figura 3.18 Encuestas según tiempo de viaje en Bus



Fuente: Proyecto Metro de Bogotá

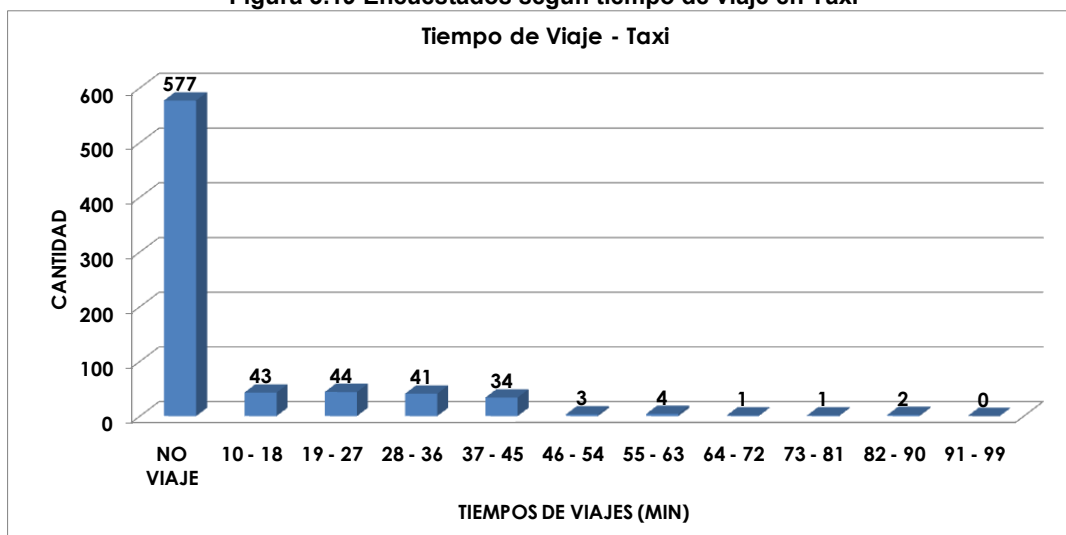
Similar al tiempo de viaje en auto, el equivalente a éste en transporte público, el taxi, tiene la muestra en su mayoría con valores del tiempo menores de 30 minutos, trayectos cortos para el efecto del presente trabajo.

Tabla 3.8 Tiempos de viaje en Taxi

Tiempo de Viaje - Taxi							Longitud Intervalo 9
No.	Intervalos Tiempo (min)		Intervalos	Xi	ni	fi	
1	0	- 9	NO VIAJE	4,5	577	0,77	
2	10	- 18	10 - 18	13,5	43	0,06	
3	19	- 27	19 - 27	22,5	44	0,06	
4	28	- 36	28 - 36	31,5	41	0,05	
5	37	- 45	37 - 45	40,5	34	0,05	
6	46	- 54	46 - 54	49,5	3	0,00	
7	55	- 63	55 - 63	58,5	4	0,01	
8	64	- 72	64 - 72	67,5	1	0,00	
9	73	- 81	73 - 81	76,5	1	0,00	
10	82	- 90	82 - 90	85,5	2	0,00	
11	91	- 99	91 - 99	94,5	0	0,00	
				Σ	750	1,00	

Fuente: Proyecto Metro de Bogotá

Figura 3.19 Encuestados según tiempo de viaje en Taxi



Fuente: Proyecto Metro de Bogotá

4 FORMULACION Y EVALUACION DE MODELOS

Con la información recolectada luego de la aplicación del experimento fue posible estimar modelos de elección discreta. Se probaron los modelos con los datos de las encuestas de preferencia revelada (PR), de preferencia declarada (PD) y una combinación de ambas.

En el caso de modelos PR los valores de los atributos fueron en la mayoría de los casos contrarios a los esperados, es decir, los signos de los parámetros no eran consistentes, de igual forma no eran significativos para ningún caso, por lo cual se decidió no incluirlos en el análisis del presente estudio. Lo mismo ocurrió con modelos de datos mixtos, ya que, aunque los modelos PD son consistentes, sus parámetros en ocasiones perdían significancia y consistencia al combinarlos con datos PR.

La razón de que estos modelos no funcionaran radicó principalmente en el hecho de no tener mediciones confiables de los modos alternativos al usado comúnmente para el viaje, no siendo posible por lo tanto obtener buenos modelos.

Por otro lado, para el caso de datos de preferencia declarada, se han considerado estimaciones dentro de la familia de los conocidos como “modelos generalizados de valor extremo” (Bierlaire, Bolduc, & McFadden, 2007), incluyendo Logit Multinomial (MNL), Logit Jerárquico (HL), y Logit Mixto (ML). Los modelos se evaluaron considerando criterios econométricos que involucran consistencia estadística y microeconómica (Ortúzar & Willumsen, 2001). Se efectuaron estimaciones en atención a las características socioeconómicas de los individuos. Así, se obtuvieron valores diferentes de las disposiciones marginales a pagar por un atributo (tiempo, confort) para diferentes tipos de usuarios.

Con la información tomada del trabajo de campo de las encuestas de PD, se procedió a estimar modelos de partición modal. Los datos recogidos permiten obtener información de los modos disponibles por los individuos y del hogar, de sus características socioeconómicas, de los atributos de las alternativas y de las características del viaje, tal como se sintetiza en la Tabla 4.1.

Tabla 4.1 Datos disponibles para estimación de modelos

Características del Individuo	Características Socioeconómicas del Hogar	Atributos de Alternativas	Características del Viaje
Genero • 1 hombre, 0 mujer Edad Estrato Socioeconómico • 1 a 6 Ocupación • Empleado • Independent e • Estudiante • Otro	Estrato Socioeconómico • 1 a 6 Número de Autos Personas en el hogar	Tiempo de Viaje (min) Costo (\$COL) Tiempo de Acceso (min) Número de Transbordos Confort -Fiabilidad	Motivo • Trabajo • Estudio • Otro Frecuencia Hora Origen y Destino Modo elegido • Auto • SITP Metro • SITP Transmilenio • Taxi • Bus

Fuente: Elaboración propia

4.1 ESPECIFICACIÓN DE MODELOS

Los atributos de las alternativas fueron definidos como sigue:

- Tiempo de viaje: Es el tiempo (minutos) mientras el vehículo está en movimiento.
- Tiempo de acceso. Es el tiempo (minutos) para acceder al vehículo, incluye caminata antes y después de abordarlo y sumado al tiempo de espera en el caso del taxi y del bus.
- Costo: Es el costo (pesos colombianos) del viaje. En el caso de transporte público y taxi es la tarifa, en el transporte privado incluye costos de combustible y estacionamiento.
- Transbordos. Este atributo aplica al transporte público y se relaciona con el número de transbordos entre vehículos durante el viaje.
- Confort-Confiabilidad. Es una variable cualitativa, medida dependiendo de la confiabilidad y regularidad de la frecuencia del transporte público y la posibilidad de ir o no sentado. Para unos casos, cuando se orientó a los estratos 4-5-6, el énfasis en la pregunta estaba en la confiabilidad del servicio, en tanto que en los estratos 1-2-3 se hizo mayor énfasis en la posibilidad de ir sentado.

A partir de los datos, se evaluaron distintos modelos, con distintas características en la definición de la matriz de covarianza, y en la forma funcional de la función de utilidad. Cada uno de los modelos es escrutado en correspondencia con criterios estadísticos y de consistencia microeconómica para finalmente escoger los que serán utilizados en la modalidad predictiva.

El proceso de evaluación de los modelos estimados implicó realizar las siguientes verificaciones:

- Consistencia de signos. Considerando la teoría microeconómica, se espera que en las variables tiempo, costo y transbordos la utilidad marginal sea negativa; por tanto, el signo de los parámetros en esos casos debe tener tal. En el caso del número de autos, debería tener una marginalidad positiva para la utilización del transporte privado. Debe indicarse que durante el análisis y selección de variables debe distinguirse aquellas que consideran relevantes o “*de política*”, como el costo, el tiempo de viaje y los transbordos, de aquellas que son complementarias (como el género).
- Nivel de significancia de los parámetros. La prueba t permite evaluar la hipótesis nula de que los parámetros sean significativamente igual a cero. Si ese es el caso, al nivel de significación escogido, se preferirá, por parsimonia, prescindir de ella. Si se desea un nivel de significancia del 95%, el valor crítico es 1,96; a su vez, para el 90% tal valor es 1,64.
- Prueba de razón de verosimilitud. Esta prueba permite evaluar dos modelos en el cual uno de ellos es una versión restringida del otro. Si la prueba estadística corrobora la hipótesis nula de que dos modelos son equivalente, se preferirá el modelo más simple, nuevamente por parsimonia. Para un nivel de significancia del 95%, el valor crítico del test chi-cuadrado para un grado de libertad es 3,85, y para dos grados de libertad es 5,99. Para otros grados de libertad puede consultarse en las tablas estadísticas correspondientes.
- Índice rho cuadrado. Si bien los resultados de este índice deben ser evaluados cuidadosamente, permiten comparar el ajuste general de dos modelos estimados a partir de la misma muestra. Se preferirá el que tenga mayor índice.

En primera instancia se estimaron modelos solo con la base de datos generada de la primera toma de información, 1125 encuestados, los cuales son usuarios de transporte privado. Posteriormente, se estiman los modelos para usuarios de transporte público que no tienen disponible vehículo particular para realizar su viaje. De esta manera, los primeros modelos estimados tienen 4 alternativas: auto, metro, BRT (Transmilenio) y taxi, mientras que el segundo juego para 787 individuos encuestados tiene 3 alternativas disponibles: bus, BRT (Transmilenio) y Metro.

Para las dos bases de datos se estiman los modelos por medio de los dos métodos descritos en capítulos anteriores: la estimación clásica o de máxima verosimilitud y la estimación bayesiana, para posteriormente realizar una comparación de las disposiciones a pagar obtenidas con los dos métodos.

4.1.1 Estimación Clásica

Para la estimación de los modelos a través de máxima verosimilitud, se utilizó el software libre Biogeme®².

² Este software se encuentra en www.epfl.ch/mbi/biogeme/ y con el se pueden formular los modelos de elección discreta de diferentes clases (Multinomial, Probit, Nested, entre otros).

Usuarios de Transporte Privado

En este esquema de modelación de PD, se definieron las variables de la siguiente manera:

V_{iq} es la utilidad sistemática de la alternativa i para el individuo q .

T_{iq} : Tiempo de viaje del modo i (min); $i = 1, 2, 3, 4$ para los modos auto, SITP-metro, taxi y SITP-Transmilenio, respectivamente.

Acc_{iq} : Tiempo de acceso del modo i (min)

$Trans_{iq}$: Número de transbordos. Aplica para los modos 2 y 4.

C_{iq} : Costo del modo i (\$)

$Sexo_q$: Variable que identifica el género del individuo, toma valor 1 para hombre y 0 para mujer.

$ConfB$: Variable muda que toma el valor de 1 si el confort es bajo, cero en otro caso.

$ConfM$: Variable muda que toma el valor de 1 si el confort es medio, cero en otro caso. Puede notarse que para este esquema de modelación se tomó como referente el nivel de confort alto.

δ_i = Variable muda que toma el valor de 1 para el modo i , cero para los restantes modos.

Z_{Est_456q} . Variable muda que identifica a los individuos que viven en uno de los estratos 4, 5 o 6; toma valor de 1 en tal caso, y 0 en otro caso. Se asocia con las personas de alto ingreso.

Z_{Est_3q} . Variable muda que identifica a los individuos que viven en el estrato 3; toma valor de 1 en tal caso, y 0 en otro caso.

Z_{Est_2q} . Variable muda que identifica a los individuos que viven en estrato 2; toma valor de 1 en tal caso, y 0 en otro caso.

A su vez, los parámetros a estimar son los siguientes.

ASC_i es la constante específica del modo i . Se ha tomado como referencia el modo auto, (modo 1) en el primer conjunto de datos y al igual que en la segunda base de datos, que para este caso es el bus, para el cual la constante modal específica toma el valor de cero.

β_C es la utilidad marginal del costo

β_T es la utilidad marginal del tiempo de viaje

$\beta_{TEst-456}$ es la utilidad marginal adicional del tiempo de viaje para los individuos de estrato alto; esto es, que pertenecen a alguno de los estratos 4, 5 o 6.

β_{TAcc} es la utilidad marginal del tiempo de acceso

$\beta_{TAccEst-456}$ es la utilidad marginal adicional del tiempo de acceso para los individuos de estrato alto; es decir, que pertenecen a alguno de los estratos 4, 5 o 6.

β_{No_A} es la utilidad marginal del número de autos. Aplica únicamente para el modo 1, auto.

β_{Sexo} Parámetro del sexo, aplica únicamente para los modos 2 y 4, SITP-metro y SITP-Transmilenio.

β_{Trans} es la utilidad marginal del número de transbordos. Aplica únicamente para los modos 2 y 4, SITP-metro y SITP-Transmilenio.

η_{iT} es un término de error aleatorio, que distribuye Normal $(0, \sigma^2)$, y que se ha añadido los modos 2 y 4, para el parámetro del tiempo de viaje. Esta componente de error permite correlacionar las observaciones repetidas de un mismo individuo.

η_{iTAcc} es el término de error aleatorio, que se ha añadido los modos 2 y 4, pero en este caso para el tiempo de acceso.

Se especificaron varios modelos Logit Multinomial (MNL), modelos con variaciones sistemáticas en los gustos (MNLVS) y un modelo logit mixto (ML) para toda la muestra, involucrando las variables antes indicadas. Finalmente se seleccionó un conjunto de ellos, de tal manera que las formas funcionales de las funciones de utilidad se especifican en la Tabla 4.2. Puede notarse que entre algunos pares de modelos se agrega una nueva variable explicatoria, de tal manera que es aplicable el test de razón de verosimilitud.

Tabla 4.2 Especificación de modelos

Modelo	Expresión
MNL1	$V_{iq} = ASC_i + \beta_T T_{iq} + \beta_C C_{iq} + \beta_{TAcc} Acc_{iq} + \delta_{2,4} \beta_{Trans} Trans_{2,4q}$
MNL11	$V_{iq} = ASC_i + \beta_T T_{iq} + \beta_C C_{iq} + \beta_{TAcc} Acc_{iq} + \delta_1 \beta_{No_A} No_A_q + \delta_{2,4} \beta_{Trans} Trans_{2,4q}$
MNL13	$V_{iq} = ASC_i + \beta_T T_{iq} + \beta_C C_{iq} + \beta_{TAcc} Acc_{iq} + \delta_1 \beta_{No_A} No_A_q + \delta_{2,4} (\beta_{Trans} Trans_{2,4q} + \beta_{Sexo} Sexo_q)$
MNLVS1	$V_{iq} = ASC_i + (\beta_T + \beta_{TEst_456} Z_{Est_456q}) T_{iq} + \beta_C C_{iq} + (\beta_{TAcc} + \beta_{TAceEst_456} Z_{Est_456q}) Acc_{iq} + \delta_1 \beta_{No_A} No_A_q + \delta_{2,4} (\beta_{Trans} Trans_{2,4q})$
MNLVS2	$V_{iq} = ASC_i + (\beta_T + \beta_{TEst_456} Z_{Est_456q}) T_{iq} + \beta_C C_{iq} + (\beta_{TAcc} + \beta_{TAceEst_456} Z_{Est_456q}) Acc_{iq} + \delta_1 \beta_{No_A} No_A_q + \delta_{2,4} (\beta_{Trans} Trans_{2,4q} + \beta_{Sexo} Sexo_q)$
MNLVS3	$V_{iq} = ASC_i + (\beta_T + \beta_{TEst_456} Z_{Est_456q}) T_{iq} + \beta_C C_{iq} + (\beta_{TAcc} + \beta_{TAceEst_456} Z_{Est_456q}) Acc_{iq} + \delta_1 (\beta_{No_A} No_A_q + \beta_{OEstud} O_{Estudq}) + \delta_{2,4} (\beta_{Trans} Trans_{2,4q} + \beta_{Sexo} Sexo_q)$
MNLVS4	$V_{iq} = ASC_i + (\beta_T + \beta_{TEst_456} Z_{Est_456q}) T_{iq} + \beta_C C_{iq} + (\beta_{TAcc} + \beta_{TAceEst_456} Z_{Est_456q}) Acc_{iq} + \delta_{2,4} (\beta_{Trans} Trans_{2,4q})$
ML1	$V_{iq} = ASC_i + (\beta_T + \beta_{TEst_456} Z_{Est_456q} + \eta_{iT}) T_{iq} + \beta_C C_{iq} + (\beta_{TAcc} + \beta_{TAceEst_456} Z_{Est_456q} + \eta_{iTAcc}) Acc_{iq} + \delta_1 \beta_{No_A} No_A_q + \delta_{2,4} (\beta_{Trans} Trans_{2,4q} + \beta_{Sexo} Sexo_q)$

Los modelos calibrados con todos los datos de ambas fuentes de información se presentan en la Tabla 4.3. En la mayoría de los casos se aprecia que los signos son consistentes y los parámetros son en su mayor parte significativos al 90%. Es importante aclarar que estos modelos fueron estimados con una muestra de 10.125 pseudo-individuos. En la columna parámetros se indican además las alternativas en cuya función de utilidad se encuentra cada uno de ellos: 1. Auto, 2. Metro, 3. Taxi, 4. SITP - BRT (Transmilenio).

Tabla 4.3 Modelos estimados para usuarios de transporte privado

	Parámetro	MNL1	MNL11	MNL13	MNLVS1	MNLVS2	MNLVS3	MNLVS4	ML1
Constantes Específicas	ASC1	0	0	0	0	0	0	0	0
	ASC2	-0,713 (-11,79)	-0,135 (-1,64)	-0,346 (-4,00)	-0,113 (-1,36)	-0,324 (-3,73)	-0,346 (-3,93)	-0,713 (-11,79)	-0,158 (-1,07)
	ASC3	-0,252 (-4,55)	0,322 (4,06)	0,304 (3,82)	0,349 (4,37)	0,334 (4,16)	0,313 (3,86)	-0,251 (-4,53)	0,068 (0,44)
	ASC4	-0,817 (8,85)	-0,247 (-2,29)	-0,484 (-4,32)	-0,226 (-2,09)	-0,463 (-4,13)	-0,472 (-4,20)	-0,816 (-8,83)	-0,0963 (-0,47)
Coefficientes de atributos	β_T (Tiempo) (1,2,3,4)	-0,0344 (-13,43)	-0,0344 (-13,37)	-0,0347 (-13,45)	-0,0263 (-7,62)	-0,0266 (-7,68)	-0,0266 (-7,68)	-0,0279 (-8,13)	-0,0401 (-5,07)
	β_C (Costo) (1,2,3,4)	-0,000255 (-14,66)	-0,000256 (-14,59)	-0,000260 (-14,77)	-0,000257 (-14,65)	-0,000261 (-14,83)	-0,000261 (-14,83)	-0,000256 (-14,68)	-0,000253 (-10,58)
	β_{TACC} (Tiempo de Acceso) (1,2,3,4)	-0,0418 (-9,46)	-0,0420 (-9,44)	-0,0425 (-9,52)	-0,0360 (-5,85)	-0,0375 (-6,07)	-0,0378 (-6,11)	-0,0326 (-5,33)	-0,0347 (-1,70)
	β_{TRANS} (Transbordo – TP) (2,4)	-0,299 (-8,19)	-0,302 (-8,23)	-0,305 (-8,28)	-0,302 (-8,23)	-0,305 (-8,28)	-0,305 (-8,29)	-0,305 (-8,29)	-0,423 (-8,27)
	β_{NO_A} (No Autos - Auto) (1)		0,376 (10,21)	0,388 (10,50)	0,393 (10,49)	0,407 (10,83)	0,410 (10,89)		0,440 (6,11)
	β_{SEX} (Sexo - TP) (2,4)			0,356 (8,34)		0,359 (8,41)	0,360 (8,41)		0,165 (1,89)
	β_{ocup_est} (Estudiante - A) (1)						-0,0908 (-1,66)		
	β_{T_E456} (Estratos 456) (1,2,3,4)				-0,0132 (-3,50)	-0,0134 (-3,54)	-0,0134 (-3,56)	-0,0103 (-2,76)	-0,0206 (-2,18)
	β_{TACC_E456} (Estratos 456) (1,2,3,4)				-0,00955 (-1,40)	-0,00808 (-1,18)	-0,00760 (-1,11)	-0,0147 (-2,17)	-0,0196 (-0,77)
	η_{IT}								-0,109 (-15,85)
	η_{ITAcc}								0,327 (18,08)
	# de parámetros	7	8	9	10	11	12	9	12
	Logverosimilitud $L(\theta)$	-6628,9	-6575,7	-6540,7	-6568,8	-6533,2	-6531,9	-6625,07	-5191,9
	Test Razón de Verosimilitud LR	650,766	775,238	827,227	771,056	842,197	844,961	658,547	4949,325
	ρ^2 Ajustado	0,046	0,053	0,058	0,054	0,059	0,059	0,046	0,252

Adicionalmente, se presenta para cada modelo el valor de las diferentes tasas marginales de sustitución entre cada atributo y el costo, que se asocia como la disposición a pagar por una mejora marginal en el valor del atributo considerado. Las tasas calculadas con el valor subjetivo del tiempo de viaje (VSTV), el valor subjetivo del tiempo de acceso (VSTAcc); estas dos, además, se evaluaron para segmentos de la población con distinto ingreso; uno para los estratos de más bajo ingreso, 1, 2 y 3, y otro para los estratos 4, 5 y 6, con mayor poder de compra. Otra tasa marginal estimada se relaciona con los

transbordos, que puede interpretarse como la penalidad en términos económicos de un transbordo o como la disposición a pagar para no realizarlo.

Tabla 4.4 Tasas marginales de sustitución

Modelo	MNL1	MNL11	MNL13	MNLVS1	MNLVS2	MNLVS3	MNLVS4	ML1
VSTV (\$/min)	134	134	133					
VSTAcc (\$/min)	163	163	164					
VSTV Estratos 23 (\$/min)				102	102	102	109	158
VSTAcc Estratos 23 (\$/min)				140	143	145	127	148
VST Estratos 456 (\$/min)				154	153	153	149	239
VSTAcc Estratos 456 (\$/min)				177	174	173	185	231
Transbordos (\$/Transbordo)	1172	1179	1173	1175	1169	1168	1168	1800
Transbordos (min/transbordo)	8,69	8,78	8,79					
Transbordos Estratos 23 (min/transbordo)				11,43	11,46	11,46	10,92	
Transbordos Estratos 456 (min/transbordo)				7,65	7,62	7,62	7,98	

Seguidamente se realiza un análisis de cada uno de los modelos obtenidos:

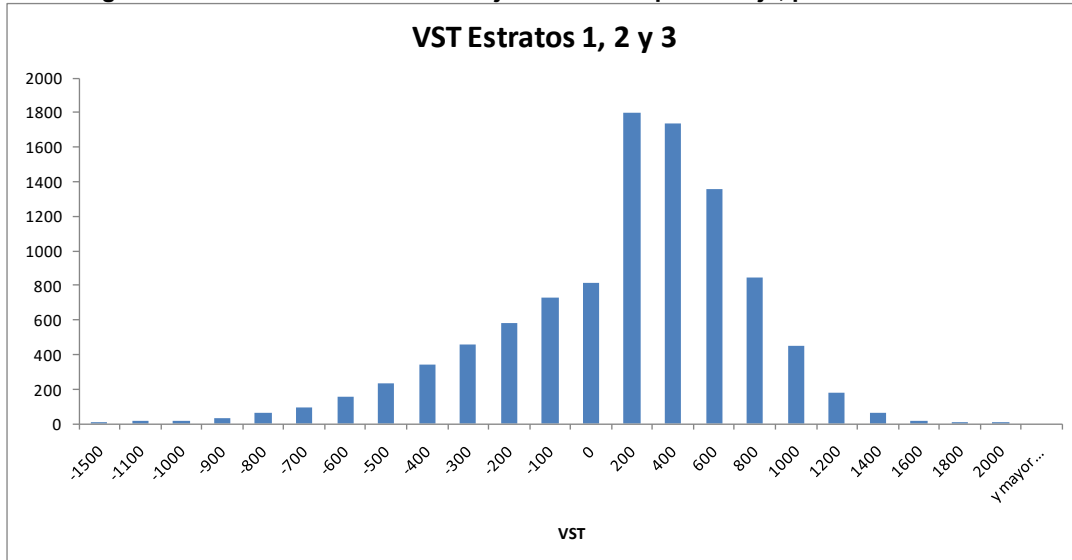
- **MNL1.** Es un modelo robusto donde todos los parámetros son significativos al 95%, y los signos corresponden a los esperados para cada uno. La estimación del valor del tiempo de viaje es de \$134/min, mientras que para el tiempo de acceso es \$163/min. El transbordo se penaliza en \$1172. Se puede ver claramente que la disposición a pagar por un ahorro en el tiempo de acceso es mayor que en el caso del tiempo de viaje, esto es de esperarse. Por otro lado, la penalidad del transbordo respecto al tiempo es de 8,69/min.
- **MNL11.** Al agregar al modelo anterior la variable número de autos mejora el poder predictivo. Las tasas marginales de sustitución obtenidas son parecidas a las del modelo anterior.
- **MNL13.** En este modelo se incorpora, además del número de autos, el género, que ha sido especificado para los modos 2 y 4 mejorando de manera importante la bondad de ajuste. Los valores del tiempo de viaje (\$133/min), del tiempo de acceso (\$164/min) y la disposición a pagar por los transbordos (\$1.173/min) son muy similares a los de los modelos anteriores.
- **MNLVS1.** Es un modelo Logit Multinomial, con similares características al MNL11, incorporando el estrato. En este caso, se tienen valores de las tasas marginales de sustitución según estrato, siendo en el caso del valor del tiempo \$102/min para los estratos 2 y 3, y \$154/min para los estratos 4, 5 y 6. Por otra parte, el valor del tiempo de acceso para los estratos 2 y 3 resultó ser \$140/min y \$177/min el de los estratos 4,5 y 6. La penalidad por cada transbordo es de \$1.175. Mientras que por minuto en el caso de los estratos bajos es de 11,43/min y para estratos altos de

7,65/min, contrario a lo que podría intuirse, que para el caso de los estratos altos las penalidades sean mayores.

- **MNLVS2.** Incorpora variaciones sistemáticas según ingreso y además la variable género, lo que mejora la bondad de ajuste. Los valores de las tasas marginales de sustitución son en el caso del valor del tiempo de viaje \$102/min para los estratos 2 y 3, y \$154/min para los estratos 4, 5 y 6. A su vez, el valor del tiempo de acceso para los estratos 2 y 3 resultó ser \$141/min y \$177/min para los estratos 4,5 y 6. La penalidad por cada transbordo es de \$1.169 y en caso de penalidad respecto al tiempo es similar al modelo anterior.
- **MNLVS3.** Incluye variaciones sistemáticas según ingreso y las variables género, y ocupación estudiante, esta última en el modo auto. El valor del parámetro no es significativo al 95%, por lo que se decide omitirlo en el próximo modelo. El valor del tiempo de viaje es \$109/min para los estratos 2 y 3, y \$153/min para los estratos 4, 5 y 6; por su parte, los valores análogos respectivos para el valor del tiempo de acceso son \$145/min y \$173/min. Cada transbordo se penaliza en \$1168.
- **MNLVS4.** Este modelo es similar al MNLV1 solo que sin incluir la variable numero de auto, ahora bien, comparando estos dos modelos se concluye que con este cambio no mejora la bondad de ajuste, por lo tanto, se incluirá entre las variables relevantes en el modelo el numero de autos. Los valores de las tasas marginales de sustitución son en el caso del valor del tiempo de viaje \$102/min para los estratos 2 y 3, y \$149/min para los estratos 4, 5 y 6. A su vez, el valor del tiempo de acceso para los estratos 2 y 3 resultó ser \$127/min y \$185/min para los estratos 4,5 y 6. La penalidad por cada transbordo es de \$1168.
- **ML1.** Este modelo considera, además de variaciones sistemáticas, una componente aleatoria para el tiempo de acceso y el tiempo de viaje. Con este modelo el procedimiento de cálculo de las tasas marginales de sustitución no es tan directo como en los casos anteriores. El comportamiento del VST no es una constante, sino por el contrario, sigue una distribución. Por su parte, el transbordo se penaliza en \$1800.

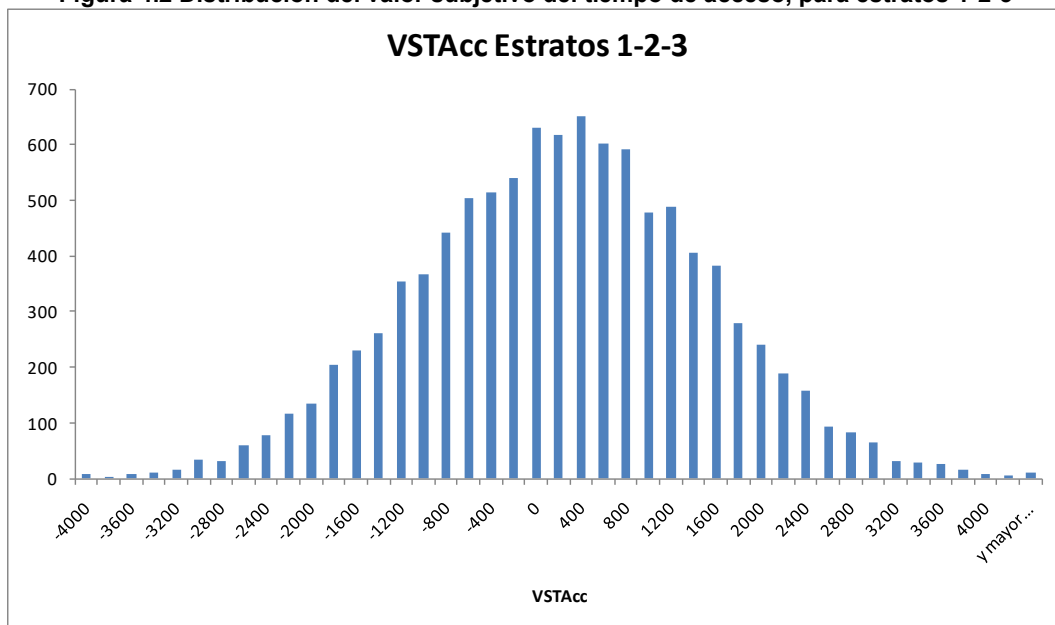
Para obtener las distribuciones de los diferentes parámetros se emplea simulación. El método consiste en obtener diferentes valores para el tiempo de viaje y el tiempo de acceso, de acuerdo con la media y la desviación estándar determinadas por el modelo en la Tabla 4.3. Esto se repite un gran número de veces, en este caso 10000 valores para cada uno de ellos, y se calcula el VST correspondiente para cada caso. En la Figura 4.1 y la Figura 4.2 se presentan las distribuciones resultantes para el VST y el VSTAcc para los estratos 1-2-3. Mientras que para los estratos 4-5-6 estas distribuciones se muestran en la Figura 4.3 y en la Figura 4.4.

Figura 4.1 Distribución del valor subjetivo del tiempo de viaje, para estratos 1-2-3



Fuente: Elaboración Propia

Figura 4.2 Distribución del valor subjetivo del tiempo de acceso, para estratos 1-2-3



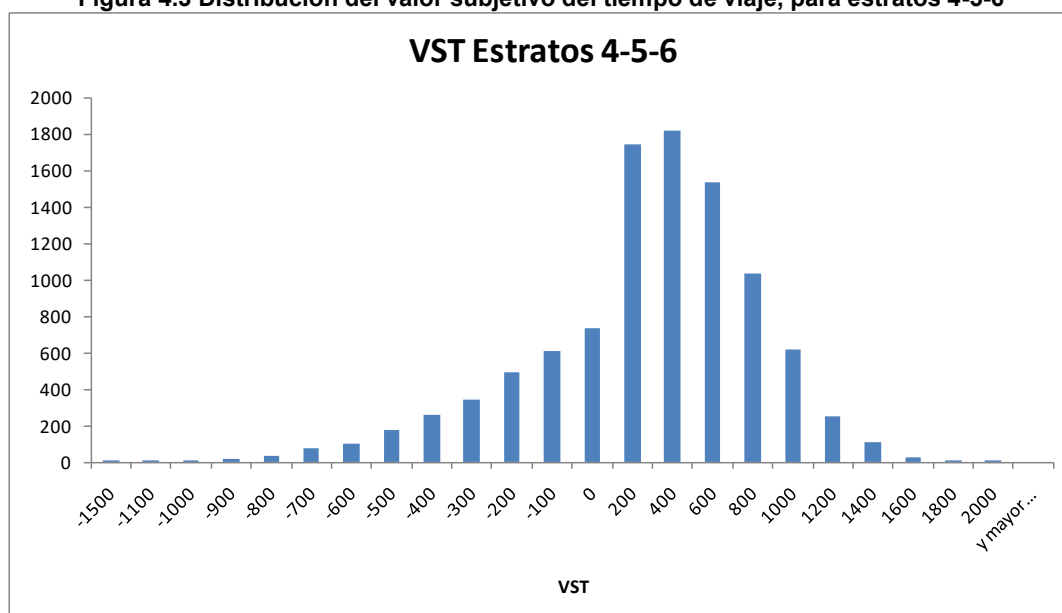
Fuente: Elaboración Propia

A pesar que los valores medios de los parámetros del tiempo de viaje y del tiempo de acceso, son contrarios a los esperados, es decir, menor valor para el parámetro del tiempo de acceso que del tiempo de viaje, se puede observar de la Figura 4.1 a la Figura 4.4 que el valor subjetivo del tiempo de viaje tiene una distribución más sesgada a valores positivos y muy pocos valores mayores a \$1.400/min. Mientras que en el caso del VST de acceso la distribución describe mejor la función normal y además tiene un porcentaje considerable de valores mayores a \$1.400/min. Cumpliéndose la premisa que en la población se tiene una percepción del VST de acceso mayor que en el caso del VST de viaje. En los usuarios de estratos bajos cerca del 36% presenta valores negativos.

En recientes investigaciones (Sillano & Ortúzar, 2005) (Hess, Bierlaire, & Polak, 2005) (Cirillo & Axhausen, 2006) se ha estudiado ampliamente la presencia de valores negativos o ceros en el VST. Algunas de las razones que plantean estos autores radican principalmente en aspectos como: aquellas personas que por alguna razón disfrutan el viaje (Cirillo & Axhausen, 2006) (les agrada la compañía o una conversación con un pasajero, le agrada el paisaje que ve diariamente en su viaje, ocurre el caso, donde el periodo de viaje es un periodo útil, que aprovechan para oír noticias o música, hablar por celular, leer), personas que por la naturaleza propia del viaje, el movimiento es el principal propósito (en este caso, hacer ejercicio o disfrutan manejar) o ya desde el punto de vista de la modelación, las funciones utilizadas para la simulación en caso de modelos logit mixto no son las adecuadas (Hess, Bierlaire, & Polak, 2005). En algunas de la investigaciones citadas, se estudia la premisa de cambiar el tipo de distribución, a una que limite los valores a un cierto rango, como es el caso de la lognormal.

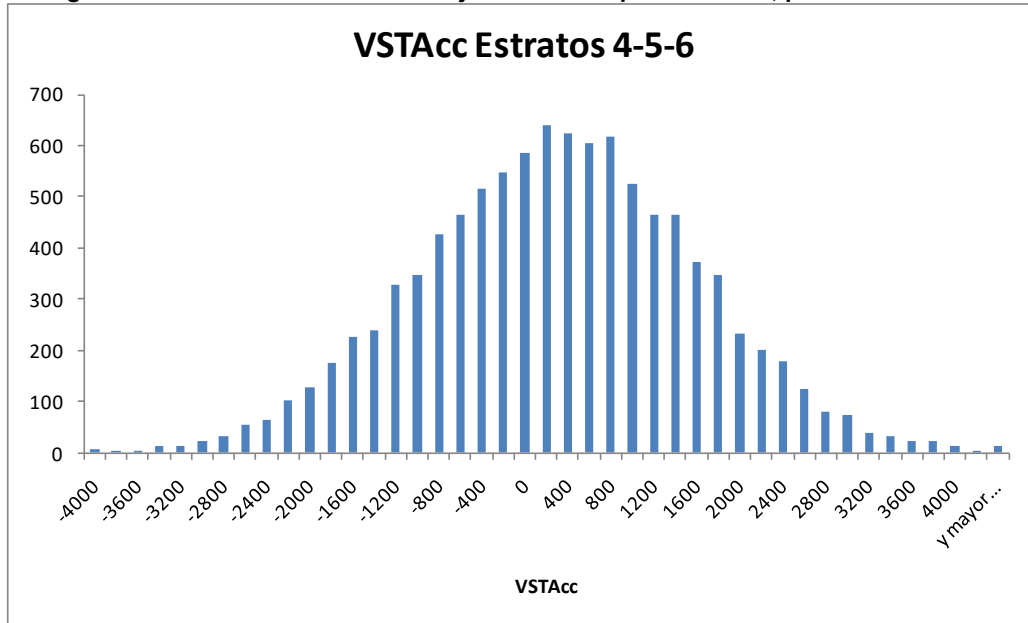
Podría existir otro tipo de variables latentes que son importantes para un porcentaje de individuos, que al no estar presentes en el modelo, distorsionan la distribución del VST para obligar matemáticamente al modelo a reproducir las elecciones. Por ejemplo la mayor sensación de seguridad en el auto respecto al transporte público.

Figura 4.3 Distribución del valor subjetivo del tiempo de viaje, para estratos 4-5-6



Fuente: Elaboración Propia

Figura 4.4 Distribución del valor subjetivo del tiempo de acceso, para estratos 4-5-6



Fuente: Elaboración Propia

Usuarios de Transporte Público

De la misma manera que en los modelos estimados para usuarios de transporte privado, se eligen aquellos modelos con mejores bondades de ajuste y parámetros significativos al 90%, por lo cual, se presentan solo tres modelos, el mejor de los modelos simples, el mejor de los modelos con variaciones sistemáticas en los gustos y el modelo logit mixto, con las mismas variables y parámetros a calcular para usuarios de transporte público, pero agregando la variable confort. En este caso los usuarios encuestados son 787, quienes tienen disponibles solo tres alternativas a saber: 1. Bus, 2. BRT (Transmilenio) y 3. Metro. Se definen dos parámetros adicionales:

β_{confB} es la utilidad marginal del confort bajo

β_{confM} es la utilidad marginal del confort medio

Las formas funcionales de la utilidad se especifican en la Tabla 4.5:

Tabla 4.5 Especificación de modelos para usuarios de transporte público

Modelo	Expresión
MNL13	$V_{iq} = ASC_i + \beta_T T_{iq} + \beta_C C_i + \beta_{TAcc} Acc_{iq} + \delta_1 \beta_{No_A} No_A_q + \delta_{2,4} (\beta_{Trans} Trans_{2,4q} + \beta_{Sexo} Sexo_q) + \beta_{confB} * ConfortB_{iq} + \beta_{confM} * ConfortM_{iq}$
MNLVS2	$V_{iq} = ASC_i + (\beta_T + \beta_{TEst_456} Z_{Est_456}) T_{iq} + \beta_C C_i + (\beta_{TAcc} + \beta_{TAccEst_456} Z_{Est_456}) Acc_{iq} + \delta_1 \beta_{No_A} No_A_q + \delta_{2,4} (\beta_{Trans} Trans_{2,4q} + \beta_{Sexo} Sexo_q) + \beta_{confB} * ConfortB_{iq} + \beta_{confM} * ConfortM_{iq}$
ML1	$V_{iq} = ASC_i + (\beta_T + \beta_{TEst_456} Z_{Est_456} + \eta_{iT}) T_{iq} + \beta_C C_i + (\beta_{TAcc} + \beta_{TAccEst_456} Z_{Est_456} + \eta_{iTAcc}) Acc_{iq} + \delta_1 \beta_{No_A} No_A_q + \delta_{2,4} (\beta_{Trans} Trans_{2,4q} + \beta_{Sexo} Sexo_q) + \beta_{confB} * ConfortB_{iq} + \beta_{confM} * ConfortM_{iq}$

Los modelos calibrados se presentan en la Tabla 4.6. Estos modelos fueron estimados con una muestra de 6.296 pseudo-individuos.

Tabla 4.6 Modelos estimados para usuarios de transporte público

	Parámetro	MNL13	MNLVS2	ML1
Constantes Específicas	ASC1	0	0	0
	ASC2	-0,247 (-4,78)	-0,286 (-5,48)	-0,0937 (-1,53)
	ASC3	-0,592 (-6,56)	-0,627 (-6,88)	-0,727 (-6,02)
Coefficientes de atributos	β_T (Tiempo) (1,2,3)	-0,0390 (-12,36)	-0,0337 (-10,29)	-0,0887 (-9,96)
	β_C (Costo) (1,2,3)	-0,00149 (-12,89)	-0,00149 (-12,88)	-0,00216 (-15,00)
	β_{TACC} (Tiempo de Acceso) (1,2,3)	-0,0378 (-7,46)	-0,0411 (-7,39)	-0,0543 (-6,06)
	β_{TRANS} (Transbordo – TP) (1,2,3)	-0,202 (-5,69)	-0,204 (-5,73)	-0,303 (-6,83)
	β_{NO_A} (No Autos - Auto) (1)	0,430 (12,89)	0,321 (8,88)	0,201 (4,27)
	β_{SEX} (Sexo - TP) (2)	-0,154 (-2,75)	-0,165 (-2,95)	-0,219 (3,38)
	β_{T_E456} (Estratos 456) (1,2,3)		-0,0232 (-5,41)	-0,0718 (-5,58)
	β_{TACC_E456} (Estratos 456) (1,2,3)		0,00905 (1,09)	0,00502 (0,41)
	$\beta_{ConfortB}$ (Confort Bajo)	-0,332 (-6,78)	-0,345 (-7,00)	-0,815 (-12,12)
	$\beta_{ConfortM}$ (Confort Medio)	-0,0936 (-2,12)	-0,102 (-2,30)	-0,397 (-7,24)
	η_{IT}			0,134 (19,60)
	η_{ITAcc}			-0,0827 (-9,90)
	# de parámetros	10	12	14
	Logverosimilitud $L(\theta)$	-6305,381	-6266,834	-5463,953
	Test Razón de Verosimilitud LR	1222,965	1300,058	2905,820
	ρ^2 Ajustado	0,087	0,092	0,208

El modelo logit simple (MNL13) muestra que todos los parámetros son significativos y con los signos esperados. Sin embargo, el parámetro del tiempo de viaje es un poco mayor al del tiempo de acceso contrario a las premisas estudiadas sobre este tema. No obstante, al analizar el modelo MNLVS2 donde se incluyen las variaciones sistemáticas en esas variables respecto al estrato socioeconómico mejora la bondad de ajuste y además se aprecia que el parámetro del acceso es mayor que el tiempo de viaje.

De igual forma que para los usuarios de transporte privado, se presentan las tasas marginales de sustitución para cada modelo en la Tabla 4.7.

En el primer modelo se comprueba lo encontrado anteriormente, donde los valores subjetivos son iguales para el tiempo y el acceso. Sin embargo, el VST del acceso en el modelo MNLVS2 es de \$28/min, mayor al VST de viaje el cual es \$23/min.

Por otro lado, las tasas marginales de sustitución del transbordo son para cada modelo \$136, \$137 y \$140, respectivamente. La penalidad del transbordo respecto al tiempo de viaje es evaluada y presentada igualmente en la Tabla 4.7.

Tabla 4.7 Tasas marginales de sustitución usuarios de transporte público

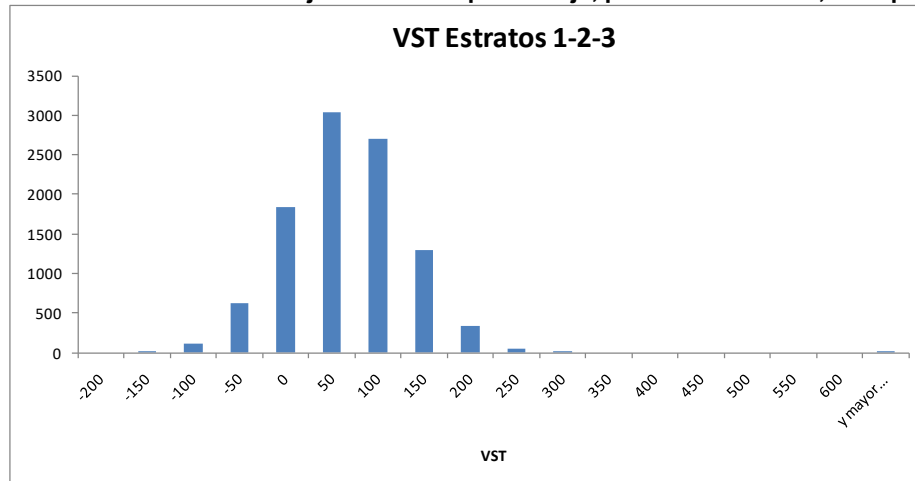
Modelo	MNL13	MNLVS2	ML1
VSTV (\$/min)	30		
VSTAcc (\$/min)	30		
VSTV Estratos 123 (\$/min)		23	
VSTAcc Estratos 123 (\$/min)		28	
VST Estratos 456 (\$/min)		38	
VSTAcc Estratos 456 (\$/min)		22	
Transbordos (\$/Transbordo)	136	137	140
Transbordos (min/transbordo)	5,18		
Transbordos Estratos 123 (min/transbordo)		6,05	
Transbordos Estratos 456 (min/transbordo)		3,58	

Cabe resaltar que al igual que en los modelos para usuarios de transporte privado la penalidad por transbordo respecto al tiempo son mayores para los estratos bajos.

En estos modelos al incluirse una variable adicional como el confort la cual es una variable latente, se estudian sólo las disposiciones a pagar por pasar de un confort al otro, ya que aproximaciones de variables latentes se sale del alcance de este proyecto. Por lo tanto, se tiene que la disposición a pagar para pasar del confort bajo al confort alto es de \$222,8. Por pasar del confort medio al alto es de \$62,81 y para pasar del confort bajo al medio se penaliza en \$3,54.

Las distribuciones de los parámetros tiempo de viaje y tiempo de acceso para el modelo logit mixto se presentan en la Figura 4.5 a la Figura 4.8. En estas gráficas se observa que para usuarios de transporte público la penalidad del valor subjetivo del tiempo de acceso es menor que para el tiempo de viaje. Podría ser consecuencia que estos usuarios no tienen disponibles los modos auto y taxi, y los tiempos de espera y de caminata para ellos son por lo general mayores que para los usuarios de transporte privado. Los porcentajes de valores negativos es menor en comparación con los usuarios de transporte privado, aproximadamente del 26% para estratos bajos y del 12% en estratos altos.

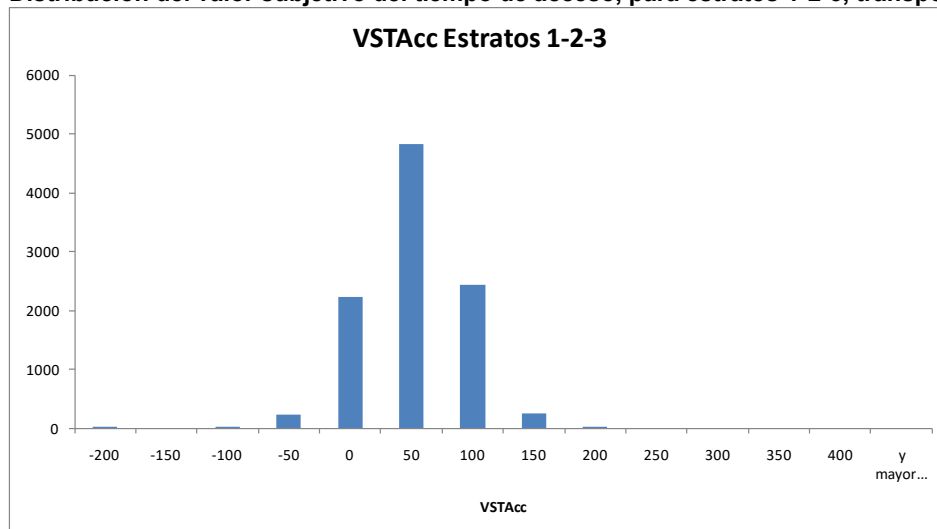
Figura 4.5 Distribución del valor subjetivo del tiempo de viaje, para estratos 1-2-3, transporte público



Fuente: Elaboración Propia

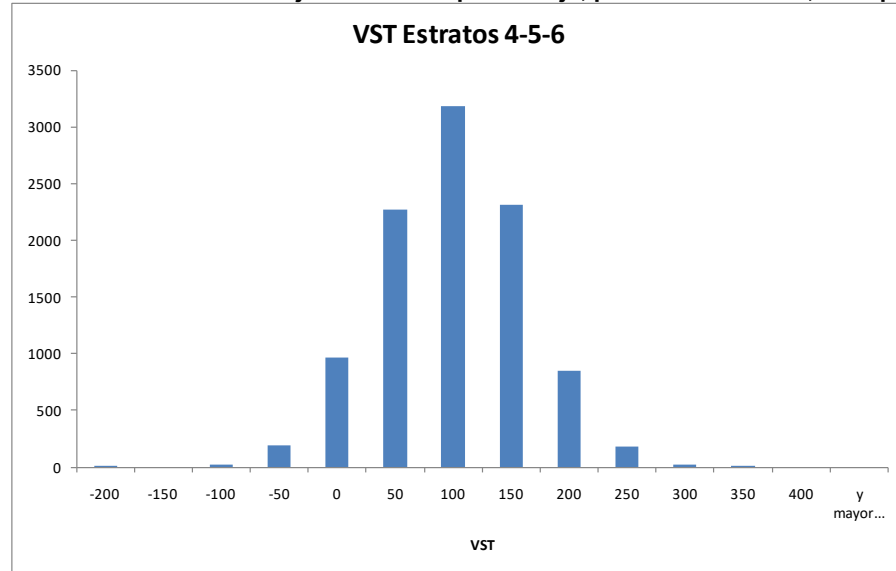
Se observa además que para este tipo de usuarios, los cuales no tienen disponible vehículo particular o en cualquier caso el taxi, las percepciones de esta población en general son muy similares, con una tendencia marcada a valores entre los \$50/min y \$200/min, valores subjetivos del tiempo razonables para estos casos.

Figura 4.6 Distribución del valor subjetivo del tiempo de acceso, para estratos 1-2-3, transporte público



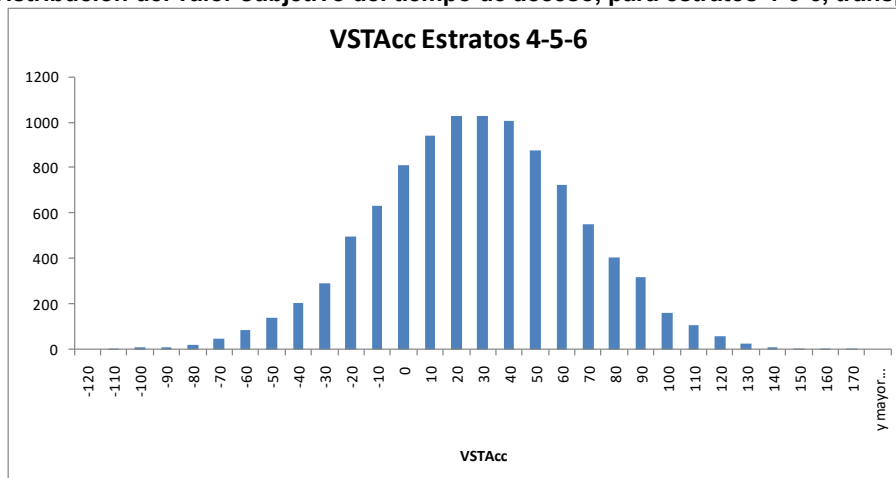
Fuente: Elaboración Propia

Figura 4.7 Distribución del valor subjetivo del tiempo de viaje, para estratos 4-5-6, transporte público



Fuente: Elaboración Propia

Figura 4.8 Distribución del valor subjetivo del tiempo de acceso, para estratos 4-5-6, transporte público



Fuente: Elaboración Propia

Por otra parte, para usuarios de estratos mayores, los cuales posiblemente tengan los modos de transporte privado disponible, se observa una mayor variabilidad en los valores subjetivos del tiempo de acceso, sin embargo, al igual que para los estratos bajos estos valores son mayores en el caso del tiempo de viaje.

4.1.2 Estimación Bayesiana

En este grupo de modelos se utilizó el software Nlogit, que permite la modelación de parámetros individuales. La base de datos se clasificó en dos partes de manera similar a los modelos con estimación clásica.

Usuarios de Transporte Privado

Se especificó un modelo logit mixto ML, involucrando las variables indicadas anteriormente, exceptuando el confort, la cual se incluye en la constante específica de los modelos. Se dividió la muestra entre los estratos socioeconómicos, los cuales para este grupo de datos son 2-3 y 4-5-6. La forma de la función de utilidad se especifica en la Tabla 4.8.

Tabla 4.8 Especificación de modelos mixed logit ML

Modelo	Expresión
ML1_Ei	$V_{iq} = ASC_i + (\beta_T + \eta_{iT})T_{iq} + \beta_C C_i + (\beta_{TACC} + \eta_{iTACC}) + \delta_1 \beta_{No_A} No_A_q + \delta_{2,4} (\beta_{Trans} Trans_{2,4q} + \beta_{Sexo} Sexo_q)$

Los modelos calibrados se presentan en la Tabla 4.9. Se puede apreciar que los signos son consistentes y los parámetros son en su mayoría significativos al 90%.

Para este grupo de encuestas el modelo se estimó con una muestra de 1125 individuos, divididos por estratos de la siguiente manera: 427 individuos de estratos 2, 3 y 698 de estratos 4, 5 y 6.

Los modos disponibles en este grupo de encuestas son: 1. Auto, 2. Metro, 3. Taxi y 4. Transmilenio. Los encuestados elegían entre dos modos en el cual siempre estaba incluido el Metro entre las alternativas disponibles.

Tabla 4.9 Modelos estimados para usuarios de transporte privado, estimación bayesiana

	Parámetro	ML1_E23	ML1_E456
Constantes Específicas	ASC1	0	0
	ASC2	-0,634 (-2,629)	-0,456 (-2,388)
	ASC3	0,561 (2,351)	0,266 (1,308)
	ASC4	-0,908 (-2,640)	-0,262 (-1,023)
Coefficientes de atributos	β_T (Tiempo) (1,2,3,4)	-0,052 (-5,995)	-0,0645 (-9,048)
	β_C (Costo) (1,2,3,4)	-0,000345 (-9,226)	-0,00032 (-10,436)
	β_{TACC} (Tiempo de Acceso) (1,2,3,4)	-0,0235 (-1,235)	-0,050 (-3,081)
	β_{TRANS} (Transbordo) (2,4)	-0,4807 (-6,173)	-0,477 (-7,031)
	β_{No_AUTOS} (No Autos) (1)	0,426 (3,251)	0,503 (5,55)
	β_{SEXO} (Sexo) (2,4)	0,426 (3,010)	0,557 (5,120)
	# de parámetros	11	11
	Logverosimilitud $L(\theta)$	-2045,892	-3195,783

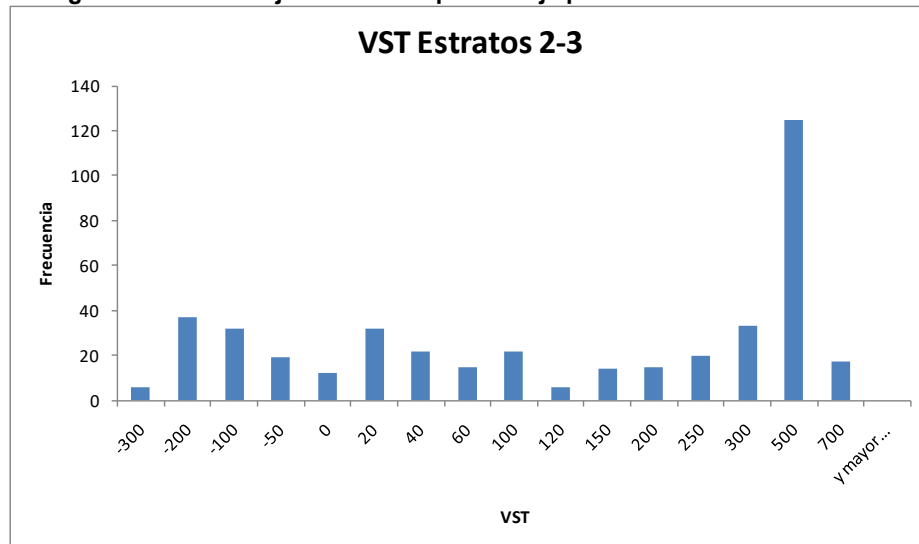
Fuente: Elaboración propia

Los dos modelos obtenidos por estrato son los siguientes:

- **ML1_E23.** Modelo para estrato dos y tres, se realizó con un total de 427 individuos encuestados. Las tasas marginales de sustitución para cada uno se calculó para el tiempo de viaje y el tiempo de acceso. Se obtuvo un valor de penalidad por el transbordo de \$1.139 valor considerablemente alto y afirma la poca disposición para aceptar el transbordo. Por otra parte, el valor medio de la penalidad del transbordo por minuto, es 9,24/min.

Ahora bien, al calcular el valor subjetivo del tiempo de viaje con los parámetros medios del tiempo y del costo sería de \$151/min, mientras que el valor subjetivo del tiempo de acceso sería de \$68/min. En todo caso, estos siguen siendo casos sin respuesta clara. Por otra parte, se puede resaltar que según investigaciones previas (Sillano & Ortúzar, 2005) a través de estas técnicas es posible identificar aquellos individuos que no responden de una manera seria y con respuestas concretas a la encuesta, presentándose una distorsión en los datos obtenidos.

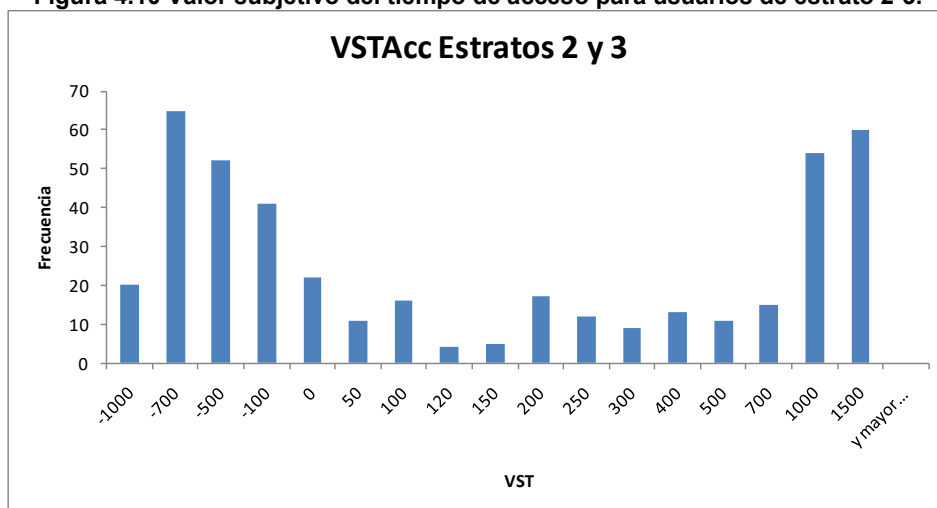
Figura 4.9 Valor subjetivo del tiempo de viaje para usuarios de estrato 2-3.



Fuente: Elaboración propia

Al analizar la Figura 4.9 y la Figura 4.10 puede notarse que la distribución del VST estimada por medio de las técnicas bayesianas no sigue la distribución Normal que fue especificada. Ello puede explicarse por el hecho de que mientras en los modelos logit mixto el VST se obtiene por medio de simulación y maximización extrayendo valores de una distribución preestablecida, en la estimación bayesiana no se utiliza tal procedimiento en tanto se obtienen parámetros que definen el comportamiento de cada individuo.

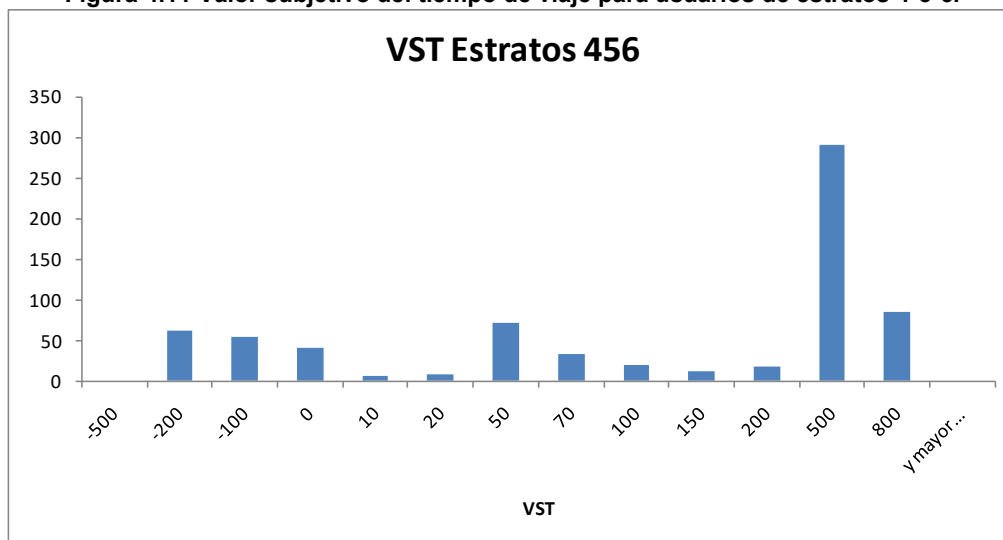
Figura 4.10 Valor subjetivo del tiempo de acceso para usuarios de estrato 2-3.



Fuente: Elaboración propia

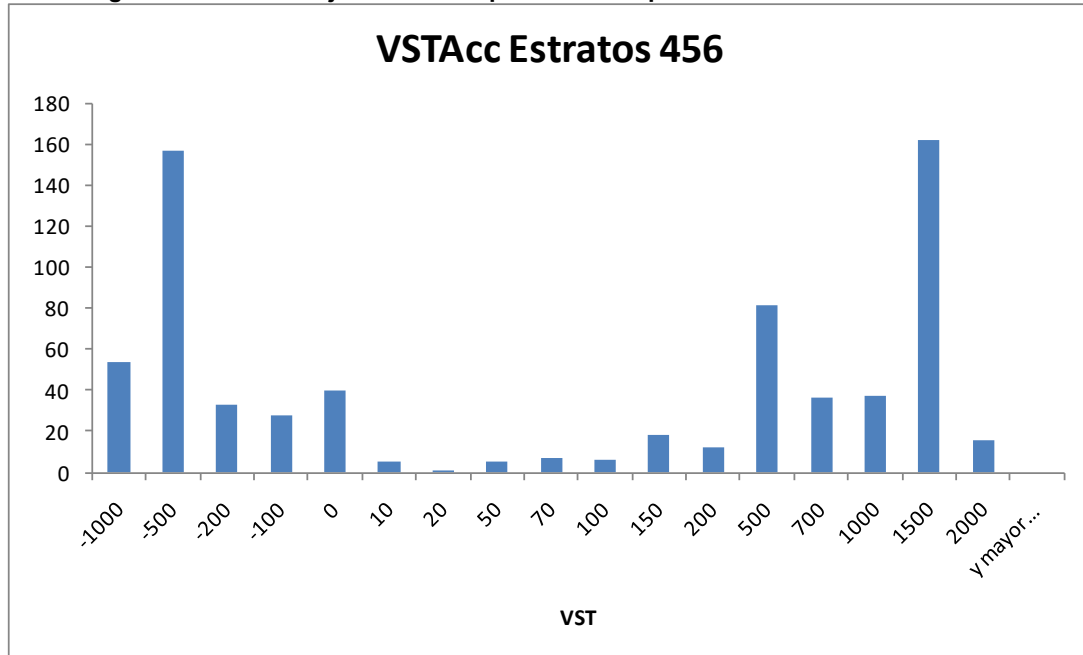
- **ML_E456.** Modelo para los estratos 4, 5 y 6, con 698 individuos encuestados. Los valores subjetivos del tiempo de viaje como valor medio \$205/min y de \$161/min para el caso del valor subjetivo del tiempo de acceso y una penalidad por transbordo de \$1.529/min.

Figura 4.11 Valor subjetivo del tiempo de viaje para usuarios de estratos 4-5-6.



Fuente: Elaboración propia

Figura 4.12 Valor subjetivo del tiempo de acceso para usuarios de estratos 4-5-6.



Fuente: Elaboración propia

Usuarios de Transporte Público

Este modelo involucra las mismas variables de los modelos anteriormente, pero agregando la variable confort en los tres niveles explicados en capítulos previos. Al igual que en los anteriores modelos se dividió la muestra entre estratos socioeconómicos. En la Tabla 4.10. se especifica la función de utilidad.

Tabla 4.10 Especificación de modelos mixed logit ML para usuarios de transporte público

Modelo	Expresión
ML1_Ei	$V_{iq} = ASC_i + \beta_T T_{iq} + \beta_C C_{iq} + \beta_{TAcc} Acc_{iq} + \delta_i \beta_{Trans} Trans_{iq} + \beta_{ConfB} ConfB + \beta_{ConfM} ConfM$

Para este grupo de encuestas el modelo se estimó con una muestra de 796 individuos, divididos por estratos de la siguiente manera: 510 individuos de estratos 1, 2 y 3 y 277 de estratos 4, 5 y 6.

Los modos disponibles en este grupo de encuestas son: 1. Metro, 2. Bus y 3. Transmilenio, y los usuarios debían elegir entre los tres, ninguno tenía disponible el taxi ni el auto particular para realizar su viaje cotidiano. Los modelos calibrados se presentan en la Tabla 4.11 donde se fija la constante específica del metro al igual que en los modelos anteriores.

Tabla 4.11 Modelos estimados para usuarios de transporte público, estimación bayesiana

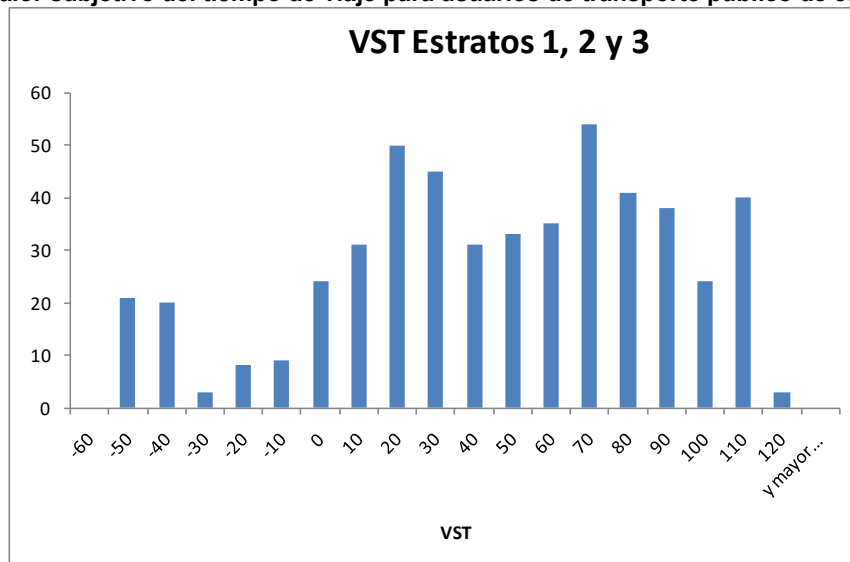
	Parámetro	ML1_E123	ML1_E456
Coefficientes de atributos	ASC1	0	0
	ASC2	-0,352 (-6,43)	-1,86 (-7,923)
	ASC3	-0,565 (-4,105)	-0,319 (-4,278)
	β_T (Tiempo) (1,2,3)	-0,101 (-11,016)	-0,144 (-8,208)
	β_C (Costo) (1,2,3)	-0,00251 (-14,522)	-0,00179 (-6,915)
	β_{TACC} (Tiempo de Acceso) (1,2,3)	-0,0547 (-6,097)	-0,0421 (-2,313)
	β_{TRANS} (Transbordo) (1,2,3)	-0,238 (-4,510)	-0,375 (-4,512)
	β_{Conf_B} (Confort Bajo) (1,2,3)	-0,637 (-8,451)	-1,225 (-9,100)
	β_{Conf_M} (Confort Medio) (1,2,3)	-0,2819 (-4,294)	-0,633 (-6,410)
	# de parámetros	10	10
	Logverosimilitud $L(\theta)$	-3646,058	-1762,86

Fuente: Elaboración propia

Los dos modelos obtenidos por estrato son los siguientes:

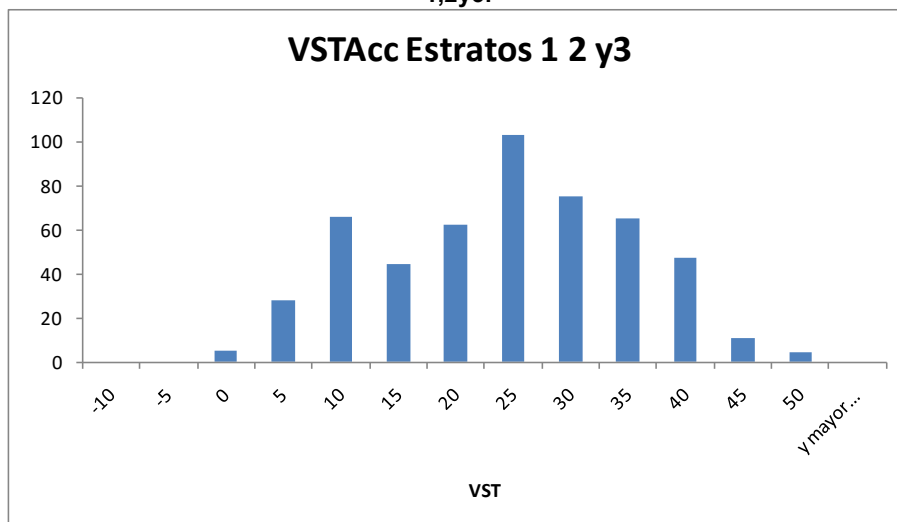
- ML1_E123.** Modelo realizado con 510 encuestados. Todos los parámetros son significativos al 90%, sin embargo la percepción de estos individuos al tiempo de acceso es menor de la esperada. Esto puede ser el resultado de que estos usuarios en su mayoría tienen tiempos de caminata y de espera considerablemente largos siendo más importante el tiempo de viaje en vehículo, al estar acostumbrados a caminar o esperar en el paradero. Esto se corrobora analizando las graficas de parámetros individuales, en la Figura 4.13 la distribución del valor subjetivo del tiempo de viaje, donde se aprecia mayor variabilidad en la población de estos valores. La penalización para el transbordo es del orden de \$95.

Figura 4.13 Valor subjetivo del tiempo de viaje para usuarios de transporte público de estratos 1, 2 y 3.



Fuente: Elaboración propia

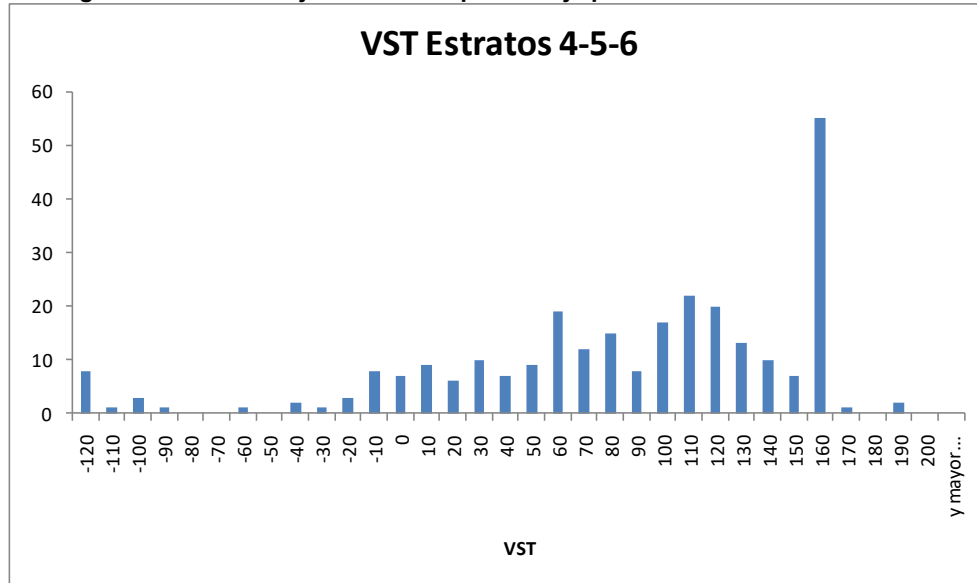
Figura 4.14 Valor subjetivo del tiempo de acceso para usuarios de transporte público de estratos 1,2y3.



Fuente: Elaboración propia

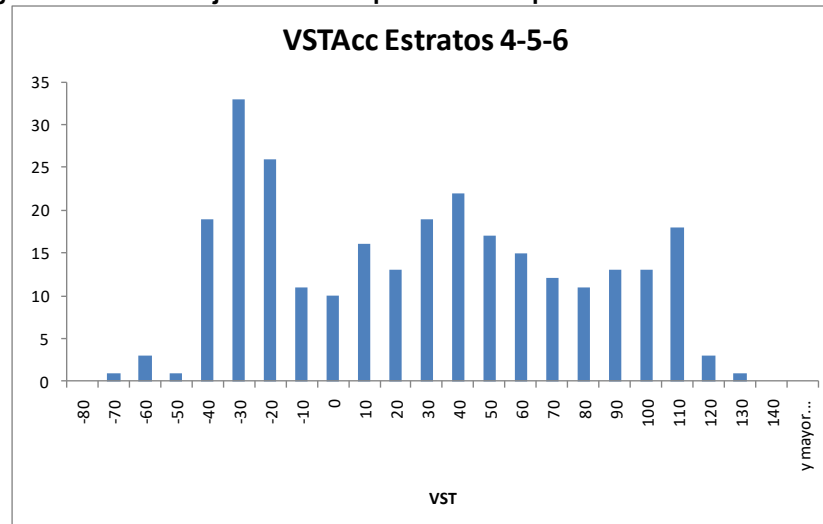
- ML_E456.** Modelo para los estratos 4, 5 y 6, con 277 individuos encuestados. Para esta muestra se observa un pico en un valor promedio de \$160/min para el caso del valor subjetivo del tiempo de viaje y pocos valores negativos un 10% del valor total de la muestra. Por otro lado, en el caso del valor subjetivo del tiempo de acceso los valores negativos son del orden de 34% ya un valor representativo para este parámetro. En la Figura 4.15 y la Figura 4.16 se presentan los histogramas para estos dos casos. En el caso del transbordo la penalidad es de \$209.

Figura 4.15 Valor subjetivo del tiempo de viaje para usuarios de estratos 456.



Fuente: Elaboración propia

Figura 4.16 Valor subjetivo del tiempo de acceso para usuarios de estratos 456.



Fuente: Elaboración propia

4.1.3 Comparación de modelos

Antes de iniciar la comparación entre las dos técnicas utilizadas en la estimación de valores subjetivos del tiempo, es menester revisar las disposiciones a pagar que se obtuvieron entre los dos tipos de usuarios estudiados a saber, usuarios de transporte privado y usuarios de transporte público, desde la perspectiva clásica.

Para la estimación de los parámetros del modelo logit mixto se empleó simulación de la manera como se describió en capítulos anteriores. En los dos casos en estudio (usuarios de transporte público y transporte privado) se utilizan 10.000 valores aleatorios tomados de la media y la desviación estándar dadas por el modelo para estimar los valores

subjetivos del tiempo de viaje y del tiempo de acceso. Sin embargo, para el caso 1 (usuarios de transporte privado) las distribuciones dibujan de una manera más clara la función escogida que en este caso ha sido la normal, presentando un porcentaje considerable de valores negativos y ceros y con una mínima diferencia entre los estratos bajos y altos. Los valores máximos y mínimos se encuentran en el mismo rango, desde \$2.000/min hasta -\$1.500/min, para VST de viaje. Sin embargo, para usuarios de estratos bajos se puede observar un pico en valores entre \$200/min y \$400/min, mientras que para estratos altos el pico está entre \$400/min y \$600/min.

Ahora bien, si comparamos estos valores con los estimados para el caso 2 (usuarios de transporte público) notamos que individuos de estratos bajos (1-2-3) tienen VST de viaje que oscilan entre -\$200/min y \$300/min, mientras que para estratos altos (4-5-6) los valores están entre -\$200/min y \$350/min. Con picos pronunciados en \$50/min y \$100/min, respectivamente. Con lo cual se podría concluir que la disposición a pagar por los ahorros en tiempo de viaje para usuarios de transporte privado podría llegar a ser hasta de 4 veces más que para usuarios de transporte público.

Esta misma comparación pero desde la perspectiva del VST de acceso, varían un poco. En el caso 1, tenemos mínimos y máximos desde -\$4.000/min hasta más de \$4.000/min, para los dos estratos analizados. Con un pico notable en valores medios de \$400/min para estratos bajos y un pico entre \$400/min y \$800/min para estratos altos. Ya en el caso 2, difiere un poco la forma de la función, ya que para estratos 1-2-3 la función tiene valores de VST desde -\$200/min hasta \$200/min, con un pico medio notable en \$50/min. Para todos los casos anteriores los valores negativos y ceros eran pocos, a diferencia del VST de acceso para estratos altos que tiene una proporción similar de valores negativos y positivos.

A través de la modelación bayesiana pudimos observar que los VST de viaje y de acceso para cada uno de los estratos analizados no dibujan una función normal como en el caso anterior. En el caso 1 se presentan VST de viaje desde -\$300/min hasta \$700/min con un notable pico en \$500/min, en estratos bajos y para estratos altos de -\$200/min hasta \$800/min con un pico menos notable en \$500/min.

Por otro lado, los VST de acceso en los estratos bajos presenta un pico entre \$1.000/min y \$1.500/min, mientras que en estratos altos este valor pico es máximo en \$1.500/min. Para los dos estratos los valores negativos y ceros pueden ser de hasta el 50% del total de la muestra. Sin embargo, esta técnica al permitir estimar las tasas marginales de sustitución por individuo de una manera directa, puede prescindirse de aquellos con parámetros positivos del tiempo y del acceso con lo cual podría mejorarse, más no eliminarse esta tendencia. Recordando, que en recientes investigaciones, la presencia de VST negativos pueden ser consecuencias de las respuestas o percepciones de los individuos (Cirillo & Axhausen, 2006) o de la función establecida para la estimación de los parámetros (Hess, Bierlaire, & Polak, 2005).

Un resultado importante de resaltar son las elasticidades respecto al tiempo y respecto al costo. Según los valores absolutos de los mostrados en la Tabla 4.12, inferiores a uno,

indican que la demanda es poco elástica respecto del tiempo y el costo. Los usuarios mas inelásticos son los de Transmilenio, como era de esperarse, pues allí es donde hay mayor cantidad de cautivos. Sin embargo, comparando para cada uno de los modos disponibles, el auto es el modo que tiene la mayor disminución de la demanda en cado de aumentar el tiempo de viaje y el costo, llegando a disminuir hasta en un 50% en caso del aumento del costo en un 100%, atributo respecto del cual hay mayor sensibilidad al compararlo con el tiempo.

Tabla 4.12 Elasticidades de los diferentes modos, respecto al tiempo y el costo.

Modo	Elasticidad/Tiempo	Elasticidad/Costo
Auto	-0,3611	-0,5060
Metro	-0,2483	-0,1937
Taxi	-0,1884	-0,3171
Transmilenio	-0,0229	-0,172

Por otro lado, los datos muestran mayor sensibilidad de demanda ante el costo y el tiempo para usuarios de auto que para usuarios de transporte público.

5 CONCLUSIONES

A través de la investigación se han propuestos modelos de elección discreta del tipo logit mixto incluyendo parámetros aleatorios individuales y comparando la información obtenida con técnicas comúnmente usadas. Para lograrlo, se estimaron inicialmente modelos logit mixtos por medio del método de máxima verosimilitud simulada, o estimación clásica y posteriormente se calcularon los modelos para parámetros individuales por medio de técnicas bayesianas.

La hipótesis básica de la investigación fue corroborada, encontrándose que es posible formular y estimar modelos de elección discreta con parámetros aleatorios con una adecuada robustez estadística y econométrica, y derivar de ellos estimaciones del valor subjetivo del tiempo de viaje y sus componentes.

Los modelos fueron estimados utilizando datos reales resultados de la aplicación de encuestas de preferencia declaradas en la ciudad de Bogotá a usuarios de transporte privado y transporte público. En las encuestas se pretendía analizar los cambios en la partición modal que pudiese ocurrir ante la introducción de un nuevo modo de transporte, específicamente el metro y un sistema integrado de transporte público.

La encuesta de preferencia declarada se aplicó, por una parte a 1125 usuarios de vehículo particular que realizan su viaje cotidiano al trabajo o al estudio, en las horas de la mañana y, por otro, a 787 de usuarios de transporte público que no tienen disponible el vehículo particular ni el taxi. Luego de procesar esta información y con la base de datos, se obtuvieron modelos logit multinomial, modelos con variaciones sistemáticas en el tiempo y el acceso y las variaciones aleatorias estimadas con el logit mixtos. Los atributos que se encontraron más relevantes en la elección modal fueron el tiempo de viaje en vehículo, el tiempo de acceso (suma del tiempo de caminata y el tiempo de espera), el costo o tarifa del viaje, el confort y el número de transbordos.

Las valoraciones subjetivas de las diferentes componentes del tiempo, así como las variables transbordo y confort se estimaron por tipo de usuario y por nivel socioeconómico. Este último se separó en dos componentes: usuarios de estratos bajos, del uno al tres y usuarios de estratos medio y alto, de cuatro a seis. Por otra parte, los usuarios con vehículo particular disponible, llamados usuarios de transporte privado, y por otro lado los usuarios de transporte público, los cuales no tienen entre sus alternativas disponibles el auto particular o el taxi.

Se muestra en la tabla los resultados obtenidos de los modelos con variaciones sistemáticas, con el fin de comparar los valores estimados según los niveles socioeconómicos.

Modelo	Transporte Privado	Transporte Público
VSTV Estratos 23 (\$/min)	102	23
VSTAcc Estratos 23 (\$/min)	140	28
VST Estratos 456 (\$/min)	154	38
VSTAcc Estratos 456 (\$/min)	177	22
Transbordos (\$/Transbordo)	1175	137

El ingreso promedio per cápita del año 2006 en la ciudad de Bogotá es de \$720.000 mensuales³, sin embargo, se ha establecido que es un valor que difiere según los estratos socioeconómicos, por lo tanto, para estratos bajos (1-2-3) el ingreso llega a estar cerca de los \$375.000 mensuales, mientras que para estratos altos (4-5-6) el promedio per cápita es de aproximadamente \$1.100.000 mensuales⁴. Tomando un día laboral normal de ocho horas de trabajo, el ingreso por minuto sería de \$27/min para estratos bajos y \$72/min para estratos altos. Consecuentemente, la disposición a pagar por los ahorros en tiempos de viaje para usuarios de transporte privado en estratos bajos es de hasta siete veces más del valor de un minuto de salario y de hasta tres veces más para usuarios de estratos altos. Para el caso de usuarios de transporte público, la valoración subjetiva de los tiempos es más cercana al valor del salario por minuto, entendiéndose que para este tipo de usuarios la disposición a pagar por los ahorros es menor y que al no tener disponible otros modos más que el transporte público la transferencia modal no es tan evidente.

En los modelos con variaciones sistemáticas al agregarle una variable adicional como sexo u ocupación las valoraciones subjetivas se mantienen constantes, no presentando aumentos ni disminuciones considerables. De tal manera que estas variables no afectan los valores subjetivos poblacionales, a pesar que las bondades del modelo mejoran.

Es evidente la fuerte correlación que hay entre ingreso y valoración del tiempo de viaje, notando además que los usuarios de transporte privado, que en general son individuos de más alto ingreso, tienen valores subjetivos mayores que los usuarios de transporte público. Es así como se encontró que el VST del tiempo de viaje en usuarios de transporte privado en promedio fue de \$133/min, el del tiempo de acceso de \$160/min y por último la penalidad por transbordo del orden de \$1.179. A su vez, para los usuarios de transporte público estos valores son en promedio de \$30/min para los tiempos de viaje y acceso y de \$130 la penalidad por el transbordo.

Los VST del tiempo utilizados en la actualidad por el proyecto Metro en la ciudad de Bogotá, guardan similitud con los mostrados en esta investigación. Estos valores suman todas las componentes del tiempo, es decir, el tiempo de viaje en vehículo, el tiempo de

³ Según Encuesta de Hogares del DANE, 2003, ajustado por inflación.

⁴ Ingresos, gastos y exclusión social en Bogotá. Secretaria de Hacienda Distrital, 2006.

caminata y el tiempo de acceso. De esta manera, para estratos 1, 2 y 3, los VST son de \$25/min, \$51/min y \$76/min, respectivamente. Y para estratos altos de \$135/min.

Refiriéndose a las estimaciones individuales del VST individuales, bien sea por el procedimiento clásico o bayesiano, los valores negativos y ceros presentes son reflejo de la presencia de aquellos factores que inevitablemente el modelador no puede observar, como son aquellos que disfrutan viajar, aquellos que no responden seriamente a las encuestas, aquellos que en su actividad cotidiana el moverse es su principal actividad, o en ciertos casos algunos individuos no entienden la formulación de las preguntas o de los atributos y sus respuestas no se asemejan a la realidad al momento de elegir un modo de transporte. Por otro parte, en investigaciones recientes también se ha estudiado la importancia de la función utilizada, donde se ha visto que la función normal presentará valores negativos, que podría solucionarse si se cambia la función a una lognormal. Sin embargo, no se deben desechar estos valores a priori sin antes analizar si estos individuos tienen algún comportamiento que imprimen un poco más de realidad al modelo en cuestión.

Pudo notarse que es ligeramente mayor la elasticidad con respecto del costo que del tiempo. En general, la elasticidad del costo está en el rango $-0,2$ a $-0,5$, y la del tiempo de $-0,1$ a $-0,4$. Se observó, además, que hay importantes diferencia en las elasticidades entre modos, siendo más inelástica ante el precio y el tiempo de viaje la demanda de Transmilenio.

En las valoraciones de los tiempos subjetivos tanto de viaje como de acceso, para el caso de la estimación clásica las percepciones estimadas son siempre mayores que en el caso bayesiano, esto era de esperarse, teniendo en cuenta que estudios recientes demuestran que la estimación por máxima verosimilitud tiende a sobreestimar tales valores en comparación con las técnicas bayesianas, como consecuencia de la maximización.

Futuras investigaciones sobre la estimación del VST permitirán verificar la evolución de las percepciones en el individuo, involucrando variables latentes. Además, comparar los resultados que aquí se obtuvieron considerando otros contextos de aplicación de los modelos de elección discreta. Resulta igualmente interesante realizar estimaciones del valor del tiempo para el transporte de carga.

6 BIBLIOGRAFÍA

- Allenby, G. (1997). An introduction to hierarchical Bayesian modeling. *Advanced Research Techniques Forum, American Marketing Association*.
- Allenby, G., & Lenk, P. (1994). Modeling household purchase behavior with logistic normal regression. *Journal of the American Statistical Association* 89 , 1218–1231.
- Anderson, D., & Wiley, J. (1992). Efficient Choice Set Designs for Estimating Cross-Effects Models. *Marketing Letters* , 357-70.
- Bates, J., & Roberts, M. (1986). Value of time research: summary of methodology and findings. *Proceedings 14th PTRC summer Annual Meeting, University of Sussex* .
- Becker, G. (1965). A theory of the allocation of time. *The economic journal* , 493-517.
- Beesley, M. (1974). Conditions for successful measurement in time valuation studies Behavioral Travel Demand Modelling and Valuation of Travel Time. *Transportation Research Board Special Report 149* , 161± 172.
- Ben-Akiva, D., Bolduc, D., & Bradley, M. (1993). Estimation of Travel Choice Models with Randomly Distributed Values of Time. *Transportation Research Record 1413* , 88-97.
- Ben-Akiva, M., & Lerman, S. (1985). Discrete choice analysis: theory and application to travel demand. *Cambridge. Mass.: The MIT Press* .
- Ben-Akiva, M., & Morikawa, T. (1990). Estimation of travel demand models from multiple data sources. *Transportation and Traffic Theory* , 461–476.
- Bhat, C. (2001). Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model. *Transportation Research B* 35 , 677–693.
- Bierlaire, M., Bolduc, D., & McFadden, D. (2007). The estimation of generalized extreme value models from choice-based samples. *Transportation Research Part B: Methodological Vol 42* , 381-394.
- Bliemer, M., & Rose, J. (2008). Efficient stated choice experiments allowing for differences in error variances across subsets of alternatives. *Transportation Research B* .
- Bliemer, M., Rose, J., & Hess, S. (2006). Approximation of Bayesian Efficiency in Experimental Choice Designs. *INSTITUTE of TRANSPORT AND LOGISTICS STUDIES* .
- Boyd, J., & Mellman, R. (1980). The Effect of Fuel Economy Standards on the U.S. Automotive Market: An Hedonic Demand Analysis. *Transportation Research A* 14 , 367-378.
- Bradley, M., & Daily, J. (1997). Estimation of logit choice models using mixed stated preference and revealed preference information. *Oxford: Pergamon* .

- Bradley, M., & Kroes, E. (1990). Forecasting issues in stated preference survey research. *69th TRB Annual meeting*. Washington D.C.
- Bunch, D., Louviere, J., & Anderson, D. (1996). *A Comparison of Experimental Design Strategies for Choice-Based Conjoint Analysis with Generic Attribute Multinomial Logit Models*.
- Cantillo, V., & Ortúzar, J. d. (2005). A semi-compensatory discrete choice model with explicit attribute thresholds of perception. *Transportation Research 39B* , 641-657.
- Cardell, N., & Dunbar, F. (1980). Measuring the Societal Impacts of Automobile. *Transportation Research A 14* , 423-434.
- Cardell, N., & Reddy, B. (1977). A Multinomial Logit Model Which Permits Variations in Tastes Across Individuals. *Working Paper, Charles River Associates, Boston*.
- Caussade, S., Ortúzar, J. d., Hensher, D., & Rizzi, L. (2005). Assessing the influence of design dimensions on stated choice experiment estimates. . *Transportation Research Part B, Vol. 39* , 621-640.
- Chesher, A., & Santos Silva, J. (2003). Taste Variation in Discrete Choice Models. *Review of economic studies* , 147-168.
- Cirillo, C., & Axhausen, K. (2006). Evidence on the distribution of values of travel time savings from a six-week diary. *Transportation Research Part A* , 444-457.
- De Serpa, A. (1971). A theory of the economics of time. *Economic Journal 81* , 828-846.
- Domencich, T., & McFadden, D. (1975). Urban Travel Demand: A Behavioural Analysis.
- Fernandez, J. (1997). Apuntes de Economía de Transporte. *Pontificia Universidad Católica de Chile* .
- Figuroa, O. (2005). Transporte urbano y globalización. Políticas y efectos en América Latina. *Revista eure* , XXXI (94), 41 - 53.
- Fowkes, A., & Wardman, M. (1988). The design of stated preference travel choice experiments with special reference to interpersonal taste variations. *Journal of Transport Economics and Policy 22* , 27-44.
- Garrido, R., & Silva, M. (2004). Low discrepancy sequences for the estimation of mixed logit. *Transportation Research B* .
- Gaudry, M., Jara-Díaz, S., & Ortúzar, J. (1989). Value of time sensitivity to model specification. *Transportation Research 23B* , 151-158.
- Godoy, G., & Ortúzar, J. (2004). On the estimation of mixed logit models. *Transportation Research Trends* , 289-310.

Golub, A., Balassiano, R., Araújo, A., & Ferreira, E. (2009). Regulation of the informal transport sector in Rio de Janeiro, Brazil: welfare impacts and policy analysis. *Transportation* , 601-616.

Greene, W. (2000). *Econometric Analysis*. NJ: Prentice Hall, Upper Saddle River.

Halton, J. (1960). On the efficiency of certain quasi random sequences of points in evaluating multi dimensional integrals. *Numer Math* 2 , 84-90.

Hensher, D. (2001). The valuation of commuter travel time savings for car drivers in New Zealand: Evaluating alternative model specifications. *Transportation* , 112-118.

Hensher, D., & Greene, W. (2003). The mixed logit model: the state of practice. . *Transportation* 30 , 133-176.

Hensher, D., Louviere, J., & Swait, J. (1999). Combining sources of preference data. *Journal of Econometrics* 89 , 197-221.

Hess, S., Bierlaire, M., & Polak, J. (2005). Estimation of value of travel-time savings using mixed logit models. *Transportation Research Part A* , 221–236.

Horswill, M. S., & Helman, S. (2003). A behavioral comparison between motorcyclists and a matched group of non-motorcycling car drivers: factors influencing accident risk. *Elsevier Science Ltd., Accident Analysis and Prevention* 35 (2003) , 589–597.

Huber, J., & Zwerina, K. (1996). The Importance of Utility Balance in Efficient Choice Designs. *Journal of Marketing Research* , 307-317.

Jara-Diaz, S., Ortúzar, J. d., & Parra, R. (1988). Valor subjetivo del tiempo considerando efecto ingreso en la partición modal. *Actas del V Congreso Panamericano de Ingeniería de Tránsito y Transporte*. Mayaguez.

Jiang, M., & Morikawa, T. (2004). Theoretical analysis on the variation of value of travel time savings. *Transportation Research Part A* 38 , 551-571.

Kocur, G., Adler, T., HYman, W., & Aunet, B. (1982). Guide to forecasting travel demand with direct utility assessment. *Urban Mass Transportation Administration, U.S. Department of Transportation. Washington D.C.*

Louviere, J., & Woodworth, G. (1983). Design and Analysis of Simulated Consumer Choice or Allocation Experiments: A Method Based on Aggregate Data. *Journal of Marketing Research* , 350-367.

Louviere, J., Hensher, D., & Swait, J. (2000). *Stated Choice Methods: Analysis and Applications*. Cambridge University Press, Cambridge .

Mackie, P., Jara-Diaz, S., & Fowkes, A. (2001). The value of travel time savings in evaluation. *Transportation Research Part E: Logistics and Transportation Review* , 91-106.

McFadden, D. (2001). Disaggregate behavioural travel demand's RUM side: a 30-year retrospective. *Travel Behavior Research* , 1-5.

McFadden, D. (1981). Structural Analysis of Discrete Data: With Econometric Applications.

McFadden, D. (1974). The measurement of urban travel demand. *Journal of public economics* 3 , 303-328.

McFadden, D., & Train, K. (2000). Mixed MNL models for discrete response. *Journal of Applied Econometrics* 15 , 447-470.

Mehndiratta, R., & Hansen, M. (1997). Analysis of Discrete Choice Data with Repeated Observations: Comparison of Three Techniques in Intercity Travel Case. *Transportation Research Record* 1607 , 69-73.

Mendieta Reyes, O. M., & Vizcaino de la Cruz, L. R. (2008). *Estudio del Fenomeno del Mototaxismo y las Variables que lo Definen en la Ciudad de Barranquilla*. Barranquilla: Universidad del Norte.

Metz, D. (2004). Travel time—variable or constant? *Journal of Transport Economics and Policy*, 38(3) .

Morey, E., & Rossman, K. (2002). Using Stated-Preference Questions to Investigate Variation in Willingness to Pay for Preserving Marble Monuments: Classic Heterogeneity and Random Parameters. *Economics Department, University of Colorado* .

Nerella S, S., & Bhat, C. R. (2004). *Numerical Analysis of Effect of Sampling of Alternatives in Discrete Choice Models*. Washington, DC 20001 USA: Transportation Research Record No. 1894, Travel Behavior and Values 2004.

Nerella, S., & Bhat, C. (2004). Numerical analysis of effect of sampling of alternatives in discrete choice models. *Transp. Res. Record* 1894 , 11-19.

Ortúzar, J. d. (1998). *Modelos de Demanda de Transporte*. Santiago de Chile: Ediciones Universidad Católica de Chile.

Ortúzar, J. d. (2000). Modelos econométricos de elección discreta. Ediciones Universidad Católica de Chile.

Ortúzar, J. d., & Willumsen, L. G. (2001). *Modelling Transport (3ª Edición Corregida)*. Chichester: John Wiley and Sons.

Paskov, S. (1994). Computing hig dimensional integrals with applications to finance. *Joint Summer Research Conference on Continuous Algorithms and Complexity*.

Paskov, S. (1997). New methodologies for Valuing Derivatives. En *Mathematics of Derivative Securities*. Cambridge University Press.

Paskov, S., & Traub, J. (1995). Faster valuation of financial derivatives. *The Jorunal of Portafolio Mangement* , 113-120.

Pérez, P., Martínez, F., & Ortúzar, J. d. (2003). Microeconomic Formulation and Estimation of a Residential Location Choice Model: Implications for the Value of Time. *Journal of Regional Science Vol 43* , 771-789.

Prasetyo, I., Fukuda, D., Yoshino, H., & Yai, T. (2003). Analysis of travel time saving benefit by understanding individual needs and value of activity time - Case study of Tokyo and Jakarta. *Traveler behavior and values 2003 transportation research record* , 12-21.

Revelt, D., & Train, K. (1998). Mixed logit with repeated choices. *Review of Economics and Statistics 80* , 647–657.

Richardson, A. (2003). Some evidence of travelers with zero value-of-time. *Transportation Research Record* .

Rose, J., & Bliemer, M. (2006). Designing Efficient Data for Stated Choice: Accounting for Socio-demographic and contextual effects in designing stated choice experiments. *11th International Conference on Travel Behaviour Research*. Kyoto, Japan,.

Sándor, Z., & Wedel, M. (2001). Designing conjoint choice experiments using managers' prior beliefs. *Journal of Marketing Research, Vol. 38* , 430-444.

Sándor, Z., & Wedel, M. (2002). Profile Construction in Experimental Choice Designs for Mixed Logit Models. *Marketing Science, 21(4)* , 455-475.

SawtoothSoftware. (1999). *The CBC/HB module for hierarchical Bayes*.

Sillano, M., & Ortúzar, J. d. (2005). Willingness-to-pay estimation with mixed logit models: some new evidence. *Environment and Planning A* , 525-550.

Srinivasan, K., & Mahmassani, H. (2005). Dynamic kernel logit model for the analysis of longitude discrete choice data: Properties and computational assessment. *Transportation Science 39* , 160–181.

Train, K. (2009). *Discrete Choice Methods with Simulation*. Cambridge University Press, Cambridge.

Train, K. (1998). Recreation Demand Models with Taste Differences Over People. *Land Economics, Vol. 74 (2)* , 230 - 239.

Train, K., & Mcfadden, D. (1978). The goods/leisure tradeoff and disaggregate work trip mode choice model. *Transportation Research 12* .

Train, K., McFadden, D., & Goett, A. (1987). Consumer attitudes and voluntary rate schedules for public utilities. *Review of Economics and Statistics LXIX* , 383–391.

Train, K., McFadden, D., & Goett, A. (1987). Consumer Attitudes and Voluntary Rate Schedules for Public Utilities. *Review of Economics and Statistics, Vol LXIX* , 383-397.

Vilain, P., & Bhandari, N. (2002). Differences in subjective and social value of time - Empirical evidence from traffic study in Croatia. *Transportation and public policy 2002 transportation research record* .

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

Welch, M., & Williams, H. (1997). The sensitivity of transport investment benefits to the evaluation of small time savings. *Journal of Transport Economics and Policy* 31 , 231-254.

Williams, H. (1977). On the formation of travel demand models and economic evaluation measures of user benefit. *Environment and Planning, Vol A. 9* , 285-344.

Zwerina, K., Huber, J., & Kuhfeld, W. (1996). A General Method for Constructing Efficient Choice Designs. *Journal of Marketing Research* .